Recent Developments in the Explanation and Prediction of Erosion and Sediment Yield (Proceedings of the Exeter Symposium, July 1982). IAHS Publ. no. 137.

Long term prediction of the extremes of river bed level fluctuations

JAN BRAŃSKI & JANUSZ ŻELAZIŃSKI
Institute of Meteorology and Water Management, 61 Podlesna Street, 01-673 Warsaw, Poland

ABSTRACT The extremes of river bed level fluctuations are controlled by erosion-accumulation processes which are closely related to flow variations. Because of problems in developing a physically-based model of such fluctuations it is useful to consider them as a random process and to develop probabilistic models for prediction. Data obtained for a number of flow gauging stations on the River Vistula have been used to develop long term records of bed level fluctuations. These fluctuations have been described using a simple probabilistic model with a linear trend component and a normal random deviation component. This model has been fitted successfully to records from five gauging stations on the Vistula and permits prediction of bed level variations up to the year 2020.

INTRODUCTION

The extremes of river bed level fluctuations are controlled by erosion-accumulation processes, which in turn are closely related to flow variations. Since the chain of causes determining river runoff in time are unknown and we do not know exactly the relationships between flow fluctuations and river bed level fluctuations, it is useful to consider the fluctuations of bed levels through time as a random process. Under such circumstances the only satisfactory solution to the prediction problem is its presentation in a probabilistic form. Since the prediction involves extreme values, it is appropriate to present the prediction in the form of a probability exceedance function. It must be recognized that river bottom and also water level fluctuations make the random process non-stationary. Research has shown that the natural erosion activity of rivers may be considerably accelerated by river regulation and channel dredging and quite frequently causes the regular lowering of the river bed and the water levels. Thus the mean value of the process is time dependent and the non-stationary character of the process results.

The problem of prediction may be therefore defined as the problem of function identification

$$G(x,t) = p(x \geqslant X)\big|_t \tag{1}$$

This is a function defining the relationship between the probability of exceedance of variable X (height of river bed) and time t.

For solution of the problem one should firstly define variable X and develop a means of empirical estimation; secondly, determine the analytical shape of function G(x,t); and thirdly, estimate the parameters of the function G(x,t) and verify the compatibility of the mathematical model with empirical data.

EMPIRICAL ESTIMATION OF RIVER BED COORDINATES

Establishment of the probability of exceedance function requires a long observational series of the investigated variable. The only available measured data providing information on fluctuations in river bed levels are the results of soundings carried out at gauging station cross sections during flow measurements. Knowledge of the water level on the day of measurement and of the height of the zero gauge datum allows the height of the bed at the points of individual soundings to be calculated. A useful reference value for this purpose is the mean bottom height at the flow equal to a medium drought (SNQL) in the summer half year period.

Estimation of bed levels defined in this way necessitates the following simple but laborious procedure; firstly, plotting of the channel cross section, secondly, plotting on the diagram the water level corresponding to flow SNQL, and thirdly, calculating the mean height of the wetted bottom cross section confined by level SNQL (Brański, 1978).

From the above calculations we can obtain a sequence of estimates of the bottom heights equal in number to the number of flow measurements taken. This sequence clearly does not include estimates of the annual extreme levels and this must affect the interpretation of the results.

MATHEMATICAL MODEL

Since their appearance in the work of Box & Jenkins (1970) ARIMA models have been widely applied as a general model of a non-stationary random process. However, the general procedure of identification, estimation and verification of an ARIMA model suggested by Box & Jenkins could not be applied in this case, mainly due to the variable timing of the process observation results. We would suggest, therefore, the following simple model, which may be considered as a special case of the ARIMA model:

$$x(t) = a + bt + \xi(t) \tag{2}$$

where: $x(t)$ = random variable dependent on time, t (stochastic process), a and b = parameters, $\xi(t)$ = stationary, normal random process with a mean value $\bar{\xi}(t) = 0$ and standard deviation σ_ξ.

Estimation of the three parameters a, b and σ_ξ provides the solution to the prediction problem, since it allows calculation for a given optional time of exceedance, of the prediction T and probability pf of the value $x(T,p)$, i.e. the value above which the probability of process implementation is equal to p:

$$x(T,p) = \bar{x}(t_0) + bt + \sigma_\xi\ tp \tag{3}$$

where $\bar{x}(t_0)$ = mean value of process at the selected moment t_0, which may be recorded as $\bar{x}(t_0) = a + bt_0$, t = quantile of random variable of a standard normal distribution N(0,1) and probability of exceedance p.

Equation (3) is acceptable only when the period of exceedance of the prediction T is longer than the ergodic period of the process, $\xi(t)$. The research has revealed that the above requirement is met if T > 5 years, which is practically always the case when the problem of long term prediction is taken into account.

The proposed model (2) is based on the following assumptions:

(a) the process x(t) is a normal process;

(b) the process x(t) is non-stationary;

(c) the mean value of the process $\bar{x}(t)$ is a linear function of time and a linear trend is developed in the process.

The above assumptions make it possible to represent the process in the form of two components: firstly, a linear function of time describing the relationship between the mean value of the process and time (a + bt) and secondly, a stationary normal random process with a zero mean value $\xi(t)$.

From the prediction viewpoint, the most important assumption concerns the occurrence of a linear trend in the investigated processes. Aside from the problem of verifying this assumption by statistical methods, the assumption is introduced for practical reasons. Two approaches to long term prediction of the mean value of a fluctuating process are possible. The first involves determining the factors controlling such variations, and using in the prediction the appropriate correlation or functional relationships, and the second involves pure extrapolation in time of the regularity found in the historical operation of the process. The first of these approaches seems to be eliminated, since for the five investigated profiles it was not possible to elucidate the factors controlling the observed trends. But analysis of the 60 years of information concerning mean bed level (and water level) fluctuation processes has permitted detection of persisting constant tendencies.

ESTIMATION OF PARAMETERS AND VERIFICATION OF THE MATHEMATICAL MODEL

In order to apply formula (3) it is essential to know the mean value of the process i.e. the height of the bed at the initial point of prediction development (x_{t0}), the coefficient of the linear trend (b), and the standard deviation of the process from the mean value (σ_ξ). Of key significance for prediction is coefficient b. It may be determined jointly with parameter a using equation (2) by the method of least squares applied to the observed series of channel bed levels. Due to a normal distribution of deviations from the mean, estimates obtained by means of least squares distinguish the minimum fluctuation variation.

The research undertaken has revealed that the coefficient of linear trend (b) is identical to that obtained from analysis of the water level series corresponding to SNQL. The trend line for the SNQL water levels is displaced parallel to the same line for the bed levels, and is separated from it by the mean value of the difference between the synchronous points representing both values. After the trend line equation is established we may calculate the last parameter of the model as the mean of the squares of the difference between the heights (ordinates) of the trend lines and the heights (ordinates) of the bed level or water level series.

When the parameters have been estimated one must verify the model by checking whether the basic assumptions are in agreement with the empirical data. The presented method cannot be applied if the following conditions are not satisfied.

(a) Process ξ i.e. the series related to the successive deviations from the trend line must represent a normal stationary random process.

(b) The ordinates of process ξ must be independent.

Condition (a) simultaneously reflects the existence of a linear trend, since if a nonlinear trend or cyclic component existed, the process ξ would be nonstationary.

For verification, the regularity of the variable ξ distribution was tested using the Kolmogorov test, autocorrelation over a distance in time exceeding five years was investigated using the Wald Wolfowitz test, and the similarity of variable ξ distributions in the first and second half of the investigation period which was also evaluated using the Wald Wolfowitz test may be considered as a substitute for a full study of the process stationarity. The theory and principles for these tests were taken from Fisz (1958).

Data from five water gauge profiles were analysed and were found to conform to the conditions outlined above at the 0.01 level of confidence. However, due to the restricted and simplified character of the tests and the small number of cases investigated the problem of model (2) adequacy should be considered as requiring further research.

IMPLEMENTATION OF THE MODEL

Prediction calculations were made using the following formula:

$$d(t)_p = (a + bt) + \overline{RD} + S_{RD}\, t_p \tag{4}$$

where (a + bt) = equation of the trend line, in which t denotes number of years from the beginning of the observational period, a = mean equalized height of water level SNQL in year t = 0, b = directional coefficient of the trend line equation, $\overline{RD}$ = mean of the differences between the height of the trend line and means of the bed levels within the SNQL channel defined from flow measurements taken in the observation period, or in other words the long term mean depth in the channel SNQL, S_{RD} = standard deviation; t_p = quantile of standard normal distribution for the probability of exceedance p.

Using formula (4) predictions of the bottom levels in five gauge profiles on the middle Vistula River were developed. In defining parameters for formula (4) the work was based on a 60 year series of water level observations and flow measurements (1919-1978).

Table 1 presents a compilation of the parameters of the prediction equations developed for the five profiles, and in Fig.1 the

TABLE 1 Parameters of the mathematical model for prediction of extreme variations in bed level

Gauge profile	*a* (*m a.m.s.l.*)	*b*	$\overline{RD}$ (*m*)	S_{RD} (*m*)
Zawichost	*131.44*	*-0.01003*	*-1.5000*	*0.6536*
Annopol	*128.37*	*+0.00015*	*-1.0663*	*0.3358*
Pulawy	*110.92*	*-0.00972*	*-1.2758*	*0.3760*
Deblin	*105.82*	*-0.00600*	*-1.9700*	*0.5767*
Warsawawa	*74.21*	*-0.03110*	*-1.5263*	*0.4338*

NOTE: (a) in the equation for trend $h(t) = a + bt$ *the value t has been calculated from the year beginning 1 January 1919. (b) Heights a.m.s.l. were defined in relation to the stipulated reference level.*

prediction of maximum and minimum level of the river bed in one of the profiles up to 2020 is presented graphically.

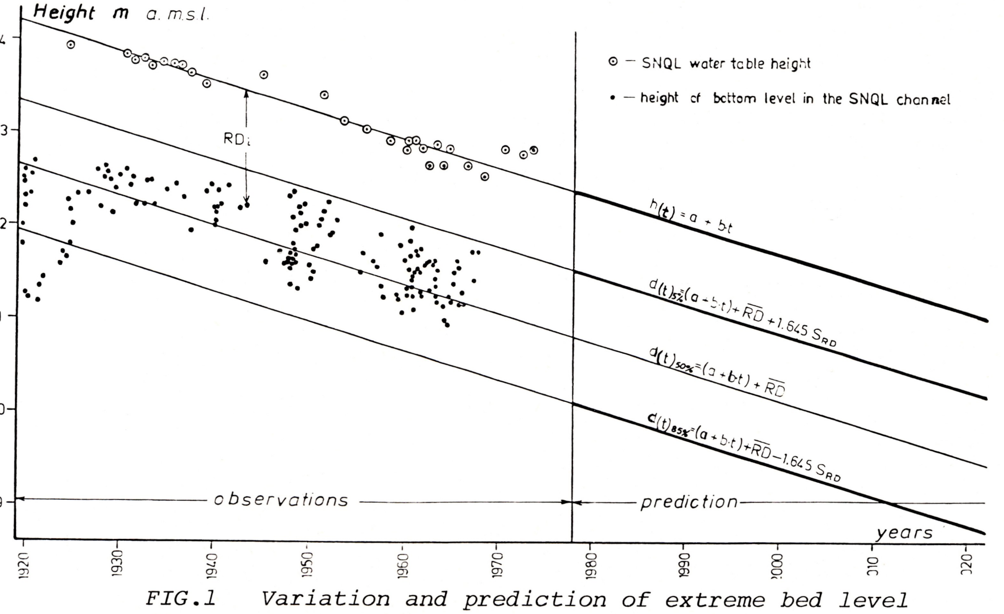

FIG.1 Variation and prediction of extreme bed level fluctuations in the River Vistula at Warsaw.

CONCLUSIONS

(a) Analysis of SNQL water level fluctuations and bed led fluctuations has shown that these are non-stationary processes and that AR (autoregression) or MA (moving average) models cannot be used for their description.

(b) An IMA (integrated moving average) model may be more useful since these models may be used for description of a wide range of non-stationary processes.

(c) The application of an ARIMA model is extremely laborious. It necessitates the use of a computer and the availability of the necessary programs needed to undertaken the successive stages of analysis.

(d) The method suggested here for the prediction of bed level fluctuations is based on the application of manual calculation techniques. This method gives satisfactory results for water gauge cross sections with a long (at least 40 year) observation period.

(e) It is also possible to use this method for other profiles, situated between water gauge profiles, but it requires detailed analysis of more comprehensive empirical data.

(f) The use of a computer would permit the development of a more complex model, which would describe in a more comprehensive way the fluctuation processes. Therefore, further work is required on the development and implementation of such a model.

REFERENCES

Box, G.E.P. & Jenkins, G.M. (1970) *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco, USA.

Brański, J. (1978) Określenie zmian położenia poziomu dna rzeki (Defining of river bottom height fluctuations). *Gosp. Wodna* no. 4.

Fisz, M. (1958) *Rachunek prawdopodobieństwa i statystyka matematyczna* (Probability calculus and mathematical statistics).