Empirical and Causal Models in Hydrology

8

V. KLEMES
National Hydrology Research Institute, Environment Canada

INTRODUCTION

Hydrology is the basic science involved with water-resource management and enters into it through the use of models of hydrologic phenomena in water-resource planning, design, and operation. It is therefore important to examine the nature of hydrologic models and the relevance of their merits and drawbacks to water-management decisions. One of the controversial issues in hydrologic modeling has been the merits and demerits of empirical models as compared to those of causal models (Kartvelishvili, 1967, 1975; Scheidegger, 1970; Mandelbrot, 1970; Klemes, 1970, 1971, 1974, 1978; Yevjevich, 1974; Jackson, 1975; Pilgrim, 1975).

To put this issue into proper perspective one must begin with the fact that while the basis of all our knowledge is, in the last analysis, empirical, the knowledge itself is not merely a sum of empirical facts but emerges from the ability of the human mind to discover relationships among these facts.

"All science is the search for unity in hidden likenesses . . . The progress of science is the discovery at each step of a new order which gives unity to what had long seemed unlike . . . For order does not display itself of itself; if it can be said to be there at all, it is not there for the mere looking . . . order must be discovered and, in a deep sense, it must be created. What we see, as we see it, is mere disorder" (Bronowski, 1972).

The relationships initially discovered are of necessity simple and lead to only limited knowledge and understanding. They generally tell us what change in one observed quantity corresponds to a change in another. Such relationships are commonly labeled "empirical"--they tell us what happens but do not derive the outcome from the dynamic mechanisms governing the process, i.e., from the "necessary relationships between objects, events, conditions, or other things at a given time and those at later times" (Bohm, 1957). Those relationships that are based on the dynamics of the process are commonly labeled "causal"; the road to their discovery is basic research. There is a definite hierarchy in terms of the explicative power of these two kinds of relationships in the sense that a causal relationship gives us more information than an empirical one. But there is no conflict; as long as an empirical relationship is valid, a causal relationship cannot negate it--it can only supplement by indicating, for instance, some limits for the validity of the empirical relation or by pointing to external (and possibly not yet empirically established) factors that may modify it.

Empirical relationships are also used as convenient summaries, or reductions, of results of complex causal chains (e.g., Darcy's law). If such summaries were not possible, every causal model would have to be developed from the absolute "first principles" (i.e., elementary empirical facts and/or axioms) known at the time. Thus, in the physical sciences a causal model would have to have been formulated in terms of molecular interactions a century ago, atomic interactions half a century ago, and interactions among subatomic particles today. Whereas this approach may be seen as a theoretical ideal of causal modeling from the point of view of science in general, it would be of little use to any particular branch of science. In fact, science has become specialized into the innumerable disciplines because of the infeasibility of such an approach. The essence of specialization is to split the (possibly) infinite causal chain into segments, each containing only a few chain links; the scope of one discipline is thus intentionally limited to seeking causal relationships among phenomena within only a relatively small range, whose lower boundary represents the discipline's "first principles" or "scientific basis," coinciding with the discipline's objective.

A discipline seldom considers its own first principles to be an integral part of the discipline itself and tends to view them as being on the other side of the "free-body cut." Accordingly, problems encountered in the first principles are regarded as a hindrance rather than a challenge and lead to an impatient desire to get rid of them rapidly without much involvement. This is not only convenient but to a great extent necessary: It is the only way things can get done, it is the root of efficiency and productivity--unless we draw the line somewhere we are bound to drift along the causal chain of things indefinitely without being able to strengthen any of its links. On the other hand, to draw the line and take an action always implies a decision based on incomplete knowledge and understanding and therefore is prone to errors and unforeseen consequences. The relationship between the disciplines of water-resource management and hydrology is a typical example of this situation.

HYDROLOGY AS A SCIENCE

The common definition of hydrology as a "science that deals with the processses governing the depletion and replenishment of the water resources of the land areas of the earth, and treats the various phases of the hydrologic cycle". (WMO and UNESCO, 1974) does not convey the true perception of hydrology by the discipline of water-resource management. Here hydrology is perceived as more like a collection of techniques that enables one to make inferences from hydrologic data about the future distribution of water resources in space and time. In other words, the emphasis is not on the study of hydrologic processes and on the understanding of the mechanisms (physical, chemical, and biological) underlying these processes, i.e., on hydrology as a science, but rather on the *prediction of states of hydrologic processes in space and time*. Hydrology is of interest, here, only insofar as it can help in the determination of these future values of hydrologic state variables. If they could be foretold from a crystal ball, hydrology would be of little use to water management. As a matter of fact, much of the hydrologic effort originating in the domain of water-resource management bears a strong resemblance to a search for such a "hydrologic crystal ball." This observation is not meant to have a pejorative connotation; the search for a crystal ball is the implicit ideal of every empirical approach--to find something simple that works. In hydrology the empirical approach has many different labels, such as operational, prescriptive, analytical, and statistical (Klemes, 1978), which all roughly correspond to what in science is more generally known as "reductionism," whose objective is "finding a wonderful new calculus that will break through the barrier of the unknown" (Ziman, 1978). The essentially reductionist approach of water-resource management to hydrology is understandable and, to a great extent, inevitable. As Ziman (1978) observed, "whatever one's philosophical attitude towards reductionism, there is an inescapable scientific necessity of trying to 'understand' and 'explain' the behaviour of any system in terms of a relatively few comprehensible elements without recourse to an elaborate extra-cerebral computation." The purpose here is to make some observations regarding the origin of reductionist pressures in hydrology and their effect on hydrology as a science.

Scientific disciplines usually evolve from the construction of empirical models to the development of causal models. This transition can occur only after the science in question has reached a fairly advanced stage of development (Bohm, 1957). Often, this transition coincides with (and, perhaps, leads to) a break of the developing science away from its parent discipline, with the establishment of this breakaway branch as a new discipline. In this regard, hydrology is a young science compared with its sister sciences, such as meteorology, climatology, geology, and mineralogy. Whereas the latter sciences separated from their parent disciplines (agriculture and mining) a long time ago and have been recognized as sciences in their own right for many decades, hydrology still remains under a strong spell of "hydraulic engineering" for which the term "water-resource management" often is only a more recent (and more ambitious) equivalent. This dependence is reflected in the status of hydrology in universities and in its main sources of research funds. In most universities throughout the world, hydrology is attached to departments of civil engineering where it is usually taught as a sideline by professors of hydraulics or fluid mechanics. Most hydrologic research has traditionally been financed as a part of the planning, design, or operation of specific engineering projects such as dams, flood protection, and navigation schemes.

The fact that hydrologists typically have engineering backgrounds, combined with the usual applied context of hydrologic analyses, tends to reinforce the reductionist bias in hydrology, the trend to find that "wonderful new calculus" that will break through the barrier of the unknown separating raw hydrologic data from information on future values of hydrologic variables. Among the best-known examples of this trend are the rational formula; the unit hydrograph; the search for various correlations, periodicities, and symmetries in hydrologic data; flood-frequency formulas; stochastic operational models; and most recently the transfer-function models.

The strong reductionist bias in hydrology can be seen not only in the proliferation of the empirical models but also in the approach to causal modeling. This perhaps is best evident in the development of the so-called conceptual hydrologic models aimed at incorporating the general pattern of physical mechanisms governing hydrologic processes. More effort is spent on trying to determine the properties of the individual links in the causal chain (the "conceptual boxes") by optimizing the fit of the model output to an observed output than is spent on the study of the physical phenomena involved. The rather low popularity of the latter line of work may seem surprising because most hydrologists would agree with the definition of hydrology cited at the beginning of this section, which explicitly points in this direction. However, when viewed in the historic perspective discussed above the situation is understandable.

A third factor has contributed to this state of affairs in recent years: the computer. It has made the pursuit of finding the "wonderful new calculus" (or, more specifically, the "perfect transfer function") much less demanding, more publication productive, and therefore more attractive and even more prestigious than basic research in hydrologic processes. It now seems clear that the computer has done a disservice to many branches of science by diverting some of the best talent into pursuing purely computational problems of little relevance to the given science. As Fiering (1976) puts it, "Fascination with automatic computation has encouraged a new set of mathematical formalisms simply because they now can be computed; we have not often enough asked ourselves whether they ought to be computed or whether they make any difference. . . ."

Having reminded the reader that contemporary hydrology has a strong reductionist bias and having stated the main reasons, one can now ask how this affects the science of hydrology and its applications to water-resource management.

Before trying to indicate some answers, the two main sources of the difficulties arising in providing information on future states of hydrologic processes--the service expected from hydrology by water-resource management--are addressed below.

The principal source of difficulties is the extreme variety, variability, and complexity of processes that affect hydrologic phenomena (Chapter 2). The different temporal and spatial scales of these processes, their direct and indirect interactions, and inherent instabilities lead to great irregularities in the fluctuation of hydrologic processes in time. These irregularities manifest themselves as noise and make prediction of future states of a process a difficult problem. The situation is aggravated by the fact that reliable hydrologic records are usually relatively short and thus grossly inadequate for making inferences about long-term future behavior of hydrologic processes. Examples of temporal (and spatial) variability of streamflow are shown in Figure 8.1.

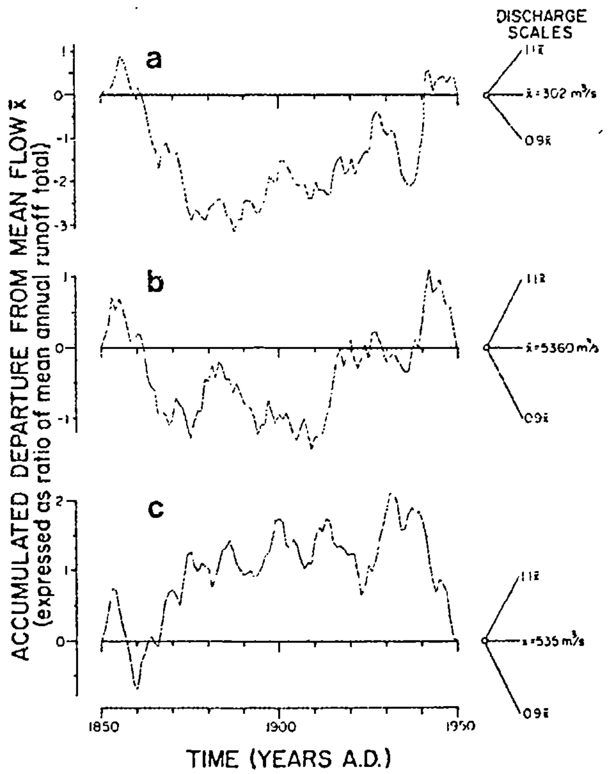

FIGURE 8.1 Accumulated departures from 100-yr mean flow x (1851-1950) for three European rivers (mean flow for any interviewing period is given by the slope of a straight line connecting its endpoints projected on the curve; it can be measured by the discharge scale shown). a, Elbe, Decin, Czechoslovakia; b, Danube, Orshava, Romania; c, Gota, Sjotorp Vanersbury, Sweden (Data sources: a, Novotny, 1963; b and c, Yevjevich, 1963).

The second major source of difficulties is that most hydrologic data pertain to variables at a point, whereas the hydrologic information desired for water-resource management usually involves variables pertaining either to a large area or to a different point for which data are not available. The difficulty in making this type of inference is proportional to the process spatial variability and irregularity that in turn increase with spatial heterogeneity of the environment in which the process evolves. Unfortunately, this spatial heterogeneity of physical variables affecting hydrologic processes is extremely high. Examples of spatial (and temporal) variability of point precipitation are shown in Figure 8.2.

MERITS AND DEMERITS OF EMPIRICAL MODELING

The main merits of empirical models are (1) the possibility of developing them without much understanding of the modeled phenomenon, (2) their simplicity achieved by short-circuiting complex causal chains, and, as a result of these two features, (3) their potential for making the collected data useable without much delay and hence their promise of high cost effectiveness of the modeling exercise.

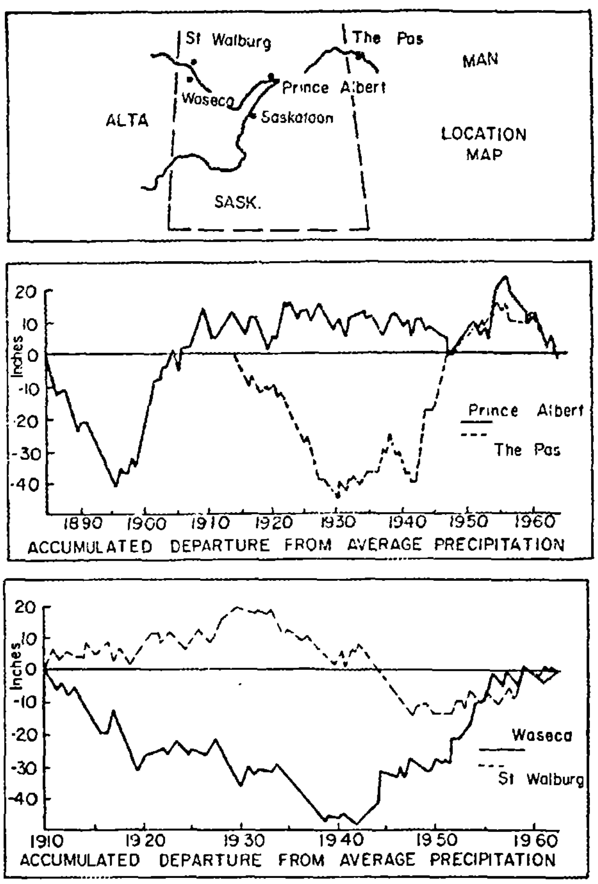

FIGURE 8.2 Accumulative departures from long-term mean precipitation for four Canadian weather stations (slope of plot = local mean) (reproduced from Water Studies Institute Report No. 2, Oct. 1965, Saskatoon, Saskatchewan, Canada, as cited in NRC Geophysics Study Committee, 1977).

The drawbacks follow naturally from the above merits but need more elaboration.

1. Because of the extreme complexity of the environment in which hydrologic processes evolve and the fragmentary (both time and space-wise) information that hydrologic data normally contain, empirical relationships based on these data usually can give only approximate results. However natural this may be, the uncertainty in the result causes displeasure to the user, who then exerts pressure on the modeler to "improve" the model. If the door to more information (in terms of more data and better understanding of the process) is closed, the only choice available to the modeler is to try to extract more information from the data at hand using some "better calculus." While sound in principle, this course of action faces several dangers.

One of them is overfitting, which amounts to regarding part of the noise in the data as information. This error is easy to commit because the demarcation line between noise and information is often blurred in the data. A well-known example from the not-so-distant past is the various harmonic analyses of hydrologic series. Their claims of discoveries in most historic records of periods with wavelengths other than those corresponding to the astronomic cycles of the earth and ranging from a few months to many decades have never been substantiated by more recent data.

Another is the danger of sidetracking into polishing some aspects of the modeling methodology that, while perhaps of value for the methodology itself, are unimportant in the context of the particular application. An outstanding example is the history of flood-frequency analysis, where much progress has been made in computation of plotting positions, in (efficient, unbiased, and consistent) estimation of population parameters, in treatment of outliers, and in goodness-of-fit testing. On the other hand, the distribution type is chosen arbitrarily, the calendar boundaries of the year are chosen arbitrarily, the typical nonhomogeneity of the flood-causing factors (some floods are caused by snowmelt runoff, some by convective storms, some by frontal storms, and some perhaps by hurricanes) and the influence of the changing character of the basin are ignored, the presumed mutual independence of flood events is usually not checked, the difference in the hydraulic behavior of the river under relatively small and extremely large floods is not considered, different hydraulic conditions along the river channel are dismissed, and the high uncertainty in the reported values of peak discharges (which are seldom directly measured) is not taken into account. When examining the mainstream of the flood-frequency literature from this point of view, one wonders what there is to be learned about the frequency of floods as opposed to the art of distribution fitting to samples of exact numbers drawn by random mechanisms from homogeneous infinite populations.

Another danger originating in the desire to improve a model without additional information on the process is to elaborate formally some aspect of the model known to be wrong and present merely for operational reasons with the understanding that the implied error "does not matter" in the given context. The case in point is, for example, the "baseflow separation" problem in the unit hydrograph technique. The concept of separate direct runoff and baseflow hydrographs is hydrologically doubtful but is convenient for estimating the water balance of an isolated flood event as long as the estimate does not have to be too accurate. The separation line is an operational device that has little, if any, hydrologic significance, and the great efforts that have been spent on the refinement of its shape have not improved unit hydrograph modeling at all.

In summary, the striving for "improvement" of an empirical model in the absence of additional information tends to be scientifically sterile and to have an extremely low benefit/cost ratio from the point of view of applications.

2. Empirical models must be regarded essentially as interpolation formulas. They have no justification outside the range of the underlying data sets and their use for extrapolation involves risk of large errors. This aspect is of prime importance in water-resource management because the hydrologic information required for most of its applications involves extensive extrapolations in space, in time, and in the range of the variable concerned (extreme values). An example of a possible consequence of such extrapolation (in time, in this case) for water-resource management can be drawn from Figure 8.1. Extrapolation of mean flow computed from one period of the historic record could lead to a serious error when extrapolated into the following period.

3. A great handicap of empirical modeling is the uncertainty about the adopted model structure. In the absence of theoretical (physically based) reasons for a specific structure, auxiliary (and, on the whole, subjective) criteria for model selection must be adopted. Mathematical convenience is a popular refuge, current fashion running a close second. The inherent danger is that the high degree of arbitrariness in the model form is either not realized or the awareness of it gradually fades away and the assumptions are treated as facts of life and later used as unalterable building blocks in other models.

A typical example is the use of long-term means in hydrology. The arithmetic mean is one measure of central tendency of a process attribute whose central tendency *can be expected* to remain constant as the process develops in time. Central to the usefulness of the long-term mean concept is the *justifiability of the expectation of the constancy* of this central tendency. This justifiability is by no means universal and depends on the nature of the process. If we are concerned, for example, with the central tendency of the amount of beer in bottles coming from a specific, well-tested bottling machine, then the expectation of constancy is justified. If, however, we deal with the amount of precipitation or runoff in individual years this may not be the case as Figures 8.1 and 8.2 indicate. We have become greatly concerned about the apparent lack of this constancy (as our increasing interest in climatic change indicates) but, nevertheless, we continue to regard long-term means of precipitation and runoff as valid concepts and use them as the basis of many hydrologic models. It does not seem to matter much that we then arrive at model structures that are hydrologically inexplicable; our hydrologic (and water-management) conscience apparently can accept this situation more easily than it can the idea of re-examining the arbitrary and hydrologically irrelevant concept of constant central tendency that has been deeply ingrained into our minds by the never ending exposure to superficial interpretations of statistics. We do not seem to realize that, from the empirical modeling viewpoint, the mean of a time series is merely one of many fits by a horizontal straight line. Why a geophysical series should be fitted by a straight line, in particular by a horizontal one that minimizes the sum of squares of vertical deviations, is difficult to see (Klemes, 1974).

4. The essential arbitrariness in the selection of the form of an empirical model exposes the modeler to the danger of adopting a basically wrong modeling approach as a result of its success in some other situation and because of some superficial similarities between the modeling problem at hand and a problem where the model proved successful. This seems to be the case with much of the applications of classical statistics in hydrology and consequently with much of the applications of statistical hydrology in water-resource management. Statistical methods in current use have been designed for the analysis of large masses of data from repeatable and controlled experiments. In empirical hydrologic modeling they are applied to hopelessly small samples generated by "unique and uncontrolled experiments," as hydrologic and other geophysical processes can be characterized. With regard to the mathematical concept of stochastic process that we routinely invoke, a given historical series is a sample of size one. It also is a sample of size one for many planning and design purposes, such as the estimation of reservoir sizes or mean flows for the economic life of a project. Yet we are arrogant enough to use this single measurement to construct distributions, "estimate probabilities," and do many other things that must make Richard von Mises turn over in his grave ("First the collective then the probability," von Mises, 1957).

5. It is generally acknowledged in the

philosophy of science as well as in science itself that the fundamental reason for empirical modeling is the availability of data combined with a lack of understanding of the relationships among the phenomena they describe. This combination is conducive to the adoption of a "let-the-data-speak-for-themselves" philosophy that has many supporters in hydrology and, which, through the development of black-box (transfer-function) modeling, has become a kind of ideology for many hydrologists. This lack of understanding of the process mechanisms, originally regarded as regretable, has been transformed into a virtue--questions about the internal workings of the system being modeled are excluded by design. This philosophy, most prominent in stochastic hydrologic analysis, has two distinct aspects.

The first is its failure to recognize that what makes the data speak is the context; without it their numerical values have little to tell and what they do tell is often misleading. For instance, it would be naive to automatically identify the mean of a series of numbers with a reasonable estimate of the mean of the variable whose states they represent. And it would be even more naive to hope that the estimate can be improved by polishing the formula for the mean. An intelligent inference about the mean of the *physical variable* concerned, e.g., a distance, can only be made by investigating *physical problems* of the following kind: Do the numbers represent measurements of the same distance or of different distances? Were they obtained by the same instrument, same person, under the same conditions? Do they represent distances between stationary or moving objects (moving with uniform or nonuniform motion)? Does it at all make sense to talk about a 'mean distance' in the given case? It is important to realize that these problems are not reducible to problems of blind-mathematical manipulation of numbers.

The second aspect is more subtle and has a Machiavellian flavor. The untenability of the "speak-for-yourself" attitude to data has often been exposed not only in the hydrologic context (Fiering, 1967; Klemes, 1971; Kartvelishvili, 1975) but also in statistical literature by such eminent scientists as Norbert Wiener, M. G. Kendall, A. Stuart, J. Neyman, and M. S. Bartlett (Klemes, 1978). It is therefore difficult to believe that the belief in this philosophy, professed by many hydrologists, is sincere. A more likely explanation is that it serves merely as a convenient and, it is hoped, dignified cover for the hydrologist's reluctance to admit his despair face to face with the enormous complexity of hydrologic processes (Karteveli-shvili, 1975, suggested that the development of an adequate causal theory of hydrologic processes may be much more demanding than was the development of the theory of relativity or the quantum theory), his lack of *hydrologic* ideas, and his overabundance of *computing* ideas combined with his good formula-manipulative skills (a natural consequence of the systems-analytical bias prevalent in graduate hydrologic programs in many, if not most, leading universities in recent years), his compliance with the merciless rules of the publish-or-perish modus operandi of contemporary science, his preference for the cozy atmosphere of his office over the inconveniences of field research, and, last but not least, his repeated failures to secure resources for long-term research of "merely academic" interest with little promise of immediate applicability, and his surrender to pressures for fast results. Many a bird can be killed with the stone of black-box modeling, and "occasionally words must serve to veil the facts. But this must happen in such a way that no one becomes aware of it; if it should be noticed, excuses must be at hand, to be produced immediately" (Machiavelli in "Instructions to Raffaelo Girolami"). Herein may well be the greatest disbenefit of empirical modeling to hydrology.

To summarize, empirical models are useful (and, indeed, indispensible as starting points in cases of insufficient data and/or understanding) as long as they are not mathematically strained beyond their intrinsically limited "carrying capacity."

WHY DO HYDROLOGIC MODELS WORK?

It can be argued that if hydrologic modeling has a strong reductionist bias and if this bias has so many dangerous consequences as claimed in the preceding section, then the performance of hydrologic models could not be as good as it appears to be. For, as many would testify, hydrologic models have been used successfully in innumerable instances in both design and operation of water-resource engineering projects.

There are a number of reasons why hydrologic models work or seem to work. The most common are listed below, approximately in decreasing order of occurrence:

1. *Model Is Empirical and Works Well As an Interpolation Formula* An empirical model is a formal statement of observed facts and thus has to give good results, i.e., results within the observational accuracy and compatible with the level of observed noise, within the range of the observations. Inside this domain its good performance is conditional on the invariance with time of the conditions reflected by the variables involved. All regression models fall into this category. A typical example is the stage-discharge relationship for a river channel cross section. While true in the above sense, its good performance hinges on the stability of river morphology in time and is limited to the range of river stages and surface slopes for which the measurements of flow velocity and cross sectional area were actually carried out. The reliability of the relationship beyond the range of the measurements cannot be deduced from the quality of its fit within that range.

2. *Model Works Well Because It Portrays Only a Small and Relatively Well-Understood Segment*

(or Component) of the Hydrologic Cycle The most typical cases are those in which the hydrologic process is locally (spatially and temporally) dominated by (i.e., can be approximately reduced to) processes of hydraulics for which relatively good causal models exist. Thus it is a hydraulic rather than hydrologic model that works. Examples include flood-routing models, flow on a hillslope, the latest addition being the various "urban-hydrology" models whose success is proportional to the degree to which the urban catchment consists of impervious surfaces, sewer pipes, and well-defined prismatic channels (hydraulic elements).

3. *Model Is Essentially Untestable and Its Good Performance Is a Matter of Faith* Into this category fall most models intended to portray aspects of the long-term behavior of hydrologic processes, especially if their characteristics have been conservatively estimated. For example, few hydrologic records will ever be long enough to make possible a conclusive refutation of the correctness of the magnitude of a 50-, 100-, or 1000-yr flood if this magnitude is inflated; no single reservoir will operate long enough under the design-release policy to enable the analyst to prove that the actual risk of failure was less than the design value of, say, 2 percent; if the value of a "probable maximum" precipitation or flood is set by an order of magnitude higher than any observed event, then its correctness is beyond reproach because of the vagueness of the definition--if a higher flood occurs it can always be classified as an "improbable event" (everybody knows that improbable events do occur--people do win millions in lotteries).

4. *Results Obtained by a Good Economic Decision Model Reflect Favorably on the Quality of a Hydrologic Model Embedded in It* For example, Slack *et al.* (1975) state ". . . the use of the normal distribution to represent the distribution of floods is generally better than either the Gumbel, lognormal, or Weibull distributions. Nothing is gained in terms of reducing expected opportunity design losses if the underlying distribution . . . is identified over and above simply using the normal as the assumed distribution." Here the first sentence may convey a wrong impression that the normal distribution is a *better hydrologic model* for flood peaks (in the sense that it provides better estimates of flood frequencies in specific cases), whereas what it really means is that it is *better in the given decision context* as is obvious from the second sentence of the quotation.

5. *Model Is Largely Irrelevant to Results Obtained with Its Aid* Claims of good performance of a specific hydrologic model are sometimes based on results that, while obtained with the aid of the model, are irrelevant to its structure and some parameters and depend on circumstances external to the model. For example, it can be claimed that a given stochastic model represents an historical flow record well because it leads to an optimum reservoir-operating policy that is essentially the same as the policy that actually would have been optimal during the period of record. However, because the optimal policy for a typical economic-loss function depends mainly on the mean inflow and is largely invariant with regard to the inflow model structure (Klemes, 1977), the claim of the good performance of the model is misleading--most models would do as long as they have the same mean and loss function (Jettmar and Young, 1975).

6. *Empirical Model Has a Form That, Without Conscious Effort of the Modeler, Happens to Describe Some Essential Aspect of the Physical Mechanism of the System* Such models do not occur frequently, but when they do they tend to acquire great popularity because their success is higher than their empirical nature would suggest. A famous example is the unit-hydrograph model. Originally formulated as a purely empirical concept, it has since been shown to represent outflow from a system of linear storages fed by a pulse inflow--a crude but physically sound model for many small catchments. Similarly, the empirically chosen autoregressive model for runoff series was later shown to represent outflow from a system of linear storages fed by an autoregressive input of a lower order. Yet another example is the empirical concept of partial runoff-contributing area that was later found to follow from a physically based model of flow on a hillslope (Chapter 1).

7. *Model's Good Reputation Is Based on Superficial Appearances* Whenever a reasonable result is obtained by a model that accommodates in a logical way a number of factors believed to be relevant to the problem on hand, the model tends to acquire a good reputation. However, it may well be that many of the factors chiefly responsible for this reputation are in fact redundant as far as the results are concerned. Conceptual hydrologic models are prone to acquiring this "false dignity." For example, Mein and Brown (1978) demonstrated that a particular conceptual-deterministic model, designed to simulate monthly flows on the basis of 13 optimized parameters and a daily time step, performed only marginally better than the same model with only 3 parameters optimized and the rest set ". . . arbitrarily high or to zero and . . . others fixed to values which can be physically justified" (in the first case the model accounted for 95.3 percent of the sum of squares of deviations, in the second for 93.6 percent). Similarly, an undeserved good reputation of a regression model may derive from high correlations that may be the result of spurious correlation as pointed out by Benson (1965) and Pilgrim (1975).

8. *Model Works for the Wrong Reasons* A prominent example from astronomy is that of the Ptolemaic geocentric planetary model that, although physically unsound, predicted the common astronomical events with a reasonable accuracy. Klemes (1974) pointed out that the same may be true of some hydrologic models, e.g., of the fractional noise model for hydrologic time series. Based on the hydrologically implausible assumption of infinite memory and a constant

mean, the model produces time series statistically similar to those exhibited by hydrologic (and other geophysical) variables that are more likely to possess finite memory and time-varying means (see Figures 8.1 and 8.2). Examples of this situation are also common in models with large numbers of optimized parameters where wrong assumptions may be compensated by physically unrealistic values of some parameters so that the model gives good results.

9. *Model Is Deemed Good by Default* Models often acquire a good reputation because they did not have a chance to fail. They may not have been used long enough for their limitations to emerge. Model building being a dynamic activity in present-day hydrology, models are continually modified, recalibrated, and otherwise improved, and "good" models are superseded by "better" models before they reveal themselves as bad models.

10. *Model That Does Not Work Is Not Publicized* Last but not least, the impression that hydrologic models work generally well arises from a tautology. We learn only about the models that seem to work well and thus can only conclude that the models reported in the literature seem to work well. Moreover, because model performance tends to be interpreted as a reflection of the modelers' (model users') own competence and skill, their reports tend to have an optimistic bias. Critical evaluations of hydrologic models are rare. A notable contribution to this area is that of Pilgrim (1975) who discussed the many weaknesses of hydrologic modeling and indicates that most of them can be traced to our limited ability to properly incorporate into the models the causal relationships governing the hydrologic cycle.

DRAWBACKS, DANGERS, AND POTENTIAL BENEFITS OF CAUSAL MODELING

The most serious drawback of causal modeling in hydrology is its well appreciated difficulty, the clear awareness of the vastness of the void separating our data and understanding of the processes involved from our goals, (Chapters 1 and 2) and thus the necessity of an extensive program of basic research. The perspective of long years of painstaking observations and hard thinking without a guarantee of significant result appeals neither to many researchers nor research managers and perhaps least of all to graduate students who should be the main source of talent. The difficulty appears perhaps even greater than it is in reality because the prospective researchers view it from the perspective of their own academic background that, as already stated, typically is slanted to hydraulic engineering and water-resource systems analysis and is inadequate in climatology, geology, biology, chemistry, and physics, whose indispensability for the task soon becomes evident to them (see Chapters 2, 3, and 6).

The attendant danger of this difficulty is the temptation of reductionism: to make shortcuts and to fill the void between the data and the goals with logically plausible assumptions that are sometimes correct but often wrong and, more often than not, individually untestable. The result is a "conceptual" model that purports to be causal but sometimes is only a disguised and somewhat structured empirical construct whose elements are regression coefficients with physically sounding names. These models have a greater potential than "blind regression" and other statistical models (even though they do not always perform better, as is shown in WMO, 1975, and Garrick *et al.*, 1978) provided that their structure correctly reflects the basic aspects of the process; however, the danger resides in their tendency to elevate the preconceived hypothetical structure on the pedestal of truth and to divert attention from the investigation of the behavior of hydrologic process into the dead-end street of parameter optimization (see point 8 in the foregoing section).

The potential of causal-hydrologic models lies in their ability to derive the behavior of a hydrologic process for a given set of states of nature (physical variables) from the dynamic mechanisms of the process, without recourse to model calibration by empirical fitting. Consequently, causal models offer a possibility to predict the behavior of a hydrologic process under conditions that did not exist during the process-recorded history, i.e., under conditions for which an empirical model cannot be constructed. Herein lies the argument against frequent claims that, because simple statistical and black-box models often outperform complex causal models, there is no need for engaging in causal modeling. It is not a question of prediction accuracy for known conditions but one of model credibility in unknown conditions. It is the difference between blind extrapolation and sound judgment. The following examples illustrate this point.

Garrick *et al.* (1978) found that simple seasonal averages derived from a historical flow record give, in one particular case, better predictions than a highly sophisticated conceptual model (SSARR). In another instance, Todini and Wallis (1977) found that a simple transfer function model (CLS) performed as well or better than the SSARR and similar models. The crux of the matter is that in these and other similar studies the physical conditions in the basin were essentially the same during periods from which data were used for model development and during those used for its testing. If, for example, flows were to be estimated for a future situation when a large inundation area within the basin would be eliminated, or a large rural area urbanized, neither of the two empirical approaches cited above could accommodate the new situation, whereas the relatively inferior SSARR model could because of its distributed nature and its built-in facility for flow routing. The fact that the "accommodation" presently achievable may be unsatisfactory demonstrates only the inadequacy of the present state of causal modeling, not an infeasibility of causal modeling in general. However, it also demonstrates one additional point: progress in causal modeling can result only from more hydrologic knowledge and not from more "causally inspired" manipulation of the little knowledge we have.

Another example is offered by Eagleson's (1972) causal model for flood-frequency distribution. In a particular instance, the model may give a fit to observed flood peaks that is inferior to a fit obtained by statistical flood-frequency fitting techniques. However, the latter techniques are useless for estimating changes in flood regime due to changed land management, while Eagleson's model can offer approximate guidance because, for a given rainfall distribution (which represents one of the "first principles" of the model and remains outside of it), it relates flood-peak frequencies dynamically to such factors as the runoff-contributing area and the conditions of flood propagation within the basin. For example, he shows that if the runoff-contributing area in one of the catchments under study increased (e.g., due to urbanization) from one third to one half of the total area, a 100-yr flood could become approximately a 10-yr flood.

As another example, approaches have been suggested in which the distribution of annual (or seasonal) runoff total is causally related to the amount of perennial snow and ice and the amount of energy available for melting (Klemes, 1971) or to a number of other climatic and physical *variables* (Eagleson, 1978). Such approaches could, for instance, provide guidance for estimating water resources affected by future climatic changes predicted in the literature (Budyko *et al.*, 1979).

The inherent ability of causal models to assess the effects of environmental changes, both natural and man-made, on water resources is by far the most important aspect of their potential in water-resource planning and management, especially in the present epoch when the rate of environmental changes is higher and is increasing more rapidly than in any other epoch of the recorded history. On a more general level, causal approach seems to represent the only means capable of increasing the low credibility of essentially untestable models. Another important aspect of causal models is their potential to point out ways to efficient short-cuts and thereby to better empirical models. For example, the kinematic wave approximation of the equations of motion suggests that the simple concept of a nonlinear-storage reservoir (or a system thereof) may have a relatively great potential in basin-runoff modeling (Laurenson, 1964; Klemes, 1973).

On the scientific level the greatest potential of causal hydrologic models is in their intrinsic ability to meet the following challenge. "There is need for a systematization of research into mathematical models of hydrologic systems in order to provide the framework for rational methodology for the use of hydrologic models. This is necessary both for the organization of research results into a body of coherent knowledge and for the ready application of research work to field problems" (Dooge, 1972). This need has become urgent during the last decade when the proliferation of models has revealed many inconsistencies and incompatibilities in the assumptions underlying different models used to portray specific aspects of one and the same phenomenon. The only framework that can reduce these inconsistencies and incompatibilities is the causal, i.e., physically based, modeling. The following example will illustrate this point.

In the analysis of rainfall-runoff relationships it is a common practice to fit independently the system's response of the catchment, the flow-routing model (models) for its river channels, the probability distributions of precipitation and runoff, the probability distributions of precipitation and runoff extremes, and the stochastic structures of the precipitation and runoff series. As a rule, the results from the individual models are incompatible with each other in the sense that they could not be produced by one and the same physical system. For example, the unit hydrograph model of a basin is incompatible with a nonlinear cascade model of the basin's river channels; a random precipitation series combined with the unit hydrograph model cannot produce a fractional-noise-type series of runoff; and a nonlinear catchment model is incompatible with a runoff-series model based on the concept of autocorrelation. These problems have been analyzed by Klemes (1978) who showed that the unifying framework for all these *ad hoc* models can hardly be anything else than a physically consistent model of the catchment mechanisms, i.e., a *causal theory of the hydrologic cycle*. A significant step in this direction has recently been made by Eagleson (1978). For the time being, his and similar attempts must be viewed chiefly as exploratory probes. This line of inquiry seems to be the most promising way out of the unenviable present situation that has been aptly compared (Dooge, 1978) to "a riot of growth reflecting a variety of scale, colour and type and . . . a cacophony of noise . . . confronting . . . a traveller lost in a jungle."

CONCLUSION

Prediction of future states (or ranges thereof) of hydrologic processes is a necessary prerequisite for scientific management of water resources. It requires adequate modeling of the hydrologic cycle in which both the empirical and the causal models have their legitimate place. However, it must be recognized that even at its best a model can be nothing more than a particular system of organization of hydrologic knowledge; at its worst it degenerates into a manipulation of conjectures. The degradation of the first case into the second is inevitable if the body of knowledge to be organized does not increase proportionally to the advancement of the methods for this organization. There are many signs that this degredation process has been taking place in hydrologic modeling in the recent past with active, though unintentional, help from water-resources management, which as a rule has placed models above the knowledge on which they should be based.

For many years, hydrologists have been discouraged, through one-sided training and short-sighted "cost-effective" research financing, from penetrating to the roots of the hydrologic processes and the mechanisms underlying the hydrologic cycle. On the other hand, they have been applauded for serving the same cheap mathematical cocktails offered under a variety of exotic names but always mixed from one or two of the three standard ingredients (often of questionable quality): point precipitation, streamflow, and groundwater level. Naturally, the same hangovers followed only to be "cured" by more of the same mixes. Unless this vicious circle is broken, the cult of mathematical embroidery of sterile concepts abolished, advancement of hydrologic knowledge (rather than fast service for water-management problems) established as the objective of hydrology, there is little hope for a substantial improvement in the scientific basis for water-resource management, an improvement that may be much needed in the years ahead.

REFERENCES

Benson, M. A. (1965). Spurious correlation in hydraulics and hydrology, *Proc. Am. Soc. Civ. Eng. 91*, 35-42.

Bohm, D. (1957). *Causality and Chance in Modern Physics*, Routledge and Kegan Paul, London.

Bronowski, J. (1972). *Science and Human Values*, Harper & Row, New York.

Budyko, M. I., K. Ya. Vinnikov, D. A. Drozdov, and N. A. Yefimova (1979). Impending climatic change, *Soviet Geography XX(7)*, 395-411 (Original in Izv. Akad. Nauk SSSR, ser. geogr., 6, 5-20, 1978).

Dooge, J. C. I. (1972). Mathematical models of hydrologic systems, in *International Symposium on Modeling Techniques in Water Resources Systems, Vol. 1*, Environment Canada, Ottawa, pp. 171-189.

Dooge, J. C. I. (1978). General report on model structure and development, in *International Symposium on Logistics and Benefits of Using Mathematical Models on Hydrological and Water Resource Systems*, IBM, Pisa, October.

Eagleson, P. S. (1972). Dynamics of flood frequency, *Water Resour. Res. 8*, 878-897.

Eagleson, P. S. (1978). Climate, soil, and vegetation, *Water Resour. Res. 14*, 705-776.

Fiering, M. B. (1967). *Streamflow Synthesis*, Harvard University Press, Cambridge, Mass.

Fiering, M. B. (1976). Reservoir planning and operation, in *Stochastic Approaches to Water Management, Vol. II*, H. W. Shen, ed., H. W. Shen, Fort Collins, Colo., pp. 17:1-17:21.

Garrick, M., C. Cunnane, and J. E. Nash (1978). A criterion of efficiency for rainfall-runoff models, *J. Hydrol. 36*, 375-381.

Jackson, B. B. (1975). Birth-death models for differential persistence, *Water Resour. Res. 11*, 75-95.

Jettmar, R. U. and G. K. Young (1975). Hydrologic estimation and economic regret, *Water Resour. Res. 11*, 648-656.

Kartvelishvili, N. A. (1967). *Teoriya veroyatnostnykh protsessov v gidrologii i regulirovanii rechnogo stoka*, Gidrometeorologicheskoe izdatel'stvo, Leningrad.

Kartvelishvili, N. A. (1975). *Stokhasticheskaya gidrologiya*, Gidrometeorologicheskoe izdatel'stvo, Leningrad.

Klemes, V. (1970). Negatively skewed distribution of runoff, in *Symposium of Wellington (N.Z.)*, IASH Publ. No. 96, pp. 219-236.

Klemes, V. (1971). Some problems in pure and applied stochastic hydrology, in *Proceedings of the Symposium on Statistical Hydrology*, Tucson, Ariz.

Klemes, V. (1973). Watershed as a semi-infinite storage reservoir, *Proc. Am. Soc. Civ. Eng. 99 (IR4)*, 477-491.

Klemes, V. (1974). The Hurst phenomenon--a puzzle? *Water Resour. Res. 10*, 675-688.

Klemes, V. (1977). Value of information in reservoir optimization, *Water Resour. Res. 13*, 837-850.

Klemes, V. (1978). Physically based stochastic hydrologic analysis, *Adv. Hydrosci. 11*, 285-356.

Klemes, V. (1979). The unreliability of reliability estimates of storage reservoir performance based on short streamflow records, in *Reliability of Water Resources Management*, Water Resources Publications, Fort Collins, Colo., pp. 193-205.

Laurenson, E. M. (1964). A catchment storage model for runoff routing, *J. Hydrol. 2*, 141-163.

Mandelbrot, B. B. (1970). Comment on "Stochastic models in hydrology" by Adrian E. Scheidegger, *Water Resour. Res. 6*, 1791.

Mein, R. G., and B. M. Brown (1978). Sensitivity of parameters in watershed models, *Water Resour. Res. 14*, 299-303.

NRC Geophysics Study Committee (1977). *Climate, Climatic Change, and Water Supply*, National Academy of Sciences, Washington, D.C., 132 pp.

Novotny, J. (1963). *Dve stolete hydrologicke rady prutokove na ceskych rekach*, Sbornik praci, No. 2, Prague.

Pilgrim, D. H. (1975). Model evaluation, testing and parameter estimation in hydrology, in *Prediction in Catchment Hydrology*, Australian Academy of Science, Canberra, pp. 305-333.

Scheidegger, A. E. (1970). Stochastic models in hydrology, *Water Resour. Res. 6*, 750-755.

Slack, J. R., J. R. Wallis, and N. C. Matalas (1975). On the value of information to flood frequency analysis, *Water Resour. Res. 11*, 629-647.

Todini, E., and J. R. Wallis (1977). Using CLS for daily or longer period rainfall-runoff modelling, in *Mathematical Models for Surface Water Hydrology*, T. A. Ciriani, U. Maione, and J. R. Wallis, eds., John Wiley & Sons, Inc., London, pp. 149-168.

von Mises, R. (1957). *Probability, Statistics, and Truth*, G. Allen and Unwin, London.

WMO (1975). *Intercomparison of Conceptual Models Used in Operational Hydrology Forecasting*, World Meteorological Organization, No. 429, Geneva.

WMO and UNESCO (1974). *International Glossary of Hydrology*, World Meteorological Organization, No. 385, Geneva.

Yevjevich, V. (1963). Fluctuations of wet and dry years, Part I, in *Hydrology Papers 1*, Colorado State University, Fort Collins, Colo.

Yevjevich, V. (1974). Determinism and stochasticity in hydrology, *J. Hydrol. 22*, 225-238.

Ziman, J. (1978). *Reliable Knowledge*, Cambridge University Press, London.