U.S. DEPARTMENT OF THE INTERIOR

U.S GEOLOGICAL SURVEY

Assessment of Folsom Lake Watershed Response to Historical and Potential Future Climate Scenarios

by

Theresa M. Carpenter[1] and Konstantine P. Georgakakos[1,2]

Open-File Report 00-333

 This report is the basis for an article submitted for publication to the *Journal of Hydrology, Special Issue on Probabilistic and Ensemble Forecasting*, that is currently under review.

1 Hydrologic Research Center, 12780 High Bluff Drive, Suite 250, San Diego, CA 92130, USA
2 Scripps Institution of Oceanography, UCSD, CA 92093-0224, USA

August 2000

Assessment of Folsom Lake watershed response to historical and potential future climate scenarios

Theresa M. Carpenter[1] and Konstantine P. Georgakakos[1,2*]

ABSTRACT

An integrated forecast-control system was designed to allow the profitable use of ensemble forecasts for the operational management of multi-purpose reservoirs. The system ingests large-scale climate model monthly precipitation through the adjustment of the marginal distribution of reservoir-catchment precipitation to reflect occurrence of monthly climate precipitation amounts in the extreme terciles of their distribution. Generation of ensemble reservoir inflow forecasts is then accomplished with due account for atmospheric-forcing and hydrologic-model uncertainties. These ensemble forecasts are ingested by the decision component of the integrated system, which generates non-inferior trade-off surfaces and, given management preferences, estimates of reservoir-management benefits over given periods. In collaboration with the Bureau of Reclamation and the California Nevada River Forecast Center, the integrated system is applied to Folsom Lake in California to evaluate the benefits for flood control, hydroelectric energy production, and low flow augmentation. In addition to retrospective studies involving the historical period 1964-1993, system simulations were performed for the future period 2001-2030, under a control (constant future greenhouse-gas concentrations assumed at the present levels) and a greenhouse-gas-increase (1-% per annum increase assumed) scenario. The present paper presents and validates ensemble 30-day reservoir-inflow forecasts under a variety of situations. Corresponding reservoir management results are presented in Yao and Georgakakos, A., this issue. Principle conclusions of this paper are that the integrated system provides reliable ensemble inflow volume forecasts at the 5-% confidence level for the majority of the deciles of forecast frequency, and that the use of climate model simulations is beneficial mainly during high flow periods. It is also found that, for future periods with potential sharp climatic increases of precipitation amount and to maintain good reliability levels, operational ensemble inflow forecasting should involve atmospheric forcing from appropriate climatic periods.

[1]Hydrologic Research Center, 12780 High Bluff Dr., Suite 250, San Diego, CA 92130
[2]Scripps Institution of Oceanography, UCSD, La Jolla, CA 92093-0224

INTRODUCTION AND OVERVIEW

The focus of this research is to quantitatively assess benefits that use of climate forecasts has for the improved management of reservoir hydrosystems, both for historical periods through retrospective studies and for future periods under a projected change of hydroclimatic conditions. The methodology is exemplified for Folsom Lake in California, a multiobjective reservoir upstream of Sacramento with significant plant objectives of flood control, hydroelectric power production and low flow augmentation. The research project is a collaboration of a forecast group (Hydrologic Research Center in San Diego, CA) and a reservoir decision-support group (Georgia Water Resources Institute in Atlanta, GA). The present paper provides information on the forecast aspects of the work while the companion paper (Yao and A. Georgakakos, 2000) emphasizes the decision-support formulation and assesses relevant benefits from reservoir operation under various climate scenarios.

Improved representations of streamflow variability in operational hydrologic forecasts have potential value to reservoir management. Experiments are designed to explore the benefits of using climate model output as a representation of climate variability for use in ensemble streamflow forecasting techniques, both for current climate and for altered future climate under warming scenarios. The primary tool for quantifying benefits is an integrated numerical system which involves components for ingesting and downscaling Global Climate Model (GCM) forecasts; generating ensemble reservoir inflow forecasts conditioned on downscaled GCM information; generating trade-off surfaces for decision support of multiobjective reservoir operations taking into consideration the forecast uncertainty in reservoir inflow; and quantifying reservoir-operation benefits for given release policies. Figure 1 shows a schematic of the components and links of the modeling system, which will be discussed in the next section. This approach was recently introduced by Georgakakos et al. 1998b to assess the utility of climate model forecasts for operational water resources management. It is important to note that, for this study, the hydrologic models used were adaptations of the US National Weather Service operational forecast models. Thus, the improvements estimated in reservoir operations are realizable with minor modifications of the existing operational forecast systems.

For the assessment of benefits, retrospective studies involving historical climate, hydrologic and reservoir-operation data together with a variety of scenarios of climate information are used. The retrospective studies are concerned with the intercomparison of the benefits from Folsom Management when: *(a)* a simple regression model based on historical snowpack information, precipitation and flow, is used to forecast inflows to Lake Folsom without any information from GCM forecasts and without explicit uncertainty modeling, and *(b)* the integrated forecast-control system is used with and without the benefit of monthly forecasts of precipitation and temperature from the Canadian Centre for Climate Modeling and Analysis coupled ocean-atmosphere GCM (CGCM1). The years 1964 - 1993 constitute the historical study period. The reservoir-management performance measures, specified in collaboration with Staff of the Bureau of Reclamation, are annual spillage, annual flood damage, low-flow augmentation, and annual energy value from hydroelectric power generation. The reliability diagrams of 30-day inflow-volume forecasts quantify forecast performance for the purposes of this study.

The principle finding is that the integrated system with GCM information produces reliable inflow volume forecasts and outperforms the simple regression forecast system without

CGCM1 information, particularly in significantly reducing maximum flood damage over the historical period while producing more energy and less spillage. It is also found that most of the improvement is due to the explicit modeling of uncertainty by the integrated system.

With these results at hand we assess the likely changes in system performance under future climate regimes by devising a simulation system for Folsom system operations during the future period 2001 - 2030. This simulation system is used as a basis for the assessments using output fields from control and 1-% greenhouse-gas-increase scenarios of the CGCM1 climate model. The principle finding is that the integrated system is able to accommodate successfully climate variability and change as long as the inflow volume forecasts conditioned on CGCM1 remain reliable and representative of the future hydroclimatic regime.

After a short description of the study catchment and available data, the next section discusses the methodology and presents a short literature review on forecast issues that are significant for this study. The mathematical formulation of the ensemble forecast approach and the measures used to validate ensemble forecasts are given in section 3. Section 4 assesses the reliability of the 30-day reservoir-inflow volume forecasts obtained for the historical and future periods and draws assessments as to the forecast system used and its limitations. Conclusions and recommendations are offered in section 5.

Catchment and Data

The three forks (North, Middle and South Fork) of the American River drain approximately 4,800 km^2 of the mountainous terrain of central California (with elevations up to 3,000 m) and join to provide inflow to Folsom Lake (Figure 2). The catchment with outlet at Folsom Lake is characterized by typical orographic rainfall patterns associated with steep terrain barriers, and with snow in the high elevations (typically above 1500 m). The climatological means of hourly precipitation, based on a sample of precipitation events for the wet period 1980-1987, show a maximum of about 2 mm/hr over the headwaters of the North Fork of the American River with pronounced variability. The historical flow records show a wet season during the winter months of November through May and significantly lower flows during the summer months of July, August, and September. The catchment average response time to significant rainfall events in the absence of snow is approximately 12 hours. Significant snow accumulation occurs during winter in the upper reaches of the three Forks.

For hydrologic modeling and to capture the high spatial rainfall variability, we subdivided the catchment area into four sub-catchments. These consist of the three Forks of the American River and of the local area after the Forks join and down to the Lake (sub-catchment boundaries are shown in Figure 2). Furthermore, and for the purpose of computing snow accumulation and ablation, snowmelt is computed in two zones for each sub-catchment: one above 1,500 m and one below 1,500 m.

Historical data of daily pan evaporation, adjusted by monthly coefficients, are used to estimate potential evapotranspiration demand over the catchment area. The monthly pan adjustment coefficients were estimated from long term mass balance considerations (e.g., see Georgakakos et al. 1995) using precipitation, pan evaporation and flow measurements with initial regional estimates provided by NOAA through the Evaporation Atlas. Because the pans are in the lower elevations, it is possible that we overestimate monthly potential evapotranspiration.

Information about the rainfall distribution over the upstream drainage area is provided by the system of automated raingauges (Figure 2). Tsintikidis et al. (2000) characterizes measurement errors associated with the presently operating raingauge system and recommend additional gauges for installation within the catchment boundaries. Their study is based on data from an intensely measured historical period (1980-1987) with flows that were above average. They conclude that the present system tends to under estimate slightly the mean areal precipitation over the entire drainage area with more significant over- and under-estimation in the North and Middle Forks. For this analysis we use the data from the operational raingauge network with the understanding that the observed rainfall may be biased over tributary catchments.

There are three gauging stations with historical daily discharge data, one on each fork of the American River. There is also a record of Lake Folsom inflows computed from water balance considerations at the Folsom Lake site. Preliminary analysis indicated that the latter record is noisy for low flows. Upstream small reservoirs in the Middle and South Forks of the American River regulate low flows and contribute to low flow uncertainty at the gauging sites of these forks (no historical records of substantial length are available for the releases from upstream reservoirs). We used the historical records at the three streamflow gauge sites to calibrate the hydrologic model for each Fork of the American River.

The annual cycles of rain plus melt, potential and actual evapotranspiration (ET), observed and simulated Lake Folsom inflow, and the saturation fraction of the upper soil (as estimated by the hydrologic model) are shown in Figure 3 for the period 1964 - 1993. The cycles are for the entire Folsom Lake drainage and have been computed from daily data processed by the hydrologic model. It is apparent that rain plus melt is significant from October through May, with maximum monthly-averaged values of more than 7 mm/d in February. The upper soils reach saturation in January-March and are very dry in August and September, when the potential ET is at a maximum. Due to soil water control, the actual ET reaches a maximum of about 2.8 mm/d in May, when the soils are wet and the potential ET is increasing. Figure 3 also shows that maximum monthly-average Lake inflow of about 5,000 cfs occurs from March to May and a minimum of about 1,000 cfs occurs in October. The model simulations over-estimate inflow in the winter and underestimate it in the spring, and to a lesser extent in the summer and fall. Errors in snow accumulation and ablation are responsible for these discrepancies.

Regulation of American River flows occurs upstream of Folsom Dam for the Middle and the South Fork. The influence of the upstream reservoirs is minimal for medium and high flows and is only significant for low flows. No real time records exist to quantify this regulation and it has been omitted in this analysis (no naturalized flow analysis was done).

METHODOLOGY AND LITERATURE REVIEW

There are two basic methodological issues discussed in this section. The first pertains to the integrated forecast-control system used to process large-scale GCM forecasts and historical hydrometeorological information and to produce ensemble reservoir-inflow forecasts. The second pertains to the approach followed to set up a simulation analog for the future period so that an integrated forecast-control methodology could be applied for that period such as it is applied for the historical period of known meteorological forcing and streamflow. We discuss these methodological issues next with reference to pertinent published studies.

Integrated Forecast-Control System

Figure 1 shows the components of the integrated system used in this work (see also Yao and A. Georgakakos, 2000). It is a sequence of interconnected modeling components for global climate, catchment hydrology and decision support for operating and planning reservoirs. The links between modeling components adjust the connecting water fluxes and estimate forecast uncertainty through the system. They consist of downscaling the GCM forecasts and generating ensembles of reservoir-inflow forecasts. System integration is accomplished through (a) the conservation of mass of water as it flows through the system components, (b) the forward propagation of forecast uncertainty, and (c) the feedback from the quantified objectives of the decision support component back to all key system components for improved performance (Figure 1). Unique aspects of this integrated system are: (a) the integration of GCM forecasts with the hydrologic and water resources management components; (b) the explicit account and use of forecast uncertainty by all the modeling components; and (c) the use of a multi-horizon reservoir control module, which operates on the basis of ensemble inflow forecasts to quantify optimal trade-off surfaces among various reservoir management objectives at given reliability levels. In this context uncertainty is used to define the range and character of forecast errors due to erroneous (or noisy) input or model-parameter estimates and model structure.

The numerical experiments based on the diagram of Figure 1 are conducted in a manner similar to real time operations. That is, using a set of initial conditions and a set of likely future meteorological forcing time series, the hydrologic model is run from the present forecast preparation time to the maximum forecast lead time several times, each time using a single set of forcing time series. Once the ensemble of reservoir inflow forecasts is produced, the model used the observed meteorological forcing time series to simulate and update its states until the next forecast preparation time at which point the ensemble forecast procedure is repeated. Use of the GCM information is made to determine which input meteorological time series would be used at each forecast preparation time. For application to Folsom Lake, GCM simulations of October to April precipitation were used in the form of distribution terciles to define most appropriate time series of precipitation, temperature and evapotranspiration input from historical years to force the hydrologic model for producing likely future flows. At each forecast preparation time, through simulation with historical data, the hydrologic model states (snow pack, soil moisture and channel water) are updated before the model is forced by the selected historical time series to produce ensemble inflow forecasts. In our methodology, the ensemble streamflow forecasts reflect hydrologic-model parametric uncertainty in addition to the uncertainty in future precipitation and temperature. The ensemble inflow forecasts feed a reservoir decision model, which uses reservoir characteristics and representations of management objectives to generate trade off scenarios at given reliability levels. Based on these scenarios and possibly on other non-quantifiable information reservoir operators make decisions on reservoir releases. These decisions are implemented by the decision model, which quantifies the benefits of reservoir operation. The benefits corresponding to the use of GCM information are compared to those obtained when, at each forecast preparation time, the historical time series of precipitation, temperature and evapotranspiration are used indiscriminately for generating ensemble inflow forecasts.

The integrated system was first used successfully by Georgakakos et al. (1998) (see also, Georgakakos, 1998) in the context of operational reservoir management for the Des Moines River basin, a tributary to the Mississippi River. When compared to current operational practices and forecast-control systems that did not use climate information, improved reservoir management benefits were obtained overall with a large reduction of maximum daily flood damage.

Prior to the Georgakakos et al. (1998) study, a system of the type in Figure 1 with fully coupled hydrologic-forecast and reservoir-control components but without GCM information processors was designed and was applied to the Des Moines River basin (Georgakakos, et al. 1995; A. Georgakakos et al. 1998). That system used the historical hydrometeorological data and developed reservoir inflow forecasts assuming that the current-year mutually dependent series of daily precipitation, potential evapotranspiration and temperature are statistically similar to those of each of the previous years for the same forecast period. Thus, GCM forecasts as to the likelihood of the current year being excessively wet or dry were not used, and the forcing from each of the historical years to the hydrologic model was considered equally likely to occur during the current year. These studies simulated the operation of the reservoir daily for a period of 25 years and they indicated that, as compared to current operational practices, the integrated system offers substantial reduction of water resources management sensitivity to climatic variability, and in particular to droughts and floods. By subdividing the historical record into three different climate periods these studies also showed that the integrated system, with reservoir control methods that use stochastic optimization under uncertain forecasts, can substantially mitigate the adverse effects of climatic change as experienced in the historical record. It was found that the value of hydrologic forecasting is higher in wetter and more variable climates for the Des Moines River region.

In the following sections we discuss the basic elements and functionality of the components of Figure 1.

GCM Forecasts

There are several GCMs, run at major centers around the world, providing seasonal and longer term forecasts for the Globe with resolution of 60,000 km^2 or coarser. The models are based on the equations for mass, momentum and energy conservation in the atmosphere complemented with several parameterized relationships to represent natural processes operating at sub-resolution scales (e.g., land-surface processes, convection and cloud processes). Typically, sea surface temperature (SST) is prescribed or forecast and ensemble forecasts of various atmospheric variables are produced starting from different initial conditions for the atmosphere. Most recently, coupled ocean-atmosphere models (without prescribed SSTs) have been developed and used for seasonal forecasts and climate simulation. On-going research and development efforts improve the physical basis, parameterizations, and resolution of GCMs.

Perhaps the most important question associated with GCM forecasts and long-term simulations concerns the degree with which they reproduce observed features of past climate on scales comparable to their spatial resolution. Several studies have addressed this issue and we only outline a characteristic few in the following. Risbey and Stone (1996) discuss the suitability of a typical GCMs in simulating large-scale and synoptic scale processes (stationary waves, jet streams and storm tracks), which are important for regional water resources. They find significant differences between the simulated and observed process features for the Sacramento basin, and they attribute such differences to a large degree in deficiencies in sub-grid scale parameterizations and to a lesser degree in model spatial resolution. They also suggest that statistical or historical-analog reconstruction of future climates present a competitive methodology to the use of GCM simulations. Sinclair and Watterson (1999) appraise the skill of an Australian GCM in replicating contemporary extratropical cyclone and anticyclone behavior. They find that the GCM reproduced present-day storm tracks realistically, although it did provide generally fewer and weaker systems over all. Use of the same GCM to predict double-CO_2 scenario cyclonic behavior suggested a 10-15% reduction in future activity of cyclones and anticyclones. Yu and Mechoso (1999) discuss simulation errors of SST and surface heat flux from the UCLA coupled GCM and intercompare with surface heat flux errors from the CGCM's atmospheric component driven by prescribed SSTs. They found that, with a few exceptions, the CGCM produces more realistic surface heat fluxes off the equator while the atmospheric GCM outperforms the CGCM at the equator. The feedback between SST and latent heat flux is opposite in the coupled model (negative) than in the observations (positive) and it is attributed to errors in the simulation of the cross-equatorial component of the surface wind.

Intercomparison studies among GCMs, and between GCMs and statistical forecast methods have also been published. Most recently such studies indicate comparable performance (e.g., Anderson et al. 1999), which allows for a forecast lead-time of skillful forecasts of a few months for North America and for El Niño events (Barnston, et al. 1999). In this work we use the simulations of a coupled GCM. Precipitation and temperature fields have been obtained from the coupled global climate model CGCM1 of the Canadian Centre for Climate Modeling and Analysis (see reference at the web site: www.cccma.bc.ec.gc.ca/cgi-bin/cgcm1). This is the climate model recommended for use in assessment studies pertaining to water resources (see assessment strategy at web site: www.nacc.usgcrp.gov/scenarios/strategy.html). The model grid spacing is approximately $3.75^0 \times 3.75^0$ in longitude and latitude and it is almost two orders of magnitude larger than the application area of the case study.

Downscaling

Even after accepting that GCMs have skill in their seasonal simulations and predictions, use of their output within the integrated system of Figure 1 requires that the information is downscaled to the spatial and temporal scales of the hydrologic model component for the catchment of interest. Furthermore, removal of any biases in the GCM forecasts for the region of interest should be accomplished at this level. As an example, Figure 4 shows the relationship of the wet season monthly precipitation between the mean areal precipitation over the Folsom Lake drainage and the few closest nodes of CGCM1. Significant scale bias exists and the variability of the mean areal precipitation is substantially different from that of CGCM1 nodes, with a cross-correlation coefficient that is less than 0.45 accounting for less than 20 percent of the observed precipitation variance. These facts have led to the development of numerous downscaling techniques (e.g., Mearns et al. 1990, Wolock et al. 1993).

Murphy (1999) distinguishes three types of downscaling: (a) relate the surface variables of interest (e.g., temperature and precipitation) to GCM simulations of same variables in nearby GCM nodes to adjust for GCM systematic simulation biases and sub-grid scale effects; (b) drive nested regional climate models (grid length of about 50 km) by the output from the GCM over the area of interest and then relate the surface variable of interest to nearby regional model simulation results; and (c) develop statistical downscaling relationships between GCM free atmosphere variables and local surface variables of interest. He inter compared these methods over Europe and found comparable skills for simulating local monthly precipitation and temperature. He also cautions that satisfactory performance over a certain period does not guarantee equally good performance over a future period under climatic change for both dynamical and statistical downscaling methods.

Statistical downscaling of GCM output, if it provides comparable results to those obtained by using nested regional climate models, is an efficient approach to downscaling. Recent results by Sailor and Li (1999), using the National Center for Atmospheric Research (NCAR) GCM and with temperature as the target variable, and Busuioc et al. (1999), using the ECHAM3 GCM of the Max Plank Institute and with precipitation as the target variable, suggest that statistical downscaling produces skillful results in diverse regions. The recent work by von Storch (1999) further shows that adding noise to the estimates obtained from GCM statistical downscaling is an appropriate way to correct for influences in local variable variance that are not controlled by the GCM large scale features. In this work, and following Georgakakos et al. (1998), we adopt a statistical downscaling methodology, which utilizes GCM simulations with monthly resolution to produce mean areal precipitation, potential evapotranspiration, and temperature forcing over the catchment of interest. The mathematical formulation is given in section 3.

Hydrologic Model

We employ hydrologic models routinely used by the National Weather Service California-Nevada River Forecast Center (CNRFC) to forecast reservoir inflows for the study basin. An adaptation (see Georgakakos, 1986) of the operational National Weather Service Sacramento soil moisture accounting model is used to simulate the total runoff inflow to the channel network for each of the four sub-catchments of the American River that constitute the Folsom Lake drainage. The modified Sacramento model is complemented with a kinematic channel routing model (Georgakakos and Bras, 1982) which routes the channel flows downstream in the channel network. The operational National Weather Service snow accumulation and ablation model (Anderson, 1973) computes snowpack properties and snowmelt (both rain-on-snow events and fair-weather-melt events) in each of two elevation zones for each of the three Forks of the American River. Staff of the CNRFC provided snow model output for the purposes of this study.

The results of calibration may be seen in Figure 5 for each of the American River Forks in terms of the degree to which the simulated flows resemble the observed flows for selected events outside the calibration period. In all cases shown, the hydrologic modeling component reproduces the flow well. When the model is forced by observed precipitation, the cross-correlation of daily observed and simulated flows is greater than 0.80 indicating at least a 64% explanation of observed-flow variance. Figure 5 and especially the North Fork results indicate somewhat of an over-estimation of daily flows. Figure 6 presents the annual cycle of monthly-averaged observed daily inflows, along with the mean and standard deviation of the daily residuals of the simulation for each month (residuals computed as simulated minus observed). The residuals indicate a slight under-simulation of observed flows from July through December, over-simulation in February and March, and under-simulation during May. Particularly large variances of the residual errors are shown for May in Figure 6. The errors in winter and spring are attributed to the snow component, which tends to generate melt flow prematurely. Forecast traces in the ensemble streamflow prediction are adjusted to account for these monthly biases and residual variances as discussed in section 3.

Operational Ensemble Flow Forecasts

To allow water resources management with some foresight and without the benefit of skillful meteorological forecasts for long lead times, hydrologists developed an ensemble forecast methodology (called ensemble streamflow prediction or ESP) to produce likely future flows (Day, 1985; Smith et al. 1991). The methodology is based on the premise that the atmospheric forcing of historical years is likely to occur in the future (a stationary climate is assumed). It is most appropriate for catchments with strong seasonal cycles of atmospheric forcing and which exhibit significant persistence in their soil water. The hydrologic model is forced with observed precipitation, potential evapotranspiration and temperature data up to the forecast preparation time and the current estimates of the soil water states of the model are obtained from the model conservation equations. Then, the model is integrated forward in time using these soil water estimates as initial conditions, and using the historical record for each year in turn as input, starting from the month and day of the forecast preparation time and extending out to the maximum forecast lead time. The result is an ensemble of equally likely streamflow prediction traces generated for the forecast horizon with temporal resolution equal to the resolution of the historical atmospheric forcing.

Nibler and Anderson (1993) and Fread et al. (1999) have recently documented operational applications of the ESP methodology and variants. For example, Fread et al. (1999) report on an operational ensemble flow forecasting system under implementation for the Des Moines River in Iowa. The ensemble-forecast system is used to provide probabilistic hydrologic forecasts into the future from days to months. It assimilates meteorological forecasts and climate predictions within the traditionally used procedure using a matching between the marginal probability distributions of precipitation and temperature historical records and climate outlooks. Perica et al. (1999) advance three models for incorporating hydrologic uncertainty in the ESP methodology. The first model adjusts the simulated streamflow directly based on the differences of marginal streamflow distributions between simulated and observed flows. The second and third models use autoregressive schemes of streamflow standardized variates to take into account autocorrelation in hydrologic simulation errors in standardized variate space. Preliminary tests showed that the error models designed improved second moment simulation statistics. A Bayesian theoretical framework for this ensemble forecast problem was also presented recently, with uncertainty in input and parameters explicitly modeled (Krzysztofowicz, 1999).

Georgakakos, et al. 1995, and Georgakakos, A., et al. 1998a document retrospective studies of the application of the ensemble forecasting methodology coupled with a reservoir control component to the management of the multi-objective Saylorville reservoir on the upper Des Moines River. A state estimator complemented the hydrologic model. The estimator was suitable for large river basins with several tributary catchments to filter the noise in the soil water estimates due to input and parametric errors for the historical period. The ensemble forecasts were then generated based on the ESP methodology. Their results show that coupling a forecast system and a reservoir control system with due account for uncertainty generates significant benefits both in flood damage reduction but also for drought management. In section 3 we develop the formalism of the ensemble forecast methodology used in this work, which accounts both for uncertainty in the GCM simulations and for hydrologic model errors, and it is suitable for use with modern reservoir management methodologies.

Validation of Ensemble Forecasts

Widespread use of ensemble forecasting in operational meteorology and in the context of climate simulations has resulted in several studies that address both methodological issues and actual case studies of validating ensemble forecasts. Thus, nonparametric statistics (Anderson and Stern, 1996), binned probability ensemble techniques (Anderson, 1999), relative operating characteristics (Mason et al. 1999) and Brier scores (Buizza et al. 1999) have been used to study forecast system performance in various areas and in terms of various target output variables. In this work we will validate the ensemble flow forecasts using reliability diagrams (see mathematical formulation in section 3). We also assess forecast benefits using the reservoir management component over long periods on the basis plant system measures agreed-upon by the Users of the system (see companion paper by Yao and A. Georgakakos, 2000).

Reservoir Management Component

Research that aims to improve reservoir management methodologies has a long history in hydrology and operations research (e.g., Loucks, 1989). System analysis methods have been used to develop reservoir control methods that account for natural climate variability, assuming statistical stationarity for climate forcing. Such methodologies have recently been combined with modern databases and graphical interfaces to form "user-friendly" decision support systems (e.g., Jamieson, 1996). Yao and Georgakakos (2000, companion paper) provide an overview of the pertinent recent literature in this area and present the essentials of the mathematical formulation used. Here we mention that within the integrated end-to-end framework depicted in Figure 1, reservoir control schemes that are capable to process individual ensemble forecast traces within a well-defined uncertainty formalism are most appropriate (e.g., Georgakakos, A., and Yao, 1993, and Yao and A. Georgakakos, 1993). Also, schemes that produce reliability-indexed trade-off surfaces among reservoir objectives are most useful for operational managers in their decision process, as this allows the consideration of non-quantifiable institutional and social objectives.

Simulations of Future Folsom Climate and Future System Benefits

Climate exhibits complex behavior in a variety of spatiotemporal scales, and predictability of climate variability and change during future decades is a topic of intense research and debate (e.g., Rind, 1999; Petersen, 2000). It has been common practice to employ global climate models or GCMs and assumed scenarios of greenhouse gas emissions to study future anthropogenic impacts on climate (IPCC, 1995). A comprehensive analysis of the reliability of such simulations in the context of reproducing the historical climatic changes and attributing those to natural and anthropogenic causes has been reported in Barnett et al. (1999). These authors conclude that control runs of GCMs may give the best estimates of natural variability on a global scale. They also state that present-day climate modeling capability is such that it is not possible to distinguish the relative contributions of specific natural and anthropogenic (e.g., greenhouse gasses and sulfate aerosols) forcings to the observed climate change. It is thus unknown what are the reliability levels of projected future climatic changes from the current-generation GCMs, and it is very likely that these levels are model specific.

Meehl et al. (2000) does show, however, that there is agreement among GCMs regarding the projected occurrence of changes in various types of extreme events. For example, higher frequencies of extreme warm days, lower frequencies of extreme cold days, a decrease in diurnal temperature ranges associated with higher nighttime temperatures, increased precipitation intensity and extremes, mid-continental summer drying, are all projected such changes.

Simulations of the coupled climate model CGCM1 from the Canadian Centre for Climate Modeling and Analysis until the end of the 21st Century were utilized in this work. Two scenarios were considered: (a) a control scenario with greenhouse-gas forcing assumed at present levels for all historical years; and (b) a scenario which assumes a 1% per year increase in greenhouse gases. An issue that arises is the application of the methodology of Figure 1 to future periods so that we may assess the impacts of a changing climate on the operation of the Folsom Lake reservoir. The integrated forecast-control system requires mutually consistent observed data of precipitation, potential evapotranspiration, temperature and flow, in order to simulate routine operations of the reservoir. Of course no such data exists for the future period. Use of the results of future CGCM1 simulations as surrogate observed data is precluded because of their poor representation of the hydroclimatic conditions in Folsom Lake drainage, and because there cannot be a direct comparison with the system used for the retrospective studies utilizing the historical record.

To create an "observed" record for future years, we matched future years with historical years having similar wet-season (October – April) CGCM1-simulated precipitation. For each year in the entire period (1964-2100), the October-April CGCM1 precipitation total was computed (sum of monthly values). The total precipitation for each of the future-years portion of the record and for each CGCM1 scenario was compared to that of each of the historical years. For each future year, the historical year with the closest total CGCM1 precipitation was identified and stored. This was done for both the CGCM1 control and greenhouse-gas simulation scenarios. It is noted that CGCM1 temperature estimates is not used to obtain a future "observed" record. Under a warming scenario, to the extent that the increase of October-April precipitation volume is associated with earlier melt, the effects of a possible temperature increase are implicitly accounted for by this procedure. Given the relatively short times of Folsom catchment response (less than a day for significant runoff), for the purposes of this study the development of the "observed" flow record was considered adequate. This is an area where additional future research is needed.

Year 2050 marked the point when this procedure for the generation of future "observed" data started to generate only a small sample of the historical years due to increasing CGCM1-projected wet climate trends for the greenhouse-gas increase scenario. Figure 7 shows simulated seasonal precipitation for the CGCM1 control and the 1-% greenhouse-gas-increase runs for the two CGCM1 nodes closest to the Folsom Lake drainage basin. The historical seasonal precipitation is also shown in the Figure. Substantial increase of seasonal precipitation is projected for the region.

For this reason and to have the same number of years in the observed historical record and in the future generated record, we defined the future period to be: 2001 - 2030. To create the "observed" daily record for each future case of CGCM1 scenario (control and greenhouse gas scenarios), the observed flow, rain plus melt, and potential evapotranspiration records of the identified historical year were used as surrogate observed records for each future year. The substitutions were done on a water year basis (October-September) to avoid discontinuities due to the pronounced natural seasonal cycle of the snow accumulation and ablation and of the flow record in this area.

Figure 8 shows the frequency distribution of the historical observed flows, along with the distributions for the resulting "observed" flows of the CGCM1 control and greenhouse-gas-increase records. Very similar distributions are obtained for the historical and CGCM1 control cases, with differences for flows with less than 1% and greater than 99.9% frequency of occurrence. The CGCM1 greenhouse gas case shows generally higher flows throughout the flow range 100 - 80,000 cfs. The annual cycles of monthly-averaged fluxes for the control run (not shown) are very similar to those of the historical period (see Figure 3). Figure 9 shows that for the greenhouse-gas-increase scenario substantial increases of rain plus melt input to the model are obtained for all but the summer months (new maximum of more than 8 mm/d in January). There are attendant increases in observed and simulated Folsom Lake inflows during the wet season (maximum monthly-averaged observed inflow is more than 7,000 cfs in January). For this scenario of future "observed" data, the model over-estimates inflows in winter and underestimates them in the other seasons.

Figure 10 shows the annual cycle of monthly-averaged inflow residual mean and standard deviations, together with the corresponding inflow observations for the greenhouse-gas-increase scenario. The character of the annual cycles is similar to that shown in Figure 6 for the historical period. The residual means and the observed flow means are larger than those of Figure 6. The standard deviations of the simulation residuals have a similar magnitude in the historical and the greenhouse-gas-increase cases. The monthly residual statistics shown in Figure 10 are used for the case of greenhouse-gas-increase scenario to rectify the ensemble Folsom Lake inflow forecasts for the future period and for this scenario (see formulation in section 3).

Once "observed" future data were generated, the procedure of Figure 1 was applied to these data and assessments of the impacts of the changed climate on the Folsom Lake operations were made. Ensemble forecasts were generated every 5 days beginning on the 1st of each month for the period 10/1/2001 through 9/30/2030 for both the CGCM1 control and greenhouse gas scenarios.

MATHEMATICAL FORMULATION

Hydrologic Models

The precipitation-runoff model used in this work consists of a snow accumulation and ablation component, a soil water accounting component, and a channel routing component, all applied to each sub-catchment of the Folsom Lake watershed. There is a total of four sub-catchments, three headwater ones and a local downstream one (Figure 1), and each headwater sub-catchment is subdivided into two elevation zones for the computation of snow accumulation and ablation. The level of 1,500 m separates the two elevation zones.

The relevant state equations may be written in general form as:

$$\frac{d\underline{s}_{ij}}{dt} = \underline{h}(\underline{s}_{ij};\underline{w}_{ij},\underline{q}_{ij}) \quad ; i=1,..,N; j=1,..,M \qquad (1)$$

$$\frac{d\underline{x}_i}{dt} = \underline{f}(\underline{x}_i;\underline{u}_i,\underline{a}_i) \qquad ; i=1,..,N \tag{2}$$

In Equation (1), $\underline{s}_{ij}$ represents the snow-model state vector consisting of the snow cover liquid water equivalent, snow-pack temperature and heat deficit for each of M (=2) elevation zones within each of the N (=4) sub-catchments of the Folsom Lake watershed (see Anderson 1973, for model formulation). The snow-model input and parameter vectors are represented by $\underline{w}_{ij}$ and $\underline{q}_{ij}$, respectively, with the input vector consisting of mean areal precipitation and temperature for elevation zone j and sub-catchment i. In Equation (2), $\underline{x}_i$ is the state vector of the soil water accounting and channel routing components consisting of the soil water content of the modified Sacramento model soil compartments and of the water content of the channel reaches of the channel routing model (see Georgakakos 1986, for model formulation). The input and parameter vectors are denoted in this case with $\underline{u}_i$ and $\underline{a}_i$, respectively, with the input vector consisting of the rain plus melt, potential evapotranspiration, and, for the case of the local sub-catchment, of the upstream total channel inflow. The model streamflow output (the total inflow to Folsom Lake) is denoted by Q_N and is given by

$$Q_N = g(\underline{x}_N;\underline{u}_N,\underline{a}_N) \tag{3}$$

Given a set of parameter values, input-vector forecasts for interval $(t_0,t_f]$, and initial state-vector estimates $\underline{s}_{ij}(t_0)$ and $\underline{x}_i(t_0)$, the state equations may be integrated forward in time to produce simulations for $Q_N(t)$ for any t in the interval $(t_0,t_f]$. In fact, in deterministic applications of the model to operational forecasting of streamflow, this is the procedure followed to obtain forecasts in real time given a set of input-vector forecasts. However, model structure and estimates of input vectors and parameters carry uncertainty, which in some cases is significant. Certainly, for long forecast lead times t_f, the input-vector estimates carry large uncertainty.

Models for Forecast Uncertainty

Stochastic process theory offers a convenient framework for uncertainty modeling and it has been used extensively in the past in the context of operational forecasting (e.g., Bras and Rodriguez-Iturbe, 1985). Thus, for second-moment computations, state estimators have been designed and implemented operationally in the US which account explicitly for parametric and input uncertainty and produce estimates of forecast variance in real time (Georgakakos and Smith, 1990; and Sperfslage and Georgakakos, 1995). In addition, using ensemble forecasting techniques that consider input-vector forecast uncertainty, likely future streamflow time series are routinely produced by the River Forecast Centers of the U.S. National Weather Service several months in advance (Schaake and Larson, 1997). In the following we cast the operational streamflow-forecasting problem in a stochastic-process framework and present the relevant formulation used in this work.

Consider a sample space W with elements w called events, outcomes of a random experiment (e.g., coin toss). In this work we will consider a stochastic vector process $\underline{y}(t;w)$ as a sequence of random vectors of sample space W indexed by time t. Thus, for a given w, $\underline{y}(t;.)$ is a time series in the ordinary sense, and it represents a vector sample path or a vector realization of the stochastic vector process $y(t;w)$. We then consider the input vectors of the hydrologic model formulation (1)-(2) as stochastic processes $\underline{w}_{ij}(t;w)$ and $\underline{u}_i(t;w)$. The parameter vectors are considered as random vectors $\underline{q}(w)$ and $\underline{a}(w)$, and so are the initial state vector conditions $\underline{s}_{ij}(t_0;w)$ and $\underline{x}_i(t_0;w)$. A sample path for each of $\underline{w}$ and $\underline{u}$, and a sample value for each of $\underline{q}$, $\underline{a}$ and initial state vectors, correspond to each choice of w. For the same choice of w, solution of the set of Equations (1)-(2) yields a value for the state vectors for any time t in the forecast interval $(t_0,t_f]$ and Equation (3) provides corresponding values for $Q_N(t)$. The ensemble of all the solutions for all possible choices of w allows the probabilistic characterization of the state vectors and forecast output of the now stochastic differential equation system (1)-(2) and (3). Under certain restrictive conditions, which pertain to allowable model uncertainty terms to ensure that the state vectors follow a Markov process, the time-transition probability density function of the state vectors may be shown to satisfy the Fokker-Planck functional differential equation (Gard, 1988). In such cases, solution of that equation (numerically this is non-trivial) provides directly the probability law of the state vector process.

In this work we choose to work with the sample path approach in an ensemble-forecasting framework. The main reasons are that *(a)* we can specify the uncertainty in the input and parameters without constraints imposed on model structure (e.g., additivity and temporal independence of input errors), and *(b)* there is strong and time-varying inter dependence of the state vector elements in time <u>for all</u> sample paths, due to the physical system nature reflected in the model and its inputs. The later makes it impractical to characterize the state vector process through its full probability density law (joint probability density for all relevant times), while for reservoir management (see companion paper by Yao and A. Georgakakos 2000), such temporal dependence of the forecast flows is critical.

The sample-path philosophy, the basis of ensemble forecasting, allows the association of a single event w with a single <u>time series</u> of forecast input vectors (as opposed to association with just a single vector input at a certain time). Thus, for a large number of such time series of forecast inputs representing the full forecast range, and for parameter values and initial conditions sampled from pre specified distributions, the sample-path solution method for the stochastic differential state equations will provide all the necessary information for the state vectors and the output streamflow in the forecast interval. Issues that arise in practice are first a small number of forecast input realizations and second the specification of the parametric and initial condition distributions. The first influences the reliability of the forecasts for extreme events, and the second the sensitivity of the model forecasts on all the uncertainty sources.

Due to model parameter and structure uncertainty, and input spatial-interpolation uncertainty, flow simulations of the model (1)-(3) are different from observed flows even when the model is forced with observed input data. The distribution of monthly simulation errors for the hydrologic model was found to be approximately Gaussian for this application with means and variances as depicted in Figure 6. These errors must be taken into consideration in the ensemble-forecast methodology.

Denote by I_m and S_m the monthly mean and standard deviation of model residual errors for month m. Also, denote by $Q_N(t;w)$ a particular sample path of forecast inflows using the ensemble forecast methodology, with t in the forecast interval $(t_0,t_f]$ in units of days. Then, the rectified sample path for the given w is obtained from

$$Q'_N(t;w) = Q_N(t;w) + x(t;w') \tag{4}$$

with $x(t;w')$ representing a sample from the Gaussian distribution $N(I_m,S_m)$ where month m corresponds to the month of day t in the forecast period. For all t in month m, $x(t;w')$ remains constant. To ensure smooth transitions across months of large potential differences in samples $x(t;w')$ and given that the response time of the basin in high flows is less than a day, we adjusted the samples as follows. For the first day t_m of a new month m within the forecast period, the sample $x(t_m;w')$ was obtained from the sample $x(t_{m-1};w')$ of the previous day t_{m-1} in month $(m-1)$ and from a new sample $x'(t_m;w'')$ for the current day t_m in month m:

$$x(t_m;w') = \frac{x'(t_m;w'') + x(t_{m-1};w')}{2} \tag{5}$$

Use of Global Climate Model Information

Available global climate model (GCM) information is in the form of simulations or forecasts of monthly quantities such as surface precipitation, temperature, etc. In this application we use simulations for the historical period from the coupled global climate model (CGCM1) of the Canadian Centre for Climate Modeling and Analysis. Assuming that the CGCM1 model simulations contain information that can condition the sample paths of the input vectors of the hydrologic model so that under some conditions the reliability of the ensemble inflow forecasts increases, it is desirable to use them.

At first, we consider the GCM computational grid point whose monthly precipitation amount has the maximum correlation to the monthly mean areal precipitation over the Folsom Lake catchment. The maximum is computed over all the wet season months and the grid point does not change through time. Denote by P_1 the CGCM1 monthly precipitation forecast at the said computational grid point for the first month of the forecast horizon. Denote by z_p the standardized monthly anomaly of P_1 based on the historical simulations of CGCM1. The upper and lower terciles of that distribution may be determined from the historical CGCM1 simulations. Then, sample paths of rain plus melt and of potential evapotranspiration are selected from years for which the standardized anomalies of the CGCM1 simulations of monthly precipitation at the grid point of interest were in the same tercile as z_p. These sample paths start at the same day in the year as the forecast preparation day and are used as input to the hydrologic model to produce the ensemble reservoir-inflow forecast for the given forecast preparation day.

Validation of Ensemble Forecasts

The validation of probabilistic forecasts may only be done in an approximate manner under the assumption of statistical stationarity of the observed climate variables over a finite historical period (e.g., 20 or 30 years). Then, the forecast frequencies for given events (e.g., exceedance of a certain quantile of flow volume) may be compared to the observed frequencies.

Denote by q_j the daily inflows to Folsom Lake with $j=1,\ldots,N$, and N denoting the total number of days in the record. Define an observed sequence of flow volumes from

$$O_{y,m,i} = \sum_{j=k}^{k+n} q_j \Delta t \tag{6}$$

where y denotes the year, m denotes the month in the year, and i denotes the day of the start date of the inflow sequence that makes up the volume. There is a one-to-one correspondence between the sequence number k and the triplet of y, m, and i. Also, the volumes $O_{y,m,i}$ may be defined at fixed day intervals (e.g., every 5 days) so that i is constrained by

$$i = lN_d \text{ and } \quad i \le D_m$$

where l is an integer in a sequence of integers, N_d is the fixed sampling interval for the volume and D_m is the maximum number of days in month m.

At first we compute standardized anomalies for the inflow volumes by month of start date. The mean value O_m^m and standard deviation O_m^s for $O_{y,m,i}$ for month m are computed respectively as:

$$O_m^m = \frac{1}{N_m} \sum_{y,i} O_{y,m,i} \tag{7}$$

and

$$O_m^s = \sqrt{\frac{1}{N_m} \sum_{y,i} O_{y,m,i}^2 - (O_m^m)^2} \tag{8}$$

where the summation indices denote that summations are over the range of y and i values for a fixed m. The standardized anomalies of inflow volume are then,

$$o_{y,m,i} = \frac{O_{y,m,i} - O_m^m}{O_m^s} \tag{9}$$

From the ranked series of inflow standardized anomalies we may then determine o^q and o_q, upper and lower quantiles (e.g., terciles, quartiles, etc.) of the series, respectively. We wish to determine the reliability of a set of probabilistic forecasts of the events:

A: the inflow volume will exceed o^q
B: the inflow volume will be below o_q

Following the sample path approach described earlier, consider a set of ensemble forecasts $F_{y,m,i}^e$ of Folsom Lake inflow volume for a given forecast preparation time (FPT), characterized by year y, month m and day i, and with e denoting a particular sample path of the ensemble forecast. The sample frequencies $f_{y,m,i}^A$ and $f_{y,m,i}^B$ of forecasting events A and B may be computed as follows,

$$f_{y,m,i}^A = \frac{C(e : F_{y,m,i}^e > o^q)}{N_e} \tag{10}$$

and

$$f^{B}_{y,m,i} = \frac{C(e : F^{e}_{y,m,i} < o_q)}{N_e} \tag{11}$$

where N_e is the number of sample paths in the ensemble forecast, and *C( )* denotes the count of the sample paths for which the inequality in parenthesis is true.

Reliability of the ensemble forecasts in this context may be assessed by comparing the forecast frequencies *f* of A or B with the actual frequency computed from the observed data for those times for which *f* was in a set interval. The ensemble forecasts would be called reliable if the forecast and observed frequencies were equal within an estimation error tolerance. Thus, one may divide the interval *[0,1]* into 10 subintervals of equal length. For each FPT for which $f^{A}_{y,m,i}$ or $f^{B}_{y,m,i}$ was in a certain subinterval of *[0,1]* determine the frequency of observed inflow volumes satisfying event *A* or *B*, respectively, and plot one against the other for both *A* and *B* events and for all the subintervals.

Large deviations from equality for a particular interval signify unreliable forecasts for that subinterval of forecast frequencies and for that event (*A* or *B*). A goodness-of-fit test may be devised for each forecast subinterval assuming that the probability of finding a possible observed inflow volume value satisfying *A* (or *B*) in the subinterval is fixed and equal to the subinterval middle value. Under this assumption (e.g., Benjamin and Cornell, 1970), the number N_I of observations in a certain subinterval of *[0,1]* follows a binomial distribution with an expected value of $N_s p$ and a variance of $N_s p(1-p)$. N_s is the number of observed inflow volume samples and *p* is the constant probability of observing a certain sample.

Define the standardized residual ν from

$$n = \frac{N_I - N_s p}{\sqrt{N_s p(1-p)}} \tag{12}$$

Then, for $N_s p$ not too small (i.e., >10) for the central limit theorem to apply, *n* is approximately normal with zero mean and variance one, and n^2 is a c^2-distrbuted random variable with one degree of freedom. Using values of the cumulative c^2 distribution we can determine critical values n^{2*} such that

$$P[n^2 \geq n^{2*}] = a \tag{13}$$

where *P[ ]* denotes the probability of the event in square brackets and with *a* typically set equal to 0.05 or 0.10.

RESULTS AND DISCUSSION

The integrated system of Figure 1 was applied to the Folsom Lake watershed for the historical years (1964-1993) and for the future years (2001-2030). In the latter case both the CGCM1 control and the greenhouse-gas-increase scenarios were used. For the historical years, a simple inflow forecast model based on the regression of snowpack with past flow, temperature and other hydrometeorological variables was also used for comparison. Such a model is used routinely by the Bureau of Reclamation to provide guidance for Folsom Lake management. It runs on the first of each month from February through May with a forecast lead-time of several months. The model was made available by the Bureau of Reclamation for this study and it provides a baseline operational forecast procedure.

For both the historical and future years the ensemble-forecast procedure outlined in the previous section was used both with and without conditioning on CGCM1 simulations. Inflow forecasts out to 60 days with daily resolution were generated every five days for the available record; simulating real time forecast conditions. These ensemble inflow forecasts were then provided to the decision component of the system (Figure 1), which managed the reservoir based on predetermined manager preferences and computed non-inferior trade-off surfaces among the various plant objectives. This component also quantified management benefits for all the cases. The forecast results are presented and assessed in the following sections in terms of 30-day inflow volumes during the wet period October-April, while the corresponding management results are presented and assessed in Yao and Georgakakos (2000).

Historical Years

Figure 11 presents mean forecasts and associated observations for monthly-total Folsom Lake inflows for the months of March and April, and for all the historical years. The operational forecasts of the simple regression model are compared to the means of the ensemble-forecast procedure for both cases of conditioning on CGCM1 simulations. The results show that for the forecast preparation time of March 1, the simple model forecasts (they do use recent observed flows) are as skillful as the mean forecasts of the ensemble forecast procedure. In 22 out of 54 cases the ensemble means of each of the ensemble-forecast cases are closer than the simple model forecasts to the observed values. In 20 cases the latter forecasts are closer to the observed values than either of the two ensemble forecast means. For high observed volumes there is no significant difference among forecast procedures in terms of skill. For low observed volumes, the means of the ensemble forecasts conditioned on CGCM1 simulations were closer to the observations in a substantially larger number of cases than the other mean forecasts (12 out of 18 cases with less than 200,000 ac-ft observed monthly volume as opposed to 6 and 4 out of 18 cases for the other two forecast methods). The results of Figure 11 also show that during the months of spring melt, mean forecasts are likely to do well in low and medium monthly-volume situations but they will underestimate high flows substantially. Similar results (not shown) have been obtained for other forecast preparation times in the wet season October-May.

Georgakakos, A., et al. (1998) has shown that forecast uncertainty can be profitably used by the decision component of Figure 1. For this reason, box plots of the ensemble Folsom Lake monthly inflows are constructed for the cases (a) without and (b) with CGCM1 conditioning, and are displayed with the corresponding observations for various forecast preparation times. The results are exemplified in Figure 12 for a 1 March forecast preparation time. A cursory review shows that the observations in all but one case are within the 5-95 percent forecast interval in both cases of using and not using CGCM1 conditioning. Differences between the two types of ensemble forecasts shown in Figures 12a and 12b are apparently small. However, reliability diagrams based on the theory developed in section 3 must be constructed to allow a detailed analysis of reliability by flow magnitude.

Figure 13 shows the reliability diagrams for the upper (upper panels) and lower (lower panels) terciles of observed 30-day Folsom Lake inflow and for the ensemble forecasts without (left panels) and with (right panels) CGCM1 conditioning. Shown in each panel are the expected values and bounds which correspond to $\alpha=0.05$ in Equation (13) for the observed frequencies for each forecast frequency range together with the actual observed frequencies (filled squares). The reliability diagrams have been constructed from ensemble forecasts with forecast preparation times every five days throughout the historical period 10/1/1964-9/30/1993, and with a maximum forecast lead-time of 30 days with daily resolution. Generally, reliable forecasts of 30-day inflow have been obtained, particularly for the upper tercile of the frequency distribution of observed flows. The difference between the reliability of ensemble forecasts without and with CGCM1 conditioning is significant for the lower tercile of the frequency distribution of the observations, and for forecasts with high forecast confidence for which conditioning on CGCM1 simulations improves forecast reliability.

A single score may be computed by producing an estimate of the weighted square deviation of the forecast frequencies from the observed frequencies for each decile. This is a reliability score and it is part of the Brier score used extensively in Meteorology (e.g., Wilks, 1995). It is appropriate for intercomparing forecasts issued under the same conditions. Lower values of the score imply more reliable forecasts. The values of the score for the unconditional ensemble forecasts of 30-, 60-, and 90-day volumes for the historical period and for the upper and lower terciles of the observed inflow volume distribution are given in Table 1. Although all the forecasts show a high degree of reliability (generally low reliability score values), substantial improvement is obtained by conditioning on CGCM forecasts, especially for longer forecast lead times.

The reliability of forecasts issued during the wet period October-April is indicated by the reliability diagrams of Figure 14. For the upper tercile, generally somewhat more reliable forecasts have been obtained both with and without CGCM1 conditioning. Somewhat less reliable inflow volume forecasts were obtained during periods when the observed volumes were in the lower tercile of their distribution at the 5% confidence level.

Future Years

Analyses analogous to that discussed in section 4.1 was performed for future years (10/1/2001 - 9/30/2030) and for both control and greenhouse-gas-increase scenarios. To study the sensitivity of the ensemble forecast methodology to the particular set of years from which the hydrologic model forcing was drawn, ensemble forecasts for the future period were produced for both cases: (a) using historical years and (b) using years from the future period and the scenario under study.

Ensemble generation with forcing from historical years

Reliability diagrams are shown in Figures 15 and 16 for the control scenario and the greenhouse-gas-increase scenario, respectively, when years from the historical period 1964-1993 were used to provide input to the hydrologic model for the generation of ensemble inflow forecasts. The control scenario results of Figure 15 indicate that for the lower-tercile inflows the control scenario forecasts are as reliable (at the 0.05 confidence level) as those of the historical period (Figure 13). Loss of reliability for the range of forecast frequencies (0.2 - 0.8) is indicated for the upper tercile cases of the control scenario when there is no conditioning on CGCM1 output. The control-scenario upper-tercile case with CGCM1 conditioning shows similar reliability behavior than the analogous historical one, with the forecast result for the range (0.7-0.8) being significantly far from the 0.05 confidence-level bounds predicted by theory.

Table 1. Reliability score of forecasts issued for historical years.

Accum. Period (Days)	*Unconditional Ensemble*	*CGCM1-Conditioned Ensemble*
Upper Tercile		
30	0.004	0.004
60	0.008	0.005
90	0.010	0.006
Lower Tercile		
30	0.011	0.014
60	0.015	0.012
90	0.016	0.011

Significant loss of reliability (as compared to the control and historical cases) is indicated for the case of greenhouse-gas-increase scenario and for observed inflow volumes in the upper tercile of their distribution. The forecast frequencies underestimate the actual ones for almost all the frequency ranges. Reliable forecasts have been obtain only for the forecast-frequency mid-ranges (0.4-0.6) and (0.4-0.5) for the cases without and with CGCM1 conditioning, respectively. Using hydrologic-model forcing from climatic periods with a different climate from the one for which forecasts are issued, and in situations where there exist strong climatic trends in precipitation (as projected by CGCM1 in this case study), results in unreliable forecasts during periods of high inflow volume. It is noted that under this greenhouse-gas-increase scenario and for inflow volumes in the lower tercile of their distribution, the ensemble forecasts with and without CGCM1 conditioning overestimate the forecast frequencies in the range (0.8-1.0). No significant difference is noted between the two CGCM1-conditioning cases.

Ensemble generation with forcing from future years

Figure 17 is analogous to Figure 16 but with hydrologic-model forcing obtained from the future period 2001-2030 for which ensemble forecasts are generated. It is apparent that the reliability of the forecasts for both the cases with and without CGCM1 conditioning is substantially increased. More reliable ensemble forecasts are obtained when conditioning on CGCM1 simulations is performed. The forecasts during periods of observed inflow volumes in the upper third (tercile) of their distribution are more reliable than those in the lower third. For lower-tercile periods and independent on CGCM1 conditioning, the ensemble forecast frequencies overestimate (at the 5% confidence level) the actual observed ones for high frequency ranges (>0.6).

CONCLUSIONS AND RECOMMENDATIONS

Application of an integrated forecast-control system appropriate to ingest climate model information for the improved management of multipurpose reservoirs (Figure 1) is documented for the Folsom Lake facility in central mountainous California. The drainage basin generating inflows to Folsom Lake is substantially smaller than the typical climate model grid scale and it has short response time for high flow events. Snow accumulation and ablation is an important process for the hydrologic budget of the physical system. The application area is characterized by highly seasonal climate with distinct wet and dry periods. Hydrologic modeling for the Folsom lake catchment with existing daily data produced good daily flow simulations for the historical period (1964-1993) with adequate reproduction of daily flow variability (64% observation-variance explained) but with seasonal bias in winter (over estimation) and spring (underestimation) due to snow model application deficiencies. An ensemble inflow forecasting methodology was used based on the operational NWS ensemble streamflow prediction technique complemented by the addition of an uncertainty term in the forecasts to account for monthly model error second moments.

The following are important conclusions from the research work pertaining to the forecast component of the integrated system of Figure 1:

(a) The mean forecasts of 30-day volume produced by the integrated system with and without climate information are comparable overall to the forecasts of an operational regression model that uses past flows, snow pack amount, and other observed hydrometeorological variables to forecast reservoir inflow volumes. Global climate model information from the Canadian coupled global climate model CGCM1 benefits the mean forecasts significantly mainly for low observed 30-day inflow volumes.

(b) Reliable ensemble forecasts (at the 5% confidence level) are generated for the historical period (1964-1993) with and without downscaled CGCM1 information, particularly for observed inflow volumes in the upper tercile of their distribution and for decile forecast-frequency ranges greater than 0.5. Improvement of the reliability for low observed inflow volumes is indicated for several decile forecast-frequency ranges when CGCM1 simulations are considered in producing ensemble hydrologic forecasts.

(c) During the period from 2001-2030 and for a 1-%-per-annum increase of greenhouse-gasses, CGCM1 simulations indicate a significant upward precipitation trend in the Folsom Lake region. Future Folsom inflows consistent with the CGCM1 simulated precipitation trends are substantially greater than those in the historical record for all cumulative frequencies of occurrence except those at the extreme tails of their distribution.

(d) When the ensemble inflow forecasts were generated from hydrologic-model forcing in years past the future period and for high observed inflow volumes, the control-scenario application of the integrated forecast-control system tends to under-forecast the low deciles and over-forecast the high deciles of the forecast-frequency ranges. Generally, for this scenario, CGCM1 information is beneficial for periods of both high (upper tercile) and low (lower tercile) observed inflow volumes.

(e) For the case of the greenhouse-gas-increase scenario, generating ensemble inflow forecasts from hydrologic forcing obtained from years with a different climate than the study years generates unreliable forecasts, especially during periods of high observed inflow volumes. Decile forecast frequency ranges substantially underestimate actual observed frequencies of high inflows. When ensemble forecast inflows are generated using hydrologic-model input from years with the same climate as the study period, then reliable forecasts are generated for most decile forecast frequency ranges, and especially for high observed inflow volume periods. For this case, conditioning on CGCM1 simulations improves the forecast performance.

It is emphasized that these conclusions pertain to the integrated forecast control system of Figure 1 with the particular downscaling and ensemble generation methods described. It is the authors' experience with a variety of approaches that the proposed system maximizes the information in the available data and models across scales while accounting for uncertainty in a consistent manner. The particular presentation in this paper used a single simulation from a particular climate model and it may be generalized to allow ingesting of an ensemble of climate simulations (work in progress).

The production of future "observed" hydroclimatic variables relies on the CGCM1 climate-model simulations. Therefore, the conclusions regarding the future scenarios are appropriate for this basin and its climate, if the precipitation as simulated by the particular climate model used is indicative of future precipitation trends over the basin. It is also anticipated that if the climatic changes simulated by the climate model were indicative of real changes, re-calibration of the operational hydrologic model would be necessary to enhance forecast reliability during these projected high-flow future periods. No such re-calibration was performed in this study, and, thus, the model error statistics used for future scenarios in this work may be artificially inflated compared to an actual implementation.

The most important next step is to use the ensemble Folsom Lake forecasts within the decision component to measure the performance of the reservoir management process under the various cases of forecasting and GCM-conditioning for present and future climate scenarios. This is accomplished in the companion paper, Yao and Georgakakos, A. (2000). With respect to the application of the integrated system to Folsom Lake, an important next step is the utilization of ensemble forecasts from various climate models (rather than a single realization from a single climate model) to condition the ensemble inflow forecasts. For the greater Sacramento River region, an appropriate subsequent study should also involve the Shasta and Oroville reservoirs and aim at the application of the integrated methodology over the entire system, treated as a hydrologically and operationally interconnected system. In such a case, the scale of dimate forecasts approaches the scale of the application area and scale errors are minimized. Application to other areas with different hydroclimatic regimes or plant management objectives is also warranted.

Acknowledgements: This work was sponsored by USGS Grant 98HQAG2118 of the Georgia Water Resources Institute as part of the National Climate Assessment, Water Sector. Additional support was provided by the NOAA California Applications Project (NA77-RJ0453) of the

Scripps Institution of Oceanography, UCSD. We thank Rob Hartman and Eric Stremm of the California-Nevada River Forecast Center, and Bill Pennington, Lloyd Peterson, David Reed and John Burke of the Bureau of Reclamation for their valuable feedback, support and assistance with data and operational models. We also gratefully acknowledge the comments and suggestions of Nick Graham of HRC and Scripps Institution of Oceanography on climate model issues, and of Aris Georgakakos and Huaming Yao of the Georgia Water Resources Institute on the use of forecasts by the decision model.

REFERENCES

Anderson, E.A., 1973: "National Weather Service River Forecast System – Snow accumulation and ablation model." NOAA Technical Memorandum NWS HYDRO-17, Office of Hydrology, National Weather Service, NOAA, Silver Spring, MD.

Anderson, J.L., 1999: A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate*, **9**, 1518-1530.

Anderson, J.L., and W.F. Stern, 1996: Evaluating the potential predictive utility of ensemble forecasts. *J. Climate*, **9**, 260-269.

Anderson, J.L., van den Dool, H., Barnston, A., Chen, W., Stern, W., and J. Ploshay, 1999: Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bull. Amer. Meteor. Soc.*, **80**, 1349-1361.

Barnett, T.P., and co-authors, 1999: Detection and attribution of recent climate change, A status report. *Bull. Amer. Meteor. Soc.*, **80**, 2631-2659.

Barnston, A.G., Glantz, M.H., and Y. He, 1999: Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997-1998 El Niño Episode and the 1998 La Niña onset. *Bull. Amer. Meteor. Soc.*, **80**, 217-243.

Benjamin, J.R., and C.A. Cornell, 1970: *Probability, Statistics, and Decision for Civil Engineers.* McGraw-Hill Co., New York, 684pp.

Bras, R.L., and I. Rordiguez-Iturbe, 1985: *Random Functions and Hydrology.* Addison-Wesley Publishing Co., Reading, Mass., 559pp.

Buizza, R., Hollingsworth, A., Lalaurette, F., and A. Ghelli, 1999: Probabilistic predictions of precipitation using the ECMWF ensemble prediction system. *Wea. Forecasting*, **14**, 168-189.

Busuioc, A., von Storch, H., and R. Schnur, 1999: Varification of GCM-generated regional seasonal precipitation for current climate and of statistical downscaling estimates under changing climate conditions. *J. Climate*, **12**, 258-272.

Day, G.N., 1985: Extended streamflow forecasting using the NWSRFS. *J. Water Resour. Plan. Manage.*, **111**(2), 157-170.

Fread, D.L., Ingram, J., Schaake, J., and D.T. Braatz, 1999: Recent experience with ensemble streamflow prediction in the Des Moines River basin. *Preprints 14th AMS Conf. on Hydrology,* American Meteorological Society, 4-7.

Gard, T.C., 1988: *Introduction to Stochastic Differential Equations*. Marcel Dekker, Inc., New York, 234 pp.

Georgakakos, A.P., H. Yao, 1993: New control concepts for uncertain water resources systems, 1, Theory. *Water Resour. Res.*, **29**(6), 1505-1516.

Georgakakos, A.P., Mullusky, M.G., Yao, H., and K.P. Georgakakos, 1998b: Impacts of climate variability on the operational forecast and management of the upper Des Moines River basin. *Water Resour. Res.*, 34(4), 799-821.

Georgakakos, K.P., 1998: Flooding attributable to El Niño. *WMO Bulletin,* **47**, 356-360.

Georgakakos, K.P., 1986: "A Generalized Stochastic Hydrometeorological Model for Flood and Flash-Flood Forecasting, 1 Formulation and 2 Case Studies," *Water Resour. Res.*, **22**(13), 2085-2106.

Georgakakos, K.P., Bae, D.H., and D.R. Cayan, 1995: "Hydroclimatology of Continental Watersheds, 1, Temporal Analyses," *Water Resour. Res.*, **31**(3), 655-675.

Georgakakos, K.P., Bae, D.-H., Mullusky, M.G., and A.P. Georgakakos, 1995: Chapter II-2, Hydrologic variability in Midwestern drainage basins; Diagnosis, prediction and control. In *Preparing for Global Change: A Midwestern Perspective*, Eds: G.R. Carmichael, G.E. Folk, and J.L. Schnoor. SPB Academic Publishing, Amsterdam, The Netherlands: 61-90.

Georgakakos, K.P. and R.L. Bras, 1982: "Real-Time, Statistically Linearized, Adaptive Flood Routing," *Water Resour. Res.*, **18**(3), 513-524.

Georgakakos, K.P., Georgakakos, A.P., and N.E. Graham, 1998a: Assessment of benefits of climate forecasts for reservoir management in the GCIP region. *GEWEX News*, **8**(3), 5-7.

Georgakakos, K.P., and G.F. Smith, 1990: On improved hydrologic forecasting – Results from a WMO real-time forecasting experiment. *J. Hydrology,* **114**, 17-45.

Georgakakos, K.P., D. Tsintikidis, J.A. Sperfslage, D.E. Smith and T. Carpenter, 1999: “Gauging Hourly Precipitation Within the Folsom Lake Watershed,” HRC Limited Distribution Report No. 9, Hydrologic Research Center, San Diego, California, 51pp.

Jamieson, D.G., (ed.), 1996: Special issue on decision-support systems. *J. Hydrol.*, **177**, 161-291.

Krzysztofowicz, R., 1999: Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Wat. Resour. Res.*, **35**, 2739-2750.

Li, Z.-X., 1999: Ensemble atmospheric GCM simulation of climate interannual variability from 1979-1994. *J. Climate,* **12**, 986-1001.

Loucks, D.P., (ed.), 1989: *Systems Analysis for Water Resources Management, Closing the Gap Between Theory and Practice.* IAHS Publication No. 180. International Assoc. of Hydrological Sciences, Institute of Hydrology, Wallingford, UK, 303pp.

Mason, S.J., and co-authors, 1999: The IRI seasonal climate prediction system and the 1997/1998 El Nino event. *Bull. Amer. Meteor. Soc.*, **80**, 1853-1873.

Mearns, L.O., Gleick, P.H., and S.H. Schneider, 1990: Climate forecasting. In *Climate Change and U.S. Water Resources,* edited by P.E. Waggoner, John Wiley, New York, 41-73.

Meehl, G.A., and co-authors, 2000: Trends in extreme weather and climate events, Issues related to modeling extremes in projections of future climate change. *Bull. Amer. Meteor. Soc.,* **81**, 427-436.

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

Perica, S., Marcus, M., Schaake, J., and D.-J. Seo, 1999: Accounting for hydrologic model errors in ensemble streamflow prediction. *Preprints 14th AMS Conf. on Hydrology,* American Meteorological Society, 12-14.

Petersen, A.C., 2000: Philosophy of climate science. *Bull. Amer. Meteor. Soc.,* **81**, 265-271.

Rind, D., 1999: Complexity and climate. *Science,* **284**, 105-107.

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

Sailor, D.J., and X. Li, 1999: A semiempirical downscaling approach for predicting regional temperature impacts associated with climate change. *J. Climate,* **12**, 103-114.

Schaake, J., and L. Larson, 1998: Ensemble streamflow prediction (ESP), Progress and research needs. *Preprints Special Symposium on Hydrology,* Am. Meteorol. Soc., Boston, Mass.

Sinclair, M.R., and I.G. Watterson, 1999: Objective assessment of extratropical weather systems in simulated climates. *J. Climate,* **12**, 3467-3485.

Smith, J.A., Day, G.N., and M.D. Kane, 1991: A nonparametric framework for long-range streamflow forecasting. *WMO/TD-428*, World Meteorol. Organ., Geneva, Switzerland, 28pp.

Sperfslage, J.A., and K.P. Georgakakos, 1996: Operational implementation of the Hydrologic Forecast System (HFS) Operation as part of the National Weather Service River Forecast System NWSRFS). HRC Technical Report No. 1. Hydrologic Research Center, San Diego, California, 213 pp.

Tsintikidis, D., Georgakakos, K.P., Sperfslage, J.A., Smith, D.E., and T.M. Carpenter, 2000: Precipitation uncertainty and raingauge network design within the Folsom Lake watershed. *J. Hydrol. Eng.*, ASCE (in review).

von Storch, H., 1999: On the use of "Inflation" in statistical downscaling. *J. Climate*, **12**, 3505-3506.

Wolock, D.M., McCabe, Jr., G.J., Tasker, G.D., and M.E. Moss, 1993: Effects of climate change on water resources in the Delaware River Basin. *Water Resour. Bull.,* **29**(3), 475-486.

Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, New York, 467 pp.

Yao, H., and A.P. Georgakakos, 1993: New control concepts for uncertain water resources systems, 2, Reservoir management. *Water Resour. Res.,* **29**(6), 1517-1526.

Yao, H., and A.P. Georgakakos, 2000: Assessment of Folsom Lake response to historical and potential future climate scenarios, 2, Reservoir management. *J. Hydrol.,* this issue.

Yu, J.-Y., and C.R. Mechoso, 1999: A discussion on the errors in the surface heat fluxes simulated by a coupled GCM. *J. Climate,* **12**, 416-426.

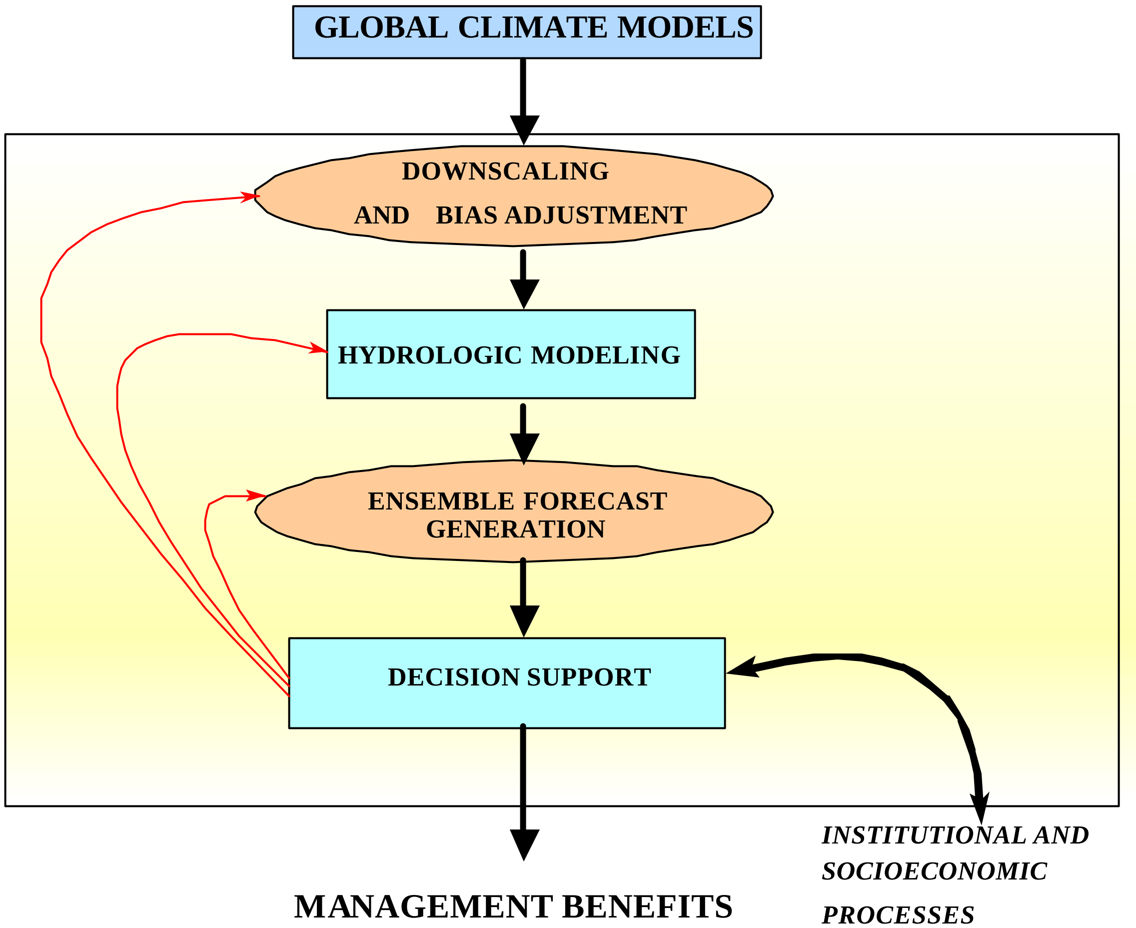


Figure 1. Integrated forecast-control system for quantifying benefits of forecast information to operational water resources management at reservoir sites.

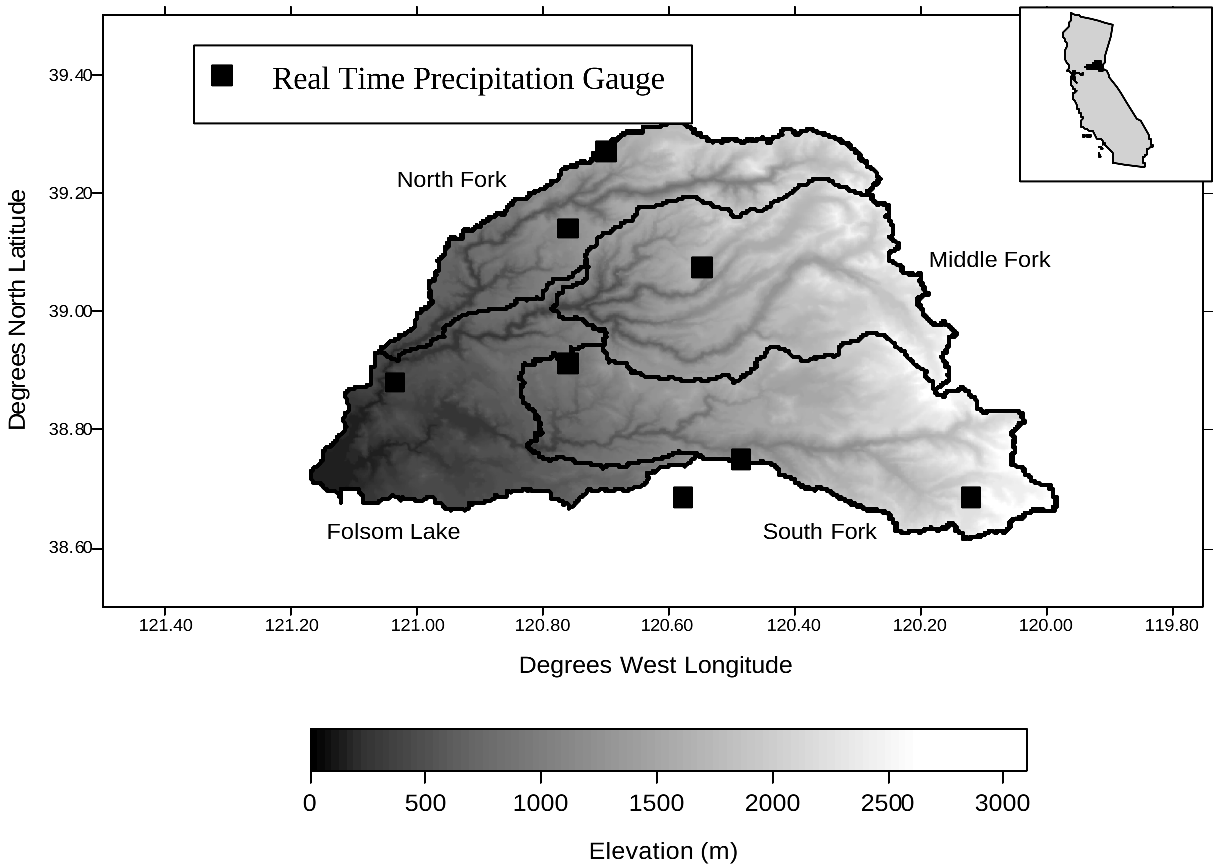


Figure 2. Folsom Lake catchment and its location in California (inset). The terrain elevation is shown together with the sub-catchments of the North, Middle and South Fork, and with the operational precipitation gauge stations.

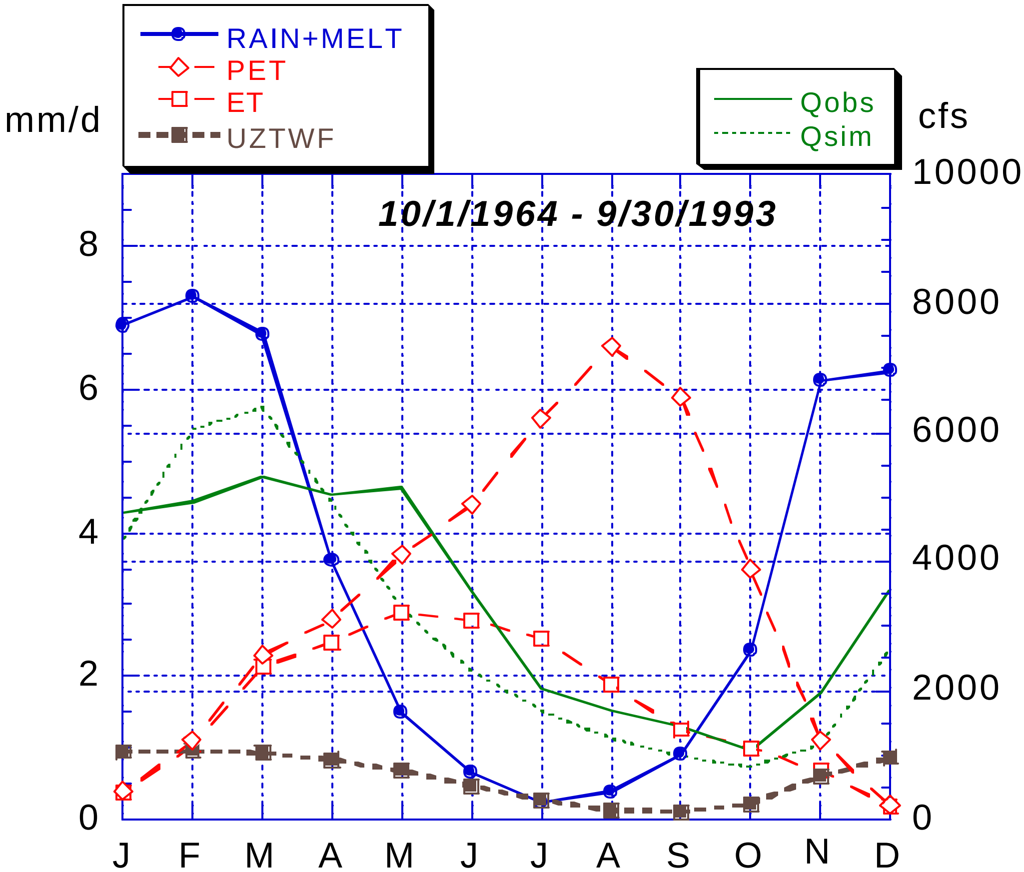


Figure 3. Annual cycles of monthly average rain plus melt (mm/d), potential and actual evapotranspiration (ET) (mm/d), and upper soil fractional saturation (dimensionless). On the right ordinate observed and simulated Lake Folsom inflow are shown in cfs. Period of record: 10/1/1964-9/30/1993.

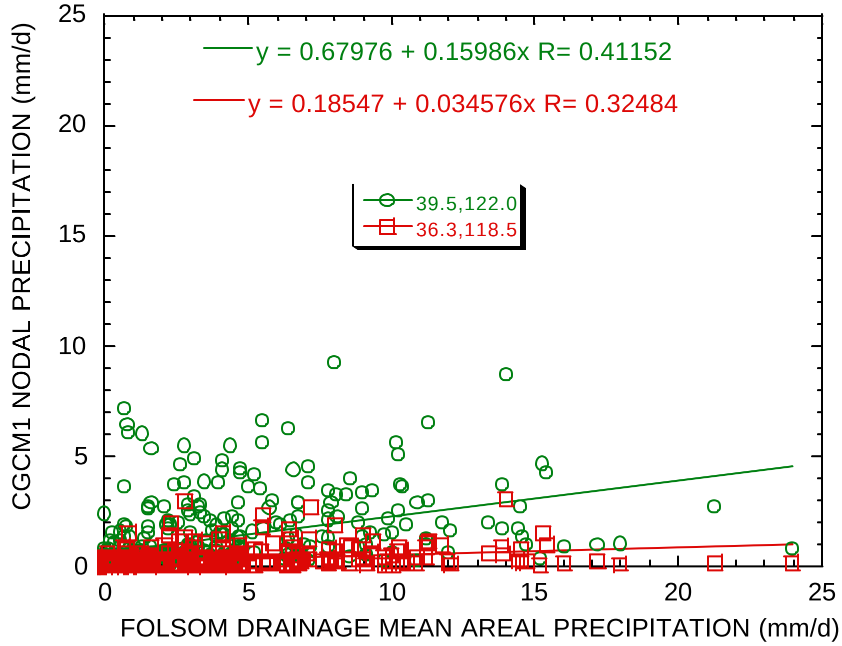


Figure 4. Association of wet-season monthly-averaged precipitation between observed mean areal precipitation over the Folsom Lake catchment and precipitation at the two closest CGCM1 nodes to the catchment. Period of record: 10/1/1964-9/30/1993.

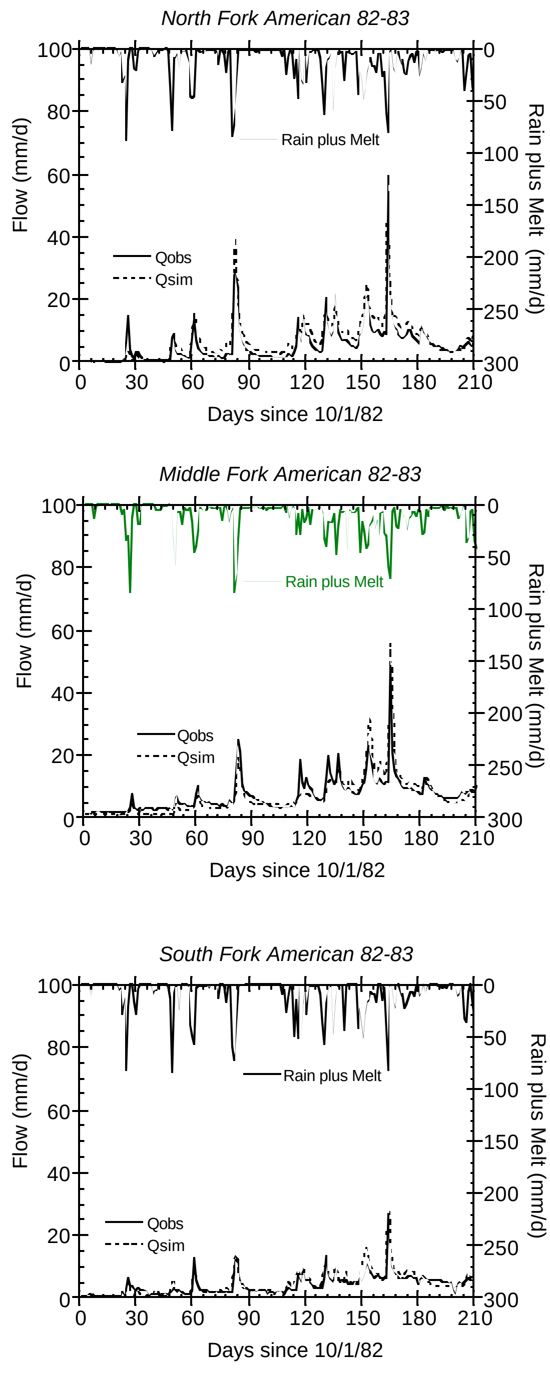


Figure 5. Observations and simulations of flow for the 1982-1983 wet season and for the gauge sites of North Fork (a), Middle Fork (b) and South Fork (c). Rain plus melt and potential evapotranspiration (PET) are shown inverted on the right ordinate.

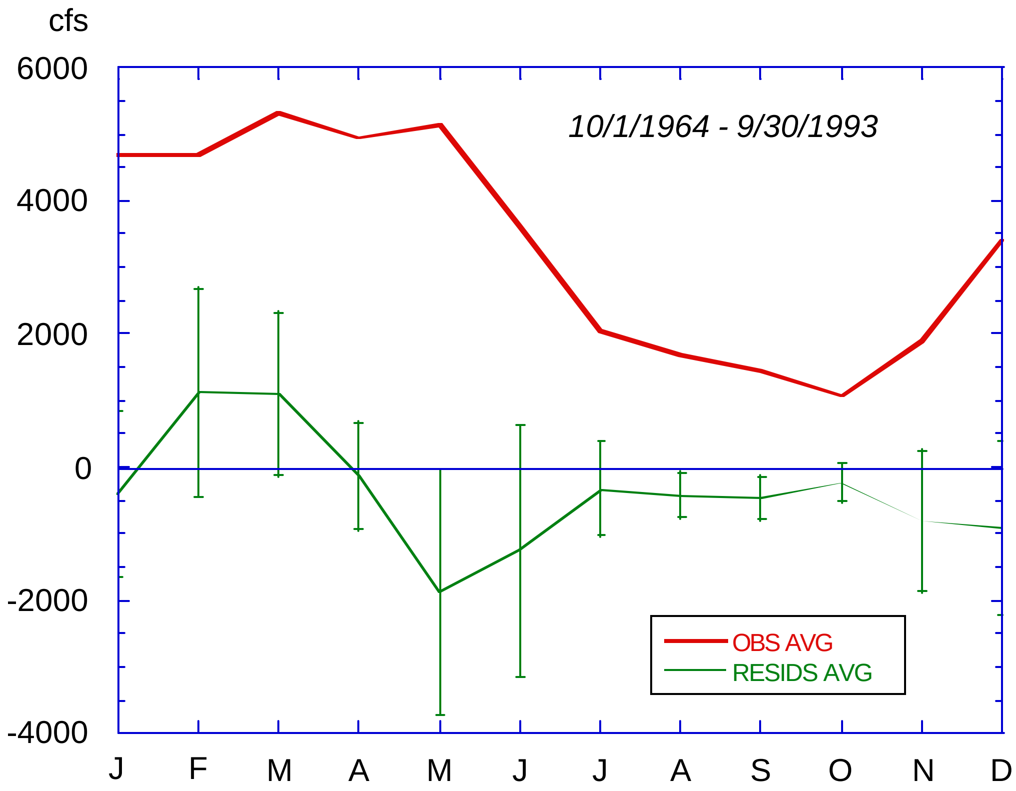


Figure 6. Annual cycle of monthly-averaged observed Folsom Lake inflow with corresponding means and standard deviations of the monthly-averaged simulation residuals (simulation-observation). Period of record: 10/1/1964-9/30/1993.

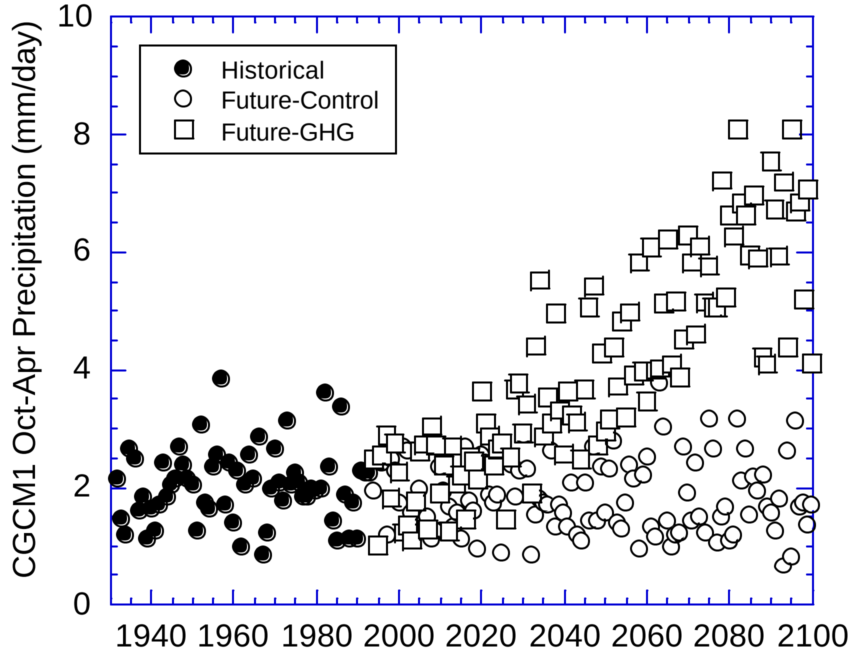


Figure 7. Monthly-averaged CGCM1-simulated precipitation for the period 1930-2100. Results from both future control and greenhouse-gas-increase scenarios are shown. Precipitation estimates are based on the CGCM1 precipitation output for the two closest nodes to the Folsom Lake catchment.

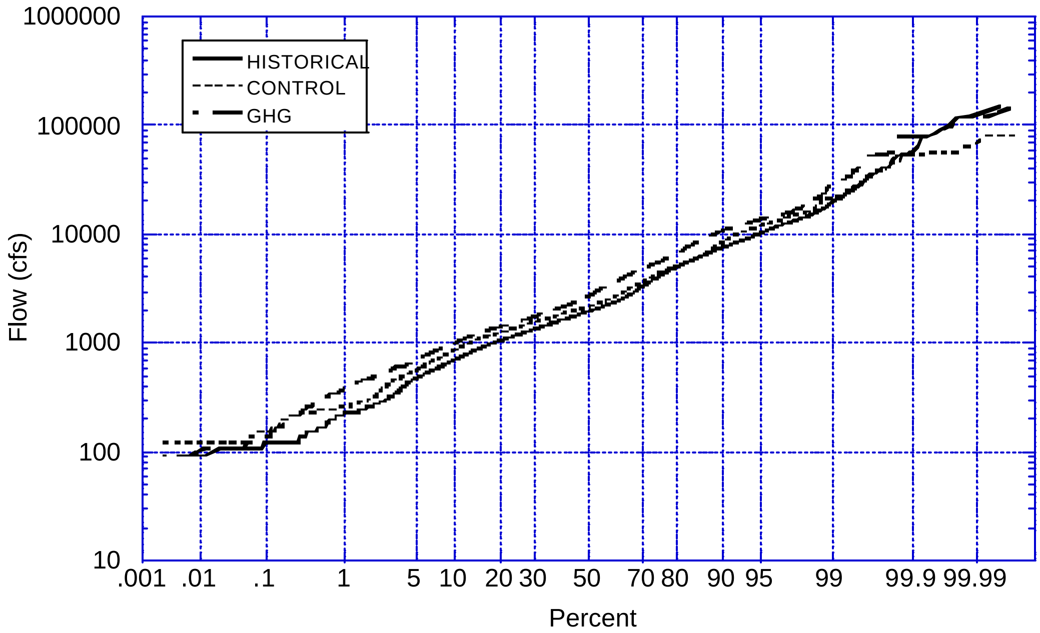


Figure 8. Cumulative frequency distributions for historical (1964-1993) and for generated future (1993-2050) Folsom Lake inflows. Inflows corresponding to both CGCM1 control and greenhouse-gas-increase scenarios are shown.

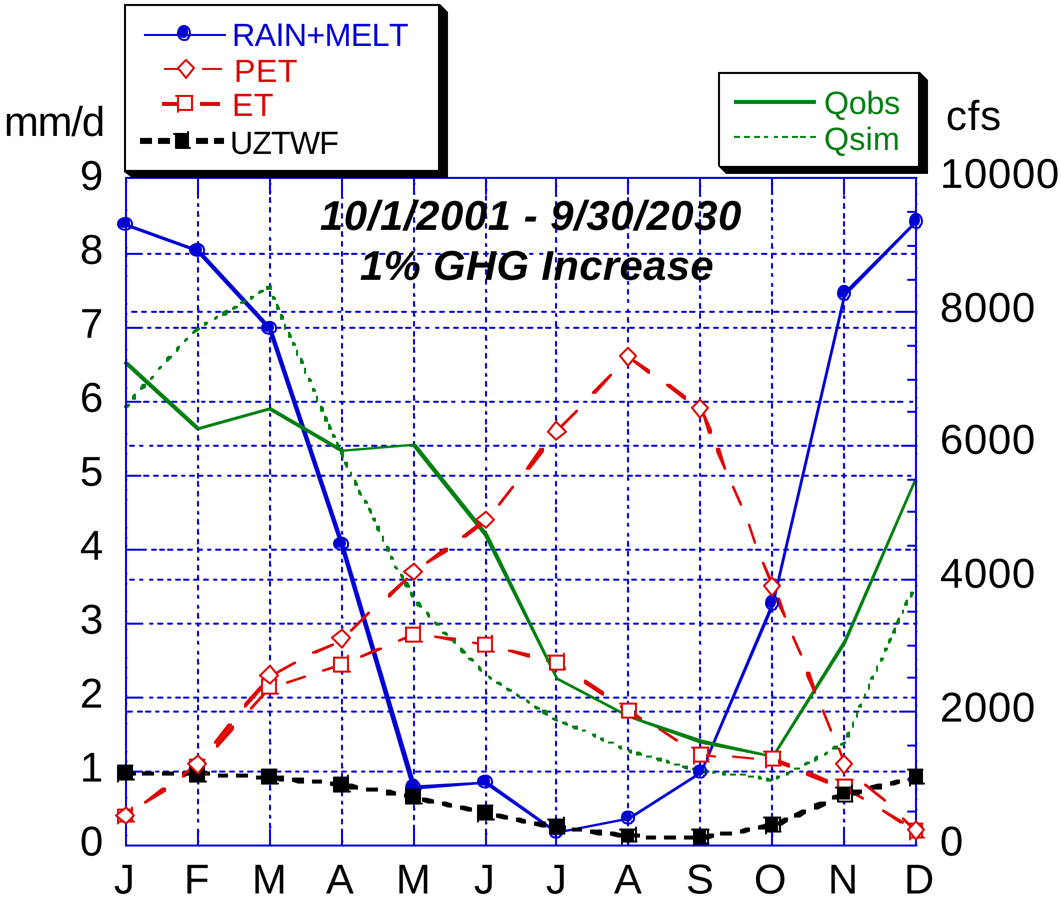


Figure 9. As in Figure 3 but for the period 10/1/2001 - 9/30/2030 and for a greenhouse-gas-increase scenario.

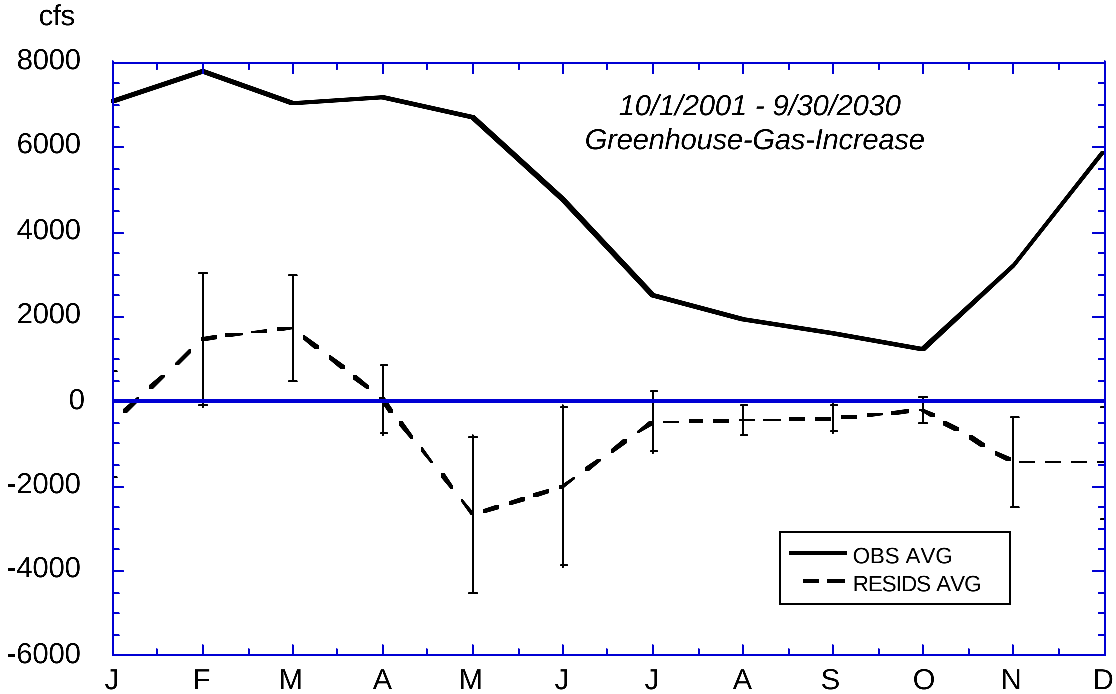


Figure 10. As in Figure 6 but for the greenhouse-gas-increase scenario and future period: 10/1/2001 - 9/30/2030.

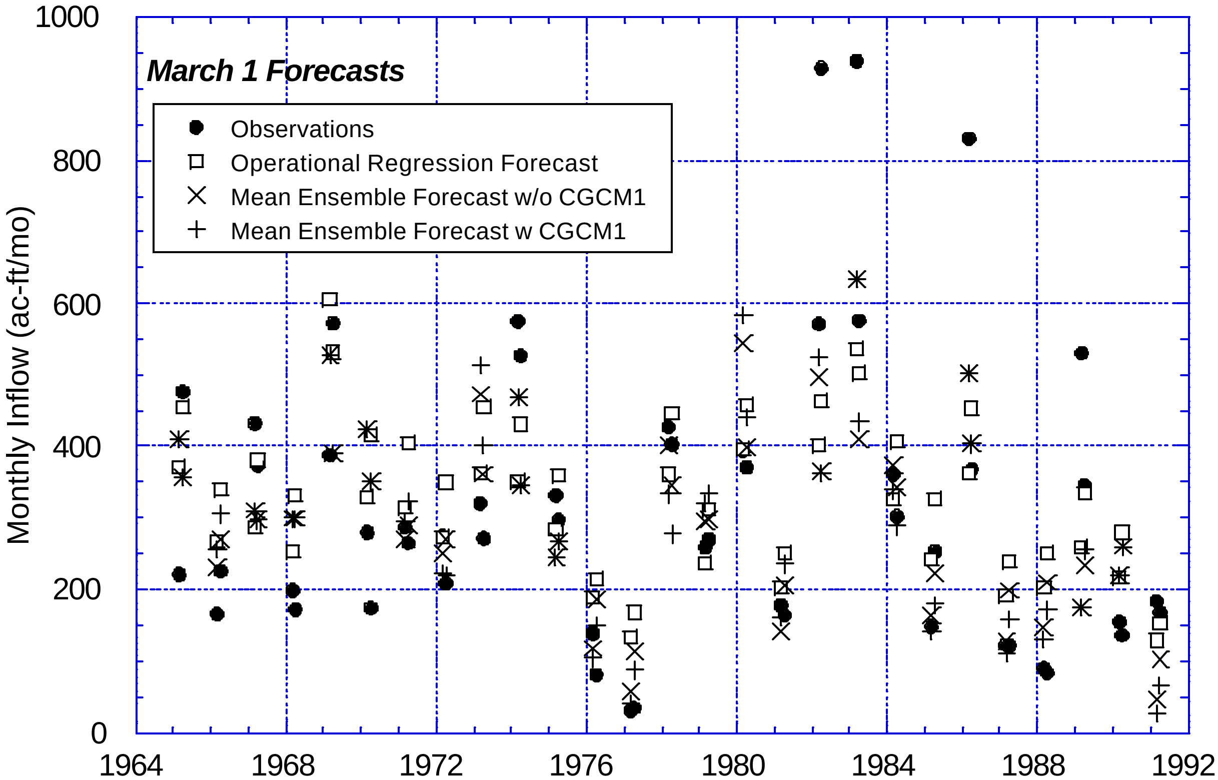
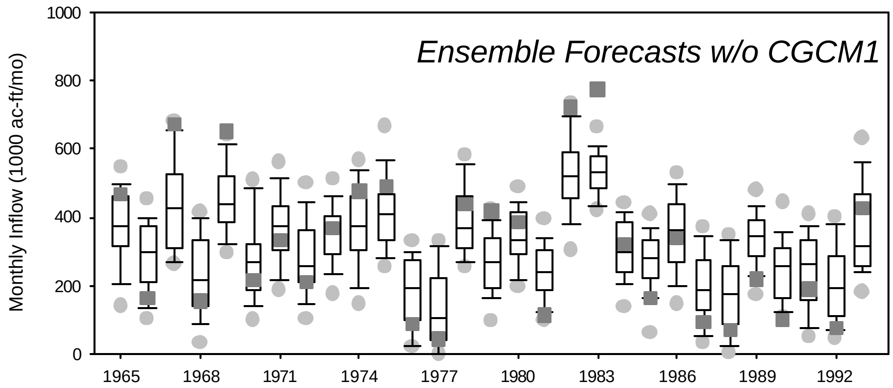
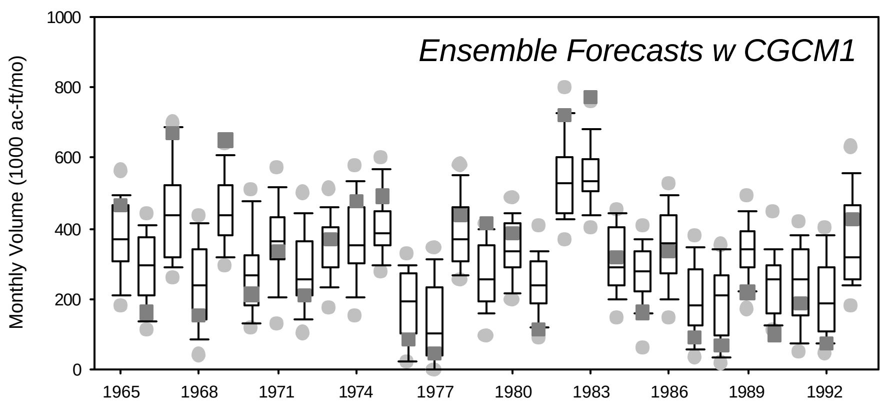


Figure 11. Monthly inflow forecasts and corresponding observations (acre-feet/month) for various models for Folsom Lake with a forecast preparation time of 1 March and with a maximum forecast lead time of two months. For ensemble forecast models the ensemble mean is shown.



(a)



(b)

Figure 12. Box plots constructed from ensemble forecasts of monthly Folsom Lake inflow (acre-feet/month) and corresponding observations. The forecast median, 25-75 percent bounds, 10-90 percent bounds and 5-95 percent bounds (filled circles) are shown for the period 1965 - 1993. Forecast preparation time is on 1 March and the maximum forecast lead time is two months. (a) Ensemble forecasts w/o CGCM1 and (b) ensemble forecasts with CGCM1.

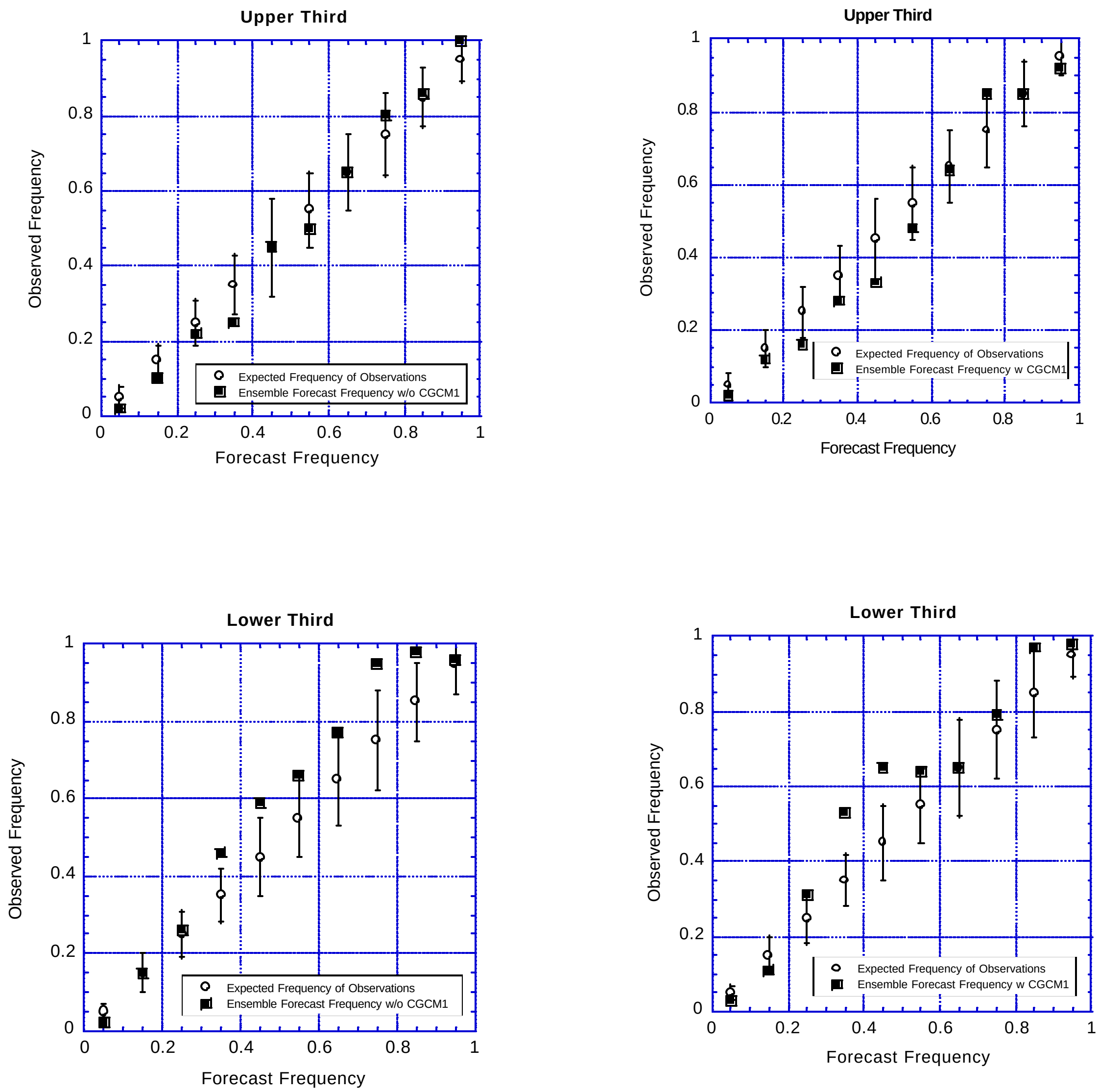


Figure 13. Reliability diagrams for ensemble forecasts of Folsom Lake 30-day inflow without (left panels) and with (right panels) CGCM1 conditioning and for periods with observed 30-day inflow in the upper (upper panels) and lower (lower panels) tercile of their frequency distribution. The expected values and the 5%-confidence bounds are shown for the observed frequencies for each forecast frequency interval. The actual observed frequencies for each forecast frequency are shown in filled squares.

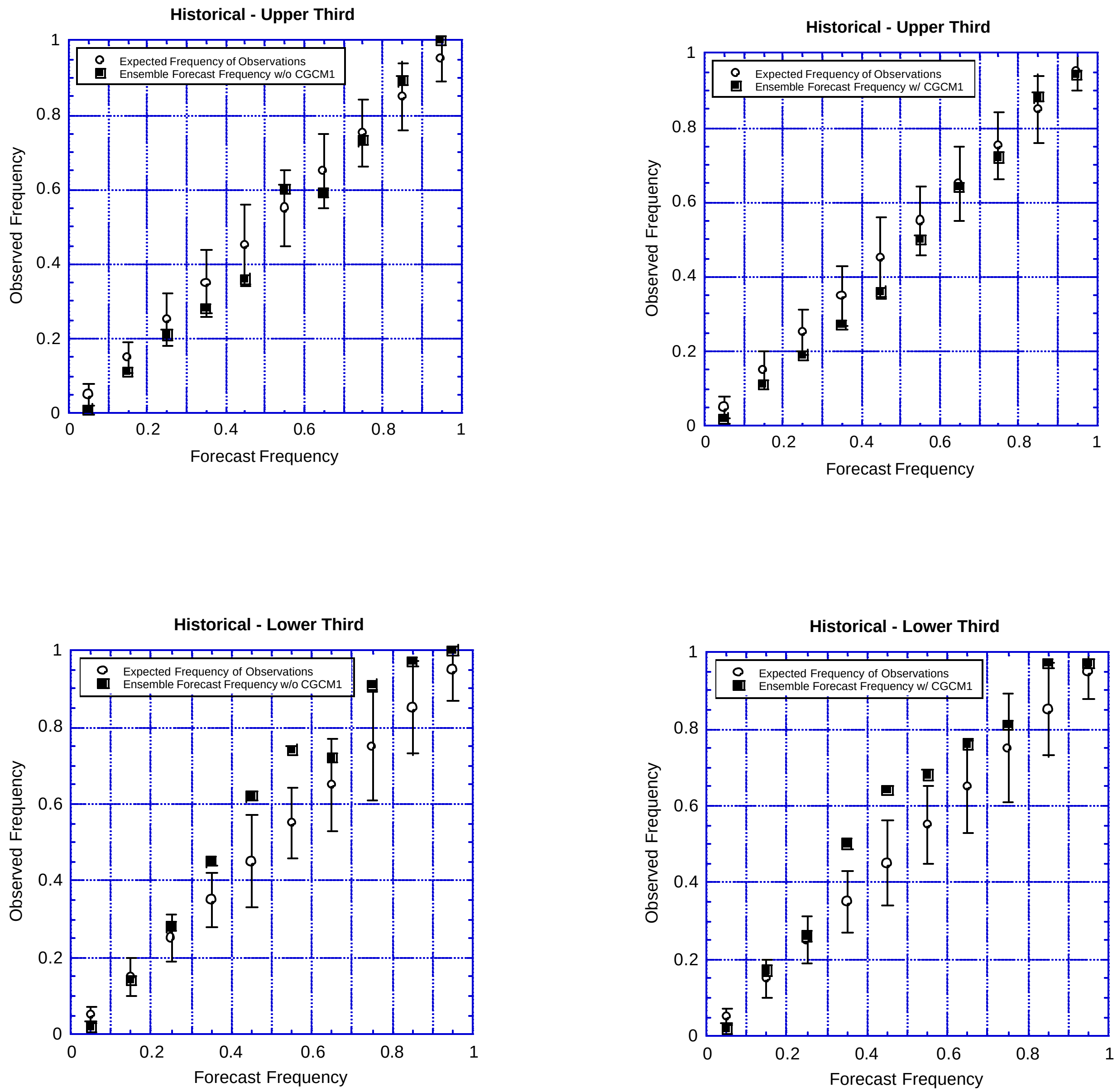


Figure 14. As in Figure 13 but for forecasts issued only during the wet period October-April.

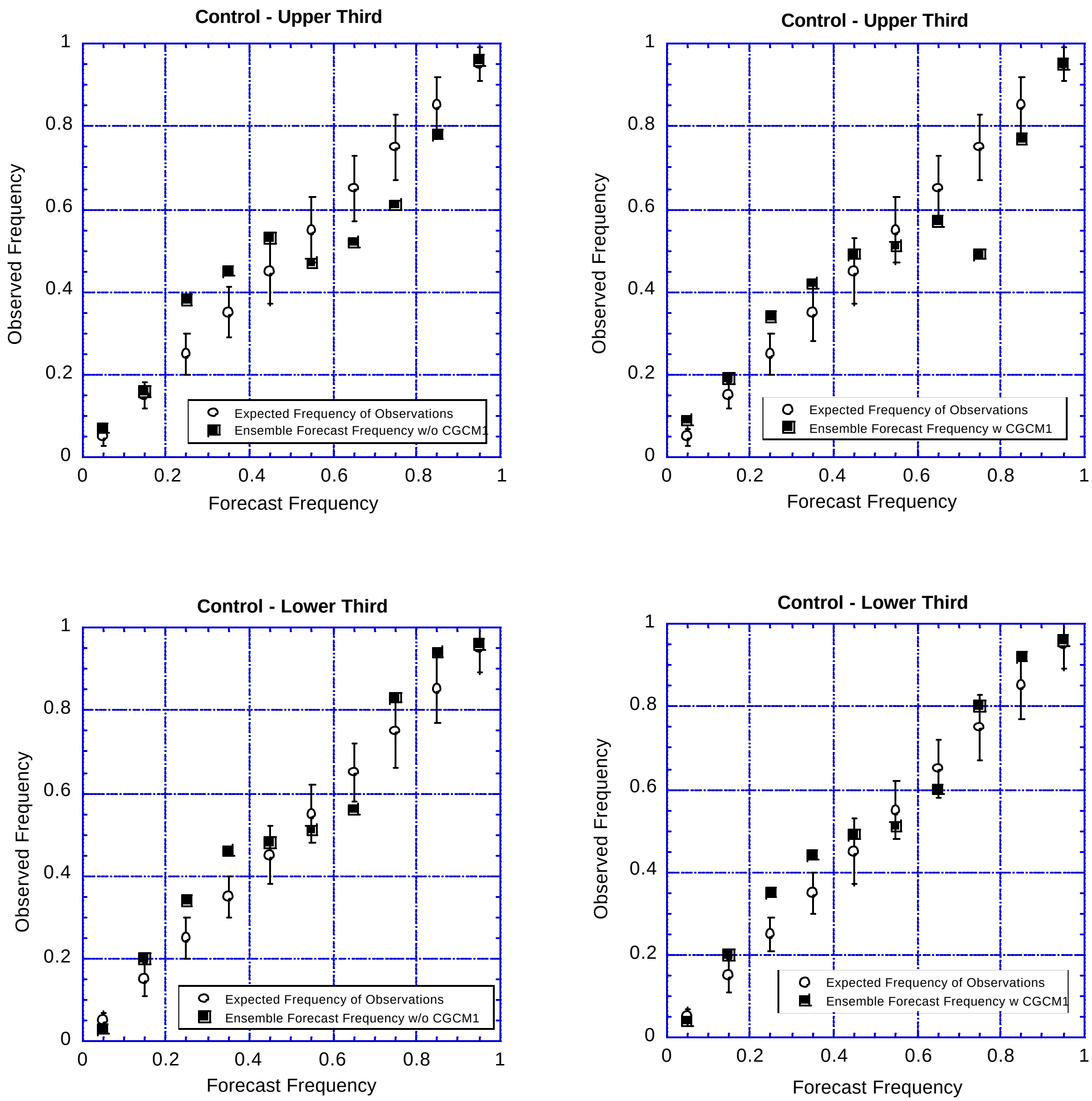


Figure 15. As in Figure 13 but for the future period (10/1/2001-9/30/2030) using a control scenario, and for hydrologic model forcing drawn from the historical period (10/1/1964-9/30/1993).

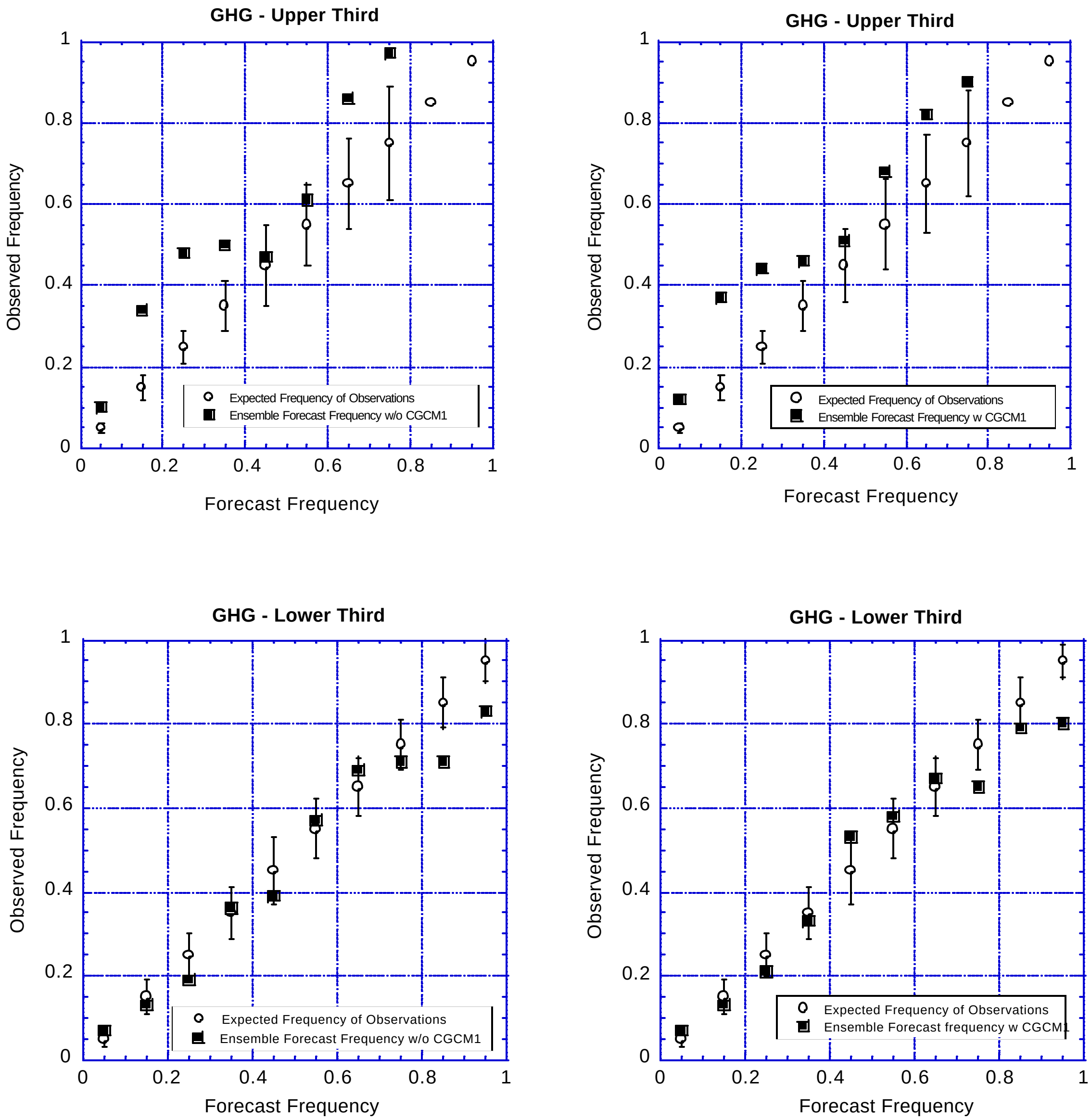


Figure 16. As in Figure 13 but for the future period (10/1/2001-9/30/2030) using a greenhouse-gas-increase scenario, and for hydrologic model forcing drawn from the historical period (10/1/1964-9/30/1993).

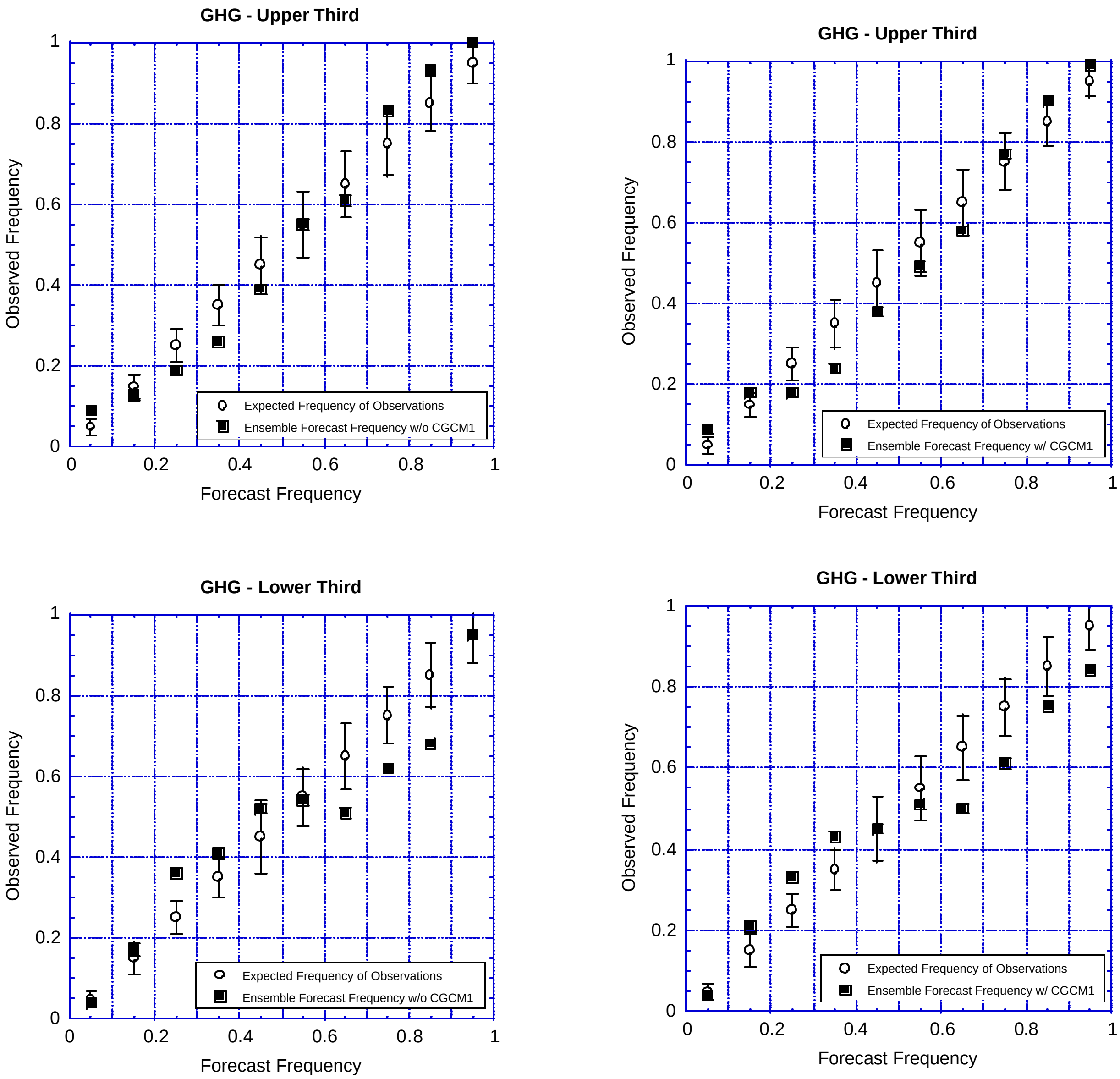


Figure 17. As in Figure 13 but for the future period (10/1/2001-9/30/2030) using a greenhouse-gas-increase scenario, and for hydrologic model forcing drawn from the future period (10/1/2001-9/30/2030).