4410 July, 1965 HY 4

Journal of the
HYDRAULICS DIVISION
Proceedings of the American Society of Civil Engineers

MATHEMATICAL MODELS OF CATCHMENT BEHAVIOR

By David R. Dawdy,[1] and Terence O'Donnell,[2] M. ASCE

INTRODUCTION

Quantitative models of catchment behavior to be useful (i.e., acceptably accurate) must inevitably be complex, yet must be feasible to operate. These requirements were incompatible until high-speed computers were available. However, the evaluation of such models is still circumscribed by the current knowledge and understanding of the processes being simulated and by the capabilities of both the computers and the computing techniques available.

The ideal model would specify completely the properties of and the processes that occur in all the relevant components of a catchment. The specification would be given in terms of physical parameters and would involve all behavioral relationships within the catchment. Given such a full specification, the hydrologic effects of a rainfall event over a catchment could be determined objectively.

Our knowledge and techniques do not permit more than a coarse approximation to this ideal. At present (as of 1965), the only reliable and accurate device for yielding the runoff resulting from a rainfall on a catchment is the catchment itself. For many problems in engineering hydrology, long records of rainfall and runoff may be sufficient; there would be no need for modeling in such cases. However, in catchments with short or nonexistent records or in catchments in which environmental changes are taking place, a reliable modeling technique is needed.

Two schools of endeavor working on modeling techniques may be distinguished as follows:

1. comprehensive simulation of catchment behavior, i.e., over-all catchment models; and

Note.—Discussion open until December 1, 1965. To extend the closing date one month, a written request must be filed with the Executive Secretary, ASCE. This paper is part of the copyrighted Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, Vol. 91, No. HY4, July, 1965.

1 Hydr. Engr., Geol. Survey, U. S. Dept. of the Interior, Menlo Park, Calif.

2 Hydr. Engr., Geol. Survey, U. S. Dept. of the Interior, Menlo Park, Calif.; on leave from Dept. of Civ. Engrg., Imperial Coll., London, England.

2. complete specification of each of the elements of catchment behavior, i.e., component models.

The two classes are highly interdependent. Progress can only be made on a complementary feedback basis between the two schools.

Over-all models, at present, treat catchment components in lumped form. The behavior of the components is generally approximated by largely empirical relationships. The construction of the components and the parameters of the relationships are adjusted until known responses, within an acceptable tolerance, are achieved from known inputs. There are usually many subjective decisions, first in the choice of components, then in specifying the behavioral relationships, and finally in the parameter adjustments. The treatment has been developed to the stage of being an effective and acceptable engineering tool. Further development of the over-all model approach will take place in two ways: (1) a less approximate representation of components and (2) an improvement in the techniques of adjusting the parameters.

The "complete specification" school will allow increasingly objective statements of the physical relation of the elements of catchment behavior to be made. Inevitably, these will be more complex than the empirical relation used in present over-all models, but they will be stated objectively in terms of certain physical properties.

The bringing together of the two schools will yield more complex but less subjective models and behavioral relations. It is not necessary for the development of the over-all models approach that we wait until the complete specification endeavors are completed. As the latter approach provides additional information, filling in the details of the picture, so will the over-all approach feed back information to show where further detailed specification is needed.

A third zone of activity is that of developing the most appropriate computing techniques for the operation of the increasingly complex quantitative models that will appear. Hand in hand with the decrease in subjectivity in the over-all models, it is hoped that there would be less subjectivity in fitting the model to a particular catchment. In particular, an efficient automatic procedure for finding numerical values of the various parameters of an over-all model would be valuable.

The work conducted by the writers has primarily been concerned with this third zone of activity, both for over-all models and for component models. The writers summarize the methods used and the results obtained and examine future plans. First, however, there is a brief review of the various mathematical models and techniques for evaluating catchment behavior.

REVIEW

A first broad classification would divide mathematical models of catchment behavior into linear and nonlinear models—a classification that largely parallels the historical picture.

Perhaps the best known linear model is that implicit in the unit hydrograph concept. Basically, unit hydrograph theory states that a rainfall excess input uniformly distributed areally over a catchment is converted to a surface runoff response at the outlet from the catchment via linear storage and translation processes. Such conversion can be represented by the convolution integral which relates any input, i(t), to a linear system and the output, q(t), from that system via the pulse response, u(t), of the system: thus,

$$q(t) = \int_0^t i(\tau)\, u(t-\tau)\, d\tau \qquad (1)$$

For a linear catchment, u(t) is the instantaneous unit hydrograph (IUH) of the catchment (i.e., the surface runoff hydrograph resulting from a unit rainfall excess deposited instantaneously and uniformly over the catchment), and q(t) is the surface runoff output caused by i(t), a rainfall excess input.

In a penetrating analysis, Dooge[3] developed, for the first time, a general theoretical basis for the unit hydrograph method. This analysis showed that any catchment suitable for unit hydrograph analysis can be represented by an equivalent ideal linear catchment model consisting of an appropriate combination of linear channels and linear reservoirs. The model would have the same IUH as the real catchment. Nash[4] developed a linear model technique, not so general as Dooge's, which permits particular numerical values of the parameters in a general two-parameter IUH equation to be found from surface runoff and rainfall excess data for any given catchment.

Valuable as the Dooge and Nash methods are in synthesizing the components of an equivalent linear model for any catchment, it is not necessary to synthesize these components if only the invariant pulse response of the system is required. It is only necessary to possess a set of input and output data. O'Donnell[5] has presented a method of finding the IUH of a catchment directly from a set of surface runoff and rainfall excess data. The method does not require any specification of a model, nor does it yield information other than the IUH. It does presuppose the linearity of catchment behavior.

In reality, catchment behavior is nonlinear. Although this has long been recognized, it is only in recent years that methods of nonlinear analysis and synthesis have been examined vis-à-vis the catchment problem. Amorocho[6] has carried out some pioneering studies, both theoretical and applied, in the use of general nonlinear analysis techniques as applied to catchment behavior. More recently, Amorocho has presented[7] a general survey of contemporary methodologies in hydrologic research. This account gives a clear exposition of systems analysis and synthesis as applied both to linear and nonlinear systems.

Perhaps the most widely known nonlinear modeling of catchment behavior is that developed by Linsley and Crawford[8] for use with a digital computer.

3 Dooge, J. C. I., "A General Theory of the Unit Hydrograph," Journal of Geophysical Research, Vol. 64, 1959, pp. 241-256.

4 Nash, J. E., "A Unit Hydrograph Study, with particular reference to British Catchments," Proceedings, Inst. of Civ. Engrs., London, Vol. 17, November, 1960, pp. 249-282.

5 O'Donnell, T., "Instantaneous Unit Hydrograph Derivation by Harmonic Analysis," Commission of Surface Waters, Publication No. 51, Internatl. Assn. of Scientific Hydrology, 1960, pp. 546-557.

6 Amorocho, J., "Measures of the Linearity of Hydrologic Systems," Journal of Geophysical Research, Vol. 68, 1963, pp. 2237-2249.

7 Amorocho, J., and Hart, W. E., "A Critique of Current Methods in Hydrologic Systems Investigation," Transactions, Amer. Geophysical Union, Vol. 45, No. 2, 1964.

8 Crawford, N. H., and Linsley, R. K., Technical Report No. 12, "The Synthesis of Continuous Streamflow Hydrographs on a Digital Computer," Dept. of Civ. Engrg., Stanford Univ., Stanford, Calif., July, 1962.

This over-all model aims at representing the whole of the land phase of the hydrologic cycle. It is programmed to produce hourly streamflow data using daily evapotranspiration and hourly precipitation data.

In using the model, the typical procedure is to select a 5 yr to 6 yr period of rainfall and runoff records for a catchment. This period is used to develop estimates of the model parameters that fit the general model to the given catchment. A second period of record is then used as a control to check the accuracy of the parameters obtained from the first period. A comparison in the control period is based on such data as total monthly flows, daily flow duration curves, and hourly hydrographs of the two maximum floods each water year.

The initial values of the model parameters are selected on the basis of previous experience and on reasonable judgment. Adjustment of the parameter values during the fitting stage is done in two ways. Most are adjusted by the operator, via a combination of experience and intuition, using clues provided by the timing and magnitude of the differences between the synthesized and recorded streamflow hydrographs. Some of the parameters are evaluated by the computer itself, using an internal looping routine of successive approximations.

A recent TVA study[9] presents a technique for digital computer evaluation of the parameters of a water yield model. Basically, the technique, a nonlinear least squares procedure, is one of successive approximations. Beginning from an initial estimated set of parameters, the errors between an output computed by the model with those parameters and an observed output are found. Partial derivatives of the computed output with respect to each of the coefficients are evaluated. A multiple regression would normally be used to associate the derivatives with the errors. This would yield regression coefficients that would be used as corrections to the parameter values, and the whole set of operations could then be repeated until an acceptable precision was reached. Instead of multiple regression, the multivariate technique of component analysis was used in the belief that the convergence to solution would be more rapid.

The nonlinear least squares fitting of observed data sets used is restrained so as not to "overdetermine" the model. The disadvantage of the method is the difficulty of testing for goodness of fit (in the statistical sense). Practical significance of the parameters has to be substituted for statistical significance.

Amorocho and Hart[7] have cautioned against an excessive reliance on synthetic models of catchment behavior, naming perhaps the most important causes of unreliability as (1) errors in the recorded data, (2) effects caused by "lumping" of components, (3) imperfections of the structure of any synthetic model, and (4) nonuniqueness of the processes of synthesizing an unknown system. In particular, the prediction of long-term output records

---

9 "A Water Yield Model for Analysis of Monthly Runoff Data," Research Paper No. 2, Tennessee Valley Authority, Office of Tributary Area Development, Knoxville, Tenn., February, 1963.

requires extremely close matching between synthetic and observed records in order to assure predictive reliability.

PARAMETER OPTIMIZATION AND SENSITIVITY

Introduction.—While bearing in mind the words of caution just noted, the writers have been exploring automatic objective methods of finding numerical values of the parameters of synthetic hydrologic models.

The successful operation of a digital computer catchment model in which the parameter values are adjusted by the operator relies, to a considerable extent, on the skilled experience and personal judgment of its operator. As previously stated, increases in detailed knowledge of the elements of the hydrologic cycle and the resulting more precise specification of their behavioral relations will lead to more sophisticated, but inevitablty more complicated, models. Without denying the power and advantages of using engineering judgment and acquired skills, it is likely that adjustment of the larger number of parameters of more complex models by subjective trial and error procedures will become impracticable.

Let $U(x_1, x_2, \ldots x_n)$ represent the dependence of a function (representing, for example, an estimate of error of prediction) and a number of parameters x_i. Such a function can be maximized or minimized (or, in general, "optimized") by finding a set of "best fit" parameters. Finding such a set of "best fit" parameter values for given physical systems, given input and output data, is a frequently met problem in many fields of activity. Optimization or "hill climbing" techniques have been developed that determine values of system parameters which maximize or minimize some dependent function of those parameters. Such techniques are completely objective. They may make many useless tests of situations that would be dismissed out of hand by an experienced and skilled human investigator, but the tremendous speed with which a computer can make such tests compensates for such inefficiency. In the present context, optimization means minimizing the errors between a synthesized streamflow and an observed record.

Existing optimization techniques run into difficulties with large numbers of parameters, and the speed of the computer used so far (a Burroughs 220) has necessitated rather long run times. However, computer optimization techniques can and will be improved, and much more powerful computers are becoming available. The advances in optimization techniques and computer capabilities might well keep pace with the developments in the "complete specification" schools of activity.

The writers' studies have been based on an over-all catchment model similar to, but simpler than, the Stanford model and on more detailed component models. Preliminary studies have been concerned with gaining experience and have been carried out with artificially generated data to free these initial studies from the effects of errors in the data. Such data have been generated by giving a set of values to the model parameters and then supplying the models with a given input.

With such compatible and error-free input and output data, studies have been made of how sensitive a model response is to changes in each of the parameters. Tests of an optimization technique as regards speed and effectiveness have been carried out using such compatible data. Beginning from a deliberately chosen "wrong" set of parameters, the rate of progress towards the known correct set has been used to develop improvements in the optimization routine.

The Optimization Technique.—Of several optimizing procedures available, one well suited to the catchment model problem is that developed by Rosenbrock.[10] The particular class of problems for which the method was developed is one in which (1) the parameters, x_i, are restricted by physical considerations and must fall within specific limits, and (2) the function, U, dependent on those parameters, and whose value is to be maximized or minimized, is such that partial derivatives of U with respect to the various x_i cannot be stated analytically in usable forms.

If there are n parameters on which the function U depends, optimization consists of a search in an n-dimensional vector space (formed by n orthogonal parameter axes and bounded by the limits set on the n parameters) until the optimum value of U is found. Rosenbrock's method is recursive in that it makes this search in a series of repetitive stages. Each stage is terminated by evaluating a new set of n orthogonal directions along which the search during the next stage is conducted. The evaluation of the new directions is based on the movements made along the n directions of the current stage. Only in the first stage are the orthogonal directions coincident with the n parameter axes. In subsequent stages, the first component of the new directions lies along the direction of fastest advance.

During each stage, movement is made along each orthogonal direction in a series of steps. A step of arbitrary length, e, is attempted first. This is treated as successful if the resulting new value of U represents an improvement of, or is equal to, the previous value. If a success, the step is allowed, and e is multiplied by $\alpha > 1$; if a failure, the step is not allowed, and e is multiplied by $-\beta$, in which $0 < \beta < 1$. A new attempt is then made. These attempts are terminated as soon as at least one successful attempt, followed by one failed attempt, has been achieved in each of the n directions. Then the new orthogonal directions used in the next stage are evaluated. An attempt in the end must succeed for each direction, because e becomes so small after repeated failures that it causes no change in U.

For the catchment models, U was defined as the sum of the squares of the differences between the recorded and synthesized runoffs for each of the intervals of the record. Other error criteria could be used, e.g., height or timing of peak flows—in fact, a combination of such criteria used in a series might well be useful in future studies.

Some modifications to Rosenbrock's computer program were made to speed up execution time. The efficiency of optimization progress is given by the reduction in U achieved for a given number of attempts. (The program keeps a running sum of the number of attempts made, both successful and failed.) It was found that the rate of improvement of U fell off as the number of stages increased. Eventually, changes in U between stages would become negligible, although U might not be very small. Instead of stopping at this point (as per Rosenbrock[10]), a modification was made to store the latest U value, to set the arbitrary steps, e, back to their start-of-run values, and to use the latest parameter values to start a new "round" of stages. This allowed

---

10 Rosenbrock, H. H., "An Automatic Method of Finding the Greatest or Least Value of a Function," The Computer Journal, Vol. 3, 1960, p. 175.

further progress to be made. A further modification was to limit the number of stages per round, because the most part of the progress in each round is made in the early stages. The program is conditioned to terminate whenever consecutive end-of-round U values differ by less than some specified small percentage. Table 1 presents typical changes in parameter and U values between rounds. Changes between stages are quite similar, in that most of the change generally occurs during the first stage of a series.

Parameter Sensitivity.—Defining U again as the sum of squares of differences in synthesized and recorded Q values, the sensitivity of the response of a model to changes in each of its parameters was examined. This was done by computing U values for both increases and decreases of 1%, 5%, and 10% in each of the parameters. A wide range of sensitivity was shown, for example, the 1% changes producing a most sensitive response nearly 100,000 times greater than the least sensitive (in the case of the over-all model).

TABLE 1.—PARAMETER OPTIMIZATION

Parameter	Correct value	Starting value	Parameter value at end of round				Residual difference, in percentage
			1	2	4	6	
K_S	10	15	10.17	10.13	10.011	10.015	0.15
f_c	0.2	0.1	0.1721	0.1700	0.1973	0.1972	1.4
k	2	3	2.931	2.113	1.983	1.970	1.5
f_o	2	1	1.952	1.943	1.972	1.967	1.7
M*	2	1	1.815	1.886	1.936	1.947	2.7
K_G	40	35	31.32	57.10	45.17	43.96	9.9
R*	0.1	0.15	0.3059	0.2615	0.1174	0.1143	14
G*	4	6	5.834	18.03	19.82	19.27	380
c_{max}	0.1	0.15	2.049	0.6363	0.5282	0.5574	460
	U value	5.07×10^{-1}	9.04×10^{-4}	1.23×10^{-4}	5.43×10^{-6}	2.91×10^{-6}	

THE OVER-ALL MODEL

Description.—The over-all model used in this study was deliberately kept simple so that emphasis could be given to the parameter sensitivity and optimization aspects of the work. Shown in blockflow diagram form in Fig. 1, the model is restricted to four storage elements having simple hydrologic characteristics. The model is open to improvement as the study proceeds, viz., by making use of increased hydrologic knowledge, by incorporating the results of component model studies, and by using more efficient optimizing procedures.

The roles of the various elements shown in Fig. 1 are as follows:

Surface Storage, R.—Augmented by rainfall, P; and depleted by evaporation, E_R, infiltration, F, and, when R exceeds a threshold, R*, channel inflow, Q_1.

Channel Storage, S.—Augmented by channel inflow, Q_1; and depleted by surface runoff at the gaging station, Q_S.

Soil Moisture Storage, M.—Augmented by infiltration, F, and capillary rise, C; and depleted by transpiration, E_M, and, when M exceeds a threshold, M*, deep percolation, D.

Ground-Water Storage, G.—Augmented by deep percolation, D, depleted by capillary rise, C, and baseflow at the gaging station, B; and if and while G exceeds G*, M is absorbed into G, C and D no longer operate, but E_M and F now act on G.

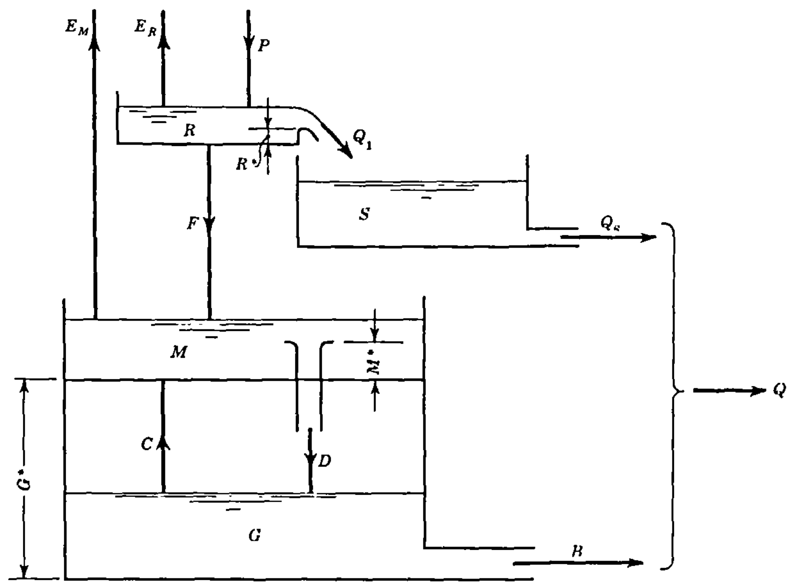

FIG. 1.—SCHEMATIC DIAGRAM OF THE OVER-ALL MODEL OF THE HYDROLOGIC CYCLE

Full details of the operating rules governing the behavior of this simple model will not be given. The rules chosen are simple ones; the most complex treatment is that which determines the infiltration procedure. The principal component model, described subsequently, was concerned with the infiltration process, and had some feedback influence on the over-all model. The study was primarily concerned with parameter optimizing techniques; this presentation will concentrate on that aspect rather than on justifying the operating rules.

In brief, there are nine parameters that control the functioning of the model. At the beginning of each interval, the volume in R lies between zero and R*, the first parameter; P is added to R; and E_R, if any, is then given first call on the sum. Next, F is calculated according to certain criteria, based on a Horton-type equation,[11] considering the rate of supply available from surface storage and the potential rate of infiltration at the start of the interval. This involves maximum and minimum infiltration rates, f_o and f_c, and an exponential die-away exponent, k (three more parameters). In preparation for the next interval, a potential rate of infiltration, f_i, is calculated for the end of the current interval. Then, Q_1 is determined to be the excess, if any, over R* left in surface storage, after E_R and F have been abstracted.

The channel storage, S is assumed to be a linear storage having a storage constant, K_S, the fifth parameter. Then, Q_S is a function of the volume in S at the beginning of the interval, of the inflow, Q_1, and of K_S. A simple budget yields the volume left in S ready for the start of the next interval.

At the beginning of an interval, M lies between zero and M*, the sixth parameter. Either E_M is removed or F is added, for one of the two will be zero depending on whether or not E_R satisfied E_P, the potential evapotranspiration. One of several alternatives is now followed depending on whether or not G, at the start of the interval, is greater than G*, the seventh parameter, and, if not, whether or not the quantity in M is now greater than M*. If G is less than G*, D is set equal to the excess, if any, over M* now in M; C is zero if D exists, otherwise it is determined as a function of demand in M, of supply in G, and of a maximum rate of rise, c_{max}, the eighth parameter.

Also, M is left at M* if D exists, or is augmented by C, if not. If G, at the beginning of the interval, is greater than G*, F, if any, acts on G directly in place of D, and C similarly in place of E_M. In this alternative, M remains at M*.

Then, G is assumed to be a linear storage having a storage constant, K_G (the ninth parameter); B is then a function of the volume in G, at the start of the interval, of the inflow D or abstraction C, and of K_S. Again, a budget yields the volume left in G ready for the start of the next interval.

At the beginning of the first interval, the volumes in each of the four storage elements and the potential infiltration rate have to be specified, in addition to a set of values of the nine parameters, for a synthesis to be initiated. Thereafter, the computations for each interval yield the four storage volumes and the potential infiltration rate for the beginning of the next interval. A completely general optimization would include the start-of-synthesis values of these five quantities together with the nine parameter values. However, by postulating a long period with no rainfall and no streamflow prior to the start of a rainfall-runoff synthesis, it is reasonable to set all four initial storages to zero and to assume that the starting potential infiltration rate has recovered to the maximum value, f_o. In this way, the number of items to be optimised was kept to nine in these preliminary studies.

The input data to the model consists of (1) P, E_P, and Q for each interval of a known record, and (2) estimates of the initial values of the nine parameters. The model then uses the P and E_P data to calculate a runoff volume

---

11 Horton, R. E., "A Simplified Method of Determining the Constants in the Infiltration-Capacity Equation," Transactions, Amer. Geophysical Union, Vol. 23, Part 2, 1942, pp. 575-577.

for each interval of the record, which, in general, will not agree with the known Q values. The optimization technique adjusts the initial parameter values such that the differences between the calculated and known Q values are eliminated.

To free the optimization studies from errors in any real data records, the model was first used to generate a synthetic set of runoff values from a set of arbitrary parameter values and a made-up record of P and E_p values. Not only does this provide error-free data—the synthetically "correct" parameter values towards which any other set of values should be optimized are now known. Because all data an synthetic, only the optimizing methods were studied with these data. Any study of hydrologic validity of the model must use measured rather than synthesized data.

Results.

Sensitivity.—Table 2 shows typical results for the sensitivity of the model response to 1% changes in each of the parameters. Table 2 compares a 240-step record that consisted of a sequence of 20 discrete rainfall events all followed by several intervals of no rainfall (but appreciable evaporation), with a 60-step record with a sequence of fewer but longer (and somewhat smaller) rainfalls and a few short periods of evaporation. In addition, the value of G* was lowered and that for f_c was raised in the 60-step case. These two cases provide runoff records in which the ratio of surface flow to baseflow is high in the 240-step case and low in the 60-step case. The sensitivity figures in Table 2 are U values, as described earlier, namely, sums of squares of differences between synthesized and correct values of total runoff. The order of the parameters in Table 2 is that of decreasing sensitivity of model response, with the sensitivity figures rounded to one place. The only appreciable difference in the two cases is that K_G has climbed three places, which is a reflection of the increased proportion of baseflow in the 60-step case.

TABLE 2.—RESPONSE SENSITIVITY RECORDS

Response Sensitivity (240-Step Record)			Response Sensitivity (60-Step Record)		
Parameter	Value	1% Sensitivity	Parameter	Value	1% Sensitivity
K_S	10	5×10^{-4}	K_S	10	3×10^{-4}
f_c	0.2	1×10^{-4}	f_c	0.3	3×10^{-5}
f_o	2	4×10^{-5}	K_G	40	1×10^{-5}
M*	2	4×10^{-5}	f_o	2	4×10^{-6}
k	2	2×10^{-5}	M*	2	4×10^{-6}
K_G	40	3×10^{-6}	k	2	2×10^{-6}
R*	0.1	1×10^{-6}	R*	0.1	1×10^{-7}
G*	4	2×10^{-7}	G*	3	4×10^{-9}
c_{max}	0.1	2×10^{-7}	c_{max}	0.1	4×10^{-9}

Optimization.—Most of the development work on the optimization aspect of the over-all model was done with a 60-step record of synthetic data. Table 1 gives the results of a test run with this data in which five complete rounds each of six stages, plus a final incomplete round of five stages, were made. The starting values of the parameters were, with one exception, 50% above or below their correct values.

It will be seen that seven of the nine parameters were optimized to within 15% of their correct values (five are within 3%), while two remain a long way away.

A revealing comparison can be made between Table 1 and Table 2, leading to a conclusion that, in general, the greater the sensitivity of the model response to a parameter, the closer and sooner will that parameter be optimized.

It is of interest to examine the final U value listed in Table 1, namely, 2.91×10^{-6}. This is the sum of the squares of 60 differences, so that on the average, the difference between a correct and a synthesized Q is 2.20×10^{-4}. The mean Q value for this 60-step record is approximately 5×10^{-1}. Thus, the Q record has been fitted, on an average, to within 0.05%. Even if all the difference between the correct and synthesized record were concentrated into 1 of the 60 steps, the error would be on the order of only 0.35%.

Any further development of automatic parameter optimization techniques must use some criterion of response sensitivity (or its equivalent) in selecting what can be considered adequately optimized parameters. Indeed, the minimization of differences from recorded data cannot be the sole criterion in interpreting the fit of any model. This is particularly relevant if the model parameters are to be related to the physical properties of real catchments.

COMPONENT MODEL STUDY

Description.—An infiltration component model was based on a combination of the Horton type equation and an approximation developed by Rubin[12] to the Philip equation. Once runoff begins, infiltration is described by an exponential function. The time at which runoff begins is a function of inherent soil characteristics, soil-moisture content, and rainfall intensity. If soil moisture is known, or if a budgeting program is used to compute soil moisture, the infiltration equation contains three parameters. They are (1) minimum infiltration capacity of the soil (Horton f_c), (2) exponential decay constant (Horton k), and (3) bubbling pressure of soil (Rubin $-H_b$). If only the previous rainfall is known, and is used as an index, the equation contains two further parameters, or maximum soil moisture content of soil column and Antecedent precipitation index.

The sensitivity of the parameters is approximately in the order of their listing. The Horton parameters are a good approximation, the Rubin parameter is a refinement and therefore somewhat less important, and the last two parameters are indexes rather than physical parameters.

Results.

Sensitivity.—A sensitivity run for the three-parameter infiltration model with synthetic data, and known "true" values, is given in Table 3. These re-

12 Rubin, J., and Steinhardt, R., "Soil Water Relations during Rain Infiltration: III. Water Uptake at Incipient Ponding," Proceedings, Soil Science Soc. of Amer., Vol. 28, No. 5, 1964.

sults are for a group of twenty storms, with the fit being made to runoff for each of the storms. For this set of data and this test, k is shown to be somewhat more sensitive than f_c. However, runs show that f_c homes in faster and closer than the other variables.

Optimization.—Table 4 shows the results of two rounds of fitting the model to the twenty storms, beginning with an arbitrary set of parameters. The final U value is approximately 0.3% of the average storm runoff. If the total difference were concentrated in the smallest runoff event, the error for that storm would be but 4%. Thus, for a simple model with few parameters, the fitted values rapidly zero in on the true values. Only about 600 attempts were made in order to have all parameters within 5%. The five-parameter model indicated that many more tries and rounds are needed to obtain similar results for more complex models. Results for the five-parameter model were

TABLE 3.—RESPONSE SENSITIVITY, INFILTRATION COMPONENT MODEL

Parameter	Value	1% sensitivity
f_c	0.05	2×10^{-4}
k	1.0	1×10^{-3}
H_b	0.06	3×10^{-5}

TABLE 4.—PARAMETER OPTIMIZATION, INFILTRATION COMPONENT MODEL

Parameter	Correct value	Starting value	Parameter of value at end of round		Residual difference, in Percent
			1	2	
f_c	0.05	1.0	0.137	0.0503	0.6
k	1.0	1.0	6.45	1.011	1.1
H_b	0.06	1.0	0.0009	0.063	5.0
	U value	621	1.47	0.0008	

intermediate between those for the three-parameter model (Table 4) and the over-all model (Table 1).

Each component study should be as physically meaningful as possible. The parameters should, whenever possible, represent measurable quantities. This is important for two reasons. First, the hydrologist can have an immediate feeling of the realism of a resulting solution. Second, if the parameters are physically meaningful, they can first be checked against field measurements, and can subsequently be estimated from field data for synthesis of data at ungaged sites.

GENERAL ANALYSIS

The results described herein are promising, but not conclusive. The infiltration component model is believed to approximate reality fairly closely, insofar as bulk parameter models can apply to drainage basins. However, as each refinement adds more parameters, with a diminishing net gain in accuracy, the added parameters are less sensitive. Therefore, although a model may appear more realistic, the fitted parameters may reflect reality less and less in their numerical values. Small errors in data may generate large errors in some of the less sensitive parameters. To all appearances, the residual errors in the computed runoff may be quite small, despite the fact that some insensitive parameters are greatly in error. A subsequent study will investigate the effects of errors in data on the derived parameters.

Even in as simple models as those so far used, methods are available for a physical check on the parameter values derived by optimization. Thus, there is some check both on the fitting procedure and on the model, if field data are used in the optimization. For instance, the minimum infiltration capacity of the soil (the saturated permeability) and the bubbling pressure can be determined in the laboratory, and some measure of the first can be made in the field. Assuming linearity, K_S and K_G can be determined from properly chosen recession hydrographs, as are the similar parameters in Linsley's model. Other parameters, such as the choice of an antecedent precipitation index, might be based on general hydrologic knowledge. Some of the more empirical parameters probably will be related to physical measurements only after considerable experience is gained in the modeling of many basins.

The closeness of fit of the computed to the observed data, in the components as well as in any over-all model, is a function of the mathematical model itself (including its hydrologic validity), the accuracy of the data used, the method of fitting the model to the data, and the criteria used for "closeness of fit." The first two are more obviously related to the closeness of fit, but the last two may be equally important. Any fitting method makes assumptions about the model fitted. Various hill-climbing or steepest ascent methods must have a so-called response surface to define the hill it must climb. The shape of the response surface will determine the workability of the optimization scheme, to an extent. Therefore, the last two criteria cited are interrelated, as the response surface is the mathematical statement of the test of closeness of fit.

The criteria of closeness of fit can vary from a purely subjective test (that looks close enough), through a semiobjective suboptimization, such as used in the Stanford model, to a very sophisticated objective test. Generally, the more objective the model, the greater the machine time used for optimization. Also, for a given test, the parameters derived will depend on what data are used. Thus, the infiltration component model may have measurements of hourly rainfall, surface runoff, and soil moisture content. If rainfall is used as an input, and runoff is used as an output, the resulting derived parameters will be best for predicting runoff. If, conversely, rainfall and soil moisture are used, the derived parameters are "best" for predicting soil moisture.

The value of the parameters probably will be different for the two fittings. Certainly, the relative sensitivities of the two sets of parameters will change.

The writers' efforts, to date, have been directed toward using as objective a test of closeness of fit as possible. The reasons for this are two-fold. As machine speed and program efficiency improve, objective fitting through machine computation should become more efficient and more accurate. Hydrologic reasoning need be applied only to final results. Also, if fitted parameters are to be used to correlate with physical properties of drainage basins, the hydrologist must not use his hydrologic knowledge to fit the values of the parameter lest he build in an assumed correlation. Therefore, at the expense of seeming inefficient in the short run, wholly objective fitting techniques have been explored exclusively.

As mentioned earlier, the criteria for closeness of fit have so far been associated solely with the amplitude of the residual differences. Timing has not been considered, although it is an important component; timing must be included eventually.

CONCLUSIONS

The simulation of the land phase of the hydrologic cycle by means of a digital computer program is feasible. As computers improve, this tool will be more extensively used. A major problem in such simulation is the fitting of the parameters in the model. At the present time, (1965) most fitting is by semiobjective means. Wholly objective "hill-climbing" methods are available for fitting, and have been found to be adaptable to many problems. As such methods are refined by the mathematicians and become more efficient, they should gain wide use in hydrologic simulation.

APPENDIX.—NOTATION

The following symbols have been adopted for use in this paper:

B = base flow runoff at the gaging station;
C = contribution to soil moisture storage by capillary rise from ground water storage;
c_{max} = maximum rate of capillary rise;
D = deep percolation from soil moisture zone to ground water storage;
E_M = transpiration from soil moisture;
E_P = potential evapotranspiration;
E_R = evaporation from surface storage;
e = arbitrary length of first step in optimization program;
F = infiltration during interval;
f_c = minimum rate of infiltration;
f_i = potential rate of infiltration at beginning of i^{th} period;
f_o = maximum rate of infiltration;
G = ground water storage;

G^* = threshold amount of ground water storage;
H_b = bubbling pressure of soil;
$i(t)$ = rainfall excess input;
K_G = linear constant for ground water storage;
K_S = linear constant for channel storage;
k = exponential die-away exponent in Horton-type infiltration equation;
M = soil moisture storage;
M^* = maximum soil moisture storage;
n = number of parameters in the computer model;
P = rainfall;
Q = runoff during interval;
Q_1 = channel inflow from surface storage;
Q_S = surface runoff at the gaging station;
$q(t)$ = surface runoff output caused by $i(t)$;
R = surface storage;
R^* = threshold surface storage, analogous to depression storage;
S = channel storage;
t = time;
U = objective function to be optimized;
$u(t)$ = instantaneous unit hydrograph (IUH) of the catchment;
X_i = i^{th} parameter in the computer model;
α = ratio of length of steps in optimization program, if previous step was a success;
β = ratio of length of steps, if previous step was a failure; and
τ = variable of integration.