JOURNAL OF THE KOREAN SOCIETY OF AGRICULTURAL ENGINEERS JULY. 2003, pp. 89-97

한국농공학회지
제45권 제4호

Effects of Model Complexity, Structure and Objective Function on Calibration Process

모형의 복잡성, 구조 및 목적함수가 모형 검정에 미치는 영향

Choi, Kyung Sook*
최 경 숙

Abstract

Using inference models developed for estimation of the parameters necessary to implement the Runoff Block of the Stormwater Management Model (SWMM), a number of alternative inference scenarios were developed to assess the influence of inference model complexity and structure on the calibration of the catchment modelling system. These inference models varied from the assumption of a spatially invariant value (catchment average) to spatially variable with each subcatchment having its own unique values. Furthermore, the influence of different measures of deviation between the recorded information and simulation predictions were considered. The results of these investigations indicate that the model performance is more influenced by model structure than complexity, and control parameter values are very much dependent on objective function selected as this factor was the most influential for both the initial estimates and the final results.

Keywords : Model complexity, Structure, Objective function, SWMM, Inference models, Calibration process

I. Introduction

In general, the calibration process involves minimisation of the deviation between monitored information and simulation predictions through repeated adjustment of parameters. Implementation of this process requires temporal and spatial information at an adequate resolution to achieve robust predictions from a catchment modelling system. Unfortunately, the available information usually is not adequate for this purpose. Especially, when a spatially distributed, physically based catchment modelling system is employed, the accurate estimation of control parameters is very difficult since modelling systems use a large number of parameters to describe complex processes within the catch−

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

ment. Therefore, it becomes necessary to either use catchment average values or to use other techniques to infer the necessary information. Developments in information technology and the availability of digital information have facilitated the later approach (see for example Choi and Ball, 2002). As presented in Choi and Ball (2002), the proposed approach is to adopt the concept implicitly implemented with the use of inferred parameters; this concept is based on the application of an inference model to determine the value of the control parameter.

Based on this concept, further extensive case studies were performed to examine the performance of the proposed calibration approach in association with different numbers of model control parameters with different structure of the model. Ten different inference scenarios were designed for this purpose. Additionally, several objective functions were also employed to examine their influences on the model calibration results.

In order to assist the calibration process, L_BFGS_B (limited memory quasi-Newton algorithm) developed by the Optimisation Technology Center, Argonne National Laboratory and Northwestern University was employed. This algorithm was particularly designed for solving large nonlinear optimisation problems with simple bounds on the variables.

Presented herein is the results of these investigations into the complexity and structure of models and objective functions used in the calibration process.

Ⅱ. Study Catchment

The Centennial Park catchment was used as a test catchment for this study. The Centennial Park catchment is located in the eastern suburbs of Sydney, Australia as shown in Fig. 1.

The total area of the catchment is 132.7ha, and the average subcatchment slope is about 5.3%. The geological composition of the catchment is Botany sands containing mainly two sand soil types, Hammondville Soil (85%) and Moore Soil (15%).

There are seven different landuses within the

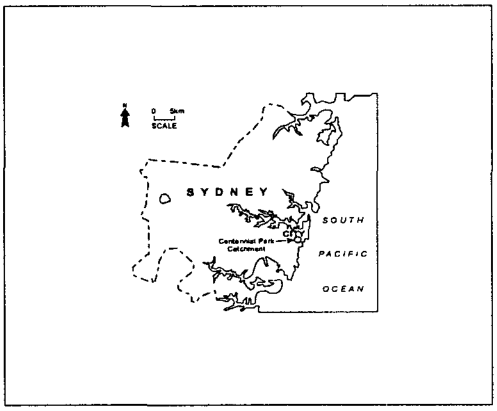

Fig. 1 Location of Centennial Park catchment

Table 1 Landuse within Centennial Park catchment

Landuse types	Area (ha)	Percentage (%)
Special building	16.5	12.4
Road and street	30.8	23.2
Low residential	21.6	16.3
Medium residential	20.8	15.7
High residential	10.8	8.1
Open space and park	28.7	21.6
Commercial/Business	3.5	2.7
Total	132.7	100

catchment. The area of each landuse within the study area and the percentages of the total area are shown in Table 1. For this study, 42 subcatchments were used with sizes varying from 0.50 ha to 27.3 ha.

Ⅲ. Methodology

1. Alternative Inference Models

In the previous study, the inference models were developed originally for selection of spatially variable model control parameters (overland flow length, impervious fraction, depression storage and Manning's roughness of impervious areas) within the Runoff Block in SWMM. These are shown in Equations (1) ~ (5).

- Subcatchment width and factor (X1)

$$W = \frac{A}{L} \quad \cdots (1)$$

$$\alpha = \frac{L}{L_f} \quad \cdots (2)$$

Where W : Subcatchment width
A : Subcatchment area
α : Factor of overland flow length (constant over all subcatchments)
L : Optimum overland flow length
L_f : Estimated overland flow length using spatial database

- Impervious fraction (X2)

$$im = \sum \frac{A_L}{A} \times im_L \quad \cdots (3)$$

Where im : Fraction of impervious area of a subcatchment
A : Subcatchment area
A_L : Area of each land use within a subcatchment
im_L: Fraction of impervious area of each landuse

- Depression storage (X3)

$$dp = \sum \frac{A_{iL}}{A_i} \times dp_L \quad \cdots (4)$$

Where dp : Depression storage
A_i : Impervious area of each subcatchment
A_{iL} : Impervious area of each landuse within a subcatchment
dp_L : Depression storage of each landuse

- Manning's roughness (X4)

$$n = \sum \frac{A_{iL}}{A_i} \times n_L \quad \cdots (5)$$

Where n : Manning's roughness coefficient
A_i : Impervious area of each subcatchment
A_{iL} : Impervious area of each landuse within a subcatchment
n_L : Manning's roughness coefficient of each landuse

Based on these equations, alternative inference models were developed. Each inference model varied with spatially variable values for each subcatchment or spatially fixed values for catchment scale as shown in Table 2. Each model, therefore, had a different complexity (i.e.

Table 2 Various inference scenarios for calibration parameters

Model	No. of parameters to be calibrated				Total
	X1	X2	X3	X4	
A	42	5	7	7	61
B	F*	5	7	7	20
C	42	5	C*	7	55
D	42	5	7	C*	55
E	F*	5	C*	7	14
F	F*	5	7	C*	14
G	42	5	C*	C*	49
H	F*	5	C*	C*	8
I	42	C*	C*	C*	45
J	C*	C*	C*	C*	4

cf) F* : Factor of X1 parameter over all subcatchments
C* : Constant value

different number of model parameters within the inference models) and structure.

In Table 2, X1 is the overland flow length parameter for each subcatchment while X2, X3, and X4 are the impervious fraction, depression storage and Manning's roughness parameters of individual landuses respectively. The number of X1 parameter is equal to therefore the number of subcatchments, while the number of X2, X3 and X4 control parameters is equal to the number of landuse types as these parameters are landuse dependent variables.

For the Centennial Park catchment, 42 subcatchments were employed for this study and there were 7 landuse types as discussed previously. Hence, the number of parameters for X1 should be 42, and the number of parameters for X2, X3 and X4 should be 7 if the models do not employ a factor or constant values for these control parameters.

The control parameter, X2, however, has only 5 parameters requiring calibration since the area of road and street, and of open space for this control parameter were excluded from the calibration process with the parameter values for these areas being fixed as 100 % and 0 % during the calibration process.

The maximum number of parameters to be considered is therefore 61 for model A, which has the maximum spatial variability in control parameter values. The minimum number of parameters is 4 for model J, which has the minimum spatial variability in control parameter values.

Through implementation of these models, two issues were investigated. These are :

- Influence of model complexity and structure on the calibration process with different objective functions.
- Effect on modelling performance of employing a factor or constant values over all subcatchments for calibration parameters such as subcatchment overland flow length, depression storage, and overland flow Manning's roughness.

2. Objective Functions

Several objective functions were selected for the calibration process to investigate the effects of different objective functions associated with different model complexity and structure on evaluation of control parameter values. The selected objective functions were Absolute Relative Errors (ARE) of runoff volume and peak flow, Sum of Square Error (SSE), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) as shown in Equations (6) ~ (11). These are common criteria for evaluation of the model performance in catchment modelling. Each model (A to J) was examined based on these six objective functions.

- Absolute Relative Error (ARE) of runoff volume

$$ARE_{runvol} = \left| \frac{R_o - R_s}{R_o} \right| \quad \cdots\cdots (6)$$

where R_o : is observed runoff volume (m^3)
R_s : is simulated runoff volume (m^3)

- Absolute Relative Error (ARE) of peak flow

$$ARE_{peak\_flow} = \left| \frac{P_o - P_s}{P_o} \right| \quad \cdots\cdots (7)$$

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

- Sum of Square Error (SSE)

$$SSE = \sum_{i=1}^{n} (Q_{oi} - Q_{si})^2 \quad \cdots\cdots (8)$$

where Q_{oi} : is observed flow rate (m^3/s)
Q_{si} : is simulated flow rate (m^3/s)
n : is number of observations in the time series

- Mean Square Error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (Q_{oi} - Q_{si})^2 \quad \cdots\cdots (9)$$

where Q_{oi} : is observed flow rate (m^3/s)
Q_{si} : is simulated flow rate (m^3/s)
n : is number of observations in the time series

- Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Q_{oi} - Q_{si})^2} \quad \cdots\cdots (10)$$

where Q_{oi} : is observed flow rate (m^3/s)
Q_{si} : is simulated flow rate (m^3/s)
n : is number of observations in the time series

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| Q_{oi} - Q_{si} \right| \quad \cdots\cdots (11)$$

where Q_{oi} : is observed flow rate (m^3/s)
Q_{si} : is simulated flow rate (m^3/s)
n : is number of observations in the time series

3. Implementation of Various Inference Models

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. Three events were selected for this

Table 3 Details of calibration events

Events	Rainfall (mm)	Runoff volume (m^3)	Peak flow (m^3/s)	AMC
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

study. Table 3 shows the details of calibration events selected.

After assessment of the simulations, the models I and J were excluded from further study as these models produced unrealistic results which could not be accepted. The reason for this phenomenon is that inference models were originally designed on the basis of land use factor, but these models do not take account this concept as they assume constant values for all land use dependent variables including impervious fraction (X2), depression storage (X3), and overland flow Manning's roughness (X4). These models cannot therefore account for the spatial variability of the catchment, which is contrary to the principle of the inference models developed using landuse factor. Consequently, only eight models (A to H) were accepted for the examination, and hence, a total of 48 alternatives (ie. eight models with six objective functions) were used for evaluation of the model performances. The results of the study will be discussed based on the initial estimates and the final results of the calibration process.

Ⅳ. Results and Discussion

1. Evaluation Based on Initial Estimates

The details of the performance of the eight models for six objective functions based on initial estimates from the calibration data are shown in Table 4.

As shown in this table, the function values of models A and B in this table are equal through the six objective functions because the structures of models A and B at the starting point are basically identical as shown in Table 2. The only difference between these models is that model B uses a factor for the subcatchment overland flow length parameter, while the model A does not. Since the initial value for this factor is 1.0, the initial estimates of this parameter for two models are the same, and hence, they provided the same function values. The models C and E, the models D and F, and the models G and H have also the same structure with different complexity as the case of the models A and B.

Table 4 Function values at starting point

Model	ARE runoff	ARE peak	SSE $(m^3/s)^2$	MSE $(m^3/s)^2$	RMSE (m^3/s)	MAE (m^3/s)
A	0.053	0.211	0.552	0.011	0.099	0.065
B	0.053	0.211	0.552	0.011	0.099	0.065
C	0.098	0.241	0.630	0.013	0.107	0.069
D	0.035	0.284	0.505	0.010	0.094	0.063
E	0.098	0.241	0.630	0.013	0.107	0.069
F	0.035	0.284	0.505	0.010	0.094	0.063
G	0.082	0.261	0.573	0.011	0.100	0.064
H	0.082	0.261	0.573	0.011	0.100	0.064

Once the optimisation process is commenced, however, these models will have different function values at each iteration point. In the case of the same number of model parameters with different structure such as the models C and D (55 parameters), and the models E and F (14 parameters), different results were produced respectively. These results indicate that model structure influences modelling performance more than model complexity does.

In general, the models D and F, having constant value for the Manning's roughness parameter, produced the lowest prediction error among the eight models for the ARE in runoff volume, SSE, MSE, RMSE, and MAE objective functions from the calibration, while the models C and E, having constant value for the depression storage parameter, showed the highest function values. These results suggest that the initial constant value for the depression storage parameter within the models C and E needs to be adjusted for most objective functions, and the initial constant value of Manning's roughness parameter within the models D and F appeared to be reasonable.

For the peak flow objective function, on the other hand, slightly different results were obtained. As shown in Table 4, the models D and F resulted in the highest prediction error for this objective function while the models A and B produced in better results compared to other models. This implies that the initial constant values of Manning's roughness parameter should be modified to minimise prediction error of peak flow.

From the above results, it was noted that the control parameter values are very much dependent on objective functions selected, and model structure influences modelling performance more than model complexity does.

Compared to objective function, however, both factors do not seem to greatly influence modelling performance for the given data set.

2. Evaluation Based on Ending Point

After the optimisation process, the results of eight models were also evaluated at ending point to investigate performances of these models. The results are summarised in Table 5. Eight models showed comparable performance through six objective functions as all models converged to the same location or nearby at ending point. Similar phenomena to the case of initial estimates were observed from the results.

From the comparison of the number of function evaluations, generally, the simpler models such as F and H showed slightly high efficiency among eight models through six objective functions. The more complex models (e.g. A, C, D and G) converged more slowly because of having more control parameters to be optimised. The complex models therefore had less efficiency to reach the

Table 5 Function values at ending point

	ARE of runoff		ARE of peak flow		SSE $(m^3/s)^2$	
Model	Ending point	Function value	Ending point	Function value	Ending point	Function value
A	6954	0.021	3721	0.180	6893	0.429
B	1260	0.021	900	0.052	3040	0.436
C	2970	0.022	6050	0.058	7205	0.419
D	4455	0.021	6270	0.128	3410	0.426
E	2884	0.021	1526	0.097	1834	0.435
F	966	0.021	1064	0.103	826	0.413
G	6811	0.023	7252	0.042	2352	0.429
H	928	0.022	664	0.068	360	0.514

	MSE $(m^3/s)^2$		RMSE (m^3/s)		MAE (m^3/s)	
Model	Ending point	Function value	Ending point	Function value	Ending point	Function value
A	4575	0.009	5002	0.090	4636	0.059
B	500	0.010	1240	0.089	2660	0.055
C	5445	0.009	5940	0.085	7260	0.061
D	2530	0.009	9185	0.086	3850	0.058
E	1428	0.010	1946	0.089	1722	0.055
F	1946	0.008	1162	0.086	1484	0.058
G	4557	0.008	6958	0.081	6027	0.053
H	672	0.009	1320	0.082	376	0.062

locations of minimum objective function values, and did not necessarily produce dominant performances. Due to this fact, Loague and Freeze (1985), Hornberger et al. (1985), Beven (1989), and Jakeman and Hornberger (1993) have suggested employing simpler models for hydrologic modelling in order to prevent over-parameterisation problems when estimating accurate parameter values with limited amount of information.

In general, there were not huge differences between the models using a factor or constant values for estimating control parameter values and the original inference models. Again, this implies that model complexity or structure does not greatly influence model performances. The same observation was made by Gan et al. (1997). For modelling purposes, however, the use of constant values for depression storage and Manning's roughness of impervious area parameters seems to be more effective than the original inference model and the models employing a factor for subcatchment overland flow length parameter, since the model G generally performed slightly better overall across the 6 objective functions.

As shown in Table 5, a wider range of function values over eight models was observed from the peak flow objective function, while smaller variations of the function values between eight models were noticed from other objective functions. This indicates that the peak flow objective function is more sensitive to the model complexity and structure compared to others.

From the 48 results, the runoff volume and peak flow objective functions showed good performance in terms of reducing prediction error in runoff volume and peak flow respectively. The SSE objective function was not effective at reducing peak flow error while the RMSE objective function appeared to be better for this purpose. Also, among all the objective functions, the RMSE objective function performed best in improving shape of the simulated hydrographs. From the above results, it was suggested that the runoff volume and peak flow objective functions are suitable for fitting volume of hydrograph and peak flow, respectively, and when spatial time series analysis is needed, the RMSE objective function is recommended along with other point error measures including SSE, MSE, RMSE, and MAE.

V. Conclusions

Eight different alternative inference models were designed by utilisation of the original inference models developed in the previous study. Using six objective functions, investigation of the influence of model complexity and structure on calibration process was performed.

Based on the results of this study, the following points were concluded:

1) Model performance based on initial estimates was influenced by model structure more than model complexity, and hence sound model structure is more important than high degree of model complexity when limited recorded information is available.

2) Compared to objective function, model complexity and structure had little influence on calibration process for the given data set although the function values at starting point were shown to be more influenced by model

structure than complexity.

3) Objective function was the most influential factor for both the initial estimates and the final results. The selection of objective functions therefore should be made carefully depending on the modelling purpose.

4) The use of constant values for depression storage and Manning's roughness parameters seems to be more effective than employing a factor for the subcatchment overland flow length parameter.

5) Peak flow objective function is more sensitive to the model complexity and structure than other objective functions selected.

6) ARE in runoff volume and peak flow and RMSE objective functions are the recommended objective functions for fitting runoff volume, peak flow and shape of hydrograph respectively.

In conclusion, the model complexity seems to influence the efficiency of the calibration process while the model structure is likely affect modelling performance. The control parameter values are however more dependent on the selection of objective functions than model complexity or structure.

References

1. Beven, K., 1989. Changing ideas in hydrology – The case of physically-based models. *Journal of Hydrology*, **105**, pp.157–172.
2. Choi, K. S. and Ball, J. E. 2002. Parameter Estimation for Urban Runoff modeling. *Urban Water Journal*. **4**, pp.31–41.
3. Gan, T. Y., Dlamini, E. M., and Biftu, G. F., 1997. Effects of model complexity and structure, data quality, and objective functions on hydrologic modelling. *Journal of Hydrology*, Elsevier Science, **192**, pp.81–103.
4. Hornberger, G. M., Beven, K. J., Cosby, B. J., and Sappington, D. E., 1985. Shenandoah watershed study: Calibration of a topography-based, variable contributing area hydrological model to a small forested catchment. *Water Resources Research*, **21**(12), pp.1841–1850.
5. Jakeman, A. J. and Hornberger, G. M., 1993. How much complexity is warranted in a rainfall–runoff model? *Water Resources Research*, **29**(8), pp.2637–2649.
6. Loague, K. M. and Freeze, R. A., 1985. A comparison of rainfall–runoff modeling techniques on small upland catchments. *Water Resources Research*, **21**(2), pp.229–248.