Environmental Modelling & Software 76 (2016) 20–36

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Calibrating hourly rainfall-runoff models with daily forcings for streamflow forecasting applications in meso-scale catchments

James C. Bennett [a, *], David E. Robertson [a], Phillip G.D. Ward [a, b], H.A. Prasantha Hapuarachchi [c], Q.J. Wang [a]

[a] CSIRO Land & Water Flagship, Highett, Victoria, Australia
[b] Monash University, Clayton, Victoria, Australia
[c] Bureau of Meteorology, Melbourne, Victoria, Australia

ARTICLE INFO

Article history:
Received 20 March 2015
Received in revised form
23 September 2015
Accepted 18 November 2015
Available online 1 December 2015

Keywords:
Rainfall-runoff calibration
Hourly data
Rainfall disaggregation
Streamflow forecasting
AWBM
GR4J
PDM
Sacramento

ABSTRACT

The absence of long sub-daily rainfall records can hamper development of continuous streamflow forecasting systems run at sub-daily time steps. We test the hypothesis that simple disaggregation of daily rainfall data to hourly data, combined with hourly streamflow data, can be used to establish efficient hourly rainfall-runoff models. The approach is tested on four rainfall-runoff models and a range of meso-scale catchments (150–3500 km^2). We also compare our disaggregation approach to a method of parameter scaling that attains an hourly parameter-set from daily data.

Simple disaggregation of daily rainfall produces hourly streamflow models that perform almost as well as those developed from hourly rainfall data. Rainfall disaggregation performs at least as well as parameter scaling, and often better. For the catchments and models we test, simple disaggregation is a very straightforward and effective way to establish hydrological models for continuous sub-daily streamflow forecasting systems when sub-daily rainfall data are unavailable.

1. Introduction

Many streamflow forecasting systems operate at sub-daily time steps to give adequate detail of streamflows, for example about the timing and magnitude of floods (e.g. Thielen et al., 2009; Demargne et al., 2014). Operating a sub-daily streamflow forecasting system requires the availability of both sub-daily forcing data (rainfall, potential evaporation) and response data (streamflows) in real-time. In addition, long records of historical data are required for forecasting systems that rely on continuous hydrological modelling, as the hydrological models have to be calibrated to reasonably long historical periods to capture variability in streamflow (e.g., Merz et al., 2009, recommended using a minimum of 5 years' data).

Catchments where sub-daily streamflow records are available without accompanying sub-daily rainfall records are common in Australia and in many other regions, particularly in developing countries. Sub-daily rainfall records require expensive monitoring (e.g. tipping bucket rain gauges), data logging and telemetry systems to adequately cover rainfall variability over a catchment. Conversely, a single streamflow gauge is often sufficient to establish a forecasting system. A sub-daily hydrological forecasting service cannot be easily extended to a catchment that has long streamflow records if there are no existing sub-daily rainfall records, even if there is a strong need for such a service. However, daily rainfall records are available over much wider geographical areas and/or for much longer periods for two reasons: i) daily rainfall can be recorded manually with simple rain gauges, and are accordingly more common than sub-daily records (see, e.g., Nalbantis, 1995); and ii) the increasing availability of daily precipitation data products (e.g. the European Climate Assessment & Dataset project, http://www.ecad.eu/).

A number of studies have shown that calibrating conceptual hydrological models to different time steps can produce markedly different parameter values (e.g., Schaake et al., 1996; Holman-Dodds et al., 1999; Littlewood and Croke, 2008; Wang et al., 2009). This makes it difficult to transfer daily parameters directly to hourly models without a substantial drop in performance.

* Corresponding author. CSIRO Land and Water, PO Box 56, Highett, VIC 3190, Australia.
E-mail address: James.Bennett@csiro.au (J.C. Bennett).

http://dx.doi.org/10.1016/j.envsoft.2015.11.006

Accordingly, few attempts have been made to use daily rainfall to inform the calibration of streamflow models run at an hourly time step in a forecasting context. The most notable of these is the study by Nalbantis (1995), who considered the case where a few hourly rainfall observations for flood events are available to supplement a more comprehensive set of daily rainfall observations. He amalgamated two sets of hydrological model parameters: 1) a parameter set calibrated at a daily time step using all daily rainfall records available and 2) a parameter set calibrated to an hourly time step for events where hourly rainfall records were available. In order to amalgamate the two parameter sets, parameters that changed with time step were standardised with theoretically derived scaling factors so they could be applied at either the daily or hourly time step. The amalgamated model markedly improved event-based and continuous forecasts generated by the hydrological model calibrated only to daily calibration data.

In this study, we consider a slightly different scenario: a catchment with an hourly streamflow record is targeted for the development of a continuous streamflow forecasting system, and rainfall gauges that monitor hourly rainfall are to be installed in the catchment. Before these hourly rain gauges are installed, only daily rainfall data are available. The scenario is not simply an academic exercise: the Bureau of Meteorology in Australia is facing similar cases as it seeks to extend its System for Continuous Hydrological Ensemble Forecasting (SCHEF - Bennett et al., 2014) to a wider range of catchments, a number of which have short sub-daily rainfall records.

We attempt to answer the question: can a simple disaggregation of daily rainfall data be used to establish hydrological models suitable for a continuous sub-daily streamflow forecasting system? To answer this question, we disaggregate daily rainfalls to hourly rainfalls. To offer the sternest test of this hypothesis, we use the simplest method of disaggregation available (simple division), though we note that more complex disaggregation methods exist (e.g., Koutsoyiannis and Onof, 2001; Segond et al., 2006). We test the disaggregation for a variety of catchments (Section 2), and use these with hourly streamflow data to calibrate and validate hydrological models in a series of experiments (Section 3). We assess the disaggregated rainfall on a range of conceptual rainfall-runoff models, and compare the results to alternative methods of deriving parameter sets for models run at an hourly time step (Section 4). We discuss our findings in the context of streamflow forecasting in Section 5, and summarise and conclude our study in Section 6.

Our experiments consider hydrological simulations forced by observed rainfall (i.e. not with forecast rainfall) Forecast rainfall may be biased (Shrestha et al., 2013), and contains errors with respect to observations (Shrestha et al., 2015), meaning rainfall-runoff models are likely to perform worse when forced with forecast rainfall than with observations. However, it is common practise in forecasting to calibrate rainfall-runoff models to observations, meaning our study is directly relevant to forecasting applications.

2. Study sites and data

We select seven catchments on the following bases: i) the availability of high-quality hourly rainfall and streamflow data of at least 5 years' duration, and ii) that catchments cover a diverse range of climatic and hydrological conditions. Catchment locations are shown in Fig. 1, and catchment characteristics are listed in Table 1. We define the catchments in this study as meso-scale (between 100 km^2 and 3500 km^2 in area). Three of the catchments had more than one streamflow record available, and we analyse a total of 17 streamflow records.

We divide each catchment into subareas in order to run semi-distributed hydrological models. Except for the Forth River catchment, the catchment delineations employed in this study are to be used in a deterministic forecasting service provided by the Bureau. Catchment delineations follow recommendations by the Bureau of Meteorology to i) ensure all subareas within a catchment are of a similar size and ii) to simulate flow at significant points-of-interest (e.g. gauge sites, reservoirs, etc.).

Rainfall and streamflow data for all catchments except the Forth River were quality controlled by the Bureau of Meteorology (http://www.bom.gov.au/waterdata/). Hydro Tasmania supplied quality-controlled streamflow data for the Forth River. Hydro Tasmania also supplied rainfall data for the Forth River catchment that were not quality controlled. We quality control rainfall data for the Forth River catchment by comparing rainfall records to nearby gauges and gridded precipitation data from the Australian Water Availability Project (AWAP, http://www.bom.gov.au/jsp/awap), then marking suspect values as null (Robertson et al., 2015).

Area-average hourly rainfalls are compiled for each subarea by applying a simple inverse-distance-squared weighting to gauged rainfalls, where distance is calculated from the gauge to the centroid of the subarea. Rainfall gauges occasionally have missing/null data, and in these cases the inverse-distance squared weighting is re-calculated and applied only to gauges without missing data. In no cases are rainfall data missing from all gauges, and the subarea rainfalls for all catchments are serially complete for the periods listed in Table 1.

Evaporation data are taken from the AWAP dataset, which uses the Priestley and Taylor (1972) method to calculate potential evaporation from observed solar radiation.

3. Methods

3.1. Rainfall disaggregation

Our first step is to aggregate the hourly subarea rainfall time series to daily data, through simple summing. This establishes daily rainfall records at each subarea (where each day commences at 0:00 AM local time), and we consider these equivalent to rainfall records that have been assembled from rain gauges that measure only daily rainfall. Daily subarea rainfall is then disaggregated to hourly rainfall through simple division: that is, we divide daily rainfall totals by 24. We refer to these as *disaggregated* rainfall, and parameter sets derived from these rainfalls as *disaggregated parameter sets*. We choose this very simple disaggregation method in the first instance to test the hypothesis that even the simplest interpolation of hourly rainfalls from daily data can provide a useful basis for calibrating hourly hydrological models.

3.2. Experimental setup

We run three calibration experiments:

1) *Disaggregated* – Models are calibrated with the disaggregated rainfall, described above, and hourly streamflow data. This is the new method we are presenting and assessing in this paper.
2) *Parameter scaling* – parameters are identified using daily streamflow and daily rainfalls, but scaled to attempt to account for changes in the characteristics of daily and hourly streamflow. This method has a long basis in theory and practice (see, e.g., Kavetski et al., 2011; Nalbantis, 1995). The methods of scaling are described in Section 3.5.
3) *Control* – observed hourly rainfall and runoff model data are used to identify model parameters. This method would usually be employed if long records of hourly rainfall are available, and

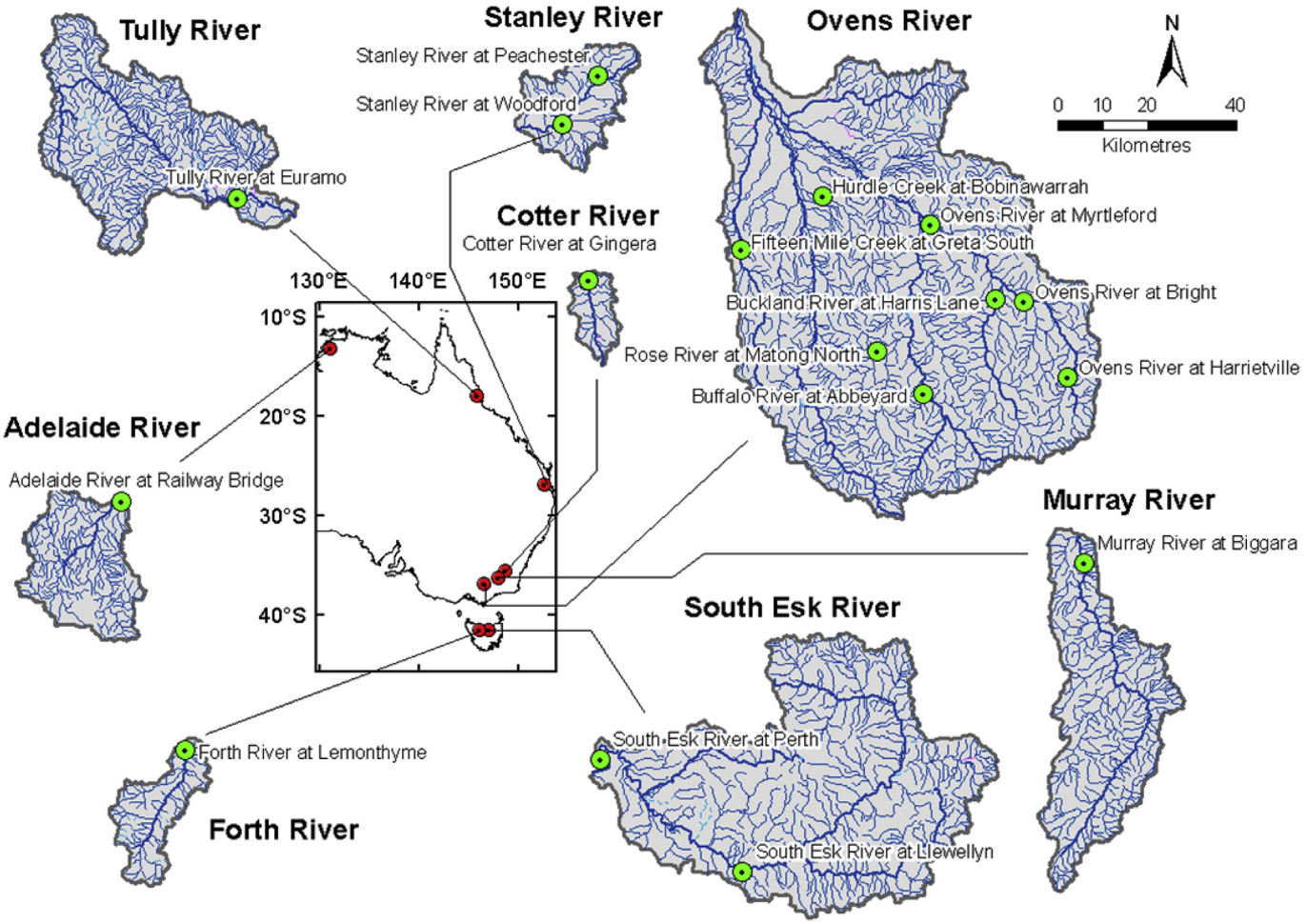

Fig. 1. River basins and streamflow gauge sites used in this study. All catchments are to the same scale.

because it makes use of both high resolution rainfall and streamflow data we expect it to produce the best performance.

3.3. Rainfall-runoff models

We test the disaggregation method in four widely used rainfall runoff models: GR4J (Génie Rural à 4 paramétres Journalier, Perrin et al., 2003), AWBM (the Australian Water Balance Model, Boughton, 2004), PDM (the Probability Distributed Model, Moore, 2007) and the Sacramento soil moisture accounting model (Burnash et al., 1973). All are simple conceptual rainfall runoff models that use various conceptual soil moisture stores and routing algorithms to simulate discharge. The basic characteristics of each model are summarised in Table 2; a conceptual diagram and description of parameters for each model is given in Appendix A.

Note that for each experiment the models are validated at the hourly time step, and to do this we configure each model to run at an hourly time step. For Sacramento and PDM, this required no changes to the configurations described in the source literature; for GR4J and AWBM (both originally designed to run at a daily time step) we apply the parameter scaling described in Section 3.5 to run them at an hourly time step. We do this so that parameter values are in the units described in the source literature, and it has no other effect on the results of our study.

3.4. Channel routing model

Hydrological models are implemented in semi-distributed form: catchments are divided into a number of subareas, and runoff is generated at each subarea. Streamflows are routed from upstream to downstream subareas with the linear Muskingum channel routing algorithm first described by Gill (1978). Storage in a reach is calculated by

$$S = k_m[x_m I + (1 - x_m)O] \tag{1}$$

and

$$k_m = \alpha L, \tag{2}$$

where I is inflow to the reach, O is outflow, L is the length of the reach, and the parameters are x_m (a translation parameter) and α (lag coefficient).

Channel routing in semi-distributed models is designed to simulate lag in streamflows due to storage of water in the river channel, a process that is distinct from the conversion of rainfall to runoff. This should free the rainfall-runoff model to concentrate on simulating the process of runoff generation.

3.5. Parameter scaling

Nalbantis (1995) provided a succinct mathematical description

Table 1
Characteristics of study catchments.

River basin	Climate	No. rain gauges	Gauge site	Drainage area (km^2)	Data available	Missing flow records (%)	Mean discharge (m^3/s (mm))	Annual rainfall (mm)	Annual potential evaporation (mm)
Adelaide	Tropical/ monsoonal	4	Adelaide River at Railway Bridge	638	01-Jan-2000 to 04-Dec-2014	0	12.5 (617)	1584	2246
Cotter	Temperate	35	Cotter River at Gingera	145	01-Jan-1990 to 31-Dec-2011	3.6	1.2 (257)	917	1177
Forth	Temperate/ montane	9	Forth River above Lemonthyme	310	01-Jan-1990 to 01-Feb-2011	0	14.3 (1458)	1777	888
Murray	Temperate	39	Murray River at Biggara	1260	31-Dec-1996 to 31-Dec-2011	2.1	11.6 (291)	886	1205
Ovens	Temperate	33	Buckland River at Harris Lane	472	01-Jan-1992 to 01-Dec-2009	0	5.1 (341)	1341	1219
			Buffalo River at Abbeyard	415	01-Jan-1992 to 01-Dec-2009	0	4.1 (311)	1302	1234
			Fifteen Mile Creek at Greta South	238	01-Jan-1992 to 01-Dec-2009	0	1.4 (189)	1126	1263
			Hurdle Creek at Bobinawarrah	165	01-Jan-1992 to 01-Dec-2009	0	0.7 (137)	942	1303
			Ovens River at Harrietville	124	01-Jan-1992 to 01-Dec-2009	0	2.3 (589)	1426	1144
			Ovens River at Bright	497	01-Jan-1992 to 01-Dec-2009	0	5.5 (346)	1318	1209
			Ovens River at Myrtleford	1232	01-Jan-1992 to 01-Dec-2009	0	13.5 (345)	1298	1230
			Rose River at Matong Rose	177	01-Jan-1992 to 01-Dec-2009	0	1.6 (282)	1134	1249
South Esk	Temperate	48	South Esk River at Llewellyn	2284	01-Jul-1996 to 01-Feb-2011	0	14.5 (200)	927	1015
			South Esk River at Perth	3279	01-Jul-1996 to 01-Feb-2011	0	17.7 (171)	854	1021
Stanley	Subtropical	15	Stanley River at Peachester	105	31-Dec-2000 to 31-Dec-2011	0.6	1.9 (578)	1328	1571
			Stanley River at Woodford	246	31-Dec-2000 to 31-Dec-2011	4.4	3.4 (438)	1175	1571
Tully	Tropical/ monsoonal	22	Tully River at Euramo	1475	01-Sep-2001 to 30-Apr-2007	0	73.3 (1569)	2810	972

Table 2
Characteristics of rainfall-runoff models used in this study.

Model	Description	No. of free Parameters	Reference
AWBM	Australian water balance model with catchment routing performed with a cascade of 2 linear reservoirs	9	AWBM structure and soil moisture accounting: Boughton (2004); cascade of 2 linear reservoirs: Moore (2007)
GR4J	Parsimonious model with two unit hydrographs and a groundwater exchange function	4	Perrin et al. (2003)
PDM	Probability distributed model	8	Moore (2007)
Sacramento	Detailed soil moisture accounting model	16	Burnash et al. (1973)

of the scaling of parameters that depend directly on time step in rainfall-runoff models. In conceptual models, water exchange between storages, $Q_{t+\Delta t}$, over time interval $[t,\ t+\Delta t]$ is sometimes described by

$$Q_{t+\Delta t} = C_{\Delta t} f(S_t) \tag{3}$$

where $C_{\Delta t}$ is a parameter that describes the maximum possible water exchange and $f(S_t)$ is a function of the water contents of the model reservoirs, S_t. Nalbantis (1995) showed that if converting from a longer time step, e.g. $\Delta t_1 = 24$ h, to a shorter time step, e.g. $\Delta t_2 = 1$ h, $C_{\Delta t}$ should be scaled by:

$$C_{\Delta t_2} = \left(\frac{\Delta t_2}{\Delta t_1}\right) C_{\Delta t_1}, \tag{4}$$

Output from simple linear reservoirs of the form

$$Q_{t+\Delta t} = K_{\Delta t} S_t, \tag{5}$$

where S_t is the quantity of water in the reservoir and $K_{\Delta t}$ is a constant that governs depletion, is commonly used in conceptual rainfall-runoff models to describe discharge from conceptual stores. To change time steps with these reservoirs, Nalbantis (1995) showed $K_{\Delta t}$ should be scaled by:

$$K_{\Delta t_2} = 1 - (1 - K_{\Delta t_1})^{\frac{\Delta t_2}{\Delta t_1}}. \tag{6}$$

Note that these scaling factors are only required in instances where fluxes have been calculated over the time step by approximate solution of the integral by the rectangle method (for more discussion of numerical implementations of water fluxes in conceptual models, see Kavetski et al., 2011). In instances where fluxes have been calculated over the time step with an exact analytical solution (as occurs, e.g., in the cascade of linear reservoirs in PDM and AWBM) parameter scaling is much simpler, and parameters

need only be scaled to keep their units consistent, as with Equation (4).

We use equations (4) and (6) to scale parameters and constants in all four rainfall-runoff models and the channel routing model (see Appendix B).

3.6. Model calibration

Rainfall-runoff and routing parameters are jointly calibrated by maximising the average of four objectives:

$$F_{obj} = \frac{F_{NSE} + F_{\ln(NSE)} + F_{Corr} + F_{Bias}}{4} \quad (7)$$

F_{NSE} is the well-known Nash–Sutcliffe efficiency (NSE), $F_{\ln(NSE)}$ is the NSE of log flows, F_{Corr} is correlation and F_{Bias} is a symmetrical measure of bias first described by Wang et al. (2011). All objectives range from 1 (perfect) to -∞. The objectives reflect the multiple purposes for which forecasts are often used: for flood warning and high flows (F_{NSE} and F_{Corr}), for management of environmental requirements during periods of low flow ($F_{\ln(NSE)}$) and for apportioning water for irrigation (F_{Bias}).

Model parameters are optimised with the Shuffled Complex Evolution algorithm (Duan et al., 1994), which attempts to find a globally optimal parameter set.

For catchments containing more than one gauge we estimate parameters for each gauge independently. That is, we simply ignore the existence of upstream gauge sites when calibrating to a particular gauge site. This simplifies our experiments, so that a given performance metric at each gauge site relates only to one parameter set.

3.7. Model validation

3.7.1. Cross-validation

All parameter sets are validated by being forced with observed hourly rainfall data, using split-sample cross-validation. The procedure is as follows:

1. The first year of data is used to warm up hydrological model states
2. The hydrological model is calibrated on the first half of the data available (less the first year)
3. Hydrological model states are warmed up with the first year of data available and data used for calibration at Step 2
4. The hydrological model is validated on the second half of the data available (less the first year)

The procedure is then reversed: calibration is performed on the second half of the data (less the first year) and validated on the first half (less the first year). Thus for a given performance metric we have two data points for each streamflow gauge. Except for the objective function values described in Section 5, all results presented in this study are cross-validated.

The cross-validation procedure means that in the Tully catchment calibration is carried out on little more than two years of data - a fairly small sample size. The effects of the small sample size on the results of the Tully are unknown, but are unlikely to be pronounced because of relatively low variability of rainfall in this catchment (it is the wettest region in Australia, and receives very consistent rainfall). For the remaining catchments, models are calibrated to a minimum of 5 years' data.

3.7.2. Performance measures

We test model performance with Pearson's correlation, relative bias, and the Nash-Sutcliffe efficiency (NSE). These are well-established measures to describe residual error and how well the models preserve patterns in observations (Bennett et al., 2013).

In forecasting applications, it is often particularly desirable for large streamflow events to be simulated well. We test this with a recently developed diagnostic called Series Distance (Ehret and Zehe, 2011). Bennett et al., 2013 classified Series Distance as belonging to the family of performance scores that measure how well data patterns are preserved by models. Series Distance emulates the process of visual inspection of event hydrographs by trained hydrologists, and gives a summary of timing and magnitude errors for simulations of larger flow events. For a given event, each data point in the simulated and observed hydrographs is categorised as being on a rising limb, a falling limb, at a peak or in a trough. The hydrographs are then smoothed with a moving average, and each point in the observed hydrograph is matched to a corresponding point in the simulated hydrograph. The vertical distance (flow magnitude) and horizontal distance (timing error) for each point can then be calculated. Vertical and horizontal distances can then be aggregated to mean absolute vertical distance, q ($m^3\ s^{-1}$), and mean absolute timing distance, t (hours).

Series Distance tends to be quite sensitive to the method of event selection. Here we apply a simple method: we define events as observed flows exceeding a threshold, defined at each gauge as the observed flow with a 10% exceedance probability. q and t can only be calculated for events where the simulated flows exceed the threshold within the duration of the observed event. An additional requirement, then, is to check how often observed events are not registered by simulations, and vice versa. We do this with the Critical Success Index (CSI; sometimes called the Threat Score). CSI is defined as

$$CSI = \frac{a}{a+b+c}, \quad (8)$$

where a is the number of hits (an event is observed and an event is simulated), b is the number of misses (an event is observed but not simulated) and c is the number of false alarms (an event is simulated but not observed). CSI ranges from 0 (poorest performance) to 1 (perfect). Because CSI takes into account both false alarms and missed events, it provides a balanced measure of a model's ability to simulate an event.

4. Results

Fig. 2 summarises overall performance characteristics under cross-validation with box-plots of correlation, relative bias, and NSE derived from all 17 gauge sites. As we expect, the control simulations are least biased (i.e., tend to be closer to zero) and have the highest (i.e., best)correlation and NSE scores of all models. The control simulations from all four models show reasonably robust performance: GR4J tends to attain slightly higher NSE values than other models, however all models achieve a median NSE of >0.7.

Somewhat surprisingly, the disaggregated parameter sets perform almost as well as the control parameter sets. Correlations for disaggregated simulations are very similar to the control simulations for all models. NSE and bias of disaggregated simulations are also similar to those of the control simulations, albeit slightly poorer for some models: NSE of the Sacramento model shows the greatest (though still modest) drop in performance from the control simulations, while GR4J calibrated with disaggregated rainfalls tends to be somewhat more positively biased than the control simulations. In those instances where performance deteriorates, there are no clear patterns associated with either catchment area or climate (not shown).

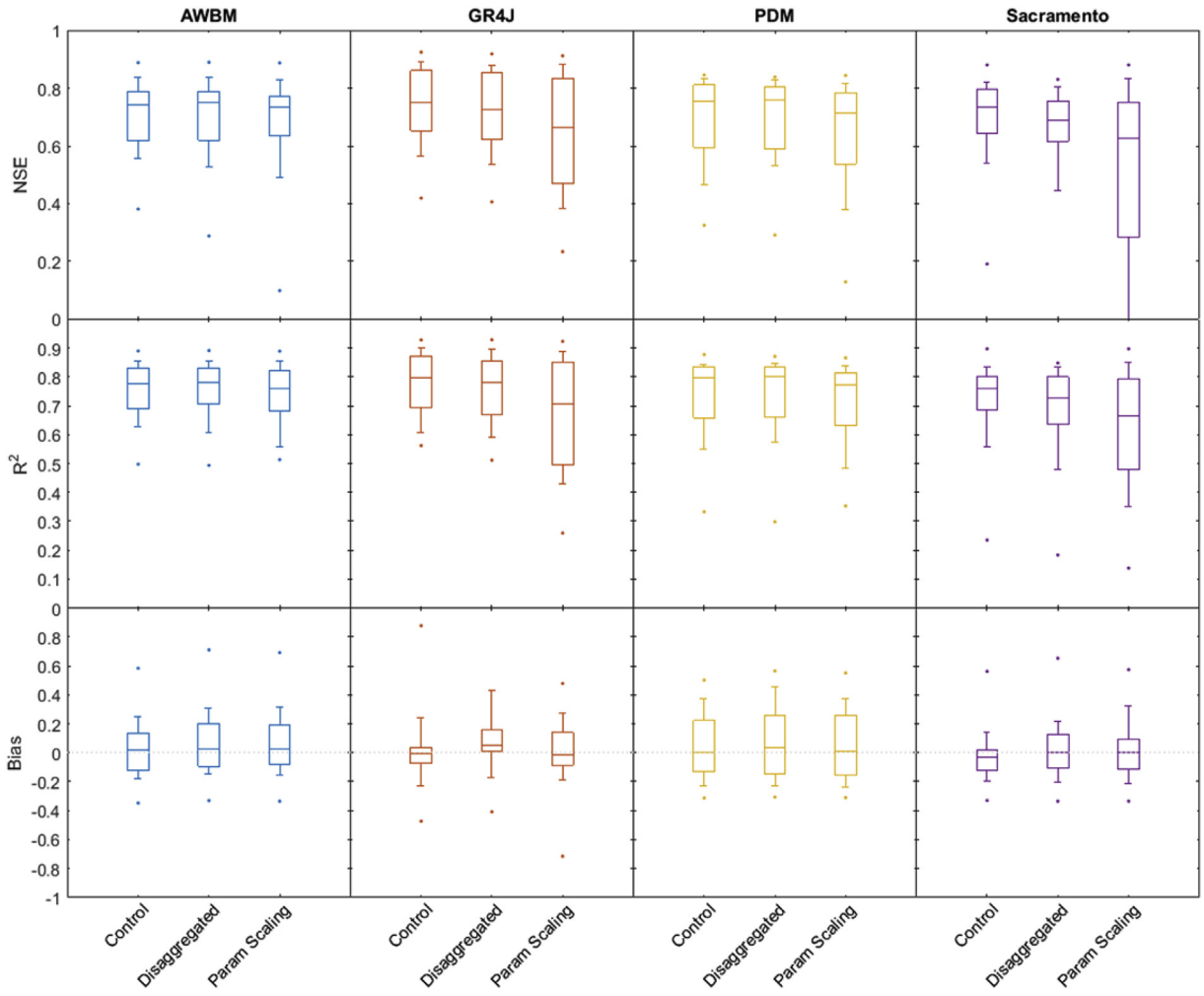

Fig. 2. Box plots of performance metrics for cross-validated hourly simulations with different calibration experiments calculated from all sites. Boxes give mean and interquartile ranges, whiskers show [0.1, 0.9] intervals, points show extrema. "Control" refers to models calibrated to hourly rainfall and hourly streamflow data, "Disaggregated" refers to models calibrated to disaggregated rainfall and hourly streamflow data, "Param Scaling" refers to models calibrated to daily streamflow and daily rainfall data with parameter scaling (see text for details).

The performance of the parameter sets derived from parameter scaling is generally poorer than the disaggregated parameter sets, though this depends on the hydrological model. Parameter scaling worked least well for the GR4J and (especially) the Sacramento models. For GR4J, this may be because parameter scaling factors are partly empirical, as we discuss when describing GR4J parameters, below. For Sacramento, we hypothesise that the relatively poor performance of parameter scaling is due to overfitting of the model due to the large number of parameters, which we discuss further in Section 5.

In forecasting, higher temporal resolution is often particularly desirable for larger flow events where end users may require detailed information about how rapidly, and when, a hydrograph is likely to rise or fall. To illustrate the effects of the different calibration methods we show events from two catchments with distinct hydrological characteristics: the temperate South Esk What is noteworthy in all cases is how closely the simulations generated with the disaggregated daily parameter sets agree with those from the control simulations. In contrast, the parameter sets generated from daily rainfall datasets, with and without parameter scaling, tended to differ markedly from the control simulation for most hydrological models. (AWBM is again an exception, showing that it is highly stable irrespective of how parameter sets are attained.).

We can show that this is generally true for a range of larger events with a recently developed diagnostic called Series Distance (Ehret and Zehe, 2011).

River (Fig. 3) and the tropical Tully River (Fig. 4). The South Esk River tends to rise more slowly than the Tully River, due in part to its larger catchment area and in part to the characteristics of rain storms in this temperate region, which tend to be less intense than the tropical rain storms in the Tully River. As might be expected, the different hydrological models vary in their ability to simulate these particular events. What is noteworthy in all cases is how closely the disaggregated simulations agree with the control simulations. Figs. 3 and 4 illustrate that simulations generated with parameter scaling also tended to agree reasonably well with the control simulations for the AWBM, GR4J and PDM models, but usually not to the same degree as the disaggregated simulations. As already noted, parameter scaling did not work as well for the Sacramento model, shown by the discrepancy between hydrographs generated by parameter scaling and the control for this model. (Note that for Sacramento, parameter scaling offered the best fit to observations for the examples in both in both Figs. 3 and 4. It is possible for parameter scaling to outperform the control for isolated events, but this not commonplace as shown by Fig. 2.)

For larger events, there is generally strong agreement between

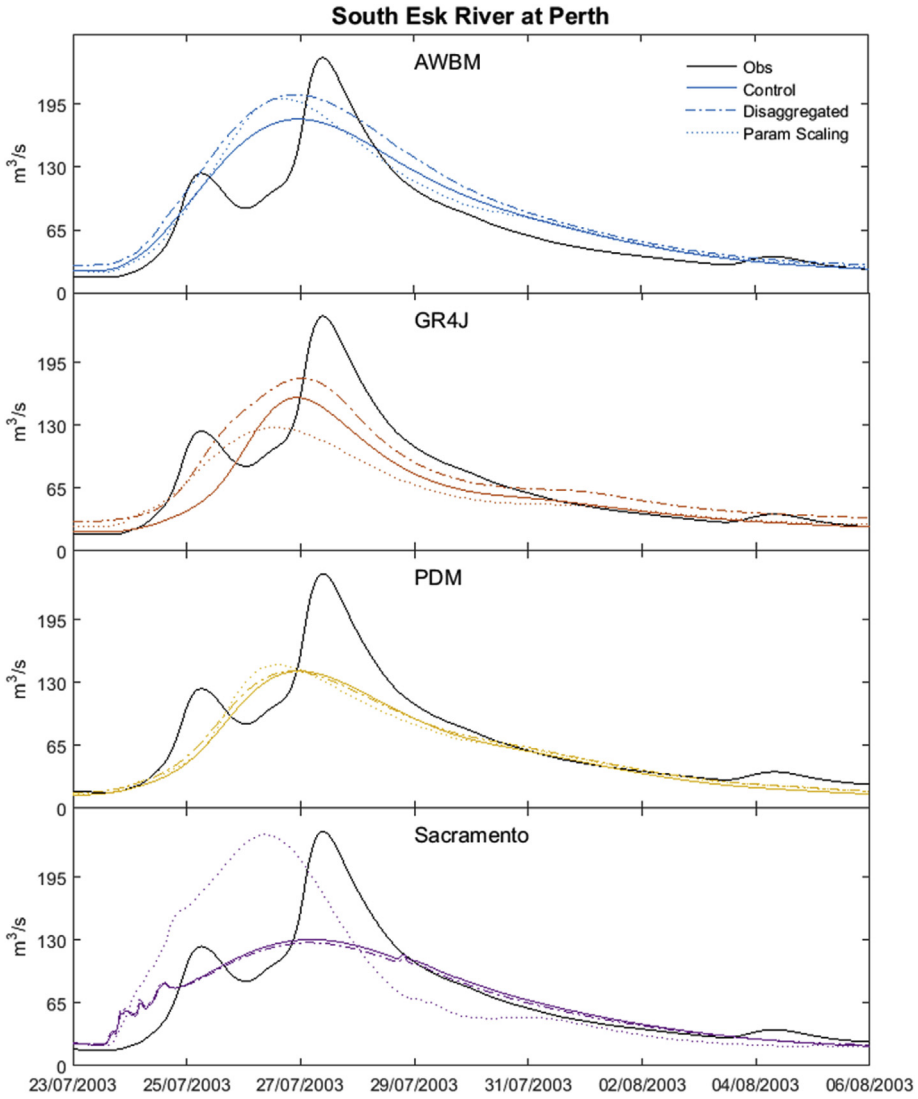

Fig. 3. Hydrographs for an example event in the temperate South Esk River at Perth.

the disaggregated simulations and the control simulations. Fig. 5 shows Series Distance scores for parameter sets derived from disaggregated rainfalls in comparison to the control simulations. Magnitude (q) and timing (t) errors for disaggregated parameter sets are similar to errors in the control simulations for all models. There is a slight decline in CSI in the disaggregated parameter sets in comparison to the control, but overall Series Distance shows that the disaggregated simulations perform very similarly to the control simulations.

Fig. 6 compare Series Distances for parameter sets derived from disaggregated rainfalls and parameter scaling for the example of the GR4J model. Mean absolute magnitude errors (q) for parameters sets derived from daily rainfalls with parameter scaling are similar to those of the disaggregated parameter sets. However, the parameters derived from parameter scaling tend to have larger timing errors (t) and are less able to discriminate events (CSI) than simulations made with the disaggregated parameter sets.

The similarities between disaggregated parameter sets and the control parameter sets are shown for GR4J in Fig. 7 (see Appendix C for the same analysis of the other models). The parameters controlling the production store (X1), groundwater exchange (X2) and reservoir routing (X3) in the disaggregated parameter sets are quite similar to those in the control simulation. In most cases the control parameter values are more similar to the disaggregated parameter values than to the parameter values calibrated with parameter scaling. There are some discrepancies between control and disaggregated values of the X4 parameter that controls the time-base of the two unit hydrographs (X4). Even here, however, the control parameters tend to be much closer to the disaggregated parameter sets than to the parameter scaling method. The scaling of the unit hydrograph exponent was calculated from empirical data (from French catchments) by Mathevet (2005), rather than theoretical scaling factors (see Appendix B). It is possible that this empirical scaling does not function well for the Australian catchments, resulting in the discrepancies in X4 parameter values between the parameter scaling method and the control. Whatever the case, calibrating with disaggregated rainfalls avoids the problem of estimating such empirical scaling factors in the first place.

Note that the other rainfall-runoff models have parameters that are more difficult to identify (in particular Sacramento – Appendix C). For these models neither the disaggregated parameters nor the scaled parameters will necessarily agree with

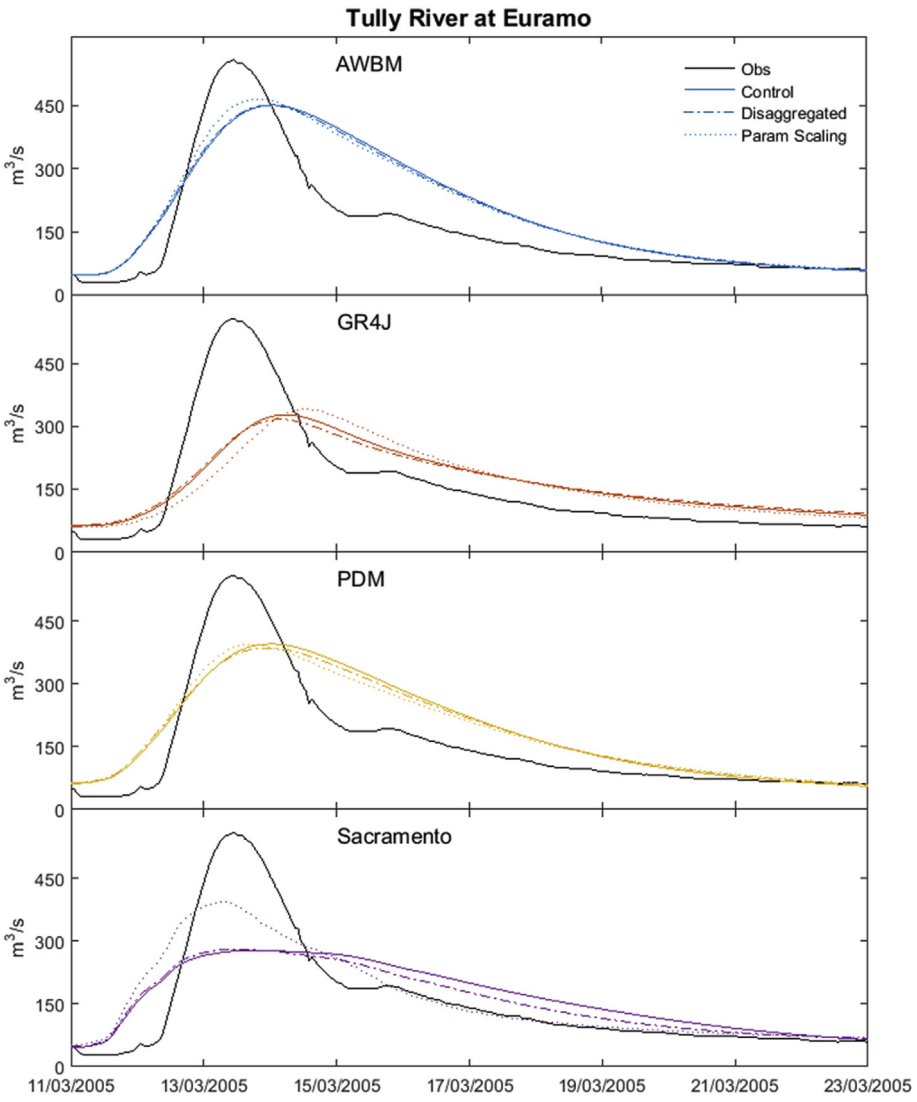

Fig. 4. Hydrographs for an example event in the tropical Tully River at Euramo.

the control parameter values, due to equifinality. We discuss this problem further in Section 5, below.

5. Discussion and further analyses

The effectiveness of the disaggregation method can be explained in two ways. First, the variance of rainfall inputs to rainfall-runoff models is often damped inside the model to replicate the processes of water transport through soil moisture stores. This makes modelled streamflow less sensitive to temporal variations in rainfall than might be at first supposed. Second, the disaggregated rainfall allows models to be optimised at the (higher) temporal resolution at which they will be used. This means that if the rainfall runoff models are not highly sensitive to temporal variations in rainfall, the optimisation function will tend to force the control models and the disaggregated rainfall models to behave alike.

We can show this by comparing model errors used to construct the objective function. The objective function can be converted to an error score by

$$ES = 1 - F_{obj}, \tag{9}$$

where F_{obj} is the objective function defined in Equation (3). The error score for a given set of parameters, θ, but different rainfall forcings, can be related by

$$ES(\theta, p^*) = ES(\theta, p) + \Delta ES, \tag{10}$$

where p is hourly rainfall forcing, p^* is disaggregated rainfall data, and

$$\Delta ES = ES(\theta, p^*) - ES(\theta, p), \tag{11}$$

If ΔES is small, then Equation (11) becomes

$$ES(\theta, p^*) \approx ES(\theta, p). \tag{12}$$

We compare $ES(\theta, p)$ and ΔES in Table 3. In all instances $ES(\theta, p)$ is larger than ΔES, usually by an order of magnitude or more,

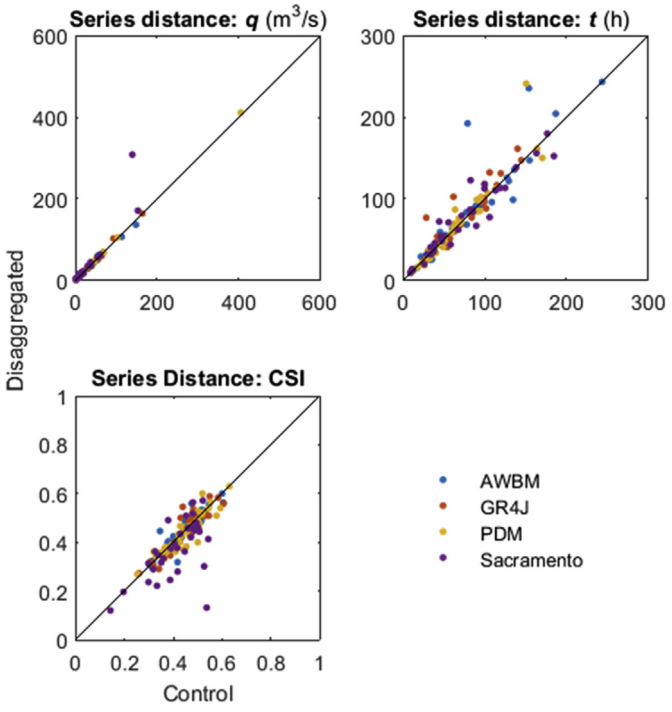

Fig. 5. Series distance scores for disaggregated models versus the control. Left hand column shows mean absolute errors in flow, middle column plots show mean absolute errors in timing, right hand column shows critical success index (CSI).

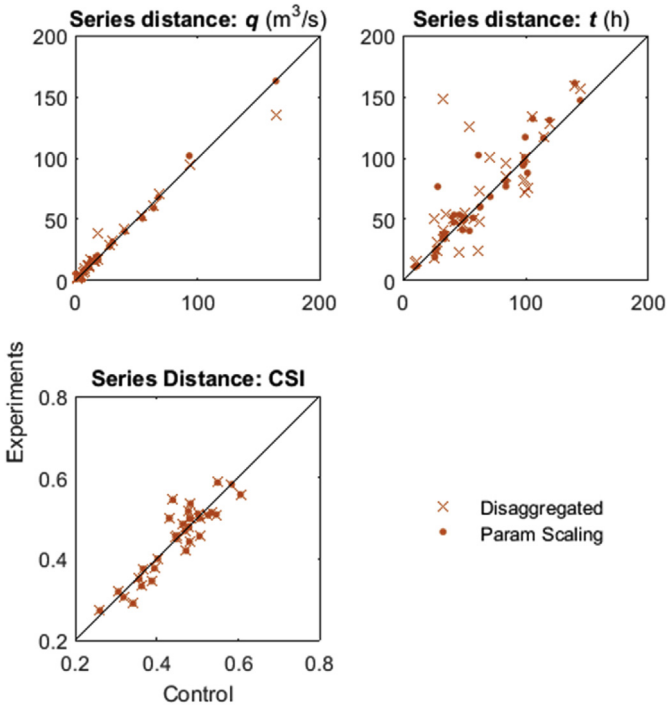

Fig. 6. Series distance scores for GR4J. Left hand column shows mean absolute errors in flow, middle column plots show mean absolute errors in timing, right hand column shows critical success index (CSI).

showing the approximation in Equation (12) generally holds. Therefore, minimising $ES(\theta, p*)$ is mostly about minimising $ES(\theta, p)$. This explains why we can use disaggregated rainfall to calibrate a hydrological model to good effect.

Even if, as we have shown, we can achieve very similar objectives with p and p^*, it does not necessarily follow that each parameter derived from p will be similar to its counterpart derived from p^*, due to the well-known problem of equifinality. Equifinality is prevalent in models.

where when one parameter replicates the function of one or more other parameters. This tends to occur in models with large numbers of parameters, e.g. Sacramento. For models with readily identifiable parameters, like GR4J, individual parameters derived from p will tend to be similar to those derived from p^*, as we show in Fig. 7. Conversely, for models where parameters are difficult to identify, a small change in the value of the objective function can result in quite different parameter values (Appendix C). However, this is not an inherent weakness of the use of the disaggregated rainfall calibration approach we describe here. Rather, it is a general weakness of models with a high degree of interference between parameters, and tends to make them less stable under changed forcings.

We have used an extremely simple method to disaggregate daily rainfall, but of course more complex methods exist (see, e.g., Koutsoyiannis and Onof, 2001; among many others). It is doubtful, however, if a more complex disaggregation method would have yielded substantially better results for the catchments in our study. We attempted a more complex disaggregation method that accounted for the diurnal cycle in rainfall (omitted from this paper for brevity), under the expectation that this would improve model performance in regions with a prominent diurnal rainfall pattern (e.g. the subtropical Stanley catchment). This more complex disaggregation method yielded no improvement in hydrological models under cross-validation over the simple method presented here. Again, this can be explained by the tendency of all hydrological models to smooth the temporal variability of rainfall.

We strongly caution against extrapolating the findings we present in this study to catchments that are either much smaller or more prone to rapid streamflow responses than those presented here. In small catchments – particularly those where streamflows rise and fall in only a few hours in response to rainfall – it is quite possible that our method will not work because information from sub-daily rainfall is likely to play an important role in the identification of model parameters. Given the results we have presented for larger catchments, however, we speculate that even in very small or flashy catchments our method would work at least as well as parameter scaling (even though neither method may work well).

Finally, we note that our results are not unprecedented in one sense: other studies have shown that rainfall-runoff models can cope remarkably well with deterioration in the quality of rainfall information in space (Andréassian et al., 2001; Vaze et al., 2011). We contend that rainfall-runoff models can cope with some deterioration in the temporal quality of rainfall as well, at least in the meso-scale catchments we have tested here.

6. Summary and conclusions

We have shown that parameter sets for hourly hydrological simulations can be derived from a simple disaggregation of daily rainfalls combined with hourly streamflow data. This simple disaggregation method produces parameter sets that perform nearly as well as parameter sets derived from hourly rainfall and hourly streamflow data (the control). This holds true for a range of runoff models and a range of meso-scale catchments. Larger events simulated using disaggregated parameter sets are similarly

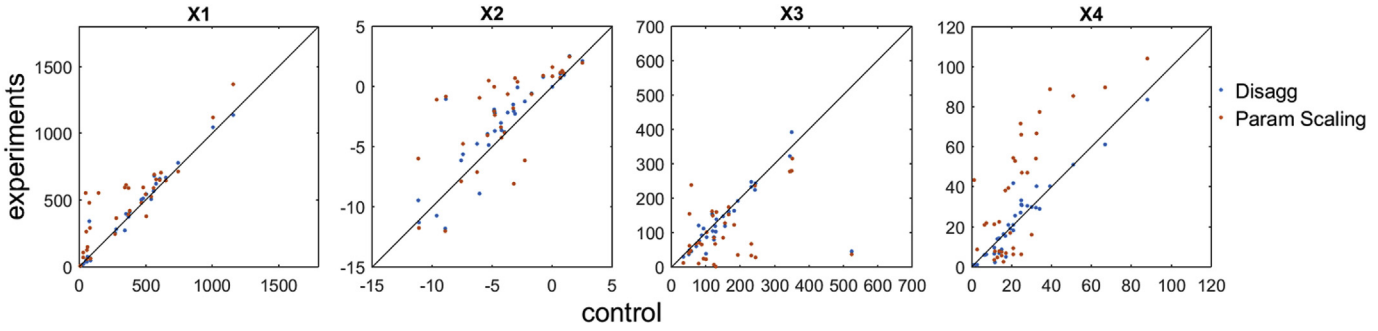

Fig. 7. GR4J parameters derived from different calibration experiments and compared to the control.

Table 3
Comparison of terms from Equation (6) for the first period of the split-sample cross-validation.

Gauge	AWBM		GR4J		PDM		Sacramento	
	$ES(\theta, p)$	ΔES	$ES(\theta, p)$	ΔES	$ES(\theta, p)$	ΔES	$ES(\theta, p)$	ΔES
Adelaide River at Railway Bridge	0.192	0.001	0.092	−0.019	0.207	0.001	0.121	−0.006
Cotter River at Gingera	0.151	−0.001	0.152	0.004	0.095	0.008	0.103	0.027
Forth River above Lemonthyme	0.159	−0.032	0.14	−0.022	0.154	−0.028	0.141	−0.015
Murray River at Biggara	0.147	0.018	0.142	0.02	0.126	0.022	0.138	0.018
Buckland River at Harris Lane	0.093	−0.002	0.077	0	0.055	0.004	0.072	0.005
Buffalo River at Abbeyard	0.104	−0.003	0.103	0	0.08	0	0.104	−0.001
Fifteen Mile Creek at Greta South	0.165	−0.01	0.158	−0.005	0.157	−0.006	0.125	0.003
Hurdle Creek at Bobinawarrah	0.173	−0.008	0.201	−0.004	0.171	−0.003	0.206	0
Ovens River at Harrietville	0.117	0	0.082	0.002	0.09	0.005	0.065	0.004
Ovens River at Bright	0.095	0	0.06	0.001	0.05	0.005	0.05	0.009
Ovens River at Myrtleford	0.103	0.001	0.063	−0.001	0.08	0.007	0.078	0.007
Rose River at Matong Rose	0.106	−0.006	0.095	−0.002	0.086	−0.001	0.065	0.005
South Esk River at Llewellyn	0.177	−0.001	0.257	0	0.175	−0.008	0.204	−0.008
South Esk River at Perth	0.169	0.001	0.254	0.005	0.169	0.007	0.225	−0.007
Stanley River at Peachester	0.244	−0.064	0.185	−0.026	0.305	−0.038	0.246	−0.019
Stanley River at Woodford	0.195	−0.002	0.155	0	0.247	0.016	0.209	0.001
Tully River at Euramo	0.061	−0.003	0.114	−0.005	0.042	0.002	0.069	0.002

accurate to the control. The strength of the results for disaggregated parameter sets is surprising, and we argue that this method is a simple solution to the problem of sub-daily rainfall data scarcity when sub-daily streamflow data are available.

The disaggregated parameter sets tend to perform slightly better than parameter sets derived from daily rainfall and streamflow data with parameter scaling. This can be partly explained by the simple fact that we are giving the models more information – i.e., hourly streamflow data – when we calibrate with disaggregated rainfalls. As we have shown, this causes the disaggregated rainfalls to have a similar influence on the calibration objective function as hourly rainfalls. Nonetheless, we are unaware of the use of such a simple disaggregation scheme elsewhere in the literature. In addition, parameter scaling factors need to be derived theoretically (Nalbantis, 1995) and sometimes honed through empirical experimentation (e.g. GR4J as used in this study). We contend that our disaggregation method is much simpler and easier to implement, and can be used with existing 'black-box' rainfall-runoff modelling software (e.g. eWater Source – www.ewater.org.au) that does not allow easy manipulation of rainfall-runoff model algorithms.

The success of the disaggregation means that it could be applied to establish sub-daily operational streamflow forecast systems more quickly than might previously have been possible in meso-scale catchments. The method is wholly reliant on the existence of sub-daily streamflow data, as we have already described. Where only daily streamflow records exist, using parameter scaling methods to transfer parameters from a daily streamflow model to an hourly streamflow model can also address the problem. In our experience, however, streamflow data are often available at sub-daily time steps, making our method broadly useful.

While we have explored this approach only for hydrological models, it may well prove useful for other environmental models where the response variable is available at higher temporal resolution than forcing variables.

7 Acknowledgements

This research has been supported by the Water Information Research and Development Alliance (WIRADA) between the Bureau of Meteorology and CSIRO Land & Water Flagship. Jayson Peterson (Entura) and Greg Carson (Hydro Tasmania) supplied data for the Forth River. David Enever and Chris Leahy (both Bureau of

Meteorology) assisted in supplying data for the remaining catchments. Jean-Michel Perraud (CSIRO) configured models for the Murray catchment. We thank David Post (CSIRO) for comments and corrections to the manuscript. Catchment maps in Fig. 1 are generated from the Bureau of Meteorology Geofabric (http://www.bom.gov.au/water/geofabric/). We are grateful to four anonymous referees, whose constructive and thorough reviews helped improve the paper substantially.

Appendices

9.1 Appendix A

Table A.1List of rainfall-runoff models and parameters

Model	Parameters
AWBM	**Loss model**
	A1, A2 (−) − Partial areas of the three soil moisture stores
	C1, C2, C3 (mm) − Capacities of the three soil moisture stores
	Surface routing
	k_1, k_2 (hour) − time constants of cascade of linear reservoirs
	Baseflow
	BFI(−) − baseflow index
	Kb (−) − baseflow recession constant
GR4J/GR4J	X_1 (mm) − capacity of the soil moisture store
	X_2 (mm) − groundwater exchange coefficient
	X_3 (mm) − capacity of the routing store
	X_4 (days) − time base of the unit hydrograph
PDM	**Probability distributed store**
	c_{min} (mm) − minimum store capacity
	c_{max} (mm) − maximum store capacity
	b (−) − exponent of Pareto distribution
	Evaporation Function
	b_e (−) − exponent in actual evaporation function. Fixed to $b_e = 1$, following Moore (2007)
	Recharge function
	k_g (hour mm (bg-1)) − groundwater recharge time constant
	b_g (−) − exponent of recharge function. Fixed to $b_g = 1$, following Moore (2007)
	S_t (mm) − soil tension storage capacity
	Surface routing
	k_1, k_2 (h) − time constants of cascade of linear reservoirs
	Groundwater storage routing
	k_b (h.mm^{m-1}) − baseflow time constant
	m (−) − exponent of baseflow non-linear storage. Fixed to $m = 3$, following Moore (2007)
Sacramento	**Soil moisture stores**
	UZTWM (mm) − upper zone tension water capacity
	UZFWM (mm) − upper zone free water capacity
	LZTWM (mm) − lower zone tension water capacity
	LZFSM (m) − lower zone supplemental free water capacity
	LZFPM (mm) − Lower zone primary free water capacity
	Lateral outflows
	UZK (d^{-1}) − Fractional daily upper zone free water withdrawal rate
	LZSK (d^{-1}) − Fractional daily supplemental withdrawal rate
	LZPK (d^{-1}) − Fractional daily primary withdrawal rate
	RSERV (−) − Fraction of lower zone free water not transferable to lower zone tension water
	Percolation
	P_{free} (−) − Fraction of percolated water going directly to lower zone free water storage
	R_{exp} (−) − Percolation equation exponent
	Z_{perc} (−) − Maximum percolation rate coefficient
	Direct Runoff
	SIDE (−) − Ratio of deep recharge to channel baseflow
	PCTIM − Minimum impervious area (unitless)
	ADIMP − Additional impervious area (unitless)
	Losses
	RIVA − Riparian vegetation area (unitless)

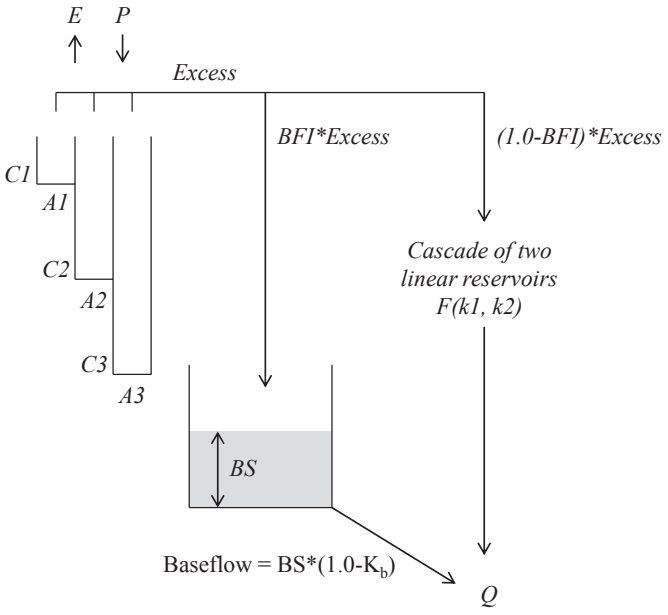

Figure A.1 AWBM model structure (adapted from Boughton, 2004).

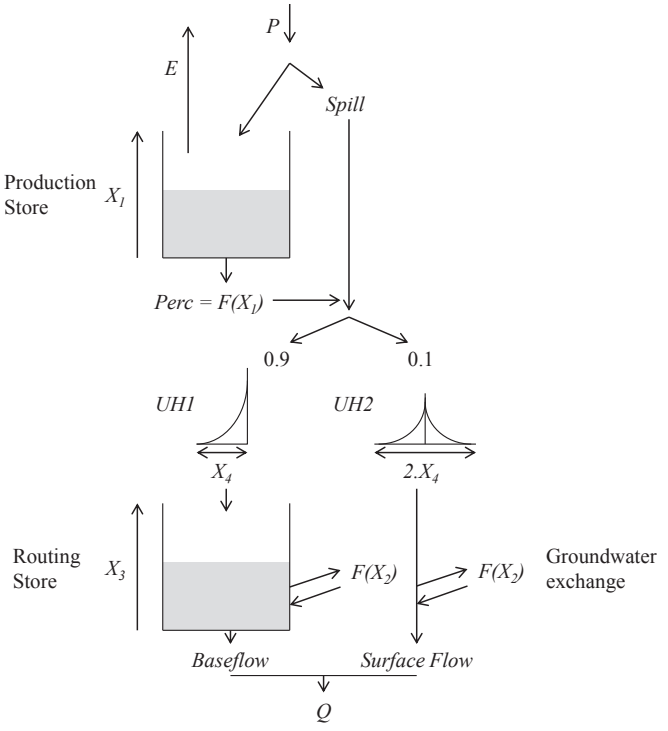

Figure A.2 GR4J model structure (adapted from Perrin et al., 2003).

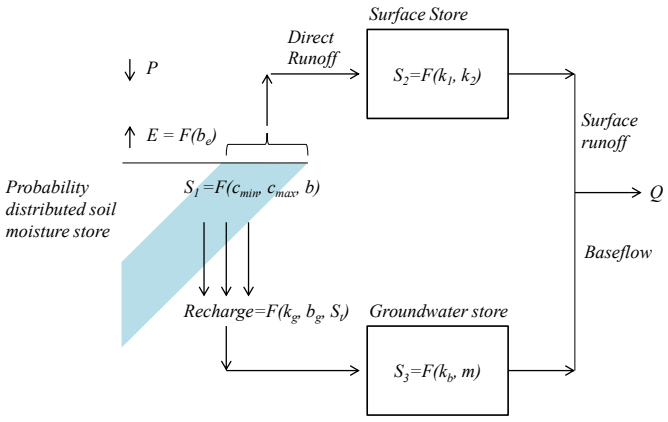

Figure A.3 PDM model structure (adapted from Moore, 2007).

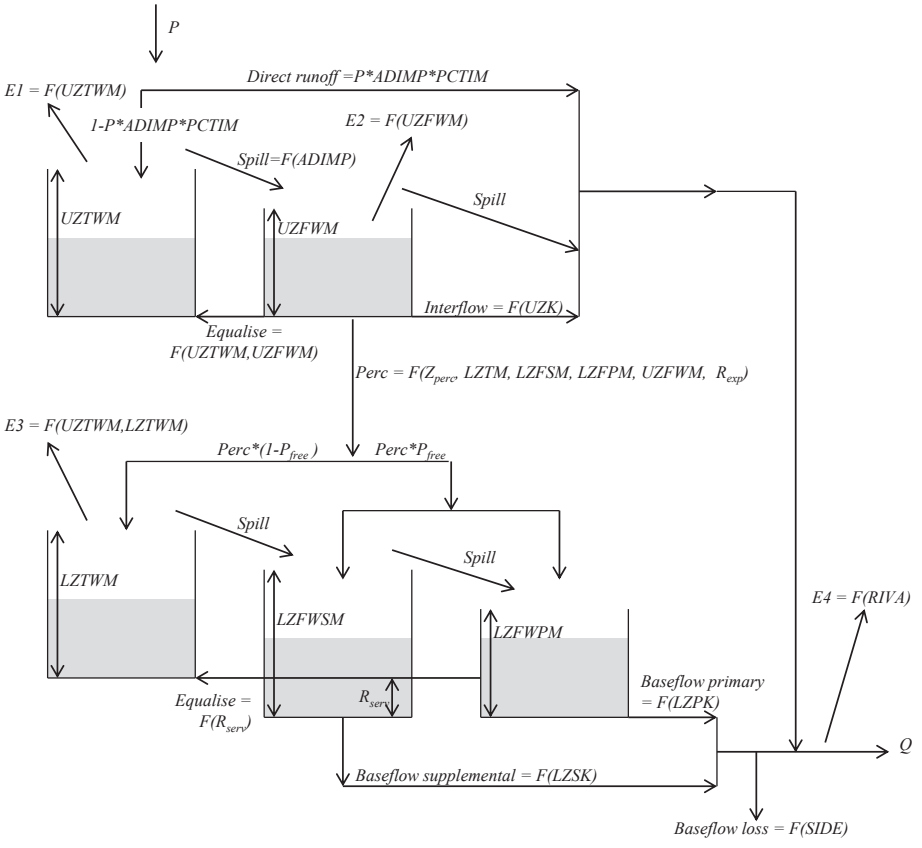

Figure A.4 Sacramento model structure.

Appendix B

Parameter scaling

AWBM. In AWBM discharge from the groundwater store, B, is regulated by the recession constant, K_b, by

$$B = G(1 - K_b), \tag{B1}$$

where G is the level of the groundwater store. When we implement AWBM calibrated to a daily time step, $\Delta t_1 = 24$ h, at an hourly time step $\Delta t_2 = 1$ h, we scale K_b according to Equation (6):

$$K_{b!_{\Delta t_2}} = K_{b!_{\Delta t_1}} - 4pt^{\frac{\Delta t_2}{\Delta t_1}} \tag{B2}$$

No other parameters require scaling. The parameters for the cascade of two linear reservoirs do depend on time step, but Moore (2007) provided an exact analytical solution over the time step so that the scaling of these parameters is implicit.

GR4J. To convert daily parameters from GR4J to be suitable for hourly modelling, a number of parameter and constants need to be scaled (following Mathevet, 2005). GR4J features two reservoirs whose outflows follow the general non-linear relation

$$Q = k.S^{\alpha}, \tag{B3}$$

where S is the level of the reservoir, and k and α are constants. When we integrate the change of storage over the time step, outflow from this reservoir can be expressed as

$$\int_0^{\Delta t} Q.dt = S\left[1 - \frac{1}{\left[1 + \left(\frac{S}{K}\right)^{\alpha-1}\right]^{\frac{1}{\alpha-1}}}\right] \tag{B4}$$

where

$$K = [k(\alpha - 1)\Delta t]^{\frac{1}{1-\alpha}} \tag{B5}$$

Equation B4 is the general form of outflow from the two GR4J reservoirs. Because Equation B4 is integrated over the time step and has an exact analytical solution, k can simply be scaled according to Equation (4), meaning for a change in time step we have

$$K_{\Delta t_2} = K_{\Delta t_1}\left[\frac{\Delta t_2}{\Delta t_1}\right]^{\frac{1}{1-\alpha}}. \tag{B6}$$

For GR4J, $\alpha = 5$, and for the production store

$$K_{\Delta t_1} = \frac{9}{4}x_1, \tag{B7}$$

where x_1 is the size of the production store (the $\frac{9}{4}$ is empirically derived.) From Equation B6, to run GR4J at the hourly time step we have

$$K_{\Delta t_2} = \left(\frac{1}{24}\right)^{-\frac{1}{4}}\frac{9}{4}x_1 \simeq 5x_1. \tag{B8}$$

For the routing store:

$$K_{\Delta t_1} = x_3, \tag{B9}$$

meaning for GR4J we have

$$K_{\Delta t_2} = \left(\frac{1}{24}\right)^{-\frac{1}{4}}x_3 \simeq 2.21x_3. \tag{B10}$$

GR4J allows losses/gains to/from ground water from the routing store and from the quick flow pathway using:

$$F = x_2\left(\frac{R}{x_3}\right)^{\frac{7}{2}} \tag{B11}$$

where R is the level of the routing store, and x_2 and x_3 are GR4J parameters. Equation B11 is of the form described by Equation (3), meaning it can simply be scaled following Equation (4), which yields:

$$F = \frac{\Delta t_2}{\Delta t_1}x_2\left(\frac{R}{x_3}\right)^{\frac{7}{2}} = 0.0417x_2\left(\frac{R}{x_3}\right)^{\frac{7}{2}}. \tag{B12}$$

Finally, GR4J applies an exponent to the S-curves of its two unit hydrographs equal to $\frac{5}{2}$. An empirical investigation by Mathevet (2005) found that an exponent of $\frac{5}{4}$ functioned better at the hourly time step, and accordingly for GR4J we use $\frac{5}{4}$.

PDM. Moore (2007) provided an exact analytical solutions for almost all of the fluxes in the PDM model, meaning that the algorithms are implicitly scaled for time step. The one exception was the recharge to groundwater, described by:

$$d_i = k_g^{-1}(S(t) - S_t), \tag{B13}$$

where $S(t)$ is the level of the surface store that is being drained, the constant S_t is the soil tension water capacity (a parameter), the constant k_g is the groundwater recharge time constant (a parameter). Equation B13 follows the general form of Equation (5), meaning that k_g can be scaled by:

$$\frac{1}{k_{g_{\Delta t_1}}} = 1 - \left(1 - \frac{1}{k_{g_{\Delta t_2}}}\right)^{\frac{\Delta t_1}{\Delta t_2}}. \tag{B14}$$

We scale $k_{g_{\Delta t_2}}$ when calibrating to daily data (rather than scaling $k_{g_{\Delta t_1}}$ when running at an hourly time step), because units of PDM parameters are given in hours.

Sacramento. Three parameters are scaled in Sacramento, *UZK*, *LZSK* and *LZPK*, using Equation (6), as described in Nalbantis (1995).

Muskingum Channel Routing.

The Muskingum parameter k_m is implicitly scaled according to equation Equation (4).

Appendix C

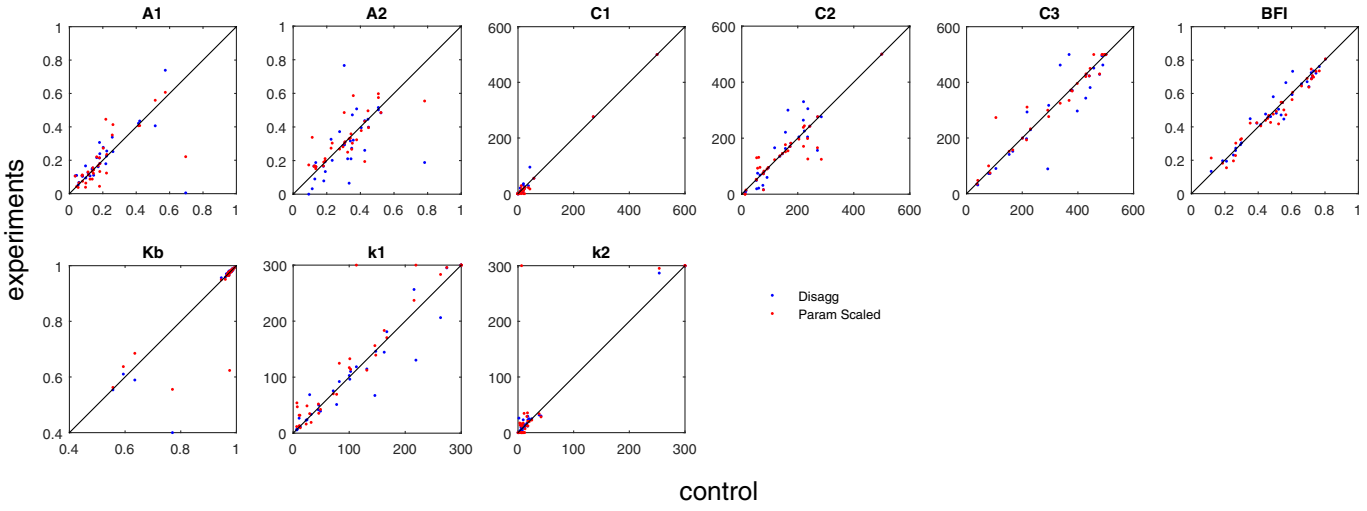

Figure C.1 AWBM parameters derived from different calibration experiments and compared to the control.

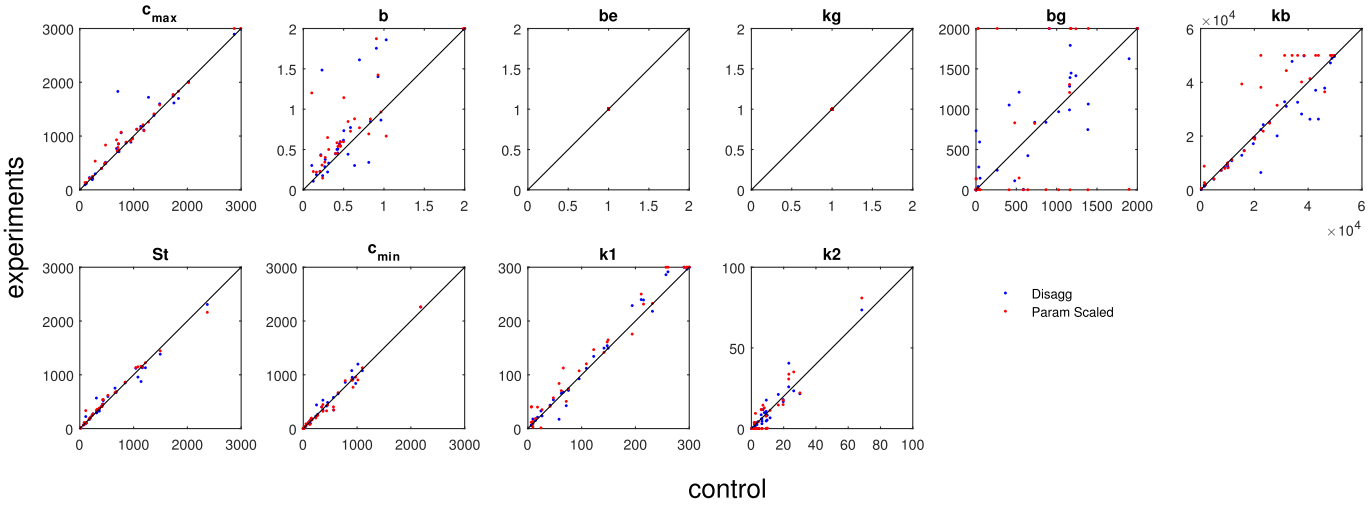

Figure C.2 PDM parameters derived from different calibration experiments and compared to the control.

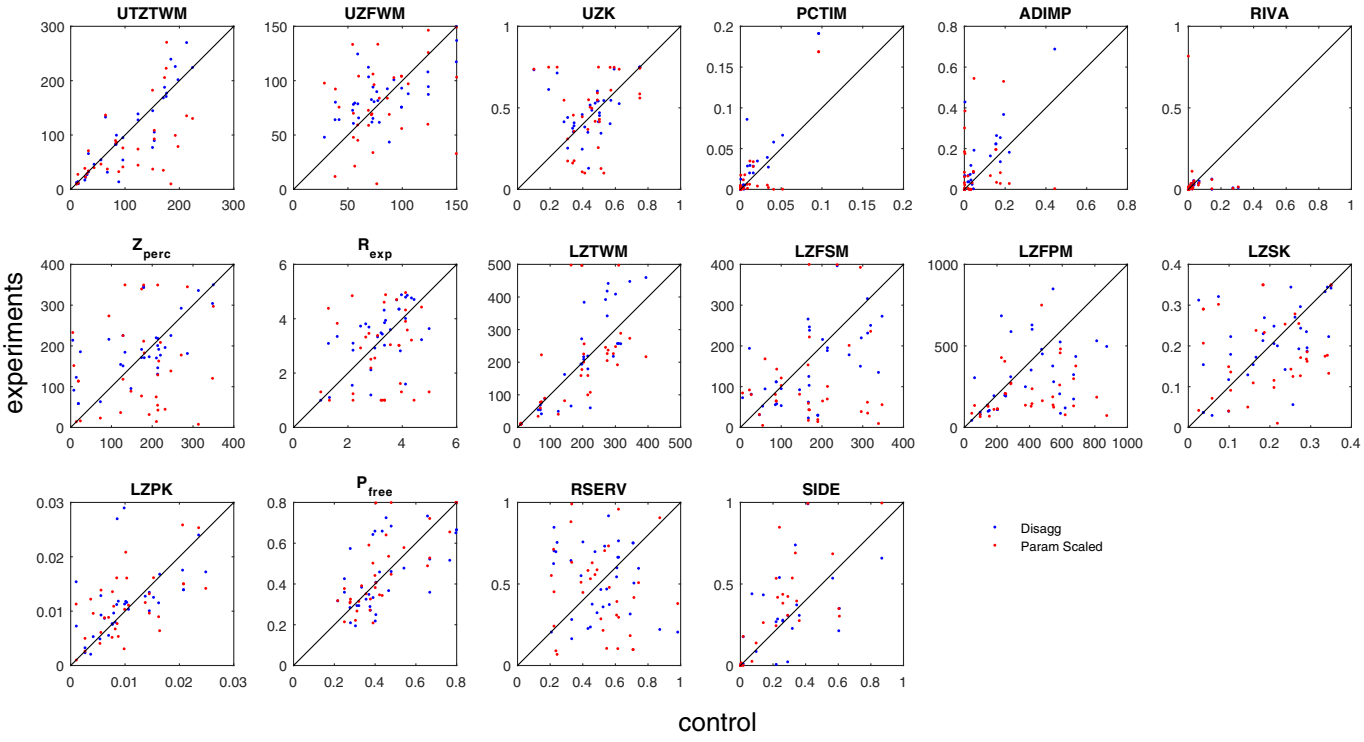

Figure C.3 Sacramento parameters derived from different calibration experiments and compared to the control.

References

Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., Lavabre, J., 2001. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. J. Hydrol. 250 (1–4), 206–223. http://dx.doi.org/10.1016/S0022-1694(01)00437-1.

Bennett, J.C., Robertson, D.E., Shrestha, D.L., Wang, Q.J., Enever, D., Hapuarachchi, P., Tuteja, N.K., 2014. A system for continuous hydrological ensemble forecasting (SCHEF) to lead times of 9 days. J. Hydrol. 519, 2832–2846. http://dx.doi.org/10.1016/j.jhydrol.2014.08.010.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Soft. 40, 1–20. http://dx.doi.org/10.1016/j.envsoft.2012.09.011.

Boughton, W.C., 2004. The australian water balance model. Environ. Model. Softw. 19, 943–956. http://dx.doi.org/10.1016/j.envsoft.2003.10.007.

Burnash, R.J.C., Ferral, R.L., McGuire, R.A., 1973. A Generalized Streamflow Simulation System-conceptual Modeling for Digital Computers. Technical Report, Joint Federal and State River Forecast Center, Sacramento, p. 204.

Demargne, J., Wu, L., Regonda, S., Brown, J., Lee, H., He, M., Seo, D.-J., Hartman, R., Herr, H.D., Fresch, M., Schaake, J., Zhu, Y., 2014. The science of NOAA's operational hydrologic ensemble forecast service. Bull. Am. Meteorol. Soc. 95, 79–98. http://dx.doi.org/10.1175/bams-d-12-00081.1.

Duan, Q., Sorooshian, S., Gupta, V.K., 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. J. Hydrol. 158 (3–4), 265–284. http://dx.doi.org/10.1016/0022-1694(94)90057-4.

Ehret, U., Zehe, E., 2011. Series distance – an intuitive metric to quantify hydrograph similarity in terms of occurrence, amplitude and timing of hydrological events. Hydrol. Earth Syst. Sci. 15, 877–896. http://dx.doi.org/10.5194/hess-15-877-2011.

Gill, M.A., 1978. Flood routing by the Muskingum method. J. Hydrol. 36 (3–4), 353–363. http://dx.doi.org/10.1016/0022-1694(78)90153-1.

Holman-Dodds, J.K., Bradley, A.A., Sturdevant-Rees, P.L., 1999. Effect of temporal sampling of precipitation on hydrologic model calibration. J. Geophys. Res. Atmos. 104 (D16), 19645–19654. http://dx.doi.org/10.1029/1999jd900121.

Kavetski, D., Fenicia, F., Clark, M.P., 2011. Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: insights from an experimental catchment. Water Resour. Res 47 (5), W05501. http://dx.doi.org/10.1029/2010wr009525.

Koutsoyiannis, D., Onof, C., 2001. Rainfall disaggregation using adjusting procedures on a Poisson cluster model. J. Hydrol. 246 (1–4), 109–122. http://dx.doi.org/10.1016/S0022-1694(01)00363-8.

Littlewood, I.G., Croke, B.F.W., 2008. Data time-step dependency of conceptual rainfall—streamflow model parameters: an empirical study with implications for regionalisation / Dépendance à l'intervalle de temps des paramètres d'un modèle conceptuel pluie—débit: étude empirique avec implications pour la régionalisation. Hydrol. Sci. J. 53 (4), 685–695. http://dx.doi.org/10.1623/hysj.53.4.685.

Mathevet, T., 2005. What is the Best Rainfall-runoff Model for the Hourly Time Step? Development and Empirical Comparison of Models on a Large Sample of Watersheds. ENGREF (Paris); Cemagref (Antony). France, Paris, p. 463.

Merz, R., Parajka, J., Blöschl, G., 2009. Scale effects in conceptual hydrological modeling. Water Resour. Res 45 (9), W09405. http://dx.doi.org/10.1029/2009wr007872.

Moore, R.J., 2007. The PDM rainfall-runoff model. Hydrol. Earth Syst. Sci. 11 (1), 483–499. http://dx.doi.org/10.5194/hess-11-483-2007.

Nalbantis, I., 1995. Use of multiple-time-step information in rainfall-runoff modelling. J. Hydrol. 165 (1–4), 135–159. http://dx.doi.org/10.1016/0022-1694(94)02567-U.

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

Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev 100 (2), 81–92. http://dx.doi.org/10.1175/1520-0493(1972)100<0081:otaosh>2.3.co;2.

Robertson, D.E., Bennett, J.C., Wang, Q.J., 2015. A strategy for quality controlling hourly rainfall observations and its impact on hourly streamflow simulations, MODSIM 2015, 21st International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, Gold Coast.

Schaake, J.C., Koren, V.I., Duan, Q.-Y., Mitchell, K., Chen, F., 1996. Simple water balance model for estimating runoff at different spatial and temporal scales. J. Geophys. Res. Atmos. 101 (D3), 7461–7475. http://dx.doi.org/10.1029/95jd02892.

Segond, M.L., Onof, C., Wheater, H.S., 2006. Spatial-temporal disaggregation of daily rainfall from a generalized linear model. J. Hydrol 331 (3-4), 674–689. http://dx.doi.org/10.1016/j.jhydrol.2006.06.019.

Shrestha, D.L., Robertson, D.E., Bennett, J.C., Wang, Q.J., 2015. Improving precipitation forecasts by generating ensembles through postprocessing. Mon. Weather Rev. http://dx.doi.org/10.1175/mwr-d-14-00329.1.

Shrestha, D.L., Robertson, D.E., Wang, Q.J., Pagano, T.C., Hapuarachchi, H.A.P., 2013. Evaluation of numerical weather prediction model precipitation forecasts for short-term streamflow forecasting purpose. Hydrol. Earth Syst. Sci 17, 1913–1931. http://dx.doi.org/10.5194/hess-17-1913-2013.

Thielen, J., Bartholmes, J., Ramos, M.-H., Ad, Roo, 2009. The european flood alert system – Part 1: concept and development. Hydrol. Earth Syst. Sci. 13, 125–140.

Vaze, J., Post, D.A., Chiew, F.H.S., Perraud, J.M., Teng, J., Viney, N.R., 2011. Conceptual Rainfall–Runoff model performance with different spatial rainfall inputs. J. Hydrometeorol. 12 (5), 1100–1112. http://dx.doi.org/10.1175/2011jhm1340.1.

Wang, Q.J., Pagano, T.C., Zhou, S.L., Hapuarachchi, H.A.P., Zhang, L., Robertson, D.E., 2011. Monthly versus daily water balance models in simulating monthly runoff. J. Hydrol. 404 (3–4), 166–175. http://dx.doi.org/10.1016/j.jhydrol.2011.04.027.

Wang, Y., He, B., Takase, K., 2009. Effects of temporal resolution on hydrological model parameters and its impact on prediction of river discharge/Effects de la résolution temporelle sur les paramètres d'un modèle hydrologique et impact sur la prévision de l'écoulement en rivière. Hydrol. Sci. J. 54 (5), 886–898. http://dx.doi.org/10.1623/hysj.54.5.886.