UPSTREAM REGULATION ADJUSTMENTS TO ENSEMBLE STREAMFLOW PREDICTIONS

ARIS P. GEORGAKAKOS AND HUAMING YAO

Georgia Water Resources Institute, Georgia Institute of Technology, Atlanta, GA 30332

In collaboration with

KONSTANTINE P. GEORGAKAKOS

Hydrologic Research Center, San Diego, CA 92130

Sponsored by NOAA/NWS/Office of Hydrologic Development

(Award No. NA08NWS4620023)

HRC TECHNICAL REPORT NO. 7

Hydrologic Research Center

12780 High Bluff Drive, Suite 250, San Diego, CA 92130, USA

30 June 2010

Page Left Blank

ACKNOWLEDGEMENTS

The research work reported herein was sponsored by the Office of Hydrologic Development of the National Weather Service, National Oceanic and Atmospheric Administration, under Award No. NA08NWS4620023. The research is a collaborative effort between staff of the Hydrologic Research Center with expertise in operational hydrometeorological prediction and staff of the Georgia Water Resources Institute with expertise in water resources systems management. The authors are grateful to Mr. Robert Hartman and the staff of the California Nevada River Forecast Center for significant contributions in selecting the case study watersheds, facilitating collaboration with upstream regulation agencies of the American River in California, and for providing historical data and hydrologic model parameters from the operational CNRFC databases. We also acknowledge the contributions of Dr. Pedro Restrepo of the Office of Hydrologic Development for providing strategic advice and reviews throughout the project tenure, and Mr. Tom Gurss of MBRFC, who provided a technical comment that is addressed in Appendix B of this report and elucidates validation issues associated with ensemble prediction of drought periods. The ideas and opinions expressed herein are those of the authors and need not represent those of the National Oceanic and Atmospheric Administration and of its sub-Agencies.

This report should be cited as follows:

Georgakakos, et al. 2010: Upstream regulation adjustments to ensemble streamflow prediction. *HRC Technical Report No. 7*. Hydrologic Research Center, San Diego, CA, 30 June 2010 (NA08NWS4620023), 63 pp.

Page Left Blank

EXECUTIVE SUMMARY

The work reported herein addresses the problem of adjusting Ensemble Streamflow Predictions (ESP), routinely issued by the US National Weather Service River Forecast Centers, to account for upstream regulation effects to downstream flows in real time. The solution methodology includes the use of regression relationships to eliminate bias in the ESP predictions for multiple lead times, the use of a conceptual reservoir regulation model to approximate the aggregate upstream storage effects, and the incorporation of sources and sinks to account for diversions in and out of the watershed of interest. A detection methodology is also advanced as an aid to determining whether upstream storage effects exist for a particular watershed application. Upstream regulation adjustments to the ESP are made to include estimates of the uncertainty in the regression and the upstream storage regulation models. The upstream regulation adjustment methodology is designed for application using only operationally available data. Observed downstream flows, and simulated and ESP flows generated by operational hydrologic models are adequate to estimate the parameters of the models developed and the uncertainty in the predictions due to these models.

The theoretical development is exemplified in two real world applications. The Middle Fork and the South Fork are both rain- and snow-fed tributaries of the American River in California that undergo significant and diverse upstream regulation. Regulation is effected through multiple storage facilities upstream and several diversions that transfer water between the two watersheds, some generating consumptive plant use from the diverted water flow. The downstream flows observed for each Fork are substantially higher than their unimpaired flows during the late summer and fall seasons due to upstream reservoir release, and their spring and early summer flows are substantially lower than the corresponding unimpaired flows due to storage filling upstream. Seasonal water imports and exports through diversion facilities, and variable policies pertaining to the operation of the upstream facilities add further complexity to the application of the upstream regulation methodologies to these watersheds. Without adjustment, the operational ESP produces biased flows for a significant part of the year in the low-to-medium flow range.

Extensive validation of the upstream regulation models developed is performed for the Middle and South Forks of the American River using retrospective analysis of daily flows for a period of twenty years (1978 – 1997). The North Fork, a minimally regulated tributary to the American River, is used

to contrast the results in cases of watersheds that are upstream regulated versus watersheds without upstream regulation. Validation focuses on the performance target variables of monthly, weekly and individual day-of-the-week flow averages at the downstream end of the watersheds for the months of September and October of each year; months with considerable upstream regulation effects on the downstream flows. ESP lead times examined range from 1 month (e.g., forecast preparation time on the first Monday in August with forecast valid time the first Monday of September) to more than 4 months (e.g., forecast preparation time the first Monday in June with forecast valid time the Sunday in the third week of October).

Twenty-year climatological means of observed versus adjusted-ESP simulated flows show the significant average improvement offered for all cases examined by the methods advanced in this work compared to the unadjusted-ESP simulated flows. Rank Probability Scores (RPS), appropriate for probabilistic forecasts, are computed together with associated skill scores that reflect percent improvement with respect to unadjusted ESP flows. Improvement is often greater than 20% and reaches up to 99 percent for monthly, weekly and daily scales. Boxplots of adjusted and unadjusted ESP are produced for the target variables together with the corresponding observed values for the range of forecast lead times. The improvement of the adjusted ESP probabilistic forecasts over the unadjusted ESP forecasts is evident, with the adjusted ESP forecasts containing the observations in most cases. Overall, the adjusted ESP forecasts for the South Fork are more reliable than those for the Middle Fork. It is also important to note that for certain historical years and for certain lead times in the Middle Fork, the adjusted ESP forecasts show small improvement (if any) over the unadjusted ESP forecasts for certain day of the week. These effects are attributed to occasional deviations of the upstream regulation policies from their long range pattern as identifiable from the historical record. The effectiveness of the regression models of the methodology to correct for biases is also demonstrated by the validation results.

The American River applications demonstrate that the methodology advanced for the adjustment of ESP to account for upstream regulation effects is effective in substantially improving the ESP both from climatological and probabilistic perspectives during the seasons when upstream regulation is significant. They also attest to the fact that operational information is adequate for the estimation of free parameters. As such, it is a good candidate for operational implementation and testing on various watersheds with significant upstream regulation effects and with various climatic and regulation characteristics.

TABLE OF CONTENTS

Page left blank

List of Tables

LIST OF FIGURES

1. Introduction

The focus of the research reported herein is the development of models and procedures for adjusting operational ensemble streamflow predictions (ESP) for the effects of upstream regulation. The ESP methodology is used routinely by the National Weather Service (NWS) River Forecast Centers (RFCs) in the United States and by forecast offices in other countries to produce an ensemble of equally likely streamflow traces for a season or longer at desired forecast points in natural and regulated watersheds. At present and for basins that undergo regulation upstream of the forecast point of interest, ESP traces reflect unimpaired flow (or full natural flow) with no account for the upstream regulation. The present work demonstrates the viability of adjusting the ESP traces to account for upstream regulation through the development of generally applicable procedures and models that are suitable for operational application and through the validation of such procedures and models with operational data from two watersheds with significant upstream regulation: the Middle and South Fork of the American River in California.

There is a variety of situations that arises in terms of upstream regulation in practice (*Woodbury 2004*). Pertinent to our focus, we classify these here in terms of *(a)* the effects on the downstream (with respect to the site of regulation) flow at the ESP forecast point (as small and large), and *(b)* whether the upstream regulation parameters are knowable with defined uncertainty or unknowable (proprietary and undisclosed plant operations). In particular, with respect to *(b)*, we refer to both the regulation node parameters (e.g., reservoir active and flood storage capacities, spillway capacities, withdrawal canal or conduit capacity) and the operating rules and targets of the regulation node (e.g., reservoir operating rule, hydroelectric energy production targets, and withdrawal targets). The former classification with respect to downstream effects influences the degree to which uncertainty in upstream regulation affects the downstream ESP, while the latter classification provides a range of uncertain situations in practice.

The term *ensemble streamflow prediction* or ESP is used herein to describe the operational US NWS procedure whereby precipitation and temperature input from past years is used as input to a well calibrated hydrologic model to produce an ensemble of streamflow traces that are considered equally likely to occur for some future period of interest. The historical input used from each historical year is for the same month, day and hour interval as the interval of interest in the future period.

In this research and as a first approach to characterizing upstream regulation uncertainty, we will only focus on situations with moderate to large downstream influence and with regulation parameters and operating rules that have knowable uncertainty. In such cases, in addition to uncertainty in regulation parameters and operating rules, uncertainty in upstream regulation exists because of real time deviations from operating rules due to anticipated weather and flow conditions (typically extreme). Our methodology aims to characterize these sources of uncertainty on the basis of historical information and to incorporate them into the ESP procedure. The value of reliable uncertainty measures for downstream regulation has been shown in several studies (e.g., *Georgakakos and Krzysztofowicz 2001; Georgakakos et al. 1998*), and the research results are expected to benefit water resources management in downstream points with significant upstream regulation influences.

The methodology consists of a procedure to identify the presence of significant upstream regulation influences to downstream point of interest using recent historical information, upon detection, a model for correcting the ensemble streamflow predictions for bias during periods when upstream regulation influences are not significant, and a model to account for upstream regulation influences that are due to storage effects and water transfers. The latter model is complemented by procedures to determine the climatological seasonal "calendar" of upstream regulation effects from historical data, and to adjust the NWS operational ensemble streamflow predictions at the downstream point of interest to incorporate estimates of the uncertainties in upstream regulation model and procedures. The next section describes the modeling framework, and it is followed by a section that provides an overview of the study basins and available hydrologic modeling and forecast data.

1.1 Modeling Framework: Typical Data, Motivation, and Overview

The need for forecast modification comes about because of the cumulative flow modification that distributed water uses (within a watershed) exert on watershed outflow. Typical water uses may include water supply for irrigation, domestic, or industrial use; small hydro-plant operation; and low flow augmentation for environmental and ecosystem sustainability. Some of the uses are consumptive, having direct impact on the quantity of available water, while others modify the timing of the natural flows. This research concerns itself with situations where such water uses are enabled by *(a)* direct water withdrawals and use without the use of storage facilities, *(b)* water transfers in or out of the watershed, and *(c)* several relatively small storage facilities distributed across the watershed. Watersheds in which seasonal or over-year water regulation occurs primarily at a few major storage facilities or other water works require more detailed modeling of the physical plants and regulation procedures and is beyond the scope of this project.

As a motivation for the proposed modeling framework, Figures 1 and 2 show the unimpaired (blue line) and observed daily outflows (red line) at the outlets of the Middle and North Fork American River watersheds, respectively, in two typical years. The unimpaired flow (also referred to as full natural flow) data series represents outflows that would have occurred in the absence of upstream regulation and is usually generated by hydrologic models driven by observed precipitation and evapotranspiration sequences.

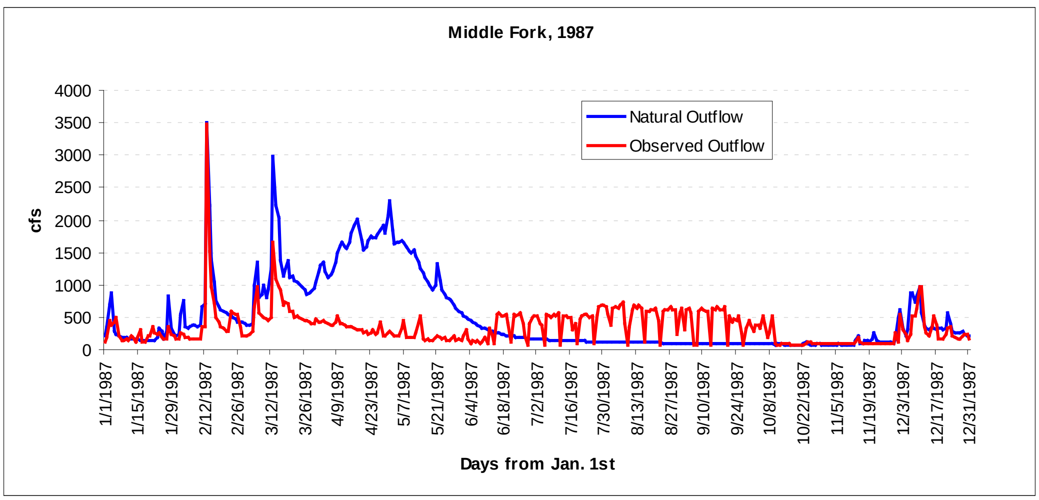


Figure 1: Comparison of Unimpaired and Observed Outflow, Middle Fork, 1987

Figures 1 and 2 illustrate *(a)* the type of information that is generally available for estimating upstream regulation and *(b)* the motivation for a general modeling approach applicable to watersheds with and without apparent upstream regulation. Typically, the available information includes:

- Daily or sub-daily flow observations at the watershed outlet;
- Unimpaired watershed outflow sequences, either model generated (based on contemporaneous data of watershed precipitation, temperature, evapotranspiration, and flow) or reconstructed from observed outflows and knowledge of existing water uses;

- Anecdotal or quantitative information (obtained by water agencies and other stakeholders) on the nature, timing, and quantities of water uses and transfers (be they exports or imports), including release rules for some of the existing storage facilities and instream flow requirements.

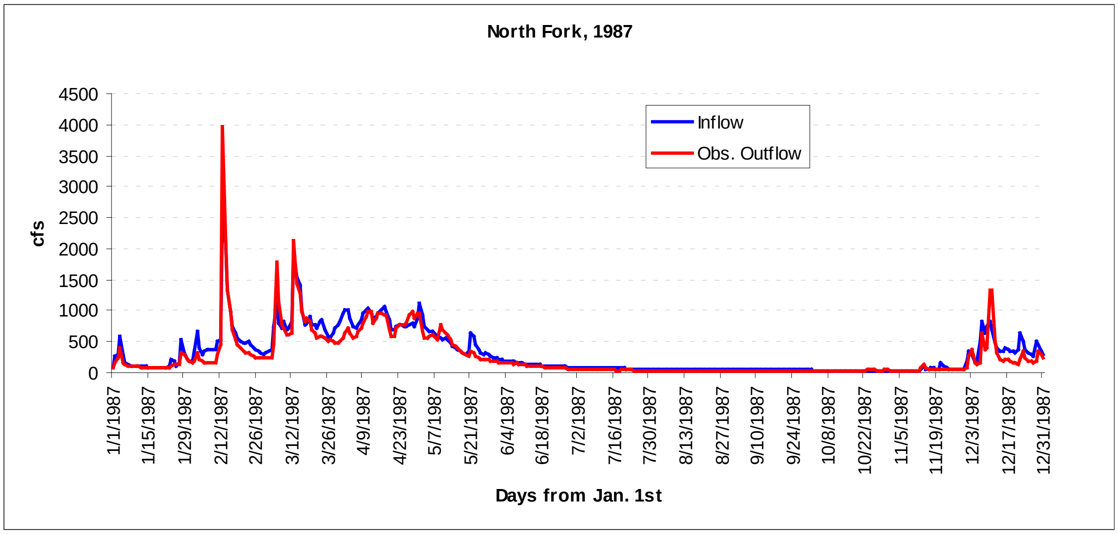


Figure 2: Comparison of Unimpaired and Observed Outflow, North Fork, 1987

The modeling approach relies primarily on the first two information types. If available, the third type of information can be used to further ascertain and validate the exploratory data analysis.

The second important observation pertaining to Figures 1 and 2 is that forecast adjustment can be beneficial for watersheds with and without upstream regulation. The North Fork watershed in Figure 2 exemplifies a case without obvious upstream regulation effects, while the Middle Fork watershed in Figure 1 clearly exhibits significant inter-seasonal regulation.

As mentioned earlier, upstream regulation also includes flow modifications that do not entail storage regulation such as direct water use, imports, and exports. Such modifications become part of the biases between observed and "natural" outflows and are not easily distinguishable from them. In such cases, Regression and/or Neural Network models are well suited for bias removal and forecast correction.

Figure 1 illustrates the effects of upstream storage regulation. In this case, one can distinguish three different time periods: a spring storage filling period where unimpaired flows consistently exceed observed flows, a summer storage release period where observed flows are clearly augmented with respect to unimpaired flows, and the rest of the year outside the storage regulation periods. In such watersheds, the extent and duration of flow augmentation depends on the available storage, which, in turn, depends on the antecedent hydrologic conditions. This type of dynamic flow modifications involve more than seasonal flow biases and cannot be fully captured by Regression and Neural Network models. Such cases are more effectively handled by an aggregate watershed storage model. Regression and Neural Network models are still potentially useful in the third period, where storage regulation effects are insignificant.

Thus, the proposed modeling framework is designed to utilize different modeling approaches depending on upstream regulation type and time of the year. Figure 3 schematically summarizes the modeling framework logic and main components.

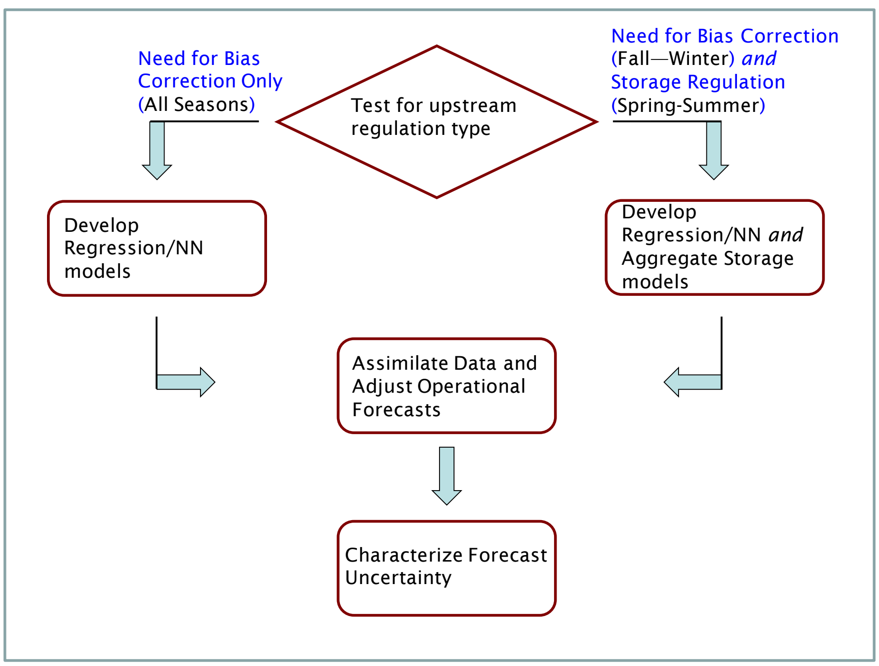


Figure 3: General Modeling Framework

1.2 CASE STUDY WATERSHEDS AND HYDROLOGIC FORECASTS

The three forks (North, Middle and South Fork) of the American River drain approximately 4,800 km^2 of the mountainous terrain of central California (with elevations up to 3,000 m) and join to provide inflow to Folsom Lake (Figure 4). The catchment with outlet at Folsom Lake is characterized by typical orographic rainfall patterns associated with steep terrain barriers, and with snow in the high elevations (typically above 1500 m). The climatological means of hourly precipitation, based on a sample of precipitation events for the wet period 1980-1987, show a maximum of about 2 mm/hr over the headwaters of the North Fork of the American River with pronounced variability. The automated operational gauge network provides estimates of the mean areal precipitation that are nearly unbiased for the entire inflow watershed, but which possess non-negligible bias for the Fork sub-catchments (*Tsintikidis et al. 2002*). The catchment average response time to significant rainfall events in the absence of snow is approximately 12 hours. Of particular interest for this work is the fact that significant and diverse upstream regulation is documented for the Middle and South Forks of the American River. Available data from the operational files of the California Nevada River Forecast Center (CNRFC) of the U.S. National Weather Service consist of: six-hourly mean areal precipitation and temperature for sub-catchments of the basin, monthly climatologies of daily potential evapotranspiration demand for each sub-catchment, and observed mean daily streamflow for all the Forks and reconstructed Folsom Lake inflow from Lake levels.

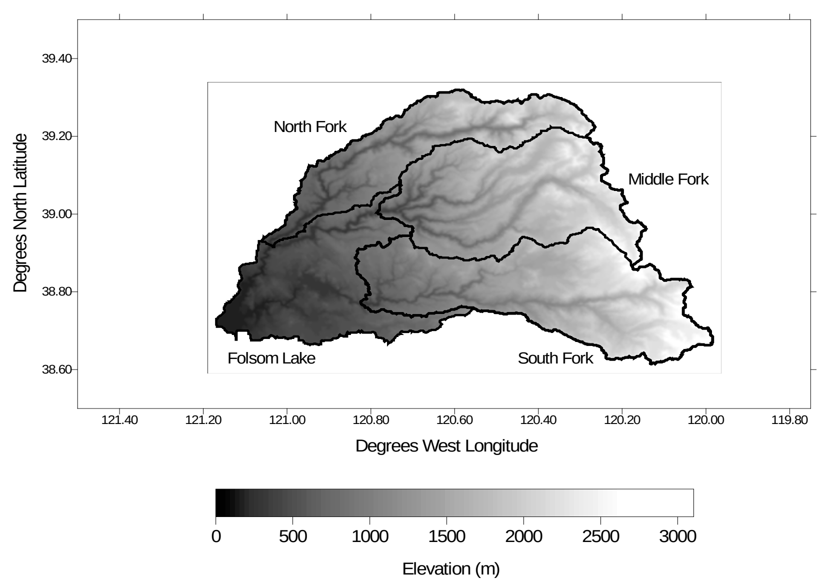


Figure 4: Folsom Lake Drainage and the Watersheds of the North, Middle and South Forks of the American River.

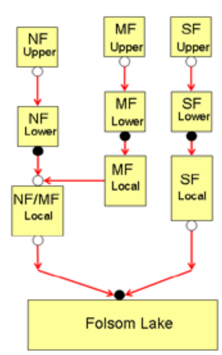
The hydrologic model used in this study to generate flow simulations and forecasts contains important features of the CNRFC operational hydrologic model, including the components for snow accumulation and ablation, soil water accounting, and channel routing. These components of the stand alone hydrologic model were designed and implemented to mirror the analogous components of the operational CNRFC forecast model. It is an adaptation of the operational model as it includes distributed channel routing in order to more accurately reproduce the timing of the flows throughout the stream network. The model consists of adaptations of the operational snow accumulation and ablation model (*Anderson 1973*) and the Sacramento soil water accounting model as described in *Georgakakos (1986)*. For channel routing, the kinematic channel routing model of *Georgakakos and Bras (1982)* is used in the form of a sequence of linear conceptual reservoirs, with parameters estimated from the CNRFC estimates of unit hydrographs applied to the sub-catchments of interest (see *Sperfslage and Georgakakos 1996* for a description of the procedure).

The hydrologic basin upstream of the Folsom Lake reservoir was subdivided into sub-basins considering stream gauge sites, significant upstream reservoir facilities, available automated precipitation and temperature sensors, and the topology of the channel network. Those sub-basins, which have significant elevation differences within their areas, are further subdivided into sub-areas (an upper and a lower sub-area in this version of the stand alone model). The snow and the soil-water models are applied to each of the sub-areas to produce rain plus melt and channel inflow volumes, respectively. These volumes are then fed into the channel routing model and are carried downstream through the channel network undergoing time distribution, advection and attenuation. The model produces outflow at all the gauging sites and all the junctions of the model-channel network, and, of course, at the basin outlet (inflow point into the reservoir). It is important to note that the stand-alone model is designed to use the same input as the operational hydrologic forecast model, and its parameters bear close relationship to the parameters of the operational hydrologic model. The values of the model parameters used by the operational model for the snow and soil-water components were used in the stand alone model as well, while (as mentioned earlier) a calibration process with available data was used to determine parameters for the channel routing model.

The configuration of the stand-alone model elements is exemplified for the Folsom Lake drainage in Figure 5. The North (NF), Middle (MF) and South (SF) Fork sub-basins are indicated, sub-divided into an upper and a lower sub-area for snow-pack, soil-water accounting and channel routing.

Channel routing occurs in each sub-area of each sub-basin and at channel network junctions the inflows are summed. Channel routing is indicated with red arrows in the Figure. There are four streamflow observation sites in the basin, shown with black filled circles. Of these, the one corresponding to the inflow point to Folsom Lake reports lake levels, which are transformed to unimpaired flows. The model used also performs channel routing to the junctions without observations (open circles) to allow for the reproduction of the observed 6-hour hydrograph.

The kinematic channel routing component of the stand alone model for each channel segment is based on a series of linear reservoirs with identical parameters. The sum of the inverse of the channel routing model parameters for all the reservoirs representing a single channel segment is equal to the travel time in the channel segment. The operational model uses unit hydrographs to reproduce channel processes. For the North, Middle and South Fork sub-basins, initial estimates of the parameters of the channel routing component of the stand-alone model were obtained by fitting the linear reservoir model to the appropriate unit hydrographs (e.g., see *Sperfslage and Georgakakos 1996* for numerical fitting procedure). Initial values of the parameters of the channel segments downstream of the Forks were based on preliminary estimates of the travel time in these segments based on drainage area size. Table 1 shows the parameter values of the snow, soil and channel components of the stand alone hydrologic model for the Folsom Lake drainage sub-basins. The nomenclature of Table 2 is used. Table 3 shows the long-term-averaged daily values of evapotranspiration demand by month (adopted from the operational parametric input files of CNRFC) used by the model for the present numerical experiments



.

Figure 5: Representation of Folsom Lake Drainage by the Hydrologic Prediction Model.

Table 1: Nominal Values of Stand Alone Model Parameters*

SNOW PARAMETERS

	NFu	*NFl*	*MFu*	*MFl*	*SFu*	*SFl*	*FL*
SCA	1.0	1.0	1.35	1.0	1.2	1.0	1.0
MFMAX	0.86	0.85	0.69	0.5	0.75	0.85	0.8
MFMIN	0.2	0.3	0.12	0.16	0.2	0.25	0.25
NMF	0.15	0.15	0.15	0.15	0.15	0.15	0.15
PLWHC	0.04	0.04	0.04	0.04	0.04	0.04	0.04
TIPM	0.25	0.25	0.25	0.25	0.25	0.25	0.25
MBASE	1.0	1.0	1.0	1.0	1.0	1.0	0.0
UADJ	0.04	0.04	0.04	0.04	0.08	0.06	0.04
DAYGM	0.1	0.1	0.1	0.1	0.1	0.1	0.1
PXTEMP	2.0	2.0	2.0	2.0	2.0	2.0	1.0
SI	900.	300.	1200.	600.	1100.	500.	200.
ELV	19.86	9.60	19.81	13.72	20.29	5.90	4.57
PADJ	1.0	1.0	1.0	1.0	1.1	1.05	0.97

SACRAMENTO MODEL PARAMETERS

	NFu	*NFl*	*MFu*	*MFl*	*SFu*	*SFl*	*Fl*
UZTWM	142.000	161.000	90.000	140.000	100.000	175.000	75.000
UZFWM	55.000	35.000	35.000	45.000	65.000	90.000	15.000
LZTWM	312.000	360.000	270.000	280.000	250.000	600.000	180.000
LZFPM	72.000	72.000	96.000	110.000	125.000	350.000	100.000
LZFSM	110.000	85.000	120.000	110.000	20.000	60.000	80.000
DU	0.075	0.070	0.105	0.115	0.040	0.050	0.062
DLPR	0.001	0.002	0.001	0.002	0.001	0.001	0.001
DLDPR	0.018	0.030	0.023	0.015	0.007	0.007	0.018
EPS	20.000	20.000	48.000	43.000	30.000	100.000	12.000
THSM	1.400	1.400	1.300	1.500	2.100	1.100	1.200
PF	0.250	0.350	0.150	0.300	0.250	0.250	0.250
XMIOU	0.000	0.000	0.000	0.000	0.000	0.000	0.100
ADIMP	0.010	0.010	0.000	0.020	0.000	0.000	0.075
PCTIM	0.000	0.000	0.005	0.005	0.000	0.000	0.065
ETADJ	1.000	1.000	1.000	1.000	1.000	1.000	1.000

KINEMATIC CHANNEL ROUTING MODEL INITIAL PARAMETERS

	NFu	*NFl*	*MFu*	*MFl*	*SFu*	*SFl*	*MF-NF*	*MF/NF-F*	*SF-F*
n_c	1	2	1	2	3	1	2	2	2
α	5.40	0.85	4.40	0.95	4.40	0.80	4.0	4.0	0.95

SUB-CATCHMENT AREAS (km^2)

	NFu	*NFl*	*MFu*	*MFl*	*SFu*	*SFl*	*Fl*
Area	325.1	550.4	713.0	533.5	898.6	632.3	1016.3

*See Table 2 for nomenclature used in this Table

Table 2: Nomenclature for Table 1

HEADINGS

For Snow, Sacramento Models, Channel Routing Model and for Areas

NFu: NORTH FORK UPPER SUB-AREA
NFl: NORTH FORK LOWER SUB-AREA
MFu: MIDDLE FORK UPPER SUB-AREA
MFl: MIDDLE FORK LOWER SUB-AREA
SFu: SOUTH FORK UPPER SUB-AREA
SFl: SOUTH FORK LOWER SUB-AREA
Fl: FOLSOM LAKE LOCAL SUB-BASIN

For Channel Routing Model

MF-NF: CHANNEL SEGMENT CONNECTING THE OUTLET OF MIDDLE FORK WITH A JUNCTION POINT DOWNSTREAM OF THE NORTH FORK OUTLET
NF/MF-F: CHANNEL SEGMENT THAT CONNECTS THE JUNCTION POINT DOWNSTREAM OF NORTH FORK OUTLET WITH FOLSOM LAKE INFLOW POINT
SF-F: CHANNEL SEGMENT THAT CONNECTS THE OUTLET OF SOUTH FORK WITH FOLSOM LAKE INFLOW POINT

SNOW MODEL PARAMETERS

SCA: SNOW CATCH ADJUSTMENT FACTOR
MFMAX: MAXIMUM MELT FACTOR (MM $DEGC^{-1}$ D^{-1})
MFMIN: MINIMUM MELT FACTOR (MM $DEGC^{-1}$ D^{-1})
NMF: MAXIMUM NEGATIVE MELT FACTOR (MME $DEGC^{-1}$ D^{-1})
PLWHC: FRACTION OF SNOW COVER FOR WATER HOLDING SNOW CAPACITY
TIPM: PARAMETER FOR ANTECEDENT TEMPERATURE INDEX COMPUTATIONS
MBASE: BASE TEMPERATURE FOR MELT COMPUTATIONS (DEGC)
UADJ: AVERAGE DAILY WIND FUNCTION FOR RAIN-ON-SNOW PERIODS (MM MB^{-1} DAY^{-1})
DAYGM: CONSTANT MELT AT SNOW-SOIL INTERFACE (MM DAY^{-1})
PXTEMP: TEMPERATURE TO DELINEATE RAIN FROM SNOW (DEGC)
SI: MAXIMUM SWE FOR 100% COVER IN SNOW DEPLETION CURVE (MM)
ELV: ELEVATION OF CENTROID OF BASIN (10^2 M)
PADJ: PRECIPITATION ADJUSTMENT FACTOR

SACRAMENTO MODEL PARAMETERS

UZTWM: UPPER ZONE TENSION WATER CAPACITY (MM)
UZFWM: UPPER ZONE FREE WATER CAPACITY (MM)
LZTWM: LOWER ZONE TENSION WATER CAPACITY (MM)
LZFPM: LOWER ZONE FREE PRIMARY WATER CAPACITY (MM)
LZFSM: LOWER ZONE FREE SUPPLEMENTARY WATER CAPACITY (MM)
DU: INTERFLOW RECESSION ($6HRS^{-1}$)
DLPR: RECESSION COEFFICIENT FOR LOWER ZONE FREE PRIMARY WATER ELEMENT ($6HRS^{-1}$)
DLDPR: RECESSION COEFFICIENT FOR LOWER ZONE FREE SUPPLEMENTARY WATER ELEMENT ($6HRS^{-1}$)
EPS: CONSTANT FACTOR IN PERCOLATION FUNCTION
THSM: EXPONENT IN PERCOLATION FUNCTION
PF: FRACTION OF PERCOLATION BYPASSING THE LOWER ZONE TENSION WATER ELEMENT
XMIOU: FRACTION OF WATER LOST TO DEEP GROUNDWATER LAYERS
ADIMP: ADDITIONAL IMPERVIOUS AREA MAXIMUM FRACTION
PCTIM: FRACTION OF PERMANENTLY IMPERVIOUS AREA
ETADJ: EVAPOTRANSPIRATION DEMAND ANNUAL ADJUSTMENT FACTOR

CHANNEL MODEL PARAMETERS

n_c: NUMBER OF LINEAR RESERVOIRS REPRESENTING THE CHANNEL SEGMENT UNDER STUDY
α: COMMON COEFFICIENT OF LINEAR RESERVOIRS WITH INVERSE DESCRIBING TRAVEL TIME ($6HRS^{-1}$)

Table 3: Daily Values of Evapotranspiration Demand Used by the Sacramento Model for Each Month (mm/d)

	NFu	NFl	MFu	MFl	SFu	SFl	Fl
J	0.760	1.280	0.760	1.280	0.780	1.300	0.860
F	0.780	1.400	1.060	1.860	1.450	2.470	1.120
M	0.820	1.800	1.470	2.520	1.670	2.940	1.640
A	1.030	2.290	1.950	3.110	1.800	3.200	2.480
M	1.800	3.640	2.550	4.110	2.280	3.850	4.150
J	3.040	6.040	4.320	6.330	3.580	7.390	4.560
J	5.260	8.220	5.400	8.650	5.760	9.160	4.640
A	5.570	8.250	6.150	9.730	5.840	8.760	4.100
S	4.100	6.550	4.770	6.950	3.270	3.790	3.220
O	1.940	3.100	2.690	3.120	1.810	2.300	2.200
N	1.140	1.690	1.190	1.440	1.360	2.050	1.230
D	0.910	1.400	0.940	1.250	1.080	1.800	0.880

1.3 Report Organization

The next section describes the procedures developed to identify whether upstream storage effects and transfers are significant. This is followed by a section on the model used to account for downstream influences of upstream storage management (one or more reservoirs) and water transfers in and out of the basin of interest. Section 4 discusses the models tested for bias adjustment, and Section 5 presents the validation of models and procedures for two regulated watersheds of the American River in California. Conclusions and recommendations are in Section 6. Appendix A presents an example of a step by step procedure for the development of ensemble forecasts based on conditioned historical analog methods, while Appendix B clarifies validation issues.

Page Left Blank

2. Detection of Regulation Type

The modeling approach of Figure 3 begins by determining the type and extent of existing regulation. While relevant anecdotal information may exist, it is desired that this determination be quantitative and will thus be based on the observed and unimpaired outflow series at the watershed outflow. As expected, the generation or reconstruction of the unimpaired outflow series involves errors, and the purpose of the proposed tests is to detect hidden biases and deterministic trends.

The first step in the testing process is to establish the *background* distribution of the errors between the unimpaired and the observed outflow data that is free of regulation effects. The most effective way to accomplish this is to identify an early watershed period for which water use and regulation were minimal. For example Figures 6, 7, and 8 display the daily error distributions for the Middle Fork winter (December to February), spring (March to May), and summer (June to November) for an early five year period (1960 to 1964; blue line) and a more recent period (1984 to 1988; red line). The figures clearly show distinct distributional shifts for spring (toward positive values) and summer (toward negative values), and less pronounced differences for winter. The statistical significance of these shifts can be detected using readily available non-parametric tests such as the Wilcoxon-Mann-Whitney rank-sum test (modified for the existence of serial data correlation) or applicable parametric tests depending on the nature of the underlying distributions. In the case of the Middle Fork, the distributional shifts for spring and summer are statistically significant (as per the W-M-W test), while the one for winter is not.

The existence of statistically significant and opposite error distribution shifts for spring and summer indicates a storage-enabled water transfer from spring to summer. If only one of these periods were detected to exhibit a significant shift, or if the shifts for both periods were significant but in the same direction, the conclusion would be that the watershed experiences water use and water transfer flow regulation that are not enabled by watershed storage facilities. In such cases, a Regression or a Neural Network (NN) forecast correction approach would be most suitable.

Figures 9, 10, and 11 display the same error distributions for the North Fork watershed. In this case, the W-M-W test cannot detect any statistical significant differences, for periods 1 and 3, but it does indicate a mild but significant (negative) shift for period 2, indicative of possible water transfers out of the watershed. The conclusion is that the North Fork watershed is not regulated, although period 2 would potentially benefit from the application of a Regression/NN type forecast correction scheme.

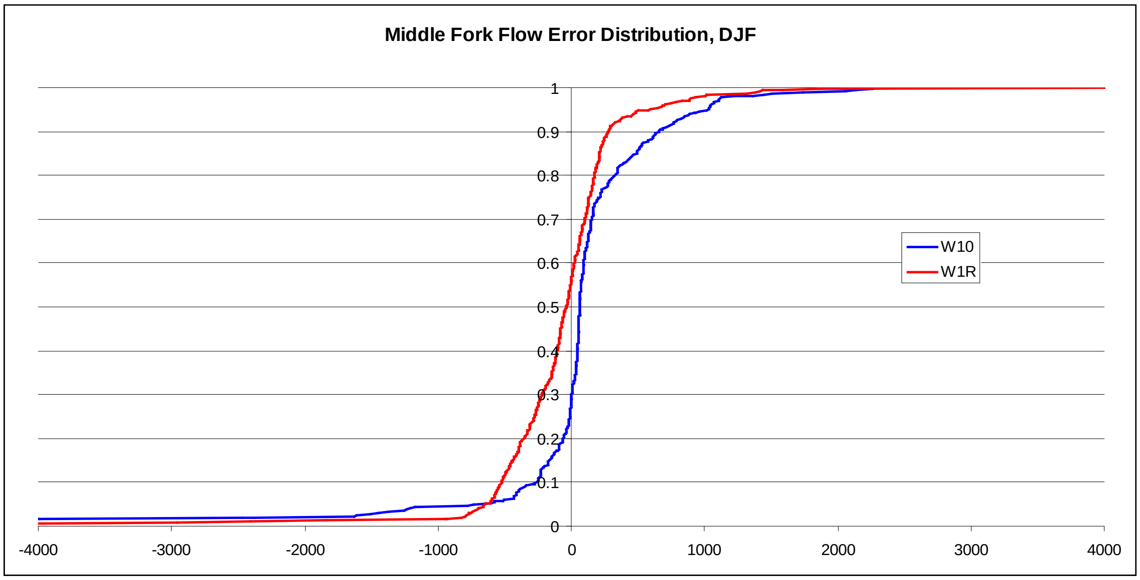


Figure 6: Error Distribution Comparison: Middle Fork, December to February

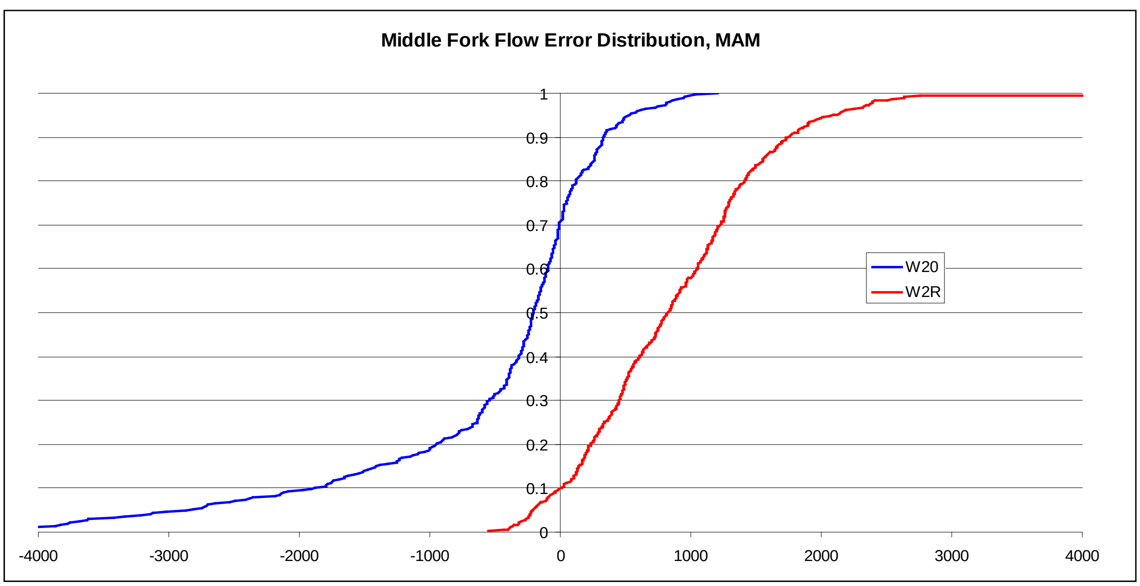


Figure 7: Error Distribution Comparison: Middle Fork, March to May

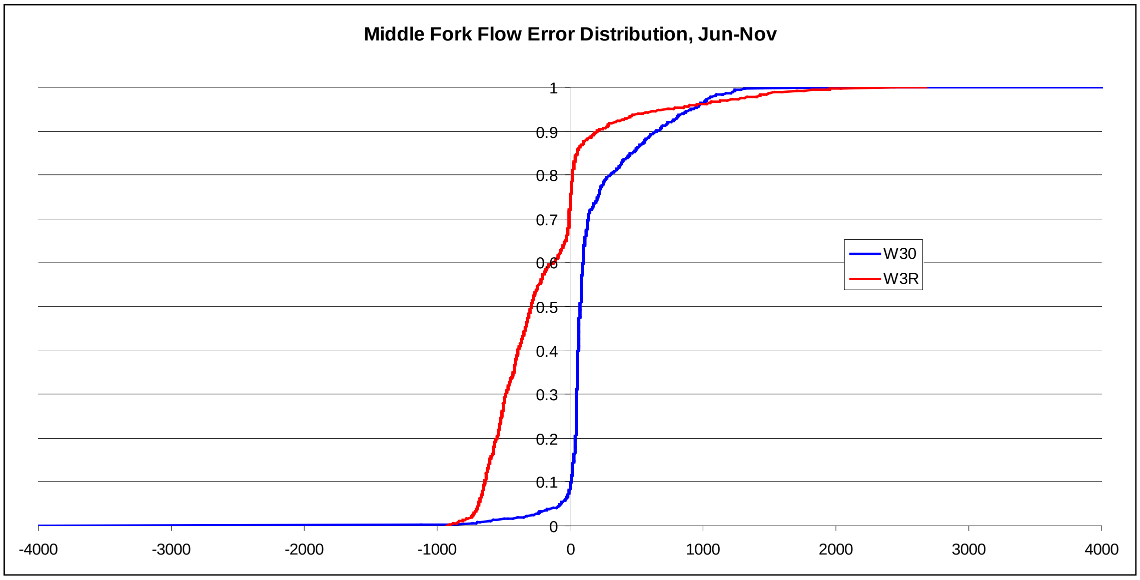


Figure 8: Error Distribution Comparison: Middle Fork, June to November

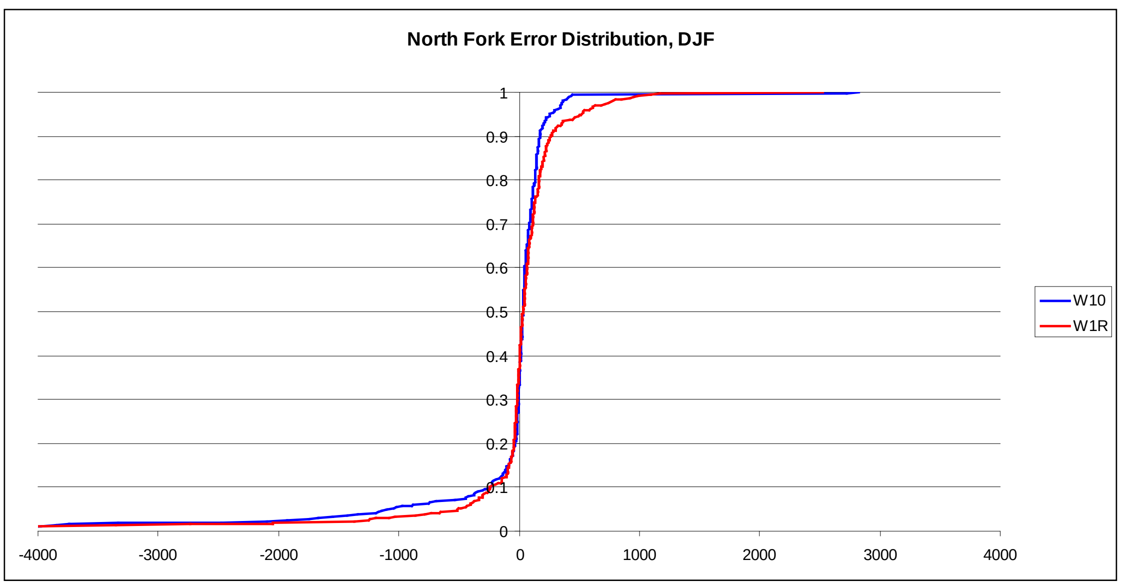


Figure 9: Error Distribution Comparison: North Fork, December to February

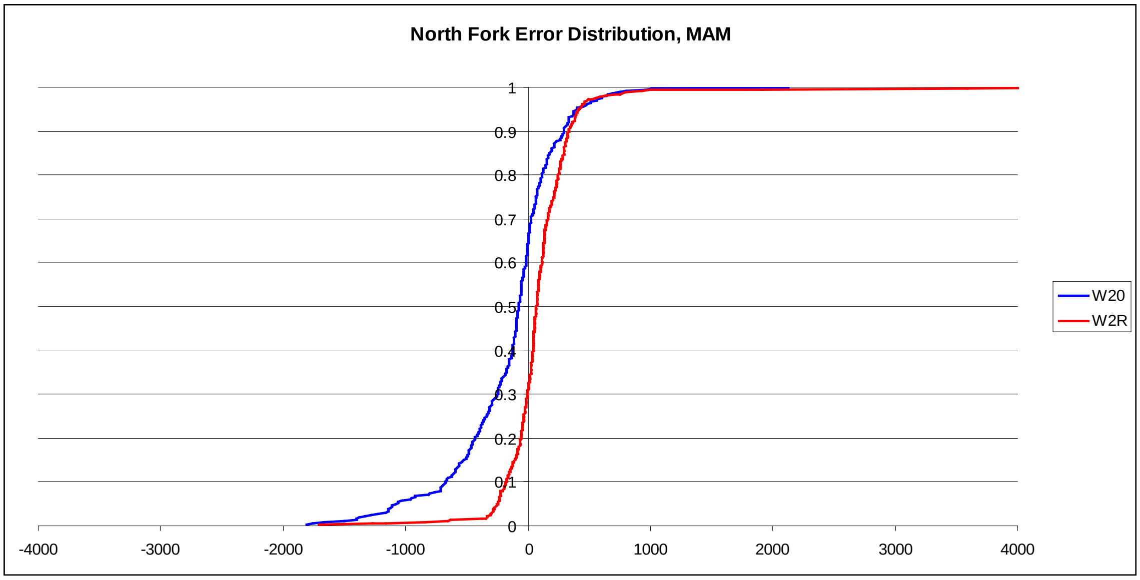


Figure 10: Error Distribution Comparison: North Fork, March to May

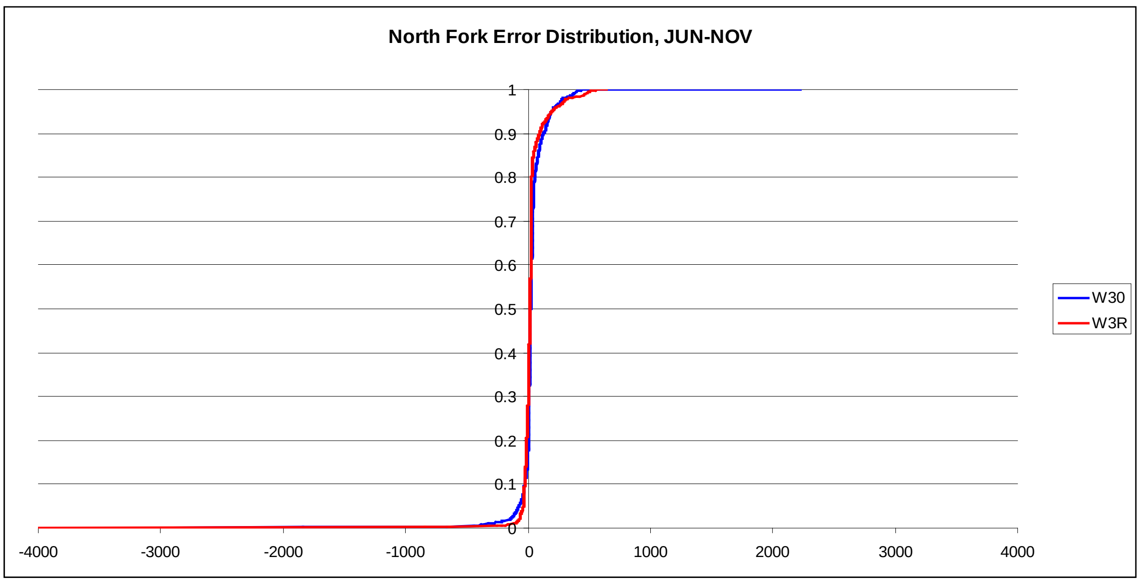


Figure 11: Error Distribution Comparison: North Fork, June to November

Lastly, the South Fork watershed presents a different but potentially common challenge for the above regulation testing approach, as no early unregulated flow period observations are available. This implies that the background error distributions (blue frequency curves in the previous figures) cannot be established, and the regulation type cannot be detected in the manner described earlier.

If unregulated period "natural" outflows cannot be reconstructed, the previous assessments can approximately be carried out by assuming that the unregulated error distributions are unbiased and centered around zero. A careful observation of the North Fork figures shows that this assumption is indeed valid, if the process by which the natural flows are generated is unbiased. Thus, for these cases, a statistically significant distributional shift could be ascertained by testing the hypothesis that the error mean of the regulated period is significantly different than zero. The conclusions on the regulation type are the same as before depending on the statistical significance of both wet and dry period shifts and their direction. For example, Figures 12, 13, and 14 depict these error distributions for the winter, spring, and summer South Fork watershed seasons. All seasons exhibit statistically significant shifts, with the spring shift being positive and the summer shift being negative. In keeping with the modeling framework logic, one would have to conclude that storage regulation is indeed taking place from spring to summer in South Fork and potentially other types of flow modifications occurring in the winter.

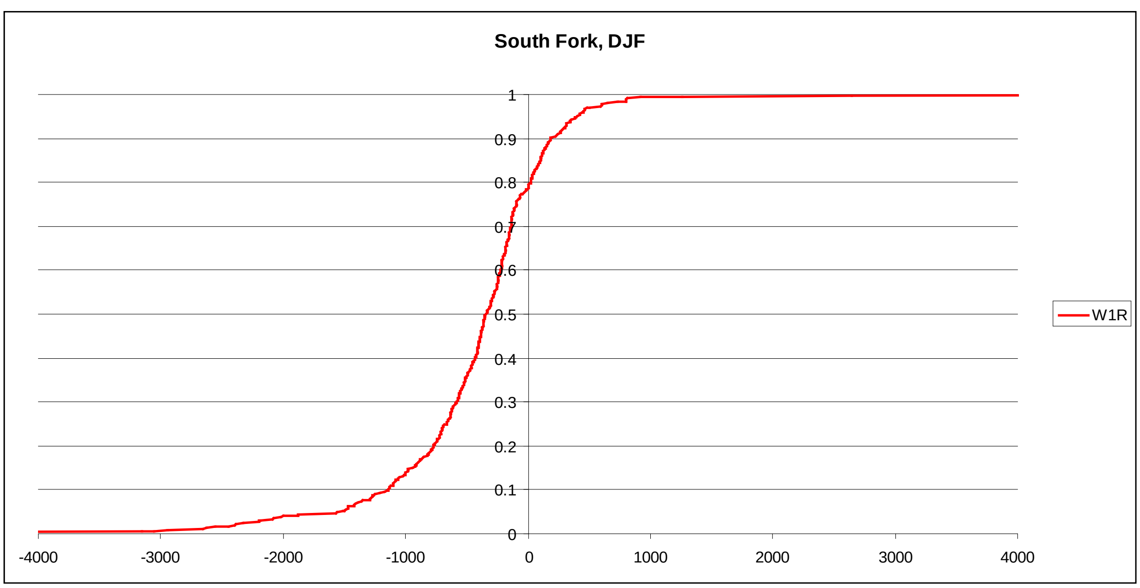


Figure 12: Error Distribution: South Fork, December to February

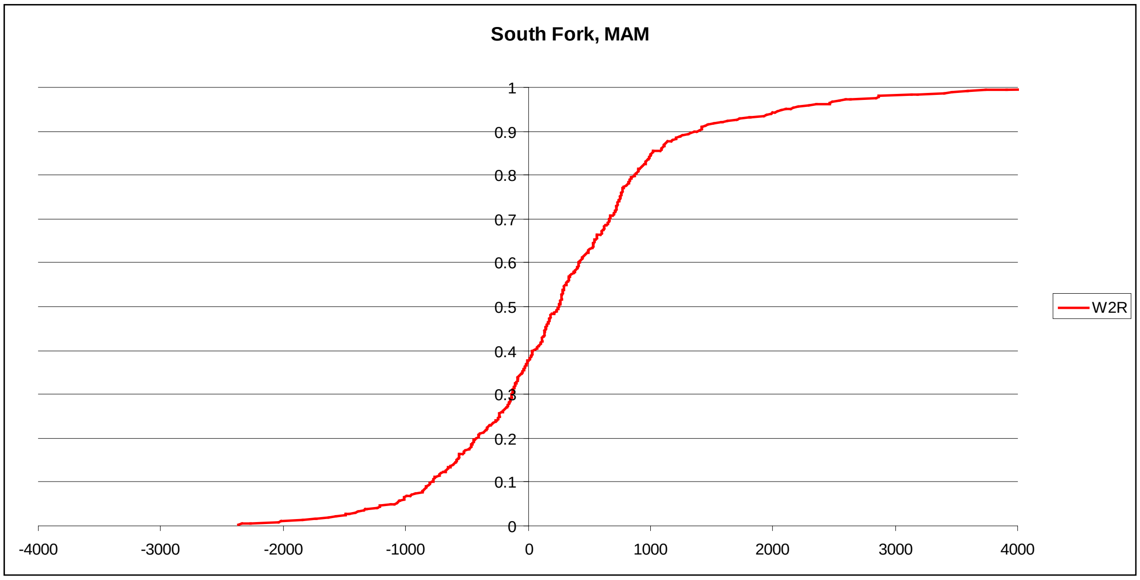


Figure 13: Error Distribution: South Fork, March to May

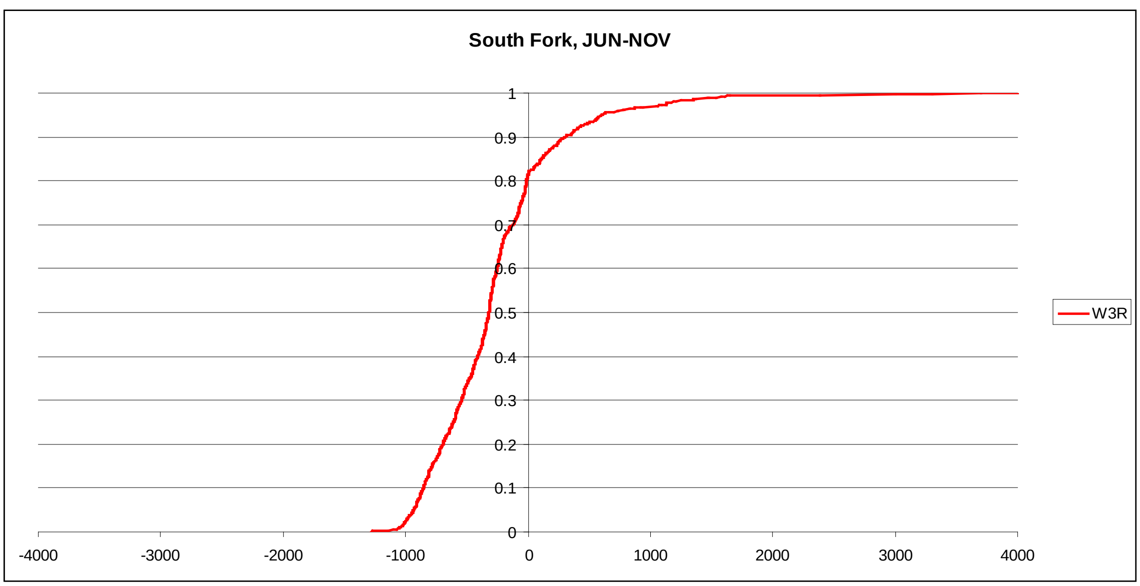


Figure 14: Error Distribution: South Fork, June to November

3. Storage and Water Transfer Effects

The need for forecast modification comes about because of the cumulative flow modification that distributed water uses (within a watershed) exert on watershed outflow. Typical water uses may include water supply for irrigation, domestic, or industrial use; small hydro-plant operation; and low flow augmentation for environmental and ecosystem sustainability. Some of the uses are consumptive, having direct impact on the *quantity* of available water, while others modify the *timing* of the natural flows. Both use types are usually enabled by small but potentially many storage facilities distributed across the watershed.

Several aspects would have to be considered in the development of physically realistic models aiming to explain watershed flow modification. These include:

(i) What is the magnitude and timing of flow modification?

(ii) Are there significant water transfers into and/or out of the watershed?

(iii) Are there significant consumptive uses within the watershed?

(iv) What is the size of the effective aggregate storage, if any?

(v) What are the operational practices that govern storage filling and depletion?

(vi) Do these practices change with different hydrologic conditions?

To address these questions and work toward model identification and development, we begin with a water balance analysis. The available data on the American River watersheds include: (1) daily hydrologic model forecasts and (2) observed watershed outflows. The data time period covers 14 years from 1984 to 1997. Data prior to this period exist but are not included because they do not reflect current water use levels and would bias model development.

Figures 15, 16, and 17 depict the 1984-1997 climatologies of the daily flow forecast (blue line), observed flow (red line), and cumulative difference (green line) at the outlet of each watershed. The following observations can be noted:

(1) Flow modification (augmentation) at the Middle and South Fork watersheds occurs

between July 1 to November 15; Flow modification is not appreciable at the North Fork watershed;

(2) The total flow augmentation volume is approximately 109 thousand acre-feet (TAF) at the Middle Fork watershed and 125 TAF at the South Fork watershed; Namely, on an average year, these water volumes would need to exist in storage (distributed throughout the watershed) at the beginning of the dry season (July 1 to November 15) to support the observed flow augmentation;

(3) Storage filling (observed flow minus flow forecast) at the Middle Fork watershed approximately commences around March 1 and is completed by the end of May, while at the South Fork watershed it commences on April 1 and is completed by June 15; The timing of the storage filling period is identified by the requirement that average inflows exceed average outflows; Storage filling appears to follow a linear rule (uniform filling) for both watersheds;

(4) For the South Fork watershed, there is a 47 TAF storage gap between the water storage requirement computed over the dry season (July 1 to November 15) and the storage that can be accumulated during reservoir filling (April 1 through June 15). This gap implies that there is significant water transfer into the South Fork to support water augmentation during dry season.

The previous observations were based on water balance considerations but are consistent with the water use facilities and practices in the case study watersheds (including the existence of storage facilities and water transfers) and provide a good basis for model formulation and calibration.

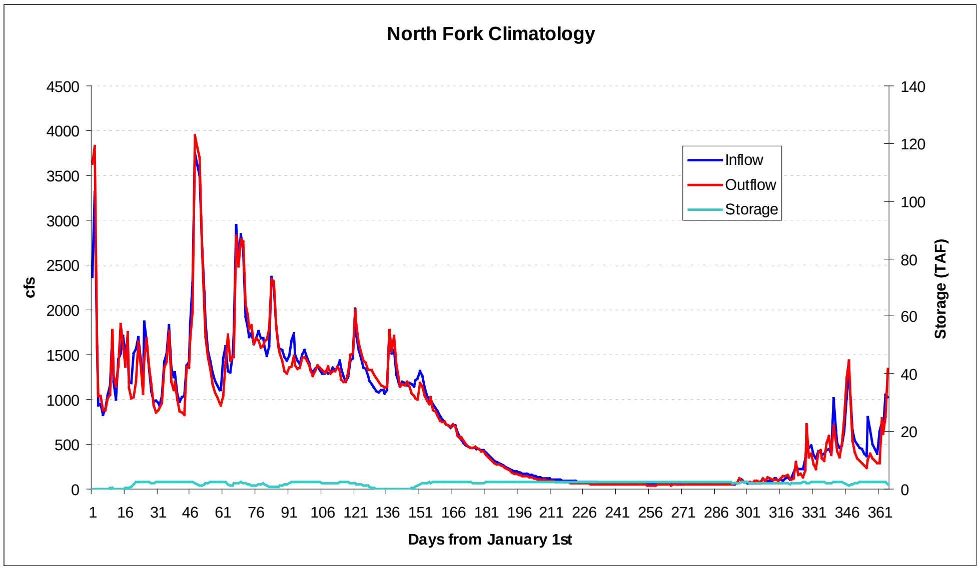


Figure 15: Water Balance Analysis for North Fork

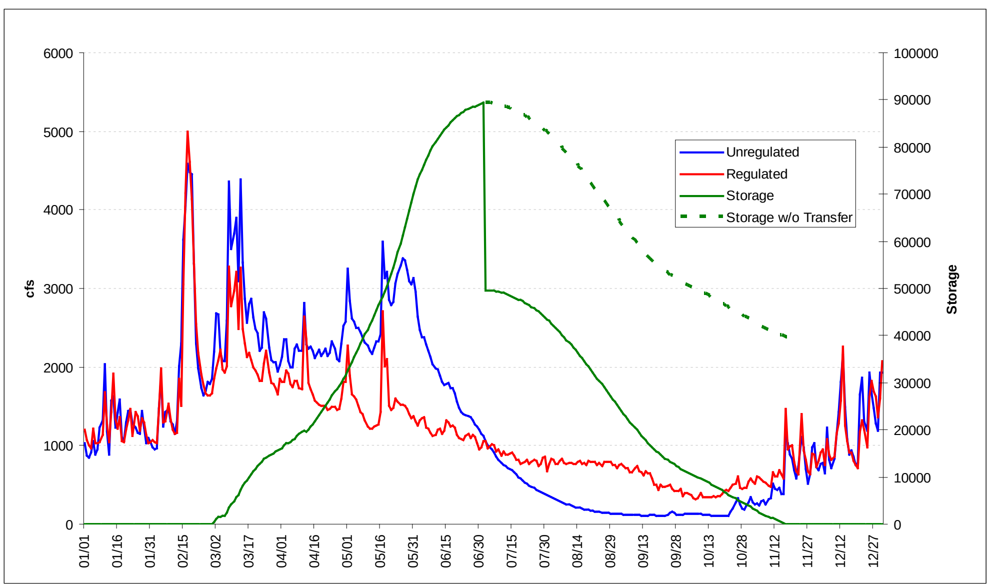


Figure 16: Water Balance Analysis for Middle Fork

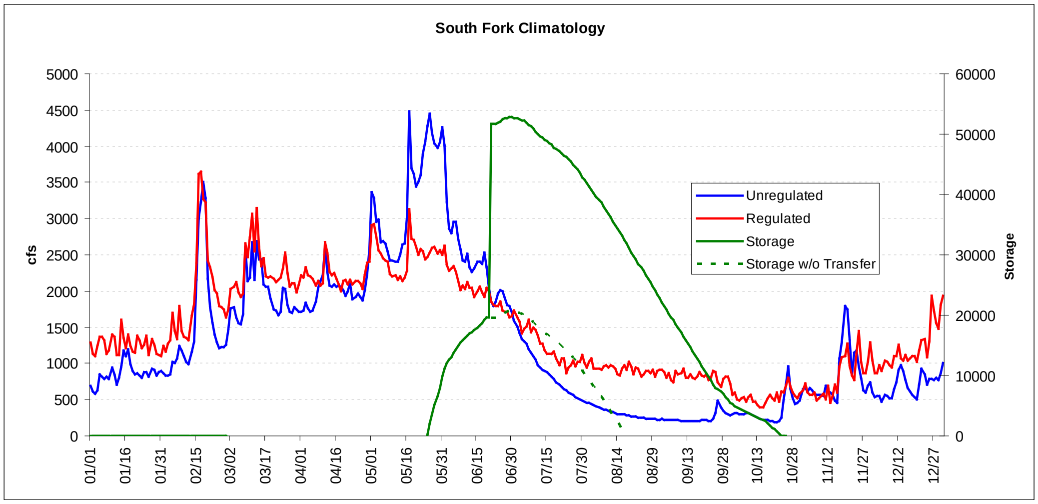


Figure 17: Water Balance Analysis for South Fork

3.1 Model Formulation

Based on the results of the preliminary water balance investigation, the observed flow augmentation can be modeled (explained) by considering an aggregate storage reservoir with the following features:

Reservoir storage water balance relationship:

$$S(k+1)=S(k) + I(k) - u(k) - L(k) - T(k) - D(k), \quad k = 0, 1, \ldots, N \qquad (1)$$

where k is the time increment used in the description of the system (day), S(k) is the reservoir storage at the beginning of time step k (with S(0), the initial storage value, considered known), I(k) is unimpaired watershed inflow during time step k, u(k) is reservoir outflow, L(k) represents storage losses due to evaporation and leakage, T(k) represents possible inter-basin transfers into or out of the watershed, D(k) represents consumptive water uses in the watershed, and N is the end of the simulation horizon.

Storage capacity constraint:

$$S(k) \leq S^{max}, \quad k=1,2, \ldots, N \qquad (2)$$

where S^{max} is the storage capacity.

Thus, if storage reaches its capacity, reservoir release is given by

$$u(k) = I(k) - L(k) - T(k) - D(k). \quad (3)$$

Reservoir filling rule:

Reservoir filling can occur in various ways. However, the most common practice is to fill storage uniformly during late spring. This process is practically easy to implement and can be modeled by the water balance relationship presented earlier. During this period, reservoir release is determined by

$$u(k) = - S^*(k+1) + S(k) + I(k) - L(k) - T(k) - D(k), \; k = 0, 1, \ldots, N_f, \quad (4)$$

where the starred storage value is the storage target for day k of the reservoir filling period. If the filling period comprises N_f days, the storage target under a uniform filling rule can be obtained from:

$$S^*(k) = k \text{ x } (S^{max}/N_f) \quad (5)$$

Reservoir release rules:

Reservoir releases are implemented to meet specific types of demands and, for this reason, can assume many different forms. For example, hydropower releases responding to week day peak power demands are different from environmental and ecological releases which do not differentiate weekdays from weekends. In the case of the American River watersheds various types of release forms can be observed in different years, indicating that release practices change from year to year or evolve as water use priorities change.

Two release rule forms that can accommodate several possibilities are as follows:

(1) Release is determined based on existing storage and flow forecasts over the dry season to support flow augmentation as best as possible (water balancing rule);

(2) Release is obtained as a function of aggregate storage through

$$u(k) = F[S(k)] \quad (6)$$

where F is a general nonlinear function to be determined (feedback release rule).

3.2 Model Calibration

The previous model requires the estimation/calibration of the following elements:

(i) Flow modification/augmentation period;

(ii) Aggregate storage requirement, S^{max};

(iii) Reservoir filling period and associated filling rule;

(iv) T(k), D(k), and L(k) for all applicable time periods k;

(v) Specification of the reservoir release rule.

Good initial estimates of the first three model elements [*(i), (ii), and (iii)*] can be readily obtained through water balance considerations as illustrated above. The sum of the water transfer and consumptive use terms [T(k) + D(k)] can also be estimated from such considerations, or directly from existing watershed records. In addition, it is important to determine the way in which water transfers, if any, are implemented. Specifically, one would need to determine if water transfers have a maximum limit, and whether they occur always or on an as-needed basis. Both of these aspects can be quantified from watershed information or from water balance investigations applied in different years. The assumptions used in the South Fork case study are that *(1)* there is a maximum water amount of 47 TAF that can be transferred into the watershed to augment the dry season flow and *(2)* this amount is added to the storage as needed at the beginning of the dry season.

Reservoir evaporation losses depend on the evaporation rate as well as reservoir water surface area. Evaporation estimates can readily be derived by data used to model evapotranspiration as part of the hydrologic forecasts. However, reservoir water surface area is a distributed quantity and cannot be readily estimated. Similarly, water balance considerations have the disadvantage that outflow observations already incorporate the effect of evaporation loss, thus making it indistinguishable (un-identifiable) from consumptive uses and/or water transfers. Thus, unless watershed storage is concentrated in a few facilities the surface area of which can be readily estimated, reservoir evaporation losses may best be considered as part of other water balance terms.

Reservoir release rules during the filling period have already been discussed assuming uniform storage filling. Comparisons of the model generated and observed flows during this period may be used to refine these rules for better correspondence. Reservoir release rules during the flow augmentation season are discussed next.

An intuitive form of the first release rule that seeks to balance available water supply with demand is as follows: At any time k of the dry season,

(1) Estimate the remaining aggregate reservoir storage, S(k), by the water balance relationship applied up to time k;

(2) Estimate the applicable consumptive uses and water transfers throughout the dry season and subtract or add their net sum to the estimated storage;

(3) Generate flow forecasts for the duration of the dry season; Determine an unimpaired flow volume expected to materialize. This quantity may correspond to the mean expected dry season unimpaired flow or an appropriate forecast percentile, depending on the operator's risk attitude. The most likely choice for this quantity is the mean flow forecast over the dry season.

(4) Compute and apply the daily release that can be sustained throughout the dry season from

Daily release u(k) = [S(k) + Dry season flow forecast – T(k) – D(k)] / (Number of remaining dry season days).

(5) Use the actually observed flow in day k, apply the water balance relationship, estimate the value of the storage S(k+1), and repeat Steps (2) to (5) throughout the dry season.

This release rule can be implemented in different ways depending on (a) the risk attitude of the reservoir operators and (b) the specific water outflow requirements over the dry season. With respect to risk attitude, different release versions can be formulated that place higher, equal, or lower preference on early versus late dry season releases. With respect to specific outflow requirements, one can establish weekly releases with specific daily patterns that reflect power generation or other water allocation targets. Due to the absence of any specific information regarding the above, this release has been implemented assuming that daily releases have equal weights during the dry season (neutral risk attitude) without any specific daily allocation preferences. Furthermore, the flow forecast mean over all of the 1984-1997 dry seasons is used as the statistic referred to in Step (3).

The second release rule is more applicable when no specific outflow patterns appear to exist, and the model identification process can mostly rely on unimpaired flow forecasts and flow observations.

The problem of determining the feedback release function $u(k) = F[S(k)]$ can be solved using a traditionally adopted approach whereby one assumes a particular functional form (e.g., piecewise linear) and calibrates its parameters to match model-generated and observed outflow sequences as best as possible. Commonly used matching criteria include bias and minimum mean square error statistics.

A simple yet practical procedure to derive the release function is implemented and tested here. In this procedure, the release function is derived using the climatology data. The procedure is summarized below:

1) Compute the storage sequence using the water balance equation with the climatology inflow and release data;

2) Extract the computed storage and release data for the dry period;

3) Sort the extracted storage and release data separately from minimum to maximum;

4) Plot the sorted release against the sorted storage data. The resulting curve is the release function.

Figures 18 and 19 showed the release functions derived using the above procedure for the Middle and South Fork, respectively. This procedure has been implemented in this project and can be used as an alternative release option. While, this approach provides a way to link the rate of release versus current storage, it cannot represent the day to day release fluctuations characteristic of hydropower operations. As mentioned earlier, in such cases, the first release rule is more suitable.

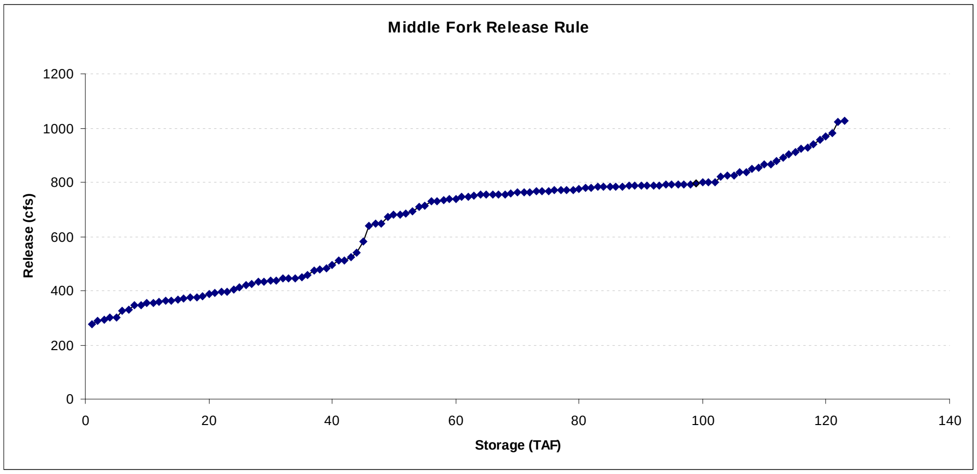


Figure 18: Feedback Release Rule for Middle Fork

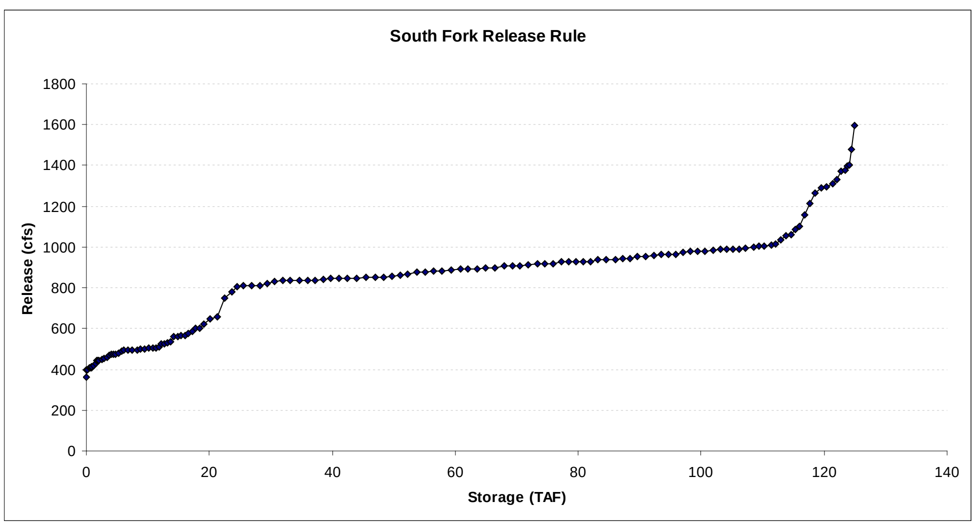


Figure 19: Feedback Release Rule for South Fork

3.3 Model Testing and Assessment

The previous model was implemented for the Middle and South Fork watersheds over the 1984-1997 time period with both release rule options. The unimpaired flow forecasts and the observed outflows constitute the data base, and the results consist of the model generated outflow and aggregate reservoir storage sequences. Model performance is assessed by computing the bias and mean square error statistics for the reservoir filling and the flow augmentation periods. For comparison purposes, the corresponding statistics of the unimpaired flow forecast sequence are also reported. These summary statistics are shown in Table 4, while typical sequences for the two case studies are shown on Figures 20, 21, 22, and 23. The main observations are summarized next.

Table 4: Regulation Model Summary Assessment Statistics

Middle Fork	Model	Filling Season (March 1 to May 31)		Dry Season (July 1 to November 15)	
		Bias	**M.S.E**	**Bias**	**M.S.E**
Rule 1	Unimpaired Forecasts	888.96	1276.65	-435.71	573.21
	Regulation Model	298.20	955.02	-85.80	350.51
Rule 2	Unimpaired Forecasts	888.96	1276.65	-435.71	573.21
	Regulation Model	298.20	955.02	-63.29	358.98

South Fork	Model	Filling Season (April 1 to June 15)		Dry Season (July 1 to November 15)	
		Bias	**M.S.E**	**Bias**	**M.S.E**
Rule 1	Unimpaired Forecasts	510.28	1347.18	-480.48	670.49
	Regulation Model	-94.88	1208.13	-80.31	472.46
Rule 2	Unimpared Forecasts	510.28	1347.18	-480.48	670.49
	Regulation Model	-94.88	1208.13	-107.24	436.79

(1) The regulation model clearly improves the performance of the unimpaired flow forecasts during both the reservoir filling as well as the dry seasons. This is evident by the significant reductions realized by both assessment measures—bias and mean square error.

(2) The regulation model performs well under both release rules, with the first release rule (seeking to balance supply and demand) yielding somewhat better bias results in the South

Fork, and the second rule (using the feedback release function) performing somewhat better in the Middle Fork. The performance with respect to the mean square error is reversed but comparable.

(3) The performance of the regulated model in the Middle and South Fork watersheds is also comparable to the performance of the hydrologic model in the North Fork. Specifically, the bias and mean square error statistics of the unimpaired forecasts for the North Fork are 19 and 560. This implies that the regulation model is able to filter out the upstream regulation discrepancies and re-instate the hydrologic model performance as realized in pristine watersheds.

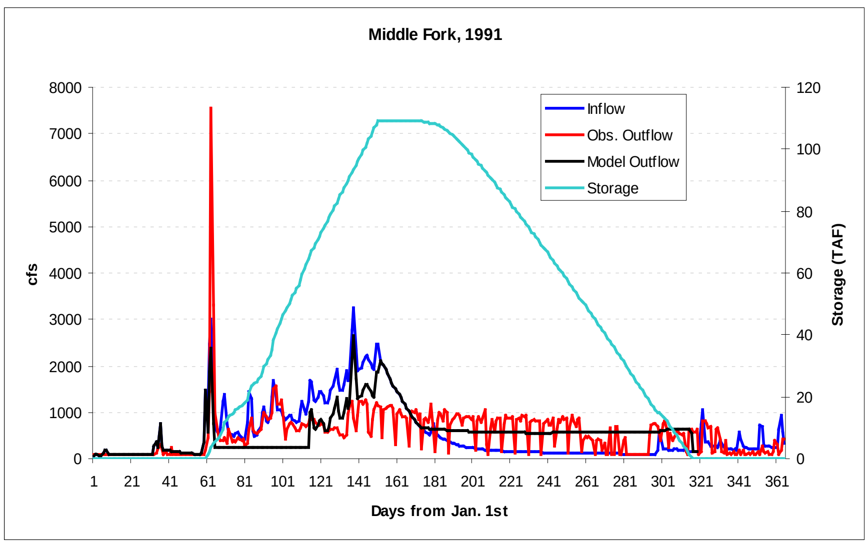


Figure 20: Middle Fork Comparison; Water Balancing Rule 1

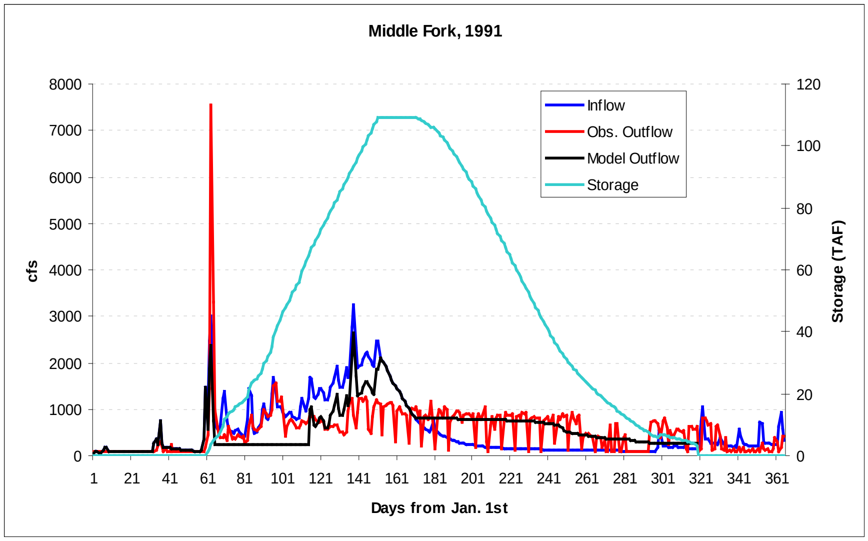


Figure 21: Middle Fork Comparison; Feedback Release Rule 2

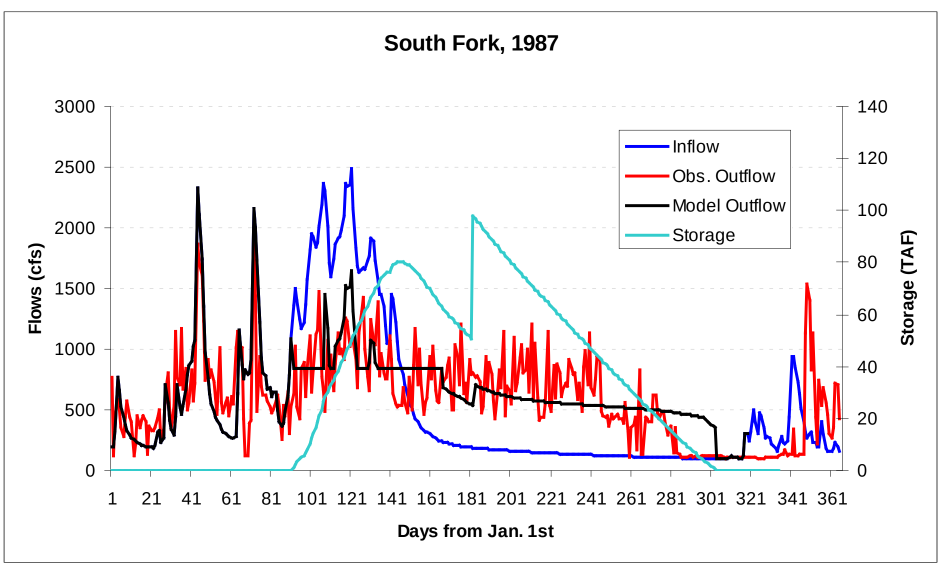


Figure 22: South Fork Comparison; Water Balancing Rule 1

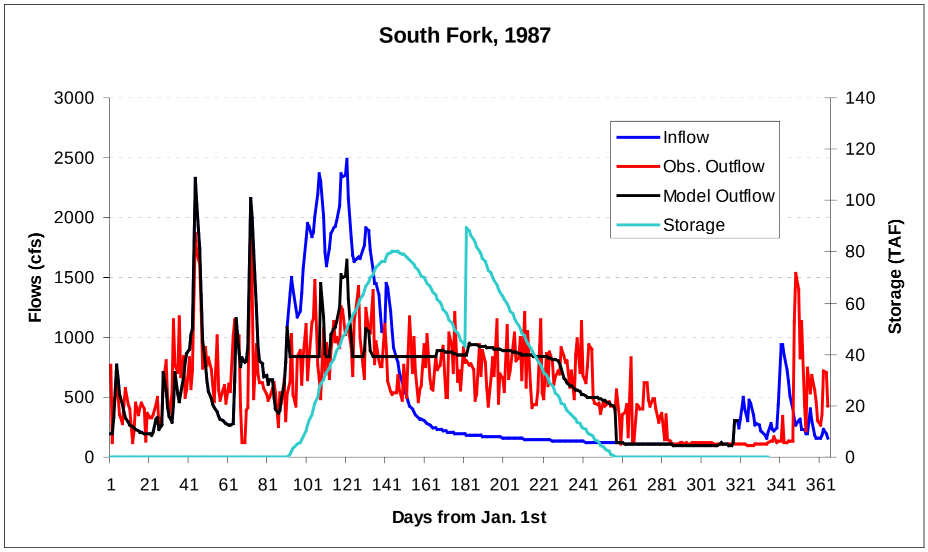


Figure 23: South Fork Comparison; Feedback Release Rule 2

This work demonstrates that upstream regulation processes can effectively be modeled by physically based models that are flexible with respect to available data and can fairly easily be integrated with operational forecasting systems. The models can be calibrated using water balance considerations combined with available watershed information, if any, on storage facilities, water transfers, and consumptive use patterns. The models were shown to produce significant improvements over unimpaired flow forecasts, filtering out the effects of upstream regulation and restoring forecast performance to that for pristine watersheds. Model features include the storage filling and release periods, storage filling rules, storage and water transfer targets, and release policies. However, a key aspect of the watershed storage is the characterization of its uncertainty at each stage of the filling and release periods which, in turn, conditions the outflow releases during the dry period.

To this end, investigations were undertaken to relate the watershed storage to key hydrologic variables such as precipitation (average and peak values) during the storage filling period, winter snow accumulation, and hydrologic model soil moisture states. No significant relationships were found with any of the above variables that would enable the estimation of the storage probability distribution at the end of the filling period. This exploratory data analysis showed that watershed storage is closely dependent on the antecedent unimpaired watershed outflows, which are available up to the time of the forecast through continuous hydrologic model simulations that use actual precipitation and temperature observations. These "natural" inflow traces can be presented in the form of a trace ensemble accounting for precipitation and temperature as well as model parameter errors.

Each historical trace ensemble combined with the observed outflow sequence gives rise to a storage estimate through a water balance computation from the beginning of the filling period to the forecast start date. However, this storage estimate includes both: *(a)* storage to be used in flow augmentation for the same watershed, as well as *(b)* possible water exports/imports to and from other watersheds. This partition can be determined by a historical analog approach considering similar historical storage circumstances. The uncertainty of this partition can be quantified by including a number of close neighbors to the estimated storage value. Once this partition is determined, forecasts can be issued using each and every watershed storage distribution value, partition scenario, and dry period release pattern corresponding to the selected historical analog year of each scenario. In this manner, ensemble forecasts include all relevant uncertainties.

Figures 24, 25, and 26 show results from a sample forecast run for the Middle Fork watershed with a forecast start date of July 1, 1987.

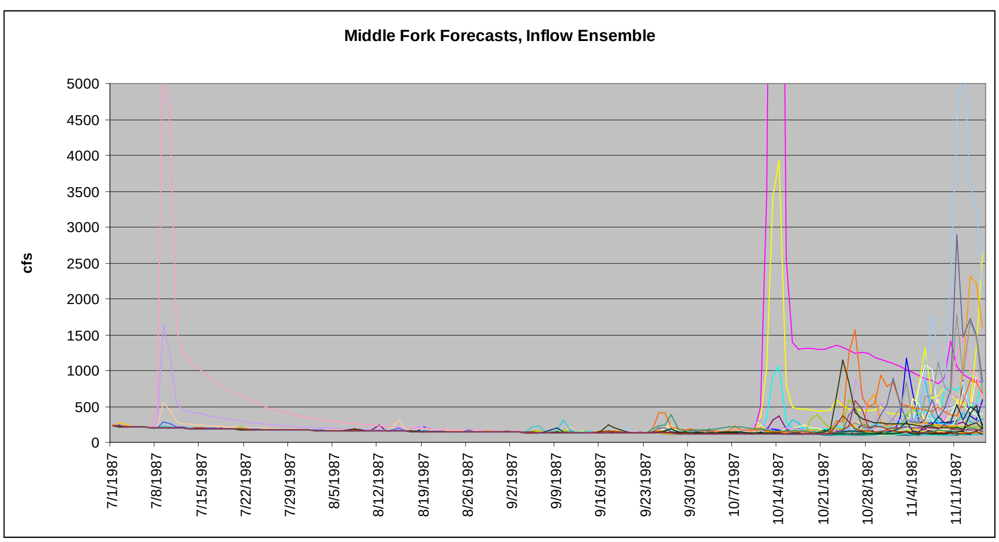


Figure 24: Natural Outflow Forecast Ensemble, Middle Fork, July 1 through November 14, 1987.

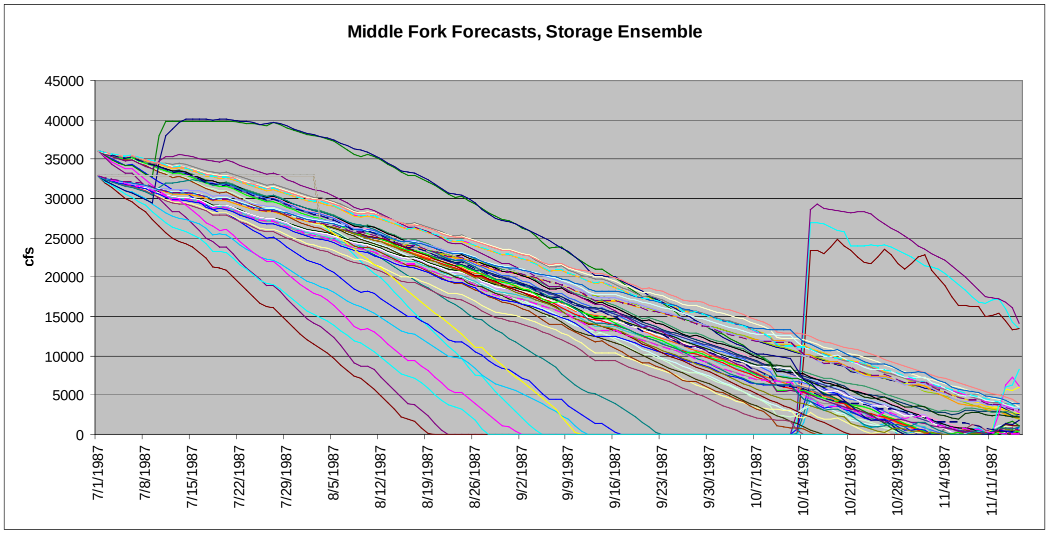


Figure 25: Storage Ensemble Forecasts for two Historical Analog Partition Scenarios, Middle Fork; July 1 through November 14, 1987.

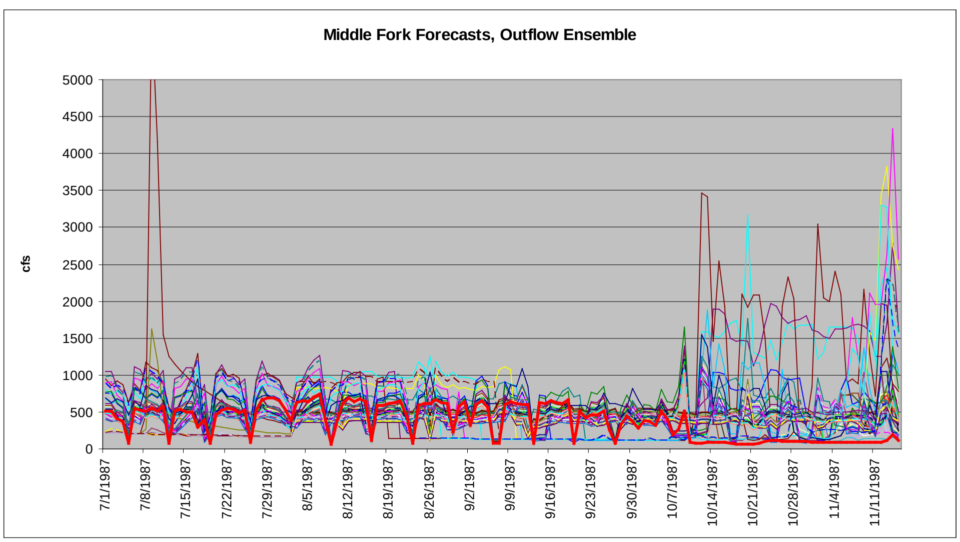


Figure 26: Flow Forecast Ensemble, Middle Fork; July 1 through November 14, 1987.

Page Left Blank

4. BIAS ADJUSTMENT

The application of Regression and Neural Network models for bias adjustment has been extensively investigated and reported in *Georgakakos et al. (2009).* The approach followed was to develop three different models for high, medium, and low flows regardless of when such flows occur in the course of the year. Subsequent investigations of dynamic seasonal water transfers through watershed storage and the modeling framework introduced earlier illustrated the need to distinguish forecast correction procedures by seasons rather than by flow magnitude.

With this motivation, Regression and NN models between unimpaired and observed flows were re-derived in this project period for each quarter of the year (DJF, MAM, JJA, and SON) and all three case study watersheds.

The following linear model was used in the regression:

$$O(k) = A_1 I(k) + A_2 I(k-1) + A_3 \;,$$

where k is the time index, O is the output variable (observed flow), I is the input variable (unimpaired flow forecast from the hydrologic model), and A_i , i=1, 2, 3, are regression coefficients.

The results are summarized in Tables 5, 6, and 7. Each table includes the regression model coefficients (A's), historical data period used to derive the model, and error statistics (correlation coefficients between predicted and observed outflows, bias, and error standard deviations). Similar error statistics of the "natural" versus observed outflows were also computed and included in the same table for comparison purpose. Four seasonal models were generated for each basin. As shown, the biases are removed from all regression models. In addition, the error standard deviations are reduced as well for all basins. It is noted that the regression models are activated only during the wet seasons while upper stream storage regulations do not exist. For the case study basins, the regression models are used for all seasons for the North Pork, used from December to March only for the Middle and South Fork.

The NN models were also developed and tested using the same input data as in the regression models. All NN models have a structure of two input nodes and one hidden layer with two hidden nodes, a total of nine parameters per model. The NN results are very similar and are not reported herein. This performance similarity between the Regression and the NN models raises questions on the NN added value in view of their many more parameters and their involved parameter estimation requirements.

It is further noted that the comparison between Regression and NN reported in *Georgakakos et al. (2009)* pertains to *one-step*-ahead prediction applications. However, the purpose of the forecast correction procedures is to be used within multi-lead forecast applications. While such model use would not impact models that only include unimpaired outflows as explanatory variables, it may greatly hinder the utility of models depending on past observed outflow values. This is because in multi-lead forecast applications future "observed" outflow values are not available and must be substituted by model generated values (output feedback). In such cases, one has to ensure that the model equations are mathematically stable so that model errors do not grow exponentially due to output feedback. Unfortunately, the nonlinear nature of the NN models is likely to cause this problem to occur, negating their value in multi-lead forecasting. On the other hand, preliminary multi-lead model runs with the regression models show that they are stable. For this reason, the regression based models are used in the final assessment.

The results reported here are based on the seasonal regression models. Considering large flow variations within the wet seasons, a model with a monthly resolution may be more appropriate and further reduce the forecasting errors.

Table 5: North Fork Regression Models and Error Statistics

Season	Data Used		Inflows Used	Outflows Used	Regression Model Coefficients			Regression Model Error Stats			Original ESP Error Stats		
					A1	A2	A3	Error Mean	Error Std	Corr Coef.	Error Mean	Error Std	Corr Coef.
1	1/1/1960	12/31/1996	2	0	0.9831	0.1476	-97.8509	0.00	680.08	0.97	-54.14	758.19	0.97
2	1/1/1960	12/31/1996	2	0	0.8463	0.1540	79.9238	0.00	531.07	0.93	-79.91	549.64	0.92
3	1/1/1960	12/31/1996	2	0	0.9865	0.0041	-43.1171	0.00	179.12	0.94	46.62	179.19	0.94
4	1/1/1960	12/31/1996	2	0	1.2837	-0.1722	-12.4276	0.00	223.71	0.96	-8.52	252.89	0.95

Table 6: Middle Fork Regression Models and Error Statistics

Season	Data Used		Inflows Used	Outflow Used	Regression Model Coefficients			Regression Model Error Stats			Original ESP Error Stats		
					A1	A2	A3	Error Mean	Error Std	Corr Coef.	Error Mean	Error Std	Corr Coef.
1	1/1/1982	12/31/1996	2	0	0.9467	0.0714	422.5355	0.00	924.22	0.95	10.29	1028.90	0.94
2	1/1/1982	12/31/1996	2	0	0.5290	0.1400	702.6582	0.00	1166.30	0.82	763.02	1324.28	0.82
3	1/1/1982	12/31/1996	2	0	0.3575	0.3214	565.6384	0.00	363.81	0.82	-52.84	790.84	0.82
4	1/1/1982	12/31/1996	2	0	0.3475	0.1221	503.5125	0.00	356.71	0.86	-290.67	410.98	0.85

Table 7: South Fork Regression Models and Error Statistics

Season	Data Used		Inflows Used	Outflows Used	Regression Model Coefficients			Regression Model Error Stats			Original ESP Error Stats		
					A1	A2	A3	Error Mean	Error Std	Corr Coef.	Error Mean	Error Std	Corr Coef.
1	1/1/1982	12/31/1996	2	0	0.9467	0.0714	422.5355	0.00	903.12	0.90	-440.10	905.73	0.90
2	1/1/1982	12/31/1996	2	0	0.5290	0.1400	702.6582	0.00	1223.45	0.79	108.47	1490.05	0.78
3	1/1/1982	12/31/1996	2	0	0.3575	0.3214	565.6384	0.00	630.69	0.89	-203.02	850.74	0.88
4	1/1/1982	12/31/1996	2	0	0.3475	0.1221	503.5125	0.00	613.63	0.61	-256.23	848.44	0.60

Page Left Blank

5. Validation of Adjusted Ensemble Streamflow Predictions

Adjustment of Ensemble Streamflow Predictions (ESP) is described in earlier sections as a result of two main models: *(a)* an aggregate watershed storage model, aiming to explain and replicate the seasonal water balance deficits and surpluses observed in regulated watersheds; and *(b)* a regression model aiming to modify hydrologic model forecasts to better correspond with observed watershed outflows. For the retrospective validation analysis described in this section, both of these approaches are included in the integrated modeling framework and are evaluated when applied in combination to the unregulated ensembles.

The retrospective runs use data from the North, Middle, and South Fork sub-catchments of the American River. We expect that the retrospective analysis results are indicative of the general model performance to other areas, as no external information was used other than what would be known routinely during forecast operations and that which is reflected in routinely available historical data. Any differences in performance between the studies reported here and studies for other regions are anticipated to be the result of specific differences in the quality of the unregulated ensemble streamflow prediction (ESP) forecasts, the type of upstream regulation (whether it involves reservoir storage or not) and in the availability of historical data. Thus, in the outset, we suggest applications to other basins be undertaken in the future to further assess the forecast benefits as well as motivate additional enhancements.

5.1 Validation Methodology

To evaluate the performance of the upstream regulation models in an operational-like environment, the authors incorporated the upstream regulation framework into the Hydrologic Research Center stand-alone ensemble streamflow prediction (ESP) modeling system for Northern California (see system details in *HRC-GWRI 2007*) and applied this system to the North, Middle and South Fork data sets. The system was applied in a retrospective manner with the current year excluded always from the ensemble members generated with forecast preparation times in the current year. The period of record used was 1978-1997 inclusive, and forecasts were generated with forecast start times on each Monday of each year with a temporal resolution of 6 hours for the unregulated ensemble forecast flows and 24 hours for the regulated ensemble forecast flows for each of the three Forks of the American River in California. Evaluation of ensemble forecast skill is done using the USGS daily observed streamflow record at the downstream measurement point of each Fork (see Figure 5). The analysis involves measurement of performance with respect to climatological averages and the Ranked Probability Score (RPS) for the target variables. The RPS is computed for terciles of the distribution of the target variables, and, for a single ensemble forecast distribution, it is estimated from (e.g., *Wilks 1995*):

$$RPS = \sum_{m=1}^{3}\left[\sum_{j=1}^{m} y_j - \sum_{j=1}^{m} o_j\right]^2 \tag{7}$$

where there are three forecast categories ($m=1,2,3$) corresponding to the low, middle and high tercile of the forecast and observed distributions, y_j signifies the forecast probability computed from the ensemble forecasts for the jth forecast category and for a particular target variable (discussed below), and o_j signifies the corresponding observed frequency for the same category. The observed frequency is specified to be equal to 1 for the forecast category that contains the current observation of the target variable and it is equal to 0 for the other forecast categories. To evaluate a collection of n forecasts, each with a ranked probability score equal to RPS_k $(k=1,..,n)$, the average $ARPS$ is used:

$$ARPS = \frac{1}{n}\sum_{k=1}^{n} RPS_k \tag{8}$$

Lastly, to compare the adjusted ESP to the unadjusted ESP, the RPS skill score, S_{RPS}, is used:

$$S_{RPS} = 1 - \frac{ARPS_{AESP}}{ARPS_{ESP}} \tag{9}$$

where subscript AESP stands for adjusted ESP.

Target variables for this validation are the streamflow volume averages for September and October, weekly averages for these two months for the first and third week in each of the months, and daily flows for days of the week (Monday, Tuesday, etc.) for the previously mentioned two weeks in each of the two months. The analysis considers forecast start times on the first Monday of June, July and August of each year, thus making for lead times from 1 to 4 months. It is noted that September and October are months with significant regulation effects in this region.

5.2 Climatological Averages

The climatological monthly averages for September and October, and the weekly averages for the first and third weeks of September and October are shown in Figure 27 for the observed flows for all Forks of the American River. Figures 28, 29 and 30 show the corresponding averages computed from the adjusted ensemble streamflow forecasts with start date on the first Monday of June, July and August, respectively. Figures 31, 32 and 33 show corresponding results from the unadjusted ensemble streamflow forecasts. The main conclusion we draw from these results is that the forecasts adjusted for upstream regulation have on average raised the original ESP flows during these months and weeks for the regulated watersheds of the Middle and South Forks of the American River, toward a better agreement with the observations in spite the long lead times involved. The improvement is more pronounced for September, when the unimpaired flow for the Middle and South Fork is very low.

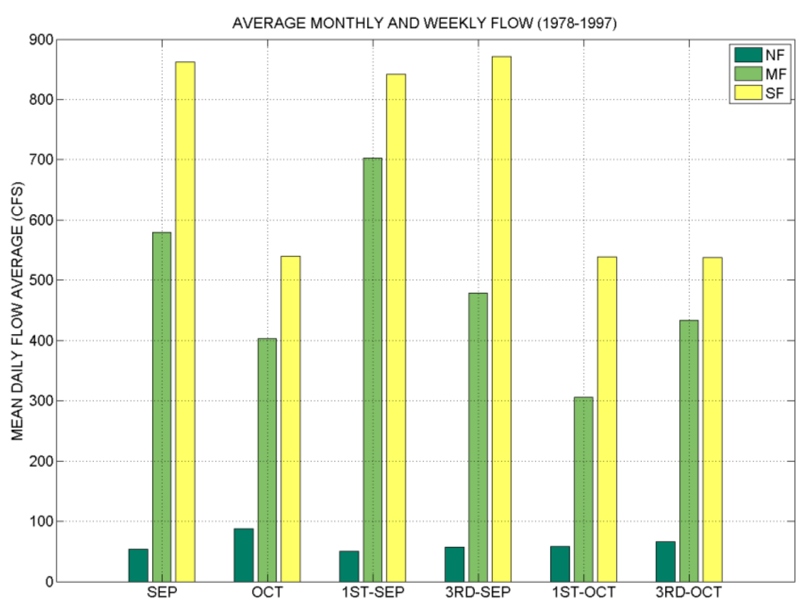


Figure 27: Average Observed Monthly and Weekly Flows in September and October.

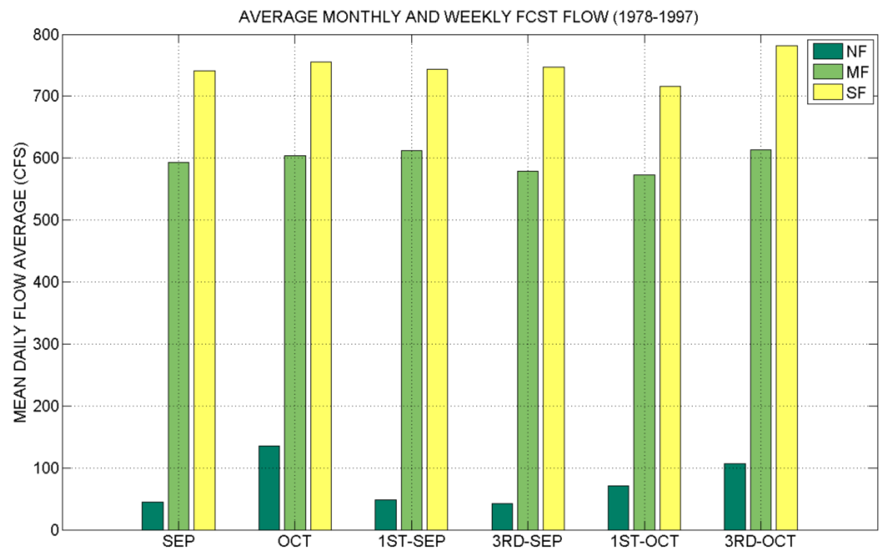


Figure 28: As in Figure 27 but from Adjusted ESP Forecasts with Start Date on the First Monday in June.

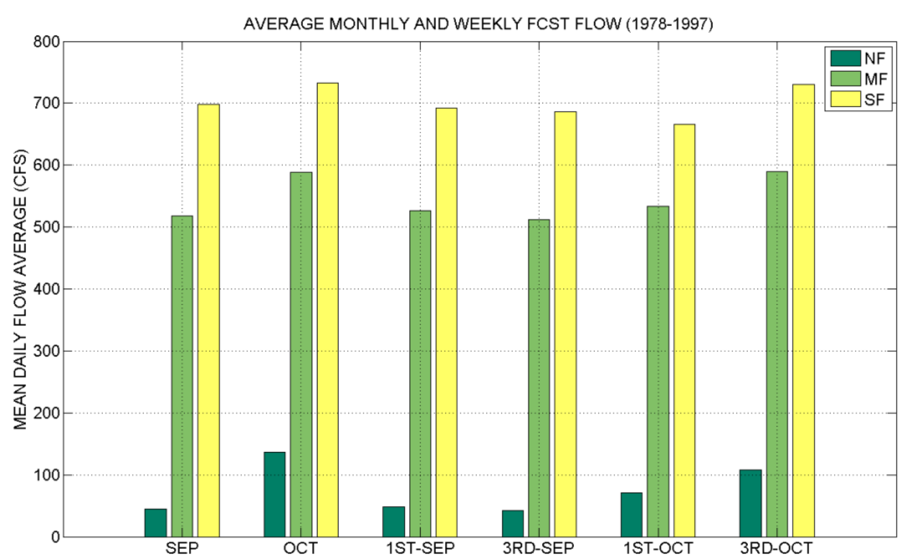


Figure 29: As in Figure 27 but from Adjusted ESP Forecasts with Start Date on the First Monday in July.

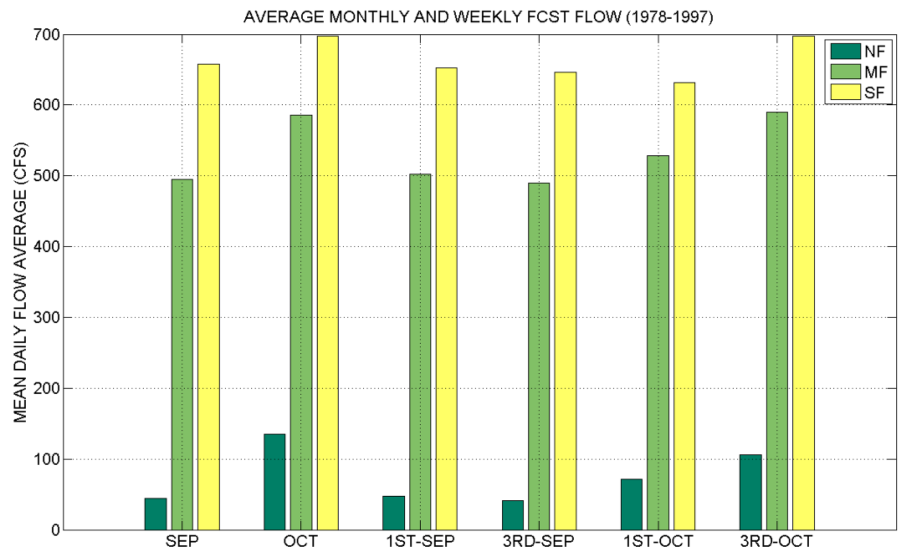


Figure 30: As in Figure 27 but from Adjusted ESP Forecasts with Start Date on the First Monday in August.

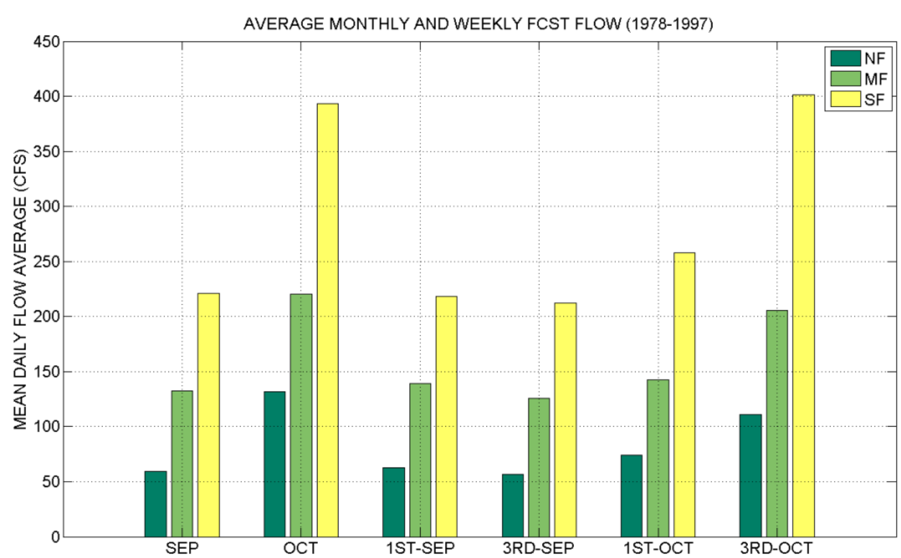


Figure 31: As in Figure 28 but for Unadjusted ESP forecasts.

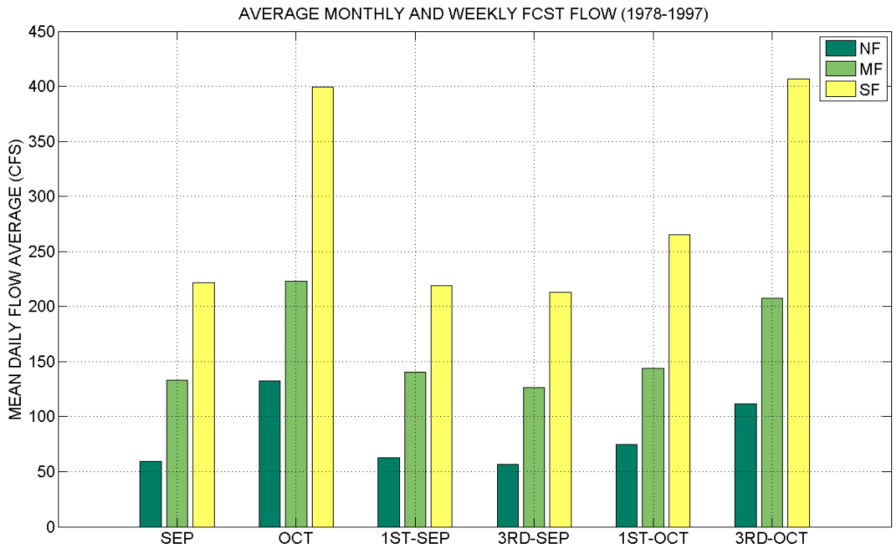


Figure 32: As in Figure 29 but for Unadjusted ESP Forecasts

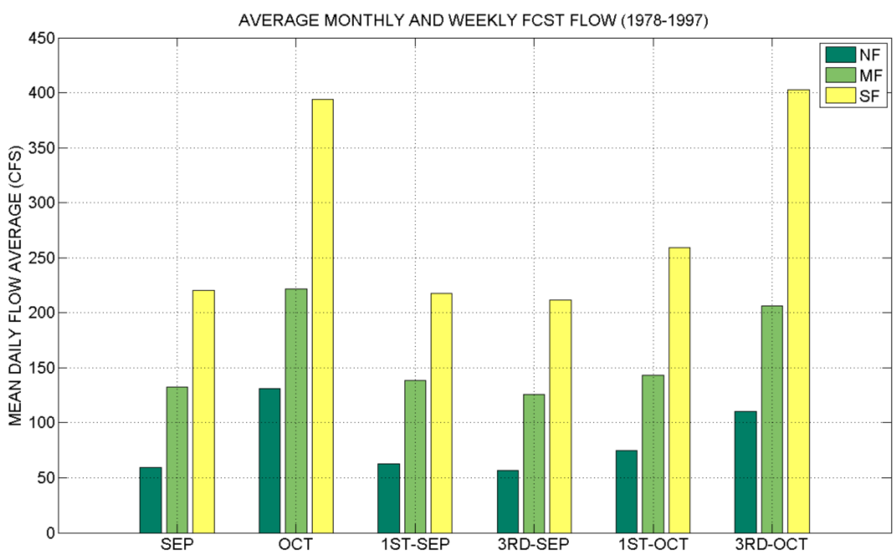


Figure 33: As in Figure 30 but for Unadjusted ESP Forecasts.

The climatological average of the day of the week mean daily flow (CFS) for the first week in September for the period 1978-1997 is shown in Figure 34 for all the Forks of the American River. The ensemble average equivalent statistics for the first week of September are shown in Figures 35, 36, and 37 for the adjusted ESP traces with forecast start time on the first Monday in June, July and August, respectively. Figures 38, 39 and 40 show the analogous statistics for the unadjusted ESP forecasts. It is apparent that the ensemble forecast statistics when the ESP traces are adjusted for estimated upstream regulation are significantly closer to the observed statistics than the unadjusted ESP averages for the Middle and South Forks, with the daily pattern of the week preserved and the magnitude of the flows being close to the observed flow magnitude. It is important to note that for the mean daily flows the improvement is even more pronounced than in the monthly and weekly averages.

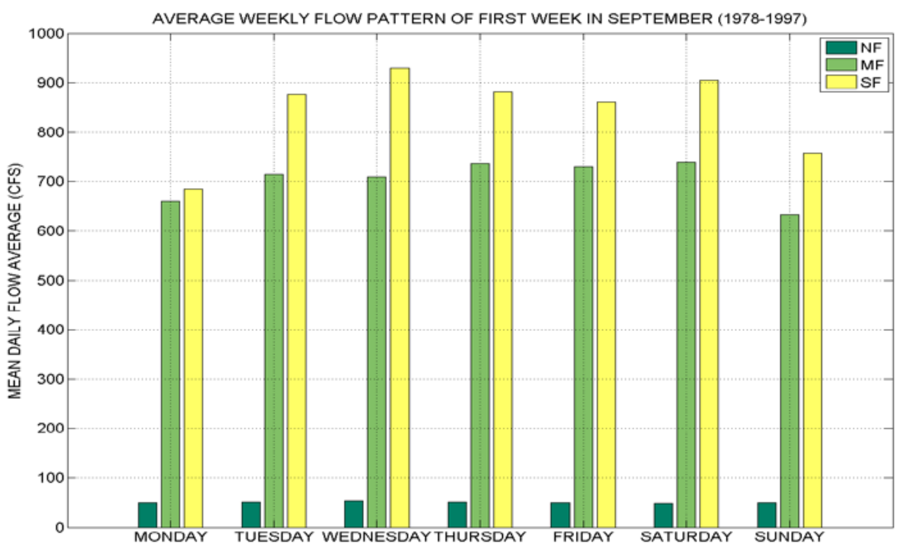


Figure 34: Climatological Average of Mean Daily Flow (cfs) for Each Day of the First Week in September for all Three Forks of the American River.

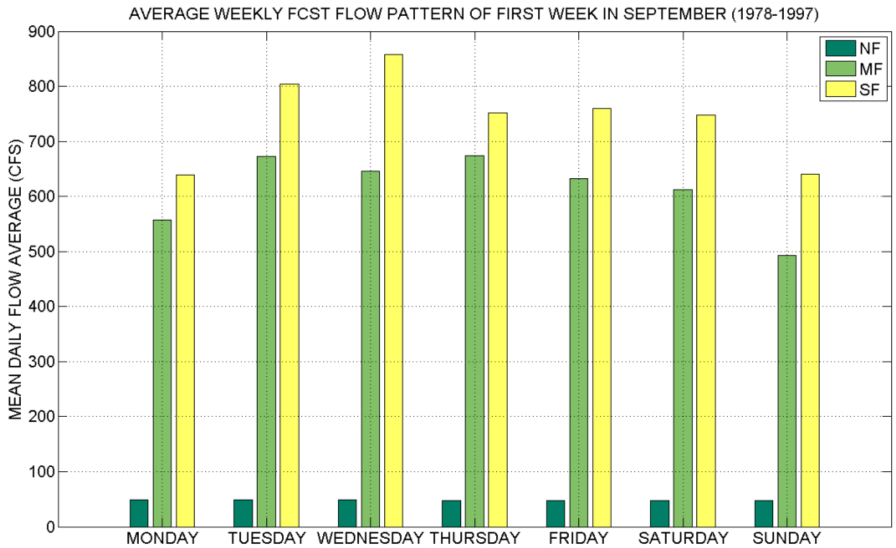


Figure 35: As in Figure 34 but for the Adjusted Ensemble Forecasts with Start Date in June.

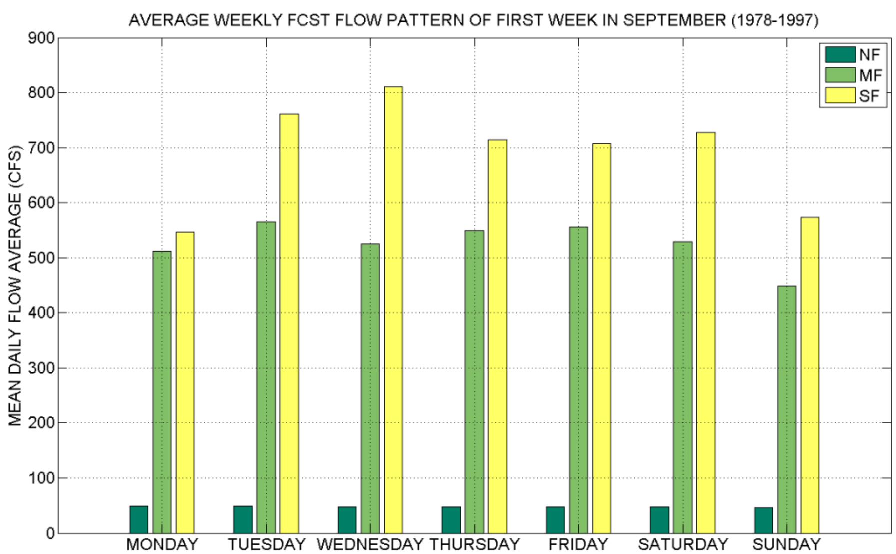


Figure 36: As in Figure 34 but for the Adjusted Ensemble Forecasts with Start Date in July.

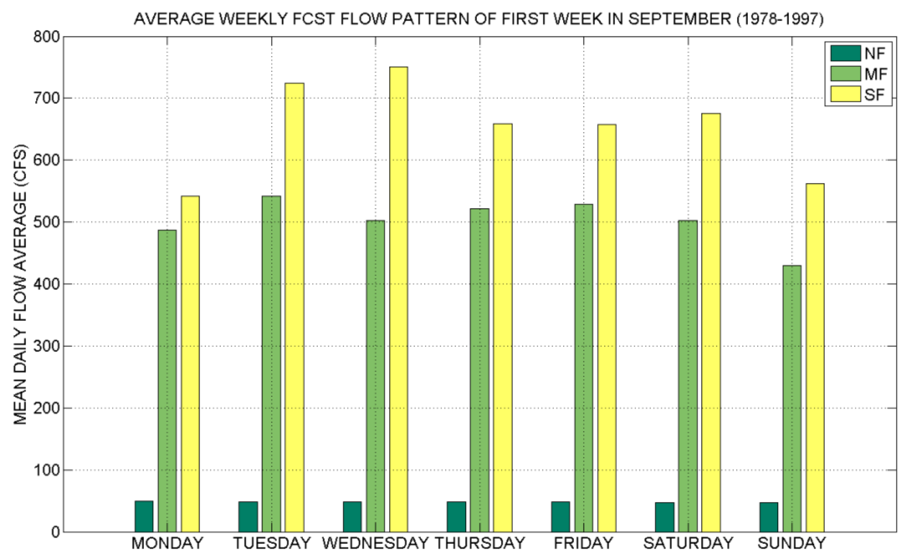


Figure 37: As in Figure 34 but for the Adjusted Ensemble Forecasts with Start Date in August.

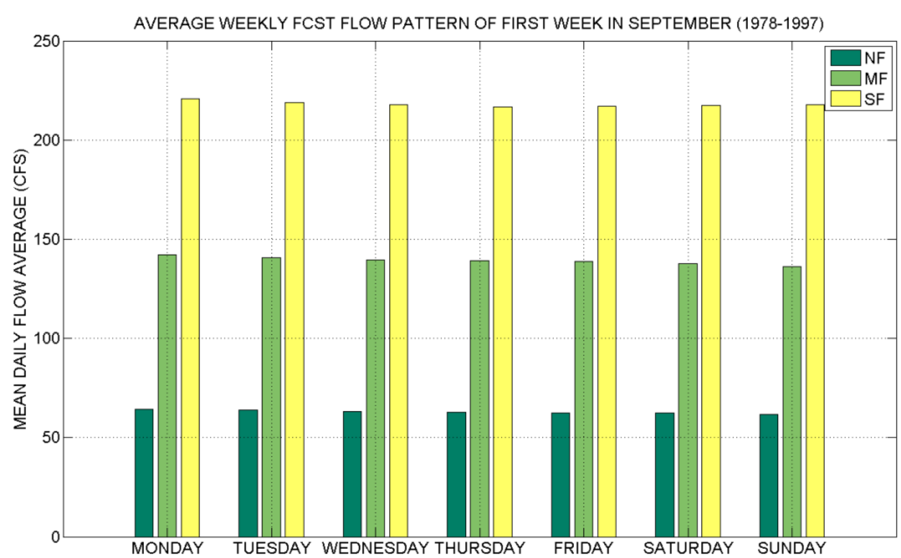


Figure 38: As in Figure 35 but for the Unadjusted ESP.

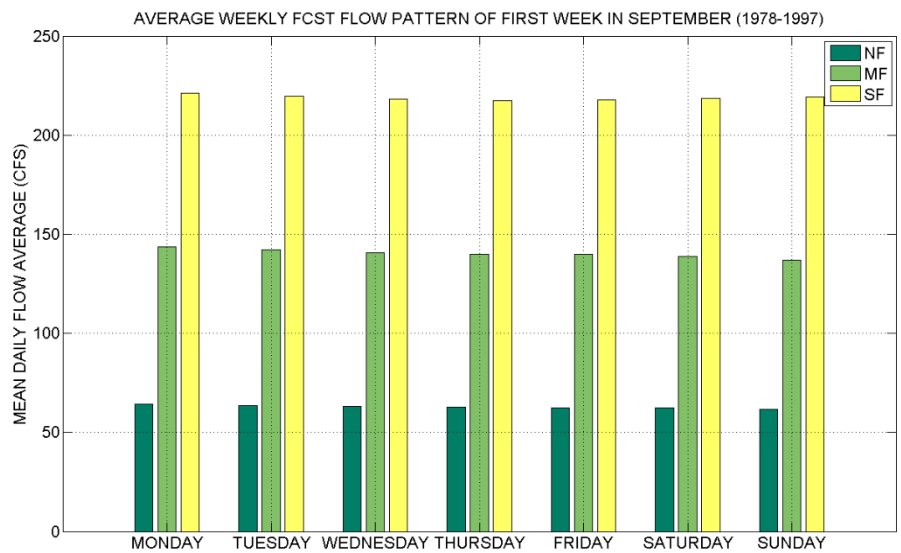


Figure 39: As in Figure 36 but for the Unadjusted ESP.

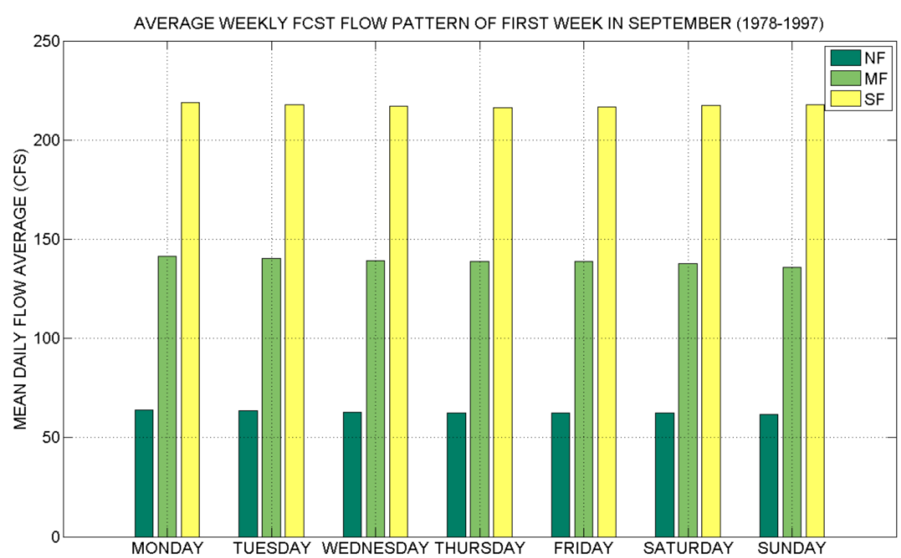


Figure 40: As in Figure 37 but for the Unadjusted ESP.

Figures 41 through 47 show results for the mean daily flow of the days of the third week of September, averaged over the years of record. The conclusion from these results is similar to the one drawn before for the first week of September. Adjustment by the methods developed leads, on average, to improved ensemble forecasts at the daily level during months of flow augmentation.

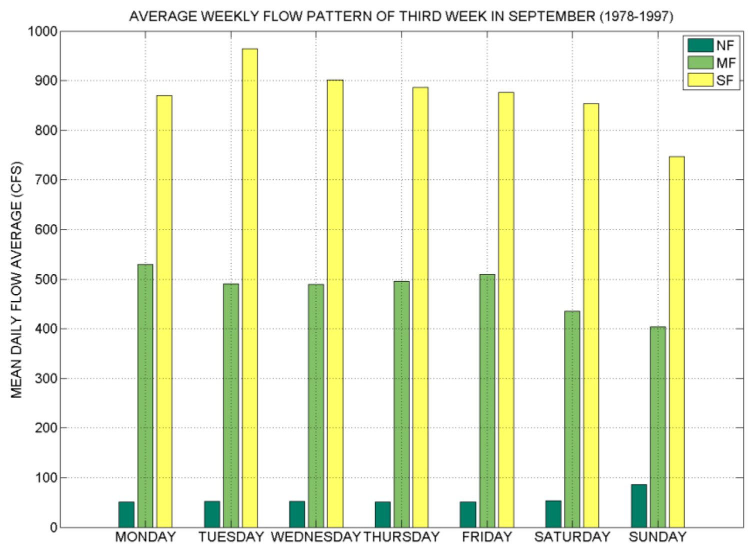


Figure 41: Climatological Average of Mean Daily Flow (cfs) for Each Day of the Third Week in September for all Three Forks of the American River.

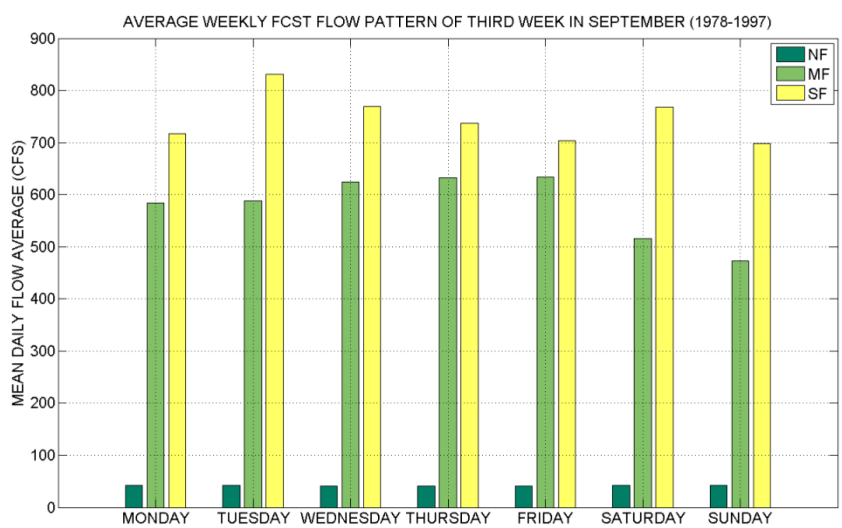


Figure 42: As in Figure 41 but for the Adjusted Ensemble Forecasts with Start Date in June.

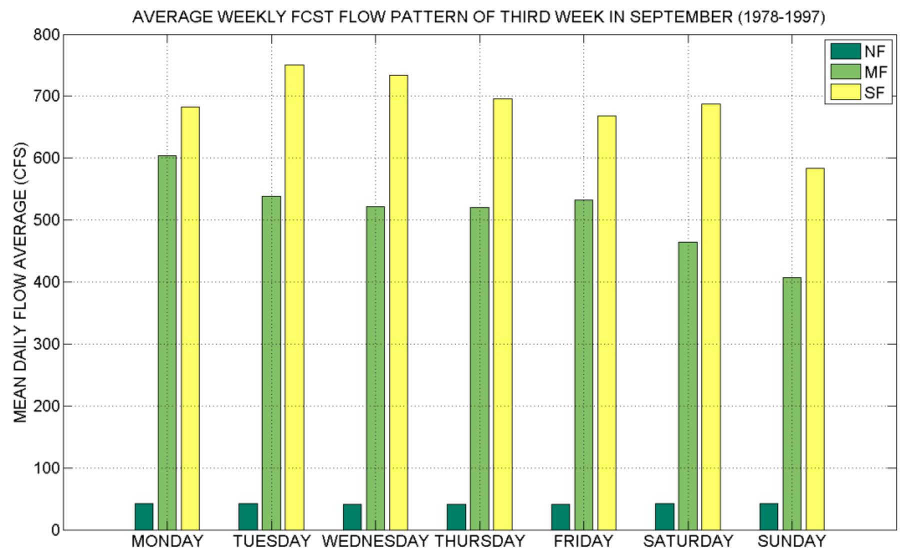


Figure 43: As in Figure 41 but for the Adjusted Ensemble Forecasts with Start Date in July.

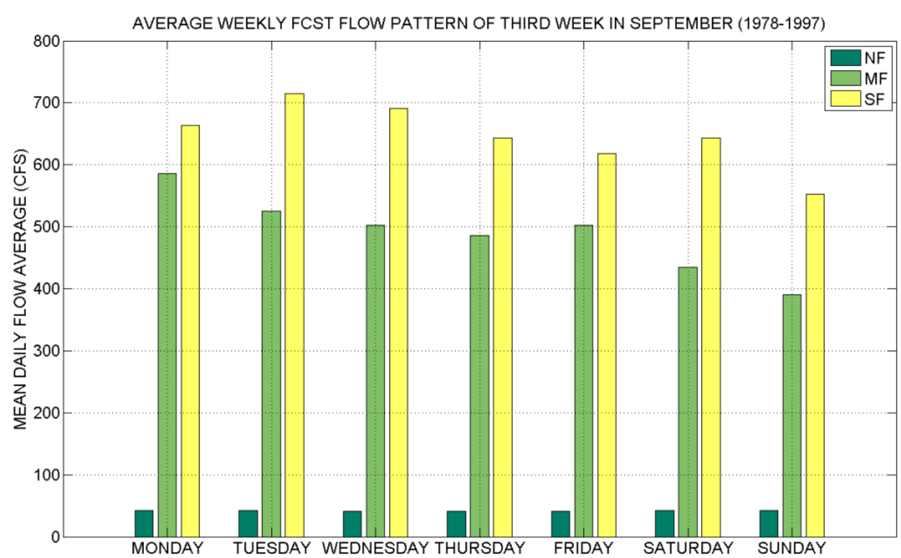


Figure 44: As in Figure 41 but for the Adjusted Ensemble Forecasts with Start Date in August.

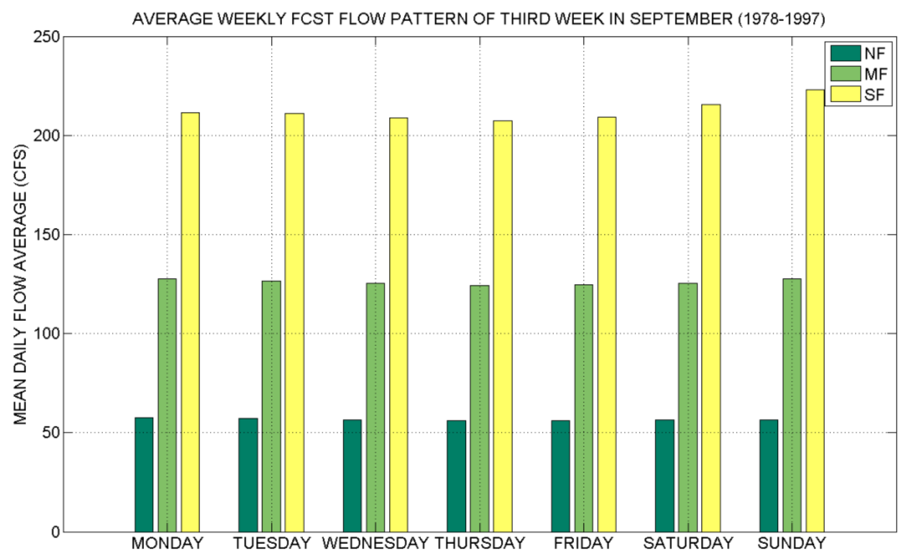


Figure 45: As in Figure 42 but for the Unadjusted ESP.

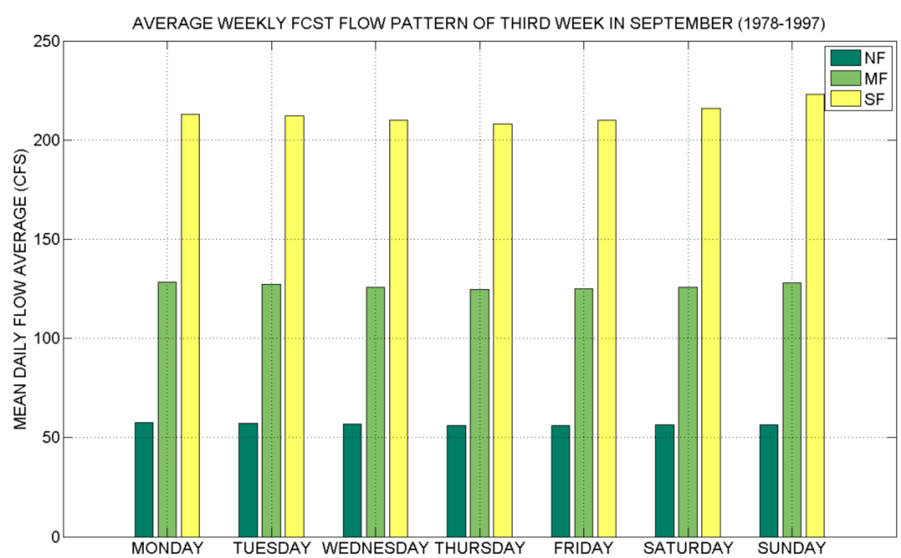


Figure 46: As in Figure 43 but for the Unadjusted ESP.

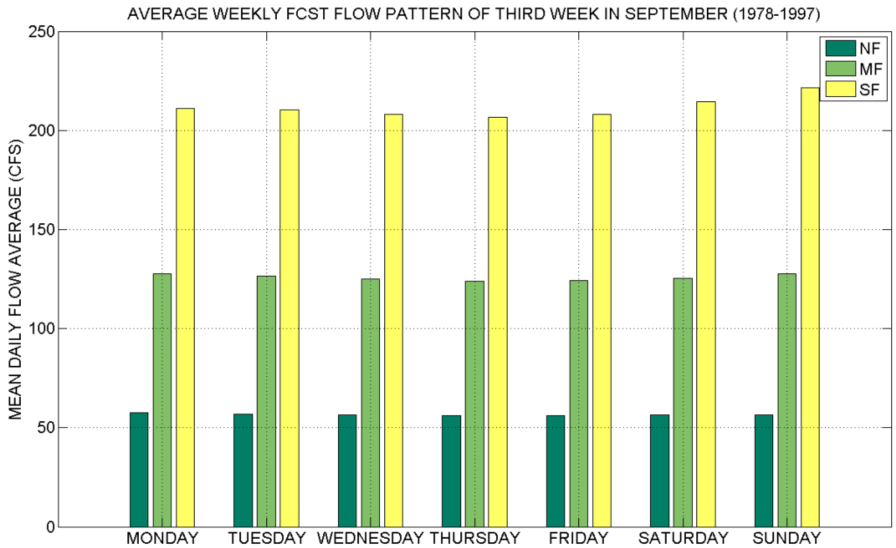


Figure 47: As in Figure 44 but for the Unadjusted ESP.

5.3 Probability Measures

The probabilistic performance of the adjusted and unadjusted ensemble forecasts with various forecast lead times is highlighted first for each of the years of record using boxplot diagrams. Figures 48, 49 and 50 exemplify the boxplots of the adjusted ensemble forecasts for each year of record for the Middle Fork of the American River for the target variable of average September mean daily flow and with start times in June, July and August, respectively. Analogous displays for the unadjusted ensemble forecasts are shown in Figures 51, 52 and 53. In all these Figures the boxplots are drawn to display the median (red line in the box), and the 25th and 75th percentiles (box lower and upper limit lines). The whiskers indicate the range of values that are lower and greater than the 25th and 75th percentiles, respectively, but are included in a distance equal to 1.5 the length of the box, below the lower limit and above the upper limit of the box. Values outside the whisker range are considered outliers and are shown with the red markers.

The substantial improvement of reliability of the unadjusted ESP forecasts for the September average flow when adjustments are made is evident. In most years the adjusted flow ensemble boxplots include the observed flow for all lead times. There are some years that the observed flow is beyond even the forecast trace outliers, indicating substantial deviation of the regulation policies upstream for that particular year.

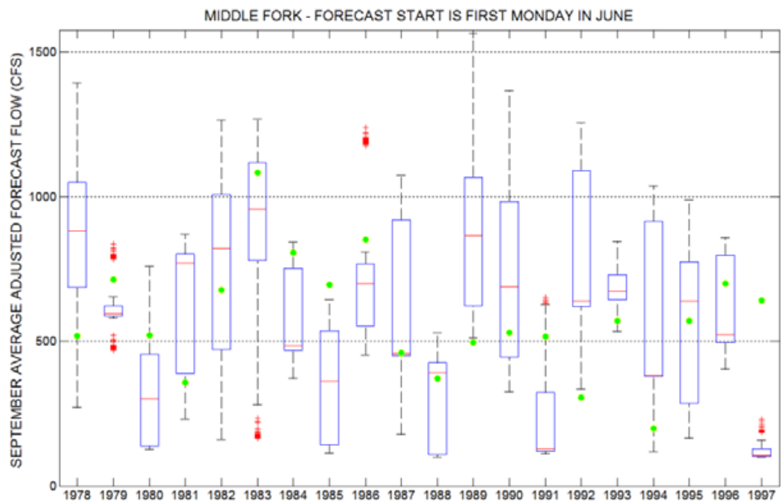


Figure 48: Boxplot Diagrams of the Ensemble Forecasts of Average September Mean Daily Flow for Each Year of Record, for the Middle Fork and for a Start Time on the First Monday in June. The Corresponding Observations are shown with Green Markers.

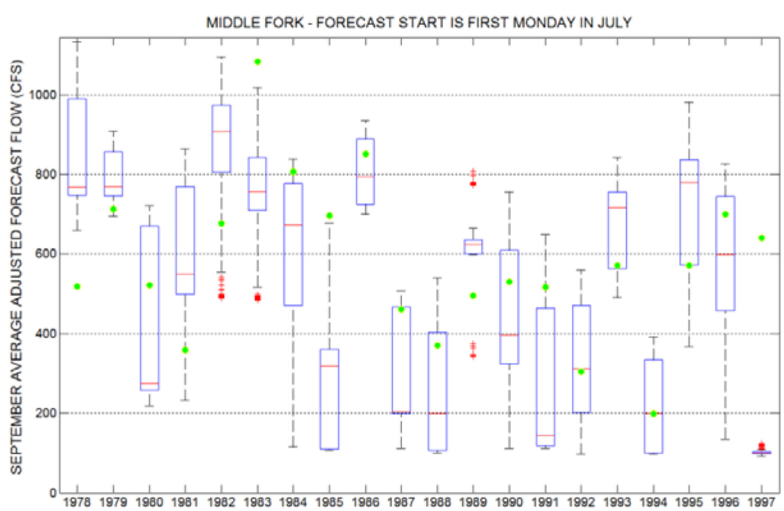


Figure 49: As in Figure 48 but with a Start Time on the First Monday in July.

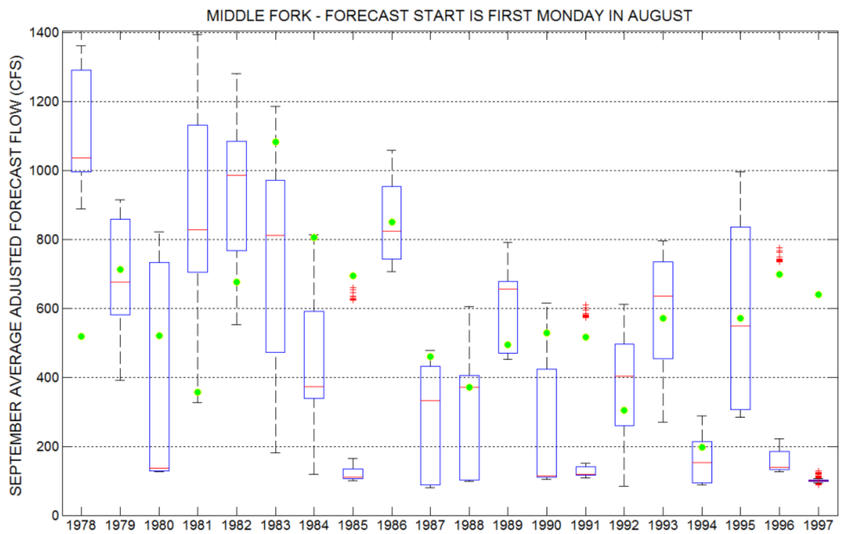


Figure 50: As in Figure 48 but with a Start Time on the First Monday in August.

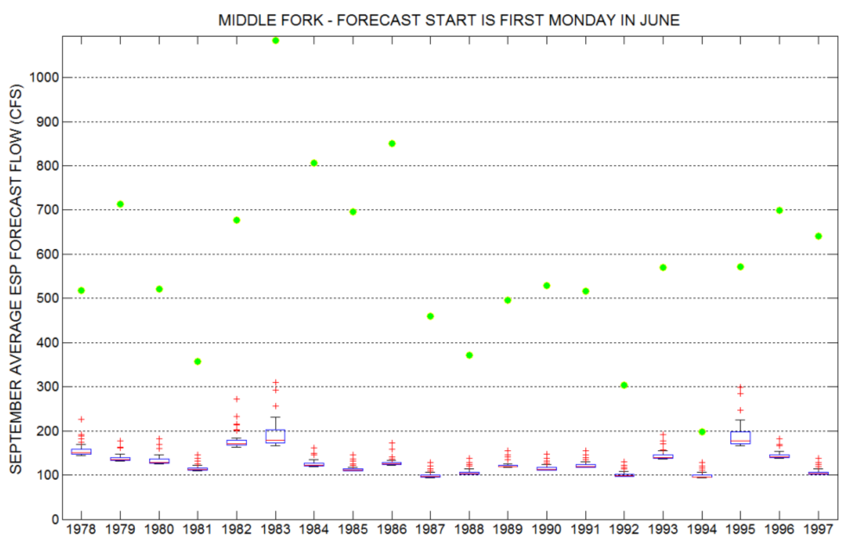


Figure 51: As in Figure 48 but for Unadjusted ESP.

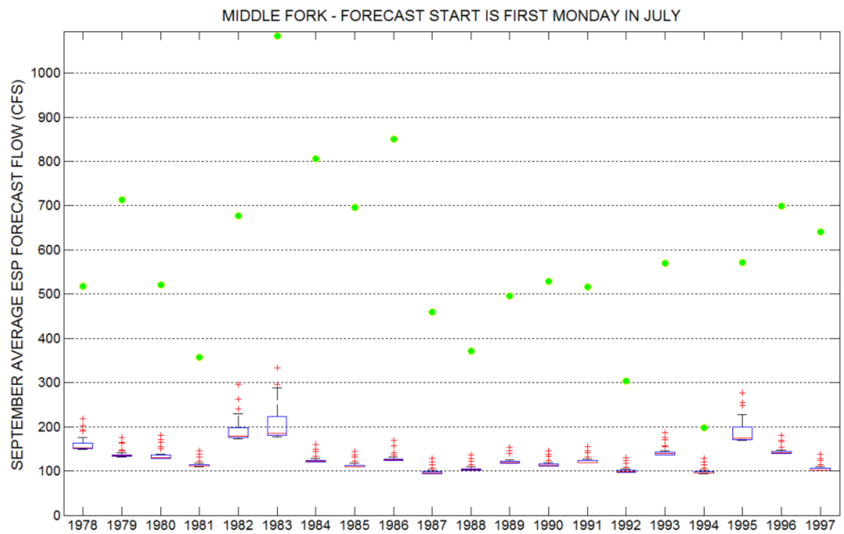


Figure 52: As in Figure 49 but for Unadjusted ESP.

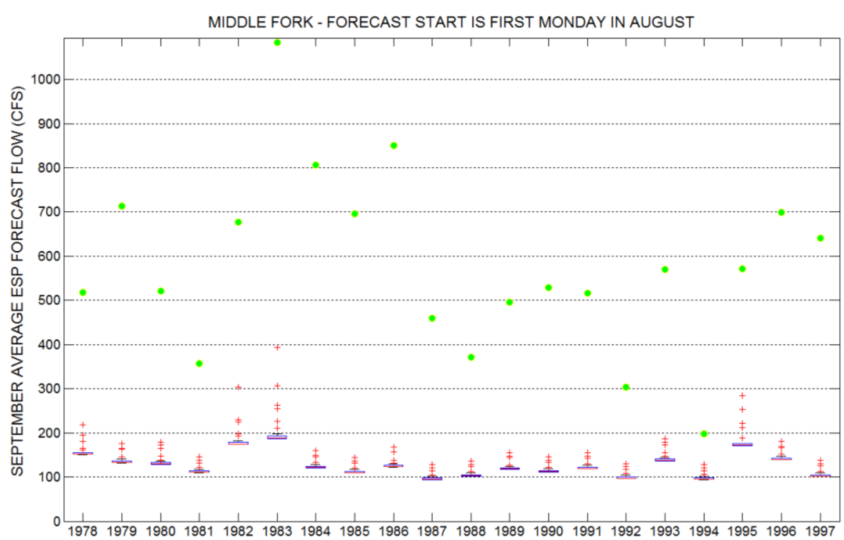


Figure 53: As in Figure 50 but for Unadjusted ESP.

An example for the weekly-scale probabilistic performance is given in Figures 54 and 55 and concerns the average mean daily flow for the third week in September of all the years of record for the South Fork and with a start date on the first Monday of July. Again the adjusted ESP forecasts possess significant reliability as compared to the unadjusted ESP forecasts.

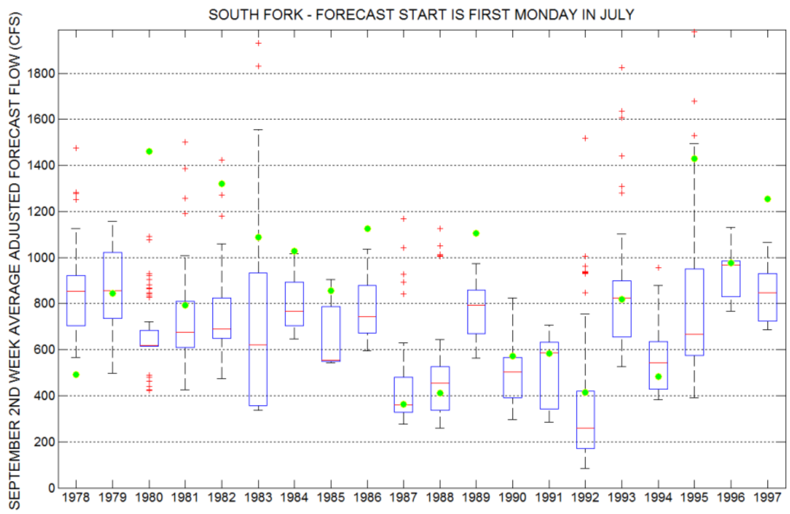


Figure 54: Boxplot Diagrams of the Ensemble Forecasts of Average Mean Daily Flow for the Third Week in September for Each Year of Record, for the South Fork and for a Start Time on the First Monday in July. The Corresponding Observations are shown with Green Markers.

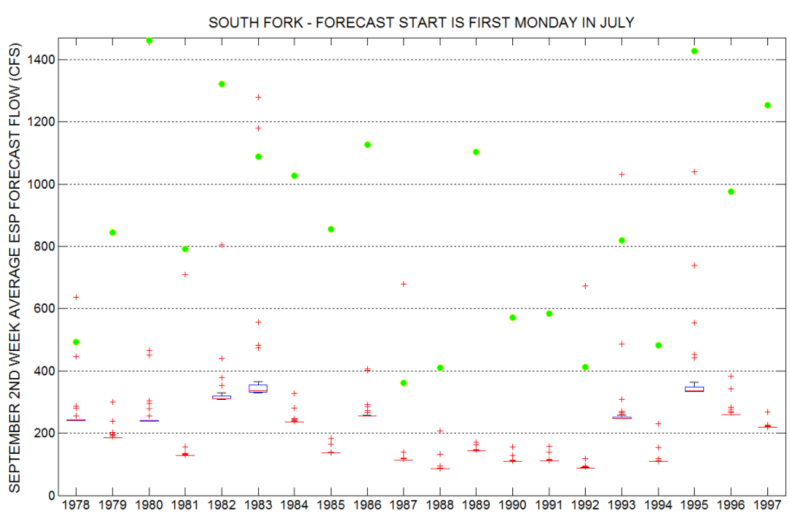


Figure 55: As in Figure 54 but for Unadjusted ESP.

Daily-scale reliability is demonstrated for the Middle and South Forks of the American River in Figures 56 and 57, respectively. The Figures show boxplots of adjusted ensemble forecasts of mean daily flow for the Monday and Wednesday of the third week in September of each year.

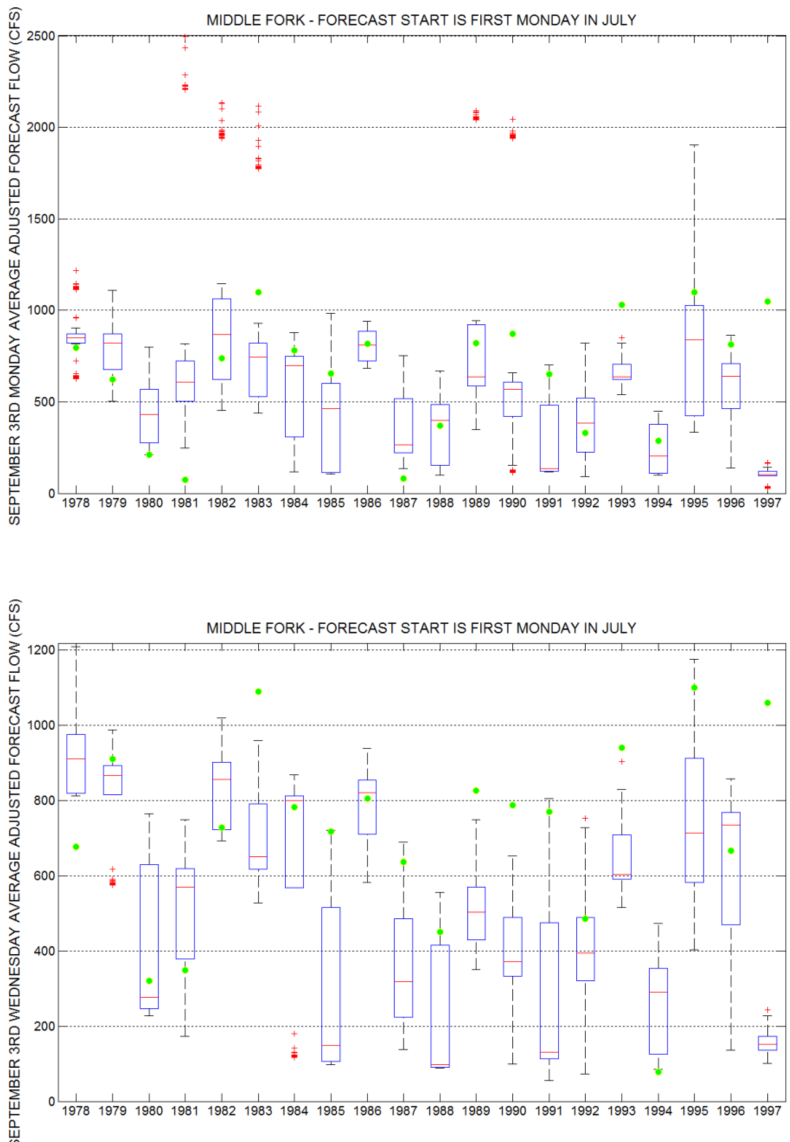


Figure 56: Boxplot Diagrams of the Ensemble Forecasts of Mean Daily Flow for the Monday (upper panel) and Wednesday (lower panel) of the First Week in September for Each Year of Record, for the Middle Fork and for a Start Time on the First Monday in July. The Corresponding Observations are shown with Green Markers.

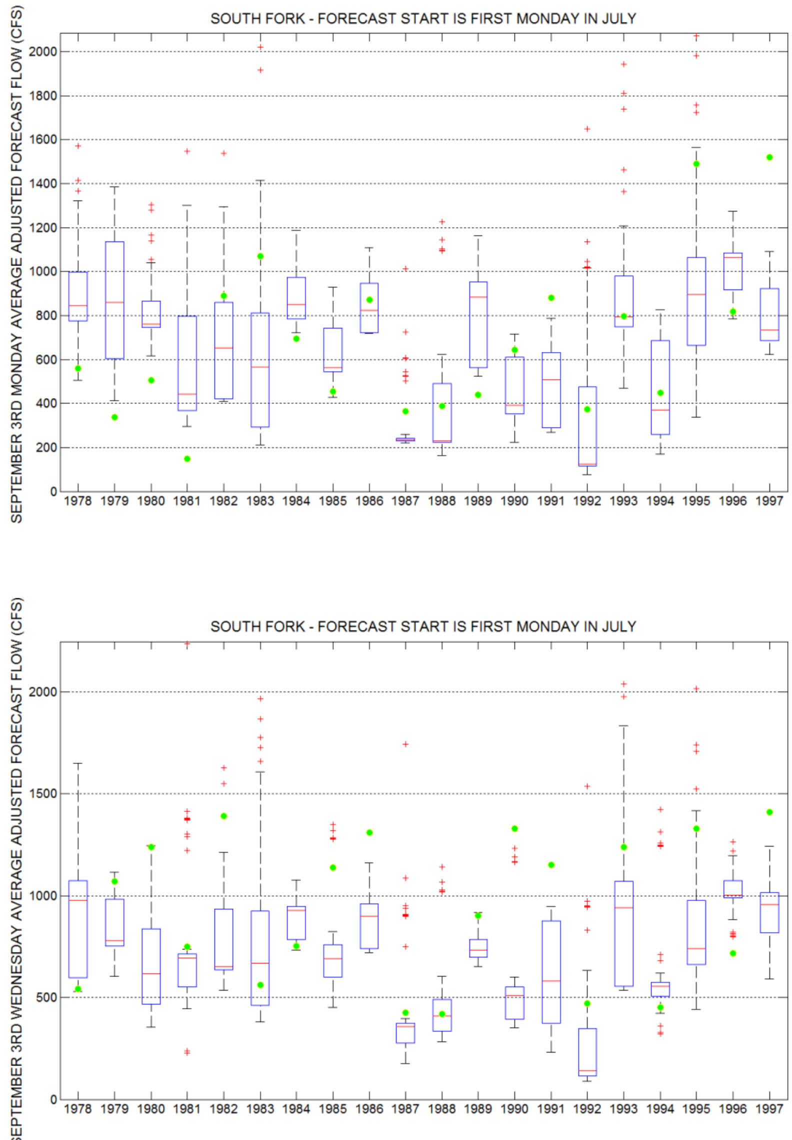


Figure 57: As in Figure 56 but for the South Fork.

The boxplots of this section have demonstrated that the adjusted ESP possesses significant reliability for the low flow periods when upstream regulation is active on all scales: monthly, weekly and daily. Skill scores based on the Rank Probability Score have been computed and quantify the improvement in reliability of the adjusted ESP with respect to the unadjusted ESP for the case of probabilistic forecasts derived from the ensemble forecasts on the basis of the observed flow terciles. These scores measure the skill of probabilistic forecast that the flow will be in the lower, middle and upper tercile of the observed flow distribution over the 20 years and for each target variable. These skill scores are presented next.

Table 8 shows the percent improvement of the adjusted ESP with respect to the unadjusted ESP on the basis of the RPS (Equation (9) expressed in percent) computed for the average mean daily flow in September (Sep Av), the first week in September (1WSep Av), the third week in September (3WSep Av), the average mean daily flow in October (Oct Av), the first week in October (1WOct Av), and the third week in October (3WOct Av). Performance is quantified for start dates on the first Monday of June, July and August. In an analogous manner, Tables 9 and 10 quantify performance with respect to the daily scale for all the days in the first and third weeks of September, respectively. Tables 8, 9 and 10 present Middle Fork results. Tables 11, 12 and 13 present analogous results for the South Fork.

Table 8: Skill Score (%) based on the Rank Probability Score (Adjusted vs. Unadjusted ESP); Middle Fork

Start Mon	*Sep Av*	*1WSep Av*	*3WSep Av*	*Oct Av*	*1WOct Av*	*3WOct Av*
June	21.0	29.1	17.9	97.0	94.2	98.3
July	42.4	16.9	29.5	96.6	95.5	98.4
August	30.0	-2.1	18.6	96.4	96.2	97.9

Table 9: Skill Score (%) based on the Rank Probability Score for Daily Scale; Middle Fork, First Week in September

Start Mon	*MON*	*TUE*	*WED*	*THU*	*FRI*	*SAT*	*SUN*
June	36.3	7.2	27.3	7.9	28.1	32.9	41.6
July	26.8	2.9	36.8	-1.3	7.9	31.3	34.3
August	21.0	-6.8	25.3	-13.8	-2.2	31.5	32.3

Table 10: Skill Score (%) based on the Rank Probability Score for Daily Scale; Middle Fork, Third Week in September

Start Mon	*MON*	*TUE*	*WED*	*THU*	*FRI*	*SAT*	*SUN*
June	15.3	30.4	3.8	5.8	-6.3	14.3	99.4
July	13.1	29.4	15.9	9.7	4.7	14.9	99.2
August	-5.9	24.0	15.2	10.6	5.2	14.2	99.1

Table 11: Skill Score (%) based on the Rank Probability Score (Adjusted vs. Unadjusted ESP); South Fork

Start Mon	*Sep Av*	*1WSep Av*	*3WSep Av*	*Oct Av*	*1WOct Av*	*3WOct Av*
June	89.6	91.6	96.4	94.5	96.7	92.6
July	95.2	91.0	98.0	93.2	92.9	93.5
August	89.9	75.2	90.6	95.1	94.1	94.3

Table 12: Skill Score (%) based on the Rank Probability Score for Daily Scale; South Fork, First Week in September

Start Mon	*MON*	*TUE*	*WED*	*THU*	*FRI*	*SAT*	*SUN*
June	94.3	78.4	75.8	96.3	97.7	76.2	78.3
July	97.6	92.1	95.0	90.3	95.5	96.0	95.4
August	91.4	20.3	90.3	76.6	92.4	80.5	91.5

Table 13: Skill Score (%) based on the Rank Probability Score for Daily Scale; South Fork, Third Week in September

Start Mon	*MON*	*TUE*	*WED*	*THU*	*FRI*	*SAT*	*SUN*
June	98.1	95.1	72.0	90.6	91.3	96.6	91.0
July	98.1	98.2	96.0	97.0	92.1	96.8	92.6
August	91.7	79.1	79.2	90.7	92.1	91.5	84.1

The results demonstrate the significant skill of the adjusted ESP forecasts with respect to the unadjusted ESP forecasts on all scales. The improvement is especially impressive for the heavily regulated South Fork of the American River for all scales (monthly, weekly and daily), ranging from 72% to more than 95%. For the Middle Fork improvement is less impressive yet substantial for both September and October. There are certain days in the first week of September at certain lead times that have negative scores (up to -13%) indicating no improvement for these days of that week, possibly due to random deviations of upstream regulation from set patterns of regulation. Overall, improvement for the Middle Fork in September is on average between 0 and 40%. In October this increases to more than 95%. These results attest to the skill of probabilistic forecasts based on adjusted ESP forecasts that consider upstream regulation effects.

Page Left Blank

6. CONCLUSIONS AND RECOMMENDATIONS

This research developed and evaluated new methods to account for upstream regulation in operational ensemble hydrologic forecasts. Forecast adjustment is particularly necessary during the dry season when full natural flow forecasts exhibit consistent biases compared to observed streamflows. The proposed methods are designed to simulate the processes of flow modification (storage and release, water transferS, and water use), and are based on routinely available information such as:

Daily or sub-daily flow observations at the watershed outlet;

Unimpaired watershed outflow sequences, either model generated (based on contemporaneous data of watershed precipitation, temperature, evapotranspiration, and flow) or reconstructed from observed outflows and knowledge of existing water uses;

Anecdotal or quantitative information (obtained by water agencies and other stakeholders) on the nature, timing, and quantities of water uses and transfers (be they exports or imports), including release rules for some of the existing storage facilities and instream flow requirements.

The present report describes the modeling framework including parameter estimation and forecast uncertainty characterization procedures. Extensive tests of the proposed methods in three California watersheds using adaptations of the operational models and procedures of the California Nevada River Forecast Center (CNRFC) demonstrate and assess the forecast improvement value in association with operational forecasting systems. The results indicate substantial improvements that often exceed 20% and in some cases reach more than 95% for monthly, weekly and daily scales. However, for some specific days of certain weeks no improvement is noted with respect to the present operational forecasting procedures, indicating that additional information on daily release patterns could yield further improvements.

The proposed methods can potentially be fine-tuned further by application to additional River Forecast Center (RFC) areas with different types of upstream regulation. Following this experience, the methods could be implemented within the operational forecast early warning system (FEWS) that RFCs use in the US. Close collaboration with and training of the RFC forecasters would be an essential component of this effort.

Page Left Blank

7. REFERENCES

Anderson, E. A., 1973: National Weather Service River Forecast System – Snow accumulation and ablation model. *NOAA Technical Memorandum NWS HYDRO-17*. Office of Hydrology, National Weather Service, NOAA, Silver Spring, MD, 217 pp.

Georgakakos, K. P., 1986: A Generalized stochastic hydrometeorological model for flood and flash-flood forecasting, 1 Formulation. *Water Resources Research* **22**(13), 2083-2095.

Georgakakos, K. P., and R. L. Bras, 1982: Real-time statistically linearized adaptive flood routing. *Water Resources Research* **18**(3), 513-524.

Georgakakos, K. P., Georgakakos, A. P., and H. Yao, 2009: Characterizing Uncertainty in Upstream Regulation Actions for Operational Ensemble Streamflow Prediction. First Progress Report NOAA NA08NWS4620023, Hydrologic Research Center, San Diego, CA, 58 pp.

HRC-GWRI - Georgakakos, K. P., N. E. Graham, T. M. Carpenter, E. Shamir, J. Wang, J. A. Sperfslage, S. Taylor, of HRC, in collaboration with A. P. Georgakakos, H. Yao, and M. Kistenmacher of GWRI, 2007: Integrated Forecast and Reservoir Management (INFORM) for Northern California: System Development and Initial Demonstration, HRC Technical Report No. 5, Hydrologic Research Center, San Diego, CA 570pp. Project sponsored by the California Energy Commission, CalFed, and NOAA, see on line version at the California Energy Commission web site: http://www.energy.ca.gov/pier/project_reports/CEC-500-2006-109.html

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

Tsintikidis, D., Georgakakos, K. P., Sperfslage, J. A., Smith, D. E., and T. M. Carpenter, 2002: Precipitation Uncertainty and Raingauge Network Design within the Folsom Lake Watershed. *ASCE Journal of Hydrologic Engineering* **7**(2), 175-184.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, San Diego, CA, 467 pp.

Page Left Blank

APPENDIX A
STEP BY STEP PROCEDURE FOR GENERATING ADJUSTED ESP TRACES FOR UPSTREAM REGULATION

This appendix provides a step by step procedure for generating adjusted ESP traces for upstream regulation. The example provided uses one ESP trace with forecast starting from July 6, 1987 for the Middle Fork. Three historical analogue release patterns are selected to generate three adjusted ESP traces. For *M* ESP traces and *n* historical analogue traces this procedure generates *M x n* adjusted ESP traces.

1. Consider forecast start times in the dry season, namely, after the time when flow augmentation begins. This time is marked by the crossover of the FNF simulation sequence below the observed regulated flows.

   Example: The FNF ESP traces are generated from the hydrologic models and used as inputs in this procedure. Figure A-1 depicts a sample plot with one ESP trace.

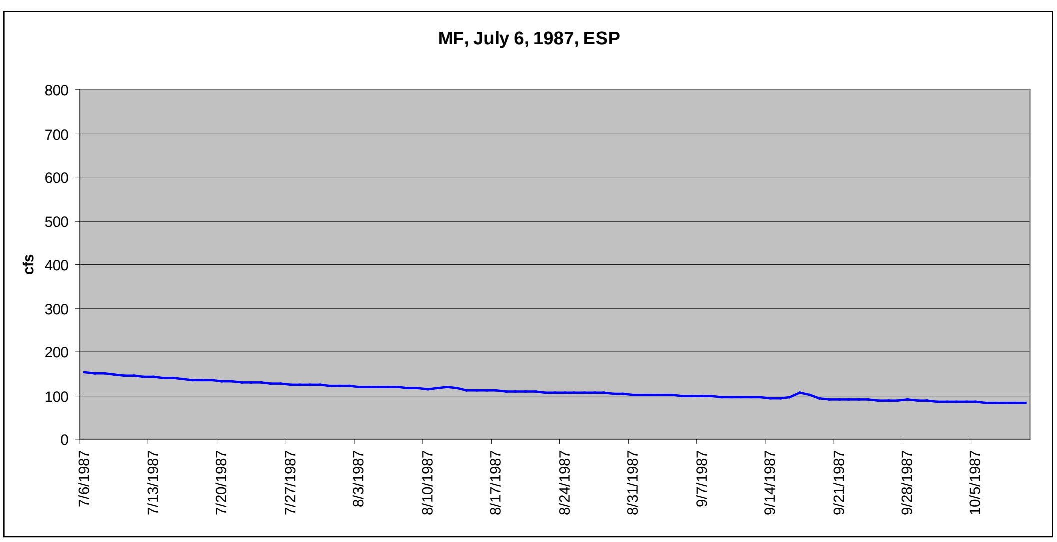


Figure A-1: Sample ESP with one trace

2. Select *n* (e.g., 5 to 7) reference years from the last *N* years (10 – 20 most recent years) based on how close the observed mean daily flow traces during the filling period (March to June) are to the current year analogous trace for the same period. *N* should be selected to include years that are representative of the current upstream regulation infrastructure and practices. "Close" is assessed by a distance measure of mean daily flows (historical from current). This distance measure can also include other quantities such as the aggregate storage (discussed below), and the dry season forecast, among others. Some sensitivity is performed here for each particular basin to determine the best set of historical analog measure.

   Example: For simplicity, only three closest reference years are picked. Based on the "close" criteria used, three years 1987, 1992, and 1996 are selected.

3. Estimate the aggregate filling period storage (as the cumulative difference between FNF and observed flows) for the current year.

   Example: This is computed using the water balance equation. The aggregated filling period storage at the beginning of the dry period (July 1, 1987) turned out to be 83424. (The unit of this storage value is the summation of daily cfs values. This number divided by the number of days will result in average cfs value.)

4. Find for each historical analogue year the fraction of aggregate storage that is dedicated to transfers and the fraction that is dedicated to flow augmentation storage.

   Example: The aggregated filling period storages and transfers for all historical years are pre-computed and saved in the input database. These values can be retrieved from the database. The fraction that is dedicated to flow augmentation storage can be computed based on the retrieved values. The fraction is 0.46, 0.31, and 0.40 for year 1987, 1992, and 1994, respectively.

5. Develop initial storage conditions for each ESP trace and each of the historical analogue years by scaling the current storage by the fraction of storage dedicated to flow augmentation to obtain n initial storage conditions for each ESP trace.

   Example: The available storage for flow augmentation is estimated using the fraction obtained in step 4. In this example, it turned out to be 37608, 25087, and 32763 for year 1987, 1992, and 1994, respectively. Note the numbers are adjusted for the forecast starting date July 6. The initial value 83424 is the storage value for July 1, 5 days earlier.

6. For each forecast ESP trace and storage initial condition (corresponding to each historical analogue year), find the weekly average volume of water available for release during the dry period (from release start to the end of the dry period). This volume is found by adding the forecast trace volume during the remaining dry period to the current augmentation storage and dividing by the number of weeks in the remaining dry period.

Example: Total volume available for the dry period is equal to the initial storage plus the total ESP forecast volume over the dry period. However, the simulation is carried on the weekly basis with daily time step. With the initial storage and the first week forecast ESP flow, the average daily releases for the first week are computed to be 531, 387, and 508 for analogue year 1987, 1992, and 1994, respectively.

7. From each of the n historical analogue years and for the dry period, develop the average fractional pattern of daily flows with respect to the average weekly flow computed in 6.

Example: The fraction patterns of the daily flows are retrieved from the database for the corresponding weeks in 1987, 1992, and 1994. The following Figures A-2 and A-3 show the daily release pattern fractions for the week of July 6.

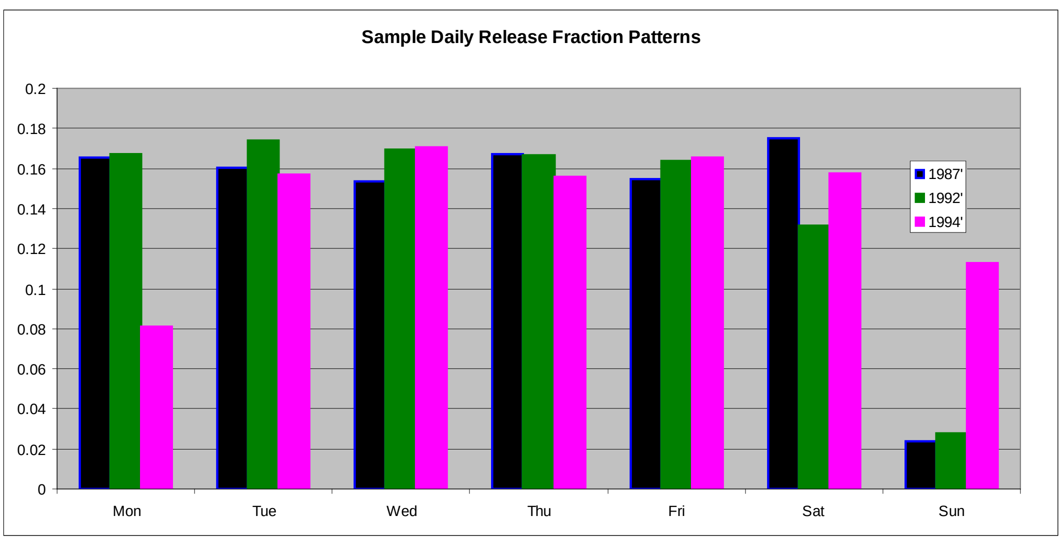


Figure A-2: Sample Daily Release Fraction Patterns of a Week

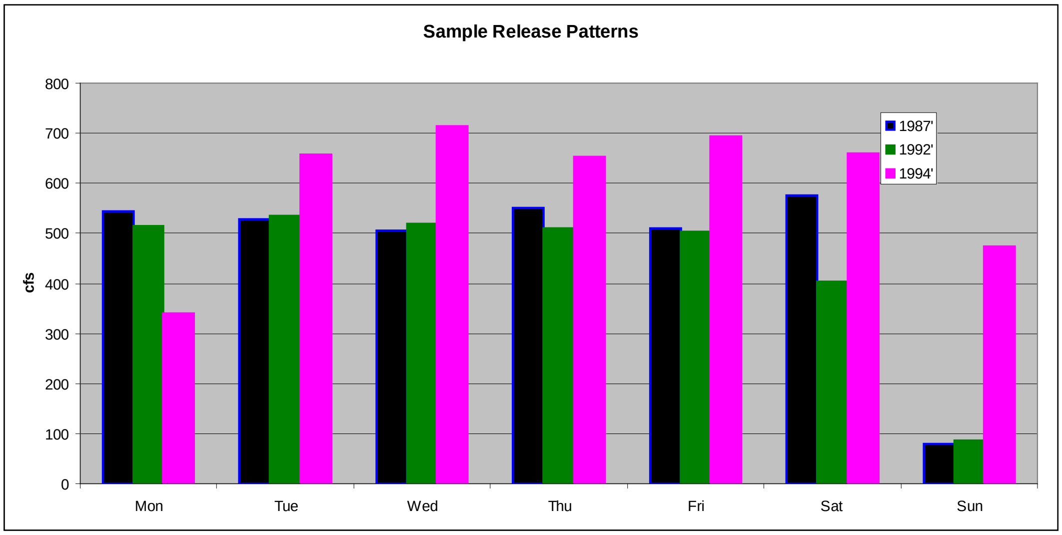


Figure A-3: Sample Daily Release Patterns of a Week

8. Apply the fractional weekly patterns to the corresponding week of the dry period releases for each forecast ESP trace. The release traces thus generated are $n \times M$ where M is the number of ESP traces. This process can also be implemented for sub-daily intervals if necessary.

Example: In this example, three adjusted EPS traces are generated. The results are shown in Figure A-4. For comparison purpose, the original ESP and observed sequences are also plotted.

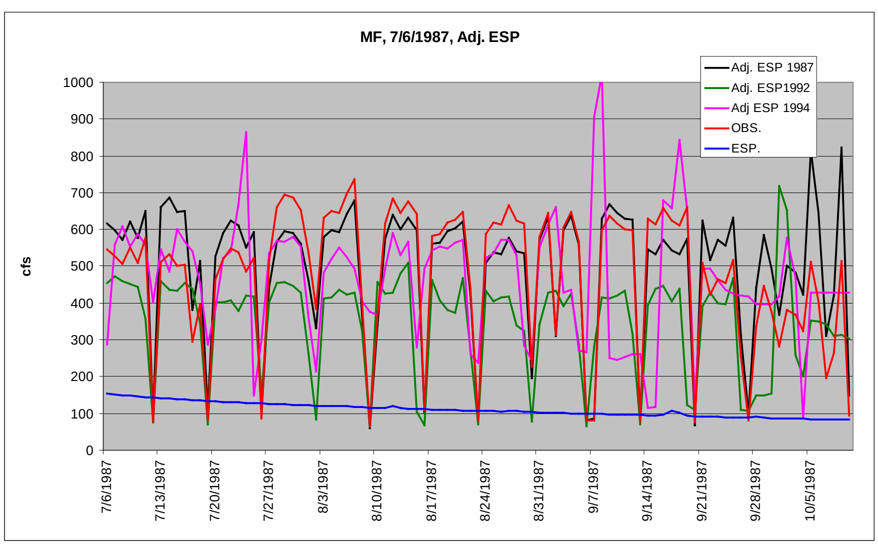


Figure A-4: Adjusted ESP Traces

9. At the next forecast time, update the value of the available flow augmentation storage (corresponding to each historical analogue year) by subtracting the cumulative difference of the observed flow minus the FNF, and repeat the process.

Page Left Blank

APPENDIX B
VALIDATION CLARIFICATIONS

Technical Comment (Tom Gurss, MBRFC):

The methods make sense. However, the conclusions leave me with no idea of accuracy. The conclusions refer to comparisons to the unadjusted natural flow forecasts, which we know have no validity compared to real regulated flows. I don't think that proves much about the accuracy. Analysis of values in the whisker plots indicate that less than half of the forecasts fell within 25-75%. I don't know whether to consider that as adequate. Obviously, it is a huge improvement over natural flow forecasts, but does that mean it adds any value? Perhaps, it would have been more meaningful to compare scores to the historical frequency distribution. Or to the scores of a model that computes flows based on using historical daily diversion/return data, instead of trying to compute the regulation. These would give a better indication of how much value is added by using the proposed method.

Technical Response (Authors):

The project goal is to improve the ensemble forecast predictions coming from the ESP procedure in an operational environment. Note that we did not aim to change the ESP methodology to a new system of possible hybrid forecasts that use model data for one part of the year and observations for another part of the year. As such we have to start from these ESP forecasts and adjust these for upstream regulation. We note that these forecasts have been shown to be very useful for simultaneous water resources management with short term (floods/hydro) AND long term (low flow augmentation/water conservation) objectives in several previous studies (e.g., *HRC-GWRI 2007*, and references therein). Our approach has been to use the uncertainty expressed in these forecasts and enhance that uncertainty with uncertainty due to the upstream regulation additional models and parameters. Thus we believe it is appropriate to compare the improvement to the ESP forecasts at monthly, weekly and day scales and for relatively long lead times (most relevant to water management and drought). The results show significant improvement with respect to several scores as shown in the report.

The problem with a comparison that would be using upstream regulation climatologies ("Or to the scores of a model that computes flows based on using historical daily diversion/return data, instead of trying to compute the regulation") is that in most cases we do not have the real transfers, water withdrawals, returns, etc., and will never have them with good enough accuracy. This is one of the project key assumptions as stated in the report. It is the method we use that allows us to estimate these quantities through aggregate water balance and the reconstruction of the storage-release rules in producing ESP adjustments that improve the ESP.

Comparing to observed regulated flow climatology is useful for longer lead times (months) but for the purpose of checking that the adjusted ESP performs reasonably close to the observed regulated flow climatology of 1/3, 1/3, 1/3 for these long lead times. Climatology may have perfect reliability but it has no resolution or skill (easiest to consider in the context of the reliability diagram for dichotomous events, whereby the climatology forecast is a point rather than a curve, see for example the point raised in *Wilks 2005*, pg. 266). As our focus has been on these lead times for our evaluation (most relevant to water management of low flows and droughts as noted below), we show sample results in Table B-1 for the MF and SF American River for the RPS skill score (% improvement) against climatology (start date of first Monday in July of each year) when the target forecast quantity is the average flow on each day of the week for the 1st week in September:

Table B-1: RPS Skill Score expressing improvement with respect to climatology for the period: 1978-1997. Forecast starts on first Monday of July of each year.

	MON	TUE	WED	THU	FRI	SAT	SUN
MF	7.8	-22.4	20.3	-27.7	-16.1	14.3	18.1
SF	11.3	13.3	-6.4	-8.6	5.2	14.1	-3.3

These results show that the procedure for the long lead times of months is compatible to the climatology on daily scales. Note also that the approach we have taken allows for periods when the climatology of the observed regulated flows is changing due to climate changes or demand changes upstream. Of course, in such cases it would not be feasible to compute the terciles of the actual climatological distribution of the observed regulated flows.

Lastly, note that the interest is on long lead times for low flow or drought events (water conservation) because unless we are able to predict several weeks (or months) ahead, the forecast has limited water resources management value, because once a system is in the drought, the options are limited. The only successful drought strategy is to anticipate the drier flows well in advance and conserve water by reducing releases early on, thereby spreading the release reductions over a longer time period and mitigating the impact.

Page Left Blank