The Improbable Probabilities of Extreme Floods and Droughts

V. KLEMEŠ

National Hydrology Research Institute,
Department of the Environment, Saskatoon, Canada

Introduction

THE CURRENT approach to the estimation of probabilities of extreme floods and droughts is based on analyis of extremes in historic streamflow or precipitation records. The main weakness of the analysis is that it takes no account of the actual climatic, hydrological and other geophysical mechanisms that produced the observed extremes. Rather, it is based on arbitrary postulates of preconceived probabilistic mechanisms so that the results do not reflect what is likely to happen in nature, but what would happen if these postulates were correct. The crucial aspect, i.e. whether they actually are correct, is not addressed at all. The usual excuse for this is that there are not enough data, the physical processes are not understood well enough and the planners, engineers, managers, etc. cannot wait with their decisions until these problems are resolved; hence, to comply with the practitioners' requests for estimates of 100-year, 1,000-year, and (lately) even 1,000,000-year floods and droughts, these crude methods have to be used.

This attitude reflects a confusion between the requirements of current decision making and the need for improving the scientific basis for future decision making. In case of hydrologic extremes, the latter has been entirely sacrificed to the former – the above mentioned excuse has been repeated for at least 50 years during which time much progress on the real difficult issues could have been made but has not, so that the present-day planners are no better off than their grandfathers were.

Probabilistic models of physical phenomena

Probabilistic modelling of physical phenomena faces one fundamental dilemma. On one hand, it is motivated by a realization that the causal structure (dynamic mechanism) of the given phenomenon cannot be explicitly described for one reason or another: it may be unknown, known but forbiddingly complex, unknowable, or it even may be considered nonexistent. On the other hand, the leading probabilists and statisticians seem to agree that a meaningful probabilistic (stochastic) model of a physical phenomenon cannot be constructed, unless the dynamics of this phenomenon are fairly well understood (some pronouncements to this effect by scientists like Wiener, Bartlett, Kendall and Stuart, and Box were quoted in Klemeš, 1978 and 1986).

The first pole of this dilemma leads, by default, to a replacement of the unknown dynamics with probabilistic assumptions which are compatible with the statistics derived from observations of the given phenomenon, but otherwise dictated by mathematical convenience and requirements of the available theoretical apparatus, sometimes even with full awareness of their conflict with reality. To give simple examples of the latter, one may mention the often invoked normality of a distribution of a phenomenon which has obvious physical bounds, or ergodicity in cases where, physically, only one realization of a 'process' can exist.

This route to probabilistic modelling is essentially a generalized curve fitting and may be acceptable (and often useful) for the reduction of data from observations, especially where they are required as inputs for various straightforward applications (e.g. mapping, smoothing, interpolation). A model of this nature may provide a good enought ad hoc description of a set of observations, but there is no reason to assume that this description adequately represents probabilities of states of a physical system on which these observations were made. This means that such a model may be completely useless and misleading for extrapolation beyond the range of observations.

It has been in recognition of this fact that statisticians have been pointing to the other pole of the dilemma, i.e. emphasizing the need for understanding the dynamics that has led to the observed states, as a necessary prerequisite for building a good probabilistic model. By 'good' model is meant one that has some measure of credibility when used for extrapolation, i.e. a model that can tell us more about the given phenomenon than the observations themselves reveal. It can be argued that only in such a case the term 'model' is justified; otherwise it is just an inflated name for an interpolation formula which, when used for extrapolation, may be an extremely unreliable and dangerous tool. In my opinion, this is the case with the overwhelming majority of approaches to the estimation of probabilities of extreme floods and droughts. In this regard, we have no hydrology of disasters; we only have a disaster of hydrology.

Probability in Hydrological Context

The general practice in contemporary hydrology is to regard the probability of a hydrological event as a limit of its relative frequency observed in the historical record. Two assumptions are crucial for this concept, in particular (1) that hydrological processes are stationary over time, and (2) that they are realizations of an ergodic stochastic process.

The first assumption (stationarity) implies that a hydrological phenomenon, e.g. streamflow at a given gauging station, fluctuates in time around a constant value in a statistically constant pattern. This further implies an assumption of a physical constancy of the mechanisms participating in the formation of the streamflow, from the regimes of precipitation and evaporation in the river basin, to geomorphological, pedological and other physical conditions. It is of course well known that this assumption is not true in general and that it diverges from reality with the length of the period considered. As the length of record increases and, with it, the nominal statistical significance of parameters derived from it, the validity of the stationarity assumption decreases. Thus, despite the preaching about the importance of long records, hydrologists are in fact more comfortable with short ones; the stationary hypothesis is much more easily defensible for, say, a 30-year record that it would be for a 300-year record. This leads to a paradox that the estimates of population parameters, population distributions and probabilities, the philosophical justification of which comes from the ideal of a very long record, makes sense only when based on a short record. Then, however, doubts arise about the validity of estimates of characteristics reflecting the long-term properties of the process. While it may be reasonable to assume, for instance, that the past 30 years may provide an adequate picture of the range of floods that one could expect in the immediate future since the physical conditions may not have changed much during that period, it is entirely different to say that, based on a distribution fitted to this 30-year record, a 1,000-year flood will be such and such.

In other words, making a statement about high extremes based on a statistical model fitted to a short record is similar to making a statement about extremes of a wavy curve based on the properties of extremes of a tangent representing it in a given point.

The second assumption (ergodicity) implies that the historic record of a hydrological phenomenon can be regarded as one of an infinite number of equally likely realizations of a stochastic process in which ensemble averages are equal to time averages. This is the most fundamental assumption on which rests the whole concept of 'probability of a hydrological phenomenon'. Without the assumption of an ensemble of realizations and the assumption that the incidence of possible events across the ensemble at a given instant is the same as was the incidence of events along the historic realization, the present concept of hydrological probability is unthinkable. And yet, from a geophysical point of view both these assumptions are not only arbitrary and unrealistic; they deliberately make a mockery of reality, of the evolutionary character of the history of a geophysical process and of the uniqueness of this history. They practically reduce all the history to a noise which could have proceeded in the reverse direction as well as in the actual one or in any other rearrangement – and the same is implied for the future. They completely negate the fact that there were specific signals associated with specific events and that many of these signals, especially those associated with the very extreme events, physically cannot be repeated either at all or at least not with the same probability in every consecutive year or instant. They provide an excellent example where a mathematical concept developed to describe one physical situation (some phenomena in thermodynamics in this case) has been applied to a situation it does not fit for no other reason than mathematical convenience.

The leap of logic by which the instantaneous probabilities are equated with the historic frequencies of occurrence is nothing else but a dismissal of any meaning of the historic process: if anything that happened in the past can happen at any instant with the same likelihood, then the history provides no meaningful information.

We are facing perhaps the greatest paradox of probabilistic statements about hydrological phenomena. They claim to give information about the future – and they arrive at this information by first suppressing most of the information from the past, by denying any significance to the order of the past events.

The automatic identification of past frequencies with present probabilities is the greatest plague of contemporary statistical and stochastic hydrology. It has become so deeply ingrained that it prevents hydrologists from seeing the fundamental difference between the two concepts. It is often difficult to put across the fact that whereas a histogram of frequencies of given quantities (or the empirical return periods) can be constructed for any function whether it has been generated by deterministic or random mechanism, it can be interpreted as a probability distribution only in the latter case. A perfect sine wave will yield a histogram of frequencies which have not the slightest probabilistic meaning. *Ergo*, automatically to interpret past frequencies as present probabilities means *a priori* to deny the possibility of any signal in the geophysical history; this certainly is not science but sterile scholasticism.

The point then arises, why are these unreasonable assumptions made if it is obvious that probabilistic statements based on them may be grossly misleading, especially when they relate to physically extreme conditions where errors can have catastrophic consequences? The answer seems to be that they provide the only conceptual framework that makes is possible to make probabilistic statements, i.e. they must be used if the objective is to make such probabilistic statements.

It is like the well known joke where a policeman sees a drunk diligently searching for something under a street lamp.

'What are you looking for?' asks the policeman.
'I have lost my key.' says the man.
'And did you lose it here?'
'Well, no, but this is the only place where I can see anything'.

Extreme Floods

In contemporary hydrological practices the esimation of probabilities of extreme floods has become a rather trivial problem which requires no understanding of either hydrology, climatology, probability theory, statistics or any other relevant science. All that is necessary is to obtain the values of the maximum annual flows from the historic record of, say, 30 years, arrange them in order of their magnitude on an internal (0, 1), fit this empirical distribution of relative cumulative frequencies with a few simple S-shaped mathematical functions, select the 'best fit' by some standard curve fitting technique and read the exceedence 'probability' of any given flow off the upper tail of this curve as the value of the corresponding decimal fraction; alternatively, the value of this decimal fraction, say 0.001, is chosen and the corresponding value of the flow is designated as the 1,000-year flood.

All this is conveniently done by computers using one of the many available software packages for flood frequency analysis. The most demanding task in the exercise probably is to enter the historic maxima into the computer. However, the most difficult task, at least for this writer, is to believe that a number obtained in this standard manner has much to do with the real probability that a flood of the given magnitude could occur in any given year.

Some of the specific reasons following from the general criticism outlined in the preceding sections are given below. A short historic record may perhaps cover a period during which the climatic and hydrological regimes were close to stationary and the observed hydrological activity really reflected mostly the random fluctuations of energy within the system. However, without considering the types of energies taking part in these 'normal' perturbations of some underlying signal, it is by no means obvious how far from equilibrium this noise can move the system. Fitting a curve, extending to infinity, through a few observed states does not in any way indicate how far the corresponding energy situation makes extrapolation along this curve physically meaningful. Maybe the system could not produce an event labelled a 1,000-year flood at all, maybe it would be capable of generating much higher floods with a much higher probability. This cannot be decided by fitting one or other mathematical function to a given set of points but only on the basis of a good understanding of the behaviour of the physical processes involved. There may well be two similar sets of annual maxima, with similar best-fit curves, but coming from different climatic regions, for instance one dominated by snow melt floods and the other caused by convective storms. Is there any reason to assume that a 1,000 or 10,000-year flood will be of a similar magnitude in both of them?

Another point which is extremely important is a possibility that long before the processes reflected in a short historic record could, through their random interactions, produce a really extreme flood, another process, perhaps not at all active during the period of record, could intervene and produce a much bigger flood. A case in point may be eastern Canada where most annual maxima are generated by snowmelt but occasionally a hurricane can stray into the region and cause a flood much larger than is conceivable via the snowmelt mechanism.

It is becoming increasingly more acceptable to extrapolate the fitted 'probability distributions' of floods beyond limits considered reasonable in the past. While one can hardly find a reference to anything larger than a 100-year flood in the literature from the turn of this century, a 1,000-year flood was quite a common concept 30–40 years ago and today a 100,000-year or a million-year flood may seriously be invoked in connection with flood protection of toxic-waste dumpsites, strategic military installations, or nuclear power plants. While it is easy enough to extrapolate the 'best fit' curves to these extremes, it is also easy to arrive at complete absurdities in this manner. For example, one could 'estimate' a million-year flood with peak flow of x m^3/s on the Manitou River on the Manitoulin Island in Lake Huron. But it may well be that under conditions necessary to produce such a flood, the whole Manitou River, and even the Manitoulin Island, would disappear so that no such peak flow could ever arise there.

Without an analysis of the physical causes of recorded floods, and of the whole geophysical, biophysical and anthropogenic context which circumscribes the potential for flood formation, results of flood frequency analysis as described above, rather than providing information useful for coping with the flood hazard, themselves represent an additional hazard that can contribute to damages caused by floods. This danger is very real since decisions made on the basis of wrong numbers presented as good estimates of flood probabilities will generally be worse than decisions made with an awareness of an impossibility to make a good estimate and with the aid of merely qualitative information on the general flooding potential. This warning is more relevant now than it may have been in the past because of the increasingly apparent nonstationarity of the climatic signal which suggests that the assumption of stationarity is no more defensible even on a very short time scale of two or three decades. Other aspects that should be considered in assessing the potential of extreme floods are the land-use changes in the basin, effects of the fluctuation of ocean heat storage on local weather patterns, effects of water storage in the basin on the flood response to precipitation, effects of volcanic eruptions on flood climates, effects of crustal movements on watershed boundaries (especially in flat regions such as, for example, the Great Lakes Basin), changes in drainage areas for small and large floods (e.g. the case of the Santa Anna River basin in

California), effects of the changing patterns of atmospheric deposition on the melting of snow and glaciers, effects of tectonic activity on the safety of large dams, effects of various periodic signals (e.g. Currie and O'Brien, 1988) and perhaps other phenomena (e.g. see Geophysics Study Committee, 1978). A more detailed criticism of current practices was given in Klemeš (1987a), while some promising new physically based approaches to assessing the flood potential can be found in Baker *et al.* (1988). Unfortunately, there has not been much innovative statistical thinking about floods besides the rudiments proposed by Todorovic and Zelenhasic (1970), and Eagleson (1972).

Extreme Droughts

The prevailing attitude to the probabilities of extreme droughts is essentially the same as to those of extreme floods and is well summarized in the following statement: 'Just as the Gumbel and the Frechet distributions are a natural choice for flood data because they are tailored to the properties of sample maxima, the Weibull distribution is an attractive choice for data which can be interpreted as minima of samples' (Beran and Rodier, 1985, p.75). Thus the only tangible difference is the type of the mathematical function of the likely 'best fit' from which the 'probabilities' are estimated. At most, the higher persistence of drought-related data is sometimes taken into account to reduce the return periods compared to those corresponding to random samples (which is how flood samples are treated).

If the current practice of estimating probabilities of extreme floods by extrapolating the upper tails of 'best-fit distributions' is bad, then estimating probabilities of droughts by extrapolating their lower tails is simply awful. While floods do exhibit some features of randomness in that they appear suddenly, may affect only small areas, and their long-term patterns indicate a high level of noise, droughts appear to represent a markedly different type of process. Typically, they develop slowly over time (they have been labelled a 'creeping' phemomenon), last much longer than floods, affect large and often enormous territories, tend to show cyclic patterns, etc. For example, the 1930s drought affected three continents and extended from north-east America through Europe to west Siberia (Shiklomanov and Markova, 1987); the current Sahel drought has lasted for about 20 years; during the past 80 years South Africa experienced four drought cycles centred around the 1910s, 1930s, 1950s and 1970s (Tyson and Dyer, 1978), etc.

From a physical point of view, droughts usually represent a process of a different order than do floods and 'symmetry' between minima and maxima of the same process generally cannot be invoked. The general relationship between the processes of floods and droughts can perhaps be best illustrated with the following simplified example. If floods were represented by the maxima of a normalized function of time, $x(t)$, then droughts would be represented much better by the minima of a normalized function $y(t) = \int x(\tau)d\tau$ than by the minima of $x(t)$ itself. Depending on the specific type of drought (meteorological, hydrological, agricultural, socio-economic, etc.), its meaning will more or less diverge from that of min y because droughts arise through accumulation of water deficits in 'hydrologic reservoirs' of different scales and governed by different 'operating mechanisms' (Klemeš, 1987b). Thus the difference is not merely between maxima and minima of a given sample, but between their probabilistic meanings. Even if two samples of historic floods and droughts could be fitted with the same distribution functions, the same relative frequencies on the two curves would be associated with different probabilities. This is not merely a matter of a smaller or greater persistence, but also of phase shifts between the two processes, their propensity to quasi-periodicity, etc. Such differences would be even more pronounced in 'socio-economic' droughts which may be of the type of min z where the (normalized) $z(t) = \int y(\tau)d\tau$. The structure of the relevant processes may become quite complex (see Klemeš and Klemeš, 1988, for details) and their probabilistic interpretation even more so. It therefore may be dangerous to rely on drought 'probabilities' obtained by 'frequency analysis' as it is routinely practiced.

Conclusions

To increase our understanding of probabilities of extreme floods and droughts and to improve their estimates, more would be gained by the study of geophysical and anthropogenic processes than by the present preoccupation with the subtleties of distribution fitting which merely diverts talent from the task of shedding light on the dark areas to a futile search for the key to nature's secrets under the proverbial street lamp. As one mathematician aptly put it in a more general context: 'We may write down equations, and nature may – at some level – obey them, but nature is not obliged to restrict herself to those solutions that our overgrown monkey intellects can write down explicitly. And so mathematics must pay attention to what really happens, rather than assume that nature conspires to make human calculations easy,' (Steward, 1988).

REFERENCES

BAKER, V R., KOCHEL, R.C. and PATTON, P.C (Eds) (1988) *Flood geomorphology*. John Wiley & Sons, New York

BERAN, M.A. and RODIER, J.A. (1985) Hydrological aspects of drought. Studies and reports in hydrology, No. 39. Unesco/WMO. Paris/Geneva

CURRIE, R G. and O'BRIEN, D.P. (1988) Periodic 18.6-year and cyclic 10 to 11 year signals in northeastern United States precipitation data. *J Climatol* 8: 255–281

EAGLESON, P.S. (1972) Dynamics of flood frequency. *Water Resour Res* **8:** 878–897.

GEOPHYSICS STUDY COMMITTEE (1978) *Geophysical predictions*. Nat Acad Sci (USA) Washington DC

KLEMEŠ, V (1978) Physically based stochastic hydrologic analysis. *Adv Hydrosci* **11:** 285–355

KLEMEŠ, V. (1986) Dilettantism in hydrology: transition or destiny? *Water Resour Res* **22:** 177S–188S.

KLEMEŠ, V. (1987a) Hydrological relevance of flood frequency analysis. In: *Hydrological frequency modeling.* Singh, V.P. (Ed.) D. Reidel, Dordrecht, pp. 1–18

KLEMEŠ, V. (1987b) Drought prediction: a hydrological perspective. In: *Planning for drought* Wilhite, D.A. and Easterling, W.E. (Eds) Westview Press, Boulder, pp. 81–94

KLEMEŠ, V. and KLEMEŠ, I (1988) Cycles in finite samples and cumulative processes of higher orders. *Water Resour Res* **24:** 93–104.

SHIKLOMANOV, I.A. and MARKOVA, O.L. (1987) A world survey of problems of water availability and transfer (in Russian). Gidrometeoizdat, Leningrad

STEWART, I (1988) Nature and the monkey intellect (Review of 'Mathematics and the Unexpected' by I. Ekeland). *New Scientist* **119:** 59

TODOROVIC, P and ZELENHASIC, E. (1970) A stochastic model for flood analysis. *Water Resour Res* **6:** 1641–1648

TYSON, P.D. and DYER, T.G.J. (1978) The predicted above-normal rainfall of the seventies and the likelihood of droughts in the eighties in South Africa. *S Afr J Sci* **74:** 372–377