

Stochastic Hydrology Revisited

Robin T. Clarke

Instituto de Pesquisas Hidráulicas - Universidade Federal do Rio Grande do Sul - Caixa Postal 15029
91501-970 Porto Alegre, RS

Artigo convidado. Aprovado em agosto.

ABSTRACT

This paper reviews some of the stochastic methods used in applied hydrology over the past thirty years, the period during which the power and availability of computers grew rapidly, and methods of time-series modelling and simulation came into use which had previously been computationally prohibitive. Where stochastic methods are used to estimate the frequencies with which extreme hydrological events (floods, droughts) will occur in the future, these methods assume that hydrological processes are stationary, so that rainfall and runoff records from past years can be used to estimate how often extreme events will occur in the future. But where there are changes in land use or climate, hydrological processes also change, and the past may not be a good guide to the future. In South America, there have been extensive changes in land use during the last thirty years, and there is increasing evidence that climate is also changing. Standard hydrological procedures, such as estimating annual events with T-year return period, and regionalization of annual floods, then become inappropriate. The paper argues that under conditions of climate and land-use change, good assessment of the future frequency of extremes must await better knowledge of the physical processes that determine the behaviour of atmosphere and the oceans.

Key Words: *stochastic; hydrology; review.*

WHAT IS "STOCHASTIC HYDROLOGY"?

A first definition of stochastic hydrology is that it is a set of procedures for computing quantities of hydrological interest, through analysis of a hydrological record (of rainfall, discharge, lake levels, volumetric water content of soil at a given depth...), the values of which are regarded as observations of a *random variable*: that is, a variable subject to probabilistic laws. This definition can be extended in at least two ways. First, it can be extended in a multivariate sense, in which relationships are established between two or more random variables (such as rainfall and runoff from a drainage basin; or daily rainfall and depths of water stored in successive 10-cm layers of soil down to rooting depth). Second, it can be extended to include a spatial dimension, in which time-sequences of observations are recorded at two or more points that are spatially separated (for example sequences of daily rainfall at each of a number of rain-gauges within a drainage basin). Whether we are dealing with observations of a single hydrological variable, or of several, the procedures of stochastic hydrology result in the calculation of quantities that are subject to *uncertainty* because they are derived from calculations using random variables.

For the purposes of this paper, no distinction will be drawn between the terms "stochastic hydrology" and "statistical hydrology". The words "stochastic" and "statistical" are sometimes used to distinguish between statistical methods applied to hydrological analyses in which values are considered to be serially dependent, and serially independent, respectively. However this distinction is unnecessary; both terms apply to the analysis of hydrological data

in which values are regarded as sequences of observations of random variables.

UNCERTAINTY AND RISK

There is some confusion in the hydrological literature concerning the terms *uncertainty* and *risk*. Uncertainty is a characteristic of *events*; the event "rainfall in the next month will exceed 100 mm" (denoted by "A") is an uncertain event, the measure of its uncertainty being the probability that the event will occur. The probability can be interpreted either in *frequentist* or *subjective* terms; from the frequentist viewpoint, the probability of the event A can be estimated from the number of times in the past that rainfall in the month considered has exceeded 100 mm. Probabilities must be estimated subjectively where no records exist from which to estimate them; an example might be the probability of the event B: "nuclear fusion will provide cheaper energy than fossil fuels within the next five years". A number of texts (e. g., Lindley, 1985) shows how subjective probabilities can be quantified.

Risk, on the other hand, is a characteristic not of events, but of *decisions* that must be made in the face of uncertainty. As an example, consider a farmer who is thinking about whether to buy irrigation equipment. If the rainfall in the next growing season is less than X mm – an uncertain event, with its uncertainty measured by its (frequentist) probability p – purchase of the irrigation equipment will be justified; however if the rainfall in the growing season is greater than X mm – also an uncertain event, with probability (1-p) – he will not need to irrigate and the

equipment, if purchased, will stand idle and depreciate. On the other hand, if he does not purchase the equipment and the rainfall is less than X mm, his yields will be depressed and his income reduced. The farmer's decisions, the uncertain events that affect them, and the consequent losses L_{ij} , can be expressed as a 2×2 table:

	Uncertain event:	
	Rain < X	Rain $\geq X$
Probabilities:	p	$1 - p$
Decisions:		
d_1 : Purchase equipment	L_{11}	L_{12}
d_2 : Do not purchase equipment	L_{21}	L_{22}

The *risk* associated with the farmer's decision d_1 is the expected loss that would result from taking it, namely $p.L_{11} + (1-p).L_{12}$. Similarly the *risk* associated with the decision d_2 is $p.L_{21} + (1-p).L_{22}$. The farmer, as a logical person, will choose whichever of the two decisions d_1 and d_2 has smaller *risk*. Risk is therefore calculated from the uncertainties (probabilities) which in this simple example have the values p and $1 - p$, together with losses (or, more generally, *utilities*). If $L_{11} = 0$ so that the farmer loses nothing if he buys the equipment when rainfall is insufficient, then the *risk* of the decision d_1 has the simple form $(1-p).L_{12}$, or:

$$\text{risk} = \text{uncertainty} \times \text{loss}.$$

In general, there are more than two possible decisions, and many more than two uncertain events that may affect the outcome of each decision. But the principle remains that the *uncertainties* in the events are measured by their probabilities, whilst the *risk* associated with each possible decision is the sum of the products of losses with probabilities.

EXAMPLES OF STOCHASTIC METHODS USED IN HYDROLOGY

Over the past quarter century, stochastic methods have been widely used in many contexts. Two main types of application can be distinguished. The first is concerned with making statements about how frequently extreme events (flood discharges in rivers, droughts, rainfall of high intensity, ...) will occur in the future. The second main type of application is concerned with estimating the future values of a hydrological variable, such as discharge, given the values observed up to the present time. These two main areas of application are sometimes described as *prediction* and *forecasting* respectively. The following sections give some examples (by no means exhaustive) of applications in these two areas.

Examples of stochastic methods for prediction

- i. Predictions of how frequently an annual maximum daily mean discharge ("annual flood") of specified magnitude will occur in the future. An observed sequence of annual floods is assumed to be serially independent, and the method consists of selecting and fitting an appropriate probability distribution and calculating its quantiles. It is also possible to use non-parametric (kernel estimation) procedures to fit the probability distribution. The problem is commonly presented in the inverse sense: given the probability of occurrence (return period) of the annual flood, calculate its magnitude. Calculations of this kind form an essential part of the design of hydraulic structures, and urban planning in flood-prone areas. There is an enormous literature concerned with the selection of an appropriate probability distribution, how to estimate its parameters, and the properties of different estimators (e. g., Stedinger et al, 1992).
- ii. Predictions of how frequently a minimum mean discharge, averaged over a given time period, will occur in the future: for example, the annual minimum of the mean discharge in seven consecutive days. The procedure is essentially the same as in 3.1.1, although the assumption of serial independence of low flows in successive years may be more questionable because periods of drought may extend from one hydrological year into the next. The predictions are used in the planning of water resources, to ensure that water demands are met with a frequency of failure that is acceptably low.
- iii. Predictions of how frequently rainfall with a given intensity and duration will occur in future. For a given duration of rainfall (such as 10 minutes, or one hour...), annual maximum intensities for this duration are calculated from the rainfall record, an appropriate probability distribution is fitted, and its quantiles calculated. The procedure is repeated for several specified durations, and the results may be expressed in the form of intensity-duration-frequency (IDF) curves. The uncertainties in rainfall intensities predicted from IDF curves can be quantified, although problems arise because predicted extremes are dependent (Buishand, 1993). These curves play an important part in urban planning and drainage network design.
- iv. Predicting the frequency of occurrence of runs of observations with a specified characteristic (such as days without rain; monthly runoff less than half the monthly average;...). Here, some form of serial dependence between successive observations is assumed, an appropriate time-series model is selected and fitted, and the model is used to simulate many sequences of the hydrological variable. The ratio of the number r of such sequences in which the run occurs, divided by the total number N of sequences simulated, gives the desired estimate of the frequency

of occurrence. These predictions may be used, for example, to determine optimum planting dates for crops in semi-arid areas (Stern and Coe, 1984). There is an extensive literature on time-series modelling for hydrological application (e. g., Salas, 1992). Where sequences of daily rainfall have been modelled by stochastic methods, discrete-state Markov chains have commonly been used to model rainfall occurrence, with the transition probabilities allowed to vary from day to day throughout the year; rainfall depths registered on wet days have been modelled using exponential or gamma distributions (Coe and Stern, 1982; Stern and Coe, 1984). McCullagh and Nelder (1989) have shown how such models can be extended to look for trends in patterns of daily rainfall.

- v. Stochastic models of rainfall for time-intervals shorter than one day have been developed by Rodriguez-Iturbe and others (Rodriguez-Iturbe et al., 1984; 1987a, 1987b; Khaliq and Cunnane, 1996) and have been used to disaggregate long records of daily rainfall into rainfalls over shorter time periods, for use in rainfall intensity studies (Glasbey et al., 1995). The models have a complex structure which require assumptions about the time intervals between rainfall events, about the number of storm cells occurring within each rainfall event, about the time-intervals separating cell arrivals at the measuring site, about the time-duration of storm cells, and about their intensity. The models have been fitted to short periods of rainfall-recorder data at Pelotas, RS, and have been used to disaggregate longer records of daily rainfall at a site nearby; however the gain in information appeared to be quite small (Damé, 2001). More promising results were obtained under South African conditions (Smithers et al., 2002).

Examples of stochastic methods for forecasting

- i. Estimation of future mean monthly flow, given the sequence of mean monthly flow observed up to the present. An important feature of forecasts given by stochastic models is that measures of their uncertainty can also be obtained in the form of *confidence limits*. Models used to forecast monthly flow may be multiple linear regressions on harmonic terms as independent variables, or ARMA time-series models of the classical Box-Jenkins type which, for monthly flows Q_t are given by

$$\phi(B)\Phi(B^{12})\nabla^d \nabla^D Q_t = \theta(B) \Theta(B^{12}) a_t$$

where $\phi(\cdot)$, $\Phi(\cdot)$, $\theta(\cdot)$, $\Theta(\cdot)$ are polynomials of order p , P , q , Q in the backwards operators B and B^{12} (such that $BQ_t = Q_{t-1}$ and $B^{12} Q_t = Q_{t-12}$), ∇ is a differencing operator defined by $\nabla = 1 - B$, and a_t is a random "shock" with zero mean and variance σ_a^2 .

The powers d , D of ∇ are integers. However the forecasts may have large variances in cases where basin response to rainfall is very rapid, and where seasonal variation in monthly flow is not well-defined.

- ii. Estimation of future mean monthly flow, given sequences both of mean monthly flow and mean monthly rainfall averaged over the drainage area. Models of the Box-Jenkins transfer-function type can be fitted and used to forecast future monthly flows. With monthly rainfall and runoff denoted by P_t , Q_t in month t , the transfer model involving seasonality can be written

$$\delta(B)\Delta(B^{12})\nabla^d \nabla^D Q_t = \omega(B) \Omega(B^{12})B^{12} P_t$$

where $\delta(\cdot)$, $\Delta(\cdot)$, $\omega(\cdot)$ and $\Omega(\cdot)$ are polynomials in B and B^{12} . However the forecasts are likely to have large variances for lead-times greater than the response time of the drainage basin. Transfer function models can also be used with several input variables, such as rainfall and upstream flows, or rainfall in several sub-regions of the drainage basin.

- iii. Stochastic models of more complicated structure have been developed for systems, such as very large river basins, with long memory, and these long-memory models can be used for forecasting. They have been shown to give more accurate forecasts (i. e., forecasts with narrower confidence intervals) than ARMA models fitted to the same data. The models may contain causative variables, such as precipitation, as independent or explanatory variables (Beran, 1994). However, the use of such long-memory models for forecasting future annual runoff is likely to be limited because of the limited records of annual runoff available to fit them. At the monthly time-scale, deviations of monthly runoff from long-term monthly means show some evidence of long-memory behaviour, so that for large basins forecasts of monthly runoff by this approach may be a possibility. Evidence for long memory can be evaluated by various procedures (Beran, 1994) but one of the simplest is to break the data sequence into sub-series of equal length N , calculate the variance V amongst the sub-series means, and plot $\log(V)$ against $\log(N)$ for different values of N . If the data sequence has long memory, the slope of a line fitted to points in this plot will have slope greater (less negative) than -1 , the value corresponding to short-memory behaviour. Figures 1 and 2 show plots obtained from deviations about monthly mean flows on the Rio Paraná at Corrientes, and monthly mean water levels for the Alto Paraguai at Ladário. Both plots show evidence of long memory.

There is a physical argument which justifies the use of these models. Granger (1980) showed that when a number of short-memory processes are aggregated, such as lag-one auto-regressive models of the form

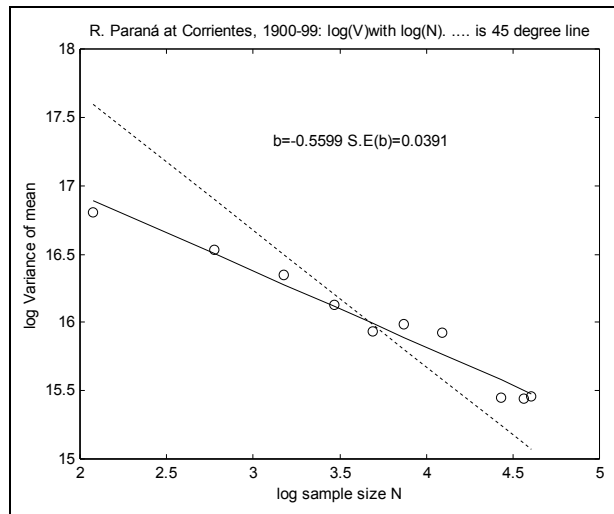


Figure 1. Plot of $\log(\text{variance})$ against $\log(\text{sample size})$, calculated from sequence of deviations from the long-term monthly means, for Rio Paraná at Corrientes, 1904-99. If the sequence had only short-term memory, the slope of the plot should be close to the 45° line with slope -1 (shown as broken line). The least squares line has slope -0.560 ± 0.039 .

$Q_t(j) = \alpha_j Q_{t-1}(j) + \varepsilon_t(j)$ for $j = 1, 2, \dots$ and with the α_j constant and such that $-1 < \alpha_j < 1$, the result is a long memory process. Now consider the response of a large drainage basin, regarded as an aggregation of a number of smaller sub-basins; after heavy rain when soil is saturated, each sub-basin will behave like a linear system, and the discrete form of a single linear reservoir with input P_t shows that the response Q_t satisfies an equation analogous to a lag-one auto-regression. Thus if sub-basins within a large drainage basin behave like single linear reservoirs when soil is saturated, the basin as a whole will show long-memory behaviour. A common model for long-memory behaviour is the auto-regressive model

$$\phi(B)\nabla^d Q_t = \theta(B)a_t$$

in which the power d is no longer an integer but satisfies $-1/2 < d < 1/2$. A model of this form is a *fractional ARIMA* (p, d, q) model.

- iv. Forecasting of future values of residuals obtained after fitting a lumped or distributed (“deterministic”) rainfall-runoff model. The applications 3.2.1 to 3.2.3 refer to stochastic flow-forecasting models which do not include descriptions of how vegetation and soil interact with rainfall to produce runoff and evaporation, but these descriptions are a fundamental feature of both lumped and distributed rainfall-runoff models. However when fitted to hydrological data, a sequence of residuals measuring lack of fit is produced which commonly contains a residue of information that is not easily extracted by modifying the rainfall-runoff model. Stochastic models have been fitted to

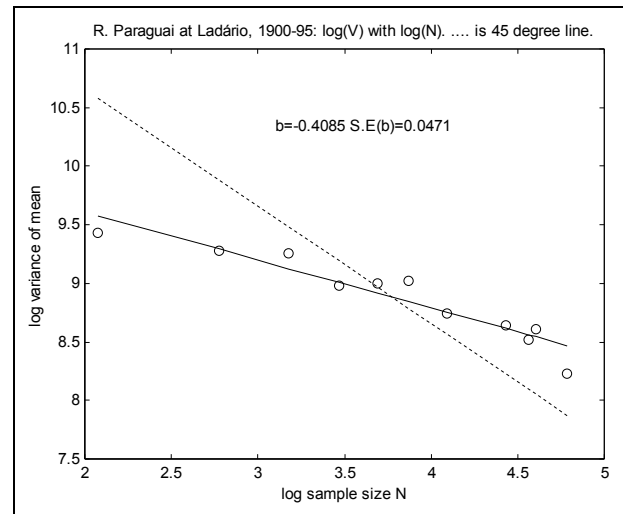


Figure 2. Plot of $\log(\text{variance})$ against $\log(\text{sample size})$, calculated from sequence of deviations from the long-term monthly means, for Alto Paraguai at Ladário, 1900-95. If the sequence had only short-term memory, the slope of the plot should be close to the 45° line with slope -1 (shown as broken line). The least squares line has slope -0.4085 ± 0.0471 .

the sequence of residuals to produce forecasts of residuals that serve to correct the forecasts of future flow obtained from the deterministic model. The ideal result from fitting both a deterministic model, followed by stochastic modelling of its residuals, would be a sequence of secondary residuals that is a completely random sequence (or *white noise*). There remains a difficulty however: whilst confidence intervals can be obtained for forecasts of future residuals given by the deterministic model, what the user really requires is confidence interval for forecast given by the deterministic model itself.

- v. Kriging is a widely used stochastic procedure for optimally interpolating a variable ($Y(s)$ say) that varies spatially over the region defined by co-ordinates s . *Cokriging* is an extension of the technique in which information on one or more additional variables $X(s)$, $Z(s)$, ... is used to improve interpolation of $Y(s)$. A particular case of co-kriging occurs where $Y(s)$ also depends upon time, $Y(s) = Y(s, t_i)$, and $X(s)$, $Z(s)$, ... are modified to become $Y(s, t_{i-1})$, $Y(s, t_{i-2})$, ... Thus co-kriging can be used to produce forecasts of the space-time process $Y(s, t_{i+1})$, $Y(s, t_{i+2})$, ... (Cressie, 1993).

THE CRITICAL ASSUMPTION OF STATIONARITY

If the modelling objective is to forecast future values of a hydrological variable, together with confidence limits for forecasts, then non-stationarity in records presents no particular difficulty. Stochastic models of the Box-Jenkins ARIMA type, in which ARMA models are fitted to a new

variable obtained by differencing the variable of interest an appropriate number of times, still provide forecasts. The same is true where a stochastic model is fitted to residuals calculated from a deterministic model (see the example 3.2.4 above).

However, where the modelling objective is to predict the frequencies of extreme events (or equivalently, the inverse problem, of predicting the magnitude of a hydrological variable that will occur with a specified return period, say 100 years), the difficulties are very much greater. Modelling for the purpose of predicting frequencies depends critically on the assumption that past hydrological behaviour will continue into the indefinite future. However, the assumption of stationarity may well be inappropriate (a) in river basins where land use is changing from forest to annual arable crops, or where urban growth is rapid; (b) where climate is changing. Figure 3 shows how annual floods are changing at two sites on the Rio Jacuí, Rio Grande do Sul; clearly it would be inappropriate to extrapolate past hydrological behaviour into the future at these sites. It is not clear how far the non-stationarity in annual flood record is a consequence of urbanisation, or of deforestation, or indeed of climate change; a careful statistical analysis of rainfall records would reveal whether it is also a contributing factor. The problem becomes marginally simpler if climate change can be ruled out; if (a) records are available that show how deforestation and impermeable urbanised areas have increased over time, and (b) upper limits to the spatial extent of deforestation and urbanisation can be reasonably specified, (c) it can be assumed that the flood regime will tend towards a steady state after these limits are reached, then methods explored by Clarke (2002a, 2002b, 2002c) for the analysis of non-stationary hydrological data with Extreme Value distributions can be used to predict frequencies of occurrence of extreme events after the steady state condition is reached.

The difficulty is much greater still where climate is changing. The report "Climate Change 2001: Impacts, Adaptation and Vulnerability" by the Intergovernmental Panel for Climate Change (IPCC) warns of climate change over the next century, envisaging "changes in the variability of climate, and changes in the frequency and intensity of some climate phenomena." Such forecasts, now being made with ever-increasing confidence, imply that the statistical stationarity necessary for many hydrologic analyses can no longer be safely assumed, and the spatial and temporal availability of water resources must be expected to change as and when regional climate changes. This is of crucial importance to a country like Brazil, in which more than 90% of its energy is provided by hydropower. Marked changes since 1970 of flow in rivers contributing to water generating hydropower from Itaipu have been reported by Müller et al. (1998); relative to mean flows before 1970, flows after 1970 in the drainage basins Paranaíba, Grande, Tietê, Paranapanema and Incremental of Itaipu were larger by 8, 18, 34, 45 and 44 per cent respectively. Smaller increases in rainfall were also detected in the same basins, of

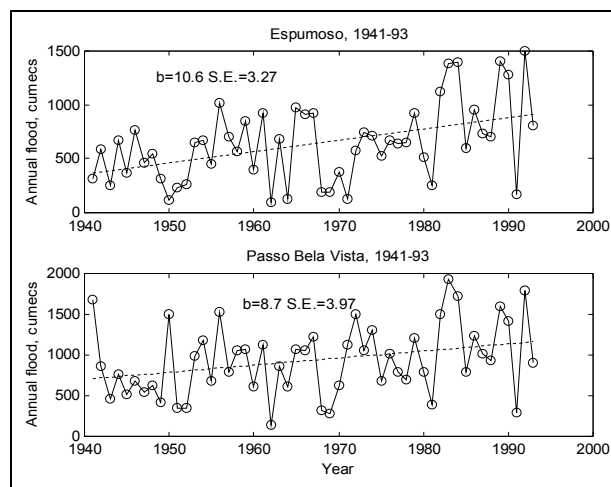


Figure 3. Records of annual floods at sites 85140000: Passo Bela Vista and 85080000 Espumoso, on the Rio Jacuí, Rio Grande do Sul.

magnitude 16, 17, 15, 16 and 8 per cent respectively. Thus annual flows in these rivers appear to show substantial non-stationarity, and estimation of annual flows with, say 100-year return period is likely to give misleading results if calculated from the entire record of data both before and after 1970. There is additional and growing evidence of non-stationarity in various South American rivers; Genta et al. (1998) found that flows in the Paraná, Paraguay, Uruguay and Negro (Uruguay) were related to an index representing sea surface temperature anomalies in the eastern equatorial Pacific Ocean. In their analysis of annual streamflow in rivers in southeastern and south-central regions of the sub-continent, Robertson and Mechoso (1998) found significant evidence of a non-linear trend, a near-decadal component, and interannual peaks associated with ENSO (el Niño – Southern Oscillation) time-scales.

It could be argued that as long as climate changes slowly and we limit our attention only to the prediction of extreme events that may occur over the next hundred years or so, existing methods based on the stationarity assumption (such as fitting Gumbel or other Extreme Value distributions to annual flood records) will not be greatly in error, particularly since the planning horizon may be no longer than 50 years. This may or may not be justified. There has been a great deal of research to estimate rises in mean global temperature (e. g., Zwiers, 2002) over the period up to 2100, and the likely increases in mean global temperatures over shorter periods, notably for the decade 2020-30, have now been reported and agreed independently by different authors (Stott and Kelleborough 2002; Knutti et al., 2002) who estimate increases in surface air temperature during this period of 0.3-1.3 K and 0.5-1.1 K respectively. Bearing in mind that these are global averages, regional values may well be larger; a greater degree of warming over land, and at higher latitudes, is expected, with possibly important consequences for regional rainfall and runoff.

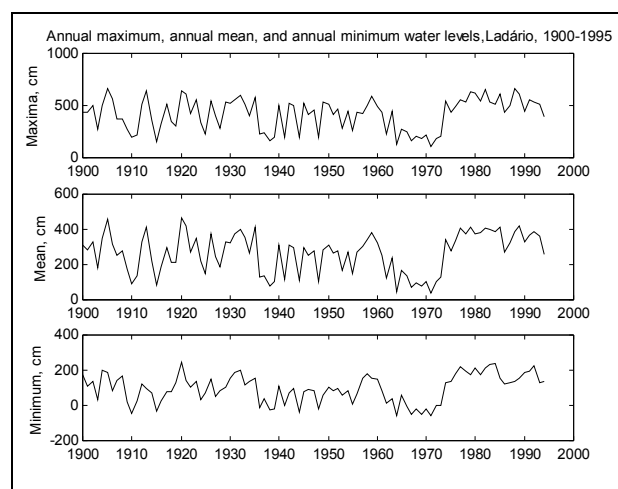


Figure 4. Annual maximum, mean and minimum water levels at Ladario, Alto Paraguai, 1900-1995.

Whilst the supposition of gradual climate change may or may not be valid, relatively abrupt changes in hydrological regime have also been documented for some Brazilian rivers. Figure 4 shows the record of annual maximum, mean, and minimum water levels at Ladario on the Alto Paraguai. The ten years or so between 1960-70 mark a period when water levels were very much lower than those before or after; indeed the whole structure of the time series appears to be different for these years as Tables 1 and 2 show. In the case of the Pantanal, the changes in hydrological regime had very severe consequences for the region's agriculture and society.

THE INCOMPATIBILITY OF THE CONCEPT OF "RETURN PERIOD" WITH NON-STATIONARITY OF HYDROLOGICAL RECORDS

Where changes in climate and/or land use are such that past hydrological behaviour of a river basin is no guide to future behaviour, the concept of return period becomes meaningless. When we say that the annual flood with return period 100 years is $X \text{ m}^3 \text{ s}^{-1}$ this means that over a very long sequence of years, an annual flood of magnitude $X \text{ m}^3 \text{ s}^{-1}$ or greater will occur, *on average*, with a frequency of once in a hundred years. But where non-stationarity exists, the idea of such "average" behaviour must be abandoned. What is the alternative?

The point of departure must be that, for planning purposes, the period of interest is the period covering the next 20 to 50 years, and in particular the probability of extremes (whether of flood flow, of low flow, or of mean flow) occurring during that particular period. To emphasise this point, our assessment of the uncertainty in an extreme event ceases to be in terms of its frequency of occurrence over the very long term, but in terms of the probability of it occurring once or more times, over the period starting at

Table 1. Mean maximum and mean minimum stages (cm), with standard errors of means, for periods 1900-60, 1961-70, and 1971-94: Ladario, Alto Paraguai.

	Period 1900-60	Period 1961-70	Period 1971-94
Annual maximum	414.1 ± 18.01	280.1 ± 41.04	480.9 ± 30.24
Annual minimum	89.67 ± 8.70	15.40 ± 21.65	143.8 ± 16.28
Difference	324.4 ± 20.0	264.7 ± 46.4	337.1 ± 34.3

the present and ending at the coming planning horizon. The solution – which will always have a large measure of uncertainty – then depends on what additional information can be supplied concerning the *physical causes* of the non-stationarity. If, for example, (a) the non-stationarity in an existing flow record can be explained by increases in deforestation or urban development, (b) no non-stationarity of rainfall regime is detected, (c) sufficient data exists on the annual rate of deforestation and urban increase, and (d) upper limits to the percentage loss of forest and to urban area can be postulated, then it may be possible to obtain a statistical relation between flow, area of forest, and urbanised area. If this relation tends towards an asymptotic form as deforestation and urbanisation increase towards their upper limits, it is possible in theory to use this asymptotic form – together with a measure of year-to-year variability – to get an approximate idea of the probability that flow exceed $X \text{ m}^3 \text{ s}^{-1}$ once, twice... N times, during the coming 20 (or 50...) years, for different values of X.

The more complex case occurs where non-stationarity in flow record is related not only to changing land-use, but also to changes in rainfall regime. At present, it seems that only very subjective methods are available (such as adopting the scenario of a sequence of extreme years of flow record, or selecting a particularly extreme sequence of rainfall years to be used as input to a rainfall-runoff mode to obtain a derived flow sequence). More soundly-based methods must await a better understanding of the ways in which atmosphere, oceans and land interact to bring about changes in rainfall regime and other climate variables; progress in this direction is being achieved by a number of authors (Robertson & Mechoso, 1998, 2002; Berri et al., 2002). Some papers in the literature have ventured to give forecasts of river flow for up to ten years ahead or even farther; however such forecasts are derived by extrapolating statistical fluctuations identified in flow records from past years, and time will tell whether the sophisticated statistical procedures now available justify forecasts so far into the future.

THE CHALLENGES FOR THE FUTURE

In the past, stochastic procedures for both predicting frequencies of extreme events and for forecasting flows in

Table 2. Serial correlation coefficients up to lag 4 of annual maximum, annual minimum and mean annual stages at Ladario, 1900-94 (a) for sequences as a whole; (b) for sequences up to 1960 only, (c) for period after 1971 inclusive.

Whole sequence (± 0.102)		Sequence to 1960 (± 0.128)		Sequence after 1971 (± 0.204)	
lag k	r_k	lag k	r_k	lag k	r_k
Annual maximum stages					
1	0.416	1	0.139	1	0.639
2	0.293	2	-0.074	2	0.419
3	0.307	3	0.037	3	0.294
4	0.164	4	-0.131	4	0.083
Annual minimum stages					
1	0.637	1	0.364	1	0.757
2	0.483	2	0.101	2	0.491
3	0.237	3	0.115	3	0.225
4	0.184	4	-0.060	4	-0.011
Annual mean stage					
1	0.523	1	0.225	1	0.706
2	0.374	2	-0.021	2	0.468
3	0.338	3	0.037	3	0.272
4	0.199	4	-0.110	4	0.096

the immediate future have used models which have little or no physical basis, serving only to “let the data speak for themselves”. However, particularly where climate is changing, statements about future hydrological regimes will require knowledge of physical processes governing the long-term behaviour of the global atmosphere, the oceans, and the interactions between the two. Stochastic methods will then abandon their role as generators of pseudo-random flow sequences in cases where serial correlation exists, and of fitting appropriate probability distributions where serial independence of data can be assumed. Instead, it will be necessary to combine statistical procedures for analysing the temporal and spatial structure of hydrological and climatological variables with mathematical descriptions of atmospheric and oceanic behaviour, for testing hypotheses about mathematical model structure, for estimating the uncertainty in estimates of parameters in the mathematical models, and for quantifying the uncertainty in forecasts obtained from them in probabilistic terms. An example of how stochastic methods are extending into climate-related studies is given by Sharma (2000) and Sharma et al. (2000) who develop a framework for rainfall probabilistic forecasting for Australian conditions, using hydro-climatic information such as ENSO and other phenomena. Increased spatial and temporal resolution of remote-sensed data, and the fact that increases in computer power allow continuous physical processes to be approximated using ever-finer discrete grids, will extend the need for efficient statistical analyses of large data sets. Thus it seems very probable that stochastic methods will no longer involve merely the computer generation of “synthetic”

flow sequences from models fitted to past flow records, but will develop a much closer relationship with physical reality through its associations with atmospheric and ocean sciences.

REFERENCES

- BERAN, J. (1994). Statistics for Long-Memory Processes. *Monographs on Statistics and Probability* No. 61: Chapman & Hall, Londres.
- BERRI, G. J., GHIETTO, M. A., GARCÍA, N. O. (2002). The Influence of ENSO in the Flows of the Upper Paraná River of South America over the Past 100 Years. *J. Hydrometeorology* 3 57-65.
- BUISHAND, T. A. (1993). Rainfall Depth-Duration-Frequency Curves: a Problem of Dependent Extremes. Capítulo 9 de *Statistics for the Environment* (Eds: V Barnett & K Feridan Turkman). John Wiley & Sons Ltd Inglaterra.
- CLARKE, R. T. (2002a). Estimating Time Trends in Gumbel-Distributed Data by Means of Generalized Linear Models. *Water Resour. Research* (aceito para publicação).
- CLARKE, R. T. (2002b). Estimating trends in data from the Weibull and a GEV distribution. *Water Resour. Research* (aceito para publicação).
- CLARKE, R. T. (2002c). Fitting and testing the significance of linear trends in Gumbel-distributed data. *Hydrological and Earth System Sciences*.
- COE, R., STERN, R. D. (1982). Fitting models to daily rainfall. *J. Appl. Meteorol.* 21, 1024-31.
- Cowpertwait P S P, O'Connell P E, Metcalfe A V, Mawdsley J A (1996). Stochastic point process modelling of rainfall. I: Single-site fitting and validation. *Jour. Hydrology* 175, 17-46.

- CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*. (2^o ed). John Wiley & Sons Inc.
- DAMÉ, R. de CASSIA FRAGA (2001). *Modelagem de dados pluviográficos e desagregação de chuva diária, na estimativa das relações intensidade-duração-frequência de ocorrência de chuva*. Tese de doutorado, IPH-UFRGS.
- GENTA, J. L., PEREZ-IRIBARREN, G., and MECHOSO, C. R. (1998). A Recent Increasing Trend in the streamflow of Rivers in Southeastern South America. *Journal of Climate* 11 2858-2862.
- GLASBEY, C. A., COOPER, G., McGECHAN, M. B. (1995). Disaggregation of daily rainfall by conditional simulation from a point-process model. *Jour. Hydrology* 165 1-9.
- GRANGER, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. *J. Econometrics* 14 227-238.
- KHALIQ, M. N., CUNNANE, C. (1996). Modelling point rainfall occurrences with the modified Bartlett-Lewis Rectangular Pulses Model. *Jour. Hydrology* 180 109-138.
- KNUTTI, R., STOCKER, T. F., JOOS, F., PLATTNER, G.-K. (2002). *Nature* 416 719-723.
- LINDLEY, D. V. (1985). *Making Decisions* (2^a edição). John Wiley & Sons Ltd Chichester Inglaterra.
- McCULLAGH, P., NELDER, J. A. (1989). Generalized Linear Models. *Monographs on Statistics and Applied Probability* No. 37, Chapman & Hall.
- MÜLLER, I. I., KRÜGER, C. M. e KAVISKI, E. (1998). Análise de estacionaridade de séries hidrológicas na bacia incremental de Itaipu. *RBRH: Revista Brasileira de Recursos Hídricos* 3(4) Out/Dez pp 51-71.
- ROBERTSON, A. W., MECHOSO, C. R. (1998). Interannual and Decadal Cycles in River Flows of Southeastern South America. *J. of Climate*, 11 2570-2581.
- ROBERTSON, A. W., MECHOSO, C. R. (2002). Interannual prediction of the Paraná river. *J. of Climate*, (a ser publicado).
- RODRIGUEZ-ITURBE, I., COX, D. R., ISHAM, V. (1987a). Some models for rainfall based on stochastic point processes. *Proc. R. Soc. London A* 410 269-288.
- RODRIGUEZ-ITURBE, I., COX, D. R., ISHAM, V. (1987b). A point process model for rainfall: further developments. *Proc. R. Soc. London A* 417 283-298.
- RODRIGUEZ-ITURBE, I., GUPTA, V. K., WAYMIRE, E. (1984). Scale considerations in the modeling of temporal rainfall. *Water Resour. Res.* 20(11) 1611-1619.
- SALAS, J. D. (1992). Analysis and Modelling of Hydrologic Time Series. Capítulo 19 de *Handbook of Hydrology* (Editor D R Maidment) McGraw Hill Inc., Nova Iorque pp 19.1 –19.72.
- SHARMA, A. (2000). Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1 – A strategy for system predictor identification. *J. Hydrology* 239 232-239.
- SHARMA, A., LUK, K. C., CORDERY, I., LALL, U. (2000). Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 2 – Predictor identification of quaterly rainfall using ocean-atmosphere information. *J. Hydrology* 240-248.
- SMITHERS, J. C., PEGRAM, G. G. S., SCHULTZE, R. E. (2002). Design rainfall estimation in South africa using Bartlett-Lewis rectangular pulse rainfall models. *J. Hydrology* 258 83-99.
- STEDINGER, J. R., VOGEL, R. M., FOUFOULA-GEORGIOU, E. (1992). Frequency Analysis of Extreme

- Events. Capítulo 18 de *Handbook of Hydrology* (Editor D R Maidment) McGraw Hill Inc., Nova Iorque pp 18.1 –18.66
- STERN, R. D., COE, R. (1984). A model fitting analysis of daily rainfall data. *J.R. Statist. Soc. A* 147, 1-34.
- STOTT, P. A., KELLEBOROUGH, J. A. (2002). *Nature* 416, 723-726.
- ZWIERS, F. W. (2002). Climate change: the 20-year forecast. *Nature* Vol 416 18 April

Reexaminando a Hidrologia Estocástica

RESUMO

Este artigo apresenta um resumo dos métodos estocásticos utilizados na hidrologia aplicada nos últimos trinta anos, período no qual a potência e disponibilidade de computadores cresceu rapidamente, tornando possível utilizar métodos de modelação e simulação de séries temporais cuja aplicação anterior era muito limitada. O uso de métodos estocásticos para a estimação das frequências de ocorrência de eventos extremos no futuro (enchentes, secas) é baseado na suposição de que os processos hidrológicos tenham caráter estacionário, no sentido de que os registros da precipitação e da vazão no passado possam ser empregados para estimar frequências futuras. Mas, na presença de mudanças do clima ou no uso do solo, os processos hidrológicos também mudam, e procedimentos padrões de hidrologia – por exemplo, a estimação de eventos anuais com período de retorno de T anos, e regionalização hidrológica – tornam-se inválidos. Na América do Sul ocorreram grandes mudanças no uso do solo nos últimos trinta anos, e estão se acumulando evidências de que o clima também está em transição. O artigo argumenta que, na presença de mudanças no clima ou uso do solo, a avaliação da frequência futura de eventos extremos requer um melhor conhecimento dos processos físicos que controlam o comportamento da atmosfera e dos oceanos.

Palavras-chave: hidrologia; estocástica; reexame.