Calibration of environmental models by genetic algorithms

Q.J. Wang, Francis H.S. Chiew and Thomas A. McMahon
Department of Civil and Environmental Engineering, University of Melbourne, Australia

Abstract The parameters in many environmental models are obtained by fitting as closely as possible the model outputs to the measured values. The fitting is often carried out using automatic optimisation techniques. Genetic algorithms are globally oriented, searching over the entire parameter space, and are therefore suitable for solving optimisation problems where the objective function responses contain multiple optima and other irregularities. This paper describes the application of a genetic algorithm to calibrate a ten-parameter conceptual daily rainfall-runoff model (modified version of HYDROLOG). The calibration results presented for two temperate catchments close to the central coast of New South Wales (eastern Australia) highlight the usefulness and limitations of the genetic algorithm. The results indicate that the genetic algorithm is not always robust and cannot always overcome the problems associated with function optimisation. However, it searches more globally compared to most other optimisation techniques and does not require the specification of a starting set of parameter values.

1. INTRODUCTION

Many environmental models have a large number of parameters. Ideally these parameters should be assigned values from direct or indirect field measurements, but in many cases this is not possible. Instead, an inverse problem is solved in which the parameters are calibrated by fitting as closely as possible the model output to the observed. An objective function is constructed as a measure of how close the model output is to the observed, and a search is conducted to find the parameter set which minimises the value of the objective function.

Models with many parameters cannot be easily optimised by standard nonlinear function optimisation techniques. The search can often be fooled into declaring convergence far short of the true optimum because of high dimensionality and irregularities contained in the objective function response such as multiple optima, unsmoothness, discontinuity, elongated ridges and flat plateaus. Some of these difficulties can be overcome by the recently developed genetic algorithms which approach the optimisation problem very differently.

Genetic algorithms are search procedures based on natural selection and genetics, combining an artificial survival of the fittest with genetic operators abstracted from nature (see Holland, 1975; and Goldberg, 1989). Genetic algorithms differ from other search techniques in that they search among a population of points and use probabilistic rather than deterministic transition rules. As a result, genetic algorithms search more globally.

This paper describes a genetic algorithm and its application to calibrating a moderately complex conceptual daily rainfall-runoff model with ten parameters (a modified version of HYDROLOG). The model calibration is carried out by first applying a genetic algorithm with 5000 objective function evaluations followed by a fine tuning using a univariant search technique. Results of the calibration on two catchments close to the central coast of New South Wales (eastern Australia) are presented to highlight the usefulness and limitations of the genetic algorithm.

2. A GENETIC ALGORITHM

Given a function $f = f(x_1, x_2, ..., x_n)$ subject to $a_i \leq x_i \leq b_i$, $i = 1, 2, ... n$, the aim is to find the set of parameter values which gives a minimum value of f. Genetic algorithms work with the coding of the parameters. A method of parameter coding that has often been used is the binary coding. An l-bit binary variable is used to represent one parameter x_i. The integer of the decoded binary variable ranges from 0 to $2^l - 1$ and are mapped linearly to the parameter range $[a_i, b_i]$. Connecting the codings of all parameters forms the coding of the overall parameter set. For example, if there are ten parameters and each parameter is represented by seven binary bits, a point in the search space is then represented by 70 bits.

The genetic algorithm used here has the following steps.

1. Locate m points randomly in the search space. Each point corresponds to one set of parameter values.
2. Find the function value for each point.
3. Rank the points so that their function values are in descending order.
4. Assign a probability value p_j to each point giving higher probability to the point of lower (better) function value. The worst point after ranking is j=1, and its probability value p_1 will be the smallest. The best point is j=m, and its probability value p_m will be the largest. The probability values for other points are linearly interpolated as

$$p_j = p_1 + \frac{p_m - p_1}{m - 1}(j - 1)$$

The sum of probability values for all points is equal to unity and the average of probability values for all points is $1/m$. A value of C/m is assigned to p_m so that the probability value for the best point is C times the average, where $C > 1$. The corresponding probability value for the worst point p_1 is then $(2 - C)/m$. To ensure that all probability values are non-negative, C should be less than or equal to 2.

5. Select two points A and B from these m points at random according to the probability distribution, p_j, $j = 1, 2, ..., m$. Having assigned higher probabilities to better points in the last step, the better points have better chances to be selected.

6. Select two bit positions, k_1 and k_2, along the overall coding of the parameter set at random, giving each bit position the same chance of being selected (if $k_1 > k_2$, their values are interchanged).

7. Form a new point by taking the values of the bits from k_1 to k_2 - 1 of the coding for point A and values of the bits from k_2 to the end and from 1 to k_1 - 1 of the coding for point B.

   For example, if the coding for point A is 110101100100
   and the coding for point B is 101010111001
   and $k_1 = 5$ and $k_2 = 11$,
   the coding for the new point will be 101001100101

8. Occasionally change some of the bits of the newly formed point. A bit value 0 will become 1 and vice versa. This occurs to each bit only at a very small probability of $p_{mutation}$.

9. Repeat steps 5-8 m times so that m new points are produced. The original m points are then replaced by the new ones, forming a new base for further search.

10. Repeat steps 2-9. The best point found so far is always recorded, and in step 2 if the newly generated m points are all inferior to the best point found so far, the latter is re-inserted into the population by replacing randomly one of the m points. Termination of the search is made by specifying a total number of objective function evaluations.

Steps 5 to 7 form the core of the method. Better points have better chances to be chosen to form new points. This is an analogy to the survival of the fittest in the theory of natural selection. The better performing individuals produce more offspring. A new point is formed by taking different blocks of bits from the codings of the two original points. This is an analogy to crossover in the theory of genetics. An offspring takes some blocks of genes from one parent and some from the other. Fit parents are likely to produce fit offspring. The combination of selection and reproduction improves the performance level of the population in the evolution process. The occasional change of bit values in step 8 is an analogy to mutation in the theory of genetics which provides a small probability for background variation.

In summary, unlike standard search techniques, genetic algorithms search among a population of points, work with a coding of the parameter set and use probabilistic transition rules. A population of m points are chosen initially at random in the search space. The objective function values are calculated at all points and compared. From these m points, two points are selected randomly, giving better points higher chances. The selected two points are subsequently used to generate a new point in a certain random manner with occasionally added random disturbance. This is repeated until m new points are generated. The generated population of points are expected to be more concentrated in the vicinity of optimum than the original points. The new population of points are then used to generate another population of points and so on, yielding points increasingly closer to the optimum.

The genetic algorithm used here is only one of the variants of the method. For example, the selection probabilities can be related directly to the function values instead of the ranking method used here. Further details on genetic algorithms can be found in Holland (1975), De Jone (1975) and Goldberg (1989). Nevertheless, Whitely (1989) and Wang (1991) have shown that the search is more robust by employing the ranking method.

For this study, a population size of $m = 100$ is used with 50 population generations (therefore totalling 5000 objective function evaluations) in each run. Each parameter range is represented by a $l = 7$ bit binary therefore discretising the parameter to 128 values. Values of C = 1.5 and $p_{mutation}$ = 0.01 are used.

3. CONCEPTUAL RAINFALL-RUNOFF MODELLING AND FUNCTION OPTIMISATION

Conceptual rainfall-runoff models are used to estimate runoff from rainfall and other climate data (usually used to calculate potential evapotranspiration). The simpler models typically have three or four parameters while the more complex ones can have more than 25 parameters. The parameter values are usually optimised to provide a good fit between the simulated and recorded flows. Key papers on function optimisation methods for calibrating rainfall-runoff models include Ibbitt and O'Donnell (1971), Johnston and Pilgrim (1976), Pickup (1977), Kuczera (1983); Sorooshian and Gupta (1983), Gupta and Sorooshian (1985), Hendrickson et al. (1988), Brazil (1989) and Duan et al. (1992, 1994).

The standard nonlinear function optimisation techniques often behave erratically when used for calibrating rainfall-runoff models. These techniques are often fooled into declaring convergence far short of the true optimum because of high dimensionality and irregularities contained in the objective function response.

Genetic algorithms use a very different approach to most of these techniques. They can be more reliable because they search more globally. Wang (1991, 1995) demonstrated the usefulness of a genetic algorithm in calibrating the Xinanjiang Model (Zhao et al., 1980; and Zhao, 1992) on several catchments in United Kingdom, Japan and Australia.

The daily conceptual model used here is a modified version of HYDROLOG. Like all conceptual rainfall-runoff models, HYDROLOG represents the catchment as a number of interconnected storages, with mathematical functions used to describe the movement of water into, between, and out of them. It attempts to represent the physical processes by using empirical equations and 'effective' parameters to describe the processes. Various versions of HYDROLOG have been used extensively to estimate runoff for a range of purposes, particularly in Australia (see Chiew and McMahon, 1994 and Chiew et al., 1995).

The version used here has 10 parameters. The model structure and the equations representing the processes are illustrated in Figure 1, with the model parameters highlighted in bold. A detailed description of HYDROLOG can be found in Porter and McMahon (1976) and Chiew and McMahon (1994).

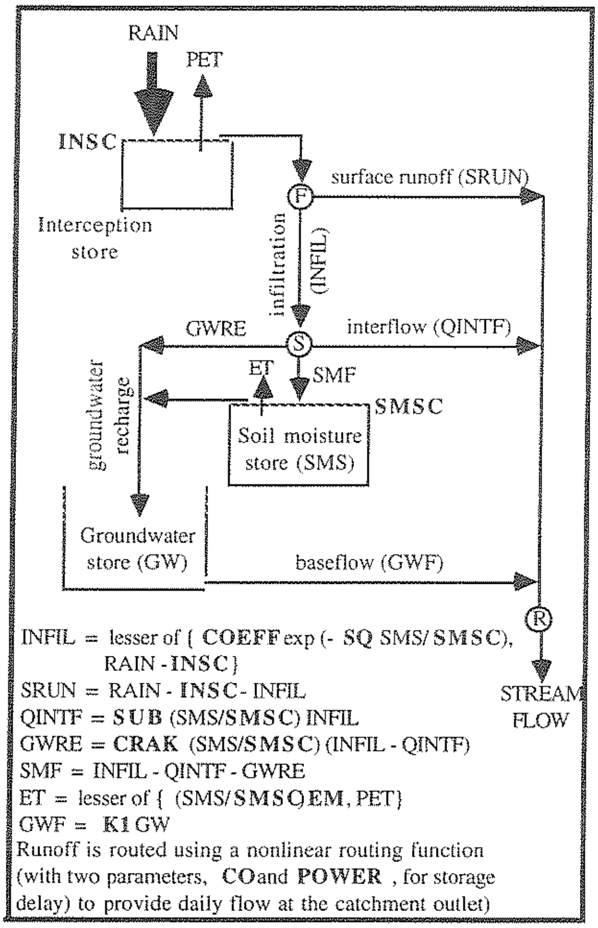

Figure 1 Structure of the modified version of the conceptual daily rainfall-runoff model, HYDROLOG

4. MODEL CALIBRATION AND RESULTS

The genetic algorithm is first used to calibrate HYDROLOG to minimise the value of the objective function, *OBJ*, defined here as the sum of squares of the differences between the daily simulated and recorded flows. When the streamflow volume is expressed in mm depth of water over the catchment area, the objective function has a unit of mm^2.

Ten calibration runs are made with different seeds for the random number generator resulting in different starting population of points and different random numbers used in the search operations. Results from the ten runs can provide an indication of the robustness of the genetic algorithm. The search is limited to 5000 objective function evaluations. The search space is defined by the range of parameter values given in Table 1.

After the 5000 objective function evaluations, a univariant (pattern) search technique (see Hookes and Jeeves, 1961; and Monro, 1971) is used (starting from the set of parameter values giving the lowest objective function value from the genetic algorithm) to improve the value of the objective function. The rationale here is that the genetic algorithm explores the large search space to locate an approximate optimum and the standard search technique then fine tunes locally. The objective function values obtained by using the combination of the genetic algorithm and the univariant search technique are also compared with those obtained by using only the univariant search technique.

The calibration results for two catchments close to the central coast of New South Wales are shown in Tables 2 and 3 to highlight the usefulness and limitations of the genetic algorithm. Table 2 shows the results for Wollombi Brook catchment close to Bulga. It has a drainage area of 1580 km^2 and is covered by forest and grass. It has a temperate climate with an average annual rainfall of 700 mm, 100 mm of which becomes runoff. Five years of data (1963 - 1967) are used. The first 50 days are used as a warming period to initialise the model stores.

Table 3 shows the calibration results for Allyn River catchment at Halton. It has a drainage area of 205 km^2 and is covered by eucalyptus and grasses. The average annual rainfall is 1200 mm, 350 mm of which becomes runoff. Eight years of data (1977 - 1984) are used. The first year is used as a warming period to be consistent with that used in another study on the catchment.

5. DISCUSSION OF RESULTS

The coefficient of efficiency, E, in all the calibration runs for the two catchments is almost 0.9. The coefficient of efficiency is defined as

$$E = \frac{F_o - OBJ}{F_o}$$

$$\text{where, } F_o = \sum_{i=1}^{n} (REC_i - \overline{REC})^2$$

REC_i is daily recorded runoff and $\overline{REC}$ is mean daily recorded runoff.

The coefficient of efficiency expresses the proportion of variance of the recorded flows that can be accounted for by the model (Nash and Sutcliffe, 1970) and provides a direct measure of the ability of the model to reproduce the recorded flows with $E = 1.0$ indicating that all the simulated flows are

Table 1 Ranges of HYDROLOG parameter values defining the search space

	INSC (mm)	COEFF (mm)	SQ	SUB	CRAK	SMSC (mm)	EM (mm)	POWER	CO	K1
Lower limit	1	20	0.1	0	0	20	5	0	0.1	0
Upper limit	5	300	8	1	2	400	20	1	100	0.5

Table 2 Results of calibration of HYDROLOG for Wollombi Brook catchment

	1	2	3	4	5	6	7	8	9	10
INSC (mm)	2.26	1.80	2.08	1.84	1.73	2.54	2.35	1.72	2.73	2.68
COEFF (mm)	195	162	171	177	161	177	206	164	180	223
SQ	6.88	6.63	6.95	6.20	5.10	7.71	6.08	5.12	7.48	6.82
SUB	.045	.013	.032	.000	.000	.000	.026	.000	.000	.067
CRAK	.987	1.07	1.09	.937	.823	1.36	.879	.806	1.35	.930
SMSC (mm)	398	400	400	374	309	400	336	314	383	367
EM (mm)	10.9	13.9	13.1	11.5	10.8	13.3	8.84	10.5	12.4	8.90
POWER	.000	.039	.055	.055	.048	.265	.081	.053	.355	.000
CO	10.3	7.99	7.12	7.19	7.51	1.67	6.22	7.34	.887	10.7
K1	.160	.177	.165	.175	.182	.159	.150	.176	.167	.134
OBJ (mm^2)	731 (754)	734 (802)	734 (794)	736 (774)	740 (822)	743 (855)	736 (788)	740 (796)	747 (768)	732 (819)

Each run consist of a search using the genetic algorithm with 5000 objective function evaluations followed by fine tuning using a univariant search. The OBJ values in brackets are objective function values reached using only the genetic algorithm.

Table 3 Results of calibration of HYDROLOG for Allyn River catchment

	1	2	3	4	5	6	7	8	9	10
INSC (mm)	5.00	1.00	1.00	5.00	5.00	5.00	1.00	2.01	1.81	5.00
COEFF (mm)	291	207	214	244	234	255	96.8	291	300	300
SQ	.349	2.90	3.45	.677	.162	.434	2.96	.473	.224	.598
SUB	.732	.103	.102	.729	.746	.748	.160	.904	.900	.752
CRAK	1.98	.279	.294	1.90	1.93	1.91	.462	2.00	2.00	2.00
SMSC (mm)	52.6	393	400	50.0	49.1	48.2	320	127	128	94.4
EM (mm)	5.77	5.43	6.83	5.00	5.00	5.00	15.3	15.1	17.6	5.00
POWER	.000	.842	.844	.000	.000	.000	.837	.000	.000	.000
CO	.100	.100	.100	.100	.100	.100	.100	.100	.100	.100
K1	.080	.195	.262	.067	.065	.063	.257	.036	.040	.074
OBJ (mm^2)	12100 (13600)	14800 (16800)	14700 (15200)	12000 (14000)	12000 (14100)	12000 (15700)	14900 (15000)	14000 (15200)	14000 (16000)	13200 (14700)

Each run consist of a search using the genetic algorithm with 5000 objective function evaluations followed by fine tuning using a univariant search. The OBJ values in brackets are objective function values reached using only the genetic algorithm.

the same as the recorded flows. Like the objective function, it is biased towards assessing the simulation of higher flows. The high values of E indicate that all calibration runs resulted in satisfactory simulations of daily runoff, particularly the higher flows.

Inspection of the parameter values in Table 2 for the Wollombi Brook catchment shows that runs 6 and 9 reached the vicinity of one optimum while the other runs reached a different and slightly better optimum in the parameter space. However, the lowest and highest objective function values differ by only two percent. In the context of rainfall-runoff modelling and in estimating daily flows, this difference is negligible. From the optimisation point of view, it is difficult for any search technique to discriminate the two optima without an exhaustive search. Further extensive search has not revealed any substantially better point indicating that the objective function values achieved in the ten runs are close or equal to the global minimum.

It may be noted from Table 2 that the fine tuning using the univariant optimisation technique improved the value of the objective function obtained using only the genetic algorithm by up to 15%.

The model calibration was also carried out using only the univariant optimisation technique. This resulted in an objective function value of 736 mm^2, which is practically the same as the values obtained using the genetic algorithm with fine tuning. It should be noted, however, that calibration using the univariant technique (and most nonlinear optimisation methods) are very dependent on the starting set of parameter values, and a poor choice of starting values can lead to poor model calibration due to the problems discussed earlier. The starting point used in this case has been chosen based on the authors' extensive experience with the HYDROLOG model. On the other hand, the genetic algorithm searches over the entire parameter space, and the user does not need to specify starting parameter values.

Unlike the Wollombi Brook catchment, the calibration runs for Allyn River catchment reached several different optima. Table 3 indicates that runs 1, 4, 5 and 6 reached the vicinity of one optimum and gave the best model calibration, while runs 2 and 3 reached another optimum and gave the poorest calibration. Runs 8 and 9 reached the vicinity of yet another optimum while runs 7 and 10 reached another two different optima. The lowest and highest objective function values differ by almost 25%.

The significantly different optima obtained with the ten calibration runs for the Allyn River catchment indicate that the genetic algorithm is not always robust. Although the genetic algorithm searches more globally and therefore has less chance of being trapped in local optima compared to most nonlinear optimisation techniques, it cannot overcome all the problems associated with function optimisation.

It may be noted from Table 3 that the fine tuning using the univariant optimisation technique improved the value of the objective function obtained using only the genetic algorithm by up to 30%.

The model calibration using only the univariant optimisation technique resulted in an objective function value of 15000 mm^2. This value is higher than those obtained from all the ten runs using the genetic algorithm with fine tuning (more than 25% higher in four cases). There is therefore merit in using the genetic algorithm to search globally for parameter sets close to the optimum.

6. SUMMARY AND CONCLUSIONS

A genetic algorithm is applied to calibrate a ten-parameter conceptual daily rainfall-runoff model (modified version of HYDROLOG). The calibration results for two temperate catchments close to the central coast of New South Wales are presented.

Calibrations on the Wollombi Brook catchment indicate that the genetic algorithm is robust with the ten calibration runs giving practically the same objective function values. However, the ten calibration runs on the Allyn River catchment resulted in significantly different optima, suggesting that the genetic algorithm cannot always overcome the problems associated with function optimisation. Nevertheless, all the calibration runs with the genetic algorithm led to lower objective function values compared to the value obtained using only the univariant search technique. In addition, unlike most standard nonlinear optimisation methods where the calibration results are dependent on the starting set of parameter values, the genetic algorithm does not require specification of starting parameter values.

In summary, genetic algorithms search globally over the entire parameter space and are therefore suited to solving optimisation problems where the objective function responses contain multiple optima and other irregularities. They are useful in calibrating environmental models, particularly when a standard optimisation technique is further used to fine tune around the optimum located by the genetic algorithm.

7. REFERENCES

Brazil, L.E., *Multilevel Calibration Strategy for Complex Hydrologic Simulation Models*, National Oceanic and Atmospheric Administration, National Weather Service, US Department of Commerce, NOAA, Silver Spring, MD, Technical Report NWS 43, 1989.

Chiew, F.H.S. and McMahon, T.A., Application of the daily rainfall-runoff model MODHYDROLOG to twenty eight Australian catchments, *J. Hydrol., 153*, 383-416, 1994

Chiew, F.H.S., Whetton, P.H., McMahon, T.A. and Pittock, A.B., Simulation of the impacts of climate change on runoff and soil moisture in Australian catchments, *J. Hydrol., 167*, 121-147, 1995.

De Jone, K.A., *An Analysis of the Behaviour of a Class of Genetic Adaptive Systems*, Doctoral Dissertation, University of Michigan, 1975.

Duan, Q., Sorooshian S. and Gupta, V.K., Effective and efficient global optimisation for conceptual rainfall-runoff models, *Wat. Resour. Res., 28*, 1015-1031, 1992.

Duan Q., Sorooshian, S. and Gupta, V.K., Optimal use of the SCE-UA global optimisation method for calibrating watershed models, *J. Hydrol., 158*, 265-284, 1994.

Goldberg, D.E., *Genetic Algorithms in Search, Optimisation and Machine Learning*, Addison-Wesley, Reading, 412 pp., 1989.

Gupta, V.K. and Sorooshian, S., The automatic calibration of conceptual catchment models using derivative-based optimisation algorithms, *Wat. Resour. Res., 21*, 437-485, 1985.

Hendrickson, J.D., Sorooshian, S. and Brazil, L.E., Comparison of Newton-type and direct search algorithms for calibration of conceptual rainfall-runoff models, *Wat. Resour. Res., 24*, 691-700, 1988.

Holland, J.H., *Adaptation in Natural and Artificial Systems*, University of Michigan Press, 183 pp., 1975.

Hookes, R. and Jeeves, T.A., 'Direct Search' solution of numerical and statistical problems, *J. Assoc. Comput. Mach., 8*, 212-229, 1961.

Ibbitt, R.P. and O'Donnell, T., Designing conceptual catchment models for automatic fitting methods, Proceedings of *International Symposium on Mathematical Models in Hydrology*, IAHS, Warsaw, 1971.

Johnston, P.R. and Pilgrim D.H., Parameter optimisation for watershed models, *Wat. Resour. Res., 12*, 477-486, 1976.

Kuczera, G., Improved parameter inference in catchment models I, Evaluating parameter uncertainty, *Wat. Resour. Res., 19*, 1151-1162, 1983.

Monro, J.C., *Direct Search Optimisation in Mathematical Modelling and a Watershed Model Application*, National Oceanic and Atmospheric Administration, National Weather Service, US Department of Commerce, NOAA, Silver Spring, MD, Technical Memo NWS HYDRO-12, 52 pp., 1971.

Nash, J.E., The form of the instantaneous unit hydrograph, *IAHS-AISH, 42*, 114-118, 1958.

Nash, J.E. and Sutcliffe, J.V., River forecasting through conceptual models, 1. A discussion of principles, *J. Hydrol., 10*, 282-290, 1970.

Pickup, G., Testing the efficiencies of algorithms and strategies for automatic calibration of rainfall-runoff models, *Hydrol. Sci. Bull., 22*, 257-274, 1977.

Porter, J.W. and McMahon, T.A., *The Monash Model: User Manual for the Daily Program HYDROLOG*, Department of Civil Engineering, Monash University, Victoria, Australia, Research Report 2/76, 41 pp., 1976.

Sorooshian, S. and Gupta, V.K., Automatic calibration of conceptual rainfall-runoff models: the question of parameter observability and uniqueness, *Wat. Resour. Res.*, 19, 251-259, 1983.

Wang, Q.J., The genetic algorithm and its application to calibrating conceptual rainfall-runoff models, *Wat. Resour. Res., 27*, 2467-2471, 1991.

Wang, Q.J., Using genetic algorithms to solve inverse problems, Paper submitted to *Engineering Optimisation*, 1995.

Whitley, D., The GENITOR algorithm and selective pressure: why rank-based allocation of reproductive trials is best, Proceedings of the *Third International Conference on Genetic Algorithms*, Morgan Kaufmann, 116-121, 1989.

Zhao, R.J., The Xinanjiang model applied in China, *J. Hydrol.*, 135, 371-381, 1992.

Zhao, R.J, Zhuang, Y.L., Fang, L.R., Liu, X.R. and Zhang, Q.S., The Xinanjiang model, Proceedings of the Oxford Symposium, April, 1980, *IAHS-AISH, 129*, 351-356, 1980.