Operational Hydrologic Simulation Modeling at the Natural Resources Conservation Service's National Water and Climate Center

Thomas Pagano, Tom Perkins, Jennifer Erxleben
Natural Resources Conservation Service, National Water and Climate Center
Portland, Oregon

Abstract

This paper describes the current status and anticipated near-term future directions of the U.S. Department of Agriculture's Natural Resources Conservation Service's (NRCS) National Water and Climate Center with respect to the use of hydrologic simulation models. It begins with a description of the water supply forecasting operations, and continues with a review of past attempts to adopt operational hydrologic simulation models. Next, the modeling environment is described, with emphasis on one model, the Precipitation Runoff Modeling System (PRMS) and its environment, the Modular Modeling System (MMS). Case studies of two basins are provided. For the sake of brevity, many of the NRCS's other simulation modeling activities are not included. Nonetheless, in the final sections, general aspects of simulation modeling in an operational environment are discussed, ranging from model calibration to the role of multiple models. The final section is relevant to any operational modeling enterprise, regardless of the specific model or methodology chosen.

History

For close to 70 years, NRCS has provided seasonal water supply outlooks for use by western U.S. water managers. These outlooks are a critical component in effective water management and are utilized by a broad spectrum of users for a variety of purposes, ranging from irrigated agriculture, flood control, municipal water supply, endangered species protection, power generation, and recreation.

The Water and Climate Services division of the NRCS National Water and Climate Center produces seasonal water supply outlooks monthly, January through June, in partnership with the National Weather Service (NWS) and local cooperating agencies, such as the Salt River Project in central Arizona. During the 2004 forecast season, four NRCS hydrologists issued over 10,000 seasonal water supply outlooks for over 630 locations. Near the start of the month, each forecaster typically has less than 3 working days to create, analyze, adjust, coordinate, and issue forecasts for over 160 points simultaneously. The geographic and climatic scope of the forecasts range from minor creeks of the semi-arid southwestern United States to glaciated basins of the Arctic Circle. Any new forecasting techniques would need to address many of the unique demands of this time-critical, yet human and computer-resource limited operational environment.

Improving these forecasts is one method of improving the sustainability of water supplies in the western United States. Increasing competition over limited resources also demands more informative forecast guidance, directly related to the user's situation. For example, while it may help a user to have an estimate of the anticipated April-July runoff volume at a specific location, his or her legal water right may be tied to the date that flow falls below 225 cubic feet per second. Such user interests are so varied and specific that it is not possible for a forecaster to maintain an armada of statistical regression equations to address (and anticipate) every user need. Instead, the forecaster could present an ensemble of plausible hydrographs from which a specific forecast would be derived by the user. A hydrologic simulation model can provide such a forecast if properly calibrated and provided with the appropriate data.

Also, a simulation model, with its representation of basin physics, can explicitly capture basin behavior during extreme years, e.g., unprecedented snowpack, and multi-year soil moisture deficits. In contrast, the current statistical forecast methodology is relatively limited and does not quantify the effects of highly unusual or even unprecedented conditions.

This paper describes the current status and anticipated near-term future directions of the NRCS National Water and Climate Center (NWCC) with respect to the use of hydrologic simulation models. The emphasis is on one model, the Precipitation Runoff Modeling System (PRMS) and its environment, the Modular Modeling System (MMS). Then general aspects of simulation modeling in an operational environment are discussed, ranging from model calibration to the role of multiple models, which are relevant to any operational modeling enterprise.

NRCS Monitoring History

Along with producing water supply forecasts, the NRCS is also responsible for operating a high-elevation hydroclimatic monitoring network. Until the early 1980s, these measurements were manually collected by snow surveyors traveling to a site on a monthly basis to use a federal snow sampler (a specially calibrated hollow aluminum tube) to measure snow water equivalent and snow depth. Increasing demands for more timely and frequent snowpack information resulted in a significant push to automate and telemeter measurements from nearby snow courses using meteor-burst communications. Thus the SNOTEL (SNOw TELemetry) network was funded and deployment began in the middle 1970s. Some of the original justification for the SNOTEL network was a demand for daily real-time measurements for use in hydrologic simulation models. Therefore the NRCS has long had an interest in adopting a simulation model for operational forecasting, and this interest has been intricately tied to the data monitoring network. Leavesley and Saindon (1985) and Marron (1986) investigated the use of PRMS in an NRCS operational setting, primarily focusing on basins in Nevada. These authors also tried to constrain model parameters so the model simulated snowpack during calibration-matched SNOTEL snow-water equivalent measurements.

Several NRCS hydrologists (Jones 1986 and Perkins 1988) operated the U.S. Army Corps of Engineers Streamflow Synthesis and Reservoir Regulation (SSARR) model on the Yellowstone and Upper Rio Grande, following the NWS's SSARR-based simulation of the Clearwater River in Idaho (Kuehl, 1979). Perkins was a former Army Corps employee and helped write part of the original SSARR computer code. These authors, likewise, compared simulated snowpack to SNOTEL measurements. Cooley (1986) of the USDA-Agricultural Research Service tested the NWS River Forecast System (NWSRFS) model on Lower Willow Creek in Montana, in cooperation with NRCS personnel. Shafer, et al., (1981) also forced the Snowmelt Runoff Model (SRM) with satellite data to produce forecasts, which was followed by more involvement in the satellite version of SRM around 1987.

All these activities built up to an internal NRCS document in 1992 comparing the results of different models and outlining a strategy for moving from forecasting prototypes to an operational system. This document identified the SSARR model as the most attractive option and committed to calibrating 200 basins in 5 years with 3 staff hydrologists. Running on a Unix 33-Mhz 386 mainframe with DOS 286 workstations, the entire enterprise was expected to cost $1.217 million. Soon after this document was released, the NRCS suffered an unexpected and significant realignment of resources; parts of the agency were reorganized, and the simulation modeling enterprise lost much of its momentum. A position at the NWCC was moved out of water supply forecasting and was devoted to simulation modeling after the 1992 report; recognizing that the agency would not have the resources to attempt operational simulation modeling after the agency reorganization, efforts of this hydrologist were turned towards more research-oriented spatially distributed snow simulation models (e.g., Garen and Marks, 1996, 2001). These snow models would eventually be a component in a next-generation spatial hydrology model likewise being developed by the research community.

A program-wide meeting of the snow survey and water supply forecasting organization was convened in 2002 in Las Vegas. At this meeting, a committee was formed to investigate the feasibility of running hydrologic simulation models in the current operational environment. With relatively fewer budget constraints, and with improved automation and data availability, the window of opportunity appeared open to at least explore the available possibilities. In addition, with the unprecedented sequence of wet and dry years at the end of the 20th century, the call arose from users asking NRCS to provide more and better information about extreme events and forecasts of within-season hydrograph behavior. Our committee formulated a plan to investigate the use of a modified version of the SRM model, as well as PRMS, the University of Washington Variable Infiltration Capacity (VIC, Wood, et al., 2001), and NWSRFS models. The implications of maintaining the status quo and/or serving as a conduit for another agency's forecasts were also identified.

Model Selection

A simulation model is a mathematical representation of processes that influence primarily the energy and water balances of a watershed. These models have a broad range of relevant scales, from continent to catchment, and have varying complexity, from highly lumped generalized conceptualizations to models with explicit representations of basin physics. No model is adequate for all circumstances; and the selection of a model (or models) involves balancing accuracy, practicality, data demands, and the ability to calibrate the model to the specific watershed.

As described by Leavesley, et al., (1983), PRMS is a modular-design, deterministic, distributed-parameter modeling system developed to evaluate the impacts of various combinations of precipitation, climate, and land use on streamflow, sediment yields, and general basin hydrology. Basin response to normal and extreme rainfall and snowmelt can be simulated to evaluate changes in water-balance relationships, flow regimes, flood peaks and volumes, soil-water relationships, sediment yields, and ground-water recharge. Parameter-optimization and sensitivity analysis capabilities are provided to fit selected model parameters and evaluate their individual and joint effects on model output. The modular design provides a flexible framework for continued model-system enhancement and hydrologic-modeling research and development.

PRMS resides within the larger MMS framework which allows the user to construct a model from individual modules, such that a model could be designed to match the situation at hand. For example, if basin hydrograph behavior is heavily influenced by groundwater, the standard PRMS subsurface water module could be replaced by a module with a more appropriate level of detail. The MMS infrastructure allows the design of individual models but it also facilitates the use of many different models on an individual basin because the input and output data formats are universal.

Data Collection and Quality Control

Accurate and representative meteorological data are key to the successful operation of hydrologic simulation models. This data plays a role during model calibration as well as real-time operations. The data demands of a forecasting agency are somewhat different than those of a group setting up a model for research purposes. Forecast models must be able to run on demand, capturing recent events less than hours after they occur. Likewise, real-time data are often of the most dubious quality, especially from automated measurement systems which can randomly produce extreme (but unlikely) values or possess gradual drift. Without automated data acquisition technology and automated, forecaster-aided intelligent data quality control, it's unlikely that the human resources of the NRCS would be able to satisfy the data demands of a single basin, much less the hundreds of basins planned.

The primary driving variables for most hydrologic simulation models are daily temperatures and precipitation amounts, although some models also ingest or assimilate snow water equivalent, snow covered area, and other variables such as surface radiation. The NRCS SNOTEL sites primarily measure current snow water equivalent, accumulated precipitation, and temperature. Many sites have recently installed soil moisture, soil temperature, and snow depth sensors. A very limited number of sites measure wind speed and direction, solar radiation, relative humidity, and/or fire fuel moisture.

The NWS also maintains a variety of networks consisting of low elevation sites, some with automated measurements others with manual measurements taken by cooperative observers (COOP). Precipitation and temperature are routinely measured although accurate snowfall and snow depth measurements are less common. Daily SNOTEL measurements generally began in the early 1980s although many of the COOP data sites have existed since the early 1900s, with widespread data available since 1948.

A recent significant advance in the availability of real-time and historical climate data is the advent of the Applied Climate Information System (ACIS, Hubbard, et al., 2004). This distributed and synchronized information network is maintained and operated by the Regional Climate Centers and the National Climate Data Center. It serves data from U.S. National

Oceanic and Atmospheric Administration (NOAA) networks including the COOP network, the Hourly Surface Airways Network, and the Historical Climatology Network. ACIS can be accessed through high-level, Web-based interfaces or directly through a Python language-based XML-RPC standard. The Python interface allows, among other things, for the user to submit a list of sites, desired dates, and variables at a command prompt and be returned to a machine readable file containing the data. Through a series of Cygwin (a UNIX emulator for Windows) shell scripts, precipitation and temperature data through yesterday are currently being retrieved from the ACIS system and the NWCC ftp server, and are being combined with streamflow data automatically downloaded from the U.S. Geological Survey (USGS) Web page to create model-ready files for forecast execution.

In addition to the real-time data, it is important to create and maintain an extremely high quality historical dataset, subjected to the most rigorous screening and data quality testing possible. The NRCS focuses most of its resources in maintaining the quality of its snow water equivalent and precipitation data. Temperature data however are largely "raw" (a recent inventory showed between 99.7% and 99.9% of historical SNOTEL temperature measurements were never altered from the original sensor value). The data possesses many outliers that must be removed and replaced with suitable alternative values for any simulation model to have any chance of accurately reproducing basin hydrologic conditions.

As mentioned in Clark and Slater (2006), one of the authors (Martyn Clark) developed a quality control software package drawing on the best aspects of at least four other major quality control approaches including "point-based and spatial checks for a) extreme values; b) internal consistency among variables (e.g., maximum temperature less than minimum temperature); c) constant temperature (e.g., 5 or more days with the same temperature are suspect); d) excessive diurnal temperature range; e) invalid relations between precipitation, snowfall, and snow depth; and f) unusual step changes or spikes in temperature time series." This procedure was used to identify suspicious values throughout the historical period of record of the SNOTEL and ACIS datasets and replace them with suitable alternatives where appropriate. This software has been transferred to the NRCS for the package to be used to screen real-time data. The NRCS is also investigating the use of the PRISM screening technology (Daly, et al., 2004).

Model Calibration

Hydrologic simulation models contain equations that describe the physical interaction of different components of the water and energy balance. Model parameters relate these abstract physical laws (or scale-dependant approximations of these laws) to the specific basin at hand. Many parameters are observable (e.g., basin area, slope, elevation, vegetation type) although some parameters are unobservable conceptualizations of basin characteristics (e.g., the nonlinearity of hydrologic response to near surface soil moisture saturation). While the ultimate goal of a model based completely on observable parameters may not be realized for several years, another key to simulation modeling success is the accurate calibration of parameters. Of particular concern to NRCS operations is the labor intensiveness of manual calibration (human guided stepwise adjustment of model parameters followed by visual inspection of model hydrograph behavior compared to the observed). Instead, the agency is seeking to measure as many parameters as possible, use automatic calibration techniques to estimate remaining parameters, and use manual calibration only when necessary as a last resort.

The spatial parameters of the PRMS model are derived using the "GIS Weasel" an ArcInfo based map and user interface driven tool to delineate, characterize, and parameterize the hydrologic response units of the model (Viger, et al., 1998). This program ingests elevation, soils, and vegetation data; and queries the user about his or her assumptions in defining a hydrologically homogeneous unit then automatically processes the spatial data to generate initial parameter estimates. A modified version of the Weasel is being tested, which uses a fixed strategy for sub-basin delineation and involves little to no human interaction with the program. Such easy, automated, and fast batch estimation of model parameters is an attractive option to agencies with limited personnel.

At this stage, many non-spatial parameters remain to be calibrated. Classically, these steps of model calibration would involve the manual adjustment of model parameters to improve the visual correspondence of the model and observed hydrographs. The danger in such calibration, especially by novice modelers, is the problem of equi-finality (the notion that many different parameter combinations would provide an equally acceptable fit to the hydrograph). While model output between two parameter sets during calibration may be nearly identical, the internal simulation of model states (e.g., the amount of snow on a watershed, the depth of water contained in soils) may be radically different. Parameter sets that "got the right answer for the wrong reason" are likely to perform poorly outside of the calibration period. Therefore it is critical to verify the intermediate states of the model during calibration.

Hay, et al., (2006) have developed an iterative multi-step automatic calibration scheme which was used to derive several initial parameter sets for operations during 2005. This procedure identifies specific parameters that influence the simulation of model states; exogenous datasets are then used to constrain model internal behavior. For example, model parameters related to solar radiation are identified in advance using sensitivity analysis. All other parameters are held constant and the UA-Shuffled Complex Evolution algorithm (Duan, et al., 1994) is used to identify the parameter combinations that give the best fit between the model's simulated solar radiation monthly climatology and observed solar radiation climatology. When an optimal parameter combination is found, the next step of the calibration related to potential evapotranspiration begins, relating the model monthly climatology to the observed. The 3^{rd} and 4^{th} rounds of calibration involve the verification of the annual water balance and the partial duration time series of peak flows above a specified threshold of low flows. The final parameter set is then returned to the first step and the calibration of solar radiation parameters is repeated. In all, the program cycles through all 4 steps 6-8 times until the program converges on an optimal parameter set that satisfies all objectives. This process remains under development and the addition of other exogenous datasets (e.g., snow covered area, snow water equivalent) is being investigated.

Model Operation

In the summer of 2004, 13 basins were identified as suitable initial candidates for an attempt to calibrate, run, and analyze PRMS models (Table 1). After exceptional fall rainfall, and record breaking streamflows following several years of extreme drought, personnel recognized a unique climatological opportunity to test the models and added 3 basins in southern and northern Utah. In the first season, 16 basins were calibrated with multiple parameter sets – the first set used all available input data, and a second used only a subset of those meteorological sites whose data are available in real-time (e.g., some NWS cooperative observers report a month's worth of data only once a month, an unacceptable timeline for real-time operations). The first dataset is likely to give the best calibration results, but may perform poorly during forecasting if many sites are missing data. The second calibration set should provide a more robust estimate of real-time performance even if the calibration is less than optimal. Third and fourth possible parameter sets are anticipated using a historical dataset with serially complete backfilled meteorological data values.

Table 1. Forecast basins and their characteristics. Latitude and longitude are the location of the streamgage. Natural Resources Conservation Service, National Water and Climate Center.

Station ID	Site Name	Latitude (North)	Longitude (West)	Drainage (Miles^2)
06024450	Big Hole River Bl Big Lake Cr at Wisdom, Montana	45.62	113.46	575
06191500	Yellowstone River at Corwin Springs, Montana	45.11	110.79	2,619
06694650	Antero Reservoir Inflow, Colorado	38.98	105.90	189
08378500	Pecos River near Pecos, New Mexico	35.71	105.68	189
08379500	Pecos River near Anton Chico, New Mexico	35.18	105.11	1,050
09112500	East River at Almont, Colorado	38.66	106.85	289
09239500	Yampa River at Steamboat Springs, Colorado	40.48	106.83	568
09251000	Yampa River near Maybell, Colorado	40.50	108.03	3,410
09299500	Whiterocks River near Whiterocks, Utah	40.59	109.93	109
09361500	Animas River at Durango, Colorado	37.28	107.88	692
09406000	Virgin River at Virgin, Utah	37.20	113.18	956
09408400	Santa Clara River near Pine Valley, Utah	37.38	113.48	18.7
12358500	Middle Fork Flathead River near West Glacier, Mont.	48.50	114.01	1,128
13010065	Snake River Ab Jackson Lake at Flagg Ranch, Wyo.	44.10	110.67	486
13105000	Salmon Falls Creek near San Jacinto, Nevada	41.94	114.69	1,450
13147900	Little Wood River Ab High Five Creek near Carey, ID	43.49	114.06	248

The spatial model calibration was completed by NWCC personnel on regular desktop computers described below. From downloading elevation data to finishing spatial calibration takes approximately 30-45 minutes per site, depending on the size and complexity of the basin. The automatic multi-objective calibration was done using the USGS Denver office's Beowulf computer cluster, taking approximately one day of computing time per basin. By October 2005, a java-based visual user interface to the internal-state calibration software was ready for testing on computers in the NWCC office.

In real-time forecast operations, models are initialized by running the model over the period of record of the input dataset (1948-2005) and saving the model states (e.g., snow covered area, snow water equivalent, and soil moisture) on the last day of the run. Forecasts are then created by forcing the model with the meteorological sequence of each historical year in turn, given the same initial model state. The result is an ensemble of equally likely possible futures, given current basin conditions. This ensemble streamflow prediction (ESP, Day 1985) technique has become a standard practice among most operational hydrologic forecast agencies. Although all historical years are run, a subjective visual analysis of input and output calibration time series was done in advance to specify a start year for acceptable traces; in many basins, the change in calibration performance was obvious when the mix of available stations changed. At the most extreme, some basins have no input data early in the period of record before any COOP or SNOTEL sites existed in the region. Inclusion of these sequences in the analysis would be clearly inappropriate.

As of May 2005, identical data collection and modeling systems were operating successfully on several computers at the NWCC as well as on a computer at the Utah NRCS snow survey data collection office and a personal home computer outside the NRCS network. All 16 basins can be run on demand for multiple parameter sets. On a standard 2.8 GHz desktop computer with 1GB of memory with a transfer rate of 1 Mbps, the data requirements for 16 basins across the western United States can be satisfied in less than 8 minutes. Improvements in database technology are likely to reduce this time, as will be necessary when more basins are adopted. Model initialization and ESP simulations for all basins, with two parameter sets per basin, are completed in 8 minutes. Currently the data collection and model operation routines are running on a scheduler four times a day (to collect late-reporting sites).

An Excel spreadsheet has been temporarily designed to ingest model output files, link to real-time streamflow data, visualize hydrograph behavior and calculate summary statistics. The user can visualize one of 18 model states (e.g., streamflow, snow covered area, temperature, and soil moisture) overlaying the real-time forecast distribution (or a subset of individual years) on top of the model simulated history and/or the observed data where available. The user can also calculate the historical and forecast peak amount, peak date, first date of crossing below or above a relative or absolute threshold, the volume above a threshold, and so on for any of the model states. Additional advanced Java-based spatio-temporal visualization tools were available by September 2005 with the transfer to the Object Modeling System (OMS), the next incarnation of MMS.

Case Study 1: Santa Clara Near Pine Valley, Utah

The streamgage on the Santa Clara River near Pine Valley drains a small, relatively high elevation unregulated watershed dominated by snowmelt. The region recently experienced an unprecedented exceptional drought. In February 2004, the basin was designated as D4, the most extreme drought classification available on the U.S. Drought Monitor (http://www.drought.unl.edu/dm/monitor.html), reserved only for events with return intervals more than 50 years. In 2002 the lowest streamflows on record were recorded; and flow from 1999-2004 was 47 percent of the long-term 1971-2000 normal rate.

Beginning on October 17, 2004, however, Utah was struck with a slow moving high-intensity storm (Bardsley and Julander, 2005). The Virgin and Santa Clara river basins experienced between 4.7 and 10.9 inches (11.938 cm. and 27.686 cm) of precipitation during October 17-23. It is estimated that one station (Gutz Peak) experienced a 24-hour precipitation amount in excess of the 1,000-year return interval. On October 21, SNOTEL soil moisture sensors rose to levels usually only reserved for full snowmelt season and persisted there for the next several months. The PRMS model soil moisture states reflected that the current year had gone well outside the range of historical variability (Figure 1). The complexity of the situation increased when in early January unprecedented snowfalls hit the region. During December 28 to January 13, 2005, sites received as much as 20 inches (50.8 cm) of new snow water equivalent on top of already record high snowpack. The snowpack was so extreme that nearby Midway Valley (NRCS Station ID: 12m23s) broke the all-year, all-time snowpack records by mid February, 2 and ½ months earlier than the previous record set in 1983; and eventually peaked out in mid April at 140 percent of the previous record, which was close to 270 percent of average. PRMS did a fair simulation of the timing and character of streamflows (Figure 2).

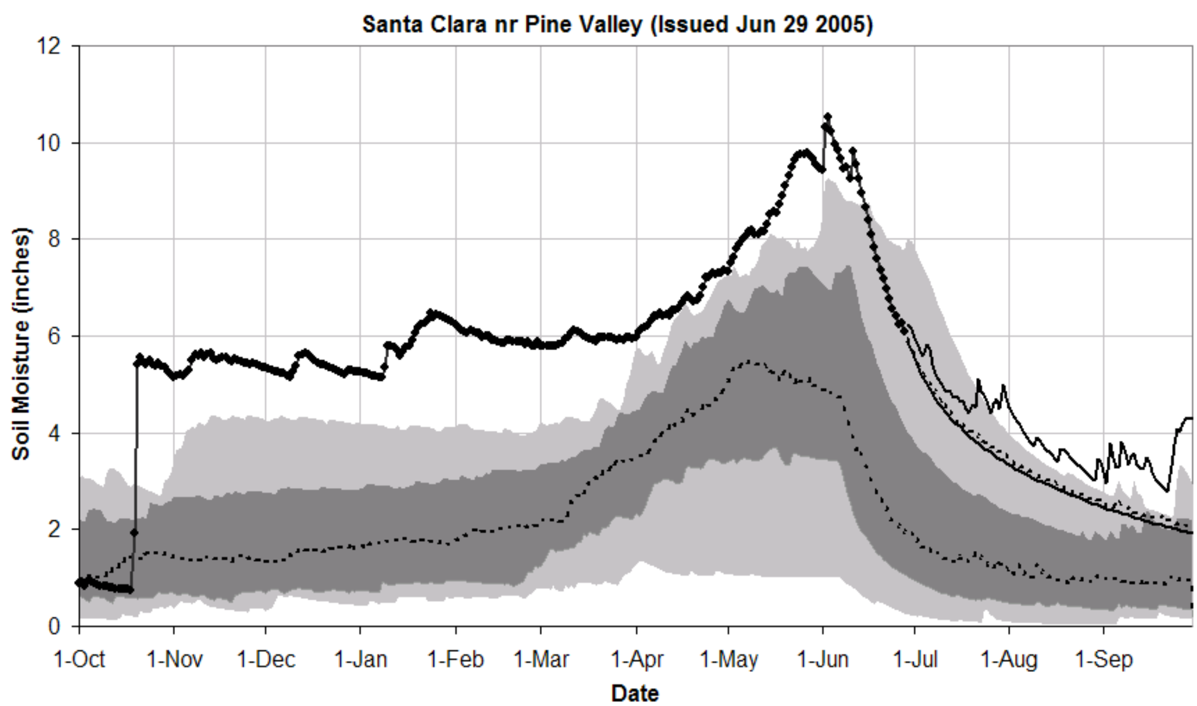

Figure 1. PRMS soil moisture simulation, Santa Clara River. Gray background indicates the model simulated climatology from 1984-2003, including the historical minimum, maximum, median and 10 and 90 percent exceedence probabilities. Heavy black dotted line shows simulated 2005 values and solid and dashed lines are the forecast 10, 50, and 90 percent exceedence probabilities issued June 29, 2005.

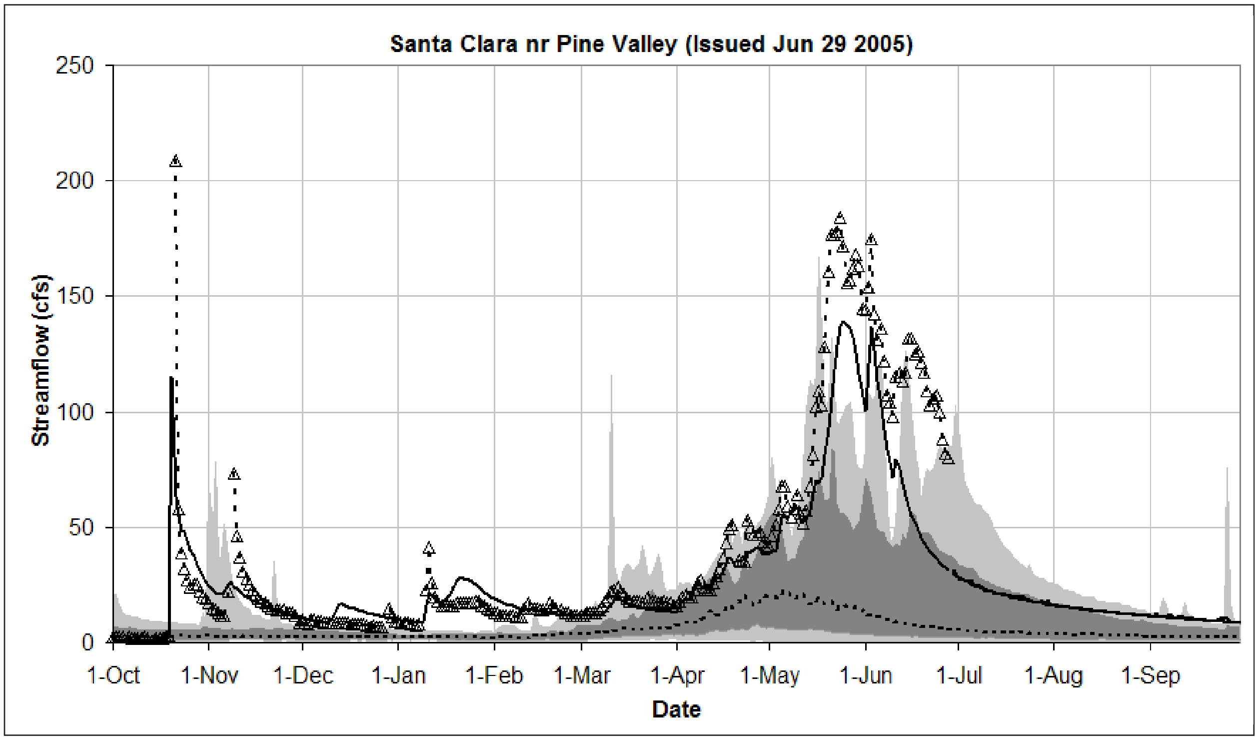

Figure 2. PRMS streamflow simulation, Santa Clara River. Gray background indicates the observed climatology from 1984-2003, including the historical minimum, maximum, median and 10 and 90 percent exceedence probabilities. Solid black line shows simulated 2005 values and the dashed line with triangles represents the observed. Data after June 29 represents the streamflow forecast exceedence levels, as Figure 1; note underdispersion of forecast ensemble spread (i.e. the forecast 10, 50 and 90 percent exceedence probabilities are overlapping).

The model captured the unprecedented winter baseflow conditions and reproduced the overall shape of the hydrograph rise in May and June. It is difficult to know the accuracy of the real-time streamflow data during very high flows although it does seem like the model had a tendency to undersimulate flows, both during the October event and during May-June. Near May 1, the model predicted a 50 percent chance of having a seasonal peak greater than 125 cubic feet per second (cfs) (3.54 cubic meters per second [cms]), whereas the NWS official forecast indicated a 50 percent chance of more than 450 cfs (12.7 cms). The spring precipitation sequence was not unusual, and the eventual peak for the season was 184 cfs (5.21 cms) observed on May 24. The NRCS has no interest in or authority to issue statements related to flooding, and at this time there is no information about whether a slight under forecast would have had a more damaging effect to the user than a larger over forecast.

Case Study 2: Little Wood River Near Caret, Idaho

The stream gage on the Little Wood River above High Five Creek near Carey, Idaho, is at an elevation of 5,320 feet (msl) (1,620 meters) and drains an area of 248 square miles (642 square kilometers). The mean elevation of the basin is 7,220 feet (2200 m). Diversions above the gage are used to irrigate 1,300 acres (526 hectares), which is less than one percent of the basin. The forecast for this location is used to manage the Little Wood Reservoir, which serves downstream irrigators, but is also important for recreation, fish and wildlife, and a small amount of hydroelectric power generation.

Prior to the 2005 water year, the Little Wood watershed was in the midst of a multi-year dry period that started in 2000. Despite a wet summer and fall, the streamflows at the High Five gage were flowing below normal. A series of nine small storms starting on March 19th and ending on May 9th, served to build upon a meager snow pack and to increase the water content of the soil profile (enhancing the expected runoff efficiency). Each of these storm events produced a half inch or less of precipitation. The period of May 15th through May 19th experienced copious amounts of precipitation over southern Idaho generally, and over the Little Wood basin specifically (>4" (10.2 cm) of precipitation). Streamflows in southern Idaho streams and rivers rose dramatically in response to these heavy precipitation events.

The PRMS model simulation matched the timing and magnitude of streamflows (Figure 3). The model tracked the two rapid rises that occurred on May 17th and May 19th, the days of heaviest precipitation. The model produced daily average flows of 1,590 cfs (45 cms) on May 17th and 1,970 cfs (55.8 cms) on May 19th, versus observed values of 1,593 cfs (45.1 cms) and 1,972 cfs (55.8 cms), respectively. It should be pointed out that the model is calibrated on daily average flows. The recorded instantaneous flows were 2,130 cfs (60.3 cms) on May 17th and 2,220 cfs (2220 cms) on May 19th, highlighting how a daily model would clearly not be sufficient if one were attempting to, for example, protect against flood damage.

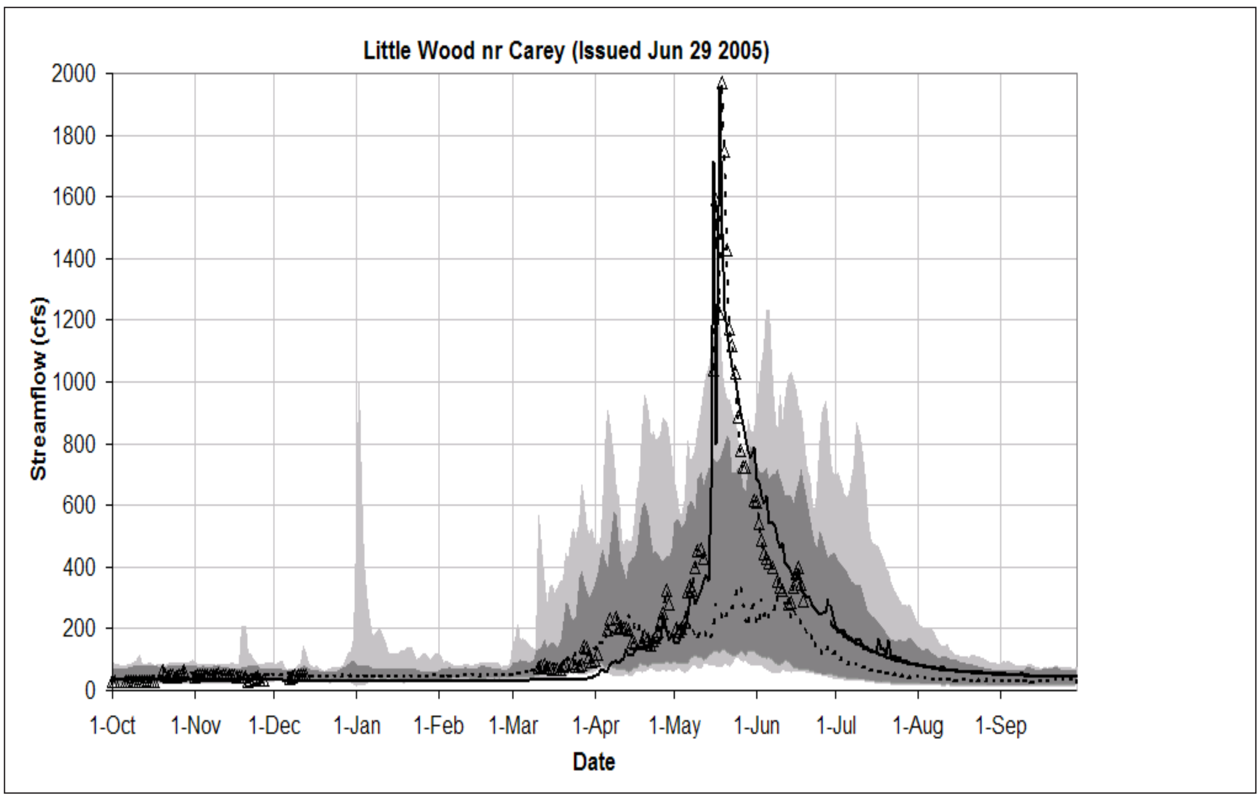

Figure 3. PRMS streamflow simulation, Little Wood River. See figure 2 for symbology. Historical and conditional years include 1987-2003. Similarly note underdispersion of forecast future flows.

While the simulation of this event was excellent, the spring precipitation event was of an extremely large magnitude and was unanticipated. It is likely that the observed streamflow would have been the edges of the conditional distribution of a forecast issued in February or March. Seasonal- and medium-range precipitation and temperature forecasts might have improved the accuracy of the streamflow forecast, but this situation is a clear example of the need to communicate the uncertainty and range of possibilities of outcomes, and to avoid reducing the forecast to "one number" or a single hydrograph trace.

Operational Concerns and Strategies

While the universe of products available from hydrologic simulation models is much more inclusive than statistical-based forecasts, simulation models are much more complex and involve setting up and maintaining. There are many opportunities for errors. Some errors are practically unavoidable, such as those due to extreme precipitation events after the forecast issue date. However, hydrologists should also try to minimize the effects of limitations in model structure, poor model calibration, and errors in input forcings. Additional opportunities exist in the post processing of forecasts and real-time adjustment of model states.

Models

It is essential to have a model or models that are complex enough to describe the hydrology of western U.S. river basins. This includes the accurate simulation of snowpack as well as soil moisture, and the spatial interpolation of forcing variables (e.g., temperature and precipitation) over complex terrain. MMS allows the ability to tailor the model structure to fit the situation at hand. In particular, it is critical for the agency to retain flexibility to adjust to evolving forecast needs and not be fixed into "one model." For example, the NRCS envisions playing a larger role in simulation modeling of water quality for agricultural processes. The MMS infrastructure makes such growth possible while maintaining the same overall architecture of data handling. If the model does not have the correct structure for the basin hydrology, the forecaster will have to hope, at best, that time-consuming subjective real-time adjustment to model states and parameters can compensate for these limitations.

Inevitably, all models are imperfect representations of reality, and each is a different perspective on a system. Operational hydrology often focuses on the use of a single tool or a single model in developing forecast guidance. In many natural science and economic settings, research consistently reveals that a consensus forecast based on the output of many tools almost always outperforms the best individual tool within the ensemble (Armstrong, 2001). The approach of creating forecasts based on an ensemble of tools (e.g., "Super-ensembles") has gained acceptance in the operational meteorological and climatological communities, and the evolution of hydrologic practice along these lines would be logical and would benefit users. Previously, the resources required to maintain many different modeling systems made such an enterprise prohibitively expensive, especially if the incremental improvement in the forecasts was small compared to the cost of maintaining many different systems. Operational meteorologists and climatologists rely heavily on automation and leverage partnerships with outside research groups (e.g., universities) running their own models; there is no reason the same approach could not be used by hydrologists.

Parameters

In relating a generalized model to a specific basin, it is necessary to estimate model parameters. As mentioned earlier, so-called "observable" parameters are preferred to non-observable parameters because human expertise is often required to subjectively estimate the non-observable ones. Much research has been done in objective automatic calibration of hydrologic models although this practice has not received widespread operational acceptance. A concern is that automatic calibration procedures "possess logic" but "lack sense," in that they can adjust parameter values to the extremes, in order to achieve the last bit of improvement in the calibration objective function. In comparison, the human forecaster can use subjective knowledge and experience to constrain certain parameter values while preventing over-fitting.

In one sense, the optimal strategy for a resource-limited agency like the NRCS would involve a hybrid of automatic and manual calibration techniques. The hydrologist would have to articulate multiple objectives that he or she would like to satisfy in terms of hydrograph behavior. Next, an objective procedure mimicking manual calibration (e.g., Hogue, et al., 2000) would return to the hydrologist a series of plausible parameter sets. The expertise of the human would then be used to narrow the range of parameter sets and/or make minor adjustments to parameter values. Of course nothing prevents the hydrologist from retaining all plausible parameter sets. The standard ESP procedure accounts for forecast uncertainty due to future climate variability but ignores uncertainty due to parameter estimation, model limitations, and data uncertainty. Running many parameter sets and aggregating the results into a "super-ensemble" would be one way of accounting for parameter uncertainty, provided that one can ensure that the parameter sets are independent samples.

Although it may be beyond the computational resources of the agency at the moment, some researchers have gone as far as suggesting that the model should be recalibrated any time a specific forecast is desired (Moore and Doherty, 2005). Rather than retaining a single "calibrated model" for all situations (e.g., peak flows, low flows, seasonal volumes), the model is recalibrated on demand to fit the exact forecast situation (e.g., the first date after the peak that flow drops below 325 cfs [9.2 cms]). This philosophy borders on treating the model like a complicated non-linear statistical tool. This approach does not satisfy the expectation that the agency should be able to provide a suite of hydrograph traces that give an equally accurate forecast, no matter what aspect of the hydrograph is being analyzed. The agency may have to decide on the most important objective, either a hydrograph based forecast of plausible daily time series, or a tool that gives the most accurate and well-calibrated probabilistic answer to the specific question being asked. Ultimately, resources may be the limiting factor in determining the calibration strategy, both in the sense that the human resources are not available to do manual calibration and computer resources are not available to do the most sophisticated forms of automatic calibration. In addition, it is not enough for a calibration approach to "have the science right." If it is not easy to install or operate, does not have an adequate user interface, or is not fully credible to the hydrologist, the likelihood of operational adoption is low.

Data

Without real-time access to high quality data, no forecast model can operate successfully. An ideal forecast system would also include a graphical interface to the data, backed by a powerful set of automated quality control routines. A forecaster's energy should be spent inspecting extreme values and "locking in" the values if they are true. This includes routines that investigate the spatial-temporal consistency and plausibility of the values, identifies the probability that the value is an outlier, and suggests replacement values. Of particular concern is the mistaken rejection of a real extreme value. The forecaster must also be able to retain control over the past data in that if a forecaster changes a value, the original value does not overwrite the forecaster edit when refreshing data, unless the data value changes at the source; in which case the forecaster should be presented with a choice to accept or reject the new value. The forecaster should also have the flexibility to decide to use all raw data, all estimated data, or some mix of original and edited values.

This, of course, assumes that the "true" data value is knowable by the forecaster. Data quality control involves the estimation of uncertain information by looking at the internal consistency of the data and the historical relationships between variables or across sites. Even a manual "ground truthing" of an automated measurement has some instrumental and representativeness errors associated with it. In this sense, the parallels between forecasting the true state of anticipated streamflow and estimating the true value of the forcing inputs (e.g., snowpack, precipitation, and temperature) are uncanny.

From the turn of the century until the 1980s, water supply forecasts were almost exclusively deterministic in that the forecast was for a specific volume at a given location and time period (e.g., "500 thousand acre-feet [617,000,000 cubic meters] for April-July on the Animas River at Durango"). In time, the need to provide users with a probabilistic forecast was recognized, so that the users could determine their respective level of risk and determine the "one number" that is most relevant to their specific operations. Although the focus of forecasters and presentation of the water supply forecast information to users still remains primarily deterministic, there is virtually no operational discussion of data values as probabilistic entities. Such a change philosophy would require a wholesale paradigm shift that is unlikely to occur in the near future. Nonetheless, it is still the responsibility of operational agencies, within the limits of their resources, to adequately and accurately represent forecast uncertainty, including data uncertainty.

Techniques exist to convert deterministic point data values into probabilistic data distributions, such as PRISM (Daly, et al., 2004). Clark and Slater (2006) also describe such a conversion and the linking of gridded precipitation conditional distribution functions with random number generators to create ensembles of adequately spatially correlated precipitation fields. These ensembles can then be used to force the hydrologic model(s) to derive an ensemble of possible model initial states (e.g., snowpack). This will create additional dispersion in the ensemble-of-ensembles forecast, accounting for the data uncertainty. As before, the success of the NRCS in accounting for these factors relies on how well the above concepts can be articulated in a user-friendly software package that does not exhaust available human and computational resources.

Assimilation and Post-Processing

The NRCS data collection and quality control effort is primarily focused on the accurate estimation of snow water equivalent data. Many models ingest precipitation and temperature data to simulate snowpack. The agency should take advantage of the information contained in the snowpack measurements to keep model states on track.

The NWS has for many years assimilated snow data, beginning with the Snow Estimation and Updating System (SEUS, McManamon, et al., 1995), continuing on with variants and modifications to the original design. The University of Washington forecast system projects the current observed snow data into a historical observed percentile which is then projected into a percentile of the model climatology (Wood, et al., 2001). Bales, et al., (2006) recently assimilated satellite snow image data into the PRMS model with limited success. Objective data assimilation is routine and widespread in meteorology, and there is also a long history of subjective modifications to hydrologic model states by the NWS (in order to match real-time simulated flows with observed flows – a significant difference in these two diminishes the credibility of a short-term forecast).

Before assimilation, the hydrologist should be inclined to ask two questions – why is the model simulation not tracking the observed and what are the ramifications of changing a model state? If the model's simulations are poor due to poor model structure or poor calibration, the forecast issue date is not the time to be adjusting model states, as the credibility of the result is already beyond repair. In an ideal situation, data assimilation should rarely produce large adjustments (e.g., when a small scale major precipitation event occurs between measurement sites). Frequent large adjustments are an indicator that the modeling system itself is fundamentally flawed and should be redeveloped.

Likewise, when a model state is changed (e.g., snowpack is "removed" from the model), what happens to the rest of the model? Did the snow never fall in the first place (in which case it should be "evaporated")? Did the snow fall and melt unexpectedly (in which case it should go into the soils or have appeared as streamflow)? The implications for the forecast are not trivial; the priming of the soils in the second example is likely to produce a wetter forecast than the first example. Changes in a model state should cascade into other model states. In an obvious example, if a zero snow water equivalent value is assimilated, the snow covered area, also, should be adjusted to zero. Slater and Clark (2006) have tested a methodology for assimilation of snow data, using the historical interrelationships among model internal states to determine how, if at all, other non-snow model states need to change.

If it is necessary for the recent hydrograph to match the near-term forecast, another possibility is to perform a statistical time series analysis of recent errors and apply a bias correction to the near-term forecast to make them align. The MIKE-SHE European hydrologic model (Refsgaard and Storm, 1995) allows the user to automatically deconstruct the recent errors into bias, magnitude, and timing, generating a time varying correction function into the future. This correction may add to the plausibility of the forecasts by giving the user the (potentially false) impression that the model is simulating basin conditions accurately, but it is simply treating the "symptoms" of what could be a more serious underlying "disease."

If a model contains unavoidable systematic biases, it may be useful to apply statistical post processing to ensure the optimal translation of model output into forecast information. This may involve, for example, converting a model forecast volume for each hydrograph trace into a percentile ranking with respect to the model climatology during calibration. This percentile is then converted into a volume with respect to the observed historical streamflow value. Other options include the fitting of model simulated and observed streamflows to statistical distributions to estimate the adjustments necessary to the forecasts. No specific technique for such bias adjustment has emerged as a clear favorite; the NWS Colorado River Basin River Forecast Center system allows the operator to fit no fewer than 5 different distributions to the historical and real-time data and has multiple error models available to post process the data. With bias adjustment, however, one must operate from the assumption that even though the model got the historical simulation “wrong,” the relative change in the model forecast relative to the model climatology was the “right” forecast of real conditions. The relevance of historical model errors to the real-time forecast errors also diminishes if real-time adjustments are made to the model (e.g., snowpack assimilation, manual changes to model states) that weren’t made during the historical model runs.

Finally, standard ESP assumes that all meteorological sequences (traces) are equally likely and therefore receive an equal weighting when deriving products (e.g., the simple average of the output traces as opposed to a weighted average). However, it is possible to indicate that, for example, if an El Niño event is underway, that the anticipated meteorological sequence is more likely to resemble other El Niño years than the opposite, La Niña years. Therefore, El Niño year traces should receive a heavier weight in output products than La Niña year traces. Werner, et al., (2004) has tested 6 climate-based trace weighting schemes for use with the NWS model. Clark, et al., (2004) has also probabilistically integrated long-term meteorological (i.e., 10-15 day) forecasts into ESP. Climate change such as regional warming may make some scenarios less relevant than others. Ideally, the NRCS hydrologic forecast should be consistent with the variety of other climatological and meteorological forecasts that exist across a range of temporal scales. As earlier, the science of trace weighting seems robust and the ability to incorporate this information will ultimately depend on the resources available to the agency.

Summary and Conclusions

Continuing a long-standing interest in the use of hydrologic simulation models for operational water supply forecasting, the NRCS has developed a preliminary prototype based on the PRMS/MMS system. As of 2005, 2 parameter sets each have been created for 16 basins across the western United States, without using any manual calibration. Data is retrieved daily from the SNOTEL and ACIS networks and is automatically processed for model ingestion. The time required to gather data, and run the model for all basins and all parameter sets is approximately 15 minutes on an ordinary desktop computer and can be done on a scheduler with no human intervention. Many scientific and technical challenges remain including, but not limited to, issues related to data quality, database design, data assimilation, uncertainty estimation, pre- and post-processing, and effective delivery and communication of results. The agency aims to build an infrastructure that is robust yet flexible enough to accommodate new developments within the research community.

Author E-mail contact: tom.pagano@por.usda.gov

References

Armstrong, J. S., 2001: Combining forecasts. In: Armstrong, J. S. (Ed.), *Principles of Forecasting.* Kluwer Academic Publishers.

Bales, R., Dressler, K., and Leavesley, G., 2006: Evaluation of gridded snow water equivalent and satellite snowcover products for mountain basins in a hydrologic model. *Hydrologic Processes*. 20(4):673-688 pp.

Bardsley, T., and Julander, R., 2005: The documentation of extreme events: Two case studies in Utah, Water Year 2005. *Proceedings of the Western Snow Conference*. Great Falls, Montana, April 11-14, 2005: 33-42 pp.

Clark, M. P., and Slater, A. G., 2006: Probabilistic quantitative precipitation estimation in complex terrain. *Journal of Hydrometeorology*, 7(1): 3-22 pp.

Clark, M., S. Gangopadhyay, L. Hay, B. Rajagopalan, and R. Wilby, 2004: The Schaake Shuffle: A Method for Reconstructing Space–Time Variability in Forecasted Precipitation and Temperature Fields. *Journal of Hydrometeorology*, 5(1): 243–262 pp.

Cooley, K. R., 1986: Evaluation of the NWSRFS Model on a Montana watershed. *Proceedings of the Western Snow Conference.* Phoenix, Arizona, April 15-17, 1986*:* 112-121 pp.

Daly, C., Gibson, W.P., Doggett, M., Smith, J. and Taylor, G., 2004: A probabilistic-spatial approach to the quality control of climate observations. *Proceedings of the 14th AMS Conf. on Applied Climatology, 84th AMS Annual Meeting Combined Preprints*, Amer. Meteorological Soc., Seattle, Washington, January 13-16, 2004, Paper 7.3.

Day, G.N., 1985: Extended Streamflow Forecsating Using NWSRFS. *Journal of Water Resources Planning and Management*, 111(2): 157-170 pp.

Duan, Q., Sorooshian, S., Gupta, V.K., 1994: Optimal Use of the SCE-UA Global Optimization Method for Calibrating Watershed Models. *Journal of Hydrology*, 158: 265-284 pp.

Garen, D.C., and Marks, D., 1996: Spatially distributed snow modelling in mountainous regions: Boise River application. *HydroGIS `96: Application of Geographic Information Systems in Hydrology and Water Resources Management*, IAHS Publication No. 235: 421-428 pp.

Garen, D.C. and Marks, D, 2001: Spatial fields of meteorological input data including forest canopy corrections for an energy budget snow simulation model. *Soil-Vegetation-Atmosphere Transfer Schemes and Large-Scale Hydrological Models,* IAHS publication, 270: 349-353 pp.

Hay, L.E., G.H. Leavesley, M.P. Clark, S.L. Markstrom, R.J. Viger, and M. Umemoto. 2006: A multi-objective step-wise automated calibration approach applied to hydrologic modeling of a snowmelt-dominated basin. *Journal of the American Water Resources Association*, 42(4) 877-890 pp.

Hogue, T.S., Sorooshian, S., Gupta, H., Holz, A., and Braatz, D., 2000: A Multi-step Automatic Calibration Scheme for River Forecasting Models, *Journal of Hydrometeorology*, 1: 524-542 pp.

Hubbard K.G., DeGaetano A.T., Robbins K.D., 2004: A modern applied climate information system. *Bulletin of the American Meteorological Society*. 86 (6) 811-812 pp.

Jones, K.J., 1986: *SSARR Modeling Application: Clarks fork of the Yellowstone River SCS-Wyoming State Engineer Cooperative Study*. SCS Final Report.

Kuehl, D. W., 1979: Volume forecasts using the SSARR model in a zone mode. *Proceedings of the Western Snow Conference,* Sparks, Nevada, April 18-20, 1979: 38-47 pp.

Leavesley, H. H., and Saindon, L. G., 1985: *Administrative Report: Evaluation of SNOTEL data for use in the precipitation-runoff modeling system.* FY 1985, Final Report.

Leavesley, G.H., Lichty, R.W., Troutman, B.M., and Saindon, L.G., 1983: *Precipitation-Runoff Modeling System: User's Manual.* U.S. Geological Survey Water-Resources Investigations Report, 83-4238, 207 p.

Marron, J. K., 1986: Parameter estimation for the Precipitation Runoff Modeling System using snow telemetry system data. *Proceedings of the Western Snow Conference,* Phoenix, Arizona, April 15-17, 1986: 154-157 pp.

McManamon, A., Hartman, R.K., and Hills, R., 1995: Implementation of the Snow Estimation and Updating System (SEUS) in the Clearwater River Basin, Idaho. *Proceedings of the Western Snow Conference,* Sparks, Nevada, April 17-20, 1995: 56-65 pp.

Moore, C. and Doherty, J., 2005: Role of the calibration process in reducing model predictive error. *Water Resources Research.* 41(5).

Perkins, T. R., 1988: Seasonal streamflow forecasting in the Upper Rio Grande Basin: Incorporating the use of SNOTEL data in the SSARR hydrologic model, *Proceedings of the Western Snow Conference,* Kalispell, Montana, April 19-21, 1988: 58-69 pp.

Refsgaard, J.C. and Storm, B., 1995: MIKE SHE, in *Computer Models of Watershed Hydrology*, Vijay P. Singh, ed., Water Resources Publications, Highlands Ranch, Colorado.

Shafer, B. A., Jones, E. B. and Frick, D. M., 1981: Snowmelt runoff simulations using the Martinec-Rango model on the South Fork Rio Grande and Conejos River in Colorado, AgRISTARS Report CP-G1-04072, Goddard Space Flight Center, Greenbelt, Maryland.

Slater, A. G., and Clark, M. P., 2006: Snow data assimilation via an ensemble Kalman filter. *Journal of Hydrometeorology*. 7(3) 478-493 pp.

Viger, R.J., Markstrom, S.L., and Leavesley, G.H., 1998: The GIS Weasel - An Interface for the Treatment of Spatial Information Used in Watershed Modeling and Water Resource Management. *Proceedings of the First Federal Interagency Hydrologic Modeling Conference*, April 19-23, 1998, Las Vegas, Nevada, II(7): 73-80 pp.

Werner, K., Brandon, D., Clark, M., and Gangopadhyay, S., 2004: Climate index weighting schemes for NWS ESP-Based seasonal volume forecasts. *Journal of Hydrometeorology*, 5(6): 1076- 1090 pp.

Wood, A.W., Kumar, A. and Lettenmaier, D., 2001: Experimental real-time seasonal hydrologic forecasting for the Columbia River basin. Presented at the 26th Annual Climate Diagnostics and Prediction Workshop, Scripps Oceanographic Institute, San Diego, California.