Journal of Hydrology 555 (2017) 586–599

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Improved error modelling for streamflow forecasting at hourly time steps by splitting hydrographs into rising and falling limbs

Ming Li [a,*], Q.J. Wang [b], David E. Robertson [c], James C. Bennett [c]

[a] CSIRO Data61, Floreat, WA, Australia
[b] Department of Infrastructure Engineering, The University of Melbourne, Parkville, Victoria, Australia
[c] CSIRO Land and Water, Clayton, Victoria, Australia

ARTICLE INFO

Article history:
Received 2 August 2017
Received in revised form 23 October 2017
Accepted 24 October 2017

Keywords:
Streamflow forecasting
Ensemble prediction
Multiple lead times
Forecast verification
Hydrograph
Hydrological uncertainty

ABSTRACT

Following the development of *Error Reduction and Representation In Stages* (ERRIS) for daily streamflow forecasting, we extend ERRIS to streamflow forecasting at an hourly time step (ERRIS-h). ERRIS applies a staged error model to reduce errors in hydrological simulations and to quantify prediction uncertainty. ERRIS produces probabilistic predictions, and is capable of propagating errors through multiple lead times to generate ensemble traces. In this study, we identify the need to model the residual distribution differently for rising and falling limbs of hydrographs when applying ERRIS to hourly streamflow forecasting. To address this need, ERRIS-h uses different distribution parameters for the two limbs.

We evaluate ERRIS-h on eight rivers in Australia. Hourly streamflow simulations are produced by forcing an initialized GR4H hourly rainfall-runoff model with observed rainfall. We apply ERRIS-h to the streamflow simulations to produce ensemble streamflow predictions with lead times up to 48 h. The ensemble streamflow predictions here can be viewed as forecasts when rainfall forecasts are perfect. In this way, we test the ability of ERRIS-h to update forecasts using the most up-to-date streamflow observations and to generate ensemble traces that reflect hydrological uncertainty.

As expected, ERRIS-h is highly effective when applied to the zeroth lead time, dramatically reducing errors in the original GR4H simulations and reliably describing forecast uncertainty. We also show that ERRIS-h ensemble forecasts have smaller errors than deterministic simulations at all lead times and are reliable in ensemble spread even at 48 h lead times.

1. Introduction

Short-term streamflow forecasts, particularly at the hourly time step, deliver detailed estimates of the volume and timing of streamflow events over a future period. Effective streamflow forecasts provide information for emergency services and support timely mitigation of the impacts of floods. Streamflow forecasts have also become a crucial component of water resources operation and planning.

Streamflow forecasts are uncertain, and the quantification of forecast uncertainty should accompany any forecast (Krzysztofowicz, 2001). Quantifying uncertainty, for example by issuing ensemble forecasts, provides additional information to end-users, enables risked-based flood warning and aids rational decision-making (e.g. Arnal et al., 2016; Raftery, 2016).

Australia is marked by hydrological extremes: it is the driest inhabited continent, but also suffers from sporadic, devastating floods. The Bureau of Meteorology (BoM) is Australia's major agency responsible for issuing streamflow forecasts and flood warnings. At present, models that underpin short-term streamflow and flood forecasts issued by the BoM are deterministic. Ensemble forecasting systems offer the promise of more accurate forecasts, as well as quantifying forecast uncertainty (e.g. Bennett et al., 2014a,b; Demargne et al., 2014). Accordingly, the BoM is seeking to transition from deterministic forecasting to ensemble forecasting.

The most common operational streamflow forecasting approach is to force an initialised conceptual rainfall-runoff model with forecast rainfall and potential evaporation (PE). Forecast rainfall is a far more important determinant of streamflow than PE (often climatology PE is used as a surrogate 'forecast') and we confine our discussion to forecast rainfall. Uncertainty in streamflow forecasts can be conveniently ascribed to two sources: uncertainty in rainfall forecasts and uncertainty in

* Corresponding author at: CSIRO Data61, Private Bag 5, Wembley, WA 6014, Australia.
E-mail address: Ming.Li@csiro.au (M. Li).

https://doi.org/10.1016/j.jhydrol.2017.10.057
0022-1694/

hydrological modelling. Many researchers have quantified uncertainty in rainfall forecasts by the use of quantitative precipitation forecasts (QPF) from numerical weather prediction (NWP) models. Some forced a hydrological model with every member of an ensemble QPF (e.g. Alfieri et al., 2012; Bartholmes and Todini, 2005; Cloke and Pappenberger, 2009; Cuo et al., 2011; Wetterhall et al., 2013). Some post-processed deterministic QPFs to interpret the full spectrum of rainfall forecast uncertainty (Bennett et al., 2014a; Robertson et al., 2013; Shrestha et al., 2015). Accounting only for rainfall uncertainty, however, usually results in streamflow forecasts that are overconfident, especially at short lead times where hydrological uncertainty is often the dominant source of uncertainty (Bennett et al., 2014a).

Hydrological uncertainty arises from imperfections in observations, model structure and model parameters that make it difficult for a hydrological model to reproduce hydrographs exactly, even with perfect forcings (i.e., observed rainfall and PE). Some methods based on the Bayesian framework have been developed to characterise modelling uncertainty using parameter uncertainty, for example the generalised likelihood uncertainty estimation (GLUE) (Beven and Binley, 1992), the hydrologic uncertainty processor (HUP) (Krzysztofowicz and Maranzano, 2004), the Bayesian total error analysis (BATEA) (Kavetski et al., 2006) and the ensemble Bayesian forecasting system (EBFS) (Herr and Krzysztofowicz, 2015). Other methods characterise modelling uncertainty by using statistical techniques to describe the behaviour of the model residuals (e.g. Evin et al., 2013; Gragne et al., 2015a,b; Schaefli et al., 2007; Solomatine and Shrestha, 2009). These methods, typically called error modelling methods, consider only the overall forecast errors without addressing the sources of errors. Error modelling methods are generally much easier to implement than the attempts to describe all separate sources of uncertainty and are particularly useful for real-time forecasting operations.

Li et al. (2016) developed an error modelling method, called *Error Reduction and Representation In Stages* (ERRIS), to quantify the hydrological model uncertainty in forecasts as well as to reduce forecast errors. ERRIS is a staged error model that uses a sequence of simple error models (instead of a single complex one) to address the statistical properties of the residuals progressively. ERRIS requires few computational resources and can be used for real-time streamflow forecasting systems. ERRIS was tested for daily streamflow simulations and was shown to improve accuracy substantially and to quantify the simulation uncertainty reliably.

This paper presents an extension of ERRIS for hourly streamflow forecasting. Hourly streamflow tends to differ from daily streamflow in important ways. For example, hourly flow and model residuals tend to be more autocorrelated, which, as we will show, can change the distribution of residuals. Further, forecast updating and uncertainty propagation must be carried out 24 times as often for hourly as for daily models; correctly propagating uncertainty over many time increments is more challenging, as we will discuss. In this study, we evaluate the suitability of ERRIS for hourly streamflow forecasting. A revised ERRIS model is introduced and evaluated on eight Australian rivers covering different climate and hydrological conditions. The forecast performance with respect to multiple lead times is evaluated by various forecast verification measures.

The manuscript is organised as follows: Section 2 reviews the original ERRIS and describes the revised model structure. Section 3 describes the background of the example application. Section 4 presents the main results and findings of this study. Section 5 provides concluding remarks and discussion.

2. Methods

2.1. Review of ERRIS

ERRIS consists of four stages; each can be considered a simple error model. Each error model successively improves the characterisation of the residuals.

All error models in ERRIS consider residuals in transformed space. For data transformation, ERRIS makes use of the log-sinh transformation (Wang et al., 2012) to normalise raw data, q, and homogenise its variance:

$$h(q) = b^{-1}\log\{\sinh(a + bq)\}. \quad (1)$$

a and b are transformation parameters. The back-transformation h^{-1} converts forecasts in the transformed space back to the original space.

We define the following notations to describe model formulation.

$q(t)$ Observed streamflow
$\tilde{q}(t)$ Simulated streamflow
$z(t) = h\{q(t)\}$ Observed streamflow in the transformed space
$\tilde{z}_i(t)$ Forecast median in the transformed space at Stage i
$\tilde{q}_i(t) = h^{-1}\{\tilde{z}_i(t)\}$ Forecast median in the original space at Stage i
$e_i(t) = z(t) - \tilde{z}_i(t)$ Model residual at Stage i

The error model at each stage contains two components: the definition of $\tilde{z}_i(t)$ and the probability distribution of $e_i(t)$.

Stage 1: Data normalization

The Stage 1 error model simply transforms data:

$$\tilde{z}_1(t) = h\{\tilde{q}(t)\} \quad (2)$$

$$e_1(t) \sim N(0, \sigma_1^2), \quad (3)$$

where σ_1 is the standard deviation of the model residual at Stage 1. The transformation ensures residuals are normally distributed and homoscedastic. It is an option to jointly estimate the hydrological model parameters with transformation parameters. Alternatively, we can simply 'bolt on' ERRIS to already calibrated hydrological models for convenience (Li et al., 2016).

Stage 2: Conditional bias-correction

To correct bias in the hydrological model, Stage 2 applies a conditional bias-correction (Bennett et al., 2016b) by

$$\tilde{z}_2(t) = \mu + d\tilde{z}_1 \quad (4)$$

$$e_2(t) \sim N(0, \sigma_2^2) \quad (5)$$

where μ and d are the bias-correction parameters and σ_2 is the standard deviation of the model residual at Stage 2. This stage is to refine the streamflow simulations when the base hydrological model has not been adequately calibrated. The stage may be skipped if the base hydrological model has been calibrated by using methods that produce conditionally unbiased simulations, for example the method of maximum likelihood or weighted least squares. ERRIS assumes that residuals are unbiased, thus if using calibration methods that may produce biased simulations (e.g. deterministic objectives such as the Nash-Sutcliffe efficiency), the bias-correction may be necessary. We note, however, that the bias-correction requires reasonably long records of observations to ensure bias can be accurately characterised (i.e., observations that cover both wet and dry periods). If only short records (e.g. <5 years) are available, we advise caution when applying the bias-correction.