JOURNAL OF THE KOREAN SOCIETY OF AGRICULTURAL ENGINEERS SEP. 2003, pp. 117-125

한국농공학회지
제45권 제5호

Errors in Recorded Information and Calibration of a Catchment Modelling System(II)
- Monitoring Calibration Approach -

기록치 오차와 유역모형의 검정(II)
- 모니터링 검정방법 -

Choi, Kyung Sook*
최 경 숙

Abstract

Since the recorded information used for operation of a catchment modelling system contain errors that influence the calibration of catchment modelling system control parameter values, the accurate estimation of these parameters is difficult. Despite these influences, existing traditional calibration approaches focus only on achieving the best "curve fitting" between simulated and recorded data, and not on generic evaluation of control parameter values. This paper introduces an Early Stopping Technique which is aimed at avoiding the procedure of curve-fitting through monitoring improvements in the objective function used for assessing the optimal parameter set. Application of this approach to the calibration of SWMM (Storm Water Management Model) on the Centennial Park catchment in Sydney, Australia is outlined.

Keywords : Calibration, Curve fitting, Early stopping technique, Monitoring information, Validation

I. Introduction

Calibration of a catchment modelling system is the process whereby the values of the control parameters for the models are selected so that the predicted catchment response to a storm event, or sequence of events, adequately reproduces recorded catchment response. These control parameters form part of the input data for operation of the catchment modelling system. Currently, two basic approaches used for evaluating control parameters when a gauged catchment is modelled are the trial and error approach, and the automatic calibration approach. The trial and error adjustment is made by visual comparison between the simulated and observed

* Consultant, Sydney Water Corporation, Australia
Tel: +82-53-950-5730
fax: +82-53-950-6752
E-mail address: ks_choi@hotmail.com

values or hydrographs until the match has improved to an acceptable accuracy, while automatic calibration uses a search engine such as an optimisation algorithm to minimise a function value which represents an objective measure of accuracy. The traditional calibration approaches, therefore, whether it is a trial-and-error technique or an optimisation technique, consist of modifying parameter values until satisfactory accuracy in the simulation is achieved.

Recorded data used for calibration usually contain errors (Pilgrim, 1975; Desbordes, 1981; Williams and Yeh, 1983). The disturbance of errors within the calibration data significantly influences estimation of control parameters. During the calibration process, therefore, it is very important that this influence is recognised and incorporated into the process so that unbiased estimates of the parameters can be obtained and, where appropriate, physically meaningful values obtained. With this situation, the calibration process should not proceed further only for the purpose of minimising the error. Existing or traditional calibration approaches, however, focus most on achieving the best "curve fitting" between simulated and monitored data, and not on generic evaluation of control parameter values. The parameter values estimated might be the best results for the calibration events but might not truly be optimised because they are derived from a curve fitting process. Refsgaard and Storm (1996) also noted that a good match through the curve fitting process does not necessarily guarantee the reliability of estimated parameter values. The final control parameter values selected in calibration, which have been assumed to be the optimum parameter set, might not be the best set but rather a poor set for validation events and subsequent application event simulations. The problem, therefore, is to find the stop point for the calibration process in order to attain meaningful parameter values and not to attain a curve fitting result.

The approach proposed in this paper is based on monitoring the calibration process and thereby avoiding the effects of errors within the calibration data, and hence reducing the potential for curve fitting of the calibration data set. An Early Stopping Technique was employed for this purpose.

II. Methodology

1. Early Stopping Technique (EST)

The Early Stopping Technique consists of three data sets: (1) calibration data, (2) monitoring data, and (3) validation data. The calibration data is used to minimise deviations between simulated and measured values through the optimisation process. The monitoring data is used to check the progress of the calibration process and to decide which point should be taken as the stopping point for the calibration. The validation data is used for final evaluation of parameter set selected from this point. The concept of this early stopping technique is illustrated in Fig. 1.

Typical error curves of monitoring data and calibration data are shown in this figure. At the beginning of optimisation process, some of the parameters are adjusted to minimise the error between calibration target and simulation output.

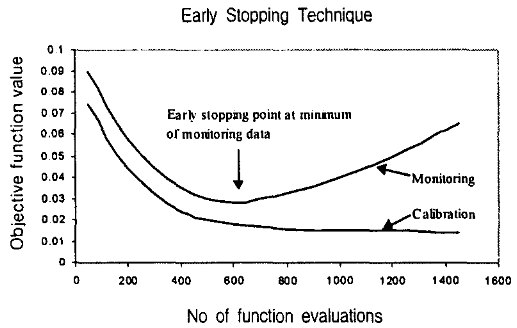

Fig. 1 Early stopping technique

As optimisation progresses, the values of parameters are getting closer to optimum values and, hence, the error in both calibration and monitoring data set will gradually decrease. As optimisation progresses further and beyond a certain point, the calibration process starts to fit the noise of the calibration data, and the process becomes one of curve fitting. Consequently, a further decrease in the calibration error occurs but an increase in the monitoring error is observed. When the monitoring data error, therefore, reaches a minimum value, the parameter set at this point is considered optimised although calibration data error can still decrease. It is at this point where the parameter values reflect the physical processes and avoid the curve fitting process. The control parameter set obtained at this early stopping point is therefore selected as most representative of the true catchment values, and this is then confirmed by the validation data, an independent data set.

2. Application of EST

The Centennial Park catchment, was selected for the case study. This catchment is also referred to as the Musgrave Avenue Stormwater Channel catchment, and is located in the eastern suburbs of Sydney, Australia. The characteristics of the catchment are described in Choi (2003). Spatial information about the catchment was constructed in a GIS database in order to achieve high resolution of spatial variability, and to attain accurate initial estimation of control parameters.

Temporal information within the catchment was available in HYDSYS which is a computer system used to store, process, analyse and report hydrometric time series database. For the application of EST, rainfall and flow information were extracted from HYDSYS in the form of instantaneous value at the end of a time interval. Both single and multiple peak events were selected with data available at a time step of 5 minutes. Three events for calibration and monitoring processes respectively and four events for validation process were used to implement this approach. The details of these

Table 1 Details of events

Events	Rainfall (mm)	Runoff (m^3)	Peak flow (m^3/s)	AMC
Calibration events				
Nov. 04, 94	4.0	1886.1	0.849	Dry
Nov. 29, 94	4.0	1203.6	0.382	Dry
Jan. 28, 95	8.0	3074.7	0.896	Dry
Monitoring events				
Oct. 21, 94	8.2	3047.4	0.726	Dry
Oct. 31, 94	5.8	1764.3	0.547	Dry
Jan. 02, 95	7.8	786.0	0.512	Rather Dry
Validation events				
Nov. 01, 94	7.8	2535.6	1.852	Dry
Dec. 22, 94	4.2	1695.9	1.413	Dry
Jan. 04, 95	8.6	3198.0	1.529	Wet
Feb. 28, 95	9.6	3451.2	2.634	Dry

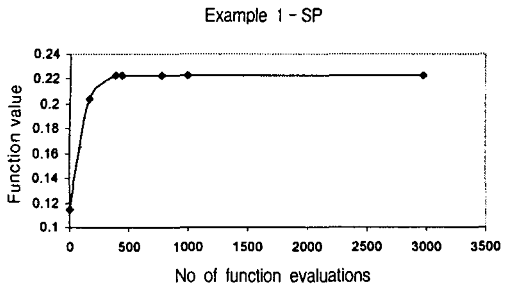

Fig. 2 Example 1 - When minimum function value occurs at SP

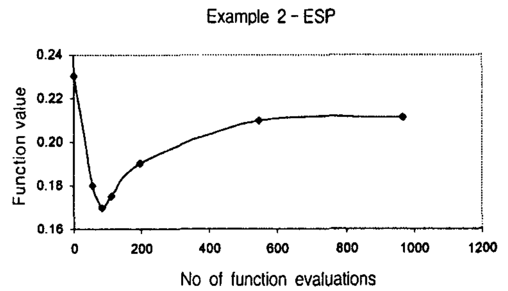

Fig. 3 Example 2 - When minimum function value occurs at ESP

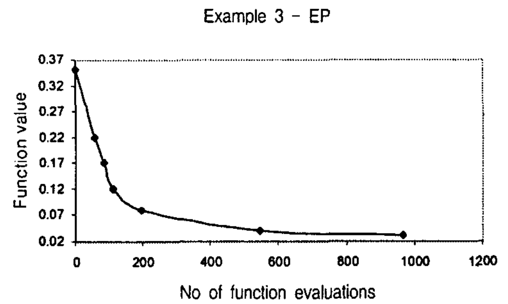

Fig. 4 Example 3 - When minimum function value occurs at EP

events are shown in Table 1.

The antecedent wetness of the catchment was categorised based on the total amount of rainfall within the proceeding 24 hours, which was adapted from Abustan (1997).

An investigation was performed to find an actual stopping point for a calibration process, and hence to find the location of the minimum function values for the validation events through monitoring process. When the minimum function value is obtained at initial iteration as shown in Fig. 2, this point was considered as Starting Point (SP). When the monitoring data error reaches a minimum function value in the middle of the process, this point was considered as Early Stopping Point (ESP). An example of ESP is shown in Fig. 3. Meanwhile, the minimum function value can also be obtained from Ending Point (EP) as shown in Fig. 4. Based on these three points, the investigation was performed by using eight models which were developed with different model complexity and structure within SWMM based on the later approach described in Choi and Ball (2002).

The selected criterion for function evaluations in this study was peak flow objective function shown in Equation (1). The reason of choosing this criteria is that the error in peak flow is considerable and highly variable, and hence, this error can be more easily monitored through EST.

Absolute Relative Error (ARE) of Peak Flow

$$ARE_{peak\_flow} = \left|\frac{P_o - P_s}{P_o}\right| \quad \cdots\cdots (1)$$

where P_o: is observed peak flow (m^3/s)
P_s: is simulated peak flow (m^3/s)

The calibration was performed by minimising the summation of errors for individual events (ie. Σe_i, where, e_i is an individual event) instead of minimising the error of each calibration event separately. As mentioned in the previous section, the monitoring of the calibration process was performed using the control parameter values obtained from each iteration of the calibration process, and then this monitoring information used for validation events to check feasibility of EST for finding right stopping point of the calibration process. The L_BFGS_B (limited memory quasi-Newton algorithm) was employed to assist in the calibration process.

Ⅲ. Results and Discussion

Presented in Table 2 are the calibration results. Since the function values of the calibration subject to decrease until EP, and hence there is no ESP within this process as shown in Fig. 4, only the function values at SP and EP were illustrated in Table 2. Each model showed a different function value as well as a different number of function evaluations depending on number of model parameters and model structure although the original concept of the models are the same. A detailed discussion of the influence of model complexity and structure on the calibration process can be found in Choi (2003).

Table 2 Calibration results

Model	No. of Calibrated parameter	SP	EP	
		Function value	No of iteration	Function value
A	61	0.211	3721	0.180
B	20	0.211	900	0.052
C	55	0.241	6050	0.058
D	55	0.284	6270	0.128
E	14	0.241	1526	0.097
F	14	0.284	1064	0.103
G	49	0.261	7252	0.042
H	8	0.261	664	0.067

From monitoring of the calibration process, it was observed that there are several patterns to the monitoring data sets. The first pattern is that an obvious ESP is shown in the monitoring data. As mentioned earlier, this point is the location containing the optimal calibration parameter sets. After this point, the parameter values are not meaningful as these values are affected by the process reducing errors in the calibration data. The parameter sets obtained from this point is, therefore, used to validate the modelling system.

In the second pattern, the monitoring data does not show clear ESP due to insensitivity of the response between monitoring data and control parameter sets achieved from the calibration process. This case demonstrates an instability in the calibration/monitoring process. It is suspected that this instability is due to one or more of the following

- Errors in the control parameters not considered part of the calibration / monitoring process (e.g. rainfall).
- Errors in the transformation within the catchment modelling system (e.g. the non-linear reservoir technique used within SWMM for simulating surface runoff).
- Errors in the monitoring data.

In the third pattern, several ESPs may be observed because of the different characteristics of the monitoring events. In other words, various shapes of function value lines may be shown depending on the characteristics of errors within the monitoring data. The first ESP was considered as the location of minimum function value in this case. Similar comments to the second case

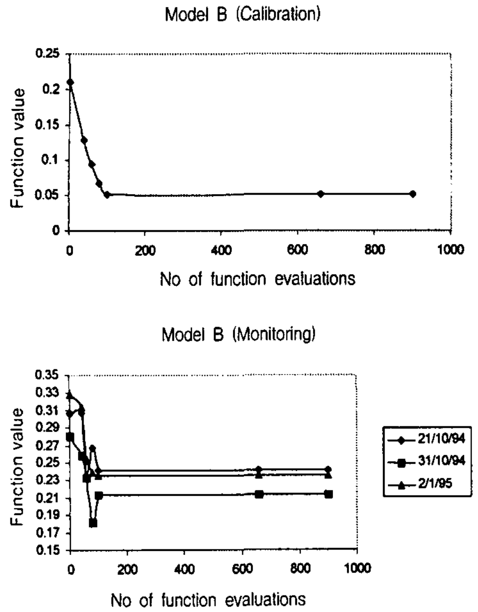

Fig. 5 Objective function value curves for the calibration and monitoring processes of Model B

Table 3 Location of the minimum values for the monitoring events

Model	Monitoring I (21/10/94)		Monitoring II (31/10/94)		Monitoring III (2/1/95)	
	No of iterat.	Func. value	No of iterat.	Func. value	No of iterat.	Func. value
A	1**	0.306	671	0.267	671	0.325
B	60	0.235	80	0.182	900*	0.234
C	110	0.203	110	0.118	110	0.183
D	4015	0.290	220	0.258	4015	0.326
E	770	0.240	1526*	0.184	420	0.200
F	1064*	0.258	42	0.172	1064*	0.249
G	2156	0.185	2009	0.105	2156	0.132
H	128	0.220	128	0.142	128	0.206

* EP ; ** SP

regarding the source of this instability are valid.

An example of the function value curves from the calibration and monitoring processes of model B are presented in Fig. 5, and the location of the minimum function values of the eight models for the three monitoring events are shown in Table 3.

As can be seen in this table, each monitoring event showed a different location of the minimum function value for the same models. Four out of a total of 24 cases (i.e., eight models×three monitoring events) had the minimum function values at EP and one case at SP. Most cases, therefore, showed the minimum function values at ESP. Although the locations of the ESP varied from the three monitoring events, these locations were not necessarily correlated with model complexity or structure while the iteration numbers of the calibration most likely depends on model complexity as the complex models usually showed higher number of iterations during the calibration process.

Using the monitoring information shown in Table 3, the validation was performed with four events. Five control parameter sets for each model, representing SP, three ESPs and EP were selected for this validation to test stability of the monitoring information. The results are summarised in Appendix A. The bold numbers in this Appendix present the minimum function values of the models. It was observed that optimal parameter sets could be located at an ESP, SP or EP. As shown in this Appendix, however, most cases produced the minimum function values at an ESP rather than at EP or SP.

The results also showed that more monitoring data gives more reliable indications of stopping points because of the higher possibility of finding true ESP. Among 32 cases of the validation results, 6 cases (i.e., the shaded line) showed identical results from three ESPs while the rest produced different function values from each

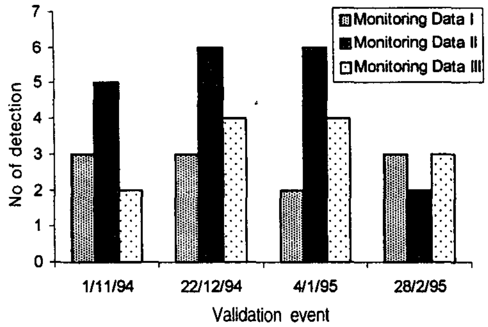

Fig. 6 Performance of each monitoring data

ESP. It can be assumed that these 6 cases have more reliable monitoring information indicating where the calibration process should stop to obtain optimal parameter sets. It can also be concluded that the more identical ESPs the calibration process has, the more reliable validation results can be achieved.

The performances of three monitoring data for finding the location of the minimum function value in the validation process are shown in Fig. 6. Different performances of monitoring data for validation process were observed.

As shown in this figure, the monitoring data II, i.e., the event on Oct. 31, 1994, showed overall better performance in detecting the location where the actual minimum function value can be obtained. This suggests that this monitoring data gave more accurate information to the validation process. However, it also can be noticed that if different validation events were selected, the accuracy of monitoring information could vary due to different noise situations in the selected validation events.

In addition, as mentioned earlier, the behaviour of each set of monitoring data was highly variable depending on characteristics of events used, and

Table 4 Occurrence rate of the minimum function value for each point for the validation events

Event	SP	ESP	EP
Nov. 01, 94	0%	87.5%	12.5%
Dec. 22, 94	0%	87.5%	12.5%
Jan. 04, 95	0%	87.5%	12.5%
Feb. 28, 95	25%	62.5%	12.5%
Mean	6.25%	81.25%	12.5%

hence this variety directly influenced the validation process. The ESP point could, therefore, have happened at any point, and this explains why unique optimum control parameters from different events cannot be achieved no matter how powerful the optimisation algorithms used for Calibration.

The occurrence rate of the minimum function values for the three points was calculated based on Appendix A, and presented in Table 4.

The control parameter sets at ESP produced the minimum function values for validation events in 81.25% of cases, while SP and EP rarely produced the minimum function values. This fact highlights the invalidity of using a control parameter set obtained from the EP for the validation process, which is a common approach. The feasibility of EST for finding a right stopping point for the calibration process was also proved from these results.

Ⅳ. Conclusions

The proposed approach attempted to avoid the problem of curve fitting through learning the noise associated with the calibration events by adoption of an Early Stopping Technique and, hence, attempted to achieve a true optimal control parameter set. The important finding of this study was that EP was not the definite location of the minimum function value for the monitoring and validation events; rather, ESP was the dominant location of the minimum function values, where the effects of errors within the calibration data are minimised, and hence optimal parameter values can be obtained. This study also provided evidence of difficulties for finding global optimum values of control parameters through the EST approach.

References

1. Abustan, I., 1997. *Modelling of phosphorous transport in urban stormwater runoff*, Ph.D. Dissertation, School of Civil and Environmental Engineering. The University of New South Wales, Australia.
2. Choi, K. S. and Ball, J. E., 2002. Parameter estimation for urban runoff modeling. *Urban Water Journal*. **4**. pp.31-41.
3. Choi, K. S., 2003. Effects of model complexity, structure, and objective function on calibration process, *Journal of the Korean Society of Agricultural Engineers*, **45**(4). pp.89-97.
4. Desbordes, M., 1981. Data and fitting problems in urban storm drainage modelling. *In Urban Storm Hydraulic and Hydrology proceedings of the Second International Conference on Urban Storm Drainage*, Ed by Ben Chie Yen, Urbana, Illinois, USA.
5. Pilgrim, D. H., 1975. Model evaluation, testing and parameter estimation in hydrology. In *Prediction in Catchment Hydrology*, Ed. Chapman, T. G. and Dunin, F. X., Australian Academy of Science, Canberra, pp.305-333.
6. Refsgaard, J. C. and Storm, B., 1996. Construction, calibration and validation of hydrological models. In *Distributed Hydrological Modelling*, Abbott, M. B and Refsgaard, J.C. eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, pp.41-54.
7. Williams, B. J. and Yeh, W.W-G., 1983. Parameter estimation in rainfall-runoff models. *Journal of Hydrology*, **63**, pp.373-393.

Appendix A

Validation Results

Validation 1 (Event 1/11/94)					
	SP	ESP 1	ESP 2	ESP 3	EP
A	0.358	0.358**	**0.344**	**0.344**	0.347
B	0.358	**0.290**	0.304	0.291*	0.291
C	0.367	0.290	0.290	0.290	**0.278**
D	0.350	0.336	**0.335**	0.336	0.337
E	0.367	**0.304**	0.327*	0.330	0.327
F	0.350	0.299*	**0.296**	0.299*	0.299
G	0.359	0.297	**0.277**	0.297	0.299
H	0.359	**0.287**	**0.287**	**0.287**	0.301
Validation 2 (Event 22/12/94)					
	SP	ESP 1	ESP 2	ESP 3	EP
A	0.337	0.337**	**0.291**	**0.291**	0.302
B	0.337	0.274	**0.249**	0.260*	0.260
C	0.319	**0.233**	**0.233**	**0.233**	0.250
D	0.302	**0.278**	0.320	**0.278**	0.287
E	0.319	0.254	0.252*	0.252*	**0.252**
F	0.302	0.287*	**0.277**	0.287*	0.287
G	0.308	0.260	**0.250**	0.260	0.277
H	0.308	**0.242**	**0.242**	**0.242**	0.242
Validation 3 (Event 4/1/95)					
	SP	ESP 1	ESP 2	ESP 3	EP
A	0.271	0.271**	**0.269**	**0.269**	0.274
B	0.271	0.213	**0.180**	0.181*	0.181
C	0.275	**0.201**	**0.201**	**0.201**	0.207
D	0.252	0.236	**0.216**	0.236	0.239
E	0.275	0.181	0.181*	**0.180**	0.181
F	0.252	0.210*	0.216	0.210*	**0.210**
G	0.252	0.180	**0.161**	0.180	0.181
H	0.252	**0.164**	**0.164**	**0.164**	0.188
Validation 4 (Event 28/2/95)					
	SP	ESP 1	ESP 2	ESP 3	EP
A	**0.270**	0.270**	0.294	0.294	0.295
B	0.270	**0.241**	0.248	0.300*	0.300
C	0.298	**0.239**	**0.239**	**0.239**	0.242
D	**0.264**	0.268	0.269	0.268	0.278
E	0.298	0.238	0.239*	**0.225**	0.239
F	0.264	0.278*	**0.236**	0.278*	0.278
G	0.265	**0.239**	0.288	**0.239**	0.274
H	0.265	0.254	0.254	0.254	**0.253**

*ESP = EP; ** ESP = SP