Hydrol. Earth Syst. Sci. Discuss., 11, 6035–6063, 2014
www.hydrol-earth-syst-sci-discuss.net/11/6035/2014/
doi:10.5194/hessd-11-6035-2014

Hydrology and Earth System Sciences
Discussions
Open Access

A strategy to overcome adverse effects of autoregressive updating of streamflow predictions

M. Li[1], Q. J. Wang[2], J. C. Bennett[2], and D. E. Robertson[2]

[1]CSIRO Computational Informatics, Floreat, Western Australia, Australia
[2]CSIRO Land and Water, Highett, Victoria, Australia

Received: 30 April 2014 – Accepted: 21 May 2014 – Published: 10 June 2014

Correspondence to: M. Li (ming.li@csiro.au)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

For streamflow forecasting applications, rainfall–runoff hydrological models are often augmented with updating procedures that correct streamflow predictions based on the latest available observations of streamflow and their departures from model simulations. The most popular approach uses autoregressive (AR) models that exploit the "memory" in hydrological model simulation errors. AR models may be applied to raw errors directly or to normalised errors. In this study, we demonstrate that AR models applied in either way can sometimes cause over-correction of predictions. In using an AR model applied to raw errors, the over-correction usually occurs when streamflow is rapidly receding. In applying an AR model to normalised errors, the over-correction usually occurs when streamflow is rapidly rising. Furthermore, when parameters of a hydrological model and an AR model are estimated jointly, the AR model applied to normalised errors sometimes degrades the stand-alone performance of the base hydrological model. This is not desirable for forecasting applications, as predictions should rely as much as possible on the base hydrological model, and updating should be applied only to correct minor errors. To overcome the adverse effects of the ordinary AR models, a restricted AR model applied to normalised errors is introduced. The new model is evaluated on a number of catchments and is shown to reduce over-correction and to improve the performance of the base hydrological model considerably.

1 Introduction

Rainfall–runoff models are widely used to generate streamflow forecasts, which provide essential information for flood warning and water resources management. For streamflow forecasting, rainfall–runoff models are often augmented by updating procedures that correct streamflow predictions based on the latest available observations of streamflow and their departures from model predictions. Model prediction errors reflect limitations of the hydrological models in reproducing physical processes as well as inaccuracies in data used to force and evaluate the models.

The most popular updating approach uses autoregressive (AR) models, which exploit the "memory" – more precisely the autocorrelation structure – of prediction errors (Kavetski et al., 2003). Essentially, AR updating uses a linear function of the known prediction errors at previous time steps to anticipate prediction errors in a forecast period. Predictions are then updated according to these anticipated errors. AR updating is conceptually simple and yet generally leads to significantly improved predictions (World Meteorological Organization, 1992). AR updating has been shown to provide equivalent performance to more sophisticated non-linear and nonparametric updating procedures (Xiong and O'Connor, 2002).

In rainfall–runoff modelling, model errors are generally heteroscedastic (i.e., they have heterogeneous variance over time) (Xu, 2001; Kavetski et al., 2003) and non-Gaussian (Bates and Campbell, 2001; Schaefli et al., 2007; Shrestha and Solomatine, 2008). In many applications (Seo et al., 2006; Bates and Campbell, 2001; Salamon and Feyen, 2010; Morawietz et al., 2011), AR models are applied to normalised errors that are considered homoscedastic and Gaussian. Normalisation is often achieved through variable transformation by using, for example, the Box–Cox transformation (Thyer et al., 2002; Bates and Campbell, 2001; Engeland et al., 2010) or, more recently, the log–sinh transformation (Wang et al., 2012; Del Giudice et al., 2013). In other applications (Schoups and Vrugt, 2010; Schaefli et al., 2007), AR models are applied directly to raw errors, but residual errors of the AR models may be explicitly specified as heteroscedastic and non-Gaussian.

There is no agreement on whether it is better to apply an AR model to normalised or raw errors. Recent work by Evin et al. (2013) found that an AR model applied to raw errors may lead to poor performance with exaggerated predictive uncertainty. They demonstrated that such instability can be mitigated by applying an AR model to standardised errors (raw errors divided by standard deviations). Here, standardisation has a similar effect to normalisation in that it homogenises the variance of the errors

HESSD
11, 6035–6063, 2014

(but does not consider the non-Gaussian distribution of errors). Conversely, Schaefli et al. (2007) pointed out that when an AR model is jointly estimated with a hydrological model, there is a clear advantage in applying an AR model to raw errors rather than normalised (or standardised) errors. Schaefli et al. (2007) found that using raw errors leads to more reliable parameter inference and uncertainty estimation, because the mean error of the predictions is close to zero and therefore the predictions are free of systematic bias. The same is not necessarily true when applying an AR model to normalised errors.

In this study, we evaluate AR models applied to both raw and normalised errors in four catchments. We show that when estimated jointly with a hydrological model, the AR model applied to normalised errors sometimes degrades the stand-alone performance of the base hydrological model. We also identify that both of these ordinary AR models can sometimes cause over-correction of predictions. We introduce a restricted AR model applied to normalised errors and demonstrate its effectiveness in overcoming the adverse effects of the ordinary AR models.

2 Autoregressive error models

2.1 Formulations

We denote the observed streamflow and modelled streamflow at time t by Q_t and $Q_{\mathrm{S},t}$, respectively. A hydrological model is a function of forcing variables (precipitation and potential evapotranspiration), initial catchment state, S_0, and a set of hydrological model parameters, θ_H. An error model is used to describe the difference between $Q_{\mathrm{S},t}$ and Q_t. In this study, we firstly examine two first-order AR error models:

i. an AR error model applied to normalised errors (referred to as *AR-Norm*) defined by:

$$Z^{(\mathrm{N})}(Q_t) = Z^{(\mathrm{N})}(Q_{\mathrm{S},t}) + \rho^{(\mathrm{N})} \left\{ Z^{(\mathrm{N})}(Q_{t-1}) - Z^{(\mathrm{N})}(Q_{\mathrm{S},t-1}) \right\} + \varepsilon_t^{(\mathrm{N})}, \quad (1)$$

HESSD
11, 6035–6063, 2014

Adverse effects of autoregressive updating of streamflow predictions

M. Li et al.

ii. an AR error model applied to raw errors (referred to as *AR-Raw*) defined by

$$Z^{(\mathrm{R})}(Q_t) = Z^{(\mathrm{R})}\left\{Q_{\mathrm{S},t} + \rho^{(\mathrm{R})}(Q_{t-1} - Q_{\mathrm{S},t-1})\right\} + \varepsilon_t^{(\mathrm{R})}, \tag{2}$$

where

$$Z(Q) = b^{-1}\log\{\sinh(a + bQ)\} \tag{3}$$

is the logarithmic, hyperbolic-sine (log–sinh) transformation (Wang et al., 2012), ρ is the lag-1 autoregression parameter and ε_t is an identically and independently distributed Gaussian deviate with mean zero and standard deviation σ. We also assume $a > 0$ and $b > 0$. We use superscript $^{(\mathrm{N})}$ and $^{(\mathrm{R})}$ to denote parameters of AR-Norm and AR-Raw models, respectively.

The formulations given by Eqs. (1) and (2) look similar, however the updating procedures differ significantly. Both models represent the lag-one autocorrelation by an AR model structure and both employ the log–sinh transformation. However, the way the log–sinh transformation is applied differs between the two models. The AR-Norm model first applies the log–sinh transformation to the observed and modelled streamflow, and then applies the autoregression parameter ρ to the errors of the transformed streamflows. In contrast, the AR-Raw model applies the autoregression parameter ρ to the raw errors to update the model prediction, and then applies the log–sinh transformation to the observed streamflow and updated model prediction.

The median of the updated streamflow prediction (referred to as *updated streamflow*), $Q_{\mathrm{U},t}$, for the AR-Norm and AR-Raw models can be derived respectively by

$$Q_{\mathrm{U},t}^{(\mathrm{N})} = (Z^{(\mathrm{N})})^{-1}\left[Z^{(\mathrm{N})}(Q_{\mathrm{S},t}) + \rho^{(\mathrm{N})}\left\{Z^{(\mathrm{N})}(Q_{t-1}) - Z^{(\mathrm{N})}(Q_{\mathrm{S},t-1})\right\}\right], \tag{4}$$

$$Q_{\mathrm{U},t}^{(\mathrm{R})} = Q_{\mathrm{S},t} + \rho^{(\mathrm{R})}(Q_{t-1} - Q_{\mathrm{S},t-1}), \tag{5}$$

where Z^{-1} is the inverse of log–sinh transformation (or back-transformation). The magnitude of the error update by the AR-Raw model, $Q_{\mathrm{U},t}^{(\mathrm{R})} - Q_{\mathrm{S},t}$, is dependent only on the

HESSD
11, 6035–6063, 2014

difference between Q_{t-1} and $Q_{\mathrm{S},t-1}$. In contrast, the magnitude of the error update by the AR-Norm model, $Q^{(\mathrm{N})}_{\mathrm{U},t} - Q_{\mathrm{S},t}$, is dependent not only on the difference between Q_{t-1} and $Q_{\mathrm{S},t-1}$, but also on $Q_{\mathrm{S},t}$. Put differently, the AR-Norm model uses errors calculated in the transformed domain, and this means that the error in the original domain can be amplified (or reduced) by the back-transformation (Eq. 4). The AR-raw model uses errors calculated in the original domain and no back-transformation is used in $Q^{(\mathrm{R})}_{\mathrm{U},t}$ (Eq. 5), meaning that the error in the original domain cannot be amplified (or reduced). In Appendix A, we show that the AR-Norm model gives greater error updates for larger values of $Q_{\mathrm{S},t}$.

2.2 Estimation

The maximum likelihood estimation is used to estimate the hydrological model parameters and the error model parameters jointly. The likelihood functions for the AR-Norm and AR-Raw models can be written respectively as

$$L\left(\theta^{(\mathrm{N})},\theta_{\mathrm{H}}\right) = \prod_t P(Q_t|Q_{\mathrm{S},t},Q_{\mathrm{S},t-1};\theta^{(\mathrm{N})},\theta_{\mathrm{H}}) = \prod_t J_{z\to Q}\phi\left[\frac{Z^{(\mathrm{N})}(Q_t)-Z^{(\mathrm{N})}(Q_{\mathrm{S},t})-\rho^{(\mathrm{N})}\left\{Z^{(\mathrm{N})}(Q_{t-1})-Z^{(\mathrm{N})}(Q_{\mathrm{S},t-1})\right\}}{\sigma^{(\mathrm{N})}}\right] \tag{6}$$

$$L\left(\theta^{(\mathrm{R})},\theta_{\mathrm{H}}\right) = \prod_t P(Q_t|Q_{\mathrm{S},t},Q_{\mathrm{S},t-1};\theta^{(\mathrm{R})},\theta_{\mathrm{H}}) = \prod_t J_{z\to Q}\phi\left[\frac{Z^{(\mathrm{R})}(Q_t)-Z^{(\mathrm{R})}\left\{Q_{\mathrm{S},t}+\rho^{(\mathrm{R})}(Q_{t-1}-Q_{\mathrm{S},t-1})\right\}}{\sigma^{(\mathrm{R})}}\right] \tag{7}$$

where $J_{z\rightarrow Q} = \{\tanh(a + bQ_t)\}^{-1}$ is the Jacobian determinant of the log–sinh transformation and $\phi(x)$ is the standard Gaussian probability density function. The probability density function is replaced by the cumulative probability function when evaluating events of zero flow occurrences (Wang and Robertson, 2011; Li et al., 2013).

3 Description of case study

We test four catchments in southeast Australia, spanning temperate to subtropical climates (Fig. 1, Table 1). The Abercrombie River intermittently experiences periods of very low (to zero) flow, while the other four rivers flow perennially (Table 1). Streamflow data are taken from the Catchment Water Yield Estimation Tool (CWYET) dataset (Vaze et al., 2011). All catchments have high-quality streamflow records with very few missing data. Rainfall and potential evaporation data are derived from the Australian Water Availability Project (AWAP) dataset (Jones et al., 2009).

We predict daily streamflow with the GR4J rainfall–runoff model (Perrin et al., 2003). We apply updating procedures to correct model predictions. We use data from 1992 to 2005 (14 years) and generate 14-fold cross-validated streamflow predictions. The data from 1990–1991 are only used to warm up the GR4J model. For a given year, we leave out the data from that year and the following year when estimating the parameters of GR4J and error models. For example, if we wish to predict flows for 1999, we leave out data from 1999 and 2000. The removal of data from the following year (2000) is designed to minimise the impact of hydrological memory on model parameter estimation. We then predict streamflows in that year (1999) from the remaining data. All results presented in this paper are based on this cross-validation instead of calibration in order to ensure the results can be generalised to independent data.

To demonstrate the problems of over-correction of errors in updating and poor stand-alone performance of the base hydrological model, we consider only streamflow predictions for one time step ahead. We will consider longer lead times in future work. Predictions are generated using observed rainfall (i.e., a perfect rainfall forecast) as

HESSD
11, 6035–6063, 2014

input. In streamflow forecasting, forecasts may be generated from rainfall information that comes from a different source (e.g., a numerical weather prediction model). Our study is aimed at streamflow forecasting applications, so we preserve the distinction between observed and forecast forcings by referring to streamflows modelled with observed rainfall as *simulations* and those modelled with forecast rainfall as *predictions*. As the forecast rainfall we use is observed rainfall, the terms *predictions* and *simulations* are interchangeable.

4 Two adverse effects of ordinary AR error models

4.1 Over-correction

The first adverse effect of the ordinary AR models is over-correction of errors in updating. By over-correction, we mean that the AR model updates the hydrological model predictions too greatly. Over-correction is difficult to define precisely, however we will demonstrate the concept with two examples: the first example illustrates over-correction by the AR-Norm model, the second example illustrates over-correction by the AR-Raw model.

To illustrate the problem of over-correction caused by the AR-Norm model, Fig. 2 presents a 1 week time series for the Mitta Mitta catchment, showing flow predictions with GR4J before error updating (referred to as flows predicted with the *base hydrological model*) and after error updating (Note that the RAR-Norm model included in Fig. 2 will be introduced and discussed in Sect. 4. The same applies to Figs. 3–7.). Figure 2 shows that the base hydrological models consistently under-estimate the flow from 23 September 2000 to 25 September 2000, and the corresponding updating procedures successfully identify the need to compensate for this under-estimation. For the AR-Norm model, however, the correction amount for 26 September 2000 is unreasonably large. Because the predicted flow on 26 September 2000 is much higher than that on the previous day, the correction based on the error in the transformed flow on

HESSD
11, 6035–6063, 2014

Adverse effects of autoregressive updating of streamflow predictions

M. Li et al.

the previous day is greatly amplified by the back-transformation, leading to the over-correction. In contrast, the AR-Raw model works better in this situation because the magnitude of the error update never exceeds the prediction error on the previous day regardless of whether the predicted flow is high or low.

Figure 3 shows that about 15–25 % AR-Norm updated predictions have an error update that is larger than the prediction error on the previous day and therefore are susceptible to over-correction. Figure 4 presents a time-series plot for the Orara catchment and shows the instances susceptible to over-correction of the AR-Norm model by the vertical inward facing tick-marks. These instances all occur when the flow rises.

A converse example is presented in Fig. 5 where the AR-Raw model causes over-correction. Here, the base hydrological model significantly under-estimates the peak on 6 July 1998. The magnitude of the error update given by the AR-Raw model cannot adjust according to the value of the prediction. As a result, the AR-Raw model updates the prediction on 7 July 1998 with a very large amount, resulting in over-estimation. In contrast, the AR-Norm model does a better job in this example, giving a smaller magnitude of error update by recognising that the hydrograph is moving downward. It is generally true that in applying the AR-Raw model, over-correction may occur when the flow is receding. Figure 6 provides more examples of the over-correction caused by the AR-Raw model from a longer time-series plot for the Abercrombie catchment. There are three clear instances of over-correction, all occurring on the time step immediately after large peaks in observed flows.

4.2 Poor stand-alone performance of the base hydrological model

The second issue with conventional AR error models is the stand-alone performance of the base hydrological model (GR4J). As noted above, the parameters of the base hydrological model are those estimated jointly with an AR model. For streamflow forecasting, we expect to obtain a reasonably accurate prediction from the base hydrological model followed by an updating procedure as an auxiliary means to improve the predictive accuracy. At long lead times (e.g., streamflow forecasts generated from mediumrange rainfall forecasts) error updating becomes less effective, and the performance of the base hydrological model is crucial for realistic forecasts. While we investigate only forecasts at a lead time of one timestep in this study, we aim to develop methods that can be applied to forecasts at longer lead times. Further, if the base hydrological model does not replicate important catchment processes realistically, the performance of the hydrological model outside the calibration period may be less robust.

Figure 7 presents the Nash–Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) calculated from the base hydrological model and the error models. When the AR-Norm model is used, the predicted flows from the base hydrological model are very poor for the Orara catchment (NSE < 0). The scatter plot in Fig. 8 shows further detail about the streamflow prediction for the Orara catchment. When the AR-Norm model is used, the base hydrological model greatly over-estimates discharge and the AR-Norm model then attempts to correct this systematic over-estimation. This is also shown in Fig. 4 where the base hydrological model has a strong tendency to over-estimate flows for a range of flow magnitudes. The base hydrological model with the AR-Norm model also performs poorly for the Abercrombie catchment (Fig. 7). In this case, the base hydrological model tends to under-estimate flows (results not shown). For the other three catchments, however, the base hydrological model with the AR-Norm model performs reasonably well.

In general, the AR-Raw base hydrological model performs as well or better than the AR-Norm base hydrological model. The AR-Raw base hydrological model is notably better than the AR-Norm base hydrological model in the Abercrombie and Orara catchments (Fig. 7). This suggests that more robust performance can be expected of base hydrological models when AR models are applied to raw errors.

We note that for both the AR-Raw model and the AR-Norm models, the updated predictions are not always better than predictions generated by the base hydrological models (Fig. 7). For the Abercrombie catchment, the updated AR-Raw predictions are not as good as the predictions generated by the AR-Raw base hydrological model. Similarly, the AR-Norm updated predictions perform worse than predictions from the

AR-Norm base hydrological model in the Tarwin catchment. This points to a tendency to overfit the parameters to the calibration period, resulting in the error model undermining the performance of the base hydrological model under cross-validation. Such a lack of robustness is highly undesirable in forecasting applications, where the hydrological models should be able to operate in conditions that differ from those experienced during calibration.

5 Restricted AR error model

Motivated to overcome the potential for over-correction, we modify the AR-Norm model by restricting the magnitude of its corrections. This is a restricted AR–Norm model, which we call the RAR-Norm model. The RAR-Norm model is defined by

$$Z^{(\mathrm{R})}(Q_t) = \begin{cases} Z^{(\mathrm{R})}(Q_{\mathrm{S},t}) + \varepsilon_t^{(\mathrm{R})} + \\ \rho^{(\mathrm{R})}\left\{Z^{(\mathrm{R})}(Q_{t-1}) - Z^{(\mathrm{R})}(Q_{\mathrm{S},t-1})\right\} & \text{if } |Q_{\mathrm{M},t} - Q_{\mathrm{S},t}| \le |Q_{t-1} - Q_{\mathrm{S},t-1}| \\ Z^{(\mathrm{R})}(Q_{\mathrm{S},t} + Q_{t-1} - Q_{\mathrm{S},t-1}) + \varepsilon_t^{(\mathrm{R})} & \text{otherwise} \end{cases} \tag{8}$$

where the superscript $^{(\mathrm{R})}$ is used to denote the parameters of the RAR-Norm model and $Q_{\mathrm{M},t} = (Z^{(\mathrm{R})})^{-1}\left[Z^{(\mathrm{R})}(Q_{\mathrm{S},t}) + \rho^{(\mathrm{R})}\left\{Z^{(\mathrm{R})}(Q_{t-1}) - Z^{(\mathrm{R})}(Q_{\mathrm{S},t-1})\right\}\right]$ is the updated streamflow prediction median given by the AR-Norm model without restriction. The actual updated streamflow prediction median of the RAR-Norm model is given by

$$Q_{\mathrm{U},t}^{(\mathrm{R})} = \begin{cases} Q_{\mathrm{M},t} & \text{if } |Q_{\mathrm{M},t} - Q_{\mathrm{S},t}| < |Q_{t-1} - Q_{\mathrm{S},t-1}| \\ Q_{\mathrm{S},t} + Q_{t-1} - Q_{\mathrm{S},t-1} & \text{otherwise} \end{cases} \tag{9}$$

The RAR-Norm model parameters may be jointly estimated with the hydrological model parameters using the maximum likelihood method, in the same as for the AR-Norm and AR-Raw models.

HESSD
11, 6035–6063, 2014

The idea behind the RAR-Norm model is simple. We use the AR-Norm model for error updating if the magnitude of the error update is not too large. Otherwise, we adopt a naïve updating scheme, which applies the raw error from the previous time step to correct the current prediction. At any time t, the magnitude of the error update is restricted to a maximum of $|Q_{t-1} - Q_{\mathrm{S},t-1}|$.

Because the RAR-Norm model imposes an upper limit on the size of the correction, it effectively reduces the tendency for over-correction. Figure 2 shows that the RAR-Norm model behaves similarly to the AR-Raw model for correcting the peak on 26 September 2000 and avoids the over-correction made by the AR-Norm model. The RAR-Norm model is also able to adjust the magnitude of the error update according to $Q_{\mathrm{S},t}$ and this is particularly useful when the hydrograph is moving downward. Figure 5 shows that when the hydrograph recedes rapidly, the RAR-Norm model produces updated streamflow similar to the AR-Norm model. In this case, the RAR-Norm model avoids the over-correction by the AR-Raw model on 7 July 1998. Similarly, the RAR-Norm works better than the AR-Raw model to avoid the three instances of over-correction for the Abercrombie catchment (Fig. 6). Overall, the RAR-Norm model takes a conservative position when flow changes rapidly, either rising or falling. When flow changes rapidly, it is difficult to anticipate the magnitude of prediction error. Accordingly the AR models are prone to over-correction in such instances.

Figure 3 provides the proportion of the instances where $|Q_{\mathrm{M},t} - Q_{\mathrm{S},t}| > |Q_{t-1} - Q_{\mathrm{S},t-1}|$. These instances are susceptible to over-correction by the AR-Norm model. The frequency of these instances varies somewhat from catchment to catchment. The RAR-Norm model identifies 15–30 % of the time series as possible instances of problematic updating, and the AR-Norm model identifies a similar number of instances (slightly fewer – they are not identical because the parameters for each model are inferred independently). This is illustrated in the Orara catchment in Fig. 4, which shows that the number of instances where $|Q_{\mathrm{M},t} - Q_{\mathrm{S},t}| > |Q_{t-1} - Q_{\mathrm{S},t-1}|$ is nearly identical for the AR-Norm model and the RAR-Norm model. In other words, the restriction defined in

HESSD
11, 6035–6063, 2014

the RAR-Norm model is largely applied to the instances where the AR-Norm model is susceptible to over-correction.

The RAR-Norm model generally improves the performance of the base hydrological model, in particular compared to the AR-Norm model (Fig. 7). The RAR-Norm base hydrological model performs similarly to, or better than, the base hydrological models of the AR-Norm and AR-Raw model. The improvement over the AR-Norm base hydrological model is especially evident for the Orara (Figs. 4 and 7) and Abercrombie catchments (Figs. 7).

In general, the updated predictions from the RAR-Norm model show similar or better predictive accuracy, as measured by NSE, than both the AR-Raw model and the AR-Norm model (Fig. 7). We note that the Orara catchment is an exception: here the AR-Raw model shows slightly better performance than both the AR-Norm and RAR-Norm models. Conversely, the RAR-Norm model shows notably better performance than both the AR-Norm and AR-Raw models in the Abercrombie catchment. This suggests the RAR-Norm model may work better in intermittently flowing catchments, although further testing is required to establish that this is true for a greater range of catchments.

Importantly, the updated predictions of the RAR-Norm model outperform the base hydrological model predictions in all catchments. This shows that for the RAR-Norm model, both the base hydrological model and the error updating perform robustly under cross-validation. This is not true of the AR-Raw model in the Abercrombie catchment or for the AR-Norm model in the Tarwin catchment. As noted above, robust performance under cross-validation, and consistent interaction between the base hydrological model and the error updating, are critically important for forecasting applications, where models should perform well in conditions that may differ substantially from those experienced during calibration.

HESSD
11, 6035–6063, 2014

6 Discussion and conclusions

For streamflow forecasting, rainfall–runoff models are often augmented with an updating procedure that corrects the prediction using information from recently observed prediction errors. The most popular updating approach uses autoregressive (AR) models that exploit the "memory" in model prediction errors. AR models may be applied to raw errors directly or to normalised errors.

We demonstrate two adverse effects of AR error updating procedures by case studies of four catchments. The first adverse effect is possible over-correction. The updating procedure may correct hydrological predictions too much at some events. The over-correction often happens at the peak or on the rise of a hydrograph for the AR-Norm model and when the hydrograph is receding for the AR-Raw model.

The second adverse effect is poor stand-alone performance of base hydrological models when the parameters of rainfall–runoff and error models are jointly estimated with the AR parameters. We show that poor base hydrological model performance is particularly prevalent in the AR-Norm model. The poor performance appears to occur in catchments with highly skewed streamflow observations (the Abercrombie, an intermittent river, and the Orara, a catchment in a subtropical climate). For example, in the Orara River, the base hydrological model tends to greatly over-estimate streamflows, and then relies on the error updating to correct the over-estimates. This is not desirable in real-time forecasting applications for two major reasons. First, modern streamflow forecasting systems often extend forecast lead-times with rainfall forecast information (e.g. Bennett et al., 2014). Updating becomes less effective at longer lead times, and predictions at longer lead times rely on the performance of the base hydrological model. Second, hydrological models are designed to simulate various components of natural systems, such as baseflow processes or overland flow. In theory, simulating these processes correctly will allow the model to perform well for climate conditions that may substantially differ from those experienced during the parameter estimation period. If the hydrological model parameters do not reflect the natural processes for a given catchment, the hydrological model may be much less robust outside the parameter estimation period.

In addition, our results for the AR-Norm and AR-Raw models indicate that the interaction between the error updating and the base hydrological model may not always be robust under cross-validation. In some catchments, the updated predictions of the AR-Norm and AR-Raw models were actually worse than the predictions generated by their respective base hydrological models. In forecasting applications, both the base hydrological model and the error updating should perform robustly outside the calibration period, as forecasts are always generated with forcing data that are independent of the calibration period.

The adverse effects of AR-Norm error correction discussed in this study are probably generic. In particular, transformations other than the log–sinh transformation may still lead to over-correction at the peak of hydrograph. The proof in Appendix A shows that if a transformation satisfies some conditions (first derivate is positive and second derivate is negative), it will tend to correct more for higher predicted flow and can cause the problem of over-correction. The conditions given by Appendix A are generally true for many other transformations used for data normalization and variance stabilization in hydrological applications, such as logarithm transformation and Box–Cox transformation with the power parameter less than 1.

We use joint parameter inference to calibrate hydrological model and error model parameters, in order to address the true nature of underlying model errors. Inferring parameters of the error model and the base hydrological model independently – i.e., first inferring parameters of the base hydrological model, holding these constant and then inferring the error model parameters – relies on simplified and often invalid error assumptions (it assumes independent, homoscedastic and Gaussian errors), but nonetheless could be a pragmatic alternative to the joint parameter inference to reduce computational demands. The over-correction of ordinary AR models is independent of the parameter inference, whether the error and base hydrological model parameters are inferred jointly or independently.

HESSD
11, 6035–6063, 2014

In order to mitigate the adverse effects of ordinary AR updating procedures, we introduce a new updating procedure called the RAR-Norm model. The RAR-Norm model is essentially a modification of the AR-Norm model. This new model is able to adjust the magnitude of the error update according to the value of the hydrological prediction, which is similar to the AR-Norm model. However, it limits the magnitude of the error update to the prediction error at the previous time step. We show that the new model indeed guards against over-correction and at the same time leads to more robust performance by the base hydrological models. In addition, the performance of the base hydrological model and the error updating are robust under cross validation. Accordingly, we contend that the RAR-Norm model is preferable to both AR-Norm and AR-Raw models for streamflow forecasting applications.

Appendix A:

We will analytically show that the AR-Norm model gives a larger magnitude of the error update for a higher predicted flow.

Firstly, we will show that the first derivate of the log–sinh transform Z defined by Eq. (3) is positive and the second derivate is negative (i.e. $Z'(Q) > 0$ and $Z''(Q) < 0$) for any $b > 0$ and any Q. Following some simple manipulation, we have

$$Z'(Q) = \frac{\cosh(a+bQ)}{\sinh(a+bQ)} > 0 \quad \text{and } Z''(Q) = \frac{-b}{\sinh^2(a+bQ)} < 0 \tag{A1}$$

Using the differentiation of inverse functions, we find the first and second derivates of the inverse transform Z^{-1}

$$[Z^{-1}]'(Q) = \frac{1}{Z'\{Z^{-1}(Q)\}} > 0 \quad \text{and } [Z^{-1}]''(Q) = \frac{-Z''\{Z^{-1}(Q)\}}{[Z'\{Z^{-1}(Q)\}]^3} > 0, \tag{A2}$$

for any $b > 0$ and any Q.

HESSD
11, 6035–6063, 2014

Next, we will derive the difference of magnitudes of the error update between low and high predicted flows. For the sake of notation simplicity, we rewrite $q = Z(Q_{\mathrm{S},t})$ and $u = \rho^{(\mathrm{T})}\{Z(Q_{t-1}) - Z(Q_{\mathrm{S},t-1})\}$ and assume that $u > 0$. Using Eq. (4), the updated streamflow can be written as $Q_{\mathrm{U},t}^{(\mathrm{N})} = Z^{-1}(q+u)$. The magnitude of the error update can be written as

$$\left|Q_{\mathrm{S},t} - Q_{\mathrm{U},t}^{(\mathrm{N})}\right| = \left|Z^{-1}(q+u) - Z^{-1}(q)\right| = \begin{cases} \int_0^u [Z^{-1}]'(x+q)\mathrm{d}x & \text{if } u > 0 \\ \int_u^0 [Z^{-1}]'(x+q)\mathrm{d}x & \text{otherwise.} \end{cases} \tag{A3}$$

Suppose that we have two predicted streamflows $Q_{\mathrm{S},t}^{(1)} \le Q_{\mathrm{S},t}^{(2)}$ and denote the normalised predicted streamflow by $q_1 = Z\left(Q_{\mathrm{S},t}^{(1)}\right)$ and $q_2 = Z\left(Q_{\mathrm{S},t}^{(2)}\right)$ and the updated streamflow by $Q_{\mathrm{U},t}^{(\mathrm{N},1)}$ and $Q_{\mathrm{U},t}^{(\mathrm{N},2)}$. Because Z is an increasing function, we have $q_1 \le q_2$. The difference in the magnitude of the error update between $Q_{\mathrm{S},t}^{(1)}$ and $Q_{\mathrm{S},t}^{(2)}$ can be derived as

$$\left|Q_{\mathrm{S},t}^{(1)} - Q_{\mathrm{U},t}^{(\mathrm{N},1)}\right| - \left|Q_{\mathrm{S},t}^{(2)} - Q_{\mathrm{U},t}^{(\mathrm{N},2)}\right| = \begin{cases} \int_0^u \left\{[Z^{-1}]'(x+q_1) - [Z^{-1}]'(x+q_2)\right\}\mathrm{d}x & \text{if } u > 0 \\ \int_u^0 \left\{[Z^{-1}]'(x+q_1) - [Z^{-1}]'(x+q_2)\right\}\mathrm{d}x & \text{otherwise.} \end{cases} \tag{A4}$$

From Eq. (A2), we have shown that $[Z^{-1}]'$ is a positive increasing function and this ensures that $[Z^{-1}]'(x+q_1) - [Z^{-1}]'(x+q_2) \le 0$. Finally we have

$$\left|Q_{\mathrm{S},t}^{(1)} - Q_{\mathrm{U},t}^{(\mathrm{N},1)}\right| \le \left|Q_{\mathrm{S},t}^{(2)} - Q_{\mathrm{U},t}^{(\mathrm{N},2)}\right|. \tag{A5}$$

Therefore, the error update at larger predicted flows is always larger than error update at lower predicted flows.

HESSD
11, 6035–6063, 2014

Acknowledgements. This work is part of the WIRADA (Water Information Research and Development Alliance) streamflow forecasting project funded under CSIRO Water for a Healthy Country Flagship. We would like to thank Durga Shrestha for valuable suggestions that led to substantial strengthening of the manuscript.

References

Bates, B. C. and Campbell, E. P.: A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall–runoff modeling, Water Resour. Res., 37, 937–947, doi:10.1029/2000wr900363, 2001.

Bennett, J. C., Robertson, D. E., Shrestha, D. L., Wang, Q. J., Enever, D., Hapuarachchi, P., and Tuteja, N. K.: A new system for continuous hydrological ensemble forecasting (SCHEF) to lead times of 9 days, J. Hydrol., submitted, 2013.

Del Giudice, D., Honti, M., Scheidegger, A., Albert, C., Reichert, P., and Rieckermann, J.: Improving uncertainty estimation in urban hydrological modeling by statistically describing bias, Hydrol. Earth Syst. Sci., 17, 4209–4225, doi:10.5194/hess-17-4209-2013, 2013.

Engeland, K., Renard, B., Steinsland, I., and Kolberg, S.: Evaluation of statistical models for forecast errors from the HBV model, J. Hydrol., 384, 142–155, doi:10.1016/j.jhydrol.2010.01.018, 2010.

Evin, G., Kavetski, D., Thyer, M., and Kuczera, G.: Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration, Water Resour. Res., 49, 4518–4524, doi:10.1002/wrcr.20284, 2013.

Jones, D. A., Wang, W., and Fawcett, R.: High-quality spatial climate data-sets for Australia, Aust. Meteorol. Oceanogr. J., 58, 233–248, 2009.

Kavetski, D., Franks, S. W., and Kuczera, G.: Confronting input uncertainty in environmental modelling, in: Calibration of Watershed Models, edited by: Duan, Q., Gupta, H. V., Sorooshian, S., Rousseau, A. N., and Turcotte, R., American Geophysical Union, Washington DC, 49–68, 2003.

Li, M., Wang, Q. J., and Bennett, J.: Accounting for seasonal dependence in hydrological model errors and prediction uncertainty, Water Resour. Res., 49, 5913–5929, doi:10.1002/wrcr.20445, 2013.

Morawietz, M., Xu, C.-Y., and Gottschalk, L.: Reliability of autoregressive error models as post-processors for probabilistic streamflow forecasts, Adv. Geosci., 29, 109–118, doi:10.5194/adgeo-29-109-2011, 2011.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.

Perrin, C., Michel, C., and Andreassian, V.: Improvement of a parsimonious model for streamflow simulation, J. Hydrol., 279, 275–289, doi:10.1016/S0022-1694(03)00225-7, 2003.

Salamon, P. and Feyen, L.: Disentangling uncertainties in distributed hydrological modeling using multiplicative error models and sequential data assimilation, Water Resour. Res., 46, W12501, doi:10.1029/2009wr009022, 2010.

Schaefli, B., Talamba, D. B., and Musy, A.: Quantifying hydrological modeling errors through a mixture of normal distributions, J. Hydrol., 332, 303–315, doi:10.1016/j.jhydrol.2006.07.005, 2007.

Schoups, G. and Vrugt, J. A.: A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors, Water Resour. Res., 46, W10531, doi:10.1029/2009wr008933, 2010.

Seo, D.-J., Herr, H. D., and Schaake, J. C.: A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction, Hydrol. Earth Syst. Sci. Discuss., 3, 1987–2035, doi:10.5194/hessd-3-1987-2006, 2006.

Shrestha, D. L. and Solomatine, D. P.: Data-driven approaches for estimating uncertainty in rainfall–runoff modelling, Int. J. River Basin Manage., 6, 109–122, doi:10.1080/15715124.2008.9635341, 2008.

Thyer, M., Kuczera, G., and Wang, Q. J.: Quantifying parameter uncertainty in stochastic models using the Box–Cox transformation, J. Hydrol., 265, 246–257, doi:10.1016/S0022-1694(02)00113-0, 2002.

Vaze, J., Perraud, J. M., Teng, J., Chiew, F. H. S., Wang, B., and Yang, Z.: Catchment Water Yield Estimation Tools (CWYET), the 34th World Congress of the International Association for Hydro-Environment Research and Engineering: 33rd Hydrology and Water Resources Symposium and 10th Conference on Hydraulics in Water Engineering, Brisbane, 2011.

Wang, Q. J. and Robertson, D. E.: Multisite probabilistic forecasting of seasonal flows for streams with zero value occurrences, Water Resour. Res., 47, W02546, doi:10.1029/2010WR009333, 2011.

HESSD
11, 6035–6063, 2014

Adverse effects of autoregressive updating of streamflow predictions

M. Li et al.

Wang, Q. J., Shrestha, D. L., Robertson, D. E., and Pokhrel, P.: A log–sinh transformation for data normalization and variance stabilization, Water Resour. Res., 48, W05514, doi:10.1029/2011WR010973, 2012.

World Meteorological Organization: Simulated Real-Time Intercomparison of Hydrological Models, World Meteorological Organization, Geneva, Switzerland, 1992.

Xiong, L. H. and O'Connor, K. M.: Comparison of four updating models for real-time river flow forecasting, Hydrolog. Sci. J., 47, 621–639, doi:10.1080/02626660209492964, 2002.

Xu, C. Y.: Statistical analysis of parameters and residuals of a conceptual water balance model – methodology and case study, Water Resour. Manag., 15, 75–92, doi:10.1023/A:1012559608269, 2001.

Table 1. Basic catchment characteristics.

Name	Gauge Site	Area (km^2)	Rainfall ($mm\,yr^{-1}$)	Streamflow ($mm\,yr^{-1}$)	Runoff coefficient	Zero flows
Abercrombie	Abercrombie River at Hadley no. 2	1447	783	63	0.08	14.4 %
Mitta Mitta	Mitta Mitta River at Hinnomunjie	1527	1283	261	0.20	0
Orara	Orara River at Bawden Bridge	1868	1176	243	0.21	0.6 %
Tarwin	Tarwin River at Meeniyan	1066	1042	202	0.19	0

HESSD
11, 6035–6063, 2014

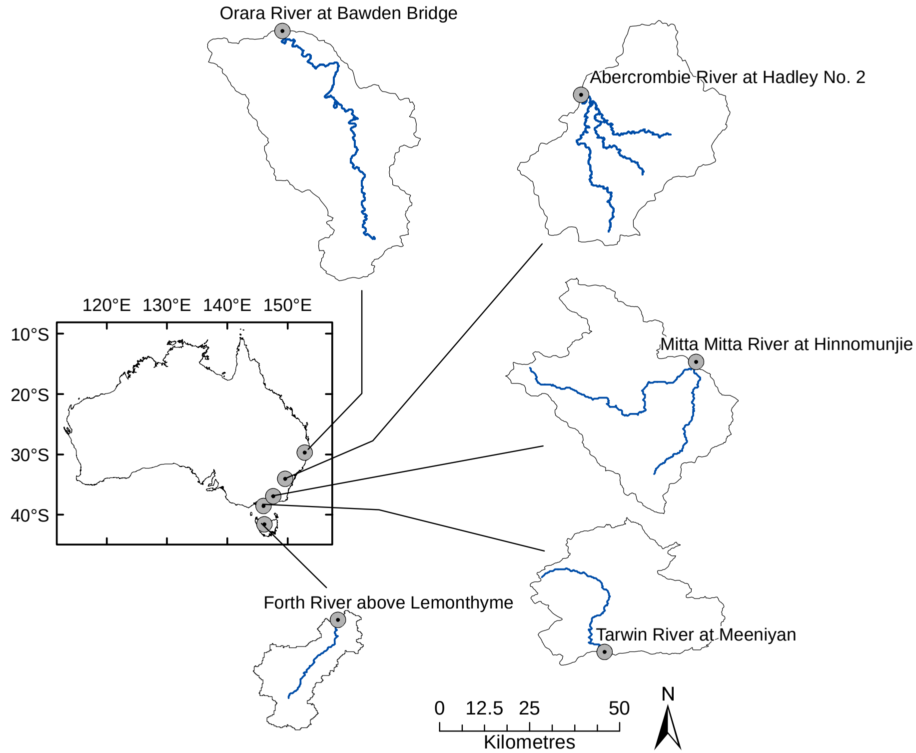


Figure 1. Map of catchments used in this study.

HESSD
11, 6035–6063, 2014

Adverse effects of autoregressive updating of streamflow predictions

M. Li et al.

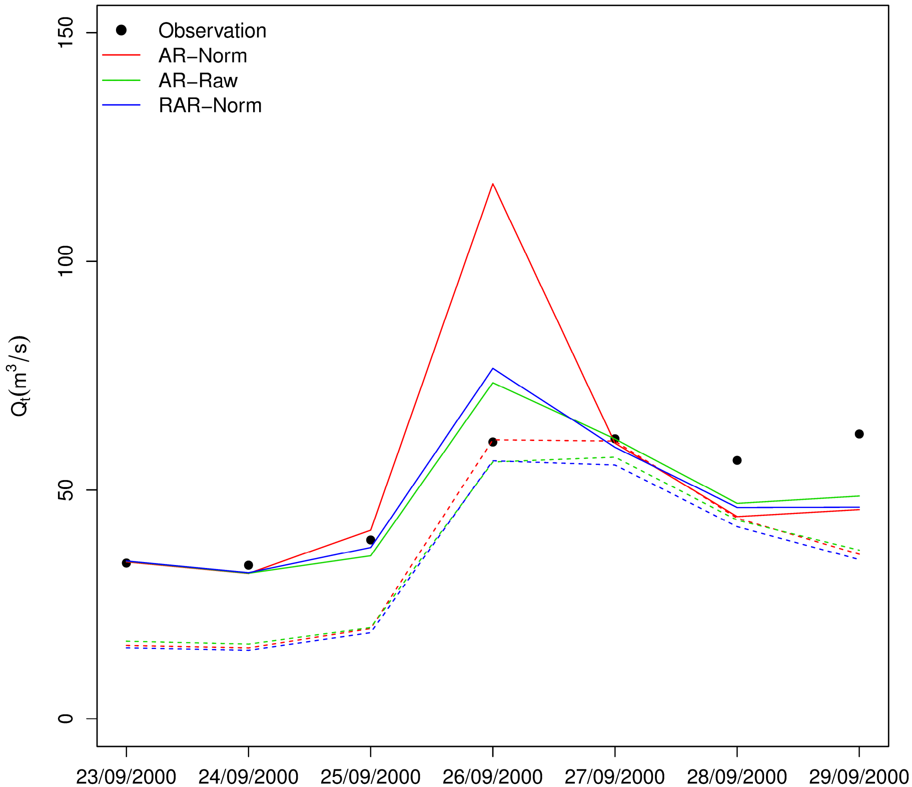


Figure 2. An example of over-correction caused by the AR-Norm model in the Mitta Mitta catchment. Dashed lines: predictions from the base hydrological model (i.e., without error updating). Solid lines: predictions with error updating.

HESSD
11, 6035–6063, 2014

Adverse effects of autoregressive updating of streamflow predictions

M. Li et al.

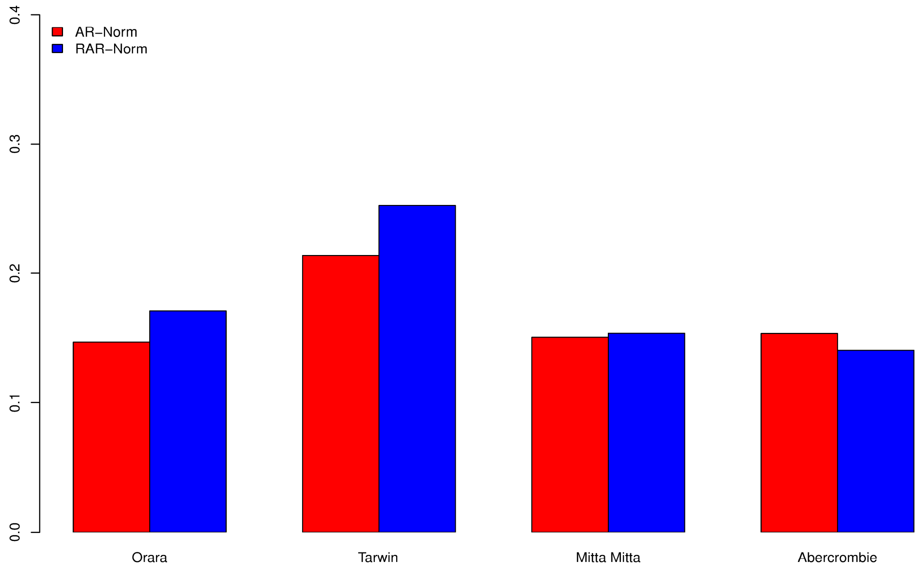
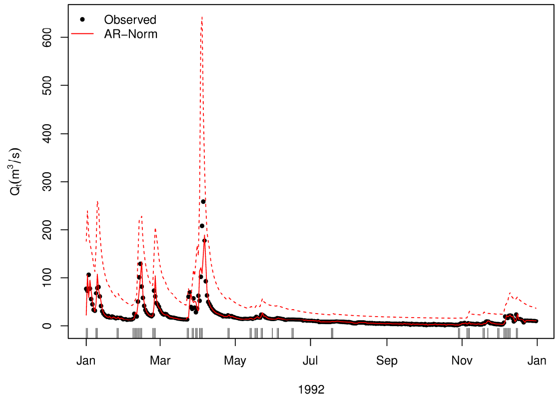


Figure 3. The fraction of instances where $|Q_{M,t} - Q_{S,t}| > |Q_{t-1} - Q_{S,t-1}|$ (i.e., instances where over-correction may occur in the AR-Norm model and where error updating is restricted in the RAR-Norm model) for the AR-Norm and RAR-Norm models.

HESSD
11, 6035–6063, 2014

Adverse effects of autoregressive updating of streamflow predictions

M. Li et al.



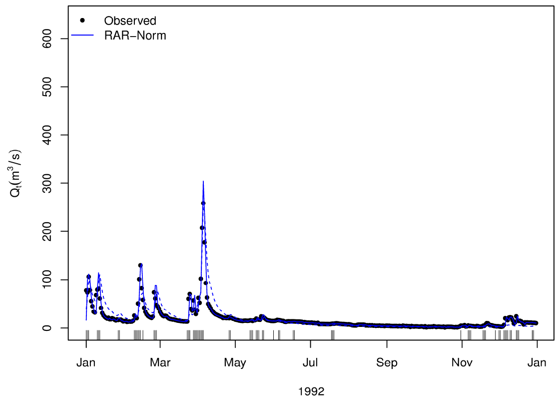


Figure 4. Predicted streamflows for the Orara catchment for an example 1 year period. Top panel shows flows predicted with AR-Norm model, bottom panel shows flows predicted with the RAR-Norm model. Dashed lines: predictions from the base hydrological model (i.e., without error updating). Solid lines: predictions with error updating. Tick marks in the x-axis denote the instance of updating where $|Q_{\mathrm{M},t} - Q_{\mathrm{S},t}| > |Q_{t-1} - Q_{\mathrm{S},t-1}|$.

HESSD
11, 6035–6063, 2014

Adverse effects of autoregressive updating of streamflow predictions

M. Li et al.

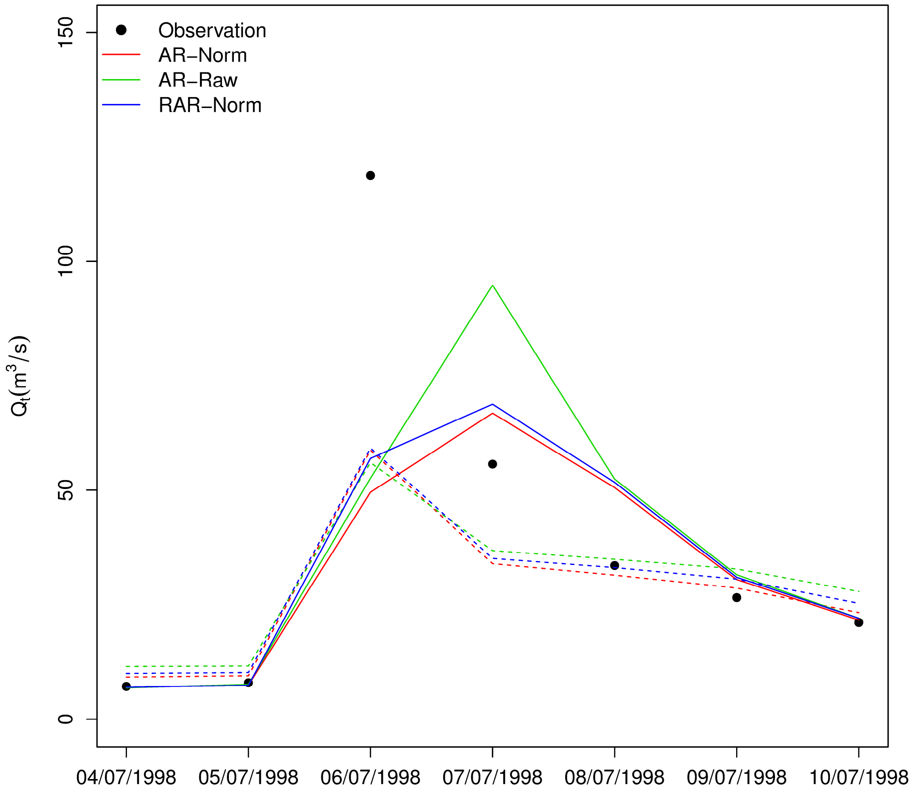


Figure 5. An example of over-correction caused by the AR-Raw model in the Mitta Mitta catchment. Dashed lines: predictions from the base hydrological model (i.e., without error updating). Solid lines: predictions with error updating.

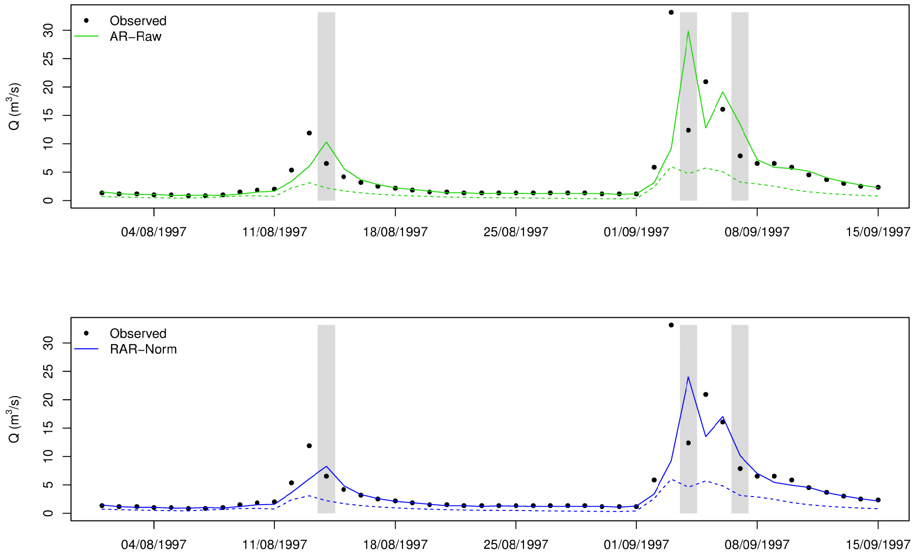


Figure 6. Predicted streamflows for the Abercrombie catchment for the period between 1 August 1997 and 15 September 1997. Top panel shows flows predicted with AR-Raw model, bottom panel shows flows predicted with the RAR-Norm model. Dashed lines: predictions from the base hydrological model (i.e., without error updating). Solid lines: predictions with error updating. Red circles denote the instances of the over-correction caused by the AR-Raw model.

HESSD
11, 6035–6063, 2014

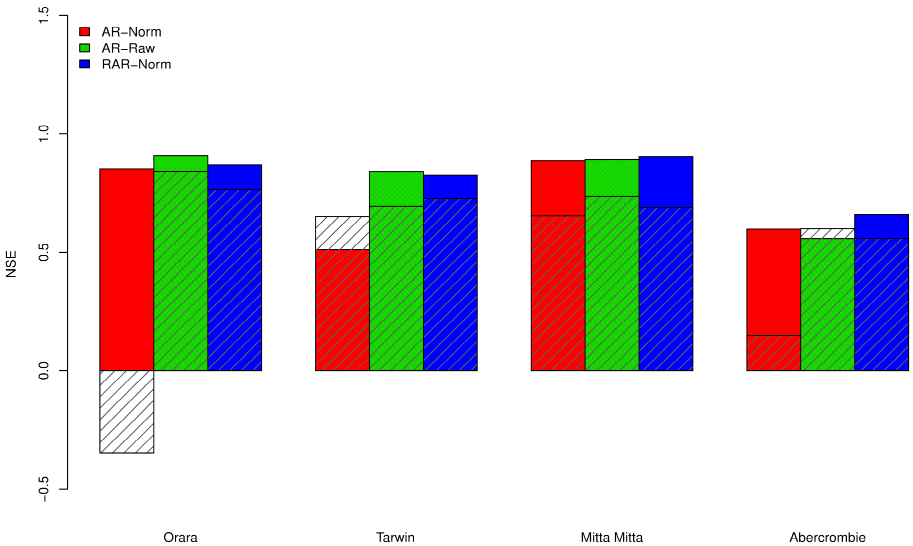


Figure 7. NSE of streamflows predicted with the AR-Norm, AR-Raw and RAR-Norm models (colours). Performance of the corresponding base hydrological models is shown by hatched blocks.

HESSD
11, 6035–6063, 2014

Adverse effects of autoregressive updating of streamflow predictions

M. Li et al.

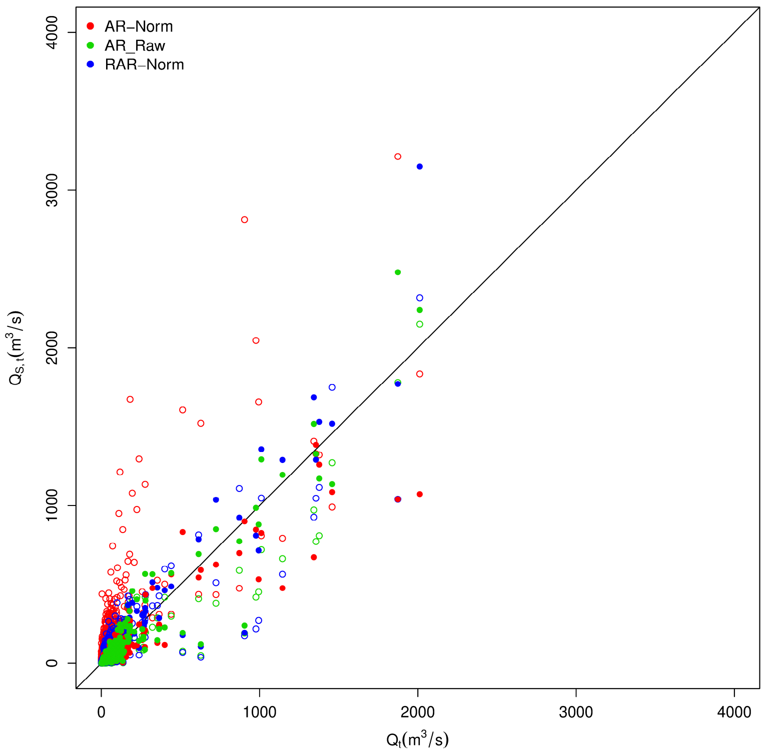


Figure 8. Comparison of the observed streamflows (Q_t) and predicted streamflows (Q_s), as predicted (1) with the base hydrological model (circles), and (2) with the base hydrological model and error updating models (dots) for the Orara catchment.

HESSD
11, 6035–6063, 2014

Adverse effects of autoregressive updating of streamflow predictions

M. Li et al.

Title Page
Abstract
Introduction
Conclusions
References
Tables
Figures
◀I
▶I
◀
▶
Back
Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion