Extreme Hydrological Events: Precipitation, Floods and Droughts (Proceedings of the Yokohama Symposium, July 1993).
IAHS Publ. no. 213, 1993.

Probability of extreme hydrometeorological events - a different approach

V. KLEMEŠ
3460 Fulton Road, Victoria, British Columbia, Canada

Abstract A criticism of the standard approach to frequency analysis is presented and a different rationale is proposed based on treating the variable of interest as a compound event and synthesizing its distribution function by enumerating all physically plausible combinations of its major components for which data are available. Since only the order of magnitude of the probabilities of the extreme events is aimed for, the fine points of small sample theory, in which the standard approach is entangled, become spurious while more room is made available for explicit consideration of the physical processes involved. The approach is illustrated by an example using data on annual maxima of daily precipitation from the coastal region of British Columbia.

INTRODUCTION

In recent years, I have criticized the standard approach to the estimation of probabilities of extreme hydrological events on several occasions (Klemeš, 1986, 1987a,b, 1988, 1989). The main point of my criticism has been the apparent confusion about the nature of the problem which has led to the pursuit of high mathematical rigour which, at best, is of marginal importance in this context while neglecting the important matters such as the hydrological information content of the data and the practical decision-oriented (rather than theoretico-statistical) purpose of the analysis.

By standard approach I mean the standard frequency analysis consisting in the fitting of mathematical probability distribution models to ordered sequences of recorded events such as flood peak flows or storm precipitation (i.e., to their empirical cumulative distribution functions - to be abbreviated as cdfs in the following) and extrapolating the tails of these models to very low exceedance probabilities. The fundamental weakness of this approach has been clearly identified almost half a century ago by the world renowned late Australian statistician and probabilist, Professor P. A. P. Moran: "... the form of the distribution is not known and any distribution used must be guessed ... since the part of the distribution we are interested in is well away from the part where observations provide some information... [this difficulty] cannot be overcome by mathematical sleight of hand" (Moran, 1957).

While, to my knowledge, nobody has ever disagreed with or challenged the validity of this statement, the bulk of the work on probabilities of hydrometeorological events has concentrated precisely on this mathematical sleight of hand, in particular on the refinement of mathematical tools for the "extraction of information" of the kind that has never been present in the historic records in the first place.

This rather absurd situation is the result of a perfectly logical development. About a hundred years ago, engineers discovered that mathematical functions known as probability distributions fitted rather well the plots of sorted hydrological data which they called duration curves. They learned that there may be something more to this than pure coincidence because, as they were told, these data could be thought of as random samples from probability distributions. Seeing that these distributions are described by exact mathematical formulae in statistical literature, they got the clever idea that the unavailable hydrological information about the duration curve tails could be supplied by these formulae if they were fitted to the data with enough mathematical rigour. Once in motion, the obsession with distribution fitting developed its own momentum and the avalanche could not be stopped. It was kept moving by engineers, geographers and others whose lack of a deeper understanding of mathematics prevented them from seeing the fallacy in their reasoning, a fallacy immediately obvious to the mathematician.

To be fair, the effort was not all in vain. Unfortunately, though predictably, it contributed merely to the art of curve fitting and to the theory of small samples from known distributions rather than to its original aim, i.e. to a better information about the probabilities of extreme hydrological events and thus to better design and planning decisions.

My position has been that the available volume of statistical and probabilistic theory is greater by several orders of magnitude than the amount whose application is warranted by the meagre information contained in samples of, say, fifty values of annual precipitation or flow maxima which represent the typical objects of the standard approach to hydrological frequency analysis. If more light is to be shed on the probabilities of hydrological extremes, then it will have to come from more information on the physics of the phenomena involved, not from more mathematics.

THE STANDARD APPROACH

A brief summary of the main conceptual difficulties and practical limitations of the standard approach to frequency analysis (of a hydrological phenomenon at a given site) will be helpful as a basis for comparison with the approach being proposed.

First, the Achilles' heel of the standard approach is its dependence on a Platonic idea of a **very long observation record**; if one does not exist in reality, a possibility of its existence must, at least, be postulated as a hypothesis plausible in principle. There simply is no other way for arriving at probabilities by the standard approach than deriving them from the frequencies in long time series.

Second, the postulated long time series from which the actual observation record is a short segment must be assumed **stationary**, otherwise this record would not be its representative sample. Given the historical evidence of the changes in climate, land use, river morphology, tectonic structure, etc., this assumption contradicts the very reality whose description it is supposed to facilitate.

Third, it is not enough to have just one such long stationary series: an infinite ensemble of similar ones must be postulated. This is necessary in order that the historical record could be regarded as a sample from a **stochastic process** without which the quantum jump from a frequency within a historical time series to an instantaneous probability is virtually impossible. This assumption implies that, in addition to the actual history, an infinite number of different potential histories must be hypothesized which, however, despite their differences, must be statistically identical to the actual one and each must be equally likely. In other words, it implies that the physical development of the Earth has been a repeatable experiment in which, however, no physical development is allowed in order to preserve the stationarity.

Fourth, one more sleight of hand is necessary to pull the probability rabbit out of the hat of a historic record: the hypothesis of **ergodicity** which states that the time averages taken along one realization of a stochastic process (i.e., along one time series) are the same as those taken across their ensemble at any point in time. This is the master stroke which eventually transforms the frequencies of sequentially generated events into their instantaneous probabilities.

Notwithstanding the mental effort required to overcome the four preceding conceptual hurdles, it all is just an intellectual massage by which the measured values of discharge, rainfall depth, etc., are metamorphosed into appropriate random variates and prepared for the serious business of determining the mathematical model of their distribution. This model, like an angel, is supposed to carry us, in one effortless leap, from the earthly relative frequencies around 10^{-1}, across the void of several orders of magnitude, to the heavenly probabilities of 10^{-4} and beyond.

Here, it is worth pausing for a moment and noting how the pursuit of the "probability" leads, step by step, away from concepts amenable to empirical scrutiny and direct physical interpretation (observation record, time from one flood to the next, etc.), to notions gradually more and more speculative and divorced from the real world. By the time one climbs up from the floods and rainstorms of a historical record to the random variables of a stationary and ergodic stochastic process, one is in a completely different world governed by different rules and obeying a different logic which not only doesn't make sense in the context of the physical reality as we perceive and measure it, but often seems to make a deliberate mockery of it; it is a world where no amount of hydrological, climatological, geological, etc., knowledge provides any guidance as to what is probable, reasonable, ridiculous, important or irrelevant. And yet, by some strange psychological mechanism, we make ourselves believe that the numbers produced by manipulating these speculative constructs which often fly directly in the face of what we positively know as being true, represent credible probabilities of the true physical events like floods.

This is possible only because the concept of probability as applied to a hydrological (and any other geophysical, i. e. unrepeatable) process is not science but merely a **rationale**. It has been transposed from mathematics where, precisely defined in terms of axioms and internally consistent rules, and free of any nonmathematical interpretations, the probability calculus is an exact **science**. The difference between rationale and science is a deep one: the former is not amenable to either verification or falsification, while the latter is. In physical sciences, results are subject to empirical falsification, in mathematical sciences to verification or refutation by proofs.

There is nothing wrong about using a rationale (speculation) where science does not provide the desired guidance (a case in point is a religious or political rationale routinely used as a guide for human actions). However, a rationale must be clearly recognized for what it is and not be confused with what it is not, namely science. When this happens and a rationale is infused with "scientific rigour", unreconcilable inconsistencies are bound to surface, reducing its credibility and usefulness (recall, for example, the medieval scholastic disputes about the number of angels that can dance on the top of a pin and the not-so-medieval but equally scholastic disputes about the best distribution type to be used in flood frequency studies in the United States).

One specific example (Klemeš, 1989) will illustrate the point. In theory, the longer the observation record, the more reliable the estimates of the distribution type and its parameters. However, at the same time, the longer a hydrological series, the lower the credibility of the assumptions of stationarity and ergodicity on which the reliability of these estimates is predicated. Thus we have a paradoxical situation whereby **the theoretical underpinnings of the standard frequency analysis are best served when they are least satisfied in practice** and when the actual information - which we seek to maximize - is minimum! For it is easier to believe that the climate has been approximately stationary over the past 20 years and that floods similar to those that occurred during this period could occur this year with about the same likelihood as they did during the immediate past, than it is to believe the same for the past 20 000 years extending into the last glacial period when no water flew in the river and the very notions of "river" and "flood" cannot even be defined.

AN ALTERNATIVE APPROACH

Philosophy

When one realizes that the principal difficulties of the standard approach come from the commitment to the idea of a very long observation record which, from the practical point of view, makes the approach impractical to start with and, from the theoretical one, leads into a dead end from which it must lift itself up by its own ergodic bootstraps, so to speak, one starts wondering what harm would be done if this cumbersome ballast of irrelevant theory is dropped entirely.

We know that meteorological and hydrological phenomena, in whose extremes we are interested, are the results of many causative factors interacting in complex ways. And historical evidence suggests that the high extremes occur more due to unusual combinations of these factors than to unusual magnitudes of the factors themselves. As long as our understanding of the formative processes and even our ability to identify and measure them were practically nonexistent, it was convenient to let nature take care of their interactions and wait for the outcome which could be, if not understood, then at least observed easily since, through its devastating impacts, it often recorded itself quite graphically of its own accord.

We have outgrown this fatalistic approach to the study of nature in many areas already and the time may be ripe to reexamine it in hydrology, meteorology and related areas. We can try to explore explicitly what we already implicitly recognize: the components of the resultant phenomena and the ways the former interact to produce the latter. Even the little we already know about these matters suggests that some components are more important than others, some have narrower ranges than others, some even have limits which cannot be crossed and, most important, that the past outcomes of their interactions are by no means the only ones possible and compatible with the physical laws as we know them.

This leads naturally to what may be called a **combinatorial approach** to frequency analysis of hydrometeorological phenomena in which their probabilities are estimated from the frequencies of a large number of physically feasible combinations of the formative factors rather than from only those few that actually took place.

However, let it be clear from the outset that neither this approach transforms the rationale of assigning probabilities to geophysical phenomena into a science! It only makes the rationale more rational because it (1) makes possible the utilization of more information from a given period, (2) corresponds more directly to the physical processes involved, and (3) is amenable to a better scientific and empirical scrutiny since it can readily incorporate new knowledge and broader (rather than only longer) empirical evidence.

Theoretical framework

The given variable of interest, for example, the maximum annual value of 24 h point precipitation, is regarded as a compound event $z = f(x_1, x_2, \dots x_n)$ where x_i, $i = 1, 2, \dots n$, are its components and f is a function whose form, as well as that of the relationships among the individual components, may follow from physical theory or, if not, can be inferred from empirical evidence as a first approximation. The historical record of z is regarded as a sample from a compound distribution and records of the x_is as a sample from an n-variate joint distribution.

This puts the proposed approach into the framework of multivariate analysis, in particular, multivariate distributions. As an authoritative source (Kotz & Johnson, 1985) suggests, "...The fitting of frequency curves, as used in the univariate case, does not extend readily to higher dimensions..."; for many problems "...the methods are ad hoc in nature and are justified by logical and common sense arguments..." and "...[their] choice must be guided by experience, conjecture and empirical validation..."

These difficulties are in fact an advantage since they make room for considering the practical purpose of the whole exercise as well as the useful inputs from the relevant physical sciences - considerations which the high rigour available for the univariate case, and imported into the standard approach, tends to squeeze out.

Pragmatic considerations

The aim of the procedure being proposed is to arrive at a credible estimate of **approximate ranges** of the magnitudes of extreme events and of their probabilities, ranges compatible with available empirical evidence and current state of understanding of the physical processes involved. This, I believe, is as much as one can realistically hope to achieve under the circumstances.

The task has two distinct parts:

a) estimation of the magnitude of an event and

b) estimation of its probability.

The first part involves a construction (modelling, synthesis) of the event of interest under specified extreme conditions. The second part boils down to the estimation of the **probability of these conditions.**

It is important to realize that, for given conditions (combination of inputs), the estimated magnitude of the event is inaccurate and may be different depending on the model employed. But these differences do not change the natural probability of the event

- this is given only by the likelihood of a simultaneous occurrence of the conditions considered in synthesizing the event. The estimate of this probability is also inaccurate which follows from approximations of the actual relationships among the individual natural conditions and from the limited information that the observation record gives us about their natural frequencies. Here, the uncertainty is even greater than in the first part of the task and we cannot realistically aim beyond a reasonable estimate of the **order of magnitude of the probability of an extreme event**. Ideally, the uncertainties in both the magnitude and the probability of an event could themselves be expressed in terms of probability distributions but this is hopelessly beyond our reach in most cases.

The elaborate theoretical-probabilistic scaffolding supporting the standard approach has only one function: to justify extrapolation of a fitted distribution model far beyond a physically justifiable range and facilitate scholastic disputations among academics. It changes nothing on the simple fact that instantaneous probabilities are approximated by empirical relative frequencies whose extrapolation is an uncertain business.

In view of this, the scaffolding will be disposed with and replaced by the guiding principle of Occam's razor which translates into a few assumptions and implementation rules as follows:

a) The objective is the estimation of the order of magnitude of the probability of extreme events.

b) The immediately past period during which the observation record can be considered approximately stationary will be regarded as an "instant" on the geological time scale and only the data from such a period will be used for the estimation of present probabilities. Treatment of cases exhibiting a clear physical or statistical evidence of time dependence or historical uniqueness will be adjusted accordingly.

c) Only the empirical distribution functions of event components will be used for the construction of a synthetic distribution function of a compound event and no theoretical distribution models will be fitted.

d) When helpful in facilitating the objective and not contradicting common sense and scientific evidence, the empirical distribution functions may be interpolated and extrapolated so as to enhance their overall trends. If N is the length of the record used, extrapolation will be limited to frequencies not exceeding 1/2N.

e) The simplest form of relationships among component variables, compatible with empirical evidence and scientific understanding, will be employed in the construction of a synthetic distribution function of the compound event.

f) In order for a synthetic distribution function to be considered credible, it must be in a reasonable agreement with the empirical distribution function of the compound event within the range of its historical frequencies.

g) The concept of "return period" is superfluous in the present context since the probability is derived directly from the empirical relative frequency in the "present geological instant".

Example

It has been a commendable practice to design facilities whose failure due to flooding could put lives of many people at risk so as to withstand precipitation and floods much higher than the highest ones on record. In most industrialized countries, the so called Probable Maximum Precipitation and Probable Maximum Flood (PMP and PMF, respectively) have been adopted as design events in such cases (Klemeš, 1991). Notwithstanding their rather problematic labels and imperfections in methods for their construction, the PMP and PMF are rational attempts to synthesize very extreme events from their known formative components as documented in historical records, combined in the most unfavourable groupings which, though very unlikely, are consistent with our understanding of climate, meteorology and hydrology. It appears that the PMP and PMF methods, stripped of their elements of arbitrariness, can serve as a basis for the combinatorial approach to frequency analysis of extremes of hydrometeorological events - a use never intended by their originators.

The common practice of PMP computation is to use only a few of the largest observed storms, determine the formative components for each of them and compute the PMP as a combination of the maximum component values (this implies an assumption of mutual independence among the components within the set of the selected storms).

As an example, the core of this procedure has been applied to the construction of a synthetic distribution function of annual maxima of daily precipitation totals, P, at Coquitlam Lake in British Columbia on the basis of their following three components: precipitable water in the air column, W, and efficiencies (also known as intensities or P/M ratios) of the synoptic (also known as convergence) and orographic precipitation components, E_s and E_o, respectively (Klemeš, 1992). A 40 year record ending in 1990 and consistent with the assumption of stationarity was available for all the four variables.

Unlike in the PMP procedure,

a) all 40 annual maxima of daily precipitation and their three components have been considered;
b) all feasible combinations of the three components have been evaluated rather than only the most extreme one; and
c) the assumption of their mutual independence was modified based on meteorological considerations (Nikleva, 1990, 1991) and empirical evidence.

The empirical distribution functions (plotted in Gaussian probability coordinates) of the three components of the observed annual maxima of the daily totals appear in Fig. 1 (cdfs of the annual maxima of the three components themselves are shown as dots). The components represent a trivariate joint distribution represented by three bivariate scattergrams shown in Fig. 2. The assumption of independence is not contradicted by correlation analysis for the pairs (E_s, E_o) and (E_s, W). A meteorological argument suggesting that maximum orographic efficiencies are unlikely to occur together with maximum values of precipitable water is empirically supported by a correlation coefficient of -0.33 (significant at a 3.4% level) between E_o and W. This dependence has been accommodated by splitting the (E_o, W) sample into two uncorrelated sub-samples N_1 and N_2, separated by a dashed line in Fig. 2c (by moving the cutoff point down from the largest W value until the differences from zero of the correlation coefficients in both sub-samples become statistically insignificant).

A synthetic distribution of the compound event P was then constructed as follows: First, each of the 34 orographic efficiencies from N_1 was added to each of the

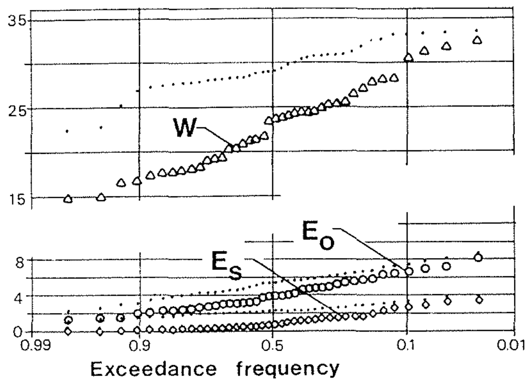

Fig. 1 Empirical distribution functions (from a 40 year record) of precipitable water, W (mm), and orographic and synoptic storm efficiencies, E_o and E_s (dimensionless), which contributed to the annual maxima of daily precipitation at Coquitlam Lake, British Columbia (cdfs of annual maxima of the three variables, max W, max E_o and max E_s, are shown by dots).

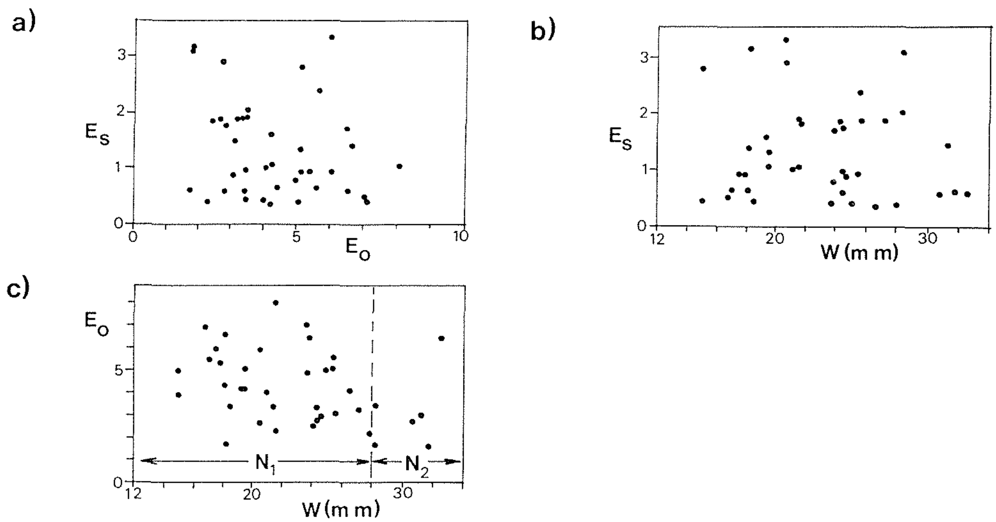

Fig. 2 Scattergrams representing the empirical trivariate joint distribution of W, E_o and E_s.

40 synoptic efficiencies (the two are independent) to form 1 360 total efficiencies; each of these was then multiplied by each of the 34 values of W from N_1 to produce the first sub-sample of 46 240 values of P. The same procedure was used to construct a sub-sample of P produced by the highest values of precipitable water: Each of the 6 orographic efficiencies from N_2 was added to each of the 40 synoptic efficiencies to form 240 total efficiencies; each of these was multiplied by each of the 6 values of W from N_2 to produce the second sub-sample of 1 440 values of P. These two sub-samples add up to a total of 47 680 synthetic values of P which specify its distribution function up to an exceedance frequency of 1/47 680. This cdf is shown as the lower curve in Fig. 3, the circles representing the empirical cdf of P.

As can be seen, the synthetic cdf agrees quite well with the empirical one. Since this agreement has been achieved without any attempt at a "best fit" and without employing any arbitrary mathematical distribution model, the synthetic cdf inspires a modest degree of credibility. It may be slightly on the low side because of the fact that some of the individual components could be higher than their historical values employed.

The upper curve in Fig. 3 may serve as an empirical upper confidence limit of the synthetic cdf. It represents a cdf of hypothetical compound events obtained from the annual maxima of the components themselves, by combining them on the basis of the same correlation structure as described above. In view of the fact that no combination of three rank-1 components occurred during the 40 year historical record, the exceedance probability of the upper curve should be less than 2% (Klemeš, 1992).

Based on the above considerations, the shaded area between the two synthetic cdfs in Fig. 3 can be regarded as a reasonable range for estimates of annual maxima of daily precipitation with exceedance probabilities down to the order of 10^{-5} for the Coquitlam Lake station.

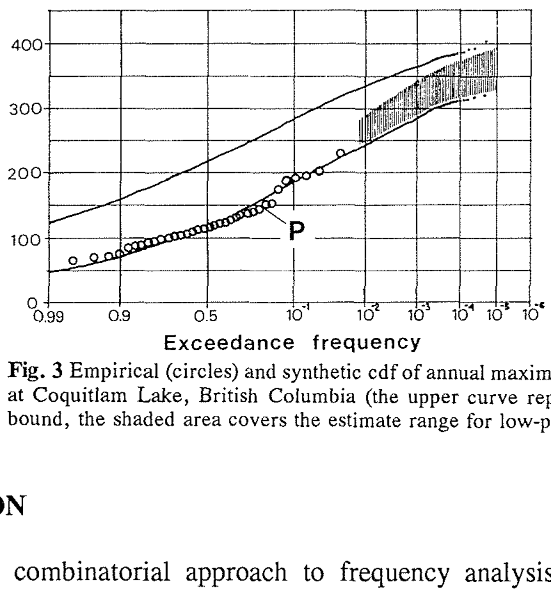

Fig. 3 Empirical (circles) and synthetic cdf of annual maxima of daily precipitation P (mm) at Coquitlam Lake, British Columbia (the upper curve represents an approximate upper bound, the shaded area covers the estimate range for low-probability events).

CONCLUSION

The proposed combinatorial approach to frequency analysis seems to offer a less arbitrary alternative to the estimation of probabilities of extreme events than does the standard approach. In spite, or rather because, of its simplicity, it also gives a better insight into the problem; for example, it allows for an explicit consideration of the influence on the distribution's upper tail of those components that have physically imposed upper limits (humidity, soil saturation, etc.) - influence that need not be obvious from the relatively few observations of the compound events themselves. By disclaiming the significance of the probability estimates beyond their orders of magnitude, this approach acknowledges the realistic limits of the exercise and renders meaningless the mathematical polish and superfluous rigour on which the standard approach thrives. Last but not least, it opens up new directions for applying the scientific method in hydrology and meteorology by formulating specific hypotheses that can be scrutinized by physical theory and tested by appropriate observations.

REFERENCES

Klemeš, V. (1986) Dilettantism in hydrology: transition or destiny? *Wat. Resour. Res.*, **22**(9), 177S - 188S.

Klemeš, V. (1987a) Empirical and causal models in hydrologic reliability analysis. In: *Engineering Reliability and Risk in Water Resources*, ed. L. Duckstein & E. J. Plate, Nijhoff, Dordrecht, 391-403.

Klemeš, V. (1987b) Hydrological and engineering relevance of flood frequency analysis. In: *Hydrologic Frequency Modeling*, ed. Singh, D. Reidel, Dordrecht, The Netherlands, 1-18.

Klemeš, V. (1988) A hydrological perspective. *J. Hydrol.*, **100**, 3-28.

Klemeš, V. (1989) The improbable probabilities of extreme floods and droughts. In: Hydrology and Disasters, ed. O. Starosolszky & O. M. Melder, James & James, London, 43-51.

Klemeš, V. (1991) International survey of the state-of-the-art of design flood estimation and approaches to dam safety assessment. BC Hydro, Vancouver, British Columbia (unpublished report).

Klemeš, V. (1992) Probability of PMP for Coquitlam Lake. BC Hydro, Vancouver, British Columbia (unpublished report).

Kotz, S. & Johnson, N. L. (Editors) (1985) *Encyclopedia of Statistical Sciences*, Vol. 6, 1-66, John Wiley, New York.

Moran, P. M. P. (1957) The Statistical Treatment of Flood Flows. *Trans. AGU*, **38**, 519-523.

Nikleva, S. (1990) Coquitlam Lake PMP: Moisture maximization and orographic precipitation. Pacific Meteorology, Inc., Richmond, British Columbia.

Nikleva, S. (1991) A synoptic meteorological approach to estimate extreme precipitation amounts. BC Hydro, Vancouver, British Columbia (unpublished report).