International Conference "Water, Environment, Energy and Society" (WEES–2009)
New Delhi, 12–16 January 2009

Coping with Model Structural Uncertainty in Medium Term Probabilistic Streamflow Forecasting Applications

Ashish Sharma
School of Civil and Environmental Engineering
The University of New South Wales
Sydney, NSW 2052, AUSTRALIA
E-mail: a.sharma@unsw.edu.au

ABSTRACT: This paper describes the rationale behind the Multiple Reservoir Inflow Forecasting System and how it is specified to issue 12 months long probabilistic forecasts across selected reservoir systems in Australia. The forecasting system enables development of 12 months long monthly flow scenarios (probabilistic forecasts) issued based on a mix of prevalent catchment and climate conditions, so as to simulate the likely evolution of the reservoir network over the coming 12 months for medium term planning and operations. The probabilistic forecasting system uses previous season flow as an indicator of catchment influences and selected sea surface temperature anomalies across varying locations and time lags as climatic indicators in its formulation. Given the considerable uncertainty that exists in choosing relevant climatic indicators, a flexible model structure (or multiple climatic indicators) is adopted, leading to the formulation of multiple predictive models, each aiming to represent the impact different climate phenomena have on the responses. Examples of such multiple models could be an El Nino Southern Oscillation indicator at a short time lag for one model, and a combination of the Inter-decadal Pacific Oscillation and a lagged Indian Ocean dipole influence for another model. The inflow scenarios that are finally predicted thus come from a collection of such models, thereby enabling an appropriate representation of the structural uncertainty that exists in this prediction problem. The above forecasting system is illustrated using two case studies representing the Sydney Catchment Authority water supply system and the Tasmanian hydroelectric system. The advantages and drawbacks of using the above mentioned model averaging logic are discussed, and modifications suggested that can improve the predictive performance and consequent reservoir operations.

INTRODUCTION

Medium to long-term probabilistic forecasting of rainfall or streamflow involves predicting (often as a conditional probability distribution) the responses being studied as a function of lead time, conditional to relevant climatic and catchment forcings. Examples of such probabilistic forecast methods are many, including those reported in (Sharma, 2000a; Sharma, 2000b; Sharma *et al.*, 2000), most such approaches providing users with the probability with which the response may fall in a designated category (such as "low", "medium" or "high") or with which it may exceed a specified threshold. A limitation of such forecasts is that they cannot be used to ascertain the risks associated with undertaking longer-term water planning. For such planning to be done, the probabilistic forecasts must link across time (to the maximum lead time they are designed for) and across the multiple nodes of the water system they represent. This allows the forecasts to be used as representative scenarios or realisations indicating how inflows may evolve over the medium to long-term planning horizon, thereby enabling water managers to identify operating policies that maximise profits and minimise risks of system failure.

This paper presents a probabilistic forecasting procedure termed the Multiple Reservoir Probabilistic Inflow forecasting System (MRFLO) that issues multiple realisations or scenarios of seasonal and monthly inflows to all nodes of a reservoir system from defined starting points for the coming twelve months. These realisations are issued taking into account the dominant climatic influences on the evolution of the streamflow in the catchment, as well as the streamflow at preceding lags to represent the persistence introduced due to the catchment storage depletion mechanism. As identification of a definite climate forcing is difficult for most catchments in the world, there is considerable uncertainty in specifying a unique predictive modeling form. A model combination or a model averaging alternative is proposed here as the basis for representing this model structural uncertainty, the individual models being formulated using a semi-independent climatic forcings, so as to enable representation of the full uncertainty one would expect in the resulting inflows. It is seen that the use of these multiple models results in more stable and consistent forecasts being issued, as compared to the case where a single model were being used. However, the combination of multiple models also results in the predictions tending to converge towards the mean observed flow value, a consequence explained using the Central Limit Theorem, as a result of which there is a tendency of low flows being overestimated and high flows underestimated on an expected value basis.

The sections that follow describe the basis of the method in greater detail and outline its application for the multiple reservoir systems operated by Hydro-Tasmania and by the Sydney Catchment Authority, two catchment systems that have rather different climatic controls and predictability associated with them. The paper is organised as follows. The next section details the logic implemented in formulating MRFLO. The section that follows sets out the details of the two applications and the methods that were used to test the suitability of the results. The next section details the results from the applications, which is followed by the conclusions drawn from the work.

METHODOLOGY

An outline of the framework used in formulating the probabilistic forecasting model is presented below. The framework consists of three main steps. The first step involves formulating key aggregate system response indicators that can be presumed to be influenced by climatic forcings that are substantially different from one another. The second step involves formulation of the probabilistic forecasting model, which entails identification of predictor variables from a list of candidate predictors that are deemed to influence the seasonal aggregate response variable under study, and, the representation of spatio-temporal dependence in the seasonal probabilistic forecasts. The final step is the disaggregation of the seasonal probabilistic forecasts of the aggregate response to monthly inflows at individual nodes of the reservoir system. More details on the individual steps alongwith related details of the two study areas—the method is applied to are presented below.

Probabilistic Forecasting Model Structure

The probabilistic forecasting procedure used here conditionally simulates multiple response variables through pre-selected system predictors. Given the uncertainty associated with selection of predictor variables, especially when predicting at long lead times, the above prediction model is formulated as an ensemble model combination or, multiple predictive models with near-independent predictive bases are formed and resulting predictions collated to form an ensemble of likely response scenarios. The steps involved in the above formulation involve selection of predictor variables from sea surface temperature anomalies and catchment indicators, and are detailed as follows:

1. For each aggregate response variable, identification of multiple (set equal to 8 in this study) pseudo-independent order 1–2 predictors of the system being forecast are identified, the predictor identification being performed using the nonparametric partial mutual information criterion (Sharma, 2000). Apart from the use of partial mutual information, which is a generic measure of dependence between variables and manages to ascertain dependence without needing to make assumptions on the nature of dependence (linear or nonlinear), consideration is given to the fact that the uncertainty in observations is nonstationary, to account for which observation standard errors that reflect the change in error characteristics over time, are used in identifying the predictors.
2. In order to represent temporal dependence in the probabilistic forecasts, the first predictor (order 1 predictor) in the above identification is selected as the preceding response value. Use of the previous response helps impart a Markov order 1 dependence in the seasonal probabilistic forecasts.
3. Identification of subsequent predictor variables that offer the best partial dependence to what is described by the leading order 1 and 2 predictors. A maximum of 4 (four) predictors were identified for each component model, that leaves us with 5 models each of which are formulated with the aim of representing relatively independent mechanisms (except for the common first predictor representing the previous season's flow) that contribute to variability in the response.
4. Identification of the optimal influence weight associated with each predictor variable, this influence weight collapsing to zero in situations where the partial usefulness of an individual variable was negligent. This influence weight was ascertained using a partial correlation based logic implemented using nonparametric kernel density estimation techniques.
5. Identification of the probability of selecting each one of the 8 models that are being used to formulate the combined probabilistic forecast. This weight was estimated based on the performance of

the model averaged forecasts in a leave-one-out cross-validation setting, the weight being estimated using a nonlinear constrained optimisation algorithm whose objective was to minimise leave-one-out cross-validation mean square error.

Note that the above procedure enables formulation of the eight component models, each of which have four predictor variables, the first of these variables being common (lag one response variable for the case of the Tasmanian study, and the aggregate or summed lag one response for the case of the Sydney Catchment Authority study). The second predictor in these five sets is ascertained using partial mutual information, with eight such predictors being chosen such that the cross-dependence between them is minimal while maintaining the partial mutual information to be high. Order 3 and 4 predictors are then chosen using partial mutual information without taking cross dependence into account, the assumption being that a unique order 2 predictor (that is relatively independent of other order two predictors) will lead to unique higher order predictors and a predictor set that satisfies the pseudo-independence argument that was raised before.

Probabilistic Forecasting Algorithm

An algorithmic description of the probabilistic forecasting procedure is as follows:

1. Form aggregate response variables $X_{s,q,t}$, $s \equiv$ site, $q \equiv$ season, $t \equiv$ year.
2. Form multiple predictive models for each aggregate variable—
   $X_{s,q,t} \mid Z_{Ms1(q,t)}$
   $X_{s,q,t} \mid Z_{Ms2(q,t)}$
   ……
   where $Z_{Msi(q,t)}$ represents an independent basis consisting of [$X_{s,q-1,t}$, selected land-ocean-atmosphere indicators].
3. Estimate probability of sampling from each model—$p(Ms1)$, $p(Ms2)$, … using cross-validation.
4. For a given realisation, for each aggregate variable, probabilistically select the model to be used.
5. Sample $X_{s,q,t} \mid Z_{Msi*(q,t)} = F^{-1}(\varepsilon_s)$ where $\text{Cor}(\varepsilon_s, \varepsilon_{s'}) \neq 0$ but specified such that $\text{Cor}(X_{s,q,t}, X_{s',q,t})$ equals the observed spatial cross-correlation.
6. Update time (q, t) and predictor vector $Z_{Msi(q,t)}$ by including the sampled $X_{s,q,t}$ as the new predictor value.

Note that the above conditional simulation was performed using a kernel density based nonparametric simulation procedure as outlined in Sharma (2000b). It should also be noted that the use of non-independent uniform random numbers (ε_s, $\varepsilon_{s'}$) is aimed at ensuring an appropriate representation of the spatial dependence that exists in the response variables, and the use of a fixed order one predictor in the multiple conditional generation models used, helps impart a Markov order 1 dependence structure in the sequences generated.

The procedure outlined in Wilks (1998) for specifying the dependence structure between the uniform random numbers (ε_s, $\varepsilon_{s'}$) involves an empirical specification of the correlation matrix in a transformed Gaussian variable space, such that the correlations between the generated responses equal the historical value. The empirical nature of this procedure makes it computationally excessive and also leads to results that are unstable when the historical correlation is high. An alternate procedure was developed in course of this study to overcome the above mentioned problem. This procedure is nonparametric and involves resampling vectors of (ε_s, $\varepsilon_{s'}$) from estimates formed based on the historical record. The steps involved in forming the above mentioned random number vectors are:

1. Denote historical estimates of the uniform random numbers as (e_s, $e_{s'}$).
2. Ascertain the rank of historical aggregate variable observations $X_{s,q,()}$ and $X_{s',q,()}$. The rank for the i'th observation is denoted as (i).
3. Estimate the empirical historical uniform random number as:

   $$e_{s_i} = \frac{(i)}{N+1}$$

   where N denotes the number of observed values.

   Consequently, if the i'th observed pair ($X_{s,q,i}$, $X_{s',q,i}$) has ranks ((i), (i')), the associated random numbers will be $\left(\frac{(i)}{N+1}, \frac{(i')}{N+1}\right)$. 4. Randomly select any of the N empirical uniform random number pairs in step 3 for use in the conditional generation of response variable values (as stated in step 5 of the algorithm).

It should be noted that while the uniform random numbers as generated above will take definite values as calculated in step 3 of the above algorithm, this has no practical limitation on the generation of response variables as the conditional generation procedure leads to new and unique realisations.

As pointed above, the Wilks (1998) procedure was found to be suitable in cases where the historical cross correlations did not become exceptionally high (greater than 0.8). With one of the two case studies results are presented for next, the historical correlations were excessive, as a result of which step 5 of the algorithm was modified to simulate a multivariate response vector consisting of the entire aggregate flow vector using a conditional multivariate probability density estimate of the aggregate responses.

Study Areas

The two study areas results are reported for were the Tasmanian Hydro-electric commission in Tasmania, Australia, and the Sydney Catchment Authority reservoir system in New South Wales, Australia. Hydro-Tasmania operations cover reservoirs spread across the state of Tasmania. The main catchments covered are represented in Figure 1.

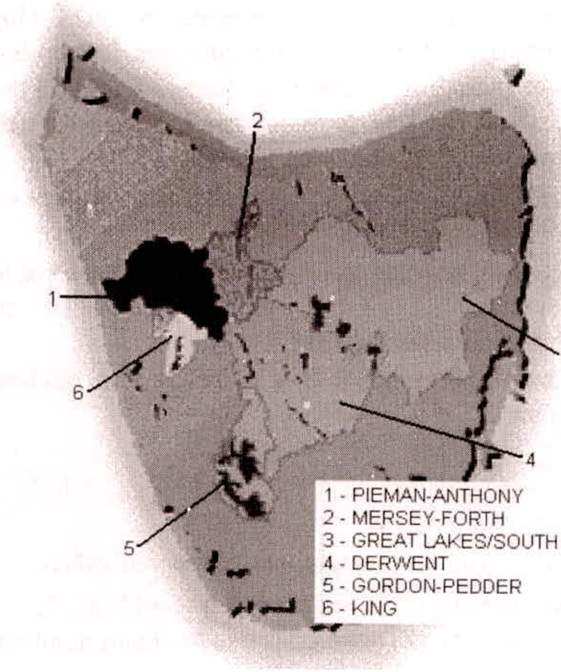


Fig. 1: Main catchments representing the Tasmanian hydro-electric system

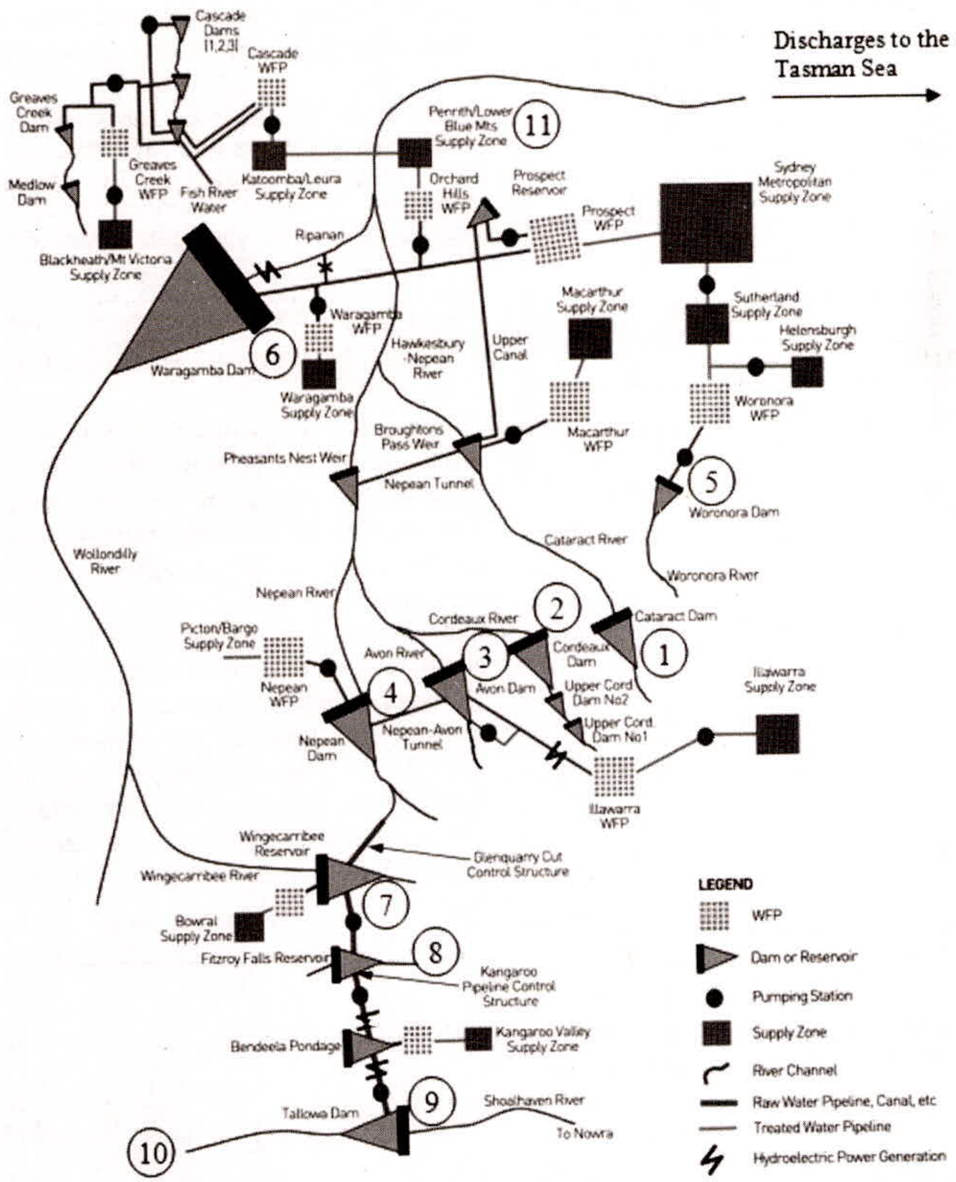
The spatial cross-correlations evaluated on a seasonal basis across the six catchment systems in Figure 1 range from 0.5 to 0.97, with the cross-correlations between the catchment systems excluding the Great Lakes (system 3) catchment falling in the high 0.8–0.97 range. This is a result of all catchments except the Great Lakes being influenced by frontal systems that originate from the western side of Tasmania, the impact of such western influences being smaller in the Great Lakes system.

As a result of the above observations, it was deemed appropriate to form two aggregate response variables representing the entire Tasmania reservoir system, these being denoted as "WCD" (catchments 1, 2, 4, 5, 6) and CGL (catchment 3) in the remainder of this paper.

The Sydney Catchment Authority water supply system (Figure 2) consists of 11 major water supply dams, all in the vicinity of Australia's largest city, Sydney. The climatic influences that effect variability across these dams are similar, because of which the cross-correlations between the seasonal inflows in these dams are high (greater than 0.8). Three regional clusters of dams can be identified, and are used in formulating the probabilistic forecasts described next. These are depicted in Figure 2, and shall be referred to as WWA, METRO and SHOAL in the results presented later. It should be pointed out that the WWA response variable includes Warragamba dam, which supplies nearly 80% of the water to Sydney, and thus has considerable importance in the formulation of the probabilistic forecasts. The cross-correlations between these three aggregate variables range from 0.77 to 0.92 (on a seasonal basis), suggesting that the differences between the climatic factors that influence variability in the catchments may not be significant.

APPLICATION AND RESULTS

The probabilistic forecasting model was applied to predict seasonal and monthly flows from four defined start points in the year (March, June, September, December) for the next 12 months. Predictions were issued on a seasonal as well as on a monthly basis, the monthly predictions being generated based on the seasonal values using a simplistic method of fragments approach. The results presented next are obtained in two application settings—cross-validation for inflow data upto 1998, and pure-forecasts for inflow data between 1999 and 2004 (2006 for the Sydney Catchment Authority application). The pure-forecasts were issued using the model that was developed based on data till 1998, hence representing a situation analogous to one in which the model will finally be used. If one were to assume that the inflows for the period represented by the pure forecasts will be similar to the flows in the pre-1999 period, one should expect the model to perform similarly in both segments. All the results presented next are based on the issuance of 100 realisations, each of them 12 months long, from the four starting points indicated before. Results are assessed on the forecasting models ability to simulate the correct spatio-temporal dependence structure, as well as reduce the predictive uncertainty through the use of the additional climate information utilised in issuing the probabilistic forecasts.



Aggregate Response Variables
WWA = Warragamba (6) + Woronora (5)
SHOAL (Shoalhaven dams) = Wingecarribee (7) + Fitzroy Falls (8) + Tallowa (9)
METRO (Metropolitan dams) = Cataract (1) + Cordeaux (2) + Avon (3) + Napean (4)

Fig. 2: Main reservoirs constituting the Sydney Catchment Authority Water Supply system. (Modified from www.sca.nsw.gov.au)

Tasmanian Application

Table 1 presents the autocorrelation attributes of the generated and observed aggregated flow variables. As can be inferred from the table, the Markov order 1 assumption serves well in representing dependence in the seasonal aggregate response.

The dependence characteristics at individual nodes are also of interest. The lag one seasonal autocorrelation at individual system nodes (for three seasonal combinations similar to Table 1) are illustrated in Figure 3. The aggregate flows are disaggregated to nodal flows using a simple method of fragments approximation (equivalent to finding the nearest neighbour of the generated aggregate flow in the historical record, and scaling nodal flows to maintain summation to the generated value). While the seasonal autocorrelation is not as well represented as it is for the case of the aggregate flow variable in Table 1, the results are acceptable given the use of a Markov order one predictor at the aggregated response variable scale. Note that further improvements are possible if a more elaborate disaggregation approach is adopted.

Table 1: Lag 1 autocorrelation in seasonal aggregate response variables. "H" and "S" denote historical or simulated, while "L1-2" etc. denote the correlations between seasons at Lead 1 and 2 (in this case MAM and JJA). AGG1 and AGG2 represent WCD and CGL response variables respectively

ACF	*L1–2*	*L2–3*	*L3–4*
H-AGG1	0.22	0.38	0.1
H-AGG2	0.58	0.47	0.1
S-AGG1	0.21	0.36	0.05
S-AGG2	0.56	0.51	0.04

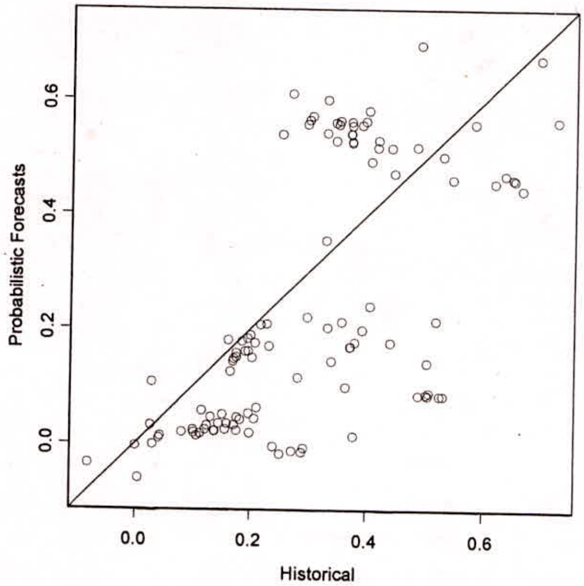


Fig. 3: Lag one seasonal autocorrelation at individual nodes (total nodes 37, for 3 seasonal pairs) of the reservoir system

Table 2 presents similar results detailing the ability to represent spatial dependence across simulations. It can be noted that while medium dependence is well simulated, high spatial dependence is difficult to simulate with the Wilks (1998) methodology adopted here.

Table 2: Historical and simulated spatial cross-correlations between aggregate seasonal responses for probabilistic forecasts issued in March

CRCOR	*L1*	*L2*	*L3*	*L4*
HIST	0.31	0.64	0.45	0.23
SIM	0.34	0.38	0.42	0.25

Table 3 presents the R-square statistics of the conditional forecasts issued in March for the coming 12 months. It should be noted that the lead 2 forecasts (4–6 months ahead) depend on the forecast issued at lead 1 (1–3 months ahead), and hence are expected to deteriorate if earlier forecasts are inaccurate.

Table 3: R^2 for seasonal aggregate response variable probabilistic forecasts issued in March for lead times of 1 to 4 seasons. Variable AGG1 represents the WCD aggregate seasonal response, while variable AGG2 represents the CGL aggregate seasonal response

R^2	*L1*	*L2*	*L3*	*L4*
AGG1	0.47	0.57	0.53	0.49
AGG2	0.56	0.59	0.49	0.51

Given that the probabilistic forecasts described here are aimed to be used to simulate system behavior, it is imperative that they be able to simulate the kind of volumes that are likely to occur. This depends strongly on the ability of the sequences to mimic the spatio-temporal dependence that is observed. Results documenting the accuracy of the sequences in simulating aggregated volumes (from 1–6, 1–9 and 1–12 months ahead) are presented in Tables 4 and 5. It is interesting to note that the accuracy of the volumes a full 12 month ahead is remarkably high. This would not be possible were spatial and temporal dependence not simulated well. The conditional distribution of the 1–12 month aggregated volumes is presented in Figure 3.

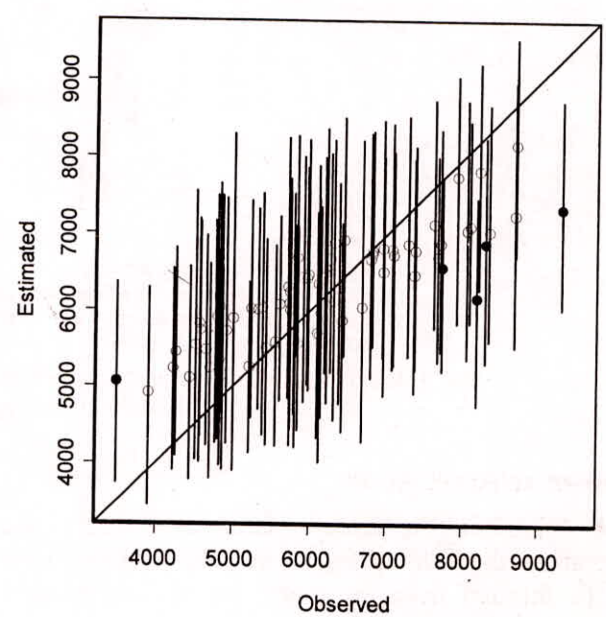


Fig. 4: Aggregate twelve months probabilistic forecast results for Net System Inflows issued in March. The vertical lines represent the 10th to 90th percentiles of the probabilistic forecasts issued. In cases where the observed flow falls within these limits (the line intersects the 45 degree line) the observed flow value is indicated by an open circle. In cases where this is not so, it is indicated by a closed circle

Table 4: R^2 results for cumulative aggregate variables up to a lead time of 4 seasons, for probabilistic forecasts issued in March. Note AGG1 and AGG2 refer to the WCD and CGL aggregate seasonal response variables, while "*L*" refers to the lead time

R^2	*L1*	*L1–2*	*L1–3*	*L1–4*
AGG1	0.47	0.6	0.6	0.6
AGG2	0.56	0.62	0.62	0.6

Table 5: R^2 for cumulative Net System Inflow probabilistic forecasts issued in March (formed by adding individual aggregate response variable values) for lead times of 1–4 seasons. Note that AGG1 + 2 represents the cumulative sum of WCD and CGL (or, the cumulative net system inflow)

R^2	*L1*	*L1–2*	*L1–3*	*L1–4*
AGG1+2	0.58	0.65	0.63	0.62

Note that while the results illustrated in Figure 4 and Tables 4 and 5 point to a high accuracy of the probabilistic forecasting system, certain limitations are also visible. The main limitation that can be spotted from the results in Figure 4 pertain to an oversimulation of low observed flows and an undersimulation of the high observed flows. This bias is a likely result of the model averaging procedure adopted, with the bias being possible to reduce if one considers an alternative where the model weights were not static as was the case here, but allowed to evolve based on past history of the accuracies associated with individual models. Work on such a formulation is currently underway.

In addition to the above results, the models were tested in a pure forecasting mode using data from 1998 to 2004. The tests involved running the probabilistic forecasts starting in each of the four seasons for the coming 12 months, and comparing results with those observed. Pooling together the results for all seasons resulted in adequate number of cases being evaluated (number of cases being number of years × number of seasons × 4 seasons representing the coming year). The overall coefficient of determination from this exercise was recorded at 0.59 (for WCD) and 0.47 (for CGL), results that are not dissimilar to those reported in Table 3. These pure forecast validation results lend more credibility to the cross-validation results reported elsewhere in this paper, and provide a strong case for the utility of the forecasting system in water management applications.

Sydney Catchment Authority Application

Table 6 presents the lag-one autocorrelation of the aggregate seasonal response Sydney Catchment Authority inflow. The table indicates that the seasonal persistence structure was well represented in the probabilistic forecasts, with all the lag one correlations being represented within tolerance limits. This is a result of the Markov order-one representation used in formulating the probabilistic forecasting model. Note that the lack of representation of the lag one autocorrelation for Quarter 4 is not surprising, given that Q4 represents a lead 4 forecast, which is heavily dependant on the quality of forecasts for quarters 2 and 3.

Table 6: Lag 1 autocorrelation of observed data and simulations for forecasts issued in quarter 1. The autocorrelation for the simulations is estimated as the average of the autocorrelations for each probabilistic forecast

Historical	*Q1*	*Q2*	*Q3*	*Q4*
WWA	–0.10	0.60	0.30	0.23
METRO	–0.31	0.42	0.05	0.15
SHOAL	–0.30	0.44	0.22	0.18
Probabilistic Forecasts				
WWA		0.51	0.18	0.08
METRO		0.32	0.01	0.07
SHOAL		0.39	0.09	0.16

Table 7 represents the spatial cross-correlations simulated across the four lead times across the three response variables. Representation of these spatial dependence attributes was a key consideration in the formulation of the probabilistic forecasting model. The multi-variate probabilistic forecasting of all responses simultaneously, along with the use of a common aggregate system predictor (the lag 1 aggregate inflow for the preceding season) results in an adequate representation of the spatial dependence attributes.

Table 8 presents the Coefficient of Determination (often represented as R^2 but used here to represent the fraction of the variance of observed flows that is explained) for the seasonal probabilistic forecasts for Leads 1 to 4. Note that the results for Lead 1 represent forecasts for Quarter 1, results for Lead 2 represent forecasts for Quarter 2 (JJA), etc.

Tables 9 and 10 present the Coefficient of Determinations for probabilistic forecasts aggregated over multiple lead times (for each response variable) and over the entire system (summation of all the responses aggregated over the indicated lead time). A good performance in these aggregate variables is indicative of a proper representation of spatio-temporal dependence between the probabilistic forecasts for each response.

Table 7: Historical and Forecast Cross Correlations for Leads 1 to 4

Historical				*Probabilistic Forecasts*			
Q1	*WWA*	*METRO*	*SHOAL*	*LEAD1*	*WWA*	*METRO*	*SHOAL*
WWA	1.00	0.84	0.80	WWA	1.00	0.81	0.78
METRO	0.84	1.00	0.78	METRO	0.81	1.00	0.76
SHOAL	0.80	0.78	1.00	SHOAL	0.78	0.76	1.00
Q2	*WWA*	*METRO*	*SHOAL*	*LEAD2*	*WWA*	*METRO*	*SHOAL*
WWA	1.00	0.92	0.86	WWA	1.00	0.89	0.83
METRO	0.92	1.00	0.86	METRO	0.89	1.00	0.80
SHOAL	0.86	0.86	1.00	SHOAL	0.83	0.80	1.00
Q3	*WWA*	*METRO*	*SHOAL*	*LEAD3*	*WWA*	*METRO*	*SHOAL*
WWA	1.00	0.89	0.82	WWA	1.00	0.84	0.79
METRO	0.89	1.00	0.80	METRO	0.84	1.00	0.74
SHOAL	0.82	0.80	1.00	SHOAL	0.79	0.74	1.00
Q4	*WWA*	*METRO*	*SHOAL*	*LEAD4*	*WWA*	*METRO*	*SHOAL*
WWA	1.00	0.91	0.76	WWA	1.00	0.87	0.73
METRO	0.91	1.00	0.86	METRO	0.87	1.00	0.82
SHOAL	0.76	0.86	1.00	SHOAL	0.73	0.82	1.00

Table 8: Coefficient of Determination (R^2) for forecasts issued in quarter 1 for individual quarters corresponding to lead times of 1 to 4 seasons

Q1 Start	*WWA*	*Metro*	*Shoal*
Lead 1	0.344	0.261	0.446
Lead 2	0.095	0.074	0.164
Lead 3	0.336	0.247	0.412
Lead 4	0.237	0.322	0.255

Table 9: Coefficient of Determination (R^2) for forecasts issued in quarter 1 that are aggregated over multiple lead times, for each response variable

Q1 Start	*WWA*	*Metro*	*Shoal*
Lead 2	0.338	0.325	0.324
Lead 3	0.334	0.297	0.312
Lead 4	0.357	0.309	0.306

Table 10: Coefficient of Determination (R^2) for forecasts issued in quarter 1 aggregated over multiple lead times and over the entire system

Q1 Start	*Lead 1*	*Lead 2*	*Lead 3*	*Lead 4*
	0.452	0.357	0.347	0.355

Forecasts for the entire year issued in Quarter 2 are presented in Figure 5. The overall R^2 statistic associated with the forecasts is 0.36, significantly lower than the corresponding statistics with respect to the Tasmanian forecasting system. However, it is notable that the representation of the temporal and spatial dependence that the model structure ensures, results in significant accuracy in the forecasts one-year out into the future. Readers are reminded that the forecasts for lead times of 2, 3 and 4 seasons are based on the forecasted inflows for the previous seasons, hence cannot be of a high quality unless dependence attributes are well modeled.

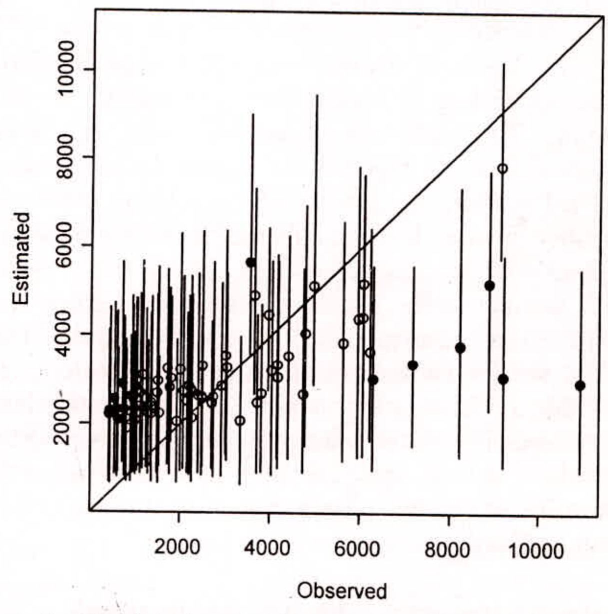


Fig. 5: Simulated and observed aggregated flows for Quarters 2, 3, 4 and 1 (of the next year) for forecasts issued in Quarter 2

As was done with the Tasmanian forecasting system, the Sydney Catchment Authority forecasts were also evaluated in a pure-forecast mode, with pure-forecasts being issued from 1999 to 2006. Coincidently, this pure forecast period coincided with one of the lowest flow periods ever recorded in the history of the Sydney Catchment Authority inflows. As a result, the bias noticeable in the forecasts of the low flows in Figure 5, was all the more evident for this period, and the accuracy that was noted was negligible (R^2 estimated for these years being less than zero). Hence, the pure forecasts did not allow a confirmation of the results that were obtained in the cross-validation mode, instead pointing out to the main limitation of the forecasting procedure, which is the bias that is introduced especially for the very low and high flows as a result of multiple model results being combined.

DISCUSSIONS

This paper presented a multi-reservoir inflow forecasting system that issues probabilistic forecasts of inflows on a monthly basis for the next 12 months at multiple nodes of a reservoir system. These inflows can then be used as inputs into a reservoir simulation approach, and optimal operating policies for the coming twelve months formulated. The probabilistic forecasting system is developed using climate information as represented by reconstructed sea surface temperature anomaly data, which, along with the previous season's inflows serves as the basis for reducing the predictive uncertainty in the forecasts issued. In this study, two applications of this forecasting system are reported. These represent the Tasmanian Hydro-electric system, the aim being to provide inflow forecasts such that knowledge on the power generation capability of the system can be ascertained beforehand, and the Sydney Catchment Authority system, where the aim is to assess likely future storages so that decisions on whether additional pumping is necessary to keep the storage at an acceptable level can be made confidently.

The capabilities and drawbacks of the forecasting procedure can be assessed based on the two forecasting applications presented here. Some observations that can be drawn from these applications are:

1. *Model averaging:* The use of multiple predictive models has advantages and disadvantages. The advantage is the representation of the full uncertainty behind the forecasting procedure in the probabilistic forecasts issued. The disadvantage is the significant bias that is introduced with respect to the very high or very low flows the system may encounter. In the two case studies presented, this bias was very evident in the results for the Sydney Catchment Authority system, where the climatic forcings used were not as strongly related to the response as they were in case of the Tasmanian inflows. Further work is needed on developing a model combination approach that can address this limitation adequately.
2. *Spatio-temporal dependence representation:* The two forecasting applications presented here used somewhat different models for representing the spatio-temporal dependence structure in their respective historical flow records. The less-spatially-correlated Tasmanian inflows used a modification of the Wilks (1998) approach as the basis for modeling spatial dependence, which the more-strongly correlated Sydney Catchment Authority inflows used a common Markov order 1 predictor (the aggregate system inflow) and a modeling procedure that predicted the full multivariate response given the same conditioning variables. These different model structures have associated advantages and disadvantages. The advantage of the Wilks (1998) approach is that each individual response can be modeled using a set of predictor variables chosen to predict the individual response by itself. The disadvantage is the limited ability of the method to simulate very high (greater than 0.8) cross-correlations in the forecasts issued. The advantage of modeling the multivariable response collectively is the simplicity the model structure presents. The disadvantage is the limited representation of the structural uncertainty present as predictor variables cannot be identified with associated individual responses.
3. *Physical interpretation of system predictors:* The logic used to identify predictor variables in either of the two applications presented was to use a statistical measure of partial dependence, known as Partial Mutual Information. The only physical consideration behind the choice of predictor variables was to restrict the lags till which they could be identified to a maximum (6 years in this study), and associate lower importance to predictors that arise from regions known to have high uncertainty associated in the sea surface temperature anomaly reconstructions. This limited physical consideration leads to a system of predictors that are difficult to explain using our knowledge of the underlying circulation that relates climate variables across the earth. This is especially

so when identifying order 3 and above predictors, which are meant to explain the response conditional to the earlier predictors that have already been identified. The use of multiple predictive models is crucial in formulating the probabilistic forecasts in such a situation. Multiple models are based on the use of multiple predictor sets, and while each set has significant selection uncertainty associated with it, collectively, the multiple sets are able to convey the full uncertainty in the climate predictions judiciously.

While the probabilistic forecasting approach presented here has the ability to produce meaningful forecasts at multiple nodes of the reservoir system for a high lead time (one year), a number of improvements are possible. The use of physically based climate simulation models can offer a completely different basis in the multi-model ensemble that formulates the probabilistic forecasting system. The use of concurrent downscaling models that work off forecasts of climate variables using General Circulation Models offers an alternative where identification of predictor variables will be restricted to the current lag, thereby reducing the structural uncertainty that is introduced by identifying such variables across multiple time lags as was done in the applications reported here (Westra *et al.*, 2008). A third alternative for improving forecasts is the use of non-static model combination weights, or weights that are dynamic in nature. The use of such dynamic model combination (Chowdhury and Sharma, 2007) can possibly help alleviate the model bias that is observed at the low or high streamflow end, by helping assign greater weights to models that perform well at each end of the flow spectrum being modeled. Work is currently underway to improve the multiple site probabilistic forecasting system by inclusion of the alternatives outlined above.

ACKNOWLEDGEMENTS

This research was funded by the Australian Research Council, Hydro Tasmania, and the Sydney Catchment Authority.

REFERENCES

Chowdhury, S. and Sharma, A. (2007). Dynamic model mixing for enhancing the predictability of hydroclimatic variables. In: Kulasiri, D. and Oxley, L. (Editors), MODSIM 2007 Congress. Modelling and Simulation Society of Australia and New Zealand, Christchurch, New Zealand.

Sharma, A. (2000). "Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1—A strategy for system predictor identification." *Journal of Hydrology*, 239(1–4): 232–239.

Sharma, A. (2000). "Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 3—A nonparametric probabilistic forecast model." *Journal of Hydrology*, 239(1–4): 249–258.

Sharma, A., Luk, K.C., Cordery, I. and Lall, U. (2000). "Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 2—Predictor identification of quarterly rainfall using ocean-atmosphere information." *Journal of Hydrology*, 239(1–4): 240–248.

Westra, S., Brown, C., Lall, U. and Sharma, A. (2008). "Multivariate streamflow forecasting using independent component analysis." *Water Resources Research*, 44, W02437(doi:10.1029/2006WR005785).

Wilks, D.S. (1998). "Multisite generalization of a daily stochastic precipitation generation model." *Journal of Hydrology*, 210: 178–191.