



# Black-Box Hydrological Models

Chong-Yu Xu, Lihua Xiong, and Vijay P. Singh

## Contents

1	Introduction .....	342
2	Antecedent Precipitation Index (API) Models .....	343
2.1	Calculations of API .....	344
2.2	Graphical API Models .....	345
2.3	Summary .....	345
3	Regression Models .....	347
3.1	Simple Linear Regression .....	347
3.2	Multiple Linear Regression Analysis .....	351
3.3	Summary .....	353
4	Time Series Models .....	353
4.1	Types of Hydrologic Time Series .....	354
4.2	Time Series Models for Stationary Data .....	354
4.3	Nonstationary Time Series Models .....	357
4.4	Summary .....	360
5	Artificial Neural Network (ANN) Models .....	361
5.1	Structure of ANN .....	361
5.2	Network Training .....	363
5.3	Summary .....	364
6	Fuzzy Logic Models .....	365
6.1	Basic Concepts of Fuzzy Systems .....	366
6.2	Operations with Fuzzy Sets .....	367

---

C.-Y. Xu

Department of Geosciences, University of Oslo, Oslo, Norway

e-mail: [chongyu.xu@geo.uio.no](mailto:chongyu.xu@geo.uio.no); [c.y.xu@geo.uio.no](mailto:c.y.xu@geo.uio.no)

L. Xiong

Department of Hydrology and Water Resources, Wuhan University, Wuhan, China

e-mail: [00011618@whu.edu.cn](mailto:00011618@whu.edu.cn)

V. P. Singh (✉)

Department of Biological and Agricultural Engineering and Zachry Department of Civil Engineering, Texas A and M University, College Station, TX, USA

e-mail: [vsingh@tamu.edu](mailto:vsingh@tamu.edu)

6.3	Types of Fuzzy Systems .....	370
6.4	Adaptive Neuro-Fuzzy Inference System (ANFIS) .....	372
6.5	Summary .....	372
7	Frequency Analysis Models .....	374
7.1	Graphical Method .....	374
7.2	Analytical Method – Frequency Factor Method .....	375
7.3	Data Sampling Methods .....	378
7.4	Outliers and Zeros .....	378
7.5	Regionalization .....	379
7.6	Summary .....	380
	References .....	381

## Abstract

This chapter discusses different types of black-box hydrological models that are based on input-output relationships rather than physical principles. They include antecedent precipitation index (API) models, regression models, time series models, artificial neural network (ANN) models, fuzzy logic models, and frequency analysis models. The purpose of this chapter is neither to provide a complete discussion of the theory of hydrological systems nor to offer a complete coverage of the studies published in the literature. Rather, the chapter is focused on presenting general theories and methods of different types of black-box models, basic model forms, and related applications in hydrology and water resources engineering.

## Keywords

Black-box · Gray-box · White-box models · Flood forecasting · Hydrology

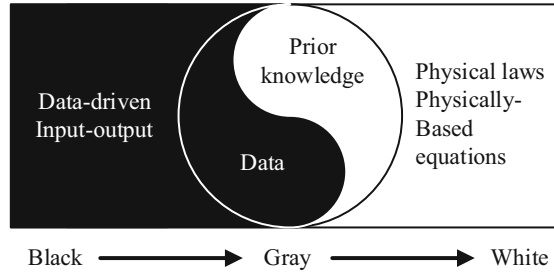
# 1 Introduction

Hydrological systems and their underlying processes are complicated. In real world, these systems are approximated by hydrological models, and a model is considered adequate if the difference between model prediction and measurement is small, satisfying a predetermined criterion.

Since the pioneering development of the rational method in the middle of the nineteenth century (Mulvaney 1850), development of hydrological models has gone through several stages, input-output models (black box), lumped conceptual models (grey box), and physically based distributed models (white box) (Fig. 1). As the simplest models, black-box hydrological models are based on input-output relations and do not describe the underlying hydrologic processes. In many cases, however, black-box models are adequate and served as a first step for modelers to conceptualize and simplify hydrologic systems. In fact, other types of hydrological models, i.e., gray-box and white-box models, more or less also owe their origin to black-box models.

Compared to other types of models, black-box models require the least input data. Precipitation and temperature are frequently used as input. Temperature is significant for hydrological modeling, especially, in climate regimes with snow.

**Fig. 1** Black-box, gray-box, and white-box models



Other hydrometeorological variables can also be possibly employed in black-box models. Some models can even be run only with output data, such as runoff data. Spatial characteristics of catchments are seldom considered but they are useful in the estimation of model parameters and can aid model interpretation and analysis.

Further, running black-box models is more easily achieved in terms of computer resources. Most black-box models do not demand high computational capability. Hence, a personal laptop or computer is enough for most cases. Many software or computer languages, such as Matlab or R, are friendly in programming black-box models.

This chapter describes different types of black-box hydrological models that are based on input-output relationships, including graphical antecedent precipitation index (API) models, regression models, time series models, artificial neural network (ANN) models, fuzzy logic models, and frequency analysis models. In the following sections, the main types of black-box models are discussed with the focus placed on the definition of a model, mathematical description and schematic representation of the elements and structures of basic model forms, and examples of applications in hydrology and water resources engineering. It should be noted that a complete discussion of the theory of hydrological systems or a complete coverage of the studies published in the literature is not a major concern of this chapter. For each type of black-box hydrological models, the basic forms are presented and a general discussion of their strengths and weaknesses is also provided.

## 2 Antecedent Precipitation Index (API) Models

The antecedent precipitation, i.e., the amount of precipitation that has occurred prior to a single storm event, plays an important role in calculating the runoff response to rainfall, especially in catchments where runoff generation is dominated by soil water or groundwater storage, i.e., runoff generation follows the principle of the “variable source area” theory (Hewlett and Hibbert 1967). The antecedent precipitation index (API) is generally defined as the weighted summation of daily precipitation amounts that is used as an index of soil moisture, and is expressed by the following equation (Kohler and Linsley 1951):

$$\text{API} = \sum_{t=-1}^{-i} P_t k^{-t} \quad (1)$$

where  $P_t$  is the amount of precipitation on the  $t$ th day prior to the occurrence of storm, and  $k$  is normally a constant.

The above exponential model is based on the assumption that the greater the time lapse between a rainfall event and a given day, the less influence the rain has on the soil moisture content of that day (Saxton et al. 1967). The value of  $i$  is usually taken as 5, 7, or 14 days (Viessman and Lewis 1996; Ali et al. 2010).

## 2.1 Calculations of API

The decrease of API is usually assumed to follow a logarithmic decay. Thus, during periods of no precipitation:

$$\text{API}_i = k \times \text{API}_{i-1} \quad (2)$$

For periods with precipitation:

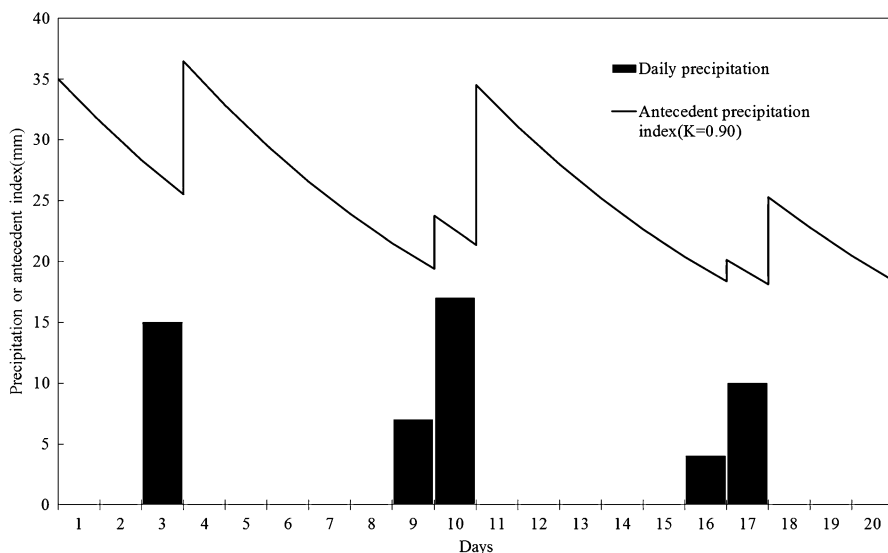
$$\text{API}_i = k \times (\text{API}_{i-1} + P_i) \quad (3)$$

This means that if any rain occurs, it should be added to the index (Fig. 2). In areas of snowfall, precipitation is applied to the model on the days when it melts rather than on the days when it falls. The value of  $k$  varies with basin physical characteristics and the meteorological condition, and a range of 0.85-0.90 over most of the eastern central parts of the United States was suggested (Viessman and Lewis 1996), which can be used as a reference for other regions.

To overcome the paradoxes that API often remains a subjectively determined and arbitrarily implemented parameter in rainfall-runoff modeling, Heggen (2001) proposed the use of a normalized antecedent precipitation index (NAPI) in place of API (Equation (4)). NAPI is defined as the ratio of the API on the day to the product of the average daily precipitation and the weighted sum of decay coefficients of the respective days before the storm.

$$\text{NAPI} = \frac{\sum_{t=-1}^i P_t k^{-t}}{\bar{P} \sum_{t=-1}^{-i} k^{-t}} \quad (4)$$

where  $\bar{P}$  is the average rainfall for antecedent days, and the other terms have been defined before. The soil moisture condition is assumed to be “dry” if  $\text{NAPI} < 0.33$ , the wet condition is defined as  $\text{NAPI} > 3$ , and the intermediate range 0.33 ~3 is the “fair” condition (Hong et al. 2007).



**Fig. 2** API relation (modified from NWSRFS User Manual Documentation at [http://www.nws.noaa.gov/ohd/hrl/nwsrfs/users\\_manual/hm/xrfsdocpdf.php](http://www.nws.noaa.gov/ohd/hrl/nwsrfs/users_manual/hm/xrfsdocpdf.php))

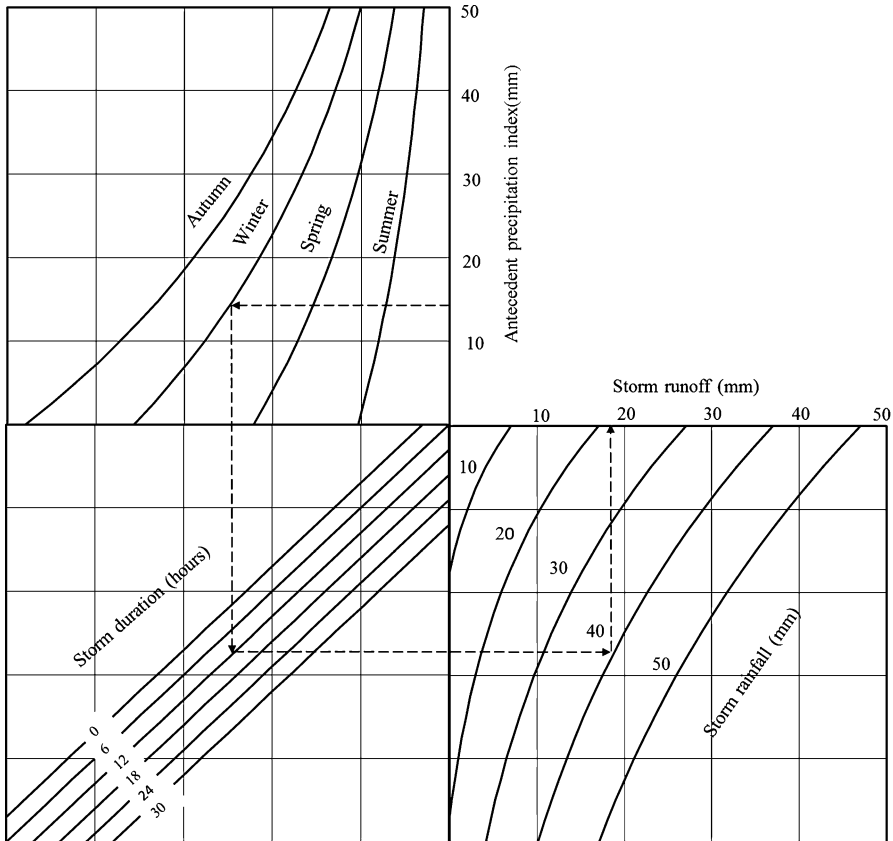
## 2.2 Graphical API Models

During the 1950–1960s, scientists were seeking techniques which would (1) simplify the relationships of rainfall and runoff, (2) require less time for the calculation and forecasting especially when the computers were not yet available, and (3) not require information of soil and surface characteristics, vegetation differences, and land use, which were usually not available.

Because of the importance of antecedent soil moisture condition to runoff generation, many indices have been used to estimate the moisture condition, such as (1) days since last rain, (2) discharge at the beginning of the storm, and (3) antecedent precipitation. API, a rough representation of the initial soil-moisture condition, generally provides better results among the three indices and can also be easily determined. Based on API, Kohler and Linsley (1951) developed a relationship between storm runoff and precipitation by a graphical method of coaxial relations (Fig. 3). The graphical method consists of 3 three-variable relations, relating storm runoff as the dependent variable to the antecedent precipitation (API), date (week number), rainfall amount, and rainfall duration as independent variables.

## 2.3 Summary

Antecedent soil moisture condition is important for watershed modeling that ultimately provides information on flood forecasting, water resources management, hydroelectric power generation, and irrigation management. Because the observed



**Fig. 3** Coaxial relation – antecedent precipitation index (modified from Sittner et al., 1969)

soil moisture data at a larger scale are usually not available, antecedent precipitation index, API, a rough representation of the initial soil-moisture condition, has been widely used in different hydrological modeling studies since the 1950s and the studies have generally shown that the use of API has potential to provide satisfactory results.

One typical example of using API as an important part of hydrological modeling is the model developed by the U.S. Soil Conservation Service (SCS) (1972) that incorporates initial losses or abstractions into a coefficient as a function of what is referred to as curve number,  $CN$ , which is a function of land use, soil type, hydrologic condition of basin, and the antecedent moisture condition (AMC) that generally is equivalent to the concept of antecedent precipitation index (API).

Other applications of API can be found in more complex rainfall-runoff models for storm runoff simulation (Saxton et al. 1967; Sittner et al. 1969; Fedora and Beschta 1989; Ali et al. 2010; Rajurkar et al. 2004; Dawson and Abrahart 2007), and in other models for landslide studies (Glade et al. 2000; Ma et al. 2014), global

runoff simulation (Hong et al. 2007), and the calculation of Forest Fire Danger Index (FFDI) (Liu et al. 2003), etc.

### 3 Regression Models

Regression (the term was first used by Pearson 1908) analysis is commonly used to describe quantitative relationships between a response variable and one or more explanatory variables. In hydrology, regression model is a useful tool for detecting relations between runoff and precipitation for the same watershed, between runoff (or precipitation) in different watersheds, between crop growth and precipitation, and so on.

An analytical problem to be solved by regression analysis involves (Riggs 1985): (1) selection of factors which are expected to influence the dependent variable; (2) describing these factors quantitatively; (3) selection of the regression model; (4) computing the regression equation, the standard error of estimate, and the significance of the regression coefficients; and (5) evaluation of results.

#### 3.1 Simple Linear Regression

Regression analysis is a statistical tool that utilizes the relation between two or more quantitative variables so that one variable (dependent variable) can be predicted from others (independent variables).

##### The simplest statistical model for simple linear regression

$$Y_i = a + bX_i + e_i \quad i = 1, \dots, n \quad (5)$$

$$\hat{Y}_i = a + bX_i \quad (6)$$

where:

$e_i$  is the error term or residual of the regression line,  $e_1, \dots, e_n$  are unobservable random variables, and are usually assumed as independent and normally distributed with mean zero and an unknown constant standard deviation,  $\sigma$ ;  $X_i$  and  $Y_i$  are the observed independent and dependent variables, respectively;  $\hat{Y}_i$  are the values estimated from the regression line;  $a$  and  $b$  are regression coefficients, where  $b$  is called the *slope* of the line and  $a$  is the *y-intercept*. The slope measures the amount  $Y$  increases/decreases when  $X$  increases/decreases by one unit. The *y-intercept* is the value of  $Y$  when  $X = 0$ .

### 3.1.1 Parameter Estimation

The goal is to find the equation of the straight line

$$\hat{Y}_i = a + bX_i$$

which would provide a “best” fit for the data points. That is to say, simple linear regression fits a straight line through the set of  $n$  points in such a way that makes the sum of squared *residuals* of the model as small as possible. In statistics, simple linear regression is the least squares estimator of a linear regression model. In other words,  $a$  (the  $y$ -intercept) and  $b$  (the slope) are solved by the following minimization problem:

$$\min_{a,b} \sum_{i=1}^n \hat{\epsilon}_i^2 = \sum_{i=1}^n (Y_i - a - bX_i)^2 \quad (7)$$

It can be shown that the values of  $a$  and  $b$  that minimize the objective function (7) are

$$b = \frac{\sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^n X_i^2 - n\bar{X}^2} = \frac{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} = \frac{\text{cov}(X, Y)}{\text{var}(X)} \quad (8)$$

$$a = \bar{Y} - b\bar{X} \quad (9)$$

### 3.1.2 Model Evaluation

**Coefficient of determination:** After fitting a line to the data-points, we want to know how much of the variability in the dependent variable ( $Y$ ) is explained by the regression. For this, the coefficient of determination ( $R^2$ ) is often used and is expressed as:

$$R^2 = \frac{\text{Explained variance}}{\text{Total variance}}$$

The variability in the dependent variable  $Y$  is quantified as a sum of squares:

- $\sum (Y_i - \bar{Y})^2$  = total sum of squares corrected for the mean = total variance
- $\sum (\hat{Y}_i - \bar{Y})^2$  = the squared deviations of the predicted values from the mean value, explained variance by the regression line
- $\sum (Y_i - \hat{Y}_i)^2$  = the sum of squares of deviation from the regression = unexplained variance



The most general definition of the coefficient of determination is (Haan, 2002),

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \tag{10}$$

$R^2$  ranges from 0 to 1, and it is normally expressed as a percentage.

**Standard error of estimate (S):** A measure of the variability of the regression line, i.e., the dispersion around the regression line is  $S$ . It tells how much variation there is in the dependent variable between the raw value and the expected value in the regression:

$$S = \frac{SSE}{n - 2} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - 2}} \tag{11}$$

where SSE is the residual sum of squares or the sum of squares due to error. This  $S$  allows us to generate the confidence interval on the regression line as well as on regression coefficients.

**Standard error (deviation) for parameters  $a$  and  $b$ :** In regression analysis, the standard errors of the least square estimators for  $a$  ( $S_a$ ) and  $b$  ( $S_b$ ) are estimated by

$$S_a = S \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \tag{12}$$

$$S_b = S \sqrt{\frac{1}{\sum_{i=1}^n (x_i - \bar{x})^2}} \tag{13}$$

### 3.1.3 Confidence Intervals

Confidence intervals are important for testing the statistical significance of the regression coefficient as well as the regression line.

**A 100(1- $\alpha$ )% confidence interval for  $a$  and  $b$  is:**

$$(\hat{a} - t_{\alpha/2}S_a, \hat{a} + t_{\alpha/2}S_a)$$

$$(\hat{b} - t_{\alpha/2}S_b, \hat{b} + t_{\alpha/2}S_b)$$

where  $\alpha$  is the significance level, and  $t_{\alpha/2}$  is the critical value of the  $t$ -distribution with degrees of freedom (d.f.) =  $n-2$ .

A 100(1- $\alpha$ )% confidence interval on the regression line is:

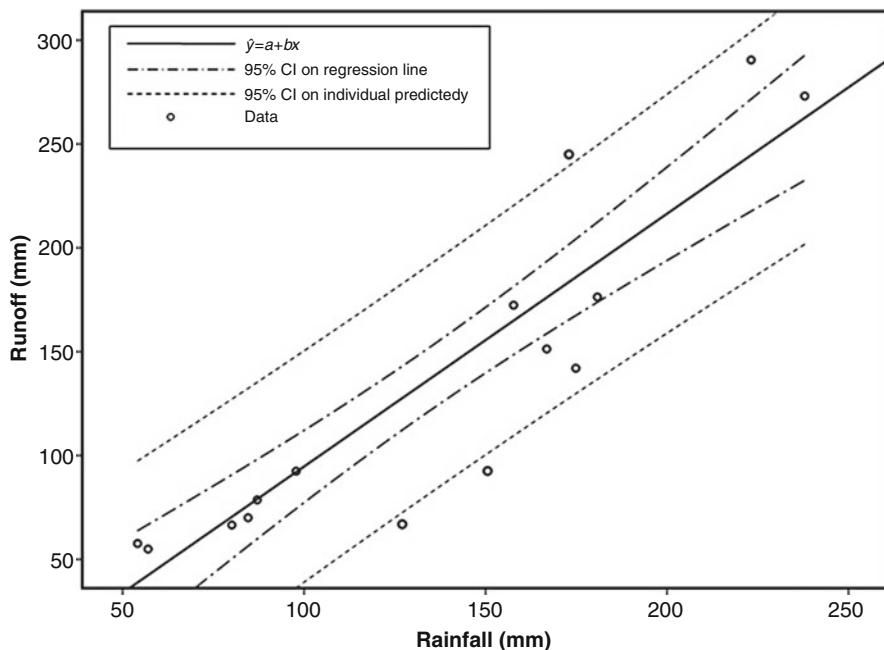
$$\left( \hat{y}_k \pm t_{\alpha/2} \sqrt{\frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \right)$$

where  $\hat{y}_k$  represents the predicted mean value of  $\hat{y}_k$ ; the confidence intervals are the narrowest at  $x_k = \bar{x}$  and widen as  $x_k$  deviates from  $\bar{x}$  as can be seen from Fig. 4.

A 100(1- $\alpha$ )% confidence interval on the individual points is:

$$\left( \hat{y}_k \pm t_{\alpha/2} \sqrt{1 + \frac{1}{n} + \frac{(x_k - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \right)$$

where the symbols are the same as above.



**Fig. 4** A typical plot of simple regression with the 95% confidence intervals and 95% prediction intervals (modified from Haan, 2002)

### 3.1.4 Significance of Coefficients

The statistical significance of parameters  $a$  and  $b$  equal to or larger/smaller than a given value (including zero) can be tested based on the  $t$ -distribution against a one- or two-sided alternative, depending on the nature of the relation that is anticipated. For example, if we want to test if  $a$  is significantly different from zero, the null hypothesis would be  $H_0: a = 0$ , and the test statistic is  $t = \frac{a-0}{s_a}$ .  $H_0$  will be rejected if  $|t| \geq t_{1-\alpha/2, n-2}$ .

Similarly, if we want to test if  $b$  is significantly different from zero, the null hypothesis would be  $H_0: b = 0$ , and the test statistic is  $t = \frac{b-0}{s_b}$ .  $H_0$  will be rejected if  $|t| \geq t_{1-\alpha/2, n-2}$ .

## 3.2 Multiple Linear Regression Analysis

The general purpose of multiple linear regression is to learn more about the relationship of a dependent or criterion variable to several independent or predictor variables.

**In general, a multiple regression procedure estimates a linear equation of the form:**

$$Y_i = a + b_1x_{i1} + b_2x_{i2} + \dots + b_px_{ip} + e_i \quad i = 1, \dots, n$$
 where  $x_{i1}, \dots, x_{ip}$  are the values of the input variables for the  $i$ th experimental run and  $Y_i$  is the corresponding response. The error terms  $e_i$  are usually assumed to be independent and normally distributed with mean zero and constant variance  $\sigma^2$ . The unknown parameters are  $a, b_i$ , and  $\sigma^2$  or  $\sigma$ .

### 3.2.1 Parameter Estimation

As in the simple linear regression, the first step in the multiple regression analysis is to obtain the least squares estimates of parameters  $a$  and  $b_i$  that minimize

$$\sum (y_i - a - x_{i1}b_1 - \dots - x_{ip}b_p)^2$$

In practice,  $n$  observations would be available on the dependent variable  $y$  and independent variables  $x_1$  to  $x_p$ . The  $p$  unknown parameters are estimated from the  $n$  observations. Thus,  $n$  must be equal to or greater than  $p$ , and in practice,  $n$  should be at least 3 or 4 times as large as  $p$ .

Most of least squares analyses of multiple linear regression models are carried out with the aid of a computer.

### 3.2.2 Evaluation of Multiple Regression Model

Similar to simple linear regression,  $R^2$ , the *coefficient of multiple determination* or *multiple coefficient of determination* is computed and used to evaluate how good the multiple regression is. The *multiple coefficient of determination* is defined as

$$R^2 = \frac{\text{Sum of squares due to regression}}{\text{Sum of squares corrected for the mean}} = \frac{\text{Regression SS}}{\text{Total SS}}$$

$$= \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (14)$$

where  $y_i$  are the observed values of the dependent variable,  $\bar{y}$  as its mean, and  $\hat{y}_i$  are the fitted values.

There are two reasons causing  $R^2$  to tend to overestimate the variance accounted for, compared to an estimate that would be obtained from the population: a large number of predictors and a small sample size. Therefore, the calculated  $R^2$  values usually need to be adjusted using Eq. (15). So, with a large sample and with few predictors, adjusted  $R^2$  should be very similar to the  $R^2$  value.

$$R^2_{\text{adjusted}} = 1 - (1 - R^2) \left( \frac{n - 1}{n - k - 1} \right) \quad (15)$$

where  $n$  is the number of data, and  $k$  is the number of independent variables used in the regression.

### 3.2.3 Stepwise Multiple Regression

In order to minimize the problem of over-parameterization, the independent variables need to be carefully chosen. The principle is that on one hand all relevant variables for theoretical or other reasons should be included, and on the other hand as few independent variables as possible (principle of parsimony) should be included, also because the more variables, the greater uncertainty, the larger type II error, and the fewer degrees of freedom.

In statistics, **stepwise regression** includes regression models in which the choice of predictive variables is carried out by an automatic procedure. The decision as to which variable should be included is determined by performing a sequence of F-tests or t-tests, but other techniques are also possible.

The main approaches are as follows:

**Forward selection:** In stage one, the independent variable best correlated with the dependent variable is included in the equation. In the second stage, the remaining independent variables with the highest partial correlation with the dependent variable, after the first independent variable is removed, is entered. The improvement of  $R^2$  value in each step is checked by the F-test (Haan 2002).

$$F_c = \frac{(1 - R_{k-1}^2) \cdot (n - k - 1)}{(1 - R_k^2) \cdot (n - k - 2)} \quad (16)$$

where  $R_{k-1}^2$  and  $R_k^2$  are the determination coefficient with  $k-1$  and  $k$ -independent variables, respectively;  $n$  is the number of data; and  $k$  is used as the number of independent variables. If  $F_c > F_{1 - \alpha, N - n - 1, N - n - 2}$ , we know that the addition of  $X_n$  is significant.

We continue until no variables "significantly" explain the residual variation. At each step, the increment of  $R^2$  is tested by the F-test.

**Backward elimination** which involves starting with all variables and eliminating independent variables one at a time until the elimination of one makes a significant difference in R-squared.

Multiple linear regression models are probably one of the most commonly used methods for hydrologic forecasting. However, various other types of statistically based regression models, e.g., nonlinear regression, principal component regression, partial least squares regression, are also used in hydrological studies (Eldaw et al. 2003; Sharda et al. 2008; Adamowski et al. 2012; Yasar et al. 2012).

### 3.3 Summary

Simple and multiple regression techniques are widely used in hydrology. Many applications of regression models can be found in the following categories: (1) hydrological forecasting, including streamflow forecasting (Yu and Liong 2007), rainfall prediction (Makarau and Jury 1997; Francis and Renwick 1998), water demand forecasting (Billings and Agthe 1998; Polebitski and Palmer 2009), etc.; (2) transferring information on hydrological behavior to ungauged catchments (Vogel and Kroll 1990; Pandey and Nguyen 1999); (3) infilling missing values of hydrologic variables, such as runoff, precipitation, temperature, soil moisture, etc. (Beauchamp 1989; Eischeid et al. 2000; Dumedah and Coulibaly 2011); and (4) regression models, which are also one of the four general categories of down-scaling methods. The relationships between large-scale and local-scale climatic fields were established by regression-based schemes (Hewitson and Crane 1996; Wilby et al. 1999; Charles et al. 2007).

In regression models, it should be noted that the regression equation does not imply a cause-and-effect relationship of the dependent variable to independent variables. Both may be influenced by some other factors that are not readily measured. However, there should be some physical tie between the variables if the results can be considered meaningful. Thus, there should be a physically plausible argument for selecting the explanatory variables to estimate the dependent variable.

Like many other statistical procedures, the regression analysis method described above is built under the assumption that the data are normally distributed, but the types of data used in hydrology commonly are not normally distributed, and some have no probability distribution at all. Hydrologists must select procedures most nearly suitable to the characteristics of data and must interpret the results accordingly.

---

## 4 Time Series Models

This section deals with basic time series models, which have become a major tool in hydrology in the era of information technology. In hydrology, time series models are usually used for building mathematical models to generate synthetic hydrologic

records, to forecast hydrologic events, to detect trends and shifts in hydrologic records, and to fill in missing data and extend records (Salas et al. 1980, Salas 1993; Haan 2002).

A time series is a series of observations of a variable in the course of time, where time is discretized to a series of time points or moments. A complete observed time series,  $y(t)$ , can be decomposed into a number of components as expressed by:

$$y(t) = y_1(t) + y_2(t) + y_3(t) + y_4(t) \quad (17)$$

where  $y_1(t)$  is the trend component,  $y_2(t)$  is the periodic component,  $y_3(t)$  is the catastrophic event, and  $y_4(t)$  is the random stochastic component. The first two terms are deterministic and can be identified and quantified fairly easily; the last two are stochastic and cannot easily be identified and quantified.

## 4.1 Types of Hydrologic Time Series

**Stationary and nonstationary series:** If the statistics of the sample (mean, variance, covariance, autocorrelation, etc.) do not change with time or length of the sample, then the time series is said to be stationary to the second-order moment, weakly stationary, or stationary in a broad sense. Otherwise, it is a nonstationary series, i.e., if a definite trend is discernible in the series or there is periodicity in a series, then the time series is nonstationary.

Generally, annual hydrologic time series are considered to be stationary, although this assumption may not be strictly correct due to large-scale climate changes and human activities. On the other hand, hydrologic time series defined at time scales smaller than a year, such as monthly and daily series, are typically nonstationary.

**White noise series:** For a stationary time series, if the process is purely random and stochastically independent, the time series is called a white noise series. It is the simplest example of a stochastic process. Such processes contain no memory by construction, that is, for every  $t$ , element  $X_t$  is independent of every other element in the process.

**Gaussian time series:** A Gaussian random process is a process (not necessarily stationary) of which all random variables are normally distributed, and of which all simultaneous distributions of random variables of the process are normal. When a Gaussian random process is weakly stationary, it is also strictly stationary, since the normal distribution is completely characterized by its first- and second-order moments.

## 4.2 Time Series Models for Stationary Data

A time series model is an empirical model for stochastically simulating and forecasting the behavior of uncertain hydrologic systems. Time series models include

stochastic models for a purely random time series with known distribution, for stationary time series, e.g., autoregressive (AR) models, moving average (MA) models, autoregressive moving average (ARMA) models, as well as for nonstationary time series, e.g., autoregressive integrated moving average (ARIMA) models, Thomas-Fiering model, etc.

#### 4.2.1 Time Series Models for Purely Random Series with Known Probability Distribution

Possibly, the simplest stochastic process to model is where the events can be assumed to occur at discrete times with the time between events constant, the events at any time are independent of the events at any other time, and the probability distribution of the event is known. Stochastic generation from a model of this type merely amounts to generating a sample of random observations from a univariate probability distribution.

Example: If  $X_t$  is a white noise series and normally distributed, i.e.,  $X_t \sim N(\mu_x, \sigma_x^2)$ , then the model can be

$$X_t = \mu_x + \sigma_x Z \quad (18)$$

where  $\mu_x$  and  $\sigma_x$  are the mean and standard deviation of the  $X_t$  series, and  $Z \sim N(0, 1)$  is a random series having standard normal distribution, which can be generated by Monte Carlo simulation.

#### 4.2.2 Autoregressive Models

The autoregressive models are used to model stationary time series when persistence (memory) is present. The general form of a  $p$ th-order autoregressive model, also called Markov type model, AR( $p$ ), is

$$\begin{aligned} y_t &= \mu + \beta_1(y_{t-1} - \mu) + \beta_2(y_{t-2} - \mu) + \cdots + \beta_p(y_{t-p} - \mu) + \varepsilon_t \\ &= \mu + \sum_{i=1}^p \beta_i(y_{t-i} - \mu) + \varepsilon_t \end{aligned} \quad (19)$$

where  $\mu$  is the mean value of the series,  $p$  is the order of AR model, written as AR( $p$ ),  $\beta_i$  are the regression coefficients, and  $\varepsilon_t$  are the noise or prediction error, normally assumed as  $N(0, \sigma_\varepsilon^2)$ . There are  $p+2$  parameters to be estimated:  $\beta_1, \beta_2, \dots, \beta_p, \mu$ , and  $\sigma_\varepsilon^2$ , the variance of residuals.

The most frequently encountered AR processes are of first or second order, and AR(0) process is white noise.

The equation for the first-order autoregressive model is:

$$y_t = \mu + \beta_1(y_{t-1} - \mu) + \varepsilon_t \quad (20)$$

Parameters  $\beta_1, \mu$ , and  $\sigma_\varepsilon$  of the model are estimated using the Yule-Walker equations. The relation between regression coefficient  $\beta$  and autocorrelation coefficient  $\rho$  is written as:

$$\rho_k = \sum_{j=1}^P \beta_j \rho_{k-j} \quad (21)$$

where  $\rho_k$  is the autocorrelation coefficient. The parameters are estimated as:

$$\hat{\beta}_1 = \rho_1; \hat{\sigma}_\varepsilon^2 = \sigma_y^2(1 - \beta_1^2); \hat{\mu} = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (22)$$

**The procedure for generating a series of values for  $y_t$  using AR(1) model is:**

- Estimate  $\mu_y$ ,  $\sigma_y$ , and  $\beta_1$  by  $\bar{y} = \mu_y$ ,  $s_y = \sigma_y$ , and  $r_1 = \beta_1$ , respectively, and  $\sigma_\varepsilon^2 = \sigma_y^2(1 - \beta_1^2)$
- Select a  $z_t$  at random from an  $N(0, 1)$  distribution
- Select an initial value for  $y_{t-1}$
- Calculate  $y_t$  based on  $\bar{y}$ ,  $s_y$ , and  $\beta_1$ , and  $y_{t-1}$  by

$$y_t = \mu_y + \beta_1(y_{t-1} - \mu_y) + z_t \sigma_y \sqrt{(1 - \beta_1^2)} \quad (23)$$

- Delete the first 50 values to get rid of the influence of the initial values

### 4.2.3 Moving-Average Models

The moving-average model of order  $q$  process, denoted by  $MA(q)$ , is formulated as follows:

$$y_t = \mu_y + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (24)$$

where  $\varepsilon_t$  is a white noise with  $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ ;  $\theta_i$  are parameters of order  $q$ , i.e., parameter  $\theta_k = 0$  for  $k > q$ .

The above equation gives the definition of the MA model: For a white noise or purely random series,  $\varepsilon_t$  is assumed to be normally distributed with zero mean and constant standard deviation. That is, a moving average model is conceptually a linear regression of the current value of the series against the previous (unobserved) white noise error terms or random shocks.

### 4.2.4 ARMA Models

Autoregressive moving-average (ARMA) models, sometimes called Box-Jenkins models (Box and Jenkins 1976), consist of two parts, an autoregressive (AR) part and a moving-average (MA) part. The model is usually then referred to as the  $ARMA(p, q)$  model, where  $p$  is the order of the autoregressive part and  $q$  is the order of the moving-average part.

In this case,  $x_t$  is a mixed process where the output is a function of past outputs and current/past inputs



$$x_t = c + \sum_{i=1}^p x_{t-i}\beta_i + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (25)$$

All notations have the same meaning as before. The error terms  $\varepsilon_t$  are generally assumed to be independent identically distributed random variables (i.i.d.) sampled from a normal distribution with zero mean:  $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ , where  $\sigma_\varepsilon^2$  is the variance of the error.

There are  $p+q+2$  parameters ( $\beta_i, i = 1, \dots, p; \theta_i, i = 1, \dots, q; c; \sigma_\varepsilon$ ). Some formulations transform the series by subtracting the mean of the series from each data point. This yields a series with a mean of zero. Whether one needs to do this or not is dependent on the software one uses to estimate the model.

In practice, the ARMA(1,1) model is often used:

$$x_t = \beta_1 x_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \quad (26)$$

Parameters  $\beta_1, \theta_1$ , and  $\sigma_\varepsilon^2$  can be estimated by solving the following equations:

$$\rho_1 = \frac{(\beta_1 - \theta_1)(1 - \theta_1\beta_1)}{1 + \theta_1^2 - 2\beta_1\theta_1} \quad (27)$$

$$\rho_2 = \beta_1\rho_1 \quad (28)$$

$$\sigma_\varepsilon^2 = \frac{(1 - \beta_1^2)\sigma_x^2}{1 - 2\beta_1\theta_1 + \theta_1^2} \quad (29)$$

### 4.3 Nonstationary Time Series Models

#### 4.3.1 ARIMA Models

The acronym ARIMA stands for "Auto-regressive integrated moving average." The ARMA models are suitable for data with the following two basic characteristics: (1) no apparent deviation from stationary assumption and (2) rapidly decreasing autocorrelation function. If these conditions are not met by a time series, a proper transformation should be performed to generate the time series with the above two conditions satisfied. This has usually been achieved by differencing, which is the essence of ARIMA models (Karamouz et al. 2012). It should be noted that the ARIMA models are nonstationary and cannot be used for synthetic generation of stationary time series, but they can be used for forecasting (Salas et al. 1980; Weeks and Boughton 1987).

#### 4.3.2 First-Order Markov Process with Periodicity: Thomas-Fiering Model

The first-order Markov model of the previous section assumes that the process is stationary in its first three moments. It is possible to generalize the model so that the

periodicity in hydrologic data is accounted for to some extent. The main application of this generalization has been in generating monthly streamflow where pronounced seasonality in monthly flows exists. In its simplest form, the method consists of the use of 12 linear regression equations. If, say, 30 years of records are available, the thirty January flows and the thirty December flows are abstracted and the January flow is regressed upon the December flow; similarly, the February flow is regressed upon the January flow, and so on for each month of the year.

$$\begin{aligned}
 q_{\text{Jan}} &= \bar{q}_{\text{Jan}} + b_{\text{Jan}}(q_{\text{Dec}} - \bar{q}_{\text{Dec}}) + \varepsilon_{\text{Jan}} \\
 q_{\text{Feb}} &= \bar{q}_{\text{Feb}} + b_{\text{Feb}}(q_{\text{Jan}} - \bar{q}_{\text{Jan}}) + \varepsilon_{\text{Feb}} \\
 &\dots\dots
 \end{aligned}$$

Fig. 5 shows a regression analysis of  $q_{j+1}$  on  $q_j$ , pairs of successive monthly flows for the months  $(j+1)$  and  $j$  over the years of record, where  $j = 1, 2, 3, \dots, 12$  (January, February, ..., December) and when  $j = 12, j+1 =$  January of next year (there would be 12 such regressions). If the regression coefficient of month  $j+1$  on  $j$  is  $b_j$ , then the regression line values of a monthly flow,  $\hat{q}_{j+1}$ , can be determined from the previous month's flow  $q_j$ , by the equation:

$$\hat{q}_{j+1} = \bar{q}_{j+1} + b_j(q_j - \bar{q}_j)$$

To account for the variability in the plotted points about the regression line reflecting the variance of measured data about the regression line, a further random component is added:

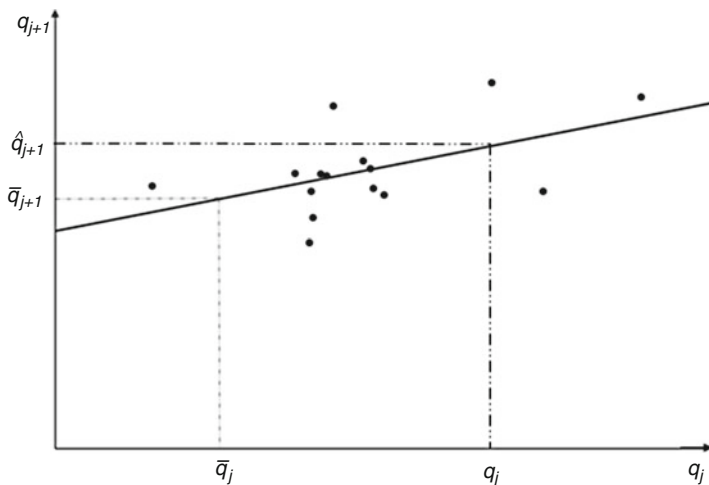


Fig. 5 Regression analysis of  $q_{j+1}$  on  $q_j$

$$\varepsilon_j = Z \cdot S_{j+1} \sqrt{(1 - r_j^2)}$$

where  $s_{j+1}$  is the standard deviation of flows in month  $j+1$ ,  $r_j$  is the correlation coefficient between flows in months  $j+1$  and  $j$  throughout the record, and  $Z = N(0, 1)$  a normally distributed random deviate with zero mean and unit standard deviation. The general form may be written as

$$\hat{q}_{j+1,i} = \bar{q}_{j+1} + b_j(q_{j,i-1} - \bar{q}_j) + Z_{j+1,i} \cdot S_{j+1} \sqrt{(1 - r_j^2)} \tag{30}$$

where  $b_j = r_j \times s_{j+1}/s_j$ . There are 36 parameters for the monthly model ( $\bar{q}$ ,  $r$ , and  $s$  for each month). Subscript  $j$  refers to a month. For monthly synthesis,  $j$  varies from 1 to 12 throughout the year. Subscript  $i$  is a serial designation from year 1 to year  $n$ . Other symbols are the same as mentioned earlier.

**The procedure for using the model is as follows:**

1. For each month,  $j = 1, 2, \dots, 12$ , calculate

(a) Mean flow  $\bar{q}_j = \frac{1}{n} \sum_i q_{j,i}; (i = j, 12 + j, 24 + j, \dots)$

(b) Standard deviation  $S_j = \sqrt{\frac{\sum_i (q_{j,i} - \bar{q}_j)^2}{n-1}}$

(c) The correlation coefficient with flow in the preceding month,

$$r_j = \frac{\sum_{i=1} (q_{j,i} - \bar{q}_j)(q_{j+1,i} - \bar{q}_{j+1})}{\sqrt{\sum_i (q_{j,i} - \bar{q}_j) \sum_i (q_{j+1,i} - \bar{q}_{j+1})}}$$

(d) The slope of the regression equation relating the month's flow to flow in the preceding month:

$$b_j = r_j \frac{S_{j+1}}{S_j}$$

2. The model is then a set of 12 regression Eq. (30)

$$\hat{q}_{j+1,i} = \bar{q}_{j+1} + b_j(q_{j,i} - \bar{q}_j) + Z_{j+1,i} \cdot s_{j+1} \sqrt{(1 - r_j^2)}$$

where  $Z$  is a random normal deviate  $N(0, 1)$

3. To generate a synthetic flow sequence, calculate (generate) a random number sequence  $\{Z_1, Z_2, \dots\}$ , and substitute in the model

### 4.3.3 ANMAX Model

Autoregressive moving-average models with exogenous inputs are denoted by ANMAX  $(p, q, b)$ , which shows a model with  $p$  autoregressive terms (AR( $p$ )),  $q$  moving-average terms (MA( $q$ )), and  $b$  exogenous input terms as a linear combination of the last  $b$  terms of a known and external time series  $d_t$  (Bailite 1980). The model formulation is as follows:

$$y_t = \varepsilon_t + \sum_{i=1}^p \beta_i(p)y_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \sum_{i=1}^b \eta_i d_{t-i} \quad (31)$$

where parameters  $\eta_1, \dots, \eta_b$  are related to the selected exogenous input. These models can be successfully utilized in cases where the historical data cannot completely cover the variations and behavior of the studied variables.

### 4.3.4 ARCH Model

Volatility (i.e., time-varying variance) clustering, in which large changes tend to follow large changes and small changes tend to follow small changes, has been well recognized in time series analysis. This phenomenon is called conditional heteroscedasticity, and can be modeled by ARCH-type (AutoRegressive Conditional Heteroscedasticity, ARCH) models, including the ARCH model introduced by Engle (1982) and their GARCH (the generalized ARCH) extension proposed by Bollerslev (1986). In these models, the key concept is the conditional variance, i.e., the variance conditional on the past.

### 4.3.5 Disaggregation Model

Disaggregation models are used to decompose time series into several subseries that are temporal or spatial fractions of the key time series. Valencia and Schaake (1973) and later extension by Mejia and Rousselle (1976) introduced the basic disaggregation model for temporal disaggregation of annual time series into seasonal time series. Disaggregation models of hydrologic time series are efficient techniques for cases where the preservation of statistical characteristics of both annual and seasonal scales is essential for the project under study. Most applications of disaggregation have been in the temporal domain, although some investigators have applied the same principle in the spatial domain.

## 4.4 Summary

Time series models are now a major tool in planning, operation, and decision making in hydrology and water resources. On one hand, time series models possess many appealing features. First, they can be used to model a time series without considering

its physical nature. Second, they can be used to extrapolate past patterns of behavior into the future. They allow a researcher, who has data only in past years, to forecast future events without having to search for other related time series data. Third, the time series approach also allows for the use of one time series to explain the behavior of another series, if the other time series data are correlated with a variable of interest and if there appears to be some cause for this correlation. On the other hand, some time series models, like ARIMA, are complex techniques, and require a great deal of experience and data. Although they often produce satisfactory results, those results depend on the researcher's level of expertise.

Machiwal and Jha (2006) reviewed both theoretical and applied research of time series models in the hydrological science. In hydrologic studies, time series models have been widely applied for detecting climatic changes (e.g., Kite 1989), investigating the long-term hydroclimatological trends (Lachtermacher and Fuller 1994), exploring the possible impact of climate change on hydrologic variables or water resources (Westmacott and Burn 1997), modeling precipitation (Janos et al. 1988), evapotranspiration (Mohan and Arumugam 1995), streamflow (Moatmari et al. 1999; Pekarova and Pekar 2006; Shao et al. 2009), groundwater (Houston 1983; Van Geer and Zuur 1997), drought (Mishra and Desai 2005), water quality (Ahmad et al. 2001), and water demand and consumption (Bougadis et al. 2005; Jorge 2007), etc.

---

## 5 Artificial Neural Network (ANN) Models

An artificial neural network (ANN) is a biologically inspired distributed computing processor system in parallel with certain performance characteristics resembling biological neural networks of the human brain, which differs from conventional computers in the way they process information (Haykin 1994). It has a distributed processing structure (Alp and Cigizoglu 2007) and consists of processing elements and connections between them with coefficients bound to the connections. Mathematically, ANNs may be treated as a universal approximator. They are able to extract the relation between inputs and outputs of a process without the physics being explicitly provided to them and to generalize the structure hidden within the whole dataset. ANN models are able to simulate nonlinear relationships through an automatic "training process" (Hsu et al. 1997). The ANN models have no limitations in the form of fixed assumptions or formal constraints and are faster compared with its conventional simulation methods, robust in noisy environments, flexible in many problems, and highly adaptive to the newer environments (Jain et al. 1999). There have been many standard ANN software that can be used to pursue intricate multipurpose nonlinear solutions.

### 5.1 Structure of ANN

ANNs are a computational model and have been developed as a generalization of mathematical models of human cognition or neural biology. An ANN is based on the following rules:

- Information processing occurs at many single elements called nodes, also referred to as units, cells, or neurons
- Signals are passed between nodes through connection links
- Each connection link has an associated weight that represents its connection strength
- Each node typically applies a nonlinear transformation called an activation function to its net input to determine its output signal

According to the absence or presence of feedback connections in a network, two types of architecture are distinguished: feedforward architecture and feedback architecture. A typical feedforward multilayer artificial neural network with a single hidden layer is illustrated in Fig. 6 (Friedman and Kandel 1999; Xiong et al. 2004).

This kind of ANNs can solve a wide variety of problems, such as classifying patterns, storing and recalling data, performing general mapping from input pattern (space) to output pattern (space), grouping similar patterns, or finding solutions to constrained optimization problems. It consists of input nodes  $\{X_i(p)\}_{i=1}^n$  (and one input to the neuron, called a bias, has a constant value of 1 and is usually represented as a separate input), hidden nodes  $\{Z_j(p)\}_{j=1}^l$  (and a bias), and output nodes  $\{Y_k(p)\}_{k=1}^m$ , where  $X$ ,  $Z$ , and  $Y$  represent the input, hidden, and output layers, respectively,  $n$ ,  $l$ ,  $m$  represent the number of the nodes in each layer, and  $p$  denotes the training pattern. The weights associated with the connections between input and hidden nodes are denoted by  $v_{ij}$ ,  $0 \leq i \leq n$ ,  $1 \leq j \leq l$ . Those between the hidden and the output nodes are denoted by  $w_{jk}$ ,  $0 \leq j \leq l$ ,  $1 \leq k \leq m$ .

For node  $Z_j$  in the hidden layer (Fig. 6), its effective aggregated input signal, denoted by  $z\_in_j$ , is calculated as:

$$z\_in_j = v_{0j} + \sum_{i=1}^n v_{ij}x_i, \quad 1 \leq j \leq l \quad (32)$$

where  $x_i$ ,  $1 \leq i \leq n$  represents the input to each node in the input layer.

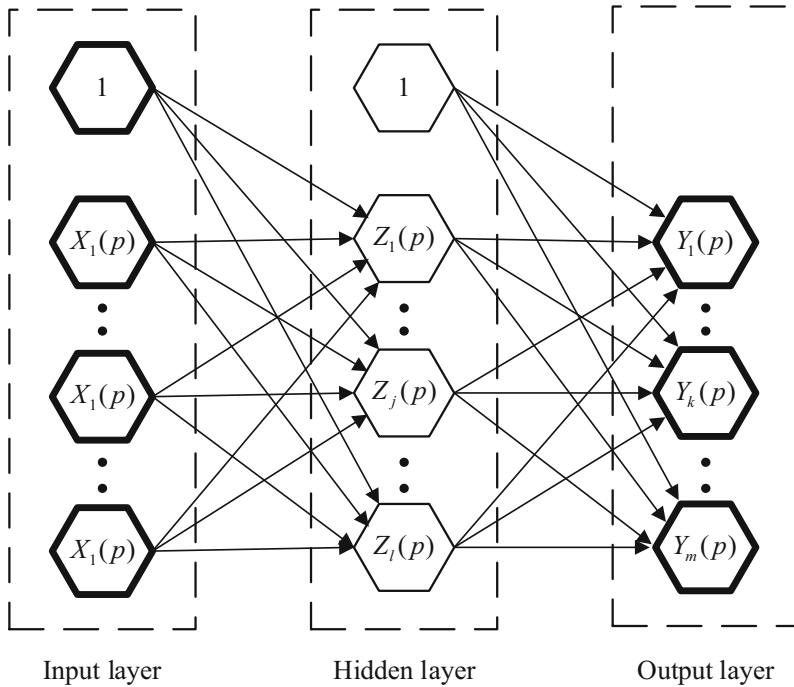
For node  $Z_j$ , its corresponding output signal, denoted by  $z_j$ , is obtained by using an activation function  $f(x)$

$$z_j = f(z\_in_j), \quad 1 \leq j \leq l \quad (33)$$

The most widely used activation function is the sigmoid function (Friedman and Kandel 1999). The sigmoid function is a bounded, monotonic, nondecreasing function that provides a graded and nonlinear response. Among several different sigmoid functions, the one most often used for ANNs is the logistic function

$$z_j = f(z\_in_j) = \frac{1}{1 + e^{-\sigma \cdot z\_in_j}} \quad (34)$$

where  $\sigma$  is an adjustable parameter used in the activation function  $f(x)$ . This function enables a network to map any nonlinear process.



**Fig. 6** A feedforward multilayer neural network with a single hidden layer (modified from Friedman and Kandel, 1999)

### 5.2 Network Training

In order for an ANN to generate an output vector  $Y = (y_1, y_2, \dots, y_p)$  that is as close as possible to the target vector  $T = (t_1, t_2, \dots, t_p)$ , a training process, also called learning, is employed to find optimal weight matrices  $v_{ij}$  and  $w_{jk}$ , which minimize a predetermined error function that usually has the form:

$$E = \sum_p \sum_l (y_l - t_l)^2 \tag{35}$$

Here,  $t_i$  is a component of the desired output  $T$ ;  $y_i$  is the corresponding ANN output;  $l$  is the number of output nodes; and  $p$  is the number of training patterns. Training is a process by which the connection weights of an ANN are adapted through a continuous process of stimulation by the environment in which the network is embedded. The learning ability of a neural network is achieved by applying a learning (training) algorithm.

Training algorithms are mainly classified into three groups (Kasabov 1996):

- (1) Supervised. The training examples comprise input vectors  $x$  and the desired output vectors  $y$ . Training is performed until the neural network “learns” to

associate each input vector  $x$  to its corresponding and desired output vector  $y$ ; for example, a neural network can learn to approximate a function  $y = f(x)$  represented by a set of training examples  $(x, y)$ . It encodes the examples in its internal structure.

- (2) Unsupervised. Only input vectors  $x$  are supplied; the neural network learns some internal features of the whole set of all the input vectors presented to it.
- (3) Reinforcement learning, sometimes called reward-penalty learning, is a combination of the above two paradigms; it is based on presenting input vector  $x$  to a neural network and looking at the output vector calculated by the network. If it is considered “good,” then a “reward” is given to the network in the sense that the existing connection weights are increased; otherwise, the network is “punished,” the connection weights, being considered as “not appropriately set,” decrease. Thus, reinforcement learning is learning with a critic, as opposed to learning with a teacher.

Learning is not an individual ability of a single neuron. It is a collective process of the whole neural network and a result of the training procedure. The connection weight matrix  $W$  has its meaning as a global pattern. It represents “knowledge” in its entirety. We do not know exactly how learning is achieved in the human brain. But learning (supervised or unsupervised) can be achieved in an artificial neural network. And there are some genetic laws of learning which have been discovered and implemented.

After training has been accomplished, it is hoped that the ANN will then be capable of generating reasonable results given new inputs. In contrast, an unsupervised training algorithm does not involve a teacher. During training, only an input data set is provided to the ANN that automatically adapts its connection weights to cluster those input patterns into classes that have similar properties. There are occasions when a combination of these two training strategies leads to reinforcement learning. A score or grade is used to rate the network performance over a series of training patterns. Most hydrologic applications have utilized supervised training. The manner in which the nodes of an ANN are structured is closely related to the algorithm that is used to train it.

### 5.3 Summary

ANNs have been utilized in many hydrologic problems and to evaluate if indeed all the strengths of ANNs have been effectively utilized in these applications. These applications include streamflow simulation (Shamseldin 1997; Hsu et al. 1995; Kişi, 2007; Wu and Chau 2011; He et al. 2014; Abrahart and See 2000; Aziz et al. 2014), water quality modeling (Rogers and Dowl 1994; Abyaneh 2014), ground water modeling (Aziz et al. 1992; Daliakopoulos et al. 2005), reservoir operation (Raman and Chandramouli 1996), water resources allocation and management (Raman and Sunilkumar 1995), evaporation estimation (Kumar et al. 2002; Shiri et al. 2014), hydrograph generation from hydrometeorological parameters (Ahmad



and Simonovic 2005), impact of climatic variations on flow discharge and dissolved organic carbon and nitrogen contents (Clair and Ehrman 1998), etc.

Zealand et al. (1999) claim that ANNs have the following beneficial model characteristics: (1) They infer solutions from data without prior knowledge of the regularities in the data. (2) ANNs are able to adapt to solutions over time to compensate for changing circumstances. (3) ANNs can generalize from previous examples to new ones, which is useful because real-world data are noisy, distorted, and often incomplete. (4) ANNs are also good at the abstraction of essential characteristics from inputs containing irrelevant data. (5) They are nonlinear, i.e., they can solve some complex problems more accurately than do linear techniques. (6) ANNs are highly parallel, containing many identical, independent operations that can be executed simultaneously, often making them faster than alternative methods.

However, ANNs also have several drawbacks for some applications. (1) Most of the ANN applications have been unable to explain the basic process in a comprehensibly meaningful way by which ANNs arrive at a decision. (2) When there is no learnable function or the data set is insufficient in size, they may fail to produce a satisfactory solution. (3) The optimum network geometry as well as the optimum internal network parameters are problem dependent and generally have to be found using a trial-and-error process. (4) The performance of an ANN deteriorates rapidly when the input vectors are far from the space of inputs used for training. (5) ANNs cannot cope with major changes in the system, because they are trained on historical data sets.

---

## 6 Fuzzy Logic Models

Fuzzy logic models are based on fuzzy logic system (Kruse et al. 1994; Klir and Yuan 1995; Kasabov 1996; Zimmermann 2001). Fuzzy logic systems, or fuzzy systems, are knowledge-based or rule-based systems. A fuzzy system is constructed from a set of fuzzy IF-THEN rules. A fuzzy IF-THEN rule is an IF-THEN statement in which some words are characterized by continuous membership functions (Wang 1997). For example, the following is an IF-THEN rule:

IF  $x$  is  $A$ , THEN  $y$  is  $B$

The functioning of fuzzy systems is based on fuzzy sets theory (Zadeh 1965). The fuzzy sets theory, as an extension of the classical sets theory, are generally used to describe imprecision or vagueness. By translation into fuzzy IF-THEN rules, subjective knowledge can be incorporated in fuzzy logic systems in a natural and transparent way. Furthermore, the major strength of fuzzy logic systems resides in their ability to infer the behaviour of complex systems purely from data (data-driven), but still providing some insight about their internal operation. Finally, fuzzy systems are flexible modeling tools, as their architecture and the inference mechanisms can be adapted to the given modeling problem.

Fuzzy logic models consist of three steps, taking inputs, applying fuzzy rules, and producing outputs. Inputs to a fuzzy system can be either exact, crisp values, or fuzzy values. Output values from a fuzzy system can be fuzzy or exact (crisp). The process of transforming a single crisp value into a fuzzy value is called fuzzification, while the process of transforming a fuzzy value into a single crisp value is called defuzzification.

## 6.1 Basic Concepts of Fuzzy Systems

### 6.1.1 Fuzzy Sets and Membership Functions

Fuzzy sets, which may be generally used to describe imprecision or vagueness, have firstly been introduced by Zadeh (1965). Fuzzy sets are sets of objects without clear boundaries; in contrast with ordinary sets where for each object it can be decided whether it belongs to the set or not, a partial membership in a fuzzy set is possible. The traditional way of representing elements  $x$  of a set  $A$  is through the characteristic function:

$$\begin{aligned}\mu_A(x) &= 1, \text{ if } x \text{ is an element of set } A, \text{ and} \\ \mu_A(x) &= 0, \text{ if } x \text{ is not an element of set } A,\end{aligned}$$

that is, an object either belongs or does not belong to a given set.

An object can belong to a set partially in fuzzy sets. The degree of membership is defined through a generalized characteristic function called membership function:

$$\mu_A(x) : U \rightarrow [0, 1]$$

where  $U$  is called the universe and  $A$  is a fuzzy subset of  $U$ .

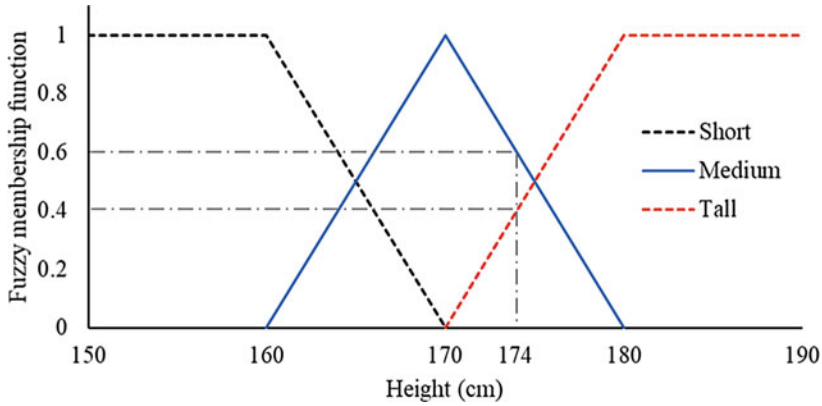
The values of the membership function are real numbers in the interval  $[0, 1]$ , where 0 means that the object is not a member of the set and 1 means that it belongs entirely to the set. Each value of the function is called a membership degree. One way of defining a membership function is through an analog function.

Fig. 7 (Kasabov 1996) shows three membership functions representing three fuzzy sets labeled as “short,” “medium,” and “tall,” all of them being fuzzy values of a variable “height.” As we can see, the value 174 cm belongs to the fuzzy set “medium” to a degree of 0.6 and at the same time to the set “tall” to a degree of 0.4.

To take another familiar example (Bárdossy et al. 1995), the set of young persons is fuzzy, as there is no generally accepted boundary between young and not young. The membership function of this set  $A$  may be defined as

$$\mu_A(x) = \begin{cases} 1 & \text{if } x \leq 25 \\ \frac{40-x}{15} & \text{if } 25 < x \leq 40 \\ 0 & \text{if } x > 40 \end{cases}$$

Fuzzy set theory can be considered as an extension of ordinary set theory; compared to the classical set theory, fuzzy set theory is very flexible in describing



**Fig. 7** Membership functions of representative three fuzzy sets for the variable “height” (modified from Kasabov, 1996)

the features of objects; it has advantages in expressing vague, uncertain, and imprecise information, which appears frequently in scientific and engineering fields (Zadeh 1965; Zimmermann 2001).

### 6.1.2 Fuzzy Rules

A fuzzy rule consists of a set of arguments in the form of fuzzy sets with membership functions and a response also in the form of a fuzzy set. For a general input vector, the rule is applied as:

$$\text{If } a_1 \text{ is } A_{i,1} \odot a_2 \text{ is } A_{i,2} \odot \dots \odot a_k \text{ is } A_{i,k}, \text{ then } B_i$$

where  $\odot$  is any logical operator, specified according to the application. Usually rules are formulated using AND/OR operators. For example, in modeling moisture movement in an unsaturated zone (Dou et al. 1999), it may consist of two premises (i.e.  $k = 2$ ).  $A_{i,1}$  may correspond to a class of upper layer moisture content (e.g., low, medium, high), and  $A_{i,2}$  to a class of lower layer moisture content (e.g., very low, medium, high, saturated), and the response  $B_i$  may be the actual quantity of water flux between the two layers.

## 6.2 Operations with Fuzzy Sets

Fuzzy set theory can be considered as an extension of ordinary set theory, i.e., the classical sets are a special case of fuzzy sets, when two membership degrees only, 0 and 1, are used, and crisp borders between the sets are defined. The following operations over two fuzzy sets  $A$  and  $B$  defined over the same universe  $U$  are the most common in fuzzy theory (Zadeh 1965; Zimmermann 2001).

*Containment,  $A \subset B$*

$A$  is contained in  $B$  (or, equivalently,  $A$  is a subset of  $B$ , or  $A$  is smaller than or equal to  $B$ ) if and only if  $\mu_A \leq \mu_B$ . In symbols

$$A \subset B \Leftrightarrow \mu_A \leq \mu_B$$

#### *Intersection, $A \cap B$*

The intersection of two fuzzy sets  $A$  and  $B$  with respective membership functions  $\mu_A(x)$  and  $\mu_B(x)$  is a fuzzy set of  $C$ , written as  $C = A \cap B$ , whose membership function is related to that of  $A$  and  $B$  by

$$\mu_C(x) = \min[\mu_A(x), \mu_B(x)], \quad x \in U$$

or in abbreviated form

$$\mu_C(x) = \mu_A(x) \wedge \mu_B(x), \quad x \in U$$

#### *Union, $A \cup B$*

The union of two fuzzy sets  $A$  and  $B$  with respective membership functions  $\mu_A(x)$  and  $\mu_B(x)$  is a fuzzy set of  $C$ , written as  $C = A \cup B$ , whose membership function is related to that of  $A$  and  $B$  by

$$\mu_C(x) = \max[\mu_A(x), \mu_B(x)], \quad x \in U$$

or in abbreviated form

$$\mu_C(x) = \mu_A(x) \vee \mu_B(x), \quad x \in U$$

#### *Equality, $A = B$*

Two fuzzy sets  $A$  and  $B$  are equal, written as  $A = B$ , if and only if

$$\mu_A(x) = \mu_B(x), \quad x \in U$$

#### *Complement*

The complement of a fuzzy set  $A$  is denoted by  $A'$  and is defined by

$$\mu_{A'}(x) = 1 - \mu_A(x), \quad x \in U$$

#### *Concentration, $\text{CON}(A)$*

$$\mu_{\text{CON}(A)}(x) = (\mu_A(x))^2, \quad x \in U$$

this operation is used as a linguistic modifier “very”

#### *Dilation, $\text{DIL}(A)$*

$$\mu_{\text{DIL}(A)}(x) = (\mu_A(x))^{0.5}, \quad x \in U$$

this operation is used as a linguistic modifier “more or less.”

*Algebraic product,  $A \cdot B$* 

The algebraic product of two fuzzy sets  $A$  and  $B$  with respective membership functions  $\mu_A(x)$  and  $\mu_B(x)$  is a fuzzy set of  $C$ , written as  $C = A \cdot B$ , whose membership function is related to those of  $A$  and  $B$  by:

$$\mu_C(x) = \mu_A(x) \cdot \mu_B(x), \quad x \in U$$

*Algebraic sum,  $A + B$* 

The algebraic sum of two fuzzy sets  $A$  and  $B$  with respective membership functions  $\mu_A(x)$  and  $\mu_B(x)$  is a fuzzy set of  $C$ , written as  $C = A + B$ , whose membership function is related to those of  $A$  and  $B$  by

$$\mu_C(x) = \mu_A(x) + \mu_B(x), \quad x \in U$$

The De Morgan laws are valid for the algebraic sum and difference.

*Bounded product*

The bounded product of two fuzzy sets  $A$  and  $B$  with respective membership functions  $\mu_A(x)$  and  $\mu_B(x)$  is a fuzzy set of  $C$ , whose membership function is related to those of  $A$  and  $B$  by

$$\mu_C(x) = \max[0, \mu_A(x) + \mu_B(x) - 1], \quad x \in U$$

*Bounded sum*

The bounded sum two fuzzy sets  $A$  and  $B$  with respective membership functions  $\mu_A(x)$  and  $\mu_B(x)$  is a fuzzy set of  $C$ , whose membership function is related to those of  $A$  and  $B$  by

$$\mu_C(x) = \max[1, \mu_A(x) + \mu_B(x)], \quad x \in U$$

*Bounded difference,  $A/ - /B$* 

The bounded difference of two fuzzy sets  $A$  and  $B$  with respective membership functions  $\mu_A(x)$  and  $\mu_B(x)$  is a fuzzy set of  $C$ , written as  $C = A/ - /B$ , whose membership function is related to those of  $A$  and  $B$  by:

$$\mu_C(x) = \min[0, \mu_A(x) - \mu_B(x)], \quad x \in U$$

*Normalization,  $\text{NORM}(A)$* 

$$\mu_{\text{NORM}(A)}(x) = \mu_A(x) / \max\{\mu_A(x)\}, \quad x \in U$$

The operations over fuzzy sets have some properties, for example, they are associative, commutative, and distributive, that is,

Associative:  $(a * b) * c = a * (b * c)$

Commutative:  $a * b = b * a$  (not valid for the bounded difference)

Distributive:  $a * (b \star c) = (a \star b) * (a \star c)$

where  $*$  and  $\star$  denote any operations from those listed above.

### 6.3 Types of Fuzzy Systems

To construct a fuzzy system, first we need to obtain a collection of fuzzy IF-THEN rules from human experts or based on domain knowledge. The next step is to combine these rules into a single system. Different fuzzy systems use different principles for this combination. There are three types of fuzzy systems that are commonly found in the literature: (i) pure fuzzy systems, (ii) Takagi-Sugeno-Kang (TSK) fuzzy systems, and (iii) fuzzy systems with fuzzifier and defuzzifier.

#### 6.3.1 Pure Fuzzy Systems

The main feature with the pure fuzzy system is that its inputs and outputs are fuzzy sets, whereas the inputs and outputs are real-valued variables in engineering systems (Fig. 8) (Wang 1997).

Each single IF-THEN rule of pure fuzzy systems has the following general form:

Rule  $m$ : IF  $(x_1 \text{ is } A_{1,m}) \text{ AND } (x_2 \text{ is } A_{2,m}) \text{ AND } \dots \text{ AND } (x_k \text{ is } A_{k,m})$  THEN  $y$  is... expressing the relation between  $k$  input variables  $x_1, x_2, \dots, x_k$  and output  $y$ . Terms  $A_{k,m}$  in the antecedents of the rules (i.e., the IF part of the rules) represent fuzzy sets (Zadeh 1965) used to partition the input space into overlapping regions.

#### 6.3.2 Takagi-Sugeno-Kang (TSK) Fuzzy Systems

In contrast to pure fuzzy systems, Takagi and Sugeno (1985) and Sugeno and Kang (1988) proposed another fuzzy system whose inputs and outputs are real-valued variables, named Takagi-Sugeno-Kang (TSK) fuzzy systems. The TSK fuzzy systems have the following general structure:

IF  $(x_1 \text{ is } A_{1,m}) \text{ AND } (x_2 \text{ is } A_{2,m}) \text{ AND } \dots \text{ AND } (x_k \text{ is } A_{k,m})$  THEN  $y = f_m(x_1, x_2, \dots, x_k)$

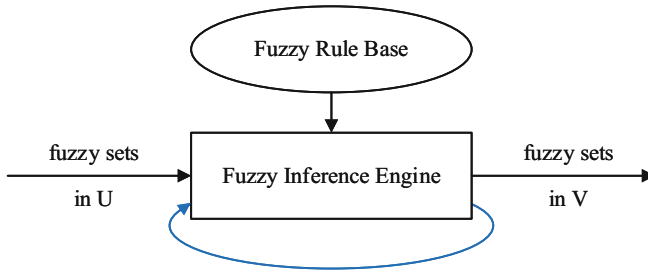
Each fuzzy rule in a TSK fuzzy inference system can be regarded as a local model of the system under consideration. The functions  $f_m$  are usually first-order polynomials, given by

$$f_m(x_1, x_2, \dots, x_k) = b_{0,m} + b_{1,m} \cdot x_1 + b_{2,m} \cdot x_2 + \dots + b_{k,m} \cdot x_k$$

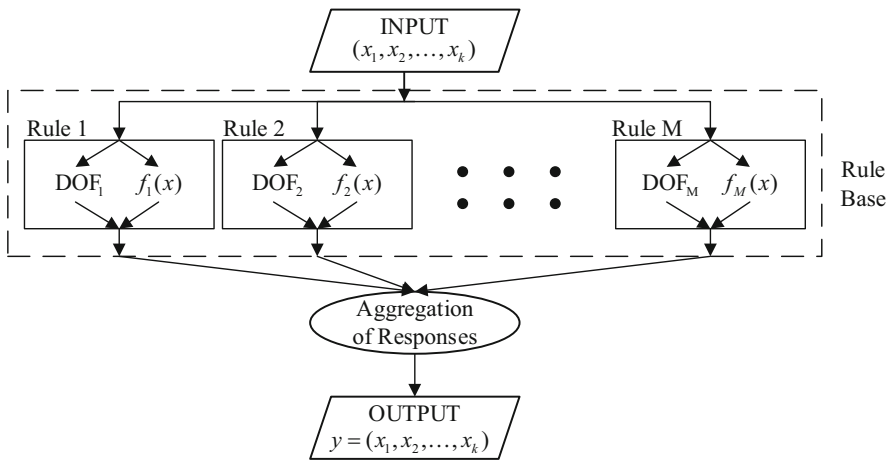
Fig. 9 shows a schematic diagram of the functioning of a typical multiple-input single-output TSK fuzzy system (Jacquin and Shamseldin 2006). The first stage in the inference process of a TSK fuzzy model is the calculation of the degree of fulfilment (DOF) of each rule. The output of each rule is obtained by the evaluation of the corresponding function  $f_m$ . Finally, the overall fuzzy model response is obtained as the weighted average of the individual rule responses.

The degree of fulfilment of a rule evaluates the compatibility of a given input vector with the antecedent of the rule (i.e., the IF part). The degree of fulfilment is normally evaluated using a T-norm, such as the algebraic product:

$$\text{DOF}_m(x_1, x_2, \dots, x_k) = \mu_{A_{1,m}}(x_1) \cdot \mu_{A_{2,m}}(x_2) \cdot \dots \cdot \mu_{A_{k,m}}(x_k)$$



**Fig. 8** Basic configuration of pure fuzzy systems (modified from Wang, 1997)



**Fig. 9** Functioning of a multiple-input single-output Takagi-Sugeno-Kang fuzzy inference system (modified from Jacquin and Shamseldin 2006)

Several types of the membership functions can be used for the fuzzy sets in the antecedents of the rules (Zimmermann 2001; Piegat 2001). The Gaussian membership functions, which have the following analytical expression:

$$\mu_{k,m}(x) = \exp \left[ -\frac{(x_k - c_{k,m})^2}{2\sigma_{k,m}^2} \right]$$

are a common choice (Chang et al. 2001; Gautam and Holz 2001; Xiong et al. 2001). In this case, each membership function has two parameters, namely the center  $c_{k,m}$  and the spread  $\sigma_{k,m}$ .

Some problems with the Takagi-Sugeno-Kang fuzzy system are listed as follows: (i) its THEN part is a mathematical formula and therefore may not provide a natural framework to represent human knowledge, and (ii) there is not much freedom left to

apply different principles in fuzzy logic, so that the versatility of fuzzy systems is not well represented in this framework.

### 6.3.3 Fuzzy Systems with Fuzzifier and Defuzzifier

In order to use pure fuzzy systems for simulating engineering systems, a simple method is to add a fuzzifier, which transforms a real-valued variable into a fuzzy set, to the input, and a defuzzifier, which transforms a fuzzy set into a real-valued variable, to the output. Thus, we get a fuzzy system with fuzzifier and defuzzifier, as shown in Fig. 10 (Wang 1997).

## 6.4 Adaptive Neuro-Fuzzy Inference System (ANFIS)

During the past four decades, significant progress has been made in the two artificial intelligence techniques, i.e., fuzzy inference system (FIS) and artificial neural networks (ANNs). A judicious integration of FIS and ANN can produce a functional neural fuzzy system capable of learning, high-level thinking, and reasoning (Jang et al. 1997; Loukas 2001). It provides an effective approach for dealing with large imprecisely defined complex systems. An ANFIS works by applying neural learning rules to identify and tune the parameters and structure of an FIS.

A typical architecture of an ANFIS, in which a circle indicates a fixed node, whereas a square indicates an adaptive node, is shown in Fig. 11 (Jang et al. 1997). For simplicity, we assume that the examined FIS has two inputs and one output.

In Fig. 11,  $x$  and  $y$  are the two crisp inputs and  $A_i$  and  $B_i$  are the linguistic labels associated with the node function. For rainfall-runoff modeling in hydrology, the input and output nodes represent rainfall process and discharge observations, respectively.

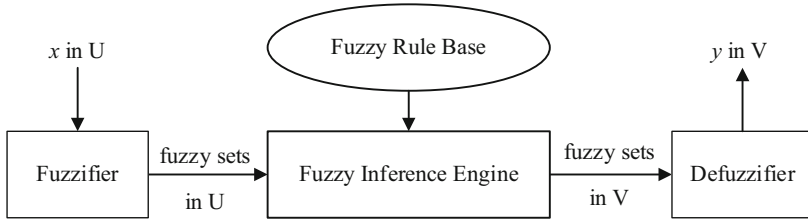
The attractive features of an ANFIS include: easy to implement, fast and accurate learning, strong generalization abilities, excellent explanation facilities through fuzzy rules, and easy to incorporate both linguistic and numeric knowledge for problem solving (Jang et al. 1997). Due to these fascinating features of the ANFIS, it is widely used in hydrological science.

## 6.5 Summary

An important contribution of fuzzy system theory is that it provides a systematic procedure for transforming a knowledge base into a nonlinear mapping. On one hand, fuzzy systems are multi-input-single-output mappings from a real-valued vector to a real-valued scalar (a multi-output mapping can be decomposed into a collection of single-output mappings), and the precise mathematical formulas of these mappings can be obtained; on the other hand, fuzzy systems are knowledge-based systems constructed from human knowledge in the form of fuzzy IF-THEN rules.

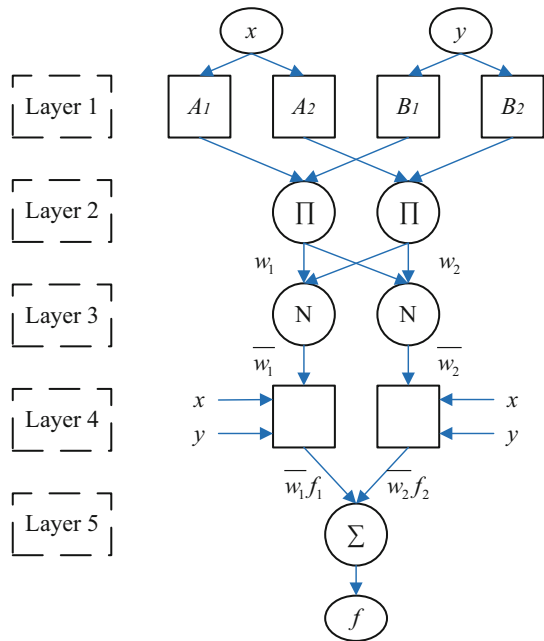
The fields of hydrology and water resources commonly involve a system of concepts, principles, and methods for dealing with modes of reasoning that are





**Fig. 10** Basic configuration of fuzzy systems with fuzzifier and defuzzifier (modified from Wang, 1997)

**Fig. 11** Architecture of the ANFIS (modified from Jang et al., 1997)



approximate rather than exact. The capability of dealing with imprecision gives fuzzy logic great potential for hydrological analysis and water resources decision making.

In hydrology, the concept of fuzzy theory and its application have found many applications in a number of research areas, such as groundwater flow in the unsaturated zone (Bárdossy and Disse 1993; Bárdossy et al. 1995; Dou et al. 1995; Schulz and Huwe 1997; Dou et al. 1999; Hong et al. 2002; Afshar et al. 2007), the interdependence between global circulation and rainfall (Pongracz et al. 2001), reconstruction of missing precipitation events (Abebe et al. 2000; Coulibaly and Evora 2007), rainfall-runoff modeling (Yu and Yang 2000; Hundecha et al. 2001; Huang et al. 2010), flood forecasting (See and Openshaw 1999; Xiong et al. 2001), flood frequency analysis (Shu and Burn 2004), reservoir

operation (Russell and Campbell 1996), water resources allocation and management (Yurdusev and Firat 2009), drought prediction (Pesti et al. 1996), and evaporation estimation (Cobaner 2011; Shiri et al. 2013, 2014).

The advantages of a fuzzy logic model include the following: (1) The model is not sensitive to parameter changes, and can be easily programmed, codes remain simple, short, and require little computer time. (2) The model is transparent and easy to understand due to its rule-based structure, which imitates the human way of thinking. (3) The fuzzy rule-based model can encode the expert's knowledge. (4) The most distinguishing property of fuzzy logic is that it deals with fuzzy propositions, that is, propositions which contain fuzzy variables and fuzzy values; thus, the fuzzy systems are especially good at dealing with nonlinear relationships. However, it must be reminded that fuzzy logic models, just like other types of the black-box models, can only describe the input-output relationships without explicit consideration of the internal hydrologic processes that lead to this transformation.

---

## 7 Frequency Analysis Models

Flood frequency estimation has been fundamental in engineering hydrology since Fuller (1914) approached the temporal variability of flood flows of extremely high return periods. The primary objective of frequency analysis is to relate the magnitude of extreme events to their frequency of occurrence through the use of probability distributions (Chow et al. 1988). The purpose of frequency analysis is to analyze past records of hydrologic variables so as to estimate future occurrence probabilities of extreme events.

The hydrologic data analyzed in frequency analysis are assumed to be independent and identically distributed (i.e., the i.i.d. assumption), if the hydrologic system producing them (e.g., a storm rain system) is considered to be stochastic, space-independent, and time-independent. The hydrologic data employed in frequency analysis should be carefully selected so that the assumptions of independence and identical distribution are satisfied. The data used in the analysis must also be evaluated in terms of the objectives, length of records available, and completeness of records. A frequency analysis can be performed using single-site data, regional data, or both. It can also include historical information and reflect physical constraints.

### 7.1 Graphical Method

Plotting position refers to the probability value assigned to each value in a random sample. It is used to calculate and graphically display the empirical frequency curve by plotting each ranked value against a probability scale. Numerous methods have been proposed for the determination of plotting positions, most of which are empirical. If  $n$  is the total number of values to be plotted and  $m$  is the rank of a value in a list ordered by descending magnitude, the exceedance probability of the  $m^{\text{th}}$  largest value is  $x_m$ ; the general form of plotting positions can be written:

$$P(X \geq x_m) = \frac{m - b}{n + 1 - 2b} \tag{36}$$

where  $b$  is a parameter commonly varying between 0 and 0.5 for various formulae. For example,  $b = 0.5$  for Hazen’s formula,  $b = 0.3$  for Chegodayev’s formula. The most popular plotting position formula is the Weibull plotting position when  $b = 0$ , which is an expected (unbiased) form of the exceedance probability at the  $m$ th largest observation for all distributions.

The procedure of graphical method includes the following:

- Select a  $Q_{\max}$  value from each year
- Arrange the data in a decreasing order, i.e.,  $Q_1 \geq Q_2 \geq Q_3 \dots$
- Assign a frequency/probability of exceedance to each  $Q_i$ . The most common method is the Weibull formula:

$$P(Q > Q_m) = \frac{m}{n + 1} \tag{37}$$

where  $n$  is the total number of data and  $m$  is the order of  $Q$ . This means that  $m(Q_{\max}) = 1$ , and  $m(Q_{\min}) = n$ .

- Plot  $Q$  versus  $P(Q > Q_m)$  or plot  $Q$  versus  $T = 1/P(Q > Q_m)$ , where  $T$  is return period
- On millimeter paper (without distribution assumption) – fit a curve
- On probability paper (with distribution assumption) – fit a straight line if the data fit the probability distribution as the probability paper presents
- Knowing the probability of  $P(Q > Q_T)$  or  $T$ , the design flow  $Q_T$  can be read from the plot or knowing the magnitude of  $Q_T$ , we can read  $P(Q > Q_T)$  or  $T$  from the plot

## 7.2 Analytical Method – Frequency Factor Method

A random variable  $X$  can be decomposed into two components and written as:

$$X = \bar{x} + \Delta X \tag{38}$$

where  $\Delta X$  is the deviation from the mean,  $\bar{x}$ . A new quantity can be defined as:  $K = \Delta X/s$ , where  $s$  is the standard deviation of the data, and the former equation can be rewritten as:

$$X = \bar{x} + sK \tag{39}$$

For a design value  $X_T$  with return period of  $T$ , Eq. (39) can be written as

$$X_T = \bar{x} + K_T s \text{ or } X_T = \bar{x}(1 + C_v K_T) \tag{40}$$

where  $C_v$  is the coefficient of variation, and  $K_T$  is the *frequency factor* depending on the *probability distribution* being used and on the *return period*,  $T$ .

Equation (40) is the working equation for frequency analysis, which can be used to calculate the design value  $X_T$  for given design level  $T$  and probability distribution. It can also be used to estimate the return period of a given  $X$  value.

Examples of design flow calculation for different probability distributions are presented below.

### 7.2.1 Example for Normal Distribution

If  $X$  is normally distributed, i.e.,  $X \sim N(\bar{x}, s^2)$ , from  $X_T = \bar{x} + K_T s$  we get

$$K_T = \frac{X_T - \bar{x}}{s}$$

That means  $K_T$  is the standardized normal variate  $Z$  ( $Z = \frac{X - \mu}{\sigma}$ ), i.e.,  $K_T = Z \sim N(0, 1)$ .  $K_T$  can then be read from the standard normal distribution table for a given  $T$  or calculated from the related equation by numerical methods.

### 7.2.2 Example for Lognormal Distribution

$X$  is said to be lognormally distributed if  $Y = \ln(X)$  is normally distributed with mean  $\mu_Y$  and variance  $\sigma_Y^2$ .

**The procedure for calculating  $X_T$  from log-normal distribution is:**

- Let  $y_i = \ln x_i$  for all  $x_i$ .
- Calculate  $\bar{y} = \frac{1}{n} \sum y_i$  and  $s_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}}$
- Read  $K_T$  from normal distribution table for a given  $T$
- Calculate  $Y_T = \bar{y} + K_T s_y$
- Calculate  $X_T = e^{Y_T}$

It should be noted that if only the mean and standard deviation ( $\bar{x}, s_x$ ) of a lognormally distributed variable  $X$  are available, then the mean and standard deviation ( $\bar{y}, s_y$ ) of the associated normally distributed variable  $Y = \ln(X)$  are calculated as:

$$\bar{y} = \ln \left( \frac{\bar{x}^2}{\sqrt{s_x^2 + \bar{x}^2}} \right), s_y = \sqrt{\ln \left( \frac{s_x^2}{\bar{x}^2} + 1 \right)}$$

### 7.2.3 Extreme Value Type I Distribution

The  $K_T$  value can either be calculated by using the equation below or read from the extreme value type I distribution table.

$$\begin{aligned}
 K_T &= -\frac{\sqrt{6}}{\pi} \left\{ 0.5772 + \ln \left[ \ln \left( \frac{T}{T-1} \right) \right] \right\}, \Rightarrow T \\
 &= \frac{1}{1 - \exp \left\{ -\exp \left[ -\left( 0.5772 + \frac{\pi K_T}{\sqrt{6}} \right) \right] \right\}}
 \end{aligned}
 \tag{41}$$

The design flow  $X_T$  can then be calculated using Eq. (40).

### 7.2.4 Pearson Type III Distribution

The Pearson type III distribution has three parameters,  $\lambda$ ,  $\beta$ , and  $\varepsilon$ , which can be estimated through the calculation of mean, standard deviation, and coefficient of skewness.

The procedure for the Pearson type III distribution is described as follows:

- Compute the mean,  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \rightarrow \lambda$
- Compute the standard deviation,  $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \rightarrow \beta$
- Compute the coefficient of skewness,  $C_s = \frac{n \sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)(n-2)s^3} \rightarrow \varepsilon$
- Compute  $K_T$  by Eq. (42) or read from the Table of  $K_T$  values for P-III distribution

$$K_T = z + (z^2 - 1)k + \frac{1}{3}(z^3 - 6z)k^2 - (z^2 - 1)k^3 + zk^4 + \frac{1}{3}k^5 \tag{42}$$

where  $K = C_s/6$

$$z = w - \frac{2.516 + 0.8029w + 0.01033w^2}{1 + 1.4328w + 0.1893w^2 + 0.00131w^3}$$

$$w = \left[ \ln \left( \frac{1}{p^2} \right) \right]^{1/2}$$

$$p = \frac{1}{T}$$

- Compute  $x_T = \bar{x} + K_T s$

### 7.2.5 Log-Pearson Type III Distribution

Similar to the lognormal distribution, the procedure for log-Pearson type-III distribution is as follows:

- Transform  $X$  to  $Y = \log(X)$  or  $\ln(X)$
- Compute the mean,  $y$

- Compute the standard deviation,  $s_y$
- Compute the coefficient of skewness,  $C_s$
- Compute  $K_T$  by Eq. (42) or read from the Table of  $K_T$  values for P-III distribution
- Compute  $y_T = \bar{y} + K_T s$
- Compute  $x_T = 10^{y_T}$  or  $x_T = e^{y_T}$

### 7.3 Data Sampling Methods

Data for frequency analysis studies can be compiled in several ways. In general, there are three data sampling methods: Annual Maximum Series, Partial Duration Series, and Annual Exceedance Series.

**Annual Maximum Series:** The Annual Maximum Series (AMS) data consist of the largest event in each year, regardless of whether the second largest event in a year exceeds the largest events of other years. An objection to using the AMS is that, in many cases, the second largest event in a year exceeds the largest event of other years.

**Partial Duration Series:** A Partial Duration Series (PDS) is a series of data which are selected so that their magnitude is greater than a predefined base value (Chow et al. 1988). Partial duration series or **peaks-over-threshold** (POT) pick all peaks above a threshold.

**Annual Exceedance Series:** If the base value of the PDS is selected so that the number of values in the series is equal to the number of years of record, the series is called an Annual Exceedance Series (AES) (Chow et al. 1988). An AES may be regarded as a special case of the PDS. Although AES is useful for some purposes, it may be difficult to verify that all the observations are independent.

The return period  $T_E$  of event magnitudes developed from an AES is related to the corresponding return period  $T$  for magnitudes derived from an AMS by (Chow 1964)

$$T_E = \left[ \ln \left( \frac{T}{T-1} \right) \right]^{-1} \quad (43)$$

### 7.4 Outliers and Zeros

*Outliers* are data points that depart significantly from the trend of the remaining data. The retention or deletion of these outliers can significantly affect the magnitude of statistical parameters computed from the data. The U.S. Water Resources Council (1982) Bulletin 17B suggests that outliers can be identified from

$$X_H = \bar{X} + K_n S_X$$

$$X_L = \bar{X} - K_n S_X$$

where  $X_H$  and  $X_L$  are the threshold values for high and low outliers, and  $K_n$  can be approximated from

$$K_n \approx 1.055 + 0.981 \log_{10} n$$

where  $n$  is the number of observations.

The detailed description of the treatment of outliers is contained in the U.S. Water Resources Council (1982) Bulletin 17B.

*Treatment of zeros:* Most hydrologic variables are bounded on the left by zero. Zero values should not simply be ignored, nor do they necessarily reflect inaccurate measurements of the minimum flow in a channel. A zero in a set of data that is being logarithmically transformed requires special handling. One solution is to add a small constant to all of the observations. Another method is to analyze the nonzero values and then adjust the relation to the full period of record. This method biases the results as the zero values are essentially ignored. A third and theoretically more sound method would be to use the theorem of total probability (For details see Haan (2002, pp168–169)).

## 7.5 Regionalization

Two broad categories of regionalization procedures have been widely used in the field of frequency analysis: the index-flood approach (Dalrymple 1960) and the multiple-regression approach (Benson 1962).

### 7.5.1 Index-Flood Method

The key assumption of an index-flood procedure is that the stations form a homogeneous region, meaning that the frequency distributions of the  $N$  stations are identical apart from a site-specific scaling factor, the index flood. We may then write

$$Q_i(F) = \mu_i q(F), \quad i = 1, 2, \dots, N \tag{44}$$

where  $Q_i(F)$ ,  $0 < F < 1$ , is the quantile function of the frequency distribution at site  $i$ ;  $\mu_i$  is the index flood (Hosking and Wallis 1997);  $N$  is the number of sites;  $q(F)$  is the regional growth curve, a dimensionless quantile function common to every site.

The index flood is estimated by  $\hat{\mu}_i = \bar{Q}_i$ , the sample mean of the data at site  $i$ , and the dimensionless rescaled data are computed by  $q_{ij} = Q_{ij} / \hat{\mu}_i$ , where  $Q_{ij}$  is the observed data at site  $i$ ,  $j = 1, 2, \dots, n_i$ , and  $n_i$  is the sample size at site  $i$ .

Hosking and Wallis (1997) suggested an index-flood method where parameters are estimated separately at each site. They considered the use of a weighted average of the at-site estimates:

$$\hat{\theta}_k^R = \sum_{i=1}^N n_i \hat{\theta}_k^{(i)} / \sum_{i=1}^N n_i \tag{45}$$

where  $\hat{\theta}_k^{(i)}$  stands for the L-moment of interest. The estimated regional quantile  $\hat{q}(F) = q(F; \hat{\theta}_1^R, \dots, \hat{\theta}_p^R)$  is obtained by substituting the estimates  $\hat{\theta}_k^{(i)}$  into  $q(F)$  (Hosking

and Wallis 1993). The quantile estimates at site  $i$  can be obtained using the estimates of  $\mu_i$  and  $q(F)$ .

$$\hat{Q}_i(F) = \hat{\mu}_i \hat{q}(F) \quad (46)$$

This index-flood procedure makes the following assumptions: (i) observations at any given site are identically distributed, and independent both serially and spatially; (ii) frequency distributions at different sites are identical apart from a scale factor; and (iii) the mathematical form of the regional growth curve is correctly specified.

### 7.5.2 Regional Regression

Regional regression models have long been used to predict flood quantiles at ungauged sites, and in a nationwide test, this method did as well as or better than more complex rainfall-runoff modeling in predicting flood quantiles (Newton and Herrin 1982).

Consider the traditional log-linear model for a statistic  $y_i$  which is to be estimated by using watershed characteristics such as drainage area and slope:

$$y_i = \alpha + \beta_1 \log(\text{area}) + \beta_2 \log(\text{slope}) + \dots + \varepsilon \quad (47)$$

A challenge in analyzing this model and estimating its parameters with available records is that one only obtains sample estimates, denoted  $\hat{y}_i$ , of the hydrologic statistics  $y_i$ . Thus, the observed error  $\varepsilon$  is a combination of (1) the time-sampling error in sample estimators of  $y_i$  (these errors at different sites can be cross-correlated if the records are concurrent) and (2) underlying model error (lack of fit) due to the failure of the model to exactly predict the true value of the  $y_i$ 's at every site. Often these problems have been ignored and standard ordinary least squares (OLS) regression has been employed (Thomas and Benson 1970).

## 7.6 Summary

Frequency analysis has been one of the earliest and most frequent uses of statistical methods in hydrology. In the earlier years (before 1960s), frequency analysis was mainly used for flood flow estimation, and nowadays, frequency analysis has been applied to almost every hydrological extreme variable, such as floods (Vogel et al. 1993; Vogel and Wilson 1996), low flows (Nathan and McMahon 1990; Lawal and Watt 1996; Durrans and Tomic 2001), rainfall events of various kinds (Pilgrim 1998; Öztekin 2007; Stedinger et al. 1993), droughts and dry spells (Lana and Burgueno 1998; Lana et al. 2008; Hallack-Alegria and Watkins 2007), etc. Hydrological frequency analysis (HFA) has been playing an essential role in the planning, design, and management of projects for flood control and water usages.

As to regional frequency analysis, recent advances mainly refer to the use of L-moments together with the index-flood method, as reported by Hosking and Wallis (1997). This methodology has been applied in modeling floods, rainfall extremes,



and low flows (Hosking et al. 1985; Vogel and Wilson 1996; Kjeldsen et al. 2001; Kumar et al. 2003; Yue and Wang 2004; Lim and Voeller 2009; Saf 2009).

The basic assumption of traditional HFA methods (both for one individual site and for a region) is that the hydrological data used are stationary, independent, and identically distributed over time. However, in the past decades, this stationarity assumption has been severely challenged because global climate change and/or large-scale human activities have altered the statistical characteristics of hydrological processes. Nonstationary frequency analysis is now a relatively new modeling approach and the number of studies is continuously increasing (Khaliq et al. 2006).

Studies of flood frequency analysis under nonstationary conditions have mostly assumed trends in time (Strupczewski et al. 2001; Renard et al. 2006; Leclerc and Ouarda 2007). In the last decade, some researchers have also explored the possibility of incorporating climate indices as external forcing into models for flood frequency analysis, assuming linear and nonlinear dependences (Sankarasubramanian and Lall 2003; El Aldouni et al. 2007; Ouarda and El-Aldouni 2011). Results have shown the feasibility of incorporating climate indices as covariates in the models, thus enabling the models to better describe changes in flood regimes over time.

**Acknowledgment** We are indebted to Yixing Yin, Yukun Hou, Qiang Zeng, and Xin-e Tao for their help in preparation of this chapter with proofreading and in supplying references, drawing figures, rewriting parts of the text, etc. We are also thankful to the two anonymous reviewers whose comments improved this chapter.

---

## References

- A.J. Abebe, D.P. Solomatine, R.G.W. Venneker, Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events. *Hydrol. Sci. J.* **45**(3), 425–436 (2000)
- R.J. Abraham, L. See, Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. *Hydrol. Process.* **14**, 2157–2172 (2000)
- H.Z. Abyaneh, Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters. *J. Environ. Health Sci. Eng.* **12**, 40 (2014.) <http://www.ijehse.com/content/12/1/40>
- J. Adamowski, H.F. Chan, S.O. Prasher, B. Ozga-Zielinski, A. Sliusarieva, Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada. *Water Resour. Res.* **48**(1) (2012). <https://doi.org/10.1029/2010WR009945>
- A. Afshar, M.A. Mariño, M. Ebtehaj, J. Moosavi, Rule-based fuzzy system for assessing ground-water vulnerability. *J. Environ. Eng.* **133**(5), 532–540 (2007)
- S. Ahmad, S.P. Simonovic, An artificial neural network model for generating hydrograph from hydro-meteorological parameters. *J. Hydrol.* **315**, 236–251 (2005)
- S. Ahmad, I.H. Khan, B.P. Parida, Performance of stochastic approaches for forecasting river water quality. *Water Res.* **35**(18), 4261–4266 (2001)
- S. Ali, N.C. Ghosh, R. Singh, Rainfall-runoff simulation using a normalized antecedent precipitation index. *Hydrol. Sci. J.* **55**(2), 266–274 (2010)
- M. Alp, H.K. Cigizoglu, Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environ. Model Softw.* **22**, 2–13 (2007)

- A. Aziz, R. Abd, K.F.V. Wong, A neural-network approach to the determination of aquifer parameters. *Ground Water* **30**(2), 164–166 (1992)
- K. Aziz, A. Rahman, G. Fang, S. Shrestha, Application of artificial neural networks in regional flood frequency analysis: a case study for Australia. *Stoch. Env. Res. Risk A.* **28**(3), 541–554 (2014)
- R.T. Bailite, Predictions from ARMAX models. *J. Econ.* **12**, 365–374 (1980)
- A. Bárdossy, M. Disse, Fuzzy rule-based models for infiltration. *Water Resour. Res.* **29**(2), 373–382 (1993)
- A. Bárdossy, A. Bronstert, B. Merz, 1-, 2- and 3-dimensional modeling of water movement in the unsaturated soil matrix using a fuzzy approach. *Adv. Water Resour.* **18**(4), 237–251 (1995)
- J.J. Beauchamp, Comparison of regression and time-series methods for synthesizing missing streamflow records. *Water Resour. Bull.* **25**, 961–975 (1989)
- M.A. Benson, Factors influencing the occurrence of floods in a humid region of diverse terrain. U.S. Geol. Surv., Water-Supply Pap., 1580-B, 64 pp. (1962)
- R.B. Billings, D.E. Agthe, State-space versus multiple regression for forecasting urban water demand. *J. Water Resour. Plann. Manag.* **124**(2), 113 (1998)
- T. Bollerslev, Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**(3), 307–327 (1986)
- J. Bougadis, K. Adamowski, R. Diduch, Short-term municipal water demand forecasting. *Hydrol. Process.* **19**, 137–148 (2005). <https://doi.org/10.1002/hyp.5763>
- G.E.P. Box, G.M. Jenkins, *Time Series Analysis: Forecasting and Control* (Holden-Day, San Francisco, 1976)
- F.J. Chang, H.F. Hu, Y.C. Chen, Counterpropagation fuzzy–neural network for streamflow reconstruction. *Hydrol. Process.* **15**(2), 219–232 (2001)
- S.P. Charles, M.A. Bari, A. Kitsios, B.C. Bates, Effect of GCM bias on downscaled precipitation and runoff projections for the Serpentine catchment, western Australia. *Int. J. Climatol.* **27**, 1673–1690 (2007)
- V.T. Chow, *Handbook of Applied Hydrology* (McGraw-Hill, New York, 1964)
- V.T. Chow, D.R. Maidment, L.W. Mays, *Applied Hydrology* (McGraw-Hill, New York, 1988)
- T.A. Clair, J.M. Ehrman, Using neural networks to assess the influence of changing seasonal climates in modifying discharge, dissolved organic carbon, and nitrogen export in eastern Canadian rivers. *Water Resour. Res.* **34**(3), 447–455 (1998)
- M. Cobaner, Evapotranspiration estimation by two different neuro-fuzzy inference systems. *J. Hydrol.* **398**, 292–302 (2011)
- P. Coulibaly, N.D. Evora, Comparison of neural network methods for infilling missing daily weather records. *J. Hydrol.* **341**, 27–41 (2007)
- I. Daliakopoulos, P. Coulibaly, I. Tsanis, Ground water level forecasting using artificial neural networks. *J. Hydrol.* **309**(1–4), 229–240 (2005)
- T. Dalrymple, Flood frequency methods. U.S. Geol. Surv. Water Supply Pap., 1543-A, 11–51 (1960)
- C.W. Dawson, R.J. Abrahart, Evaluation of two different methods for the antecedent precipitation index in neural network river stage forecasting. *Geophys. Res. Abstr.* **07522**, 9 (2007)
- C. Dou, W. Woldt, I. Bogardi, M. Dahab, Steady State Groundwater Flow Simulation With Imprecise Parameters. *Water Resour. Res.* **31**(11), 2709–2719 (1995)
- C. Dou, W. Woldt, I. Bogardi, Fuzzy rule-based approach to describe solute transport in the unsaturated zone. *J. Hydrol.* **220**(1), 74–85 (1999)
- G. Dumedah, P. Coulibaly, Evaluation of statistical methods for infilling missing values in high-resolution soil moisture data. *J. Hydrol.* **400**(1–2), 95–102 (2011)
- S.R. Durrans, S. Tomic, Comparison of parametric tail estimators for low-flow frequency analysis. *J. Am. Water Resour. Assoc.* **37**(5), 1203–1214 (2001)
- J.K. Eischeid, P.A. Pasteris, H.F. Diaz, M.S. Plantico, N.J. Lott, Creating a serially complete, national daily time series of temperature and precipitation for the Western United States. *J. Appl. Meteorol.* **39**, 1580–1591 (2000)

- S. El Aldouni, T. Ouarda, X. Zhang, R. Roy, B. Bobee, Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. *Water Resour. Res.* **43**, 1–13 (2007)
- A.K. Eldaw, J.D. Salas, L.A. Garcia, Long-range forecasting of the Nile River flows using climatic forcing. *J. Appl. Meteorol.* **42**(7), 890–904 (2003)
- R.F. Engle, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kindom inflation. *Econometrica* **50**(4), 987–1007 (1982)
- M.A. Fedora, R.L. Beschta, Storm runoff simulation using an antecedent precipitation index (API) model. *J. Hydrol.* **112**, 121–133 (1989)
- R.I.C.C. Francis, J.A. Renwick, A regression-based assessment of the predictability of New Zealand climate anomalies. *Theor. Appl. Climatol.* **60**, 21–36 (1998)
- M. Friedman, A. Kandel, *Introduction to Pattern Recognition: Statistical, Structural, Neural, and Fuzzy Logic Approaches*, Series in Machine Perception Artificial Intelligence, vol 32 (World Scientific, Singapore, 1999)
- W.E. Fuller, Flood flows. *Trans. ASCE* **77**(1293), 564–617 (1914)
- D. Gautam, K. Holz, Rainfall-runoff modelling using adaptive neuro-fuzzy systems. *J. Hydroinf.* **3**, 3–10 (2001)
- T. Glade, M.J. Crozier, P. Smith, Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical (Antecedent Daily Rainfall Model). *Pure Appl. Geophys.* **157**(6/8), 1059–1079 (2000)
- C.T. Haan, *Statistical Methods in Hydrology*, 2nd edn. (Iowa State University Press, Ames, 2002.) 496 pp
- M. Hallack-Alegria, D.W. Watkins, Annual and warm season drought intensity-duration-frequency analysis for Sonora, Mexico. *J. Clim.* **20**(9), 1897–1909 (2007)
- S. Haykin, *Neural Networks* (MacMillan, London, 1994)
- Z. He, X. Wen, H. Liu, J. Du, A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. *J. Hydrol.* **509**, 379–386 (2014)
- R.J. Heggen, Normalized antecedent precipitation index. *J. Hydrol. Eng. ASCE* **6**(5), 377–381 (2001)
- B.C. Hewitson, R.G. Crane, Climate downscaling: techniques and application. *Clim. Res.* **7**, 85–95 (1996)
- J.D. Hewlett, A.R. Hibbert, Factors affecting the response of small watersheds to precipitation in humid regions, in *Forest Hydrology*, ed. by W. E. Sopper, H. W. Lull, (Pergamon Press, Oxford, 1967), pp. 275–290.
- Y.S. Hong, M.R. Rosen, R.R. Reeves, Dynamic fuzzy modeling of storm water infiltration in urban fractured aquifers. *J. Hydrol. Eng.* **7**(5), 380–391 (2002)
- Y. Hong, R.F. Adler, F. Hossain, S. Curtis, G.J. Huffman, A first approach to global runoff simulation using satellite rainfall estimation. *Water Resour. Res.* **43**, W08502 (2007). <https://doi.org/10.1029/2006WR005739>
- J.R.M. Hosking, J.R. Wallis, *Regional frequency analysis: an approach based on L-moments* (Cambridge University Press, Cambridge, 1997)
- J.R.M. Hosking, J.R. Wallis, E.F. Wood, An appraisal of the regional flood frequency procedure in the UK Flood Studies Report. *Hydrol. Sci. J.* **30**, 85–109 (1985)
- J.R.M. Hosking, J.R. Wallis, Some statistics useful in regional frequency analysis: *Water Resource Research*, **29**(2), 271–281 (1993)
- J.F.T. Houston, Groundwater system simulation by time series techniques. *Ground Water* **21**, 301–310 (1983)
- K.-L. Hsu, H.V. Gupta, S. Sorooshian, Artificial neural network modeling of the rainfall-runoff process. *Water Resour. Res.* **31**, 2517–2530 (1995)
- K.L. Hsu, X. Gao, S. Sorooshian, H.V. Gupta, Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteorol.* **36**(9), 1176–1190 (1997)
- Y. Huang, X. Chen, et al., A fuzzy-based simulation method for modelling hydrological processes under uncertainty. *Hydrol. Process.* **24**(25), 3718–3732 (2010)

- Y. Hundecha, A. Bardossy, H.W. Werner, Development of a fuzzy logic-based rainfall-runoff model. *Hydrol. Sci. J.* **46**(3), 363–376 (2001)
- A.P. Jacquin, A.Y. Shamseldin, Development of rainfall-runoff models using Takagi–Sugeno fuzzy inference systems. *J. Hydrol.* **329**, 154–173 (2006)
- S.K. Jain, A. Das, D.K. Srivastava, Application of ANN for Reservoir Inflow Prediction and Operation. *J. Water Resour. Plan. Man. ASCE*, **125**(5), 263–271 (1999)
- J.S.R. Jang, C.T. Sun, E. Mizutani, *Neuro-Fuzzy and Soft Computing* (Prentice-Hall, Englewood Cliffs, NJ, 1997)
- B. Janos, D. Lucien, H.R. Omar, Practical generation of synthetic rainfall event time series in a semi-arid climatic zone. *J. Hydrol.* **103**, 357–373 (1988)
- C. Jorge, Forecasting water consumption in Spain using univariate time series models. *Proc IEEE Span. Comput. Intell. Soc.* **2007**, 415–423 (2007)
- M. Karamouz, S. Sara Nazif, M. Falahi, *Hydrology and Hydroclimatology: Principles and Applications* (CRC Press, Boca Raton, 2012)
- N.K. Kasabov, *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering* (MIT Press, Cambridge, MA, 1996)
- M.N. Khaliq, T.B.M.J. Ouarda, J.-C. Ondo, P. Gachon, B. Bobée, Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *J. Hydrol.* **329**, 534–552 (2006)
- Ö. Kişi, Streamflow forecasting using different artificial neural network algorithms. *J. Hydrol. Eng.* **12**(5), 532–539 (2007)
- G. Kite, Use of time series analyses to detect climatic change. *J. Hydrol.* **111**, 259–279 (1989)
- T.R. Kjeldsen, J.C. Smithers, R.E. Schulze, Flood frequency analysis at ungauged sites in the KwaZulu-Natal Province, South Africa. *Water SA* **27**(3), 315–324 (2001)
- G.J. Klir, B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications* (Prentice Hall, Upper Saddle River, 1995)
- M. A. Kohler, R. K. Linsley, Predicting the Runoff from Storm Rainfall, U.S. Weather Bureau Research Paper No. 34 (1951)
- R. Kruse, J.E. Gebhardt, F. Klowon, *Foundations of Fuzzy Systems* (Wiley, New York, 1994)
- M. Kumar, N. Raghuvanshi, R. Singh, W. Wallender, W. Pruitt, Estimating evapotranspiration using artificial neural network. *J. Irrig. Drain. Eng.* **128**(4), 224–233 (2002)
- R. Kumar, C. Chatterjee, S. Kumar, A.K. Lohani, R.D. Singh, Development of regional flood frequency relationships using L-moments for Middle Ganga Plains Subzone 1(f) of India. *Water Resour. Manag.* **17**(4), 243–257 (2003)
- G. Lachtermacher, J.D. Fuller, Backpropagation in hydrological time series forecasting, in *Stochastic and Statistical Methods in Hydrology and Environmental Engineering*, (Springer, Cham, 1994), pp. 229–242
- X. Lana, A. Burgueno, Daily dry-wet behavior in Catalonia (NE Spain) from the viewpoint of the first and second order Markov chains. *Int. J. Clim.* **18**, 793–815 (1998)
- X. Lana, M.D. Martínez, A. Burgueño, C. Serra, Return period maps of dry spells for Catalonia (northeastern Spain) based on the Weibull. *Hydrol. Sci. J.* **53**, 48–64 (2008)
- S.A. Lawal, W.E. Watt, Non-zero lower limit in low flow frequency analysis? *Water Resour. Bull.* **32**(6), 1159–1166 (1996)
- M. Leclerc, T.B.M.J. Ouarda, Non-stationary regional flood frequency analysis at ungauged sites. *J. Hydrol.* **343**, 254–264 (2007)
- Y.H. Lim, D.L. Voeller, Regional flood estimations in Red River using L-moment-based index-flood and Bulletin 17B Procedures. *J. Hydrol. Eng.* **14**(9), 1002–1016 (2009)
- S. Liu, L. Leslie, M. Speer, R. Bunker, Predicting forest fire danger using improved model derived soil-moisture and antecedent precipitation. International Congress on Modelling and Simulation, Townsville, Australia, July 14–17, 2003
- Y.L. Loukas, Adaptive Neuro-Fuzzy Inference System, An Instant and Architecture-Free Predictor for Improved QSAR Studies. *J. Med. Chem.* **44**(17), 2772–2783 (2001)

- T. Ma, C. Li, Z. Lu, B. Wang, An effective antecedent precipitation model derived from the power-law relationship between landslide occurrence and rainfall level. *Geomorphology* **216**, 187–192 (2014)
- D. Machiwal, M.K. Jha, Time series analysis of hydrologic data for water resources planning and management: A review. *J. Hydrol. Hydromech.* **54**(3), 237–257 (2006)
- A. Makarau, M.R. Jury, Predictability of Zimbabwe summer rainfall. *Int. J. Climatol.* **17**, 1421–1432 (1997)
- J.M. Mejia, J. Rousselle, Disaggregation Models in hydrology revisited. *Water Resour. Res.* **12**, 185–186 (1976). <https://doi.org/10.1029/WR012i002p00185>
- A.K. Mishra, V.R. Desai, Drought forecasting using stochastic models. *Stoch. Environ. Res. Risk Assess.* **19**, 326–339 (2005)
- A. Moatmari, M. Longoni, R. Rosso, A seasonal long memory stochastic model for the simulation of daily river flows. *Phys. Chem. Earth (B)* **24**(4), 319–324 (1999)
- S. Mohan, N. Arumugam, Forecasting weekly reference crop evapotranspiration series. *Hydrol. Sci. J.* **40**(6), 689–720 (1995)
- T.J. Mulvaney, On the use of self-registering rain and flood gauges. *Trans. Inst. Civ. Eng. Ireland* **4**(2), 1–8 (1850)
- R.J. Nathan, T.A. McMahon, Practical Aspects of Low-Flow Frequency Analysis. *Water Resour. Res.* **26**(9), 2135–2141 (1990)
- D.W. Newton, J.C. Herrin, Assessment of commonly used methods of estimating flood frequency. Transportation Research Record, Series 896, Washington, DC, pp. 10–30 (1982)
- T. Ouarda, S. El-Aldoumi, Bayesian nonstationarity frequency analysis of hydrological variables. *J. Amer. Water Resour. Assoc.* **47**, 496–505 (2011)
- T. Öztekin, Wakeby distribution for representing annual extreme and partial duration rainfall series. *Meteorol. Appl.* **14**, 381–387 (2007)
- G.R. Pandey, V.T.V. Nguyen, A comparative study of regression based methods in regional flood frequency analysis. *J. Hydrol.* **225**, 92–101 (1999)
- K. Pearson, On the generalized probable error in multiple normal correlation. *Biometrika* **6**, 59–68 (1908)
- P. Pekarova, J. Pekar, Long-term discharge prediction for the Turnu Severin station (the Danube) using a linear autoregressive model. *Hydrol. Process.* **20**, 1217–1228 (2006)
- G. Pesti, B.P. Shrestha, L. Duckstein, I. Bogárdi, A fuzzy rule-based approach to drought assessment. *Water Resour. Res.* **32**(6), 1741–1747 (1996)
- A. Piegat, *Fuzzy Modeling and Control. Studies in Fuzziness and Soft Computing (Book 69)* (Springer, Berlin, 2001)
- D. H. Pilgrim (ed.), *Australian Rainfall and Runoff: A Guide to Flood Estimation, Volumes I and II* (Institution of Engineers Australia, Canberra, 1998)
- A.S. Polebitski, R.N. Palmer, Seasonal residential water demand forecasting for census tracts. *J. Water Resour. Plan. Manag.* **136**(1), 27–36 (2009)
- R. Pongracz, J. Bartholy, I. Bogardi, Fuzzy rule-based prediction of monthly precipitation. *Phys. Chem. Earth BHydrol. Oceans Atmos.* **26**(9), 663–667 (2001)
- M.P. Rajurkar, U.C. Kothiyari, U.C. Chaube, Modeling of the daily rainfall-runoff relationship with artificial neural network. *J. Hydrol.* **285**, 96–113 (2004)
- H. Raman, V. Chandramouli, Deriving a general operating policy for reservoirs using neural networks. *J. Water Resour. Plan. Manag.* **122**(5), 342–347 (1996)
- H. Raman, N. Sunilkumar, Multi-variate modeling of water resources time series using artificial neural networks. *Hydrol. Sci. J.* **40**, 145–163 (1995)
- B. Renard, M. Lang, P. Bois, Statistical analysis of extreme events in a non-stationary context via a Bayesian framework: case study with peak-over-threshold data. *Stoch. Env. Res. Risk Assess.* **21**, 97–112 (2006)
- H.C. Riggs, *Streamflow characteristics: Developments in water science* 22 (Elsevier, Amsterdam, 1985)

- L.L. Rogers, F.U. Dowla, Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. *Water Resour. Res.* **30**(2), 457–481 (1994)
- S.O. Russell, P.F. Campbell, Reservoir operating rules with fuzzy programming. *J. Water Resour. Plann. Manag.* **122**(3), 165–170 (1996)
- B. Saf, Regional flood frequency analysis using L-moments for the west Mediterranean region of Turkey. *Water Resour. Manag.* **23**(3), 531–551 (2009)
- J.D. Salas, Analysis and modeling of hydrologic time series, Chapter 19, in *Handbook of Hydrology*, ed. by D. R. Maidment (McGraw-Hill, New York, 1993)
- J.D. Salas, J.W. Deulleur, V. Yevjevich, W.L. Lane, *Applied Modeling of Hydrologic Time Series* (Water Resources Publications, Littleton, 1980)
- A. Sankarasubramanian, U. Lall, Flood quantiles in a changing climate: seasonal forecasts and causal relations. *Water Resour. Res.* **39**, 4.1–4.12 (2003)
- K.E. Saxton, A.M. ASCE, A.T. Lenz, F. ASCE, Antecedent retention indexes predict soil moisture. *J. Hydraul. Div. ASCE* **93**(HY4, Proc. Paper 5351), 223–241 (1967)
- K. Schulz, B. Huwe, Water flow modeling in the unsaturatedzone with imprecise parameters using a fuzzy approach. *J. Hydrol.* **201**, 211–229, (1997). [https://doi.org/10.1016/S0022-1694\(97\)00038-3](https://doi.org/10.1016/S0022-1694(97)00038-3).
- L. See, S. Openshaw, Applying soft computing approaches to river level forecasting. *Hydrol. Sci. J.* **44**(5), 763–778 (1999)
- A.Y. Shamseldin, Application of a neural network technique to rainfall-runoff modeling. *J. Hydrol.* **199**, 272–294 (1997)
- V.N. Sharda, S.O. Prasher, R.M. Patel, P.R. Ojasvi & Chandra Prakash. Performance of Multivariate Adaptive Regression Splines (MARS) in predicting runoff in mid-Himalayan micro-watersheds with limited data / Performances de régressions par splines multiples et adaptives (MARS) pour la prévision d'écoulement au sein de micro-bassins versants Himalayens d'altitudes intermédiaires avec peu de données. *Hydrol. Sci. J.* **53**(6), 1165–1175 (2008). <https://doi.org/10.1623/hysj.53.6.1165>
- Q.X. Shao, H. Wong, M. Li, W.C. Ip, Streamflow forecasting using functional-coefficient time series model with periodic variation. *J. Hydrol.* **368**, 88–95 (2009)
- J. Shiri, A.H. Nazemi, A.A. Sadreddini, G. Landaras, O. Kisi, A.F. Fard, P. Marti, Global cross-station assessment of neuro-fuzzy models for estimating daily reference evapotranspiration. *J. Hydrol.* **480**, 46–57 (2013)
- J. Shiri, P. Marti, A.H. Nazemi, A.A. Sadreddini, O. Kisi, Local vs. external training of neuro-fuzzy and neural networks models for estimating reference evapotranspiration assessed through k-fold testing. *Hydrol. Res.* **46**(1), 72–88 (2014). <https://doi.org/10.2166/nh.2013.112>
- C. Shu, D.H. Burn, Homogeneous pooling group delineation for flood frequency analysis using a fuzzy expert system with genetic enhancement. *J. Hydrol.* **291**(1-2), 132–149 (2004)
- K. Schulz, B. Huwe, Water flow modeling in the unsaturatedzone with imprecise parameters using a fuzzy approach. *J. Hydrol.* **201**, 211–229 (1997). [https://doi.org/10.1016/S0022-1694\(97\)00038-3](https://doi.org/10.1016/S0022-1694(97)00038-3)
- W.T. Sittner, C.E. Schauss, J.C. Monroe, Continuous hydrograph synthesis with an API type hydrological model. *Water Resour. Res.* **5**(5), 1007–1022 (1969)
- J.R. Stedinger, R.M. Vogel, E. Foufoula-Georgiou, Frequency analysis of extreme events, Chapter 18, in *Handbook of Applied Hydrology*, ed. by D. R. Maidment (McGraw-Hill, New York, 1993), pp. 1–66
- W.G. Strupczewski, V.P. Singh, H.T. Mitosek, Non-stationary approach to at-site flood frequency modelling, III. Flood analysis of Polish rivers. *J. Hydrol.* **248**, 152–167 (2001)
- M. Sugeno, G.T. Kang, Structure identification of fuzzy model. *Fuzzy Sets Syst.* **28**, 15–33 (1988)
- T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst. Man Cybern.* **SMC-15**(1), 116–132 (1985)
- D.M. Thomas, M.A. Benson, Generalization of streamflow characteristics from drainage-basin characteristics, US Geological Survey, Water Supply Paper, 1975 (1970)

- United States Water Resources Council, Guidelines for determining flood flow frequency Bull. 17B, U.S. Water Resour. Coun. Hydrol. Comm., Washington, DC (1982)
- R.D. Valencia, J.C. Schaake, Disaggregation processes in stochastic hydrology. *Water Resour. Res.* **9**, 580–585 (1973). <https://doi.org/10.1029/WR009i003p00580>
- F.C. Van Geer, A.F. Zuur, An extension of Box-Jenkins transfer/noise models for spatial interpolation of groundwater head series. *J. Hydrol.* **192**, 65–80 (1997)
- W. Viessman Jr., G.L. Lewis, *Introduction to Hydrology*, 4th edn. (Harper Collins, New York, 1996)
- R.M. Vogel, C.N. Kroll, Generalized low-flow frequency relationships for ungaged sites in Massachusetts. *Water Resour. Bull.* **26**(2), 241–253 (1990)
- R.M. Vogel, I. Wilson, Probability distribution of annual maximum, mean, and minimum streamflows in the United States. *ASCE J. Hydrol. Eng.* **1**(2), 69–76 (1996)
- R.M. Vogel, W.O. Thomas Jr., T.A. McMahon, Flood-flow frequency model selection in Southwestern United States. *J. Water Resour. Plann. Manag.* **119**(3), 353–366 (1993)
- L.X. Wang, *A Course in Fuzzy Systems and Control* (Prentice-Hall, Englewood Cliffs, 1997)
- W.D. Weeks, W.C. Boughton, Tests of ARMA model forms for rainfall-runoff modeling. *J. Hydrol.* **91**, 29–47 (1987)
- J.R. Westmacott, D.H. Burn, Climate change effects on the hydrologic regime within the Churchill-Nelson River Basin. *J. Hydrol.* **202**(1-4), 263–279 (1997)
- R.L. Wilby, L.E. Hay, G.H. Leavesley, A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado. *J. Hydrol.* **225**, 67–91 (1999)
- C.L. Wu, K.W. Chau, Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *J. Hydrol.* **399**, 394–409 (2011)
- L. Xiong, A.Y. Shamseldin, K.M. O’Connor, A non-linear combination of the forecasts of rainfall-runoff models by the first-order Takagi–Sugeno fuzzy system. *J. Hydrol.* **245**, 196–217 (2001)
- L. Xiong, K.M. O’Connor, S. Guo, Comparison of three updating schemes using artificial neural network in flow forecasting. *Hydrol. Earth. Sys. Sci.* **8**(2), 247–255 (2004)
- A. Yasar, M. Bilgili, E. Simsek, Water demand forecasting based on stepwise multiple nonlinear regression analysis. *Arab. J. Sci. Eng.* **37**(8), 2333–2341 (2012)
- X.Y. Yu, S.Y. Liang, Forecasting of hydrologic time series with ridge regression in feature space. *J. Hydrol.* **332**, 290–302 (2007)
- P.S. Yu, T.C. Yang, Fuzzy multi-objective function for rainfall-runoff model calibration. *J. Hydrol.* **238**, 1), 1–1),14 (2000)
- S. Yue, C.Y. Wang, Possible regional probability distribution type of Canadian annual streamflow by L-moments. *Water Resour. Manag.* **18**(5), 425–438 (2004)
- M.A. Yurdusev, M. Firat, Adaptive neuro fuzzy inference system approach for municipal water consumption modeling, An application to Izmir, Turkey. *J. Hydrol.* **365**(3-4), 225–234 (2009)
- L.A. Zadeh, Fuzzy sets. *Inf. Control.* **8**, 338–353 (1965)
- C.M. Zealand, D.H. Burn, S.P. Simonovic, Short term streamflow forecasting using artificial neural networks. *J. Hydrol.* **214**, 32–48 (1999)
- H.J. Zimmermann, *Fuzzy Set Theory and Its Applications*, 4th edn. (Kluwer Academic Publishers, Boston, 2001)