SEARCH FOR PHYSICALLY BASED RUNOFF MODEL—A HYDROLOGIC EL DORADO?[a]

Original article

By David A. Woolhiser,[1] Member, ASCE

ABSTRACT: The search for physically based distributed runoff models has been underway for more than half a century. A great deal of progress has been made and we have a much better understanding of hydrological processes today than engineers and scientists did in the 1930s. Physically based models are being used more frequently in engineering practice, but questions are being raised regarding their "physical basis" and accuracy, and it has been suggested that simpler models are superior. Selected papers presenting tests of physically based models or comparisons with simpler models are reviewed. It is shown that, on a relatively small scale, these papers are overly pessimistic due to problems of hydrologic measurement and interpretation. However, there are great difficulties involved in scaling up to larger watersheds and we must realize that significant uncertainties are involved in predicting surface runoff.

INTRODUCTION

I am greatly honored to be named the Hunter Rouse Lecturer for 1994 and to join the highly distinguished group of previous recipients. Unlike some of them, I was never a student of Hunter Rouse, nor were our contacts of a professional nature. In the late 1970s, Hunter Rouse spent the summers in Fort Collins where he taught a graduate course at Colorado State University. At that time we had a small lunch room at the Engineering Research Center and I shared a table with him on many occasions. I recall several pleasant conversations on a variety of topics.

I am struck by the large element of chance involved in receiving an award such as this. Clearly, the factors that condition the probability relate to my good fortune in having outstanding teachers, students, and colleagues, and in being in the right place at the right time. I have been privileged in this regard throughout my career, and I am most grateful for the guidance, encouragement, and friendship I have received. I also wish to acknowledge the unwavering support of my wife, Kathryn.

Much of my career has been devoted to the search for a physically based, distributed surface runoff model and I think that quite a bit of progress has been made. Models of this sort are being used more and more in engineering practice. However, in the past 10 years, several papers have appeared, which seem to show that the predictive ability of these models leaves a great deal to be desired. Indeed, some question whether such a model is possible or desirable. One of the books that I read during my recent stay in Córdoba, Spain, was *The Loss of El Dorado* (Naipaul 1969), a most interesting book in which the author describes some 16th Century Spanish expeditions to search for El Dorado.

> El Dorado, though, is essentially a Spanish delusion. There had been a golden man, *el dorado*, the gilded one, in what is now Columbia: a chief who once a year rolled in turpentine, was covered with gold dust and then dived into a lake. But the tribe of the golden man had been conquered a generation before Columbus came to the New World. It was an Indian memory that the Spanish pursued; and the memory was confused with the legend, among jungle Indians, of the Peru the Spaniards had already conquered. (Naipaul 1969)

This is far from a perfect analogy because no one is expecting to receive great wealth by developing a physically based, distributed runoff model. However, one might raise the question: "Are we searching for something that is unattainable?" In attempting to answer this question, I have reread several review papers dealing with this topic. Included among them are three of my own (Woolhiser 1973, 1975; Woolhiser et al. 1990a). It has been most interesting to follow the evolution of my own attitudes toward these models over the past two decades.

My attitudes are clearly colored by my interests, past research assignments, and experience, so it is appropriate that I summarize them. I have been most interested in surface runoff phenomena on small, agricultural watersheds and have had experience in field measurements of precipitation and surface runoff in Arizona, Wisconsin, Missouri, Iowa, and Colorado, and in indoor and outdoor experiments at Colorado State University. I have done some mathematical modeling of surface runoff and of erosion and chemical transport in small watersheds and have had a limited amount of experience in obtaining data on erosion and water quality in the laboratory and the field.

SOME DEFINITIONS

First, it seems appropriate to consider what we mean when we speak of a physically based distributed runoff model. I assume that it is generally accepted that: "Models, formal or intellectual on the one hand, or material on the other, are . . . a central necessity of scientific procedure." [Rosenblueth and Wiener (1945), p. 316]. From this, it certainly follows that models are essential in engineering practice, where models of physical systems are combined with economic models in an attempt to maximize net social benefits. After all, a model is just our best description of "how things work." A mathematical model belongs to the formal or intellectual class of models and consists of a set $L_1, L_2, \ldots, L_n$ of general laws or theoretical principles and a set $C_1, C_2, \ldots, C_m$ of statements of empirical circumstances (Hempel 1963). The general laws in a runoff model include conservation of mass and possibly the conservation of momentum, while the empirical statements might include Manning's friction relationship for overland or open-channel flow, Darcy's law for flow in porous media, and the system geometry. The most general partial differential equations are written in three space dimensions and time, with the term "distributed" referring to the space dimensions. Al-

---

[a] Presented at the August 1–5, 1994, ASCE Hydraulics Division Conference, held at Buffalo, N.Y.

[1] Fac. Affiliate, Dept. of Civ. Engrg., Colorado State Univ., Fort Collins, CO 80523; formerly Sr. Res. Sci., Dept. of Civ. Engrg., Colorado State Univ., Fort Collins, CO.

Note. Discussion open until August 1, 1996. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on February 21, 1995. This paper is part of the *Journal of Hydraulic Engineering*, Vol. 122, No. 3, March, 1996. ©ASCE, ISSN 0733-9429/96/0003-0122–0129/$4.00 + $.50 per page. Paper No. 10149.

though the equations are continuous, we usually resort to discrete (finite-difference or finite-element) methods to solve them. Thus, even if the physical entities represented by the parameters vary smoothly in space and are constant in time, the parameters are actually lumped to some extent in the model. This is not important when the computational scale is smaller than the scale of the variations, but when the computational scale becomes too large, important variations in the parameter fields will be missed or badly distorted.

We must keep in mind that all models are simplifications or abstractions of reality and all models are to some extent wrong. In fact, if they aren't simpler in some sense than the real-world object, they aren't useful! For this reason we neglect certain aspects of the problem because they are considered to be unimportant. These simplifications should be based on sound physical reasoning or strong empirical evidence obtained from field studies or appropriate material models. We have already made significant simplifications when we describe the surface topography by stream tubes, cascades of planes, rectangular grids, or triangular irregular networks; use Manning's equation to describe the dynamics of overland and open-channel flow; and use Darcy's law for unsaturated and saturated flow in the soil, vadose, and ground-water zones. Other types of simplification include linearizing nonlinear equations (Harley et al. 1970), using the kinematic approximations for overland flow (Woolhiser and Liggett 1967), or neglecting molecular diffusion and hydrodynamic dispersion in the transport of chemicals through porous media. A very common simplification is to reduce the dimensionality by treating streamflow as a problem in one space dimension. One extremely important simplification is due to our inability to describe the system geometry and properties in sufficient detail as we deal with larger (and generally more complicated) watersheds. Because of difficulties and the cost of measurement, we simply cannot provide a detailed, three-dimensional description of the surface microtopography, the hydraulic characteristics of the soils, and the underlying geologic materials. For this reason, we make some sweeping assumptions, i.e., the soil consists of layers of known thickness and the vadose and ground-water zones are homogeneous or have some regular stratification.

In my opinion, the term physically based, distributed model is an imprecise term, which was originally used to describe models that are not pipe and bucket compartment models or general linear or nonlinear system models. A better approach might be to recognize that, instead of specific classes of models, we have an array, with the simpler models being higher-order abstractions of the more detailed models (Woolhiser 1975). As such, they share some common properties, but all inevitably involve distortion. The simpler models are more general and involve fewer parameters, but this generality has been purchased at the price of a less detailed description. We would expect that when the computational scales of the physically based model and the scales of spatial variability of the real system are commensurate, the parameters would have some physical significance, provided that important mechanisms have not been ignored. However, as the system being modeled becomes too large and the computational scale becomes large relative to the scale of variability of the real system, the parameters will most certainly lose some of their significance, although there may be some regularities that can be used.

BACKGROUND

The search for a physically based, distributed runoff model has been underway for a long time (more than 50 years). Hjelmfelt and Amerman (1980) discovered a paper by Merrill Bernard, written and presented in 1937, in which he used a rectangular grid to represent the topography of a small watershed and used an essentially kinematic routing scheme to represent overland flow. Infiltration was described by an empirical equation. He observed the kinematic shock phenomenon [not so named until Lighthill and Whitham (1955)]. Of course, all the computations had to be done by hand, so his work had little impact and was forgotten for more than 40 years. With the development of the digital computer in the 1950s, interest was renewed in the computational aspects of watershed modeling. Several early papers dealing with the mathematical modeling of overland flow include those by: Behlke (1957), Liggett (1959), Morgali and Linsley (1965), Schaake (1965), and Brakensiek (1966).

I completed my undergraduate studies at the University of Wisconsin in 1955 and took a position at The University of Arizona where I was responsible for installing instruments and analyzing data from a 7.5 ha. (18.5 sq mi) experimental watershed near Tucson. My experience with this semiarid watershed convinced me that a distributed model was essential to describe the important hydrologic phenomena. After three years I joined the Agricultural Research Service (ARS) of the U.S. Department of Agriculture (USDA) and worked at the 153 km^2 (59 sq mi) Walnut Gulch Experimental Watershed at Tombstone, Arizona for nine months. My experiences there strengthened my belief that the current procedures such as the Soil Conservation Service (SCS) curve number and the unit hydrograph simply could not accommodate the high spatial variability of the rainfall and the interaction between runoff generated in the upland areas with the infiltration losses in the alluvial channels. In 1959, I was transferred to Madison, Wisconsin, where I worked for Minshall (1960), a very painstaking experimental hydrologist who was the first to provide strong experimental evidence of the nonlinear nature of runoff from small watersheds. I learned a great deal from him about the importance of accurate measurements, careful calibration and maintenance of equipment, and checking of computations. I first met the Hunter Rouse Lecturer for 1989, Jim Liggett, at The University of Wisconsin where it was my good fortune to take his course in advanced fluid mechanics. I was also introduced to the excellent book *Engineering Hydraulics* (1950), edited by Hunter Rouse. I was transferred to Columbia, Missouri in 1961 where I helped design and install the instrumentation for the Treynor, Iowa experimental watersheds, which were set up to determine the effects of level terraces on the hydrologic regime, including erosion losses and both water quantity and quality. One of my assignments at this time was to read through several annual reports of watershed studies that were begun in the Midwest in the 1930s. These reports were enlightening because they showed how the engineers had set up the studies to try to understand the physical processes involved in the transformation of rainfall to runoff. They recognized the many different mechanisms of surface runoff generation and made careful measurements of soil properties, even to the extent of digging trenches to sample and map soil-profile variability along hillslopes. The engineers who started these studies were highly talented and were well-supported with personnel and equipment. However, with the outbreak of World War II, most of the staff went to war and the funding was drastically cut. Many studies were discontinued and the objectives of the remaining studies were curtailed.

An interesting finding from some studies in Illinois was that the runoff from a watershed could not be adequately described merely by adding the products of the areas of soil mapping units and the runoff per unit area of small plots located within them. This finding was no doubt related not only to interflow and perched ground-water contributions to runoff and to complex runoff-runon interactions along hillslopes, but also to the small-scale spatial variability of soils. This observation was based model is based on the kinematic routing of runoff generated by the Hortonian mechanism, while the saturated hydraulic conductivity of the forested Hubbard Brook catchment was much greater than the greatest rainfall intensity, so one would not expect the model to apply. The Mahantango Watershed is also marginal in this regard, because it is known that there are other mechanisms, saturated overland flow and interflow, that are significant. The two largest events (where the Hortonian runoff could be significant) were eliminated from the data because the durations were too long. A careful examination of the Nash-Sutcliffe statistics reveals that the quasi-physically based model actually performed better than the other two for peak and volume at R-5, in spite of the fact that parameters for the other two models were estimated from the calibration data set while the parameters for the Q-P model were estimated a priori. Further, the poor time-to-peak efficiency statistic for the Q-P model was strongly influenced by one long storm where runoff continued for many hours and the model predicted the peak at the early part of the storm, while the real peak (not much different in magnitude) occurred many hours later! This type of error has no practical or theoretical significance. Another factor that must be considered in this study is that a broadcrested weir is used to measure runoff from R-5. The weir pond is large enough to significantly affect the outflow hydrograph from the watershed, yet no "ponding corrections" (Brakensiek et al. 1979) were made. The regression model and the unit hydrograph model were automatically calibrated to the effects of the pond, while the Q-P model attempts to predict inflow to the pond. In addition, the storage in the weir pond below the notch elevation is greater than the volume of many of the events in the runoff record, and there is no record of the water level in the weir pond before each runoff event. Consequently there is a great deal of noise in the data for the small runoff events. Finally, only one rain gauge was used to provide input to the models, and the runoff per unit area of many of the runoff events was smaller than the resolution of the rain gauge. As we shall see, this is more important for the Q-P model than for the other two. We must conclude that the Q-P model with parameters estimated a priori is actually better than the two models for the watershed to which it applies, although the efficiency statistics are not impressive. The maximum forecasting efficiencies for the volume and peak rate achieved by Loague and Freeze (1985) for R-5 were 0.25 and 0.71 (after calibration), respectively.

Goodrich (1990) obtained much more promising results for a physically based model. He used the model KINEROSR, a research version of KINEROS (Woolhiser et al. 1990b). In this model, topography is represented by cascades of planes and channels; infiltration by the Smith and Parlange (1978) approximation to the Richards equation; and runoff and channel flow are modeled by a four-point implicit finite-difference approximation to the kinematic wave equation. KINEROSR includes a provision to model the small-scale spatial variability of saturated hydraulic conductivity by modeling runoff from plane elements as the sum of hydrographs from n parallel planes, each having a hydraulic conductivity equal to the mean of the ith increment of a lognormal cumulative distribution function, $i = 1, \ldots, n$. Goodrich (1990) applied the model to four small watersheds in the Walnut Gulch Experimental Watershed operated by the ARS, USDA in southeastern Arizona. Rainfall and runoff data from watersheds LH-106, 0.36 ha; LH-102, 1.46 ha; LH-104, 4.40 ha; and WG-11, 631 ha were divided into calibration and verification sets. All parameters were identified based on published information and limited field studies. Initial soil water content was estimated at each rain gauge using a water balance model CREAMS (Knisel 1980). The relative values of saturated hydraulic conductivities K_s, coefficient of variation of the conductivities C_v, and Manning's n for overland flow planes were assumed to be correct, but three multipliers of these values were fitted in the calibration phase to approximately maximize the Nash-Sutcliffe forecasting efficiencies for runoff volume and peak. The forecasting verification efficiencies attained for runoff volume and peak flow rate, respectively, were: LH-106, 0.98 and 0.79; LH-102, 0.93 and 0.93; LH-104, 0.99 and 0.96; and WG-11, 0.49 and 0.16. Clearly these efficiencies are very good for the three smaller watersheds, and the efficiencies for WG-11 are much better than those attained earlier. Unfortunately, Goodrich (1990) did not compare these results with those of simpler models, so we do not have a direct comparison. However, Hughes and Beater (1989) used data from six Walnut Gulch Watersheds (43 large events) with a lumped and semidistributed conceptual model. Their best forecast efficiencies were −0.02 and 0.01 for the lumped and semidistributed version of their model where no parameter adjustment was allowed for verification events. We have no evidence here for the superiority of "simpler" models at this scale!

To put the results of Goodrich (1990) into perspective, however, we must consider the related work of Michaud and Sorooshian (1994a), who used the same distributed runoff model (KINEROSR) on WG-1, the entire 150 km^2 Walnut Gulch Experimental Watershed. Their objective was to determine whether distributed runoff models would be useful for real-time flash-flood prediction in semiarid regions. Accordingly they used rain gauge densities similar to those found at flash-flood warning sites (one gauge per 20 km^2) instead of all appropriate gauges on the watershed. They compared the accuracy of the KINEROSR predictions with those of a simple distributed model based on the Soil Conservation Service (SCS) method and a simple lumped model (also based on the SCS method).

Michaud and Sorooshian (1994a) concluded that: "None of the models investigated here were able to accurately simulate peak flows or runoff volumes for individual events. Models showed somewhat more skill in predicting time to peak and the peak to volume ratio." In a related study, Michaud and Sorooshian (1994b) showed that the rainfall errors due to too few rain gauges were responsible for about half of the simulation errors. They also pointed out that KINEROSR, which has nonlinear routing algorithms, may be more sensitive to rainfall errors as demonstrated by Singh and Woolhiser (1976).

What factors are responsible for the deterioration of model accuracy as watershed size increases? It is not possible to examine all possibilities in detail. Certainly, the distortions involved as the scale increases are a major factor. Michaud and Sorooshian (1994a) reported that the length of the pervious planes used in their watershed discretization ranged from 76 to 1.701 m. It is clear that elements of this size include significant channel networks, which would have different hydraulic properties and infiltration properties than the hillslopes contributing runoff to them. One can no longer claim much (if any) physical significance given plane elements of this size. What we are really doing is hypothesizing that the runoff generation and routing dynamics in the complex real system are similar to the runoff generation and routing characteristics of flow from a plane element. At smaller scales, this may not be a bad assumption. Indeed, Goodrich (1990) developed an algorithm for combining plane and channel elements into single plane elements that have nearly the same hydrologic response. His procedure involved maintaining basin area, mean overland flow length, average slope and soil parameters, and a characteristic response time (equilibrium storage divided by rainfall rate) by adjusting Manning's n. However, his results were for a much smaller scale (<631 ha), so it cannot be assumed that they also hold at a larger scale.

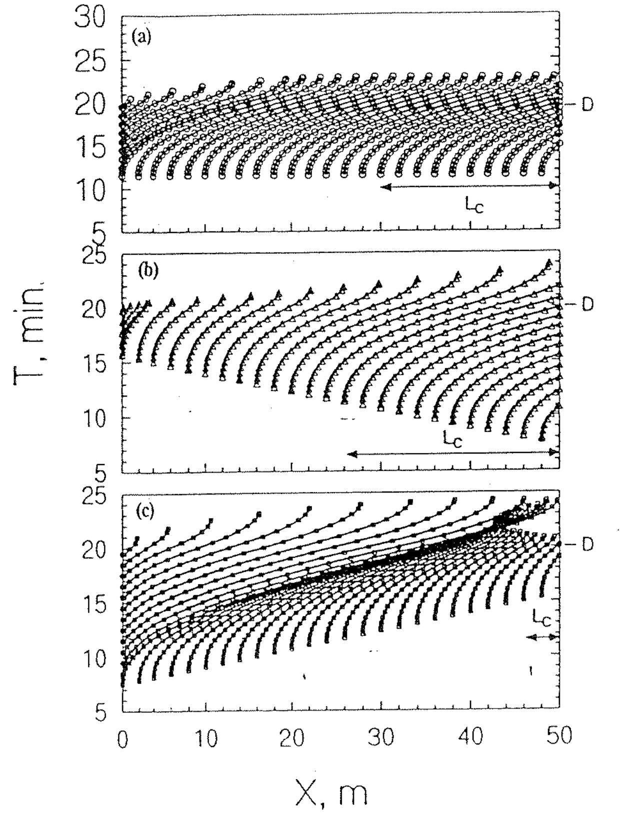

FIG. 1. Characteristic Nets for Hillslope Example of Smith and Hebbert (1979) with Rainfall Intensity of 88.8 mm/h for 20 min [from Woolhiser et al. (in press, 1996)]: (a) Case 4, Uniform K_s; (b) Case 5, Decreasing K_s; (c) Case 6, Increasing K_s

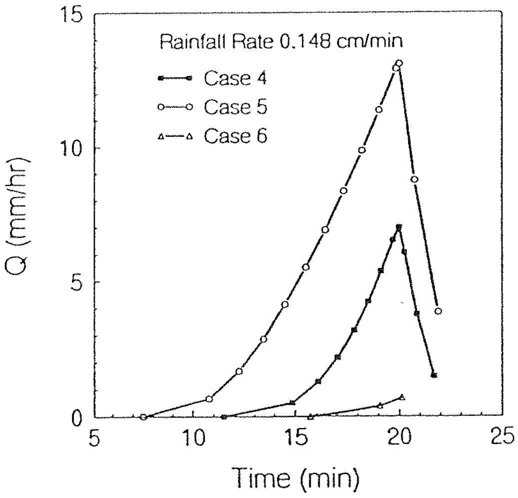

FIG. 2. Runoff Hydrographs for Hillslope Example of Smith and Hebbert (1979); Rainfall of 88.8 mm/h for 20 min [from Woolhiser et al. (in press, 1996)]

eral researchers have investigated this problem using somewhat different approaches. Freeze (1980) examined hillslope runoff response with spatially varied hydraulic conductivities. He analyzed a hillslope 100 m long and 200 m wide with saturated hydraulic conductivity varying on each 10 m × 10 m element. He used the Smith and Parlange (1978) infiltration equation for Hortonian runoff generation. However, there was no interaction between elements in the downstream direction (i.e., no runoff runon), and Freeze (1980) used a simple translation of runoff from each element to obtain hydrographs. Smith and Hebbert (1979) used a distributed finite-difference model to study the effects of spatial variability on hillslope runoff. Both Freeze (1980) and Smith and Hebbert (1979) found that spatial variability was very important for small storms, and that runoff response could not be closely approximated using average saturated conductivities.

To appreciate some of the subtleties involved, it is helpful to formulate the surface runoff problem in a kinematic characteristic framework (Woolhiser et al., in press, 1996). Characteristic nets are shown in Fig. 1 for the hillslope example of Smith and Hebbert (1979) with the low rainfall intensity (88.8 mm/h) and duration, $D = 20$ min, for the cases of uniform K_s and K_s decreasing and increasing downslope. Hydrographs for these three cases are shown in Fig. 2 (case 4 has uniform saturated conductivity, in case 5 the saturated conductivity decreases downslope, and in case 6 the saturated conductivity increases downslope). The extreme variability of the hydrographs can be easily explained by examining the characteristic nets. We see that for all three cases, only a portion of the hillslope, L_c, contributes runoff to the downstream boundary because infiltration after the rainfall ends depletes the water en route. The case where K_s increases is the most extreme with only about 5 m of the hillslope contributing. Random variation of K_s around a mean value with the same standard deviation would lead to hydrographs with greater volume than the uniform case, but smaller than the decreasing K_s case. The response variability decreases as rainfall rates increase. Obviously, if there are trends in K_s in the field, a model that does not take them into account will provide a very poor prediction.

Another possibly important aspect of hydrology is the degree of interaction between runoff water and infiltration after rainfall has stopped. In the KINEROS model there is a parameter, h_c (RECS in FORTRAN), which is conceptually related to the effect of the microtopography on infiltration during recession. If $h_c = 0$, the entire wetted area is subject to infiltration loss when rain stops. If $h_c > 0$, the effective infiltrating area per unit wetted area is linearly related to the ratio of the local average depth, h, to h_c if $h < h_c$. If $h_c > 0$ and $h > h_c$, the entire wetted area is subject to infiltration. This parameter was added when it was found that the model tended to underestimate recession hydrographs for plot data. On natural surfaces there is a concentration of flow into rills and we should expect this to affect recession infiltration rates. In testing on natural watersheds at a small scale, we haven't found the results to be very sensitive to h_c. However, it appears that there could be situations where h_c is very important. Fig. 3 shows the characteristic net obtained by setting $h_c = 1$ cm for the case with K_s increasing downslope, and comparative hydrographs are shown in Fig. 4. Under this condition, more of the water in transient storage on the hillslope at the time rainfall stops can

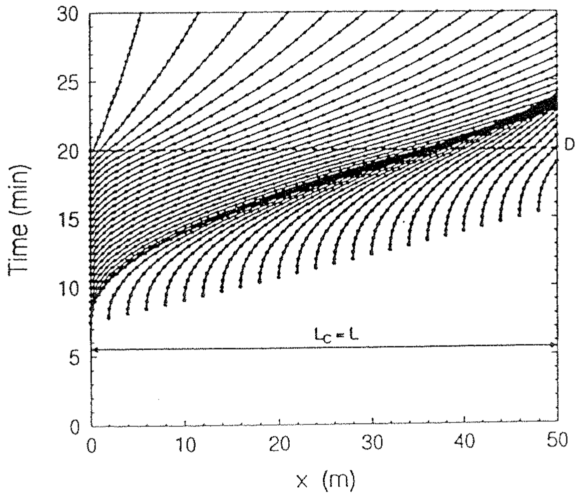

FIG. 3. Characteristic Net for Saturated Conductivity Increasing Downslope and $h_c = 1$ cm [from Woolhiser et al. (in press, 1996)]

Minshall, N. E. (1960). "Predicting storm runoff on small experimental watersheds." *J. Hydr. Div.*, ASCE, 86(B), 17–38.

Morgali, J. R., and Linsley, R. K. (1965). "Computer analysis of overland flow." *J. Hydr. Div.*, ASCE 91(3), 81–100.

Naipaul, V. S. (1969). *The loss of el dorado*. Penguin Books, Ltd., Middlesex, England.

Rosenblueth, A., and Wiener, N. (1945). "The role of models in science." *Philosophy of Sci.*, XII(4), 316–321.

Schaake, J. C. Jr. (1965). "Synthesis of the inlet hydrograph." *Tech. Rep. 3*, Storm Drain. Res. Proj., Dept. of Sanitary Engrg. and Water Resour., The Johns Hopkins Univ., Baltimore, Md.

Simons, D. B. (1992). "Future trends and needs in hydraulics." *J. Hydr. Engrg.*, ASCE, 118(12), 1607–1620.

Singh, V. P., and Woolhiser, D. A. (1976). "Sensitivity of linear and nonlinear surface runoff models to input errors." *J. of Hydro.*, Amsterdam, The Netherlands, Vol. 29, 243–249.

Smith, R. E., and Hebbert, R. H. B. (1979). "A Monte Carlo analysis of the hydrologic effects of spatial variability of infiltration." *Water Resour. Res.*, 15(2), 419–429.

Smith, R. E., and Parlange, J.-Y. (1978). "A parameter-efficient hydrologic infiltration model." *Water Resour. Res.*, 14(33), 533–538.

Smith, R. E., and Woolhiser, D. A. (1971). "Overland flow on an infiltrating surface." *Water Resour. Res.*, 7(4), 899–913.

Snyder, J. K., and Woolhiser, D. A. (1985). "Effects of infiltration on chemical transport into overland flow." *Trans. ASAE*, 28(5), 1450–1457.

Woolhiser, D. A. (1973). "Hydrologic and watershed modeling—state of the art." *Trans. ASAE*, 16(3), 553–559.

Woolhiser, D. A. (1975). "The watershed approach to understanding our environment." *J. Envir. Quality*, 4(1), 17–21.

Woolhiser, D. A., and Liggett, J. A. (1967). "Unsteady, one-dimensional flow on a plane—the rising hydrograph." *Water Resour. Res.*, 3(3), 753–771.

Woolhiser, D. A., Jury, W. A., and Nielsen, D. R. (1990a). "Effects of spatial and temporal variability on water quality model development." *Proc., Int. Symp. on Water Quality Modeling of Agric. Non-Point Sources, Part 2; ARS-81*, U.S. Govt. Printing Ofc., Washington, D.C., 505–522.

Woolhiser, D. A., Smith, R. E., and Goodrich, D. C. (1990b). "KINEROS, a kinematic runoff and erosion model: documentation and user manual." *ARS-77*, U.S. Govt. Printing Ofc., Washington, D.C.