University of Washington
Department of Civil and Environmental Engineering

USING CLIMATE MODEL ENSEMBLE FORECASTS for SEASONAL HYDROLOGIC PREDICTION

Andrew Wood

Water Resources Series
Technical Report No.177
January 2004

Seattle, Washington
98195

Department of Civil Engineering
University of Washington
Seattle, Washington 98195

USING CLIMATE MODEL ENSEMBLE FORECASTS for SEASONAL HYDROLOGIC PREDICTION

Andrew Wood

Water Resources Series
Technical Report No. 177

January 2004

UNIVERSITY OF WASHINGTON

Abstract

USING CLIMATE MODEL ENSEMBLE FORECASTS FOR SEASONAL HYDROLOGIC PREDICTION

Andrew W. Wood

Chair of the Supervisory Committee

Professor Dennis P. Lettenmaier
Department of Civil and Environmental Engineering

Seasonal hydrologic forecasting has long played an invaluable role in the development and use of water resources. Despite notable advances in the science and practice of climate prediction, current approaches of hydrologists and water managers largely fail to incorporate seasonal climate forecast information that has become operationally available during the last decade. This study is motivated by the view that a combination of hydrologic and climate prediction methods affords a new opportunity to improve hydrologic forecast skill. A relatively direct statistical approach for achieving this combination (i.e., downscaling) was formulated that used ensemble climate model forecasts with a six month lead time produced by the NCEP/CPC Global Spectral Model (GSM) as input to the macroscale Variable Infiltration Capacity hydrologic model to produce ensemble runoff and streamflow forecasts. The approach involved the bias correction of climate model precipitation and temperature fields, and spatial and temporal disaggregation from monthly climate model scale (about 2 degrees latitude by longitude) fields to daily hydrology model scale (1/8 degrees) inputs. A qualitative evaluation of the approach in the eastern U.S. suggested that it was successful in translating climate forecast signals to local hydrologic variables and streamflow, but that the dominant influence on forecast results tended to be persistence in initial hydrologic conditions. The suitability of the statistical downscaling approach for supporting hydrologic simulation was then assessed (using a continuous retrospective 20-year climate simulation from the DOE Parallel Climate Model) relative to dynamical downscaling via a regional, meso-scale climate model. The statistical approach generally outperformed the dynamical approach, in that the dynamical approach alone required additional bias-correction to reproduce the retrospective hydrology as well as the statistical

approach. Finally, using 21 years of retrospective forecasts for the western U.S., the skill of the GSM-based hydrologic forecasts was assessed relative to NWS Extended Streamflow Prediction (ESP) method forecasts. Because of unexceptional GSM climate forecasts, the GSM-based and ESP hydrologic forecasts generally showed similar skill. During strong ENSO anomalies, however, GSM-based forecasts yielded higher forecast skill in the Sacramento-San Joachin and Columbia River basins, but lower skill in the Colorado and upper Rio Grande River basins.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number

LIST OF TABLES

ACKNOWLEDGEMENTS

This publication was supported in part by a grant to the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) at the University of Washington, under NOAA Cooperative Agreements NA17RJ1232 (Contribution 911) and NA67RJ0155 (Contribution 924), and as part of the GEWEX Continental-Scale International Project (GCIP). In addition, the U.S. Department of Energy Office of Science's (BER) Accelerated Climate Prediction Initiative provided funding for the research of Chapter III, which was performed at the University of Washington under Grant No. 354967-AQO and at Pacific Northwest National Laboratories under support from the US Dept. of Energy Office of Biological and Environmental Research.

This work owes much to the insight and assistance of the faculty, staff and graduate students in the Civil and Environmental Engineering Department at the University of Washington. In particular, Dennis Lettenmaier has provided invaluable guidance and assistance for almost a decade, and my work has benefited immeasurably from his critical eye. Rick Palmer and Steve Burges have been consistently supportive and have supplied many thoughtful insights over the years, on both professional and personal subjects. Renate Staub has been an enthusiastic provider of wisdom on a wide and useful range of topics. Outside of the department, reading committee member Arun Kumar has helpfully supplied much of the data at the heart of the dissertation, in addition to feedback on matters beyond my expertise. Finally, I acknowledge my three siblings, who to their eternal credit have asked many good questions over the years (and gamely listened to the answers), and have been an unconditional source of support. I greatly appreciate the consideration and effort of those in this small group – they have both enabled and enriched my pursuit of this degree.

DEDICATION

This dissertation is dedicated to my parents, for their love, support and guidance throughout my life; to Elizabeth, for going the extra mile when the finish proved elusive; and to Patrick, whose timely arrival made it all the more interesting.

I. INTRODUCTION

Hydrologic predictability varies widely between events striking at the shortest scales in time and space and those developing at the longest. At shorter end of this continuum, the flash flood captures the imagination for the rapidity of its streamflow transition from dry to wet extremes, and for the immediacy of its consequences, often exacerbated by a lack of prior warning, for humans and infrastructure. The timing of onset and the evolution of a flash flood are very difficult to predict because the processes involved, in the land surface and the atmosphere, persist only at time scales on the order of mere minutes to hours. Fortunately, most other challenges in water management, and those having the more lasting effects on human society, involve atmospheric and hydrologic variability over relatively larger scales in time and space. The Great Flood of 1927 along the Mississippi River, which at its peak carried three times the flow of the Mississippi River flood in 1993, followed months of extremely wet, cold climate which swelled tributaries in at least a dozen states throughout the midwestern U.S. (Barry, 1997). Severe droughts, such as the 1930s U.S. Dust Bowl drought that occasioned John Steinbeck's novel "The Grapes of Wrath" and the 1970s drought in the Sahel of northwest Africa, generally result from climate anomalies lasting years to decades. Both flood and drought extremes can be manifest at the sub-continental scale.

This dissertation focuses on monthly to seasonal hydrologic forecasting, which in many ways is a more tractable problem than flash flood or decade-long drought scale prediction. The governing geophysical phenomena for monthly to seasonal climate and hydrologic variations derive from ocean-atmosphere and land-atmosphere interactions that have much greater inertia than the combination of convective precipitation and infiltration excess runoff processes that lead to most flash floods. The slower evolution of the particular ocean and land features (e.g., oceanic thermal gradients and snow pack or soil moisture) that serve as boundary conditions for the atmosphere affords a greater opportunity for observations, and for prediction at lead times long enough to be of practical use.

Hydrologists have employed a variety of methods for seasonal streamflow forecasting. The most basic forecast of historically observed streamflow averages (climatology) is readily surpassed in accuracy by methods that capture the persistence in streamflow itself and/or the predictive value of the physical drivers of streamflow (primarily snowpack, where relevant, and soil moisture). Since at least the first half of the 20th century, graphical curve methods and so-called "index methods" (both essentially regression) were used to forecast seasonal streamflow volumes (e.g., Parshall, 1948; MacLean, 1948) and derivatives of these methods still underpin the approaches of U.S. agencies charged with streamflow forecasting such as the Natural Resource Conservation Service (NRCS) (Garen, 1992). By the 1970s, statistical methods adopted for seasonal streamflow forecasting had evolved in complexity, and included, for instance, stochastic autoregressive (e.g., Burges and Johnson, 1973) and moving average or combined models (Box and Jenkins, 1976).

The rise of computing in the 1950s, however, fostered the implementation of digital conceptual (also called physical) hydrologic models, among the first of which was the Stanford Watershed Model (SWM) of Crawford and Linsley (1966). In contrast to the NRCS's statistical forecasting framework, the dynamical simulation approach took root in the National Weather Service (NWS) with the NWS River Forecasting System (NWSRFS) model (Anderson, 1973), an offspring of SWM. NWSRFS is currently used in a procedure called Extended Streamflow Prediction (ESP: Twedt *et al*., 1977), in which historically observed streamflow sequences are combined with an initial condition estimate to generate a probability distribution of forecast outcomes. The basic concept of ESP is common to most physical approaches to hydrologic prediction in that it relies on the accurate specification of both an initial hydrologic state and of future climate. The first task is now reasonably well accomplished using physical hydrologic models (the more detailed of which simulate the land surface water and energy balance with spatially distributed representations of vegetation and soil water storage, snowpack, and associated moisture and energy fluxes), given adequate meteorological inputs prior to the forecast date, and supported where possible by timely snow or soil moisture observations.

For the second task, the specification of future climate, however, hydrologists traditionally have used climatology in one form or another. Only in the last decade has a body of work been initiated that links with the forecasting progress made in the atmospheric and oceanographic

sciences. Scientists have long studied relationships between seasonal climate patterns and ocean phenomena, by some accounts since the turn of the 20th century (Goddard *et al.*, 2001). In the last several decades, much insight has been gained into the role of ocean thermal dynamics and associated atmospheric circulation patterns in determining regional climate (e.g., Livezey, *et al*., 1997; Shukla, 1998; Palmer and Anderson, 1994), in part due to improvements in monitoring (both remote and direct) of the ocean and atmosphere, with consequent improvements in seasonal climate forecasting (McPhaden *et al*., 1998). The most widely recognized of these ocean-atmosphere "teleconnections", and the source of most seasonal climate predictability, is the El-Nino Southern Oscillation (ENSO) (Philander, 1990), although other ocean-atmosphere dynamics such as the Pacific Decadal Oscillation (PDO) are also thought to constrain North American climate (Mantua *et al.,* 1997; Hamlet and Lettenmaier, 1999b). The sea surface temperature (SST) defined ENSO state evolves gradually enough that it can be forecasted with moderate skill at lead times of months to seasons, either with statistical methods or dynamic coupled ocean-atmosphere general circulation models (OAGCMs) (Barnston *et al*., 1996, 1999; Livezey, 1990; Kumar *et al*., 1996; Livezey *et al*., 1996). A combination of statistical and dynamical techniques for forecasting SSTs and associated seasonal climate has been adopted at a number of research centers (e.g., National Centers for Environmental Prediction (NCEP), NASA Seasonal to Interannual Prediction Project, Columbia University's International Research Institute and the European Center for Medium Range Weather Forecasting), resulting in sophisticated operational systems for seasonal climate forecasting.

Climate forecasts are generally presented with the sub-continental region (e.g., the southwestern U.S.) as the smallest spatial unit, in part because of the long-recognized difficulty of atmospheric models in reproducing observed climate at smaller scales (e.g., less than ~10^7 km^2) (IPCC, 1996; Anderson and Stern, 1996) and because of computational constraints. In addition to the scale problem, the land surface variables of greatest interest to hydrologists and to society, such as surface precipitation and runoff, are generally predicted less reliably in climate models than are features of large scale circulation (see, e.g., Risbey and Stone, 1996). As a result, connecting climate forecasts to the scales and features of the hydrosphere in which human society is often most interested -- regional and smaller, for land surface variables – has been problematic.

Complementing hydrologists' interest in bridging the scale gap (generally termed downscaling), the climate research community has sponsored a number of initiatives to foster use of climate simulations for hydrologic and water resources objectives. The primary objective of the World Climate Research Program's supported Global Energy and Water Cycle Experiment (GEWEX) Continental-Scale International Project (GCIP), for instance, is "to demonstrate skill in predicting water resources on timescales of up to seasonal and annual", with a particular emphasis on improving "the utility of hydrologic predictions for water resources management" (National Research Council, 1998).

Atmospheric scientists, and to a lesser extent, hydrologists have developed a number of downscaling methods. These include dynamical approaches that use finer resolution (mesoscale) atmospheric models nested within boundaries at which fluxes are taken from OAGCMs (e.g., Cocke and LaRow, 2000; Giorgi and Mearns, 1991), statistical approaches (Wilby and Wigley, 1997; Wilby *et al*., 1998) and climate-analog approaches (IPCC, 1996; Leung *et al*., 1999). Comparisons of dynamical and statistical methods (Murphy, 1999; Kidson and Thompson, 1998; Hay *et al*., 2002) indicate that in addition to being numerically cumbersome, dynamic downscaling must be accompanied by further statistical adjustment to address model biases. Of these approaches, hydrologists have tended to opt for simple strategies such as OAGCM-conditioned compositing, in which OAGCM output is used to guide the construction or weighting of an ensemble of historically observed meteorological time series, which subsequently are used as input to a hydrologic model (e.g., Georgakakos *et al*., 1998; Leung *et al*., 1999). A variation on this approach is to derive the conditioning signal from SST-classified climate modes, such as PDO and ENSO (Hamlet and Lettenmaier, 1999a-b). With notable exceptions (e.g., Wilby *et al*., 2000; Kim *et al*., 2000; Hay *et al*., 2002; Hamlet and Lettenmaier, 1999b), few researchers have examined in much detail the utility of downscaling methodologies for supporting hydrologic simulation, particularly for forecasting purposes.

The relative dearth of research utilizing climate model forecasts to improve hydrologic forecasting at monthly to seasonal lead times scale motivates the science and engineering questions that this dissertation seeks to address:

1. Can we implement a relatively direct statistical approach for using climate model ensemble forecasts as forcings for hydrologic ensemble forecasts, given that the approach

must (a) eliminate or reduce climate model regional biases and (b) downscale the outputs from the multi-degree spatial scale (of the climate model) to scales of interest for hydrologic forecasting, while (c) preserving climate model forecast information?

2. How would such an approach perform relative to more physically based (but computationally intensive) method for deriving hydrologic forecasts, such as dynamical downscaling?

3. Where in the western U.S. and at what times of the year does the forecast approach yield improved skill in predicting streamflow and other hydrologic variables relative to methods currently used in practice?

Research addressing these questions is the focus of the following three chapters. Chapter II (published as: Wood, A.W., Maurer, E.P., Kumar, A. and D.P. Lettenmaier, 2002. Long range experimental hydrologic forecasting for the eastern U.S., *Journal of Geophysical Research*, Vol. 107, D20, October) presents an approach for translating climate model ensemble forecasts into hydrologic forecasts, and demonstrates it for a case study in the eastern U.S. Using a retrospective climate model simulation, Chapter III (accepted for publication in a special issue of the journal *Climatic Change*: Wood, A.W., Leung, L. R., V. Sridhar and D.P. Lettenmaier, Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs) assesses the suitability of the approach through comparison with variations that alternatively omit elements of it, and/or include an additional dynamical downscaling step. Chapter IV (submitted to *Journal of Climate*) describes the implementation of the forecasting approach over the western U.S. and, via retrospective analysis, estimates the skill of the forecasts and the degree to which the skill arises from persistence of initial conditions or from climate model forecast accuracy. Taken as a whole, the research is intended to offer insights into the potential for climate and hydrologic modeling, with simple downscaling techniques, and to form the basis for advanced seasonal hydrologic forecasting throughout the western U.S.

II. AN ENSEMBLE-BASED HYDROLOGIC FORECASTING APPROACH

This chapter has been published in its current form in the Journal of Geophysical Research: Wood, A.W., Maurer, E.P., Kumar, A. and D.P. Lettenmaier, 2002. Long range experimental hydrologic forecasting for the eastern U.S., *Journal of Geophysical Research*, Vol. 107, D20, October.

1. INTRODUCTION

Great strides have been made over the last decade in understanding of climate teleconnections, as manifested by ocean-atmosphere phenomena resulting in large part from thermal inertia of the oceans (see, e.g. Livezey, *et al*., 1997; Shukla, 1998; Koster *et al*., 1999), such as El Niño-Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO) and the North Atlantic Oscillation. Exploitation of understanding of these phenomena has resulted in demonstrable improvements in long-lead (months to years) climate forecasting. These forecasts are based on coupled ocean-atmosphere general circulation models (OAGCMs) (Barnston *et al*., 1999; Livezey, 1990; Kumar *et al*., 1996; Livezey *et al*., 1996) or statistical methods such as canonical correlation analysis (e.g., Barnston *et al*., 1996). A recent trend has been to use ensemble forecasting approaches in which a global land-atmosphere-ocean model (initialized with atmospheric, land surface and ocean conditions at forecast time) is run into the future for forecast horizons of months to years, using prescribed sea surface temperatures (SSTs) derived using one of a variety of forecast methods. Although the atmosphere is essentially chaotic, the prescribed SSTs effectively constrain the evolution of model forecasts. By perturbing the initial atmospheric conditions and repeating the simulation a number of times, an ensemble of forecasts is constructed which represents the range of global atmospheric conditions that may occur over the forecast period.

Much of the research in this area has focused on atmospheric simulation outputs having the large sub-continental region (e.g., the southwestern U.S.) as the minimum scale, in part because of the long-recognized difficulty of atmospheric models in reproducing observed climate at smaller

scales (e.g., less than 10^7 km^2) (IPCC, 1996, Section 6). As a result, connecting climate forecasts to the scales and features of the hydrosphere in which human society is often most interested -- regional and smaller, for land surface variables – has been problematic. In addition to the scale problem, the land surface variables of greatest interest to society, such as surface precipitation and runoff, are generally predicted less reliably than features of large scale circulation (see, e.g., Risbey and Stone, 1996). In parallel with global scale predictions, however, a number of downscaling methods have evolved including dynamical approaches that use finer resolution (mesoscale) atmospheric models (e.g., Cocke and LaRow, 2000; Giorgi and Mearns, 1991), statistical approaches (Wilby and Wigley, 1997; Wilby *et al*., 1998) and climate-analog approaches (IPCC, 1996; Leung *et al*., 1999; Georgakakos *et al.,* 1998). Recent comparisons of dynamical and statistical methods are given in Murphy (1999) and Kidson and Thompson (1998).

At much smaller scales, hydrologists have long been concerned with understanding and reproducing the dynamics of the land surface water and energy balance. Hydrologic study has mostly focused on the local scale of catchments or basins (on the order of 10^2-10^3 km^2) at which water management is effected. Much applied hydrologic prediction work, for instance efforts to abstract the physics of runoff generation and groundwater behavior, has been intended to benefit water resources end uses, such as irrigation, water supply, hydropower generation, fisheries management and navigation. An intersection of the interests of hydrologists and climate modelers has occurred over the last decade as the difference in spatial scales has narrowed. While general circulation models now often operate at spatial resolutions of one to three degrees, macroscale hydrologic models (e.g., those of Liang *et al*., 1994; Leavesley and Stannard, 1995; Beven and Kirkby, 1979) have increased in scale and geographical coverage so that modeling of continental scale river basins (e.g., the Columbia, the Mississippi) is now possible. Furthermore, the land surface parameterizations in coupled land-atmosphere-ocean models increasingly resemble or borrow from macroscale hydrology model representations, and vice versa (Koster *et al*., 2000; Ducharne *et al*., 2000). In consequence, an operational linkage of hydrologic and climate forecasting models is now being pursued at a number of research centers.

Although the motivation for development of macroscale hydrologic models has been, in part, to improve representation of the land surface in coupled land-atmosphere-ocean models, they can also be implemented using one-way forcing from OAGCMs, such as ensemble climate forecasts.

While conceptually simpler than operation in a fully coupled mode, the one-way linkage is still hampered by the need to address regional biases in OAGCM climate simulation outputs. These biases are substantial enough to preclude direct use in hydrologic modeling of OAGCM output fields such as surface precipitation and temperature (Leung *et al*., 1999; Chen *et al*., 1996; Roads *et al*., 1999).

Hydrologists have tended to surmount this difficulty using simple strategies such as OAGCM-conditioned compositing, in which OAGCM output is used to guide the construction or weighting of an ensemble of historically observed meteorological time series, which subsequently is used as input to a hydrologic model (e.g., Georgakakos *et al.,* 1998; Leung *et al*., 1999). In theory, at least, the probabilistic assessment of differences between streamflow ensembles resulting from these hydrologic simulation and streamflow ensembles based on observed climatology (inputs without conditioning) may then support recommendations for operational decisions by water resources system managers. A variation on this approach is to derive the conditioning signal from a typecast of the present year related to SST-classified climate modes. Hamlet and Lettenmaier (1999a, b), for instance, demonstrated a simplified method of long range forecasting (up to a one year lead) for the Columbia River basin. The method utilized resampling of previous observed hydrometeorological data for years with apparently analogous climatic characteristics, determined by ENSO and PDO-based compositing. The shortcoming of this method is that it requires partitioning of the historic record into climate categories, which can result in statistical problems when the number of years in a given class is small. Furthermore, there is an implicit assumption that the classification method is stable over time.

One way climate model-hydrology model linkages were explored by Kim *et al*. (2000), who applied a mesoscale regional climate model over northern California for downscaling one member of an OAGCM-based 3-month forecast ensemble during the 1997/98 El Nino event. Using the spatially distributed TOPMODEL (Beven and Kirkby, 1979) for hydrologic simulation, they found that the largest hydrologic forecast errors resulted from GCM errors in precipitation prediction. For a smaller catchment in Colorado, Wilby *et al*. (2000) compared statistical and dynamical downscaling methods, using a regional climate model, for translating National Centers for Environmental Prediction (NCEP) / National Center for Atmospheric Research (NCAR) reanalysis (Kalnay *et al*., 1996) output into local precipitation and temperature forcing time series

for a hydrologic model. Their results underscored the need for bias correcting climate model outputs and confirmed the view that while statistical and dynamical approaches yield similar downscaling skill, statistical techniques are less computationally demanding.

This paper describes an exploratory hydrologic forecast system that uses monthly ensemble climate forecasts of monthly total precipitation (Ptot) and monthly average temperature (Tavg) for six-month lead times produced by the NCEP Global Spectral Model (GSM), an OAGCM. We test a relatively simple approach for linking global ensemble forecasts from coupled ocean-land atmosphere models with macroscale hydrologic models, with the intent of improving hydrologic prediction capabilities for soil moisture, runoff and streamflow.

The region chosen for the study was the eastern U.S, defined as the area east of the Mississippi River drainage plus the Ohio River basin, but excluding the Laurentian Great Lakes basin, for a forecast period from May to October, 2000. This period was selected because a severe drought was anticipated for the southeastern U.S. as a result of much below normal soil moisture and streamflow during late winter and early spring, 2000 (reflected as early as December 1999 in federal agency outlooks such as the National Drought Mitigation Center's weekly Drought Monitor and the Climate Prediction Center's U.S. Drought Monitor). During this period, SST anomalies in the tropical Pacific were returning to near-normal from a prior ENSO cold phase (La Nina) episode. Lingering effects of the cold phase, which in the southeastern U.S. have been correlated with dryness, had the potential to compound the existing soil moisture deficits. We also evaluated the method for a study period (beginning in November 1997) during which SST anomalies reflected a strong El Nino event.

2. APPROACH

Our forecast approach uses GSM's surface forecast fields (Ptot and Tavg) to create daily forcing ensembles for the Variable Infiltration Capacity (VIC) macroscale physical hydrology model. Hydrologic model forecasts are produced by first initializing VIC model states with a spin-up period based on observed meteorology prior to the start of forecast, and then driving the model with ensemble forecast meteorology through the end of the forecast period. This basic framework is illustrated in Figure 2.1.

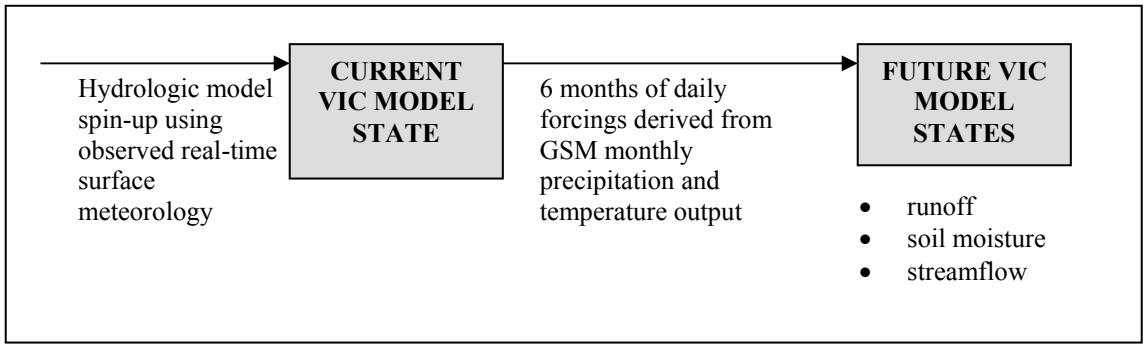

Figure 2.1 Experimental long-lead hydrologic forecasting approach

2.1 Global Spectral Model Ensemble Generation

Each month, NCEP's Climate Modeling Branch generates a 20-member ensemble of six month lead climate forecasts, simulated with GSM. The forecasts are accompanied by a 210-member ensemble of climate hindcasts (also six months long, matching the calendar months of the forecasts) representing the period 1979-99 (21 years). The 20 forecast ensemble members are produced by using 20 different atmospheric initializations with predicted SSTs in the tropical Pacific Ocean as of the date of the forecasts. The hindcast ensemble generation process is similar, except that the ensemble members are produced by using 10 different atmospheric initializations with observed SSTs for each of the 21 years in the 1979-99 hindcast. The different atmospheric initializations are in each case drawn from a sequence of atmospheric analysis fields at the beginning of the forecast initialization month (the month prior to the six month forecast period), spaced 12 hours apart. This process is repeated every month, and Ptot and Tavg, among other variables, are archived. In the forecast runs, predicted SSTs over the tropical Pacific domain are specified based on the NCEP OAGCM (Ji *et al*., 1998). At the time of our research, GSM forecasts were run at T42 horizontal resolution (2.8125 degrees latitude/longitude).

GSM is actually run at a time steps on the order of an hour or less, and therefore in principal the temporal disaggregation is not necessary. Use of monthly ensembles, however, greatly reduces the data distribution and handling overhead. Because our downscaling approach (Section 2.3.2) imposes plausible daily temporal structure to the monthly GSM forecast products as part of the same process that spatially disaggregates the GSM products to 1/8 degree spatial resolution, the

use of the monthly GSM output is not only adequate for purposes of our hydrologic forecast objectives, but also streamlines the process considerably.

2.2 VIC Macroscale Hydrology Model

The VIC model (Liang *et al*., 1994; 1996; 1999) is a semi-distributed grid-based hydrological model which parameterizes the dominant hydrometeorological processes taking place at the land surface-atmosphere interface. A mosaic representation of land surface cover and parameterizations for infiltration and the spatial variability of precipitation account for sub-grid scale heterogeneities in key hydrological processes. The model uses three soil layers and one vegetation layer with energy and moisture fluxes exchanged between the layers. The model has been applied to such large continental rivers as the Columbia (Nijssen *et al*., 1997), the Arkansas-Red (Abdulla *et al*., 1996), and the Mississippi (Maurer *et al*. 1999; Cherkauer and Lettenmaier, 1999), and, as part of the Land Data Assimilation System (LDAS) project (Mitchell *et al*., 2000), to the continental U.S. (Wood *et al*., 1998). A more complete description of model processes can be found in Liang *et al*. (1994; 1996). Routing of runoff generated within a grid cell is routed to the stream gauge locations using methods described by Lohmann *et al*. (1998a, b). The VIC model uses vegetation and soil parameters produced for use by LDAS and described in Maurer *et al*. (2001).

2.2.1 Forcings

VIC model forcings are used both in driving the hydrologic model during a one year spin-up period, and, via resampling, in assembling the daily forecast sequences. Because the meteorological variables most widely available in long-term data archives are daily precipitation and daily temperature minimum and maximum, we estimate most of the other forcing variables required by the VIC model (e.g., downward solar and longwave radiation, humidity) from this minimum set of variables using methods described in Maurer *et al*. (1999). Wind speed data are taken from the NCEP/NCAR reanalysis (Kalnay *et al*., 1996), which are available to within a month of real time. The observational data are typically taken from National Climatic Data Center (NCDC) Cooperative Observer (Co-op) Stations, which are available on the web within 2 to 4 months of real time. For real time data (bridging the gap between the end of the available CO-OP data and the forecast date), we used data available from the LDAS project. The LDAS

precipitation data are so-called Stage IV observations, a combination of radar and station data produced by NCEP. Temperature data are from the Eta Data Assimilation System (EDAS), and are essentially an analysis product from the NCEP Eta weather forecast model, run over the continental U.S. LDAS also produces a real-time wind data set (another EDAS product). All LDAS data are produced on a geographic 1/8 degree grid, for the area from latitude 25N-53N, longitude 67W-125W. Typical monthly biases in this product revealed by our preliminary verification were a spatially averaged (over the study domain) bias of -5 percent in monthly precipitation totals, -1.5 ^{o}C in average maximum temperature and +1.5 ^{o}C in average minimum temperature. LDAS wind speeds appear to be significantly higher than those produced by the NCEP/NCAR reanalysis, although recent EDAS modifications appear to have reduced the discrepancy somewhat. Currently, we use the LDAS product without adjustment.

2.2.2 Eastern U.S. application

The VIC model was implemented at 1/8 degree latitude/longitude resolution (~150 km^2 cell area) over the domain shown in Figure 2.2. The domain includes the Ohio River basin, which drains the easternmost portion of the Mississippi River basin, and an East Coast region which includes 24 coastal drainage basins, 17 of which flow east-southeast to the Atlantic Ocean, and 7 of which drain southward to the Gulf of Mexico. Within the model domain, runoff in smaller subbasins was routed to produce streamflow estimates at U.S. Geological Survey (USGS) river gauging station locations.

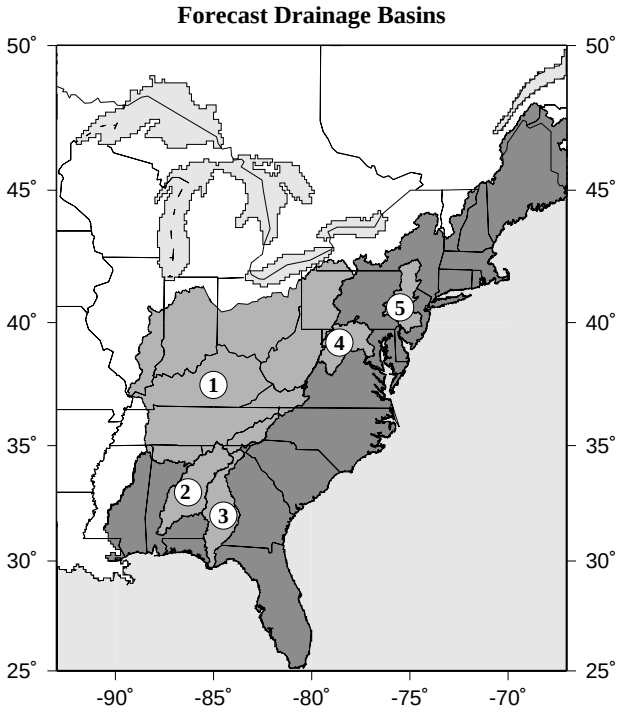

Figure 2.2 Hydrologic forecasting model domain, including the Ohio River basin (light gray) and the East Coast drainages (dark gray). Runoff in numbered basins was routed to produce streamflow: (1) Ohio; (2) Alabama-Coosa-Tallapoosa (ACT); (3) Apalachicola-Chattahoochee-Flint (ACF); (4) Potomac; and (5) Delaware River.

The Ohio River basin and East Coast models simulate areas of about 600,000 and 1.1 million km^2, respectively. Subbasins calibrated for streamflow forecasting included the Ohio River, the Delaware River, the Potomac River, the Apalachicola-Chattahoochee-Flint (ACF) River system and the Alabama-Coosa-Tallapoosa (ACT) River system. Principal streamflow routing nodes were the Apalachicola River at Sumatra, FL (USGS station 12359170), the Alabama River at Claiborne L&D near Monroeville, AL (USGS station 02428400), the Potomac River near Washington, D.C. - Little Falls (USGS station 01646500) and the Delaware River at Trenton, NJ (USGS station 01463500), and the Ohio River at Metropolis, IL (USGS station 03611500).

The model forcing data spanned the period 1950-current, where the current date evolved during the experiment. Model calibration was accomplished by varying parameters related to infiltration and subsurface drainage, with the aim of reproducing monthly streamflow volumes while

preserving the general features of the daily response (e.g., daily average flow peaks and recessions). Sample calibration results for each basin are shown in Figure 2.3.

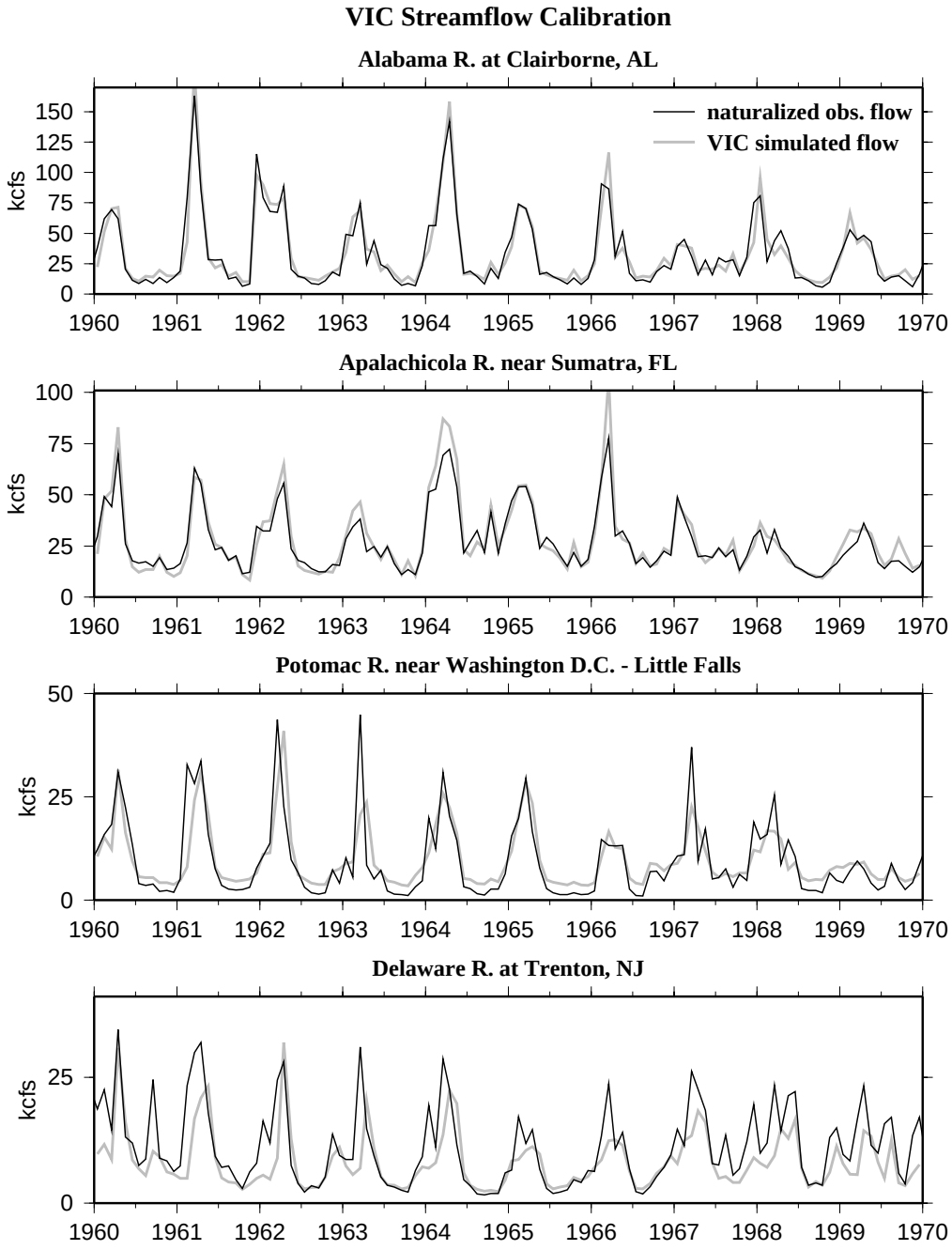

Figure 2.3 Calibration results for five streamflow forecasting basin gauging locations

2.3 Hydrologic Forecasting Approach

To translate long range climate predictions to the realm of hydrology, regional biases and the temporal and spatial scale mismatch between models must be resolved. This section describes our bias-correction and downscaling approach and its demonstration as a proof of concept exercise.

2.3.1 Unbiasing of climate model ensembles

The premise of the bias correction step is that despite biases in GSM-simulated climate, the GSM forecasts may have a useful signal if interpreted relative to the GSM climatology rather than the observed climatology. The GSM climatology is defined by the monthly distributions (for months one to six in the forecast period, separately) of simulated GSM Ptot and Tavg taken from the GSM hindcast simulations (i.e., the 210 simulated values for each of the six forecast period months, for each variable). The monthly observed climatology spans the same time period as the GSM output (1979-99) and was created from CO-OP station daily observations averaged to a monthly timestep, and to the GSM grid resolution; hence the observed monthly distributions for Ptot and Tavg are defined by only 21 values per variable. Bias correction is achieved by replacing GSM forecast values for Tavg and Ptot with values having the same percentiles (non-exceedence probabilities) with respect to the observed climatology that the original GSM values had with respect to the GSM climatology, for a given month. The forecasts are subsequently expressed as anomalies (temperature shift, and precipitation percentage) with respect to the observed monthly means for the 21-year climatology period. Bias correction is performed at the GSM scale, and each GSM cell (23 cells spanned the study region) is treated individually, defining its own set of monthly distributions.

For example, bias-correcting a monthly Tavg forecast for January – June requires the following steps:

1. The January GSM Tavg is assigned a non-exceedence probability (or percentile) within the 210-value GSM climatology distribution for January Tavg.

2. A January Tavg having the same non-exceedence probability in the observed climatology is then calculated.

3. Steps 1 and 2 are repeated for Tavg in months February through June, and the entire process is repeated for each of the ensemble forecast members.

4. Finally, the bias-corrected forecasts are expressed as additive (for Tavg) and multiplicative (for Ptot) anomalies.

In the precipitation and temperature bias correction scheme, when either the GSM output or the associated percentile falls above or below the range of empirical Weibull percentiles (equal to 1/N+1 and N/N+1, where N is the number of members from which the probability distribution is estimated), theoretical probability distributions are fit to the data to extend the empirical distributions. This becomes necessary because the historical climatology is defined by the 21 years of historical observations, whereas the model ensembles consist of a larger 210-member data set. For low precipitation, an Extreme Value Type III (Weibull) function was used, with a minimum lower bound of zero; whereas for extreme high precipitation, an Extreme Value Type I (Gumbel) distribution was employed. For temperature, a normal distribution was used for both minimum and maximum.

The need for bias correction is demonstrated by an example that compares GSM's precipitation and temperature climatology for one typical cell (centered on latitude 37.97N, longitude 87.19W, in the Ohio River basin) with observed values (Figure 2.4). As alluded to in the introduction, biases of the magnitude shown (e.g., up to 15 degrees Celsius for Tavg in August) are occasionally found in climate model simulations of surface variables, particularly if one examines the output for individual cells rather than for a region or continent. The lack of agreement with observations stems in part from poor climate model resolution of sub-grid scale land surface related heterogeneities, such as orography or soil wetness. Such biases preclude the use of the climate model output as a direct input to the hydrology model.

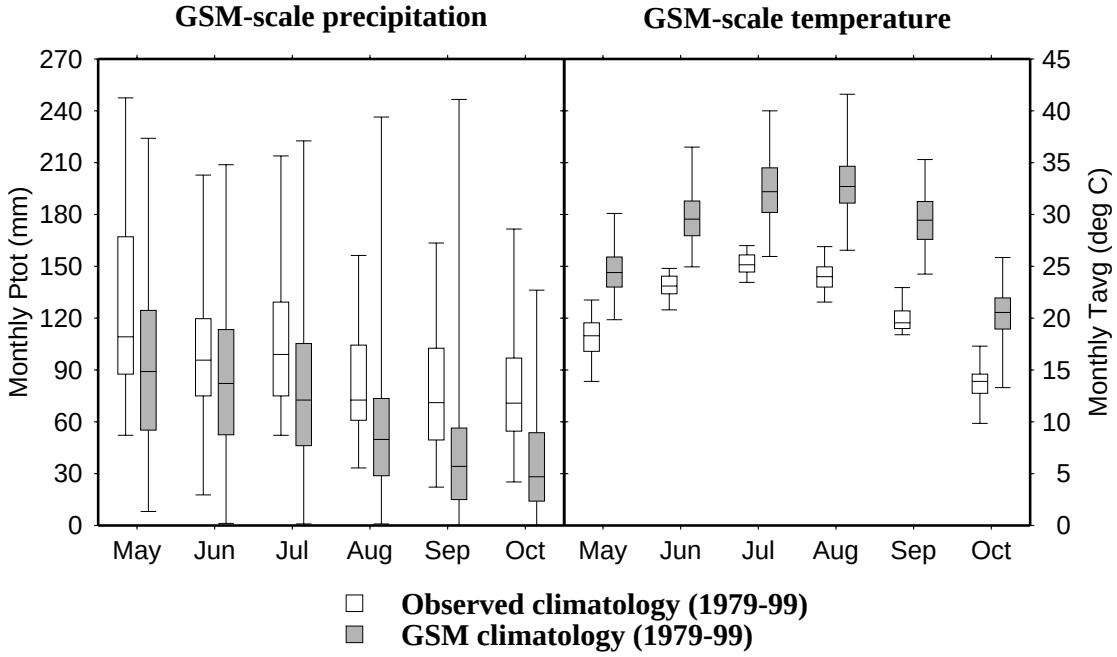

Figure 2.4 April 2000 GSM climatology for monthly total precipitation and average temperature, compared with observations averaged over the corresponding geographic area. The data are for the GSM computational cell centered on latitude 37.97N, longitude 87.19W, in the Ohio River basin.

2.3.2 Downscaling of long range ensemble forecasts

Following bias correction, the monthly GSM-scale forecast anomalies are translated to the spatial and temporal scale of VIC model inputs. The Tavg and Ptot anomalies are spatially interpolated to the 1/8 degree VIC cell centers, and applied to the monthly observed 1979-99 1/8 degree cell means (derived from Co-op station observations as described in Section 2.2.1), to create monthly forecast sequences at the VIC model scale, in the following manner:

$$T_{VICfcst}(m, e) = T_{VICmean}(m) + T_{ANOMfcst}(m, e)$$

$$P_{VICfcst}(m, e) = P_{VICmean}(m) * P_{ANOMfcst}(m, e)$$

Here $T_{VICfcst}(m, e)$ is the forecast monthly Tavg for a given VIC cell in month m (m = 1 to 6) of a forecast ensemble member e (e = 1 to 20). $T_{VICmean}(m)$ is the observed 1979-99 mean Tavg for

month m, and $T_{ANOMfcst}$(m, e) is the additive Tavg forecast anomaly for month m and ensemble member e. Likewise, $P_{VICfcst}$(m, e) is the forecast monthly Ptot for a given VIC cell in month m of a forecast ensemble member e, $P_{VICmean}$(m) is the observed 1979-99 mean Ptot for month m, and $P_{ANOMfcst}$(m, e) is the multiplicative Ptot forecast anomaly for month m and ensemble member e. The addition of temperature anomalies will hereafter be referred to as shifting and the multiplication by precipitation anomalies as scaling.

The final step in preparing the forecasts for input to the VIC model is to replace the monthly mean sequences by daily sequences. For each month (e.g., January) in each forecast ensemble, one year from the climatology period is randomly selected (e.g., 1988). For each VIC cell, the observed daily values of precipitation for the selected year and month (e.g., 1988, January) are scaled so that the monthly total precipitation is equal to the forecast Ptot for the ensemble member and month. The resulting values of daily precipitation become the daily sequence for that month of the particular forecast ensemble member. Daily Tmin and Tmax from the same selected year (e.g., 1988) are shifted equally so that their average, (Tmin + Tmax)/2, reproduces the monthly forecast Tavg for the ensemble member and month; and the resulting values of Tmin and Tmax become the daily sequence for that month of the particular forecast ensemble member. Daily wind speed is taken without adjustment from the VIC daily values for the selected year and month, forming the fourth daily forcing used by the VIC model. The same year is used to select the daily data for a given month of an ensemble forecast member in every cell of a study area (the Ohio River basin and East Coast). Using the same year-month combination for resampling over the large scale hydrologic units helps to preserve a degree of spatial synchronization in the weather components driving hydrologic response. The random sampling of a climatology year for selection of daily sequences is repeated for each month in each forecast ensemble member.

We performed a test of this method using observed total monthly precipitation and average temperature timeseries for 1979-99, aggregated to the GSM scale, as raw forcings over the Ohio River drainage area. These large scale forcings were processed (using the interpolation and temporal disaggregation steps) into daily VIC scale forcings, with which we simulated streamflow. Figure 2.5 shows that the method is able to reproduce the mean and variance of the basin streamflow climatology without introducing substantial method-related bias.

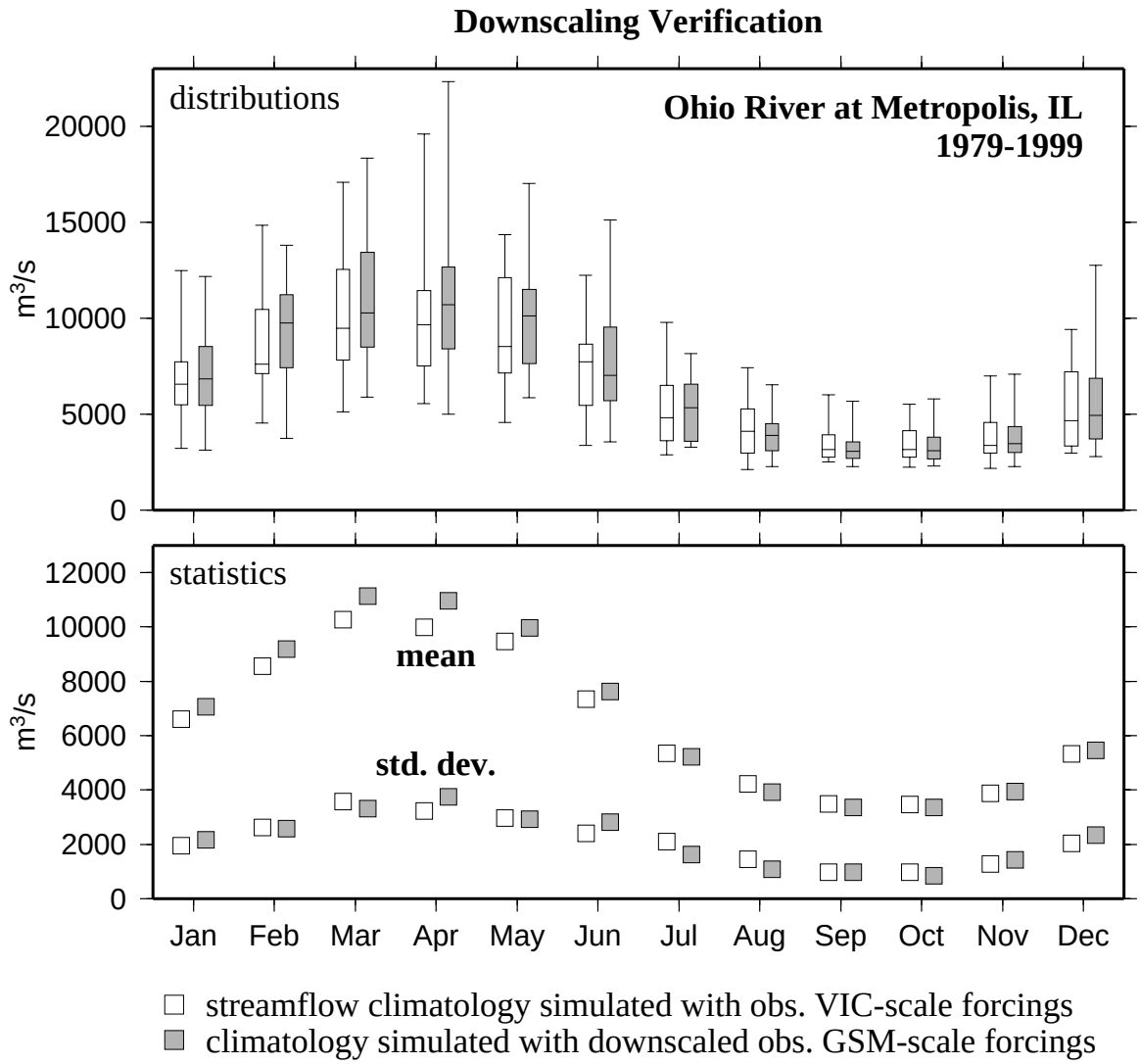

Figure 2.5 Climatology period (1979-99) streamflow distribution simulated from daily VIC 1/8 degree observations, compared to a parallel simulation from monthly GSM-scale (2.8125 degree) spatially averaged observations, after the downscaling and disaggregation procedure

2.3.3 Producing hydrologic forecasts

Near the 10th day of each month, the GSM ensemble forecasts become available and we process the monthly GSM output into format suitable for input to VIC. By the 20th day of the month during the period April - September 2000, the hydrology model state was initialized through that current date using a one year spin-up simulation, the forcings for which were the gridded observational data described in Section 2.2.1. Once the current hydrology model moisture states were obtained, the 20 forecast ensemble members were run to produce an ensemble of 6-month long hydrologic forecasts, beginning the following month.

In addition to the forecast ensemble, we also generated a hydrologic ensemble hindcast by applying the procedures described in Sections 2.3.1-2.3.2 to the ensemble members of the GSM hindcast in place of the forecast. This hydrologic ensemble hindcast yields a hydrologic climatology for 1979-99 derived by the methods used to derive the hydrologic ensemble forecast. Rather than comparing the forecast ensemble results directly to the empirial probability distribution of observed streamflow or of model-simulated fields (e.g, grid cell runoff or soil moisture) based on observations, we compared the hydrologic ensemble forecasts with the hydrologic ensemble hindcast. In this experiment, we wanted to ensure that any incidental forecast error associated with the approach would also arise in the climatology distributions to which the forecasts were compared. Upon completion of each forecast or hindcast run, monthly total precipitation, evaporation and runoff (surface plus baseflow), and monthly average soil moisture and temperature were archived, and the daily streamflow routing was performed for the selected subbasins (shown in Figure 2.2).

2.3.4 Retrospective ENSO-event forecast simulation

As an additional test of the method, we performed a retrospective comparison of the 10-member hindcast ensemble associated with the November 1997 SSTs (just prior to the strong 1997-1998 El Nino), which was extracted from the 210 member climatology ensemble for November. The purpose of this analysis was to evaluate forecast performance made from a month exhibiting strong SST anomalies in the tropical Pacific Ocean -- that is, conditions favorable for skillful climate forecasting. The NINO3 index, which measures the deviation from normal of the sea surface temperature in the Eastern Pacific, and which is high during an El Nino event, reached its

highest value in decades in winter, 1997. We treated the 10-member ensemble for 1997 as a surrogate for an actual forecast made at that time, even though the SSTs used in the hindcast from which it was drawn were prescribed according to observations, rather than projected. In that particular month, however, forecast skill for tropical Pacific SSTs was relatively high, thus the prescribed SSTs and forecast SSTs would not have differed as greatly as in other periods. The hydrologic forecast ensemble based on the 10 member El Nino ensemble members was compared with ensembles yielded from use of the entire 210 member November hindcast.

3. RESULTS

We evaluated the results of the experiment using two types of output: a) spatially distributed variables such as surface forcings, hydrologic model runoff and soil moisture; and b) streamflow at selected locations. These outputs were generated for the six monthly forecast dates beginning in April, 2000. We report here a representative sample of the forecast results for three starting dates – April 20, June 20 and August 20 – and observed conditions for May, July and September. For spatial forecast results, the forecast ensemble medians are plotted as a percentile of GSM climatology ensembles derived from GSM hindcasts. These percentiles are verified against the observation-based, retrospective forecast fields, shown as percentiles of the observation-based, retrospective climatology. For streamflow forecast results, GSM forecast and hindcast (climatology) distributions (discussed in Section 2.3.3.) are shown in addition to observations that have become available since the forecasts were made.

3.1 Spatial Analyses

Figure 2.6 shows observation-based gridded precipitation and temperature fields and corresponding simulated soil moisture and runoff. Data and model deficiencies notwithstanding, these are treated as surrogate observations for the summer 2000 period. The broad features of the results for simulated soil moisture and runoff are consistent with the signals in precipitation and temperature, modulated by the simulated antecedent soil moisture conditions (characterized by general deficits in the southeast throughout the study period). Soil moisture and runoff percentiles were quite similar, which reflects the VIC model tendency for precipitation inputs to elevate runoff and baseflow in concert with soil moisture, especially considered from a monthly

standpoint. Against these, we contrast Figures 2.7a-b, which show GSM-based forecasts of the same fields.

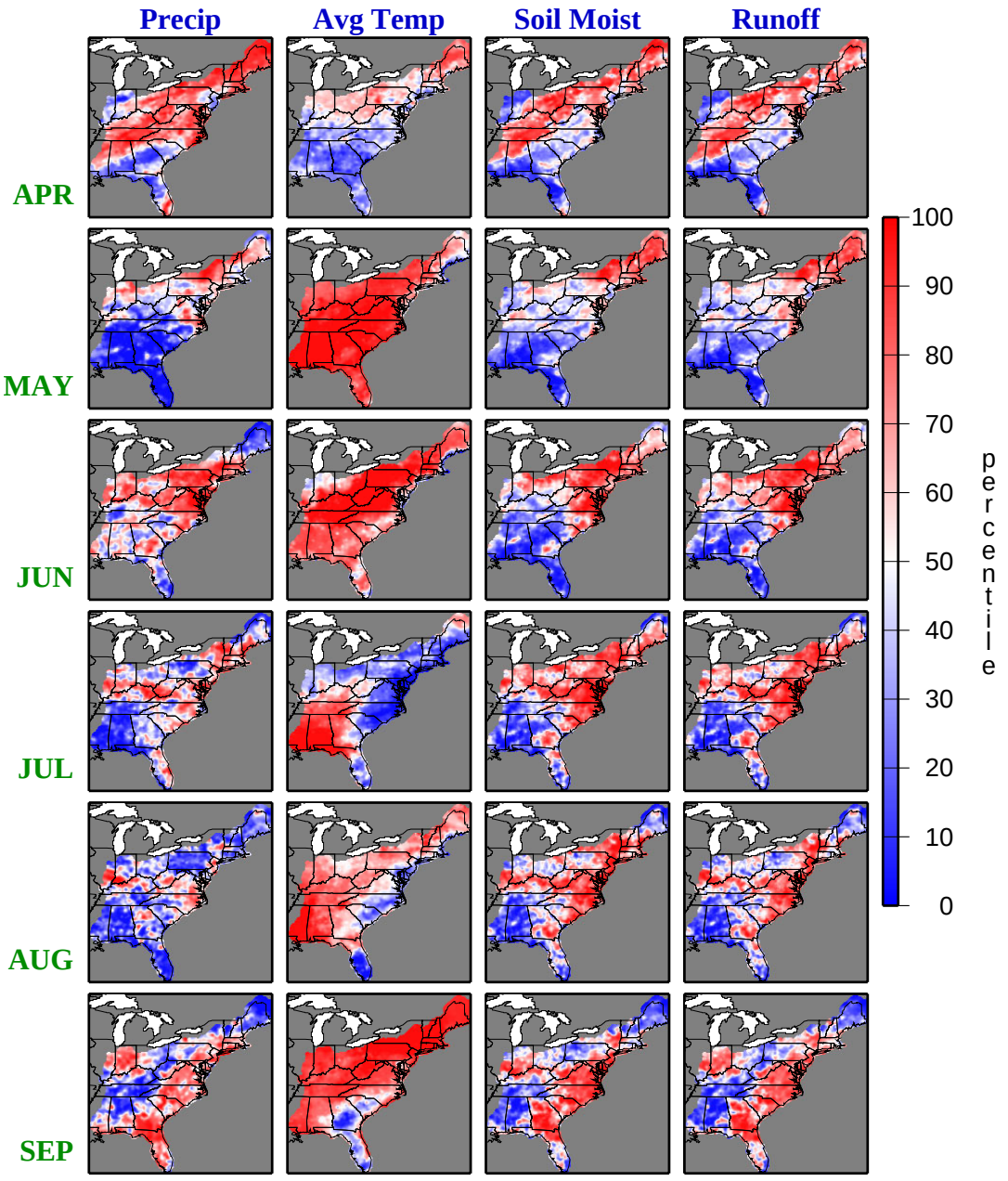

Figure 2.6 Observed climatology for April through September, 2000, defined as the monthly gridded observations of total precipitation and average temperature, and associated simulated analyses of average soil moisture and total runoff. These are shown as percentiles of the variables' observed and simulated (21 year) climatological distributions, respectively.

In the observational analyses, extremely low May precipitation coupled with high temperatures deepened drought conditions throughout the Ohio River valley and southeast, while the northeastern U.S. experienced slightly higher than average precipitation (with respect to the 1979-99 climatology) (Figure 2.6). By July, temperatures along the East Coast and the northern Ohio River valley were cooler, and precipitation had risen to above normal in many locations, while the relative dryness and heat persisted from the Gulf Coast to the southern Ohio River valley. September brought high temperatures everywhere except Florida and Georgia, and the region of low precipitation shifted north, while the drought eased along the Gulf Coast. In response to these forcings, anomalously low soil moisture and runoff, which had been general over the entire domain south of New England, recovered gradually along the East Coast and the northern Ohio River valley. The center of the drought-stricken region, which initially included Florida, shifted west toward Alabama and Louisiana.

The April ensemble forecasts (spanning the period May-October) showed above normal precipitation in the southeastern U.S. in May, and on the East Coast (excepting Florida) in July, but then slightly below normal precipitation everywhere except Florida, in September (Figure 2.7a). The median forecast was for temperatures slightly above normal everywhere except Florida in May, then more strongly above normal west of the Atlantic states in July and September. Normal to cool conditions were forecast in the Atlantic states during the study period. Consequently, initially dry soil moisture and below normal runoff were predicted to recover (at least in the median forecast) in Florida and the mid-Atlantic states by July, but linger in parts of Alabama and Louisiana. In southern New England and the Ohio River valley, however, the median forecast called for a continuation of dryness and heat, hence low soil moisture and runoff.

The June forecasts (Figure 2.7b, top, shows July and September) anticipated, in the median, above normal precipitation in the southeastern U.S. in July, and on the East Coast and Ohio River valley (except in New England) in September, and cooler temperatures everywhere in both months. Even so, the forecasts indicated that the dry soil moisture and below normal runoff would fail to recover fully in the southern and Gulf states, while in southern New England and the Ohio River valley, they would to transition to above normal levels.

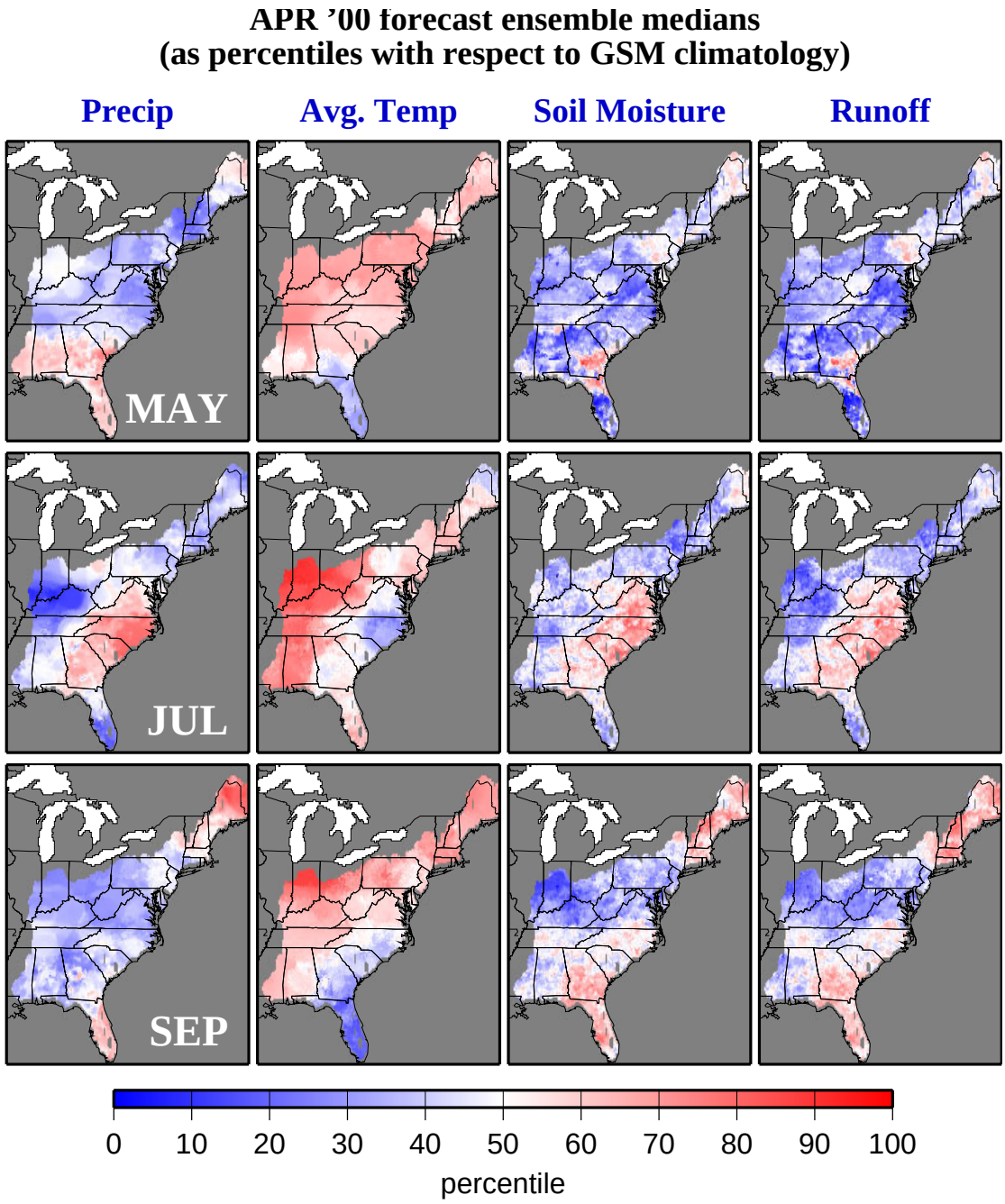

Figure 2.7a April 2000 GSM forecast ensemble medians for May, July and September monthly total precipitation and average temperature, and GSM forecast-based (VIC simulated) ensemble medians of average soil moisture and total runoff, shown as percentiles of the 21-year GSM hindcast climatology distribution for each respective variable

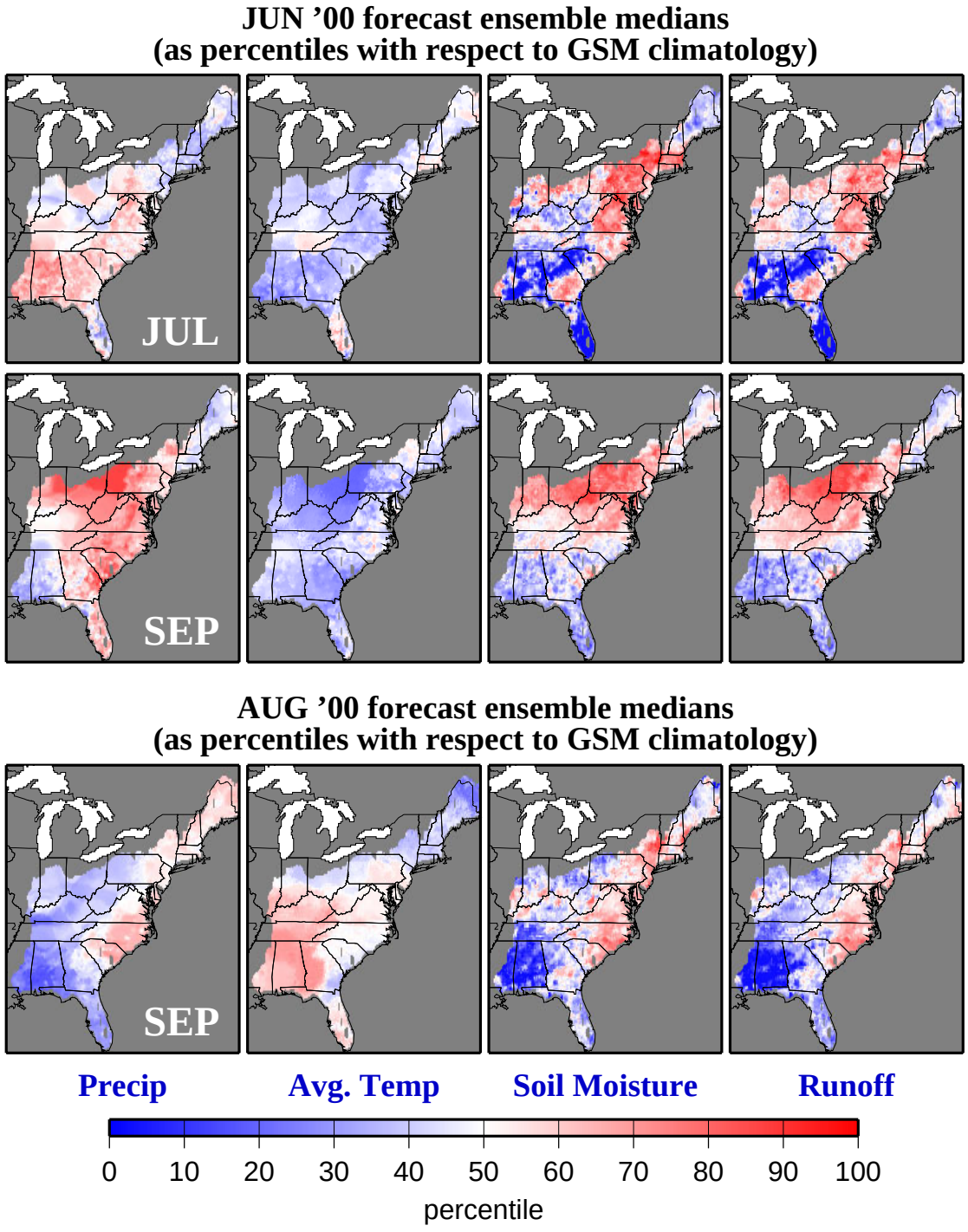

Figure 2.7b (top) June and (bottom) August 2000 GSM forecast ensemble medians for July and September monthly total precipitation and average temperature, and GSM forecast-based (VIC simulated) ensemble medians of average soil moisture and total runoff, shown as percentiles of the 21-year GSM hindcast climatology distribution for each respective variable

The median August forecasts (Figure 2.7b, bottom, for September) called for dry and hot conditions in Florida, the Gulf states and the Ohio River valley, with above normal precipitation and cooler to normal temperatures from the mid-Atlantic states to New England. These circumstances would serve to aggravate the low August soil moisture centered on Alabama.

3.2 Streamflow Evaluation

Predicted streamflows reflected the condition of soil moisture and runoff, although the deviation from normal was generally small relative to the variability exhibited by both the climatological and forecast ensemble distributions. Figures 2.8 and 2.9 show the location of the streamflow sites (upper right) and the forecast to climatology ensemble comparisons for the same three forecast start dates detailed in Figure 2.7.

Figure 2.8, for the Potomac River near Washington, D.C. - Little Falls, shows, for the April forecast, forecast streamflow distributions which are similar to the hindcast climatology distributions, but a bit lower in May and September, when it can be seen from Figure 2.7a that the entire watershed is below normal. The observed streamflows in May were lower than either distribution suggested, but the forecast correctly indicates the direction of the anomaly. In the June and August, the forecasts and climatology distributions had similar medians except for July streamflow in the June forecast, when the forecast was erroneously higher, reflecting the above normal precipitation forecast. The forecasts in June and August agreed fairly well with observations, which were not far from normal for the June to September period.

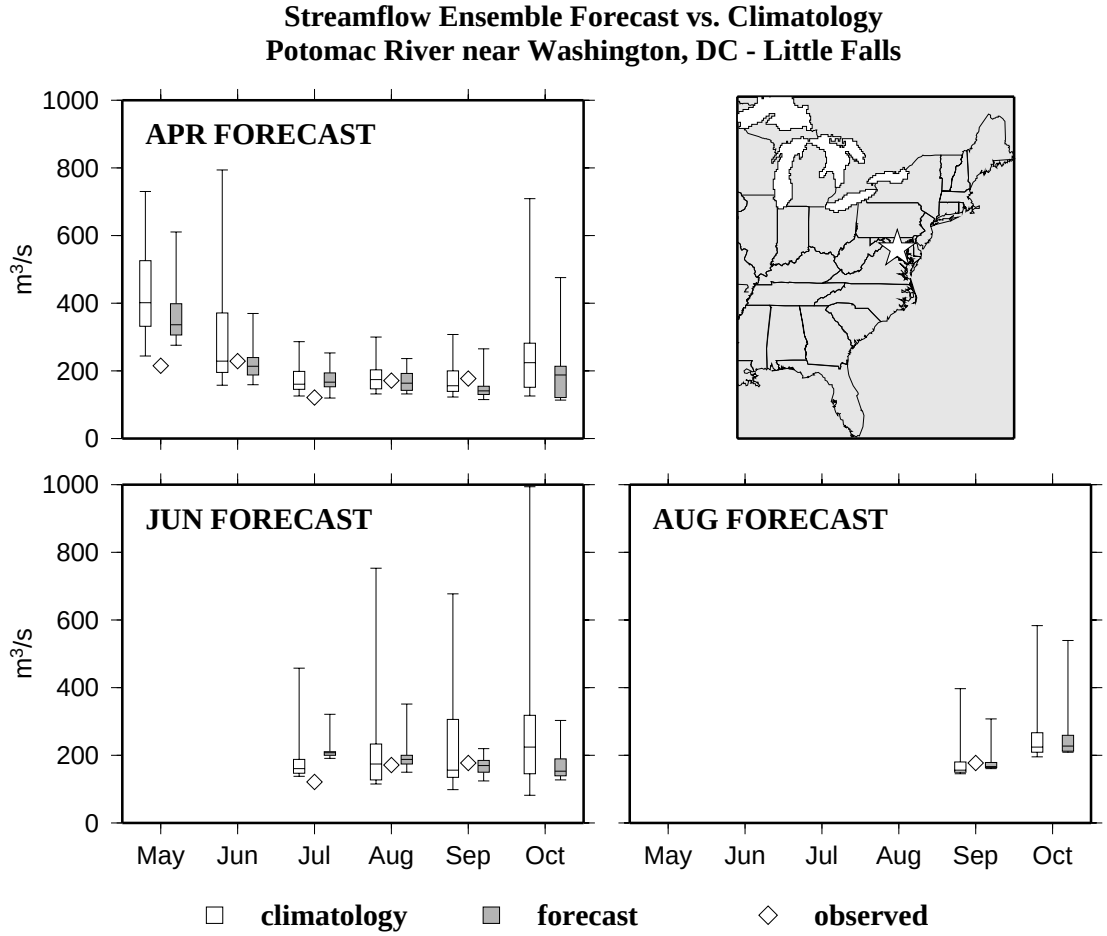

Figure 2.8 April, June and August 2000 monthly average streamflow forecast and hindcast (climatology) ensembles compared with observed values, for the Potomac River near Washington, D.C. - Little Falls (location shown in the upper right)

In Figure 2.9, showing the April forecast for the Alabama River at Claiborne L&D, the forecast ensemble distributions were slightly lower than the climatology ensembles in the first two months of the forecast period, but thereafter were similar. The forecasts gave a slight indication that streamflows would be low early in summer, but no indication of the severity of the streamflow anomaly that was observed. This result was consistent with the above normal simulated antecedent soil moisture for April (Figure 2.6) in half of the watershed, and the forecast of above normal precipitation in July (Figure 2.7a). Subsequent forecasts in June and August, however, were significantly lower than the climatology ensemble distributions, responding to the dry initial conditions for those forecasts (Figure 2.6) and near or below normal precipitation thereafter

(Figure 2.7b). The June and August forecasts distinctly anticipated the severe declines in summer streamflow that were observed.

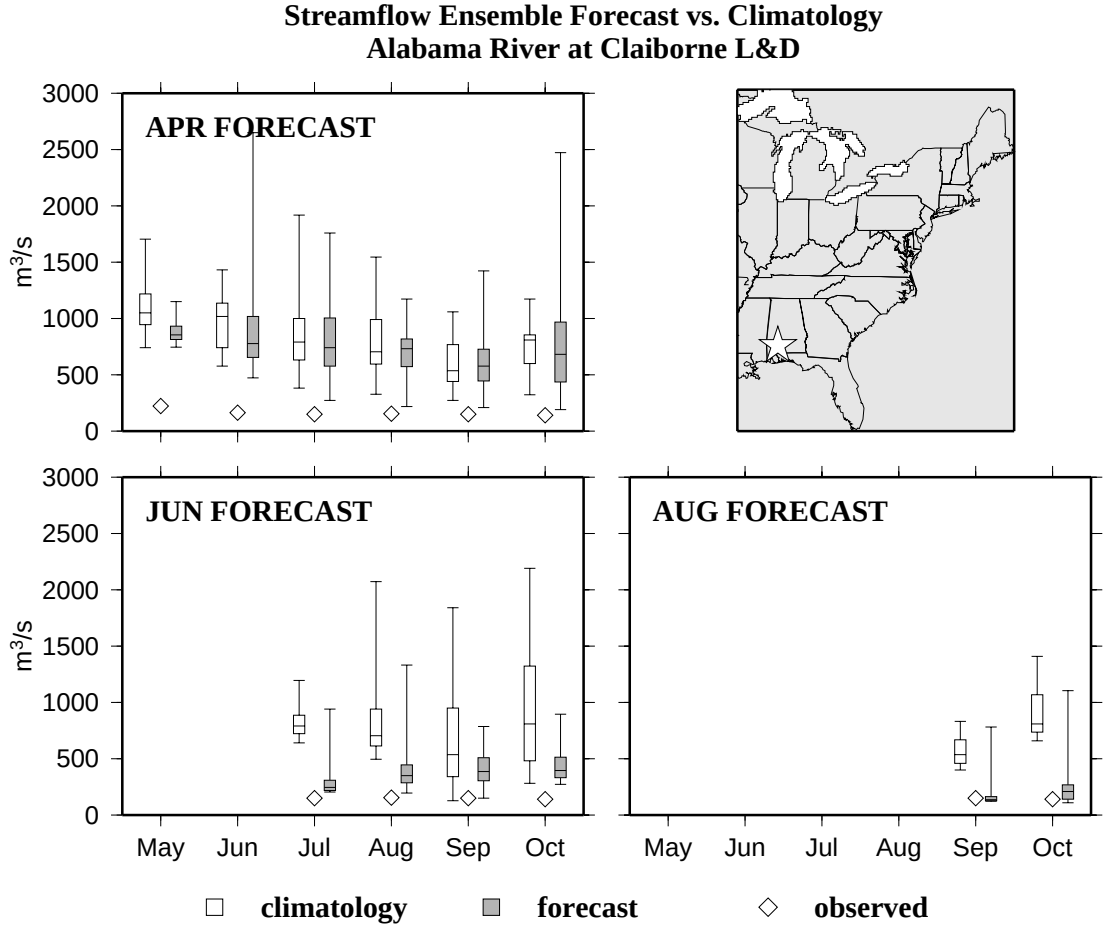

Figure 2.9 April, June and August 2000 monthly average streamflow forecast and hindcast (climatology) ensembles compared with observed values, for the Alabama River at Claiborne L&D, AL (location shown in the upper right)

3.3 1997 El Nino Period Forecast Results

Consistent with the expectation of an anomalously wet winter season in the Southeast associated with the El Nino climate phase (Changnon, 1999; Barnston *et al*., 1999), the gridded observed precipitation and temperature and associated hindcast hydrologic simulations of soil moisture and runoff for November 1997 (not shown here) showed wetter than normal (again, with respect to the 1979-1999 climatological period) precipitation in the southeastern U.S. and the Atlantic

drainages, with drier than average conditions in the Ohio River basin -- a pattern echoed in the relative soil moistures and runoff. At the same time, temperatures throughout the study domain were relatively and uniformly cool. In December, the southeastern states were again very wet while the rest of the domain received about normal precipitation, and the entire region was still relatively cool, but less so than in the previous month. In January through March, the entire domain received above average precipitation, but by April drier weather appeared to move north from Florida, eventually extending to the entire southeast. During this time, with the exception of a cool March in the southeast, temperatures were mostly above normal throughout the domain. The net result for soil moisture and runoff for fall and winter 1997-98 was wetter than average in the southeastern U.S. and along the East Coast, with the largest soil moisture anomaly in the southeastern U.S. in December 1997 to March 1998, and in the northeastern U.S. in March and April, 1998.

The spatial (perfect SST) forecasts made in November 1997 appear to indicate correctly the direction of soil moisture and runoff anomalies – i.e., the median forecast was for wetter than normal conditions. Exceptions were the northeastern U.S. and the Ohio River basin, where the median forecasts were wetter than average as early as December 1997, and drier than average in April and May, 1998, whereas the retrospective analysis showed the opposite condition. Also, the forecast percentiles tended toward the median relative to the hindcast percentiles, perhaps as a result of the finer resolution afforded the probability scale by the use of 210 ensemble members for the climatology compared with the 21 ensemble members used to provide statistical context for the forecasts. A comparison of months 1-3 (December 1997 to February 1998) of the hindcast and the forecast is shown in Figure 2.10.

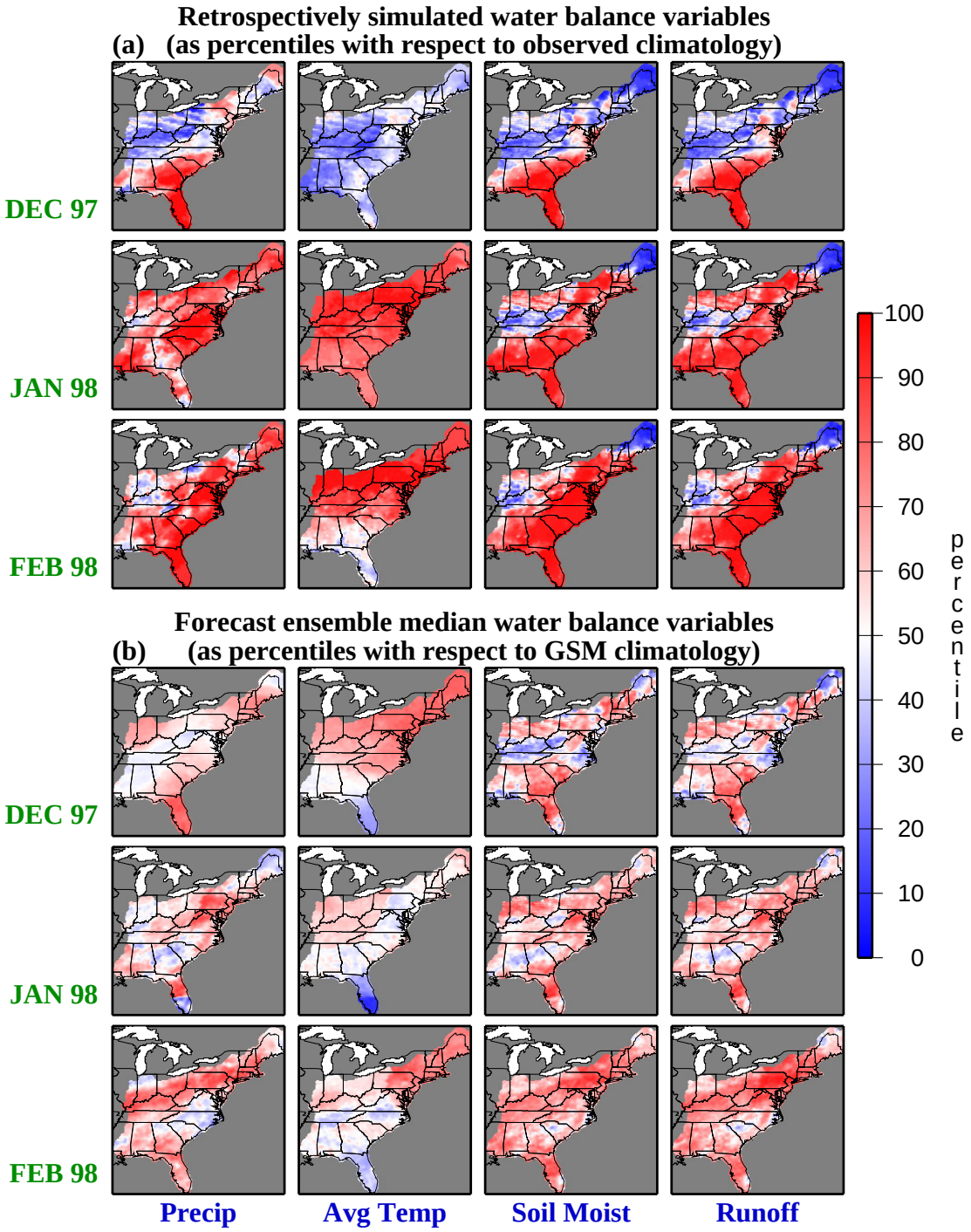

Figure 2.10 (a) Dec. 1997 - Feb. 1998 gridded observed monthly total precipitation and average temperature, and associated analyses of average soil moisture and total runoff, shown as percentiles of the observed climatology; (b) November 1997 GSM-derived forecasts for the same period, shown as percentiles of the GSM-based hindcast climatology

A sample of streamflow results for the El Nino period forecast is given in Figure 2.11, for the Potomac River near Washington, D.C. - Little Falls. The Potomac River watershed's slightly below normal precipitation in December and near normal soil moisture (Figure 2.10) led to a forecast ensemble with a similar median to the climatology ensemble in December. Both ensembles were slightly higher than the observed streamflow. The forecasts' above normal precipitation in January and February produced above normal forecast distributions for streamflow in those months, although the forecasts, as in the Alabama River example shown in Figure 2.9, failed to anticipate the magnitude of the anomaly actually observed, which continued throughout the forecast period.

4. DISCUSSION

We evaluated the spatial forecasts from a qualitative standpoint only, broadly assessing the consistency of anomaly direction in the climate forecasts and resulting hydrology forecasts without undertaking to determine the skill of the climate model and resulting hydrologic forecasts quantitatively (hence the term "skill" is here used loosely to indicate general consistency of forecasts with expected values). We did not focus on climate model forecast skill because our primary purpose was to develop a framework within which ensemble climate forecasts could be used for hydrological purposes. A secondary objective was to determine whether the climate model forecast signal or hydrologic (soil moisture) persistence would dominate in situations where the climate forecast anomalies were significant. We conclude from this exercise that the downscaling procedure successfully transfers the climate forecast signals to the hydrologic variables. It is especially encouraging that the hydrologic model, which performs a non-linear transformation of temperature, precipitation and other inputs to streamflow, was able to retrieve the observed streamflow climatology from the downscaled, GSM-scale observed precipitation and temperature climatology.

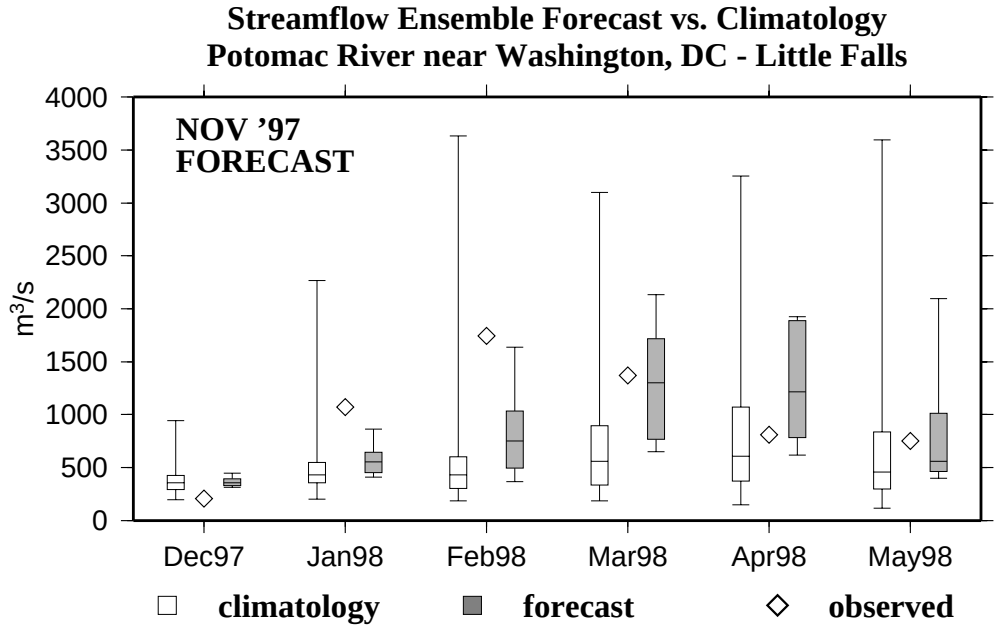

Figure 2.11 El Nino condition streamflow forecast for the Potomac River near Washington, D.C. - Little Falls, using a 10 member GSM hindcast ensemble for November 1997 as a forecast surrogate, compared with the GSM hindcast ensembles for November 1979-99 (a climatology), and with observed streamflows

One feature of this approach that may require further evaluation is the spatial interpolation of bias-corrected GSM-scale output anomalies, rather than their associated probability values, directly to the VIC 1/8 degree resolution grid cell centers. Interpolation of anomaly quantiles (in probability space) to the VIC resolution would yield a joint probability anomaly for the thousands of VIC cells in each basin far exceeding the original GSM anomaly quantile, so the latter approach was chosen as a method of bridging the scale gap between GSM output and VIC input. It also appears that the downscaling procedure may increase the variability of the forecasts somewhat by producing high outliers in the precipitation fields (as a result of the rescaling of sampled daily patterns to match monthly anomalies), and refinements will be pursued in future applications to resolve this problem.

For the summer 2000 study period and region, the climate model forecasts had a mixed performance, as estimated spatially from the precipitation, temperature, runoff and soil moisture ensemble medians and from the streamflow comparisons. For example, in the April and June month 1 median forecasts (Figure 2.7), the southeast is slightly wetter than the climatology median, while our retrospective analysis shows that the region continued to be dry. In each case,

however, the persistence of antecedent low soil moisture from the hydrology model maintained low soil moisture and runoff several months into the forecast period, so that the forecast results still agreed reasonably well with the retrospective analysis for lead times of several months. Here, the weak and at times incorrect direction of the climate model forecasts was balanced by a degree of skill derived from persistence in the hydrologic states.

For the El Nino-condition forecasts of November 1997 (Figure 2.10), in contrast to the summer 2000 results, the hydrologic forecast results appeared to be determined by the climate model forecast signal as well as by persistence in the antecedent hydrologic model state. The initial soil moisture signature, characterized by a decreasing gradient in soil moisture and runoff percentile from south to north (which resulted from the antecedent conditions) was largely erased by the normal to wetter than average precipitation forecasts coupled with normal to cooler than average temperature forecasts (after the December warmth in the northeast). By February (forecast month 3), the wet anomaly over the entire region was consistent with, although weaker than, the anomaly revealed in the hindcast analysis.

One issue that bears mention is the high variance in forecast ensemble and climatology ensemble spatial fields (for a given point), hence in the streamflow ensembles (Figures 2.8, 2.9 and 2.11). Given a high variance, a large shift in the mean of the forecast distribution is required to produce a statistical difference in the forecast outcome, hence the discrimination of the forecast system (in the sense described by Wilks, 1995) is weaker than it would be for a narrower distribution of ensembles. A wide forecast distribution prohibits water managers from making decisions that effectively rule out one end or the other (or, in some cases, both) of the climatological distribution of expected hydrologic conditions for the forecast period. Future work in refining methods to bias-correct and downscale climate forecasts for use in hydrologic prediction must therefore take care to minimize the addition of method-related uncertainty.

Further exploration of the approach should also include a broader range of climate and land surface conditions than were here examined. Where snowpack plays a major role in the seasonal cycle, for example, contributions of the hydrologic and climate components of the forecasts are expected to be significantly different at different times of the year. An effort to determine *a priori* where and when long range forecasts are likely to have skill, based purely on land surface and climate considerations, would be useful.

Our experience in this study suggests that future work to apply the climate-hydrology model forecasting approach in real-time and to assess the outputs quantitatively should also concentrate on improvements in two areas. First, the accurate estimation of initial hydrologic conditions for the forecasts is critical for capturing the influence of land surface anomalies that persist into the forecast period; hence an improvement in real-time access to meteorological data for hydrologic simulation of initial conditions would increase the accuracy of the forecasts. Second, quantitative evaluation of the climate forecasts would benefit from the existence of retrospective meteorological forecast datasets generated with the identical methods (to the extent possible) used in producing the current forecasts, for a climatology period of several decades. The retrospective perfect SST-based surrogate forecast of the type explored here hints only at an upper bound on forecast performance, rather than an estimate of performance consistent with the forecasts produced in real time.

Nonetheless, from an end user standpoint, the forecasting approach appears to have potential utility for conditioning water resources related outlooks, particularly when there is either a strong anomaly signal in the climate forecasts or highly anomalous antecedent conditions in the hydrologic model state. A quantitative exploration into the suitability of the spatial and streamflow forecasts for particular water resources applications appears to be warranted.

III. HYDROLOGIC IMPLICATIONS OF DYNAMICAL AND STATISTICAL APPROACHES TO DOWNSCALING CLIMATE MODEL OUTPUTS

This paper is accepted for publication in the journal *Climatic Change* in its current form: Wood, A.W., Leung, L. R., V. Sridhar and D.P. Lettenmaier, 2004, Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs.

1. INTRODUCTION

Improved understanding of the interactions between ocean, land, and atmosphere has led to definitive advances in the ability to forecast weather and climate using complex models of the ocean-land-atmosphere system (e.g., Betts *et al*., 1997; Livezey, *et al*., 1997; Shukla, 1998; Koster *et al*., 1999). Despite improved skill in weather and climate forecasts, hydrologists struggle with how best to use forecast information in applications such as water resource planning, management and conservation as well as irrigation and drainage for sustainable development. The lack of spatial specificity and accuracy has rendered weather and climate forecasts inadequate for hydrologic applications that have serious ramifications to stakeholders and the society at large (Stern and Easterling, 1999).

One factor that has limited the use of climate forecast information in hydrological prediction is the scale mismatch between climate model output and the spatial scale at which hydrological models are applied – typically some subdivision, either natural (subcatchment) or gridding of a watershed (e.g., Lettenmaier *et al*., 1999; Wood *et al*., 2002; Wilby *et al.,* 2000). Various studies have evaluated downscaling methods designed to bridge this gap, particularly in terms of their ability to reproduce surface temperature and precipitation fields (IPCC, 2001; Leung *et al*., 2003). The methods that have been most widely used include dynamical modeling by nesting a regional climate model (RCM – see Leung *et al.,* 2004) within a general circulation model (GCM) (Cocke and LaRow, 2000; Leung *et al*., 1999; Giorgi and Mearns, 1991; Kim *et al*., 2000; Yarnal *et al*., 2000), statistical or empirical transfer functions that relate local climate to GCM output

(Hewitson and Crane, 1996; Wilby and Wigley, 1997; Wilby *et al*., 1998) and climate-anolog procedures (IPCC, 1996). Still other methods (e.g., Charles *et al*., 1999) combine dynamical and statistical procedures. While dynamical and statistical downscaling approaches yield similar reproductions of current climate, they can nonetheless differ significantly in their projections of future climate conditions. Studies by Murphy (1999), Kidson and Thompson (1998) and Wilby *et al*. (2000) further suggest the need to bias correct climate model output to assure meaningful results in applications like hydrologic and water resources assessments.

The papers in this special issue report results from the pilot phase of the Department of Energy Accelerated Climate Prediction Initiative (ACPI), which used GCM scenarios of future climate produced by the DOE-NCAR Parallel Climate Model (PCM; Washington *et al.,* 2000, Dai *et al*., 2004). A variety of methods were used in ACPI projects to downscale PCM output. Dettinger *et al*. (2004) studied water resources impacts of climate change projected by PCM in several subbasins of the Sacramento-San Joaquin River basin after adjustment of historical climate model output to match daily observed precipitation and temperature statistics. Payne *et al*. (2004) utilized variations of the probability mapping methods described by Wood *et al.* (2002) for spatial downscaling and bias correction of both global and regional climate model outputs in their investigation of water resources impacts of climate change in the Columbia River Basin (CRB). Van Rheenen *et al*. (2004) and Christensen *et al*. (2004) used similar approaches in studies of the Sacramento-San Joaquin and Colorado River basins, respectively. Vail and Wigmosta (2004) studied the impacts to fisheries in the Yakima River Basin using PCM climate scenarios downscaled by the regional climate model (RCM) of Leung *et al.* (2004) and subsequently bias corrected using statistical methods. All of these studies used different methods to downscale and bias correct the global or regional model outputs in order to produce realistic simulations of hydrologic conditions of the current climate. It is worth noting that this is a de facto minimum standard of any useful downscaling method for hydrologic applications: the historic observed conditions must be reproducible.

Few studies have evaluated the differences among various downscaling methods based on their implications for hydrological predictions (Crane *et al*., 2002; Wilby, 2000). There remain critical questions, for instance, about the value of dynamic downscaling, given that biases inevitably remain that must be removed, usually by subsequent application of statistical methods. With

these questions in mind, we evaluated six different methods of downscaling from global or regional models to the still finer scale of a grid based hydrological model (specifically, the Variable Infiltration Capacity, or VIC model). Included are three statistical downscaling methods -- linear interpolation (LI), spatial disaggregation (SD) and bias-corrected spatial disaggregation (BCSD) -- applied either directly to PCM outputs or to dynamically downscaled (to intermediate resolution) PCM output, i.e., to the output of a regional climate model. The six methods were compared through application to a retrospective climate simulation, and those that performed the best were also applied to a future climate scenario. The spatial domain of all comparisons was the Columbia River Basin (CRB) of the U.S. Pacific Northwest (PNW) region (see Payne *et al.,* 2004 for background).

2. APPROACH

The general approach was to simulate land surface energy and water fluxes using the VIC macroscale hydrological model (see Section 2.2), driven by meteorological outputs from PCM with and without intervening dynamical downscaling using a regional climate model. Results of a twenty-year climate-hydrology scenario were evaluated by comparison with a retrospective observational analysis of surface climate and hydrologic conditions. Implications of the more successful of the approaches were also explored for a future climate run. The observational analysis is discussed in Section 2.1, the models and simulations in Section 2.2, and the downscaling methods in Section 2.3. Sampling error issues are discussed in Section 2.4

2.1 Observational Analysis

The observed climatological and hydrological fields used to evaluate the downscaled climate model outputs were taken or derived from the hydroclimatic retrospective analysis of Maurer *et al*. (2002), which is based on a 1/8 degree hydrologic simulation of land surface energy and water variables run at a 3 hour timestep over the continental U.S. for the period 1950-2000. We used average monthly temperature and soil moisture, total monthly precipitation, runoff and evaporation, and basin-averaged monthly snow water equivalent. The climate variables (precipitation and temperature) for the 20-year retrospective period (approximately 1976-96) were taken directly from the Maurer *et al*. (2002) dataset, and the hydrologic ones were generated via a retrospective simulation of the 20 year period (with a 2-year hydrologic model spin-up,

producing an initial model state from which all retrospective runs were started), driven by climate variables taken from the same gridded observations.

2.2 Models and Simulations

We used output from two climate models: PCM and the RCM of Leung *et al*. (2004) for that portion of their domains included within the VIC model's 1/8 degree representation of the PNW. The PNW domain is divided by the Cascade Mountain range into coastal basins draining to the west, and all of the CRB. Almost all of the region has a winter-dominant precipitation regime in which most of the annual precipitation is derived from frontal systems originating in the North Pacific, the majority of the moisture from which falls on the west slopes of the Cascades. Although much drier, over the CRB to the east winter precipitation occurs mostly as snow (in the mountains of British Columbia, Canada, and Idaho and Montana), much of which contributes to a strong seasonal runoff peak in the late spring and early summer. Figure 3.1 shows the study domain, along with the PCM and RCM grid alignments (T42 or 2.8125 degrees latitude/longitude for PCM, and ½ degrees latitude/longitude for RCM). Also shown are the PCM and RCM average annual precipitation and temperature climatologies for the period 1975-95, at the resolution of each climate model, and the observed 1/8 degree climatology described in section 2.1. The figure shows the effect of RCM's higher spatial resolution relative to PCM (PCM represents the PNW region with about 20 grid cells, while RCM uses about 500, and the observed 1/8 degree climatology has about 6400).

The climate scenarios and climate model simulations used in the study are described in greater detail elsewhere in this issue (Dai, *et al*., 2004 for PCM and Leung, *et al*., 2004 for RCM), but in brief, they resulted from retrospective historical simulation and future climate simulations, based on a observed historical greenhouse gas and aerosol emissions for the historical run, and "business as usual" (BAU) global emissions future climate. Because the RCM simulations were of length 20-years (using a subset of longer PCM sequences to represent boundary conditions), all analyses were based on the 20-year simulations for both PCM and RCM for the periods designated "RCM subset" in Table 3.1 to avoid sample length differences.

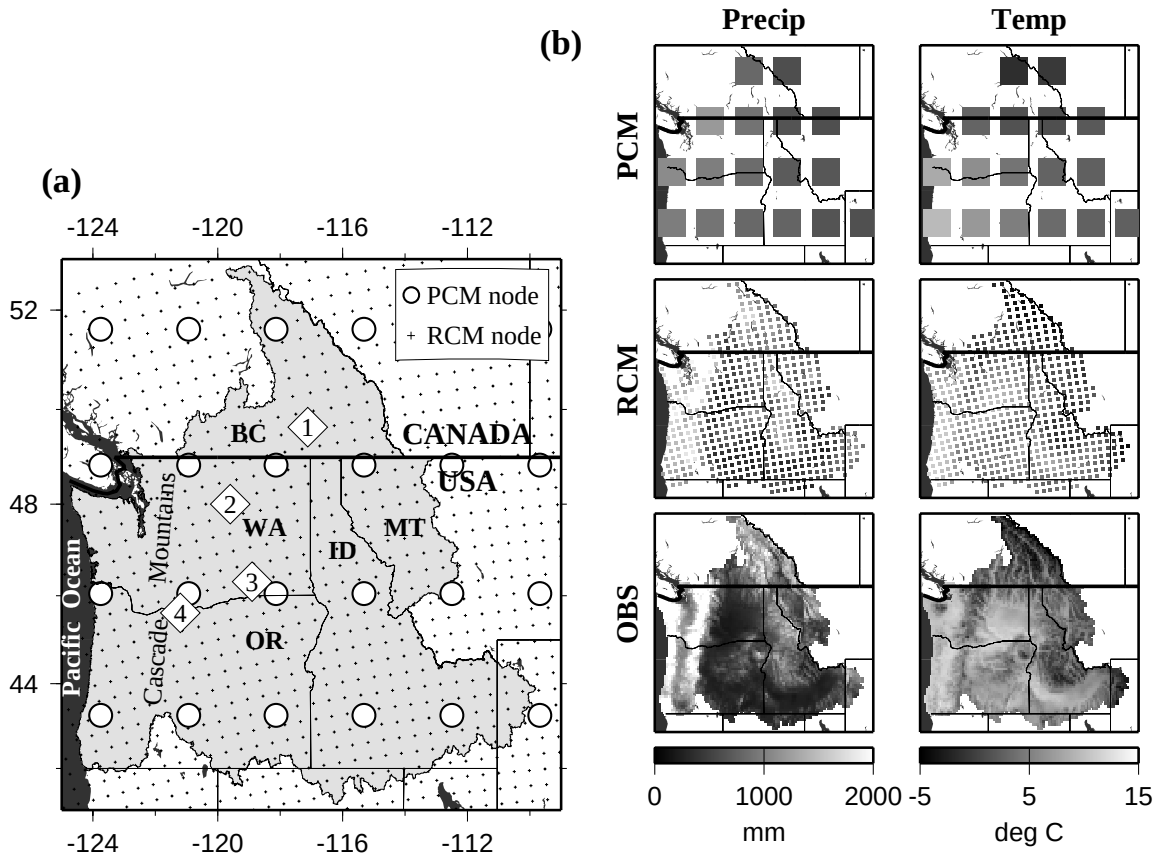

Figure 3.1 (a) CRB Domain with PCM and RCM model grids and four streamflow simulation locations (diamonds: 1-Corra Linn; 2-Chief Joseph; 3-Ice Harbor; and 4-The Dalles). (b) CRB annual average 1979-95 model climatologies for total precipitation and average temperature at PCM's T42 and RCM's ½ degree resolutions, compared with the 1/8 degree observed climatology of Maurer *et al*., (2002).

Table 3.1 Simulations used in this study

Run	**Description**	**Run Period**	**RCM Subset**
B06.22	Historical (greenhouse CO_2+aerosols forcing)	1870-2000	10/1975 - 9/1995
B06.44	Climate Change (future scenario forcing)	1995-2099	7/2040 – 6/2060

The hydrologic model used in this study, the Variable Infiltration Capacity (VIC) model of Liang *et al*., (1994; 1996; 1999) is a semi-distributed grid-based hydrological model which parameterizes the dominant hydrometeorological processes taking place at the land surface-atmosphere interface. The VIC model has been implemented previously for the CRB, and the calibration procedure and results are described in Nijssen *et al.* (1997) and Payne *et al*. (2004). VIC model climate inputs for this study were daily precipitation, maximum and minimum temperature, and daily average wind speed, for each 1/8 degree model grid cell (other forcing variables – specifically downward solar and longwave radiation, and dew point – were derived using methods described by Maurer *et al*., 2002). Because VIC was run at a finer spatial resolution than the climate models, a downscaling step (methods described in the next section) to bridge the resolution gap between climate model and VIC was implemented, whether PCM or RCM output was used. The VIC model was applied to the entire PNW study domain of Figure 3.1, although the hydrologic analysis was confined to the CRB drainage upstream of The Dalles, OR, a domain identical to that used by Payne *et al*. (2004). The primary difference between the VIC model used for this study and the Payne *et al*. (2004) implementation was grid resolution: we used 1/8 degrees longitude and latitude rather than 1/4 degrees to afford a greater resolution gap for the downscaling method evaluation. Streamflow results are reported for four locations shown in Figure 3.1: Kootenai River at Corra Linn Dam, Columbia River at Chief Joseph Dam, Snake River at Ice Harbor Dam, and the Columbia River at The Dalles, OR. These reflect, roughly, streamflow effects in the Canadian portion of the basin, the middle and upper Columbia River, the Snake River drainage, and the entire basin. Figure 3.2 shows simulated streamflow at these locations when VIC was driven by observed precipitation and temperature. The simulations generally reproduce the observed long term monthly mean hydrograph and also capture interannual flow variation.

2.3 Downscaling Methods

For each of the climate model runs summarized in Table 3.1, we compared six approaches for downscaling climate model output: three model output post-processing approaches – linear interpolation (LI), spatial disaggregation (SD) and bias-corrected spatial disaggregation (BCSD) -- each applied to PCM output directly and to RCM output, which represents an intermediate dynamical downscaling step. To distinguish between three direct PCM output methods and the

three RCM output methods, a prefix of PCM- or RCM- is used with the post-processing method designator. Regardless of the method, the climate model output fields that were downscaled were the same: monthly mean temperature (T_{avg}) and total precipitation (P_{tot}), at the climate model resolution. Each downscaling procedure reproduced these as input fields for the VIC model at 1/8 degree resolution, and an additional step was taken to disaggregate the monthly fields into daily time series required by VIC (daily precipitation, maximum and minimum temperature, and daily average wind speed). This final disaggregation step is identical to that used in Wood *et al*. (2002), and is summarized briefly in Section 2.3.1 below. Sections 2.3.2 and 2.3.3 summarize the LI and SD approaches, primarily focusing on their differences from BCSD.

2.3.1 Bias correction of climate model output, followed by spatial disaggregation (BCSD)

For direct use of PCM output, T_{avg} and P_{tot} forcings from each climate model cell centered within the study region were treated individually for purposes of bias correction. For bias removal, a quantile-based mapping (e.g., the empirical transformation of Panofsky and Brier, 1968) was constructed from the PCM model climatology to the observed monthly climatology for each variable (T_{avg} and P_{tot}). The observed climatology was derived from Maurer *et al*. (2002) for the period 1975-95, re-gridded and averaged to the PCM grid resolution. The PCM climatology was taken from modeled T_{avg} and P_{tot} from the B06.22 simulation for the same period. The mapping from PCM to observed climatology was subsequently applied to the PCM raw output, translating it to a plausible range with respect to historical observations. The mapping was performed at the resolution of the PCM output, hence the adjustments vary spatially at the PCM grid scale and by month. For the BAU scenarios, the PCM cell-specific temperature shift (monthly averages relative to the B06.22 retrospective run monthly averages) were removed from the uncorrected PCM output before, and replaced after, the bias-correction step. For the BAU runs, this step was needed because the BAU temperature distribution was quite different from that of the climate model historic run. When the temperature shift was removed, the spread of the BAU run temperature distribution was near the historical range, enabling the bias-correction step to be applied with little extrapolation. The basic assumption of this approach is that the variability of the BAU run temperature distributions will remain similar to the retrospective run variability, despite the BAU mean shift.

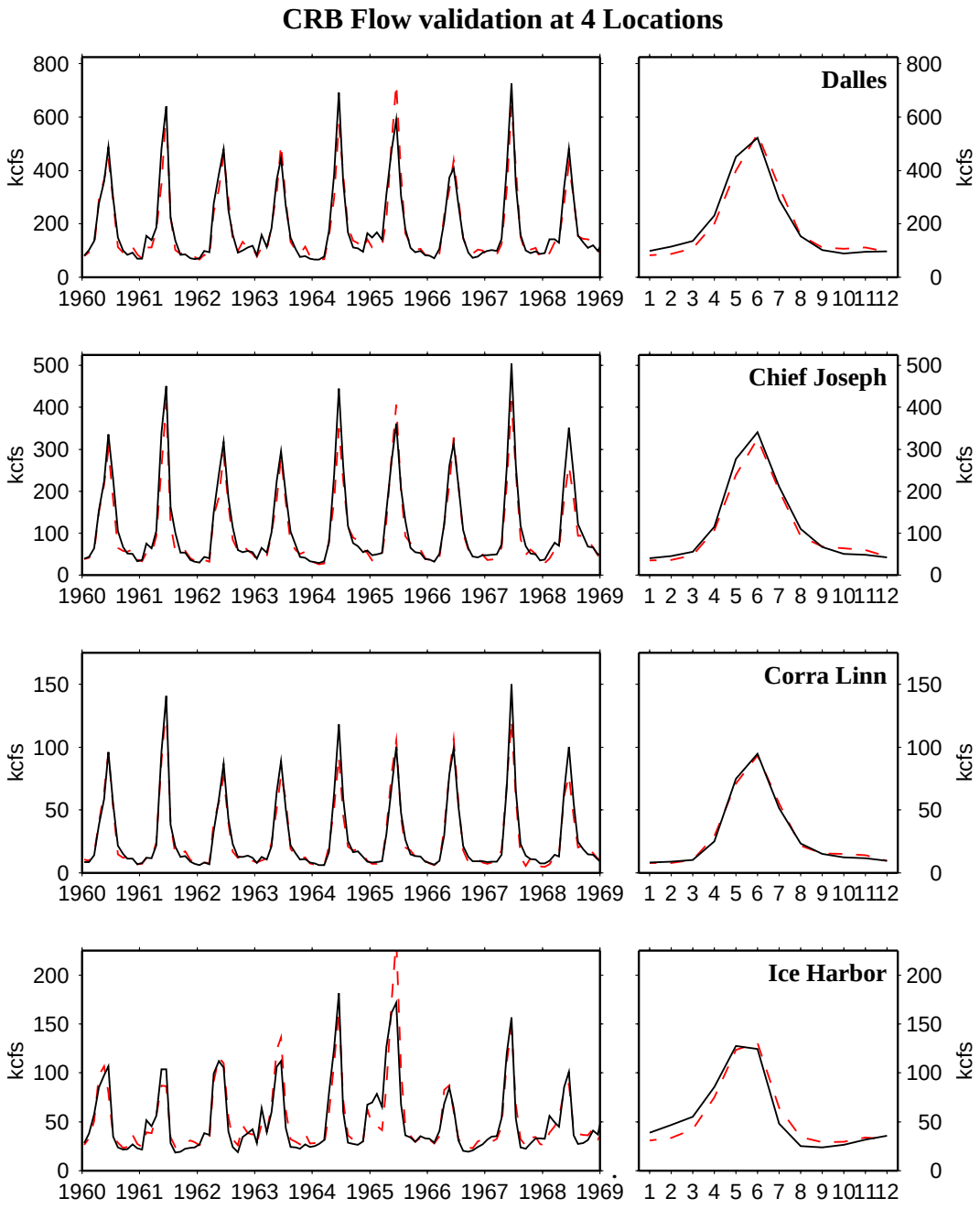

Figure 3.2 VIC model validation results for four streamflow routing locations (solid line is observations; dashed line is simulation). A subset timeseries from the validation period is shown at left, the monthly mean hydrographs at right.

Spatial disaggregation imposed sub-PCM grid scale spatial variability on the bias-corrected, PCM-scale forcings. The monthly time step, bias-corrected PCM-scale BAU scenario time series were spatially interpolated to the hydrology model grid cell centers. Anomaly fields (multiplicative for precipitation, and additive for temperature, different for each calendar month), developed from the observed climatological monthly means (for T_{avg} and P_{tot}) were applied to the resulting 1/8 degree monthly variable fields as follows: (a) observed monthly mean T_{avg} and P_{tot} 1975-95 averages were aggregated to the climate model scale (T42 or ½ degree), and then interpolated back to the 1/8 degree scale, exactly as the climate model scale forcings were interpolated; (b) the differences (for temperature) or ratios (for precipitation) between the 1/8 degree monthly mean T_{avg} and P_{tot} and the interpolated monthly mean fields were calculated to create the anomaly fields. The mean monthly sets of anomaly fields so constructed, when applied to timeseries of interpolated climate model-derived fields, added spatial variability to the smooth 1/8 degree field created by the interpolation step. The spatial disaggregation created VIC scale monthly forcing time series corresponding to the PCM scale time series, but reflecting VIC-scale spatial structure.

Finally, a temporal disaggregation step was used to form daily time step inputs for the VIC model. The monthly forcing time series were replicated using scaled or shifted daily patterns sampled from the historic record, at the hydrology model resolution. Month-long daily patterns of precipitation and temperature (more specifically T_{min} and T_{max}, with T_{avg} defined subsequently as their average) were sampled for each monthly timeseries by picking a year from the 50-year climatology period at random. Each sampling year was used for the entire CRB domain to preserve a degree of synchronization in the weather components driving hydrologic response. The daily patterns were then scaled (for precipitation) and shifted (for temperature) to match the monthly timeseries (in T_{avg} and P_{tot}) created by applying the interpolated, bias-corrected PCM anomalies to the VIC cell climatological means. Various screening methods were applied to the precipitation patterns to ensure that rescaling did not result in unrealistic values. The same temporal disaggregation step was applied in all six methods, to avoid confounding the results by differences in the derivation of daily weather patterns. The rationale for use of monthly rather than daily or sub-daily climate model outputs is discussed in Wood *et al*. (2002).

Application of the BCSD method to RCM output was as described above, with the following differences:

- Instead of PCM output for T_{avg} and P_{tot}, RCM's dynamically downscaled monthly T_{avg} and P_{tot} at ½ degree were used. A ½ degree observed climatology was developed for bias-correction by aggregation from the Maurer *et al*. (2002) archive.

- The spatial disaggregation began with bias-corrected ½ degree monthly fields rather than the 2.8 degree (PCM resolution) fields.

Note that RCM is driven with a more comprehensive set of PCM output fields than the limited surface variables used in our hydrologic downscaling (see Leung *et al*., 2004 for details).

2.3.2 Spatial disaggregation of climate model output, without bias correction (SD)

The SD approach was similar to the BCSD method, except that the PCM or RCM output fields were interpolated to the 1/8-degree VIC model grid without the intervening bias-correction step.

2.3.3 Spatial linear interpolation of climate model output (LI)

The LI procedure was similar to BCSD, except that the PCM output fields or RCM output fields were linearly interpolated to the hydrologic model grid cells without the intervening bias-correction step, and without spatial disaggregation. The LI approach is intended to provide a baseline for comparison with the other methods because it adds the least additional information to the raw output of the climate model. More elaborate interpolation approaches exist that draw from ancillary information sources (e.g., using elevation data to estimate precipitation gradients across an interpolation space, as in Hutchinson, 1995), to yield more intelligent distribution of the interpolated data. These methods arguably fall closer to the category of spatial disaggregation (SD), and are not considered in this paper, given our inclusion of a separate SD method.

2.4 Method Discussion

The success of two of the techniques – BC and SD – depends on the stability over time of the probability distributions used to correct climate model bias and to impose spatial variability, respectively. In the retrospective assessment, the probability distributions were estimated from

the same simulation period that was being downscaled. Hence the distributions at first glance appear to yield unbiased results after downscaling – although it should be understood that even in this case, the historic period provides only an estimate to the underlying statistical populations of the variables being downscaled, and there is bias associated with the short record length from which the probability distributions were estimated. Had the corrections been applied to another retrospective period, the effect of bias resulting from the relatively short record length used to estimate the probability distributions would have been more apparent If the timeseries of the variables in question are stationary and the variance is relatively small, this should be a somewhat minor issue, as small samples (e.g., N=20, as in this paper) will not produce large sampling biases, at least in the estimation of the means of derived hydrologic variables. On the other hand, for variables with larger bias, and/or estimation of variables near the tails of the probability distributions, the problem is potentially important – notwithstanding that it can be alleviated by use of a longer period of coincident historical observations and climate simulation.

As noted in Section 2.3.1, for the BAU climate downscaling, the sampling bias issue is compounded by the need to estimate unbiased probability distributions of *future* climate with which to adjust the climate model biases. We recognized through exploratory data analysis that the PCM precipitation distribution changes are small enough that the retrospective period can be used to estimate these distributions, but the shift in the climate model temperature distributions cannot be ignored. Our approach to minimizing bias in future temperature distributions and the required assumptions is described in section 2.3.1.

A thorough investigation of the sampling bias issues associated with estimation of probability distributions from short record lengths is beyond the scope of this paper. Instead, we report here a brief investigation of the implications. The goal is to estimate the extent to which biases in results could arise from errors in estimating probability distributions of the underlying variables. The discussion applies to the means of derived variables, and to the BC method only.

Using a Monte Carlo framework, 500 pairs of samples of 20 years of monthly P_{tot} and T_{avg} (each year drawn randomly with replacement) were taken from the observed 1951-99 PCM-scale record over the CRB domain, and from the spatial means, monthly probability distributions of sampling errors were estimated. These are shown in Figure 3.3 (panels a and b) by the 95 percent confidence limits, together with the mean of the 49-year period. For precipitation, the largest

errors are roughly 20-25 percent in winter and spring, while the largest temperature errors are 1-2 degrees in winter, and 0.9 degrees in spring. It should be noted that winter and spring precipitation and temperature are the dominant meteorological variables affecting streamflow.

In this limited investigation, it was not feasible to carry this analysis through the hydrologic simulation. Instead, we estimated the effects of the precipitation and temperature errors on mean monthly basin-averaged runoff using 2000 randomly drawn non-consecutive 20-year samples from the simulated runoff associated with the full 49-year period. Figure 3.3 (c) shows the variation in May-August runoff as a function of December-March precipitation and temperature. While the May-August runoff is relatively insensitive to temperature variations, absolute changes (in mm) were approximately half the absolute winter precipitation changes (in mm). Figure 3.3 (d) shows the variation in the runoff fraction May/June as a function of April-June precipitation and temperature. The runoff fraction ranges from approximately 0.85, for the retrospective climate, to approximately 1.15 for a shift in peak flow commonly associated with moderate climate warming in this region. The implications of precipitation and temperature sampling error for runoff, at the extremes, are that a 1-degree Celsius warm bias in spring temperature, about the same magnitude as the 0.975 non-exceedence sampling error, is sufficient to produce the entire shift, and a winter precipitation negative bias could produce a 20 percent summer runoff reduction. On the other hand, as the results in Section 3 show, these runoff biases are minor relative to the distortions arising from a failure to bias correct climate model output, even using a short correction period.

These results are not comprehensive, but illustrate that the 20 year scenario lengths used here are on the shorter end of the size needed to produce robust correction distributions for application to different scenario periods. Again, the magnitude of these errors can be reduced by increasing the length of the retrospective period, although it should be noted that the rate of error reduction is expected to go roughly as the ½ power of the record length, so even using the entire 49-year period of historic observations (and assuming RCM runs of this length were made) would only reduce the errors by a factor of about 1.6.

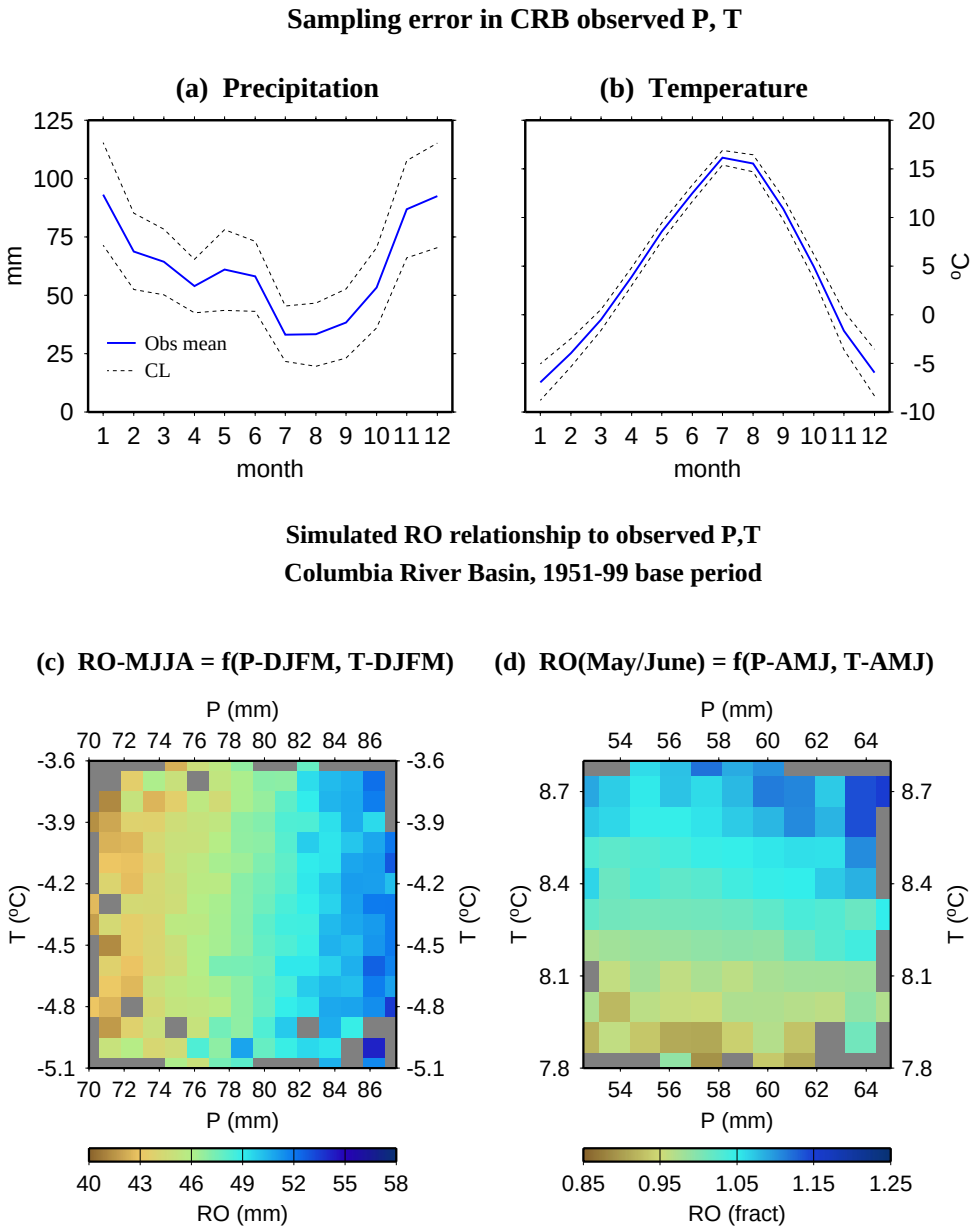

Figure 3.3 Sampling error 95 percent confidence limits for monthly precipitation (a) and temperature (b); c) dependence of May-August runoff (Q) on December – March precipitation and temperature; and d) dependence of ratio of May runoff to June Runoff on April-June precipitation and temperature.

3. RESULTS

The results for the three different downscaling approaches tested in this study are organized in the following way. Results are presented first for the retrospective climate simulation, followed by the future climate simulation, and each is compared to the observational analysis described in Section 2.1. Results include (a) temporally averaged spatial climate fields (monthly total precipitation and monthly average air temperature) for December and July (which reflect winter and summer conditions); (b) associated spatially averaged variables (monthly precipitation, temperature, evapotranspiration, snow water equivalent, runoff and soil moisture); and (c) monthly average streamflow (runoff routed through a stream network) at four locations in the CRB shown in Figure 3.1.

3.1 Retrospective Analysis (October 1975 – September 1995)

3.1.1 Spatial analyses of precipitation, temperature and snow water equivalent

For December and July, average monthly total precipitation and averaged temperature (Figures 3.4 and 3.5) were compared with the PCM and RCM-derived retrospective simulation (B06.22) results. The main features of the observed climatology for precipitation (top row, Figure 3.4) are a spatial divide between higher precipitation to the west of the Cascade Mountains (which run north-south at about longitude 121-122W – see Figure 3.1) and lower precipitation to the east, and a temporal divide between high and low precipitation in winter (December) and summer (July), respectively. A second order feature is associated with the higher precipitation areas in Canada, Idaho and Montana, which correspond primarily to higher elevations. The LI results show that PCM simulates the west-east gradient toward lower precipitation in the December, and somewhat reproduces a spatially smoothed version of observed precipitation with the reverse in the July, but not surprisingly fails to capture any elevation-dependent features (hence at the local scale is biased almost everywhere, even though the basin average has only moderate bias). The RCM better resolves these spatial features, but shows a wet tendency in December (except in coastal areas and British Columbia, where it is too dry), and a dry tendency in July. Spatial disaggregation (SD) alone leads to better representation of precipitation for PCM, greatly reducing the December local precipitation biases and nearing RCM's performance with LI. Because of RCM's better resolution, RCM-SD closely resembles RCM-LI, although SD appears

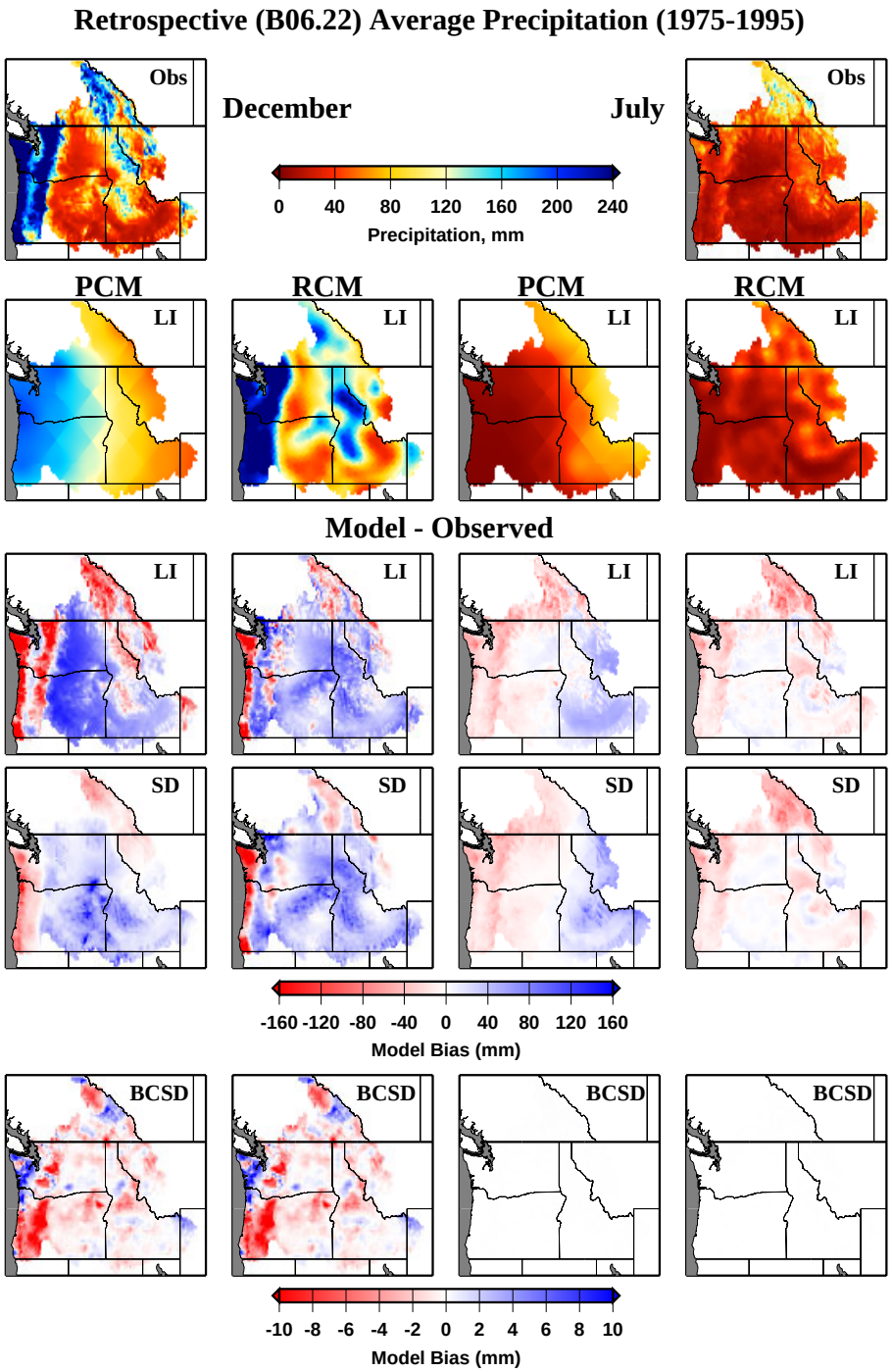

Figure 3.4 December and July total precipitation for the PCM and RCM-driven retrospective simulations (1975-95), and for each downscaling method, as compared with the observed climatology (top row) for the same period. LI method values are shown in the second row, below which are differences from observed values for the LI, SD and BCSD methods.

to exacerbate RCM-LI local biases in some areas while improving them in others. BCSD reduces differences between observed and simulated monthly average precipitation in both December and July to within 1-2 percent of observed.

The main feature of the observed climatology for temperature (top row, Figure 3.5) is a cooling gradient, present both in winter (December) and summer (July), from southwest to northeast, which is moderated primarily by elevation, and secondarily by humidity effects associated with lower precipitation east of the Cascades Mountains (which, along with the Snake River plain in the southeast, is clearly identifiable in the observed climatology). The interpolated PCM-LI and RCM-LI results both capture the primary gradient, but RCM is clearly superior in resolving the temperature range and spatial distribution across the basin. Both models show cold and warm biases in the lower and higher elevation areas, respectively, but these are much stronger in PCM, particularly in December. For PCM, the SD method alone removes much of the spatial elevation related bias for July and December, leaving broad scale biases of a few degrees or less. (It should be noted, however, that this apparent agreement is somewhat deceptive, as the hydrological model is quite sensitive even to temperature biases of this magnitude). For RCM, SD may also smooth biases arising from the finer resolution of the observed climatology, but RCM's initial biases, for the most part, remain. BCSD improves the results to the point that the PCM and RCM monthly average temperature simulations match the observed means to within a few tenths of a degree in December, and a few hundredths of a degree in July.

For the CRB domain (upstream of The Dalles, OR, bordered to the west by the Cascade Mountains rather than the Pacific Ocean), the simulated average April 1 snow water equivalent (SWE), which reflects the effects of winter and spring temperature and precipitation, is shown in Figure 3.6. PCM-LI completely fails to capture both the magnitude and spatial distribution of SWE, and although interpolated RCM distributes snow correctly, it has a high bias, particularly in the Snake River plain, western Montana and eastern Oregon. With SD, PCM derived results improve greatly, despite leaving a high bias in eastern Oregon and central Idaho and a low bias in BC. For RCM, SD makes little difference. For both PCM and RCM, the BCSD method eliminates most of the bias inherent in using both RCM and PCM outputs directly, although some small differences between the two are evident, and some very localized biases remain (such as in the northern part of the study domain).

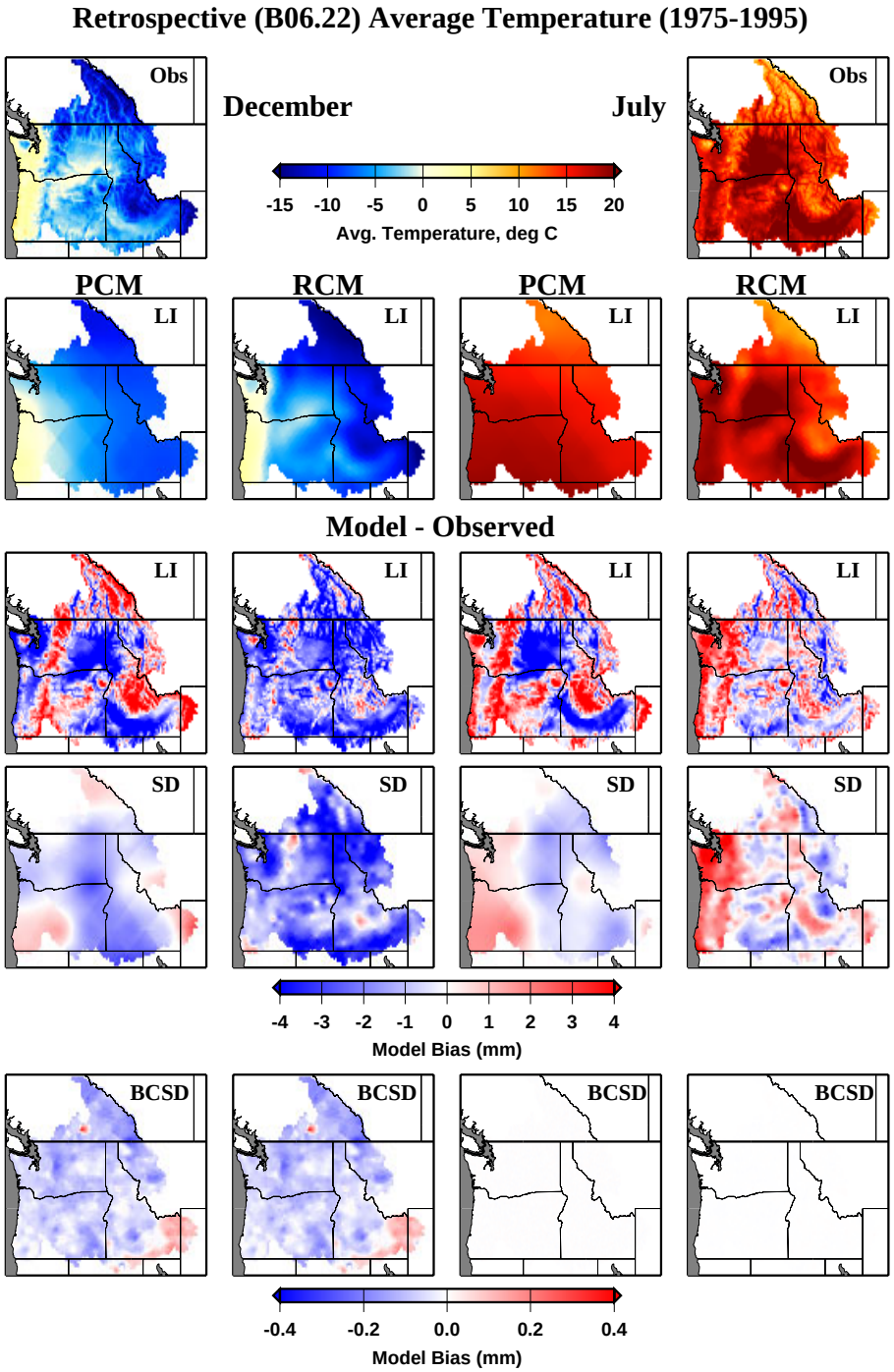

Figure 3.5 December and July average temperature for the PCM and RCM-driven retrospective simulations (1975-95), and for each downscaling method, as compared with the observed climatology (top row) for the same period. LI method values are shown in the second row, below which are differences from observed values for the LI, SD and BCSD methods

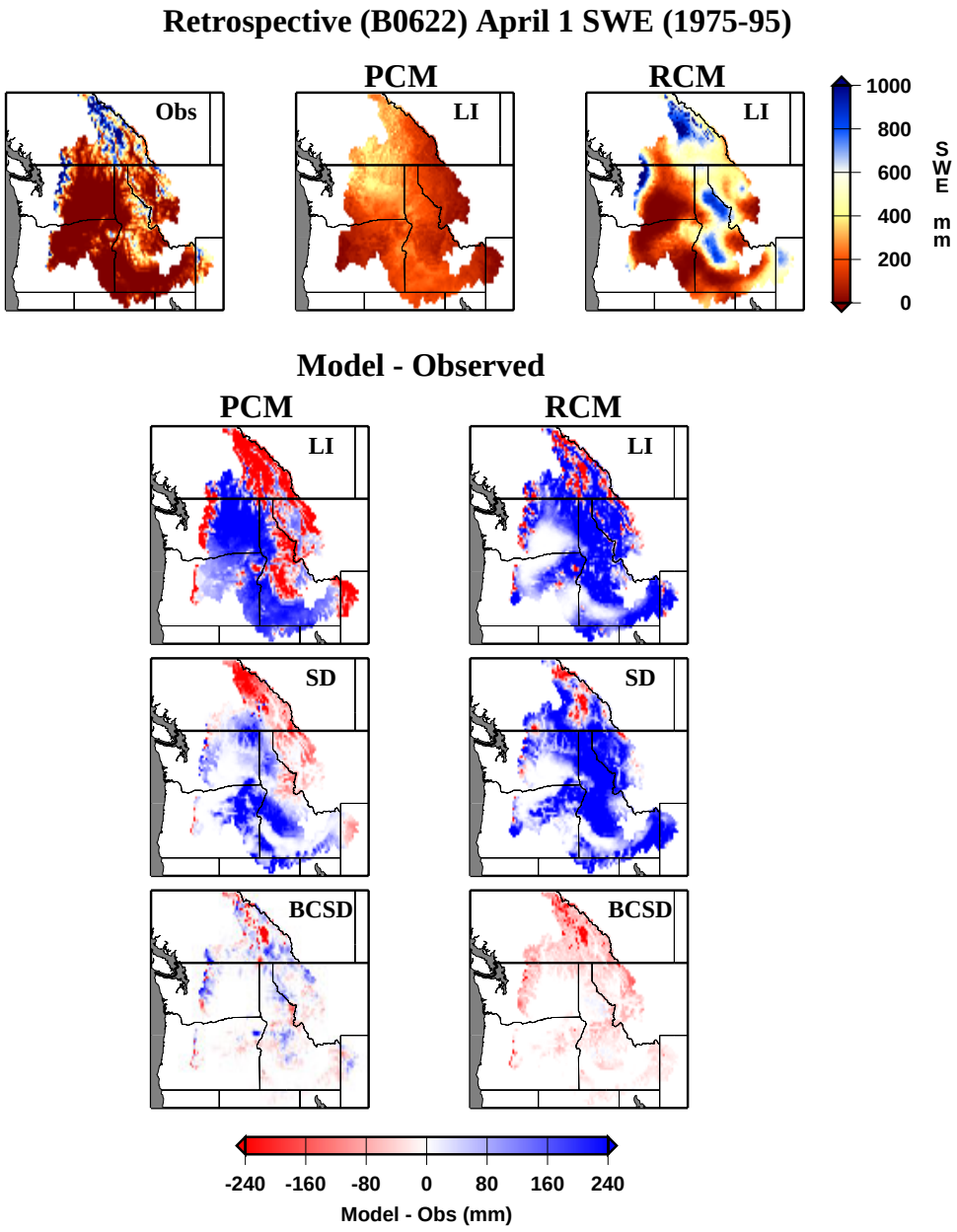

Figure 3.6 Average April 1 SWE simulation for the retrospective 1975-95 climate simulations, compared with the observed (simulated by the VIC model forced with observations) climatology for the same period (top left). LI method values are shown to the right of the observed values; rows 2-4 contain differences from observed values for the LI, SD and BCSD methods.

3.1.2 Basin-average monthly analysis

The basin average monthly analysis (Figure 3.7) shows that the LI and SD methods produce nearly the same basin-wide precipitation and temperature signal for each model. In dynamical downscaling, however, there are no physically-based mechanisms that constrain the simulation to preserve PCM's basin mean precipitation or temperature. In this example, while the RCM changes the PCM temperature signal only slightly (leaving a cold winter and spring bias), it worsens the bias in the seasonality of precipitation, particularly the high bias in fall and winter. For RCM, with the LI and SD methods, the high bias in fall and winter precipitation leads to an oversimulation of SWE, soil moisture and summer evaporation. For PCM, the results are varied, with SWE and runoff undersimulated for the LI method and oversimulated for the SD method, but soil moisture and summer evaporation are oversimulated for both. SD has little effect on the interpolated results for RCM, but as before, greatly changes the interpolated results for PCM. The BCSD method, by definition, forces the mean and variance of the PCM and RCM output to equal the observed distribution, so for precipitation and temperature, the PCM and RCM BCSD results cannot be discriminated (in top row panels of Figure 3.7) from the observed climatology. For hydrologic variables, however, the exacting monthly corrections of precipitation and temperature alone do not eliminate all biases relative to the observed hydroclimatology. Note that PCM and RCM BCSD runoff shifts slightly earlier in the year, SWE is reduced, and soil moisture is slightly out of phase as compared with values simulated directly using VIC forced with gridded observations.

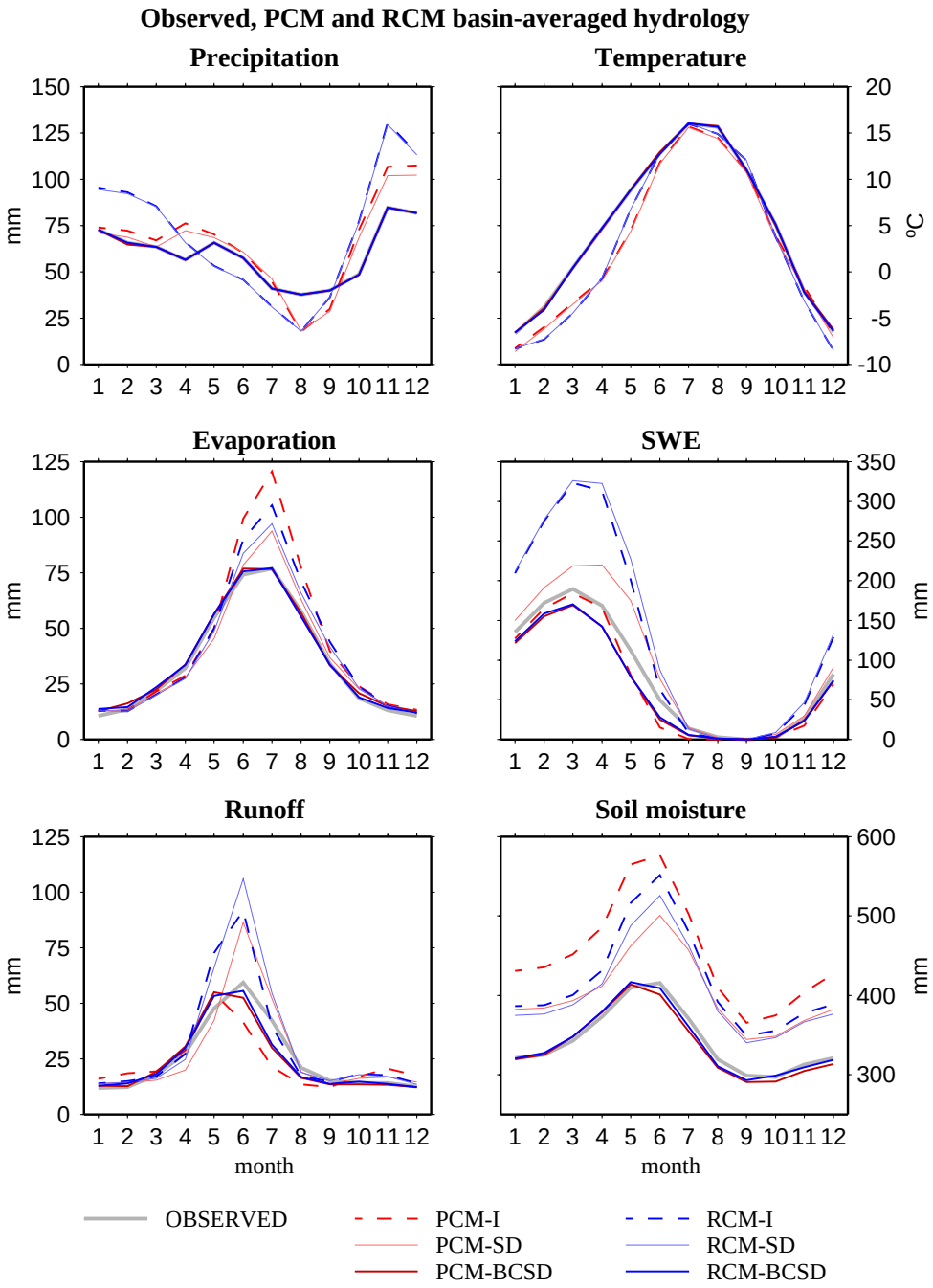

Figure 3.7 Columbia River basin averages of climate and hydrology variables for the retrospective 1975-95 climate simulations, compared with the observed climatology (i.e., observed precipitation and temperature, and simulated hydrologic variables based on these observations) for the same period (note that the PCM and RCM BCSD methods produce monthly mean precipitation and temperature that are indistinguishable in the figure).

3.1.3 Monthly average streamflow

The plots of monthly average streamflow in Figure 3.8 show the implications of the downscaling methods for different parts of the basin. At The Dalles, RCM's high precipitation bias leads to oversimulation of runoff for both the SD and LI methods; whereas for PCM, the LI results are reasonably close to observed, while the SD method over-simulates runoff. The same is true for PCM at Ice Harbor, while at Corra Linn and Chief Joseph, SD greatly improves runoff simulation. For RCM, the SD and LI methods yield similar streamflows, and these are much improved relative to PCM streamflows at Corra Linn and Chief Joseph, but are worse (due to oversimulation) at Ice Harbor. The SD step in all cases reduces the difference between interpolated PCM streamflow and RCM-LI streamflow. This reflects the fact that RCM inherited large scale bias from PCM and therefore streamflows simulated using the RCM-LI outputs are similar to those simulated by PCM-SD. The BCSD method greatly improves streamflow simulation relative to the other methods for both PCM and RCM, at all four sites, although the small bias toward earlier runoff remains. The PCM and RCM BCSD results are essentially identical, more or less by construct.

3.2 BAU Analysis (July 2040 – June 2060)

3.2.1 Spatial analyses of precipitation, temperature and snow water equivalent

Because of the large biases resulting from the LI and SD methods for the retrospective climate period, results for BCSD only are discussed for the BAU climate. LI spatial plots are shown, however, to help illustrate the differences between the BAU and retrospective scenarios.

For the BAU simulations, after BCSD, the primary changes in precipitation (top row, Figure 3.9), as compared with the observed climatology and with the retrospective climate simulation results (second row, Figure 3.4), are an intensification of precipitation in the northwest and northeast parts of the domain, a drying in the southeast in December, and a moderate drying over most of the region in July. The BCSD results for both models are quite similar, but RCM simulates greater precipitation in some areas in December (particularly west of the Cascade Mountains and in the mountains of Idaho) and less precipitation in the eastern part of the basin in July.

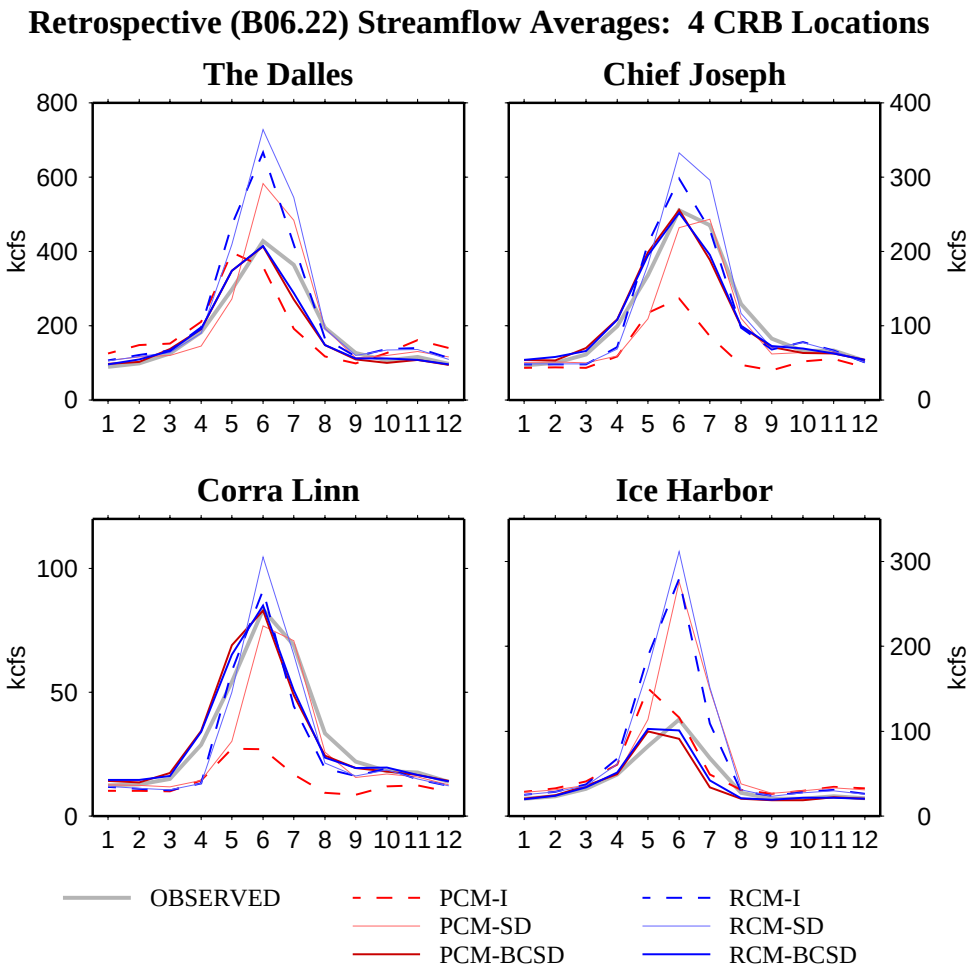

Figure 3.8 Streamflow at four locations (see Figure 3.1) for the retrospective 1975-95 climate simulations, compared with the observed (simulated by the VIC model driven with observations) climatology for the same period.

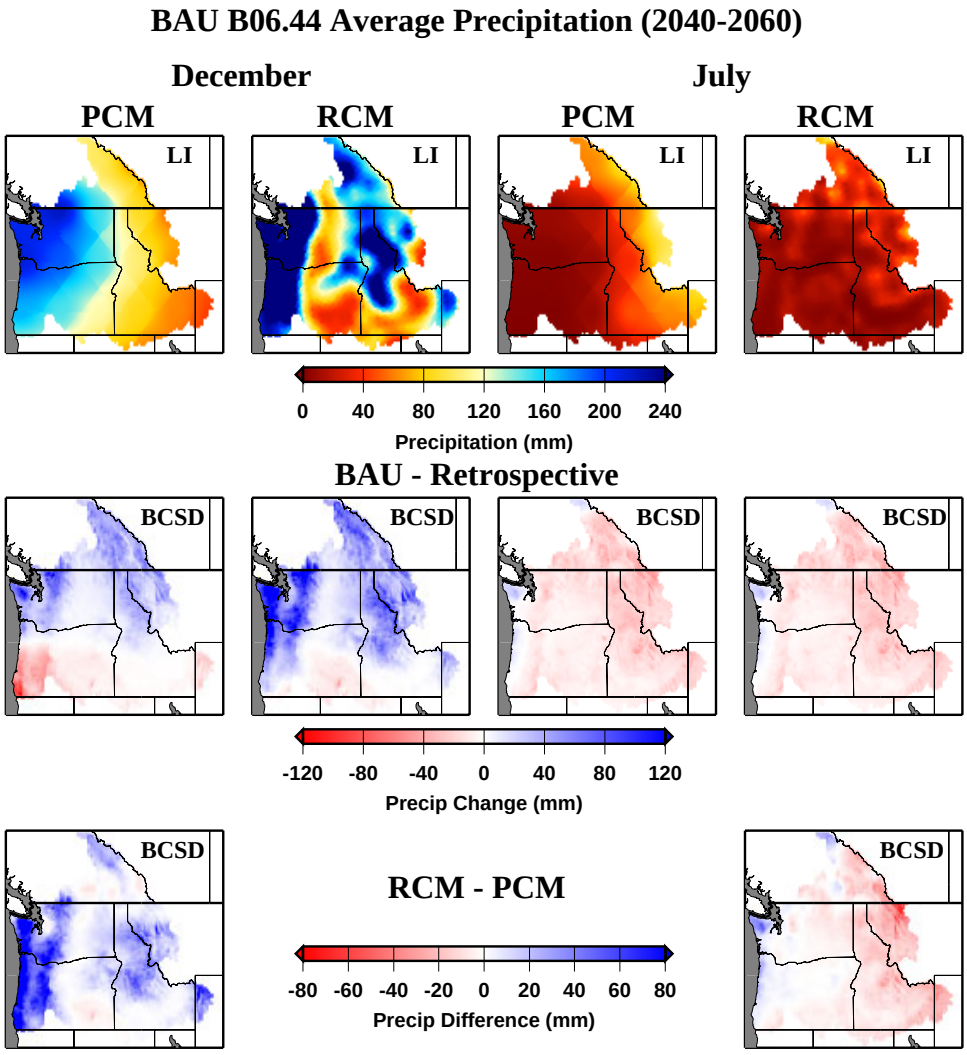

Figure 3.9 December and July total precipitation for the PCM and RCM-driven BAU future climate simulations (2040-60). (top row) With LI only; (second row) differences in BAU BCSD results for PCM and RCM from their retrospective BCSD results; (third row) differences between RCM-BCSD and PCM-BCSD results.

For temperature (Figure 3.10), the BAU simulations from PCM and RCM preserved the spatial patterns as each model's retrospective simulation (Figure 3.5), but were uniformly about 3 and 1.8 degrees Celsius warmer in winter and July, respectively, a difference that PCM-BCSD and

RCM-BCSD simulated almost identically. The two approaches nonetheless lead to differences of up to ½ degree Celsius in places.

For precipitation and temperature, the RCM-PCM differences with BCSD appear in most cases to be reasonably consistent with tendencies present in the retrospective simulations before any adjustment (e.g., comparing RCM-PCM differences in the LI rows of Figures 3.4 and 3.5). For example, the RCM's BAU climate is wetter than PCM's to the west of the Cascade Mountains, where it is also wetter in the retrospective LI approach. The Snake River basin RCM BAU climate is warmer than PCM in July, while in the retrospective simulations, that area is both warmer and drier before bias correction. In winter, the reverse is true (as it is for the eastern rim of the domain). These differences are consistent with the RCM's colder, wetter bias in those areas in the retrospective simulations. Although these differences are damped out in the retrospective BCSD, they filter through BCSD in the BAU simulations.

For the BAU climate April 1 SWE (Figure 3.11), the BCSD method with RCM and PCM yielded significantly less snow for each model, relative to their retrospective SWE results. RCM had less SWE relative to PCM except in the northern tip of the basin, where the RCM BAU was both colder and wetter than the PCM BAU in December.

3.2.2 Monthly basin average and streamflow analyses

In the monthly analysis of basin average variables (Figure 3.12), the BAU RCM and PCM with BCSD approaches both had increased spring and decreased summer precipitation, although PCM precipitation was greater than RCM, except in fall. BAU temperature increases were nearly identical for the two approaches, hence different hydrologic results for the basin-averages variables followed more for precipitation differences. Soil moisture and evaporation were higher in PCM-BCSD, moderating the effect of PCM's higher precipitation on runoff, which was only slightly higher for PCM than RCM. For both approaches, the peak runoff came about one month earlier, but for RCM-BCSD, volume also decreased. Basin-average BAU SWE declined relative to retrospective SWE, but without much difference between approaches (despite the spatial differences in Figure 3.11).

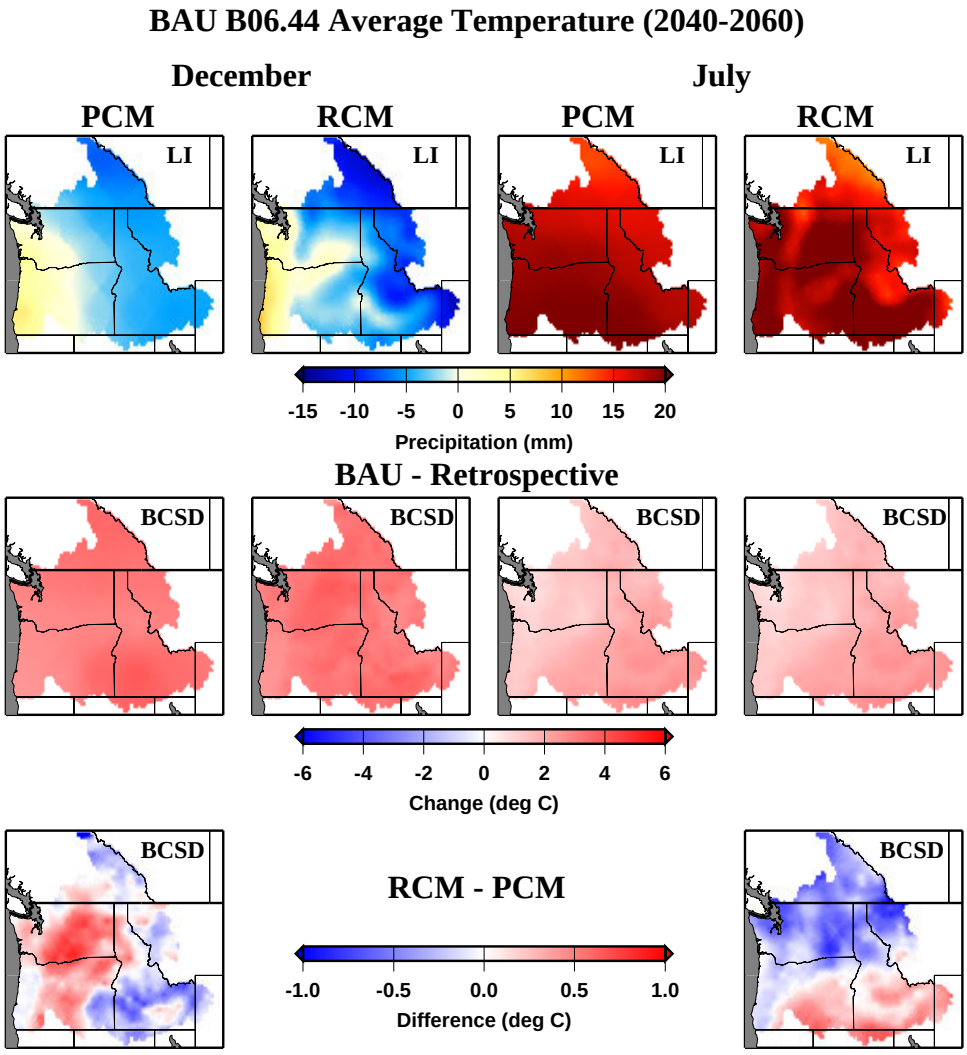

Figure 3.10 December and July average temperature for the PCM and RCM-driven BAU future climate simulations (2040-60). (top row) With LI only; (second row) differences in BAU BCSD results for PCM and RCM from their retrospective BCSD results; (third row) differences between RCM-BCSD and PCM-BCSD results.

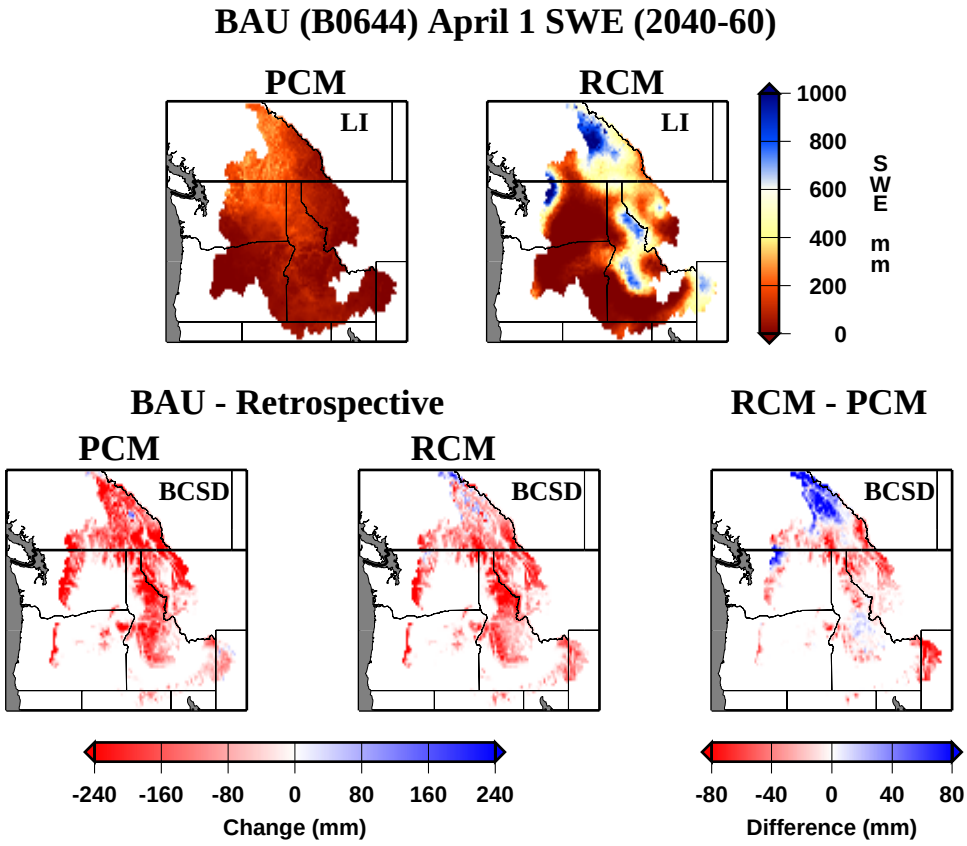

Figure 3.11 Average April 1 SWE simulation for the PCM and RCM-driven BAU future climate simulations (2040-60). (top row) LI method results; (second row, left) PCM and RCM BCSD differences between BAU and retrospective results; (second row, right) differences between RCM-BCSD and PCM-BCSD results.

Relative to the retrospective results, the BAU streamflow (Figure 3.13) shows an even larger seasonality shift toward higher winter-spring flows, and lower summer ones, from the observed climatology. Although equally shifted, the RCM BCSD streamflows are more sensitive to climate warming (with decreases in volume in addition to the shift) than the PCM-BCSD flows. This effect is exaggerated at Ice Harbor (near the mouth of the Snake River), where flow does not benefit from the RCM-BCSD climate's relatively higher April 1 snowpack in the Canadian headwaters of the CRB.

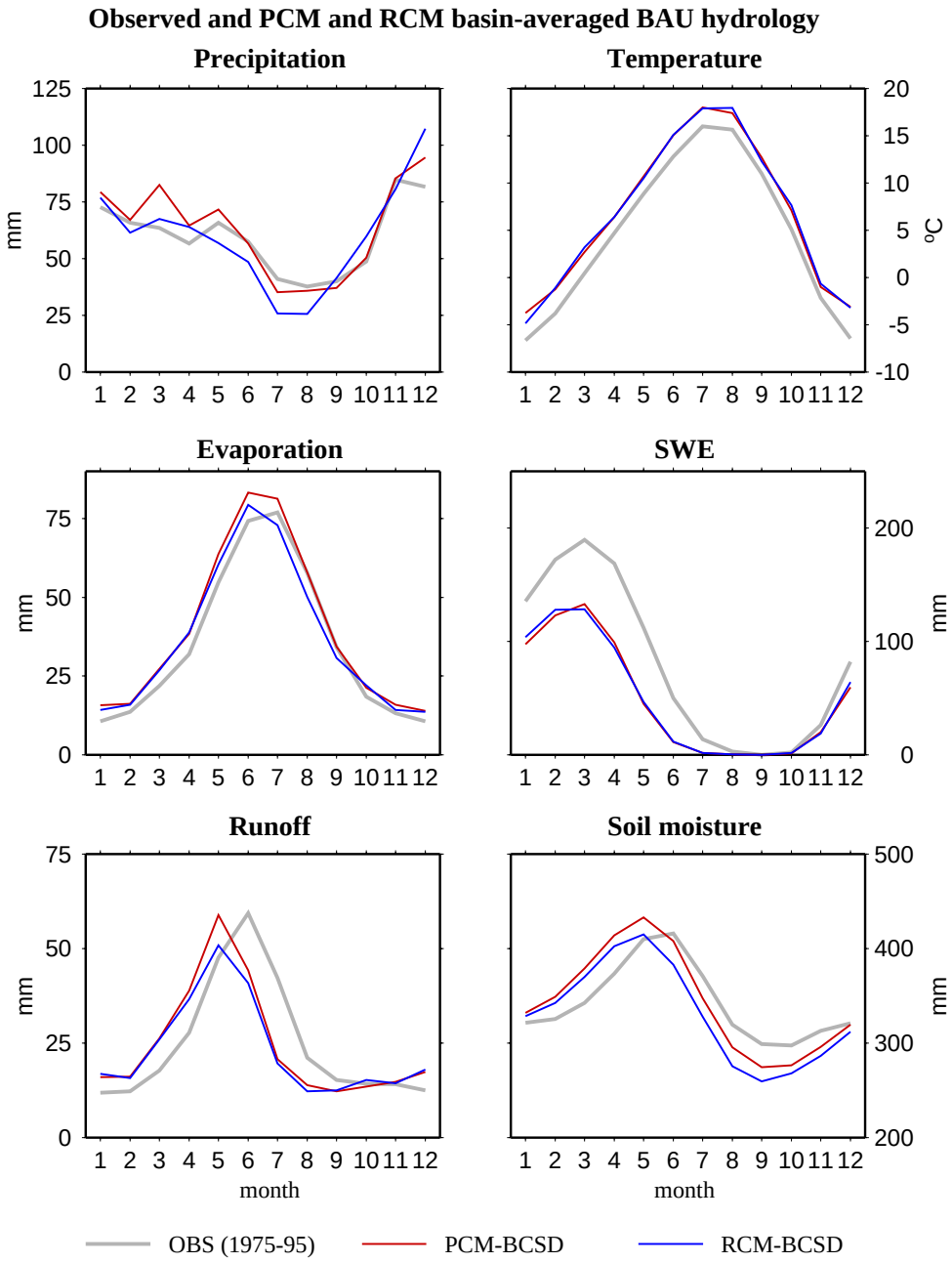

Figure 3.12 Columbia River basin averages of climate and hydrology variables for the PCM and RCM-driven BAU future (2040-60) climate simulations, compared with the observed 1975-95 climatology (i.e., observed precipitation and temperature, and simulated hydrologic variables based on these observations).

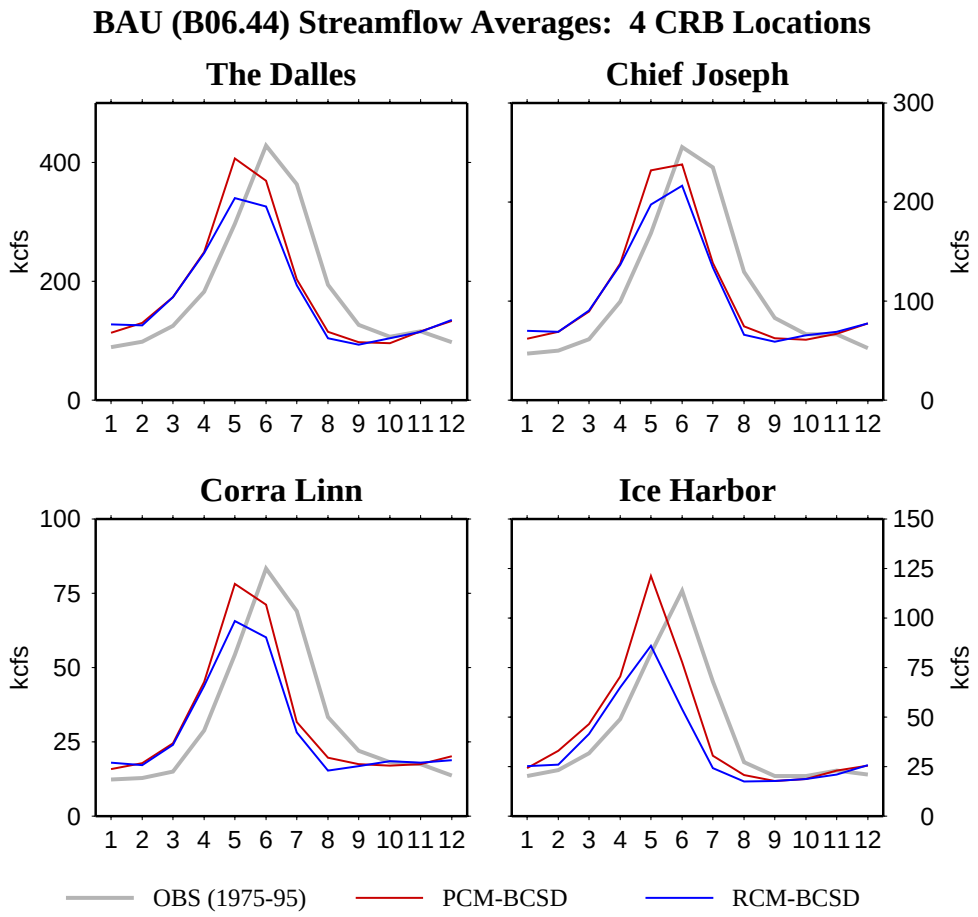

Figure 3.13 Streamflow at four locations (see Figure 3.1) for the PCM and RCM-driven BAU future (2040-60) climate simulations, compared with the observed 1975-95 climatology (simulated by the VIC model driven with observations).

4. CONCLUSIONS

The foregoing results of spatial analyses for December and July, and monthly analyses of basin averaged climate and hydrology variables and streamflow were chosen to characterize spatial and temporal differences arising from six different approaches to downscaling climate model output. We recognize that there are some difficulties in diagnosing hydrologic effects from one-month "summer" and "winter" snapshots of climate variables, even coupled with the continuous

monthly analyses of basin averaged variables. Nonetheless, we draw the following conclusions from the retrospective analysis:

With BCSD (in contrast to the two other post-processing choices), prior dynamical downscaling does not lead to large differences in the resulting hydrologic simulations.

- Linear interpolation of PCM or RCM output was insufficient to support plausible hydrologic simulations, even over large areas, despite the fact that RCM moderated PCM-derived hydrologic biases relative to the 1/8 degree observed climatology.

- If large-scale climate model outputs are relatively unbiased, applying spatial disaggregation (SD) to impose subgrid spatial variability improves hydrologic simulations, but substantial local biases will remain. After SD, for example, hydrologic (e.g., SWE and runoff) and streamflow simulations derived from the PCM output produced similar results to the finer scale RCM outputs after LI. Simply stated, correctly simulating basin mean precipitation and temperature is not sufficient to produce realistic hydrologic simulations in river basins of the western U.S. Conversely, correctly simulating the spatial distribution is also insufficient unless biases in the basin mean are removed first.

- If the climate fields are biased, SD alone may exacerbate biases locally (while leaving the basin average bias unchanged), particularly for precipitation. The result is that the downscaled climate variables may be unsuitable for use in hydrologic simulation.

Hydrologic simulation is sufficiently sensitive to biases in the basin mean and spatial distribution of precipitation and temperature at the monthly level, that nearly all local bias must be removed from climate inputs to achieve plausible hydrologic simulations. This is particularly true where seasonal snowpack transfers moisture input to the soil column and runoff from one season to the next. A primary conclusion of the retrospective study is that, although the BCSD method is successful in reproducing observed hydrology using biased climate model simulation outputs from both PCM and RCM, the monthly temporal scale used in correction of climate model precipitation and temperature, separately, fails to rectify more subtle differences between climate model simulation and observed climate. Interdependencies between precipitation and

temperature (for example, the frequency of wet-warm and wet-cold winters) are not addressed by the BCSD method, nor are the characteristics of seasonal distributions of precipitation and temperature arising from temporal autocorrelation in climate variables. Although the RCM may augment PCM in simulating these dynamics, after the BCSD method, the RCM and PCM retrospective hydrology simulations (with residual biases) were nonetheless nearly identical.

Like seasonal climate variability, interannual climate variability is only represented by the downscaling methods described in this paper via those characteristics that are directly transmitted to the downscaled values. For instance, while the methods ensure that the long term model-based monthly climatology (and to a large extent) the hydrology will resemble the observed hydroclimatology, it does not guarantee that the interseasonal or interannual sequencing of different climate regimes (e.g., wet/dry periods such as the 1988 drought or 1993 midwestern U.S. flooding) in a retrospective climate model simulation will be accurately simulated. In fact, extratropical interannual climate variability, particularly for precipitation, is not predicted well by GCMs (Lau *et al*., 1996), even given observed ocean boundary forcings, far less in free-running climate integrations that form the basis for many climate change studies. Given retrospective boundary conditions, however, many climate models (PCM included, as noted in Zhu *et al*., 2003) simulate long-term average annual and seasonal climate characteristics (for means and other statistics) reasonably well, which is the rationale for using these models in climate impact assessments. Approaches that combine climate model estimates of changes in characteristics which climate models simulate well with observation-derived information about poorly simulated climate characteristics (e.g., interannual variability, subgrid spatial variability) may be a fruitful area for future investigations.

When applied to the future (BAU) climate scenarios, the BCSD method yielded the only consistently plausible streamflow simulations, whether or not dynamical downscaling was also used. A significant conclusion, however, is that dynamical downscaling of the climate model scenarios before applying the BCSD method yielded results showing greater hydrologic sensitivity to climate change in the CRB than PCM-BCSD without dynamical downscaling. The RCM's initial biases in the spatial simulation of temperature and precipitation that were removed for the retrospective scenarios appeared to provide tendencies that produced the model differences for the BAU climate. Charles *et al*. (1999) note that for climate change assessments,

the inclusion in statistical downscaling approaches of an atmospheric moisture prediction variable (in this case, one other than surface precipitation -- perhaps one more confidently simulated by climate models) can lead to convergence in the results of statistical and dynamical approaches.

The greater hydrologic sensitivity found using RCM-BCSD compared to PCM-BCSD may imply that the RCM has a role to play in climate change analysis. The hydrologic differences are the combined result of differences between the PCM and RCM simulated warming signals, and differences in their precipitation characteristics. Leung *et al*. (2004) showed larger warming in the BAU scenario at the higher elevations that may be associated with snow-albedo feedback effects (a dependence also found in observations by Beniston *et al*., 1997, and other regional climate simulations, including those of Giorgi *et al*., 1997; Leung and Ghan, 1999; and Kim, 2001). These effects can only be simulated using a fully coupled land-atmosphere model: one way coupling of large scale climate models with high resolution hydrologic models cannot recover the effects of the missing regional scale climate change signatures. Where such regional signatures can be shown to be important (*and* can be accurately represented) in a subgrid scheme, they argue for higher resolution modeling or subgrid treatments in fully coupled land-atmosphere models for the study of climate change effects.

IV. WESTERN U.S. HYDROLOGIC FORECAST APPLICATION AND RETROSPECTIVE SKILL ASSESSMENT

This chapter has been submitted to the *Journal of Climate* (as Wood, A.W., A. Kumar and D.P. Lettenmaier, 2004: A retrospective assessment of NCEP GSM-based ensemble hydrologic forecasting in the western U.S.) in the form reproduced below.

1. INTRODUCTION

Managed water resources systems are designed to protect water supply and other water use objectives from the vagaries of hydrologic extremes, namely flooding and drought. Although constructed storage and diversification of supply sources (e.g., surface water and groundwater, linked from various locations) are the primary means for providing reliable water supplies, water managers increasingly are finding that traditional approaches to meeting reliability targets via construction of storage are so costly that more advanced means of managing water resources must be considered as well. Primary among these are more accurate prediction of supply and demand over a range of lead times. Perhaps the most basic supply (i.e., streamflow) prediction is the forecast of historically observed streamflow averages (the deterministic "climatology forecast"). Greater forecast skill, however, has long been provided by methods ranging from simple regression or index methods (occasionally in graphical form, as demonstrated in Hall and Martinec, 1985), relating, for instance, early spring snow depth and/or streamflow with summer streamflow volumes (e.g., Huber and Robertson, 1982; Lettenmaier and Garen, 1979) to elaborate mathematical and statistical time series modeling approaches (Box and Jenkins, 1976). The early 1970s saw the introduction of deterministic, conceptual computerized hydrologic models (see Linsley *et al*., 1975, for an overview), which led to the development of streamflow forecasting models currently run by the National Weather Service (NWS) in a probabilistic framework called Extended (or more recently, "Ensemble") Streamflow Prediction (ESP: Twedt *et al*., 1977; Day, 1985).

In the last several decades, advances in understanding of systematic land-atmosphere-ocean interactions have helped provide the basis for seasonal prediction of precipitation and temperature, two key drivers of land surface hydrology, at seasonal lead times (Goddard *et al*., 2001 provides a comprehensive review). The important role of the thermal inertia of the world's oceans in determining continental climate, primarily through its control of synoptic atmospheric systems and the associated strength and path of moisture transport from the oceans onto the land, is well enough understood that seasonal climate forecasts based on such dynamics are now widespread (Changnon, 1999; Barnston *et al.,* 1994; 1999; Goddard *et al*., 2001; Latif *et al*., 1998). Perhaps the most recognized of these climate-determining "teleconnections" is the El-Nino Southern Oscillation (ENSO), which has been found to have a robust influence on North American climate (Trenberth, 1997; Livezey *et al*., 1997; Piechota and Dracup, 1996), although other ocean-atmosphere dynamics such as the Pacific Decadal Oscillation (PDO) are also thought to moderate ENSO effects (Mantua, 1997). A combination of statistical and dynamical methods for forecasting ocean temperature and associated atmospheric effects has been adopted at a number of operational and research weather and climate centers (e.g., National Centers for Environmental Prediction, NCEP; NASA Seasonal to Interannual Prediction Project, NSIPP; and the International Research Institute, IRI), resulting in the evolution of an operational capability for seasonal climate forecasting. At NCEP, for example, operational seasonal forecasts are based on a merging of forecasts based on several statistical methods in addition to ensemble output from the Global Spectral Model (GSM), an ocean-land-atmosphere general circulation model (OAGCM), the current version of which is called the Seasonal Forecast Model (SFM: Kanamitsu, *et al*., 2002).

Despite the operational availability of seasonal climate forecasts, hydrologic applications have largely been restricted to research settings and for the most part have not found their way into operational procedures (with the possible exception of modifications to ESP at NWS described by Perica *et al*., 2000). Aside from questions concerning climate model forecast skill, downscaling climate model output to the river basin scale has been a major obstacle. A wealth of literature (e.g, Wilby *et al*., 1997; 1998; Murphy, 1999) addresses the downscaling problem from an atmospheric science perspective, but fewer studies have addressed the challenges of downscaling climate model outputs for the purpose of hydrologic simulation. Hay *et al*. (2002), for example, compared the suitability of statistical and dynamical approaches for downscaling climate model

simulations. Wilby *et al*. (2000) performed a similar study, using NCEP/NCAR reanalysis simulations to derive hydrologic model forcings for the Animas River basin in Colorado. More recently, Clark *et al*. (2004) suggests the use of non-parametric methods to restore plausible spatial and temporal structure to climate model-based ensemble forecasts where it is lacking. Shaman *et al*. (2003) produced seasonal surface wetness forecasts by sampling an observed climatology to reflect distributional shifts in probabilistic climate forecasts.

Both Clark *et al*. (2004) and Shaman *et al*. (2003) resolve the downscaling problem by imposing observed local climatological structure onto climate model forecast signals, using resampling approaches to generate daily or shorter weather sequences. This general strategy was also adopted in Wood *et al*. (2002; reproduced as Chapter II), who surmounted the downscaling obstacle using a climate-model scale monthly bias-correction step followed by statistical spatial and temporal disaggregation. The process interprets climate model forecasts (of precipitation and temperature) relative to their respective climatological probability distributions (estimated from retrospective climate model simulations), which allows a mapping to the observed climatology in a way that eliminates most or all temporal and spatial climate model bias at the daily to monthly time scale. Using 20 years of a historical climate model simulation, Wood *et al*. (2004) demonstrated that this statistical downscaling approach compared favorably with dynamical downscaling in that the same retrospective simulation downscaled via a regional climate model was found to require additional bias-correction to yield similar results – a result consistent with the findings of Hay *et al*. (2002). In the East Coast U.S. summer 2000 drought example evaluated by Wood *et al*. (2002), hydrologic forecast skill arose mostly from predictability in the evolution of initial hydrologic conditions (primarily soil moisture); whereas an El Nino example appeared to show that additional skill accrued from the climate model forecasts. The results of Wood *et al*. (2002) were mostly qualitative, but nonetheless suggested that a more quantitative assessment of the climate model-based hydrologic forecasting approach was warranted.

In this paper, we apply the approach of Wood *et al*. (2002) over the western U.S. and determine the extent to which a hydrologic forecasting approach based on downscaling of global ensemble climate forecasts can yield hydrologic prediction skill in excess of that achievable using climatological meteorology forecasts. Using a 21-year GSM retrospective forecast (also termed "hindcast") dataset, we compare the GSM-based hydrologic forecasts with climatological

ensemble forecasts and with ESP forecasts. Because recent work suggests that the forecast skill of the NCEP seasonal (climate) forecasting model is higher in strong warm and cold ENSO periods than under ENSO neutral conditions (Kanamitsu *et al*., 2002), we also assess composites of GSM- and ESP-based forecasts that have strong ENSO anomalies in the forecast initiation month. Our conclusions focus on streamflow at selected locations throughout the western U.S. and basin-averages of monthly climate and hydrologic variables for the five USGS hydrologic regions in the western U.S. (Figure 4.1).

2. APPROACH

We created retrospective six-month lead forecasts of hydrologic variables with a macroscale hydrologic model driven by land surface variables derived from a) a global OAGCM, after downscaling; and b) observed meteorology, via the ESP method. The retrospective forecasts span the period 1979-1999, and are for four initiation months – January (JAN), April (APR), July (JUL) and October (OCT). GSM and ESP forecast skill for monthly basin-averaged variables -- temperature and precipitation, snowpack, runoff and soil moisture -- and for streamflow is evaluated relative to the skill of forecasting the climatological distribution of each, taken from a retrospective observational analysis. Note that the hydrologic model was forced by spatially distributed land surface variables (primarily precipitation and temperature) at the 1/8 or 1/4 degree spatial resolution; hence the assessment of forecast skill for basin-averaged forcings and derived hydrologic variables is a convenience to provide additional summary insight. From an operational standpoint, the primary focus would be streamflow forecast accuracy.

2.1 Observational Analysis

The western U.S. domain of the study was divided into five major regions (Figure 4.1): the Pacific Northwest (PNW), the major river basin within which is the Columbia River basin (CRB); California (CA), containing the Sacramento-San Joaquin River basin (SSJB); the Colorado River basin (CORB), the Great Basin (GB) and the upper Rio Grande River basin (RGB). Over this domain, the observation-derived surface forcings and derived hydrological fields used to assess the downscaled climate model outputs were created as described in Maurer *et al*. (2002), which presents a 1/8 degree meteorological analysis and associated hydrologic simulation of land surface energy and water variables run at a 3 hour time step over the continental US (as well as

part of Canada and Mexico) for the period 1950-2000. The 1/8 degree climate variables (daily precipitation, minimum and maximum temperatures and wind speed) for the 21-year retrospective hindcast assessment period (and a 2-year prior hydrologic model spin-up period) were taken directly from the Maurer *et al*. (2002) dataset, and the hydrologic variables were generated via retrospective daily hydrologic simulation driven by the climate variables. The ¼ degree climate variables (used for simulations in the larger CRB, CORB and GB) are aggregated from the 1/8 degree climate variables.

The hydrologic model (section 2.2.1) was calibrated and validated using observed naturalized streamflows (water management effects removed), to the extent that these data were available. For the "observational" verification dataset (the baseline for assessing hindcast ensemble performance), however, we used simulated historic streamflow, based on inputs of the retrospective forcing dataset. This choice eliminated the confounding effects of streamflow simulation errors, and provided consistency with the use as a baseline of simulated historical values (from the retrospective hydroclimatology) for spatial variables (e.g., soil moisture or grid-cell runoff), observations for which are, to our knowledge, nonexistent.

2.2 Models and Simulations

2.2.1 Hydrologic model

The Variable Infiltration Capacity (VIC) model of Liang *et al*. (1994; 1996; see also Cherkauer *et al*., 2003) is a semi-distributed grid-based hydrological model which parameterizes the dominant hydrometeorological processes taking place at the land surface-atmosphere interface. For this study, the VIC model simulated the daily water balance (although with a sub-daily time step for certain moisture and energy-related calculations, and for snow simulation), and required as meteorological inputs (forcings) daily precipitation, maximum and minimum temperature, and daily average wind speed, for each model grid cell. Grid-cell runoff and baseflow was routed via a separate channel routing model to produce streamflow at selected points within the simulation domain. The VIC model has been implemented previously, at various resolutions, for the entire study domain (Figure 4.1) and details are available in the following references: CRB (Nijssen *et al*., 1997; Payne, *et al*., 2004); SSJB (Van Rheenen *et al*., 2004); CORB (Christensen *et al*., 2004); GB and RGB (Maurer *et al*., 2002).

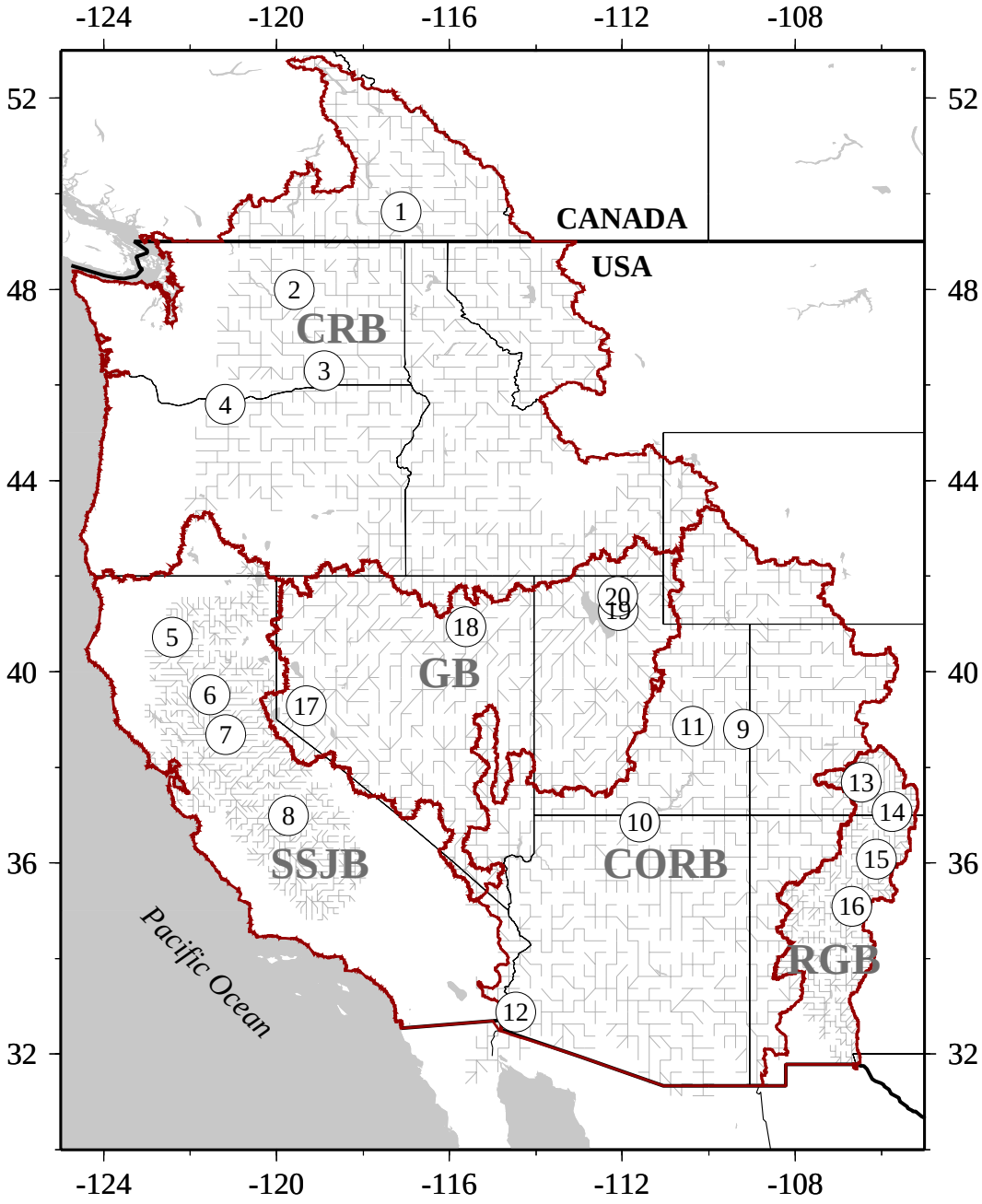

Figure 4.1 VIC hydrologic model calibration and streamflow forecasting sites selected for this paper. The numbers correspond to stations listed in Table 4.1; the heavy lines delineate the five basin areas (CRB, SSJB, CORB, GB and RGB) used for averaging climatic and hydrologic variables; and the light gray lines show the channel routing network used to transform VIC cell runoff to streamflow (only the network areas upstream of the stations were used).

Although the model forcing data span the period 1949-2000, the period of analysis used here was 1977-2000, which included the hindcast evaluation period plus two prior years (1977-78) for model spin-up. Model calibration was accomplished as a separate exercise, for most basins as part of the cited previous implementations (see references above for details). In general, the VIC model is calibrated by varying parameters related to infiltration and subsurface drainage, with the aim of reproducing monthly streamflow volumes while preserving the general features of the daily response (e.g., daily average flow peaks and recessions). Monthly validation results from each basin are shown in Figure 4.2 for five of the 25 streamflow forecast points (one in each basin) used in the study (listed in Table 4.1).

2.2.2 Climate model

The hydrologic forecasting approach used 6-month climate model ensemble forecast fields (monthly total precipitation and average temperature – $\mathbf{P_{tot}}$ and $\mathbf{T_{avg}}$) produced by the NCEP Global Spectral Model (GSM). GSM is one component of a forecasting system used by CPC to generate six-month lead forecasts of global surface precipitation and temperature, as well as other atmospheric variables. Each month, GSM generates a 20-member ensemble of 6-month lead climate forecasts. The 20 forecast ensemble members are produced by using 20 different atmospheric initializations with predicted SSTs in the tropical Pacific Ocean as of the date of the forecasts (Ji *et al*., 1998). The forecasts are accompanied by a 210-member ensemble of climate hindcasts (also six months long, matching the calendar months of the forecasts) for the period 1979-99 (21 years). The hindcast ensemble generation process is similar to that used to produce the forecasts, except that the ensemble members are produced by using 10 different atmospheric initializations with observed SSTs for each of the 21 years in the 1979-99 hindcast. GSM forecast spatial resolution is currently (as of 2003) T62 (approximately 1.8 degrees latitude/longitude).

For this study, hindcasts for JAN, APR, JUL and OCT were used to derive hydrologic hindcasts (10-member ensembles) in every year of the climatology period, 1979-1999. Note that because the GSM hindcasts were created using retrospectively analyzed SSTs, their skill represents the upper bound of the skill afforded using current SST-forecasting techniques (essentially perfect foresight of SST is used in this analysis). We are unaware of the existence of a retrospective hindcast dataset that uses forecasted SSTs (as is the case for the real-time forecasts).

Table 4.1 Streamflow forecast locations. Abbreviations in italics indicate that naturalized monthly flows were available for calibration.

# ABBREV	USGS # or CDEC[1] ID	NAME	Drainage Area (km^2)
1 *CORRA*	12322730	Kootenai R. at Corra Linn Dam, BC	45,843
2 *CHIEF*	12437990	Columbia R. at Chief Joseph Dam, WA	195,286
3 *ICEHA*	13352980	Snake R. at Ice Harbor Dam, WA	281,015
4 *DALLE*	14105700	Columbia R. at the Dalles, OR	613,830
5 *SHAST*	SHA (USBR[2])	Sacramento R. at Shasta Dam, CA	17,262
6 *OROVI*	FTO (CADWR)	Feather R. at Oroville, CA	9,386
7 *FOLSO*	AMF (CADWR)	American R at Folsom, CA	4,856
8 *SANJO*	11251000 / SJF	San Joaquin R. below Friant, CA	4,341
9 *GUNNI*	09152500	Gunnison R. at Grand Junction, CO	20,534
10 *SANJU*	09379500	San Juan R. near Bluff, UT	59,570
11 *GREEN*	09315000	Green R. at Green River, UT	116,162
12 *COLOR*	09429500	Colorado R. below Imperial Dam, CA-AZ	477,855
13 *DELNO*	08220000	Rio Grande R. near Del Norte, CO	3,419
14 *LOBAT*	08251500	Rio Grande R. near Lobatos, CO	19,943
15 CHAMA	08290000	Rio Chama R. near Chamita, NM	7,884
16 RIOGR	08330000	Rio Grande R. at Albuquerque, NM	37,555
17 FTCHU	10312000	Carson R. near Ft. Churchill, NV	3,372
18 HELKO	10318500	Humboldt R. near Elko, NV	7,198
19 PLAIN	10141000	Weber R. near Plain City, UT	5,390
20 CORIN	10126000	Bear R. near Corinne, UT	18,205

[1]CA Department of Water Resources (CADWR) CA Data Exchange Center
[2]US Bureau of Reclamation

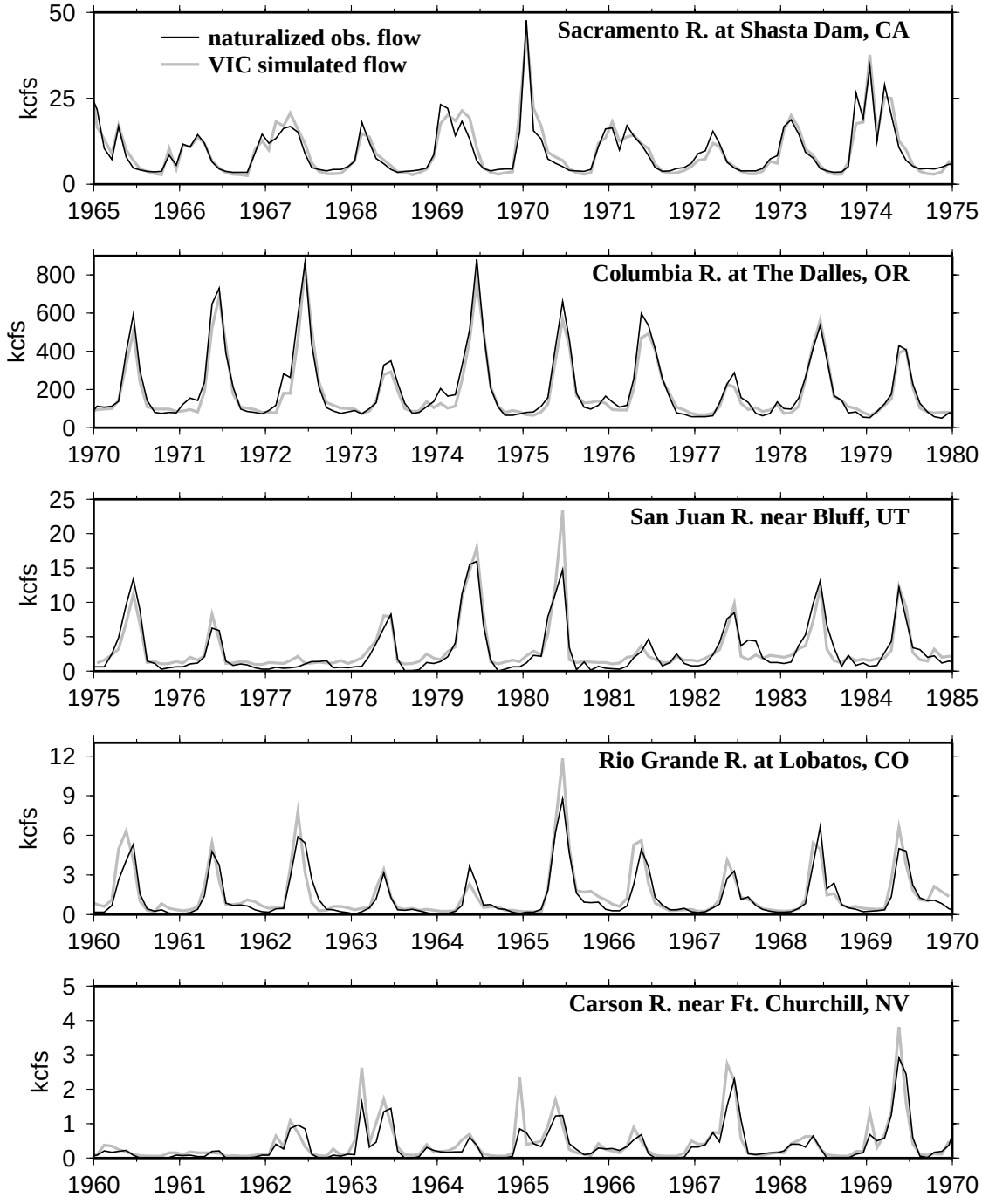

Figure 4.2 VIC monthly streamflow simulations and observations for a representative site in each major basin. Monthly averages of daily simulated flow time series are compared with naturalized monthly flows (first four locations) and with measured (i.e., unreconstructed) flows (for Carson R.).

2.3 Ensemble Forecast Methods

For each of the four forecast initiation months and for the five basins of Figure 4.1, we produced forecasts using two methods: ESP and the downscaling (via the method outlined in Wood *et al*., 2002) of GSM climate forecast ensembles.

For both forecast approaches, we initialized the hydrologic model with a 1.5-2 year spin-up simulation (using observed surface forcings) that yielded the current moisture states (in snowpack and soil) for the start date of the forecasts. Because the GSM forecasts are available near the beginning of each month but lack information for the initial month, the start date for each hydrologic forecast was 5 days before the beginning of the following month (GSM forecast month 1), which approximates the lead time needed to produce and process the hydrologic forecasts (ahead of month 1) if the procedure were implemented in real time. From the hydrologic forecast start date forward, the hydrologic model was forced with ensembles of surface variables derived via one of the two methods, producing daily output for the six month forecast period. Daily runoff and baseflow were routed to produce streamflow, and the streamflow and other variables were averaged to a monthly time step for skill evaluation.

2.3.1 Bias correction and downscaling of climate model (GSM) forecasts

The method adopted for producing daily hydrologic model inputs from climate model output, in this case monthly mean temperature ($\mathbf{T_{avg}}$) and total precipitation ($\mathbf{P_{tot}}$), is presented in Wood *et al*. (2002). In brief, monthly GSM outputs for each climate model cell centered within the study region were first bias-corrected on a cell-by-cell basis. Bias-correction was effected by evaluating the GSM ensemble hindcast variables ($\mathbf{T_{avg}}$ and $\mathbf{P_{tot}}$) as percentiles relative to the GSM model climatology, and then extracting the percentiles' associated variable values instead from the observed climatology -- a transformation described in Panofsky and Brier (1963). The observed climatology was the observational analysis (section 2.1) for the period 1979-99 (i.e., the retrospective climate model run period), averaged to the GSM grid resolution. The GSM climatology comprised the monthly $\mathbf{T_{avg}}$ and $\mathbf{P_{tot}}$ distributions from the GSM hindcast simulations for the same period. The mapping from GSM to observed climatology translated the hindcasts to

a plausible range with respect to observations. The adjustments varied spatially at the GSM grid scale (at which the mapping was performed) as well as by calendar month.

Following bias correction, the monthly GSM-scale forecast anomalies were disaggregated to the spatial and temporal scale of VIC model inputs via a simple statistical downscaling step. The monthly time step, bias-corrected GSM-scale hindcast $\mathbf{T}_{\mathbf{avg}}$ and $\mathbf{P}_{\mathbf{tot}}$ sequences were expressed as anomalies with respect to GSM-scale 1979-99 monthly means. The anomalies were then spatially interpolated to the 1/8 or 1/4 degree VIC cell centers and applied to the monthly observed 1979-99 1/8 or 1/4 degree cell means (from the observed climatology described in Section 2.2.1), to create monthly forecast sequences at the VIC model scale. A final temporal disaggregation step, the details of which are presented in Wood *et al*. (2002), was used to form daily time step inputs for the VIC model. Like the Clark *et al*. (2004) approach, the statistical and temporal disaggregation creates VIC scale daily forcing time series corresponding to the GSM-scale sequences, but exhibiting the VIC-scale spatial and temporal correlation characteristics.

2.3.2 Extended Streamflow Prediction (ESP) forecasts

The ESP method involves the estimation of the hydrologic state at the time of forecasts using hydrologic simulations driven by recent observed surface forcings up to the time of forecast, followed by the simulation of a hydrologic ensemble driven by an ensemble of forecast surface forcings over the forecast period. For consistency with the GSM hindcast set, we drew the ESP initiation dates and meteorological sequences from the same 21-year period spanned by the GSM hindcasts (1979-1999). JAN ESP hindcast ensembles, for example, were initialized in every January from 1979-1999, and each hindcast had 21 meteorological sequences or ensemble members. For an operational application of ESP, forcing sequences would likely be chosen based on other strategies, e.g., use of the longest period for which surface forcings could be estimated.

3. RESULTS

Results were assessed using a simple root mean square error (RMSE)-based skill score of the form:

$$SSrmse = 1 - \frac{RMSE(forecast)}{RMSE(reference\ forecast)}$$

The skill score format is described in Wilks (1995), and equals one for perfect forecasts, zero for forecasts no better than the reference forecast, and is unbounded below zero for forecasts that are worse than the reference forecast. Using the skill score, the GSM-based ensemble forecast skill was measured relative to reference forecasts consisting of (1) a climatological ensemble and (2) the ESP-based ensemble.

The climatological ensemble (CLIM) comprised the monthly distributions of the different climate and hydrologic variables, and streamflow, from the hindcast period, 1979-99, taken from the observational analysis of section 2.1. For the climate variables, $\mathbf{P_{tot}}$ and $\mathbf{T_{avg}}$, CLIM is equivalent to the unconditional ESP, while for the hydrologic variables and streamflow, the ESP ensembles benefit (unlike CLIM; like GSM) from knowledge of initial hydrologic conditions.

The statistical significance of the skill scores for streamflow and basin averages was assessed using a Monte Carlo procedure in which a 500-member distribution for each skill score was generated and the 95 percent confidence limits were noted (i.e., at p=0.025 and p=0.975). The skill score distributions are particular to each month, basin and variable, each streamflow site and statistic, and each forecast start date. Both conditional and unconditional forecasts skill scores were evaluated with respect to the unconditional confidence interval limits. The skill score probability distributions were generated by randomly varying the sequencing of the observational time series (using random resampling by year, with replacement) against which the forecast time series (e.g., for the basin averaged or streamflow variables) were verified. This procedure created a distribution of skill scores resulting from an experimental process that has no inherent skill, hence values higher or lower than the confidence limits should have skill that is unlikely to have arisen by chance. The Monte Carlo method was chosen because the small and varying sample sizes involved in the RMSE calculation precluded a straightforward analytical calculation of significance.

3.1 Unconditional Forecasts

The unconditional (i.e., using all 21 retrospective assessment years) forecast results are presented in Figure 4.3 and Table 4.2. Figure 4.3 shows the performance of the all-years GSM-based forecasts for basin-wide averages of monthly $\mathbf{P_{tot}}$, $\mathbf{T_{avg}}$, total runoff (**RO**), average soil moisture (**SM**) and snow water equivalent (**SWE**), relative to the CLIM ensembles (top) and to the ESP

ensembles (bottom), for each forecast initiation date. Based on the significance analysis, skill results that are not significant are hashed. Those which could not be calculated (i.e., for **SWE** in summer) are shown in gray. Table 4.2 summarizes the associated streamflow forecast skill results for cumulative flow in months 1-3 ($\mathbf{Q_3}$) and months 1-6 ($\mathbf{Q_6}$) of the forecast, where month 1 immediately follows the forecast initiation month. Statistically significant and **better** (worse) GSM skill scores are **bolded** (underlined).

As noted previously, the basin-averaged variable results are shown to afford a sense of hydrologic forecast performance at the regional scale, with length scales comparable those at which climate forecasts show significant spatial variation. Correspondence with the streamflow forecast results is expected when and where the skill of the basin average forecasts is very high, but otherwise, skillful basin average forecasts do not guarantee skillful streamflow forecasts. One reason is the discrepancy in the areas of the streamflow drainage basins and the five large reporting basins (particularly for some RGB, GB and SSJB drainages). Another factor is the different averaging periods used for streamflow and the basin averages (1-month versus 3- and 6-month), coupled with the monthly varying importance of variables in the hydrologic cycle for generating streamflow. For example, a **RO** forecast may be skillful in June and July but not May, but May have more **RO** than June and July together, leading to insignificant streamflow forecast skill for the May-July average. The basin-averaged hydrologic forecast performance is only suggestive of the potential for arbitrary streamflow forecasts within each region (perhaps better tailored in time and space than those presented here) to derive skill from the GSM climate forecasts.

3.1.1 JAN forecast (February-July)

With the exception of CRB February-March $\mathbf{T_{avg}}$ and CORB March $\mathbf{P_{tot}}$, the downscaled GSM JAN climate forecasts exhibit no significant skill. Nonetheless, significant hydrologic forecast skill exists at 4-6 month lead times for SM and SWE (entire domain), relative to CLIM (Figure 4.3a, top), as a result of initial conditions. For **RO**, forecast skill varies by basin: it is present at 5 month lead times for CORB and RGB, and several months for the other three basins, centered on their primary runoff season (e.g., April-July). GSM hydrologic forecast skill relative to ESP, in contrast, is much reduced (Figure 4.3a, bottom). For **SM** and **SWE**, CORB continues to show moderate skill at 4-5 month lead times. To a lesser extent, this is true of SSJB and GB as well, although significant skill at 1-2 month lead time is gone. GSM forecast **RO** skill is significantly

worse than ESP in the CRB (April) and RGB (February-March), but still better in a few months for SSJB (May-June) and CORB and GB (June).

The residual GSM CORB **SWE** and SM skill is attributable to large systematic errors in the ESP ensembles resulting from combination of several wet-warm Februarys (e.g., 1998) with above average initial **SWE** (e.g., 1979). The associated February melt elevates **SM** (raising ESP **SM** error) and also **RO**, although the effect on **RO** is buffered by the consequent restoration of winter soil moisture deficits. As a result, GSM **SWE** and **SM** forecasts are significantly better than ESP forecasts while **RO** forecasts are not significantly different.

Because runoff generation is a dynamic, non-linear process in which the sensitivities to precipitation and temperature are not independent, and vary by month, significant skill for $\mathbf{P_{tot}}$ and/or $\mathbf{T_{avg}}$ separately does not guarantee **RO** skill either in concurrent months or even at a lag. The degree to which GSM correctly simulates the observed joint distributions of precipitation and temperature (e.g., wet-warm versus wet-cold periods) can also be an important factor.

Results for JAN streamflow forecast skill (Table 4.2) are reasonably consistent with the basin average **RO** forecast skill maps. Relative to CLIM, CRB has significant skill at one site for $\mathbf{Q_3}$ flows, and two for $\mathbf{Q_6}$ flows, whereas the other basins have skill at 2-4 sites for $\mathbf{Q_3}$ flows, and all sites for $\mathbf{Q_6}$ flows. Compared to ESP, however, GSM streamflow forecast skill is mostly negligible or significantly worse (as in the CRB), with the main exception being several GB $\mathbf{Q_6}$ flow forecasts (which had RMSE reductions of about 10-20 percent relative to ESP). The latter is likely to be a result of a relatively skillful GSM June **RO** forecast.

Table 4.2 GSM unconditional streamflow forecast skill with respect to CLIM and to ESP, for four forecast initiation dates and two streamflow statistics, cumulative 3 and 6 month flows (wrt C: with respect to CLIM; wrt E: with respect to ESP). **Bolded** (underlined) values show **superior** (inferior) GSM skill that is statistically significant.

	JAN forecast				APR forecast				JUL forecast				OCT forecast			
	Feb-Apr		Feb-Jul		May-Jul		May-Oct		Aug-Oct		Aug-Jan		Nov-Jan		Nov-Apr	
river (site)	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E
1. Kootenai R (CORRA)	-0.37	-1.36	0.19	-0.88	**0.57**	0.07	**0.52**	-0.04	**0.35**	0.11	**0.21**	0.06	0.02	-0.44	-0.01	-0.51
2. Columbia R. (CHIEF)	-0.37	-1.63	0.16	-1.21	**0.62**	-0.03	**0.60**	-0.11	**0.45**	0.07	**0.27**	-0.01	-0.17	-0.60	-0.17	-0.65
3. Snake R (ICEHA)	**0.24**	-0.25	**0.51**	-0.23	**0.70**	-0.07	**0.70**	-0.03	**0.55**	-0.07	**0.33**	-0.12	**0.18**	0.01	**0.11**	-0.06
4. Columbia R. (DALLE)	-0.06	-0.19	**0.51**	0.00	**0.69**	0.03	**0.67**	-0.06	**0.55**	0.02	**0.34**	-0.06	-0.06	-0.39	-0.09	-0.49
5. Sacramento R (SHAST)	**0.33**	0.00	**0.37**	0.07	**0.68**	0.01	**0.66**	-0.05	**0.35**	**0.25**	0.19	0.17	0.06	0.06	0.09	0.07
6. Feather R. (OROVI)	**0.30**	0.01	**0.38**	0.08	**0.81**	0.03	**0.78**	-0.05	**0.56**	**0.33**	0.21	0.20	0.13	0.07	0.14	0.10
7. American R. (FOLSO)	**0.38**	0.00	**0.41**	0.06	**0.72**	-0.01	**0.71**	-0.04	**0.62**	**0.40**	0.23	0.22	0.18	0.18	0.14	0.11
8. San Joaquin R. (SANJO)	**0.34**	-0.09	**0.45**	**0.11**	**0.85**	-0.08	**0.83**	-0.13	**0.09**	-0.58	0.22	0.16	0.24	0.11	0.19	0.11
9. Gunnison R. (GUNNI)	-0.11	-0.02	**0.31**	-0.05	**0.76**	0.00	**0.74**	-0.04	0.27	0.11	0.15	0.08	**0.36**	-0.62	-0.28	-1.18
10. San Juan R. (SANJU)	**0.16**	0.05	**0.41**	-0.24	**0.82**	-0.17	**0.73**	-0.28	0.06	-0.14	0.12	-0.02	**0.32**	-0.16	**0.04**	-0.44
11. Green R. (GREEN)	0.32	0.14	**0.46**	-0.06	**0.68**	-0.01	**0.66**	-0.04	0.18	-0.14	0.19	-0.12	0.16	-0.21	0.14	-0.15
12. Colorado R. (IMPER)	**0.33**	0.09	**0.40**	-0.11	**0.75**	-0.05	**0.72**	-0.06	**0.31**	-0.12	0.16	-0.11	**0.34**	-0.60	0.03	-0.22
13. Rio Grande R. (DELNO)	0.19	0.09	**0.41**	-0.04	**0.83**	-0.12	**0.72**	-0.03	**0.35**	0.08	**0.39**	0.15	**0.72**	-0.54	**0.45**	-0.13
14. Rio Grande R. (LOBAT)	0.10	0.02	**0.45**	-0.02	**0.86**	-0.13	**0.76**	-0.09	0.28	0.04	**0.32**	0.10	**0.64**	-0.36	**0.37**	-0.16
15. Rio Chama R. (CHAMA)	**0.27**	-0.17	**0.24**	-0.24	**0.80**	-0.35	**0.69**	-0.52	-0.11	-0.23	0.13	0.05	**0.26**	-0.01	0.10	0.00
16. Rio Grande R. (RIOGR)	**0.12**	-0.30	**0.28**	-0.25	**0.81**	-0.29	**0.71**	-0.33	0.09	-0.11	**0.19**	0.04	**0.48**	0.00	**0.22**	-0.12
17. Carson R. (FTCHU)	**0.40**	0.16	**0.47**	**0.15**	**0.85**	**0.12**	**0.84**	**0.08**	**0.38**	0.19	0.32	0.25	0.31	0.25	0.28	0.19
18. Humboldt R. (HELKO)	**0.51**	0.12	**0.55**	0.04	**0.66**	0.02	**0.66**	0.02	0.29	0.21	0.03	0.00	-0.16	-0.30	0.11	0.12
19. Weber R. (PLAIN)	**0.47**	0.01	**0.56**	**0.12**	**0.64**	0.13	**0.65**	0.14	0.29	0.19	0.19	0.11	0.14	0.04	0.28	0.14
20. Bear R. (CORIN)	0.20	0.11	**0.48**	**0.16**	**0.72**	**0.08**	**0.72**	0.08	0.22	0.09	0.24	0.04	**0.32**	-0.10	0.11	0.14

3.1.2 APR forecast (May-October)

In the APR forecast (Figure 4.3b), climate variables ($\mathbf{P_{tot}}$ and $\mathbf{T_{avg}}$) again show little or no skill. The GSM hydrologic forecasts relative to CLIM (Figure 4.3b, top), however, have high skill levels for **SWE** for the remainder of the snow season in each basin: up to 6 months for CRB, 3 months (RGB), 2 months (CORB), and 1 month (SSJB and GB). **SM** forecast skill extends for all basins for 2-6 months, and **RO** skill for 2-5 months, longest for SSJB and shortest for GB. Relative to ESP (Figure 4.3b, bottom), GSM hydrologic forecast skill is again greatly diminished, although for several basins, despite the lack of climate forecast skill, the GSM forecasts have slight skill advantages in **SWE** and **RO** at 1-3 month lead times. SSJB, as in the JAN forecasts, has an extended period of significant skill (May-August), but this is followed by significantly worse skill than ESP. The high apparent significance of the GSM August CORB **SWE** forecast is likely to result from unstable statistics when **SWE** is near zero, and matters little hydrologically.

The GSM APR streamflow forecasts (Table 4.2) are, unsurprisingly, highly skillful at all sites for the $\mathbf{Q_3}$ and $\mathbf{Q_6}$ flows, relative to CLIM (primarily because of initial snowpack water storage). With respect to ESP, however, the only locations with significantly better GSM forecast skill are several GB sites. Elsewhere, GSM forecast skill is insignificant or significantly worse than ESP forecast skill, particularly in the RGB. The GSM GB streamflow forecast skill may be associated with June **RO** skill, but there is lack of streamflow forecast skill for CORB, even given better **RO** forecast results as compared to GB (i.e., better GSM **RO** for May as well as June). The negative results for RGB relative to ESP are also not clearly associated with **RO** results of Figure 4.3b.

In the CORB, the ESP **RO** forecast error is dominated by the observed climate anomalies from two ensemble years, the most extreme of which (1995) has an exceedingly wet-cold May-June. The meteorological sequences for this year (and the other), appearing in all 21 ESP ensembles, inflate the RMSE of **SWE** and **RO** in May and June, hence the better GSM forecast skill for these variables (note that the buffering soil moisture effect of the JAN forecasts is not present in APR, since initial soil moisture deficits are smaller than in winter). At the monthly level, total runoff and streamflow are highly correlated (albeit with some differences in large drainages related to channel routing). The GSM $\mathbf{Q_3}$ flow forecasts, however, exhibit extremes in some years that are

comparable to ESP $\mathbf{Q_3}$ extremes, resulting in part from differences between GSM and observations in the monthly autocorrelations of precipitation and of runoff anomalies (in this case, observed correlations for $\mathbf{P_{tot}}$ and **RO** from May to June are -0.22 and -0.66, respectively, whereas GSM correlations are 0.08 and -0.20). As a result, GSM streamflow sequences have a greater tendency to compound sequential anomalies in these high flow months, and the GSM monthly **RO** forecast skill advantages in June and July, separately, do not lead to a GSM forecast skill advantage in $\mathbf{Q_3}$ flows. Although we do not apply any here, approaches for disaggregating runoff volume forecasts (e.g., Pei *et al*., 1987) may be useful for adjusting biased monthly runoff autocorrelations.

The CORB **RO/SWE/**$\mathbf{Q_3}$ results would likely differ were a larger ESP ensemble used or a more restrictive screening of GSM forecast anomalies employed – that is, they appear to be determined in part by methodological choices (e.g., the period of analysis, the screening of precipitation outliers). The APR and JAN CORB forecast skill diagnoses exemplify the season-, basin- and flow site-specific variation in ensemble error characteristics that cloud the interpretation of some of the remaining results, particularly where the relative advantage of GSM or ESP forecasts is weak. We do not include here a diagnosis of every case, but rather suggest that the reader view only the stronger of the GSM forecast skill significance results (positive or negative) as being robust to potential methodological variations.

3.1.3 JUL forecast (August-January)

GSM JUL climate forecasts (Figure 4.3c) show significant skill for $\mathbf{P_{tot}}$ only in August (CRB) and $\mathbf{T_{avg}}$ only in November (CORB) and January (CRB). Because initial **SWE** is still present at higher elevations in CRB and SSJB, short lead forecast skill relative to CLIM (Figure 4.3c, top) is present in these two basins for **SWE** and **RO**. The most extensive forecast skill is for **SM**, ranging from 1 month lead for RGB to 6 months for CRB. With respect to ESP (Figure 4.3c, bottom), GSM hydrologic forecasts have skill that is at best barely better than the climate forecasts, and significantly worse skill at times, e.g., August **RO** for SSJB.

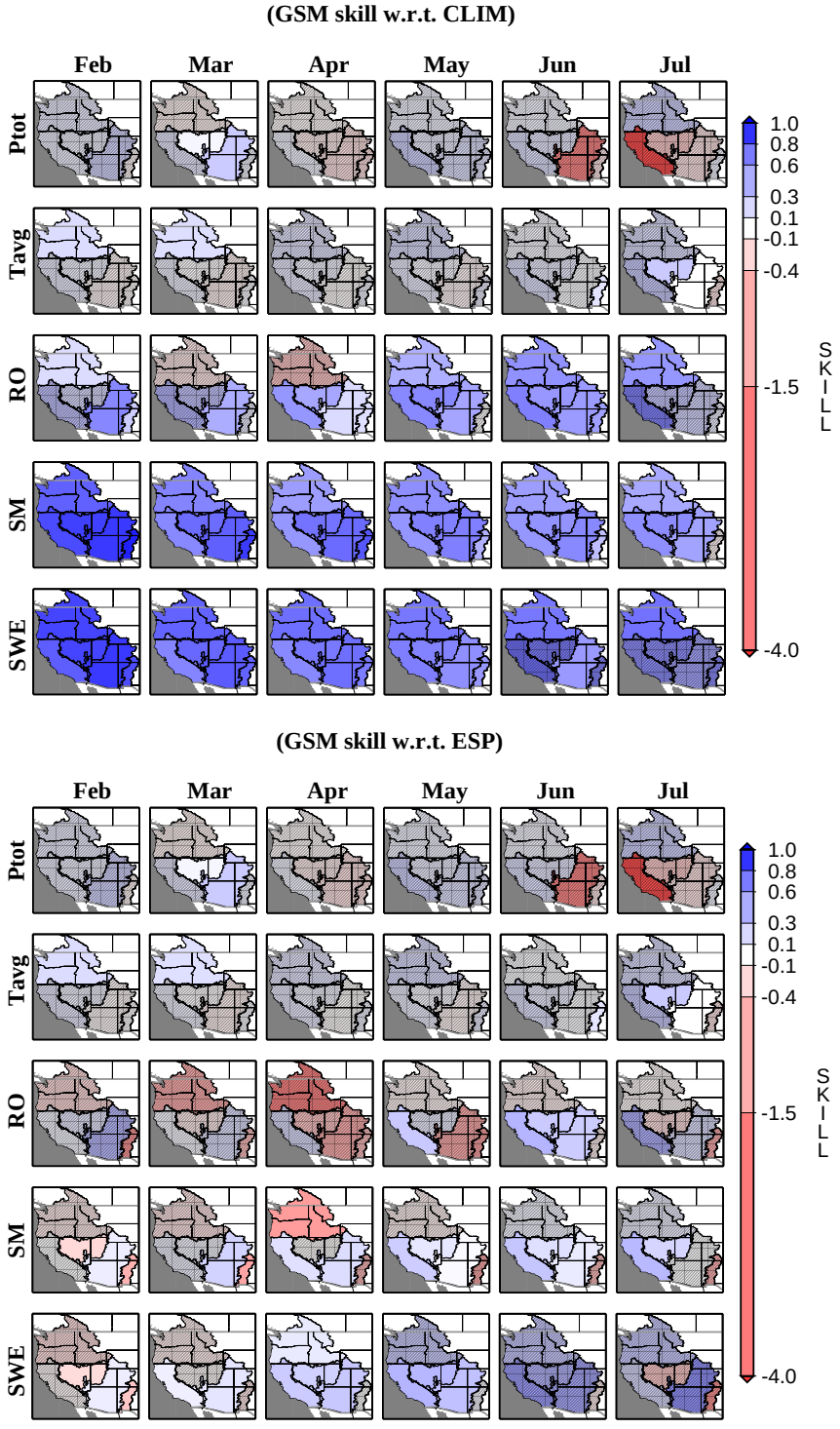

Figure 4.3a JAN GSM forecast average *SSrmse* (top: with respect to CLIM; bottom: with respect to ESP) for all forecast years in the period 1979-1999.

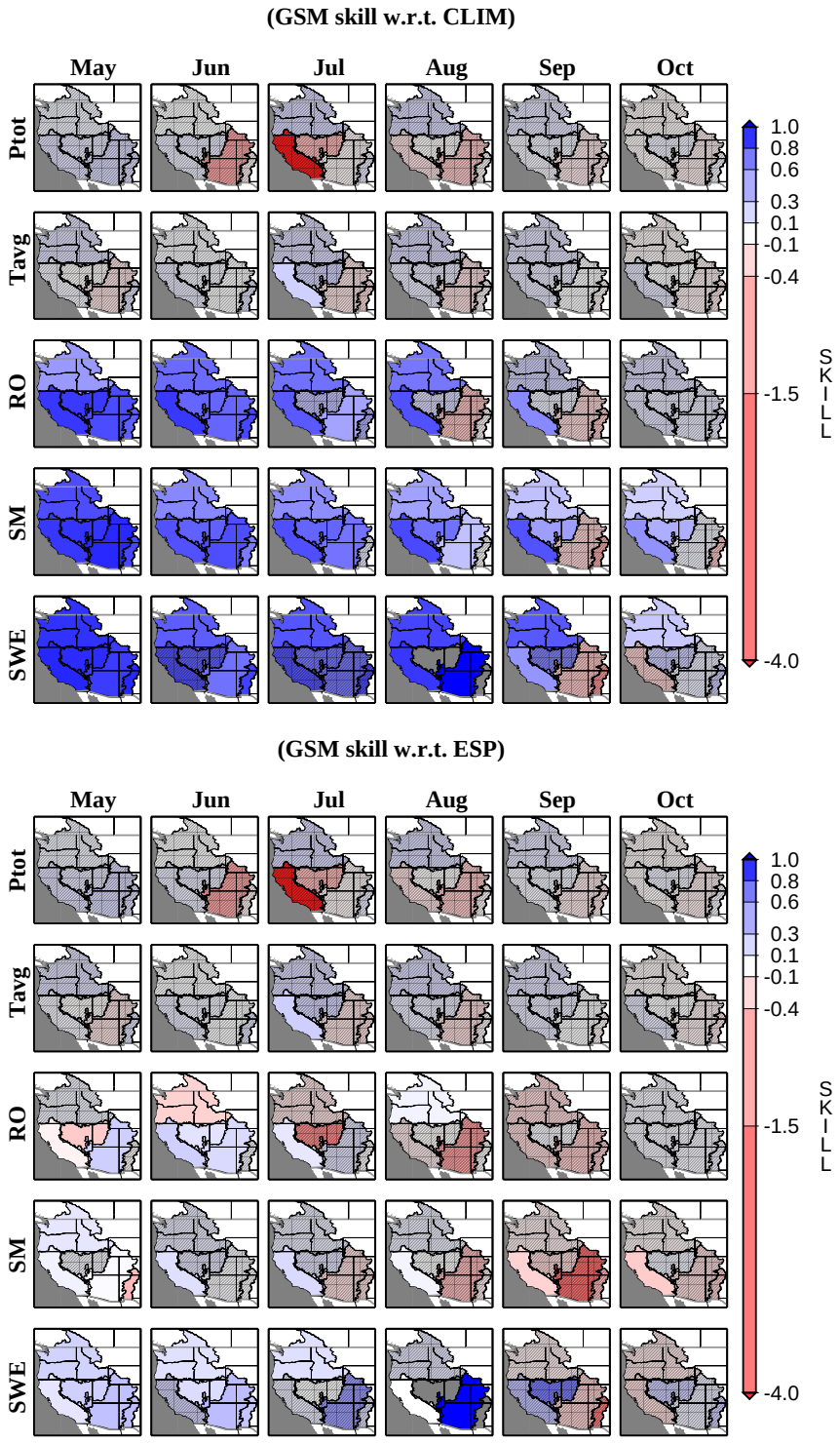

Figure 4.3b APR GSM forecast average *SSrmse* (top: with respect to CLIM; bottom: with respect to ESP) for all forecast years in the period 1979-1999.

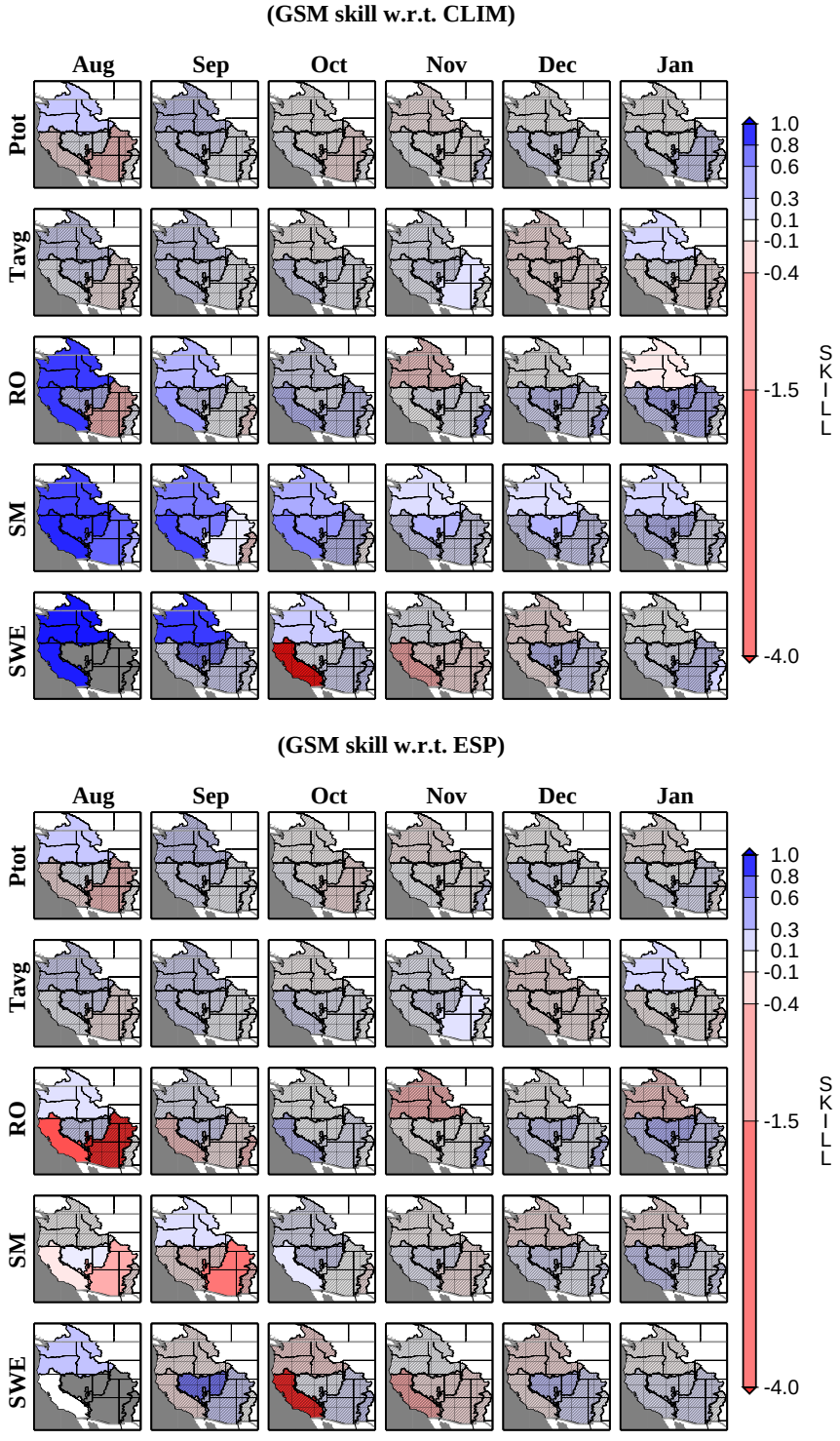

Figure 4.3c JUL GSM forecast average *SSrmse* (top: with respect to CLIM; bottom: with respect to ESP) for all forecast years in the period 1979-1999.

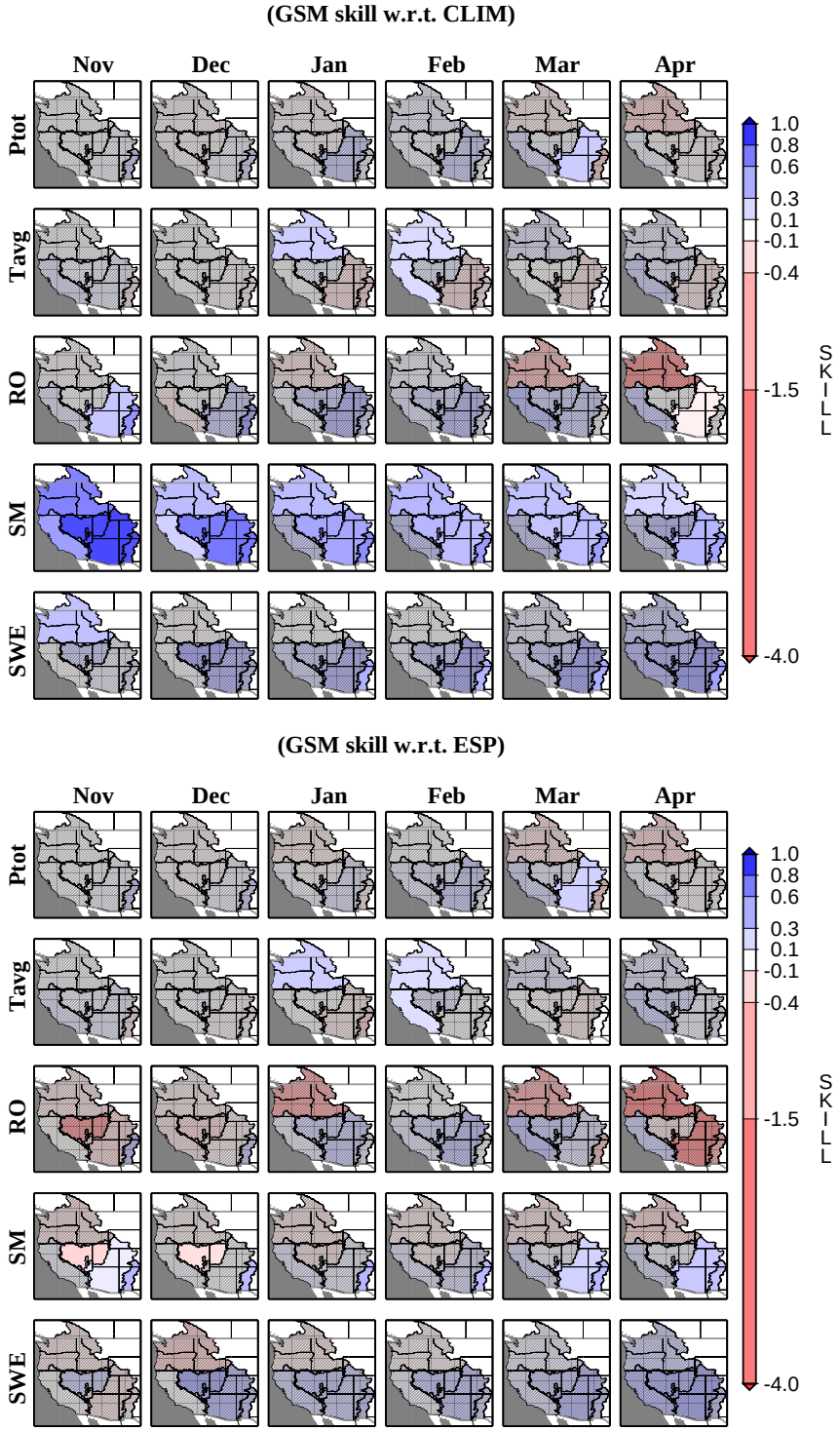

Figure 4.3d OCT GSM forecast average *SSrmse* (top: with respect to CLIM; bottom: with respect to ESP) for all forecast years in the period 1979-1999.

Relative to CLIM, the associated GSM $\mathbf{Q_3}$ flow forecasts (Table 4.2) show greater skill in the CRB and SSJB basins, while the RGB has greater skill for the $\mathbf{Q_6}$ flows only. These results appear to be consistent with the **RO** map, on which RGB has skill in November. With respect to ESP, the GSM $\mathbf{Q_6}$ but not $\mathbf{Q_3}$ flow forecasts show skill in SSJB, perhaps as associated with **RO** skill scores that are positive, despite not reaching significant levels.

3.1.4 OCT forecast (November-April)

The GSM OCT climate forecasts (Figure 4.3d) show generally no more skill than in the other forecast sets, excepting a slight winter $\mathbf{T_{avg}}$ (CRB) and March $\mathbf{P_{tot}}$ (CORB) skill advantage, both of which also appear in the JAN forecast. The notable feature of the hydrologic forecasts with respect to CLIM (Figure 4.3d, top) is the significant skill for **SM** at short and long leads, for most basins, while skill for other variables is negligible. Taken relative to ESP (Figure 4.3d, bottom), however, **SM** skill remains only for RGB (at all leads) and CORB (at 1 and 5-6 month leads), and the other variables again show no skill or significantly worse skill (e.g., GB **SM**).

With respect to CLIM, the GSM streamflow forecasts (Table 4.2) show little or no skill in the CRB, SSJB and GB, whereas the RGB flow forecasts are skillful for both $\mathbf{Q_3}$ and $\mathbf{Q_6}$ flows, and the CORB forecasts are skillful mostly for $\mathbf{Q_3}$ flow forecasts. Compared to ESP, however, the GSM flow forecasts show mostly negligible skill, and a few locations have significant negative skill. In this case, the streamflow results are reasonably consistent with those for areal **RO**.

3.2 Conditional (ENSO-defined) Forecasts

The conditional forecast results are presented in Figure 4.4 and Table 4.3. Figure 4.4 shows forecast performance for the strong ENSO years -- i.e., having an absolute Nino3.4 SST anomaly greater than 1.0 in the forecast initiation month -- for GSM-based forecasts for $\mathbf{P_{tot}}$, $\mathbf{T_{avg}}$, **RO**, **SM** and **SWE**, relative to the ESP ensembles (CLIM not shown), for each forecast initiation date. Out of 21 possible years, eight JAN forecasts were so classified, and four, five and six years were eligible in the APR, JUL and OCT conditional forecasts, respectively. Table 4.3 summarizes the associated streamflow forecast skill results for cumulative 3- and 6-month flows (denoted $\mathbf{Q_3}$ and $\mathbf{Q_6}$) beginning in the month following each initiation date. Conventions for showing significance are the same as in Figure 4.3 and Table 4.2.

3.2.1 JAN forecast (February-July)

The GSM JAN strong ENSO climate forecasts (Figure 4.4a, top) are better than the unconditional JAN forecasts, showing significant $\mathbf{T_{avg}}$ skill in most of the basin for a four month lead, with occasional skill in February-March $\mathbf{P_{tot}}$ (although improved skill in March is followed by negative skill in April, in CORB and RGB). The main changes in hydrologic forecast skill from the unconditional case (Figure 4.3a, bottom) include both skill improvements for **RO** (CRB May, SSJB Feb, GB May) and **RO** skill decreases (CORB Apr-May, GB Apr, RGB Mar-Jul). Also, **SM** skill improves in SSJB while worsening in CORB, and **SWE** skill increases most of the basins, but decreases in CORB (Mar) and RGB (May). Streamflow forecasts (Table 4.3), particularly for $\mathbf{Q_6}$ flows, improve in CRB and SSJB, but worsen in CORB and RGB, which is reasonably consistent with monthly **RO** forecast skill changes. Note that the conditional GSM and ESP ensembles, like the unconditional ensembles, had 10 and 21 members, respectively, but the number of ensembles used to calculate the skill score for the conditional forecasts was smaller, hence the results in this section are more vulnerable to sampling influences.

3.2.2 APR forecast (May-October)

Relative to the unconditional APR forecasts (Figure 4.3b), the GSM APR conditional climate forecasts (Figure 4.4a, bottom) show both better and poorer $\mathbf{P_{tot}}$ skill (varying by month and basin), but better late summer $\mathbf{T_{avg}}$ skill in a number of basins. Hydrologic forecast skill for **RO** and **SM** generally declines compared to the unconditional ensembles, except in isolated months for a few basins (e.g., CRB August and October **RO**). Conditional GSM **SWE** forecast skill is somewhat better in months 1 and 6, and generally worse at other times, than the unconditional forecast skill. The results for streamflow (Table 4.3) are mixed in each basin, and the most notable result is worsening forecast skill in CORB, RGB and GB, relative to the unconditional APR forecasts.

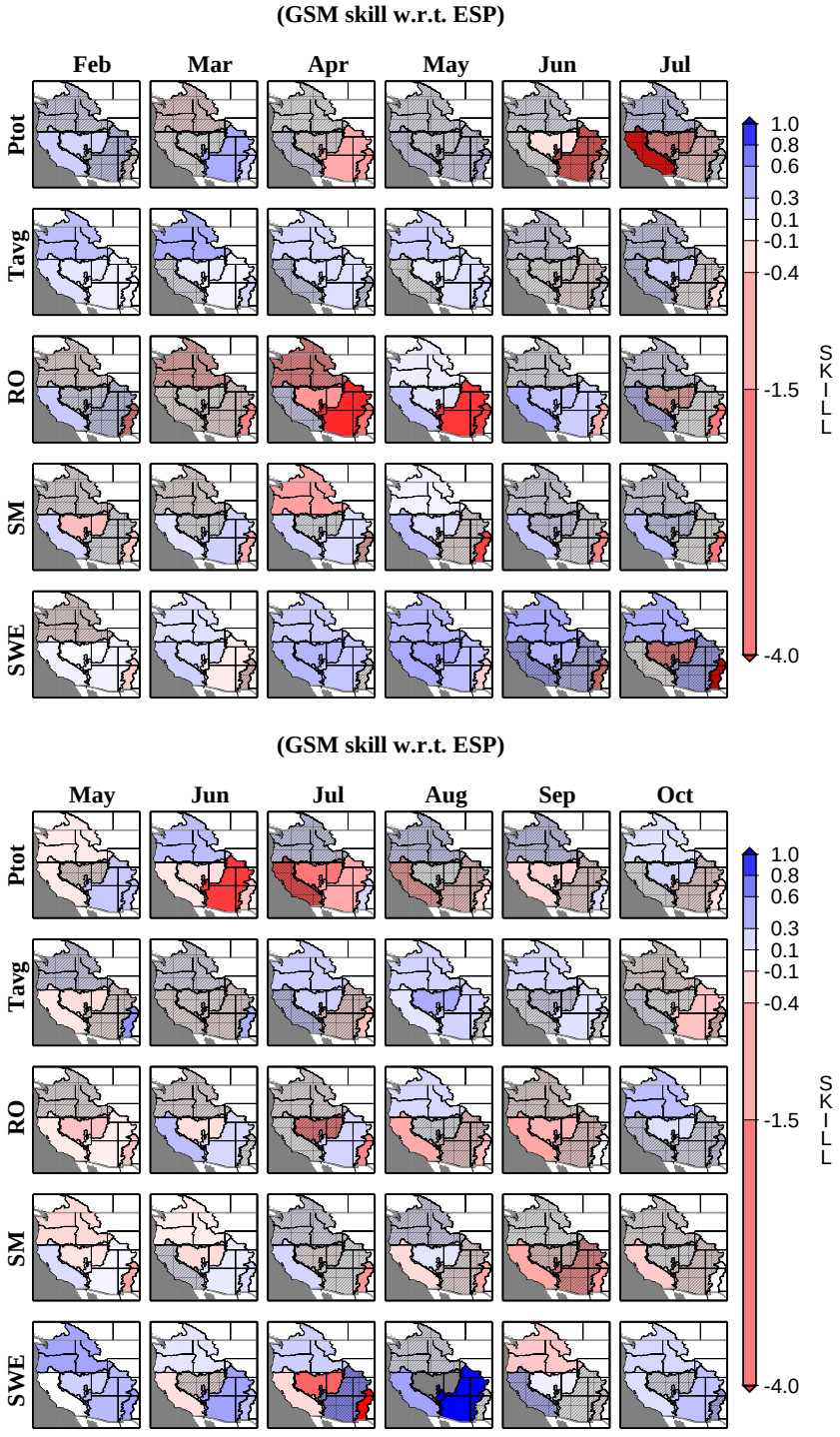

Figure 4.4a GSM forecast average *SSrmse* (with respect to ESP) for forecasts made when absolute NINO3.4 SST anomalies in the forecast month are larger than 1.0 degree Celsius. (top) JAN and (bottom) APR forecasts.

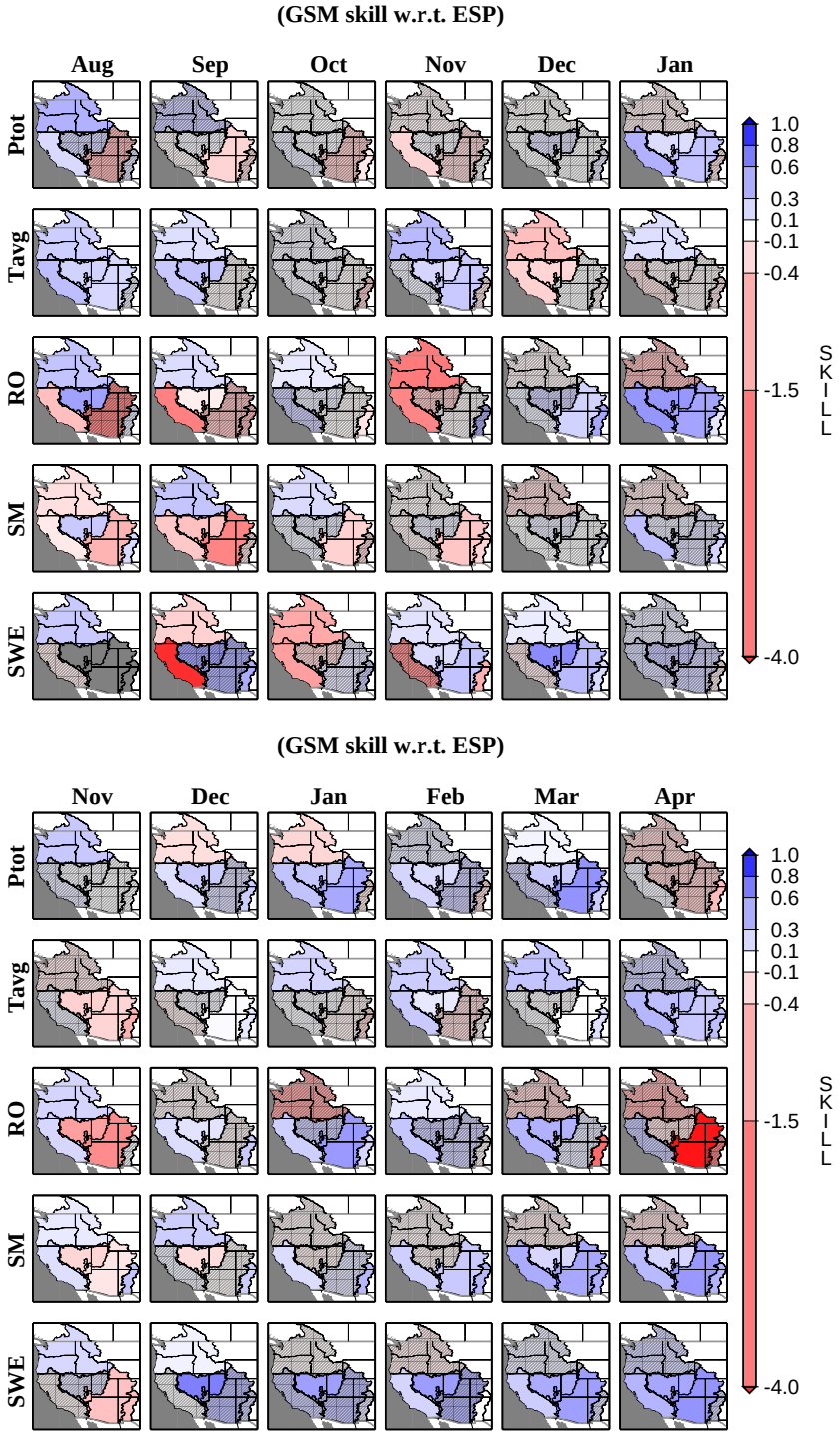

Figure 4.4b GSM forecast average *SSrmse* (with respect to ESP) for forecasts made when absolute NINO3.4 SST anomalies in the forecast month are larger than 1.0 degree Celsius. (top) JUL and (bottom) OCT forecasts.

Table 4.3 GSM conditional streamflow forecast skill (for strong ENSO years) with respect to CLIM and to ENSP, for four forecast initiation dates and two streamflow statistics, cumulative 3 and 6 month flows (wrtC: with respect to CLIM; wrtE: with respect to ESP). **Bolded** (underlined) values show **superior** (inferior) GSM skill that is statistically significant.

	JAN forecast				APR forecast				JUL forecast				OCT forecast			
	Feb-Apr		Feb-Jul		May-Jul		May-Oct		Aug-Oct		Aug-Jan		Nov-Jan		Nov-Apr	
river (site)	wrt C	wrt E	Wrt C	wrt E	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E	wrt C	wrt E
1. Kootenai R (CORRA)	-0.16	-0.51	**0.38**	-0.17	**0.66**	**0.19**	**0.57**	-0.02	**0.49**	**0.22**	**0.33**	0.10	**0.06**	-0.48	**0.11**	-0.48
2. Columbia R. (CHIEF)	-0.18	-0.51	**0.51**	-0.17	**0.68**	-0.03	**0.68**	-0.07	**0.56**	**0.19**	**0.38**	0.03	-0.07	-0.61	-0.01	-0.53
3. Snake R (ICEHA)	**0.26**	-0.16	**0.47**	-0.22	**0.68**	-0.22	**0.68**	-0.18	**0.60**	-0.03	**0.39**	-0.12	**0.36**	**0.10**	**0.28**	**0.08**
4. Columbia R. (DALLE)	-0.01	0.02	**0.49**	**0.19**	**0.73**	0.00	**0.72**	-0.05	**0.62**	**0.09**	**0.40**	-0.11	**0.08**	-0.33	**0.08**	-0.39
5. Sacramento R (SHAST)	**0.38**	0.02	**0.41**	0.10	**0.74**	-0.01	**0.68**	-0.14	**0.36**	0.15	**0.27**	**0.25**	0.14	**0.09**	**0.22**	**0.17**
6. Feather R. (OROVI)	**0.38**	0.02	**0.46**	**0.14**	**0.84**	-0.08	**0.78**	-0.24	**0.57**	**0.26**	**0.24**	**0.21**	**0.29**	**0.19**	**0.36**	**0.30**
7. American R. (FOLSO)	**0.48**	0.05	**0.50**	**0.15**	**0.86**	**0.06**	**0.85**	0.03	**0.39**	0.07	**0.28**	**0.26**	**0.33**	**0.33**	**0.31**	**0.26**
8. San Joaquin R. (SANJO)	**0.43**	-0.14	**0.56**	**0.20**	**0.93**	-0.01	**0.91**	-0.15	**0.16**	-0.41	**0.25**	**0.15**	**0.38**	**0.23**	**0.41**	**0.32**
9. Gunnison R. (GUNNI)	-0.03	-0.23	0.09	-0.23	**0.69**	-0.12	**0.68**	-0.15	**0.28**	0.04	**0.13**	0.01	0.16	-0.85	-0.53	-1.14
10. San Juan R. (SANJU)	-0.07	-0.26	0.19	-0.63	**0.73**	-0.29	**0.62**	-0.34	0.04	-0.26	0.08	-0.11	0.20	-0.24	-0.23	-0.78
11. Green R. (GREEN)	0.28	-0.03	**0.39**	-0.13	**0.68**	-0.20	**0.67**	-0.16	0.11	-0.14	0.17	-0.03	0.30	-0.11	0.04	-0.32
12. Colorado R. (IMPER)	0.25	-0.12	0.20	-0.34	**0.69**	-0.23	**0.68**	-0.16	0.19	-0.22	0.18	-0.16	**0.26**	-0.84	-0.04	-0.41
13. Rio Grande R. (DELNO)	0.17	0.05	0.14	-0.39	**0.77**	-0.30	**0.67**	-0.11	**0.33**	-0.06	**0.36**	-0.04	**0.70**	-0.53	**0.30**	-0.26
14. Rio Grande R. (LOBAT)	0.12	-0.03	0.22	-0.35	**0.82**	-0.32	**0.73**	-0.13	**0.31**	-0.10	**0.34**	-0.08	**0.60**	-0.47	0.17	-0.30
15. Rio Chama R. (CHAMA)	0.09	-0.37	-0.07	-0.61	**0.79**	-0.68	**0.66**	-0.58	-0.21	-0.40	0.03	-0.08	0.13	-0.14	-0.21	-0.13
16. Rio Grande R. (RIOGR)	-0.07	-0.56	-0.14	-0.75	**0.75**	-0.58	**0.65**	-0.38	0.06	-0.25	0.14	-0.15	**0.39**	-0.06	-0.16	-0.41
17. Carson R. (FTCHU)	**0.42**	0.08	**0.53**	**0.15**	**0.92**	**0.09**	**0.90**	-0.02	0.24	0.03	0.34	0.25	**0.39**	0.34	**0.38**	0.25
18. Humboldt R. (HELKO)	**0.55**	0.04	**0.57**	-0.09	**0.51**	-0.31	**0.53**	-0.35	0.16	0.10	-0.03	-0.07	-0.13	-0.35	**0.30**	**0.32**
19. Weber R. (PLAIN)	**0.62**	**0.14**	**0.62**	**0.18**	**0.65**	0.00	**0.65**	0.06	0.20	0.10	0.18	0.09	**0.31**	**0.16**	**0.48**	**0.34**
20. Bear R. (CORIN)	**0.19**	0.09	**0.47**	0.17	**0.76**	0.01	**0.74**	**0.10**	0.16	0.04	0.21	0.05	**0.33**	-0.19	0.08	0.16

3.2.3 JUL forecast (August-January)

The GSM JUL $\mathbf{T_{avg}}$ forecasts (Figure 4.4b, top) improve upon the unconditional JUL forecast (Figure 4.3c), and $\mathbf{P_{tot}}$ also has occasional skill increases, the most important coming in January, for SSJB, CORB and GB. Consequently, GSM **RO** forecast skill increases at short leads in CRB and in January for the other four basins. The **RO** improvements are somewhat represented in streamflow (Table 4.3), which has better conditional than unconditional forecast skill in the CRB for the $\mathbf{Q_3}$ flows, and particularly in SSJB for the $\mathbf{Q_6}$ flows. For CORB and RGB, however, streamflow skill diminishes, possibly as a result of negative **SM** skill in several months.

3.2.4 OCT forecast (November-April)

The conditional OCT GSM climate forecasts (Figure 4.4c, bottom) have generally better skill than the unconditional forecasts (Figure 4.3d), but this result varies by month and location: $\mathbf{T_{avg}}$ forecast skill improves, mainly in the CRB (where $\mathbf{P_{tot}}$ skill changes are mixed); and $\mathbf{P_{tot}}$ skill increases in SSJB, CORB and GB. In those basins, because better $\mathbf{P_{tot}}$ skill is found at times when precipitation is important hydrologically, **RO**, **SM** and **SWE** also improve in at least a few months. For streamflows (Table 4.3), the conditional forecasts particularly benefited SSJB (both $\mathbf{Q_3}$ and $\mathbf{Q_6}$ flows) and GB, and to a lesser extent CRB.

4. DISCUSSION AND CONCLUSIONS

The foregoing analysis was designed to assess quantitatively the added value of climate model forecast inputs from the NCEP GSM to a hydrologic forecast system for the western U.S., relative to ensembles of historical meteorology (ESP). Because previous work has indicated that GSM climate forecast skill is improved during extreme ENSO events, the analysis was extended to a strong-ENSO composite ensemble forecast. The assessment was based on forecasts of streamflow at 20 locations and of monthly basin averages for snow water equivalent (**SWE**), total runoff (**RO**) and soil moisture (**SM**).

Although a number of methodological issues (discussed below) bear on the results, the analysis supports the following general conclusions.

- The large discrepancy in GSM forecast skill when compared to with respect to CLIM (a baseline climatology forecast) and with respect to ESP (which benefited from the initial condition information) underscored the well known finding that substantial hydrologic forecast skill is possible based on the predictable evolution of initial hydrologic conditions. For **RO** and **SWE**, initial condition skill was highest in winter and spring, extending up to 6 months in some locations, and it was generally present throughout the year for **SM**.

- For streamflow, skill tended to mirror that of **RO** when **RO** forecast skill was high, but otherwise the varying monthly importance of runoff for streamflow generation complicated this association.

- The unconditional GSM climate forecasts did not generally have enough skill (particularly in precipitation) to improve on the skill of the ESP approach. Intermittent or minor skill improvements in some basins for **SWE**, **SM** and **RO** did not necessarily translate into streamflow forecast improvements.

- The conditional (ENSO composite) results verified that GSM climate forecast skill is enhanced in strong ENSO phases, particularly for temperature. For precipitation, results were mixed, with SSJB and CRB benefiting while CORB and RGB saw forecast performance become significantly worse. The superior skill for temperature in some locations led to better **SWE** forecast than were achieved via ESP, in some cases further translating into better **RO** and **SM** skill. Basin average **RO** forecast skill increases (e.g., SSJB) or decreases (e.g., CORB and RGB) for the conditional forecasts were large enough and temporally extensive enough to be reflected in streamflow (for better or worse).

That the climate model offers a significant skill improvement under certain conditions is encouraging, yet there are several reasons why the skill improvement may be less significant than the results of section 3 indicate. The assessment is unabashedly *a posteriori*, searching for skill in every possible location and month. Each combination of basin, forecast month, initiation month, and variable combination is taken to be independent -- a simplifying, but demonstrably false assumption. By ignoring interdependence, the levels of statistical significance are likely to be set

too low for any given combination, and the broad survey creates a greater chance that statistically significant results may arise spuriously. Also, the GSM forecast assessment is based on retrospectively produced forecasts (hindcasts) driven by observed SST fields. The actual forecasts must rely on forecast SSTs, which have inferior skill, thus the results likely overstate the skill of the GSM forecast, particularly at long lead times. Lastly, some of the skill discrepancy between the GSM and ESP forecasts may result from differing variance in the 10- and 21-member ensembles, as well as the vulnerability of skill scores based on the small ensembles to the influence of outlying meteorological sequences. This last concern is particularly relevant for the strong ENSO forecast results, which draw on fewer (but equally sized) ensembles. The GSM-based ensembles tend to have slightly lower variance due to the way they are downscaled, which would convey an advantage in the RMSE statistic that is the basis for the skill score. As noted earlier, we suggest a conservative appraisal of the results in which only the strongest are viewed as being potentially robust to methodological choices and misrepresentative sampling.

The lesson from this assessment is that the while climate model forecasts presently suffer from a general lack of skill, there appear to be locations, times of year and conditions for which they improve hydrologic forecasts relative to ESP, and (with careful screening to avoid the opposite possibility) could be useful for hydrologic forecasting. The routine lack of skill, however, is of sufficient concern that a pursuit of hydrologic forecast improvements (i.e., beyond ESP) using alternative climate forecasts (such as statistical and hybrid dynamical/statistical products) is likely to be more fruitful in the near term.

V. CONCLUSIONS AND RECOMMENDATIONS

The primary objectives of this dissertation were to develop, implement and evaluate a hydrologic forecasting approach based on the direct statistical downscaling of climate model ensemble forecasts for use in the production of hydrologic and streamflow forecasts at monthly to seasonal lead times. The dissertation's central premise was that such an approach, when used to link land-ocean-atmosphere and physical hydrology models, could integrate the potential forecast skill arising both from ocean-atmosphere teleconnections and from the persistence of initial hydrologic conditions. The evaluation of the hydrologic forecasting approach followed two main avenues. The first compared the performance of the downscaling method relative to dynamical downscaling with a meso-scale regional circulation model, an alternate downscaling method that demonstrably improves surface climate simulation, largely by dint of the regional model's superior resolution of land surface features on which local climate depends. The second avenue determined whether the climate model-based hydrologic forecasts improve upon current forecast methods, in light of recent indications that the climate model forecasts have superior skill during strong El Nino and La Nina events.

The main element of the hydrologic forecasting approach is the direct statistical downscaling of seasonal ensemble climate forecasts generated by a general circulation model (at spatial resolutions much coarser than those used by the hydrologic model) to the river basin scale. In Chapter II, a downscaling method was developed that employed probability mapping methods to transform monthly climate model resolution outputs to daily hydrologic model resolution inputs -- in this case, from the coarser 1.9 degrees spatial resolution of the NCEP Global Spectral Model's (GSM) to the 1/8 and 1/4 degree resolution of the Variable Infiltration Capacity (VIC) hydrologic model. The initial demonstration of the approach for the eastern U.S. (for the drought of summer 2000 and the El Nino event of winter 1997) showed that:

- the downscaling and disaggregation approach reproduced the observed streamflow climatology with only minor biases when forced with observed monthly climate variables

aggregated to the climate model scale (which confirmed that various simplifying assumptions incorporated in the spatial and temporal disaggregation were justified);

- hydrologic forecast skill derived from use of GSM ensemble forcings for the summer 2000 drought event appeared to derive largely from hydrologic initial conditions rather than from climate forecast skill; and

- the climate forecast contribution to hydrologic forecast skill during a strong ENSO (warm) episode (for November, 1997) was enhanced as compared the summer 2000 forecasts.

Although the findings of the eastern U.S. demonstration were mostly qualitative, the apparent success of the hydrologic forecasting approach argued for a thorough, quantitative assessment of the forecasts' skill, a challenge that was undertaken in Chapter IV.

The statistical downscaling strategy introduced and evaluated in Chapter II is simpler than most recently published downscaling methods -- either statistical or dynamical -- and is therefore well-suited to the context of ensemble forecasting. In Chapter III, further investigations of this strategy were undertaken to assess its performance relative to dynamical downscaling. The central analysis was based on a twenty-year retrospective (1975-1995) climate simulation produced by the NCAR-DOE Parallel Climate Model (PCM), although a future (2040-2060) PCM climate change scenario was also downscaled and evaluated. The statistical downscaling method developed in Chapter II (i.e., bias correction with subsequent spatial disaggregation) and two variations (linear interpolation and spatial disaggregation), each applied both with and without prior dynamic downscaling (using the Pacific Northwest National Laboratories Regional Climate Model, RCM), were evaluated. The major findings were:

- only the statistical method was successful in reproducing the main features of the observed hydrometeorology from the retrospective climate simulation, whether applied to PCM output directly or with the intervening step of dynamical downscaling;

- linearly interpolated RCM output (i.e., without further downscaling or bias-correction) resulted in unacceptably large biases in the hydrologic simulations, although it

constituted a much more realistic climate forcing dataset than did linearly interpolated PCM output; and

- the climate sequences derived via the statistical downscaling method were more hydrologically useful than those generated by the alternatives evaluated, even if dynamical downscaling was used.

Although these conclusions are based on a multi-decade, continuous retrospective climate simulation, consideration of the Chapter III results in light of Chapter II suggests that they are valid for seasonal climate forecast applications as well.

For seasonal to interannual hydrologic forecasting, the rationale for forcing the hydrologic model with surface variables derived from climate model forecast fields is that climate model forecast skill should augment forecast skill derived solely from initial hydrologic conditions (as captured, for instance, in ESP methods and persistence-based forecasts). To distinguish the degree to which this is true with the GSM-VIC implementation of the forecasting strategy, the approach developed in Chapter II was implemented in Chapter IV over the western U.S., for a retrospective (21-years, 1979-1999) series of forecasts. The resulting GSM-based forecasts were evaluated relative to two baselines: a climatological forecast (i.e., treating the observed distributions of the variables as a forecast ensemble) and ESP forecast. From this work, it was apparent that the GSM-based forecasts provided significant skill for hydrological variables (and streamflow) relative to the climatological forecast, particularly in winter and spring. This result simply underscored the ability of the forecasting approach to benefit from knowledge of initial hydrologic conditions. With respect to ESP forecasts (a more stringent benchmark), in contrast, the GSM-based forecast skill increases were generally negligible when an unconditional use was made of the available GSM hindcast ensembles. For a composite containing only strong ENSO years, however, the GSM temperature forecasts appeared to have more skill in several months of each forecast period, throughout the western U.S. domain, but precipitation forecast skill increases were rarer, appearing mostly in isolated months, and at times offset by skill decreases. These differences yielded both statistically significant improvements and declines in streamflow forecasts relative to ESP, depending on the season and region. The region-specific findings were:

- the most convincing streamflow skill increases were for California (except for April forecasts), followed by the Columbia River basin (January, July and October forecasts) and the Great Basin (for forecasts made in October); yet

- for the Colorado and Rio Grande River basins, where GSM precipitation forecast skill was poorest (worse than climatology for the strong ENSO forecasts), streamflow forecast skill was significantly degraded.

Although the small sample of GSM and ESP forecasts on which the analysis was based made these results somewhat vulnerable to sampling misrepresentation and methodological choices, the general conclusion is that the GSM-based hydrologic forecasts lead to marginal skill improvements only during strong ENSO events, in parts of the western U.S.

The quantitative assessment of ensemble climate forecasts, used for seasonal hydrologic forecasting via a statistical bias correction and downscaling approach, shows that at best, the skill added relative to the current ESP forecasting approach is insufficient to justify its use operationally. Published climate forecasting skill assessments for other global forecast models implemented for seasonal climate forecasting suggest that these results, which were particular to GSM, may in fact be more general. Nonetheless, relative to weather forecasting, which has seen steady improvement in skill over the decades, operational seasonal climate forecasting is in its infancy, and there appear to be potential sources of predictability that are not fully exploited in the current generation of models. The results for strong ENSO event conditions are encouraging in this respect. As the skill of the climate forecasts improves in the future, the hydrologic forecasting approach presented here should serve as a useful framework for hydrologic forecast evaluation. A cautionary note, however, is that future assessments of climate-model based forecast approaches will encounter difficulties, similar to some of those that surfaced in this research, related to the limited availability of retrospective forecasts and small forecast ensemble size. The exploration of a wider range of skill assessment metrics (perhaps including non-parametric and linear error scores that are less vulnerable to small sample biases), however, may strengthen a future assessment's conclusions in the face of such obstacles.

In the near term, this work suggests the potential of two alternative avenues for improvements in seasonal hydrologic forecast skill beyond that achieved with the climate model ensembles. First,

given that most of the hydrologic forecast skill was derived from the predictable evolution of hydrologic initial conditions rather than climate forecast skill, it makes sense to exploit as much information as possible from observations that could help improve estimates of hydrologic state variables prior to and at the time of a forecast. One such approach (a logical extension of current practices) is to use the hydrologic model to assimilate newly available snow remote sensing products, coupled with better use of surface measurements (such as the SNOTEL network of snow water equivalent observations). Second, other operational seasonal climate forecast products (primarily hybrid, "consensus" products that combine dynamical and statistical techniques) exist that may have greater skill than the purely dynamical (climate) model forecasts. Although the hybrid forecasts' quasi-subjective, evolving nature and limited retrospective availability hinders any robust assessment of their error characteristics, their more skillful verifications to date are promising.

REFERENCES

Abdulla, F.A., Lettenmaier, D.P., Wood, E.F. and J.A. Smith, 1996: Application of a macroscale hydrologic model to estimate the water balance of the Arkansas-Red River Basin. *J. Geophysical Research*, 101(D3), 744-7549.

Anderson, E.A., 1973: National Weather Service River Forecast System, Snow Accumulation And Ablation Model, NOAA Tech Memo NWS HYDRO-17, U.S. Department of Commerce, Silver Spring, MD (NTIS Acquisition No. COM-74-10728), November.

Anderson, J.L. and W.F. Stern, 1996: Evaluating the Potential Predictive Utility of Ensemble Forecasts. *J. Climate* 9, No. 2, 260-269.

Barnston, A.G. and Y. He, 1996: Skill of Canonical Correlation analysis forecasts of 3-month mean surface climate in Hawaii and Alaska. *J. Climate* 9, October.

Barnston, A.G., Van den Dool, H.M., Zebiak, S.E., Barnett, T.P., Ji, M., Rodenhuis, D.R., Cane, M.A., Leetmaa, A., Graham, N.E., Ropelewski, C.F., Kousky, V.E., O'Lenic, E.A. and R.E. Livezey,1994: Long-lead seasonal forecasts -- Where do we stand? *Bull. Amer. Met. Soc.* 75, 2097-2114, November.

Barnston, A.G., Leetmaa, A., Kousky, V., Livezey, R., O'Lenic, E.A., Van den Dool, H., Wagner, A.J. and D.A. Unger, 1999: NCEP Forecasts of the El Nino of 1997-98 and its U.S. impacts. *Bull. Amer. Met. Soc.,* September.

Barry, J.M., 1997: Rising Tide: The Great Mississippi Flood of 1927 and How It Changed America, Simon and Shuster, New York, NY.

Beniston, M., Diaz, H.F. and R.S. Bradley, 1997: Climatic change at high elevation sites: An overview. *Climatic Change* 36, 233-251.

Betts, A.K., Chen, F., Mitchell, K. and Z.I. Janjic, 1997: Assessment of the land surface and boundary layer models in two operational versions of the NCEP Eta model using FIFE data, *Mon. Weather Rev.* 125, 2896-2916.

Beven, K.J. and M. Kirkby, 1979: A physically based, variable contributing area model of basin hydrology. *Hydrologic Processes*, 6, 279-298.

Box, G.E.P. and G.M. Jenkins, 1976: Time Series Analysis Forecasting and Control, Holden-Day, Oakland, CA.

Burges, S.J. and A.E. Johnson, 1973: Probabilistic short-term river yield forecasts, *ASCE J. Irrig. and Drainage Div.*, V. 99, No. IR2, 143-155.

Changnon, S., 1999: Impacts of 1997-98 El Nino-Generated weather in the United States. *Bull. Amer. Met. Soc.* 80:9, September.

Charles, S.P., Bates, B.C., Whetton, P.H. and J.P. Hughes, 1999: Validation of a downscaling model for changed climate conditions in southwestern Australia. *Clim. Res*. 12, 1-14.

Chen, M., Dickinson, R.E., Zeng, X. and A.N. Hahmann, 1996: Comparison of precipitation observed over the continental United States to that simulated by a climate model. *J. Climate* 9, 2233-2249, September.

Cherkauer, K. A., L. C. Bowling and D. P. Lettenmaier, 2003, Variable Infiltration Capacity (VIC) cold land process model updates, *Global and Planetary Change* 38(1-2), 151-159.

Cherkauer, K. and D.P. Lettenmaier, 1999: Hydrological effects of frozen soils in the upper Mississippi River basin. *J. Geophysical Research*, GCIP II Special Issue.

Christensen, N.S., Wood, A.W., Lettenmaier, D.P. and R.N. Palmer, 2004: Effects of climate change on the hydrology and water resources of the Colorado River basin, *Climatic Change* (accepted)

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B. and R. Wilby, 2004: The Schaake Shuffle: A method for reconstructing space-time variability in forecasted precipitation and temperature fields, *J. of Hydrometeorology* (submitted).

Cocke, S. and T.E. LaRow, 2000: Seasonal predictions using a regional spectral model embedded within a coupled ocean-atmosphere model. *Mon. Weath. Rev.* 128, 689-708, March.

Crane, R.G., Yarnal, B., Barron, E.J. and B.C. Hewitson, 2002: Scale interactions and regional climate: examples from the Susquehanna River Basin, *Human and Ecological Risk Assessment* 8, 147-158.

Crawford and Linsley, 1966: Digital Simulation in Hydrology: Stanford Watershed Model IV, Stanford Univ., Dept. Civ. Eng. Tech. Rep. 39.

Dai, A., Washington, W.M., Meehl, G.A., Bettge, T.W. and W.G. Strand, 2004: Climate change simulations using a new ocean initialization method, *Climatic Change* (accepted).

Day, G. N., 1985: Extended Streamflow Forecasting using NWSRFS, *ASCE J. Water Resources, Planning and Mgmt.* 111(2), 157-170.

Dettinger, M.D., Cayan, D.R., Meyer, M.K. and A.E. Jeton, 2004: Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basins, Sierra Nevada, California, 1900-2099, *Climatic Change* (accepted).

Ducharne, A., Koster, R.D., Suarez, M.J., Stieglitz, M. and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model. 2. Parameter estimation and model demonstration. *J. Geophysical Research*, 105:D20, 24823-38.

Garen, D.C., 1992: Improved Techniques in Regression-Based Streamflow Volume Forecasting, *ASCE J. Water Resources, Planning and Mgmt.* 118(6), pp. 654-670

Georgakakos, K.P., Georgakakos, A.P. and N.E. Graham, 1998: Assessment of benefits of climate forecasts of reservoir management in the GCIP region. GEWEX News, 8:3, August.

Giorgi, F. and L.O. Mearns, 1991: Approaches to the simulation of regional climate change: A review. *Reviews in Geophysics*, 29, 191-216.

Giorgi, F., Hurrell, J.W., Marinucci, M.R. and M. Beniston, 1997: Elevation dependency of the surface climate signal: A model study, *J. Climate* 10, 288-296.

Goddard, L., S.J. Mason , S.E. Zebiak , C.F. Ropelewski , R. Basher , M.A. Cane, 2001: Current approaches to seasonal to interannual climate predictions, Int. J. of Climatology, 21, 1111-1152.

Hamlet, A.F. and D.P. Lettenmaier, 1999a: Effects of climate change on hydrology and water resources in the Columbia River basin, *J. AWRA* , 35 (6), 1597-1623.

Hamlet, A.F. and D.P. Lettenmaier, 1999b: Columbia River streamflow forecasting based on ENSO and PDO climate signals, *ASCE Jour. of Water Res. Plan. and Mgmt*., 125 (6), pp 333-341, Nov/Dec.

Hay, L.E., M.P. Clark, R.L. Wilby, W.J. Gutowski, Jr., G.H. Leavesley, Z. Pan, R.W. Arritt and E.S. Takle, 2002: Use of Regional Climate Model output for hydrologic simulations, *J. Hydrometeorology*, Vol. 3, October, 571-590.

Hewitson, B.C. and R.G. Crane, 1996: Climate downscaling: techniques and application, *Climate Research* 7, 85-95.

Huber, A.L. and D.C. Robertson, 1982: Regression models in water supply forecasting, *Proc. 50th Annual Western Snow Conference*, Reno NV, April 19-23.

Hutchinson, M.F., 1995: Interpolating mean rainfall using thin plate smoothing splines, *Int. J. Geogr. Inf. Syst.* **9**,385-403.

IPCC (Intergovernmental Panel on Climate Change), 1996: Climate Change 1995: The Science of Climate Change, Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change, J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg and K. Maskell, eds., WMO/UNEP, Cambridge University Press, 572 pp.

IPCC, 2001: Climate Change 2001: The Scientific Basis, Houghton, J.T. and Y. Ding, eds., Cambridge: Cambridge UP.

Ji, M., Behringer, D. W. and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part II: The coupled model. *Mon. Wea. Rev*., 126, 1022-1034.

Kalnay, E., Kanamitso, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woolen, J., Zhu, Y., Leetma, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janoviak, J., Mo, K.C., Jenne, R. and D. Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project, *Bull. Amer. Meteor. Soc.* 77, 437-471.

Kanamitsu, M., A. Kumar, H-M. H. Juang, J-K Schemm, W. Wang, F. Yang, S-Y . Hong, P. Peng, W. Chen, S. Moorthi and M. Ji, 2002: NCEP Dynamical Seasonal Forecast System 2000. *Bull. Amer. Meteor. Soc.* 83:7, July, 1019-1037.

Kidson, J.W. and C.S. Thompson, 1998: A comparison of statistical and model-based downscaling techniques for estimating local climate variations, *J. Climate* 11, 735-753.

Kim, J., 2001: A nested modeling study of elevation-dependent climate change signals in California induced by increased atmospheric CO_2, *Geophys. Res. Lett.* 28, 2951.

Kim, J., Miller, N.L., Farrara, J.D. and S-Y. Hong, 2000: A seasonal precipitation and stream flow hindcast and prediction study in the western United States during the 1997/98 winter season using a dynamic downscaling system. *J. of Hydrometeorology*, 1:4, 311-329.

Koster, R.D., Suarez, M.J. and M. Heiser, 1999: Variance and predictability of precipitation at seasonal-to-interannual timescales. *J. Hydrometeorology*, 1, 26-46.

Koster, R.D., Suarez, M.J., Ducharne, A., Stieglitz, M. and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model. 1. Model structure. *J. Geophysical Research*, 105:D20, 24809-22.

Kumar, A., Hoerling, M., Ji, M., Leetmaa A. and P. Sardeshmukh, 1996: Assessing GCM's suitability for making seasonal predictions. *J. Climate*, 9:1, 115-129.

Latif, M., Anderson, D.L.T., Barnett, T.P., Cane, M.A., Kleeman, R., Leetmaa, A., O'Brien, J., Rosati, A. and E. Schneider,1998: A review of the predictability and prediction of ENSO. *J. Geophys. Res.* 103: 14375-14393.

Lettenmaier, D.P. and D. Garen, 1979: Evaluation of streamflow forecasting methods, *Proc. 47th Annual Western Snow Conference*, Sparks, NV, April 18-20.

Lau, W. K.-M., J. H. Kim and Y. Sud, 1996: Intercomparison of hydrologic processes in AMIP GCMs. *Bull. Amer. Met. Soc.* **77**, 2209-2226.

Leavesley, G.H. and L.G. Stannard, 1995: The precipitation-runoff modeling system - PRMS. In: Singh, V.P. (ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Highlands Ranch, Colorado, 281-310.

Lettenmaier, D.P., Wood, A.W., Palmer, R.N., Wood, E.F. and E.Z. Stakhiv, 1999: Water resources implications of global warming: a US regional perspective, *Climatic Change* 43, 537-79.

Leung, L.R. and S.J. Ghan, 1999: Pacific Northwest climate sensitivity simulated by a regional climate model driven by a GCM. Part I: Control Simulations, *J. Climate* 12, 2010-2030.

Leung, L.R., Hamlet, A.F., Lettenmaier, D.P. and A. Kumar, 1999: Simulations of the ENSO hydroclimate signals in the Pacific Northwest Columbia River basin, *Bulletin of the AMS*, 80 (11), pp 2313-2329, December.

Leung, L.R., Mearns, L.O., Giorgi, F., and R. Wilby, 2003: Workshop on regional climate research: Needs and opportunities, *Bull. Amer. Met. Soc*. 84(1), 89-95.

Leung, L.R., Qian, Y., Bian, X., Washington, W.M., Han, J. and J.O. Roads, 2004: Mid-century ensemble regional climate change scenarios for the western United States, *Climatic Change* (accepted).

Liang, X., D.P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for General Circulation Models, *J. Geophys. Res.* 99(D7), 14415-28.

Liang, X., E. F. Wood, and D. P. Lettenmaier, 1996: Surface soil moisture parameterization of the VIC-2L model: Evaluation and modifications, *Global and Planetary Change*, 13, 195-206.

Liang, X., Wood, E. F. and D. P. Lettenmaier, 1999: Modeling ground heat flux in land surface parameterization schemes. *J. Geophys. Res*. 104(D8), 9581-9600.

Linsley, R.K., M.A. Kohler and J.L.H. Paulhus, 1975: Hydrology for Engineers, 2nd edition, McGraw-Hill, pp 319-337.

Livezey R. E., Masutani, M., Leetmaa, A., Rui, H.-L., Ji, M. and A. Kumar, 1997: Teleconnective response of the Pacific/North American region atmosphere to large central equatorial Pacific SST anomalies, *J. Climate* 10,1787-1820.

Livezey, R. E., M. Masutani and M. Ji, 1996: SST-forced seasonal simulation and prediction skill for versions of the NCEP/MRF model. Bull. Amer. Metor. Soc., 77, 507-517.

Livezey, R., 1990: Variability of skill of long-range forecasts and implications for their use and value. *Bull. Amer. Meteor. Soc.*, 71, 300-309.

Lohmann, D. and many others, 1998a: "The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) Phase-2c Red-Arkansas River Basin Experiment: 3. Spatial and Temporal Analysis of Water Balance Fluxes", *J. Global and Planetary Change*, 19, 161-179.

Lohmann, D., Raschke, E., Nijssen, B. and D.P. Lettenmaier, 1998b: Regional scale hydrology: II. Application of the VIC-2L model to the Weser River, Germany. Hydrological Sciences-*Journal des Sciences Hydrologiques*, 43, 143-158.

Mantua, N., S. Hare, Y. Zhang, J. M. Wallace, R. Francis, 1997: A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production, *Bulletin of the American Meteorological Society*, Vol 78, 1069-1079, June.

Hall, D. K. and Martinec, J., 1985: Remote Sensing of Ice and Snow, Chapman & Hall Ltd., London - New York, 189 pp.

Maurer, E.P., Lettenmaier, D.P. and J.O. Roads, 1999: Water balance of the Mississippi River basin from a macroscale hydrologic model and NCEP/NCAR reanalysis. EOS, *Transactions*, 80(46) pp. F409-410.

Maurer, E.P., O'Donnell, G.M., Lettenmaier, D.P. and J.O. Roads, 2001: Evaluation of NCEP/NCAR reanalysis water and energy budgets using macroscale hydrologic simulations as a benchmark. In: *Observations and Modeling of the Land Surface Hydrological Processes*, AGU series in Water Science and Applications, V. Lakshmi, J. Albertson, and J. Schaake (eds).

Maurer, E.P., Wood, A.W., Adam, J.C., Lettenmaier, D.P. and B. Nijssen, 2002: A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States, *J. Climate* 15, 3237-3251.

McLean, D.A., 1948: Adjusting snow-runoff correlation curves by use of autumn precipitation, *Proc. Western Snow Conf., 16th Annual Meeting*, Reno, NV, April, p. 105-116.

McPhaden, M.J., A.J. Busallacchi, R. Cheney, J.-R. Donguy, K.S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G.T. Mitchum, P.P. Niiler, J. Picaut, R.W. Reynolds, N. Smith and K. Takeuchi, 1998): The Tropical Ocean Global Atmosphere observing system: A decade of progress. *J. Geophysical Research*, 103(C7), 14,169–14,240.

Mitchell, K., and many co-authors, 2000: Recent GCIP-sponsored advancements in coupled land-surface modeling and data assimilation in the NCEP Eta mesoscale model. *Preprints of the 15th AMS Conf. on Hydrology*, Long Beach, CA, Paper P1.22

Murphy, J., 1999: An evaluation of statistical and dynamical techniques for downscaling local climate. *J. of Climate*, 12:8, 2256-2284, August.

National Research Council, 1998: Global Energy and Water Cycle Experiment (GEWEX) Continental-Scale International Project: A Review of Progress and Opportunities, Report of the GEWEX Panel, Commission on Geosciences, Environment, and Resources, 112 pp., National Academy Press, Washington, D.C.

Nijssen, B., D.P. Lettenmaier, X. Liang, S.W. Wetzel, E.F. Wood, 1997: Streamflow simulation for continental-scale river basins. *Water Res. Research*., 33 (4), 711-724, April.

Palmer T.N. and D.L.T. Anderson, 1994: The prospect for seasonal forecasting – a review paper. *Q. J. R. Meteorol. Soc.*, 120, 755-793.

Panofsky, H.A. and G.W. Brier, 1968: Some Applications of Statistics to Meteorology, The Pennsylvania State University, University Park, 224 pages.

Parshall, R.L., 1948: Forecast of runoff based on the water content of the watershed snow cover corrected by a factor involving fall flow of the stream, *Proc. Western Snow Conf., 16th Annual Meeting*, Reno, NV, April, p. 157-161.

Payne, J.T., A.W. Wood, A.F. Hamlet, R.N. Palmer and D.P. Lettenmaier, 2004: Mitigating the effects of climate change on the water resources of the Columbia River basin, *Climatic Change* (accepted).

Pei, D., S.J. Burges and J.R. Stedinger, 1987: Runoff volume forecasts conditioned on a total seasonal runoff forecast, *Water Resources Research* **23**:1, 9-14, January.

Perica, S., J.C. Schaake and D.-J. Seo, 2000: Hydrologic application of global ensemble precipitation forecasts, *Proc. Conf. on Hydrology, Amer. Met. Soc*., Long Beach, January.

Piechota, T.C., Dracup, J.A., 1996: Drought and Regional Hydrologic Variations in the United States: Associations with the El Niño/Southern Oscillation. *Water Resources Research*, 32(5), 1359-1373.

Philander, S.G., 1990: El Nino, La Nina, and the Southern Oscillation. Academic Press, Inc., San Diego, CA. 293 pp.

Risbey, J.S. and P.H. Stone, 1996: A case study of the adequacy of GCM simulations for input to regional climate change assessments. *J. Climate*, 9, 1441-1467.

Roads, J.O., Chen, S.C., Kanamitsu, M. and H. Juang, 1999: Surface water characteristics in NCEP global spectral model and reanalysis. *J. Geophysical Research*. 104:D16, 19307-27, August.

Shaman, J., Stieglitz, M., Zebiak, S. and M. Cane, 2003: A local forecast of land surface wetness conditions derived from seasonal climate predictions, Journal of Hydrometeorology, Vol. 4, No. 3, 611-626.

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

Stern, P.C. and W.E. Easterling, eds., 1999: Making Climate Forecasts Matter. National Research Council Report, National Academy Press, Washington, D.C.

Trenberth, K. E., 1997: The definition of El Niño. *Bull. Amer. Meteor. Soc*. 78: 2771-2777.

Twedt, T.M., Schaake, J.C., Jr. and E.L. Peck, 1977: National Weather Service Extended Streamflow Prediction, *Proc. Western Snow Conf.*, Albuquerque, New Mexico, 9 Pages, April.

Vail, L. and M.S. Wigmosta, 2004: Yakima River assessment study, a focus on water-management effects at the sub-basin scale, *Climatic Change* (accepted).

Van Rheenen, N.T., A.W. Wood, R.N. Palmer and D.P. Lettenmaier, 2004: Potential implications of PCM climate change scenarios for Sacramento - San Joaquin River basin hydrology and water resources, *Climatic Change* (accepted).

Washington, W.M., Weatherly, J.W., Meehl, G.A., Semtner, A.J., Bettge, T.W., Craig, A.P., Strand, W.G., Arblaster, J., Wayland, V.B., James, R. and Y. Zhang, 2000: Parallel climate model (PCM) control and transient simulations, *Climate Dynamics* 16 (10-11), 755-774.

Wilby, R.L. and T.M.L. Wigley, 1997: Downscaling general circulation model output: a review of methods and limitations, *Progress in Physical Geography*, 21, 530-548.

Wilby, R.L., Hay, L. E., Gutowski Jr., W. J., Arritt, R. W., Takle, E. S., Pan, Z., Leavesley, G. H. and M. P. Clark, 2000. Hydrological responses to dynamically and statistically downscaled climate model output. *Geophysical Research Letters*, 27:8, pp. 1199-1202, April.

Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewistson, B.C., Main, J. and D.S. Wilks, 1998: Statistical downscaling of general circulation model output: A comparison of methods, *Water Resour. Res.* 34, 2995-3008.

Wilks, D.S., 1995: Statistical methods in the atmospheric sciences, Academic Press, Inc., San Diego.

Wood, A.W., Leung, L. R., V. Sridhar and D.P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Climatic Change* (accepted)

Wood, A.W., Maurer, E.P., Kumar, A. and D.P. Lettenmaier, 2002: Long range experimental hydrologic forecasting for the eastern U.S., *J. of Geophysical Research*, Vol. 107, D20, 4429, October.

Wood, E.F., Liang, X., Lohmann, D. and D.P. Letenmaier, 1998: The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) Phase-2(c) Red-

Arkansas River Experiment: 1. Experiment description and summary intercomparisons, *J. Global and Planetary Change*, 19, 115-135.

Yarnal, B., Lakhtakia, M.N., Yu, Z., White, R.A., Pollard, D., Miller, D. A. and W.M. Lapenta, 2000: A linked meteorological and hydrological model system: the Susquehanna River Basin Experiment (SRBEX), *Global and Planetary Change* 25, 149-161.

Zhu, C., D.W. Pierce, T.P. Barnett, A.W. Wood and D.P. Lettenmaier, 2004: Evaluation of Hydrologically Relevant PCM Climate Variables and Large-scale Variability over the Continental U.S., *Climatic Change* (accepted).

VITA

EDUCATION

Ph.D.	University of Washington, Civil and Environmental Engineering, 2003
M.S.E.	University of Washington, Civil and Environmental Engineering, 1995
B.A.	Amherst College, English, 1988

PROFESSIONAL AND RESEARCH EXPERIENCE

December 2002 -- present
Research Scientist, Dept. of Civil and Environmental Engineering, Hydrology Group, University of Washington, Seattle

Responsibilities include:

- Coordination of efforts to development, evaluate and implement a seasonal western U.S. hydrologic and streamflow forecasting system
- Advising of Masters students and a technician working on related or subsidiary projects
- Preparation and/or delivery of presentations of technical results to academic audiences, and to sponsors
- Development and maintenance of links to federal forecasting agencies
- Development of web site for distribution of forecast system results
- Consultation with other academic researchers who are pursuing similar efforts
- Occasional field work focusing on the remote sensing of snow pack properties

September 1998 – December 2002
Research Assistant, Dept. of Civil and Environmental Engineering, University of Washington, Seattle

In pursuit of doctorate degree in hydrologic engineering under Dr. Lettenmaier (Hydrology Group), responsibilities included:

- Development, evaluation and real-time implementation of climate model-based seasonal hydrologic and streamflow forecasts
- application of a physical hydrologic model (VIC) to western U.S. river basins
- development and evaluation of datasets for initialization of the hydrologic model during the spin-up period for real-time forecasts
- snow field survey work in Idaho and snowpack modeling, for research on retrieving snow water equivalent maps from passive microwave satellite radiometry (AMSR)
- modeling of the hydrologic effects of global warming as simulated by the Dept. of Energy PCM climate model, for the western U.S.
- funding proposal preparation

June 1996 – June 1998
Visiting Fellow, U.S. Army Corps of Engineers Institute for Water Resources (IWR), Policy and Special Studies Division, Alexandria, VA

During Intergov. Personnel Act placement from U. of Washington, duties included:

- model-based (HSPF) evaluation of potential influence of wetland storage on Mississippi River basin flooding
- model-based (STELLA) assessment of probability of long-term (5-year) flooding in Devil's Lake, ND; and presentation of findings in an interagency workshop, Grand Forks, ND
- presentation of Devil's Lake model results to Associate Secretary of the Army, Civil Works; and to the U.S. Congressman and Senators from ND
- funding proposal preparation

April 1996 – June 1996

Research Staff, Dept. of Civil and Environmental Engineering, University of Washington, Seattle

As research staff to Dr. Lettenmaier (Hydrology Group), implemented existing HEC-PRM model for the Columbia River basin and assessed potential future operational changes due to altered agricultural policy and instream flow requirements.

March 1994 – April 1995

Research Assistant, Dept. of Civil and Environmental Engineering, University of Washington, Seattle

In pursuit of an M.S.E. degree in systems engineering under Rick Palmer (Systems Engineering and Water Resources) and afterward, work included:

- development of a reservoir system model (for long range planning studies) of the Willamette River basin
- model-based assessment of implications of global warming for water resources systems including Boston municipal water supply, Savannah River system, and the ACF-ACT River system, and Tacoma (Green River system)

PUBLICATIONS AND PRESENTATIONS

PUBLICATIONS (PEER REVIEWED)

Wood, A.W., L.R. Leung, V. Sridhar and D.P. Lettenmaier, 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model surface outputs, *Climatic Change* (accepted).

Wood, A.W., A. Kumar and D.P. Lettenmaier, 2004: A retrospective assessment of NCEP GSM-based ensemble hydrologic forecasting in the western U.S., *J. Climate* (submitted).

Christensen, N.S., Wood, A.W., Lettenmaier, D.P. and R.N. Palmer, 2004. Effects of Climate Change on the Hydrology and Water Resources of the Colorado River Basin, *Climatic Change* (accepted)

Payne, J.T., A.W. Wood, A.F. Hamlet, R.N. Palmer, and D.P. Lettenmaier, 2004. Mitigating the effects of climate change on the water resources of the Columbia River basin, *Climatic Change* (accepted)

Van Rheenen, N.T., A.W. Wood, R.N. Palmer and D.P. Lettenmaier, 2004. Potential Implications of PCM Climate Change Scenarios for Sacramento - San Joaquin River Basin Hydrology and Water Resources, *Climatic Change* (accepted).

Zhu, C., D.W. Pierce, T.P. Barnett, A.W. Wood, and D.P. Lettenmaier, 2004. Evaluation of Hydrologically Relevant PCM Climate Variables and Large-scale Variability over the Continental U.S., *Climatic Change* (accepted).

Wood, A.W., Maurer, E.P., Kumar, A. and D.P. Lettenmaier, 2002. Long Range Experimental Hydrologic Forecasting for the Eastern U.S., *J. Geophys. Res.,* 107(D20), doi:10.1029/2001JD000659.

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen, 2002. A Long-Term Hydrologically-Based Data Set of Land Surface Fluxes and States for the Conterminous United States, *J. Climate* 15(22), 3237-3251.

Chen, C-T., Nijssen, B., Jianjun, G., Tsang, L., Wood, A.W., Hwang, J-N. and D.P. Lettenmaier, 2001. Passive microwave remote sensing of snow constrained by hydrological simulations. *IEEE Transactions on Geoscience and Remote Sensing*, 39:8, August, 1744 -1756.

Lettenmaier, D.P., Wood, A.W, Palmer, R.N., Wood, E.F. and E.Z. Stakhiv, 1999. Water Resources Implications of Global Warming: A U.S. Regional Perspective. *Climatic Change* 43, no.3, November, 537-579.

Wood, A.W., R.N. Palmer, D.P. Lettenmaier and E.Z. Stakhiv, 1997. Assessing climate change implications for water resources planning. *Climatic Change* 37, 203-228.

Wood, A.W., R.N. Palmer and K. Petroff, 1997. An Assessment of Zero-Tolerance Regulation in King County. *ASCE J. Water Resources Planning and Management*, Vol. 123, No. 4, July/August, 239-245.

Wood, A.W., D.P. Lettenmaier and R.N. Palmer, 1997. Assessing climate change implications for water resources planning, in Climate Change and Water Resources Planning Criteria, K.D. Frederick, D.C. Major and E.Z. Stakhiv, eds., Kluwer Academic Publishers.

SELECTED CONFERENCE PRESENTATIONS

Wood, A.W. and D.P. Lettenmaier, 2003. Relative contribution of initial condition and climate forecast skill to seasonal hydrologic forecast accuracy, GEWEX Americas Prediction Project PI's meeting, Seattle, WA, July.

Wood, A.W., A.F. Hamlet and D.P. Lettenmaier, 2003. A west-wide seasonal to interannual hydrologic forecast system, GEWEX Americas Prediction Project PI's meeting, Seattle, WA, July.

Wood, A.W. and D. P. Lettenmaier, 2003. Comparing hydrologic forecast uncertainty due to initial condition error versus climate forecast error, European Geophysical Society XXVIII General Assembly, Nice, France, April

Wood, A.W. and D. P. Lettenmaier, D.P., 2003. Retrospective Assessment of Seasonal Hydrologic Forecasts in the Western U.S., AMS Annual Meeting, Long Beach, CA, Feb.

Wood, A.W. and D. P. Lettenmaier, 2002. Experimental real-time seasonal hydrologic forecasting, AMS Conference on Applied Climatology, Portland, OR, May.

Wood, A.W., D.P. Lettenmaier, A. Kumar and E.L. Miles, 2002. A blueprint for west-wide seasonal hydrologic forecasting, 2002. AMS/GAPP Mississippi River Climate and Hydrology Conference, New Orleans, May.

Wood, A.W., N. Christensen, N.T. Van Rheenen, J.T. Payne, A.F. Hamlet, R.N. Palmer, and D.P. Lettenmaier, 2002. PCM climate change scenario implications for western U.S. hydrology and water resources, European Geophysical Society XXVII General Assembly, Nice, France, April

Wood, A.W. and D. P. Lettenmaier, 2002. Experimental Real-time Seasonal Hydrologic Forecasting, European Geophysical Society XXVII General Assembly, Nice, France, April

Lettenmaier, D.P., A.W. Wood and A.F. Hamlet, 2001. End-to-End Streamflow Forecasting Products for Large River Basins, GEWEX Americas Prediction Project PI's meeting, Seattle, WA, July 21, 2001.

Maurer, E.P., A.W. Wood, J.C. Adam, B. Nijssen, and D.P. Lettenmaier, 2001. Derived Continental U.S. Land Surface Hydrologic Fluxes and State Variables, 1950-2000, EOS, Transactions, 82(47) p. F409. (presented at AGU Fall Meeting, 12/11/2001).

C. Zhu, A.W. Wood, D.P. Lettenmaier, 2001. Evaluation of Hydrologically Relevant PCM Precipitation Characteristics over the Continental U.S., Fall AGU Meeting, San Francisco, Dec.

Wood, A.W., D.P. Lettenmaier, N. Christensen, 2001. PCM Climate Change Scenario Implications for Western U.S. Hydrology, Fall AGU Meeting, San Francisco, Dec.

Wood, A.W., A. Kumar and D. Lettenmaier, 2001. Experimental Real-time Seasonal Hydrologic Forecasting for the Columbia River basin. 26th Annual Climate Diagnostics and Prediction Workshop, Scripps Oceanographic Institute, San Diego, CA, October.

E.P. Maurer, A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen, 2001. A long-term land surface hydrologic data set for the continental U.S., 26th Annual Climate Diagnostics and Prediction Workshop, San Diego, Oct.

Wood, A.W., A. Kumar and D.P. Lettenmaier, 2001. Experimental real-time seasonal hydrologic forecasting for the Columbia River basin. Presented at Climate Impacts Group Annual Meeting, NOAA-PMEL, Seattle, June.

Wood, A.W., E.P. Maurer, D.P. Lettenmaier, 2001. Experimental hydrologic forecasting for East Coast US river basins during summer 2000, European Geophysical Union XXVI General Assembly, Nice, France, March.

Wood, A.W., E.P. Maurer, A. Kumar and D.P. Lettenmaier, 2000. Experimental hydrologic forecasting for East Coast US river basins during summer 2000, 25th Annual Climate Diagnostics and Prediction Workshop, Palisades, Oct. 2000

Wood, A.W., E.P. Maurer, A. Kumar and D.P. Lettenmaier, 2000. Experimental Hydrologic Forecasting for East Coast US River Basins during Summer 2000, University of Washington/University of British Columbia Joint Hydrology Meeting, Seattle, October.

Wood, A.W., S. Schwartz, D.P. Lettenmaier, and D.P. Lettenmaier, Experimental streamflow forecasting for water resources management in the Ohio River basin, EOS, Transactions of the American Geophysical Union, 80(46) (Supplement): 364, 1999

Wood, A.W. and D.P. Lettenmaier, 1998. Implications of Global Warming for U.S. Water Resources. EOS Transactions of the American Geophysical Union (Presented at Fall AGU Meeting), 1998

Wood, A.W., 1998. GCIP and water resources: an assessment of precipitation forecast use potential in the Ohio River basin, University of Washington/University of British Columbia Joint Hydrology Meeting, Seattle, October.

INVITED SEMINARS

Climate Impacts Group, Joint Institute for the Study of the Atmosphere and the Oceans, University of Washington, Seattle, June 2003

Climate Impacts Group, Joint Institute for the Study of the Atmosphere and the Oceans, University of Washington, Seattle, November 2001

FELLOWSHIPS AND HONORS

VALLE Scholarship and Scandinavian Exchange Program, University of Washington, 1998-1999
Intergovernmental Personnel Act Fellowship, U.S. Army Corps of Engineers Institute for Water Resources, Policy Studies Division, Alexandria, VA, 1996-1998
Chi Epsilon, University of Washington chapter, 1995
National Merit Scholar, 1984

PROFESSIONAL SOCIETIES

American Geophysical Union
American Society of Civil Engineers
American Meteorological Society

COMPUTING

Operating systems: Windows and UNIX (HP-UX / Linux / Solaris / FreeBSD)
Programming languages: C, FORTRAN, Perl, various UNIX shell utilities
Applications: MS Office and other Windows app's.; Splus, Matlab, StarOffice, Gempak/GARP, Generic Mapping Tools, ARC/INFO (GIS) and other UNIX app's
Data Formats: GRIB, NetCDF; also experience processing large (up to GB size) datasets

PROFESSIONAL REGISTRATION

Engineer-in-Training, Washington, 1995, No. 20716