International Congress on Environmental Modelling and Software

3rd International Congress on Environmental Modelling and Software - Burlington, Vermont, USA - July 2006

Jul 1st, 12:00 AM

Model Averaging, Equifinality and Uncertainty Estimation in the Modelling of Ungauged Catchments

J. P. C. Reichl

F. H. S. Chiew

A. W. Western

Model Averaging, Equifinality and Uncertainty Estimation in the Modelling of Ungauged Catchments

J.P.C.Reichl[1], **F.H.S. Chiew**[1,2] **and A.W. Western**[1]

[1]*The University of Melbourne, Australia*

[2]*CSIRO Land and Water, Australia*

email: j.reichl@civenv.unimelb.edu.au

Abstract: The problem of estimating runoff in ungauged catchments remains an important but elusive one. Previous studies suggest that there are two important properties common to rainfall-runoff models: over-parameterisation, leading to parameter covariance; and equifinality, the existence of multiple parameter sets which reproduce the streamflow adequately. Both reduce parameter identifiability, impeding identification of relationships between model parameter values and catchment characteristics that would otherwise be useful for regionalisation. This study investigates the use of a model averaging framework to circumvent this problem. Multiple behavioural parameter sets arising from Monte-Carlo simulation are used from each catchment in order to retain information about data and parameter uncertainty, and to estimate the uncertainty of prediction. The model averaging is based simply on each catchment's physical similarity to a target ungauged catchment. Ungauged prediction results are assessed based on the Nash-Sutcliffe Efficiency (E). The model averaging schemes are compared to local cross-validation results (the benchmark), 'nearest neighbour' calibration (parameter sets taken from calibration of the geographically closest gauged catchment) and 'regression method' (each parameter value is estimated using regressions between optimised parameter values in gauged catchments and catchment characteristics). The study is carried out using the conceptual daily rainfall-runoff model SIMHYD on 44 catchments in south-east Australia. The results indicate that the model averaging approach shows promise for estimating streamflow in ungauged catchments. The streamflow simulations are significantly better when model parameter sets are retained using the model averaging approach than when parameter values are estimated using the nearest neighbour and regression approaches. The results from the model averaging approach are also better than the cross-validation results in over 40 % of the catchments. It is likely that more detailed analyses of the choice of weights and descriptors of catchment similarity will lead to even better modelling results.

Keywords: Runoff estimation; regionalisation; ungauged catchment; model averaging; SIMHYD

1. INTRODUCTION

The ability to estimate streamflow time series in ungauged catchments is important to natural resource management, but has to date been elusive. Attempts at applying 'regional' information to infer ungauged hydrological behaviour (known as regionalisation) have met with little success. Recent focus has been on finding correlations between conceptual rainfall-runoff model parameters and catchment physical characteristics, be they measured or estimated (Chiew et al., 2005, Merz and Bloschl, 2004). This approach (referred to as 'parameter regression') has been thwarted, largely due to the many sources of error and uncertainty in the modelling of complex natural systems, and the implications of this for estimation of meaningful parameter values. Errors typically arise from input data, model structure, parameter choice and observed output data (Thyer et al., 2005). Uncertainty in data and model structure, in addition to parameter covariance (ever-present in conceptual rainfall-runoff models) often result in parameter equifinality (Beven and Freer, 2001). All of the aforementioned factors have hampered efforts at regionalisation.

This paper investigates the potential of a model averaging approach to regionalisation. Rather than searching for a relationship between 'optimal' parameter values and catchment physical characteristics, the model averaging approach relates gauged catchments to ungauged based on measures of physical similarity (McIntyre et al., 2004). This paper investigates the behaviour of the SIMHYD model parameters, retaining multiple 'behavioural' parameter sets, and thus information about parameter uncertainty (Beven and Freer, 2001, Wagener, 2003), within the model averaging framework.

The potential for use of this technique is investigated in a case study of 44 catchments in south-east Australia, and compares regionalisation results with local cross-validation, 'nearest neighbour' and 'parameter regression' regionalisation techniques.

2. REGIONALISATION

Various approaches have been used to predict hydrological behaviour in ungauged catchments. Most are based on the idea of common physical properties. Merz and Bloschl (2004) conducted a thorough study of 308 catchments in Austria, comparing 8 regionalisation techniques, including the use of parameter sets from the closest upstream and downstream catchments, a parameter regression approach and the use of a 'global' parameter set. They found that, "apparently, spatial proximity is a better surrogate of unknown controls on runoff dynamics than catchment characteristics", since the upstream/downstream approach performed significantly better than the parameter regression approach.

Peel et al. (2000) in a study of 331 catchments in Australia found statistically significant (but not strong) correlations between most SIMHYD model parameters and catchment characteristics. Chiew et al. (2005) investigated the potential for regionalisation of SIMHYD model parameters based on these correlations. Some results from the study of Chiew et al. (2005) will be compared with the regionalisation results described in this paper.

Vogel (2005), in addition to a thorough review of regionalisation techniques, described a variation on the regression approach, whereby the calibration of individual catchments aimed at optimising local streamflow reproduction while at the same time optimising the regional relationships between parameter values and catchment characteristics. It was shown that while this approach led to highly correlated relationships, these relationships were very similar to the weaker relationships found when calibration optimised streamflow reproduction alone, and thus produced extremely similar (and quite poor) results in predicting ungauged streamflow.

McIntyre et al. (2004) described an alternate approach to the problem of regionalisation. In a study of 36 catchments in the UK, intact parameter sets from calibrated gauged catchments were regionalised to predict streamflow in ungauged catchments. Each gauged catchment was assigned a prior likelihood of accurately predicting streamflow in the ungauged catchment. This likelihood was based on physical similarity, and allows for the information contained in an intact parameter set to be retained. This approach had some success, achieving regionalisation results which were far better than the regression approach, and in some cases as good as the local calibration results.

3. CASE STUDY

In order to investigate the potential of the model averaging approach in estimating streamflow in ungauged catchments, a case study of 44 catchments in southeast Australia was undertaken.

The study catchments are located in Victoria, in the south-east of Australia.

The data consist of daily rainfall time series, monthly streamflow time series and mean monthly areal potential evapotranspiration. The data is a subset of the Australian dataset collated for an Australian Land and Water Resources Audit project (Peel et al., 2000). All catchments have at least 10 years of streamflow data for model calibration and testing.

The catchments range in area from 50 to 850 km^2, mean annual rainfall from 550 to 2100 mm and mean annual runoff from 40 to 1370 mm.

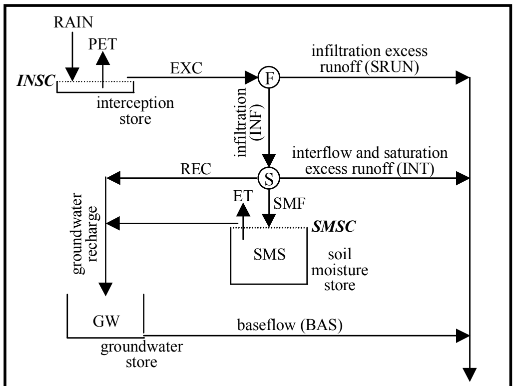
Five catchment characteristics (obtained from Lowe and Nathan, 2005) are used as measures of catchment similarity (details in Section 3.2.1):

- Mean annual rainfall (*RAIN*)
- Catchment area (*AREA*)
- Percentage forest cover (*TREE*)
- Soil permeability (*KS*)
- Stream frequency (*SFREQ*)

3.1 SIMHYD Conceptual Rainfall-Runoff Model

SIMHYD is a lumped conceptual daily rainfall-runoff model. It is driven by daily rainfall and potential evapotranspiration (PET), and simulates daily streamflow. It has been tested and used extensively across Australia (Chiew et al., 2002).

Figure 1 shows the structure of SIMHYD and the algorithms controlling the inflow of water from precipitation, through several stores, and the outflow through ET and runoff (for full details of SIMHYD see Chiew et al., 2002). SIMHYD has seven model parameters.



PET = areal potential evapotranspiration (input data)
EXC = RAIN - ***INSC***, EXC > 0
INF = lesser of { COEFF exp (-***SQ*** x SMS/***SMSC***) , EXC }
SRUN = EXC - INF
INT = ***SUB*** x SMS/***SMSC*** x INF
REC = ***CRAK*** x SMS/***SMSC*** x (INF - INT)
SMF = INF - INT - REC
ET = lesser of { 10 x SMS/***SMSC*** , PET }
BAS = ***K*** x GW

Model parameters

INSC interception store capacity (mm)
COEFF maximum infiltration loss (mm)
SQ infiltration loss exponent
SMSC soil moisture store capacity (mm)
SUB constant of proportionality in interflow equation
CRAK constant of proportionality in groundwater recharge equation
K baseflow linear recession parameter

Figure 1. Lumped conceptual rainfall-runoff model SIMHYD

3.1.1 Equifinality and SIMHYD parameterisation

The concept of equifinality, while not new, has significant implications for regionalisation of model parameters. Beven and Freer (2001) state that “it may be endemic to mechanistic modelling of complex environmental systems that there are many different model structures and many different parameter sets within a chosen model structure that may be behavioural or acceptable in reproducing the observed behaviour of that system”.

This paper carries out an investigation into how this statement relates to the SIMHYD model. The results of 40 000 Monte-Carlo simulations run on the 44 catchments are assessed in terms of the sum of squared errors of monthly streamflow (see Equation 1, where sse is the objective function, EST_t and REC_t are the estimated and recorded runoff at time-step t, and T is the total number of time-steps). Figure 2 shows the results for catchment 401210. For each parameter (only two are shown here) it is seen that across the entire range sampled (which reflects the range in which that parameter physically makes sense), very good (low objective function values) and very bad (high objective function values) are achieved.

$$sse = \sum_{t=1}^{T}\left(EST_t - REC_t\right)^2 \qquad (1)$$

This indicates that it is not the individual parameter value which determines the quality of the simulation. Rather, it is the set of parameters as a whole. This equifinality of SIMHYD (and other conceptual rainfall-runoff models), amounts to large parameter uncertainty. It suggests that this is the major reason for the failure of attempts at parameter regression against physical characteristics.

Since there are many ‘behavioural’ parameter sets for each catchment, Monte-Carlo simulations are an effective way of calculating confidence limits of predicted time series (Beven and Freer, 2001, Wagener, 2003).

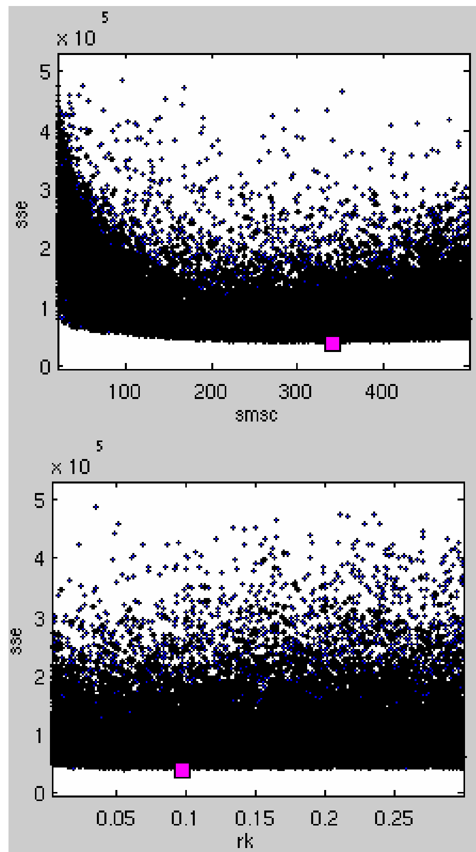


Figure 2. Plots showing model sum of squared errors across the range of two SIMHYD parameters for catchment 401210. The large square represents the best performing parameter set

3.2 Streamflow prediction in ungauged catchments

In this case study each catchment is treated as being ungauged in turn, and information from the other 43 gauged catchments is regionalised in various ways. This allows for assessment of the quality of the simulation against the observed streamflow time series for that catchment. Four approaches to regionalising rainfall-runoff parameters are compared in this paper.

- Nearest Neighbour approach – This approach takes the calibrated model parameters from the geographically nearest gauged catchment and uses those parameters to simulate streamflow in the target ungauged catchment.;
- Parameter Regression approach – In this approach, each optimised model parameter is related to various physical and climatic catchment characteristics, and the relationships are used to estimate the model parameter values for the ungauged catchment. This study uses the results of Chiew et al. (2005), where multiple linear regressions relate each optimised parameter values to various catchment characteristics.

  It should be noted that the regression relationships used here are derived from an Australia-wide dataset. Relationships derived specifically for these 44 catchments may yield better results;
- Model Averaging approach – In this approach, a weight is assigned to each catchment based on its physical similarity to the target gauged catchment (essentially a likelihood that that catchment will accurately represent the target catchment – see Section 3.2.1). Simulations are run using each parameter set separately, with the input precipitation and PET data from the target 'ungauged' catchment. The output time series are then obtained by adding the "weighted" results from all the simulations (described in more detail in Section 3.2.1);
- Model Averaging-Threshold approach – This is a variation of the Model Averaging approach. It uses only the information from those catchments which fall within a threshold of similarity. In this study the threshold is chosen to be D > 2 (see Equation 2). The rationale here is that probably there comes a point at which the catchment is so dissimilar to the target catchment that its parameter set is not useful at all in representing the behaviour of the target catchment.

Within both of the Model Averaging approaches five variations are used. The 100 best, 50 best, 20 best and 2 best parameter sets from the Monte-Carlo simulations are retained for each catchment, as well as the single set obtained from an automatic pattern-search calibration. In all cases the Nash-Sutcliffe Efficiency *(E)* (Nash and Sutcliffe, 1970) of monthly streamflows are used to assess the model performance.

3.2.1 Model Averaging

McIntyre et al. (2004) used with some success a measure of dissimilarity taken from the Institute of Hydrology Flood Estimation Handbook (1999). In order to assess the potential of a similar approach for use in Australian catchments, this paper uses the same measure of dissimilarity as a starting point.

Catchment dissimilarity (D) is described by Equation 2, with a higher D value denoting less similarity between catchments.

$$D_i = \sqrt{\frac{1}{2}\left(\frac{\ln A_i - \ln A_u}{\sigma(\ln A_u)}\right)^2 + \left(\frac{\ln R_i - \ln R_u}{\sigma(\ln R_u)}\right)^2 + \left(\frac{B_i - B_u}{\sigma(B_u)}\right)^2} \quad (2)$$

D_i is the dissimilarity of the ith gauged catchment with respect to the target ungauged catchment, and $A_{i,\ u}$, $R_{i,\ u}$ and $B_{i,\ u}$ are the area, standardised mean annual rainfall and baseflow coefficient of the ith gauged catchment and the target ungauged catchment respectively.

Here baseflow index cannot be calculated via the normal means (Kalman filtering of streamflow, see Maybeck, 1979), as, clearly, the streamflow is not available for an ungauged catchment. Instead, a slightly cruder method has been adopted from Lowe and Nathan (2005), who established a strong relationship (using data from Victorian catchments) between baseflow index (from Kalman filtering) and four catchment characteristics, namely soil permeability (*KS*), percentage forest cover (*TREE*), standardised mean annual rainfall (*RAIN*) and stream frequency (*SFREQ*, the density of stream junctions). This allows for estimation of baseflow index without access to streamflow data.

The catchment weights are defined by Equation 3, with Wc_i being the weight assigned to the ith catchment.

$$Wc_i = \frac{1}{\sum_{i=1}^{n}(1 - D_i / D_{MAX})}(1 - D_i / D_{MAX}) \quad (3)$$

In the Model Averaging cases where many parameter sets from each gauged catchment are retained, these are also weighted. Each is assigned a prior likelihood of adequately representing the catchment from whence it came. This is based on each set's performance in the calibration period, reflected by the E value returned, as described by Equation 4.

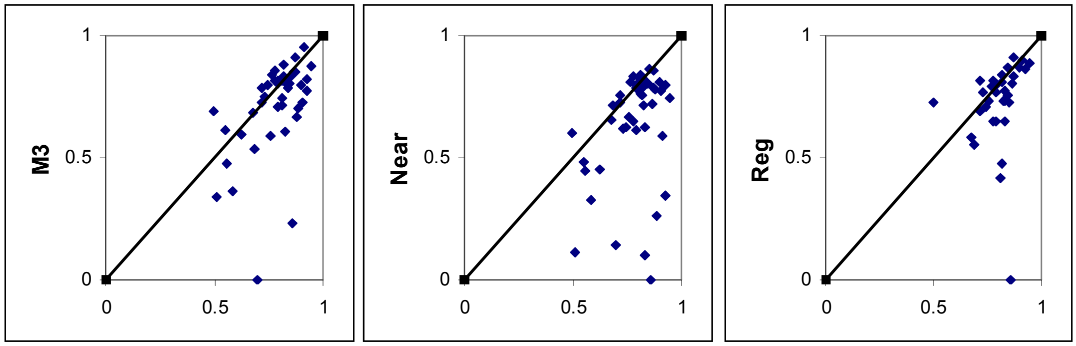


Figure 3. E values for (a) M3 Model Averaging; (b) Nearest Neighbour; and (c) Parameter Regression approaches versus Local Cross-Validation

$$Wp_{i,j} = \frac{E_{i,j}}{\sum_{j=1}^{m} E} \tag{4}$$

$Wp_{i,j}$ is the weight assigned to the jth parameter set of the ith gauged catchment.

$$S_t = \sum_{i=1}^{m} Wc_i \left\{ \sum_{j=1}^{n} Wp_{i,j} M\left(\Theta_{i,j}, Y_t\right) \right\} \tag{5}$$

Equation 5 describes the assembly of the streamflow time series, where S_t is the streamflow output at time step t, m is the number of gauged catchments, n is the number of retained parameter sets per catchment, $\Theta_{i,j}$ is the parameter set j for the ith gauged catchment, and Y_t is the input data at time step t. M in this case is the structure of the model SIMHYD, as only one model structure is being used.

3.2.2 Results

The Nash-Sutcliffe Efficiency (E) for various modelling results are summarised in Figures 3 and 4 and Table 1. "Cal" in Figure 4 shows results for SIMHYD when calibrated against the entire time series. In "Val" in Figures 3 and 4 and Table 1, the data are divided into three parts, with the parameter values optimised using two parts of the data used to estimate the flows in the third part (cross-validation). "Near" refers to the nearest gauged catchment. M1 is Model Averaging using the 100 best parameter sets from each catchment. M2 is the same as M1 but with the threshold of similarity applied (Model Averaging - Threshold), and M3 is Model Averaging - Threshold using the 2 best parameter sets from each catchment.

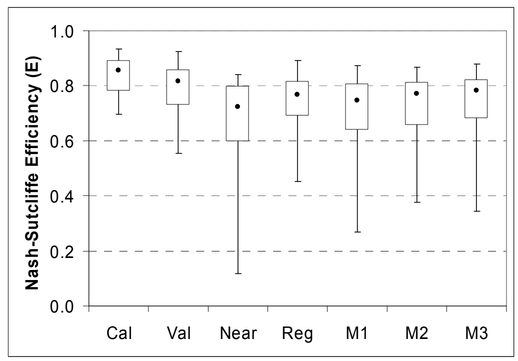


Figure 4. Nash-Sutcliffe Efficiencies for Local Calibration, Local Cross-Validation, Nearest Neighbour, Regression, and three Model Averaging approaches. The marker represents the 50th percentile, the box the 25th-75th percentile, and the bars the 5th-95th percentile

Table 1. Comparison of E values of Local Validation, Nearest Neighbour, Parameter Regression and M3 (the best of the Model Averaging approaches)

	Loc Val	Nearest	Regression	M3
Total Mean	0.78	0.63	0.71	0.71
%>Nearest	80	N/A	58	73
%>Loc Val	N/A	20	27	41
%best regionalisation	N/A	21	27	52

The results indicate that the regionalisation of parameter sets using the Model Averaging method generally leads to better flow simulations than when using parameter values estimated using the Nearest Neighbour or Regression methods. This is highlighted by the higher E values (Figures 3 and 4), and by Table 1, which shows that in over

40 % of catchments the M3 result is better than that for the Local Cross-Validation, and better than the other regionalisation techniques in over 50 % of cases.

There is little difference between the Model Averaging methods with and without the similarity threshold imposed (eg. M1 vs M2). It is likely that this is an artefact of the relatively small sample of catchments, or that the similarity threshold is not discriminatory enough. While it is possibile that the best approach is to include all catchments, regardless of similarity, it seems counter-intuitive and unlikely.

4. CONCLUSION

This paper compares four approaches for regionalising parameter values for use in ungauged catchments, by applying the daily conceptual rainfall-runoff model SIMHYD to 44 catchments in south-east Australia.

The results indicate that the model averaging approach shows promise for estimating streamflow in ungauged catchments. The streamflow simulations are significantly better when model parameter sets are regionalised using the model averaging approach than when parameter values are estimatead using the nearest neighbour and regression approaches. This is due to the utilisation of parameter equifinality and maintenance of parameter inter-relationships. The results from the model averaging approach are also better than the cross-validation results in over 40 % of the catchments.

It is likely that more detailed analyses of the choice of weights, descriptors of catchment similarities and thresholds will lead to even better modelling results.

5. REFERENCES

Beven, K., and J. Freer, Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *Journal of Hydrology*, 249, 11-29, 2001

Chiew, F.H.S., M.C. Peel and A.W. Western, Application and testing of the simple rainfall-runoff model SIMHYD, *Mathematical models of small watershed hydrology and application (Eds: V.P.Singh and D.K.Frevert), Water Resources Publication, Littleton, Colorado, USA*, 335-367, 2002

Chiew, F.H.S., and L. Siriwardena, Estimation of SIMHYD parameter values for application in ungauged catchments, *In Zerger, A. and Argent, R.M. (eds) MODSIM 2005 International Congress on modelling and simulation, Modelling and Simulation Society of Australia and New Zealand*, 2005

Vogel, R.M., Regional calibration of a watershed model, *in Watershed Models, Chapter 3 (eds: V.P. Singh and D.F. Frevert), CRC Press*, 549-567, 2005

Institute of Hydrology (IH), *Flood Estimation Handbook, Wallingford, UK*, 1999

Lowe, L., and R. Nathan, A robust procedure for transposing gauged streamflows to ungauged catchments, *29th Hydrology and Water Resources Symposium, Engineers Australia*, 2005

McIntyre, N., H. Lee, H. Wheater, and A. Young, Tools and approaches for evaluating uncertainty in streamflow predictions in ungauged UK catchments, *In Pahl-Wostl, C., Schmidt, S., Rizzoli, A., and Jakeman, A., (eds) Complexity and Integrated Resources Management, Transactions of the 2nd Biennial Meeting of the iEMSs*, 2004

Maybeck, P., Stochastic models, estimation and control, *Academic Press Inc.*, Vol 1, Chapter 1, 1979

Merz, R., and G. Bloschl, Regionalisation of catchment model parameters, *Journal of Hydrology*, 287, 95-123, 2004

Nash, J.E., and J.V. Sutcliffe, River flow forecasting through conceptual models Part 1 – A discussion of principles, *Journal of Hydrology*, 10(3), 282-290, 1970

Peel, M.C., F.H.S. Chiew, A.W. Western, T.A. McMahon, Extension of monthly unimpaired streamflow data and regionalisation of parameter values to estimate streamflow in ungauged catchments, *Report prepared for the National Land and Water Resources Audit, in Australian Natural Resources Atlas*, http:audit.ea.gov.au/anra/water/docs/national/streamflow/streamflow.pdf

Thyer, M., G. Kuczera, D. Kavetski, and S. Franks, Characterising model error in conceptual rainfall-runoff models using storm-dependent parameters, *In Zerger, A. and Argent, R.M. (eds) MODSIM 2005 International Congress on modelling and simulation, Modelling and Simulation Society of Australia and New Zealand*, 2005

Wagener, T., Evaluation of Catchment Models, *Hydrological Processes*, 17, 3375-3378, 2003