Stochastic Streamflow Generation Incorporating Paleo-Reconstruction

James Prairie[1], Balaji Rajagopalan[2], and Terrance Fulp[3]

[1]Bureau of Reclamation, University of Colorado, 421 UCB, Boulder, CO 80309; PH (303) 492-8572; FAX (303) 492-1347; email: jprairie@uc.usbr.gov
[2]Department of Civil, Environmental and Architectural Engineering, University of Colorado, 421 UCB, Boulder, CO 80309; PH (303) 492-5968; FAX (303) 492-3717; email: balajir@spot.colorado.edu
[3] Boulder Canyon Area Office Manager, Bureau of Reclamation, Lower Colorado Region, PO Box 61470, Boulder City, NV 89006-1470; PH (702) 292-8418; FAX (702) 293-8156; email: tfulp@lc.usbr.gov

Abstract

The Colorado River Basin experienced the worst drought on record from 2000-2004. Though this drought was unprecedented in the observed streamflow record (1906-present) reconstructed streamflows dating back to 1490 generated from tree-ring chronologies have shown droughts of greater magnitude and duration. Yet, the decision to adopt the information from tree-rings is not without question. Alternate techniques to reconstruct streamflows based on tree-rings can display different magnitudes of past flow. Though the magnitudes are different these various reconstructions do show similar system state, i.e., wet or dry state.

We present a technique to combine reconstructed streamflow state information with the observed records flow magnitude. This a achieved using a nonhomogeneous Markov chain model with kernel smoothing coupled with a K-nearest neighbor sampling algorithm. The technique is demonstrated for the Lees Ferry stream gauge on the Colorado River. The coupled models retain the ability to generate basic statistics similar to the observed record while also capturing the state properties of the reconstructed streamflows.

Introduction

Water resource planners must consider streamflow variability to provide effective long term planning and management. Incorporating this variability has traditionally been achieved through generation of stochastic streamflows. Stochastic streamflows represent plausible alternate streamflows comparable to observed data available in a river basin. These observed data are typically limited in time, limiting variability in the stochastic streamflows, particularly with the frequency of extremes.

This limitation was highlighted in the Colorado River Basin when an unprecedented drought occurred from 2000-2004. Though this drought was unprecedented in the observed record, streamflow reconstructions generated from tree-ring chronologies have shown droughts of greater magnitude and duration. The most recent streamflow reconstruction for the Colorado River was completed by Woodhouse et al. (2006). The streamflow reconstruction (Figure 1) dates from water year 1490 until 1997 while the observed flows only date from 1906-2003. To put the recent (2000-2004) five-year drought in context the 5-year running means are presented. It is evident the recent drought has not been seen in the observed period while the reconstructed streamflows have seen similar or worse droughts of 5 year duration four times over the approximate 500 year period indicating these droughts are not unprecedented. Yet the decision to adopt the streamflow values from tree-ring reconstructions is not without question.

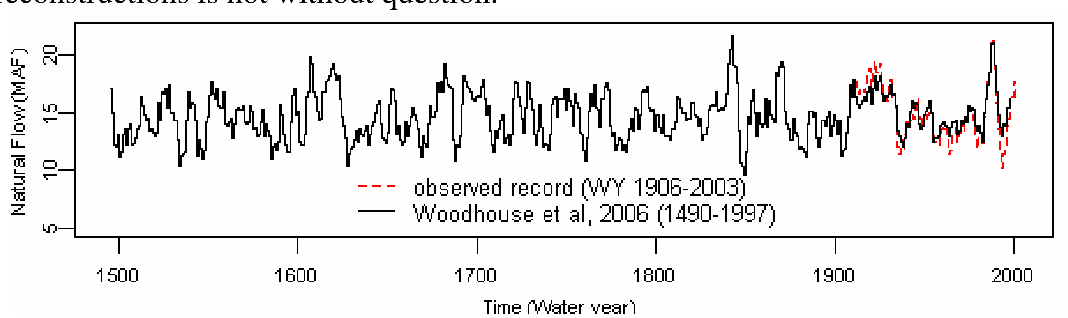

Figure 1. Five-year running means for historic and reconstructed streamflow.

River basin planners have long understood the need to model a variety of possible streamflow scenarios. In the Colorado River Basin current stochastic generation of streamflows models the possibility of drought or surplus under the assumption that the worst that could be experienced has occurred in the observed record, which extends from 1906 through present. Paleo-reconstructed streamflows have been available for many decades but their regular use in planning and decision making has been hampered by uncertainty in the magnitude of the flows generated from the reconstructions. Hidalgo et al. (2000) demonstrated that alternate procedures in tree-ring based reconstructions significantly impact the magnitudes exhibited in streamflow reconstructions. This seed of doubt has made incorporating reconstructed streamflow data into a planning model contentious and open to criticism. Though, few argue streamflow reconstructions show duration and frequency of drought that differs from the short observed record. Incorporating the duration and frequency of drought and surplus exhibited by streamflow reconstructions is essential to generating stochastic hydrology which is not limited to observed flows duration and frequency for drought and surplus.

We propose a framework that combines the magnitudes of flow seen in the observed record with the duration and frequencies observed from the reconstructed streamflows. Excluding the magnitude information of the reconstructed streamflows but retaining the duration and frequency of the system states allows the incorporation of important information furnished through the reconstructions, which is remarkably consistent irrelevant of the techniques and procedures used to generate the reconstruction. Generating stochastic hydrology with these two characteristics from the observed and reconstructed data eliminates barriers that have prevented the use of important information provided though reconstructed hydrologies.

The proposed framework is introduced with a description of the data sets (observed and reconstructed streamflow data) incorporated within. Next a brief description of the framework introduces the two models incorporated in the framework. The section concludes with an algorithm describing the use of the framework in generating simulations. Results are presented after application of the framework in generation of streamflows for the Colorado River at Lee's Ferry. Streamflows are generated using reconstructed flows (1490-1997) and observed flows (1906-1997). The stochastic streamflows are analyzed through a suite of statistics to illustrate the advantages of incorporating the reconstructed flow data. The summary and conclusion, including steps required to extend this research for basin-wide planning studies, close the paper.

Data Sets

Two data sets were combined in development of the proposed framework. First are the natural flow data developed by the Bureau of Reclamation. Natural flows are available annually from water year 1906-2003. Naturalized streamflows are computed by removing anthropogenic impacts (i.e., reservoir regulation, consumptive water use, etc.) from the recorded historic flows[1]. Prairie and Callejo (2005) present a detailed description of methods and data used for the computation of natural flows in the Colorado River Basin.

The second data set consists of streamflows reconstructed from tree-ring information. Tree-ring widths are influence by available soil moisture and are shown to correlate well with annual runoff. Tree-ring data are collected from a series of trees at multiple locations where tree growth is moisture-limited. Two core samples are taken from each tree. The core samples are first cross dated and the ring widths are measured. A standard series of techniques (Stokes and Smiley, 1968; Swetnam et al., 1985) are employed to process the ring width series. Typically, the series is first detrended to remove the effects of reduced ring width with aging. Next the series of

---

[1] The natural flow data and additional reports describing these data are available at http://www.usbr.gov/lc/region/g4000/NaturalFlow/index.html.

ring widths from various cores at a single location are combined with a weighted mean to develop a site chronology (Cook et al., 1990).

The site chronology is related to observed streamflow with a reconstruction model. The reconstruction models are developed by adding predictors in a forward stepwise method (Weisberg, 1985) to a multiple linear regression that correlates the chronologies with observed streamflow. For the Colorado River at Lees Ferry gauge the chosen streamflow reconstruction from Woodhouse et al. (2006) describes approximately 84% of the annual variance.

As with any proxy method to determine streamflow, tree-ring information is not without fault. Correlation of tree-ring width and higher flows has shown ring growth tends to be influenced by variables other than moisture at these higher flows degrading the ability for tree-rings proxy data in representing high flows (Woodhouse and Brown, 2001). Further, different datasets and techniques to process tree-ring information result in difference in reconstructed flows at a single given location (Hidalgo et al., 2000).

For instance, multiple data sets of tree-ring reconstructed streamflows are available for the Colorado River Basin at Lees Ferry (Woodhouse et al. 2006). Each reconstruction was developed with a different set of chronologies and processed with different operational techniques in development of a final reconstructed streamflow. Hidalgo et al. (2000) demonstrated that the choice of techniques can influence the magnitudes of flow in the resulting reconstruction. The variability in streamflow reconstruction resulting from alternate datasets and techniques at a single location are shown in Figure 2. Four different reconstructions at Colorado River at Lees Ferry are shown. The reconstructions include the original reconstruction performed by Stockton and Jacoby (1976), one completed by Hidalgo et al. (2000), one completed by Hirschboeck and Meko (2005) for the Salt River Project, and the recently completed reconstruction by Woodhouse et al. (2006). Each reconstruction used a different set of chronologies and different methods to develop and process their chronologies. Of particular interest is the increased severity of drought and reduced overall mean displayed by the Hidalgo reconstruction. Unfortunately, these alternate streamflow reconstructions have not helped instill confidence in using these data in decision support systems because policy makers cannot decide which reconstruction best represents the unobserved past. To date policy makers have been more comfortable using the limited observed streamflows when setting policy in the Colorado River Basin.

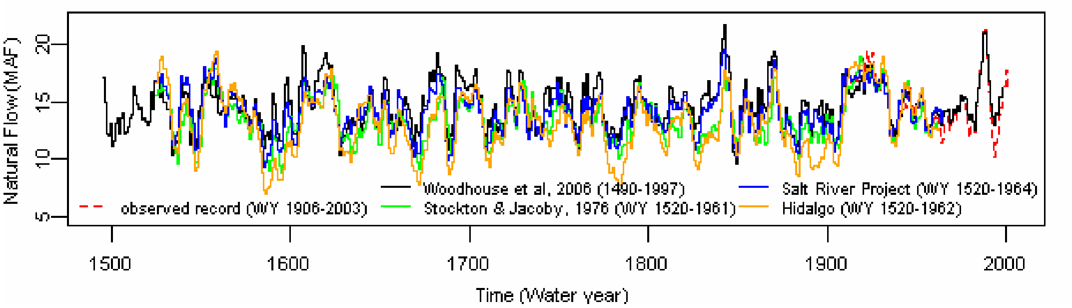

Figure 2. Five-year running means for recent and previous streamflow reconstructions at Lees Ferry.

Though basin managers may not yet be ready to incorporate reconstructed streamflows in DSS, a step towards incorporating limited information from reconstructions is proposed. An alternate view of reconstructions is to use the reconstructions to display the state of the streamflow. For example the streamflow series can be divided into high and low flow states with the division occurring at the median flow. The system state information for the previous four reconstructions is shown in Figure 3. It is evident the magnitude of the various reconstructions varies, for example the late 1500's show low flows in all the reconstructions though the severity of these low flows differs across the reconstructions. These differences support the skepticism that basin planners hold toward incorporating reconstructions as they are uncertain which reconstruction to implement. Though the magnitudes differ the state of the system (i.e., dry or wet) does show similarity across the various reconstructions. In the late 1500's all the reconstructions display a dry state followed by a wet state. Though the states of the system are fairly consistent across the reconstructions there are some differences. These arise from the different chronologies and techniques used to generate the reconstructions. Of concern with the various reconstructions is that they display a reduced long-term mean compared with the observed record and the choice of reconstruction will influence the DSS. To many a reduced long-term flow is appropriate, but it is difficult for basin planner to implement these lower mean hydrologies; indicated by the lack of use of reconstructions by Federal agencies. Much of this stems from the allocation of water being a highly politicized process which is influence by these long-term DSS.

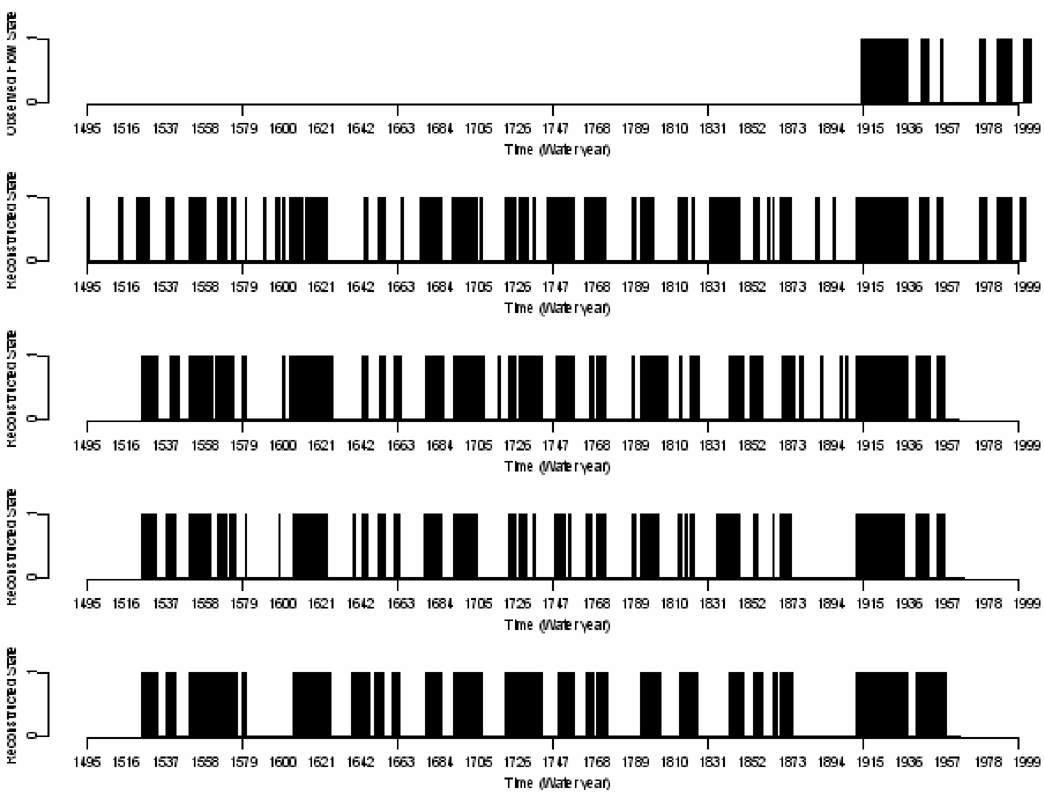

Figure 3. Five-year running means streamflow state for historic and reconstructed.

We propose a method to combine strengths of the observed and reconstructed data set. Basin planners have grown to the trust the magnitudes of the observed flow because these magnitudes result from a simple process that has gained wide acceptance (natural flow computation). But the reconstructed streamflows contain state information that can improve modeling the frequency and duration of drought as a result of incorporating a data set that is five times longer.

We used the latest reconstructions completed in Woodhouse et al. (2006) and chose the Lee's Ferry reconstruction with the highest calibration period accuracy (84 percent). This calibration used the full-pool of available standard chronologies (30 predictors) for the Upper Colorado Basin. Further details of the data used and the reconstruction procedures are detailed in Woodhouse et al. (2006). These data include annual water year flows from 1490-1997.

Framework Description

The two data sets described in the previous section are combined by coupling a nonhomogeneous Markov chain with kernel smoothing, which models system state, with a K-nearest neighbor algorithm to resample observed flows conditioned on the system state. The framework description is shown Figure 4.

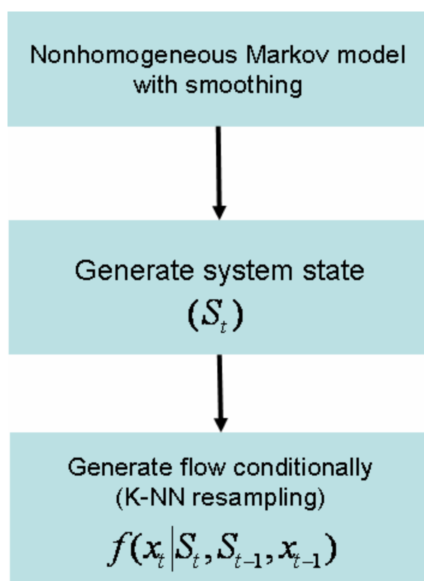

Figure 4. Modeling framework description.

Nonhomogeneous Markov chain with kernal smoothing (NHMCKS)

Nonhomogeneous Markov chains were used by Rajagopalan et al. (1996) in development of daily precipitation. The use of nonhomogeneous Markov chains allows modeling of nonstationary transition probabilities. The transition probabilities are directly estimated from the data. In a two state Markov chain the probabilities found through a counting process estimate the probability of a dry year flowed by a wet year, P_{dw}, and the probability of a wet year followed by a dry year, P_{wd}. The probability of a dry year followed by a dry year can be found as $P_{dd}(t) = 1 - P_{dw}(t)$ likewise, the probability of a wet year followed by a wet year can be found as $P_{ww}(t) = 1 - P_{wd}(t)$.

The state occurrence process is presented in Figure 5. From this process 4 types of data can be extracted. These include (1) the year indices $t_{d_1}, t_{d_2}, ..., t_{d_{nd}}$ for nd years; (2) the year indices $t_{w_1}, t_{w_2}, ..., t_{w_{nw}}$ for nw years; (3) the year indices $t_{dw_1}, t_{dw_2}, ..., t_{dw_{mdw}}$ for ndw years on which the transition from dry to wet occurs; and (4) the year indices $t_{wd_1}, t_{wd_2}, ..., t_{wd_{mwd}}$ for nwd years on which the transition from wet to dry occurs. The year index represents years available in the reconstructed streamflows. The transition probabilities for a given year are estimated from the data with a discrete nonparametric kernel estimator.

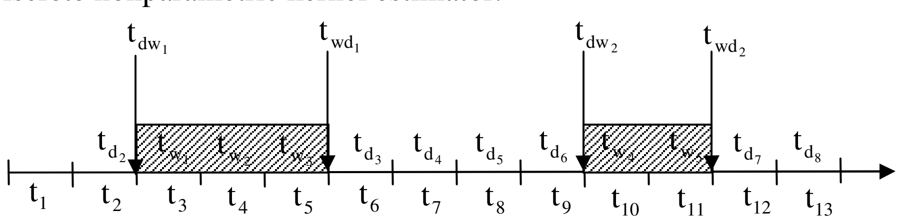

Figure 5. State occurrence process.

Estimating transition probabilities with a discrete nonparametric kernel function differs from traditional Markov chain methods. Traditional methods estimate the transition probabilities from the available record as the ratio of transitions from dry to wet (wet to dry) over the number of dry (wet) years in the available record. The kernel estimator localizes the estimation around each year. The assumption is years close to the current year influence that year more the years further away. The kernel weights years closer more while years further way are given less weight. The kernel estimators for transition probabilities $P_{dw}(t)$ and $P_{wd}(t)$ are given as:

.
$$P_{dw}(t) = \frac{\sum_{i=1}^{ndw} K\left(\frac{t - t_{dw_i}}{h_{dw}}\right)}{\sum_{i=1}^{nd} K\left(\frac{t - t_{d_i}}{h_{dw}}\right)} \tag{1}$$

$$P_{wd}(t) = \frac{\sum_{i=1}^{nwd} K\left(\frac{t - t_{wd_i}}{h_{wd}}\right)}{\sum_{i=1}^{nw} K\left(\frac{t - t_{w_i}}{h_{wd}}\right)} \tag{2}$$

where ndw = the number of transitions from dry to wet in the reconstructed record; nwd = the number of transitions from wet to dry in the reconstructed record; nd = the number of dry days in the reconstructed record; nw = the number of wet days in the reconstructed record; $K(\,)$ = the kernel function; $h_{(\,)}$ = the kernel bandwidth; t = water year of interest; and the $t_{(\,)}$ s are as described earlier for Figure 5. The local neighbors included in the kernel function range from $t - h_{(\,)}$ to $t + h_{(\,)}$. The weight for each neighbor around t is determined by a discrete Kernel function developed by Rajagopalan and Lall (1995) given as:

$$K(x) = \frac{3h}{(1 - 4h^2)}(1 - x^2) \quad \text{for} \quad |x| \le 1 \tag{3}$$

where $x = t - t_{(\,)} / h_{(\,)}$ measures the distance for event $t_{(\,)}$ from the year of interest t within the bandwidth $h_{(\,)}$, where $h_{(\,)}$ is an integer. The Kernel function presented in (3) weights the values within $h_{(\,)}$ such that the sum of the weights equal 1, are always positive, and are symmetric are the point of interest t.

The transition probability estimators (1) and (2) are fully defined once the bandwidth $h_{(\,)}$ is determined for each. An objective method based on a least square cross validation (LSCV) procedure (Scott, 1992) is used to choose the bandwidth. A bandwidth is chosen such that it minimizes the LSCV function given as:

$$\mathrm{LSCV}(h) = \frac{1}{n}\sum_{i=1}^{n}[1 - \hat{P}_{-t_i}(t_i)]^2 \tag{4}$$

where n = the number of observations (ndw or nwd); $\hat{P}_{-t_i}(t_i)$ = the estimate of the transition probability ($\hat{P}_{dw}$ or $\hat{P}_{wd}$) at year t dropping the information on year t. The 1 in (4) results from an assumption that the prior probability of transition is 1 for the years on which a transition has occurred. Values from 1 to 25 years are applied for h when searching for the function (4) minimum.

Once h_{dw} and h_{wd} are objectively determined this formulation is fully defined and applied to generate a transition probability matrix (TPM) for each year in the record. These TPM's are used in a Markov model to generate a time series of system states.

The order of the Markov model is determined through AIC criteria as presented by Gates and Tong (1976) for Markov models. Two state and three state systems were both explored. A two state system better represented the characteristics of the reconstructed streamflow and was therefore used. This results from the fact that the limited observed streamflows provide increased variability during K-NN resampling (described next) for flow magnitudes when distributed across 9 categories, which is required for a 3 state system, rather than only the 4 categories required for a 2 state system.

Flow magnitudes

Once a time series of system states are generated from the Markov chain model a K-NN (Lall and Sharma, 1996) scheme is used to resample an observed flow = x_t conditioned on the current system state = S_t, previous system state = S_{t-1}, and previous flow = x_t.Where t represents the year of interest. This conditional probability is defined as:

$$f(x_t | S_t, S_{t-1}, x_{t-1}) \tag{5}$$

This is accomplished by first splitting the observed data into S states. In a two state system S = 2 and represents a wet or dry state. The choice of S can be determined in various ways. The observed data can be broken up based on equally spaced categories by using the data median as the breakpoint. To determine if the observed data displays natural breakpoints the empirical cumulative density function (CDF) can be plotted. A CDF of the observed data at the Colorado River at Lees Ferry did not indicate any obvious breakpoints in the observed data therefore the data was divided into equal lengths using the median as the breakpoint. Based on the TPM for a 2 state system the observed data are divided into four categories representing the four transition states, dry-dry, dry-wet, wet-dry, and wet-wet.

The K-NN are sampled from the appropriate category; where once the category is identified all the values in the category are weighted based on their Euclidian distance from the previous years flow. The flow values closest to the pervious flow are given a higher weight and those farthest are given a lower weight. A weighted flow value is randomly resampled and the flow for the next year in the observed record from the resampled flow is assigned as the current flow, thereby incorporating a portion of the lag 1 correlation present in the observed data.

We next present the full algorithm combining these two models.

Algorithm

Combining the NHMCKS model with the K-NN algorithm for simulation of annual flow proceeds as follows.

1. Determine the planning horizon for the simulations. For example a simulation may simulate future flows from 2008-2038 indicating a 30 yr planning horizon.

2. Randomly resample blocks of data equal to the planning horizon from the reconstructed streamflows.
3. Generate flow states $S(t)$ where $t = 1,2,...,30$ for each resampled block using the TPM that were generated from the NHMCKS.
4. Generate flow magnitudes $x(t)$ for each t by resampling an observed flow using the conditional K-NN method. Where the conditional probability is defined as $f(x_t | S_t, S_{t-1}, x_{t-1})$
5. Repeat steps 2 through 4 to obtain as many simulations are required.

Performance evaluation

The performance of the proposed framework is evaluated at the Colorado River at Lees Ferry, Arizona with the data sets described earlier, i.e., the reconstructed flows from Woodhouse et al. (2006) extending annually from 1490-1997 and the natural flows provided by Reclamation extending annually from 1906-2003.

A suite of basic statistics were viewed including annual (i) mean, (ii) standard deviation, (iii) coefficient of skew, (iv) maximum, (v) minimum, and (vi) lag-1 autocorrelation. Drought and surplus statistics include the longest surplus (LS), longest drought (LD), maximum surplus (MS) volume, maximum deficit (MD) volume, average length surplus (avgLS), average length drought (avgLD), average surplus (avgS), and deficit (avgD) volume. Droughts are defined as values below the median of the observed record, while surpluses are values above the median of the observed record. Figure 6 graphically depicts the drought and surplus statistics definition.

The results are displayed as boxplots where the box represents the interquartile range and whiskers extend to the 5th and 95th percentile of the simulations (note this is different from the standard boxplot definition). Outliers are shown as points beyond the whiskers and include values beyond the 5th and 95th percentiles. The statistics of the observed data are represented as a triangle. Performance on a given statistic is judged as good when the observed statistic falls within the interquartile range of the boxplots, while increased variability is indicated by a wider boxplot.

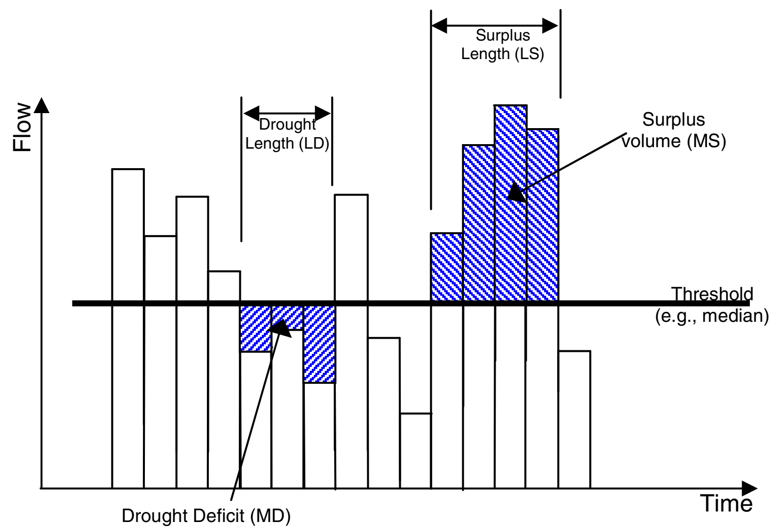

Figure 6. Definition of surplus and drought statistics.

Results

We applied the framework presented and generated 500 simulations. The basic statistics in Figure 7 are preserved well as all are captured within the interquartile range. The maximum and minimum are bounded by the observed record as expected since the resampling scheme only resamples the observed data.

To demonstrate the influence of incorporating the reconstructed streamflows with the Markov chain model we additionally generated 500 simulations using only the observed natural flows and a traditional K-NN model (Lall and Sharma, 1996). The generation of hydrologic state information to condition the generation of streamflows was excluded. Figure 8 presents the drought statistics when resampling with only a traditional K-NN. The observed record (1906-2003) statistic is represented as a triangle and the subset of reconstructed record (1907-1997) is represented as a circle. The LS and LD capture the observed statistics within the interquartile range as expected, but not the longer drought and surpluses seen in the reconstructed streamflows. These extend beyond or at the whiskers indicating these longer droughts are not well represented. The MS tends to slightly under represent the observed statistic while the MD tends to over represent this statistic. Again, the reconstructed streamflow is under represented for both. The average statistics all capture the observed within the interquartile range while the reconstructed flows are under represented for all the average statistics except avgS.

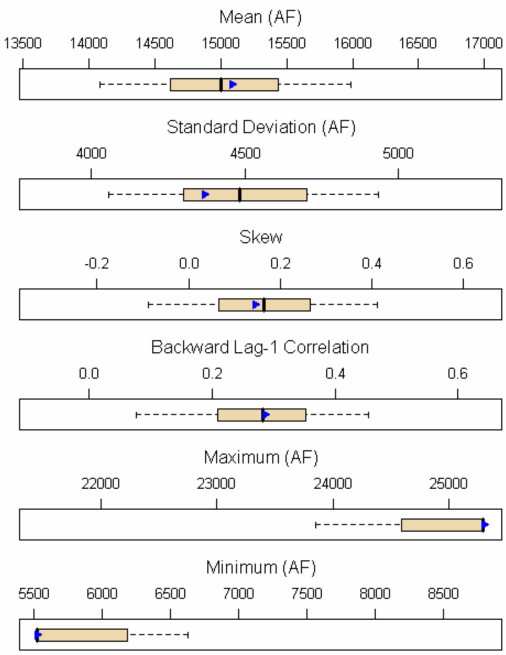

Figure 7. Basic statistics for Lees Ferry on the Colorado River, AZ.

Using the proposed framework, conditioning with the reconstructed flows while resampling with the traditional K-NN, the drought and surplus statistics (Figure 9) better represent the properties of the reconstructed streamflows. Incorporation of the reconstruction state information generates droughts and surplus lengths that are longer than the lengths seem when using only the traditional K-NN to resample the observed streamflows. The average drought and surplus statistics all show a shift towards higher average statistics as seen in the reconstructed streamflow statistics.

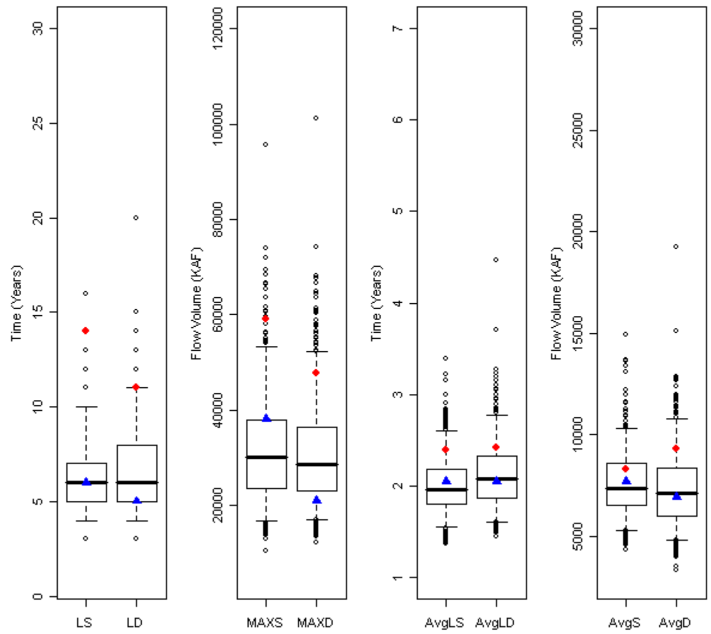

Figure 8. Drought statistics with only traditional K-NN resampling.

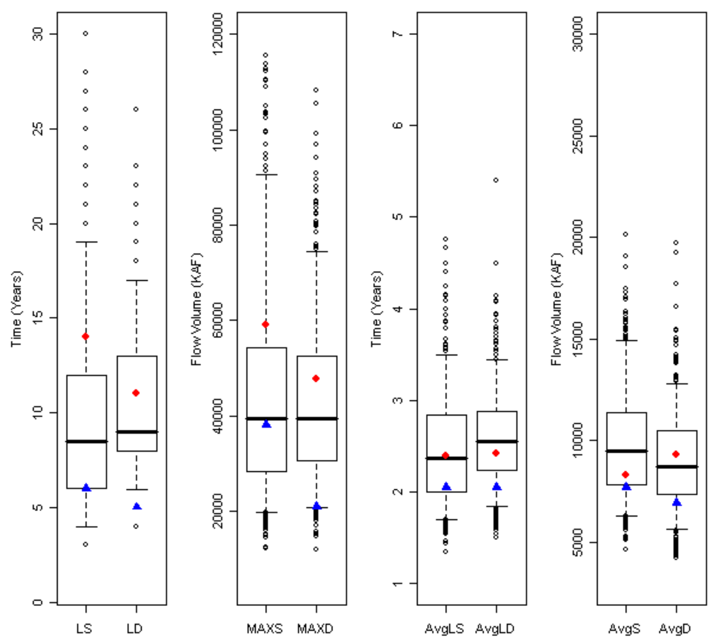

Figure 9. Drought statistics with proposed framework combining the Markov chain model with traditional K-NN resampling.

The surplus and drought statistics are based on a pre-selected threshold that determines whether a value represents a drought or surplus year. The choice of this threshold (median for the analysis shown here) will impact representation of both the drought and surplus statistics. In the Colorado River Basin a critical sequence of flows used in basin operations is a series of droughts connected over 12 years with years between each drought that are above the median. These multiple droughts with a surplus year interspersed are not represented with the previous drought statistics.

To address this, we can instead determine required basin storage for a given streamflow sequence and specified basin yield without requiring a subjective threshold. The required basin storage is found with the sequent peak algorithm (Loucks et al. 1981) defined as:

$$S_i^{'} = \begin{cases} S_{i-1}^{'} + d - y_i \\ 0 \end{cases} \tag{6}$$

$$S_c = \max\left[S_i^{'}, \ldots, S_N^{'}\right] \tag{7}$$

Where S_i is the storage at time step i, d is the demand or yield, y_i is the streamflow from a sequence at time i, and S_c is the storage capacity. The algorithm is run for all values in the inflow sequence from 1 to N, where N is the number of years in the inflow sequence. The algorithm is run for various demand (yield) levels and plotted (Figure 10), with the historic flow shown as a triangle and results from each trace of the 500 simulations shown as boxplots. Figure 10 indicates that at a demand of 16.5 MAF the historic inflow sequence requires 325 MAF of storage capacity to reliably meet the demand. The simulations display a range of storage capacities required to meet the given demand based on the 500 paleo-conditioned streamflow sequences generated with our proposed method. On average the simulations require the same storage capacity as the observed flows. The inter quartile range spans from a required storage of 250 MAF to 375 MAF.

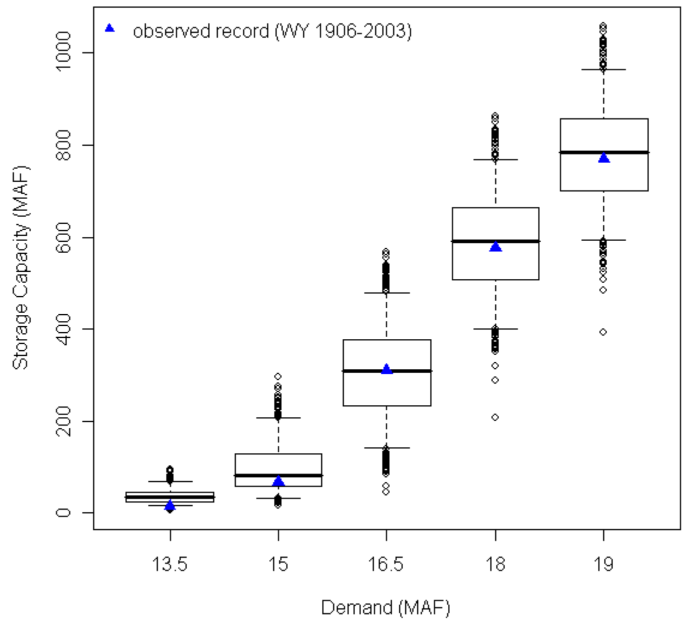

Figure 10. Sequent peak algorithm results.

A reliability estimate of meeting a demand for a given storage can be determined from the boxplots. The boxplot, plotted as a probability density function (PDF) (Figure 11a) or a cumulative density function (CDF) (Figure 11b), can easily be used to find the reliability. For example, a demand of 16.5 cannot reliability be met 99% of the time for a storage capacity of 60 MAF (the approximate current storage capacity of the Colorado River basin). The reliability probability is 1-area under the PDF curve (cross hatched area) or 1-0.99 = 0.01. The reliability of alternate storage capacities can be found from Figure 11a or 11b in a similar manner.

It is important to note that the sequent peak method assumes that the demand is unchanging through time and the demand must be meet in all years. These assumptions limit the ability to read policy decisions from these Figures. These Figures are primarily for screening of simulated flows and comparison with the observed flows. To fully appreciate the actual operations of a river basin a decision support system that incorporates variable demand schedules, proper topographic layout for river system reservoir, diversion points, and operating policy must be used.

In the next steps of this research the paleo-conditioned flows are used in a decision support system, which accounts for the aforementioned requirements, to better understand the implication to output variables when alternate hydrologies, such as these paleo-conditioned flows, are used.

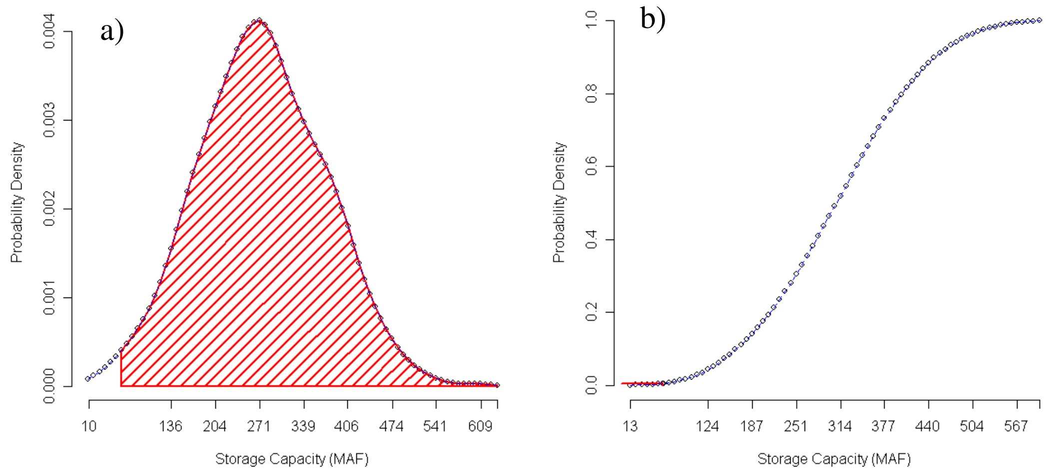

Figure 11. a) PDF and b) CDF for 16.5 MAF demand boxplot from Figure 10.

Summary and Discussion

A technique is presented to incorporate state information derived from reconstructed streamflows while preserving the distributional properties of an observed record. A nonhomogeneous Markov chain with kernel smoothing is used to compute transition probability matrices over a planning horizon based on reconstructed flows. System states are computed from the transition probability matrices. Lastly, a value from the observed record is conditionally resampled on the previous simulated flow, the previous simulated state and the current simulated state with a K-NN algorithm.

This technique was demonstrated for the Colorado River at Lees Ferry and was able to preserve the distributional properties of the observed record while also simulating drought and surplus sequences similar to those observed in the reconstructed streamflows.

The incorporation of hydrologic state information from the reconstructed streamflows, which extend 500 years into the past, greatly enhances risk and reliability estimates and allows consideration of climate variability in river basin modeling.

The annual streamflows generated with the proposed framework can be spatially and temporally disaggregated (Prairie, et al. 2006) to represent ensembles at multiple gauges in a river basin. These ensembles can then be applied in a decision support system (Prairie, 2006), allowing decision makers to consider the impact of climate variability.

Acknowledgements

Funding for this research by the Bureau of Reclamation's Lower Colorado regional office via grant #04PG303326 is gratefully acknowledged. Continued support by Kib Jacobson of the Bureau of Reclamation's Upper Colorado regional office is appreciated. Thanks are also due to the Center for Advanced Decision Support in Water and Environmental Systems (CADSWES) at the Univ. of Colorado, Boulder for use of its facilities and computational support.

References

Cook, E.R., K. Briffa, S. Shiyatov, and V. Mazepa (1990), Tree-ring standardization and growth-trend estimation, *in Methods of Dendrochronlogy: Applcations in the Environmental Sciences*, edited by E.R. Cook and L.A. Kairiukstis, pp. 153-162, Springer, New York.

Gates, P., and H. Tong (1976), On Markov chain modeling to some weather data, *Journal of Applied Meteorology*, 15, 1145-1151.

Hidalgo, H.G., T.C. Piechota, and J.A. Dracup (2000), Alternative principal components regression procedures for dendrohydrologic reconstructions, *Water Resources Research*, 36(11), 3241-3249.

Hirschboeck, K.K. and D. M. Meko (2005), *A Tree-Ring Based Assessment of Synchronous Extreme Streamflow Episodes in the Upper Colorado & Salt-Verde-Tonto River Basins- Final Report*, Laboratory of Tree-Ring Research, University of Arizona.

Lall, U. and A. Sharma (1996), A nearest neighbor bootstrap for resampling hydrologic time series, *Water Resources Research*, 32(3), 679-693.

Loucks, D.P., J.R. Stedinger, and D.A. Haith (1981), *Water Resource System Planning and Analysis*, Prentice-Hall, Inc., New Jersey.

Prairie, J. (2006), *Stochastic nonparametric framework for basin wide streamflow and salinity modeling: application for the Colorado River basin*, Ph.D. Dissertation, University of Colorado, Boulder, Colorado.

Prairie, J. and R. Callejo (2005), *Natural flow and salt computation methods*, U.S. Dept. of Interior, Salt Lake City, Utah.

Prairie, J.R., B. Rajagopalan, U. Lall, and T.J. Fulp (2006), A stochastic nonparametric technique for space-time disaggregation of streamflows, *Water Resources Research* (in press).

Rajagopalan, B., U. Lall (1995), A kernel estimator for discrete distributions, *Journal of Nonparametric Statistics*, 4, 409-426.

Rajagopalan, B., U. Lall, and D.G. Tarboton (1996), Nonhomogeneous Markov model for daily precipitation, Journal of Hydrologic Engineering, 1(1), 33-39.

Scott, D.W. (1992), *Multivariate density estimation: Theory, practice, and visualization*. Wiley Ser. In Probability and Math. Statistics. John Wiley and Sons, New York, N.Y.

Stockton, C.W. and G.C. Jacoby (1976), Long-term surface-water supply and streamflow trends in the Upper Colorado River Basin, *Lake Powell Res. Proj. Bull.* 18, Natl. Sci. Found., Arlington, Va.

Stokes, M.A. and T.L. Smiley (1968), *An Introduction to Tree-Ring Dating*, Univ. of Ariz. Press, Tucson.

Swetnam, T.W., M.A. Thompson, and E.K. Sutherland (1985), *Using Dendrochronology to Measure Radial Growth of Defoliated Trees, Agric. Handb.*, vol. 639, For. Serv., U.S. Dep. Of Agric., Washington, D.C.

Weisberg, S. (1985), *Applied Linear Regression*, 2nd ed., John Wiley, Hoboken, N.J.

Woodhouse, C.A. and P.M. Brown (2001), Tree-ring evidence for Great Plains drought, *Tree-Ring Research*, 57, 89-103.

Woodhouse, C.A., S.T. Gray, and D.M. Meko (2006), Updated streamflow reconstructions for the Upper Colorado River Basin, *Water Resource Research*, 42, W05415, doi:10.1029/2005WR004455.