Conceptual Rainfall–Runoff Model Performance with Different Spatial Rainfall Inputs

J. Vaze, D. A. Post, F. H. S. Chiew, J.-M. Perraud, J. Teng, and N. R. Viney

CSIRO Water for a Healthy Country National Research Flagship, CSIRO Land and Water, Canberra, Australian Capital Territory, Australia

(Manuscript received 10 August 2010, in final form 20 January 2011)

ABSTRACT

Different methods have been used to obtain the daily rainfall time series required to drive conceptual rainfall–runoff models, depending on data availability, time constraints, and modeling objectives. This paper investigates the implications of different rainfall inputs on the calibration and simulation of 4 rainfall–runoff models using data from 240 catchments across southeast Australia. The first modeling experiment compares results from using a single lumped daily rainfall series for each catchment obtained from three methods: single rainfall station, Thiessen average, and average of interpolated rainfall surface. The results indicate considerable improvements in the modeled daily runoff and mean annual runoff in the model calibration and model simulation over an independent test period with better spatial representation of rainfall. The second experiment compares modeling using a single lumped daily rainfall series and modeling in all grid cells within a catchment using different rainfall inputs for each grid cell. The results show only marginal improvement in the "distributed" application compared to the single rainfall series, and only in two of the four models for the larger catchments. Where a single lumped catchment-average daily rainfall series is used, care should be taken to obtain a rainfall series that best represents the spatial rainfall distribution across the catchment. However, there is little advantage in driving a conceptual rainfall–runoff model with different rainfall inputs from different parts of the catchment compared to using a single lumped rainfall series, where only estimates of runoff at the catchment outlet is required.

1. Introduction

Conceptual rainfall–runoff models have been widely used for catchment water balance studies across the world. The observed runoff used to calibrate and validate these models is recorded at the catchment outlet and as such is an aggregated response of spatially variable rainfall across the catchment. There are uncertainties associated with the rainfall data, and the measured point rainfall data are usually available only at limited locations within a catchment or close to the catchment. This has been recognized in hydrologic literature as it poses a major problem for calibrating hydrological models.

This poses a major issue when using this single rainfall time series, especially for large catchments, as we are assuming that this point rainfall data series is representative of the entire catchment. This issue can be overcome in part by using spatially distributed rainfall across the catchment extracted from climate surfaces generated from a number of recorded point rainfall measurements covering a much larger area. One example is the 0.05° SILO gridded daily rainfall and climate data across Australia (Jeffrey et al. 2001), interpolated from point measurements made by the Australian Bureau of Meteorology (BoM).

It is well documented that rainfall data quality can have a large impact on both the parameter values and overall quality of hydrologic model simulations (Andréassian et al. 2001; Chaubey et al. 1999; Oudin et al. 2006; Bras and Rodríguez-Iturbe 1976a,b; Wilson et al. 1979; Faures et al. 1995; Segond et al. 2007). In addition, it has been demonstrated that the use of more rain gauges can produce better hydrologic model response (Anctil et al. 2006), although it appears from the Anctil study that the optimal results may be achieved with less than the maximum number of rain gauges. Most of the studies reported in literature are carried out using datasets from one or a limited number of catchments using a single model.
This paper examines the variation in the quality of calibration and simulation outputs for four conceptual rainfall–runoff models [Simplified Hydrolog (SIMHYD); Sacramento; Soil Moisture Accounting And Routing Model (SMARG); and Identification of Unit Hydrographs and Component Flows from Rainfall, Evaporation, and Streamflow Data (IHACRES)] when using daily rainfall time series derived in four different ways for 240 catchments in southeast Australia. The 240 catchments are selected such that they are unregulated, have long records of observed daily streamflow data (1970–2006), and have gone through negligible change in terms of land use and development over the modeling period and with catchment areas ranging from 50 to 2000 km².

The primary aim of the study is to investigate the improvement in model performance with improved spatial representation of rainfall. The second aim of the study is to assess the improvement in model performance that can be achieved in moving from a single lumped rainfall series to represent a catchment to a “distributed”

<table>
<thead>
<tr>
<th>Method</th>
<th>Rainfall</th>
<th>Modeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeling with a single lumped catchment daily rainfall series</td>
<td>M1</td>
<td>Single rain gauge closest to catchment centroid</td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>Thiessen weighted average of rain gauges within and around the catchment</td>
</tr>
<tr>
<td></td>
<td>M3</td>
<td>Arithmetic average of 0.05° interpolated rainfall surface in the catchment</td>
</tr>
<tr>
<td>Modeling at each 0.05° grid cell in a catchment with a different daily rainfall series</td>
<td>M4</td>
<td>Different rainfall series for each 0.05° grid cell</td>
</tr>
</tbody>
</table>
rainfall representation over the catchment. In the latter, the rainfall–runoff model is applied to 0.05° grid cells across the catchment with different rainfall inputs for each grid cell, but with the same model parameter values for all the grid cells.

2. Models, data, and methods

a. Rainfall–runoff models

Four widely used lumped conceptual rainfall–runoff models—SMARG, SIMHYD, Sacramento, and IHACRES—are used in this study. These models have been applied in numerous studies both within Australia and internationally. The input data into the models are daily rainfall and potential evapotranspiration (PET), and the models simulate daily runoff.

The SMARG model is a lumped conceptual rainfall–runoff model that consists of two components in sequence, namely, a water balance component and a routing component (Goswami et al. 2002; O’Connell et al. 1970). In the SMARG model, the input variables, rainfall and evaporation, are transformed into discharge through

<table>
<thead>
<tr>
<th>Rainfall (mm yr⁻¹)</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
<th>Method 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>380</td>
<td>436</td>
<td>439</td>
<td>439</td>
</tr>
<tr>
<td>25th percentile</td>
<td>658</td>
<td>674</td>
<td>720</td>
<td>720</td>
</tr>
<tr>
<td>Median</td>
<td>823</td>
<td>817</td>
<td>856</td>
<td>856</td>
</tr>
<tr>
<td>75th percentile</td>
<td>1011</td>
<td>1007</td>
<td>1065</td>
<td>1065</td>
</tr>
<tr>
<td>Maximum</td>
<td>1849</td>
<td>1833</td>
<td>2049</td>
<td>2049</td>
</tr>
</tbody>
</table>

FIG. 2. Comparison of method 1 with method 2 calibration NSE for all 240 catchments for the four models (catchments <500 km² are shown with a gray square, catchments between 500 and 1000 km² are shown using a × symbol, and catchments >1000 km² are shown using a + symbol).
a series of steps that, in a very simplified manner, mimic some of the known physical processes in the rainfall–runoff transformation. SMARG has nine parameters, which include five water balance parameters and four routing parameters (Vaze et al. 2004; Tuteja et al. 2007). In this study only eight of the parameters are optimized. The SMARG model has been exhaustively tested on data from a large number of catchments and it has been implemented successfully in a number of regionalization and river flow forecasting studies (Goswami et al. 2002; Tuteja et al. 2003).

SIMHYD is a simple lumped conceptual daily rainfall–runoff model with seven parameters (Chiew et al. 2002). For the purposes of the current study the model has been simplified to a six-parameter model by setting the values of two parameters to default values across all catchments, but introducing a new parameter, which is the time-lag parameter of an hourly Muskingum routing scheme. SIMHYD has been used successfully across Australia for various applications, including the estimation of runoff in the National Land and Water Resources Audit (Peel et al. 2000), the estimation of climate change impact on runoff (Chiew and McMahon 2002; Chiew et al. 2009), and in various regionalization studies (Zhang and Chiew 2009).

The Sacramento model is also a lumped conceptual daily rainfall–runoff model (Burnash et al. 1973), but it is considerably more complex than the other three models used in this study. The Sacramento model has been used widely across the world and in Australia, in particular as part of the river system model [Integrated Quantity and Quality Model (IQQM); Simons et al. 1996; Vaze et al. 2011a] implementations in New South Wales and Queensland. The Sacramento model has also been used in several regionalization studies (Gan and Burges 2006; Vaze and Teng 2011). The Sacramento model has 18 parameters, but in the application here, only 14 parameters

**FIG. 3.** As in Fig. 2, but for the comparison of method 2 with method 3.
are optimized with the other four parameters set to default values.

IHACRES is a lumped conceptual rainfall–runoff model based on unit hydrograph principles. The model is defined by seven parameters, which together predict the daily hydrologic response of a catchment. The model consists of a nonlinear module to convert rainfall to effective rainfall and a linear module to route this effective rainfall to streamflow. The model has been previously used in both land use and climate change studies, as well as in regionalization to ungauged catchments (Post and Jakeman 1999).

b. Streamflow data

Daily streamflow data from 240 unregulated medium sized catchments (50–2000 km²) from southeast Australia are used in this study (Fig. 1). The study area in southeast Australia is about 1.4 million km² and covers about 20% of mainland Australia. The region generates more than half of Australia’s agricultural income, and more than half of Australia’s population lives in the southern and eastern parts of the region. The climate varies considerably across the region, from temperate near the coast to semiarid and arid further inland toward the northwest and west. All the catchments are in the higher runoff generation areas in the southeast and eastern perimeter of the region (Fig. 1). The mean annual streamflow ranges from 30–450 mm (10th to 90th percentile values) with a median of 120 mm, and the annual runoff coefficient ranges from 0.06–0.36 with a median of 0.15. The streamflow data has been compiled and quality controlled for large regional-scale modeling (Vaze et al. 2011b).

c. Rainfall and PET data

Four different methods are used to obtain the daily rainfall series for the rainfall–runoff modeling in each of

![Fig. 4. As in Fig. 2, but for the comparison of method 3 with method 4.](image-url)
the 240 catchments (see Table 1). The rainfall data in methods 1 and 2 come from the rain gauges in Fig. 1. Method 1 uses the rainfall data from a single rain gauge closest to the centroid of the catchment. Method 2 uses a Thiessen polygon weighting of all rain gauges in and around the catchments (within 100 km) to calculate a single lumped catchment-average daily rainfall series for the modeling. In method 2, the daily rainfall series for 10% of the 240 catchments come from a single rain gauge (therefore the same as method 1), 30% from two rain gauges, 30% from three rain gauges, 20% from four rain gauges, and 10% from five or more rain gauges.

Methods 3 and 4 use the SILO rainfall surface for 0.05° grid cells \((5 \times 5)\) km. Unlike method 2, the SILO rainfall surface is derived from all available BoM rainfall stations using a sophisticated thin-spline interpolation, which takes into account rainfall variations with elevation. Method 3 aggregates the SILO interpolated rainfall surface from all the 0.05° grid cells in the catchment to obtain a single daily rainfall series for the modeling. Unlike methods 1, 2, and 3, which use a single input rainfall series for a catchment, method 4 applies the rainfall–runoff model to each 0.05° grid cell in a catchment, each with a different daily rainfall series. The same model parameter values are used for all the grid cells. The modeled daily runoffs from all the grid cells are aggregated (without any spatial routing) to obtain the modeled daily runoff for the catchment. Like the other methods the model is calibrated against observed streamflow at the catchment outlet. Method 4 is referred to as “distributed rainfall” modeling.

These four rainfall input methods are used for the modeling experiments throughout this study. The rainfall statistics for the four methods is provided in Table 2. The daily PET series for the rainfall–runoff modeling is obtained from the 0.05° climate data (incoming solar radiation, maximum and minimum temperature, and actual vapor pressure) using Morton’s wet environment (or equilibrium evaporation or areal potential evaporation) algorithms (Morton 1983; Chiew and McMahon 1991). The PET used here is conceptually the upper limit of actual evaporation in the rainfall–runoff models. The same PET series is used for all the modeling experiments.

d. Model calibration

The models are calibrated against the 1984–2006 observed daily streamflow data from the 240 catchments. In the model calibration, the model parameters are optimized to maximize the Nash–Sutcliffe efficiency (NSE)–bias objective function, which is a weighted combination of Nash–Sutcliffe (Nash and Sutcliffe 1970) efficiency and a logarithmic function of bias given by

<table>
<thead>
<tr>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;500 km² (141 catchments)</td>
</tr>
<tr>
<td>500–1,000 km² (70 catchments)</td>
</tr>
<tr>
<td>1,000 km² (23 catchments)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Median (and 10th and 90th percentiles) of improvement in NSE between methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMARG</td>
<td>M1 0.56 (0.50–0.61) M2 0.59 (0.53–0.63) M2–M1 0.01 (0.00–0.11)</td>
</tr>
<tr>
<td>SIMHYD</td>
<td>M1 0.62 (0.56–0.67) M2 0.68 (0.62–0.72) M2–M1 0.07 (0.00–0.17)</td>
</tr>
<tr>
<td>Sacramento</td>
<td>M1 0.70 (0.65–0.75) M2 0.76 (0.70–0.80) M2–M1 0.08 (0.00–0.20)</td>
</tr>
<tr>
<td>HICRES</td>
<td>M1 0.72 (0.67–0.78) M2 0.78 (0.73–0.83) M2–M1 0.05 (0.00–0.13)</td>
</tr>
</tbody>
</table>

Table 3. Median calibration NSE of the 240 catchments for the 4 models and median (10th and 90th percentiles of improvement in NSE between each method and the next are given in parentheses).
where NSE is the Nash–Sutcliffe efficiency and $B$ is the bias (total modeled error divided by observed total streamflow) (Viney et al. 2009). The coefficients of this equation control the severity and shape of the resulting constraint penalty. A 1-yr warm-up period is used at the start of the calibration period. With the above NSE-bias objective function, the models are effectively calibrated to essentially maximize NSE while ensuring that the bias is small (total modeled runoff is not too different to the total observed runoff). The model parameters are optimized using the shuffled complex evolution global optimization method (Duan et al. 1993) followed by a local optimization method (Rosenbrock 1960).

The NSE expresses the proportion of variance of the observed runoff that can be accounted for by the model and provides a direct measure of the ability of the model to reproduce the observed runoff, with $NSE = 1.0$ indicating that all the modeled daily runoff is the same as all the observed daily runoff.

e. Model simulation

The optimized parameter values from the above model calibration against 1984–2006 streamflow data are also used to model runoff over 1970–83, and the 1970–83 modeled runoff is compared to the observed runoff to assess the model performance over an independent test period.

3. Results and discussion

a. Calibration results

Figure 2 compares the calibration NSE values for method 2 rainfall input versus method 1 rainfall input for the four rainfall–runoff models for the 240 catchments.
Figure 3 compares the NSE values for method 3 versus method 2, and Fig. 4 compares the NSE values for method 4 versus method 3. In the plots, different symbols are used to distinguish catchments below 500 km² (141 catchments), between 500 and 1000 km² (76 catchments), and above 1000 km² (23 catchments). Table 3 summarizes the median NSE values from all the catchments within each of the three area categories for the four methods and the improvement of each method over another. As in previous applications, the four rainfall–runoff models can generally be calibrated to reproduce the observed daily runoff, with NSE values greater than 0.5 in 67%, 77%, 92%, and 93% of the catchments in methods 1–4, respectively.

Methods 1, 2, and 3 use a single daily rainfall series to represent the entire catchment, with improved consideration of spatial rainfall as we move from method 1 (single rainfall station), to method 2 (Thiessen average of several rainfall stations), to method 3 (average of interpolated rainfall surface). There is a clear improvement in the model calibration for all the four rainfall–runoff models and for practically all the 240 catchments with better consideration of spatial rainfall, with method 3 consistently giving higher NSE values compared to method 2 (Fig. 3), and method 2 consistently giving higher NSE values than method 1 (Fig. 2).

The median increase in the NSE value in method 2 over method 1 is about 0.01 in the four models for catchments less than 500 km², 0.05–0.07 for catchments between 500 and 1000 km², and 0.07–0.11 for catchments larger than 1000 km². The bigger improvement in the larger catchments is likely because larger catchments require more rainfall stations to adequately capture the spatial rainfall variability than smaller catchments. The median increase in the NSE value in method 3 over method 2 is 0.08–0.12 for catchments less than 500 km², 0.07–0.12 for catchments...
for catchments between 500 and 1000 km$^2$, and 0.03–0.05 for catchments larger than 1000 km$^2$. The smaller improvement in the larger catchments could be due to the larger catchments having more rainfall stations to obtain a Thiessen average daily rainfall series (method 2) compared to the smaller catchments, and therefore a catchment-average rainfall series from an interpolated rainfall surface (method 3) would benefit larger catchments less than the smaller catchments. All four rainfall–runoff models show improvements in the model calibration in method 3 over method 2 and in method 2 over method 1.

Both methods 3 and 4 use the same SILO 0.05° gridcell interpolated rainfall data. In method 3, the rainfalls from all the 0.05° grid cells in a catchment are averaged to obtain a single daily rainfall series for the catchment. In method 4, the rainfall–runoff model is applied to each 0.05° grid cell (different rainfall series for each grid cell) and the modeled runoffs in all the grid cells are aggregated to obtain the modeled daily runoff series for the catchment. There is little difference in the calibration NSE values between method 3 and the distributed consideration of spatial rainfall in method 4 in most of the catchments, except for the larger catchments (>1000 km$^2$), where there is marginal improvement in method 4 over method 3 in the SMARG modeling (median NSE improvement of 0.05) and very marginal improvement in the SIMHYD and Sacramento modeling (median NSE improvement of 0.01 and 0.02; Fig. 4 and Table 3). The marginal improvement in method 4 over method 3 observed for the larger catchments with these three models is likely due to the greater need for distributed consideration of spatial rainfall gradients in larger catchments. These improvements might be greater if a fully distributed

FIG. 7. As in Fig. 2, but for the comparison of method 3 with method 4 simulation.
modeling approach considering explicit cell-to-cell routing was used.

The improvement in method 4 over method 3 (distributed rainfall consideration through model application to multiple grid cells versus modeling with a single rainfall series), where it occurs, is considerably smaller than the improvement from method 3 over method 2 and method 2 over method 1 (using a single rainfall series, but with improved representation of spatial rainfall).

b. Simulation results

Similar to the calibration results in section 3a, Fig. 5 compares the simulation NSE values (for the independent 1970–83 test period using the optimized parameter values from the calibration against 1984–2006 data) for method 2 versus method 1, Fig. 6 compares the NSE values for method 3 versus method 2, and Fig. 7 compares the NSE values for method 4 versus method 3. Table 4 summarizes the median NSE values from all the catchments in the three area categories for the four methods and the improvement of each method over another. The NSE values for the daily runoff modeled over the independent test period are slightly lower than the calibration NSE values. Nevertheless, the modeled daily runoff generally match the observed daily runoff reasonably well, with NSE values greater than 0.5 in 52%, 59%, 81%, and 85% of the catchments for methods 1–4, respectively.

The results for the model simulation over the independent test period are very similar to the model calibration results, with the better representation of spatial rainfall in method 3 consistently giving higher NSE values than method 2, which in turn is consistently better than method 1. However, there is considerably more variability in the results, with some catchments showing much higher NSE values with better spatial rainfall representation in the model simulation compared to model calibration, while there are some catchments with lower NSE values with better spatial rainfall representation in the model simulation compared to little or no catchments with lower NSE values in the model calibration (cf. Fig. 5 vs Fig. 2 and Fig. 6 vs Fig. 3). Overall the median improvements in the NSE values for method 3 compared to method 2, and method 2 compared to method 1 are generally slightly higher in the model simulation than in the model calibration (cf. Tables 4 and 3). While the greater variability in the model simulation results compared to the model calibration results can be expected, the modeling results here indicate that the improvements in model calibration using a better spatial representation of rainfall also translates to model simulation over another independent period.

<table>
<thead>
<tr>
<th>Area</th>
<th>Median NSE from all catchments</th>
<th>Median (and 10th and 90th percentiles) of improvement in NSE between each method and the next</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;500 km²</td>
<td>MA: 0.48 Sacramento: 0.51 M2: 0.55</td>
<td>M2-M1: 0.02 (0.01–0.14) M3-M2: 0.02 (0.02–0.10) M4-M3: 0.02 (0.00–0.06) M4-M2: 0.02 (0.00–0.04)</td>
</tr>
<tr>
<td>(141 catchments)</td>
<td></td>
<td>0.35 (0.32–0.40) 0.66 (0.64–0.68) 0.67 (0.65–0.69) 0.67 (0.65–0.70)</td>
</tr>
<tr>
<td>500–1000 km²</td>
<td>MA: 0.48 Sacramento: 0.55 M2: 0.66</td>
<td>M2-M1: 0.02 (0.02–0.22) M3-M2: 0.06 (0.02–0.21) M4-M3: 0.02 (0.00–0.07) M4-M2: 0.02 (0.00–0.04)</td>
</tr>
<tr>
<td>(117 catchments)</td>
<td></td>
<td>0.66 (0.64–0.68) 0.70 (0.67–0.73) 0.71 (0.69–0.73) 0.71 (0.69–0.73)</td>
</tr>
<tr>
<td>&gt;1000 km²</td>
<td>MA: 0.48 Sacramento: 0.56 M2: 0.70</td>
<td>M2-M1: 0.02 (0.01–0.14) M3-M2: 0.02 (0.01–0.09) M4-M3: 0.02 (0.00–0.07) M4-M2: 0.02 (0.00–0.04)</td>
</tr>
<tr>
<td>(17 catchments)</td>
<td></td>
<td>0.70 (0.68–0.72) 0.73 (0.70–0.75) 0.74 (0.72–0.76) 0.74 (0.72–0.76)</td>
</tr>
</tbody>
</table>
As in the model calibration, there is little difference between method 4 and method 3 (distributed rainfall for multiple grid cells versus modeling with a single aggregated rainfall series) in the IHACRES model (median NSE improvement of 0.0 for catchment smaller than 1000 km² and an improvement of 0.01 for catchments larger than 1000 km²). As in the model calibration, the NSE values from method 4 are slightly higher than the NSE values from method 3 for SMARG, SIMHYD, and Sacramento models, and the differences appear to be bigger in the model simulation than in the model calibration (cf. Fig. 7 vs Fig. 4).

c. Modeled mean annual runoff

The scatterplots in Fig. 8 compare the modeled and observed mean annual runoff over the simulation period for the 240 catchments for the four models and four rainfall input methods. The model calibration results are not shown because the calibration against the NSE-bias objective function already forces the model to closely match the observed mean annual runoff (the modeled mean annual runoff in the model calibration is within 5% of the observed mean annual runoff in practically all the catchments).
4. Conclusions

The results indicate that the models can estimate the mean annual runoff reasonably well for the independent test/simulation period. The modeled mean annual runoff is within 10% of the observed value in more than half of the catchments and within 30% of the observed value in almost all the catchments. The results also indicate that the better spatial rainfall representation generally gives better estimates of mean annual runoff for the SMARG, SIMHYD, and particularly the Sacramento model (the modeled values are closer to the observed values for method 3 compared to method 2 and for method 2 compared to method 1, best seen in the smaller RMSE values in the better spatial rainfall representation). The greater improvements in RMSE from method 2 to method 3 for the Sacramento model probably reflect the ability of the model to utilize more effectively the additional information contained in the improved rainfall estimates. Conversely, the simpler IHACRES model is not able to effectively utilize the additional information in the improved rainfall inputs and the RMSE actually increases going from method 2 to method 3. As in the earlier results, there is little difference between method 4 and method 3.

Acknowledgments. This study is undertaken as part of the Catchment Water Yield Estimation Tool (CWYET) Project carried out in the CSIRO Water for Healthy Country National Research Flagship and eWater Cooperative Research Centre. We thank Lu Zhang and Nick Potter for reviewing the manuscript.

REFERENCES

——, M. C. Peel, and A. W. Western, 2002: Application and testing of the simple rainfall-runoff model SIMHYD. Mathematical Models of Small Watershed Hydrology and Applications, V. P.


