18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Selecting predictors for seasonal streamflow predictions using a Bayesian joint probability (BJP) modelling approach

Robertson, D.E. [1] **and Q.J. Wang** [1]

[1] *CSIRO Land and Water, Highett, Victoria*
Email: david.robertson@csiro.au

Abstract: Forecasts of future seasonal streamflows are valuable to a range of water managers and users, including irrigators, urban and rural water supply authorities, environmental managers and hydroelectricity generators. Such forecasts can inform planning and management decisions to maximize returns on investments and available water resources, and to ensure security of water supplies. Historically, routine forecasts of seasonal streamflows have not been available in Australia.

There are two main sources of predictability in Australian streamflows. Strong serial correlations in streamflows arise due to soil and groundwater storages extending the time between the incidence of rainfall and any resulting streamflow. Thus, indicators of initial catchment conditions may be good predictors of future streamflows. Future rainfall and climate also influence future streamflows. Many indices of large-scale climate anomalies, such as the Southern Oscillation Index and Indian Ocean Dipole Mode Index, show significant concurrent and lagged correlations with rainfall and streamflows and therefore may be useful predictors of streamflows too. However, there has been no systematic investigation into how to best process a large array of candidate predictors to produce skilful and reliable seasonal streamflow predictions for different locations and seasons.

This paper introduces a method to select predictors of streamflows for the recently developed Bayesian joint probability (BJP) modelling approach to seasonal streamflow prediction at multiple sites. The predictor selection method seeks to identify reliable predictors that produce skillful predictions. An important outcome should be that the selected predictors are consistent with our understanding of the physical hydrological and climate systems. The method selects predictors that give the largest improvement in prediction accuracy and are supported by statistical evidence. The prediction accuracy is assessed using a skill score based on mean squared error in probability (SS_{RMSEP}), while the evidence supporting predictor selection is assessed using the log pseudo Bayes factor (log PsBF).

The predictor selection method is tested on two catchments in south eastern Australia: the Goulburn and the Murrumbidgee. Predictions of three month streamflow totals are made on the first day of each month using separate models for each prediction date. In each catchment, streamflow predictions are made jointly for three stream gauging stations. A base model is established using only streamflows for the month preceding prediction as predictors. These antecedent streamflows are used to represent the initial catchment conditions. Additional predictors are selected for each prediction date from a set of 42 candidates that includes a number of climate indices at various lag times. The predictive performance measures used in selecting predictors are derived from cross validation predictive probability density functions computed using importance resampling. The performance of the predictors is assessed using graphical means. The physical interpretability is appraised by comparing findings to the reported literature and knowledge of climate experts.

The results show that climate indices can make improvements to the accuracy of streamflow predictions most of the year with the greatest improvement during spring. The predictors selected for the Goulburn catchment are dominated by climate anomalies in the Indian Ocean, while those selected for the Murrumbidgee catchment are dominated by climate anomalies in the Pacific Ocean. Reports in the literature support the finding that the two catchments are influenced by different climate systems. During autumn, the selected predictors seem to be augmenting the representation of the initial catchment conditions and therefore further investigation into the best indicators of initial catchment conditions are required. Future work is also required to test the robustness of the proposed predictor selection method using double cross-validation and applying it to catchments elsewhere in Australia.

Keywords: *Streamflow prediction, Bayesian methods, predictor selection*

1. INTRODUCTION

Forecasts of future seasonal streamflows are valuable to a range of water managers and users, including irrigators, urban and rural water supply authorities, environmental managers and hydroelectricity generators. Such forecasts can inform planning and management decisions to maximize returns on investments and available water resources and to ensure security of supply. Historically, routine forecasts of seasonal streamflows have not been available in Australia. The Bureau of Meteorology are now extending their seasonal climate prediction service to include water availability and in the first instance predictions of streamflows [*Plummer et al.*, 2009].

There are two main sources of predictability in Australian streamflows. Streamflows in many parts of Australia display high serial correlation (or persistence) [*Chiew et al.*, 1998]. This persistence arises due to soil and groundwater storages delaying responses in rainfall-runoff processes, giving streamflow data a memory of several months [*Chiew et al.*, 1998]. Therefore, indicators of the initial catchment conditions provide a source of predictability for future streamflows.

The magnitude of future streamflows is also related to future rainfall and climate conditions. Many studies have found concurrent and lagged correlations between large-scale climate indices, such as the Southern Oscillation Index and Indian Ocean Dipole Mode Index, and rainfall [for example, *McBride and Nicholls*, 1983] and streamflows [for example, *Chiew et al.*, 1998]. The strength of the observed correlations suggests that large-scale climate indices may be useful predictors of future streamflows. However, the literature reports a large array of climate indices that display correlations with rainfall and streamflows and there has been no systematic investigation into how to best process all these potential predictors to produce skilful and reliable seasonal streamflow predictions for different locations and seasons.

There are several possible approaches to using the potential predictors to produce streamflow predictions, including hierarchical modelling, model averaging and predictor/model selection. Where there are a very large number of potential predictors, methods such as hierarchical modelling and model averaging become computationally challenging for operational use, and therefore the selection of a subset of predictors is necessary. The subset of predictors may be selected using certain measures of predictive performance. However, in dealing with a large number of potential predictors and only limited data, there is a need to minimize the influence of chance features in the available data on the selection of predictors. An important outcome of predictor selection should be that the selected predictors are consistent with our physical understanding of hydrological and climate systems.

This paper describes a method of predictor selection for seasonal streamflow prediction using the Bayesian joint probability (BJP) modelling approach. The method selects predictors that give the largest improvement in prediction accuracy and are supported by the statistical evidence. The prediction accuracy is assessed using a skill score based on mean squared error in probability (SS_{RMSEP}), while the evidence supporting predictor selection is assessed using the log pseudo Bayes factor (log PsBF). The method is tested through two case studies involving the prediction of streamflows in south eastern Australia. Streamflow predictions are made on the first day of each month for a lead time of three months. Antecedent streamflows are used as base predictors, and additional predictors are selected from a list of 42 candidate climate indices. The selected predictors are examined for their consistency with understanding of the contributing hydrological and climate systems.

2. METHODS

2.1. Bayesian Joint Probability (BJP) Modelling Approach

The BJP modeling approach was used produce predictions of future streamflows [*Wang et al.*, 2009]. The BJP modelling approach assumes that a set of predictands ($\boldsymbol{y}(2)$), in this case future streamflows, and predictors ($\boldsymbol{y}(1)$) is described by a Box-Cox transformed multivariate normal distribution.

$$\boldsymbol{z} = [\boldsymbol{z}(1), \boldsymbol{z}(2)] \sim \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\sigma}\, \boldsymbol{R} \boldsymbol{\sigma}^{\mathrm{T}})$$

where $\boldsymbol{z}$ is a vector of Box-Cox transformed future streamflows and their predictors. Model parameters (θ) include vectors of the Box-Cox transformation parameter (λ), the mean (μ) and standard deviation (σ), and the correlation coefficient matrix ($\boldsymbol{R}$). The posterior distribution of model parameters ($p(\theta \mid \boldsymbol{y}_{OBS}^{n}, \boldsymbol{y}_{OBS}^{n-1}, \ldots, \boldsymbol{y}_{OBS}^{1})$, where $\boldsymbol{y}_{OBS}^{t}$ is a vector of historically observed cases for year $t = 1, 2, \ldots, n$), is obtained from Bayesian inference, performed using Markov chain Monte Carlo sampling. Probabilistic predictions of future seasonal streamflows are produced by conditioning the transformed multivariate normal distribution on predictor values.

The BJP modeling approach overcomes many of the limitations of previous statistical techniques applied to streamflow prediction [for example, *Piechota et al.*, 2001]. The Box-Cox transformed multivariate normal distribution has considerable flexibility for modelling a wide range of predictors and predictands. The BJP modeling approach can be used to produce joint probabilistic predictions of streamflows at multiple sites that preserve inter-site correlations. The BJP modeling approach allows the use of data that contains non-concurrent and missing records in both parameter inference mode and prediction mode. The model flexibility and data handling ability means that the BJP modelling approach is potentially of wide practical application.

2.2. Streamflow prediction models

For this study, predictions of three month streamflow totals for multiple sites were made on the first day of each month. A base model (M_0) was established using only streamflows for the month preceding the prediction as predictors. To produce predictions for all prediction dates, 12 separate models were established, one for each month. The base model was then expanded in steps to new models (M_1) that include climate indices as additional predictors.

The performance of each model in predictive mode was assessed using leave-one-out cross-validation predictions. Leave-one-out cross-validation involves establishing the model using all historical data except one case and then predicting the streamflow for the missing case. This procedure is repeated for all cases in the historical record to provide an overall assessment of model performance.

Establishing a model involved Bayesian inference of model parameters using Markov chain Monte Carlo sampling. Importance resampling was used as a fast method to calculate the leave-one-out cross-validation posterior parameter densities [*Gelfand*, 1995; *Vehtari and Lampinen*, 2002]. A total of 10,000 sets of parameters were sampled from the full posterior parameter distribution using the Metropolis algorithm. Importance resampling was then used to sample 1000 sets of parameters that represented the leave-one-out cross-validation posterior parameter densities for each case in the historical record. The streamflow predictions for each case were numerically represented by a sample of 1000 sets of values, one generated for each of the 1000 sets of parameter values.

For each model, the cross validation joint predictive probability density function of streamflows at multiple sites is given by

$$\begin{aligned} f^t(\boldsymbol{y}^t(2)) &= p(\boldsymbol{y}(2)\,|\,\boldsymbol{y}(1) = \boldsymbol{y}^t_{OBS}(1);\, \boldsymbol{Y}^{(t)}_{OBS}(1),\, \boldsymbol{Y}^{(t)}_{OBS}(2)) \\ &= \int p\left(\boldsymbol{y}(2)\,\middle|\,\boldsymbol{y}(1) = \boldsymbol{y}^t_{OBS}(1);\, \boldsymbol{\theta}\right) \cdot p\left(\boldsymbol{\theta}\,\middle|\,\boldsymbol{Y}^{(t)}_{OBS}(1), \boldsymbol{Y}^{(t)}_{OBS}(2)\right) \cdot d\boldsymbol{\theta} \end{aligned}$$

where $\boldsymbol{Y}^{(t)}_{OBS}(1)$ denotes the observed predictors for all the historical cases $[\boldsymbol{y}^1_{OBS}(1), \boldsymbol{y}^2_{OBS}(1), \ldots, \boldsymbol{y}^n_{OBS}(1)]$ except $\boldsymbol{y}^t_{OBS}(1)$, and $\boldsymbol{Y}^{(t)}_{OBS}(2)$ the observed predictands (streamflows) for all the historical cases $[\boldsymbol{y}^1_{OBS}(2), \boldsymbol{y}^2_{OBS}(2), \ldots, \boldsymbol{y}^n_{OBS}(2)]$ except $\boldsymbol{y}^t_{OBS}(2)$.

The marginal distributions of $f^t(\boldsymbol{y}^t(2))$ give predictive probability functions for streamflows at individual sites, denoted here as $f^t(y^t)$ for the density distributions and $F^t(y^t)$ for the cumulative distributions with median predictions y^t_{MED}.

2.3. Predictor selection criteria

Different mathematical models generally give different predictions. Two measures were used here to assess improvements in predictions by model M_1 over a base model M_0. The first measure is the RMSEP (root mean squared error in probability) skill score of model M_1 median predictions in reference to model M_0 median predictions. For each streamflow site, it is defined as

$$SS_{RMSEP10} = \frac{RMSEP_{M_1} - RMSEP_{M_0}}{0 - RMSEP_{M_0}}$$

where $RMSEP = \left(\frac{1}{n} \sum_{t=1}^{n} (F_{CLI}(y_{MED}^{t}) - F_{CLI}(y_{OBS}^{t}))^2 \right)^{1/2}$

$F_{CLI}(.)$ being the observed historical cumulative distribution (climatology) of the streamflows at that site. An overall $SS_{RMSEP10}$ value was obtained by averaging the skill scores for all the streamflow sites. The *RMSEP* is a measure of prediction error. The $SS_{RMSEP10}$ measures the reduction in prediction error, or the improvement in the prediction accuracy, of M_1 over M_0. The *RMSEP* is similar to the Linear Error in Probability Space (*LEPS*) score [*Potts et al.*, 1996]. However, the *RMSEP* is suited to a more traditional skill score formulation than *LEPS*. Skill scores calculated using both *LEPS* and *RMSEP* are similar over the range of 0% to 100% (Wang and Robertson, unpublished data). Note that the *RMSEP* used here is based on median predictions only. In a follow-up study [*Robertson and Wang*, 2009], it is applied to the full distributions of predictions.

The second measure is the log pseudo Bayes factor defined as

$$\log PsBF_{10} = \log \prod_{t=1}^{n} \frac{f_{M_1}^{t}(\boldsymbol{y}^{t}(2) = \boldsymbol{y}_{OBS}^{t}(2))}{f_{M_0}^{t}(\boldsymbol{y}^{t}(2) = \boldsymbol{y}_{OBS}^{t}(2))}$$

The $\log PsBF_{10}$ assesses the statistical evidence supporting M_1 over M_0. The pseudo Bayes factor differs from the traditional Bayes factor in that it is calculated using the cross validation predictive density rather than the prior predictive density [*Gelfand*, 1995]. The pseudo Bayes factor is therefore less sensitive to the prior parameter distribution [*Vehtari and Lampinen*, 2002].

Preliminary investigations showed some seasonal patterns in $SS_{RMSEP10}$ and $\log PsBF_{10}$ but also random fluctuations. Weighted neighbor averaging was used to smooth out some of the noise on the assumption that there is an underlying seasonal continuity in climate drivers. For each prediction date, mon, weighted neighbor averages of $SS_{RMSEP10}$ and $\log PsBF_{10}$ were calculated as

$$\overline{SS}_{RMSEP10}^{mon} = 0.2 SS_{RMSEP10}^{mon-1} + 0.6 SS_{RMSEP10}^{mon} + 0.2 SS_{RMSEP10}^{mon+1}$$

$$\overline{\log PsBF}_{10}^{mon} = 0.2 \log PsBF_{10}^{mon-1} + 0.6 \log PsBF_{10}^{mon} + 0.2 \log PsBF_{10}^{mon+1}$$

Predictor selection seeks to find the predictors that give the largest improvement in the prediction accuracy, provided that such improvement is reasonably supported by statistical evidence. Specifically, the predictor that give the largest improvement in $\overline{SS}_{RMSEP10}$ and satisfy $\overline{\log PsBF}_{10} \geq 0.5$ is selected.

2.4. Analysis of selected predictors

To examine the performance of the different predictors, a predictor selection plot of the $\overline{SS}_{RMSEP10}$ values of all candidate predictors with $\overline{\log PsBF}_{10} \geq 0.5$ for all prediction dates was produced. The consistency of the selected predictors with understanding of the physical climate and hydrological systems was assessed by examining the literature for supporting or contradictory evidence and through discussions with climate experts.

3. DATA

3.1. Streamflow data

The predictor selection procedure was tested on two locations in south eastern Australia. The locations were chosen in catchments with considerable consumptive water use, where streamflow predictions were potentially of some value, and with streams that are not ephemeral, due to a current limitations in the BJP modelling approach. The two locations chosen were the Goulburn catchment in Victoria and the Murrumbidgee catchment in NSW. At each location, three gauging stations were identified on streams that had unimpaired catchments and long, relatively complete streamflow records (Table 1). Stations in close proximity to each other were chosen to ensure that they responded to the same climate influences (Figure 1).

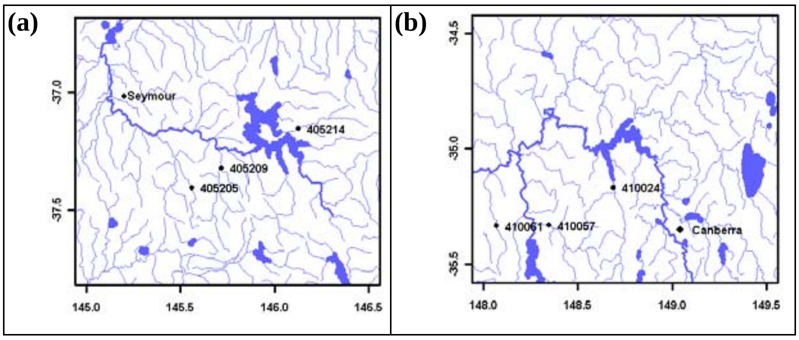

Figure 1. Location of gauging stations in the (a) Goulburn and (b) Murrumbidgee catchments

Table 1. Study gauging stations

Location	Gauging Station	Station Name	Catchment area (km^2)	Available record	Mean Annual Flow (mm)
Goulburn	405205	Murrindindi River at Murrindindi	619	1946-2008	494
Goulburn	405209	Acheron River at Taggerty	368	1948-2008	495
Goulburn	405214	Delatite River at Tonga Bridge	694	1955-2008	310
Murrumbidgee	410024	Goodradigbee River at Wee Jasper	990	1914-2007	275
Murrumbidgee	410057	Goobarragandra River at Lacmalac	673	1944-2007	399
Murrumbidgee	410061	Adelong Creek at Batlow Road	155	1947-2007	240

3.2. Additional predictors

From the literature, 12 indices of large-scale climate anomalies were identified as potential predictors of streamflows (Table 2). All of these indices have been found to have high concurrent or lagged correlations with monthly rainfall in different parts of Australia. Monthly values of all 12 indices lagged by up to three months were considered to be potential predictors. These predictors were considered to characterize the oncoming climate conditions. Antecedent rainfall was also included as an additional measure of the initial catchment condition. Antecedent catchment rainfall lagged by up to six months was considered to be a potential predictor, due to the potentially very long delays in rainfall-runoff processes that may occur. In total, 42 candidate predictors were considered. Preliminary data analysis showed that the climate indices based on direct observations, specifically the Southern Oscillation Index and catchment rainfall, were considerably noisier that other climate indices. To reduce this noise, three month averages were used instead of monthly values.

Table 2. Indices identified as potential additional streamflow predictors

Predictors	Period of record	Data source
Indian Ocean West Pole Index	1854-2008	NCAR, ERSST.v3 [*Smith et al.*, 2008]
Indian Ocean East Pole Index	1854-2008	NCAR, ERSST.v3 [*Smith et al.*, 2008]
Indian Ocean Dipole Mode Index	1854-2008	NCAR, ERSST.v3 [*Smith et al.*, 2008]
Indonesia Index	1854-2008	NCAR, ERSST.v3 [*Smith et al.*, 2008]
NINO3	1950-2008	NECP – SST anomalies
NINO3.4	1950-2008	NECP – SST anomalies
NINO4	1950-2008	NECP – SST anomalies
Southern Oscillation Index	1876-2008	Bureau of Meteorology
ENSO Modoki	1854-2008	NCAR, ERSST.v3 [*Smith et al.*, 2008]
Thermocline	1980-2008	Bureau of Meteorology
Southern Annular Mode	1957-2008	Marshall's[2003] data
Tasman Sea Index	1854-2008	NCAR, ERSST.v3 [*Smith et al.*, 2008]
Catchment rainfall	1900-2008	SILO data

4. RESULTS

4.1. Predictor selection for the Goulburn catchment

Figure 2a presents the predictor selection plot for the Goulburn catchment, showing the $\overline{SS}_{RMSEP10}$ for those candidate predictors with $\overline{\log PsBF_{10}} \geq 0.5$. The plot shows that additional predictors increase the prediction accuracy for all dates with the exception of those made in June, July, August and December. The greatest increase in prediction accuracy occurs during the spring months.

The best predictors also show consistency with knowledge of the physical processes influencing streamflow. The Indian Ocean Dipole Mode Index, lagged by one month, is the best predictor for predictions made during the spring months. This finding concurs with the conclusions of recent studies [for example, *Ummenhofer et al.*, 2009] that south eastern Australian rainfall is influence by Indian Ocean sea surface temperature anomalies in spring. The Indonesian Index, lagged by three months, is the best predictor for predictions made during January, February and March. This is consistent with the finding of Cai and Cowan [2008] that rainfall in northern Victoria is correlated with Indonesian sea surface temperature anomalies. Catchment rainfall lagged by 6 months is the best predictor for predictions made during April and May. One possible explanation of this result is that the antecedent streamflow contains both base flow and quick flow signals, while the lagged rainfall provides an indicator of just the base flow component. During these months the catchment is wetting up and therefore the total streamflow for the month prior to the prediction does not completely capture the initial catchment condition and the lagged rainfall is acting as an additional indicator.

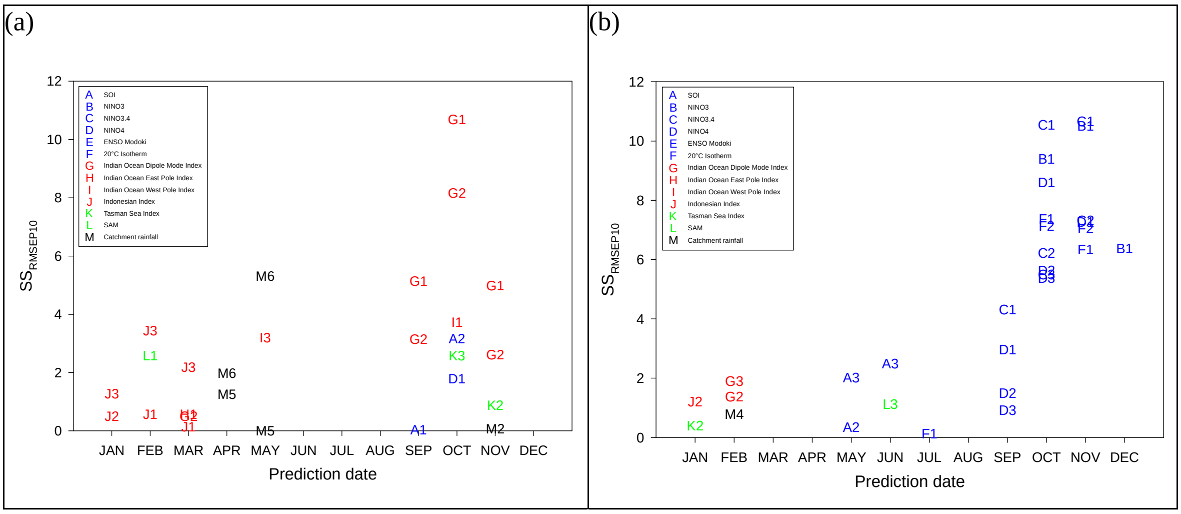

Figure 2. Predictor selection plot for the first additional predictor for (a) the Goulburn catchment and (b) the Murrumbidgee catchment.

Each symbol is a candidate predictor, the letter signifies the climate index and the number the lag. For example A3 is SOI lagged by 3 months.

4.2. Predictor selection for the Murrumbidgee catchment

Figure 2b presents the predictor selection plot for the Murrumbidgee catchment. The plot shows that the largest improvement in prediction accuracy occurs between September and December, with only a small improvement for the remainder of the year. The majority of the best predictors are indicators of climate anomalies in the Pacific Ocean that are related to the El Nino – Southern Oscillation. During the period between September and December, there is consistency between the best predictors and understanding of physical processes. Several authors have shown that the correlations between the El Nino – Southern Oscillation process and seasonal rainfall in central New South Wales are strongest during spring [for example, *McBride and Nicholls*, 1983]

For predictions made between January and July, there is little consistency of best additional predictors with understanding of the physical processes. While correlations have been observed between climate predictors and rainfall during this period, the examination of modeled soil moisture data suggests that incident rainfall during this period will tend to replenish soil and groundwater storages rather than contributing directly to streamflows (not shown). Therefore, the best additional predictors for this period may be augmenting the representation of the initial catchment condition, or may be representing chance features within the data. This suggests further investigation into the best indicator of the initial catchment condition is necessary.

5. CONCLUSIONS

Seasonal predictions of streamflows are valuable to a wide range of users. Predictability in seasonal streamflows can be sourced from indicators of initial catchment and future climate conditions. However, it is necessary to identify which indicators are useful predictors of seasonal streamflows. This paper has described a method to select predictors for seasonal streamflow prediction using the Bayesian joint probability modelling approach. The method selects those predictors that result in the greatest improvement in predictive accuracy, provided that such improvement is reasonably supported by statistical evidence. Specifically, the method selects those predictors that have the largest neighbor weighted average $\overline{SS}_{RMSEP10}$ with $\overline{\log PsBF}_{10} \geq 0.5$.

The proposed predictor selection method was tested on the Goulburn and Murrumbidgee catchment in south eastern Australia. The results show that climate indices can make improvements to the accuracy of streamflow predictions for most of the year with the greatest improvement during spring. The predictors selected for the Goulburn catchment are dominated by climate anomalies in the Indian Ocean, while those selected for the Murrumbidgee catchment are dominated by climate anomalies in the Pacific Ocean. Reports in the literature support the finding that the two catchments are influenced by different climate systems. During autumn the selected predictors seem to be augmenting the representation of the initial catchment conditions and therefore further investigation into the best indicators of initial catchment conditions are required. Future work is also required to test the robustness of the proposed predictor selection method using double cross-validation and applying it to catchments elsewhere in Australia.

ACKNOWLEDGMENTS

This work is part of the Water Information Research and Development Alliance between CSIRO's Water for a Healthy Country Flagship and the Bureau of Meteorology and is also part of the South Eastern Australian Climate Initiative. The research is also supported by the CSIRO CEO Science Leader Scheme. We would like to thank a number of colleagues at the Australian Bureau of Meteorology for useful discussions. We would also like to thank Andrew Frost and Eun-Pa Lim of the Australian Bureau of Meteorology for reviewing the paper.

REFERENCES

Cai, W., and T. Cowan (2008), Dynamics of late autumn rainfall reduction over southeastern Australia, *Geophysical Research Letters*, *35*(9), 5.

Chiew, F. H. S., T. C. Piechota, J. A. Dracup, and T. A. McMahon (1998), El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting, *Journal of Hydrology*, *204*(1-4), 138-149.

Gelfand, A. E. (1995), Model determination using sampling-based methods, in *Markov Chain Monte Carlo in Practice*, edited by W. R. Gilks, et al., pp. 145-159, Chapman & Hall/CRC, London.

Marshall, G. J. (2003), Trends in the southern annular mode from observations and reanalyses, *Journal of Climate*, *16*(24), 4134-4143.

McBride, J. L., and N. Nicholls (1983), Seasonal relationships between Australian rainfall and the Southern Oscillation, *Monthly Weather Review*, *111*(10), 1998-2004.

Piechota, T. C., F. H. S. Chiew, J. A. Dracup, and T. A. McMahon (2001), Development of Exceedance Probability Streamflow Forecast, *Journal of Hydrologic Engineering*, *6*(1), 20-28.

Plummer, N., N. K. Tuteja, Q. J. Wang, E. Wang, D. E. Robertson, S. Zhou, A. Schepen, O. Alves, B. Timbal, and K. Puri (2009), A seasonal water availability prediction service: Opportunities and Challenges, in *18th World IMACS/MODSIM Congress*, edited, Modelling and Simulation Society of Australia and New Zealand Inc., Cairns.

Potts, J. M., C. K. Folland, I. T. Jolliffe, and D. Sexton (1996), Revised "LEPS" scores for assessing climate model simulations and long-range forecasts, *Journal of Climate*, *9*(1), 34-53.

Robertson, D. E., and Q. J. Wang (2009), A Bayesian joint probability approach to seasonal prediction of streamflows: Predictor selection and skill assessment in *H2009: 32nd Hydrology and Water Resources Symposium*, edited, Newcastle.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore (2008), Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006), *Journal of Climate*, *21*(10), 2283-2296.

Ummenhofer, C. C., M. H. England, P. C. McIntosh, G. A. Meyers, M. J. Pook, J. S. Risbey, A. S. Gupta, and A. S. Taschetto (2009), What causes southeast Australia's worst droughts?, *Geophysical Research Letters*, *36*, 5.

Vehtari, A., and J. Lampinen (2002), Bayesian model assessment and comparison using cross-validation predictive densities, *Neural Computation*, *14*(10), 2439-2468.

Wang, Q. J., D. E. Robertson, and F. H. S. Chiew (2009), A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites, *Water Resour. Res.*, *45*.