Journal of Floodplain Management
Floodplain Management Association
FEB. 2004 Vol. 4, No.1

Theodore V. Hromadka II [1]

A Review of Hydrologic Modeling History: Accuracy in Results

Abstract

A review of the history of surface runoff hydrologic models indicates that modeling accuracy is subject to a wide number of variables and conditions. Such variability still continues to impact modeling accuracy in current hydrologic studies. In this paper is provided a brief review of some of the more intense efforts to quantify and qualify modeling accuracy and modeling calibration issues identified in the recent past. These cited case studies may provide an additional resource to the floodplain manager when evaluating floodplain mapping decisions based upon the results from hydrologic models.

A Brief Review of Stormflow Estimation Techniques

Stormflow Determination Methods

When studying a watershed for severe storm runoff characteristics, the usual procedure is to collect data on precipitation, soil types, stream discharge, and other hydrologic and geologic characteristics. This data may then be evaluated in accordance with theory presented in standard texts. Although precipitation and streamflow data are available at selected locations throughout the country (for example the U.S. Weather Service and the U.S. Geological Survey), sufficient data are usually unavailable for local watersheds to develop precise hydrologic calculations. More importantly, the long-term effects on flood hydrology due to urbanization of the watershed are usually not precisely represented by the available data. For these reasons, synthetic flood hydrology methods are usually required. And since the introduction of digital computers, literally hundreds of hydrologic models have been produced.

Method for Development of Synthetic Flood Frequency Estimates

The uses of flood flow frequency data range from the specification of flood insurance risk relationships to the commonly occurring problem of designing flood control facilities. Typically, however, stream gauge data are usually unavailable at the study site; consequently, some type of method is needed to synthesize a flood frequency curve for ungauged streams.

The various types of procedures used to develop flow frequency estimates at ungauged locations can be grouped as follows: (1) Data transfer methods, (2) Statistical methods, (3) Empirical equations, and (4) Simulation models.

---

[1] *Professor of Mathematics, Environmental Studies, and Geological Sciences, California State University, Fullerton, CA*

Because flood flow frequency information is used for various purposes, the hydrologist must be aware of the limitations and factors involved which are associated with each of the groupings of methods. For example, flood flow frequency estimates used for design of flood control facilities often are conservative in that the design discharges are high for the corresponding return frequency. In this fashion, the designer compensates for the unknown reliability of the design flow rate and provides for a factor of safety. For flood insurance studies, however, use of the computed flood flow frequency estimate may be desirable in order to avoid excessively high costs for the corresponding benefit (see U.S. Army Corps of Engineers Training Document No. 11, TD-11, 1980).

Detailed discussions of the several categories of flood flow frequency analysis procedures are contained in TD-11. In that publication, the four groupings of methods are further defined into eight categories as follows:

(I) statistical estimation of peak flowrates
(II) statistical estimation of moments
(III) index flood estimation methods
(IV) transfer methods
(V) empirical equations
(VI) single event methods
(VII) multiple discrete event methods
(VIII) continuous simulation methods.

Advantages and disadvantages of methods in each of these eight categories are discussed in the following paragraphs.

Category I: Statistical estimation of peak flowrate (Q_P) methods use regression equations for determining a specific return frequency of flowrate by correlating stream gauge data to watershed characteristics. Ungauged stream flowrate estimates can then be obtained from the regression equations. Table 1 (TD-11, 1980) compares the advantages and disadvantages associated with this category of methods.

Category II: The statistical estimation of moments procedure extends the procedures of Category I by correlating the statistical moments of the frequency function developed from the stream gauge data to watershed characteristics. Table 2 (TD-11, 1980) lists the advantages and disadvantages of this category of methods.

Category III: Index flood estimation methods (see Table 3) are analogous to the above two categories except that a selected index flood, such as the mean annual event, is used for the development of the necessary statistical relationships for events other than the index event.

Category IV: Transfer methods (Table 4) usually refer to the relationships used to estimate flowrates immediately upstream or downstream of a stream gauge location. However TD-11 broadens this category to include procedures for the direct transfer of peak flood flow frequency values or frequency functions from similar gauge locations to the subject study point.

Table 1. Statistical Estimation of Q_P (Category I)
Reference: TD-11 (1980)

Applicability/Advantages	Limitations/Disadvantages
• Procedures are based on accepted statistical methods. • Procedures are available for most of the country. • Reliability of the prediction equations is known for gauged areas used in derivation. • Estimates are reliable for hydrologically similar basins as those used in the derivation. • Once developed, the procedure is quick and easy to use. • Permits direct calculation of specific peak flood flow frequency estimates that are individually and statistically derived. • Procedures may be used in conjunction with other procedures such as to provide calibration relationships for simulation models. • Provides a quick check for reasonableness for situations requiring use of other procedures.	• Requires knowledge of both statistics and hydrology in derivation and utilization. • Procedures require numerous regression analyses and are time consuming to develop. • Only provides estimates of specific peak flood flow frequency relationships. • Cannot evaluate effects resulting from modifications in the system (physical works and alternative land use patterns). • Procedures are often misused by application for areas with different stream patterns and other hydrologic characteristics from the gauged locations used in the derivation. • Cannot adequately evaluate hydrologically unique areas in the region. • Easy to use therefore may be used where other methods would be more appropriate. • Derivation requires several hydrologically similar gauged basins in the region. • Does not assume a distribution; hence reliability confidence limits cannot be calculated.

Table 2. Statistical Estimation Of Q_p (Category II) Reference: TD-11 (1980)

Applicability/Advantages	Limitations/Disadvantages
• Procedures are based on accepted statistical methods. • The entire frequency function is developed from the three moments; means, standard deviation and skew. • Reliability of the prediction equations is known for gauged areas used in derivation. • Estimates are as reliable for hydrologically similar basins as those used in derivation. • Once developed, the procedure is quick and easy to use. • Procedures may be used in conjunction with other procedures, such as, to provide calibration results for simulation models. • Provides a quick check for reasonableness for situations requiring use of other procedures.	• Requires knowledge of both statistics and hydrology in derivation and utilization. • Procedure requires regression analysis for the two or three moments of the frequency. • May be time consuming to develop. • Does not calculate specific flood flow frequency events. • Only provides estimates of peak flood flow frequency relationships. • Cannot evaluate effects resulting from modifications in the system (physical works and alternate land use patterns). • Cannot adequately evaluate many complex river systems. • Cannot evaluate hydrologically unique areas in the region. • Ease of use may result in improper application. • Derivation requires several hydrologically similar gauged basins in the region.

Table 3. **Index Flood Estimate (Category III)**
Reference: TD-11 (1980)

Applicability/Advantages	Limitations/Disadvantages
• Procedure is easier to develop than other statistical methods, and has only one regression analysis.	• Procedure yields same variance (slope of frequency curve) for all applications.
• Procedures are commonly used and based on accepted statistical methods.	• Probably least accurate of the statistical procedures.
• Reliability of prediction equation for index flood is known for derivation.	• Requires knowledge of both statistics and hydrology in derivation and utilization.
• Estimates are reliable for hydrologically similar basins as those used in derivation.	• May be time consuming to develop.
• Once developed, the procedure is quick and easy to use.	• Only provides estimates of peak flood flow frequency relationships.
• Procedures may be used in conjunction with other procedures, such as, to provide calibration results for simulation models.	• Cannot evaluate effects resulting from modifications in the system (physical works and alternative land use patterns).
• Provides a quick check for situations requiring use of other procedures.	• Cannot adequately evaluate many complex river systems.
	• Cannot evaluate hydrologically unique areas in the region.
	• Ease of use may result in improper application.
	• Derivation requires several hydrologically similar gauged basins in the region.

Category V: Empirical equations are often used for the estimation of peak flowrates. The well-known rational method is an important example of this category. Table 5 (TD-11) compares the advantages and disadvantages of this group of methods.

Category VI: Single event methods are the most widely used approach for developing runoff hydrographs which are subsequently used to develop a flood flow frequency curve. Incorporated in this category are the design storm methods which attempt to relate runoff and rainfall frequency curves. Table 6 from TD-11 examines several features of this category of methods.

Category VII: By considering a series of important record storm events with a single event method, an approximate flood frequency curve can be developed. The multiple discrete event category (see Table 7) of models serves as a blend of the single event category of models and the concept of continuous simulation.

Category VIII: Continuous simulation (or continuous record) models attempt to develop a continuous streamflow record based on a continuous rainfall record. Although in concept this category (see Table 8) of models appears to be plausible, the success of these methods has not been clearly established due to the lack of evidence that this approach out performs the much simpler and more often used unit hydrograph procedures of Category VI.

Table 4. **Transfer Methods (Category IV)**
Reference: TD-11 (1980)

Applicability/Advantages	Limitations/Disadvantages
(WRC Transfer of Q_P)	(WRC Transfer of Q_P)
• Procedure us easy and quick to use.	• Ease of use may result in improper application.
• Provides reliable estimates immediately upstream and downstream of gauge location if hydrologic characteristics are consistent.	• Can only be utilized immediately upstream and downstream of gauged area where hydrologic characteristics are consistent.
• Procedure is commonly used and generally acceptable.	
(Direct Transfer)	(Direct Transfer)
• Provides quick estimate where time constraints are binding and other procedures are not applicable.	• Estimates are not accurate enough for most analysis requirements.
• Can readily be used as a check for reasonableness of results from other procedures.	• Cannot be used for modified basin conditions.
• Provides valuable insight as to the regional slope characteristics of the flood flow frequency relationships.	• Can only be used as check in areas where hydrologic characteristics are nearly similar and with drainage areas within the same order of magnitude.

Table 5. Empirical Equations (Category V)
Reference: TD-11 (1980)

Applicability/Advantages	Limitations/Disadvantages
• Provides quick means of estimating peak discharge frequency for small areas.	• Generally are not applicable for areas greater than one square mile.
• Concepts can be understood by nonhydrologists.	• Estimate only the peak discharge frequency relationships.
• Suitable for many types of municipal engineering analyses (storm sewers, culverts, small organizations impacts, etc.).	• Cannot be used to design storage facilities.
• Familiarity of procedures and use had led to politically acceptable solutions for small areas.	• Cannot adequately evaluate complex systems where timing and combining of flood hydrographs are important.
• Can be used as a check for reasonableness of more applicable procedures in small areas.	

Table 6. **Single Event Simulation (Category VI)**
Reference: TD-11 (1980)

Applicability/Advantages	Limitations/Disadvantages
• Generates other hydrologic information rather than peak discharges (volumes, time to peak, rate of rise, etc.).	• Balanced flood concept is difficult to understand.
• Generates balanced floods as opposed to historically generated events which may be biased.	• Modeling requires more time, data, and resources (costs) than statistical procedures.
• Enables evaluation of complex systems and modifications to the watersheds.	• Hydrologists must understand the concepts utilized by the model.
• Provides good documentation for quick future use.	• Requires calibration to assure rainfall frequency approximates runoff frequency.
• Uses fewer parameters than most continuous simulation models.	• Unit hydrograph assumes a linear relationship with runoff.
• Approximates the hydrologic runoff process as opposed to statistical methods.	• Requires data processing capabilities.
• Procedures are more economical than continuous simulation procedures.	• Procedures greatly simplify the hydrologic process.
• Calibration procedures are easier than continuous simulation models.	• Procedures are generally limited to basins greater than one square mile.
• Models may be calibrated to either simple or complex systems.	• Parameters are difficult to obtain for existing and modified conditions.
	• Difficult to obtain antecedent moisture conditions.
	• Depth-area of rainfall varies with drainage area size.

Table 7. Multiple Discrete Events (Category VII)
Reference: TD-11 (1980)

Applicability/Advantages

- Concepts are easier to understand than those associated with hypothetical frequency events.
- Antecedent moisture conditions are determined.
- Depth-area precipitation problems are eliminated.
- Evaluates fewer events than continuous simulation models.
- Enables evaluations of complex systems and physical modifications in the watershed.
- Uses fewer parameters than continuous simulation models.
- Approximates hydrologic process as opposed to statistical methods.
- Provides good documentation for future use.

Limitations/Disadvantages

- Requires numerous storm analyses and subsequent event analyses.
- Important events may be overlooked.
- Results may be biased by historic records.
- Procedures use simplified hydrologic process.
- Requires data processing capabilities.
- Parameters are difficult to obtain.
- Unit hydrograph assumes linear relationship with runoff.
- Requires calibration which is more time consuming than single event due to the large number of events that are processed.
- Procedure is significantly more expensive than single event modeling.
- Procedures generally not feasible for small study areas, short time constraints, etc.

Table 8. Continuous Simulation (Category VIII)
Reference: TD-11 (1980)

Applicability/Advantages	Limitations/Disadvantages
• Concepts are easily understood.	• The calibration process is extensive and generally must be performed by qualified experienced hydrologists.
• Concepts are more physically based than other procedures.	• Procedures are expensive and time consuming to use, impractical for moderate or small resources allocated projects.
• Antecedent moisture conditions are automatically accounted for.	• The results may be biased by the use of historic rainfall data.
• Can be used in unique basins where other procedures such as statistical procedure are not applicable.	• The procedures require large analytical processing capabilities.
• Process analyses in single computer runs as numerous discrete events.	• The models typically require a large amount of data to properly define the parameters.
• Can automatically determine annual peak floods at various locations even if their frequencies are different.	
• Can model the effects of complex systems and physical works.	

A Brief Review Of Some Stormflow Estimation Uncertainty Issues

Watershed Modeling Uncertainty

Watershed runoff is a function of rainfall intensity, the storm duration, the infiltration capacity of the soil, the cover of the soil, type of vegetation, area of the watershed and related shape factors, distribution of the storm with respect to space and time, watershed stream system topology, connectivity and branching, watershed geometry, stream system hydraulics, overland flow characteristics, and several other factors. Because of the dozens of variables which are included in a completely deterministic model of watershed runoff and due to the uncertainty which is associated to the spatial and temporal values of each of the various mathematical definitions, urban hydrologists need to include a measure of uncertainty in predicting surface runoff quantities.

With the widespread use of minicomputers and inexpensive microcomputers, the use of deterministic models is commonplace. These models attempt to simulate several of the most important hydrologic variables that strongly influence the watershed runoff quantities produced from severe design storm events. Generally speaking, the design storm (e.g., single event) and continuous simulation models include approximations for runoff hydrograph generation (coupled with models for estimating interception, evapotranspiration, interflow, and infiltration), channel routing, and detention basin routing. The computer program user then combines these processes into a link-node schematic of the watershed. Because each of the hydrologic processes involves several parameters, the resulting output of the model, the runoff hydrograph, may be a function of several dozen parameters. In a procedure called calibration, many or all of the parameters are estimated by attempts to duplicate significant historical runoff hydrographs. However, Wood (1976) notes that the watershed model parameter interaction can result in considerable difficulty in optimizing the parameter set. In a similar deterministic modeling approach for soil systems and soil water movement, Guymon et al. (1981) found that just the normal range of uncertainty associated with laboratory measurement of groundwater flow hydraulic parameters can produce considerable variation in the model output. A detailed analysis of the sensitivity corresponding to a watershed model is given by Mein and Brown (1978). Because of the vast spectrum of rainfall-runoff models available today, it is appropriate to review some of the comments noted in the literature as to the relative success of rainfall-runoff models in solving the runoff estimation problem in a purely deterministic setting.

Some Concerns in Deterministic Rainfall-Runoff Model Performance

Due to the need for developing runoff hydrographs for design purposes, statistical methods such as those contained in model categories I-V are usually precluded in watershed hydrologic studies. Consequently, the categories of models available are essentially restricted to categories VI, VII, and VIII. The "single event" models directly transform a design storm (hypothetical causative input) into a flood hydrograph. The "multiple discrete event" models transform an annual series of selected discrete rainfall events (usually one storm for each year) into an annual series of runoff hydrographs whose peak flowrates are used for subsequent statistical analysis. The "continuous record" or "continuous simulation" model results in a continuous record of synthetic runoff hydrographs for statistical synthesis. Each of the above three categories of deterministic models contain various versions and modifications which range widely in complexity, data requirements, and computational effort.

In general, the well-known unit hydrograph design storm approach has continued widespread support among practitioners and governmental agencies involved in flood control design. Such general purpose models include the U.S. Department of Agriculture, Soil Conservation Service or SCS model (1975) and the U.S. Army Corps of Engineers (HEC) hydrology computer program package (see TD-15, 1982). In a recent survey of hydrologic model usage by Federal and State governmental agencies and private engineering firms (U.S. Department of Transportation, Federal Highway Administration Hydraulic Engineering Circular No. 19, October, 1984), it was found that "practically no use is made of watershed models for discrete event and continuous hydrograph simulation." In comparison, however, design storm methods were used from 24 to 34 times more frequently than the discrete event or continuous simulation models by Federal agencies and the private sector, respectively. The frequent use of design storm methods appears to be due to several reasons: (1) design storm methods are considerably simpler to use than discrete event and continuous simulation models; (2) it has not been established in general that the more complex models provide an improvement in computational accuracy over design storm models; and (3) the level of complexity typically embodied in the continuous simulation class of models does not appear to be appropriate for the catchment rainfall-runoff data which is typically available. Consequently, the design storm approach continues to be the most often selected for flood control and drainage design studies.

A criterion for classifying a model as being simple or complex is given by Beard and Chang (1979) as the "difficulty or reliability of model calibration.... Perhaps the simplest type of model that produces a flood hydrograph is the unit hydrograph model"...and... "can be derived to some extent from physical drainage features but fairly easily and fairly reliably calibrated through successive approximations by relating the time distribution of average basin rainfall excess to the time distribution of runoff." In comparison, the "most complicated type of model is one the represents each significant element of the hydrologic process by a mathematical algorithm. This is represented by the Stanford Watershed Model and requires extensive data and effort to calibrate."

The literature contains several reports of problems in calibrating complex models, especially in parameter optimization. Additionally, it has not been clearly established whether complex models, such as in the continuous simulation or discrete event classes of models, provide an increase in accuracy over a simple single event unit hydrograph model. There are only a few papers and reports in the literature that provide a comparison in hydrologic model performance. From these references, it appears that a simple unit hydrograph model oftentimes provides estimates of runoff quantities which are comparable to considerably more complex rainfall-runoff models.

In their paper, Beard and Chang (1979) write that in the case of the unit hydrograph model, "the function of runoff versus rainfall excess is considered to be linear, whereas it usually is not in nature. Also, the variations in shapes of unit hydrographs are not derivable directly from physical factors. However, models of this general nature are usually as representative of physical conditions as can reasonably be validated by available data, and there is little advantage in extending the degree of model sophistication beyond validation capability."

Schilling and Fuchs (1986) write "that the spatial resolution of rain data input is of paramount importance to the accuracy of the simulated hydrograph" due to "the high spatial variability of storms" and "the amplification of rainfall sampling errors by the nonlinear transformation" of rainfall into runoff. Their recommendations are that a rainfall-runoff model should employ a simplified surface flow model if there are many subbasins; a simple runoff coefficient loss rate; and a diffusion (zero inertia) or storage channel routing technique.

In attempting to define the modeling processes by the available field data forms, Hornberger et al (1985) find that "Hydrological quantities measured in the field tend to be either integral variables (e.g., stream discharge, which reflects an integrated catchment response) or point estimates of variables that are likely to exhibit marked spatial and/or temporal variation (e.g., soil hydraulic conductivity)." Hence, the precise definition of the physics in a modeling sense becomes a problem that is "poorly posed in the mathematical sense." Typically, the submodel parameters cannot be estimated precisely due to the large associated estimation error. "Such difficulties often indicate that the structural complexity of the model is greater than is warranted on the basis of the calibration data set." It was also noted by Hornberger et al (1985) that success in rainfall-runoff modeling "has proved elusive because of the complexity of the processes, the difficulty of performing controlled experiments, and the spatial and temporal variability of catchment characteristics and precipitation." They concluded that "Even the most physically based models...cannot reflect the true complexity and heterogeneity of the processes occurring in the field. Catchment hydrology is still very much an empirical science."

Schilling and Fuchs (1986) note that errors in rainfall-runoff modeling occur for several reasons, including:

"1. The input data, consisting of rainfall and antecedent conditions, vary throughout the watershed and cannot be precisely measured.
1. The physical laws of fluid motion are simplified.
2. Model parameter estimates may be in error."

By reducing the rainfall data set resolution from a grid of 81 rain gauges to a single catchment-centered rain gauge in an 1,800 acre catchment (Fig. 1), variations in runoff volumes and peak flows "is well above 100 percent over the entire range of storms implying that the spatial resolution of rainfall has a dominant influence on the reliability of computed runoff." It is also noted that "errors in the rainfall input are amplified by the rainfall-runoff transformation" so that "a rainfall depth error of 30 percent results in a volume error of 60 percent and a peak flow error of 80 percent." Schilling and Fuchs (1986) also write that "it is inappropriate to use a sophisticated runoff model to achieve a desired level of modeling accuracy if the spatial resolution of rain input is low."

Similarly, Beard and Chang (1979) write that in their study of 14 urban catchments, complex models such as continuous simulation typically have 20 to 40 parameters and functions that must be derived from recorded rainfall-runoff data. "Inasmuch as rainfall data are for scattered point locations and storm rainfall is highly variable in time and space, available data are generically inadequate...for reliably calibrating the various interrelated functions of these complex models." Additionally, "changes in the model that would result from urbanization could not be reliably determined." Beard and Chang (1979) write that the application "of these complex models to evaluating changes in flood frequencies usually requires simulation of about 50 years of streamflow at each location under each alternative watershed condition."

A Brief Review Of Some Stormflow Model Calibration Issues

Garen and Burges (1981) noted the difficulties in rainfall measurement for use in the Stanford Watershed Model, because the K1 parameter (rainfall adjustment factor) and UZSN parameter (upper level storage) had the dominant impact on the model sensitivity. This is especially noteworthy because Dawdy and O'Donnell (1965) concluded that insensitive model coefficients could not be calibrated accurately. Thus, they could not be used to measure physical effects of watershed changes.

In the extensive study by Loague and Freeze (1985), three event-based rainfall-runoff models (a regression model, a unit hydrograph model, and a kinematic wave quasi-physically based model) were used on three data sets of 269 storm events from three small upland catchments. In that paper, the term "quasi-physically based" or QPB is used for the kinematic wave model. The three catchments were 25 acres, 2.8 square-miles, and 35 acres in size, and were extensively monitored with rain gauge, stream gauge, neutron probe, and soil parameter site testing.

For example, the 25 acre site instrumentation (Fig. 2) contained 35 neutron probe access sites, 26 soil parameter sites (all equally spaced), an on-site rain gauge, and a stream gauge. The QPB model (Fig. 3) utilized 22 overland flow planes and four channel segments. In comparative tests between the three modeling approaches to measured rainfall-runoff data it was concluded that all models performed poorly and that the QPB performance was only slightly improved by calibration of its most sensitive parameter, hydraulic conductivity. They write that the "conclusion one if forced to draw...is that the QPB model does not represent reality very well; in other words, there is considerable model error present. We suspect this is the case with most, if not all conceptual models currently in use." Additionally, "the fact that simpler, less data intensive models provided as good or better predictions that a QPB is food for thought."

Based on the literature, a major difficulty in the use, calibration, and development of rainfall-runoff models appears to be the lack of precise rainfall data and the high model sensitivity to (and magnification of) rainfall measurement errors. Nash and Sutcliffe (1970) write that "As there is little point in applying exact laws to approximate boundary conditions, this, and the limited ranges of the variables encountered, suggest the use of simplified empirical relations."

It is noteworthy to consider the HEC Research Note No. 6 (1979) where the Hydrocomp HSP continuous simulation model was applied to the West Branch DuPage River in Illinois. Personnel from Hydrocomp, HEC (U.S. Army Corps of Engineers, Hydrologic Engineering Center) and COE (U.S. Army Corps of Engineers) participated in this study which started with a nearly complete hydrologic/meteorologic data base. The report stated that "It took one person six months to assemble and analyze additional data, and to learn how to use the model. Another six months were spent in calibration and long-record simulation." This time allocation applies to only a 28.5 square-mile basin. The quality of the final model is indicated by the average absolute monthly volume error of 32.1 and 28.1 percent for calibration and verification periods, respectively. Figure 4 shows a typical comparison of modeled and measured results. Peak flow rate absolute errors were 26 and 36 percent for calibration and verification periods, respectively. It was concluded that "Discharge frequency under changing urban conditions is a problem that could be handled by simpler, quicker, less costly approaches requiring much less data; e.g., design storms or several historical events used as input to a single-event model, or a continuous model with a less complex soil-moisture accounting algorithm."

In another study, HEC Technical Paper No. 59 (Abbott, 1978) compared six hydrologic models, plus two variants of one and a variant of another, in a preliminary evaluation of their relative capabilities, accuracy and ease of application on a 5.5 square-mile urban watershed near Oakland, California. Four continuous simulation models were tested: Storage Treatment Overflow Runoff Model (STORM), Hydrocomp Simulation Program (HSP), Streamflow Synthesis and Reservoir Regulation (SSARR), and Continuous Flood Hydrographs (HEC-IC). Single-storm event comparisons were made using STORM, HSP, SSARR, Storm Water Management Model (SWMM), Massachusetts Institute of Technology Catchment Model (MITCAT) and the HEC-1 unit hydrograph model (single area analysis). Each model was calibrated with the first 40 percent of a 42 month record, and the resulting calibration coefficients were used in simulating the remaining record. The study results showed that the more complex models did not produce better results in developing watershed runoff quantities than the simple models for this test watershed (see Fig. 5).

In the absence of more encouraging results in the use of complex hydrology models, the widespread use and continued acceptance of simpler rainfall-runoff models such as unit hydrograph methods for the estimation of watershed runoff quantities is understandable. For a new rainfall-runoff modeling approach to achieve widespread acceptance, it must clearly demonstrate a superiority in performance. For example, Hall (1984) writes that some predetermined criterion of "goodness-of-fit" is typically used to assess a new model's capability in reproducing historic storm event runoff quantities. The new model is first calibrated to observed rainfall-runoff data and then "verified" using storm events excluded from the calibration storm event data set. This type of split-sample testing (for example, TP-59, 1978; Loague and Freeze, 1985) has been found to be a standard in comparing rainfall-runoff model performance.

A second set of criteria must be evaluated when using a new rainfall-runoff model for design storm flood estimation. Model parameters must be correlated to watershed characteristics, or regional values of the parameters must be established. More specifically, the model parameters used as the dependent variables must provide a relationship between the return frequency of runoff and the return frequency of the input rainfall. Acceptance of any new modeling technique typically depends upon the models ease of use and reproducibility of the results by different engineers and hydrologists. Hall (1984) concludes that "until the additional steps required to develop a rainfall-runoff model into a flood estimation method are more widely appreciated, this apparent reluctance to accept innovation is liable to remain a feature of design practice."

The lack of success in concluding a purely deterministic rainfall-runoff modeling approach for developing watershed runoff quantities has motivated the proliferation of dozens of complex, conceptual or so-called physically-based models. However, based upon the available literature, the weight of evidence indicates that use of simpler models such as the well-known unit hydrograph approach will continue to be the most widely used modeling technique. It appears as though the simpler models are able to represent a considerable amount of the explainable phenomena that frequently occurs, and the improvement in modeling accuracy due to inclusion of additional complexity is oftentimes overwhelmed by the scale of uncertainty which cannot be reduced. In a study of stochastic hydrologic methods, Klemes and Bulu (1979) write that often modelers "sidestep the real problem of modeling – the problem of how well a model is likely to reflect the future events – and divert the user to a more tractable, though less useful, problem of how to construct a model that will reproduce the past events. In so doing they expect, and perhaps rightly so, that by the time the prospective modeler has dug himself out of the heaps of technicalities, he either will have forgotten what the true purpose of modeling is or will have invested so much effort into the modeling game that he would prefer to avoid questions about its relevance." According to Gburek (1971), "...a model system is merely a researcher's idea of how a physical system interacts and behaves, and in the case of watershed research, watershed models are usually extremely simplified mathematical descriptions of a complex situation...until each internal submodel of the overall model can be independently verified, the model remains strictly a hypothesis with respect to its internal locations and transformations...".

The current thrust in development of rainfall-runoff models is towards being physically based in that they model all the several components of the hydrologic cycle in rainfall-runoff processes. However the resulting products "...are simplified nonlinear, lumped parameter, time-invariant, discontinuous representations of a complex nonlinear, distributed parameter, time-variant and continuous system" (Sorooshian and Gupta, 1983). The use of a lumped parameter approach means that a characteristic or representative value of a parameter is assumed to apply for the entire watershed, for each parameter used in the model. The invariant parameter assumption assumes that all parameters are constant with respect to seasonal moisture changes. Rain gauge data are also lumped by some selected procedure which ignores the time and spatial variations of rainfall over the watershed, and between storm events. Watt and Kidd (1975) write that the differences between physically based and so-called "black-box" models, (e.g., unit hydrograph models), become less obvious when applied to a field situation. The authors conclude that the considerations of whether the model is physically based or is a black box model "should carry very little weight in the selection process."

Another major issue involving use of rainfall-runoff models is that each of these models requires a calibration of the model parameters be performed in order to obtain an optimum parameter set. However, currently there is no proven technique to obtain this true optimum parameter set.

A brief summary of the success and failures in calibration of model parameters is contained in Sorooshian and Gupta (1983) who write

> "In a recent paper, Alley et al. (1980) stated that 'many of these models have been developed as intellectual exercises rather than useful tools for practicing engineers'. They stressed the need for a balance between (1) processes and (2) the operational characteristics of the model affecting its utility for practical applications. Moore and Clarke (1981) expressed a similar concern by stating that 'it is no exaggeration to say that the present state of rainfall-runoff modeling is extremely fragmented'. Among the reasons they provided in support of the above statement are (1) the difficulty in the selection (i.e., among the many models available) of the 'right model' by a potential user and (2) the difficulty encountered in the calibration of the selected model, using an 'automatic' approach. With respect to the latter difficulty they reference the work of Johnston and Pilgrim (1976) and Pickup (1977) with the Boughton model. The most important conclusion of the work of Johnston and Pilgrim was their inability, in over two years of full-time effort, to find a 'true optimum' parameter set for a nine-parameter version of the Boughton model on the Lidsdale 2 catchment in Australia. Perhaps more disturbing is the fact that even under ideal conditions (created by assuming a perfect set of parameters and using synthetic data), Pickup (1977) was unable (using an automatic approach) to obtain the 'true' values of the Boughton model's parameters. Worth mentioning is the fact that Ibbitt (1970), working with a version of the Stanford watershed model, experienced the same difficulty."

The study of Johnston and Pilgrim (1976) highlighted the complexities associated to determining the optimum parameter set for a conceptual model, and although the Boughton model was used, it was concluded that "most of the findings are applicable to all rainfall-runoff models." Their study identified nine levels of difficulty in optimizing a parameter set, most of which are related to parameter interdependence and the use of a specific objective function to optimize the parameters. They conclude that "until more confidence can be placed in the derivation of truly optimum values, some doubt must remain on the potential usefulness of rainfall-runoff models." When attempting to calibrate a simulation model to model-produced runoff data, Gupta and Sorooshian (1983) reported that "even when calibrated under ideal conditions, it is often impossible to obtain unique estimates for the parameters."

In another examination of the 13-parameter Boughton model, Mein and Brown (1978) examine the conceptual rainfall-runoff model's sensitivity to variations in each parameter of the 'optimized' parameter set. They conclude that "relationships derived between any given parameter value and measurable watershed characteristics would be imprecise, i.e., they would have wide confidence limits" and that "one could not be confident therefore in changing a particular parameter value of this model and then claiming that this alteration represented the effect of some proposed land use change. On the other hand, the model performed quite well in predicting flows with these insensitive parameters, showing that individual parameter precision is not a prerequisite to satisfactory output performance."

Dawdy and Bergmann (1969) identify two categories of error which impact rainfall-runoff models, namely, errors in the estimation of an optimum parameter set and errors resulting due to the unknown variability and intensity of rainfall and storm volume over the watershed. The second error category "places a limit of accuracy upon simulation results," even given the true long-term parameter set. The study concluded that for the test 9.7 square-mile California watershed, using data from a single rain gauge whose data had been adjusted to represent mean basin conditions, the prediction of flood peaks could not be made better than about 20 to 25 percent using a rainfall-runoff simulation model.

Ideally, a dense network of rain gauges within the watershed should be used to determine the spatial and temporal variation in storm rainfalls for each storm event. However, usually only one or two gauges are available, and often not within the watershed. "Even if measurements from a single gauge may be assumed to be representative of overall basin precipitation in an expected value sense, other statistical properties of point rainfall, mainly variability, will differ considerably from the corresponding properties of average basin rainfall. The result can be serious errors in runoff prediction and large biases in parameter estimates obtained by calibration of the model" (Troutman, 1982).

Indeed, rainfall measurement errors at the rain gauges themselves provide a source of concern (see for example, Kelway, 1975). "For single rainfall events, where the totally catch exceeded 12mm (0.5 inch), the error ranged between 0 and 75 percent, depending on wind characteristics during the storm," (Neff, 1977).

Another source of difficulty in the determination of the true optimum parameter set is the optimization procedure used during the calibration process, that is, the so-called objective function which is to be minimized. "The choice of the set of data and of the objective function to be used for any given model is a subjective decision which influences the values of the model parameters and the performance of the model," (Diskin and Simon, 1977).

Pilgrim (1986) writes that "Another approach uses a watershed model to simulate either a long flow record from continuously recorded rainfall, or a series of historical floods from the rainfall recorded in the major storms on the basin. While they are attractive theoretically, none of these approaches is used widely at present, and it is unlikely that any will make serious inroads on the use of a single design flood in the foreseeable future."

Pilgrim notes that "There has been a tendency for researchers to develop complex models of what they assume or imagine happens on real watersheds based on limited data. The enshrinement of procedures in sophisticated models may then lead to general acceptance that nature does actually behave in the assumed manner."

Conclusions

In conclusion, this author agrees with the statements from the several cited authors that complex, multi-parameter models are oftentimes not warranted in view of the success of simpler models. Moreover, less complex models appear to make it easier to quantify model output uncertainty and to offer less chance for error by the user. A brief review of recent but historic efforts to evaluate uncertainty in surface runoff hydrologic models is provided in this paper. Also discussed are calibration issues related to such hydrologic models. The issues identified still continue to apply to hydrologic modeling efforts conducted today for floodplain management purposes. Such uncertainty in modeling results may be considered in the development of floodplain management decisions.

References

1. Alley, W.M., Dawdy, D.R. and Schaake, J.C., Parametric-Deterministic Urban Watershed Model, Journal of The Hydraulics Division, ASCE, Vol. 106, HY 5, pp 679-690, 1980.

2. Beard, L. and Chang, S., Urbanization Impact on Streamflow, Journal of The Hydraulics Division, ASCE, Vol. 105, HY 6, 1979.

3. Dawdy, D. and O'Donnell, T., Mathematical Models of Catchment Behavior, Journal of The Hydraulics Division, ASCE, Vol. 5, No. 5, October 1969.

4. Diskin, M.H. and Simon, E., A Procedure for the Selection of Objective Functions for Hydrologic Simulation Models, Journal of Hydrology, 34, pp 129-149, 1977.

5. Gburek, W.J., Discussion of Hydrologic Consequences of Rainfall Augmentation, Journal of The Hydraulics Division, ASCE, Vol. 97, HY12, pp 2114-2115, 1971.

6. Guymon, G.L., Harr, M.E., Berg, R.L., and Hromadka II, T.V., A Probabilistic-Deterministic Analysis of One-Dimensional Ice Segregation in a Freezing Soil Column, Cold Regions Science and Technology, (5), 1981.

7. Hall, M.J., Urban Hydrology, Elsevier, London, 1984.

8. Hornberger, G.H., Beven, K.J., Cosby, B.J. and Sappington, D.E., Shenandoah Water Shed Study: Calibration of A Topography-Based, Variable Contributing Area Hydrologic Model to a Small Forested Catchment, Water Resources Research, Vol. 21, No. 12, pp 1841-1850, 1985.

9. Johnston, P. and Pilgrim, D., Parameter Optimization for Watershed Models, Water Resources Research, Vol. 12, No. 3, pp 477-486, 1976.

10. Kelway, P., The Rainfall Recorder Problem, Journal of Hydrology, 26, pp 55-77, 1975.

11. Klemes, V. and Bulu, A., Limited Confidence in Confidence Limits Derived by Operational Stochastic Hydrologic Models, Journal of Hydrology, 42, pp 9-22, 1979.

12. Loague, K. and Freeze, R., A Comparison of Rainfall-Runoff Modeling Techniques on Small Upland Catchments, Water Resources Research, Vol. 21, No. 2, pp 229-248, 1985.

13. Mien, R.G. and Brown, B.M., Sensitivity of Optimized Parameters in Watershed Models, Water Resources Research, Vol. 14, No. 2, 1978.

14. Moore, R.J. and Clarke, R.T., A Distribution Function Approach to Rainfall-Runoff Models, Water Resources Bull., 14, pp 777-802, 1978.

15. Nash, J. and Sutcliffe, J., River Flow Forecasting Through Conceptual Models Part 1 – A Discussion of Principles, Journal of Hydrology, Vol. 10, pp 282-290, 1970.

16. Neff, E., How Much Rain Does a Rain Gage Gage?, Journal of Hydrology, 35, pp 213-220, 1977.

17. Pickup, G., Testing the Efficiencies of Algorithms and Strategies for Automatic Calibration of Rainfall-Runoff Models, Hydrol. Sci. Bull., 22(2), pp 257-274, 1977.

18. Pilgrim, D.H., Bridging the Gap Between Flood Research and Design Practice, Water Resources Research, Vol. 22, No. 9, pp 1655-1765, 1986.

19. Schilling, W. and Fuchs, L., Errors in Stormwater Modeling – A Quantitative Assessment, Journal of The Hydraulics Division, ASCE, Vol. 112, HY 2, pp 111-123, 1986.

20. Sorooshian, S. and Gupta, V., Automatic Calibration of Conceptual Rainfall-Runoff Models: The Question of Parameter Observability and Uniqueness, Water Resources Research, Vol. 19, No. 1, pp 260-268, 1983.

21. Troutman, B., An Analysis of Input in Perception-Runoff Models Using Regression With Errors in the Independent Variables, Water Resources Research, Vol. 18, No. 4, pp 947-964, 1982.

22. U.S. Army Corps of Engineers, The Hydrologic Engineering Center, Adoption of Flood Flow Frequency Estimates at Ungaged Locations, Training Document No. 11, 1980.

23. U.S. Army Corps of Engineers, The Hydrologic Engineering Center, Continuous Hydrologic Simulation of The West Branch DuPage River Above West Chicago: An Application of Hydrocomp's HSP, Research Note No. 6, 1979.

24. U.S. Army Corps of Engineers, The Hydrologic Engineering Center, Hydrologic Analysis of Ungaged Watersheds Using HEC-1, Training Document No. 15, 1982.

25. U.S. Army Corps of Engineers, The Hydrologic Engineering Center, Testing of Several Runoff Models on an Urban Watershed, Technical Paper No. 59, 1978.

26. U.S. Department of Transportation, Federal Highway Administration, Hydrology, Hydraulic Engineering Circular No. 19, 1984.

27. Watt, W. and Kidd, C., Quurm-A Realistic Urban Runoff Model, Journal of Hydrology, 27 (1975) 225-235, Elsevier Scientific Publishing Company, Amsterdam-Printed in the Netherlands.

28. Wood, E.F., An Analysis of the Effects of parameter Uncertainty in Deterministic Hydrologic Models, Water Resources Research, Vol. 12, 1976.