

**Seasonal to Inter-Annual Streamflow Simulation and
Forecasting on the Upper Colorado River Basin and Implications
for Water Resources Management**

by

Cameron W. Bracken

B.S., Humboldt State University, 2009

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Master of Science

Department of Civil, Environmental and Architectural Engineering

2011

This thesis entitled:
Seasonal to Inter-Annual Streamflow Simulation and Forecasting on the Upper Colorado River
Basin and Implications for Water Resources Management
written by Cameron W. Bracken
has been approved for the Department of Civil, Environmental and Architectural Engineering

Balaji Rajagopalan

Edith Zagona

Subhrendu Gangopadhyay

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Bracken, Cameron W. (M.S., Civil Engineering)

Seasonal to Inter-Annual Streamflow Simulation and Forecasting on the Upper Colorado River Basin and Implications for Water Resources Management

Thesis directed by Prof. Balaji Rajagopalan and Prof. Edith Zagona

The past decade has been one of unprecedented drought in the Upper Colorado River Basin (UCRB), having wide impacts on basin hydrology and management. The drought has raised many questions about the sustainability of the UCRB water supply and emphasized the need for accurate forecasts and simulations of streamflow. The UCRB is home to many of the largest reservoirs in the United States including Lake Powell. For large reservoirs such as these, the ability to forecasts on a seasonal to inter-annual time scale is particularly important for management.

The first chapter of this thesis develops a seasonal forecasting framework for generating ensembles of peak season flow at long lead times. The framework uses large scale climate predictors as well as snowpack and soil moisture conditions to produce ensemble forecasts of Lees Ferry natural flow. The natural flow forecasts are then disaggregated spatially and temporally to produce monthly forecasts at four key UCRB sites.

One question raised by the 2000-2010 drought is, how likely is it to occur again? Traditional time series models would estimate the risk of such a drought as very low. Along these lines a Hidden Markov (HM) time series model was developed that is able to capture a distinct regime switching behavior in the Lees Ferry natural flow record. The model suggests that the likelihood of a drought of the current magnitude is much more likely than suggested by traditional autoregressive methods. The HM model is also useful for ensemble forecasting. Here it is used to generate skillful inter-annual forecasts of UCRB flow.

The availability of skillful seasonal and inter-annual streamflow forecasts does not guarantee their usefulness in water supply forecasting. To evaluate this, flow sequences are used as input to the Midterm Operations Model (MTOM), a new operational water supply model for the Col-

orado River Basin. The MTOM was developed as part of this research in collaboration with the Bureau of Reclamation. Ensemble forecasts of reservoir operations are developed and compared along side the existing model for the region.

Dedication

To my family and to my friends, past and present, thanks for being awesome, I love you all.

Acknowledgements

Thank you first of all to my advisors, Balaji and Edie. Your input and guidance has been invaluable. Thanks also to my third committee member Subhrendu Gangopadhyay.

Thank you to my long time office mates Ken and Alan, I thoroughly enjoyed working close to you if not closely with you over the past two years. I would have been stuck 1000 times over without your advice and problem solving skills.

Thank you so much to all the folks at Bureau of Reclamation for putting up with questions and data requests and for providing so much valuable input. In particular Katrina for being my main contact and for being so accommodating; also Heather, Shana, Jim, Carley, Russ, Rick and anyone else involved behind the scenes.

Last but certainly not least, thank you to all the parties who funded my degree: The Department of Hydrology, Water Resources and Environmental Fluid Mechanics at CU Boulder, Western Water Assessment and the Bureau of Reclamation.

Contents

Preface	1
1 A Multi-Site Seasonal Ensemble Streamflow Forecasting Technique	3
1.1 Introduction and Background	4
1.2 Study Area	7
1.3 Data	7
1.3.1 Streamflow Data and Index Gage	7
1.3.2 Large-scale climate data	9
1.3.3 SWE data	9
1.3.4 PDSI data	9
1.4 Proposed Integrated Framework	10
1.4.1 Predictor Suite for Index Gage seasonal streamflow	10
1.4.2 Multimodel Selection for I Gage Seasonal Flow Forecast	10
1.4.3 Multimodel Ensemble Forecast of Index Gage Seasonal Streamflow	13
1.4.4 Spatial and Temporal Disaggregation of the Index Gauge Forecast	14
1.4.5 Verification and Validation	16
1.5 Results	17
1.5.1 Predictor Identification	18
1.5.2 Multi Model Selection	19
1.5.3 Forecast Skill	20
1.6 Summary	25

Bibliography **28**

2	A Nonstationary Hidden Markov Model for Stochastic Streamflow Simulation and Short Term Forecasting in the Upper Colorado River Basin	31
2.1	Introduction	31
2.2	Methodology	35
2.2.1	Model Formulation	35
2.2.2	Moments of the Gamma HM	37
2.2.3	Parameter Estimation and Model Order Selection	38
2.2.4	Global Decoding	39
2.3	Applications of Hidden Markov Models	39
2.3.1	Optimal Model Order and Stationary Distributions	39
2.3.2	Simulation Procedure	41
2.3.3	Forecast Distributions	42
2.3.4	Modified Forecast Procedure	43
2.4	Results	44
2.4.1	Simulation	44
2.4.2	Forecasting	47
2.4.3	Three State Forecasts	50
2.5	Conclusions	51

Bibliography **53**

3	A New Probabilistic Model for Reservoir Operations in the Colorado River Basin	56
3.1	Introduction	56
3.2	Study Area and Data	59
3.2.1	Study Area	59
3.2.2	Flow Data	61

3.2.3	Climate Data for Seasonal Forecasts	62
3.2.4	Snow Data	62
3.2.5	Additional MTOM Data	62
3.3	Methodology	63
3.3.1	CBRFC coordinated forecasts	64
3.3.2	Seasonal Forecast Model	64
3.3.3	Resampled data	67
3.3.4	Hidden Markov Model Timeseries Forecast	67
3.3.5	Generating Unregulated Traces	68
3.3.6	Constructing a continuous trace	69
3.3.7	MTOM model runs	69
3.4	Verification	70
3.4.1	Seasonal Forecast Model Verification	70
3.4.2	HM Verification	71
3.4.3	Inflow Trace Verification and Control Runs	71
3.5	Results	72
3.6	Conclusions	79
	Bibliography	84
	Final Remarks and Future Work	87
	Appendix	88
A	Chapter 1 supporting materials	89
B	Chapter 3 supporting materials	93
B.1	Seasonal Forecast Verification	93

Tables

1.1	Potential predictors and their regions in lon:lat	18
1.2	Selected multimodels for each lead time. “1” indicates the presence of a predictor and “0” indicates the absence of a predictor. A selected predictor refers to the most recent data available at the forecast time (eg. May GPH for an Apr 1st forecast). *This Jan. 1st prediction used the Apr. 1st and Nov. 1st GPH as separate predictors.	19
1.3	Skills of spring season streamflow forecast at different lead times.	24
1.4	Model correlations with historical flow for available Coordinated forecast years . . .	25
2.1	BIC and AIC for both Gamma and Normal HM models.	39
2.2	BIC and AIC for both Gamma and Normal HM models.	40
2.3	Median Correlation (MC) and Median Ranked probability skill score (RPSS) of forecasts for 1980–2010. Dry RPSS, Ave RPSS and Wet RPSS represent the median RPSS for years in the lower, middle and upper tercile respectively.	48
3.1	UCRB reservoirs and corresponding natural flow sites	60
3.2	Natural flow sites.	61
3.3	RPSS values for the index gauge for each lead time.	71
3.4	RPSS and MC (in parentheses) after disaggregation and drop-one cross validation for each lead time. At 95% confidence 0.21 is a significant correlation. Postive skills and significant correlations are shown in bold.	82

3.5	Median Correlation (MC) and Median Ranked probability skill score (RPSS) of forecasts for 1980–2010. Dry RPSS, Ave RPSS and Wet RPSS represent the median RPSS for years in the lower, middle and upper tercile respectively.	83
B.1	Selected Models for each lead time	98

Figures

1.1	Study area.	8
1.2	Leave-one CV forecasts of index gauge spring streamflow issued on (a) Apr 1st and (b) Jan 1st. The box corresponds to the interquartile range, the horizontal line inside the boxes is the median, whiskers extend to 5th and 95th percentile of ensemble and solid line joins the true value for each year. The dashed horizontal lines correspond to the 33rd, 50th and 66th percentiles of the data.	21
1.3	(a) same as Figure 1.2(a) but for Lees Ferry. (b) Same as Figure 1.2(b) but for Lees Ferry.	22
1.4	(a) same as Figure 1.3(a) but for May streamflow at Lees Ferry. (b) same as Figure 1.3(b) but for May streamflow.	23
1.5	Boxplot of RPSS of spring streamflow forecast at the four locations and the index gauge. Forecasts are based on dropping 10% of the observations.	24
1.6	Spring flow forecast at Lees Ferry issued on (a) Jan. 1st and (b) Apr. 1st in a retroactive mode. Also, May flow forecast at Lees Ferry (c) after spatial and temporal disaggregation, issued on Apr. 1st. The dashed line represents the coordinated forecast and the dotted line represents the ESP forecast.	27
2.1	Lees Ferry annual natural flow time series.	32
2.2	Stationary Distributions of HM2G (Top) and HM2N (Bottom).	41

2.3	Global Decoding of the HM2G model using the Viterbi algorithm. Plotted below is the Lees Ferry annual time series with. The horizontal lines indicate the mean of the periods above where the model was in a particular state.	42
2.4	Basic simulation statistics of HM2G, HM2N and AR1 models. Boxplots represent the spread of the simulations and the horizontal line represents the observed value.	45
2.5	Observed (solid line) and boxplots of simulated pdf with the HM2G model.	46
2.6	Observed and simulated (gray region) PDF of run lengths. The gray region extends to the 5th and 95th percentiles of the simulated PDFs.	46
2.7	Mean count (Top) and cumulative mean count (Bottom) of run lengths. Note that these are not probabilities but rather the expected number of run lengths of a given size (or greater than a given size) in a single simulation.	47
2.8	Quartiles of simulated HM2G and AR Global Wavelet Power Spectrum (the normal HM model is similar to the Gamma HM model).	48
2.9	One step ahead forecasts from 1980 to 2010 for the HM2G (Top) and the HM2N (Bottom) with the actual data (solid line).	49
2.10	Boxplots of RPSS values for the forecast period 1980-2010. The horizontal line at 0 represents the skill of the climatological forecast.	49
2.11	Forecast Distributions for 1984 (Left) and 2004 (right). The historical value is shown as a vertical line.	50
2.12	One step ahead forecasts from 1980 to 2010 for the HM3G (Top) and the HM3N (Bottom) with the actual data (solid line).	51
3.1	Map of UCRB with reservoir locations corresponding to Table and a schematic of network connectivity in the MTOM.	60
3.2	The Upper Colorado River Basin with locations of snow data sites used in this study indicated by an open circle.	63
3.3	Schematic of disaggregation process using proportion disaggregation.	65

3.4	Construction of single trace for each lead time.	70
3.5	Unregulated inflow traces (shown as boxplots with whiskers extending to the 10th and 90th percentile of the data) for (a) April 1, Navajo, (b) January 1, Flaming Gorge and (c) November 1, Lake Powell. The green line represents the 24MS most probable trace and the blue line represents the observed unregulated inflow.	73
3.6	Navajo pool elevation for the (a) November, (b) January and (c) April lead times. Boxplots indicate the MTOM results, with whiskers extending to the 10th and 90th percentile of the data. The green line is the 24MS output, the dashed green lines are the 10th and 90th 24MS runs and the blue line is the observed.	74
3.7	Same as Figure 3.6 but for Flaming Gorge Pool elevation.	76
3.8	Same as Figure 3.6 but for Lake Powell Pool elevation.	77
3.9	Same as Figure 3.6 but for Flaming Gorge Outflow.	78
3.10	Same as Figure 3.6 but for Powell Outflow. Note that the control is identical to the 24MS in many months.	80
A.1	Correlation maps of index gage spring (Apr-Jul) streamflow with Feb.-Mar. GPH (top left), March ZW (top right), Jan.-Mar. SST (bottom left) and Mar. MW (bottom right).	90
A.2	Composite maps of Oct.-Mar. 500 mb vector winds for 6 wettest years (top left) and 6 driest years (top right) years and preceding Fall PDSI for wet (bottom left) and dry (bottom right) years.	91
A.3	Boxplots of cross correlation of spring season streamflow forecasts issued on Apr 1st between sites: C=Cisco, B=Bluff, G=GRUT, LF=Lees Ferry. The solid line represents the observed statistics.	92

B.1	Sample ensemble forecasts for (a) Apr1, (b) Feb1 and (c) Nov1 seasonal flow volumes at Gunnison River Near Grand Junction (Site 6). Boxplots extend to the 5th and 95th percentiles of each ensemble. MC stands for the median correlation, the correlation of the median of the ensembles with the historical record.	95
B.2	Same as Figure 3 but for June flows.	96
B.3	Same as Figure B.1 but for retroactive forecasts overlaid with the CBRFC coordinated forecast where available. The blue solid lines are the 10th and 90th percentile and the dashed blue line is the 50th percentile of the coordinated forecast.	97

Preface

This thesis is written in manuscript form. That is, each chapter is self contained in a format that is acceptable for submission to an academic journal. The thesis consists of this preface, three main chapters and concluding chapter containing final remarks and future work.

Chapter 1 has already been published in Water Resources Research and is reproduced here by permission of the American Geophysical Union. The chapter describes a seasonal forecasting framework for generating ensembles of peak season flow at long lead times. The framework uses large scale climate predictors as well as snowpack and soil moisture conditions to produce ensemble forecasts of Lees Ferry natural flow at lead times starting in November. The natural flow forecasts are then disaggregated spatially and temporally to produce monthly forecasts at four key UCRB sites. The forecasts are shown to have positive skills even in November when no snow information is available. The study was proof of concept but was expanded upon and shown to be useful on a larger scale in Chapter 3.

Chapter 2 outlines the development of a nonstationary Hidden Markov model with Gamma component distributions, as opposed to Normal distributions which is widely used in the literature, for stochastic simulation and short term forecasting. The HM model is able to capture a long term persistence and distinct regime switching behavior in the Lees Ferry natural flow record. The model is used for both simulation and forecasting of flows in the UCRB.

Chapter 3 brings together forecasts generated with the techniques described in the first two chapters. The seasonal forecast is expanded to cover all natural flow sites in the UCRB. These are then aggregated and combined with historical data and HM forecasts to produce input traces

the Midterm Operations Model (MTOM), a new operational water supply model for the Colorado River Basin. The MTOM was developed as part of this research in collaboration with the Bureau of Reclamation. The MTOM is run for three lead times November 2007, January 2008 and April 2008. The results are examined alongside historical data and output from the 24-month study, the existing water supply forecasting model in the UCRB.

Chapter 4 makes concluding remarks and suggests further research.

Chapter 1

A Multi-Site Seasonal Ensemble Streamflow Forecasting Technique

Reproduced by permission of American Geophysical Union¹.

Abstract. We present a technique for providing seasonal ensemble streamflow forecasts at several locations simultaneously on a river network. The framework is an integration of two recent approaches; the nonparametric multimodel ensemble forecast technique (*Regonda et al.*, 2006) and the nonparametric space-time disaggregation technique (*Prairie et al.*, 2007). The four main components of the proposed framework are: (i) an index gauge streamflow is constructed as the sum of flows at all the desired spatial locations, (ii) potential predictors of the spring season (April-Jul) streamflow at this index gauge are identified from the large scale ocean-atmosphere-land system including snow water equivalent, (iii) the multimodel ensemble forecast approach (*Regonda et al.*, 2006) is used to generate the ensemble flow forecast at the index gauge, (iv) the ensembles are disaggregated using a nonparametric space-time disaggregation technique (*Prairie et al.*, 2007) resulting in forecast ensembles at the desired locations and for all the months within the season. We demonstrate the utility of this technique in skillful forecast of spring seasonal streamflows at four locations in the Upper Colorado River Basin at different lead times. Where applicable, we compare the forecasts to the Colorado Basin River Forecast Center's (CBRFC) Ensemble Streamflow Prediction (ESP) and the National Resource Conservation Service, 'coordinated' forecast, which is a combination of the ESP, Statistical Water Supply (SWS), a principal component regression technique, and modeler knowledge. We find that overall the proposed method is equally skillful to

¹ Original Citation: C. Bracken, B. Rajagopalan, and J. Prairie (2010), A multisite seasonal ensemble streamflow forecasting technique, *Water Resour. Res.*, 46, W03532, doi:10.1029/2009WR007965.

existing operational models while tending to better predict wet years. The forecasts from this approach can be a valuable input for efficient planning and management of water resources in the basin.

1.1 Introduction and Background

The recent protracted dry period (2000-2008) in the Upper Colorado River Basin (UCRB) has had various impacts on basin hydrology and management. For example, Lake Powell has seen its lowest levels since its filling in 1980 (*Brandon, 2005*). In particular the recent drought has emphasized the need for accurate streamflow predictions at longer lead times than usual. Accurate forecasts at several spatial locations are also desirable for efficient basin-wide reservoir management.

Nearly 80% of the streamflow in the UCRB and in many of western US river basins is due to snowmelt and as a result streamflow models have long been dominated by this information. In particular, the April 1st snow water equivalent in the basin is a potent predictor of the subsequent spring snowmelt runoff. However, many important planning decisions are made during the winter (Nov-Mar) requiring a skillful projection of the spring streamflow at a time when the snow information is incomplete. This poses a challenging problem of providing skillful basin-wide streamflow forecasts at longer lead time.

There is increasing evidence that large scale climate features in the Pacific have strong influence on the hydroclimatology of western US including the UCRB; on winter snow (*Clark et al., 2001; Cayan, 1996*); surface temperatures (*Redmond and Koch, 1991; Higgins et al., 2002; Gershunov and Barnett, 1998*) and streamflow (*Kahya and Dracup, 1993, 1994; Dracup and Kahya, 1994; Piechota et al., 1997; Maurer and Lettenmaier, 2004; McCabe and Dettinger, 1999, 2002; Hidalgo and Dracup, 2003; Brown and Comrie, 2004; Hunter et al., 2006; Soukup et al., 2009*). These links enhance the prospects for long lead streamflow forecast. These links were identified and incorporated in a statistical modeling approach to generate skillful forecasts of spring streamflow at long lead times during early winter on several river basins in the western US, e.g., Truckee and Carson river basins (*Grantz*

et al., 2005), Gunnison river basin (*Regonda et al.*, 2006), Columbia River (*Hamlet and Lettenmaier*, 1999; *Clark et al.*, 2001) and the Yakima river basin (*Opitz-Stapleton et al.*, 2007).

The Colorado Basin River Forecast Center (CBRFC) and the Natural Resource Conservation Service (NRCS) are jointly charged with the task of predicting streamflows in the CRB. The current CBRFC models include Statistical Water Supply (SWS) and Ensemble Streamflow Prediction (ESP) (*Brandon*, 2005). The SWS is a regression based method that relates observed data (precipitation, snow water equivalent, monthly flow volume, and climate indices) with future streamflow. The ESP is an empirically based method that includes antecedent streamflow, soil moisture, reservoir information, snowpack states and climate data in forecasting streamflows. The ESP produces ensemble forecasts from historical data based on current conditions thus ensemble size and scope is limited to that of the historical data; for example if 20 years of data are available then only 20 ensemble member can be generated. The NRCS model uses a principle components regression technique that includes snowpack states, fall and spring precipitation, base flow, and climate indices (e.g., Southern Oscillation Index). (*Pagano et al.*, 2004). Together, the CBRFC and NRCS issue a ‘coordinated forecast’ based on their models and modeler knowledge. The coordinated forecast is used in the Bureau of Reclamation’s (BOR) “24-Month Study”. The lack of a rich variety in the ensemble forecasts and consequently the uncertainty estimation are some of the main shortcomings of this approach.

Water resources management in a basin requires streamflow forecasts at several locations on the river network. To address this, *Regonda et al.* (2006) developed a multimodel ensemble streamflow technique that includes a principal component analysis on the basin streamflow. This method performs an orthogonal transformation of the streamflow at multiple locations in the basin. Predictors of the leading principal component in the land-ocean-atmosphere system are identified and a local polynomial based statistical method is used to generate an ensemble forecast of the leading principal component. These are lastly back transformed to obtain ensemble forecasts at multiple locations in the basin. The authors applied this technique for streamflow forecasting at six locations in the Gunnison River basin (a tributary of the Colorado River) and demonstrated

significant skill at longer leads. While this approach is quite good it suffers from two key drawbacks. i. The forecasted flows at the spatial locations do not satisfy the summability criteria, i.e., the flow at any location should be the sum of flows at locations above it. This is important for water budgeting and essential for decision making when these flows are used to drive a decision support system). ii. The streamflow forecasts are for the spring season total, but the decision making process requires streamflow for all the months within the season that satisfy the summability criteria temporally (the monthly forecasts should add up to the seasonal) and spatially (as mentioned in i.).

To overcome the drawbacks of the coordinated forecasts and the multi-site forecasting method of *Regonda et al.* (2006), we propose a framework that integrates the multimodel ensemble forecast technique of *Regonda et al.* (2006) using large scale climate information with a nonparametric space-time disaggregation technique to address the summability. The four components of the proposed framework are: (i) an index gauge streamflow is constructed as the sum of flows at all the desired spatial locations, (ii) potential predictors of the spring season (April-Jul) streamflow at this index gauge are identified from the large scale ocean-atmosphere-land system including snow water equivalent, (iii) multimodel ensemble forecast approach (*Regonda et al.*, 2006) is used to generate the ensemble flow forecast at the index gauge and, (iv) the ensembles are disaggregated using a nonparametric space-time disaggregation technique (*Prairie et al.*, 2007) resulting in forecast ensembles at the desired locations and for all the months within the season. By using an index gauge in this manner we only need to develop predictors for a single “site.” Not only is this approach parsimonious but spatial and temporal correlations are also preserved (*Prairie et al.*, 2007). We demonstrate this framework to forecast monthly streamflow at four key sites in the UCRB. The paper is organized as follows. Preliminary information on the study area and data sets used are first described. Next, the proposed framework is explained, including the algorithmic details of two main components, the multimodel ensemble (MME) forecast and the disaggregation procedure. Results from application to the UCRB are then presented followed by a summary and discussion.

1.2 Study Area

The CRB includes parts of seven states in the western United States with an area of approximately 650,000 square kilometers. The basin includes widely varying topography with elevations ranging from 61 to 4328 meters. Most of the flow in the basin is a result of snowmelt from the UCRB, while most of the water use occurs in the semi-arid and desert regions of Lower Colorado River Basin (LCRB). To demonstrate the proposed forecast framework, we chose four key locations on the UCRB network shown in Figure 1.1, Colorado River near Cisco, Utah (Cisco); Green River at Green River, Utah (GRUT); San Juan River near Bluff, Utah (Bluff) and, Colorado River at Lees Ferry, Arizona (Lees Ferry). Lees Ferry gage, 26 kilometers below Glen Canyon Dam, is the key gauge through which 90% of the Colorado River flow passes through and is approximately a mile upstream of Lee Ferry the official demarcation point of the Upper and Lower Basins for operational and management purposes.

1.3 Data

1.3.1 Streamflow Data and Index Gage

The natural streamflow data for the Colorado River Basin are developed by the Bureau of Reclamation (Reclamation) and updated regularly. Naturalized streamflows are computed by correcting for anthropogenic impacts (i.e., reservoir regulation, consumptive water use, etc.) from the recorded historic flows. *Prairie and Callejo* (2005) present a detailed description of methods and data used for the computation of natural flows in the Colorado River Basin. We used the monthly natural streamflow at the four locations for the period 1949-2005. The flows at the four sites for each month are added to create the “index gauge” monthly streamflow. Almost all of the annual flow at these locations and in the basin occurs during the spring (Apr-Jul) from spring melt of winter snow, much like other river basins in the western US.

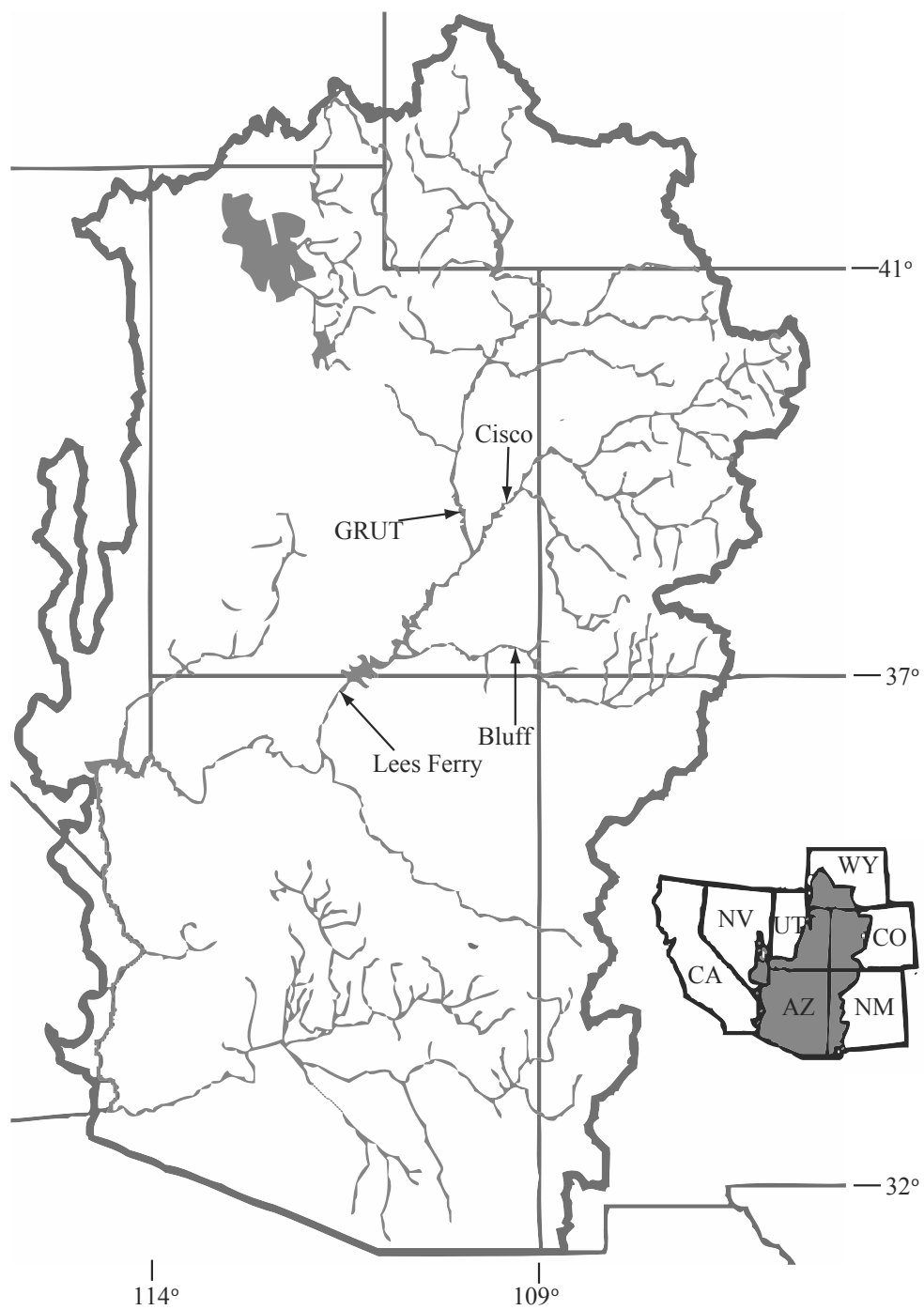


Figure 1.1: Study area.

1.3.2 Large-scale climate data

Ocean-atmospheric circulation variables that capture the large-scale climate forcings are available from NOAA's Climate Diagnostics Center web site . In particular, the variables used were 500 mb geopotential height (GPH), zonal (ZW) and meridional winds (MW) and sea surface temperature (SST). These variables are provided on a 2.2° by 2.2° grid spanning the globe from the NCEP-NCAR reanalysis project (*Kalnay et al.*, 2011) for the period 1949 to present.

1.3.3 SWE data

Snow water equivalent (SWE) data, which quantifies the amount of water present in a snowpack, is obtained from snow course surveys by the Natural Resources Conservation Service (NRCS) from their website . Data was obtained at 10 sites in the UCRB (for the years 1948-2005) for each monthly (February, March and April 1st) SWE. Principal Component Analysis was performed for each month and the leading principal component (which explains most of the variance) was used as the 'basin SWE' for each month. This is similar to using a spatial average but more robust – in this case, we found the leading principal component and the spatial average time series to be highly correlated.

1.3.4 PDSI data

Antecedent summer and fall season land conditions can play an important role in the variability of the following spring streamflow. This was shown in *Regonda et al.* (2006) for the Gunnison River Basin where they found a significant reduction in spring streamflow due to infiltration, relative to the snowpack, in years that succeed a dry summer and fall season and vice-versa. Thus including this in the forecast can improve the skills, especially in such anomalous years. While soil moisture would be the best variable to capture the antecedent land surface conditions, PDSI is shown to be a good surrogate (*Dai and Trenberth*, 2004). Basin averaged PDSI data from October (for the years of 1948–2004) was used as a surrogate for antecedent moisture conditions.

1.4 Proposed Integrated Framework

The proposed framework applies three steps to the index gauge; (i) Predictor identification, (ii) multimodel ensemble forecast (iii) and nonparametric spatial and temporal disaggregation resulting in forecast ensembles at the desired locations and for all the months within the season. The components of each step are described below.

1.4.1 Predictor Suite for Index Gauge seasonal streamflow

The index gauge spring streamflow is correlated with global ocean, atmosphere and land variables (i.e., GPH, SST, ZW, MW and PDSI from preceding seasons). Regions of high correlations are identified for each variable and the spatial average of data points in these regions are computed to create a suite of potential predictors. The on-line tool developed by Physical Sciences Division, National Oceanic and Atmospheric Administration (NOAA) is used for this purpose.

1.4.2 Multimodel Selection for I Gauge Seasonal Flow Forecast

The MME methodology consists of two distinct steps; (i) selection of the multi models each with its own subset of predictors and (ii) combining forecasts from the individual models. The general form of the forecast model is:

$$Y = f(\mathbf{x}) + \varepsilon \quad (1.1)$$

where Y is the index gauge spring season streamflow, $\mathbf{x}$ is a suite of predictors and ε is the residual that is assumed to be Normally distributed with mean 0 and variance $\varepsilon \sim N(0, \sigma)$.

If f is a global function based on the entire data and linear, then the model is a traditional linear regression. The theory behind this approach, procedures for parameter estimation and hypothesis testing are very well developed (*Helsel and Hirsch, 1995; Rao and Toutenburg, 1999*) and widely used. However, they do have some drawbacks including, (i) the assumption of a Normal distribution of the errors and the variables and (ii) fitting a global relationship (e.g., a linear equation in the case of linear regression) between the variables. If the linear model is found inadequate,

higher order models (quadratic, cubic, etc.) have to be considered, which can be difficult to fit in the case of short data. Also if the variables are not Normally distributed, which is often the case in practice, suitable transformations have to be obtained to transform them to a Normal distribution. All of this can make the process unwieldy. Thus, a more flexible framework would be desirable.

Local estimation methods (also known as nonparametric methods) provide an attractive alternative. There are several approaches for local functional estimation applied to hydrologic problems, see (*Lall, 1995*). Of these, the Locally Weighted Polynomial regression (LWP) is simple and robust, and has been used in a variety of hydrologic and hydroclimate applications with good results, e.g., for streamflow forecasting on the Truckee and Carson river basins (*Grantz et al., 2005*), salinity modeling on the Upper Colorado River basin (*Prairie, 2006*), forecasting of Thailand summer rainfall (*Singhrattana et al., 2005*), and spatial interpolation of rainfall in a watershed model (*Hwang, 2005*). Given these experiences, we adopt the LWP method in this research as *Regonda et al. (2006)* also did.

The ‘best model’ in this form is described by the size of the neighborhood, the order of the polynomial and the subset of predictor variables. This is identified using objective criteria. A brief description of the implementation steps is presented below, which is largely abstracted from *Regonda et al. (2006)*. For details on local polynomial estimation we refer the reader to Loader, 1999.

- (1) Select a subset of predictors
- (2) Select a degree of polynomial to fit, typically, $p=1$ (linear fit) or 2 (quadratic) is quite adequate.
- (3) Select a size of the neighborhood, $K = \alpha N$ (where $\alpha = (0,1)$ and N =number of observations).
- (4) For a point x identify the K nearest neighbors in the data; Euclidean distance is typically used, other metrics can be used such as Mahalanobis distance (*Yates et al., 2003*).

- (5) Weighted least squares estimation is employed to fit a polynomial of order p to the K nearest neighbors. Use the fit to obtain the estimate, $\hat{Y}$.
- (6) Repeat steps 3-4 for all the data points
- (7) Compute the objective criteria, Generalized Cross Validation Estimation (GCV)

$$\text{GCV}(K, p) = \frac{\sum_{i=1}^N \frac{e_i}{N}}{\left(1 - \frac{q}{N}\right)^2} \quad (1.2)$$

Where e_i is the model residual ($Y_i - \hat{Y}$) for the i th data point, N is the number of data points, and q is the number of parameters in the local polynomial model. Generalized Cross Validation (GCV) provides a good estimate of predictive risk of the model, unlike other statistics, which are goodness of fit measures (*Craven and Wahba, 1979*).

- (8) Repeat steps 1 through 7, thus, obtaining the GCV values for a suite of predictor subset, K and p combination. The model with the least GCV score is selected as the 'best' model. This is akin to the stepwise regression method in a traditional linear regression context (*Rao and Toutenburg, 1999; Walpole et al., 2002*) where in, an objective function such as Mallows's Cp statistic, adjusted R², AIC (Akaike Information Criteria), or an F-test is calculated from the fitted model for several predictor combinations. The GCV based model selection is preferable since it provides a better estimate of predictive risk as mentioned above.
- (9) All combinations with GCV values within a prescribed threshold (user specified as typically within 5% of the least GCV value) are selected as admissible constituting the pool of candidate models (i.e., multimodels). Recent studies show that multimodel ensemble forecasts which include multiple 'best' models tend to perform much better than a single model forecast (*Hagedorn et al., 2005; Krishnamurti et al., 1999, 2000; Rajagopalan et al., 2002*). Combinations with predictor variables significantly correlated amongst each other (i.e., multicollinear) are removed from the multimodel pool, as they lead to over fitting and

poor predictive skill (Wadsworth, 1990), thus resulting in a set of multimodels each with different predictor subsets.

Note that if $\alpha = p = 1$ and if an ordinary least squares estimation is used to fit the model, it collapses to a traditional linear regression. Therefore, LWP can be viewed as a more flexible approach that includes traditional linear regression as a subset.

1.4.3 Multimodel Ensemble Forecast of Index Gage Seasonal Streamflow

The multimodel ensemble forecast generation is described below, again following *Regonda et al.* (2006). Suppose we desire a multimodel ensemble forecast for a point $\mathbf{x}_j$, the steps are as follows.

- (1) From one of the multimodels identified in the preceding section, obtain the model prediction and the estimate of the error variance () (Loader, 1999). Ensembles z are then generated by sampling a given number (we chose 250) of random Normal Deviates and adding these to the prediction $\hat{Y}_i$:

$$z_{i,j} = \hat{Y}_i(\mathbf{x}_j) + N(0, \sigma_{e_j}) \quad (1.3)$$

where $N(0, \sigma_{e_j})$ is a normally distributed random variable with mean 0 and standard deviation σ_{e_j} . This approach assumes Normality of the residuals, a standard assumption from regression theory. For the application here we found the residuals to satisfy this assumption of Normality.

- (2) Repeat step 1 for all the models in the multimodel suite. Thus, obtaining an ensemble forecast of size n from each model.
- (3) Weight each of the multimodels based on $1/\text{GCV}$ criteria. This way the model with the lowest GCV value is most heavily favored.
- (4) Randomly choose a model based on the above set of weights.

- (5) Randomly choose one of the n ensemble members.
- (6) Repeat steps 4-5 n times to obtain a multimodel ensemble forecast.

1.4.4 Spatial and Temporal Disaggregation of the Index Gauge Forecast

The seasonal forecast of the index gauge generated above needs to be disaggregated in space (to the four locations) and in time (to the four months of the season) resulting in an ensemble forecast for each month at all the four locations. This is achieved by employing a nonparametric space-time disaggregation technique proposed by *Prairie et al.* (2007) to the seasonal forecast. All the motivational and technical details of the disaggregation method are comprehensively described in *Prairie et al.* (2007). Below we provide a brief description of its implementation in the current application.

The disaggregation procedure can be thought of as sampling from the conditional probability density function (PDF), $f(\mathbf{x}|\mathbf{z})$, where $\mathbf{x}$ is a d dimensional vector of flows and $\mathbf{z}$ is the aggregate flow, but with the (additivity) constraint that the values in $\mathbf{x}$ add up to $\mathbf{z}$. This is achieved by an orthonormal rotation of the data $\mathbf{x}$ to $\mathbf{Y}$ and the simulation is performed in the rotated space and back rotated (*Tarboton and Sharma, 1998; Prairie et al., 2007*). The steps are described below for a temporal disaggregation (seasonal to monthly at the index gauge), they are identical for the spatial disaggregation.

- (1) The first step is to generate the orthonormal rotation matrix, $\mathbf{R}(d)$ using the Gram-Schmidt algorithm. Note that the $d \times d$ matrix $\mathbf{R}$ is only a function of the dimension d and has the property $\mathbf{R}^T = \mathbf{R}^{-1}$ by definition, where T denotes transpose. This process is described in detail in the appendix of *Tarboton and Sharma (1998)*, the reader is referred there for the details.
- (2) The matrix $\mathbf{X}$ of the historical monthly streamflow at gauge I is rotated to $\mathbf{Y}$ by the rotation matrix $\mathbf{R}$ as,

$$\mathbf{Y} = \mathbf{RX}. \quad (1.4)$$

Both $\mathbf{X}$ and $\mathbf{Y}$ are of dimension N (number of years) rows by $d(=4$ months in the season) columns.

The development of the $\mathbf{R}$ matrix is detailed in the appendix of Tarboton et al. (1998). The rotated matrix $\mathbf{Y}$ has its last column $\mathbf{y}_d = \mathbf{z} / \sqrt{d} = \mathbf{z}'$, where $\mathbf{z}$ is the vector of the aggregate flows (i.e., the seasonal totals). If we denote the first $d - 1$ columns as $\mathbf{U}$ then

$$\mathbf{Y} = [\mathbf{U}, \mathbf{z}'] \quad (1.5)$$

- (3) For a given seasonal flow forecast say z_{sim} , K nearest neighbors are identified from the historical seasonal flow values in $\mathbf{z}$. One of the K nearest neighbor is resampled using a weight function

$$W(k) = \frac{1}{k \sum_{i=1}^K \frac{1}{i}} \quad (1.6)$$

where $k = 1, 2, \dots, K$. This weight function gives more weight to the nearest neighbors and less to the farthest neighbors. For further discussion on the choice of the weight function readers are referred to *Lall and Sharma* (1996). The number of nearest neighbors, K is based on the heuristic scheme $K = \sqrt{N}$ where N equals the sample size (*Lall and Sharma*, 1996), following the asymptotic arguments of *Fukunaga* (1990).

- (4) Suppose the selected neighbor corresponds to historic year j , the new vector $\mathbf{y}_{sim}$ is constructed as

$$\mathbf{y}_{sim} = \left[u_j, z_{sim} / \sqrt{d} \right]. \quad (1.7)$$

- (5) This is back transformed to the original space as

$$\mathbf{x}_{sim} = \mathbf{R}^T \mathbf{y}_{sim}. \quad (1.8)$$

The vector $\mathbf{x}_{sim}$ now contains $d = 4$ values, the disaggregated monthly flow of the seasonal total flow, $\mathbf{z}_{sim}$.

The disaggregated monthly flows are subjected to the same procedure to obtain monthly flows at the four locations. This ensures the additivity criteria, for each month the flows at the index gauge are a sum of the flows at the four locations. Applying this to all the multimodel ensemble members results in a multimodel ensemble forecast for each month at the four locations. We also note that the sampling the residuals from a Normal distribution coupled with the disaggregation method, provide the ability to generate values outside the range of the historical observations this was shown in *Regonda et al.* (2006) and *Prairie et al.* (2007).

1.4.5 Verification and Validation

Since we generate an ensemble forecast (i.e., a PDF), the skill of the forecast needs to be evaluated in probabilistic terms. One such common measure is the Ranked Probability Skill Scores (RPSS) (*Wilks*, 1995). It measures the accuracy of multi-category probability forecasts relative to a climatological forecast. Typically, the flows are divided into k mutually exclusive and collectively exhaustive categories for which the proportion of ensembles falling in each category constitutes the forecast probabilities $(p_1, p_2, \dots, p_k)$. The observational vector $(d_1, d_2, \dots, d_k)$ is obtained for each forecast, where d_k is unity if the observation falls in the k th category and zero otherwise. The ranked probability skill score (RPSS) is defined as follows:

$$RPS = \sum_{j=1}^k \left[\left(\sum_{s=1}^j p_s - \sum_{s=1}^j d_s \right)^2 \right] \quad (1.9)$$

$$RPSS = 1 - \frac{RPS(\text{forecast})}{RPS(\text{climatology})}. \quad (1.10)$$

In this research, the streamflows are divided into three categories, at the tercile boundaries, 33rd and 66th percentile of the historical observations. Values below the 33rd percentile represent 'dry', above 66th percentile 'wet', and 'near normal' otherwise. Of course, the climatological

forecast for each of the tercile categories is $1/3$.

The RPSS ranges from negative infinity to positive unity. Negative RPSS values indicate the forecast accuracy to be worse than climatology, positive to be higher than climatology, zero to be equal to that of climatology, and a perfect categorical forecast yields an RPSS value of unity. In this application the RPSS is calculated for each year and the median value is reported. Correlation between the median value of the ensemble and the observed (MC) are also computed to test the performance of the median forecast. The forecast skills were computed for three types of forecasts:

- (1) Leave-one out forecast. In this each year is dropped and the multi model ensemble for that year is generated from the rest, e.g., (*Grantz et al.*, 2005; *Regonda et al.*, 2006; *Singhrattna et al.*, 2005).
- (2) Leaving one year out at a time may not stress the model adequately. Here we drop 10% of the observations and forecasts for the dropped years are made using the rest of the observations. The skill scores are computed for the forecasts. This is repeated a number of times, obtaining an ensemble of skill scores. This validation method provides a mechanism for sensitivity analysis.
- (3) To be able to compare with forecasts from CBRFC, 'retroactive' forecasts have to be performed. In this, a forecast for a given year is made using all the data prior to that year imitating a real time situation.

1.5 Results

We chose four lead times, Nov 1st, Jan 1st, Feb 1st and Apr 1st, to predict the spring streamflows. Separate predictors are identified for each lead time and the above methodology is applied to obtain the ensemble forecasts. Next, we present a representative set of results.

1.5.1 Predictor Identification

The index gauge spring season streamflow was correlated with large-scale ocean, atmosphere and land variables from preceding months. The combination of months that showed strong correlation are Feb-Mar GPH, Mar ZNW, Jan-Mar SST and Mar MDW (see Appendix A). The correlation maps are consistent with prior findings for the river flows in western US (*Grantz et al.*, 2005; *Regonda et al.*, 2006, for example). The regions with high correlations were identified and a spatially averaged time series was computed as potential predictors, these are detailed in Table 1.1. For SWE the leading principal component (similar to a basin average) was used, this is available from Feb 1st.

Table 1.1: Potential predictors and their regions in lon:lat

Variable	Lead Time	Season	Negative Region	Positive Region	Max Cor.
GPH	Nov. 1st	oct	40,42N:137,140E	17,19N:178,182E	0.39
ZNW	Nov. 1st	oct	40,42N:-108,-106E	30,32N:132,135E	0.43
MDW	Nov. 1st	oct	40,42N:124,126E	40,42N:124,126E	0.31
SST	Nov. 1st	oct	-17,-15N:-96,-92E	38,40N:170,173E	-0.26
GPH	Jan. 1st	oct-dec	42,45N:-115,-110E	52,58N:-160,-155E	-0.56
ZNW	Jan. 1st	oct-dec	54,55.5N:-119,-114E	32,34N:-110,-106E	-0.58
MDW	Jan. 1st	oct-dec	45,50N:-135,-130E	42,45N:-91,-88E	-0.57
SST	Jan. 1st	oct-dec	28,31N:161,164E	46,49N:-166,-163E	-0.63
GPH	Feb. 1st	oct-jan	60,62N:128,132E	61,64N:-153,-150E	0.24
ZNW	Feb. 1st	oct-jan	72,74N:126,128E	28,31N:-119,-116E	0.42
MDW	Feb. 1st	oct-jan	66,68N:-108,-105E	38,42N:-90,-85E	0.30
SST	Feb. 1st	oct-jan	25,28N:160,163E	46,48N:-167,-164E	-0.64
GPH	Apr. 1st	feb-mar	40,42N:137,140E	65,70N:160,165E	-0.47
ZNW	Apr. 1st	mar	56,58N:172,176E	25,30N:155,160E	0.46
MDW	Apr. 1st	mar	10,12N:-177,-174E	56,60N:142,144E	-0.42
SST	Apr. 1st	jan-mar	24.5,26N:158,162E	48,51N:-165,-160E	-0.54

To understand the physical relationship between the large scale variables and the spring streamflow, composite maps of the 6 wettest and driest years are developed (see auxiliary materials). During wet years the anomalous wind propagation during the winter in the basin is from the ocean and from a southwest direction. This brings moisture resulting in more snowfall in the basin and consequently, more spring streamflow. In the dry years it is the opposite, i.e., the

basin experiences northerly dry winds; hence, less snow and consequently low spring streamflow. These composite map features are consistent with the correlation maps. The land conditions from antecedent fall season also plays a role in the spring streamflow. Wetter conditions in the fall favor less infiltration during the spring snow melt enabling enhanced streamflow. This can be seen in the PDSI composite and vice-versa during dry years.

1.5.2 Multi Model Selection

Using the predictors identified in Table 1 and the methodology described in the previous section, multimodels were selected for different lead times and listed in Table 2. It is interesting to note that the number of multimodels decreases with lead time. This is intuitive, in that on Jan 1st SWE is not available and the forecasts have to be made only from climate information. Consequently, individual models have greater uncertainty, and more models qualify as candidates for the multimodel pool. However, on April 1st, SWE information is complete and the best predictor of the ensuing spring streamflow; therefore, fewer models with other predictor variables are needed, necessitating a smaller number of multimodels. Similar observations were made by *Regonda et al.* (2006).

Table 1.2: Selected multimodels for each lead time. “1” indicates the presence of a predictor and “0” indicates the absence of a predictor. A selected predictor refers to the most recent data available at the forecast time (eg. May GPH for an Apr 1st forecast). *This Jan. 1st prediction used the Apr. 1st and Nov. 1st GPH as separate predictors.

Lead Time	Predictors	PDSI	GPH	ZNW	MDW	SST	SWE
Apr. 1st	2	1	0	0	0	0	1
Feb. 1st	2	0	0	0	0	1	1
Feb. 1st	2	1	0	0	0	0	1
Feb. 1st	3	0	0	0	1	1	1
Jan. 1st*	3	0	1	1	0	0	0
Jan. 1st	2	0	1	1	0	0	0
Nov. 1st	2	1	1	0	0	0	0
Nov. 1st	1	1	0	0	0	0	0

1.5.3 Forecast Skill

The leave-one out cross validated ensemble forecast of the index gauge spring season streamflow issued on April 1st and Jan 1st are shown as boxplots in Figure 1.2. The box height corresponds to the interquartile range, the whiskers depict the 5th and 95th percentiles and the horizontal line is the median. The true observations for each year are joined by a line. The dashed horizontal lines correspond to the 33rd, 50th and 66th percentiles of the data. The forecast ensembles capture the true values very well at both lead times, as attested by high RPSS scores. The high skill of the Jan 1st forecast is noteworthy as this forecast is made entirely from climate information. These forecasts were disaggregated to ensemble forecasts at the four locations, Figure 1.3 shows the ensemble forecast for Lees Ferry. We find that the forecasts capture the observed flows quite well and the high skill in forecasting the index gauge (Figure 1.2) is translated to the disaggregated forecasts. The results were similar at other locations. The skill scores of the seasonal forecast at different locations and lead times are shown in Table 1.3. The skill at the Bluff location is low in comparison to the rest. This is due to the fact that flows at Bluff are quite small compared to the other three locations and the disaggregation method tend to under perform in such situations also seen by *Prairie et al.* (2007). One alternative is to perform the disaggregation in two steps where the Bluff flows are generated in the second step. Temporal disaggregation of the seasonal forecasts to monthly forecasts was also skillful, this can be seen in the boxplot of May monthly ensemble forecast at Lees Ferry in Figure 1.4. The noteworthy aspect is that the skills are retained through the spatial and temporal disaggregation process. Furthermore, the forecasts are significantly skillful at longer lead times; therefore, water managers can obtain a good idea of the spring streamflow months in advance which will be useful for efficient water resources management. As to be expected forecast skills decrease with increase in lead time but improve upon climatology, nonetheless.

To challenge the forecasting system, we performed the forecasts by dropping 10% of the observations and repeating the forecast 100 times. The RPSS skills on the dropped observations for

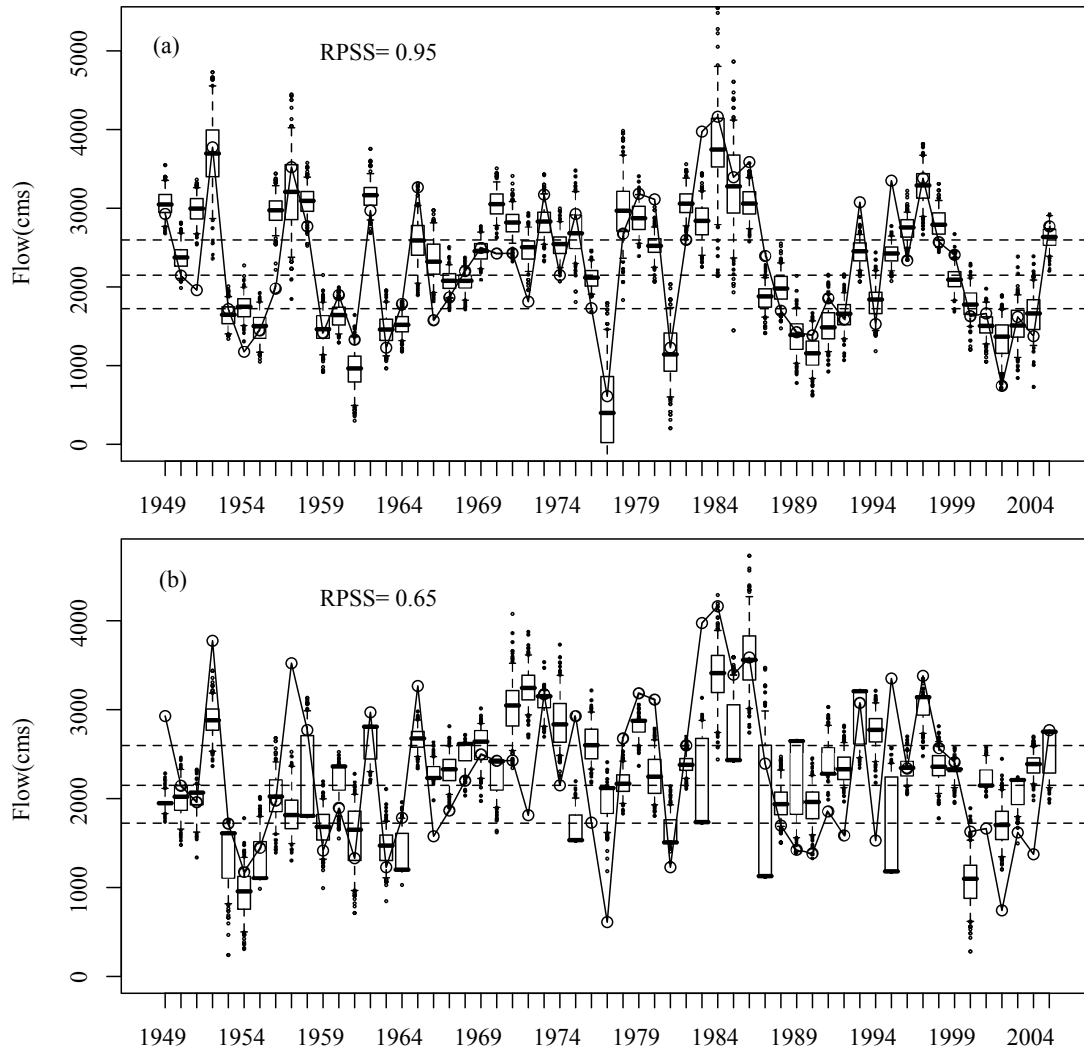


Figure 1.2: Leave-one CV forecasts of index gauge spring streamflow issued on (a) Apr 1st and (b) Jan 1st. The box corresponds to the interquartile range, the horizontal line inside the boxes is the median, whiskers extend to 5th and 95th percentile of ensemble and solid line joins the true value for each year. The dashed horizontal lines correspond to the 33rd, 50th and 66th percentiles of the data.

the seasonal flow forecast issued on Apr 1st are shown as boxplots in Figure 1.5. There is considerable variability in the skill scores due to sampling, but the median skill scores are quite high at all the locations except Bluff for the reasons mentioned above. Similar results were observed for the monthly forecasts.

One of the advantages of the disaggregation method is that it can capture the spatial cor-

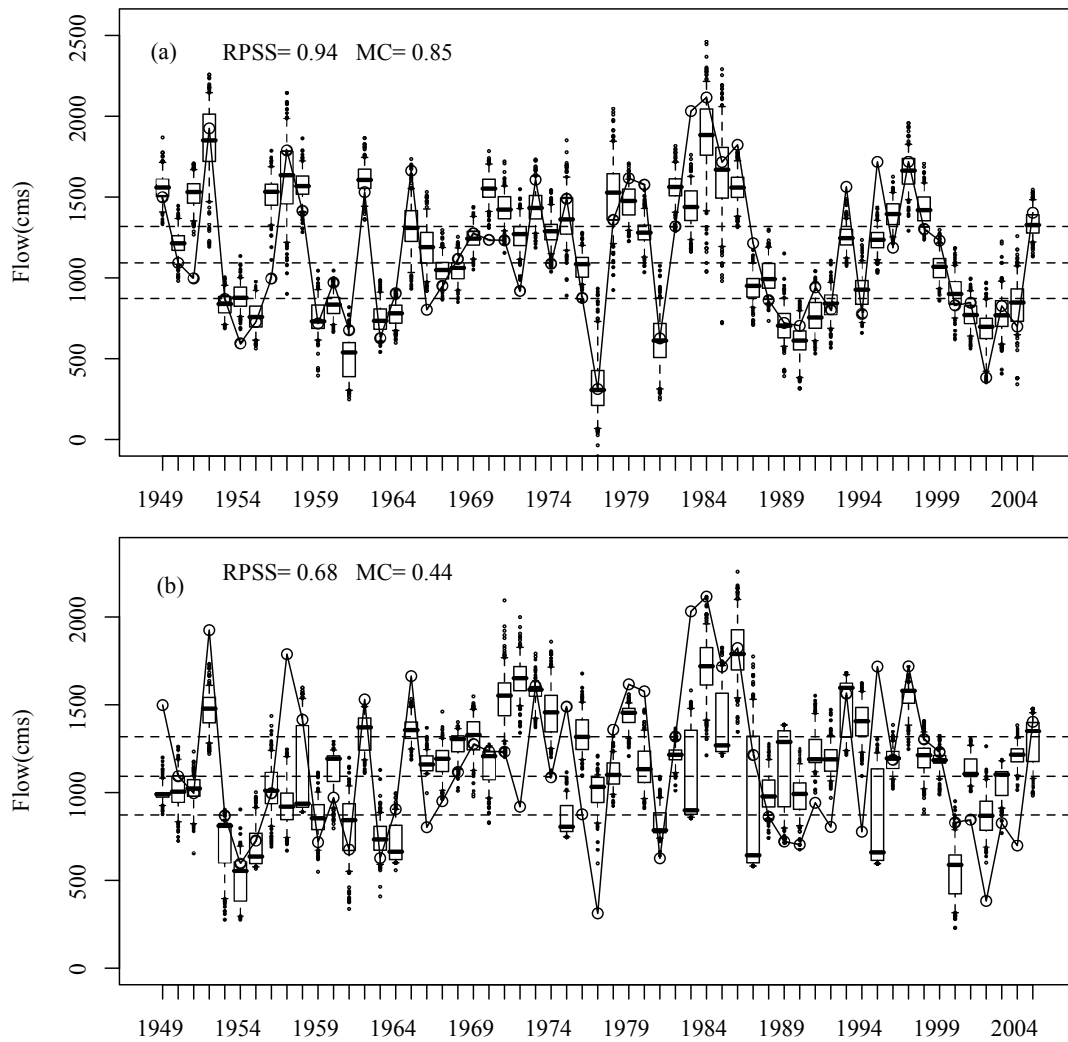


Figure 1.3: (a) same as Figure 1.2(a) but for Lees Ferry. (b) Same as Figure 1.2(b) but for Lees Ferry.

relation in a parsimonious manner. We evaluate this by comparing the ensemble leave-one out cross validated forecasted flows with the historical values. The spatial correlations are very well captured (see auxiliary materials).

In practice forecasts are made one year at a time using all the prior data, i.e., a retroactive forecast. Creating a retroactive forecast enables us to compare with coordinated forecasts and the ESP forecasts, which are issued in this manner. We performed retroactive forecasts for the period 1990-2005 for which coordinated forecasts are available for comparison. Figure 1.6(a), (b) shows

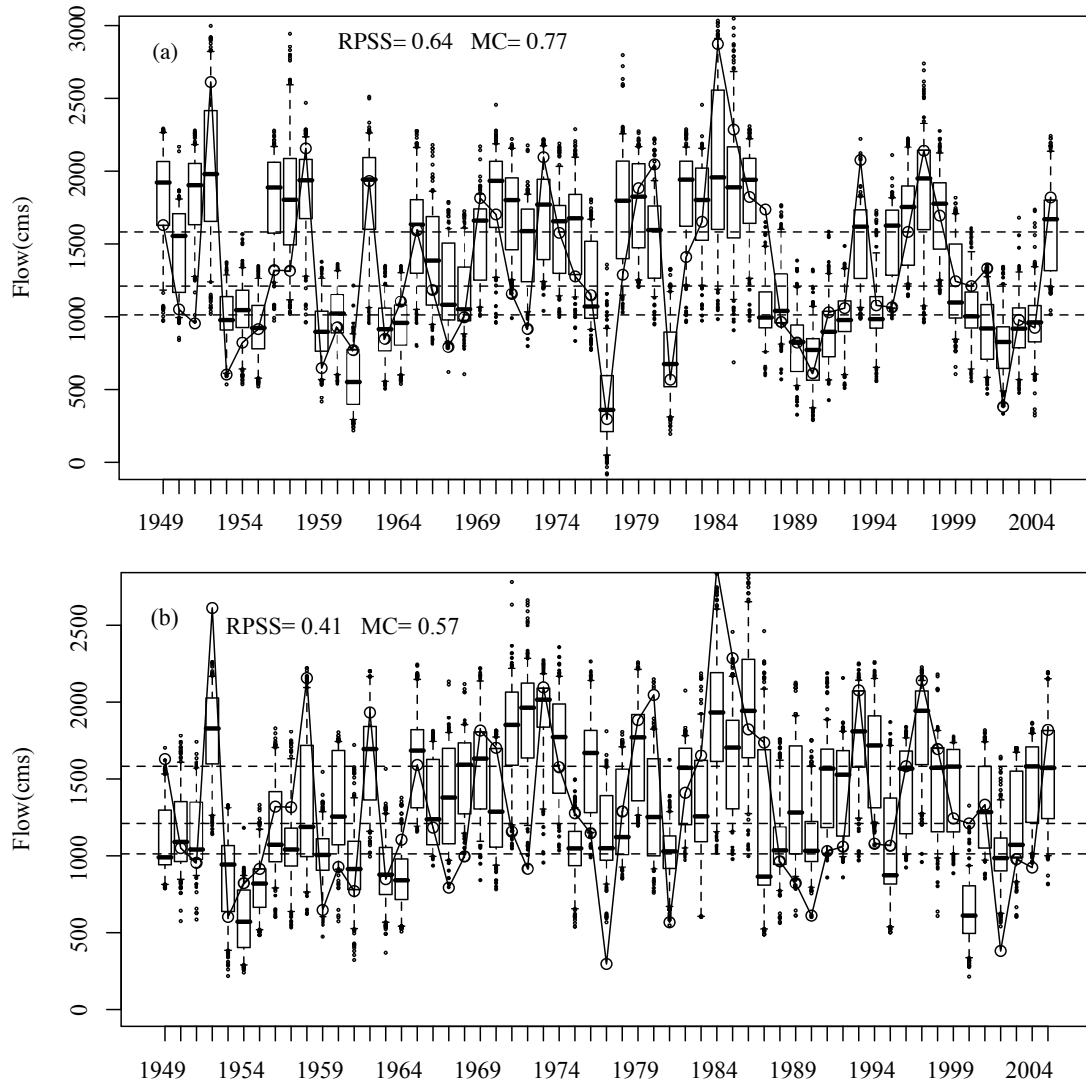


Figure 1.4: (a) same as Figure 1.3(a) but for May streamflow at Lees Ferry. (b) same as Figure 1.3(b) but for May streamflow.

the forecasts of Lees Ferry seasonal streamflow issued on Jan 1st and Apr 1st, respectively, Figure 1.6(c) shows the forecasts of May streamflow at Lees Ferry issued on Apr 1st. In all these, the forecast from our approach is comparable to the coordinated forecasts. The correlation between the historical flow and the Coordinated forecast was computed for the years available and compared to the median ensemble forecast correlation for the corresponding year Table 1.4. In wet years our approach tends to be better but does not capture extreme dry years such as 2002 as well

Table 1.3: Skills of spring season streamflow forecast at different lead times.

Validation mode	Site	Apr. 1st RPSS	Feb. 1st RPSS	Jan. 1st RPSS	Nov. 1st RPSS	Apr. 1st MC	Feb. 1st MC	Jan. 1st MC	Nov. 1st MC
Leave-one	Index	0.95	0.85	0.65	0.28	-	-	-	-
Leave-one	Cisco	0.37	0.58	0.70	0.34	0.56	0.75	0.44	0.42
Leave-one	GRUT	0.63	0.49	0.12	0.21	0.76	0.77	0.44	0.37
Leave-one	Bluff	0.73	0.06	0.40	0.19	0.78	0.53	0.41	0.26
Leave-one	Lees Ferry	0.68	0.71	0.68	0.31	0.63	0.79	0.44	0.42
Retroactive	Index	0.80	0.85	0.29	0.10	-	-	-	-
Retroactive	Cisco	0.97	0.73	0.62	0.78	0.81	0.72	0.31	0.52
Retroactive	GRUT	0.54	0.46	0.17	0.58	0.92	0.79	0.13	0.48
Retroactive	Bluff	0.37	0.40	0.15	0.15	0.41	0.58	0.52	0.35
Retroactive	Lees Ferry	0.87	0.73	0.62	0.60	0.88	0.77	0.33	0.57

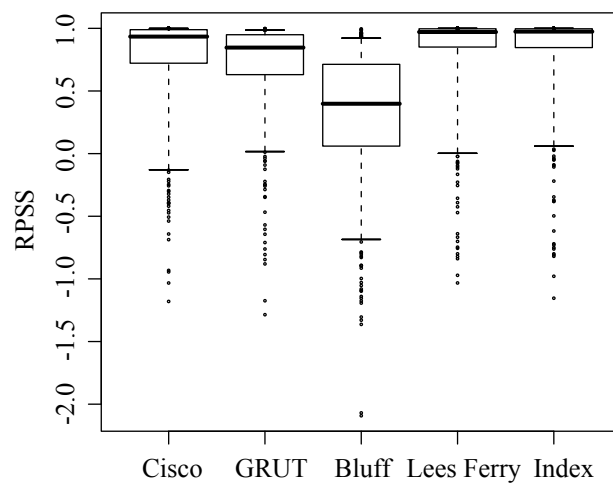


Figure 1.5: Boxplot of RPSS of spring streamflow forecast at the four locations and the index gauge. Forecasts are based on dropping 10% of the observations.

as the coordinated forecast. In dry years land surface conditions play a substantial role, which is better captured by the physical models hence, increased skill in the CF relative to the statistical approach. This is corroborated in *Regonda et al.* (2006). Note that though we provide these comparisons, systematic comparisons need to be made with forecasts from all the techniques. We also suggest a multimodel ensemble combination approach to objectively blend these forecasts.

Table 1.4: Model correlations with historical flow for available Coordinated forecast years

Figure	Coordinated Median	Ensemble
6(a)	0.404	0.571
6(b)	0.955	0.962
6(c)	0.846	0.962

1.6 Summary

We present a parsimonious framework to provide seasonal ensemble streamflow forecasts at several locations simultaneously in a river network that preserves summability. The framework is an integration of two recent approaches; (i) a nonparametric multimodel ensemble forecast technique (*Regonda et al., 2006*) and (ii) a nonparametric space-time disaggregation technique (*Prairie et al., 2007*). The approach generates an ensemble forecast for the seasonal streamflow of an ‘index gauge’, which is constructed as the sum of all the spatial flows, using large scale ocean-land-atmospheric features and a multimodel combination method. These forecasts are then disaggregated in space and time resulting in an ensemble forecast for all the months and at all the locations, thereby capturing the temporal and spatial correlations.

We demonstrate the framework by generating streamflow forecast at four locations in the Upper Colorado River basin. The noteworthy finding is that the high skill annual forecasts at the index gauge are translated to monthly, multi-location forecasts; achieving this was not obvious when we embarked on this study. The forecasts also showed significant skill at longer lead times; thus, providing crucial advance knowledge of the spring streamflow in the river basin enabling efficient planning and management. Furthermore, the skills from this simpler framework are comparable to the coordinated predictions that are currently used. The combination of these two forecasts is the logical next step to improve the forecast skills further.

Acknowledgements

The first author is thankful for a summer fellowship in 2007 funded by the NSF REU program that was instrumental in this research. The authors would like to thank three anonymous reviewers and the editors for their helpful and insightful comments in improving the manuscript.

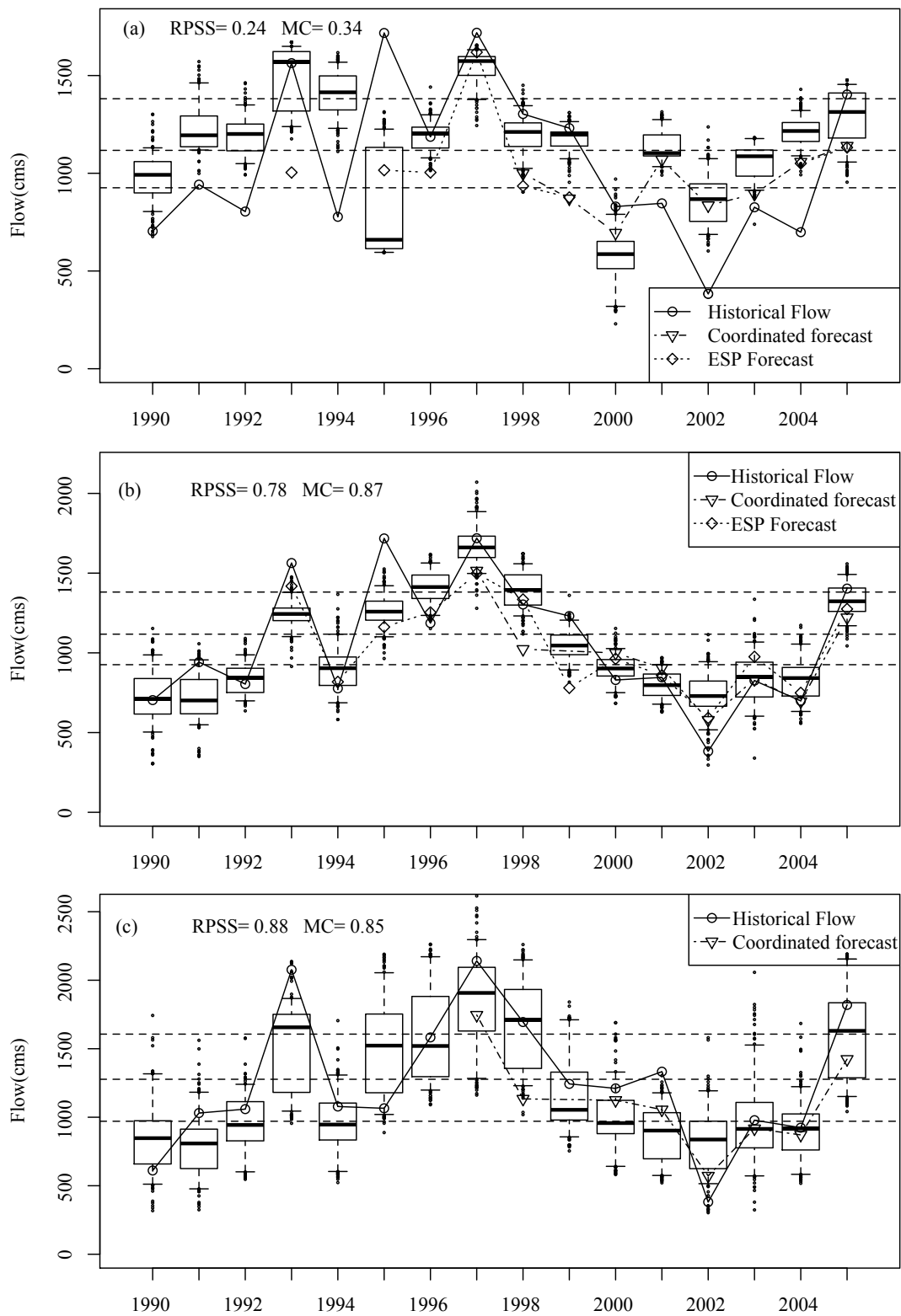


Figure 1.6: Spring flow forecast at Lees Ferry issued on (a) Jan. 1st and (b) Apr. 1st in a retroactive mode. Also, May flow forecast at Lees Ferry (c) after spatial and temporal disaggregation, issued on Apr. 1st. The dashed line represents the coordinated forecast and the dotted line represents the ESP forecast.

Bibliography

- Brandon, D. G. (2005), Using NWSRFS ESP for Making Early Outlooks of Seasonal Runoff Volumes into Lake Powell, in *Hydrology of Arid and Semi-Arid Regions*.
- Brown, D., and A. Comrie (2004), A winter precipitation 'dipole' in the western United States associated with multidecadal ENSO variability, *Geophys. Res. Lett.*, 31(9), –.
- Cayan, D. R. (1996), Interannual climate variability and snowpack in the western United States, *Journal of Climate*, 9(5), 928–948.
- Clark, M. P., M. C. Serreze, and G. J. McCabe (2001), Historical effects of El Nino and La Nina events on the seasonal evolution of the montane snowpack in the Columbia and Colorado River Basins, *Water Resources Research*, 37(3), 741–757.
- Craven, P., and G. Wahba (1979), Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation, *Numerical Mathematics*, 31, 377–403.
- Dai, A., and K. Trenberth (2004), A global dataset of Palmer Drought Severity Index for 1870-2002: Relationship with soil moisture and effects of surface warming, *Journal of Hydrometeorology*.
- Dracup, J. A., and E. Kahya (1994), The relationships between U.S. streamflow and La Niña Events, *Water Resources Research*, 30(7), 2133.
- Fukunaga, K. (1990), Introduction to statistical pattern recognition, Academic Press.
- Gershunov, A., and T. P. Barnett (1998), ENSO Influence on Intraseasonal Extreme Rainfall and Temperature Frequencies in the Contiguous United States: Observations and Model Results, [http://dx.doi.org/10.1175/1520-0442\(1998\)011;1575:EIOIER;2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1998)011;1575:EIOIER;2.0.CO;2).
- Grantz, K., B. Rajagopalan, M. Clark, and E. Zagana (2005), A Technique for Incorporating Large-Scale Climate Information in Basin-Scale Ensemble Streamflow Forecasts, *Water Resources Research*, p. W10410.
- Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting. Part I: Basic concept, *Tellus, Ser. A*, 57, 219–233.
- Hamlet, A., and D. Lettenmaier (1999), Columbia River streamflow forecasting based on ENSO and PDO climate signals, *Journal of . . .*
- Helsel, D., and R. Hirsch (1995), Statistical methods in water resources, Elsevier, New York.

- Hidalgo, H., and J. Dracup (2003), ENSO and PDO effects on hydroclimatic variations of the Upper Colorado River basin, *Journal of Hydrometeorology*, 4(1), 5–23.
- Higgins, R., A. Leetmaa, and V. Kousky (2002), Relationships between climate variability and winter temperature extremes in the United States, *Journal of Climate*, 15(13), 1555–1572.
- Hunter, T., G. Tootle, and T. Piechota (2006), Oceanic-atmospheric variability and western U.S. snowfall, *Geophys. Res. Lett.*, 33(13).
- Hwang, Y. (2005), Impact of input uncertainty in ensemble streamflow generation, Ph.D. thesis, University of Colorado at Boulder.
- Kahya, E., and J. Dracup (1993), US streamflow patterns in relation to the El Nino/southern oscillation, *Water Resources Research*, 29, 2491–2503.
- Kahya, E., and J. Dracup (1994), The influences of type 1 El Nino and La Nina events on streamflows in the Pacific southwest of the United States, *Journal of Climate*.
- Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph (2011), The NCEP/NCAR 40-Year Reanalysis Project, *Bull. Amer. Meteor. Soc.*, 77(3), 437–471, doi: 10.1175/1520-0477(1996)077<0437:TNYP;2.0.CO;2.
- Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran (1999), Improved weather and seasonal climate forecasts from multi-model superensemble, *Science*, 285, 1548–1550.
- Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. E. LaRow, D. R. Bachiochi, C. E. Williford, S. Gadgil, and S. Surendran (2000), Multimodel ensemble forecasts for weather and seasonal climate, *Journal of Climatology*, 13, 4196–4216.
- Lall, U. (1995), Nonparametric Function Estimation: Recent Hydrologic Contributions, *Reviews of Geophysics*, pp. 1093–1099.
- Lall, U., and A. Sharma (1996), A nearest neighbor bootstrap for resampling hydrologic time series, *Water Resources Research*, 32(3), 679–693.
- Loader, C. (1999), *Local Regression and Likelihood*, Statistics and Computing, Springer, New York.
- Maurer, E., and D. Lettenmaier (2004), Variability and potential sources of predictability of North American runoff, *Water Resources Research*.
- McCabe, G., and M. Dettinger (1999), Decadal variations in the strength of ENSO teleconnections with precipitation in the western United States, *International Journal of Climatology*.
- McCabe, G., and M. Dettinger (2002), Primary modes and predictability of year-to-year snowpack variations in the western United States from teleconnections with Pacific Ocean climate, *Journal of Hydrometeorology*.
- Opitz-Stapleton, S., S. Gangopadhyay, and B. Rajagopalan (2007), Generating streamflow forecasts for the Yakima River Basin using large-scale climate predictors, *Journal of Hydrology*, 341, 131–143.

- Pagano, T., D. Garen, and S. Sorooshian (2004), Evaluation of official western US seasonal water supply outlooks, 1922-2002, *Journal of Hydrometeorology*, 5, 896–908.
- Piechota, T., J. Dracup, and R. Fovell (1997), Western US streamflow and atmospheric circulation patterns during El Nino Southern Oscillation, *Journal of Hydrology*, 201, 249–271.
- Prairie, J. (2006), Stochastic nonparametric framework for basin wide streamflow and salinity modeling: Application for the Colorado River Basin, Ph.D. thesis, University of Colorado at Boulder.
- Prairie, J., and R. Callejo (2005), Natural Flow and Salt Computation Methods, Calendar Years 1971-1995, *Bureau of Reclamation*, pp. 1–112.
- Prairie, J., B. Rajagopalan, U. Lall, and T. Fulp (2007), A stochastic nonparametric technique for space-time disaggregation of streamflows, *Water Resources Research*, 43, 1–10.
- Rajagopalan, B., U. Lall, and S. Zebiak (2002), Optimal categorical climate forecasts through multiple GCM ensemble combination and regularization, *Monthly Weather Review*, 130, 1792–1811.
- Rao, C. R., and H. Toutenburg (1999), Linear models: least squares and alternatives, Springer Verlag.
- Redmond, K. T., and R. W. Koch (1991), Surface Climate and Streamflow Variability in the Western United States and Their Relationship to Large-Scale Circulation Indices, *Water Resources Research*, 27(9), 2381–2399.
- Regonda, S. K., B. Rajagopalan, M. Clark, and E. Zagana (2006), A multimodel ensemble forecast framework: Application to spring seasonal flows in the Gunnison River Basin, *Water Resources Research*, 42, W09404.
- Singhrratna, N., B. Rajagopalan, M. Clark, and K. K. Kumar (2005), Forecasting Thailand summer monsoon rainfall, *International Journal of Climatology*, 25, 649–664.
- Soukup, T. L., O. A. Aziz, G. A. Tootle, T. C. Piechota, and S. S. Wulff (2009), Long lead-time streamflow forecasting of the North Platte River incorporating oceanic-atmospheric climate variability, *Journal of Hydrology*, 368, 131–142.
- Tarboton, D., and A. Sharma (1998), Disaggregation procedures for stochastic hydrology based on nonparametric density estimation, *Water Resources Research*.
- Wadsworth, H. M. (Ed.) (1990), *Handbook of Statistical Methods for Engineers and Scientists*, McGraw-Hill.
- Walpole, R. E., R. H. Myers, S. L. Myers, and K. Yee (2002), Probability and statistics for engineers and scientists.
- Wilks, D. (1995), *Statistical Methods in the Atmospheric Sciences*, Academic Press, San Diego.
- Yates, D., S. Gangopadhyay, B. Rajagopalan, and K. Strzepek (2003), A technique for generating regional climate scenarios using a nearest-neighbor algorithm, *Water Resources Research*, 39(7).

Chapter 2

A Nonstationary Hidden Markov Model for Stochastic Streamflow Simulation and Short Term Forecasting in the Upper Colorado River Basin

Abstract. Upper Colorado River Basin annual flow exhibits very low autocorrelation but regime shifting behavior causing long departures from the historical average flow producing sustained wet and dry periods. Traditional stochastic time series models do not capture this feature thereby misrepresenting the water resources system risk and consequently impacting the management and planning efforts. To address this, we developed a nonstationary Hidden Markov (HM) model with Gamma component distributions, as opposed to Normal distributions which is widely used in literature, for stochastic simulation and short term forecasting. Global decoding from this model reveals and captures strong underlying persistent structure in the Lees Ferry flow time series. In addition to capturing the shifting mean, simulations from this model have a 20% greater chance than a first order Auto Regressive model (AR1), the best time series model for this data, of simulating wet and dry runs of 6 or more years. Relative to AR1 the HM model also captures the spectral features quite well. When applied to short term forecasting (i.e. of 1-2 years) they show higher skill relative to climatology but also to an AR1 model.

2.1 Introduction

The past decade has been one of unprecedented drought in the Upper Colorado River Basin (UCRB) (www.usbr.gov/uc/feature/drought.html). Traditional time series methods for simulating and forecasting (*Salas et al.*, 1980) annual flow at Less Ferry, the outlet of the UCRB, generates

such drought with extremely small probability, leading to an underestimation of water resources system risk. This problem with traditional time series methods applied to the annual natural flow data at Lees Ferry is due to an interesting dilemma that it presents – in that the data exhibits distinct regime-like behavior with long departures from the mean annual flow though still maintains a very weak autocorrelation structure (lag 1 is barely significant) (Figure 2.1). Traditional time series models are based on autocorrelation structure, thus, a weak autocorrelation leads to weak persistence and lower probability of long wet and dry spells. These long departures from the mean are important for multi-year reservoir planning in the UCRB since they stress the system far more than single wet or dry years. This behavior is similar to that suggested by *Akntuž and Rasmussen* (2005) where persistence structure in the data (in terms of climate regimes) is not fully described by the autocorrelation function.

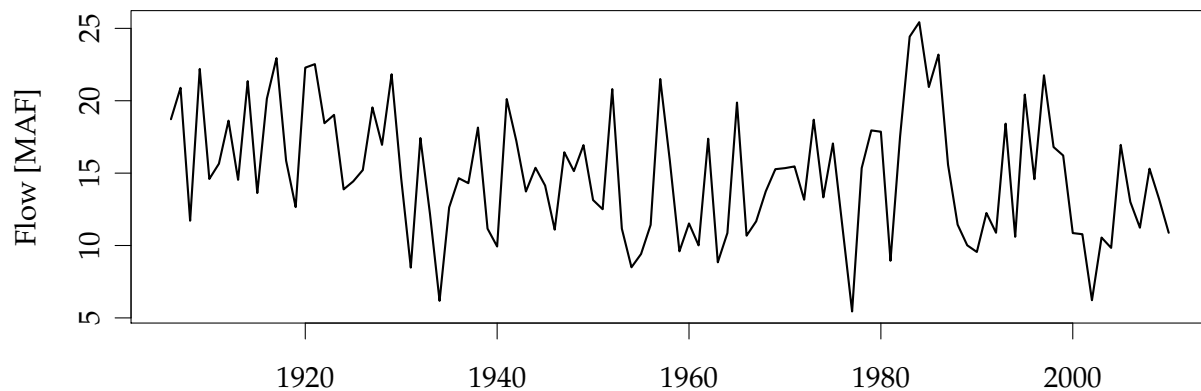


Figure 2.1: Lees Ferry annual natural flow time series.

Links between large scale climate drivers and UCRB hydrology are increasingly apparent (*Timlsena et al.*, 2009; *McCabe and Betancourt*, 2007; *Hunter et al.*, 2006; *Grantz et al.*, 2005; *Hidalgo*, 2003; *Piechota and Dracup*, 1996; *Nash*, 1991). These links are thought to create the type of long period regime switching behavior observed in the data. Though as the literature indicates, these links are not necessarily strait forward to explain, especially at the annual level or aggregated over the entire basin.

Many models have been proposed for annual flow time series that have applications to both simulation and forecasting. Most basic is the lag 1 autoregressive model (AR1) (*Salas et al.*, 1980) within the general autoregressive moving average (ARMA) framework. In fact, this is the model which the Akaike Information Criterion (AIC) criterion suggests as the best for the Lees Ferry data. These models have a number of drawbacks including the assumption of normality and stationarity. This is typically corrected for by transformation but does not guarantee that statistics will be preserved after transformation. AR models are also known to poorly reproduce the spectral properties of time series (*Kwon et al.*, 2007, ; *Nowak et al.*, 2011).

Another commonly used method for simulation in the UCRB is the index sequential method (ISM) (*Ouarda and Labadie*, 1997). This method is one used by the Bureau of Reclamation to produce flow sequences for the Colorado River Simulation System (CRSS) (*Bureau of Reclamation*, 1987). The ISM can directly reproduce observed statistics but (1) it does not maintain the integrity of a complete data sequence, (2) it cannot simulate unobserved values (3) nor can it be used for forecasting (since a forecast would simply be climatology). Improvements to this have been proposed in the K-nearest neighbor resampling framework (*Lall and Sharma*, 1996; *Prairie et al.*, 2008). These capture the distributional properties much better but suffer from their inability to capture the spectral and persistence properties (i.e., long excursions of wet and dry spells) that are important for management.

Spectral methods such as Fourier analysis (*Salas et al.*, 1980) and wavelet autoregressive methods (WARM) (*Kwon et al.*, 2007, and references therein) improve upon traditional methods by capturing the observed spectrum. These methods involve breaking a timeseries down into orthogonal components, modeling each component separately and adding the components back together. WARM can reproduce spectral features in a data and can capture a prescribed nonstationarity (*Nowak et al.*, 2011).

Some time series models can explicitly capture regime switching behavior without decomposition. The Shifting Level (SL) proposed by *Boes and Salas* (1978) is one such example. *Fortin et al.* (2004) provides an overview of other similar models. *Salas and Boes* (1980) describe the case

of the Nile River Basin where the shifting means can produce spurious autocorrelation. In fact the Lees Ferry data exhibits somewhat similar behavior, the serial correlation from 1906-1981 is 0.11 and from 1982-2010 it is 0.48 while for the whole period it is 0.26. This is indicative of a regime switching behavior in terms of autocorrelation but the behavior is also apparent in the mean. The original SL model did not estimate the nonstationary mean and was therefore not useful for forecasting. A reformulated SL model was described by *Fortin et al.* (2004) and shown to be a special case of a class of models known as Hidden Markov models.

Hidden Markov (HM) models (also known as Markov switching models or dependent mixture models) have wide applicability in hydrology for both simulation and forecasting though most of the applications have been to rainfall data (*Jackson, 1975; Zucchini and Guttorp, 1991; Thyer and Kuczera, 2000; Mehrotra and Sharma, 2005*). In applications to streamflow, HM models (*Zucchini and Guttorp, 1991; Akntuğ and Rasmussen, 2005; Gelati et al., 2010*) are attractive because of their ability to simulate long persistence and regime switching behavior in hydrologic time series. Regime changes, especially on the annual scale have been attributed to regime shifts driven by large scale changes in the climate features. The above applications to annual flow have focused on data with strong autocorrelation.

Akntuğ and Rasmussen (2005) discuss the correspondence of HM models with AR. Though different in structure, they suggest that an $HM(m)$ model has a similar autocorrelation structure to an order $AR(m+1)$. This result is indeed important for datasets that exhibit significant autocorrelation at high lags but leaves open the question of data with very weak autocorrelation such as Lees Ferry. *Akntuğ and Rasmussen* (2005) also use a homogeneous stationary hidden Markov model to simulate annual runoff for the Niagara River. The Niagara River exhibits strong autocorrelation and so is ideal to be simulated with HM models or higher order ARMA models. As mentioned previously this is a very different problem than in the UCRB. *Gelati et al.* (2010) use a stationary nonhomogeneous HM model conditioned on ENSO indices. They make single step quarterly forecasts and longer term simulations of runoff. Given this extensive literature, we believe that HM model would be a suitable approach for UCRB annual flow, this motivates the present research.

The goal of this paper is to develop a nonstationary Hidden Markov (HM) model with Gamma component distributions for simulation and forecasting annual flow in the UCRB. In particular we are interested in the ability of HM models to capture regime switching behavior but maintain low autocorrelation.

The paper is structured as follows. A brief overview of the model formulation, parameter estimation is given and moments of a Gamma HM model are derived. The application of HM models to Lees's Ferry Natural flow data is described for both simulation and short forecasting. Results of the simulation and forecasting procedures are then presented followed by the conclusions of this study.

2.2 Methodology

2.2.1 Model Formulation

As mentioned above, Hidden Markov (HM) models are also known as Markov switching models, Markov mixture models or dependent mixture models. An order m HM model transitions or switches between m 'hidden' states according to a discrete Markov chain with transition probability matrix Γ . These states are typically described as climate regimes (*Thyer and Kuczera, 2000; Akntuğ and Rasmussen, 2005; Gelati et al., 2010*). Each state prescribes a probability distribution known as a component distribution. The parameters of the component distributions are dependent on the state of the Markov process. We avoid using the terms "wet" and "dry" to describe the states of two state model as this can be slightly misleading since low flow can be generated in the "wet" state and visa-versa. We assume that the component distributions follow gamma distributions.

The notation for HM models in the literature is somewhat non-standard; we will adopt the notation of *Zucchini (2009)*. For an observed sequence $X_t, t = 1, 2, \dots, T$, the general form of a HM model is

$$\Pr(C_t | \mathbf{C}^{(t-1)}) = \Pr(C_t | C_{t-1}), t = 2, 3, \dots, T \quad (2.1)$$

$$\Pr(X_t | \mathbf{X}^{(t-1)}, \mathbf{C}^{(t)}) = \Pr(X_t | C_t), t \in \mathbb{N} \quad (2.2)$$

Where C_t is the unobserved or “hidden” sequence that follows a simple first order Markov process. $\mathbf{C}_t$ denotes the sequence $C_1, C_2, \dots, C_T$. The transition probabilities, i.e. the conditional probabilities of transition from one hidden state to another are defined as

$$\gamma_{jk} = \Pr(C_{i+1} = k | C_i = j) \quad (2.3)$$

Or in matrix form

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ \vdots & \ddots & \vdots \\ \gamma_{m1} & \cdots & \gamma_{mm} \end{bmatrix} \quad (2.4)$$

In this formulation, the observed sequence X_t is dependent only on the current hidden state C_t . Note that in general X_t is not a Markov process (Zucchini, 2009).

The hidden sequence C_t determines the state dependent probability distribution

$$p_i(x) = \Pr(X_t = x | C_t = i) \quad (2.5)$$

For our purposes p_i will represent a probability density function but it may similarly represent a probability mass function in the discrete case. In the literature, Gaussian component distributions have been commonly used either directly or after transforming the data (Jackson, 1975; Thyer and Kuczera, 2000; Akntuğ and Rasmussen, 2005; Gelati et al., 2010). The Gamma distribution is commonly used in the hydrologic modeling because of its lower bound of zero Salas et al. (1980). Here we investigate the use of a Gamma component distribution. In the case of a Gamma component distribution

$$p_i(x) = g(x; k_i, \beta_i) = \frac{\beta_i^{k_i}}{\Gamma(k_i)} x^{k_i-1} e^{-\beta_i x} \text{ for } x \geq 0. \quad (2.6)$$

Where k_i is the state dependent shape parameter, β_i is the state dependent rate parameter and Γ is the gamma function. The result is analogous for normal component distributions.

We use a nonstationary version of the model described by *Akntuž and Rasmussen* (2005). Our model does not assume that the initial distribution is the stationary distribution and therefore allows the expected state to change in time. The stationary distribution, δ , can be computed conveniently from the identity (*Zucchini*, 2009)

$$\delta(\mathbf{I}_m - \mathbf{\Gamma} + \mathbf{U}) = \mathbf{1}_m \quad (2.7)$$

where $\mathbf{I}_m$ is the $m \times m$ identity matrix, $\mathbf{U}$ is an $m \times m$ matrix of ones and $\mathbf{1}_m$ is an m dimension row vector of ones.

2.2.2 Moments of the Gamma HM

We provide moments and the autocorrelation function of the Gamma HM since it is not widely used. Corresponding formulas for Normal HM models are given in *Akntuž and Rasmussen* (2005) and *Frühwirth-Schnatter* (2006). Let δ denote the stationary distribution of the HM model then:

$$\mathbb{E}(X_t) = \sum_{i=1}^m \frac{\delta_i k_i}{\beta_i} \quad (2.8)$$

$$\text{Var}(X_t) = \sum_{i=1}^m \delta_i \left[a_i^2 + \frac{k_i}{\beta_i^2} \right] \quad (2.9)$$

where

$$a_i = \frac{k_i}{\beta_i} - \mathbb{E}(X_t) \quad (2.10)$$

and

$$\text{Skew}(X_t) = \text{Var}(X_t)^{-3/2} \sum_{i=1}^m \delta_i a_i \left[a_i^2 + 3 \left(\frac{k_i}{\beta_i^2} \right)^2 \right]. \quad (2.11)$$

The autocorrelation function is

$$\rho(k) = \frac{\sum_{i=1}^m \sum_{j=1}^m \frac{\delta_i k_i k_j \gamma_{ij}(k)}{\beta_i \beta_j} - [\mathbb{E}(X_t)]^2}{\text{Var}(X_t)} \quad (2.12)$$

Where $\gamma_{ij}(k)$ is the i, j entry of $\mathbf{\Gamma}^k$.

2.2.3 Parameter Estimation and Model Order Selection

Many methods exist to estimate the parameters of HM models. Commonly used techniques include direct maximization of the likelihood function (*Zucchini, 2009; Akntuğ and Rasmussen, 2005*) and Bayesian estimation procedures (*Thyer and Kuczera, 2000, 2003*). Another common method is known as the Expectation Maximization (EM) algorithm (*Dempster et al., 1977*). The EM algorithm provides a good compromise between the efficiency of direct maximization and the robustness of Bayesian techniques. The implementation of the EM algorithm in this context is known as the Baum-Welch algorithm. This algorithm estimates the parameters of the component distributions (Λ), the transition probabilities (Γ) and the initial distribution (δ) simultaneously.

Let θ represent all the parameters to be estimated, $(\Lambda, \Gamma, \delta)$. The EM algorithm requires initial guesses of parameters. We found that the final parameter estimates can be somewhat sensitive to initial parameter estimates; Through experimentation we decided on the following criteria for initial guesses:

- (1) $\Gamma = \mathbf{U}/m$, where $\mathbf{U}$ is an $m \times m$ matrix of ones.
- (2) The initial parameters are estimated by fitting a single component distribution to the entire data and then the same estimates are used for all the component distributions
- (3) The initial distribution is first estimated as $\delta = (1, \mathbf{0}_{m-1})$ where $\mathbf{0}_{m-1}$ is an $m - 1$ dimension row vector of zeros. If the choice for δ does not yield m distinct component distributions after employing the EM algorithm, then try $\mathbf{1}_m/m$ where $\mathbf{1}_m$ is an m dimension row vector of ones.

The HiddenMarkov package in R provides an implementation of the Baum-Welch algorithm that was used in this study (*Harte, 2011*). The details of the procedure will not be discussed here but we refer the interested readers to (*Zucchini, 2009*).

2.2.4 Global Decoding

Global decoding of the hidden states was done using the Viterbi algorithm (*Forney, 1973*). The Viterbi algorithm is a recursive procedure which maximizes the conditional probability

$$\Pr(\mathbf{C}^{(T)} = \mathbf{c}^{(T)} | \mathbf{X}^{(T)} = \mathbf{x}^{(T)}). \quad (2.13)$$

The resulting sequence $c_1, c_2, \dots, c_T$ is the most likely sequence of states, known as the global decoding. We again refer the interested readers to (*Zucchini, 2009*) for complete details of the procedure.

2.3 Applications of Hidden Markov Models

This section describes the application of HM models to Lees's Ferry natural flow data and their use for both forecasting and simulation.

2.3.1 Optimal Model Order and Stationary Distributions

Both the Bayesian Information Criteria (BIC) (*Schwarz, 1978*) and the Akaike Information Criteria (AIC) (*Akaike, 1974*) was calculated to determine the optimal model order for both normal and gamma component distributions (Table 2.1). In both cases the second order model was clearly the best in terms of both AIC and BIC. The difference between the HM2G and the HM2N is quite small so we evaluate the properties of both models. The model parameters are shown in Table 2.2. The stationary distributions of each model are shown in Figure 2.2.

Table 2.1: BIC and AIC for both Gamma and Normal HM models.

Model	BIC	AIC
HM2G	621.34	605.42
HM3G	647.26	615.42
HM2N	620.91	604.98
HM3N	643.73	611.89

Table 2.2: BIC and AIC for both Gamma and Normal HM models.

Parameter	Model
HM2G	
$\mathbf{k}$	$\begin{bmatrix} 12.28512 \\ 22.02217 \end{bmatrix}$
β	$\begin{bmatrix} 0.902059 \\ 1.204543 \end{bmatrix}$
Γ	$\begin{bmatrix} 0.98 & 0.02 \\ 0.08 & 0.92 \end{bmatrix}$
δ	$[0, 1]$
HM2N	
μ	$\begin{bmatrix} 13.57519 \\ 18.28946 \end{bmatrix}$
σ	$\begin{bmatrix} 3.728448 \\ 3.899349 \end{bmatrix}$
Γ	$\begin{bmatrix} 0.98 & 0.02 \\ 0.09 & 0.91 \end{bmatrix}$
δ	$[0, 1]$

Since both models have nearly identical transition probability matrices the stationary distributions are also nearly identical. The shapes of both stationary distributions are also nearly identical which gives some confidence in the the parameter estimates. Also of note is the nearly normal shape of both the distributions indicating that even with a nearly normally distributed data it is still possible to have underlying regimes that are very different.

The most likely sequence of hidden states is very revealing of the underlying persistence present in the system. Figure 2.3 shows the global decoding and the original time series. Also shown is the mean of the period over which the model was in a particular state. The strong tendency of the system to persist in a particular state is apparent. We see that the states also correspond to wet and dry epochs in the data. The period in the early 1900's corresponds the time in which the Colorado River Compact was signed which is known to be a period of above average flow.

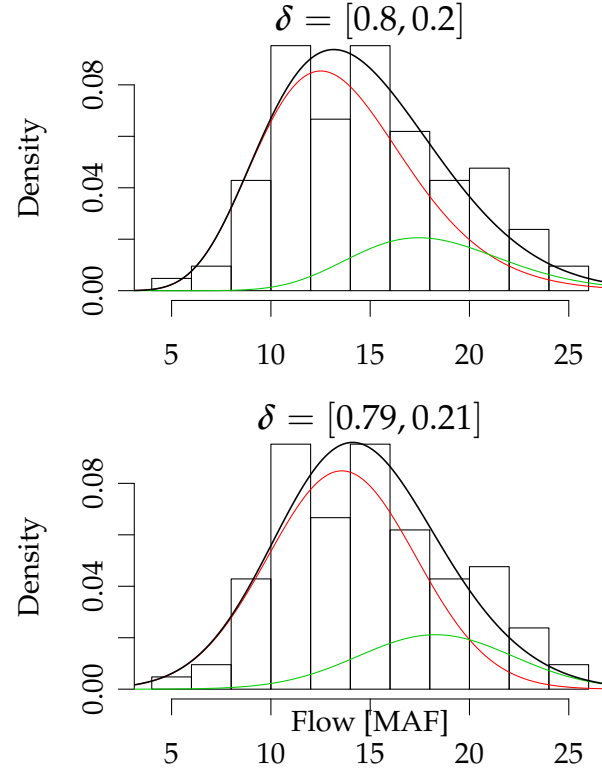


Figure 2.2: Stationary Distributions of HM2G (Top) and HM2N (Bottom).

2.3.2 Simulation Procedure

Once the model is fit, simulations can be easily obtained. The procedure for simulation is:

- (1) Determine the initial state from the initial distribution, δ .
- (2) Draw a uniform random number to determine the hidden state transition from the transition probability matrix, Γ .
- (3) Draw a random number from the component distribution corresponding to the current state.
- (4) Repeat 2 and 3 for the length of the observed data set.
- (5) Repeat 1–4 1200 times.

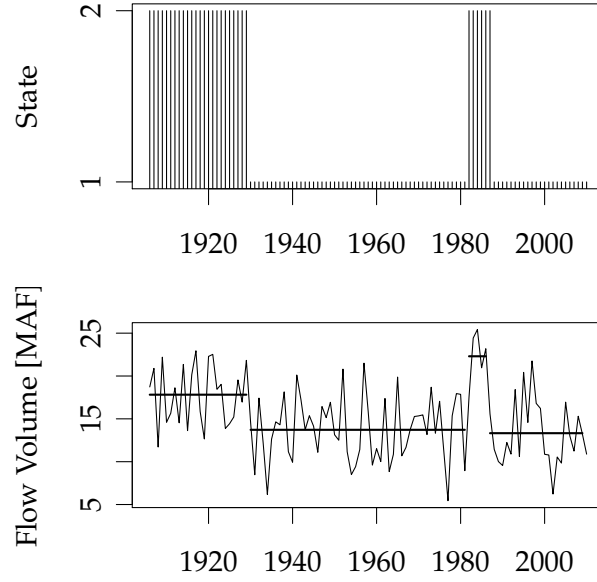


Figure 2.3: Global Decoding of the HM2G model using the Viterbi algorithm. Plotted below is the Lees Ferry annual time series with. The horizontal lines indicate the mean of the periods above where the model was in a particular state.

2.3.3 Forecast Distributions

Though they are not calibrated to do so, HM models can be used to conveniently produce forecast distributions for future time steps. A forecast distribution is simply a weighted combination of the component distributions of the HM (Zucchini, 2009)

$$\Pr(X_{T+h} = x | \mathbf{X}^{(T)} = \mathbf{x}^{(T)}) = \sum_{i=1}^m \varepsilon_i(h) p_i(x) \quad (2.14)$$

Where $\varepsilon_i(h)$ is the i th entry of

$$\boldsymbol{\alpha}_T \boldsymbol{\Gamma}^h / \boldsymbol{\alpha}_T \mathbf{1}_m' \quad (2.15)$$

The elements of $\boldsymbol{\alpha}_T$ are known as the forward probabilities and are defined as

$$\boldsymbol{\alpha}_T = \delta P(x_1) \boldsymbol{\Gamma} P(x_2) \dots \boldsymbol{\Gamma} P(x_T) = \delta P(x_1) \prod_{s=2}^T \boldsymbol{\Gamma} P(x_s) \quad (2.16)$$

where $P(x_i) = \text{diag}(p_1(x_i), \dots, p_m(x_i))$. Using Equation 2.14, forecast distributions can be generated for any h future time step.

These forecast distributions are useful in the UCRB for second year forecasts where climatology is the only information available presently. We make forecasts from April 1st when the snowpack is a good indicator of spring runoff. We make our forecasts assuming perfect knowledge of the current year's runoff such that a single step ahead forecast is actually for the 'second year'. This assumption is justified given the skill of April 1st runoff forecasts (*Bracken et al., 2010*) and the lack of sensitivity of the HM model to the addition of a single point. In an operational setting we would simply replace the current year flow with the median April 1st forecast from the ensemble (*Bracken et al., 2010*).

The forecasts are verified in two ways: (1) the ranked probability skill score (RPSS) (*Wilks, 1995*) is calculated for each year and (2) the correlation of the median of the forecast distributions with the observed data is calculated (referred to as the median correlation or MC). An RPSS value of 1 indicates a perfect categorical forecast, 0 indicates no different from climatology and negative values suggest worse than climatological forecast.

2.3.4 Modified Forecast Procedure

The forecast distributions are useful for assessing risk but many management models cannot input data as a PDF. Most models require a sequence or trace of input. Simply sampling from the forecast distributions will generate a trace but not one that maintains the persistent properties of the HM model. With this motivation we present a forecasting procedure that can generate realistic traces:

- (1) Assume the system is in the state that has a component mean closest to the current system magnitude.
- (2) Transition to a future state via the t.p.m. and the assumed state.
- (3) Sample from the future state distribution

- (4) Repeat steps 2–3 for the desired forecast length, h .

This method simply traverses through the forecast distributions in a way that maintains the persistence in the HM model.

2.4 Results

We present the parameter estimates and results from both simulation and some examples of forecasting.

2.4.1 Simulation

Following *Guimarães and Santos* (2011) we generate 1200 series for each model of length 105 (the same length of the data). Figure 2.4 shows the basic distributional statistics from the HM2G, HM2N and AR1 models, shown as boxplots along with observed value as horizontal line. The basic statistics are captured very well by all the models. The HM2G model tends to slightly oversimulate the skew while HM2N tends to undersimulate it. The minimum value is captured the best by the HM2G model where the AR1 and HM2N models can simulate negative values. This is one direct benefit of using a gamma underlying distribution. On the other hand, the HM2G slightly oversimulates the maximum which could be undesirable in some situations. Interestingly both HM models slightly under simulate the lag-1 autocorrelation as they are not designed to directly capture this.

Figure 2.5 shows the boxplot of the probability density function (PDF) of simulations from HM2G along with the observed. This too is captured very well – similar performance was seen from HM2N model as well (figure not shown).

The ability of HM models to simulate longer spell lengths is of interest here as it is one indicator of regime switching behavior. We have provided two views of this ability by displaying the frequency distribution and the mean counts of the spell lengths. Figure 2.6 shows a PDF of run lengths (in any tercile) from the HM2G simulations (shown in grey) along with the PDF of the

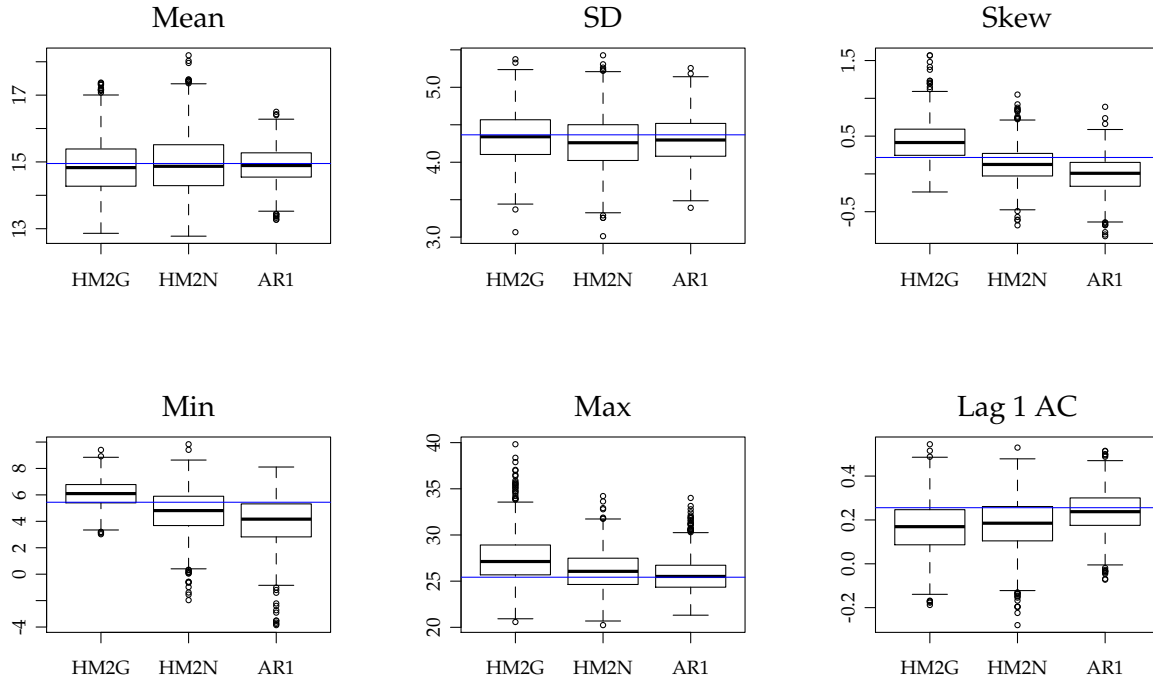


Figure 2.4: Basic simulation statistics of HM2G, HM2N and AR1 models. Boxplots represent the spread of the simulations and the horizontal line represents the observed value.

spell length from paleo reconstructed data from *Meko et al.* (2007). We used this to estimate the historical distribution of run lengths as it represents a greater range of variability of the system than the observed. Another view of the run lengths is provided in Figure 2.6. The mean counts (i.e. frequency) and the cumulative mean counts are provided across all the simulations for the three models (Figure 2.7). The cumulative mean count (and to some extent the mean count) show the increased probability of longer run lengths using either the HM2G or the HM2N models. For example both HM models will produce a run length of 6 or greater in 8 out of 10 simulations where an AR1 model will only produce it 6 out of 10 simulations – this translates into persistent wet and dry epochs that impact water resources system reliability.

As discussed in *Kwon et al.* (2007), the spectral signature is an important quantity to consider in simulation. Figure 2.8 shows the distinctly different spectral signature of the HM models com-

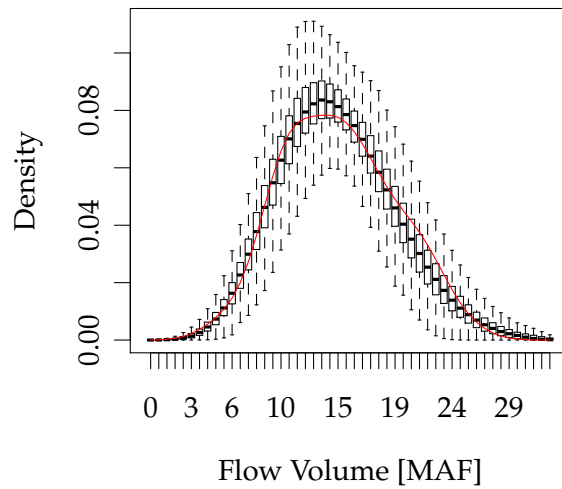


Figure 2.5: Observed (solid line) and boxplots of simulated pdf with the HM2G model.

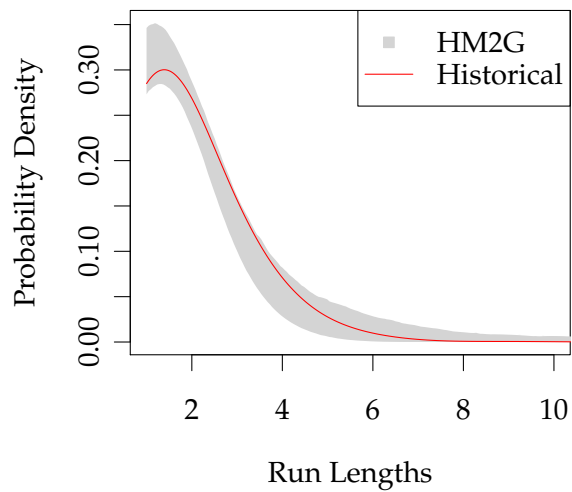


Figure 2.6: Observed and simulated (gray region) PDF of run lengths. The gray region extends to the 5th and 95th percentiles of the simulated PDFs.

pared to an AR1 model. The HM models show much greater power in long periods, nearly the opposite signature from and AR1 model. The model is not tuned to the spectral signal of the data

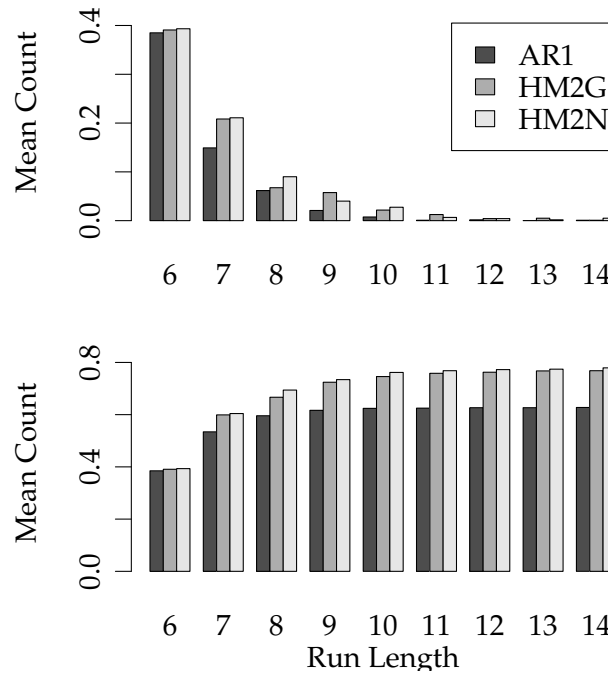


Figure 2.7: Mean count (Top) and cumulative mean count (Bottom) of run lengths. Note that these are not probabilities but rather the expected number of run lengths of a given size (or greater than a given size) in a single simulation.

so we do not expect to capture the spectral peaks – but the low frequency power that is seen in the simulations is remarkable.

2.4.2 Forecasting

Single step ahead forecasts as described above using the modified forecast procedure did slightly better than simply using the forecast distributions, with the added benefit that these forecasts can now be used as input to traditional management models. Forecasts for the HM2G and the HM2N are shown in Figure 2.9. The HM2N model is more responsive than the HM2G model to changes in the time series, producing slightly better forecasts overall. Table 3.5 shows the forecasting statistics (RPSS and MC). The HM2N model performs the best in all statistics, having the highest median RPSS value and median correlation (though it is not significant). The AR1 model actually does worse than climatology over 50% of the time. The HM models both tend to do better

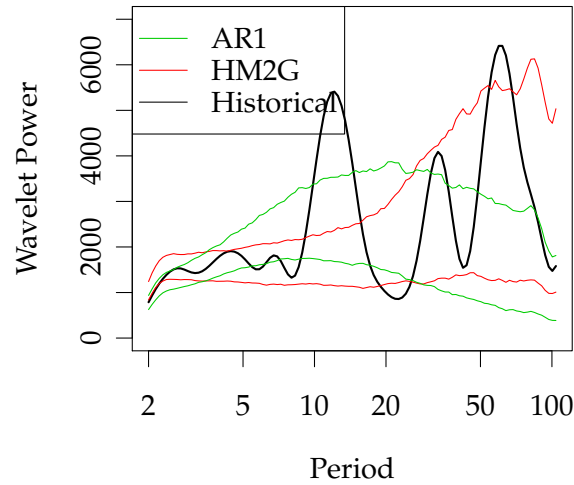


Figure 2.8: Quartiles of simulated HM2G and AR Global Wavelet Power Spectrum (the normal HM model is similar to the Gamma HM model).

during wet periods while the AR1 model tends to do better during dry periods. Boxplots of the RPSS values for each year are shown in Figure 2.10

Table 2.3: Median Correlation (MC) and Median Ranked probability skill score (RPSS) of forecasts for 1980–2010. Dry RPSS, Ave RPSS and Wet RPSS represent the median RPSS for years in the lower, middle and upper tercile respectively.

Model	MC	RPSS	Dry RPSS	Ave RPSS	Wet RPSS
HM2G	0.31	0.21	0.26	−0.11	0.05
HM2N	0.24	0.17	0.20	0.11	0.08
AR1	0.07	−0.03	0.07	−0.03	−0.25

Figure 2.11 shows example forecasts for 1984 (wet) and 2004 (dry). This view shows how the forecast distributions shift relative to climatology. The HM models have the ability to simulate non-normal shapes though typically the forecast distribution is essentially zero. In practice, we found that with the Lees Ferry data, the forecast distribution converged on the stationary distribution after about 20 future time steps.

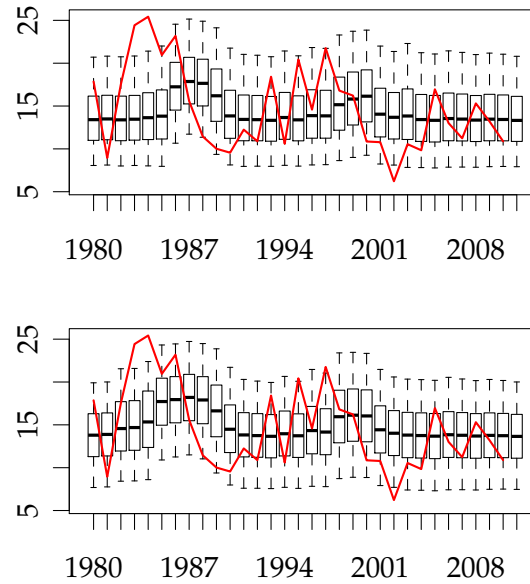


Figure 2.9: One step ahead forecasts from 1980 to 2010 for the HM2G (Top) and the HM2N (Bottom) with the actual data (solid line).

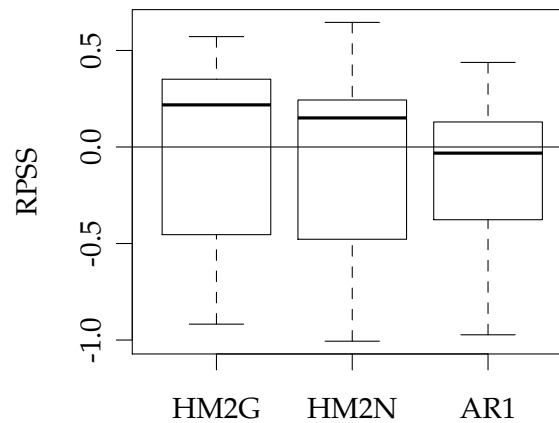


Figure 2.10: Boxplots of RPSS values for the forecast period 1980-2010. The horizontal line at 0 represents the skill of the climatological forecast.

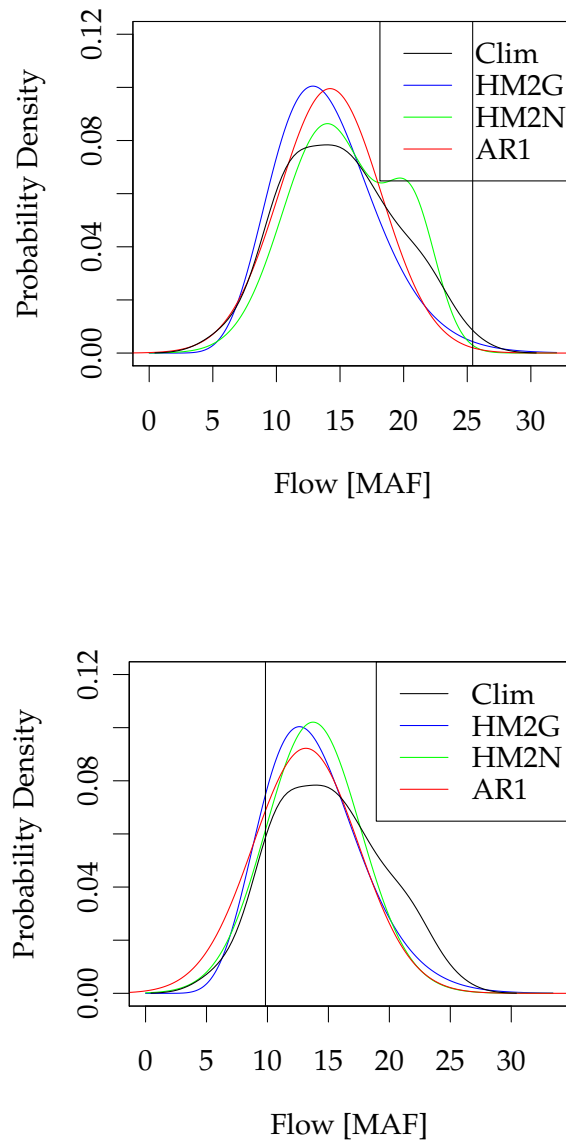


Figure 2.11: Forecast Distributions for 1984 (Left) and 2004 (right). The historical value is shown as a vertical line.

2.4.3 Three State Forecasts

Forecasts with a two state HM model tend to perform well for dry periods but not as well for average and wet periods. This is partially due to the heavy weighting on the “dry” component distribution. We investigated the use of a three state model for forecasting to further resolve the

“dry” pdf. The results were worse overall than a two state forecast but the three state normal HM model (HM3N) tended to be more responsive in the most recent drought period (Figure 2.12).

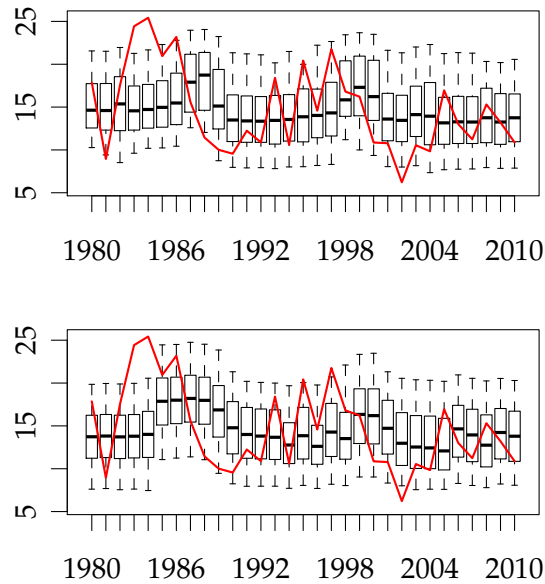


Figure 2.12: One step ahead forecasts from 1980 to 2010 for the HM3G (Top) and the HM3N (Bottom) with the actual data (solid line).

2.5 Conclusions

We discuss the application Hidden Markov time series models to natural flow from the Upper Colorado River Basin. This data exhibits very low autocorrelation but regime shifting behavior is present that can cause long departures from the historical average flow. We suggest the low autocorrelation could be due to the shifting regimes. We show that hidden Markov models are able to capture the regime like behavior of the system while maintaining low autocorrelation. We compare nonstationary HM models with gamma and normal component distributions. The optimal model order based on both the BIC and AIC is 2. The gamma (HM2G) and normal (HM2N) models have nearly identical AIC/BIC and so we examine the behavior of both compared to the more traditional AR1 model.

In simulation mode, the HM models are able to capture all the basic statistics including the

lag 1 autocorrelation. The historical PDF is captured as well as the distribution of run lengths. In addition to capturing the basic statistics, HM models have 20% greater chance than an AR1 model of simulating runs of 6 or more. The models are able to capture shifting mean. The ability of an HM model to capture long period regime shifting behavior is also highlighted in the wavelet spectrum where the HM models show a much higher power than an AR1 model at longer periods. The HM2G model may be better for simulation since it will not generate negative values.

In forecast mode, the HM models show modest skill over climatology but are able to outperform an AR1 model. The HM2N does better at forecasting than the HM2G model. The HM models tend to do better at forecasting during wet periods as opposed to an AR1 model which does better during dry periods.

A global decoding of both models is perhaps the most revealing result of this study. The decoding reveals and captures a strong underlying persistent structure in the Lee's Ferry time series. This finding suggests that simply capturing the autocorrelation function in a simulation may not be enough to describe the system.

The next step in in this investigation could be to relate the global decoding to some combination of atmospheric processes. These forecasts are also currently being tested as input to a decision model for the UCRB.

Bibliography

- Akaike, H. (1974), A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Akntuž, B., and P. F. Rasmussen (2005), A Markov switching model for annual hydrologic time series, *Water Resources Research*, 41.
- Boes, D. C., and J. D. Salas (1978), Nonstationarity of the mean and the hurst Phenomenon, *Water Resources Research*, 14(1), 135.
- Bracken, C., B. Rajagopalan, and J. Prairie (2010), A multisite seasonal ensemble streamflow forecasting technique, *Water Resources Research*, 46(3), W03,532.
- Bureau of Reclamation (1987), Colorado River simulation system, system overview, *Tech. rep.*, U.S. Dept. of the Interior, Denver.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), pp. 1–38.
- Forney, G. D. (1973), The viterbi algorithm, *Proceedings of the IEEE*, 61(3), 268–278.
- Fortin, V., L. Perrault, and J. Salas (2004), Retrospective analysis and forecasting of streamflows using a shifting level model, *Journal of Hydrology*, 296(1-4), 135–163.
- Frühwirth-Schnatter, S. (2006), Finite Markov Mixture Modeling, in *Finite Mixture and Markov Switching Models*, pp. 301–318, Springer New York, 10.1007/978-0-387-35768-3_10.
- Gelati, E., H. Madsen, and D. Rosberg (2010), Markov-switching model for nonstationary runoff conditioned on El Niño information, *Water Resources Research*.
- Grantz, K., B. Rajagopalan, M. Clark, and E. Zagona (2005), A Technique for Incorporating Large-Scale Climate Information in Basin-Scale Ensemble Streamflow Forecasts, *Water Resources Research*, p. W10410.
- Guimarães, R., and E. G. Santos (2011), Principles of Stochastic Generation of Hydrologic Time Series for Reservoir Planning and Design: A Case Study, *Journal of Hydrologic Engineering*, 1(1), 237.
- Harte, D. (2011), *HiddenMarkov: Hidden Markov Models*, Statistics Research Associates, Wellington, R package version 1.4-4.

- Hidalgo, H. (2003), ENSO and PDO effects on hydroclimatic variations of the Upper Colorado River Basin, *Journal of Hydrometeorology*.
- Hunter, T., G. Tootle, and T. Piechota (2006), Oceanic-atmospheric variability and western U.S. snowfall, *Geophys. Res. Lett.*, 33(13).
- Jackson, B. B. (1975), Markov mixture models for drought lengths, *Water Resources Research*, 11(1), 64.
- Kwon, H.-H., U. Lall, and A. F. Khalil (2007), Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature, *Water Resources Research*, 43(5), W05,407–.
- McCabe, G., and J. Betancourt (2007), Associations of Decadal to Multidecadal SeaSurface Temperature Variability with Upper Colorado River Flow¹, *Journal of the American Water Resources Association*, AWRA.
- Mehrotra, R., and A. Sharma (2005), A nonparametric nonhomogeneous hidden Markov model for downscaling of multisite daily rainfall occurrences, *J. Geophys. Res.*, 110(D16), D16,108.
- Meko, D. M., C. A. Woodhouse, C. A. Baisan, T. Knight, J. J. Lukas, M. K. Hughes, and M. W. Salzer (2007), Medieval drought in the upper Colorado River Basin, *Geophys. Res. Lett.*, 34(10), L10,705–.
- Nash, L. (1991), Sensitivity of streamflow in the Colorado basin to climatic changes, *Journal of Hydrology*.
- Ouarda, T., and J. Labadie (1997), Indexed Sequential Hydrologic Modeling For Hydropower Capacity Estimation, *Journal of the American Water Resources Association*, AWRA.
- Piechota, T. C., and J. A. Dracup (1996), Drought and Regional Hydrologic Variation in the United States: Associations with the El Niño-Southern Oscillation, *Water Resources Research*, 32(5), 1359–1373.
- Prairie, J., K. Nowak, B. Rajagopalan, U. Lall, and T. Fulp (2008), A stochastic nonparametric approach for streamflow generation combining observational and paleoreconstructed data, *Water Resources Research*, 44(6), 1–11.
- Salas, J., and D. Boes (1980), Shifting level modelling of hydrologic series, *Advances in Water Resources*, 3(2), 59–63.
- Salas, J. D., J. D. a. Yevjevich, and W. L. Lane (1980), *Applied Modeling of Hydrologic Time Series*, Water Resources Publications.
- Schwarz, G. (1978), Estimating the dimension of a model, *The annals of statistics*.
- Thyer, M., and G. Kuczera (2000), Modeling long-term persistence in hydroclimatic time series using a hidden state Markov model, *Water Resources Research*, 36(11), 3301–3310.
- Thyer, M., and G. Kuczera (2003), A hidden Markov model for modelling long-term persistence in multi-site rainfall time series 1. Model calibration using a Bayesian approach, *Journal of Hydrology*, 275(1-2), 12–26.

- Timlsena, J., T. Piechota, G. Tootle, and A. Singh (2009), Associations of interdecadal/interannual climate variability and long-term colorado river basin streamflow, *Journal of Hydrology*, 365(3-4), 289–301.
- Wilks, D. (1995), *Statistical Methods in the Atmospheric Sciences*, Academic Press, San Diego.
- Zucchini, W. (2009), *Hidden Markov models for time series: an introduction using R*.
- Zucchini, W., and P. Guttorp (1991), A hidden Markov model for space-time precipitation, *Water Resources Research*, 27(8), 1917–1923.

Chapter 3

A New Probabilistic Model for Reservoir Operations in the Colorado River Basin

Abstract. A new monthly reservoir operations model for the Colorado River Basin, The Mid-Term Operations Model (MTOM), allows for stochastic inflow input. The primary feature of the MTOM is the ability to generate reservoir outflows without the need for operator input through the use of operational rules. We generate 1000 inflow traces for this model spanning 32 months. The inflow traces are built as a combination of a seasonal forecast model, historical data and an Hidden Markov time series model for the out years. The MTOM is run for three start dates, November 1, 2007, January 1 2008 and April 1 2008. We show a variety of reservoir scenarios of both outflow and reservoir elevation. The result indicate that errors in the forecast are due both to inflow forecast error and operational rule error though in most cases a single cause cannot be separated. We provide evidence that inter-annual forecasts can provide better operational forecasts than the existing model for reservoir operations in the UCRB. We conclude that the MTOM is a valuable tool for assessing reservoir conditions in the Colorado River Basin and its worth can only increase with further development.

3.1 Introduction

In the last ten years water supply management in the Upper Colorado River Basin (UCRB) has come under increasing scrutiny and regulation. Drought (*Fulp, 2005*), climate change (*Rajagopalan et al., 2009*), environmental flows and increasing demands (*Bureau of Reclamation, 2008*) have all put the reliability of the system into question. In 2007 the “Interim Guidelines” were

passed as a response to some of these factors (*Bureau of Reclamation, 2007*). These guidelines lay a framework for coordinated operations of Lake Powell and Lake Mead which depend heavily on water supply forecasts. These factors all underscore the need for skillful forecasts at long lead times at many locations throughout the UCRB.

Every month the United States Bureau of Reclamation (USBR) provides an outlook of reservoir operations in the Upper Colorado River Basin (UCRB) for the following 24 months (<http://www.usbr.gov/uc/water/crsp/studies/>). The outlooks are produced with a RiverWare model known as the 24-Month study (24MS). The inputs to the 24MS are unregulated reservoir inflows and outflows, initial pool elevations and storage. The outputs are reservoir storage and elevations for the next 24 months (in some cases the model is run for up to 32 months). The model takes unregulated flow input since it has very few explicit demands. In early 2010, a model known as the Expanded 24-Month Study (E24MS) as put into use, replacing the original model. This model explicitly represents many demands in the Lower Colorado River Basin (LCRB). Demand inputs are generated by various daily and monthly models run by the LC region.

The unregulated inflow forecasts for the 24MS (generated by the CBRFC) are currently used for only the first 12 months of the model (approximately). After the forecast period, climatology is input and the model will always tend toward average conditions. As a result, there less confidence among operators or stakeholders in the model output during out years.

The E24MS like the 24MS before it, can only input a single hydrologic trace (though the inflow is available as an ensemble) and therefore its output is not probabilistic. The USBR is able to produce a range of possible outputs by running the model with multiple traces representing the 10th, 50th and 90th percentiles of flow. This process has been criticized for a number of reasons, one being the considerable effort required to run a single trace, another being that running a 90th percentile trace for example is not a realistic hydrologic scenario.

Also in early 2010, development began on a probabilistic version of the E24MS, know as the Midterm Operations Model (MTOM). The main feature of this model is built in rules that mimic the operations of each reservoir. This makes it unnecessary to input reservoir outflows. As a

result of the built in operations, an arbitrary number of inflow traces may be run to provide a truly probabilistic look at reservoir conditions over the next 24–32 months.

Unregulated inflow forecasts are generated by the Colorado Basin River Forecast Center (CBRFC). Forecasts for peak season (April-July) are first issued in January when early snowpack measurements are available. Many recent studies have shown that climate features in the Pacific have an influence on streamflow in the west and in the UCRB (*Kahya and Dracup, 1993, 1994; Dracup and Kahya, 1994; Piechota et al., 1997; Maurer and Lettenmaier, 2004; McCabe and Dettinger, 1999, 2002; Hidalgo and Dracup, 2003; Brown and Comrie, 2004; Hunter et al., 2006; Soukup et al., 2009*). These climate links can be used in statistical models to produce long lead forecasts when no snowpack information is available. Previous studies used statistical frameworks to produce long lead streamflow forecasts in the Truckee and Carson river basins (*Grantz et al., 2005*), Gunnison river basin (*Regonda et al., 2006*), Columbia River (*Hamlet and Lettenmaier, 1999; Clark et al., 2001*), the Yakima river basin (*Opitz-Stapleton et al., 2007*), and the UCRB (*Bracken et al., 2010*).

The nonparametric multisite ensemble forecast framework developed by Bracken et al. (2010) showed skill at long lead times in forecasts of peak season streamflows. They demonstrated their approach on four key sites in the Upper Colorado River Basin (Colorado River near Cisco, Utah, Green River at Green River, Utah, San Juan River near Bluff, Utah, and Colorado River at Lees Ferry, Arizona). In this forecasts of an “index gauge” that is composed of the flow at all the sites in the network is first generated based on a number of large scale climate predictors using local polynomial functional estimation approach Large scale climate predictors are identified for the index gauge flow at several lead times and local polynomials is used to obtain the best model. The generated seasonal forecast of the index gauge is then disaggregated using the approach of Prairie et al. (2007) to obtain forecasts at the four locations and for each month of the peak season. As mentioned, significant skills were obtained at the four locations. While promising, these results leave open the question of whether the skills translate to all the locations on the UCRB that are needed for planning and management of the water resources system.

As mentioned previously, climatology is used as input to the 24MS during the second year.

Largely this is due to the lack of information available for making forecast at this time scale. The use of HM models for forecasting streamflow is limited in the literature, recent applications are *Gelati et al.* (2010) and *Bracken et al.* (2011). This type of model captures an underlying regime persistence in UCRB flow that can be used to produce skillful inter-annual or second year forecasts.

The goals of this study are to (1) combine seasonal and inter-annual forecasting methods to generate multiple forecast traces of UCRB reservoir inflow three different lead times, (2) drive the new MTOM model with the forecast traces and (3) compare the output of the MTOM with the 24MS, in both the first year and out years. Concerning the third goal, there are two conflicting factors to consider, the ability of the operating rules to reproduce operations that are input in the 24MS and the skill in the inflow traces, both of which we would like to evaluate. In comparing these models, we would not like to conclude that one is better than the other. Rather we would like to demonstrate some of the similarities and differences between these two distinct models.

This paper is structured as follows: In the methodology section, we describe the process of generating 32 month inflow traces based on a combination of forecast models and historic data. We describe each of the forecast models as well as the the precise inputs to the MTOM. In the verification section we provide evidence of of the skill of each of the forecast models. The results section provides example output from the MTOM compared to the corresponding 24MS output and historical data followed by conclusions and future work.

3.2 Study Area and Data

3.2.1 Study Area

The UCRB has a drainage area of 279,300 square kilometers. The terrain varies over 4000 m from east to west. Lees Ferry, just below Lake Powell, is located at the outlet of the Upper Basin. The sites used in the MTOM (Figure 3.1a) are all the major reservoirs in the UCRB. Most of the sites also correspond to USBR natural flow nodes and USGS gauges, the names of which are given in Table 3.1.

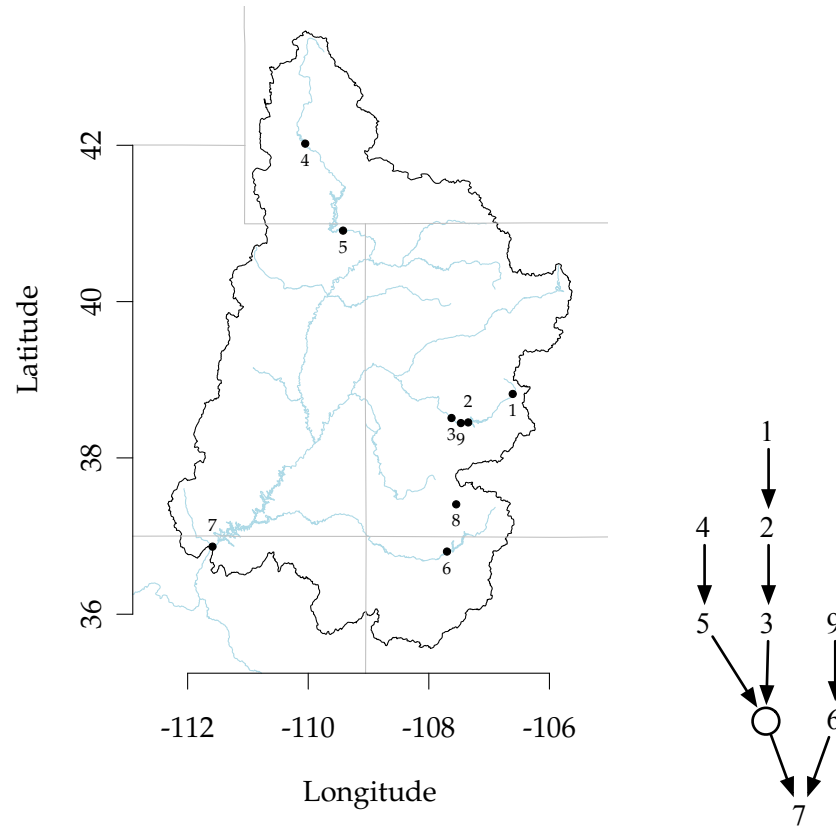


Figure 3.1: Map of UCRB with reservoir locations corresponding to Table and a schematic of network connectivity in the MTOM.

Table 3.1: UCRB reservoirs and corresponding natural flow sites

Node	Reservoir	USGS Gauge	Natural Flow Site Name
1	Taylor Park	09109000	Taylor River Below Taylor Park Reservoir, CO
2	Blue Mesa	09124700	Gunnison River Above Blue Mesa Reservoir, CO
3	Crystal	09127800	Gunnison River At Crystal Reservoir, CO
4	Fontenelle	09211200	Green R Bel Fontenelle Res, WY
5	Flaming Gorge	09234500	Green River Near Greendale, UT
6	Navajo	09355500	San Juan River Near Archuleta, NM
7	Lees Ferry	09380000	Colorado R At Lees Ferry, AZ
8	Morrow Point	—	—
9	Vallecito	—	—

3.2.2 Flow Data

We obtained the most recent monthly natural flow data for 20 sites in the UCRB developed by the Bureau of Reclamation in the UCRB that currently extends to 2007 (Table 3.2). This data set is developed and updated regularly by the United States Bureau of Reclamation. Naturalized streamflow are computed by removing anthropogenic impacts (i.e., reservoir regulation, consumptive water use, etc.) from the recorded historic flows. *Prairie and Callejo* (2005) present a detailed description of methods and data used for the computation of natural flows in the CRB. The data is available from (<http://www.usbr.gov/lc/region/g4000/NaturalFlow/index.html>). The data is provided as both intervening (gains since the last upstream gauge) or as total flow (sum of all upstream intervening flow at a gauge).

Table 3.2: Natural flow sites.

Node	USGS Gauge	Site Name
1	09072500	Colorado River At Glenwood Springs, CO
2	09095500	Colorado River Near Cameo, CO
3	09109000	Taylor River Below Taylor Park Reservoir, CO
4	09124700	Gunnison River Above Blue Mesa Reservoir, CO
5	09127800	Gunnison River At Crystal Reservoir, CO
6	09152500	Gunnison River Near Grand Junction, CO
7	09180000	Dolores River Near Cisco, UT
8	09180500	Colorado River Near Cisco, UT
9	09211200	Green R Bel Fontenelle Res, WY
10	09217000	Green R. Near Green River, WY
11	09234500	Green River Near Greendale, UT
12	09251000	Yampa River Near Maybell, CO
13	09260000	Little Snake River Near Lily, CO
14	09302000	Duchesne River Near Randlett, UT
15	09306500	White River Near Watson, UT
16	09315000	Green River At Green River, UT
17	09328500	San Rafael River Near Green River, UT
18	09355500	San Juan River Near Archuleta, NM
19	09379500	San Juan River Near Bluff, UT
20	09380000	Colorado R At Lees Ferry, AZ

We also obtain historic unregulated inflow data from the Bureau of Reclamation for all the

reservoirs in Table 3.1.

3.2.3 Climate Data for Seasonal Forecasts

For accurate seasonal forecasts at long lead times we use climate predictors of The climate variables are the same as (*Bracken et al.*, 2010) though we obtained the most recent observations for all of the predictors (through 2010 where available). The variables used were zonal (ZNW) and meridional (MDW) wind, Sea Surface Temperatures (SST), geopotential height (GPH) and Palmer Drought Severity Index (PDSI, a surrogate for soil moisture) – all from the NOAA Earth Science Research Laboratory as predictors of large scale climate. Regions of high correlation with Lees Ferry Flow were determined using the ESRL linear correlation tool (<http://www.esrl.noaa.gov/psd/data/correlation/>). The variables are averaged over these regions and the resulting time series is used as a predictor. This analysis is repeated for each lead time to obtain a suite of climate predictors. This technique is described in detail in *Grantz et al.* (2005) and *Regonda et al.* (2006).

3.2.4 Snow Data

The amount of snow water equivalent (SWE) data was greatly increased over that used in *Bracken et al.* (2010). *Bracken et al.* (2010) used 10 representative sites obtained from Natural Resources Conservation Service (NRCS) (<http://www.wcc.nrcs.usda.gov/snow>). We use the 86 sites (Figure 3.2) that go into the UCRB snowpack report (<http://www.usbr.gov/uc/water/notice/snowpack.html>). As a result of using more data we were able to generate a snowpack predictor for January 1st as well as improve the snow predictors for all other lead times.

3.2.5 Additional MTOM Data

In addition to inflow provided by the forecast models, historical data was required for initialization. Initial reservoir states historical outflows were obtained from archived 24MS reports (<http://www.usbr.gov/lc/region/g4000/24mo/index.html>). Lower basin demand data was obtained from the USBR.

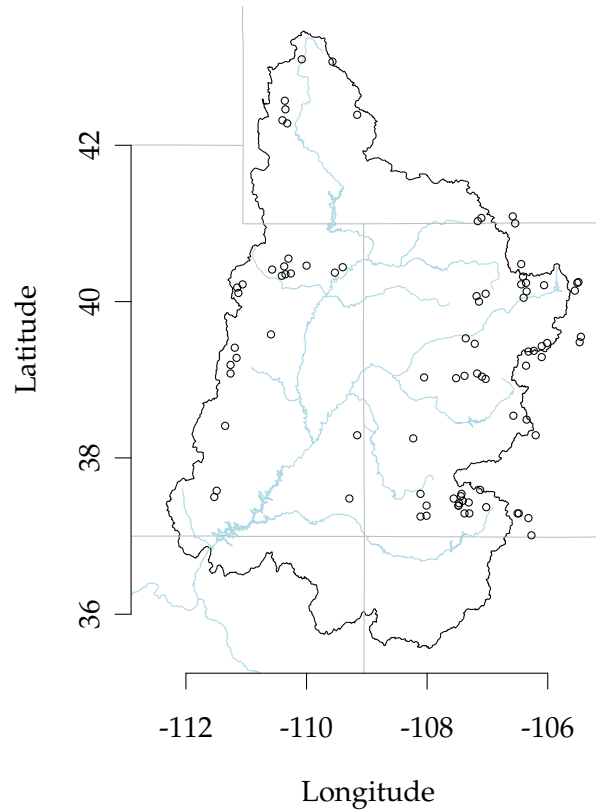


Figure 3.2: The Upper Colorado River Basin with locations of snow data sites used in this study indicated by an open circle.

3.3 Methodology

Forecast traces are developed for model runs starting on November 1st, 2007, January 1st, 2008 and April 1, 2008. The traces are a combination of CBRFC coordinated forecasts (for the November 1st and January 1st lead time), a seasonal forecast model based on (*Bracken et al., 2010*), historical resampling and a disaggregated forecast from a hidden Markov model. This section will describe the three components.

3.3.1 CBRFC coordinated forecasts

Each month the CBRFC produces a coordinated forecast which is a combination of The National Resource Conservation Service's (NRCS) statistical water supply (SWS) model, the CBRFC's ensemble streamflow prediction (ESP) model and expert knowledge. We do not specifically forecast early season (November - March) flows so for the November 1st and January 1st start dates we use the Coordinated forecast values to supplement pre-April 1st inflows.

3.3.2 Seasonal Forecast Model

For April - July of the first year of the model we use the framework as (*Bracken et al.*, 2010, see their section 4) with some notable changes. Most importantly we (1) expand the disaggregation to include all 20 natural flow nodes (Table 3.2) in the UCRB and (2) we include the flexible disaggregation method of *Nowak et al.* (2010) instead of computationally intensive the disaggregation method of *Prairie et al.* (2007). The first change is necessary in order to later aggregate the natural flows to the correct location for the MTOM. The second change allows us to use a large number of ensemble members efficiently.

As in *Bracken et al.* (2010) we created an "index" gauge peak season flow which the sum of peak season flows at all the 20 natural flow locations; this is forecasted using large scale climate variables in a multi-model ensemble approach, for full details the reader is referred to *Bracken et al.* (2010) and *Regonda et al.* (2006). The ensemble forecasts of seasonal flow are then disaggregated to ensemble monthly intervening flows at all the 20 locations.

The nonparametric disaggregation method of *Nowak et al.* (2010) is known as proportion disaggregation is a computationally simple approach that preserves the summability criteria of the network (intervening flows upstream sum to downstream total flow at Lees Ferry). The method is able to simultaneously conduct space and time disaggregation through the use of proportion matrices (matrices whose contents sum to unity). Figure 3.3 shows schematically how the index gauge peak season flow is split simultaneously to in space (at all twenty sites) and time (for all

the months in the peak season). In space-time disaggregation there is one proportion matrix for each year in the historical record. For a given seasonal flow value to be disaggregated, the method identifies K nearest neighbors of the flow value and selects a historical proportion vector which is multiplied by the total flow to obtain space-time disaggregated values. For more details, the interested reader is referred to *Nowak et al. (2010)*.

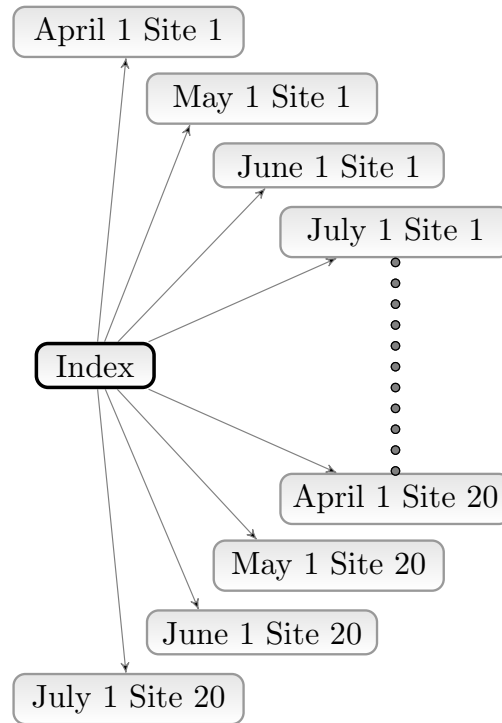


Figure 3.3: Schematic of disaggregation process using proportion disaggregation.

Once the intervening natural flow ensembles are generated, they must be aggregated to the locations in the MTOM. This is a strait forward process for most of the MTOM locations. Two locations, Vallecito and Morrow Point, do not exist in the natural flow data. Morrow point can be estimated as a fraction of Crystal inflow since it is relatively close and there are not many hydrologic gains in that reach. Vallecito is generated by relating a separate natural flow record. Flow at this site is generated by regressing Navajo inflows with the Los Pinos natural flow record. The magnitude of flow at this site is relatively low and error introduced by using a regression is expected to be insignificant.

Below we describe the complete procedure for the seasonal forecast model to generate natural inflow traces for the MTOM:

- (1) 1. Create an “index” gauge flow by summing all the total seasonal flow at each site, as mentioned above. Identify large scale climate predictors of index gauge flow. This is done by correlating it with global climate variables at different lead times and creating average time series from regions of high correlation with physical regions (*Grantz et al., 2005; Bracken et al., 2010*).
- (2) 2. Identify a set of the “best” set of models using locally weighted polynomial (Loader 1999) for each lead time. The best models are defined as those that have low generalized cross validation value (GCV) (*Craven and Wahba, 1979*) and do not exhibit multicollinearity. A suite of best models is selected using their GCV values - typically, all models within a small range of the lowest GCV value are selected (see *Bracken et al. (2010); Regonda et al. (2006)*, for details). The functional form of the models is the typical regression form

$$y = f(\mathbf{x}) + \varepsilon \quad (3.1)$$

where $\mathbf{x}$ is the vector of predictor variables, f is the function that is estimated locally using locally weighted polynomial approach and ε is the error with the standard assumption of Normality and identical and independently distributed residuals (i.i.d) based on regression theory. The details of this approach and applications can be found in the aforementioned references. Also, see *Rajagopalan et al. (2010)* and *Lall (1995)* for a general review of local functional estimation methods.

- (3) Identify a set of the best locally weighted polynomial models *Loader (1999)* for each lead time. The best models are defined as those that minimize the generalized cross validation value (GCV) (*Craven and Wahba, 1979*) and do not exhibit multicollinearity *Regonda et al. (2006)*.
- (4) Generate ensemble forecasts of index gauge flow by weighting the best models according

to their GCV value (lower GCV gets higher weight). Multimodel ensemble predictions are made by randomly selecting a model based on the GCV-based weights, obtaining the mean forecast from the model (of the form in equation 3.1 above) and generating a Normal deviate with the appropriate error variance. This is repeated to generate the ensemble.

- (5) Disaggregate the total seasonal flow at to the index gauge to intervening monthly flow at all the 20 sites.
- (6) Aggregate the 20 natural flow such that they correspond to the 9 reservoir inputs in the MTOM, using special consideration for Morrow Point and Vallecito.

3.3.3 Resampled data

For the August - December period of the first year, we use a KNN resampling of historical data. We resample historic April - July volumes closest to the forecasted volume for a given trace and then use the following August - December monthly values for the trace.

3.3.4 Hidden Markov Model Timeseries Forecast

For the out years (January 2009 on), we forecast natural flow using a Hidden Markov (HM) model. HM models (a.k.a. Markov switching models and dependent mixture models) have been widely used in hydrologic applications. Recent application have been to rainfall (*Jackson, 1975; Zucchini and Guttorp, 1991; Thyer and Kuczera, 2000; Mehrotra and Sharma, 2005*) and streamflow (*Zucchini and Guttorp, 1991; Akntuğ and Rasmussen, 2005; Gelati et al., 2010*). HM models assume the system switches between m states based on an underlying or “hidden” discrete Markov process. States are defined by probability distributions (typically normal) each with their own parameters.

A recent study by *Bracken et al. (2011)*, uses HM models for forecasting of the Lees Ferry annual natural flow series. HM models perform better for forecasting than traditional autoregressive models. In addition they are able to capture a strong underlying persistence (defined by the global decoding) but still maintain a low autocorrelation. We use this model to generate

forecasts of annual flow, which are then disaggregated to monthly flow sequences using the same disaggregation method as above.

With HM models it is possible to generate forecast distributions for any number of future time steps. The only complication is that forecasts are generated as probability distributions for any future time step, and the MTOM requires traces. We used the forecast procedure described by *Bracken et al.* (2011). Annual HM forecasts for out years are generated as follows:

- (1) Fit an HM model to the historical data plus the median of the seasonal forecast model to represent the current year.
- (2) Generate forecast distributions for h years ahead.
- (3) Assume the system is in the state that has a component mean closest to the current system magnitude.
- (4) Transition to a future state via the t.p.m. and the assumed state.
- (5) Sample from the future state distribution
- (6) Repeat steps 2–3 for the desired forecast length, h and number of forecast traces.

The annual flows are then disaggregated using the method of (*Nowak et al.*, 2010). As a final step, the flows are bias corrected (scaled and shifted to have the same standard deviation and mean as the observed data). This helps alleviate the tendency of the HM model to under or over predict due to the two state model.

3.3.5 Generating Unregulated Traces

At this point we have constructed complete traces for each input site in the MTOM for the entire 32 month runs, but the data is natural and we need unregulated. Since we cannot explicitly model all the demands in the UCRB we must again resort to linear regression. We regress the natural flow record against the unregulated flow record for each site and each month. The natural

flow traces are then converted to unregulated flow using these regression equations. In nearly all cases, the regression simply translates to a constant shift since demands are relatively constant in the UCRB. The cases where a simple shift was not enough were limited to very low flow months and/or sites and are not expected to introduce significant error.

3.3.6 Constructing a continuous trace

A 32 month trace starting on April 1st for a single site is broken down as follows:

- (1) **Nov - Feb, Year 1:** If the model start date is before April 1st, use the appropriate amount of CBRFC forecast data as described in Section 3.3.1
- (2) **Apr - Jul, Year 1:** Use the data from the seasonal forecast model described in Section 3.3.2.
- (3) **Aug - Dec, Year 1:** Use resampled historical data as described in Section 3.3.3
- (4) **Jan - Dec, Year 2:** Use disaggregated HM model forecasts generated as traces by the forecast algorithm described in Section 3.3.4
- (5) **Jan - Dec, Year 2:** Same as above but use the second step ahead forecast.
- (6) **Jan - Dec, Year 2:** Same as above but use the third step ahead forecast (if necessary to generate 32 months of input).

3.3.7 MTOM model runs

We run the MTOM model starting on November 1st 2007, January 1st 2008 and April 1 2008. This start date allows for verification of the entire forecast period. In addition, many operations of UCRB reservoirs have recently (in the last 10 years) changed; using a more recent start date will enable more accurate comparison of 24MS output to MTOM output. Figure 3.4 has a schematic of the data inputs for each of the lead times.

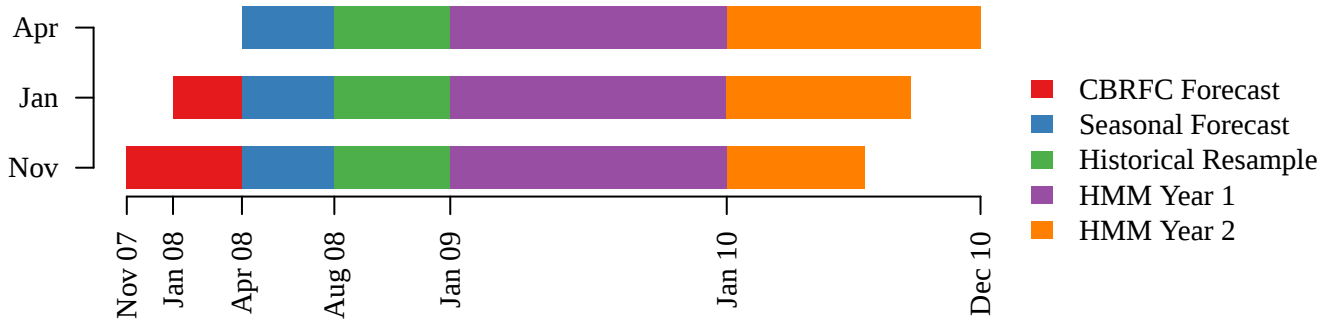


Figure 3.4: Construction of single trace for each lead time.

3.4 Verification

3.4.1 Seasonal Forecast Model Verification

To verify the forecasts of the seasonal forecast model, we calculate two metrics. The ranked probability skill score (RPSS) is a categorical measure of skill and is used to compare ensemble forecast to historical values. It takes values from $-\infty$ to unity, any number greater than unity indicating skill over climatology. The other metric the median correlation (MC) measures correlation between the median of each ensemble and the historical values, a high correlation indicating responsiveness in the forecast.

Table 3.4 shows the RPSS and MC for all the locations and all months in the peak season for the three lead times mentioned above. In general we see the skills decreasing as lead time increases, as to be expected. Some sites such as the Duchesne River Near Randlett, UT (14) have consistently positive skill back to the Nov 1 lead time. June had the highest overall skill of any month, likely because the average peak flow occurs in June. The sites included in *Bracken et al.* (2010) – “key gauges” (8, 16, 19 and 20, outlined in bold) represent the major contributions to Lees Ferry flow and thus important for planning purposes. The skills at these gauges are high as they have a strong response to large-scale climate forcing. Another trend we see is that the skills tend to be better as we move toward the outlet of the basin. That is consistent with the result that the larger tributaries perform better and that by aggregating the flows the impact of large scale

climate signal emerges more strongly. See Appendix B for additional seasonal forecast verification information.

With the exception of few sites the skills of the disaggregated flows are positive from Nov 1st onwards with increasing magnitudes moving towards Apr 1st (Table 3.3). This indicates that the skill in forecasting streamflow at the index gauge is translated throughout the upper basin, an important contribution that will have significant impact on water resources planning and management. The locations with poor skills are mainly in smaller tributaries and this could be improved by disaggregating in chunks based on regions (such as the San Juan basin) to better capture regional behavior.

Table 3.3: RPSS values for the index gauge for each lead time.

	Apr1	Feb1	Jan1	Nov1
Leave-one out	0.85	0.74	0.49	0.30
Retroactive	0.62	0.58	0.55	0.52

3.4.2 HM Verification

Table 3.5 shows the verification for the HM forecast. Also shown are the skills of an AR1 forecast. It can be seen that the HM model outperforms the AR1 model. The HM model tends to do better during wet periods while the AR1 model tends to do better during dry periods. See Chapter 2 for complete details on the model selection procedure and the verification process.

3.4.3 Inflow Trace Verification and Control Runs

Figure 3.5 shows generated unregulated inflow traces (shown as boxplots) at three different sites (Navajo, Flaming Gorge and Lake Powell) at three different lead times (November, January and April). The green line represents the 24MS most probable trace and the blue line represents the observed unregulated inflow. The three plots are a good representation of the general range of values generated in the inflow sequences. These traces represent a wide range of possible

hydrologic outcomes.

In addition to the inflow traces we conducted control runs of the MTOM to isolate sources of error. A control run consists of the 24MS most probable input trace run through the MTOM. If both models were identical the control run output would be identical to the 24MS output and somewhat similar to the median of the MTOM output.

3.5 Results

The results of the MTOM, as in the 24MS, are a suite of reservoir variables including (but not limited to) inflow, outflow/release, storage, pool elevation, power and evaporation. We limit our analysis to pool elevation and outflow. Pool elevation is essentially a scaled measure of storage. It has a fairly slow response to changing inputs and therefore is a good measure of long term performance of both rules and inflow forecasts. Outflow is a good measure of how the operating rules in the MTOM are mimic actual operations.

Figure 3.6 shows Pool Elevation at Navajo reservoir. Boxplots indicate the MTOM results, with whiskers extending to the 10th and 90th percentile of the data. The green line is the 24MS output, the dashed green lines are the 10th and 90th 24MS runs, the blue line is the observed and the purple dashed line is the control run output. In this case we see that for the first model year, the MTOM, the 24MS and the control run are in good agreement, only starting to diverge in the out year. The benefit of true probabilistic inflow traces is apparent here; The MTOM presents a wide range of possible scenarios. The extreme low pool elevations seen in the MTOM output are unlikely but represent traces of consistently low (and also negative) inflow. In the second year, we see a good agreement between the 24MS and the control run while the MTOM median remains closer to the observed. This is in indication of the benefit of second year forecasts. Both the 24MS and the control run use climatology in the out year and so in this case to not capture the pool elevation as well as the MTOM. It should be noted here that the 10th and 90th “traces” from the 24 month study are not traces in the hydrologic sense but (in the first year) are simply the 10th and 90th percentile of the inflow forecast.

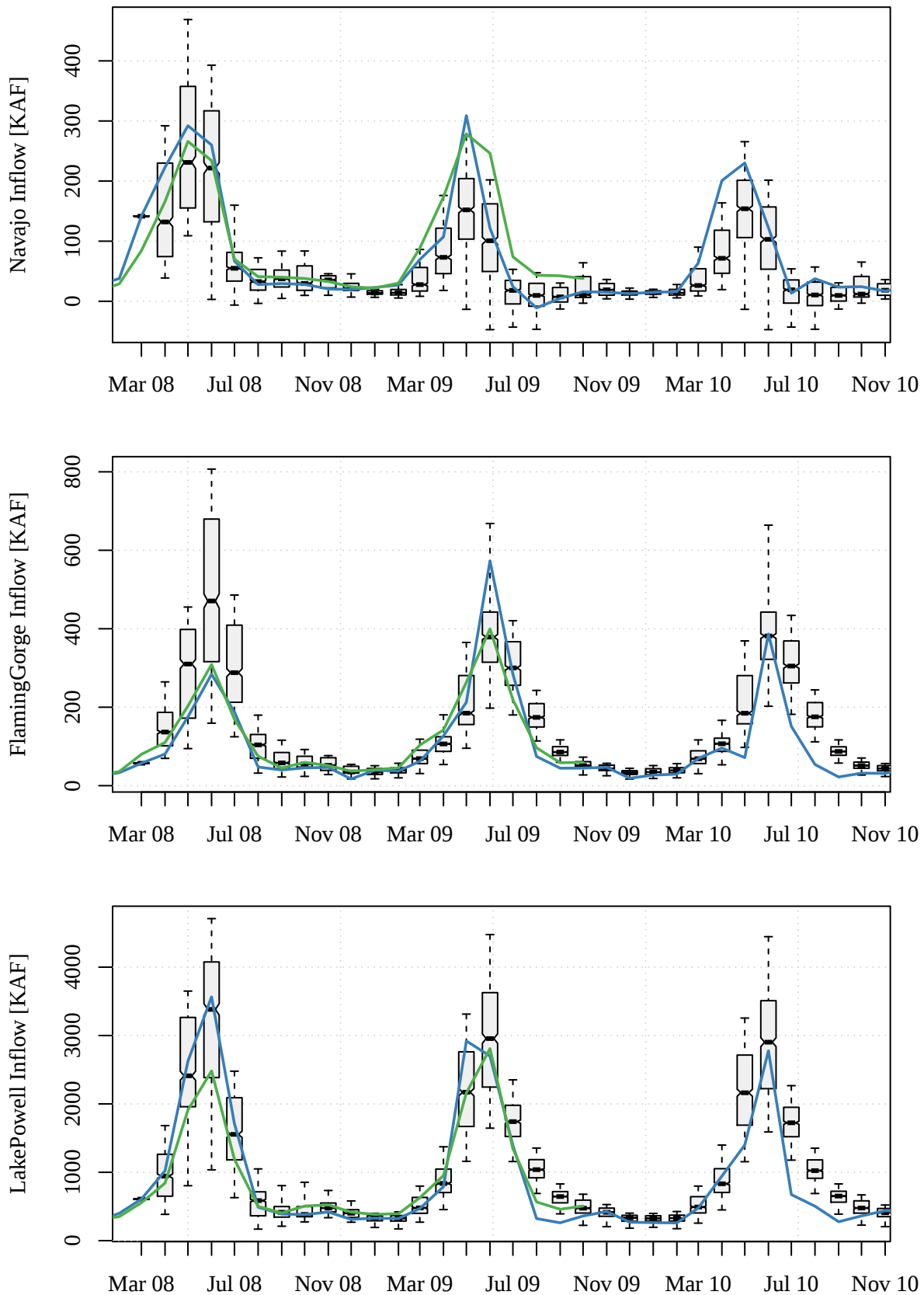


Figure 3.5: Unregulated inflow traces (shown as boxplots with whiskers extending to the 10th and 90th percentile of the data) for (a) April 1, Navajo, (b) January 1, Flaming Gorge and (c) November 1, Lake Powell. The green line represents the 24MS most probable trace and the blue line represents the observed unregulated inflow.

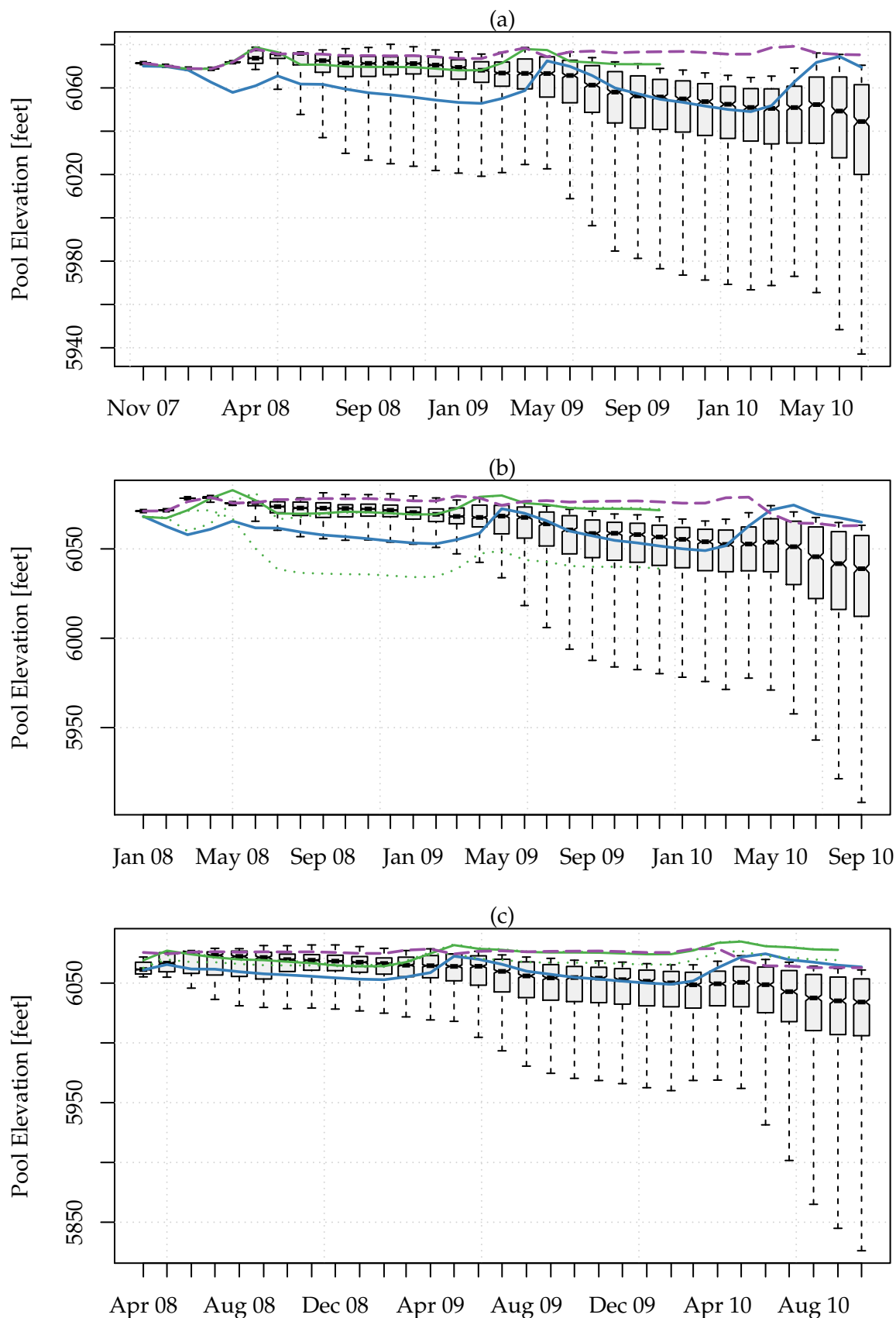


Figure 3.6: Navajo pool elevation for the (a) November, (b) January and (c) April lead times. Boxplots indicate the MTOM results, with whiskers extending to the 10th and 90th percentile of the data. The green line is the 24MS output, the dashed green lines are the 10th and 90th 24MS runs and the blue line is the observed.

Figure 3.6 shows Pool Elevation at Flaming Gorge reservoir. This is a case where runs starting in November and January were not able to capture the second year operations and the subsequent jump in pool elevation. Not until the April lead time do we see a shift in the forecasted pool elevation toward the observed. The inflow for the April run was over predicted at Flaming Gorge so we see a large errors in the pool elevation during the first year. Over or under prediction at some sites is a consequence of the disaggregation process. Disaggregation tends to favor overall system behavior, so errors can be introduced when a particular site diverges from the system as a whole.

Lake Powell pool elevation is shown in Figure 3.8. The MTOM, 24MS, control and observed tend to agree in the first year for the November and January runs. The control run nearly matches the median of the MTOM output for all three runs. This indicates that the main source of error here is operational rules. This does not mean that the rules that govern Lake Powell's release are the main source of error. Lake Powell is the basin outlet and so aggregates the error from operations every UC reservoir. This is further corroborated by the fact that the 2008 unregulated inflow forecast to Powell is exceptionally good (Figure 3.5).

In addition to the control run, outflow is a good indicator of the ability of the operational rules in the MTOM to reproduce operations. Figure 3.9 shows the outflow at Flaming Gorge reservoir. Figure 3.9 (a) and (b) show how error is introduced from the rules not reproducing reservoir outflows; The June peak of the first year is too high creating subsequent low base flows. In Figure 3.9 (c) the main source of error in the first year is the inflow forecast (as discussed earlier) though the outflow in the second year are reproduced very well.

One of the most telling results of the MTOM are shown in Figure 3.10. Boxplots are collapsed together because Powell's operations are broken into tiers. This set of plots clearly shows the benefit of improved second year forecasts in terms of reproducing the outflow magnitude at Powell. Operations in the out years are forecasted nearly exactly compared to observed while the 24MS over predicts outflow because climatology is used. It should be noted that the source of the difference between the 24MS and the MTOM in the out years is the actual inflow to Lake Powell. The combined error of all the UC reservoir operations cause the actual inflow to Powell to

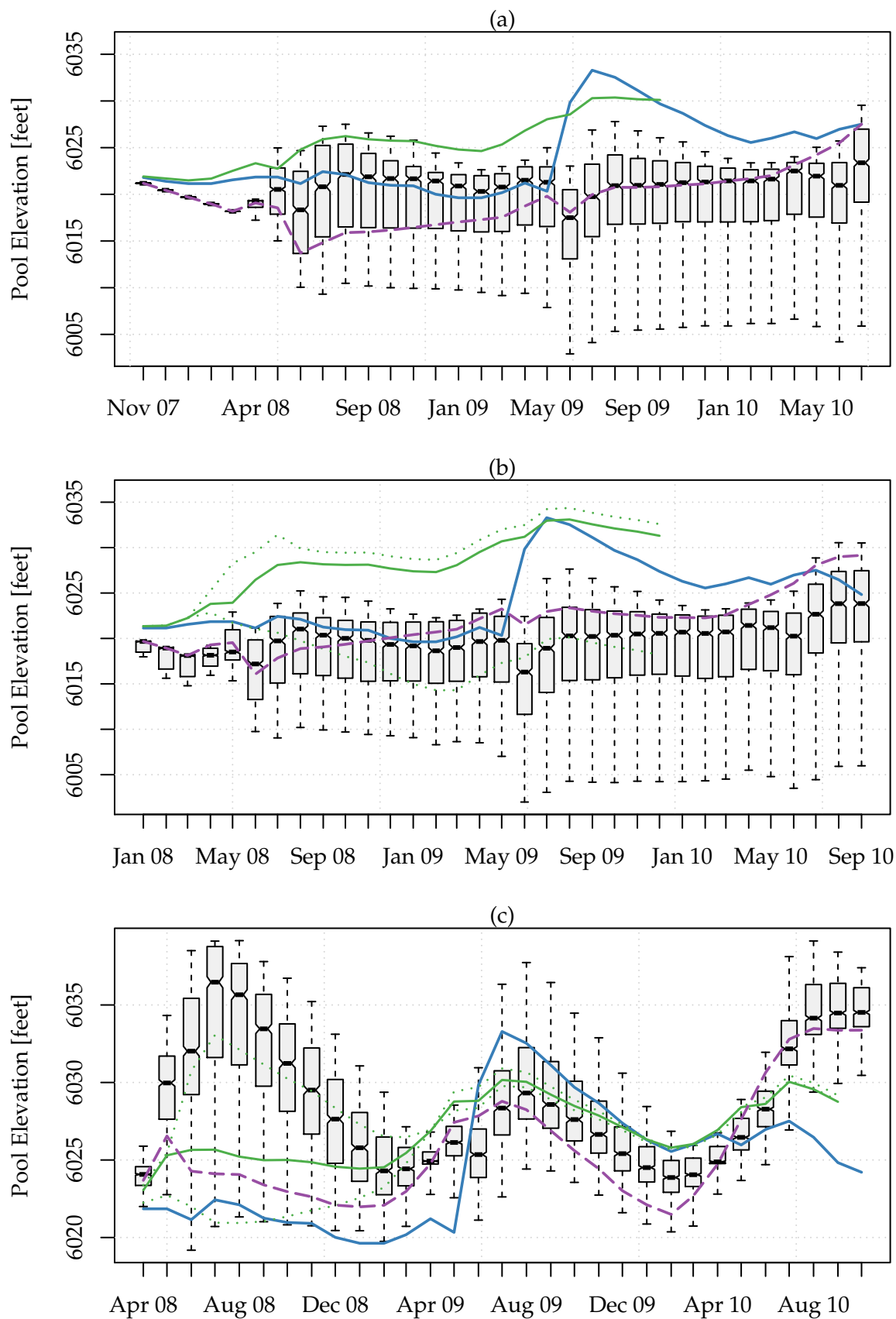


Figure 3.7: Same as Figure 3.6 but for Flaming Gorge Pool elevation.

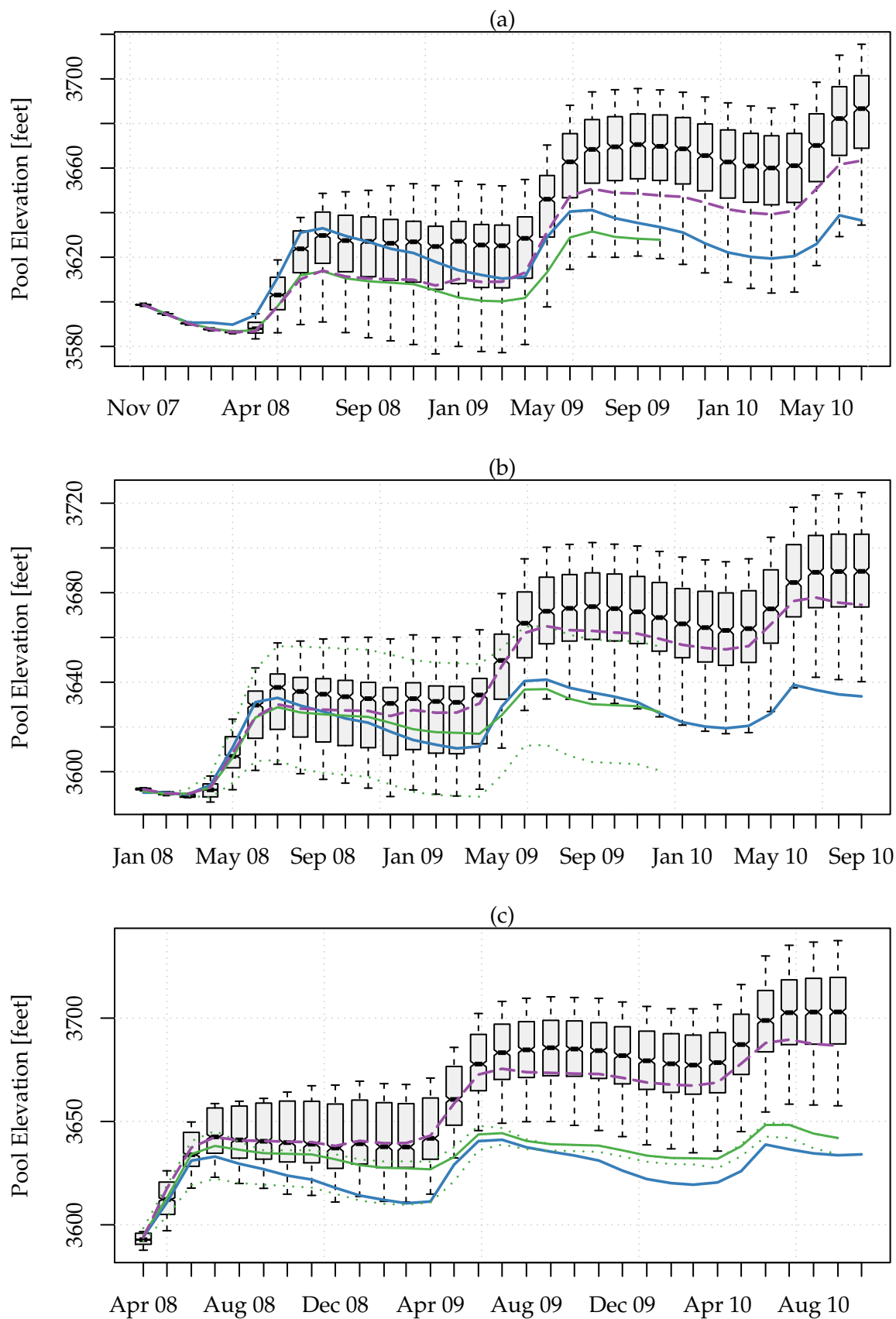


Figure 3.8: Same as Figure 3.6 but for Lake Powell Pool elevation.

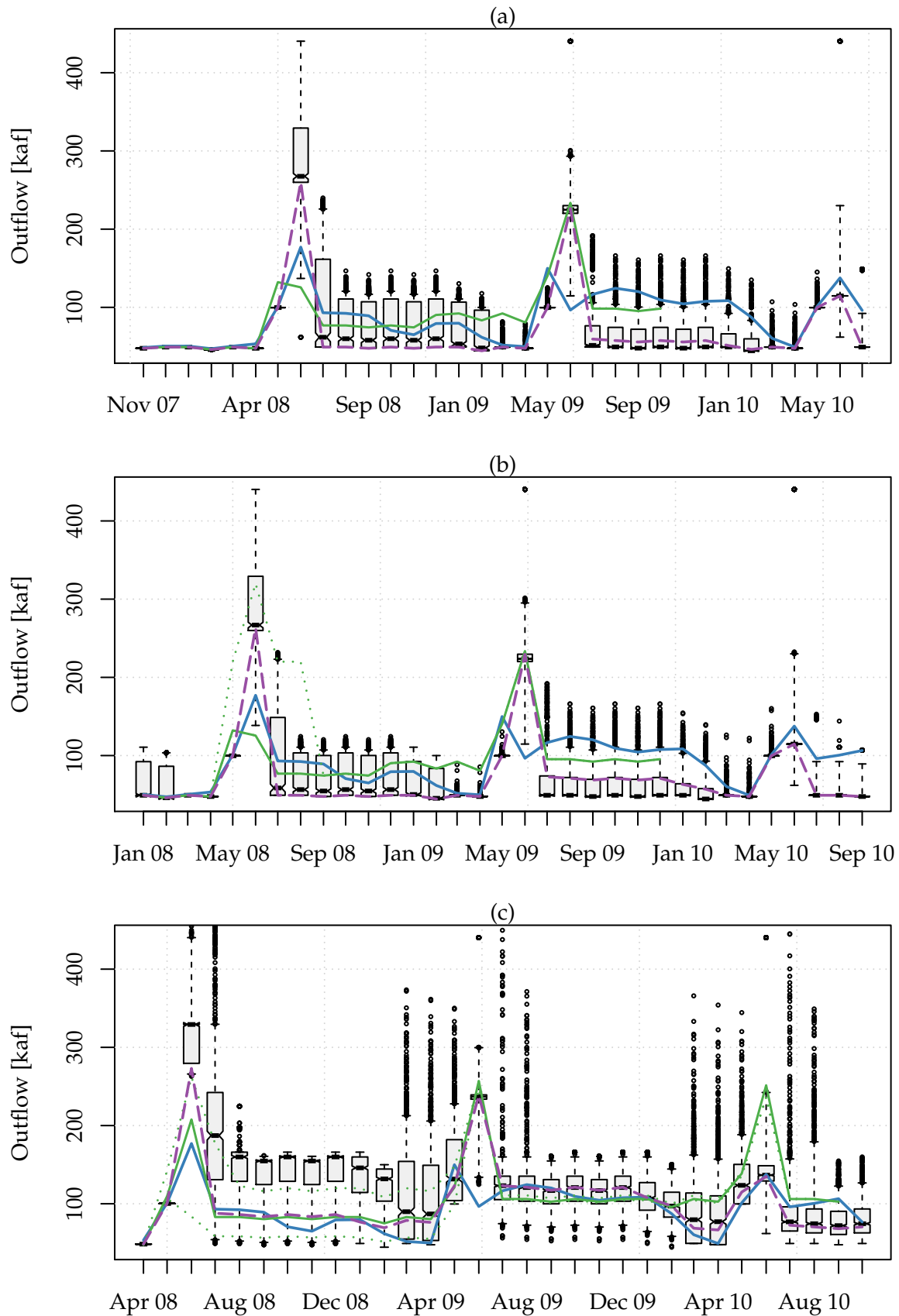


Figure 3.9: Same as Figure 3.6 but for Flaming Gorge Outflow.

be over predicted. The rules try to faithfully reproduce outflow, causing Powell's pool elevation to steadily rise (Figure 3.8).

3.6 Conclusions

We have presented a methodology for combining forecast methods and historical data to produce inflow traces that are suitable for input to a new forecast model for reservoir operations in the Upper Colorado River Basin (UCRB). The forecasts use an accurate seasonal forecast for the first year and Hidden Markov model forecast for the out years, combined with historical data and some existing forecasts to generate 32 months of input. The Midterm Operations Model (MTOM) has the ability to produce monthly ensemble forecasts of reservoir conditions.

We ran the model for start dates, November 1st 2007, January 1st 2008 and April 1st 2008. For these start dates the MTOM wide range of future hydrologic outcomes compared to the 24-Month Study. Ensemble output, for example, can be used to assess the probability of equalization or the probability of a shortage occurring.

We present a range of situations, some in which the 24MS clearly performs well, and others where the MTOM does. For example, the MTOM reproduced Lake Powell's outflow in the out years but because of compounded errors from all of the Upper Basin reservoirs poorly reproduced the pool elevation. In some situations (such as Flaming Gorge), for these particular model start dates, the inflow forecasts caused errors in reservoir releases as the rules tried to compensate for high inflows.

Overall the benefit of an accurate seasonal forecast is somewhat washed out by the skill in the initial conditions (all traces start out the same). We provide evidence that a skillful second year forecast can improve out year forecasts through the exact benefits of an improved forecast is difficult to pinpoint. In some cases, such as Flaming Gorge, it is relatively easy to identify the source of error as the inflow forecast in the first year. Likewise, the inaccuracy in Powell's pool elevation is likely due to a compounding of error in the entire Upper Basin. The control run helps to identify rule inaccuracy or the inflow forecast as the primary source of error but in most cases

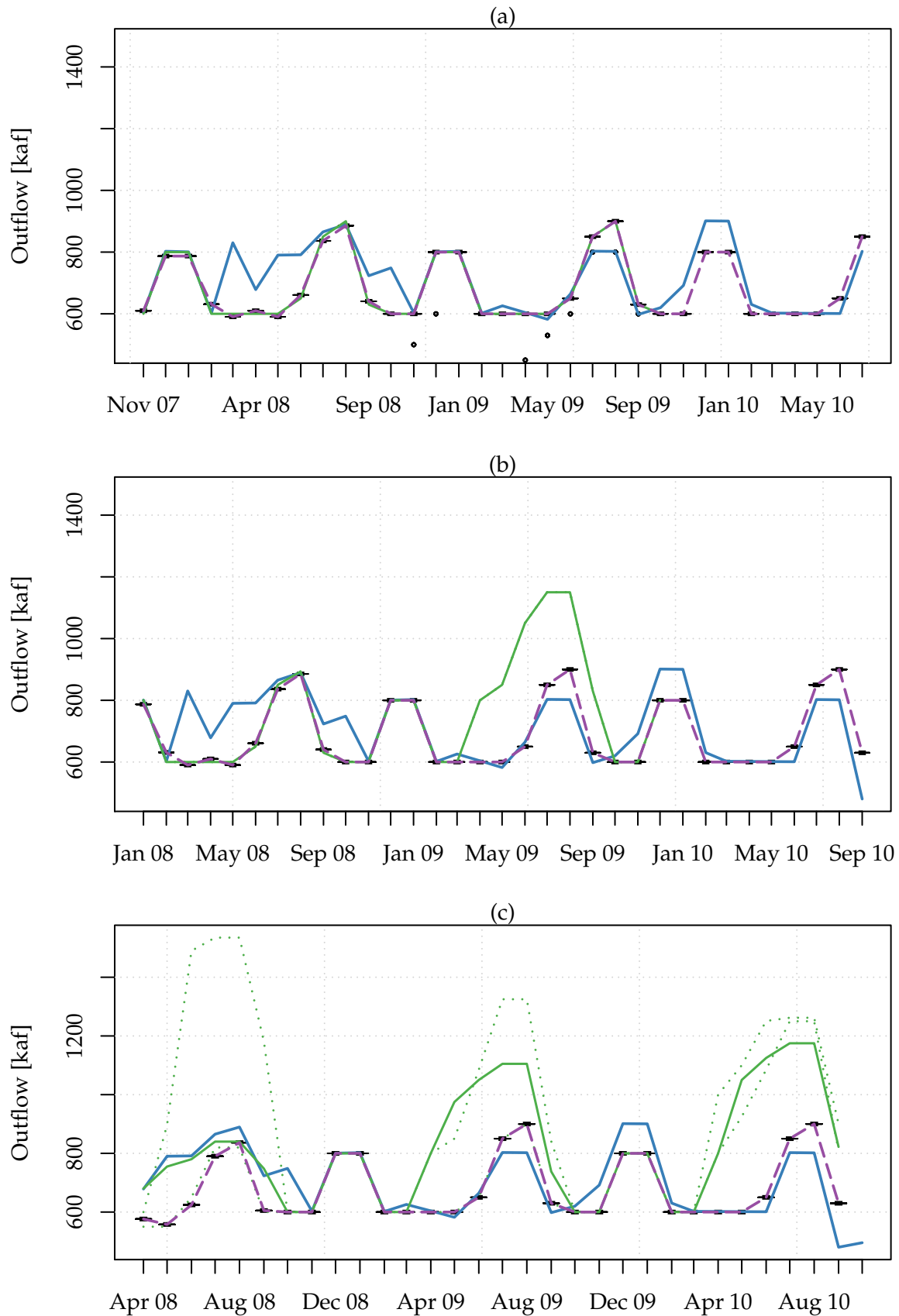


Figure 3.10: Same as Figure 3.6 but for Powell Outflow. Note that the control is identical to the 24MS in many months.

both causes cannot be separated.

These results indicate that the MTOM appears to be a valuable tool for probabilistically evaluating reservoir conditions in the UCRB. It should be stressed that this model was only run for three lead times (within a 5 month period). More evaluation is certainly necessary under many hydrologic scenarios. A rigorous evaluation of the operating rules can be done using the control run approach. Hypothetically if operations could be eliminated as a source of error, the benefit of improved forecasts could be examined more closely.

The most important future application of the MTOM is to improving reservoir operations. Eventually the model could be used to conduct sensitivity tests and optimization experiments for the purpose of better managing water in the UCRB.

Table 3.4: RPSS and MC (in parentheses) after disaggregation and drop-one cross validation for each lead time. At 95% confidence 0.21 is a significant correlation. Postive skills and significant correlations are shown in bold.

Node	April 1					Jan 1					Nov 1				
	April	May	June	July		April	May	June	July		April	May	June	July	
1	0.14 (0.24)	0.23 (0.57)	0.57 (0.75)	0.47 (0.62)		0.01 (0.09)	0.19 (0.44)	0.43 (0.61)	0.29 (0.46)		-0.14 (0.17)	0.07 (0.30)	0.13 (0.33)	0.18 (0.37)	
2	0.22 (0.28)	0.32 (0.64)	0.69 (0.77)	0.66 (0.63)		0.11 (0.15)	0.22 (0.47)	0.47 (0.67)	0.37 (0.49)		-0.04 (0.33)	0.10 (0.38)	0.03 (0.37)	0.23 (0.41)	
3	0.13 (-0.00)	0.26 (0.58)	0.56 (0.77)	0.62 (0.61)		-0.01 (-0.03)	0.09 (0.24)	0.26 (0.63)	0.41 (0.46)		0.01 (-0.04)	0.01 (-0.21)	0.08 (0.28)	0.26 (0.29)	
4	0.15 (0.45)	0.36 (0.67)	0.80 (0.78)	0.75 (0.65)		0.12 (0.39)	0.25 (0.52)	0.50 (0.64)	0.47 (0.49)		-0.03 (0.06)	0.14 (0.22)	0.01 (0.35)	0.18 (0.33)	
5	0.26 (0.48)	0.41 (0.64)	0.74 (0.77)	0.71 (0.62)		0.21 (0.42)	0.30 (0.56)	0.49 (0.64)	0.48 (0.45)		0.07 (0.09)	0.19 (0.31)	0.05 (0.37)	0.26 (0.34)	
6	0.13 (0.45)	0.44 (0.73)	0.77 (0.75)	0.80 (0.61)		0.05 (0.47)	0.32 (0.63)	0.42 (0.59)	0.55 (0.46)		-0.05 (0.33)	0.29 (0.44)	0.15 (0.39)	0.22 (0.36)	
7	0.40 (0.66)	0.40 (0.65)	0.51 (0.67)	0.51 (0.63)		0.21 (0.47)	0.29 (0.53)	0.32 (0.44)	0.41 (0.47)		0.09 (0.27)	0.26 (0.42)	0.31 (0.40)	0.38 (0.38)	
8	0.33 (0.52)	0.38 (0.71)	0.78 (0.78)	0.72 (0.64)		0.15 (0.51)	0.29 (0.60)	0.48 (0.65)	0.45 (0.50)		-0.03 (0.34)	0.13 (0.45)	0.08 (0.40)	0.35 (0.43)	
9	0.35 (0.55)	0.42 (0.62)	0.34 (0.50)	0.30 (0.50)		0.21 (0.41)	0.26 (0.44)	0.09 (0.16)	-0.08 (0.36)		0.07 (0.03)	0.07 (0.28)	-0.08 (-0.08)	-0.08 (0.17)	
10	0.45 (0.60)	0.37 (0.63)	0.45 (0.51)	0.40 (0.53)		0.37 (0.43)	0.18 (0.44)	0.20 (0.18)	-0.02 (0.36)		0.28 (0.10)	-0.04 (0.28)	0.01 (-0.07)	-0.02 (0.16)	
11	0.33 (0.61)	0.46 (0.74)	0.39 (0.57)	0.38 (0.52)		0.27 (0.48)	0.24 (0.61)	0.15 (0.28)	-0.01 (0.38)		0.18 (0.19)	0.16 (0.42)	0.00 (0.01)	-0.02 (0.22)	
12	0.23 (0.52)	0.43 (0.70)	0.68 (0.72)	0.57 (0.59)		0.03 (0.38)	0.34 (0.56)	0.42 (0.59)	0.36 (0.46)		-0.11 (0.30)	0.28 (0.42)	0.08 (0.37)	0.25 (0.38)	
13	-0.02 (0.22)	0.45 (0.63)	0.41 (0.68)	0.32 (0.43)		0.01 (0.17)	0.37 (0.52)	0.28 (0.66)	0.22 (0.27)		0.06 (0.30)	0.30 (0.47)	0.12 (0.45)	0.06 (0.04)	
14	0.40 (0.60)	0.66 (0.71)	0.48 (0.72)	0.43 (0.44)		0.24 (0.56)	0.32 (0.61)	0.30 (0.57)	0.32 (0.40)		0.19 (0.38)	0.19 (0.49)	0.25 (0.39)	0.22 (0.24)	
15	0.04 (0.40)	0.30 (0.65)	0.63 (0.76)	0.74 (0.59)		0.10 (0.35)	0.24 (0.55)	0.36 (0.66)	0.45 (0.50)		0.10 (0.39)	0.14 (0.45)	0.10 (0.46)	0.39 (0.49)	
16	0.38 (0.60)	0.60 (0.82)	0.50 (0.77)	0.53 (0.62)		0.20 (0.50)	0.44 (0.67)	0.33 (0.61)	0.08 (0.50)		0.06 (0.31)	0.20 (0.52)	0.14 (0.35)	0.04 (0.33)	
17	0.35 (0.52)	0.37 (0.58)	0.80 (0.73)	0.14 (0.54)		0.07 (0.45)	0.29 (0.54)	0.37 (0.60)	0.10 (0.42)		0.00 (0.28)	0.26 (0.38)	0.19 (0.40)	0.10 (0.36)	
18	0.24 (0.54)	0.36 (0.63)	0.36 (0.68)	0.36 (0.57)		-0.06 (0.41)	0.21 (0.51)	0.11 (0.46)	0.25 (0.44)		-0.33 (0.17)	0.11 (0.31)	0.01 (0.25)	0.20 (0.18)	
19	0.35 (0.56)	0.27 (0.62)	0.38 (0.69)	0.55 (0.58)		0.04 (0.39)	0.22 (0.51)	0.27 (0.43)	0.31 (0.46)		-0.24 (0.22)	0.17 (0.30)	0.22 (0.30)	0.23 (0.30)	
20	0.27 (0.58)	0.49 (0.79)	0.81 (0.82)	0.74 (0.66)		0.23 (0.55)	0.31 (0.66)	0.55 (0.69)	0.38 (0.53)		0.17 (0.34)	0.30 (0.50)	0.10 (0.45)	0.33 (0.41)	

Table 3.5: Median Correlation (MC) and Median Ranked probability skill score (RPSS) of forecasts for 1980–2010. Dry RPSS, Ave RPSS and Wet RPSS represent the median RPSS for years in the lower, middle and upper tercile respectively.

Model	MC	RPSS	Dry RPSS	Ave RPSS	Wet RPSS
HM2	0.31	0.21	0.26	−0.11	0.05
AR1	0.07	−0.03	0.07	−0.03	−0.25

Bibliography

- Akntuž, B., and P. F. Rasmussen (2005), A Markov switching model for annual hydrologic time series, *Water Resources Research*, 41.
- Bracken, C., B. Rajagopalan, and J. Prairie (2010), A multisite seasonal ensemble streamflow forecasting technique, *Water Resources Research*, 46(3), W03,532.
- Bracken, C., B. Rajagopalan, and E. Zagona (2011), A Nonstationary Hidden Markov Model for Stochastic Streamflow Simulation and Short Term Forecasting in the Upper Colorado River Basin, *Water Resources Research*, submitted.
- Brown, D., and A. Comrie (2004), A winter precipitation 'dipole' in the western United States associated with multidecadal ENSO variability, *Geophys. Res. Lett.*, 31(9), –.
- Bureau of Reclamation (2007), Final environmental impact statement Colorado River interim guidelines for lower basin shortages and coordinated operations for lakes Powell and Mead, p. 6 Chapters.
- Bureau of Reclamation (2008), Upper Colorado River Basin Consumptive Uses and Losses Report: 2006–2010, pp. 1–17.
- Clark, M. P., M. C. Serreze, and G. J. McCabe (2001), Historical effects of El Nino and La Nina events on the seasonal evolution of the montane snowpack in the Columbia and Colorado River Basins, *Water Resources Research*, 37(3), 741–757.
- Craven, P., and G. Wahba (1979), Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation, *Numerical Mathematics*, 31, 377–403.
- Dracup, J. A., and E. Kahya (1994), The relationships between U.S. streamflow and La Niña Events, *Water Resources Research*, 30(7), 2133.
- Fulp, T. (2005), How low can it go, *Southwest Hydrology*, pp. 16–28.
- Gelati, E., H. Madsen, and D. Rosberg (2010), Markov-switching model for nonstationary runoff conditioned on El Niño information, *Water Resources Research*.
- Grantz, K., B. Rajagopalan, M. Clark, and E. Zagona (2005), A Technique for Incorporating Large-Scale Climate Information in Basin-Scale Ensemble Streamflow Forecasts, *Water Resources Research*, p. W10410.

- Hamlet, A., and D. Lettenmaier (1999), Columbia River streamflow forecasting based on ENSO and PDO climate signals, *Journal of . . .*
- Hidalgo, H., and J. Dracup (2003), ENSO and PDO effects on hydroclimatic variations of the Upper Colorado River basin, *Journal of Hydrometeorology*, 4(1), 5–23.
- Hunter, T., G. Tootle, and T. Piechota (2006), Oceanic-atmospheric variability and western U.S. snowfall, *Geophys. Res. Lett.*, 33(13).
- Jackson, B. B. (1975), Markov mixture models for drought lengths, *Water Resources Research*, 11(1), 64.
- Kahya, E., and J. Dracup (1993), US streamflow patterns in relation to the El Nino/southern oscillation, *Water Resources Research*, 29, 2491–2503.
- Kahya, E., and J. Dracup (1994), The influences of type 1 El Nino and La Nina events on streamflows in the Pacific southwest of the United States, *Journal of Climate*.
- Lall, U. (1995), Nonparametric Function Estimation: Recent Hydrologic Contributions, *Reviews of Geophysics*, pp. 1093–1099.
- Loader, C. (1999), *Local Regression and Likelihood*, Statistics and Computing, Springer, New York.
- Maurer, E., and D. Lettenmaier (2004), Variability and potential sources of predictability of North American runoff, *Water Resources Research*.
- McCabe, G., and M. Dettinger (1999), Decadal variations in the strength of ENSO teleconnections with precipitation in the western United States, *International Journal of Climatology*.
- McCabe, G., and M. Dettinger (2002), Primary modes and predictability of year-to-year snowpack variations in the western United States from teleconnections with Pacific Ocean climate, *Journal of Hydrometeorology*.
- Mehrotra, R., and A. Sharma (2005), A nonparametric nonhomogeneous hidden Markov model for downscaling of multisite daily rainfall occurrences, *J. Geophys. Res.*, 110(D16), D16,108.
- Nowak, K., J. Prairie, B. Rajagopalan, and U. Lall (2010), A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow, *Water Resources Research*, 46(8), W08,529.
- Opitz-Stapleton, S., S. Gangopadhyay, and B. Rajagopalan (2007), Generating streamflow forecasts for the Yakima River Basin using large-scale climate predictors, *Journal of Hydrology*, 341, 131–143.
- Piechota, T., J. Dracup, and R. Fovell (1997), Western US streamflow and atmospheric circulation patterns during El Nino Southern Oscillation, *Journal of Hydrology*, 201, 249–271.
- Prairie, J., and R. Callejo (2005), Natural Flow and Salt Computation Methods, Calendar Years 1971–1995, *Bureau of Reclamation*, pp. 1–112.
- Prairie, J., B. Rajagopalan, U. Lall, and T. Fulp (2007), A stochastic nonparametric technique for space-time disaggregation of streamflows, *Water Resources Research*, 43, 1–10.

- Rajagopalan, B., K. Nowak, J. Prairie, M. Hoerling, B. Harding, J. Barsugli, A. Ray, and B. Udall (2009), Water supply risk on the Colorado River: Can management mitigate?, *Water Resources Research*, 45(8), W08,201.
- Rajagopalan, B., J. Salas, and U. Lall (2010), Stochastic methods for modeling precipitation and streamflow, in *Advances in Data-based Approaches for Hydrologic Modeling and Forecasting*, edited by B. Sivakumar and R. Berndtsson, World Scientific, Singapore.
- Regonda, S. K., B. Rajagopalan, M. Clark, and E. Zagona (2006), A multimodel ensemble forecast framework: Application to spring seasonal flows in the Gunnison River Basin, *Water Resources Research*, 42, W09,404.
- Soukup, T. L., O. A. Aziz, G. A. Tootle, T. C. Piechota, and S. S. Wulff (2009), Long lead-time stream-flow forecasting of the North Platte River incorporating oceanic-atmospheric climate variability, *Journal of Hydrology*, 368, 131–142.
- Thyer, M., and G. Kuczera (2000), Modeling long-term persistence in hydroclimatic time series using a hidden state Markov model, *Water Resources Research*, 36(11), 3301–3310.
- Zucchini, W., and P. Guttorp (1991), A hidden Markov model for space-time precipitation, *Water Resources Research*, 27(8), 1917–1923.

Final Remarks and Future Work

This thesis attempts to address many relevant questions related to water availability and management in the Upper Colorado River Basin (UCRB). The main topics covered were seasonal forecasting, interannual forecasting, annual flow simulation, and water supply management modeling.

Chapter 1 presented a proof-of-concept framework for generating accurate long lead forecasts of April-July runoff at four key sites in the UCRB. Current forecast for this region currently only begin in January when the first snowpack information is widely available. Using climate information in addition to the typical snowpack and soil moisture data, forecasts were made at starting in November. This work was promising but many questions remained, namely could this framework be expanded to additional sites and could these forecasts be made useful in an operations model. These questions were subsequently addressed in Chapter 3.

Chapter 2 presented a Hidden Markov (HM) model for simulation and forecasting annual flows. The HM model was able to reveal some remarkable persistent regime-switching behavior in the Lees Ferry natural flow time series. Using the best fit HM models with Gamma and Normal component distributions, simulations of annual flow were made. The simulations suggested a high chance of long drought lengths, relative to an AR model, indicating that the 2000-2010 drought may not be as unlikely as previously thought. The HM model was also used to produce annual flow forecasts and were shown to have skill over climatology and over the best fit AR model.

There are many potential opportunities to expand and improve on the HM model devel-

oped in Chapter 2. Global decoding sequences can be related to atmospheric variables, providing insight into the driving mechanisms of the system. Models can also be developed that use exogenous input (such as ENSO) in the model fitting process.

Chapter 3 introduced a new model for probabilistic reservoir operations in the CRB known as the Midterm Operations Model (MTOM). This model is designed to be run without the input of reservoir operators. Currently operators in the CRB must manually input reservoir outflows into a model known as the 24 Month Study. A methodology was developed to generate ensemble input to the MTOM using a combination of the Seasonal forecast from Chapter 1 and the HM forecasts from Chapter 2. The MTOM was run for three lead times. The ability of the model to reproduce operations as well as forecast reservoir conditions was assessed. The results were varied; The benefit of the seasonal forecast model was difficult to distinguish from the skill in the initial reservoir conditions. In some cases the second year inflow forecast helped produce better reservoir forecasts in the out years. In many cases the sources of error were a combination of errors from operating rules (outflows) and inflow forecast errors.

The next step of the MTOM is a rigorous assessment of the sources of error in the operating rules. Only then can the benefit of improved forecasts be determined precisely. The MTOM also has the potential to become a tool for optimizing CRB reservoir operations.

Appendix A

Chapter 1 supporting materials

Three images are included in these auxiliary materials to accompany Chapter 1. They are not crucial to the paper but help to strengthen the validity of method. The first two files were generated using the on-line correlation tool developed by Physical Sciences Division, National Oceanic and Atmospheric Administration (<http://www.cdc.noaa.gov/Correlation/>). The third file shows cross correlations that were calculated from disaggregated ensembles from the framework described in the paper.

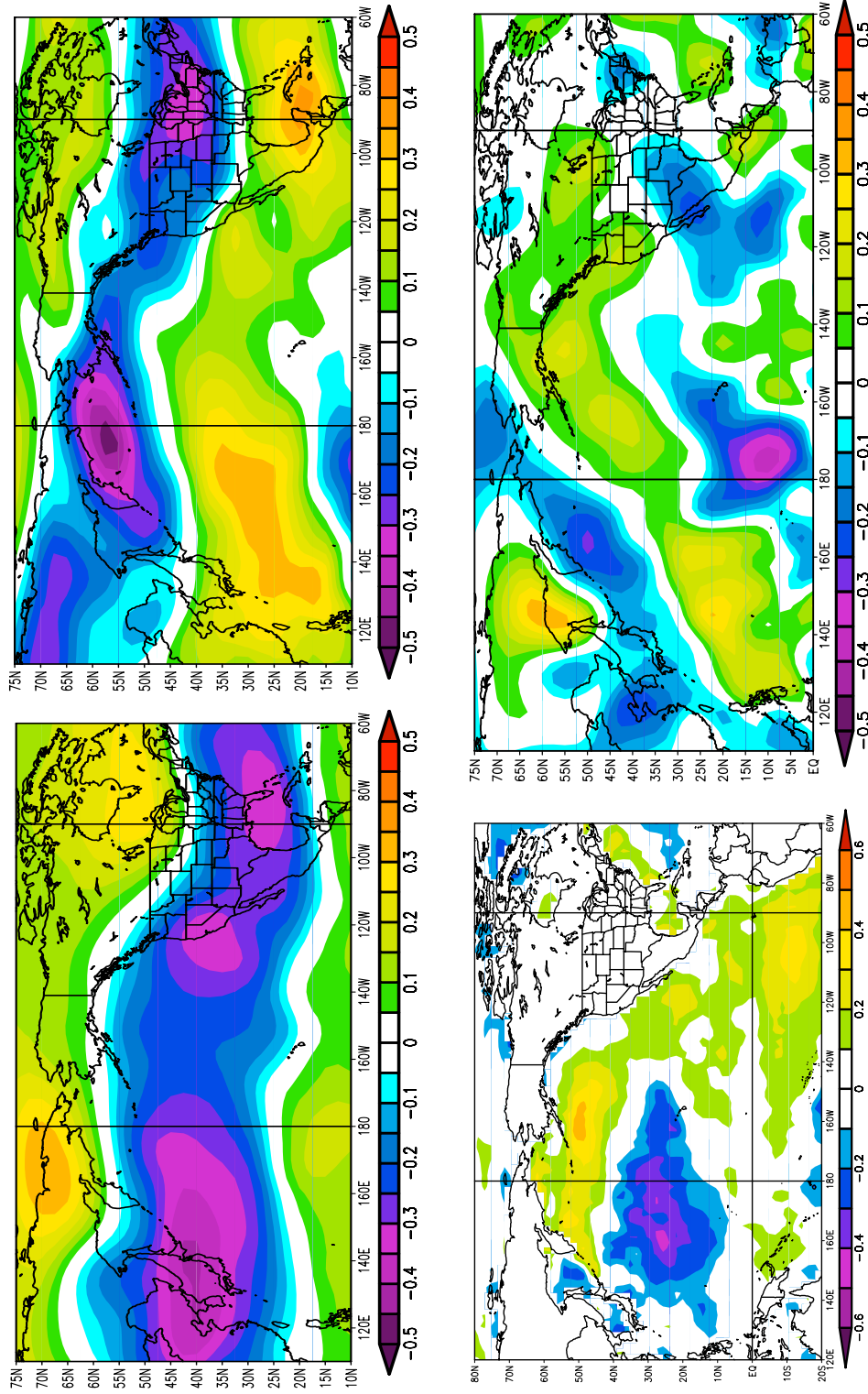


Figure A.1: Correlation maps of index gage spring (Apr-Jul) streamflow with Feb.-Mar. GPH (top left), March ZW (top right), Jan.-Mar. SST (bottom left) and Mar. MW (bottom right).

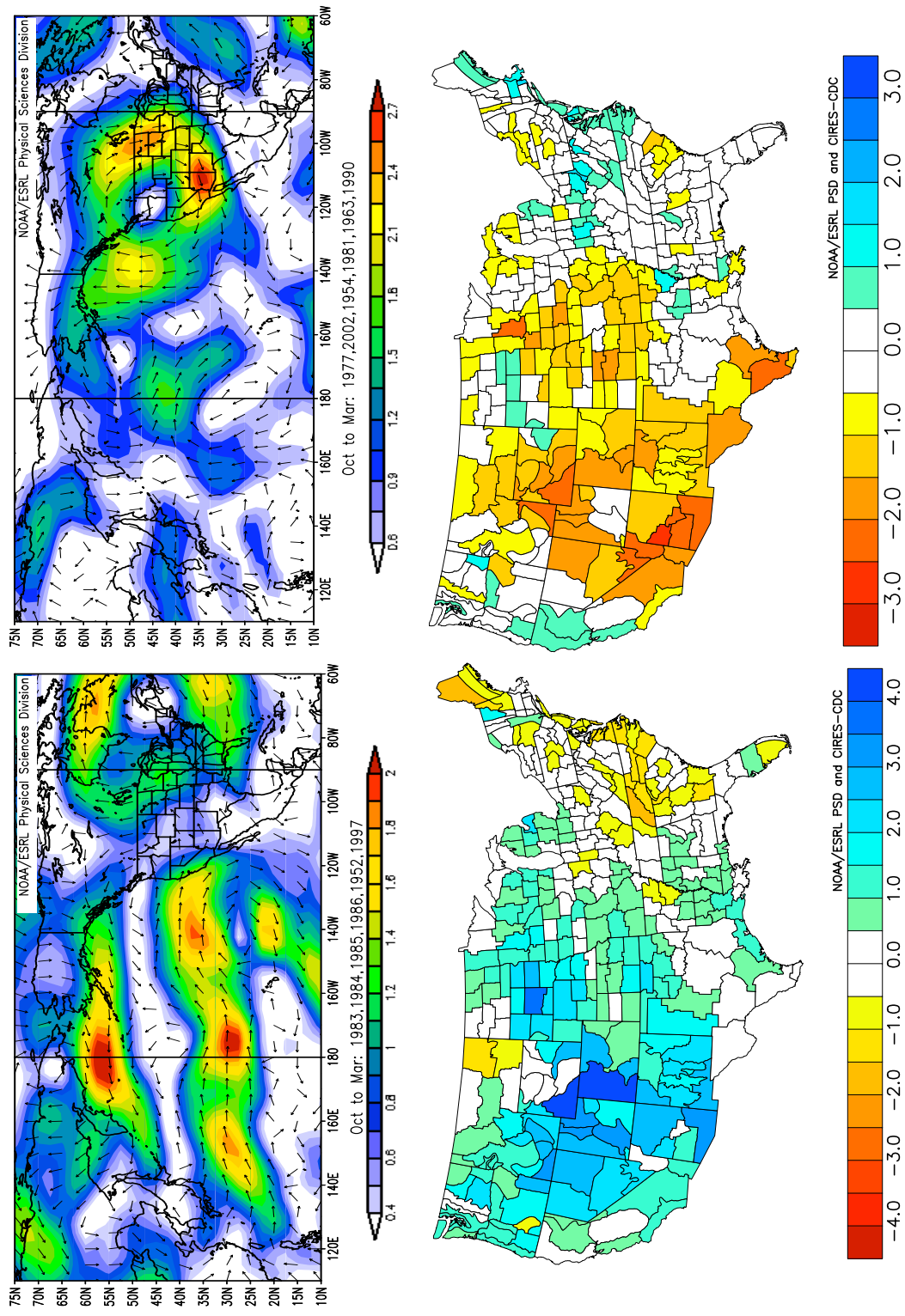


Figure A.2: Composite maps of Oct.-Mar. 500 mb vector winds for 6 wettest years (top left) and 6 driest years (top right) years and preceding Fall PDSI for wet (bottom left) and dry (bottom right) years.

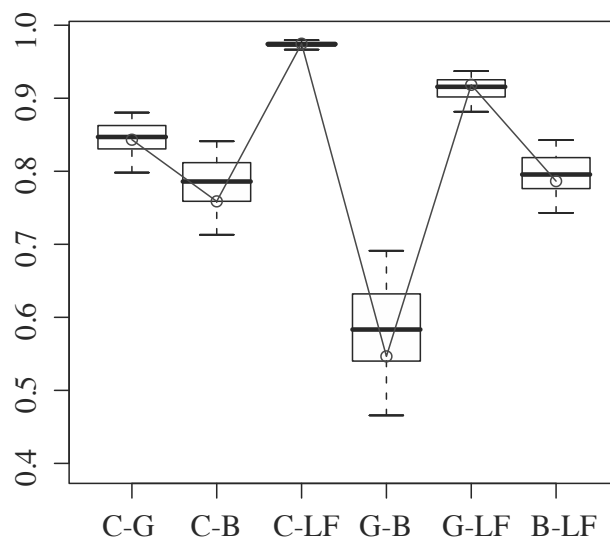


Figure A.3: Boxplots of cross correlation of spring season streamflow forecasts issued on Apr 1st between sites: C=Cisco, B=Bluff, G=GRUT, LF=Lees Ferry. The solid line represents the observed statistics.

Appendix B

Chapter 3 supporting materials

B.1 Seasonal Forecast Verification

The predictors in the various “best” multi-models at all the lead times are shown in Table . It can be seen that the number of “best” models decrease with decrease in lead time. For example, for March 1st and April 1st forecast have two models while the earlier lead times have five models. This is consistent in that at long lead times when basin snow information is absent or partial at best, predictability is obtained mainly through climate variables, while at March and April 1st the snow information is complete and hence, the role of climate variables is minimal. Also of interest is that the PDSI enters the best models for shorter lead times, the reasoning is that it captures the role of antecedent land surface, such that, warmer preceding summer and fall conditions can enable increased losses and thus reduce streamflow. This role of land surface was also identified by *Regonda et al.* (2006) in the Gunnison River Basin.

As a representative sample, we display the results for the location Gunnison River near Grand Junction (Site 6), results from other locations are quite similar. Also this site was chosen because it was not presented in the four sites used by Bracken et al. (2010) and it is an important site for water resources planning in the UCRB. Figure B.1 shows the boxplot of ensemble forecast of the peak season volume at the three lead times along with the historical natural flow (in red). The skill scores are shown in the figure captions and it can be seen that there is a significant positive skill for the seasonal flow forecast issued on November 1st (about six months lead time) which increases as the lead time decreases. Figure B.2 shows the ensemble forecast of June (the peak

month in the season) streamflow at the same lead times and here too the skills for Nov 1st forecast are quite impressive. These plots show the general pattern seen across almost all the sites, that is, ensemble variability increases and skill decreases as lead time increases. Furthermore, they also capture the high and low flow variability quite well, especially for Jan 1st and Apr 1st forecasts, which are very useful for management. Figure B.3 shows the forecasts in retroactive mode, where in the model fitting is updated for each year with the most recent data, this is the approach used by CBRFC in their forecasting framework using watershed models. Also shown in this figure are the CBRFC coordinated forecasts for comparison. The forecasts from our multi-model ensemble performs very well in comparison to the CBRFC forecasts, also, it captures the high and low flow years very well.

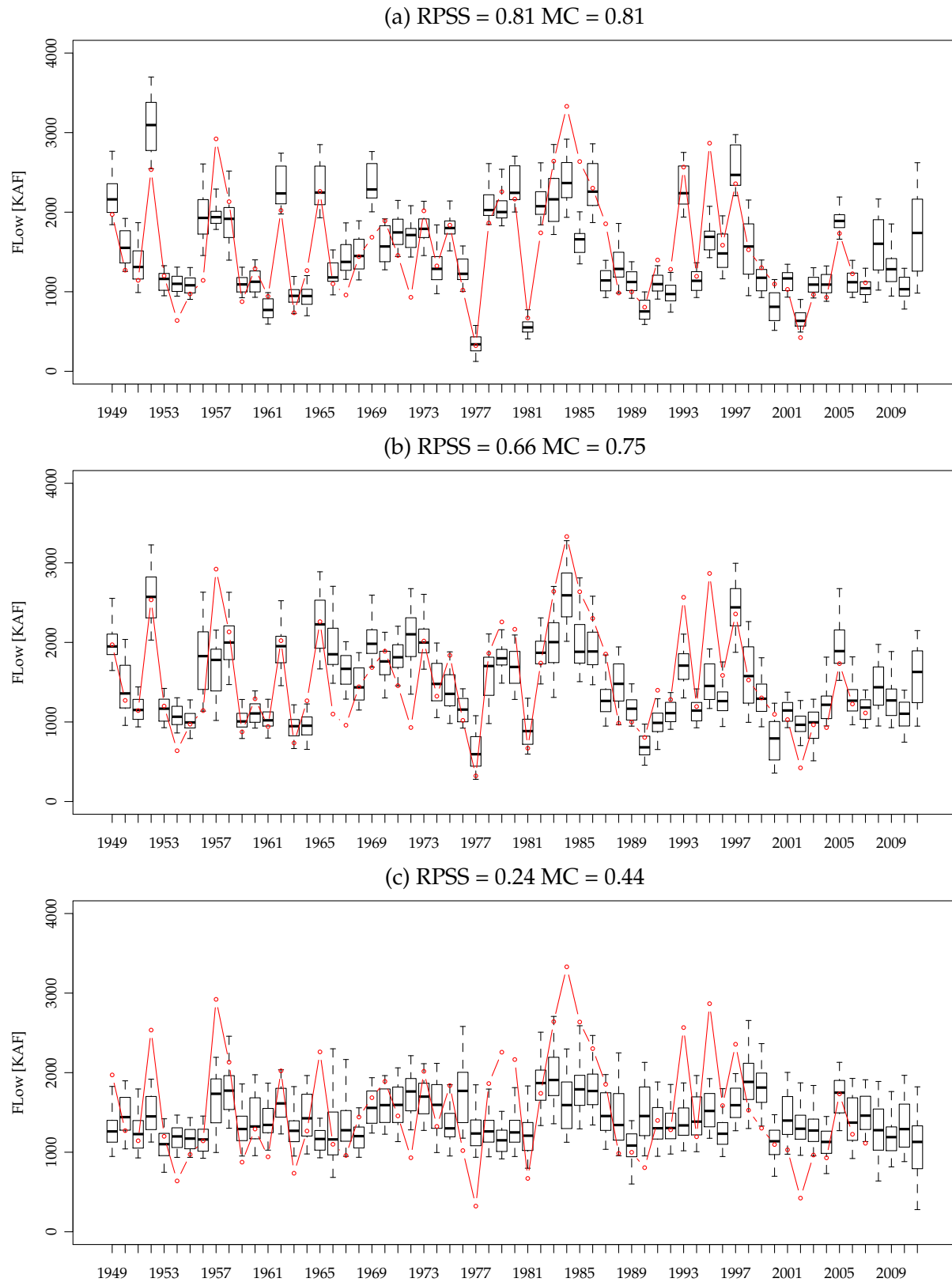


Figure B.1: Sample ensemble forecasts for (a) Apr1, (b) Feb1 and (c) Nov1 seasonal flow volumes at Gunnison River Near Grand Junction (Site 6). Boxplots extend to the 5th and 95th percentiles of each ensemble. MC stands for the median correlation, the correlation of the median of the ensembles with the historical record.

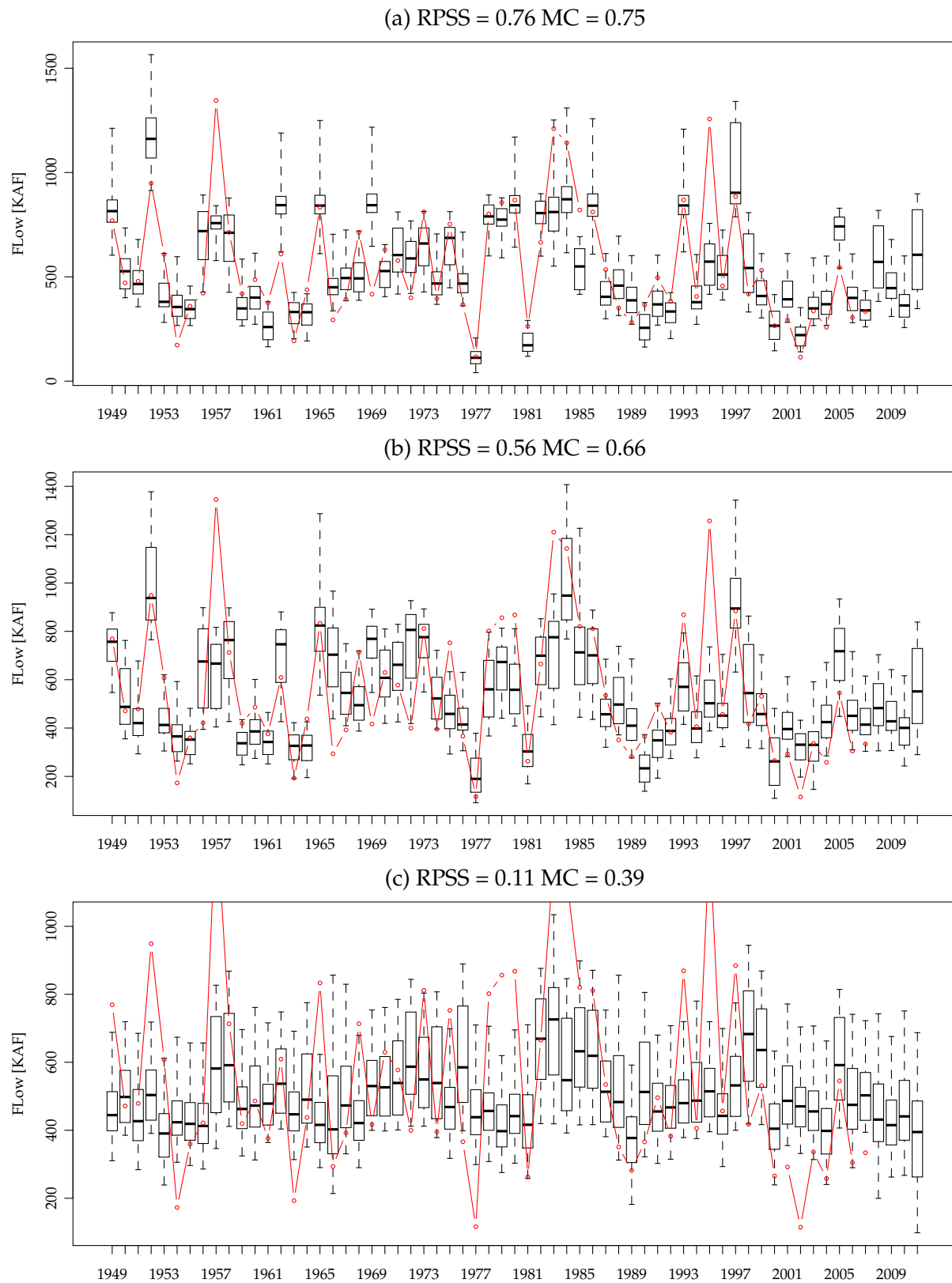


Figure B.2: Same as Figure 3 but for June flows.

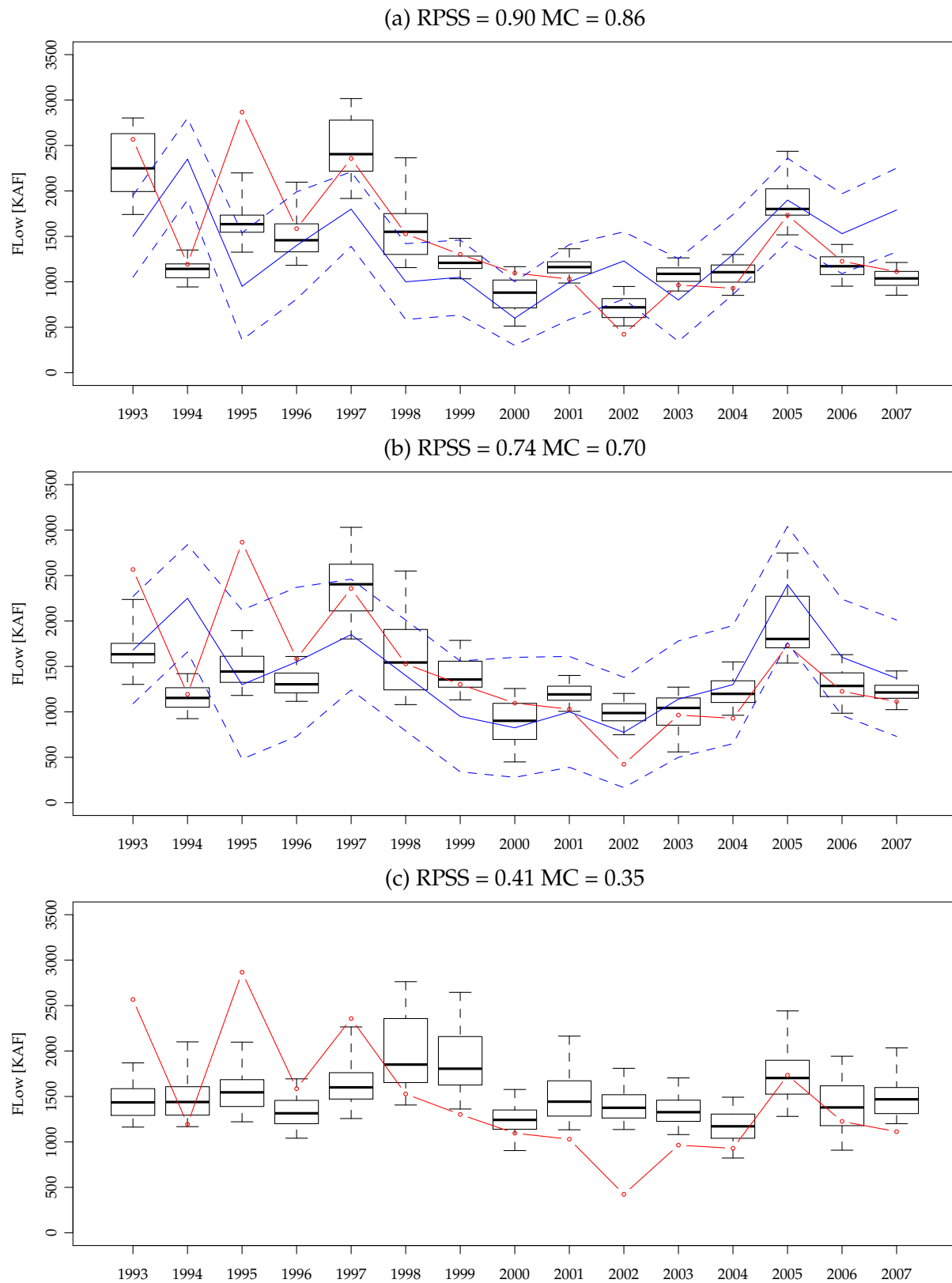


Figure B.3: Same as Figure B.1 but for retroactive forecasts overlaid with the CBRFC coordinated forecast where available. The blue solid lines are the 10th and 90th percentile and the dashed blue line is the 50th percentile of the coordinated forecast.

Table B.1: Selected Models for each lead time

	GPH	ZNW	MDW	SST	PDSI	SWE
November 1						
1	0	0	1	1	0	0
2	0	1	0	0	0	0
3	0	0	0	0	1	0
4	0	0	0	1	0	0
5	0	0	1	0	0	0
6	1	0	0	0	0	0
January 1						
1	0	0	0	1	0	1
2	0	0	0	0	0	1
3	0	0	0	0	1	0
4	0	0	1	0	0	0
5	0	0	0	1	0	0
February 1						
1	0	0	0	1	0	1
2	0	0	0	0	0	1
3	0	0	0	0	1	0
March 1						
1	0	0	0	1	0	1
2	0	0	0	0	0	1
April 1						
1	0	0	0	1	0	1
2	0	0	0	0	0	1