HYDROLOGICAL AND ENGINEERING RELEVANCE OF FLOOD FREQUENCY ANALYSIS

V. Klemeš
National Hydrology Research Institute
Department of Environment
Saskatoon, Saskatchewan S7N 3H5
Canada

"The engineer may be asked to evaluate the risks involved in a certain project and to take effective measures for the protection of lives and property in a catastrophic eventuality. To do this intelligently, the very basis for the design is often missing or hard to obtain: this is an adequate knowledge of the mechanics of the feared phenomena. It is indeed not the engineer's task to study natural phenomena as such: this is the task of the scientist."

Adrian E. Scheidegger (1975)

ABSTRACT. The mainstream of contemporary flood frequency analysis (FFA) qualifies neither as a hydrological science nor as an engineering discipline. It is not a hydrological science because it does not analyze the frequencies of floods - it merely postulates that flood records are random samples from simple probability distributions. It is not an engineering discipline because it pays most attention to, and exerts most energy on, formal polishing of concepts which are crude by their very nature and whose basic assumptions, which have an overriding influence on design parameters, are arbitrary and are dictated by expediency rather than scientific knowledge. At best, much of flood frequency analysis is just a part of small sample theory in disguise, the term "flood" being used merely as a name for the numbers employed; at worst, it is a pretentious game draining resources both from hydrology and engineering research, and a cheap opportunity to satisfy the need of academics to publish papers and supply easy topics for graduate students who know little beyond elementary statistics, probability theory, and computer programming.

1. INTRODUCTION

Flood frequency analysis (FFA) is regarded as a respectable scientific discipline. However, it has the typical features of a pseudoscience, in particular, (1) its most crucial predictions cannot be tested, (2) in response to criticism its exponents often resort to demagoguery.

V. P. Singh (ed.), Hydrologic Frequency Modeling, 1–18.

The first point follows from the basic concept which FFA employs - the return period defined as an average period between two successive attainments or exceedances of a given (maximum annual) flood peak discharge. Thus to verify an estimate of the most popular and basic FFA result, the "100-year flood", one would have to have many hundreds of years of record in order to be able to obtain a meaningful average period of such a length.

The second point is more subtle and often goes unnoticed. The reason is that, as in every well thought out demagoguery, the argument employs commonly accepted standards and beliefs or even facts but uses them to evade rather than meet a criticism. Typically, instead of showing where the challenger's criticism is wrong, the defender will change the issue. We all are familiar with such practices on the scene of international politics where totalitarian governments resort to them whenever challenged on some of their indefensible policies. For example, a critique of the suppression of a free access in a communist country to noncommunist press and literature may be countered by boasting about a universal access to a free medical care in that country, etc. The defense of the practices of FFA has a form of the following circular argument: (a) When a hydrologist criticizes FFA for its assumptions trivializing the complex and inadequately understood processes resulting in extreme floods, and for the consequent lack of realism of its results, the FF analysist will assume the posture of a down-to-earth engineering practitioner and say that the engineer cannot wait until everything is known and needs a result, however crude, now because he must make decisions on flood-related measures now. (b) When an engineer summons his courage and asks why then, given the simplifications and distortions of reality implied in the basic assumptions, should he employ all the sophisticated mathematical machinery to get that crude result, and why not use some simple procedure instead, the FF analysist will immediately change into a meticulous theoretician and extol the merits of mathematical rigour in the quest for extraction of the maximum amount of information from the data. (c) And should a statistician ask him how does he know that his rigorously fitted distribution models give information relevant to flood frequencies, the FF analysist would transform into a seasoned hydrologist and put him to shame for not being aware of the empirically established time-honoured ability of those models to represent hydrological data and of the good service they have rendered generations of hydrologists in such important tasks as flood frequency analysis, etc. At this point the argument returns to item (a) above.

In short, the contemporary FFA is well on its way to intellectual self-sufficiency and immunity to criticisms from any of the three directions from which it originally started; it is becoming an end in itself. It now often appears that streamflow and flood data are no longer gathered to provide statistical information on real floods for real-life engineering needs, but to supply numbers for Flood Frequency Analysis.

Taking cognizance of the few islands of sanity that still exist in the sea of FFA, one may relax its classification as pseudoscience and say that, at best, it is a discipline suffering from the "maladies of cookbookery and mathematistry." This diagnosis has been borrowed from G.E.P. Box who originally applied it to the state of health of statistics. He described the two maladies as follows:

"The symptoms of the former are a tendency to force all problems into the molds of one or two routine techniques, insufficient thought being given to the real objectives of the investigation or to the relevance of the assumptions implied by the imposed methods...Mathematistry is characterized by development of theory for theory's sake, which, since it seldom touches down with practice, has a tendency to redefine the problem rather than to solve it. Typically, there has once been a statistical problem with scientific relevance but this has long since been lost sight of." (Box 1976).

In our case, the legitimate problem has been the actual frequency of floods which has not only a scientific relevance but also a practical relevance in the context of the economics of flood protection. However, in FFA this legitimate problem has been redefined as an analysis of small random samples from a handful of simple probability distributions, and forced into the mold of curve-fitting techniques. In the following sections, this state of affairs will be examined in the light of Box's criticism, i.e. from the point of view of hydrological relevance and the the relevance to the original engineering problem of flood-protection.

2. SCIENCE, HYDROLOGY, STATISTICS AND MATHEMATICS

The real challenge of statistics is to discover, through analysis of empirical data, patterns in quantitative characteristics of specific objects or events encountered in nature. In this sense, statistics is simply a method of scientific inquiry and it has been through such inquiry that most of the statistical techniques (summarily called mathematical statistics) have been developed. As a rule, a statistical technique has been designed as an aid to the solution of a specific scientific problem, e.g. an assessment of the effectiveness of drugs, the control of the quality of a product, etc. This is why many of the best statisticians have repeatedly emphasized that, in order to make a meaningful analysis, the statistician must have a deep insight into the nature of the scientific problem. Several such statements were quoted in Klemeš (1978). To give just one example, it will suffice to turn to Box's paper quoted above where he says that "...statisticians must learn how to be good scientists."

In the context of flood frequency analysis it would mean that the statistician should learn how to be a good hydrologist. This, however, might turn out to be a difficult task. Those from whom he could learn hydrology are not only very few but seldom hold teaching positions. To his bewilderment, he might discover that, for instance, most courses on "Advanced Hydrology" contain little more than amateurish expositions of elementary linear algebra, probability and statistics. For, paradoxically, it has become almost axiomatic that to be a good hydrologist means to learn how to be a mediocre mathematician or statistician. This attitude is the result of a common misconception about the role of mathematics in science, in particular that, rather than being a method for a logically consistent and parsimonious formulation of scientific insights and exploration of consequences of scientific hypotheses, it is a generator of new scientific knowledge about the outside world; that correct mathematical manipulations and proofs lead ipso facto to correct statements about nature. This misconception has been a permanent source of frustration to scientists as well as to mathematicians who over and again keep encountering it among their students as well as among practicing scientists. Thus, for example, Richard Feynman, the physicist, says:

"Let me say also something that people who worry about mathematical proofs and inconsistencies seem not to know. There is no way of showing mathematically that a physical conclusion is wrong or inconsistent. All that can be shown is that the mathematical assumptions are wrong. If we find that certain mathematical assumptions lead to a logically inconsistent description of nature, we change the assumptions, not Nature" (Feynman, 1971, lecture 13, p. 8).

George E.P. Box, the statistician, expressed his frustration with the same phenomenon as follows:

"In such areas as sociology, psychology, education, and even, I sadly say, engineering, investigators who are not themselves statisticians sometimes take mathematistry seriously. Overawed by what they do not understand, they mistakenly distrust their own common sense and adopt inappropriate procedures derived by mathematicians with no scientific experience" (Box, 1976).

3. HYDROLOGY AND FFA FUNDAMENTALS

3.1. Frequency and Return Period

Frequency of floods, or of some of their quantitative characteristics like peak flows, volumes etc., is a legitimate object of scientific analysis. However, *the fundamental problem with FFA as practiced today is that it does not analyze the observed frequencies of floods.* In fact, the more interesting they are, the less attention they get. The only piece of factual information that FFA provides is that it tells us how many times (or in how many years) a given flood peak occurred or was exceeded during the period of record. However, even this information is presented in a form which implies an interpretation that can often be wrong. Firstly, the exclusive use of relative frequencies tends to suppress the importance of the length of the record from which these frequencies were computed and implies a degree of generality which may be unjustified. There is no a priori reason why, for instance, a flood peak that occurred twice in a specific 20-year period should typically occur five times in any 50-year period or, say, about ten times in the 100-year design period starting in 1990. Secondly, this appearance of generality is reinforced by the neglect of the influence of the actual historic dates of the flood record. Typically, a regional FFA may be based on records which not only differ substantially in their lengths but also in their dating although it is common knowledge that different periods in history often had different hydrologic regimes. Thirdly, the introduction of the

concept of return period as an average period between two successive exceedances may further contribute to misinformation, the more so because an average return period is a formally correct number even in cases where it conveys a completely wrong impression, i.e. where the phenomenon has no pronounced central tendency which the average is supposed to characterize. Thus, an example given in Klemeš (1987) shows that one can easily be led to classify as a 2 year flood, i.e. as a flood that was exceeded on the average every second year, a flood which in reality was not reached for 17 consecutive years but subsequently was exceeded in all but one of the following 12 years, only because it was exceeded twenty times within a 40-year historic record.

However, this last example notwithstanding, it is in connection with small and frequent floods where FFA is hydrologically most securely grounded and where it can most likely produce realistic results. This follows from the fact that data on small floods are abundant so that there exists solid empirical evidence regarding their actual return periods, both in different geographic regions and different historic periods. This evidence suggests that there indeed are many rivers where occurrences of small floods have a pattern of random events with a fairly stable frequency which can be meaningfully characterized by the average return period. But this is no excuse for forcing into this mold rivers deviating from this pattern or for automatically postulating it as a general pattern valid for cases arbitrarily far beyond the range of empirical evidence, both of which is routinely done in FFA.

It is important to note that, in order to extend the concept of return period beyond the range where the actual periods between the relevant floods are documented in historic records and thus can be averaged by calculating their arithmetic mean, FFA redefines the return period as the reciprocal of relative frequency. While the two definitions are mathematically equivalent, their "truth content" may be vastly different since the latter definition does not indicate from how many actual periods the (average) return period has been computed. Its most "appealing" feature is that it can produce, without even blushing, an "average" return period of a flood peak based on its one occurrence in a given record, that means *when in reality not a single period between this peak and its closest equal or higher neighbour is documented.* This heroic feat is, so to speak, one of the basic pillars on which the "scientific" edifice of FFA rests.

3.2. Randomness and Probability

In FFA it is assumed that the observed floods constitute a random sample from some mathematically defined probability distribution of floods. This assumption may be reasonable as a null hypothesis. However, since it provides the sole basis for the interpretation of results of FFA as well as for their applications in engineering decisions it should be carefully tested and continuously re-examined in the light of available scientific evidence. Instead, if it is tested at all, (as, for instance, has become a standard practice in Environment Canada in recent years), only crude statistical tests are generally used which are rather ineffective for the small samples usually available. The simple truth is that, while the independence hypothesis often cannot be rejected, neither could be hypotheses of many types of dependencies which suggest themselves on hydrological and other geophysical grounds. For example, there is evidence that maximum annual flood peaks often are related to annual runoff totals or mean flows (e.g. Sangal and Kallio, 1977; Panu et al., 1984) which in turn often exhibit more or less complex patterns of serial dependence traceable to various geophysical causes. However, such leads are seldom seriously explored (a rare recent exception to this rule appears in Booy and Morgan, 1985) and FF analysts are not only too happy to accept the independence assumption but they often postulate or tacitly imply it despite strong evidence to the contrary and justify this attitude by (what else!) the need to perform FFA. Only in the better cases are flood samples that do not pass a randomness test excluded from FFA (e.g. Panu et al., 1984). The philosophy of contemporary FFA is simple: if floods are to be considered for the analysis they *must* be random; or at least they *must be declared* random - otherwise they don't qualify. This might look puzzling only to some frivolous scientists like, for instance, Richard Feynman for whom Nature is just an object for academic nit-picking (see the earlier quotation) and who even seems to rejoice whenever something goes wrong - he is, for example, on record saying that "the thing that doesn't fit is the thing that's the most interesting, the part that doesn't go according to what you expected" (Feynman, 1983). Such a perverse attitude can of course be taken only by somebody divorced from reality, somebody who does not have the heavy responsibility of the FF analyst to aid engineering planning and design.

The crucial point is that, in "design and planning concerned with future events whose time or magnitude cannot be forecast, we *must* resort to statements of probability, or frequency..." (Linsley et al., 1975, p. 338; emphasis added). In order to meet this challenge, floods *must be declared* purely random events since otherwise the quantum jump from (relative) frequency to probability *cannot* be made and it does not follow that "Return period and probability are reciprocals" (ibid., p. 340). Since, while "return period" is at least a mathematically valid notion for every recurring event, be it a win in a lottery, a change in the political party in power, or the appearance of Halley's comet, it certainly is not valid to assume that the occurrences of each of these events in any given year have constant probabilities equal to the reciprocals of their respective return periods. In the case of floods, it simply has been found *convenient* to resort to statements of probability and to calculate flood risk by simple multiplication of elementary probabilities; given this concensus, we then must declare them random events regardless of the actual status of their frequencies in the range between a lottery and a Halley's comet. While this may be quite a legitimate attitude in the context of decision making (which will be discussed later), it certainly cannot be regarded as science. We simply are facing here the old problem identified by Yevjevich (1968) in the context of maximum probable precipitation (and flood), in particular a confusion between "engineering expediency" and "scientific truth" which also is at the root of the continuing controversy about

the U.S. Water Resources Council's (1977) endorsement of the log-Pearson III distribution as a basis of FFA in the United States.

3.3. Statistics Used as an Instrument to Justify a Preconceived Idea

Because of the practical importance of adequate flood protection measures, it would be desirable to obtain accurate estimates of the actual probabilities of flood occurrences. Since FFA postulates that floods are random events, it is not surprising that this task is redefined as a necessity "to obtain as accurate an estimate of p as possible" (Linsley et al., 1975, p. 350), where p is the probability of exceedance of a random event whose relative frequency is the same as that of a given flood in a given historic record. This end is pursued with the aid of mathematical statistics which is employed for the justification of the assumptions of randomness, homogeneity, stationarity, goodness of fit and other properties necessary for the accuracy of the above probability p. The justification rests on what may be called "numerical default" of statistical tests which relies heavily on the smallness of available flood samples. Rather than being a handicap, a small sample size has become a godsend which virtually guarantees non-rejectability of any hypothesis. Thus the first half of the success consists in "testing" the right hypothesis. The second half is assured by disregarding the main original purpose of a statistical test which is to arrive at hints of the presence or absence of relationships where these are not known a priori and where the only clue is their possible effect on the numbers.

Statistical test are meant for search in the dark, so to speak, and are pointless in bright daylight. For instance, one hardly needs a statistical test based on the distribution of the sizes of dogs to test a null hypothesis that a large animal is a dog when it is obvious that it actually is a horse. However, the approach of the FF analyst is first to close his eyes and then to say that, for the purpose of Dog Frequency Analysis, the animal can be considered a dog since this hypothesis cannot be rejected at the 5% significance level. On the other hand, if a St. Bernard was a puppy when the dog size distribution model was fitted, the FF analyst would likely exclude him as an outlier when he grew up even if he had a first class pedigree. Because it obviously is counterproductive to look and see if decisions made with closed eyes are both more convenient and can be better justified by rigorous mathematics. Once the status of a random sample from a simply distributed homogeneous population is scientifically "proven" in this fashion, great care is taken to adhere to the rigour of mathematical statistics in order that the estimates of p be obtained with the greatest accuracy, so that the "largest amount of information" (presumably on the size of big dogs or floods, as the case may be) is extracted from the data.

The remarkable feature of FFA is not that it deals with the shadowy world of "*as if* representations of reality" as Willeke (1979) put it in his stimulating paper but that so many otherwise rational people have been taking its mathematical games so seriously for so long, given that they could well have turned their attention to the study of the real features of flood frequencies (if they had a scientific interest in them), or to real engineering problems related to flood protection, or simply to theoretical work in mathematical statistics. What is even more striking is the apparent seriousness with which are regarded the most questionable results of FFA, namely those related to floods with relative frequencies lower than 1/100 and based entirely on wild extrapolations of concepts which are often difficult to take seriously even when only interpolation is involved.

4. EXTRAPOLATIONS OF FLOOD FREQUENCY MODELS

4.1. The General Problem

"With respect to flood frequency analysis, it's all in the tails, so to speak" (Greis, 1983). Extrapolation is the very raison d'être of flood frequency analysis and the assumption that historic floods are a random sample from a simple distribution is the only basis on which the extrapolated "probabilities" of extreme floods can be justified in principle. If this assumption is granted everything else is in order and the whole FFA is just one perfectly legitimate subarea of the small sample theory. And it is fair to say that several valuable advancements of the small sample theory have been stimulated by, and made within the framework of, FFA. In this context, computer simulation has proved extremely useful because it makes possible to test the theoretical predictions in ideal situations that strictly conform to the underlying assumptions. For the study of frequencies of actual floods, simulation results also have been extremely useful, though in a negative sense. They have shown that the estimates of parameters, inferences about population distribution types and about the behaviour of the extremes, etc., face fundamental limitations when attempted on the basis of small samples, even under the ideal conditions of simulation experiments which are not, and never will be, attainable in real flood samples. It is therefore important to give a serious thought to the additional obstacles that may be "superimposed" on simulation experiments by Nature and to re-examine the practices of FFA, notably its cavalier attitude to extrapolation, in this light.

4.2. The Physical Problem

Mathematical models of the behaviour of natural phenomena must fit the empirical data. Once fitted, a mathematical model can easily be extrapolated beyond the range of observations which it was designed to fit. However, there is no a priori reason why its correspondence with the manifestations of the physical prototype should hold beyond the range of observations. Science has documented many examples where an accurate fit in the observed range soon becomes inaccurate beyond it and eventually may become completely wrong; often there even occurs a sudden departure point beyond which the lack of the correspondence is not merely a matter of inaccuracy which can be removed when data in that range become available, but where the fit of the model cannot be repaired because its structure cannot accommodate the manifestations of the phenomenon in that region. Examples of this pattern are so many

that they suggest a general hypothesis that *no model developed for small range of a phenomenon will be valid for its entire range.* This is just another way of saying that different physical forces dominate a given phenomenon in different ranges of its states. To name a few specific cases, one may mention the eventual sudden collapse of Hooke's Law, the transition of laminar flow into turbulent flow, the sudden change of state of most substances with a gradual change in temperature, or the inapplicability of simple vector algebra for dealing with very high velocities.

For this reason alone it seems unlikely that a distribution model that fits the relative frequencies of a few, and mostly small, floods within a 20- to 50-year period, can be used for a meaningful prediction of a thousand- or a million-year flood which are now coming increasingly into fashion.

Another point relevant to such extreme floods is how far it is physically justified to extrapolate any distribution model with an infinite tail. One can sometimes hear that there is no limit since there always could be, say, a 1 mm more rain than there is in any conceivable rainstorm. This is a fallacious argument following from the inherent inability of the human mind to stop extrapolating which, since antiquity, has led to various logical paradoxes, e.g. the paradox of Achilles' inability to overtake the tortoise. This limitation of human mental faculties may easily lead a hypothetical Dog Frequency analyst to calculate the "probability" of a horse-size and even elephant-size dog; and the FF analyst to (scientifically!) estimate the peak flow of a million-year flood without ever thinking of a possibility that during such a deluge the whole area might be submerged under a stagnant pool of water with the flow being close to zero.

4.3. The Hydrological Problem

The main problem is that, in the climate of contemporary hydrology, it has become virtually impossible to discuss the hydrological problems associated with extreme floods. The several decades of FFA indoctrination has blurred most hydrologists' minds to the point that the only images that the notion of "extreme flood" can conjure up are those of distribution tails, quantile estimates, confidence limits, asymptotic distributions, extreme value distributions, outliers, etc. In short, hydrologists now often are able to think about floods only as about numbers to be subjected to statistical computations.

Flood frequency theorists often cannot comprehend that the difference between the occurrence of small and large floods is more fundamental than the difference between drawing a card marked with a small number and with a large one. Unable to see beyond a sample of random numbers each of which contains the same amount of information about the parent distribution, they cannot appreciate that the smallest floods from a given historic period may contain little or no information relevant to the likelihood of a very large flood, that the processes governing the formation of small and large floods may be vastly different. For example, one can often see that the whole shape of a distribution model fitted to a flood peak sample is dominated by the two or three smallest numbers which exert a commanding influence on the upper tail of the distribution and hence on the "probability" of the highest floods - in fact, the bigger the flood and the more hydrologically remote from the nature of the smallest one, the more its "probability" is influenced by the latter. An example given in Klemeš (1986) shows that, in a 25-year sample, almost negligible changes in the three smallest annual peak flows (which can hardly be called floods at all) can easily change the return period of a large flood from 100 to 10 000 years. Taking such an effect seriously - as is the rule in FFA - implies that the analyst believes that, say, a 2 cm difference in snow cover or a few hours difference in the duration of a mid-winter thaw in a small basin in Ontario has a controlling influence on the likelihood of a hurricane originating in the eastern Carribean and finding its way into the Toronto area. It is not argued here that such a connection is impossible in principle but that its implication in the result cited above would normally not even occur to the FF analyst, not to say that it should disturb him or lead him to investigate its substance.

There are other hydrological problems with the extrapolation of flood frequency data by models fitted to a few flood peaks from the lower part of the spectrum. For example, there are cases where, above a certain threshold not reached during the historic period, the flood generating area may suddenly increase and the likelihood of a very large flood may be much higher than the extrapolation of a model fitted to small floods would indicate.

One case in point is the Santa Anna River basin in California where floods from about a third of the basin (controlled by Lake Elsinor) normally do not reach the Santa Anna River but could do so in extreme conditions. Another similar case is documented in Sumas River in the western border area between Canada and the USA where in extreme conditions large overflow into Sumas River may occur from Nooksack River in the US. Such an overflow did not occur during the past 30 years during which records have been kept in Canada. Based on these records, the 100- and 500-year floods can be extrapolated as 52 m^3/s and 64 m^3/s, respectively. However, the U.S. authorities, considering this overflow, quote Q_{100} as about 270 m^3/s and Q_{500} as 540 m^3/s. Several such overflows occurred in the first half of the century and the largest that was actually measured (in 1932) was 127 m^3/s. A complementary effect takes place in the overflowing river where the peaks of extreme floods are correspondingly reduced. Diversion of flow from one basin to another during extremely high flows is quite a likely phenomenon in rivers with flat headwaters controlled by lakes as is often the case in Canada, e.g. the Ottawa and Gatineau Rivers in Quebec, Indian Brook in Newfoundland, etc.

Other aspects that may invalidate the upper tails of flood distribution models include inundation of large detention areas normally out of reach of flooding, floods caused by a collapse of morainic dams as is often the case in alpine streams (e.g. the 1983 flood on Nostetuko and Homathko Rivers in British Columbia), backwater effects of a tributary or a recipient, changes in the natural detention capacity of river valleys by geological processes, climatic anomalies caused by volcanic

eruptions, processes in the ocean, anthropogenic climatic and land use changes, etc. Frequencies of extreme floods provide a typical example of *a common situation where the key to a solution of a hydrologic problem lies outside hydrology*. The excellent work of Victor R. Baker and his collaborators (Baker, 1982, and references given herein) demonstrates how, for example, relatively simple geological insights can often be more illuminating in regard of extreme floods than volumes on various maximum likelihood, parametric and nonparametric, and other parahydrologic FF analyses.

In summary, from a hydrological point of view, very extreme floods and their causes tend to be outliers by definition, i.e. very little, if any, information about their likelihood is contained in the frequencies of relatively small floods of which the bulk of a typical flood sample is composed. Extrapolating distribution models fitted to these samples is tantamount to extrapolating the small flood dynamics beyond the range where it can physically function. For example, a 10 000-year flood extrapolated for a small Ontario basin on the basis of, say, a 30-year record of its ordinary snowmelt floods may be a nonsense because, whatever the amount of snow in the basin, there may never be enough energy available during a single day to melt snow at the rate implied in the corresponding flood flow. On the other hand, a summer storm or a hurricane may well cause such a flood quite easily and many times within a single century.

Thus the hydrological problem of a proper FFA is to predict the as yet unobserved outliers by analyzing the plausible dynamics of extreme-flood formation, rather than to worry about exclusion or inclusion of numerical outliers in empirical samples and about their effects on the fitted distribution tails. For it is quite conceivable that for an occurrence of an extreme flood, the tail of Halley's comet may be more important than the tails of all the standard distributions combined.

This may seem like an advocacy of the often ridiculed and presumably long discredited "probable maximum flood". This impression is partly correct, but only as far as the motivation behind the concept is concerned. The procedures for its implementation have many weaknesses and the name of the product is rather unfortunate though its defamation has in part been the result of a snobbery which, impressed by the presumed ability of FFA to quantify the probabilities of real large floods, fails to acknowledge the obvious fact that the term "probable" can legitimately be used in a nontechnical sense and that it may be more honest to do so than to hide arbitrary and often meaningless numbers behind its technical quantitative facade. However, the whole scholastic argument about the "probable" and the "maximum" could easily be avoided if labels like "unusually large flood", or "extreme flood", or "conceivable catastrophic flood" were used. For reasons which will be discussed in the next section, it may often be preferable to give just an idea of the magnitude of a flood that, based on current knowledge of the physical phenomenon involved, could conceivably occur on rare occasions than to base important decisions on seemingly exact values of *probabilities* which *probably* have little or no meaning.

5. BACK TO BASICS

5.1. What the Engineer Really Needs

The basic thing that the engineer needs in connection with flood-related planning and design is protection - not so much from floods but from accusations that his flood protection measures were (1) underdesigned (in cases where flood damage was not completely avoided), or (2) over-designed (in cases where it was). While this dilemma is not unique to flood-related engineering, it is there where it is exposed in perhaps the most dramatic manner. As a way out of this no-win situation, engineers have always been striving to develop "objective" approaches to design. Their common underlying feature is the replacement of an individual judgement with a collective one. This can be achieved by a variety of means ranging from an informal consensus of a professional organization often presented in the form of guides or handbooks containing recommended procedures, to more formal arrangements like standardization, regulations or even legislation.

The formation of a collective opinion at any level can be achieved only on the basis of some generally accepted principles. Among these, "maximization of public good" is probably the most common one but it is difficult to implement. For practical purposes, the latter has often been divided into a tangible and an intangible component in the hope that at least the first one can be objectively assessed. This has led to cost-benefit analysis, gradually generalized into "economic optimization", which has become perhaps the most important instrument of the planning and design methodology. Its universal appeal derives from the assumption that its result can attain a status of "scientific truth" free of subjective bias. This assumption is reasonable only to the extent to which economic optimization is based on relevant scientific facts and employs scientific methods. In theory, this caveat should naturally lead to a strong commitment to scientific research through which scientific facts can be obtained. In practice, pressures to perform economic optimizations and to maximize the credibility of their results often lead to pseudoscience, typically to a substitution of irrelevant scientific facts for the relevant ones which are difficult to handle or are inadequately understood, or to diverting attention from inadequacies in the substance to methodological subtleties. Flood Frequency Analysis is a perfect example of a result of such pressures.

However, the drift into irrelevance could be easily stopped if engineers did not let themselves be intimidated by pseudoscientific snobbery but exposed it to a scrutiny of sound engineering judgement, of the common sense to which Box appeals in the earlier quotation. It would then soon become obvious that the scientific content of economic optimization is rather modest. This is so even in quite simple and straightforward cases and becomes an illusion in complex situations, especially those involving modification of the environment as is the rule with civil-engineering water-related projects. Here, economic optimization is merely a conceptually appealing rationale for arriving at an acceptable collective opinion. Its appeal comes from its formal resemblance of the straightforward mathematical problem of finding an extremum of the sum of two monotonous functions, one of which is decreasing and the other increasing. The other element of its appeal is the fact that we have been conditioned to homogenizing various non-commensurate utilities by expressing them in monetary equivalents. Thus what we treat as "scientific facts" are often entities that have been made facts by definition or concensus. For example, we can freely mix apples and oranges when we define both as commodities whose only attributes are their unit prices. Similarly, it is we who define what should be included into the "cost" and into the "benefit" brackets and how far their effects should be traced; we treat as objective facts the monetary equivalents while we know that they are a result of many subjective and arbitrary manipulations like taxes, tariffs, interest rates, market fluctuations, speculations, government subsidies, agreements, etc. etc.

In view of this, it is far more desirable to acknowledge explicitly that economic optimization is just an agreed upon rationale and try to standardize it on the basis of simple procedures than to keep pretending that it is a science and try to perfect it by refining the methodology for manipulating its "scientific facts". This would relieve the pressure for the cultivation of the supporting pseudo-sciences such as Flood Frequency Analysis which then itself could be recognized for what it is, namely a rationale useful for performing a cost-benefit analysis, and standardize it within a simple framework. "Bulletin 17" (U.S. WRC, 1977) has been a bold step in this direction and should be viewed in this context. It is unfortunate that the hydrologic community has let itself be tricked into futile arguments about its "scientific" merits and demerits which is a pseudoproblem par excellence.

5.2. Flood Frequency Analysis as a Rationale

"Until this problem of estimating the magnitude of floods with return period of at least 1,000 years [is solved], no satisfactory quantitative hydro-economic decisions based on balancing costs and benefits can be made" (Alexander, 1973). This is the kind of reasoning that has been responsible for pushing FFA into the realm of pseudoscience. The fallacy implied in the above statement is that a "satisfactory" decision is taken to mean an "objective", or "scientifically based" decision. The crucial point is that a satisfactory decision does not necessarily have to be based on scientific knowledge. In many cases such knowledge is simply not available and decisions have to be made regardless. They are judged satisfactory or otherwise according to the degree to which they satisfy, to use a fashionable term, the currently accepted paradigm, not according to their scientific content.

Paradigms may, and necessarily do, have a strong component of myth - in modern times a significant part of this myth component has been a belief that science can alleviate the burden of responsibility in decision making, that "scientific results" may in a way replace decisions. This flight from responsibility and from courage to take a stand then pushes into a "scientific" framework things that have no scientific content thus giving rise to various pseudosciences

(Berlinski, 1976). Facing the ever growing complexities of life and being more and more aware of the multitudes of risks and potentially disastrous consequences of wrong decisions, it is becoming increasingly more "desirable" to be able to wash one's hands and hide behind science and computers. In other words, the subjective element of decisions is pushed one or more steps back, and out of reach of public scrutiny (Bazelon, 1979).

This trend is fundamentally very unhealthy and dangerous and, although in flood-related decisions it may be less dangerous than elsewhere, there is no need for its continuing cultivation. Given the fact that neither the economical nor the hydrological component of flood-related economic analyses have a solid scientific basis, it would be more honest, convenient and practical to anchor the procedures to standards decided upon within the profession than to continue trying to endow the transparent "as if" game with more respectability via more mathematistry. The weaknesses on the economical side have been abundantly discussed by specialists (e.g. Loentief, 1982; Eichner, 1985) and some of the hydrological weaknesses were pointed out earlier in this paper. Some practical reasons for stopping to pretend that FFA is a science are discussed below.

Science is a continuous search for "...order which gives unity to what had long seem unlike... For order does not display itself of itself... [but] must be discovered and, in a deep sense, it must be created. What we see, as we see it, is mere disorder" (Bronowski, 1972). FFA is obviously not in this category. It does not search for patterns in flood occurrences but is only too happy that its basic assumption - that there is no element of pattern involved - "cannot be rejected". *FFA does not use predictions to test the validity of its hypotheses but uses hypotheses to make untestable predictions.* Its only scientific content is the empirical frequencies of floods in a given period of record. Everything beyond and above, including the interpretation of the reciprocal of these frequencies as return periods, equating them with probabilities, representing the latter by various distribution models, etc. is just a superstructure of hypotheses. These hypotheses are not inspired by the study of the process involved but dictated by a need of the decision maker to quantify flood risk in a way acilitating a unique result of cost-benefit analysis of flood related projects.

Thus the driving force behind FFA is not a scientific discovery but a necessity to make a decision. This necessity is an objective reality imposed by society. Given that, the need for some kind of rationalization of the decision is undisputable and quite independent of whether and to what extent the rationalization can be based on scientific knowledge of the problem. Where facts and scientific knowledge are not available the rationalization must be based on other foundations, for example on assumptions, on recourse to a higher authority, tradition, custom, etc. This is perfectly legitimate and beyond challenge. What is not legitimate is to pretend that the latter entities are elements of scientific knowledge, or that they somehow will be transformed into them by being forced into a scientific format either by decree or by subjecting them to treatment by scientific

methods. Although this has been pointed out on many occasions (e.g. Hall, 1971; Feynman, 1983), it is still common practice because of the high prestige which the "scientific status" enjoys in the contemporary paradigm.

Once we accept that the present FFA is merely a rationale which accommodates the empirical evidence that large floods occur less often than small floods, and formalizes it in a framework convenient to cost-benefit analysis of design, several things become obvious, in particular: (1) That the whole edifice of models and techniques based on the assumption that floods are random samples from simple distributions is only an "as-if" representation of reality (Willeke, 1979); (2) That there is no theoretical reason why some distribution models should fit the data from our small samples better than others and that the reasons why certain models have been used are purely practical (Moran, 1959); (3) That "the part of the distribution we are interested in is well away from the part where the observations provide some information about the shape. It is therefore easy to construct two different distributions both of which fit the observations closely but for which the tails are of quite different shape" (Moran, 1957, 1959); because this problem arises from the shortness of our records it "cannot be overcome by mathematical sleight of hand" (Moran, 1957). This was well demonstrated by Wallis and Matalas (1973) who showed by a simulation experiment that, for a small sample, surprisingly often the "best fit model" is different than the parent distribution. Not surprisingly, this reference is seldom cited in FFA literature; (4) That, therefore, "... in so far as extrapolation made beyond the range of the recorded floods is concerned, the most elaborate analytical procedure of curve fitting gives results that are no more reliable than those obtained by a simple extension by eye of the flood frequency curve on any kind of probability paper" (Thomas, 1948); (5) That, if we have to extrapolate, then we should pay much more attention to the largest recorded floods than to the small ones since they presumably better reflect the type of conditions influencing the tail behaviour which we want to estimate (Harold Thomas' rationale underlying the Wakeby distribution; also see Reich and Renard, 1981); and finally (6) That, in the light of the above points, the "mathematical prestidigitation" (Moran, 1959) which has become the main preoccupation of contemporary FFA is a preposterous scholasticism that has no value for hydrology or engineering.

This conclusion is further supported by the following:
(a) Flow measurements under flood conditions are very difficult and unreliable so that the upper sections of flow rating curves are subject to large errors. Moreover, flood peak flows are almost never measured but extrapolated from rating curves or by indirect methods. Measured flows are seldom higher than 50% of peak flows. Cases where the maximum measured flow is under 20% of estimated peak flows are not uncommon. For example, on Yakoun River at Port Clements in British Columbia, the annual maxima in the four years 1962-1965 ranged from 396 to 612 m^3/s while the maximum measured flow during this period was 72 m^3/s (for examples from Newfoundland see Panu et al., 1984). (b) Flood peaks are often estimated from single daily staff gauge readings. A

recent study done for coastal streams in British Columbia indicates that errors due to this single cause can be as high as 90%; in 10% of all the cases analyzed it was greater than 50%. (c) However, in FFA, peak flow data are treated as exact deterministic numbers and the only errors recognized are those between these "exact" numbers and their model estimates. These then serve as a measure of "confidence" in the estimates of the various T-year floods which is given in "exact" probabilistic terms. *Few FF theorists seem to question the wisdom of striving for a high confidence in estimating numbers which can claim only a very low confidence level themselves.* (d) From the point of view of the final product, i.e. the result of the cost-benefit analysis, it often makes little difference what distribution model is fitted to the data (Slack et al., 1975). This is the consequence of three things; one, that the influence of the body of the distribution far outstrips that of its tail - an observation made by Moran (1959) in connection with storage design; two, that this body can be equally well fitted with many different models (see item 3 above); and three, that all the fitted models are made to "preserve" the basic parameters of the sample, especially the mean and the variance, which have a dominant role in the standard economic analyses (Klemeš, 1977). (e) It is generally the case that cost-benefit analysis of risk plays an important role only in projects exposed to damage from relatively small and frequent floods. Whenever damages of catastrophic proportions are envisaged - typically involving danger to human life, major disruptions in the functioning of society's infrastructure, danger to strategic defence installations or to disposal sites of highly toxic or radioactive wastes, etc. - flood protection levels are invariably set by policy decisions which are outside the realm of an "objective" cost-benefit analysis. Here, the notions as 100-year, 1000-year or million-year flood have mostly qualitative and public-relations significance and could well be replaced by terms like dangerous flood, extremely dangerous flood, flood of catastrophic proportions, etc. The quantification in terms of the return period can often be even misleading and convey a greater feeling of safety than it actually implies; for example, the flood stage of a "1000-year" flood may often be only a few centimetres higher than that of a "100-year" flood although in the "public-relations" perception a "1000-year" protection seems by an order of magnitude higher than a "100-year" protection. Hence, there is little justification for "mathematical prestidigitation" where the decision is determined largely by the emotional impact of large rounded numbers of the decimal system in very much the same way as in the case of setting the values of lottery prices where a "million-dollar prize" always carries the same fascination which seems to be almost unaffected by inflation.

As a result it can be safely concluded that the chief and only claim in favour of mathematical sophistication in FFA, in particular, of the importance of "exactness" and "accuracy" in engineering decision making, is false. Engineers don't need the "science" of FFA but a simple agreed upon framework in which the rationale of the "as-if" representation of reality can be implemented. Perfection of curve fitting procedures, of estimation and inference problems in small sample theory may be an interesting and legitimate pursuit by themselves - but they are neither *engineering* nor *hydrology* and there is no valid reason why they should continue to drain the much needed resources and talent from these areas.

6. ACKNOWLEDGEMENTS

Many practicing hydrologists provided materials illustrating the points made in this paper, and helped me to sharpen the focus through discussions of the problems raised herein. Special thanks are due to the following individuals: E.G. Smith, and G. Tofte (Environment Canada, Pacific and Yukon Region, Vancouver, B.C.), P.J. Pilon, R. Condie, D.A. Smith and B.P. Sangal (Environment Canada, Ottawa, Ontario), R.R. Wyman, R.Y. McNeil, W. Obedkoff and C.H. Coulson (Ministry of Environment, Victoria, B.C.), M. Miles (Miles & Associates, Ltd., Victoria, B.C.), J.Y. Ding (Ministry of Natural Resources, Toronto, Ontario), J.R. Beck (U.S. Geological Survey, Sacramento, California), and H.E. Kubik (Hydrologic Engineering Center, U.S. Army Corps of Engineers, Davis, California). This does not necessarily imply their endorsement of the opinions expressed herein which are my own.

7. REFERENCES

Alexander, G.N. 1973. 'Estimation of the 10,000-Year Flood.' In: *Proceedings, 11th Congress on Large Dams*, Q.41, R.70, 1327-1350, International Commission on Large Dams, Madrid.

Baker, Victor R. 1982. 'Geology, Determinism, and Risk Assessment.' In: *Scientific Basis of Water Resource Management*, 109-117, National Academy Press, Washington, D.C.

Bazelon, D.L. 1979. 'Risk and Responsibility'. *Science*, **205**, 277-280.

Berlinski, D. 1976. *On Systems Analysis*. MIT Press, Cambridge, Mass.

Booy, C., and Morgan, D.R. 1985. 'The Effect of Clustering of Flood Peaks on a Flood Risk Analysis for the Red River.' Canadian Journal of Civil Engineering, 12, 150-165.

Box, George, E.P. 1976. 'Science and Statistics.' *Journal of the American Statistical Association*, 71(356), 791-799.

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

Eichner, A.S. 1985. 'The Lack of Progress in Economics.' *Nature*, 313, 427-428.

Feynman, R.P. 1971. *Lectures on Gravitation*. California Institute of Technology, Pasadena, California.

Feynman, R.P. 1983. 'The Pleasure of Finding Things Out.' NOVA Program No. 1002, WGBH Transcripts, Boston.

Greis, Noel P. 1983. 'Flood Frequency Analysis: A Review of 1979-1982.' *Reviews of Geophysics and Space Analysis*, 21(3), 699-706.

Hall, Warren A. 1971. 'Biological Hydrologic Systems.' In: *Proceedings, The Third International Seminar for Hydrology Professors, Biological Effects in the Hydrological Cycle*, 1-7,

Purdue University, West Lafayette, Indiana.
Klemeš, V. 1977. 'Value of Information in Reservoir Optimization.' *Water Resources Research*, 13(5), 837-850.
Klemeš, V. 1978. 'Physically Based Stochastic Hydrologic Analysis.' In: V.T. Chow (Editor), *Advances in Hydroscience*, 11, 285-355, Academic Press, New York.
Klemeš, V. 1986. 'Dilettantism in Hydrology - Transition or Destiny?' *Water Resources Research*, 22(9 Supplement), 177S-188S
Klemeš, V. 1987. 'Empirical and Causal Models in Hydrologic Reliability Analysis.' In: L. Duckstein and E. Plate (Editors), *Engineering Reliability and Risk in Water Resources*. Proceedings, NATO Advanced Study Institute, The University of Arizona, Tucson, May 19-June 2, 1985.
Leontief, W. 1982. 'Academic Economics.' *Science*, 217, 104 + 107.
Linsley, Ray K., Kohler, Max A., and Paulhus, Joseph L.H. 1975. Hydrology for Engineers, McGraw-Hill, New York.
Moran, P.A.P. 1957. 'The Statistical Treatment of Flood Flows.' *Transactions of American Geophysical Union*, 38, 519-523.
Moran, P.A.P. 1959. *The Theory of Storage*. Methuen, London.
Panu, U.S., Smith, D.A. and Ambler, D.C. 1984. *Regional Flood Frequency Analysis for the Island of Newfoundland*. Department of Environment, St. John's, Newfoundland; and Environment Canada, Atlantic Region, Dartmouth, Nova Scotia.
Reich, Brian M. and Renard, Kenneth G. 1981. 'Application of Advances in Flood Frequency Analysis.' *Water Resources Bulletin*, 17(1), 69-74.
Sangal, B.P. and Kallio, R.W. 1977. *Magnitude and Frequency of Floods in Southern Ontario*. Inland Waters Directorate, Environment Canada Ottawa.
Scheidegger, Adrian, E. 1975. *Physical Aspects of Natural Catastrophes*. Elsevier, Amsterdam.
Slack, J.R., Wallis, J.R. and Matalas, N.C. 1975. 'On the Value of Information to Flood Frequency Analysis.' *Water Resources Research*, 11(5), 629-647.
Thomas, Harold A., Jr. 1948. 'Frequency of Minor Floods.' *Journal of the Boston Society of Civil Engineers*, XXXV(4), 425-442.
United States Water Resources Council, Hydrology Committee 1977. *Guidelines for Determining Flood Frequency Information*, Bulletin 17A, Washington, D.C.
Wallis, J.R. and Matalas, N.C. 1973. 'Relative Importance of Decision Variables in Flood Frequency Analysis.' In: *Design of Water Resource Projects with Inadequate Data*. UNESCO, Studies and Reports in Hydrology, No. 16, 435-442.
Willeke, Gene E. 1979. 'Myths and Uses of Hydrometeorology in Forecasting.' In: *Improved Hydrological Forecasting - Why and How*. Proceedings of the Engineering Foundation Conference, 117-124. American Society of Civil Engineers, New York.
Yevjevich, V. 1968. 'Misconceptions in Hydrology and their Consequences.' *Water Resources Research*, 4(4), 225-232.

STATISTICAL FLOOD FREQUENCY ANALYSIS – AN OVERVIEW

Arun Kumar
Professor of Civil Engg.
Delhi College of
Engineering,Delhi,India

Subhash Chander
Professor of Civil Engg.
Indian Institute of
Technology,New Delhi,India

ABSTRACT. The paper critically examines various issues involved in statistical frequency analysis. Emphasis has been laid on areas where further research is required. It is highlighted that the Water Resources Council modified report of 1981 has largely reiterated their earlier findings thereby ignoring criticism labelled by various researchers on their recommendations. In the light of recent studies, it is advocated that WRC,U.S.A recommendations need reevaluation.

INTRODUCTION

Statistical flood frequency analysis has probably been one of the most active areas of hydrological research for the last thirty or more years. Since fifties almost every issue of journals related to Water Resources research contains papers on this topic. With such a voluminous research material available, one would like to ask whether significant progress has been made in understanding or in arriving at more acceptable methodologies for frequency analysis.The answer to this question is rather difficult. The questions such as (i) which parent distribution the data may follow? (ii) how to account for sampling error while identifying the distribution (iii) what should be the most suitable parameter estimation technique and many other similar questions remain unresolved. It looks that as long as the flood frequency analysis remain within the statistical domain,the problems of this nature will remain persisting. Being aware that statistical flood frequency analysis has inherent shortcoming of not explaining in any rigorous manner the flood phenomenon itself, researchers have tried many approaches other than statistical. For example, the use of geomorphologic unit hydrograph (GUH) in regional flood frequency analysis (Rodriguez-Iturbe and Valdes,1979, Diaz-

V. P. Singh (ed.), Hydrologic Frequency Modeling, 19–35.