David McInerney, Mark Thyer, Dmitri Kavetski, Richard Laugesen, Narendra Tuteja, and George Kuczera

The MuTHRE Model for High Quality Sub-seasonal Streamflow Forecasts

Proceedings of the Hydrology and Water Resources Symposium (HWRS 2021), 2021, pp.1-9

ISBN number for HWRS 2021 is 978-1-925627-53-4

The MuTHRE Model for High Quality Sub-seasonal Streamflow Forecasts

David McInerney[(1*)], Mark Thyer[(1)], Dmitri Kavetski[(1)], Richard Laugesen[(2)], Narendra Tuteja[(2)] and George Kuczera[(3)]

[(1)] School of Civil, Environmental and Mining Engineering, University of Adelaide, SA, Australia

[(2)] Bureau of Meteorology, Canberra, ACT, Australia

[(3)] School of Engineering, University of Newcastle, Callaghan, NSW, Australia

[*] david.mcinerney@adelaide.edu.au

ABSTRACT

Sub-seasonal streamflow forecasts, with lead times up to 30 days, can provide valuable information for water management, including reservoir operation to meet environmental flow, irrigation demands, and managing flood protection storage. A key aim is to produce "seamless" probabilistic forecasts, with high quality performance across the full range of lead times (1-30 days) and time scales (daily to monthly).

This paper demonstrates that the Multi-Temporal Hydrological Residual Error (MuTHRE) model can address the challenge of "seamless" sub-seasonal forecasting. The MuTHRE model is designed to capture key features of hydrological errors, namely seasonality, dynamic biases due to hydrological non-stationarity, and extreme errors poorly represented by the common Gaussian distribution.

The MuTHRE model is evaluated comprehensively over 11 catchments in the Murray-Darling Basin using multiple performance metrics, across a range of lead times, months and years, and at daily and monthly time scales. It is shown to provide "high" improvements, in terms of reliability for short lead times (up to 10 days), in dry months, and dry years. Forecast performance also improved in terms of sharpness. Importantly, improvements are consistent across multiple time scales (daily and monthly).

This study highlights the benefits of modelling multiple temporal characteristics of hydrological errors, and demonstrates the power of the MuTHRE model for producing seamless sub-seasonal streamflow forecasts that can be utilized for a wide range of applications.

INTRODUCTION

Water management and operations across large river basins have historically focused on releasing and delivering water for consumptive purposes (e.g. irrigation), under relatively controlled and predictable flow conditions. Prolonged dry conditions and water scarcity in recent decades in many major river basins around the world supporting large populations, have led to the development of integrated water resource management plans that set the amount of water that can be taken from the basin each year, while leaving enough environmental water for the rivers, lakes and wetlands and the plants and animals that depend on them (e.g. Murray-Darling Basin Plan, Hart 2016). Environmental water management is complex, requiring the release of large volumes of environmental water from storages to be delivered over long distances at sub-seasonal or longer time scales to achieve a range of environmental targets and outcomes, under both regulated and unregulated conditions. To ensure that future water delivery optimises consumptive as well as environmental outcomes, new forecasting and planning tools, and streamlined processes are necessary especially at sub-seasonal time scale. In particular, considerable benefit can be obtained by producing probabilistic sub-seasonal forecasts which are "seamless" in time; i.e. from a single product that is reliable and sharp across a range of lead times and aggregation time scales (White et al. 2017).

Streamflow forecasts are subject to uncertainty in rainfall, associated with predicting future rainfall, and hydrological errors, associated with uncertainty in model structure, initial conditions and parameters. In order to represent both sources of uncertainty in streamflow forecasts, the "ensemble dressing" approach is often implemented, whereby (i) replicates of forecast rainfall are propagated through a rainfall-runoff model, and (ii) a residual error model is used to add hydrological errors to each streamflow replicate.

A key challenge is the development of the residual error model, which must capture relevant features of hydrological errors. It is well-known that hydrological errors are heteroscedastic (larger errors for larger flows) and persistent (similar errors for consecutive times), and these features are typically represented in residual error models (e.g., Kuczera 1983; Bates and Campbell 2001; McInerney et al. 2017). However, other important features which are less commonly represented include

- *Seasonal variability*, due to hydrological models being unable to appropriately capture seasonal variations in streamflow (e.g., Woldemeskel et al. 2018);
- *Dynamic biases,* i.e., shifts in the mean of hydrological errors over longer time periods (e.g. month to year) due to hydrological non-stationarity (e.g., Westra et al. 2014);
- *Non-Gaussian errors*. The random component (innovation) of residual error models are commonly assumed to follow a Gaussian distribution (Seber and Wild 1989; Bates and Campbell 2001; McInerney et al. 2018). However, recent studies have found that non-Gaussian distributions better capture extreme errors (e.g., Li et al. 2016).

The Multi-Temporal Hydrological Residual Error (MuTHRE) model is the first residual error model (to the best of the authors' knowledge) which represents these three temporal error characteristics (i.e., seasonal variability, dynamic biases and non-Gaussian innovations).

The aims of this conference paper are to

(i) Demonstrate the ability of the MuTHRE model for producing seamless sub-seasonal streamflow forecasts, and

(ii) Describe the benefits of seamless sub-seasonal forecasts for practitioners

The remainder of this paper is structured as follows. We begin by presenting the *Theory* of the MuTHRE model, then describe the *Case Study* and its *Results*, and finish with a discussion on the *Benefits of Seamless Sub-seasonal Streamflow Forecasts*, and a summary of our key *Conclusions*.

THEORY

MuTHRE model for representing hydrological uncertainty

The MuTHRE model represents hydrological uncertainty in streamflow q_t (at time step t) through a probability model Q_t (i.e. $q_t \sim Q_t$). The probability model combines a deterministic term q_t^{det} and a residual error term η_t in transformed space,

$$z(Q_t;\boldsymbol{\theta}_z) = z(q_t^{\text{det}};\boldsymbol{\theta}_z) + \eta_t \tag{1}$$

where z represents the Box-Cox transformation (Box and Cox 1964), with power parameter $\lambda = 0.2$ (McInerney et al. 2017).

The deterministic term is

$$q_t^{\text{det}} = h(\boldsymbol{\theta}_h;\mathbf{x}_t,\mathbf{s}_{t-1}) \tag{2}$$

where h is a rainfall-runoff model with parameters $\boldsymbol{\theta}_h$, inputs $\mathbf{x}_t$ and initial conditions $\mathbf{s}_{t-1}$,

The residual error term follows as AR(1) model

$$\eta_t = \phi_\eta(\eta_{t-1} - \mu_{t-1}) + \mu_t + y_t \tag{3}$$

The residual error mean (in transformed space) is

$$\mu_t = \mu_{d(t)}^{(\text{s})} + \mu_t^{(\text{b})} \tag{4}$$

Equations (3) and (4) include the following multi-temporal components of the MuTHRE model:

- Seasonality component: $\mu_{d(t)}^{(\text{s})}$ varies according to the day-of-year $d(t)$,
- Dynamic bias component: $\mu_t^{(\text{b})}$ is intended to represent recent hydrological errors,
- Non-Gaussian innovations: y_t are represented by a two-component mixed-Gaussian distribution

$$y_t \sim \mathcal{N}_{\text{mix}}\left(0,\sigma_{y1}^2,0,\sigma_{y2}^2,w_{y1}\right) \tag{5}$$

where the component means are set to zero, σ_{y1} and σ_{y2} are the component standard deviations, and w_{y1} is the weight of component 1.

See McInerney et al. (2020) for equations used for the seasonality and dynamic bias terms ($\mu_{d(t)}^{(\text{s})}$ and $\mu_t^{(\text{b})}$), and a description of the approach used to calibrate hydrological and residual error model parameters.

Streamflow forecasts accounting for rainfall forecast uncertainty

Propagating forecast rainfall replicates through the deterministic rainfall-runoff model produces replicates for q_t^{det}. Equations (3)-(5) can be used to generate samples for η_t. These terms (q_t^{det} and η_t) are then used in the rearranged form of equation (1)

$$q_t = z^{-1}\left(z(q_t^{\text{det}}) + \eta_t\right) \tag{6}$$

to produce a single post-processed streamflow forecast replicate. This is repeated for all times and rainfall replicates.

CASE STUDY

Hydrological data and model

Forecasts from the MuTHRE model are generated in 11 catchments in the Murray Darling Basin (see Figure 1 for location of catchments and Table 1 for their properties). Daily time series of observed rainfall, PET and streamflow between 1991 and 2002 from the Australian Bureau of Meteorology's Hydrological Reference Stations (HRS) dataset are used (Zhang et al. 2014). Rainfall forecasts are obtained from the Australian Community Climate Earth-System Simulator - Seasonal (ACCESS-S) (Hudson et al. 2017), and subsequently post-processed to remove biases and improve reliability (Schepen et al. 2018). A subset of 100 forecast replicates beginning on the first day of each month and with lead times up to 28 days are used.

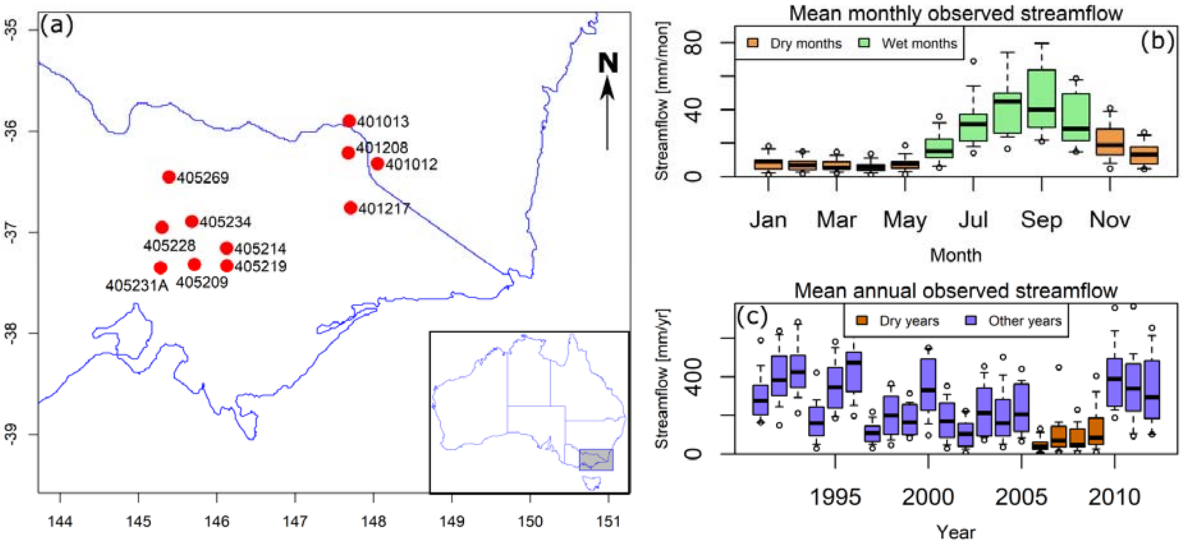

Figure 1. (a) Locations of the 11 case study catchments, (b) Mean monthly observed streamflow with wet/dry month classifications, and (c) Mean annual observed streamflow with dry/other year classifications. Boxplots in (b) and (c) represent distribution over all catchments, with boxes for inter-quartile range and whiskers for 10th/90th percentiles.

Table 1. Case study catchments and their properties

Site name	Site ID	Area (km^2)	Mean rain (mm/yr)	Mean flow (mm/yr)	Runoff ratio
Murray River at Biggara	401012	1257.3	1117	370	0.33
Jingellic Creek at Jingellic	401013	390.0	876	112	0.13
Cudgewa Creek at Berringama	401208	351.1	1127	209	0.19
Gibbo River at Gibbo Park	401217	389.8	1138	273	0.24
Delatite River at TongaBridge	405214	368.0	959	248	0.26
King Parrot Creek at Flowerdale	405231A	181.0	999	187	0.19
Seven Creeks river at Kialla West	405269	1513.4	655	93	0.14
Seven at D/S Polly McQuinns Weir	405234	147.6	852	226	0.27
Hughes Creek at Tarcombe Rd	405228	474.8	760	116	0.15
Acheron River at Taggerty	405209	629.4	1234	443	0.36
Goulburn River at Dohertys	405219	700.2	1156	424	0.37

We use the GR4J conceptual rainfall-runoff model (Perrin, Michel, and Andreassian 2003), which has four calibration parameters, as the determistic model h in equation (2). A leave-one-year out cross-validation procedure is used to calibrate and evaluate forecasts.

Forecast evaluation

We evaluate forecast performance in terms of reliability and sharpness. Reliability of forecasts (i.e. the degree of statistical consistency between observations and the forecast distribution) is quantified using the metric of Evin et al. (2014). Sharpness (i.e., uncetainty in the forecast distribution) is quantified as a skill score by the average ratio of the 90% limits of the forecast distribution and the 90% limits of the climatology (McInerney et al. 2020). Lower metric values indicate better performance.

Forecasts are evaluated over (i) multiple time scales from daily to aggregated monthly forecasts, and (ii) multiple stratification types, including by lead time, month and year.

Comparison with "baseline" error model

Forecasts from the MuTHRE model are compared against a "baseline" model which does not include the multi-temporal components. The baseline model is obtained by setting $\mu_{d(t)}^{(s)} = 0$ (i.e. removing seasonality) and $\mu_t^{(b)} = 0$ (i.e. removing dynamic biases) in equation (4), and $w_{y1} = 0$ in equation (5) (i.e. reverting to a standard Gaussian).

"Practical significance" tests are used to compare the two models across the range of catchments (McInerney et al. 2019). This determines whether any differences (e.g. improvements) in metric values are of practical relevance, defined as a difference of more than 20% of the metric value for the baseline model. We note that catchment 405234 is a sub-catchment of 405269, and the forecasts from these two catchments are not entirely independent; however, the impact of this dependence on the model comparison is likely minor due to the relative sizes of these two catchments.

RESULTS

Time series

Figure 2 provides an illustration of daily and cumulative streamflow forecasts in the Biggara catchment (401012) during August 2002. For daily forecasts, the observations lie within the 90% predictive limits for both climatology (Figure 2a) and the MuTHRE model (Figure 2b). However, forecasts from the MuTHRE model are much sharper than climatology. These forecasts are sharpest for short lead times, but are still considerably sharper than climatology for longer lead times. Similarly, 90% limits for cumulative forecasts from the MuTHRE model (Figure 2d) capture the observed values, and are much sharper than climatology.

Performance metrics

Daily forecasts

Figure 3 compares performance of *daily* forecasts from the MuTHRE and baseline models for different lead times. The MuTHRE model provides consistent good reliability over all lead times (Figure 3a), and practically significant improvements over the baseline model for short lead times (10 out of the first 11 days). MuTHRE forecasts are much sharper than climatology, especially for short lead times (e.g. median value of 0.2 for lead time of 1 day corresponds to 80% reduction in uncertainty). Compared to the baseline model, sharpness improves for all lead times (Figure 3b), although these are not classified as practically significant.

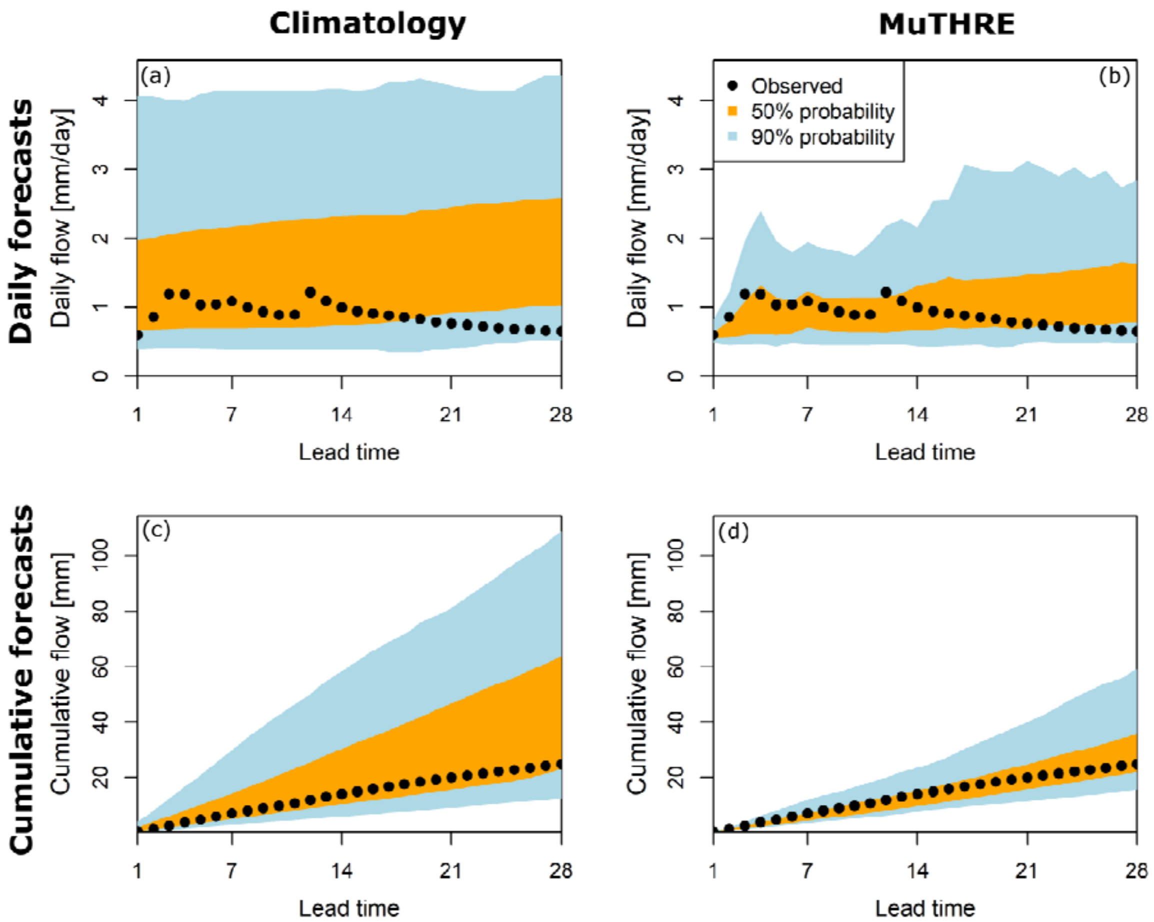

Figure 2. Predictive limits of forecasts in the Biggara catchment (401012) during August 2002. Climatology (left) is compared with the MuTHRE model (right). Results are shown for both daily forecasts (top) and cumulative forecasts (bottom).

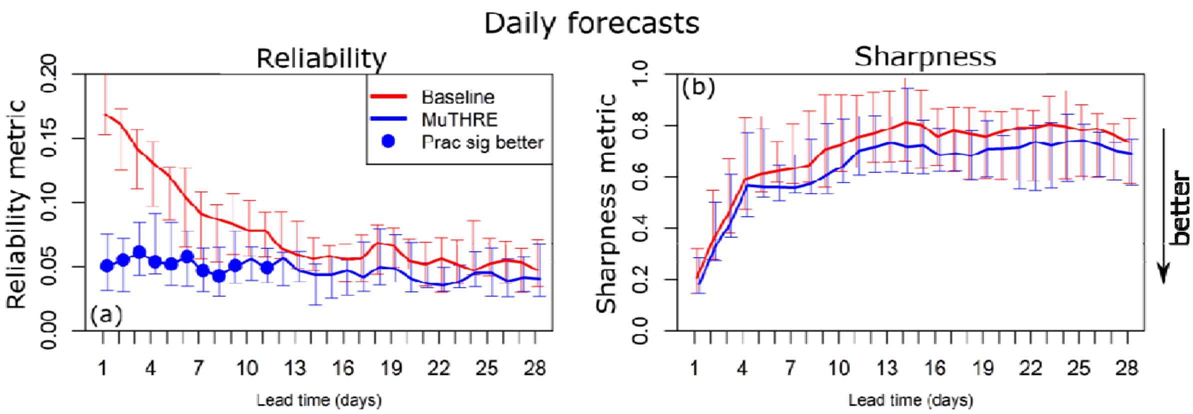

Figure 3. (a) Reliability and (b) Sharpness metrics when MuTHRE and baseline forecasts are stratified by lead time. Lines represent median metric values calculated over the 11 catchments, whiskers represent 90% limits, and circles indicates lead times for which the MuTHRE model produces practically significantly better performance than the baseline model.

Monthly forecasts

Figure 4 compares the reliability of *monthly* forecasts when stratified by month and year. The MuTHRE model provides consistent reliability over all months (Figure 4a), with practically significant improvements over the baseline model in 6/7 dry months (November-May). The MuTHRE model also provides improvements in reliability when stratified by year – these are largest in dry years (2006-2009).

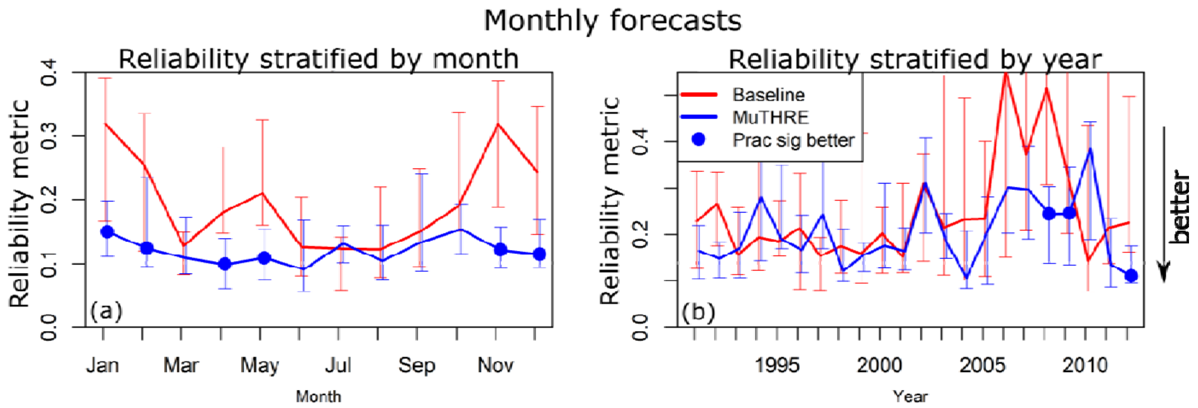

Figure 4. Reliability metrics when MuTHRE and baseline forecasts are stratified by (a) month, and (b) year.

BENEFITS OF SEAMLESS SUB-SEASONAL STREAMFLOW FORECASTS

The results presented in this paper show that the MuTHRE model produces seamless sub-seasonal forecasts, with consistent reliability over lead times (1-30 days), timescales (daily to monthly), and all months and years.

Consistent reliability enables water resource managers and planners to confidently utilize sub-seasonal forecasts for decision support in a wide range of applications. Examples include

- The easy integration of streamflow forecasts into exisiting river system models, such as eWater Source (Welsh et al. 2013). These models are typically run at the daily scale, but are used for managing multiple resource management decisions over a range of time scales, so it is important that streamfow forecasts are reliable at both daily and longer time scales. River system models are commonly run with historical streamflow inputs, so utilizing reliable and sharp sub-seasonal streamflow forecasts would enable improved decision making through better quantifying uncertainty.
- Improved reservoir management for rural water supply systems with demands for irrigation water use and environmental flows. Utilising forecasts for long timescales (e.g. weeks/months) and lead times (up to 1 month) can improve the management of these systems (Murray-Darling Basin Authority 2019). For example, if a high streamflow event is forecasted, the water supply authority can delay/avoid releasing water for environmental flows, and prevent wasting water.
- Forecast informed flood control. Multi-purpose reservoirs serve as both water supply and downstream flood protection services. Sub-seasonal forecasts can inform the management of the systems. For example, large volumes of streamflow are forecast, reservoir operators can release water in advance to provide additional flood storage and reduce risks of flooding.
- Operation of urban water supply systems, which benefit from aggregated monthly forecasts (Zhao and Zhao 2014). Reliable forecasts can inform managers about whether urban demand can be met from river flows, or whether water needs to be transferred between multiple reservoirs or sourced from costly desalination.

CONCLUSIONS

The Multi-Temporal Hydrological Residual Error (MuTHRE) model has been developed for seamless sub-seasonal streamflow forecasting. This model accounts for charactersitics of hydrological errors at multiple time scales, including seasonality, dynamic biases, and extreme errors.

The MuTHRE model has been applied to 11 case study catchments in the Murray Darling Basin, using the GR4J rainfall-runoff model and post-processed ACCESS-S rainfall forecasts. It is shown to produce reliable forecasts over all lead times, aggregation scales (daily and monthly), and when stratified by lead time, month and year. These forecasts are much sharper than climatology. When compared to a baseline model (which does not include multi-temporal components), the MuTHRE model provides large improvements in relaibility for short lead times, dry months and dry years, as well as improvements in sharpness.

Sub-seasonal streamflow forecasts are valuable for a wide range of water management applications. The consistent high quality performance of the MuTHRE model over multiple lead times (1-30 days) and time scales (daily to monthly) provides confidence in the suitability of forecasts for multiple practical applications, including their use in river system models to optimize water delivery for irrigation and enorinmental outcomes.

Further information on the MuTHRE model, including comprehensive analysis of forecast performance and detailed algorithms, can be found in McInerney et al. (2020).

REFERENCES

Bates, B. C., and E. P. Campbell. 2001. 'A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling', *Water Resources Research*, 37: 937-47, doi:10.1029/2000WR900363.

Box, George EP, and David R Cox. 1964. 'An analysis of transformations', *Journal of the Royal Statistical Society. Series B (Methodological)*: 211-52.

Cloke, H. L., and F. Pappenberger. 2009. 'Ensemble flood forecasting: A review', *Journal of Hydrology*, 375: 613-26.

Evin, G., M. Thyer, D. Kavetski, D. McInerney, and G. Kuczera. 2014. 'Comparison of joint versus postprocessor approaches for hydrological uncertainty estimation accounting for error autocorrelation and heteroscedasticity', *Water Resources Research*, 50: 2350-75.

Hart, Barry T. 2016. 'The Australian Murray–Darling basin plan: challenges in its implementation (part 1)', *International Journal of Water Resources Development*, 32: 819-34.

Hudson, D., O. Alves, H. H. Hendon, E. Lim, G. Liu, J.J. Luo, C. MacLachlan, A. G. Marshall, L. Shi, G. Wang, R. Wedd, G. Young, M. Zhao, and X. Zhou. 2017. 'ACCESS-S1 The new Bureau of Meteorology multi-week to seasonal prediction system', *Journal of Southern Hemisphere Earth System Sciences*, 67: 132-59.

Kuczera, G. 1983. 'Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty', *Water Resources Research*, 19: 1151-62, doi:10.029/WR019i005p01151.

Li, M., Q. J. Wang, J. C. Bennett, and D. E. Robertson. 2016. 'Error reduction and representation in stages (ERRIS) in hydrological modelling for ensemble streamflow forecasting', *Hydrol. Earth Syst. Sci.*, 20: 3561-79.

McInerney, David, Dmitri Kavetski, Mark Thyer, Julien Lerat, and George Kuczera. 2019. 'Benefits of Explicit Treatment of Zero Flows in Probabilistic Hydrological Modeling of Ephemeral Catchments', *Water Resources Research*, 55: 11035-60.

McInerney, David, M Thyer, D Kavetski, B Bennett, J. Lerat, M. Gibbs, and G Kuczera. 2018. 'A simplified approach to produce probabilistic hydrological model predictions', *Environmental Modelling & Software*, 109: 306-14.

McInerney, David, Mark Thyer, Dmitri Kavetski, Richard Laugesen, Narendra Tuteja, and George Kuczera. 2020. 'Multi-temporal Hydrological Residual Error Modeling for Seamless Subseasonal Streamflow Forecasting', *Water Resources Research*, 56: e2019WR026979.

McInerney, David, Mark Thyer, Dmitri Kavetski, Julien Lerat, and George Kuczera. 2017. 'Improving probabilistic prediction of daily streamflow by identifying Pareto optimal approaches for modeling heteroscedastic residual errors', *Water Resources Research*, 53.

Murray-Darling Basin Authority. 2019. 'Joint management of the River Murray', Accessed 11/12/2019. https://www.mdba.gov.au/river-murray-system/joint-management-river-murray.

Perrin, C., C. Michel, and V. Andreassian. 2003. 'Improvement of a parsimonious model for streamflow simulation', *Journal of Hydrology*, 279: 275-89, doi:10.1016/S0022-1694(03)00225-7.

Schepen, A., T. Zhao, Q. J. Wang, and D. E. Robertson. 2018. 'A Bayesian modelling method for post-processing daily sub-seasonal to seasonal rainfall forecasts from global climate models and evaluation for 12 Australian catchments', *Hydrol. Earth Syst. Sci.*, 22: 1615-28.

Seber, GAF, and CJ Wild. 1989. "Nonlinear regression. 1989." In.: Wiley, New York.

Welsh, Wendy D., Jai Vaze, Dushmanta Dutta, David Rassam, Joel M. Rahman, Ian D. Jolly, Peter Wallbrink, Geoffrey M. Podger, Matthew Bethune, Matthew J. Hardy, Jin Teng, and Julien Lerat. 2013. 'An integrated modelling framework for regulated river systems', *Environmental Modelling & Software*, 39: 81-102.

Westra, Seth, Mark Thyer, Michael Leonard, Dmitri Kavetski, and Martin Lambert. 2014. 'A strategy for diagnosing and interpreting hydrological model nonstationarity', *Water Resources Research*, 50: 5090-113.

White, Christopher J., Henrik Carlsen, Andrew W. Robertson, Richard J.T. Klein, Jeffrey K. Lazo, Arun Kumar, Frederic Vitart, Erin Coughlan de Perez, Andrea J. Ray, Virginia Murray, Sukaina Bharwani, Dave MacLeod, Rachel James, Lora Fleming, Andrew P. Morse, Bernd Eggen, Richard Graham, Erik Kjellström, Emily Becker, Kathleen V. Pegion, Neil J. Holbrook, Darryn McEvoy, Michael Depledge, Sarah Perkins-Kirkpatrick, Timothy J. Brown, Roger Street, Lindsey Jones, Tomas A. Remenyi, Indi Hodgson-Johnston, Carlo Buontempo, Rob Lamb, Holger Meinke, Berit Arheimer, and Stephen E. Zebiak. 2017. 'Potential applications of subseasonal-to-seasonal (S2S) predictions', *Meteorological Applications*, 24: 315-25.

Woldemeskel, F., D. McInerney, J. Lerat, M. Thyer, D. Kavetski, D. Shin, N. Tuteja, and G. Kuczera. 2018. 'Evaluating residual error approaches for post-processing monthly and seasonal streamflow forecasts', *Hydrol. Earth Syst. Sci. Discuss.*, 2018: 1-40.

Zhang, Sophie Xiaoyong, Mohammed Bari, Gnanathikkam Amirthanathan, David Kent, Andrew MacDonald, and Daehyok Shin. 2014. "Hydrologic reference stations to monitor climate-driven streamflow variability and trends." In *Hydrology and Water Resources Symposium 2014*, 1048. Engineers Australia.

Zhao, Tongtiegang, and Jianshi Zhao. 2014. 'Joint and respective effects of long- and short-term forecast uncertainties on reservoir operations', *Journal of Hydrology*, 517: 83-94.