An Evaluation of Three Rainfall-Runoff Models

by
I. D. MOORE, B.E., Grad.I.E.Aust.
Post Graduate Student, Monash University
and
R. G. MEIN, B.E.(Agr.), M.Eng.Sc., Ph.D., M.I.E.Aust.
Senior Lecturer, Department of Civil Engineering, Monash University

SUMMARY Three mathematical rainfall-runoff models, the Boughton model (with a modification to include base flow), the Monash model, and the Stanford Watershed model (slightly modified for daily input data) were applied to four catchments in Australia on a daily basis. The procedure used was to fit the model parameters on several years of data and to test the model performance on a different data period. Evaluation of the models is made on results, ease of use, and computational time required.

1 INTRODUCTION

In the last 10-15 years a considerable number of mathematical rainfall-runoff models have been developed for use in the study of the land phase of the hydrologic cycle and for prediction of runoff from rainfall.

A common feature of many such models is the representation of the hydrologic processes which occur on a catchment by a number of conceptual storages with mathematical functions to describe the movement of water into, between, and out of them. It is usual to claim that some or all of such storages and functions are related directly to the physical catchment and consequently that the model(s) can be used to predict the effect of changes in the hydrologic response of a catchment to rainfall. A major feature is the potential for computing runoff hydrographs on catchments with no streamgauging records.

The data required as input for rainfall-runoff models varies; the internal time step in the model is an important factor. For this project it was felt that since a large amount of rainfall and runoff data is held as daily values, an evaluation should be made of selected models which either use, or could be adapted to use, daily data. To examine whether the climatic nature of the catchment has any effect on model performance, catchments with a wide variety of climate should be represented.

It is common to find in the literature that models have been tested by the model developers on a few catchments. It is less common to find that a model has been independently tested by a user not familiar with the model. It is rare to find any systematic application and comparison of several models on the same catchments. The objective of this project is directed toward the latter in an attempt to show potential users of such models some idea of the effort required to use each model and of the results that can be obtained.

2 MODELS SELECTED

The three models selected represent a variety of catchment simulation methods ranging from simple to complex. The Boughton Model, modified to include base flow, was selected because of its overall simplicity and limited data requirements. The Stanford Watershed Model and the Monash Model were chosen because they attempt to model the major physical processes in some detail. The Stanford Model is well known for the quality of simulation that has been achieved. The Monash Model appears to have potential to achieve good simulation because it is based on theoretical concepts, as a contrast to many of the Stanford's empirical routines.

Programs for all three models were available at the beginning of the project, the Department of Civil Engineering at Monash University being involved in the development of two of them. Despite this association of the Department with the programs, an independent evaluation was possible because of the non-involvement of the senior author in their development.

(a) Boughton Model (Ref.1)

The Boughton Model as modified by McMahon and Mein (Ref.2), with a further small modification, was selected for this study. This version simulates baseflow by dividing the lower zone soil store into two substores, both of which contribute to baseflow with linear recessions. The lower sub-store must fill before the upper substore can hold water. This version is based on the assumption that there is no deep groundwater loss, which does not hold in all cases; in order to make the model more general a deep seepage component has been added. Deep seepage is assumed to be a constant percentage of the moisture level of the lower zone store.

(b) The Stanford Watershed Model

The Stanford Model operates on a fifteen minute time interval and uses hourly rainfall and daily potential evaporation as the basic inputs (Ref.3). In this form it is unsuitable for use in this study; the computer processor time requirements are prohibitive and the use of daily data does not warrant such a small time increment.

The complete land phase of the model was modified to operate on a daily time increment. At this level detailed simulation of overland flow could not be justified so the calculations involving storage delay were omitted. The channel phase was unaltered, but it was only used in its simplest form. Results from the model after the modifications, tested using one year of test data from the U.S.A., were almost identical with the results from the original version. The values of four of the optimized parameters were altered, however.

(c) Monash Model (Porter, Refs.4,5)

The Monash Model in its daily form consists of four storages. Three, namely interception, depression and soil storage have fixed capacities, while the fourth, the groundwater store, has no fixed capacity. A storage routing routine based on Laurenson's technique is also employed. The infiltration component is based on Philip's theory and groundwater discharge is modelled using a nonlinear storage-discharge relationship.

3 MODEL EVALUATION

(a) Evaluation Criteria

Although there are many other factors which can be considered in the comparison of the models, the standard of simulation achieved by each model has the greatest influence on the evaluation. However, it is necessary to specify on which hydrograph characteristics most importance is being placed because it may be found that different models have different merits in flow reproduction. For instance, a model may reproduce peak flows satisfactorily, but be poor in its prediction of mean daily flow. Rather than base the entire evaluation on one or two performance indices, it is probably better to compute several and to evaluate the model performances on a subjective assessment of all of them.

An important factor not always recognized is the ease of use of the model. It is one thing for a researcher to develop a model and use it; it is quite another for someone else to obtain the program and get it working. Full program documentation is necessary. This should at least include a listing, a description of the program and its data requirements, and a sample data set and results. Then there is the time and effort required to become familiar with the model and to run it with catchment data to optimize the parameters. The degree of skill required to interpret the results during the optimizing process also affects the time taken to fit the model parameters to a catchment.

The computational effort necessary to run a model may be a limiting factor for potential users, depending on the memory requirements and processor time required. While this factor may be only of secondary importance to the developer of a catchment model, the computer "cost" must be taken into account by a user proposing to employ a model on a routine basis.

The factors discussed above are not independent of each other. For example, a complex model may be more difficult to use and cost more to run than a simple model; it may also get better results. It is one aim of this study to obtain relative values for each of these factors to enable potential users to make a choice.

(b) Evaluation procedure

The evaluation procedure was a split record technique, one section of the data was used to optimize model parameters while the remainder was used to independently evaluate model performances.

One catchment (North Para River) was used to become familiar with the operation of the models. Different parameter combinations and optimization techniques were tested to gauge their effect on the simulation, and as a result it became evident that the following optimization procedure was satisfactory:

(i) An initial parameter set was selected based on the available data,

(ii) Individual parameters or groups of parameters were altered until the simulated and recorded mean flows were about the same. The ability to do this was largely a function of the operator's experience with the models,

(iii) The steepest ascent method (Ref.6) was applied using the sum of the squares of the errors in the daily flows as the objective function.

This is a generalized procedure only as no one strategy was rigidly adhered to and no one objective function was consistently used at the expense of other indices of model performance. Optimization of model parameters continued until either the improvement in the simulation from one run to the next became negligible or the number of runs became excessive.

Using the optimum parameter set arrived at in the above manner and the remaining part of the record, flows were simulated and compared to the recorded flows. A large number of performance indices were evaluated but only the following are presented in this paper:

(i) Mean daily discharge for both the simulated and recorded flows,

(ii) Standard deviation of both the daily and monthly recorded and simulated flows,

(iii) Correlation coefficient between the simulated and recorded flows,

(iv) Sum of the squares of the errors in the daily simulated flows.

None of these indices by themselves adequately evaluate the performance of a model but viewed together they do convey the correct impression in regard to the relative standard of simulation.

(c) Catchment Selection

For the results of this study to be of practical value the data should be of comparable standard to the data available for a significant number of Australian catchments. The following criteria formed the basis for the selection of catchments used in this study:

(i) Different climatic zones should be represented to enable a variety of flow conditions to be studied,

(ii) Catchments should be of similar size to enable comparisons between catchments. Inclusion of one large catchment to test the performances of the routing routines would be an advantage,

(iii) Rainfall, evaporation and streamflow data should have concurrent record lengths exceeding ten years duration,

(iv) The catchments on which the original versions of the models were developed should not be included to ensure independence of the test.

Ten catchments were finally selected but evaporation data for six of them proved later to be inadequate. The remaining four, used in this study, were the catchments of the Thomson River (Vic.),

TABLE I

CATCHMENT CHARACTERISTICS

Catchment	Area 10^3ha	Mean Annual Rainfall (mm)	% Miss-ing	Stream-flow (mm)	% Miss-ing	% Run-off
Thomson	52	1402	4.4	534	N.A.	38
Shoalhaven	277	874	18.8	255	21.4	29
King	45	2999	0.8	2218	2.2	74
North Para	38	542	7.1	45.5	0.5	8.4

the Shoalhaven River (N.S.W.), the King River (Tas.) and the North Para River (S.A.). A summary of catchment characteristics is presented in Table I.

4 RESULTS

(a) Familiarization

Prior to the start of the project programs for all of the models were available within the Department of Civil Engineering at Monash University. The only documentation used for any of the models consisted of published papers and Institutional Research Reports. For the relatively simple Boughton Model such documentation was adequate, and no difficulty was experienced in proceeding to apply the model. The more complex Stanford Watershed model had no documentation apart from Ref.3 and similar reports. Comment cards within the program were not adequate to fully define all parameters. The only documentation for the Monash model was unpublished (Ref.4), was not user oriented, and consequently it would be difficult for a new user to understand the program structure.

TABLE II

COMPUTER PROCESSOR TIME AND THE NUMBER OF RUNS

Model	Thomson River		Shoalhaven River		King River		Average
	Processor Time*	No. of Runs	Processor Time*	No. of Runs	Processor Time*	No. of Runs	Processor Time*
Boughton	1.559	188	1.662	237	1.775	101	1.665
Monash	16.040	210	17.704	64	11.262	64	15.002
Stanford	38.378	146	48.675	167	26.440	212	37.831

*Average time (seconds) required to process one year of data on the Burrough's B6700 computer at Monash University.

(b) Parameter Optimization

The difficulty involved in attempting to use a consistent strategy to optimize model parameters is that, for each model, different parameters may be dominant for different catchments, for different seasonal conditions on a catchment, for different initial conditions, and may also depend on the initial values assigned to the parameters themselves. That is, the best strategy depends on where one starts on the "response surface" of the objective function.

Most of the parameters in all of the models were found to have a definite "sensitivity range" which depended on the hydrologic conditions. Inside this range, small changes in the parameter value produced significant changes in the objective functions. Outside this range, large changes in parameter values had little effect. To be successful with the steepest ascent procedure initial

TABLE III

THOMSON RIVER RESULTS

Statistic	Recorded	Boughton Model	Monash Model	Stanford Model
OPTIMIZATION PERIOD (1955-1962)				
Mean daily flow (m^3/sec.)	9.87	10.05	9.31	9.93
Standard Deviation (m^3/sec.day) M	251.34	259.45	228.62	232.32
Standard Deviation (m^3/sec.day) D	11.20	10.08	8.72	9.54
Correlation Coefficient M		0.921	0.926	0.939
Correlation Coefficient D		0.797	0.868	0.876
Sum of squares of errors ($m^6/sec^2 \times 10^5$)		1.379	0.944	0.852
TEST PERIOD (1963-1971)				
Mean daily flow (m^3/sec.)	7.80	8.95	7.71	8.38
Standard Deviation (m^3/sec.day) M	197.53	287.33	216.62	208.66
Standard Deviation (m^3/sec.day) D	8.41	10.58	7.88	7.88
Correlation Coefficient M		0.887	0.912	0.913
Correlation Coefficient D		0.738	0.837	0.831
Sum of squares of errors ($m^6/sec^2 \times 10^5$)		1.730	0.718	0.756

M = Monthly D = Daily

TABLE IV

SHOALHAVEN RIVER RESULTS

Statistic	Recorded	Boughton Model	Monash Model	Stanford Model
OPTIMIZATION PERIOD (1958-1963)				
Mean daily flow (m^3/sec.)	35.32	36.57	35.20	35.89
Standard Deviation (m^3/sec.day) M	1303.97	1234.69	1135.43	1134.67
Standard Deviation (m^3/sec.day) D	121.42	90.21	102.18	112.32
Correlation Coefficient M		0.904	0.945	0.934
Correlation Coefficient D		0.817	0.864	0.821
Sum of squares of errors ($m^6/sec^2 \times 10^6$)		10.893	8.207	10.893
TEST PERIOD (1964-1971)				
Mean daily flow (m^3/sec.)	11.03	8.81	9.78	9.73
Standard Deviation (m^3/sec.day) M	623.98	504.65	431.80	453.69
Standard Deviation (m^3/sec.day) D	46.72	20.67	33.97	37.45
Correlation Coefficient M		0.866	0.895	0.911
Correlation Coefficient D		0.674	0.828	0.758
Sum of squares of errors ($m^6/sec^2 \times 10^6$)		3.835	2.075	2.724

M = Monthly D = Daily

TABLE V

KING RIVER RESULTS

Statistic		Recorded	Boughton Model	Monash Model	Stanford Model
OPTIMIZATION PERIOD (1957-1963)					
Mean daily flow (m^3/sec)		30.18	30.35	30.44	30.38
Standard Deviation (m^3/sec.day)	M	633.19	662.48	602.83	617.66
	D	40.27	42.07	38.83	38.67
Correlation Coefficient	M		0.939	0.970	0.968
	D		0.887	0.918	0.911
Sum of squares of errors ($m^6/sec^2 \times 10^5$)			9.829	6.609	7.180
TEST PERIOD (1964-1972)					
Mean daily flow (m^3/sec)		32.85	32.72	32.53	33.16
Standard Deviation (m^3/sec.day)	M	655.19	663.93	598.91	634.94
	D	39.71	41.99	38.27	38.50
Correlation Coefficient	M		0.942	0.977	0.974
	D		0.880	0.912	0.912
Sum of squares of errors ($m^6/sec^2 \times 10^5$)			13.109	8.816	8.946

M = Monthly D = Daily

TABLE VI

NORTH PARA RIVER RESULTS

Statistic		Recorded	Boughton Model	Monash Model	Stanford Model
OPTIMIZATION PERIOD (1957-1962)					
Mean daily flow (m^3/sec)		0.42	0.40	0.42	0.46
Standard Deviation (m^3/sec.day)	M	25.82	26.49	25.76	27.29
	D	1.43	1.01	1.47	1.15
Correlation Coefficient	M		0.962	0.969	0.960
	D		0.718	0.869	0.788
Sum of squares of errors ($m^6/sec^2 \times 10^5$)			0.217	0.121	0.170
TEST PERIOD (1963-1970)					
Mean daily flow (m^3/sec)		0.65	0.53	0.60	0.71
Standard Deviation (m^3/sec.day)	M	41.26	34.76	43.98	42.72
	D	2.52	1.28	2.64	1.72
Correlation Coefficient	M		0.884	0.844	0.811
	D		0.631	0.696	0.635
Sum of squares of errors ($m^6/sec^2 \times 10^5$)			1.150	1.187	1.115

M = Monthly D = Daily

parameter values ideally should be within the "sensitivity range".

As discussed in Section 3(b) familiarization with the models was achieved by applying each of them to the North Para River. The number of iterations required to optimize the parameters, such that further improvement in the results is difficult, is given in Table II. Although there are anomalies in the table it was found that the Monash tended to be easier to optimize than either of the other two. The Boughton and Stanford models, overall, were comparable in the number of iterations required, even though the Stanford model has many more parameters to fit.

(c) Independent Testing

The results for both the optimization period and the test period achieved by each model for each river are given in Tables III, IV, V and VI.

The first point to make is that no one model is superior on all of the four rivers. The Stanford and Monash models performed equally well on the King River. On the Thomson River the Stanford model was fractionally better than the Monash model; the use of a groundwater recession parameter which varied with soil moisture store level seemed to be to the advantage of the Stanford model. For the Shoalhaven River the Monash model gave slightly better simulation than the Stanford model on a monthly basis, and better still on a daily basis. The North Para River was the most difficult to simulate, because it is an intermittent stream. Although Table VI doesn't show it the Boughton model performed better than the others for years in which flows were relatively high, but poorer for years of low flows. Overall, for the North Para River, the Boughton model was marginally superior on a monthly basis. On a daily basis the Monash model was the best performer for this catchment. A slight modification of the function controlling diversion of moisture to the inactive groundwater store was necessary to achieve good simulation with the Stanford model.

(d) Running costs

The computer processor time required to run the models for one year of data on each catchment is given in Table II. The difference in the figures for each model is due primarily to the varying complexities of the channel routing routines - the Boughton model has none, the Monash model routine operates on a daily time step and the Stanford model has routing routines which operate on an hourly time step.

As far as the user is concerned the data preparation for input to each model is easiest for the Boughton model, more extensive for the Monash model, and most time consuming for the Stanford model. A rough figure would be 1:4:5 for the time taken to prepare input data respectively for the Boughton, Monash and Stanford models.

5 CONCLUSIONS

At the time of writing, evaluation of the results was not completed, but the following conclusions can be made:

(i) Each of the models has advantages over the others for specific applications depending on the catchment hydrology, the budgetry constraints, and whether daily or monthly flows are required.

(ii) The Boughton model performed almost as well as the other two on a monthly basis and its running costs and data preparation efforts are considerably less. On a daily basis the standard of simulation is considered to be poor.

(iii) The Monash model and Stanford model produced comparable results in this study, but the former requires less computer and user time.

(iv) The baseflow routines of the Stanford model give it an advantage on catchments where base flow is important.

6 ACKNOWLEDGEMENTS

The authors are indebted to the following organisations who supplied the hydrologic data used in this study; Bureau of Meteorology, Irrigation and Water Supply Commission (Qld.), Water Conservation and Irrigation Commission (N.S.W.), State Rivers and Water Supply Commission (Vic.), Engineering and Water Supply Department (S.A.), Department of the Northern Territory (N.T.), and Hydroelectric Commission (Tas.). The assistance given by the officers of those organisations is gratefully acknowledged. The copy of the Stanford Watershed model was obtained from Prof. S.J. Burges, University of Washington, U.S.A.

7 REFERENCES

1. BOUGHTON, W.C. A Mathematical Model for Relating Runoff to Rainfall with Daily Data. Civ. Engg. Trans. I.E. Aust., Vol. CE8, No.1, April, 1966, pp.83-93.

2. McMAHON, T.A. and MEIN, R.G. Storage-Yield Estimates with Inadequate Streamflow Data. Proc. Symp. on the Design of Water Resources Projects with Inadequate Data. Madrid 1973. Vol.2, pp.355-367.

3. CRAWFORD, N.H. and LINSLEY, R.K. Digital Simulation in Hydrology: Stanford Watershed Model IV. Technical Report No. 39, Dept. Civil Engineering, Stanford Univeriity, 1966.

4. PORTER, J.W. The Synthesis of Continuous Streamflow. Thesis (Ph.D.), Monash University, 1972.

5. PORTER, J.W. and McMAHON, T.A. A Model for the Simulation of Streamflow from Climatic Records. Journal of Hydrology, Vol.13, No.4, 1971, pp.297-324.

6. BOUGHTON, W.C. A Mathematical Catchment Model for Estimating Runoff. Journal Hydrology (N.Z.), Vol.7, No.2, 1968, pp.75-100.