THEME-5

HYDROLOGIC EXTREMES

STATISTICAL FREQUENCY ANALYSIS OF HYDROLOGIC EXTREMES : A REVIEW

Fahim Ashkar & Bernard Bobee***
* University of Moncton, New Brunswick, Canada
** University of Quebec, Quebec, Canada

In hydrology, one often needs to estimate the frequency of occurrence of some extreme event such as a flood X_T corresponding to a return period of T years, where T is relatively high (the 100-year flood, for example). An adequate knowledge of such extreme events may be needed for the mangement of a hydro-electrical facility, the establishing of the flow capacity of an urban sewer network, the planning of a hydraulic work on a water stream, etc.

Statistical frequency analysis allows the estimation of the probability of occurrence of such extreme events. In the "annual maximum series" (AMS) approach, estimation of X_T is done on the basis of the maximum yearly values of the hydrological variable of interest (river discharge, snow depth, rainfall intensity, etc.). These "annual values" of the random variable X define a sample of size n, where n is the number of years of record of the hydrologic variable. The event X_T being estimated is such that Prob $(X \geq X_T)$ 1/T so that T is the average number of years separating two events $X \geq X_T$. Another approach for estimating X_T uses an average of m values of the hydrologic variable per year, which are those that exceed a certain truncation level. This is the "peaks over threshold" (POT) series approach.

Some comparisons have been made in the hydrological literature of methodologies for analyzing AMS and POT series. These have resulted in recommendations concerning "appropriate" procedures, such as statistical distribution and fitting method, to be used to fit the values of the hydrological variable of interest, and subsequently to estimate hydrological extremes. However, these recommendations still do not satisfy the needs of practitioners because they contain many points that need statistical justification, or that require further discussion. Although numerous distributions and estimation procedures have been proposed, the comparison of their performance has generally not been possible due to the difference in the criteria and the components of the simulation plans that have been used by the various investigators.

One of the most difficult problems in the statistical analysis of hydrologic extremes is that there is no theoretical procedure, or physical basis, for deciding which distributon is to be used, especially with small sample sizes commonly found in hydrology. What makes the problem more difficult is that interest lies in the tails of the distribution ("extremes"), in which area frequency estimates can differ substantially under alternate models, all of which may fit the data well in the central area.

We review from a historical perspective some of the commonly used procedures for frequency estimation of hydrologic extremes, with emphasis on the AMS and POT methodologies. We reflect on some of the classical questions in this area, and discuss some of the misleading concepts that have been used, such as "robustness to the choice of distribution", "robustness to outliers" and the notion of "separation of skewness". We try to identify some of the reasons for the current state of confusion, and recommend possible strategies for resolving some of the existing problems.

A COMPARISON OF FIVE DISTRIBUTIONS USED IN FLOOD FREQUENCY ANALYSIS IN BANGLADESH

Md. Abdul Karim & Jahir Uddin Chowdhury
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh

Log-Normal (LN), Gumbel (EV1), Generalized Extreme Value (GEV), Pearson Type 3 (P3) and Log-Pearson Type 3 (LP3) distributions are used in the frequency analysis of annual maximum (AM) flood flows in Bangladesh. An attempt has been made to determine which distribution represents the statistical characateristics of observed AM discharge series relatively better. It has been based upon statistical behaviour analysis and goodness-of-fit analysis using AM discharge data from 31 gauge stations for the period 1965 to 1989. LN, EV1, GEV, P3 and LP3 distributions have been fitted to the data by employing method of maximum likelihood, method of probability weighted moment (PWM), method of PWM, method of moment and method of moment in log-space respectively.

The study consists of five methods of comparison. The first is the comparison of the L-kurtosis versus L-coefficient of skewness (L-CS) diagramme obtained from observed flood samples with that obtained from similarly sized synthetic flood samples randomly generated from a distribution. In the second method the sampling distribution of L-CSs obtained from observed samples is compared with that of L-CSs from generated samples. The third comparison is between the sampling distribution of the observed largest peaks from normalized samples of equal length and that of the generated largest peaks. The fourth method uses the probability plot correlation coefficients between ordered observations and the corresponding fitted quantiles. The fifth is the comparison of average relative root mean square errors based upon deviations between observations and corresponding fitted quantiles.

The first three comparisons indicate that EV1 and P3 distributions are not suitable for AM discharge series in Bangladesh. Based upon five comparisons, a ranking has been made among the remaining distributions. The distributions with rank 1,2 and 3 are GEV, LP3 and LN respectively.

FUNCTIONAL RELATIONSHIPS AND ASYMPTOTIC PROPERTIES OF DISTRIBUTIONS OF INTEREST IN HYDROLOGIC FREQUENCY ANALYSIS

F. Ashkar, B.Bobee** & T.B.M.J.Ouarda**
** University of Moncton, Moncton, Canada*
*** University of Quebec, Quebec, Canada*

Statistical hydrologic frequency analysis is the most common of the two main approaches to the estimation of design flood events, the other approach being streamflow simulation from consideration of precipitation input. The procedure of hydrologic frequency analysis involves fitting a theoretical probability distribution to a series of flows, water levels or rainfall. The theoretical distribution must be adequately chosen to reflect the nature of the phenomenon and the characteristics of the data being modeled (positive or negative skewness, upper and/or lower bounds, etc.). A number of statistical distributions, with various numbers of parameters, have been proposed and used in a number of countries for the fitting of samples of hydrologic flood data (daily average flow, or maximum annual discharge, for example). These distributions are reviewed in this paper, and the advantages and disadvantages of each distribution are briefly discussed, along with their main characteristics. The functional relationships between these most commonly used distributions are highlighted. Proofs are derived for all the relationships that are established, and all necessary transformations of variables are identified. Special cases for each distribution are also discussed. This paper includes also a study and classification of the distributions according to their asymptotic properties (characteristics of the right tail of the distribution). One, two, three, and four parameter distributions are considered in this study. The final results are summarized in a diagram outlining the functional relationship between the variates of the different distributions, and a table detailing the probability density function (pdf) and the different forms and parameters of each distribution as well as the domains for its variables and parameters. Another table classifying the asymptotic properties of these distributions is also presented.

FREQUENCY ANALYSIS OF UPPER CAUVERY FLOOD DATA BY L-MOMENTS

A. Ramachandra Rao & Khaled Hamed
Purdue University, Indiana, USA

The objectives of the present study are to investigate the hydrological homogeneity of Upper Cauvery annual maximum flow data and to select a suitable distribution for the frequency analysis. The L-moments method is used in this analysis. The Upper Cauvery river basin is shown to be hydrologically heterogeneous. The 3 parameter log normal and the generalized extreme value distributions are recommended for the frequency analysis of data in this region.

FLOOD FORECASTING USING ARX MODELS

N.K. Goel & Vu Van Hong***

* University of Roorkee, Roorkee, India
** Institute of Hydraulic Design, Nhatrang, Vietnam

Since time immemorial floods have been causing loss of human lives, valuable property and crops in most parts of the world. The flood damages can be reduced drastically by taking various structural and nonstructural measures. The National Water Policy of India, 1987, gave special emphasis on nonstructural measures such as flood forecasting and warning and flood plain zoning for the minimisation of losses due to floods.

During last two decades, stochastic models have been increasingly employed for flood forecasting, as these models can take into consideration the error betwen the observed and forecasted values for their parameter estimation and updating.

In the present study, autoregressive models with exogeneous inputs (ARX) have been applied for flood forecasting for Kolar river at Satrana (India). For flood stage forecasting, exogeneous inputs of previous stages and mean areal rainfall were considered.

DETERMINATION OF SHELTER HEIGHT IN STORM SURGE FLOOD RISK AREA OF BANGLADESH COAST

Jahir Uddin Chowdhury
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh

The disaster due to floods in the coastal area caused by storm surges during tropical cyclones is a great concern for Bangladesh. The storm surge during the cyclone of April 29, 1991 devastated about 2000 sq. km. of area. The loss of human lives was approximately 138,000. A plan has been drawn up to construct shelters to provide refuge to the people during floods. Height of the stilt of the shelters has been designed such that storm surges can pass under. This paper briefly discusses how the required stilt height was determined for the High Risk Area (HRA).

Based upon the geography of the sea coast and the geometry of the continental shelf, the coastline of Bangladesh was divided into 5 parts. Maximum surge heights at every coast were predicted by using a formula which required the length of the continental shelf and the maximum wind speed as input data. A model of the distribution of annual maximum cyclonic wind speed provided wind speed data. Surge heights correspondig to various return periods were estimated by utilizing a conditional probability approach. Design surge heights at 50-year return period alongwith 90% confidence intervals in the 5 coasts are 3.7 ± 0.8, 5.8 ± 1.3, 6.5 ± 1.4 5.1 ± 1.1 and 4.3 ± 1.0 m.

The HRA extends from the coastline upto an inland limit where the depth of flooding by storm surge may reach 1 m. It covers an area of approximately 8100 sq.km. Human lives are at risk in this area. Using a probabilistic analysis of death statistics and wind speed data over a period of 30 years, the expected death of human beings without shelters has been estimated to be approximately 24,000. After providing allowance for the wind generated waves riding atop the storm surge wave, three heights for the stilts of the shelters have been selected. The stilt heights are 7.00 5.25 and 3.50 m depending upon the coastal location and the distance of a shelter from the sea coast.

FLOOD PREDICTION THROUGH PARTIAL DURATION SERIES

Arie Ben-Zvi
Israel Hydrological Service, Jerusalem, Israel

Flood magnitudes, associated with given recurrence intervals, are predicted through probability distribution functions which are fitted to data extracted from relevant records. For most of the applications, the maximal momentary discharge, or the maximal flood volume, which occured in each year of the records, are extracted. The other extraction strategy is of the relatively high peak discharges, or the voluminous events, regardless their years of occurrence. Series extracted through the latter strategy are called partial duration series.

Containing all the high magnitudes, partial duration series provide the entire information which is required for determination of recurrence intervals of rare events.

Yet, the absence of a fixed relationship between the sample size and length of records makes it necessary to transfer the serial frequency into temporal terms. This transfer might add a certain difficulty in the application of the series.

The selection and fit of proper distribution functions are of a major concern in hydrology. Candidate distributions are examined for complying with suitable assumptions and for fitting to data of case studies. Yet, no distribution function and no fitting procedure have gained a unanimous application. A number of governments have adopted regulations about these activities when they are required for formal studies of flood problems. The superiority of the adopted functions and of the fitting procedures are still controversial. The present article would contribute to the controversy about data extraction strategy, distribution selection and fitting procedures.

Various issues such as selection of the threshould discharges, choice of best fit distribution for the flood peak, volume and times of occurrence have been examined on data recorded at eight hydrometric stations located in the semi-arid and the arid regions of Israel. Their watershed areas range from 131 to 2830 km^2, records lengths from 20 to 41 years, total number of events from 66 to 271 and mean number of events range from 2.2 to 7.5 year. For this case study, the negative binomial distribution fits the distribution of number of events in a year better than the Poisson distribution does.

The rate of exceedance of peak discharges is determined from a plot of the conditional mean exceedance against the corresponding threshold discharges. The plots cover the complete duration series of peak discharges, while the selected threshold values coincide with the recorded peaks. The graphs indicate, in general, a slow rate for the lower threshold values an uadulating stable rate for middle values, and a fast rate for the high threshold values.

For each one of the stations, nested partial duration series has been defined with an association to all of the recorded peaks as threshold values. Goodness of fit of the Generalized Pareto distribution has been determined through the Anderson - Darling test. The test was limited to series containing at least 5 members which maintain the rate of exceedance concluded for the upper tail of the peak discharges. Ranges of good fit have been found for all of the stations. The best fits are obtained for series associated with middle valued thresholds. For most of the stations, the significance level of rejection of the best fit is higher than 0.25. The sizes of the series, for which the best fit are obtained, range from 14 to 70 events, and the mean number of events range from 0.36 to 1.92 year.

The records of those stations have been used also for an analysis of magnitude distribution of accumulated flood volumes beyond design discharges. A number of design discharges have been assigned for records of each station, and series of accumulated volumes have been computed. Each one of these series has been analyzed through the same procedures as the series of peak discharges had been. A good fit to these series has been found for the Generalized Pareto distribution.

The outcomes of the case studies indicate that the foregoing considerations provide us with good and practical procedures for quantitative predictions of flood magnitudes. The scientific merits of each one of these procedures are discussed in other works, whereas the purpose of the present one is to demonstrate their suitability and simpliciity for application. We believe that this purpose is achieved.

DROUGHT : A HYDROLOGIC EXTREME

P.K. Majumdar & J.S. Rawat

National Institute of Hydrology HRC, Belgaum, India

Drought, Definition and Difficulty are the three nouns, which are conjugate to each other. Definition of drought is difficult due to wide variation in the users and accordingly their area of interest, but the impact of drought remains same everywhere i.e. ultimate degradation of national economy.

Basically drought is a national hazard, striking the socio-economic structure of a country. African sahel, parts of India, parts of Brazil, Queensland in Australia, central America and Mexico has under gone the bad impacts of drought in recent past. Over the years India has felt

the impact of drought so frequently in terms of years and regions that it has become habitant of maintaining the doctor - patient relationship with the situation. Scientists, Technocrats and the concerned authorities are at least now able to console the people by dictating it as an hydrologic extreme.

Hydrology is an earth science, which creates passage for technology to develop and manage the water resources to optimal utility. It has been argued since several years that floods and drought are the two extremes of the hydrological phenomenon. "What makes drought as an hydrologic extreme?" Is there really any direct relation between drought and hydrology?

Drought is the result of deficiency in hydro-availability in one or other out of several forms, in such a magnitude that it creates problem for maintaining the accustomed human activities of the region. This ultimately generates a value for water from 'water is life' to 'accustomed level of water is life'. The word accustomed is totally a psychological creation. On the other hand flood has nothing to do with this human psychology in terms of impacts and its mitigation.

Likewise in the present paper try has been made to introduce a physical meaning of drought based on available state of art and experience. Its impact upon people and the economy has been reviewed with special emphasis to some of the recent past and present case studies. The value of prediction in the case of present so called hydrologic extreme has been assessed. Finally paper is concluded in terms of value based recommendations.

A STOCHASTIC CHARACTERIZATION OF PALMER DROUGHT SEVERITY INDEX

G.V. Loganathan, S.Mostaghimi, M.K.Tchaou & V.K.Lohani

Virginia Tech, Blacksburg, USA

The Palmer Drought Severity Index (PDSI) combines the key factors in defining a drought. A time varying Markov chain for monthly PDSI transitions and an annual homogeneous Markov chain to assess the long term drought proneness of a region are investigated. The Markov chain characterization becomes natural because of the class delineation adopted in the PDSI definition. An Auto Regressive Moving Average model (ARMA) with time varying coefficients is used to model the PDSI values themselves along with a multiplicative Auto Regressive Integrated Moving Average (ARIMA) model. The performances of the Markov Chain and Time Series approaches are compared utilizing the data for a climatic region in Virginia, U.S.A. A split length record comparison of observed and predicted PDSI values show promise in three month ahead forecasts for planning purposes. The long term PDSI class frequencies computed by Karl (1986, Journal of Climate and Applied Meteorology) are matched in an almost identical fashion by the steady state probabilities attained by the annual Markov chain. The theorem of total probability is utilized to assess the long term monthly PDSI class probabilities with the amonthly Markov chain which are also matched n an identical manner by the empirical frequencies computed from ninety six years of monthly records. It is believed that the results should be quite useful for planning purposes.

DROUGHT IN THE PENNAR BASIN USING STREAM FLOWS

B.R.K.Murthy & P.Venkatraman***

* K.S.R.M. College of Engineering, Cuddapah, India

** S.V.U. College of Engineering, Tirupati, India

Stream flow is one of the important hydrological parameters as it represents the run off from a basin or catchment and determines the quantity of water available in various surface water resources. The precipitation deficiency is ultimately reflected in the resulted stream flow. The low stream flows are indicative of drought situations. When the flows are not sufficient enough to meet the required demand of water, it is considered that drought has set in.

In order to manage the drought one needs to know the characteristics of the drought, its duration, areal extent and its severity. In this paper a methodology to identify these parameters from the available historic/forecasted flow data is proposed following a modified procedure of Herbert et. al (1966) originally suggested for analysis of drought using rainfall data. This procedure is applied

to 55 years of stream flow series (1928-29 to 1983-84) at Somasila and to 28 years at Upper Pennar Project (1958-59 to 1985-86). The identified droughts by the proposed methodology are concurrent with historically realized drought, thus proving the viability of the methodology in identification of drought condition.

ESTIMATION OF EXTREME FLOODS WITH PARTICULAR REFERENCE TO SOUTHERN AFRICA

D. Stephenson
University of Witwatersand, Wits, South Africa

The prediction of extreme floods associated with different recurrence intervals has always ben a difficult taks. Even the alternative of estimating maximum probable floods is fraught with uncertainty. The methodology is not yet standardized and the datums and factors to consider are not all incorporated in any single procedure for estimating these floods.

The estimates are even more difficult in areas with limited data. Developing areas in particular have sparse hydrological networks and the specturms of floods gauged in these areas provide limited, if not at times misleading, information. Rain data are often more reliable and more extensive and therefore many methods for estimating floods use rainfall data.

The paper compares alternative methods such as the rational method, unit hydrograph method, soil conservation method, and empirical methods. Owing to wide variations in topography and climate over large catchments it is demonstrated that many of these methods cannot be scaled up or extrapolated. It appears catchment modelling using hydraulic techniques is the most reliable where extrapolation is required.

Various empirical methods of estimating extreme floods are compared and it is shown that for large catchments available data illustrates the floods converge to common figures whichever data are used.

ROBUSTNESS OF FLOOD FREQUENCY DISTRIBUTIONS IN DEALING WITH THE OUTLIERS

K. Srinivasan, B.S. Thandaveswara *& M.G. Devamane***
* Indian Institute of Technology, Madras, India
** Walchand Engg. College, Sangali, India

An outlier in a set of data is defined as an observation of a subset of observations that appears to be inconsistent with the remainder of that set of data. These outliers distort the fitted distribution parameters to a considerable extent. These may be present at both the high and the low ends of the flood sepcturm and need to be detected and objectively modified to decrease the bias in design flood estimates.

This study examines the effect of outliers on parameter estimates and flood peak quantiles, for various flood frequency distributions listed in the literature. The roubustness of the different distributions in handling the outliers is discussed, based on examples of annual flood peak (extreme value) series from different hydrologic regimes. The discussion is confined only to at-site flood frequency analysis.

ROLE OF RAINFALL INPUT IN THE REAL TIME FLOOD FORECASTING

Mandakinee Majumdar & Krashnendra Singh
Indian Institute of Technology. Kharagpur, W.B. India

It is observed that the real time forecasts based on channel routing of flood hydrographs are generally more accurate compared to the ones based on rainfall runoff models. This is mainly due to the difficulties encountered in faithfully modelling the temporal and spatial variability of within the basin processes. Nevertheless, the rainfall as a model input is indispensable for the system where there is no suitable upstream station or there is substantial ungaged lateral inflow contribution into the reach.

The rainfall runoff process nonlinearity is predominantly because of the fashion in which rainfall contributes to various runoff components such as runoff, interflow, etc. whereas the hydraulic processes involved can adequately be represented by linear models. The effects of various rainfall transformations, namely, a low pass

filter and a threshold type in conjunction with linear transfer function type model are studied in the past. In the present paper a real time flood forecasting study on river Damodar in India has been reported employing transfer function type models.

The lead times of 3 hours and 24 hours are considered. The results include inflow forecasts.

ASSESSEMNT OF EXTRME FLOODS IN INDIAN TROPICAL REGION

P.R. Rao

Central Water Commission, New Delhi, India

Storm transposition and moisture maximisation techniques have serious limitations in their application to catchments of Indian Tropical region. Methods appropriate to Indian Tropics are to be devised to arrive at design storm depth and flood estimates and adopted after due process of testing and validation. Attempts to achieve standardisation are presented, with an illustrative case study, bringing out the issues that need to be debated and resolved.

DELINEATION OF GROUNDWATER RESOURCES IN DROUGHT PRONE AREA OF UPPER GODDUVANKA RIVER BASIN, CHITTOOR DISTRICT, ANDHRA PRADESH USING ELECTRICAL RESISTIVITY METHOD

T.V.K. Reddy, S.S. Reddy** & N. Janardhana Raju**

* S.V. University, Tirupati, India

** Engineering College, Pondicherry, India

The Upper Godduvanka river basin rises from the gamkonda hill ranges and flows in north east direction joining in Ponnai river. The Upper Godduvanka river basin has an areal extent of 478 Km^2. The basin lies in between north latitude 13° 8' - 13° 23' and east longitude 78° 43' 15" - 78° 59' 50". The study area includes the sub-basins of Bahu, Penrevulaeru and Bangarupalemeru.

With the ever increasing demand for water supply and the inadequate surface water especially in the drought prone areas, attention is turned towards groundwater resources. The present study area depends mostly on groundwater resources for the development of the region. The basin comprises of crystalline rocks of Archean age. The major rock type exposed in the study area is granite which is intruded by dolerite dykes. The river course is mostly concealed by alluvium and soil. The identification of structurally favourable zones of groundwater viz., lineaments, faults, fractures and joints have gained importance in the development of groundwater potential zones in hardrock. Of all the surface geophysical methods, the electrical resistivity method has been applied most widely for groundwater investigations. The search for groundwater in hard rock areas can be done by applying geophysical methods during different stages of exploration.

Electrical resistivity method has been used to delineate groundwater resources in drought prone area of Upper Godduvanka river basin, Chittoor district, Andhra Pradesh. 71 vertical electrical soundings were conducted using Wenner configuration. The data was interprerted by partial curve matching with the help of master curves. Contour map of apparent resistivity for the separation a = 30 m is prepared to delineate high and low resistivity zones. Based on the results, three vertical geoelectrical cross sections are prepared to know the distribution of groundwater resources along the respective profiles. Depth to basement contour map and finally probable groundwater potential zones are delineated for future exploitation. The study indicate that the groundwater resources are mainly confirmed to the sand and weathered granite zones of the study area.
