Stochastic models in hydrology

V. Yevjevich
Research Professor, School of Engineering and Applied Sciences, George Washington University, Washington, DC 20052, USA

Abstract: A stochastic approach to the analysis of hydrologic processes is defined along with a discussion of causes of tendency, periodicity and stochasticity in hydrologic series. Sources of temporal non-stationarity are described along with objectives and methods of analysis of processes and, in general, of information extraction from data. Transferred information as measured by correlation coefficients is compared with the transferable information as measured by entropy coefficients. Various multivariate approaches to hydrologic stochastic modeling are classified in light of complexities of spatial/temporal hydrologic processes. Alternatives of time series structural decomposition and modeling are compared. A special approach to modeling of space properties further contributes to approximate simulations of spatial/temporal processes over large areas. Several aspects of stochastic models in hydrology are concisely reviewed.

Key words: Stochastic hydrology, multivariate analysis, information extraction, information transfer, structure of time series, time series analysis, spatial characteristics, simulation of processes.

1 Determinism and stochasticity

Determinism and stochasticity constitute the two basic approaches to investigation of nature. Axioms of determinism are based on cause-effect relationships. Usually they are described by mathematical equations. Axioms of stochasticity lead to standpoints that relationships often cannot be expressed in simple or complex cause-effect mathematical forms. Instead, the "effect" variables are observed and their properties investigated by using methods of stochastic processes and mathematical statistics (Yevjevich 1974).

Figure 1 presents schematically three cases: (1) a pure deterministic relationship (left graph) as one extreme, (2) a pure stochastic case of cause-effect relationship (center graph) as the other extreme; and (3) transitions (right graph) between the two extremes. Ordinates of these graphs are partial effects on the resulting total "effect" variable by individual causal factors which are given on the abscissa. The left graph of Fig. 1 has a limited number of causal factors which jointly produce the full "effect" variable. It represents the classical case of a deterministic relationship. Errors in measurements act as additional factors, often as random noise. The center graph of Fig. 1 represents the case of effect being dependent on an infinite number of causes, each of them with an infinitesimally small partial effect. Here, no mathematical expression is feasible for a description of the cause-effect relationship. The effect is then conceived and investigated as a random variable.
The most current case of cause-effect relationship in the geophysical sciences is represented by the right graph of Fig. 1. The partial effects of a large number of causal factors are unequal, and assumed sorted in a descending order. A correlative association function takes into account the partial effects of a limited number of causal variables only. Partial effects of the large number of remaining causal factors are replaced by a random variable. It is composed of all neglected causal factors (which have not been identified, are not identifiable or are not economically observable), plus errors in measurement of variables included in the relationship.

A correlative equation is always composed of two parts: the mathematical regression equation between the $n + 1$ variables and the random term

$$Y = f(X_1, X_2, ..., X_n) + e$$  

(1)

It may be conceived as a deterministic-stochastic analysis of cause-effect relationship. This case and the pure stochastic case (Fig. 1, center) are assumed here to represent stochastic processes and models in hydrology.

2 Causal factors which produce hydrologic space-time processes

Causal factors which produce continuous or intermittent hydrologic space-time processes as effects are essentially provided by three large sources. They are: (1) astronomic motions of bodies in the solar system acting through variation in input of solar energy to places on the Earth or through tides; (2) thermal processes and movements of fluids in the atmosphere, oceans and surface and subsurface continental environments; and (3) anthropogenic influences.

The first source of causal factors (Fig. 2, left side) basically produce: the short-range periodic solar energy inputs to places on the Earth's surface (with day and year as cycles), the lunar/solar almost-periodic tides, and the long-range almost-periodic astronomic movements (the three Milankovich cycles, the Earth/Moon orbital eccentricity, the precession of equinoxes and the obliquity of Earth's axis). The additional astronomic causal factors may come from the periodic-stochastic process of sunspot activity, which has an average cycle of 11.3 years but is subject to some random time-dependent fluctuation.

Causal factors of the Earth's environments, which affect hydrologic space-time processes (Fig. 2, center), include many thermal and fluid motion random processes on the Earth. The atmosphere is most important in creating randomness because it is a light non-conservative fluid. It not only generates basic randomness, but also acts in transmitting and modifying periodicities and almost-periodicities of solar energy inputs and other causes of periodicity to other environments. In turn these environments further contribute stochasticity and modify randomness and periodicities produced by the atmosphere.