Evaluation of spatially variable control parameters in a complex catchment modelling system: a genetic algorithm application

Tianjun Fang and James E. Ball

ABSTRACT

Successful implementation of a catchment modelling system requires careful consideration of the system calibration which involves evaluation of many spatially and temporally variable control parameters. Evaluation of spatially variable control parameters has been an issue of increasing concern arising from an increased awareness of the inappropriateness of assuming catchment averaged values. Presented herein is the application of **a** real-value coding genetic algorithm (GA) for evaluation of spatially variable control parameters for implementation with the Storm Water Management Model (SWMM). It was found that a real-value coding GA using multiple storms calibration was a robust search technique that was capable of identifying the most promising range of values for spatially variable control parameters. As the selection of appropriate GA operators is an important aspect of the GA efficiency, a comprehensive investigation of the GA operators in a high-dimensional search space was conducted. It was found that a uniform crossover operation was superior to both one-point and two-point crossover operations over the whole range of crossover probabilities, and the optimal uniform crossover and mutation probabilities for the complex system considered were in the range of 0.75–0.90 and 0.01–0.1, respectively.

Key words | calibration, catchment modelling, genetic algorithm, hydroinformatics, urban

Tianjun Fang
Water Research Laboratory,
School of Civil and Environmental Engineering,
The University of New South Wales,
Sydney, NSW 2052,
Australia

James E. Ball (corresponding author)
Faculty of Engineering,
University of Technology Sydney,
Broadway, NSW 2007,
Australia
E-mail: *james.ball@uts.edu.au*

INTRODUCTION

Genetic algorithms (GAs) based on the mechanics of natural selection have been applied to a variety of problems in the field of water resources. Several studies have focused on developing methodologies for applying GAs to water distribution pipe networks (see, for example, Murphy *et al.* 1993; Simpson *et al.* 1994; Dandy *et al.* 1996), while Ritzel *et al.* (1994) used GAs to solve a multi-objective groundwater pollution containment problem. In addition, there have been several applications of GAs to reservoir systems (Oliveira & Loucks 1997; Wardlaw & Sharif 1999).

Genetic algorithms have been applied also for the calibration of catchment modelling systems (Wang 1991; Franchini 1996; Liong *et al.* 1995). Wang (1991) and Franchini (1996) used lumped conceptual rainfall–runoff models where the parameters were assumed to have average values across the catchment. The application of a GA in the physically distributed storm water management model (SWMM) was presented by Liong *et al.* (1995). Though SWMM is associated with a large number of spatially variable parameters for the description of subcatchment characteristics, Liong *et al.* (1995) only considered eight control parameters where each control parameter was set as a factor based on the default values across subcatchments. As a result, the study of Liong *et al.* can be considered as using a lumped approach whereby the control parameters for each subcatchment

doi: 10.2166/hydro.2007.026

were adjusted by the same percentage during the calibration process.

Previous studies of the application of GAs for calibrating catchment modelling systems concentrated on identifying a unique optimal or near-optimal parameter set (Wang 1991; Franchini 1996; Liong *et al.* 1995). As a result of the need for physically distributed catchment modelling systems, such as SWMM, the calibration process is associated with the evaluation of a large number of spatially variable control parameters and hence the search for an optimal or near-optimal set of control parameters in a high-dimensional search space. Many studies, however, have demonstrated the difficulties associated with high-dimensional search spaces due to the uncertainty of model system structure, errors associated with the system input data and the observed data, and interactions between control parameters (see, for example, Kuczera 1983; Sorooshian *et al.* 1983; Beven & Binley 1992).

These difficulties have resulted in the development of the concept of "equifinality". This concept recognises that alternative behavioural sets of control parameters with a catchment modelling system are capable of producing reasonable estimates of the catchment response as measured by objective functions defining the goodness of fit (Beven & Binley 1992).

The potential of multiple control parameter sets which result in similar system performance increases as the number of parameters increases. At the same time, an increase in the number of control parameters introduces additional difficulties in the search for behavioural sets of control parameters. In an attempt to address this issue, Spear *et al.* (1994) developed a tree-structured density estimation technique to identify small densely populated regions within the parameter space where the desired performance could be achieved. However, this approach is based on each control parameter being independent, which is not a valid assumption for many physically distributed catchment modelling systems. Therefore, the tree-structured density estimation approach is not applicable or practical for a complex catchment modelling system with a large number of spatially variable control parameters.

Presented herein are the results obtained from an investigation into the use of a real-value coding GA with multiple storm events for the seeking of the behavioural sets of control parameter values associated with a physically distributed catchment modelling system. While the main purpose of the study was to identify behavioural sets of control parameters, and hence the most promising ranges for spatially variable control parameters in a complex multi-dimensional system, a comprehensive evaluation of the GA operators for this complex multi-dimensional system was conducted also as the selection of appropriate operators is an important aspect of the efficiency of a GA.

GENETIC ALGORITHMS

Concepts

Inspired by the principles of natural evolution, Holland (1975) first developed a search technique based on random generated points that he referred to as a genetic algorithm. Genetic algorithms have a remarkable capability to keep a balance between exploration and exploitation of the search space (Michalewicz 1994). Genetic algorithms initially explore an entire search space randomly to find a promising region. As long as a promising region is found, GAs exploit the best solutions in the promising region while continuing to explore for other promising regions. The application of GAs does not guarantee delineation of the optimal solution, but does guarantee a high probability of finding the better solutions.

A comparison between GAs and other optimisation methods was presented in Goldberg (1989). The major difference between GAs and other standard search techniques is that GAs deal with a population of possible solutions rather than a single solution. A possible solution to a problem is represented by a suitably encoded string of model control parameters, called a chromosome. A collection of chromosomes is called a population. The main issues associated with the application of GAs consist of representation, or how the control parameters are represented in a chromosome, selection, or how individual chromosomes are selected for future usage, crossover, and mutation, or how a chromosome may be adjusted to ensure the search space is investigated fully.

Representation

The first process in the operation of GAs is to design a suitable coding (or representation) for the control parameter values. Binary coding is the most common encoding technique and requires parameters to be encoded into a string of binary bits. In the application of GAs to calibrate catchment modelling systems, the studies undertaken by Wang (1991) and Franchini (1996) both used binary coding.

Gray coding uses adjacent variable values where the code occurs as only one binary digit. It was developed to overcome a problem called "Hamming Cliffs" that exists in binary coding and has been used in a number of studies in the water resources field (see, for example, Dandy *et al.* 1996; Ng & Perera 2003).

Recently, there has been growing interest in real-value coding for GAs. In real-value coding, each chromosome is coded as a vector of real numbers with the same length for the parameter vector. Usually the real numbers are constrained within a predetermined range.

Advantages of real-value coding have been described by, for example, Janikow & Michalewicz (1991), Michalewicz (1994) and Yoon & Shoemaker (1999), and can be summarised as:

- Real-value coding is capable of representing quite large search domains. Other encoding methods require prohibitively long representation dealing with large domains.
- Real-value coding has the representation space approximately equivalent to the problem space so that operators can be easily and efficiently implemented.
- No encoding or decoding of the parameter space is required for real-value coding.
- Real-value coding can be more consistent, and achieve higher performance in terms of speed and accuracy.

Wardlaw & Sharif (1999) investigated the performance of a GA with different operators and encoding methods for optimal reservoir system operation. They found that the real-value coding clearly outperformed binary and gray coding over a wide range of crossover probabilities. Real-value coding has been used successfully by Oliveira & Loucks (1997) and Jain *et al.* (2005).

It is worth noting that a risk associated with the application of GAs is that strong chromosomes may dominate the population. A real-value coding GA can dramatically mitigate this risk by the following procedures:

- a large number of initial population size generated by the GA, which uniformly distribute within the initial parameter space, and
- multi-points crossover and adaptable population size during the GA process.

Selection schedule

Selection is a process that determines the number of chromosomes for participation in the reproduction process to form a new generation. The most common selection method is the fitness-proportionate method that selects chromosomes in proportion to their fitness, relative to the average fitness of the whole population. This selection might permit extremely fit individuals to take over the population and lead to a loss of population diversity and premature convergence.

An alternative to fitness-proportionate selection is tournament selection. In tournament selection, chromosomes are chosen randomly from the population. The fittest chromosome from a comparison of their fitness is then selected to participate in the reproduction process. The procedure is repeated until the desired number of chromosomes are selected for reproduction. Of the alternative tournament selection methods, binary tournament selection is the most commonly used method.

Goldberg & Deb (1991) compared different selection methods including proportionate selection, fitness ranking, tournament selection and steady state selection. They concluded that no one selection method was superior to the others. Falkenauer (1998), however, found that tournament selection has the advantage of maintaining adequate selection pressure while avoiding the drawbacks associated with fitness-proportionate and ranking methods.

Crossover and mutation

Crossover is a process where two chromosomes from the parent population exchange segments to produce offspring chromosomes, with each offspring inheriting some of the characteristics of each parent. Typically, a crossover operator will maintain the "good" components of the parent chromosomes (i.e. common components) while exploring new regions through exchange of non-common components.

Within the crossover operation, the number of crossover points determines how many segments of a chromosome are exchanged. Generally one-point, two-point or uniform crossover points are used. Uniform crossover operates on the individual genes of the two parent chromosomes rather than swapping the segments of the two parent chromosomes. The performance of uniform crossover has been investigated by Spears & De Jong (1991) and Qi & Palmieri (1993). Spears & De Jong (1991) presented a positive view of uniform crossover points as a result of its recombination potential and its exploration power. However, for binary and gray coding, the use of a uniform crossover point has the potential to break a good pattern in the parent's chromosomes. This is not an issue for real-value coding GAs since each gene represents the normalized value of a control parameter.

Mutation occurs by changing the value of one or more randomly selected genes. For binary and gray coding systems, the value of a chosen gene is changed by flipping the binary numbers. In real-value coding system, the value of a selected gene is adjusted within a designed range. Michalewicz (1994) proposed two mutation methods for real-value coding GAs, namely uniform and non-uniform mutation. In the application of uniform mutation, the selected gene is substituted by a random value generated within the corresponding range. For application of a non-uniform mutation, randomly selected genes within a chromosome can be modified by an amount which decreases as the number of generations progresses. Michalewicz (1994) argued that non-uniform mutation represents the fine-tuning capability of a GA.

Population size

Finally, the population size plays an important role in the application of a GA. If the population size is too small, there is an increased risk of premature convergence due to the insufficient initial search space. If the population size is too large, the efficiency of the GA is inhibited by exploration of too large a search space while, if the population is too small, the efficiency of a GA is inhibited by constraints on the search space.

EQUIFINALITY IN A COMPLEX SYSTEM

The effects of spatially variable control parameters on catchment responses have been examined previously by, for example, Freeze (1980) and Beven (1989). It has been found that the search for a unique optimal parameter set is neither realistic nor applicable in physically distributed catchment modelling systems due to uncertainty in the model process, interactions between parameters, as well as errors in input and observed variables. A single parameter set is not capable of reproducing the heterogeneity of responses generated by spatially variable parameters representing the characteristics of catchments (Beven 1989).

The concept of equifinality, which was introduced by Beven & Binley (1992), recognises that a group of "equivalent" control parameter sets within a modelling system can result in a similar level of performance where the level of performance is determined from the differences between the catchment responses and the observable variables. Similar concepts have appeared as "behavioural" parameter sets (Spear *et al.* 1994) or "acceptable" parameter sets (Klepper & Rouse 1991). Following this concept, a set of control parameter values can be classed as behavioural if the corresponding performance meets a specific threshold criterion.

In this study, the performance of the alternative sets of control parameters was evaluated by using the Root Mean Square Error (RMSE) as the fitness measure with those sets where the RMSE $< 10\%$ being classified as behavioural sets. The RMSE can be expressed as

$$\mathrm{RMSE} = \mathrm{SQRT}\left[\left(\sum_{i=1}^{n}(Q_{\mathrm{ti}} - Q_{\mathrm{si}})^2\right)/n\right] \tag{1}$$

where Q_{ti} and Q_{si} are observed and simulated discharges respectively, and n is the number of observations in the time series.

The equifinality concept has been applied to a wide range of problems inclusive of the calibration of rainfall–runoff models (Beven 1993; Freer *et al.* 1996; Beven & Freer

2001; McMichael *et al.* 2006), flood frequency estimation (Romanowicz & Beven 1998; Cameron *et al.* 2001), calibration of land surface models (McCabe *et al.* 2005) and the assessment of land use change effects (Eckhardt *et al.* 2003). In these studies, the initial sets of control parameters were generated from uniform random samples. It has been demonstrated in these studies that there is a difficulty in obtaining sufficient behavioural sets of control parameters based on a large number of simulations, e.g. Spear *et al.* (1994) developed only 20 behavioural sets from 2.6 million simulations. As a result of this, Spear *et al.* (1994) developed a tree-structured density estimation technique to identify small densely populated regions within the control parameter space. This tree-structured density estimation technique was developed to provide a guide for selection of the ranges of control parameter values, which could dramatically increase the likelihood of finding behavioural control parameter sets. This technique is not applicable or practical for complex catchment modelling systems where strong interaction exists among a large number of spatially variable control parameters. To mitigate this problem, a real-value coding GA was used herein to identify the most promising ranges of control parameter values in a high-dimensional search space.

METHODOLOGY DEVELOPMENT

Storm Water Management Model (SWMM)

The US EPA's SWMM is a physically distributed catchment modelling system (Huber & Dickinson 1988). Typically, there are 10 control parameters associated with the RUNOFF block for each subcatchment, namely the subcatchment width (W_s), impervious percentage (%imp) and the percentage of impervious area with zero depression storage, depression storage of the impervious (d_i) and pervious areas (d_p), Manning's roughness for impervious (n_i) and pervious areas (n_p) and three infiltration parameters for Horton's infiltration Equation (F_0, F_u and F_k). Following the concept of Dayarante & Perera (1999), Fang & Ball (2005) considered two classifications of storm events: these classifications were those events that would be expected to produce runoff only from impervious areas and those storm events that may result in runoff from both the pervious and impervious areas. Considering only those events likely to result in runoff from the impervious areas of the catchment, Fang & Ball (2005) determined that the four sensitive parameters were impervious percentage, subcatchment width, impervious Manning's roughness and depression storage of impervious area. This study considered only these four sensitive parameters for each subcatchment.

GA coupled with SWMM

The approach of a real-value coding GA coupled with SWMM was used. The GA code developed by Anderson (1995) was modified for this purpose. The main features of the modified GA are described below:

- real-value coding;
- binary tournament selection with complete replacement for chromosomes;
- adaptable variable population size; and
- crossover (without elitism) operations with uniform mutation at specified probabilities

The linkage of the GA with SWMM is shown in Figure 1. As shown in this figure, the initial steps in the procedures are:

- The GA generates the initial chromosomes x_1, x_2, $x_3, \ldots, x_n$, which are uniformly distributed in an initial search space θ.
- A chromosome x_i consists of 168 genes, where each gene represents a control parameter value for a subcatchment.
- Within each gene, the control parameter value is normalised to a range of [0, 1].

If the initial range of a parameter is set as the feasible range, the behavioural sets of control parameter values should be distributed within this range. There is no knowledge or skills to guarantee that the initial ranges of spatially variable control parameters are appropriate. If the initial range of a parameter is larger than its feasible range, it is expected that most of the behavioural parameters are located in a smaller range than the initial range.

There is a potential that a behavioural set identified for one storm sequence may not be a behavioural set for

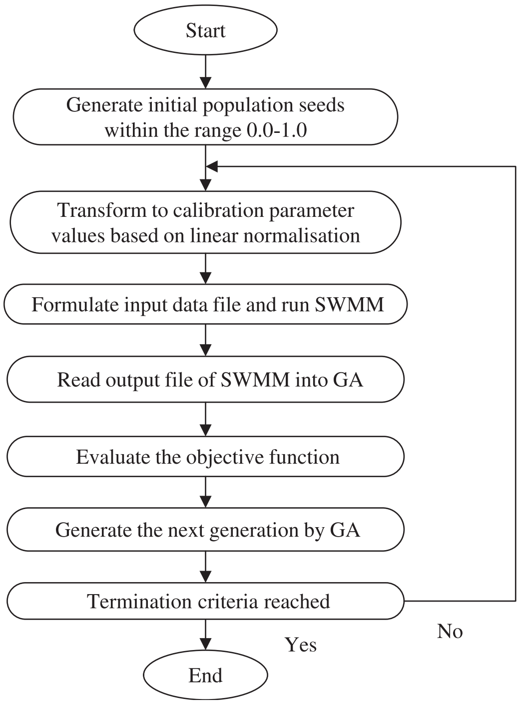

Figure 1 | The process of GA_SWMM.

Figure 2 | The subcatchment boundaries of CPC.

alternative storm sequences. Therefore, three storm events representing different storm characteristics were calibrated in sequence. The feasible ranges of a control parameter for these three storm events were identified by the GA as θ_1, θ_2 and θ_3. The most promising range of this control parameter therefore is represented as

$$\hat{\theta} = \theta_1 \cap \theta_2 \cap \theta_3. \quad (2)$$

THE STUDY CATCHMENT AND DATA

The Centennial Park Catchment (CPC) is located in the eastern suburbs of Sydney, Australia. This urbanised catchment has an area of 132.7 ha and is served by separate drainage systems. The length of stormwater drain is approximately 5.2 km from the upstream portions of the catchment to the gauging station located at the outlet of Musgrave Avenue Stormwater Channel into the Centennial Park pond system. The geological composition of the catchment is Botany sands that consist of Hammondville Soil (85%) and Moore Soil (15%). The catchment is divided into 42 subcatchments with the size of each varying from 0.91 ha to 26.5 ha, based on land use, conduit diameter and topographic characteristics, as shown in Figure 2.

The initial values of subcatchment width were determined by dividing each subcatchment area by the maximum overland flow length of this subcatchment while the initial values of the impervious percentage and other control parameters of the subcatchments were based on Abustan's (1997) study. These initial ranges for the control parameters are presented in Table 1.

As mentioned previously, three storm events with single and multiple peaks were used. Details of these storm events are shown in Table 2. The antecedent wetness of the catchment was categorised based on Abustan's (1997) study.

Table 1 | The initial ranges of calibration parameters

Parameter	W_s	%imp	n_i	d_i
Lower	−30%	−20%	0.01	0.0
Upper	+30%	+20%	0.03	0.5

Table 2 | Characteristics of the storm events

Events	Rainfall (mm)	Peak flow (m^3/s)	Duration (min)	Catch_wetness
5 Jan. 1998	13.6	3.095	150	Dry
14 Dec. 1998	9.0	3.735	80	Dry
24 Feb. 1999	21.4	2.196	230	Wet

EVALUATION OF THE GA OPERATORS

Knowledge about the proper selection of GA operators is only fragmentary and based on empirical studies (Michalewicz 1994). It has not demonstrated that a single global GA operator set can solve effectively a variety of optimisation problems. Three GA operators, namely population size, crossover and mutation operators, were investigated in this study. The storm event occurring on 5 January 1998 was used for this purpose.

Population size

Population sizes of 300, 500, 1000, 2000 and 5000 chromosomes were investigated. The other GA operators used were a two-point crossover with a probability of 0.8 and a uniform mutation at a probability of 0.01. The relationship between the average RMSE for the best 20 chromosomes and the number of generations is shown in Figure 3.

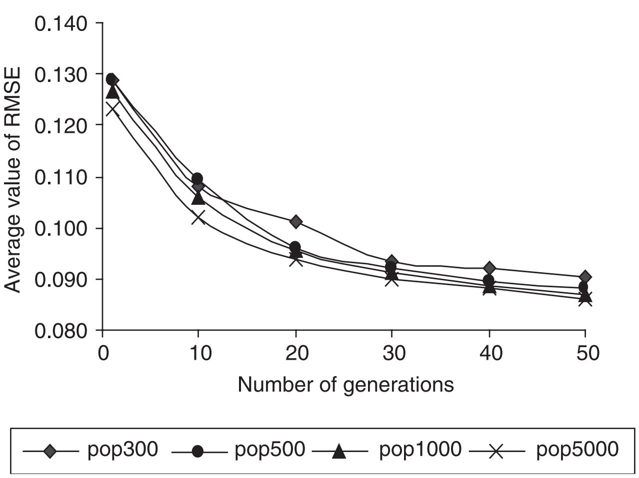

Figure 3 | Performance based on different population sizes.

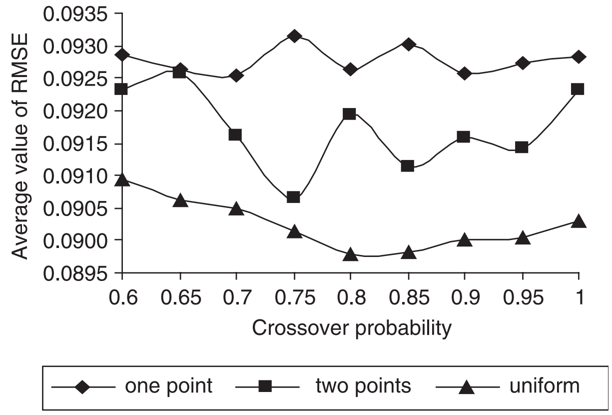

Figure 4 | Performance of three types of crossover.

In the initial population, the average value of RMSE for the best 20 parameter sets was getting smaller as the population size increased. However, the average value of RMSE for the best 20 parameter sets between the population size of 1000, 2000 and 5000 did not show much difference after 50 generations. Therefore, a population size of 1000 was chosen in this study.

Crossover and mutation operators

The performance of three categories of crossover operation was examined, with these categories being one-point, two-point and uniform crossover. Crossover probabilities varying from 0.6–1.0 at an interval of 0.05 were considered also. The population size and the mutation probability were set at 1000 and 0.01, respectively. The average values of the

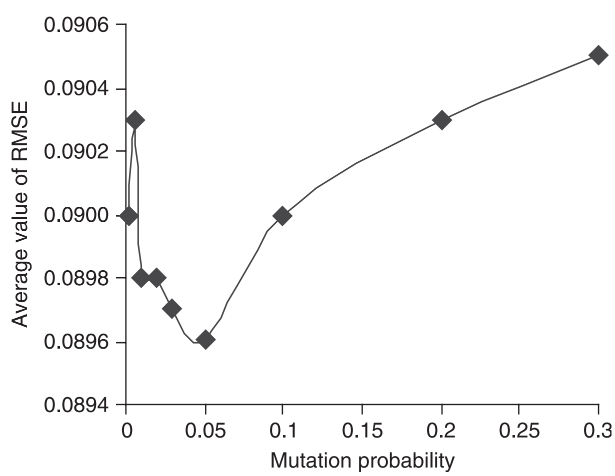

Figure 5 | Performance of uniform mutation.

Table 3 | Performance of behavioural sets of control parameter values

Storm event	5 Jan. 1998	14 Dec. 1998	24 Feb. 1999
Average RMSE	0.0783	0.0880	0.0715
Standard deviation of RMSE	0.0012	0.0010	0.0014

RMSE for the best 20 parameter sets after 50 generations are shown in Figure 4. As shown in Figure 4, a uniform crossover operation outperformed the other two types of crossover operation across the whole range of crossover probabilities. The best performance of the uniform crossover occurred at a probability in the range of 0.75–0.9. Hence, probabilities for uniform crossover applied with problems of this type should be set within this range.

The performance of the mutation operator as part of a uniform crossover was investigated. Mutation probabilities were considered within the range 0.001–0.3 at varying intervals (smaller intervals for the low range), while crossover probabilities were fixed at 0.80. As shown in Figure 5, it was found that the best mutation probability was in the range of 0.01–0.1.

CASE STUDY

The real-value coding GA approach based on the use of multiple storms calibration was used to identify the most promising ranges for spatially variable control parameters associated with a catchment modelling system applied to an urban catchment. The study catchment was the Centennial Park catchment described earlier, while the details of the GA operators were:

- a population size of 1000;
- a uniform crossover with a crossover probability 0.8; and
- a uniform mutation with probability 0.01.

The cumulative number of "behavioural" sets of control parameter values for the three storm events (5 January 1998, 14 December 1998 and 24 February 1999) after 50 generations were 898, 896 and 897, respectively; as noted earlier, a set of control parameter values was considered to be "behavioural" if the RMSE $< 10\%$ during any of the 50 generations. Shown in Table 3 are the characteristics of the optimisation objective function (i.e. the RMSE) for the behavioural sets of control parameters.

Frequency distributions of the normalised values of the width and the impervious percentage of subcatchment 3 for the 898 "behavioural" parameter sets obtained for the storm event of 5 January 1998 are illustrated in Figure 6. As shown in Figure 6(a), the normalised values of the width for subcatchment 3 for the "behavioural" sets were in the range of 0.79–1.0. The most likely values were in the range 0.95–1.0, which covered 69.3% of the behavioural sets. There was only one outlier and its normalised value was 0.325 (not shown in Figure 6 (a)).

The normalised values of the impervious percentage of subcatchment 3 for all "behavioural" parameter sets were

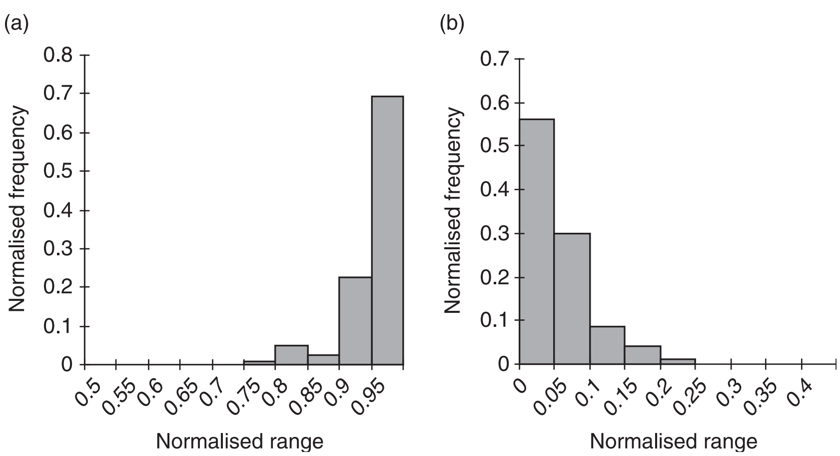

Figure 6 | The frequency distribution plots of the normalised values of width (a) and impervious percentage (b) of subcatchment three for 898 "behavioural" parameter sets by calibrating the storm event of 5 January 1998.

within the range 0.0–0.26, as shown in Figure 6 (b). As shown in this figure, 56.2% of behavioural sets were in the range of 0.0–0.05. Interpretation of this result strongly suggests that a real-value coding GA is capable of identifying the feasible range of a parameter in a high-dimensional search space.

The most promising ranges of subcathment width, impervious percentage and impervious Manning's roughness for each subcatchment identified with the multiple storm calibration are shown in Figures 7–9. From consideration of Figure 7, it can be seen that the most promising ranges of the subcatchment width for eight subcatchments were in the range 0.0–1.0, which suggests that the initial ranges of these parameters were appropriate. Sixteen of the subcatchment widths were dramatically narrowed to a range of 0.5–1.0. For example, the most promising range of the width of subcatchment 3 was within

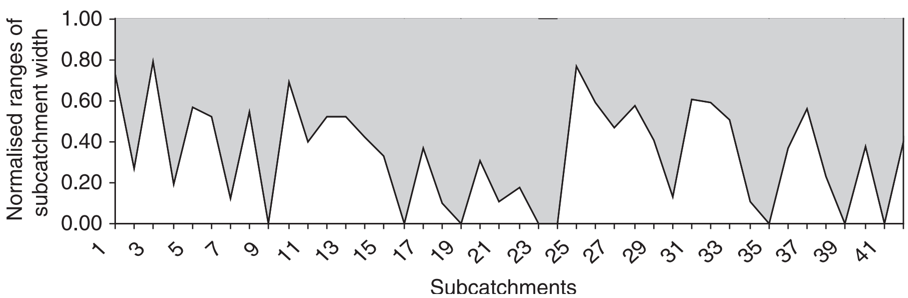

Figure 7 | The most promising ranges of subcatchment width.

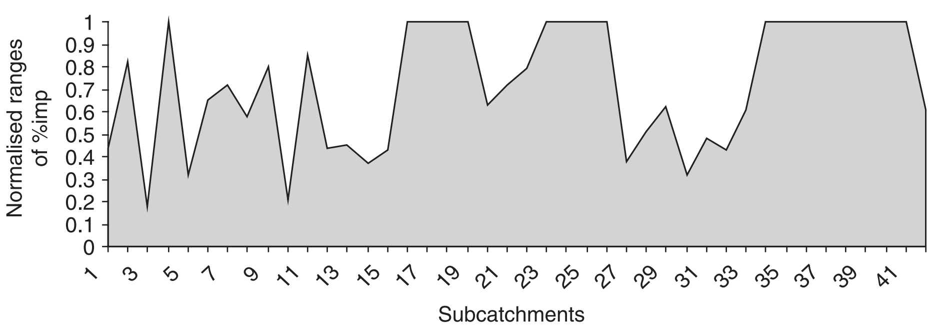

Figure 8 | The most promising ranges of impervious percentage.

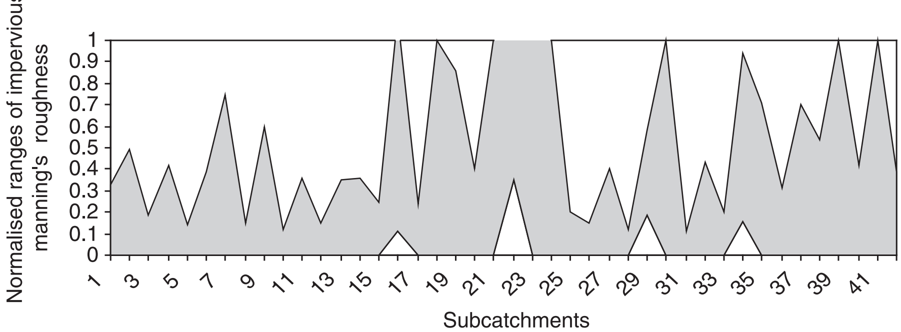

Figure 9 | The most promising ranges of impervious Manning's *n*.

the range of 0.79–1.0, which indicates that the most promising range of this parameter has been decreased to a range of 121–135 m from the initial range of 73–135 m.

As shown in Figure 8, the initial ranges of the imperviousness parameter for 17 subcatchments were appropriate. For another 13 subcatchments, however, the most promising range was dramatically decreased to a range of 0.0–0.5. In particular, the most promising range of the impervious percentage for subcatchment 3 was in the range of 0.0–0.26; this resulted in a decrease in the range of values from a range of 52–78% to a range of 52–59%.

In a similar manner, from Figure 9, it can be seen that the most promising range of the impervious Manning's roughness for seven subcatchments was in the range 0.0–1.0, which indicated that these parameters for all "behavioural" parameter sets were distributed within the initial range. For 23 subcatchments, however, the most promising range for this parameter was narrowed to a range of 0.0–0.5.

In general, the values of depression storage for impervious area for "behavioural" sets were found to be distributed within the initial range for all subcatchments. One reason is that the initial range for this parameter already represented the parameter's feasible range. Another potential reason is that calibration of SWMM is less sensitive to this type of control parameter than the other three control parameters.

CONCLUSIONS

An application of a real-value coding GA with the physically based distributed catchment model SWMM was investigated. It was found that a uniform crossover operation outperformed one-point and two-point crossover operations across the whole range of crossover probabilities. Furthermore, for implementation of the real-value coding GA in a complex system, it was found that the optimal uniform crossover and mutation probabilities were in the range 0.75–0.90 and 0.01–0.1, respectively.

The real-value coding GA approach based on the use of multiple storms was found to be a robust technique to identify the most promising ranges for spatially variable parameters. It was found that the initial ranges for eight subcatcment widths, seventeen impervious percentages and seven impervious Manning's roughness were appropriate since the behavioural parameter sets were distributed across the whole of the initial range. On the other hand, the most promising range for 16 subcatcment widths was identified in the range 0.5–1.0, which is a decrease from the initial range of values considered. In a similar manner, for 13 subcatchments, the most promising range for the impervious percentage was decreased to a range of 0.0–0.5. Finally, for the impervious Manning's roughness, the most promising range for 23 subcatchments was narrowed to 0.0–0.5.

In general, the parameter representing depression storage for an impervious area was found to be distributed across the initial range for all subcatchments. One reason is that the initial range for this parameter already represents the feasible range for this parameter. Another reason is that calibration of SWMM is less sensitive to this parameter than the other three parameters.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Department of Civil and Environmental Engineering at the University of Adelaide for the provision of the base GA software used in this study.

REFERENCES

Abustan, I. 1997 *Modelling of Phosphorous Transport in Urban Stormwater Runoff*, Doctoral Dissertation, School of Civil and Environmental Engineering, The University of New South Wales, Sydney, Australia.

Anderson, A. 1995 *Genetic Algorithm Software in f77.*, Report, School of Civil and Environmental Engineering, University of Adelaide, Adelaide, Australia.

Beven, K. 1989 Changing ideas in hydrology – the case of physically based models. *J. Hydrol.* **10**, 157–172.

Beven, K. & Binley, A. 1992 The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.* **6**, 279–298.

Beven, K. & Freer, J. 2001 Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *J. Hydrol.* **29** (1), 11–29.

Beven, K. J. 1993 Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv. Wat. Res.* **16**, 41–51.

Cameron, D. S., Beven, K. J., Tawn, J., Blazkova, S. & Naden, P. 2001 Flood frequency estimation by continuous simulation for a gauged upland catchment with uncertainty. *J. Hydrol.* **219**, 169–187.

Dandy, G. C., Simpson, A. R. & Murphy, L. J. 1996 An improved genetic algorithm for pipe network optimisation. *Wat. Res. Res.* **32** (2), 449–458.

Dayarante, S. T. & Perera, B. J. C. 1999 Parameter optimisation of urban stormwater drainage models. In *Proc. 8th Int. Conf. on Urban Storm Drainage, Sydney, Australia*, pp. 1768–1775.

Eckhardt, K., Breuer, L. & Frede, H. G. 2003 Parameter uncertainty and the significance of simulated land use change effects. *J. Hydrol.* **273**, 164–176.

Fang, T. & Ball, J. E. 2005 Implementation of a genetic algorithm for sensitivity analysis on spatially variable parameters. In *Proc. 9th Annual Environmental Research Event, Hobart, Australia*.

Falkenauer, E. 1998 *Genetic Algorithms and Grouping Problems*. John Wiley and Sons, New York.

Franchini, M. 1996 Using a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall-runoff models. *Hydrol. Sci. J.* **41** (1), 21–40.

Freer, J., Benen, K. J. & Ambroise, B. 1996 Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Wat. Res. Res.* **32** (7), 2161–2173.

Freeze, R. A. 1980 A stochastic-conceptual analysis of rainfall-runoff processes on a hillslope. *Wat. Res. Res.* **16**, 391–408.

Goldberg, D. E. 1989 *Genetic algorithms in search, optimization and machine learning*. Addison-Weasley, Reading, MA, USA.

Goldberg, D. E. & Deb, K. 1991 A comparative analysis of selection schemes used in genetic algorithms. In *Foundation of Genetic Algorithms* (ed. in B. M. Spatz), Morgan Kaufmann, San Mateo, CA. pp. 69–93.

Holland, J. H. 1975 *Adaptation in Natural and Artificial Systems*. MIT Press, Cambridge, MA.

Huber, W. C. & Dickinson, R. E. 1988 *Storm Water Management Model (SWMM): Version 4, User's Manual*. *Report No. EPA/600/3-88/001a*. Environmental Protection Agency, Athens, GA.

Jain, A., Srinivasalu, S. & Bhattacharjya, R. K. 2005 Determination of an optimal unit pulse response function using real-coded genetic algorithm. *J. Hydrol.* **303** (1), 199–214.

Janikow, C. Z. & Michalewicz, Z. 1991 An experimental comparison of binary and floating point representations in genetic algorithms. In *Proc. 4th International Conference on Genetic Algorithms*, Morgan Kaufmann, San Maeto, CA. pp. 31–36.

Klepper, O. & Rouse, D. I. 1991 A procedure to reduce parameter uncertainty for complex-models by comparison with real system output illustrated on a potato growth-model. *Agric. Syst.* **36** (4), 375–395.

Kuczera, G. 1983 Improved parameter inference in catchment models: 2 Combining different kinds of hydrologic data and testing their compatibility. *Wat. Res. Res.* **19**, 1163–1172.

Liong, S. Y., Chan, W. T. & Shreeram, J. 1995 Peak-flow forecasting with genetic algorithm and SWMM. *ASCE J. Hydraul. Engng.* **121** (8), 613–617.

McCabe, M. F., Franks, S. W. & Kalma, J. D. 2005 Calibration of a land surface model using multiple data sets. *J. Hydrol.* **302**, 209–222.

McMichael, C. E., Hope, A. S. & Loaiciga, H. A. 2006 Distributed hydrological modelling in California semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation. *J. Hydrol.* **317** (3–4), 307–324.

Michalewicz, Z. 1994 *Genetic algorithms + data structure = evolution programs*. Springer-Verlag, New York.

Murphy, L. J., Simpson, A. R. & Dandy, G. C. 1993 Design of a network using genetic algorithms. *Water* **20**, 40–42.

Ng, A. W. M. & Perera, B. J. C. 2003 Selection of genetic algorithm operators for river water quality model calibration. *Engng. Appl. Artif. Intell.* **16**, 529–541.

Oliveira, R. & Loucks, D. P. 1997 Operating rules for multireservoir systems. *Wat. Res. Res.* **33** (4), 839–852.

Qi, X. & Palmieri, F. 1993 The diversification role of crossover in the genetic algorithms. In *Proc. 5th International Conference on Genetic Algorithms*, Morgan Kaufmann, San Maeto, CA. pp. 132–137.

Ritzel, B. J., Eheart, J. W. & Ranjihan, S. 1994 Using genetic algorithms to solve a multiple objective groundwater pollution problem. *Wat. Res. Res.* **30** (5), 1589–1603.

Romanowicz, R. & Beven, K. 1998 Dynamic real-time prediction of flood inundation probabilities. *Hydrol. Sci. J.* **43** (2), 181–196.

Simpson, A. R., Dandy, G. C. & Murphy, L. J. 1994 Genetic algorithms compared with other techniques for pipe optimisation. *ASCE J. Wat. Res. Plann Mngmnt.* **120** (4), 423–443.

Sorooshian, S., Gupta, V. K. & Fulton, J. L. 1983 Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models: influence of calibration data variability and length on model credibility. *Wat. Res. Res.* **19**, 251–259.

Spear, R. C., Grieb, T. M. & Shang, N. 1994 Parameter uncertainty and interaction in complex environmental models. *Wat. Res. Res.* **30** (11), 3159–3169.

Spears, W. M. & De Jong, K. A. 1991 On the virtues of parameterised uniform crossover. In *Proc. 4th International Conference on Genetic Algorithms*, Morgan Kaufmann, San Maeto, CA. pp. 230–236.

Wang, Q. J. 1991 The genetic algorithm and its application to calibrating conceptual rainfall runoff models. *Wat. Res. Res.* **27** (9), 2467–2471.

Wardlaw, R. & Sharif, M. 1999 Evaluation of genetic algorithms for optimal reservoir operation. *ASCE J. Wat. Res. Plann. Mngmnt.* **125** (1), 25–33.

Yoon, J. H. & Shoemaker, C. A. 1999 Comparison of optimization methods for ground-water bioremediation. *ASCE J. Wat. Res. Plann. Mngmnt.* **125** (1), 54–63.