Uppsala University
Department of Earth Sciences
Hydrology

TEXTBOOK

OF

HYDROLOGIC MODELS

(Lärobok i Avrinningsmodeller)

(edition 2002)

Chong-yu Xu
2002-10-20

Postal address	Visiting address	Telephone	Fax	Electronic mail
Villavägen 16	Earth Sciences Centre	+46 - 18 471 22 72 (direct)	+46 - 18 55 11 24	Chong-yu.Xu@hyd.uu.se
SE-752 36 UPPSALA	Villavägen 16	+46 - 18 471 25 28 (secretary)		http://www.hyd.uu.se/chong-yu/
Sweden	Uppsala			

PREFACE

According to the official document of UNESCO (1985) - 'Teaching aids in Hydrology': "The main purpose of using hydrological models in the teaching process is not to duplicate the complicated hydrological process in detail by a sophisticated model, but to demonstrate the principal elements of the process, their combination into a simple or comprehensive model, and the importance of the model in solving typical problems of engineering hydrology".

The course “Hydrological analysis and modelling” at the Department of Earth Sciences, Uppsala University consists of two parts, the first part deals with statistical analysis of hydrological data and the second part deals with modelling of runoff processes. This monograph serves as the lecture book for the second part. In particular, it is concerned with hydrologic models of precipitation-runoff process on catchment scale (groundwater modelling and urban storm flow modelling are not included). We will discuss their uses, formulations and methodology for model evaluation. The subject matter is divided into eight chapters.

Chapter 1 is to give students a basic knowledge about modelling in hydrology. That includes: (1) the reasons of using models in problem solving in hydrology; (2) some definitions commonly used in the literature of hydrological modelling; (3) classification of runoff models; and (4) objectives of hydrologic models.

Chapter 2 is to give an introduction to the students about the concept of the time series analysis and stochastic models. By this chapter, students will learn how to analyse a given time series of a hydrological variable, how stochastic models are formulated and what are their application fields.

Chapters 3 through 5 show how the principal hydrological processes, such as, precipitation (rainfall and snowfall), evapotranspiration, streamflow, infiltration, etc., are generally treated in different kinds of runoff models.

Chapter 6 is to study the methodology of model evaluation, which includes the issues in model selection, in model calibration (parameter estimation), in model verification, and estimation of its range of applicability. Although it is important to recognise that all four evaluations are of equal fundamental importance, estimation of the parameters and verification of model performance will receive more attention in this chapter.

Some topics in optimisation are discussed in Chapter 7. Since the automatic optimisation procedures have attracted increasing interests in the field of conceptual catchment modelling, and they are, nowadays, widely used to minimise differences between selected features of modelled and observed streamflows by systematic trial alterations in the values of the model parameters. Although it is not necessary that the user himself writes programs, since a great number of these programmes are stored in many computer libraries. Nevertheless it is useful to have an idea of the principles of these algorithms in order to judge the results obtained, and moreover to choose the algorithm in such a way that it is adapted to the problem proposed.

In chapter 8, some particular catchment models are discussed in more detail to illustrate how such models are formulated and what are their uses. The proposed models for discussion include, but not limited to, the HBV model (a simple conceptual

deterministic model), WASMOD model (a simple stochastic-conceptual snow and water balance model), the TOPMODEL (a relatively simple physically-based model), and the SHE model (a physically-based, distributed-parameter model).

Moreover, this course includes a certain amount of individual computer work, which is not specified in this text: preparation of the inputs, running of different models and analysis of the outputs, etc. Several computer exercises will be done. By these means, the students learn to use hydrological models and develop them into powerful tools in the process of decision making.

The content presented in the above chapters may be refined during the lecture time according to the interests of the students, and the orders of the appearance of the chapters may be changed.

I accept the full responsibility for any omissions, shortcomings, or mistakes that remain. I would benefit and be obliged if readers would transmit to me any errors, omissions, or criticisms.

This lecture note is intended for internal use only.

Chong-yu Xu
Docent in Hydrology
2002-08-05
New edition

CONTENTS

CHAPTER 1
MODELLING IN HYDROLOGY

1.1. WHY HYDROLOGICAL MODELS ARE NEEDED?

Recently, mathematical models have taken over the most important tasks in problem solving in hydrology (UNESCO, 1985). Many discussions regarding modelling have appeared in the scientific literature, but the rationale for model building was perhaps best expressed by Rosenblueth and Wiener (1945):

> No substantial part of the universe is so simple that it can be grasped and controlled without abstraction. Abstraction consists in replacing the parts of the universe under consideration by a model of similar but simpler structure. Models, formal or intellectual on the one hand, or material on the other, are thus a central necessity of scientific procedure.

Most hydrologic systems are extremely complex, and we cannot hope to understand them in all detail. Therefore, abstraction is necessary if we are to understand or control some aspects of their behaviour. Indeed, man has found through experience that understanding and predicting the behaviour of any significant part of his environment requires abstraction.

The catchment hydrologic models have been developed for many different reasons and therefore have many different forms. However, they are in general designed to meet one of the two primary objectives. One objective of catchment modelling is to gain a better understanding of the hydrologic phenomena operating in a catchment and of how changes in the catchment may affect these phenomena. Another objective of catchment modelling is the generation of synthetic sequences of hydrologic data for facility design or for use in forecasting. They are also providing valuable for studying the potential impacts of changes in landuse or climate. The variety of uses and the rapid increase both in scientific understanding and in technical support, from data collection systems and computer technology, have produced an enormous range in levels of sophistication.

1.2. HISTORICAL PERSPECTIVE

The development and application of hydrological models have gone through a long time period, the remarkable dates in the history of the development of hydrological models are:

- **In the 19th century**: The origins of rainfall-runoff modelling in the broad sense can be found in the middle of the 19th century arising in response to three types of engineering problems: (1) urban sewer design, (2) land reclamation drainage systems design, and (3) reservoir spillway design. In all three problems the design discharge was the major parameter of interest. The concept of the rational method for determining flood peak discharge from measurements of rainfall depths owes its origins to Mulvaney (1850), an Irish engineer who was

concerned with land drainage. Some Americans attribute first mention of the formula to one of their engineers engaged upon sewer design (Kuichiling, 1889). The method to give the peak flow Q_p is:

$$Q_p = CiA \tag{1.1}$$

Where C is the coefficient of runoff (dependent on catchment characteristics) *i* is the intensity of rainfall in time T_c and A is the area of catchment. T_c is the *time of concentration*, the time required for rain falling at the farthest point of the catchment to flow to the measuring point of the river. The well-known rational formula may be seen as the first generation of hydrologic models, where Q_p is the output variable, *i* and *A* are input variables, and C is the model parameter. By its main assumption, i.e., rainfall intensity and catchment characteristics are uniformly distributed in space and time, the use of rational formula is limited to small urban catchments.

- **In the 1920s**: During the 1920's, when the need for a corresponding formula for large catchments was perceived, many modifications were introduced in the rational method in order to cope with the non-uniform distribution, in space and time, of rainfall and catchment characteristics. The modified rational method, based on the concept of isochrones or lines of equal travel time, can be seen as the first basic rainfall-runoff model based on a transfer function whose shape and parameter were derived by means of topographic maps and the use of Manning's formula to evaluate the different travel times.
- **In the 1930s**: A major step forward in hydrological analysis was the concept of the unit hydrograph introduced by the American engineer Sherman in 1932 on the basis of superposition principle. Although not yet known at the time, the superposition principle implied many assumptions, i.e., the catchment behaves like a causative, linear time invariant system with respect to the rainfall/surface runoff transformation. The use of unit hydrograph made it possible to calculate not only the flood peak discharge (as the rational method does) but also the whole hydrograph (the volume of surface runoff produced by the rainfall event). At the end of 1930s and during the 1940s a number of techniques were proposed in order to improve the objectivity of the method and results, and the techniques of statistical analysis were invoked. A discussion on the different approaches and the relevant bibliography can be found in a report by Dooge (1973).
- **In the 1950s**: The real breakthrough came in the 1950s (Todini, 1988) when hydrologists became aware of system engineering approaches used for the analysis of complex dynamic systems. They finally realised that the unit hydrograph was the solution of a causative, linear time invariant system and that the use of mathematical techniques such as Laplace, Fourier and Z transforms could lead to the derivations of the response function from the analysis of input and output data. This was the period when conceptual models originated.
  The derivation of the unit hydrograph in discretized form (the unit graph) from sampled data (known as the inverse problem) still remained a big problem by that time, due to the non-particularly linear behaviour of the system and the generally large errors in input and output data. To overcome this problem, hydrologists found that shapes of the unit hydrograph could be provided on the basis of the solution of more or less simplified differential equations, such as for instance those describing the time behaviour of the storage in a reservoir or in a cascade of reservoirs (Nash, 1958, 1960). The unit hydrograph could then be

expressed in terms of few parameters to be estimated from catchment characteristics or by means of statistical procedures: moments, regression, maximum likelihood, etc. a bloom of these models gave rise an unbelievable variety of solutions: a cascade of linear reservoirs, linear channels, linear channels and reservoirs, nonlinear reservoirs (Prasad, 1967). However, in deriving the unit graph shape from actual data, very few advances were made before the work of Tikhonov (1963a,b) and the introduction of continuity and regularization constraints in the estimation phase (Eagleson et al, 1965; Natale and Todini, 1977) more realistic and reliable estimates of the unit hydrograph were obtained.

- **From the 1960s**: Many other approaches to rainfall-runoff modelling were considered in the 1960s. In search for a more physical interpretation of the process one could represent the behaviour of single components of the hydrologic cycle, at the catchment scale, by using a number of interconnected conceptual elements, each of which represented the purpose of a particular subsystem. A large number of conceptual, lumped, rainfall-runoff models appeared thereafter include: Dawdy and O'Donnell (1965), Stanford Model IV (Crawford and Linsley, 1966), Sacramento Model (Burnash et al., 1973), the HBV model (Bergström and Forsman, 1973), the Tank model (WMO, 1975) which represented differently the interconnected subsystems and were considered the leading models of 1960s and 1970s.
- **In the 1970s**:
  - Box and Jenkins (1970) provided hydrologists with an alternative model type – i.e. the autoregressive moving average (ARMA) model and other forms of time series stochastic models.
  - The real-time forecasting models as an answer for the need of warning in flood prone areas, and as a tool for reservoirs or hydraulic structure management were developed. Generally based on recent updating and recalibrating techniques such as Kalman filters (Kalman, 1960; Kalman and Bucy, 1961; Todini, 1978; Todini and Wallis, 1978; O'Connell, 1980; Wood, 1980; Wood and O'Connell, 1985).
  - One of the remark model developed in the late 1970s is the TOPMODEL (Beven and Kirkby, 1979) that is based on the idea that topography exerts a dominant control on flow routing through upland catchments is called. TOPMODEL calculates not only the streamflow hydrograph but information that is useful for linking hydrological calculations to hydrochemical models.
- **In the 1980s:** To meet the need of forecasting (1) the effects of land-use changes, (2) the effects of spatially variable inputs and outputs, (3) the movements of pollutants and sediments, and (4) the hydrological response of ungauged catchments where no data are available for calibration of a lumped model, the physically-based distributed-parameter models were developed. The most sophisticated models take a three-dimensional view of water exchange, with meshes superimposed vertically. These techniques have opened the way for major advances in modelling by linking them with elevation models (DTM/DEM) derived from maps, or with other data derived from raster-based satellite imagery, which may indicate vegetation cover, soil moisture patterns and lines of subsurface drainage. The Systéme Hydrologique Européen (SHE) model developed during the 1980s in a multinational programme stimulated by

the European Community is a good example of such models (Abbott et al., 1986).

- **From the late 1980s:** The evolution of continental-scale hydrology has placed new demands on hydrologic modellers. The macro-scale hydrological models were developed on the basis of the following motivations. First, for a variety of operational and planning purposes, water resource managers responsible for large regions need to estimate the spatial variability of resources over large areas, at a spatial resolution finer than can be provided by observed data alone. Second, hydrologists and water managers are interested in the effects of land-use and climate variability and change over a large geographic domain. Third, there is an increasing need of using hydrologic models as a base to estimate point and non-point sources of pollution loading to streams. Fourth, hydrologists and atmospheric modellers have perceived weaknesses in the representation of hydrological processes in regional and global atmospheric models. Examples of GIS supported macro-scale hydrological models include those developed by Vörösmarty et al. (1989), the VIC model (Wood et al., 1992) and the Macro-PDM (Arnell, 1999). These models are state-of-the-art tools in assessing regional and continental scale water resources.
- *Nowadays, mathematical models have taken over the most important tasks in problem solving in hydrology.*

1.3. HYDROLOGIC SYSTEM ANALYSIS AND MODELLING

We begin by defining some terms as they are to be used throughout this course.

- *a hydrological system*: A more general definition is given by Dooge (1973). In a simplified way it can be said as a set of physical, chemical and/or biological processes acting upon an input variable or variables, to convert it (them) into an output variable (or variables).

- *a variable*: is understood to be a characteristic of a system which may be measured, which assumes different values when measured at different times. Daily rainfall, runoff, evaporation, temperature, infiltration, soil moisture, etc. are some of examples.

- *a parameter*: is a quantity characterising a system. It may or may not remain constant in time (in most cases of modelling we consider it as time constant).

- *a model:* is a simplified representation of a complex system. Consequently, a model always describes the basic and most important components of a complex system, or as pointed out by Dooge (1977), a model involves similarity without identity and it simulates some, but not all the characteristics of the prototype system.

The watershed can be considered as a hydrologic system. The system boundary is drawn around the watershed by projecting the watershed divide vertically upwards and downwards to horizontal planes at the top and bottom (Fig1.1). Rainfall is the input, distributed in space over the upper plane; streamflow is the output, concentrated in space at the watershed outlet. Evaporation and subsurface flow are also outputs. By using the system concept, effort is directed to the construction of a model relating inputs and outputs rather than to the extremely difficult task of exact representation of the system details, which may not be significant from a practical point of view or may not be known. Nevertheless, knowledge of the physical system helps in developing a good model and verifying its accuracy.

The objective of hydrologic system analysis is to study the system operation and predict its output. A hydrologic system model is an approximation of the actual system;

its inputs and outputs are measurable hydrologic variables and its structure is the concept of system transformation.

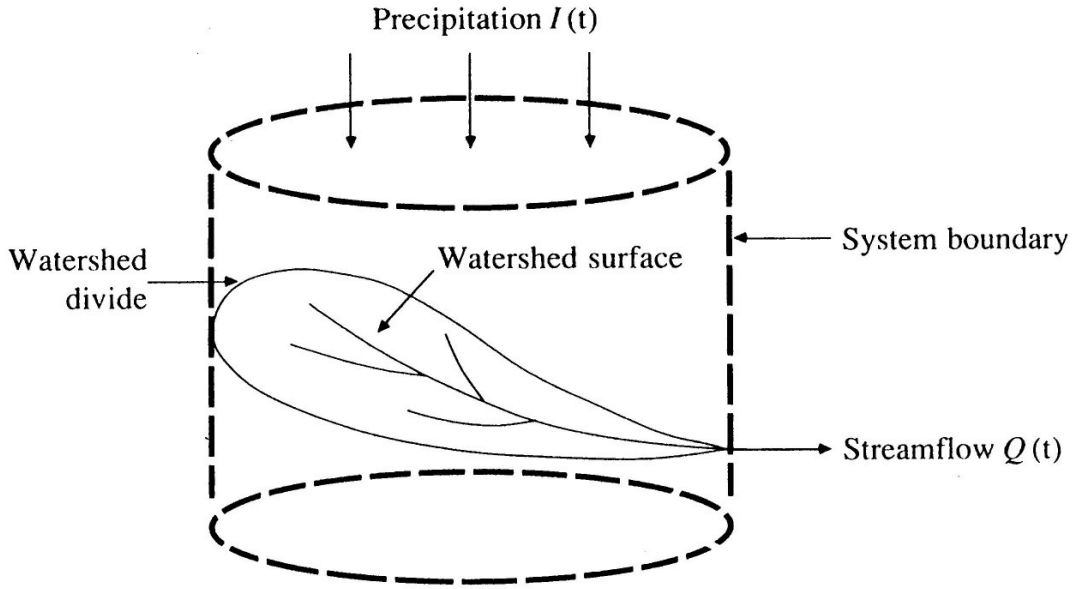


Fig.1.1 The watershed as a hydrologic system (from Chow et al, 1988)

A general model of the hydrologic system may be derived as follows. Let the input and output be expressed as functions of time, I(t) and Q(t) respectively, for t belonging to the time range T under consideration. The system performs a transformation of the input into the output represented by

$$Q(t) = \Omega I(t) \tag{1.2}$$

which is called the transformation equation of the system. The system Ω is a transfer function between the input and the output. If this relationship can be expressed by an algebraic equation, then Ω is an algebraic operator. For example, if

$$Q(t) = C I(t) \tag{1.3}$$

where C is a constant, then the transfer function is the operator

$$\Omega = \frac{Q(t)}{I(t)} = C \tag{1.4}$$

If the transformation is described by a differential equation, then the transfer function serves as a differential operator. For example, a linear reservoir has its storage S related to its outflow Q by

$$S = kQ \tag{1.5}$$

where k is a constant having the dimensions of time. By continuity, the time rate of change of storage *dS/dt* is equal to the difference between the input and the output

$$\frac{dS}{dt} = I(t) - Q(t) \tag{1.6}$$

Eliminating S between the two equations and rearranging,

$$k\frac{dQ}{dt} + Q(t) = I(t) \tag{1.7}$$

so

$$\Omega = \frac{Q(t)}{I(t)} = \frac{1}{1+kD} \tag{1.8}$$

where D is the differential operator *d/dt*. If the transformation equation has been determined and can be solved, it yields the output as a function of the input. Equation (1.8) describes a linear system if *k* is a constant. If *k* is a function of the input *I* or the output *Q* then (1.8) describes a nonlinear system which is much more difficult to solve.

1.4. CLASSIFICATION OF HYDROLOGIC MODELS

Hydrologic models can be variously classified. One of the classification methods used by Singh (1988) is used here which distinguishes hydrologic models as (1) material and (2) symbolic or formal. (Fig.1.2)

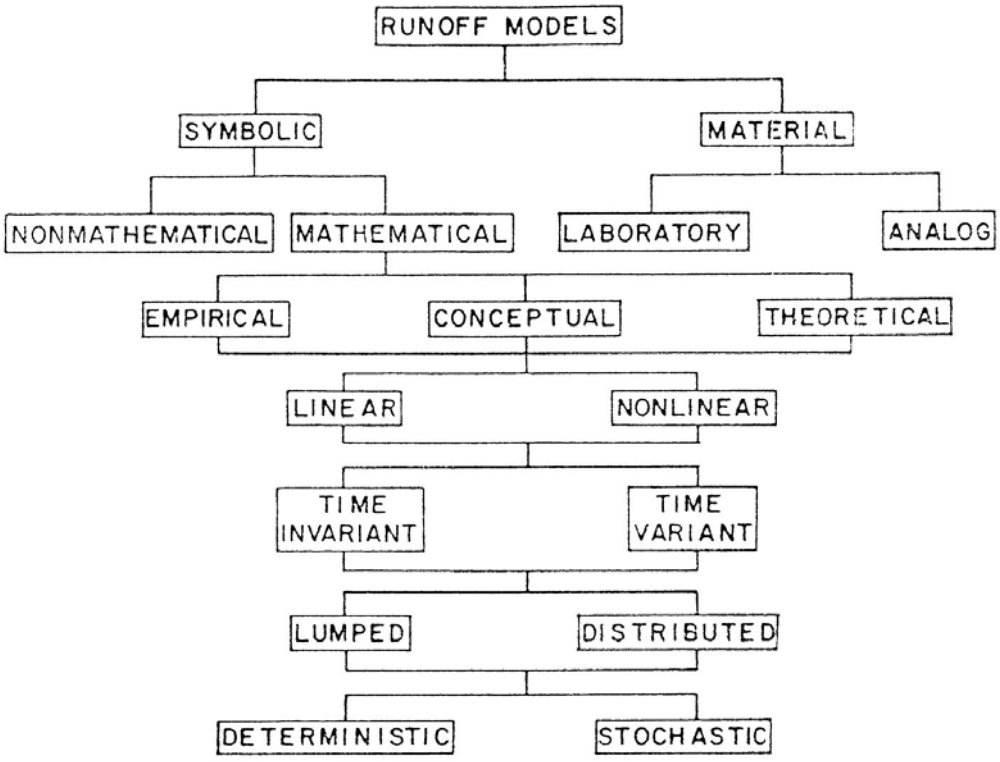


Fig.1.2. A classification of hydrologic models

1.4.1 Material models

A *material model* (also called a physical model in the literature, e.g. Chow et al, 1988) is the representation of the real system by another system, which has similar properties but is much easier to work with. Material (physical) models can be classified as iconic, scale, or "look-alike" models and analog models. A scale model represents the system on a reduced scale and bears a physical resemblance to the prototype system. Examples in this class may include laboratory watersheds, lysimeters, and hydraulic model of a dam spillway. Analog models measure different physical substances than the prototype (i.e. use another physical system having properties similar to those of the prototype), such as flow of electric current which represents the flow of water. An analog model does not physically resemble the prototype but depends on the correspondence between the symbolic models describing the prototype and the analog system.

Material models are useful in the following cases:

1). They may assist the researcher in replacing a phenomenon in an unfamiliar field.

2). A material model may permit experiments to be conducted under more favourable conditions than would be normally available with the prototype system.

A material model that does not involve a change in scale may still be valuable because experiments can be carried out more conveniently or can be repeated at will. Some experimental watershed systems installed in the NOPEX project area can be considered to be of prototype scale.

1.4.2 Symbolic or Formal models

A formal model (also called an abstract model in the literature, e.g. Chow et al., 1988) is a symbolic expression in logical terms of an idealised, relatively simple situation sharing the structural properties of the original system. Symbolic models can be variously expressed, in this course we are concerned with symbolic models of mathematical nature.

A mathematical model expresses the system behaviour by a set of equations, perhaps together with logical statements expressing relationships between variables and parameters. Equation (1.9) is an example of a mathematical model,

$$y_t = f^*(x_t, x_{t-1}, x_{t-2}, \cdots; y_{t-1}, y_{t-2}, \cdots; a_1, a_2, \cdots) + \varepsilon_t \qquad (1.9)$$

where x_t is the input variable, $f^*(\cdot)$ is a function of specified form and a_i, i=1,2, ..., are measured or estimated parameters, and ε_t is a residual expressing lack of fit between observed output y_t and fitted output $f^*(\cdot)$. In order to classify models it is necessary to consider what features they have in common and the respects in which they differ. The feature that all mathematical models have in common is that the observed output variable y_t (often discharge from a basin) derived from its fitted values $f^*(\cdot)$ by a residual amount ε_t; the respects in which they differ are the assumptions made about $f^*(\cdot)$ and assumptions made about ε_t. The most important terms, which are often seen in the hydrological literature, are explained in the following paragraphs.

1.4.3 The distinction between theoretical, conceptual and empirical models

Theoretical models (sometimes called white-box models or physically-based models) presumably are the consequences of the most important laws governing the phenomena. A theoretical model has a logical structure similar to the real-world system and may be helpful under changed circumstances. Examples of theoretical models may include watershed runoff models based on St. Venant equations, infiltration models based on two phase flow theory of porous media (Morel-Seytoux, 1978), evaporation models based on theories of turbulence and diffusion (Brutsaert and Mawdsley, 1976), and groundwater models based on fundamental transport equations (Freeze, 1971). An example of physically-based models is the SHE model (Abbott et al., 1986).

Empirical models (sometimes called black-box models or input output models) do not aid in physical understanding. They contain parameters that may have little direct physical significance and can be estimated only by using concurrent measurements of input and output. Examples are stochastic time series models. In many situations, empirical models can yield accurate answers and can, therefore, serve a useful tool in decision-making. The ARMA (autoregressive moving average model) and other time series models are examples of this class.

Conceptual models (sometimes called grey-box models) are intermediate between theoretical and empirical models. Hydrologic models are here considered as conceptual if the form of the function of equation (1.9) is, suggested by consideration of the physical processes acting upon the input variable(s) to produce the output variable(s). Generally, conceptual models consider physical laws but in highly simplified form. They are very many models belong to this class; an example which is familiar for us is the HBV model.

All three types of mathematical models are useful but in somewhat different circumstances. Each has its own effectiveness, depending upon the objective of study, the degree of complexity of the problem, and the degree of accuracy desired. There is no conflict between these models; they represent different levels of approximation of reality.

1.4.4 The distinction between linearity and non-linearity in the system-theory sense and in the statistical regression sense.

Models whether theoretical, conceptual or empirical may be linear or non-linear. Usage of the term linearity has at least two meanings. A model is linear in the system-theory sense (LST) if the principle of superposition holds: that is, given that y1(t), y2(t) are the outputs corresponding to inputs x1(t), x2(t), a model is LST if the output corresponding to input x1(t)+x2(t) is y1(t)+y2(t). This is the sense in which linearity is most widely used in the literature. However, linearity has an alternative meaning; the model is linear in the statistical regression sense (LSR) if it is linear in the parameters to be estimated, and it is in this sense that it is used by mathematical modellers in fields other than hydrology. Thus if input x(t) and output y(t) were related by the equation $y = a + bx + cx^2$, this model is linear in statistical regression sense, but non-linear in the system-theory sense; the converse is true for $y = a + x/b$.

1.4.5 The distinction between time-invariant and time-variant models

A model is time-invariant if its input-output relationship does not change with time. The form of the output depends only on the form of the input and not on the time at

which the input is applied. Models do not have this property are called time-variant. Most hydrologic systems are time-variant due to variations in solar activity during the day and seasonal variations during the year. For simplicity, they are assumed to be time-invariant.

1.4.6 The distinction between lumped and distributed models

In terms of spatial discretization or resolution we can identify an ascending scale of sophistication beginning with lumped models treating the complete basin as a homogeneous whole, through semi-distributed models, which attempt to calculate flow contributions from separate areas or sub-basins that are treated as homogeneous within themselves, to fully distributed models, in which the whole basin is divided into elementary unit areas like a grid net and flows are passed from one grid point (node) to another as water drains through the basin (Fig. 1.3). Becker and Serban (1990) further distinguished spatial variability of the models into geometrically-distributed models, which express spatial variability in terms of the orientation of the network points one to another and their distance apart (Fig.1.3), and probability-distributed models describe the spatial variability without reference to the geometrical configuration of the points in the network at which an input variable such as rainfall is measured, or for which a model parameter is to be measured or estimated. For example, the Stanford watershed model (Crawford and Linsley, 1966) is of this type. It is assumed that infiltration capacity at any time varies over the segment. For lack of better information this variation is assumed to be linear.

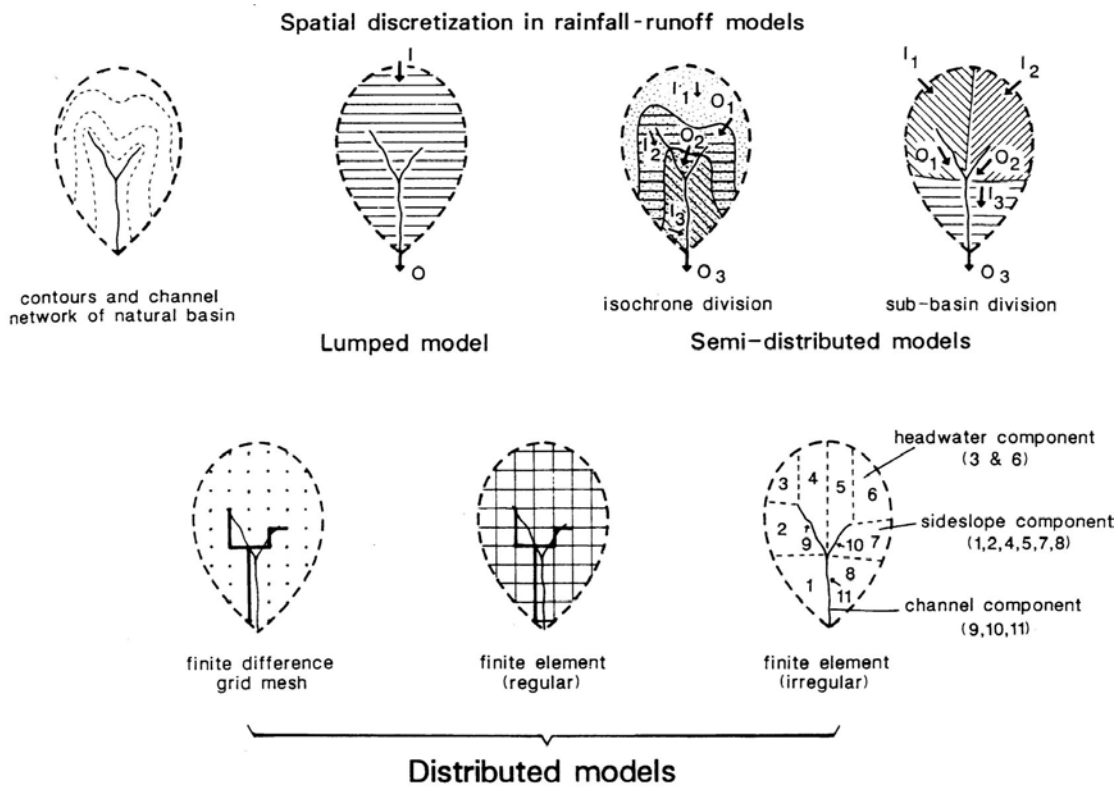


Fig.1.3 Graphic representation of geometrically – distributed and lumped models. (from Jones, 1997). I is input and O is output.

1.4.7 The distinction between deterministic and stochastic models

If any of the variables x_t, y_t, ε_t in equation (1.8) are regarded as random variables having distributions in probability, then the model is a stochastic model: stochastic, rather than statistical (probabilistic), to emphasise the time-dependence of the hydrological variables related by the model. If all variables in equation (1.8) are regarded as free from random variation, so that none is thought of as having a distribution in probability, then the model is here regarded as deterministic.

1.4.8 Summary on classification

The two most often used classification methods are that according to the description of the physically processes hydrological models may be classified as conceptual and physically based, and according to the spatial description of catchment processes as lumped and distributed. In this respect, two typical model types are lumped conceptual and the distributed physically based ones. Typical examples of lumped conceptual model codes are the Stanford watershed model (Crawford and Linsley, 1966), the HBV model (Bergström, 1976) and the Sacramento (Burnash, 1995). Typical models of distributed physically based are the SHE (Abbott et al., 1986a,b), the IHDM (Beven et al., 1987) and the Thales (Grayson et al., 1992a,b). A code such as TOPMODEL (Beven and Kirkby, 1979) may by characterized as conceptual distributed.

1.5 THE USE OF HYDROLOGIC MODELS

Physically-based or theoretical models are often use in research purpose to gain a better understanding of the hydrologic phenomena operating in a catchment and of how changes in the catchment may affect these phenomena. The hydrologic phenomena they calculate are generally defined by the laws of continuity, energy and momentum. As such these models are seldom used to generate synthetic data.

All other type of models, vary from deterministic form, using much information about the physical processes involved, to “black box” forms, where physical processes are not involved, are used in operational purpose to generate synthetic sequences of hydrologic data for facility design or for use in forecasting.

1.6 METHODOLOGY FOR USING HYDROLOGIC MODELS

Dooge (1972) outlined a rational methodology for the use of hydrologic models. This methodology consists of a number of steps. These, with slight modifications, are as follows (see also Singh, 1988):

1. Define the problem.
2. Specify the objective.
3. Study the data available.
4. Determine the computing facilities available.
5. Specify the economic and social constraints.
6. Choose a particular class of hydrologic models.
7. Select a particular type of model from the given class.
8. Calibrate the model (that is, optimise the parameters).
9. Evaluate the performance of the model.
10. Use the model for prediction purposes.
11. Embed the model in a more general model.

In a systematic application of the methodology, steps 6-9 could be iterated until a satisfactory model was obtained. Step 9 is, however, crucial in this entire operation. Only when a model has been objectively calibrated and evaluated can it be applied to a specific problem with assurance that the best use is being made of the data and that something is known about the order of magnitude of the accuracy of prediction. There exists a multitude of hydrologic models. However, a rational methodology for their choice is yet to be developed.

REFERENCES OF CHPATER 1

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J.: 1986. An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of a physically-based distributed modelling system, and 2: Structure of a physically-based distributed modelling system. *J. Hydrol*. 87, 45-59 and 61-77.

Arnell, N.W., 1992. Factors controlling the effects of climate change on river flow regimes in a humid temperate environment, *J. Hydrol*., 132: 321-342.

Becker, A. and P Serban, 1990. *Hydrological models for water-resources system design and operation.* Geneva, WMO Operational Hydrology Report No 34:80pp.

Bergström, S., 1976. Development and application of a conceptual runoff model for Scandinavian catchments. SMHI report, Nr. RHO 7.

Beven, K.J. and M.J. Kirkby, 1979. A physically based, variable contributing area model of basin hydrology, Hydrological Sciences Bulletin 24: 43-69.

Beven, K.J., Calver, A. And Morris, E.M., 1987. The Institute of Hydrology distributed model. Institute of Hydrology Report 98, Wallingford, UK.

Brutsaert, W. and J.A. Mawdsley, 1976. The applicability of planetary boundary layer theory to calculate regional evapotranspiration. Water Resour. Res. 12: 852-858.

Burnash, R.J.C., Ferral, R.L. and McGuire, R.A., 1973. A generalised streamflow simulation system – conceptual modelling for digital computers, US Department of Commerce, national Weather Service and State of California, Department of Water resources.

Chow, V.T., et al., 1988. Applied Hydrology. McGraw-Hill Series in Water Resources.

Crawford, N.H. and R.K. Linsley, 1966. Digital simulation in hydrology. Stanford Watershed Model IV. Stanford, Dept of Civil Engineering, University of California, Technical Report No 39.

Dawdy, D.R. and O'Donnell, T., 1965. Mathematical models of catchment behaviour. Proceedings of American Society of Civil Engineers: Journal of the Hydraulics Divisions of the ASCE 91(HY4): 123-137.

Dooge, J.C., 1972. Mathematical models of hydrologic systems. Proceedings of the International Symposium on Modelling Techniques in Water Resources Systems, Ottawa, Canada, vol.1, 171-189.

Dooge, J.C., 1977. Problems and methods of rainfall-runoff modelling. In: T.A. Ciriani, V. Malone and J.R. Wallis (Eds). Mathematical models for surface water hydrology. John Wiley and Sons Ltd. London, England.

Dooge, J.C.I., 1973. Linear theory of hydrologic system. Technical Bulletin No 1468. Agricultural Research Service, USDA. Washington D.C.

Eagleson, P.S., Mejia, R. and March, F., 1965. The computation of optimum realizable unit hydrographs from rainfall and runoff data. Hydrodynamic Laboratory Report No 84, MIT.1970 Dynamic Hydrology, New-York, McGraw-Hill.

Freeze, R.A., 1971. Three-dimensional, transient, saturated-unsaturated flow in a groundwater basin. Water Resour. Res. 7(2):347-366.

Grayson, R.B., Moore, I.D. and McHahon, T.A., 1992a,b. Physically based hydrologic modelling. Water Resources Research 28(10), 2639-2658.

Jones, J.A.A., 1997. *Global Hydrology*, Pearson Education Limited, Edinburgh Gate, Harlow, England.

Kalman, R.E. and Bucy, R., 1961. New results in linear filtering and prediction. J. Basic Eng. (ASME), 83D, 95-108.

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. J. Basic Eng. (ASME), 82D, 35-45.

Kuichling, E. 1889. The relation between the rainfall and the discharge of sewers in populous districts. ASCE Trans., 20: 1.56.

Morel-Seytoux, H.J., 1978. Derivation of equations for variable rainfall infiltration. Water Resour. Res. 14(4): 561-568.

Mulvaney, T.J., 1850. On the use of self-registering rain and flood gauges. Trans. Inst. Civ. Eng. Ireland, 4(2): 1-8.

Nash, J.E., 1958. The form of the instantaneous unit hydrograph. IUGG Gen. Assem. Toronto, Vol III Publ No 45, IAHS, Gentbrugge, p114-121.

Nash, J.E., 1960. A unit hydrograph study with particular reference to British catchments. Proc Inst Civ Eng.

Natale, L. and Todini, E., 1977. A constrainted parameter estimation technique for linear models in hydrology. Mathematical models for Surface Water Hydrology, Ciridni, T.A. et al. (Eds), Wiley, New York.

O'Connell, P.E., 1980. Real time hydrological forecasting and control, Rep. Inst Hydrol., Wallingford.

Prasad, R., 1967. A non linear hydrologic system response model. J. Hydraul. Div., ASCE HY4.

Rosenblueth, A. and N. Wiener, 1945. Role of models in science. Philosophy of Science. 7(4):316-321.

Sherman, L.K., 1932. Streamflow from rainfall by the unit-graph method. Engineering News Record, 108: 501-505.

Singh, V.P., 1988. Hydrologic Systems. Volume 1: Rainfall-runoff modelling. Prentice Hall, New Jersey.

Tikhonov, A.N., 1963a. Solution of incorrectly formulated problems and the regularization method. Sov. Math. 4, 1035-1038.

Tikhonov, A.N., 1963b. Regularization of incorrectly posed problems. Sov. Math. 4, 1624-1627.

Todini, E. and Wallis, J.R., 1978. A real-time rainfall-runoff model for an on-line flood warning system. Proc. AGU Chapman Conf. Appl. Kalman Filtering Tech. Hydrol., Hydraul., Water Resour., Dept Civ Eng University of Pittsburgh, Pa.

Todini, E., 1978. Mutually interactive state parameter (MISP) estimation. Proc. AGU Chapman Conf. Appl. Kalman Filtering Tech. Hydrol., Hydraul., Water Resour., Dept Civ Eng University of Pittsburgh, Pa.

Todini, E., 1988. Rainfall-runoff modelling – past, present and future. Journal of Hydrology 100: 341-352.

UNESCO, 1985. Teaching aids in hydrology, Universitaires de France, Vendome.

Vörösmarty, C.J., Moore, B., Grace, A.L., et al., 1989. Continental scale models of water balance and fluvial transport: an application to South America. Global Biogeochem. Cycles 3, 241–265.

WMO, 1975. Intercomparison of conceptual models used in operational hydrological forecasting. Oper. Hydrol. Report 7, WMO No 429, Geneva.

Wood, E.F. and O'Connell, P.E., 1985. Real-time forecasting. Hydrological Forecasting, Anderson, M.G. and Burt, T.P. (Eds).

Wood, E.F., 1980. Recent developments in real-time forecasting/control of water resources systems (Vansteenkiste, G.E. ed), Amsterdam: North Holland Publishing.

Wood, E.F., Lettenmaier, D.P., Zartarian, V.G., 1992. A land-surface hydrology parameterisation with subgrid variability for general circulation models. J. Geophys. Res. 97, D3, 2717-2728.

CHAPTER 2
TIME SERIES ANALYSIS AND STOCHASTIC MODELLING

2.1 INTRODUCTION

Methods and procedures of time series analysis and stochastic modelling will be discussed in the chapter, while the remaining chapters of this monograph deal with problems and approaches used in modelling hydrologic systems and components. In general, they describe the physical processes involved in the movement of water onto, over, and through the soil surface. Quite often the hydrologic problems we face do not require a detailed discussion of the physical process, but only a time series representation of these processes. Stochastic models may be used to represent, in simplified form, these hydrologic time series.

Unlike the models that to be discussed in the remaining chapters, stochastic modelling places emphasis on the statistical characteristics of hydrologic processes. Some background in probability and statistics is necessary to fully understand this chapter. However, references and examples throughout the chapter should give readers with a more limited background an appreciation of the role of stochastic models in hydrology.

The material presented in this chapter can be divided into four major parts. The first part is a discussion of the statistical properties and components of a time series. The next part of the chapter is a discussion of the methods for identifying and modelling of different components of a hydrologic time series. The third part of the chapter is a discussion of different kinds of stochastic models that are available. The last part of the chapters is a presentation of the application fields of stochastic models.

Since this chapter concentrates on the basic concepts of stochastic processes and not on models of specific processes, details of such models may not be described. Many such models are described in the listed references.

2.2 TIME SERIES

The measurements or numerical values of any variable that changes with time constitute a time series. In many instances, the pattern of changes can be ascribed to an obvious cause and is readily understood and explained, but if there are several causes for variation in the time series values, it becomes difficult to identify the several individual effects. In Fig.2.1, the top graph shows a series of observations changing with time along the abscissa; the ordinate axis represents the changing values of y with time, t. From visual inspection of the series, there are three discernible features in the pattern of the observations. Firstly, there is a regular gradual overall increase in the size of values; this trend, plotted as a separate component $y_1(t)$, indicates a linear increase in the average size of y with time. The second obvious regular pattern in the composite series is a cyclical variation, represented separately by $y_2(t)$, the periodic component. The third notable feature of the series may be considered the most outstanding, the single high peak half way along the series. This typically results from a rare catastrophic event which does not from part of a recognisable pattern. The definition of the function $y_3(t)$ needs very careful consideration and may not be possible. The remaining hidden

feature of the series is the random stochastic component, $y_4(t)$, which represents an irregular but continuing variation within the measured values and may have some persistence. It may be due to instrumental of observational sampling errors or it may come from random unexplainable fluctuations in a natural physical process. A time series is said to be a random or stochastic process if it contains a stochastic component. Therefore, most hydrologic time series may be thought of as stochastic processes since they contain both deterministic and stochastic components. If a time series contains only random/stochastic component is said to be a purely random or stochastic process.

The complete observed series, y(t), can therefore be expressed by:

$$y(t) = y_1(t) + y_2(t) + y_3(t) + y_4(t) \tag{2.1}$$

The first two terms are deterministic in form and can be identified and quantified fairly easily; the last two are stochastic with major random elements, and some minor persistence effects, less easily identified and quantified.

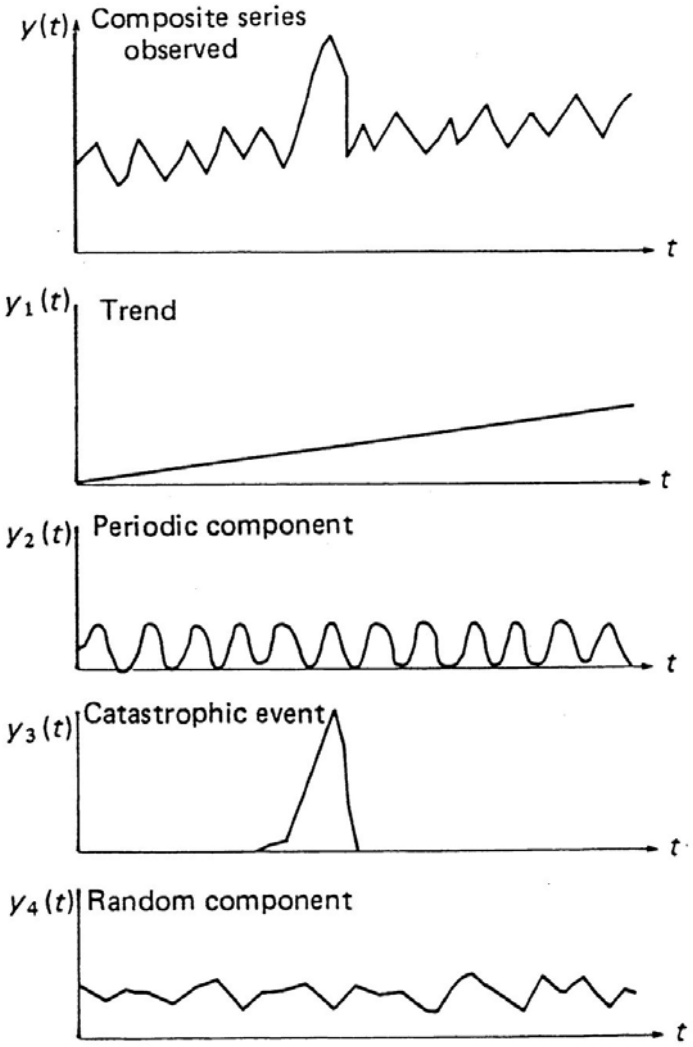


Fig.2.1 The time series components.

2.3 PROPERTIES OF TIME SERIES

The purpose of a stochastic model is to represent important statistical properties of one or more time series. Indeed, different types of stochastic models are often studied in terms of the statistical properties of time series they generate. Examples of these properties include: trend, serial correlation, covariance, cross-correlation, etc. Therefore, before reviewing the different types of stochastic models used in hydrology, some distribution properties of stochastic processes will be discussed. The following basic statistics are usually used for expressing the properties/characteristics of a time series.

Name	*Sample estimation*	*Notation for population*	
Mean	$E(X_t) = \overline{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$	μ	(2.2)
Variance	$S^2 = \frac{1}{n-1}\sum_{t=1}^{n}(X_t - \overline{X})^2$	σ^2	(2.3)
Covariance	$\mathrm{cov}(X_t, X_{t+L}) = \frac{1}{n-L}\sum_{t=1}^{n-L}(X_t - \overline{X})(X_{t+L} - \overline{X})$	λ_L	(2.4)

Where L is the time lag.

Stationary time series

If the statistics of the sample (mean, variance, covariance, etc.) as calculated by equations (2.2)-(2.4) are not functions of the *timing* or the *length* of the sample, then the time series is said to be stationary to the second order moment, weekly stationary, or stationary in the broad sense. Mathematically one can write as:

$$E(X_t) = \mu$$
$$Var(X_t) = \sigma^2$$
$$Cov(X_t, X_{t+L}) = \lambda_L$$

In hydrology, moments of the third and higher orders are rarely considered because of the unreliability of their estimates. Second order stationarity, also called covariance stationarity, is usually sufficient in hydrology. A process is strictly stationary when the distribution of X_t does not depend on time and when all simultaneous distributions of the random variables of the process are only dependent on their mutual time-lag. In another words, a process is said to be strictly stationary if its n-th (n for any integers) order moments do not depend on time and are dependent only on their time lag.

Nonstationary time series

If the values of the statistics of the sample (mean, variance, covariance, etc.) as calculated by equations (2.2)-(2.4) are dependent on the *timing* or the *length* of the sample, i.e. if a definite trend is discernible in the series, then it is a non-stationary series. Similarly, periodicity in a series means that it is non-stationary. Mathematically one can write as:

$$E(X_t) = \mu_t$$

$$Var(X_t) = \sigma_t^2$$
$$Cov(X_t, X_{t+L}) = \lambda_{L,t}$$

White noise time series

For a stationary ties series, if the process is purely random and stochastically independent, the time series is called a white noise series. Mathematically one can write as:

$$E(X_t) = \mu$$
$$Var(X_t) = \sigma^2$$
$$Cov(X_t, X_{t+L}) = 0 \qquad \text{for all } L \neq 0$$

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 weekly stationary, it is also strictly stationary, since the normal distribution is completely characterised by its first and second order moments.

2.4 ANALYSIS OF HYDROLOGIC TIME SERIES

Records of rainfall and river flow form suitable data sequences that can be studied by the methods of *time series analysis*. The tools of this specialized topic in mathematical statistics provide valuable assistance to engineers in solving problems involving the frequency of occurrences of major hydrological events. In particular, when only a relatively short data record is available, the formulation of a time series model of those data can enable long sequences of comparable data to be generated to provide the basis for better estimates of hydrological behaviour. In addition, the time series analysis of rainfall, evaporation, runoff and other sequential records of hydrological variables can assist in the evaluation of any irregularities in those records. Cross-correlation of different hydrological time series may help in the understanding of hydrological processes.

Tasks of time series analysis include:

(1) identification of the several components of a time series,
(2) mathematical description (modelling) different components identified.

If a hydrological time series is represented by X_1, X_2, X_3, ..., X_t, ..., then symbolically, one can represent the structure of the X_t by:

$$X_t \Leftrightarrow [T_t, P_t, E_t]$$

where T_t is the trend component, P_t is the periodic component and E_t is the stochastic component. The first two components are specific deterministic features and contain no element of randomness. The third, stochastic, component contains both random

fluctuations and the self-correlated persistence within the data series. These three components form a basic model for time series analysis.

The aims of time series analysis include but not limited to:

(1) description and understanding of the mechanism,
(2) Monte-Carlo simulation,
(3) forecasting future evolution,

Basic to stochastic analysis is the assumption that the process is stationary. The modelling of a time series is much easier if it is stationary, so identification, quantification and removal of any non-stationary components in a data series is undertaken, leaving a stationary series to be modelled.

2.4.1 Trend component

This may be caused by long-term climatic change or, in river flow, by gradual changes in a catchment's response to rainfall owing to land use changes. Sometimes, the presence of a trend cannot be readily identified.

Methods of trend identification:

Different statistical methods, both nonparametric tests and parametric tests, for identifying trend in time-series are available in the literature. Two commonly used methods for identifying the trend are discussed briefly in this section.

(1) Mann-Kendall test

The test uses the raw (un-smoothed) hydrologic data to detect possible trends. The Kendall statistic was originally devised by Mann (1945) as a non-parametric test for trend. Later the exact distribution of the test statistic was derived by Kendall (1975).

The Mann-Kendall test is based on the test statistic S defined as follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}(x_j - x_i) \tag{2.5}$$

where the x_j are the sequential data values, n is the length of the data set, and

$$\operatorname{sgn}(\theta) = \begin{cases} 1 & \text{if } \theta > 0 \\ 0 & \text{if } \theta = 0 \\ -1 & \text{if } \theta < 0 \end{cases} \tag{2.6}$$

Mann (1945) and Kendall (1975) have documented that when $n \geq 8$, the statistic S is approximately normally distributed with the mean and the variance as follows:

$$E(S) = 0 \tag{2.7}$$

$$V(S) = \frac{n(n-1)(2n+5) - \sum_{p=1}^{q} t_p(t_p-1)(2t_p+5)}{18} \tag{2.8}$$

where n = number of data
t_p = the number of ties for the pth value (number of data in the pth group)
q = the number of tied values (number of groups with equal values/ties)

The standardised Mann-Kendall test statistic Z_{MK} is computed by

$$Z_{MK} = \begin{cases} \dfrac{S-1}{\sqrt{Var(s)}} & S > 0 \\ 0 & S = 0 \\ \dfrac{S+1}{\sqrt{Var(s)}} & S < 0 \end{cases} \tag{2.9}$$

The standardised MK statistic Z follows the standard normal distribution with mean of zero and variance of one.
The hypothesis that there has not trend will be rejected if

$$|Z_{MK}| > Z_{1-\alpha/2} \tag{2.10}$$

where $Z_{1-\alpha/2}$ is the value read from a standard normal distribution table with α being the significance level of the test.

(2) Linear regression method

Linear regression method can be used to identify if there exists a linear trend in a hydrologic time series. The procedure consists of two steps, fitting a linear regression equation with the time T as independent variable and the hydrologic data, Y as dependent variable, i.e.

$$Y = \alpha + \beta \cdot T \tag{2.11}$$

and testing the statistical significance of the regression coefficient β.

Test of hypothesis concerning β can be made by noting that $(\beta - \beta_o)/S_\beta$ has t distribution with n-2 degrees of freedom. Thus the hypothesis $H_o : \beta = \beta_o$ versus $H_a : \beta \neq \beta_o$ is tested by computing

$$t = \frac{\beta - \beta_o}{S_\beta} \tag{2.12}$$

where S_β is the standard deviation of the coefficient β with

$$S_\beta = S \Big/ \sqrt{\sum_{i=1}^{n} (T_i - \overline{T})^2} \tag{2.13}$$

and

$$S = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \qquad (2.14)$$

where S is the standard error of the regression, Y_i and $\hat{Y}_i$ are observed and estimated hydrologic variable from the regression equation, respectively.

The hypothesis H_o, i.e. no trend, is rejected if $|t| > t_{1-\alpha/2,\,n-2}$

Models for trend:

The shape of the trend depends on the background of the phenomenon studied. Any smooth trend that is discernible may be quantified and then subtracted from the sample series. Common models for trend may take the following forms:

$T_t = a + bt$ (a linear trend, as in Fig.2.1) (2.15)

or

$T_t = a + bt + ct^2 + dt^3 + ...$ (a non-linear trend) (2.16)

The coefficients a, b, c, d, ... are usually evaluated by least-squares fitting. The number of terms required in a polynomial trend being primarily imposed by the interpretation of the studied phenomenon. The number of terms is usually based on statistical analysis, which determines the terms contributing significantly to the description and the interpretation of the time series. Restriction is made to the significant terms because of the principle of parsimony concerning the number of unknown parameters (constants) used in the model. One wishes to use as small a number of parameters as possible, because in most cases the addition of a complementary parameter decreases the accuracy of the other parameters. Also prediction- and control procedures are negatively influenced by an exaggerated number of parameters. This principle of parsimony is not only important with respect to the selection of the trend function but also with respect to other parts of the model.

2.4.2 Periodic component

In most annual series of data, there is no cyclical variation in the annual observations, but in the sequences of monthly data distinct periodic seasonal effects are at once apparent. The existence of periodic components may be investigated quantitatively by (1) Fourier analysis, (2) spectral analysis, and (3) autocorrelation analysis. Of which, the autocorrelation analysis method is widely used by hydrologists and will be discussed briefly in this section.

Identification of periodic component by autocorrelation analysis:

The procedure consists of two steps, calculating the autocorrelation coefficients and testing their statistical significance. For a series of data, X_t, the autocorrelation coefficient r_L between X_t and X_{t+L} are calculated and plotted against values of L (known as the lag), for all pairs of data L time units apart in the series:

$$r_L = \frac{1}{n-L}\sum_{t=1}^{n-L}(X_t - \overline{X})(X_{t+L} - \overline{X}) \Big/ \frac{1}{n}\sum_{t=1}^{n}(X_t - \overline{X})^2 \tag{2.17}$$

where $\overline{X}$ is the mean of the sample of n values of X_t and L is usually taken for values from zero up to n/4. A plot of r_L versus L forms the correlogram. The characteristics of a time series can be seen from the correlogram. Examples of correlograms are given in Fig.2.2. Calculation of equation (2.17) for different L gives the following cases:

- If L = 0, $r_L = 1$. That is, the correlation of an observation with itself is one.

- If $r_L \approx 0$ for all L ≠ 0, the process is said to be a purely random process. This indicates that the observations are linearly independent of each other. The correlogram for such a complete random time series is shown in Fig.2.2(a).

- If $r_L \neq 0$ for some L ≠ 0, but after L > τ, then $r_\tau \approx 0$, the time series is still referred to as simply a random one (not purely random) since it has a 'memory' up to L = τ. When $r_\tau \approx 0$, the process is said to have no memory for what occurred prior to time t-τ. The correlogram for such a non-independent stochastic process is shown in Fig.2.2(b). This is representative of an auto regressive process. Typically, such a correlogram could be produced from a series described by the Autoregressive model:

  $$X_t = a_1 X_{t-1} + a_2 X_{t-2} + a_3 X_{t-3} + \cdots + \varepsilon_t \tag{2.18}$$

  where a_i are related to the autocorrelation coefficients r_i and ε_t is a random independent element.

- In the case of data containing a cyclic (deterministic) component, then $r_L \neq 0$ for all L ≠ 0, the correlogram would appear as in Fig.2.2(c). Where T is the period of the cycle.

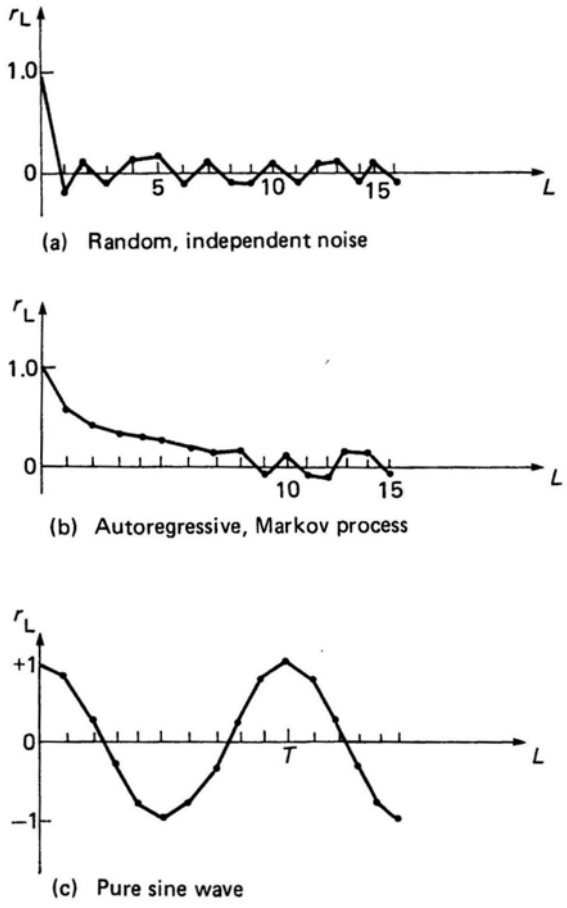


Fig.2.2 Examples of correlograms.

Modelling of periodic component:

A periodic function P_t is a function such that

$P_{t+T} = P_t$ for all t

The smallest value of T is called the period. The dimension of T is time, T thus being a number of time-units (years, months, days or hours, etc.) and we also have

$P_{t+nT} = P_t$ for all t and for all integer n.

The frequency is defined as the number of periods per time-unit:

$$frequency = \frac{1}{period}$$

Trigonometric functions are simple periodic functions. For example,

$\alpha \sin(\omega t + \beta)$

has a period of $2\pi/\omega$, because

$\alpha \sin[\omega(t+2\pi/\omega)+\beta] = \alpha \sin(\omega t+2\pi+\beta) = \alpha \sin(\omega t + \beta)$

The pulsation or angular frequency is defined as

$$\omega = \frac{2\pi}{period} = 2\pi \cdot frequency$$

the constant α is termed the amplitude and β the phase (with respect to the origin) of the sine-function.

A simple model for the periodic component may be defined as (for more discussions refer to the literature of Time Series Analysis):

$$P_t = m + C\sin(2\pi t / T) \tag{2.19}$$

where C is the amplitude of the sine wave about a level m and of wavelength T.

The serial (auto) correlation coefficients for such a P_t are given by:

$$r_L = \cos(2\pi L / T) \tag{2.20}$$

The cosine curve repeats every T time units throughout the correlogram with $r_L = 1$ for L = 0, T, 2T, 3T, …. Thus periodicities in a time series are exposed by regular cycles in the corresponding correlograms.

Once the significant periodicities, P_t, have been identified and quantified by μ_t (the means) and σ_t (the standard deviations) they can be removed from the original times series along with any trend, T_t, so that a new series of data, E_t, is formed:

$$E_t = \frac{X_t - T_t - m_t}{s_t} \tag{2.21}$$

Simple models for periodic component in hydrology can be seen in the literature. For example, in many regions, typical monthly potential evapotranspiration variation during the year can be modelled more or less by a sinusoidal function, with a couple of parameters to tune the annual mean and the amplitude (Xu and Vandewiele, 1995).

This behavior leads to the idea to model ep_t by a truncated Fourier series

$$ep_t = \{a + b\sin[(2\pi/12)(t-c)]\}^+$$

where again t is time in month. The plus sign at the end is necessary for avoiding negative values of ep_t which otherwise may occur in rare cases. Again parameters a, b and c are characteristics of the basin.

2.4.3 Stochastic component

E_t represents the remaining stochastic component of the time series free from non-stationary trend and periodicity and usually taken to be sufficiently stationary for the next stage in simple time series analysis. This E_t component is analysed to explain and

quantify any persistence (serial (auto) correlation) in the data and any residual independent randomness. It is first standardized by:

$$Z_t = \frac{E_t - \overline{E}}{s_E} \tag{2.22}$$

where $\overline{E}$ and s_E are the mean and standard deviation of the E_t series. The series, Z_t, then has zero mean and unit standard deviation. The autocorrelation coefficients of Z_t are calculated and the resultant correlogram is examined for evidence and recognition of a correlation and/or random structure.

For example, in Fig.2.3a for a monthly flow, the correlogram of the Z_t stationary series (with the periodicities removed) has distinctive features that can be recognised. Comparing it with Fig.2.2, the Z_t correlogram resembles that of an auto regressive (Markov) process. For a first order Markov model

$$Z_t = r_1 Z_{t-1} + e_t \tag{2.23}$$

where r_1 is the autocorrelation coefficient of lag 1 of the Z_t series and e_t is a random independent residual. A series of the residuals e_t may then be formed from the Z_t series and its known lag 1 autocorrelation coefficient, r_1:

$$e_t = Z_t - r_1 Z_{t-1} \tag{2.24}$$

The correlogram of residuals is finally computed and drawn (Fig.2.3b). For this data this resembles the correlogram of 'white noise', i.e. independently distributed random values. If there are still signs of autoregression in the e_t correlogram, a second-order Markov model is tried, and the order is increased until a random e_t correlogram is obtained. The frequency distribution diagram of the first order e_t values (Fig.2.3c) demonstrates an approximate approach to the normal (Gussian) distribution.

At this stage, the final definition of the recognisable components of the time series has been accomplished including the distribution of the random residuals. As part of the analysis, the fitted models should be tested by the accepted statistical methods applied to times series. Once the models have been formulated and quantified to satisfactory confidence limits, the total mathematical representation of the time series can be used for solving hydrological problems by synthesizing non-historic data series having the same statistical properties as the original data series.

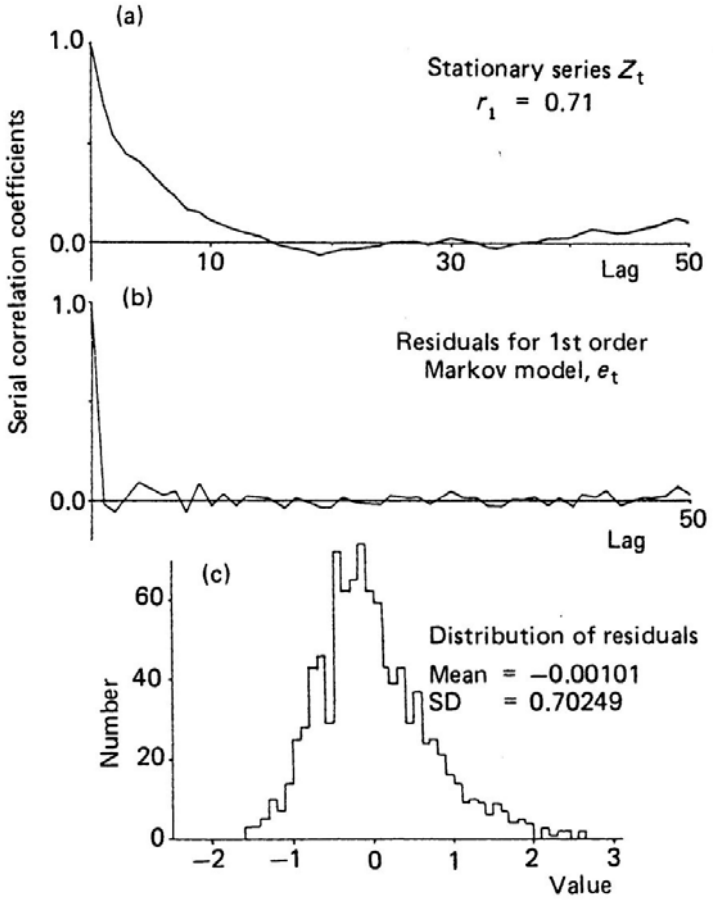


Fig.2.3 River Thames at Teddingtom Weis (82 years of monthly flows, from Shaw, 1988)

2.5 TIME SERIES SYNTHESIS

The production of a synthetic data series simply reverses the procedure of the time series analysis. First, for as many data items as are required, a comparable sequence of random numbers, drawn from the e_t distribution, is generated using a standard computer package. Second, the corresponding synthetic Z_t values are recursively calculated using equation 2.23 (starting the series with the last value of the historic Z_t series as the Z_{t-1} value). Third, the E_t series then derives from equation 2.22 in reverse:

$$E_t = Z_t s_E + \overline{E} \tag{2.25}$$

The periodic component P_t represented by m_t and s_t for time period *t* is then added to the E_t values to give:

$$X_t = T_t + E_t s_t + m_t \qquad \text{(from equation 2.21)} \tag{2.26}$$

The incorporation of the trend component T_t then produces a synthetic series of X_t having similar statistical properties to the historic data series.

2.6 SOME STOCHASTIC MODELS

Ultimately design decisions must be based on a stochastic model or a combination of stochastic and deterministic models. This is because any system must be designed to operate in the future. Deterministic models are not available for generating future watershed inputs in the form of precipitation, solar radiation, etc., nor is it likely that deterministic models for these inputs will be available in the near future. Stochastic models must be used for these inputs.

2.6.1 Purely random stochastic models

Possibly the simplest stochastic process to model is where the events can be assumed to occur at discrete times with the time between events constants, 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. For example, random observations for any normal distribution can be generated from the relationship,

$$y = \sigma R_N + \mu \tag{2.27}$$

where R_N is a standard random normal deviate (i.e. a random observation from a standard normal distribution) and μ and σ are the parameters of the desired normal distribution of Y. Computer routines are available for generating standard random normal distribution.

2.6.2 Autoregressive models

Where persistence is present, synthetic sequences cannot be constructed by taking a succession of sample values from a probability distribution, since this will not take account of the relation between each number of sequences and those preceding it. Consider a second order stationary time series, such as an annual time series, made up of a deterministic part and a random part. The deterministic part is selected so as to reflect the persistence effect, while it is assumed that the random part has a zero mean and a constant variance. One of the models to simulate such a series is the Autoregressive model. The general form of an autoregressive model is

$$(y_t - \mu) = \beta_1(y_{t-1} - \mu) + \beta_2(y_{t-2} - \mu) + \ldots + \beta_k(y_{t-k} - \mu) + \varepsilon_t \tag{2.28}$$

where μ is mean value of the series, β is the regression coefficient, the $\{y_1, y_2, \ldots, y_t, \ldots\}$ is the observed sequence and the random variables ε_t are usually assumed to be Normally and independently distributed with zero mean and variance σ_ε^2. In order to determining the order k of autoregression required to describe the persistence adequately, it is necessary to estimate k+2 parameters: $\beta_1, \beta_2, \ldots \beta_k, \mu$ and the variance of residuals σ_ε^2. Efficient methods for estimating these parameters have been described by Kendall and Stuart (1968), Jenkins and Watts (1968), and illustrated in the hydrological context by Carlson et al, (1970), see also Clarke (1973, page 44).

The first order autoregression

$$y_t - \mu = \beta_1(y_{t-1} - \mu) + \varepsilon_t \tag{2.29}$$

has found particular application in hydrology. When equation (2.29) is used to model annual discharge series, the model states that the value of y in one time period is dependent only on the value of y in the preceding time period plus a random component. It is also assumed that ε_t is independent of y_t.

Equation (2.29) is the well-known first order Markov Model in the literature. It has three parameters to be estimated: μ, β_1, and σ_ε^2.

For the moment method of parameter estimation, parameter μ can be computed from the time series as the arithmetic mean of the observed data.

As for β_1, the Yule-Walker equation (Delleur, 1991) shows that

$$\rho_k = \sum_{j=1}^{P} \beta_j \rho_{k-j}, \qquad k > 0 \tag{2.30}$$

the above equation, written for k = 1, 2, …, yields a set of equations. Where ρ_k is the autocorrelation coefficient for time lag k. As the autocorrelation coefficients ρ_1, ρ_2, …, can be estimated from the data using equation (2.17), these equations can be solved for the autoregressive parameters β_1, β_2, …, β_p. This is the estimation of parameters by the method of moments. For example, for the first order autoregressive model, AR(1), the Yule-Walker equations yield

$$\rho_1 = \beta_1 \cdot \rho_o = \beta_1 \text{ since } \rho_o = 1 \tag{2.31}$$

in the similarly way we can derive the equations for computing β_1 and β_2 for the AR(2) model as

$$\beta_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2} \qquad \beta_2 = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} \tag{2.32}$$

It can be shown that σ_ε^2 is related to σ_y^2 (the variance of the y_t series) by:

$$\sigma_\varepsilon^2 = \sigma_y^2(1-\beta_1^2) \tag{2.33}$$

If the distribution of y is N(μ_y, σ_y^2) then distribution of ε is N(0, σ_ε^2). Random values y_t can now be generated by selecting ε_t randomly from a N(0, σ_ε^2) distribution. If z is N(0,1) then $z\sigma_\varepsilon$ or $z\sigma_y\sqrt{1-\beta_1^2}$ is N(0, σ_ε^2). Thus, a model for generating Y's that are N(μ_y, σ_y^2) and follow the first order Markov model is

$$y_t = \mu_y + \beta_1(y_{t-1} - \mu_y) + z_t\sigma_y\sqrt{(1-\beta_1^2)} \tag{2.34}$$

The procedure for generating a value for y_t is:

(1) estimate μ_y, σ_y, and β_1 by $\bar{y}, s_x$, and r_1 (eq.2.17) respectively,
(2) select a z_t at random from a N(0, 1) distribution, and
(3) calculate y_t by eq. (2.34) based on $\bar{y}, s_x$, and β_1, and y_{t-1}.

The first value of y_t, i.e. y_1, might be selected at random from a N(μ_y, σ_y^2). To eliminate the effect of y_1 on the generated sequence, the first 50 or 100 generated values might be discarded.

Equation (2.34) has been widely used for generating annual runoff from watersheds (Fiering and Jackon, 1971, see also Haan, 1976).

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

$$q_{jan} = \bar{q}_{jan} + b_{jan}(q_{dec} - \bar{q}_{dec}) + \varepsilon_{jan}$$
$$q_{feb} = \bar{q}_{feb} + b_{feb}(q_{jan} - \bar{q}_{jan}) + \varepsilon_{feb}$$
… …

Fig.2.4 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, ..., 12 (Jan, Feb, ... Dec) and when j = 12, j+1 = 1 = Jan (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 months flow q_j, by the equation:

$$\hat{q}_{j+1} = \bar{q}_{j+1} + b_j(q_j - \bar{q}_j) \qquad (2.35)$$

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

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

where s_{j+1} is the standard deviation of the 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 written as

$$\hat{q}_{j+1,i} = \overline{q}_{j+1} + b_j(q_{j,i-1} - \overline{q}_j) + Z_{j+1,i} \cdot s_{j+1}\sqrt{(1-r_j^2)} \tag{2.36}$$

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

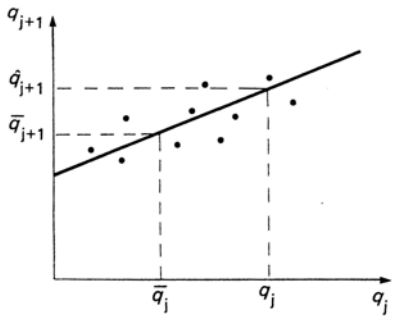


Fig.2.4 Thomas-Fiering model

The procedure for using the model is as follows:

(1) For each month, j = 1, 2, … 12, calculate

(a) the mean flow $\overline{q}_j = \frac{1}{n}\sum_i q_{j,i}$; $\quad (i = j, 12+j, 24+j, \cdots)$

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

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

$$r_j = \frac{\sum_{i=1} (q_{j,i} - \overline{q}_j)(q_{j+1,i} - \overline{q}_{j+1})}{\sqrt{\sum_i (q_{j,i} - \overline{q}_j)^2 \sum_i (q_{j+1,i} - \overline{q}_{j+1})^2}}$$

(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 set of twelve regression equations

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

where Z is a random Normal deviate N(0, 1).

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

2.6.4 Moving average models

The model form:

The moving average has frequently been used to smooth various types of hydrologic time series such as daily or weekly air temperature, evaporation rates, wind speed, etc. The moving average process used in the stochastic generation hydrologic data is somewhat different. In this use, the moving average process describes the deviations of a sequence of events from their mean value.

A process $\{x_t\}$ defined as

$$x_t = e_t + \phi_1 e_{t-1} + \phi_2 e_{t-2} + \ldots + \phi_q e_{t-q} \tag{2.37}$$

where $\{e_t\}$ is an uncorrelated stationary process, is called a moving average process of order q, denoted MA(q)-process.
It can also be written as

$$x_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q} \tag{2.38}$$

with $\phi_1 = -\theta_1$, $\phi_2 = -\theta_2$, ..., $\phi_q = -\theta_q$.

The properties of the moving average process:

The autocovariance of the process is obtained by forming the product $x_i \cdot x_{t-k}$ and taking the expectation:

$$\gamma_k = E\left[\left(e_t - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}\right)\left(e_{t-k} - \theta_1 e_{t-k-1} - \ldots - \theta_q e_{t-k-q}\right)\right] \tag{2.39}$$

For k = 0 we obtain the variance of the process

$$\sigma^2 = \gamma_o = \sigma_e^2(1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2) = \sigma_e^2 \sum_{j=0}^{q} \theta_j^2 \tag{2.40}$$

with the convention $\theta_o = -1$

$$\gamma_k = \sigma_e^2(-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \ldots + \theta_{q-k}\theta_q) = \sigma_e^2 \sum_{j=0}^{q-k} \theta_j\theta_{j+k} \qquad \text{for } k \le q \tag{2.41}$$

$$\gamma_k = 0 \qquad \text{for } k > q \tag{2.42}$$

The autocorrelation function is then

$$\rho_k = \frac{\gamma_k}{\gamma_o} = \frac{\sum_{j=0}^{q-k} \theta_j \theta_{j+k}}{\sum_{j=0}^{q} \theta_j^2}, \qquad k \le q$$
$$= 0 \qquad k > q \tag{2.43}$$

Equations (2.40) and (2.41) can be used for the estimation of the parameters by method of moments. For this purpose they are rewritten as follows:

$$\sigma_e^2 = \frac{\gamma_o}{1 + \theta_1^2 + \theta_2^2 + ... + \theta_q^2} \tag{2.44}$$

$$\theta_j = -(\frac{\gamma_j}{\sigma_e} - \theta_1\theta_{j+1} + \theta_2\theta_{j+2} + ... + \theta_{q-j}\theta_q) \tag{2.45}$$

Equ. (2.44) and (2.45) are used recursively. For example for the MA(1) model

$$x_t = e_t - \theta_1 e_{t-1} \tag{2.46}$$

we have

$$\hat{\sigma}_e^2 = \frac{\hat{\gamma}_o}{1 + \hat{\theta}_1^2} \qquad \hat{\theta}_1 = \frac{\hat{\gamma}_1}{\hat{\sigma}_e^2} \tag{2.47}$$

where $\hat{\gamma}_o$ and $\hat{\gamma}_1$ are estimates of the auto-covariance and computed from the data.

2.6.5 ARMA models

Model form:

In stochastic hydrology ARMA models are known as Auto-Regressive Moving Average (ARMA) models. They combine any direct autocorrelation properties of a data series with the smoothing effects of an updated running mean through the series. The two components of the model for a data series x_t, e.g. annual river flows, are described by:

Auto-regression (AR(p))

$$x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_p x_{t-p} + e_t \tag{2.48}$$

Moving average (MA(q))

$$x_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - ... - \theta_q e_{t-q} \tag{2.49}$$

where e_t are random numbers with zero mean and variance σ_e^2.

The Auto-regressive moving average (ARMA(p, q)) model is defined as:

$$x_t = \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_p x_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q} \tag{2.50}$$

One of the merits of the ARMA process is that, in general, it is possible to fit a model with a small number of parameters, i.e. $p+q$. This number is generally smaller than the number of parameters that would be necessary using either an AR model or a MA model. This principle is called the parsimony of parameters.
The first order model ARMA(1, 1) is:

$$x_t = \beta_1 x_{t-1} + e_t + \theta_1 e_{t-1} \qquad (-1 < \beta_1 < 1) \text{ and } (-1 < \theta_1 < 1) \tag{2.51}$$

Properties of ARMA model:

Consider in the ARMA(1, 1) model which has been used extensively in hydrology:

$$x_t = \beta_1 x_{t-1} + e_t + \theta_1 e_{t-1} \tag{2.52}$$

Multiplying both sides of (2.52) by x_{t-k}

$$x_{t-k} x_t = \beta_1 x_{t-k} x_{t-1} + x_{t-k} e_t + \theta_1 x_{t-k} e_{t-1}$$

and taking the expectation of both sides we obtain the autocovariance

$$\gamma_k = \beta_1 \gamma_{k-1} + E(x_{t-k}\, e_t) - \theta_1 E(x_{t-k}\, e_{t-1}) \tag{2.53}$$

For k = 0, equ (2.53) becomes

$$\gamma_o = \beta_1 \gamma_1 + E(x_t\, e_t) - \theta_1 E(x_t\, e_{t-1})$$

but

$$E(x_t e_t) = E\left[\beta_1 x_{t-1} e_t + e_t^2 + \theta_1 e_{t-1} e_t\right] = \sigma_e^2$$

and

$$\begin{aligned} E(x_t e_{t-1}) &= E\left[\beta_1 x_{t-1} e_{t-1} + e_t e_{t-1} + \theta_1 e_{t-1}^2\right] \\ &= E\beta_1 [x_{t-1} e_{t-1}] - \theta_1 \sigma_e^2 \\ &= (\beta_1 - \theta_1)\sigma_e^2 \end{aligned} \tag{2.54}$$

Thus

$$\gamma_o = \beta_1 \gamma_1 + \sigma_e^2 - \theta_1 (\beta_1 - \theta_1)\sigma_e^2 \tag{2.55}$$

For k = 1 equ (2.53) becomes

$$\gamma_1 = \beta_1 \gamma_0 + 0 - \theta_1 \sigma_e^2$$

Combining with the previous equation

$$\gamma_o = \beta_1^2\gamma_o - \beta_1\theta_1\sigma_e^2 + \sigma_e^2 - \theta_1(\beta_1 - \theta_1)\sigma_e^2$$

or

$$\gamma_o = \frac{1 + \theta_1^2 - 2\beta_1\theta_1}{1 - \beta_1^2} \tag{2.56}$$

and

$$\gamma_1 = \frac{(\beta_1 - \theta_1)(1 - \beta_1\theta_1)}{1 - \beta_1^2}\sigma_e^2 \tag{2.57}$$

For $k \geq 2$

$$\gamma_k = \beta_1\gamma_{k-1} \qquad k \geq 2 \tag{2.58}$$

the autocorrelation function (ACF) is obtained by dividing (2.56), (2.57) and (2.58) by γ to obtain

$$\rho_k = 1 \qquad k = 0 \tag{2.59a}$$

$$= \frac{(\beta_1 - \theta_1)(1 - \beta_1\theta_1)}{1 + \theta_1 - 2\beta_1\theta_1} \qquad k = 1 \tag{2.59b}$$

$$= \beta_1\gamma_{k-1} \qquad k \geq 2 \tag{2.59c}$$

Observe that the MA parameter θ_1 enters only in the expression for ρ_1. For ρ_2 and beyond the behaviour of the autocorrelation is identical to that of the AR(1) model.

Estimates of the parameters θ_1 and β_1 can be obtained from equations (2.59b) and (2.59c), since the serial (auto) correlation coefficients ρ_1 and ρ_2 can be computed from data. More efficient methods of estimating ARMA parameters are to be found in advanced texts (e.g. Box & Jenkins, 1970).

In general for an ARMA(p, q) model the autocovariance is

$$\gamma_k = \beta_1\gamma_{k-1} + \ldots + \beta_p\gamma_{k-p} + E[x_{t-k}e_t] - \theta_1 E[x_{t-k}e_{t-1}] - \ldots - \theta_q E[x_{t-k}e_{t-q}] \qquad k < q+1 \tag{2.60a}$$

$$\gamma_k = \beta_1\gamma_{k-1} + \ldots + \beta_p\gamma_{k-p} \qquad k \geq q+1 \tag{2.60b}$$

and the ACF is

$$\rho_k = \beta_1\rho_{k-1} + \ldots + \beta_p\rho_{k-p} \qquad k \geq q+1 \tag{2.61}$$

after lag q+1 the ACF tails off as for an AR(p) process. For the first q lags, the ACF depends on AR and MA parameters.

Hydrologic justification of ARMA models

A physical justification of ARMA models for annual streamflow simulation is as follows. Consider a watershed with annual precipitation X_t, infiltration aX_t and evapotranspiration bX_t. The surface runoff is $(1\text{-}a\text{-}b)X_t = dX_t$. (See Fig 2.5).

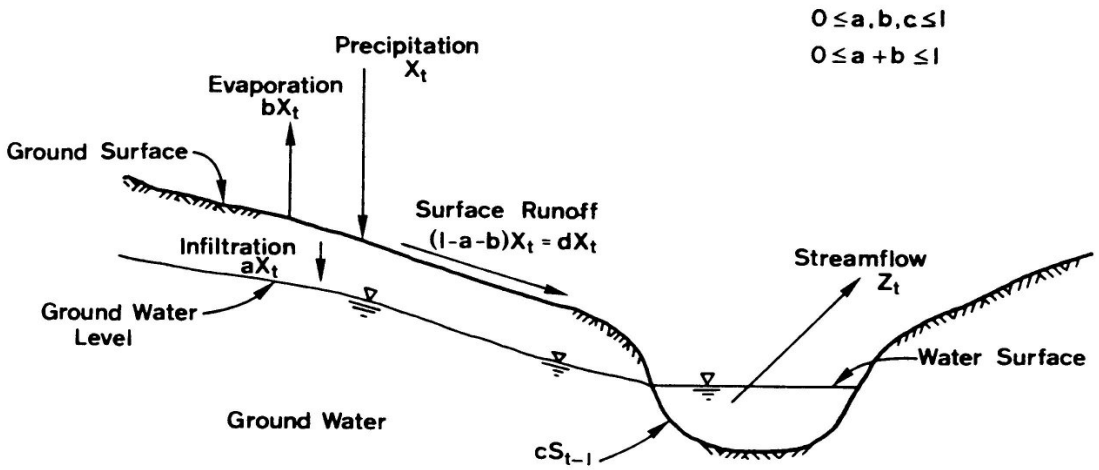


Fig.2.5 Conceptual representation of the precipitation-streamflow process after Salas and Smith (1980)

Let the groundwater contribution to the stream be cS_{t-1}.
Thus,

$$Z_t = cS_{t-1} + dX_t \tag{2.62}$$

The conservation of mass for the groundwater storage is

$$S_t = S_{t-1} + aX_t - cS_{t-1} \tag{2.63}$$

or

$$S_t = (1-c)S_{t-1} + aX_t \tag{2.64}$$

Rewriting (2.62)

$$Z_{t-1} = cS_{t-2} + dX_{t-1}$$

or

$$S_{t-2} = \frac{1}{c}Z_{t-1} + \frac{d}{c}X_{t-1} \tag{2.65}$$

and rewriting (2.64) as

$$S_{t-1} = (1-c)S_{t-2} + aX_{t-1} \tag{2.66}$$

Combining (2.62), (2.66) and (2.65) we obtain

$$Z_t = c(1-c)S_{t-2} + acX_{t-1} + dX_t$$

$$Z_t = (1-c)Z_{t-1} - d(1-c)X_{t-1} + acX_{t-1} + dX_t$$

$$Z_t = (1-c)Z_{t-1} + dX_t - [d(1-c) - ac]X_{t-1} \tag{2.67}$$

which has the form of an ARMA (1, 1), i.e. equation (2.52) model when the precipitation, X_t is an independent series and when (1-c) = β_1, d = 1, and [d(1-c)-ac)] = θ_1..

2.6.6 Daily data generation models

The synthetic generation of series of daily events is an extremely complicated problem for certain types of data. Data which can be considered nearly independent from one day to the next are not particularly difficult and can be handled by any of the previously described processes. However, daily processes such as temperature, solar energy, and streamflow have characteristics that are much more difficult to model. Streamflow, for example, is extremely difficult.

The high degree of persistence, due to the drainage of flood water from the channel system within which it has been stored, makes streamflow difficult to model on a daily basis. During the recession, correlation between the flow for period and that either preceding or following is very high. The magnitude of the autocorrelation (slope of the recession) is a function of many things such as the irregularity (roughness) of the channel, slope of the channel, size of the channel, temperature of the water, sediment content, and the amount and condition of vegetation on channel banks. Changes in these factors can cause the autocorrelation coefficients to vary from event to event, season to season and even year to year. Moreover, streamflow is made up of two components of entirely different character. One component is surface runoff which is a nonlinear response due to the high degree of control that solar energy, vegetation growth, evapotranspiration and soil moisture exercise on flow characteristics. The other component is groundwater flow which is much more linear in response because it acts primarily like drainage from one or more reservoirs. The magnitude of the different components varies considerably from one site to another. It can be entirely surface runoff, for example, where streams have small headwater catchments and are in soils of very low permeability, to entirely subsurface runoff such as is experienced in some sand soil or coastal plain soil areas.

These characteristics of streamflow make the synthetic generation of daily data extremely difficult. Few studies (Weiss, 1973, 1977, O'Connell, 1977; see also Haan et al., 1982), nevertheless, have been made to use shot-noise model to represent daily flow records as a stochastic process. Fitting of such a model to daily hydrologic data is quite complex and can be a laborious task. No details will be given here. In many respects, the best model of daily data may be obtained from catchment rainfall-runoff models as discussed in other chapters of this book.

2.6.7 Miscellaneous models

The models of some hydrologic processes are such that they cannot be classified into any of the previous categories. Several rainfall models fall into this group. Since these are quite important from the standpoint of stochastic models, they are mentioned here. However, since they were developed to model a specific process and are not general models of runoff process, they will not be described extensively. Readers who are interested in this specific model types, can easily found many examples in the literature with the keyword of 'rainfall models'.

2.7 THE USES OF STOCHASTIC MODELS

(1) To make predictions of frequencies of extreme events

Stochastic models have been used to make predictions about the frequency of occurrence of certain extreme events of interest to the hydrologist. Models such as that given by equation (2.29) are selected, and the residual ε_t is taken to be random variable with probability distribution whose parameters are specified. The parameters are estimated from data; so-called "synthetic" sequence $\{y_t\}$ can then be constructed, and the frequency with which the extreme event occurs in them can be taken as an estimate of the "true" frequency with which it would occur in the long run.

(2) For the investigation of system operating rules

A further use for synthetic sequences generated by stochastic models is in reservoir operation, such as the investigation of the suitability of proposed operating rules for the release of water from complex systems of interconnected reservoirs. By using the generated sequence as inputs to the reservoir system operated according to the proposed rules, the frequency with which demands fail to be met can be estimated. This may lead to revision of the proposed release rules; the modified rules may be tested by a similar procedure.

(3) To provide short-term forecasts

Stochastic models have been used to make forecasts. Given the values x_t, x_{t-1}, x_{t-2}, ...; y_t, y_{t-1}, y_{t-2}, ... assumed by the input and output variables up to time t, stochastic models have been constructed from this data for forecasting the output from the system at future times, t+1, t+2, ..., t+k, ... In statistical terminology, k is the lead-time of the forecast. Many stochastic models have a particular advantage for forecasting purposes in that they provide, as a by-product of the procedure for estimating model parameters, confidence limits for forecasts (i.e. a pair of values, one less than the forecast and one greater, such that there is a given probability P that these values will bracket the observed value of the variable at time t+k). Confidence limits therefore express the uncertainty in forecasts; the wider apart the confidence limits, the less reliable the forecast. Furthermore, the greater the lead-time k for which forecasts is required, the greater will be the width of the confidence interval, since the distant future in more uncertain than the immediate.

(4) To "extend" records of short duration, by correlation

Stochastic models have been used to "extend" records of basin discharge where this record is short. For example, suppose that it is required to estimate the instantaneous

peak discharge with a return period of T years (i.e. such that it would recur with frequency once in T years, in the long run). One approach to this problem is to examine the discharge record at the site for which the estimate is required, to abstract the maximum instantaneous discharge for each year of record, and to represent the distribution of annual maximum instantaneous discharge by a suitable probability density function. The abscissa, Y_o, say, that is exceeded by a proportion 1/T of the distribution then estimates the T-year flood.

It, however, frequently happens that the length of discharge record available is short, say ten years or fewer. On the other hand, a much longer record of discharge may be available for another gauging site, such that the peak discharges at the two sites are correlated. In certain circumstances, it is then permissible to represent the relation between the annual maximum discharges at the two sites by a regression equation and to use this fitted equation to estimate the annual maximum instantaneous discharges for the site with short record.

(5) To provide synthetic sequences of basin input

Suppose that the model has been developed for a system consisting of a basin with rainfall as input variable, streamflow as output variable. If a stochastic model were developed from which a synthetic sequence of rainfall could be generated having statistical properties resembling those of the historic rainfall sequence, the synthetic rainfall sequence could be used as input to the main model for transformation to the synthetic discharge sequence. The discharge so derived could then be examined for the frequency of extreme events.

This approach to the study of the frequency of extreme discharge events is essentially an alternative to that described in paragraph (1) above. In the latter, a synthetic sequence is derived from a stochastic model of the discharge alone; in the former, a synthetic discharge sequence is derived by using a model to convert a synthetic sequence of rainfall into discharge.

REFERENCES OF CHAPTER 2

Box, G.E.P. and Jenkins, G.M., 1970. Time series analysis, Forecasting, and control. Halden-day, San Francisco. 543p.

Clarke, R.T., 1973. A review of some mathematical models used in hydrology, with observations on their calibration and use. J. Hydrol. 19: 1-20.

Clarke, R.T., 1994. Statistical Modelling in Hydrology, John Wiley & Sons Ltd.

Delleur, J.W., 1991. Time Series Analysis Applied to Hydrology. V.U.B.-Hydrologie (19), Free University Brussels.

Fiering, M.B. and Jackson, B.B., 1971. Synthetic Streamflows. Water Resources Monograph 1, Amer. Geophys. Union, Washington, DC, 1-98.

Haan, C.T., 1976. Statistical Methods in Hydrology. The Iowa University Press, Ames, p378.

Haan, C.T., Johnson, H.P. and Brakenslek, D.L., 1982. Hydrologic Modeling of Small Watersheds. An ASAE Monograph Number 5 In a Series published by American Society of Agricultural Engineers.

Jenkins, G.M. and Watts, D.G., 1968. Spectral analysis and its applications, Holden Day, 525pp.

Kendall, M.G. and Stuart, A., 1968. The advanced theory of statistics. Vol3: design and Analysis, and time series, Hafner, 557pp.

Kendall, M.G., 1975. Rank Correlation Methods. Griffin, London.

Mann, H.B., 1945. Nonparametric tests against trend. Econometrica 13, 245-259.

O'Connell, P.E., Shot noise models in synthetic hydrology. In: Mathematical models for surface water hydrology by Ciriani, Maione and Wallis. John Wiley and Sons. Pp19-26.

Salas, J.D. and Smith, R.A., 1980. Physical basis of stochastic models of annual flows. Water Resources Research.

Shaw, E.M., 1988. *Hydrology in Practice 2nd edition*. Van Nostrand Reinhold (UK) Co. Ltd.

Weiss, G., 1973. Shot noise models for synthetic generation of multisite daily streamflow data. Symposium UNESCO World Meteorological Organisation, Intr. Assoc. of Hydrol. Sci., Madrid.

Weiss, G., 1977. Shot noise models for the generation of synthetic streamflow data. Water Resources Research 13(1).

Xu, C-Y and Vandewiele, G.L., 1995. Parsimonious monthly rainfall-runoff models for humid basins with different input requirements, Advance in Water Resources 18: 39-48.

CHAPTER 3
PRECIPITATION IN CATCHMENT MODELS

3.1 INTRODUCTION

In general precipitation is the largest quantity in the hydrological cycle and in a water balance equation. Precipitation is the main input to most hydrologic models. The accuracy of measurement and computation of precipitation from a network of stations determines to a considerable extent the reliability of water balance computations. No reliable water balance computation is possible with insufficient knowledge of the spatial rainfall patterns. The main reason for errors in the areal mean of precipitation is the high spatial-temporal variability of precipitation and the resulting complicated statistical structure of precipitation data.

The proposed use of a hydrologic model dictates the needed detail and complexity of precipitation input. Economic considerations usually determine whether the desired sampling detail is actually achieved. For example, data from a single standard rain-gage may be sufficient to determine average annual or seasonal rainfall on a small watershed. A single recording rain-gage may provide enough information to predict average annual erosion and surface water yield. A network of recording gages is needed to describe the variation of precipitation in time and space. Data from a network of recording gages may be needed to estimate flood peaks, erosion, and sedimentation from individual events, or spatial variability of runoff production. Other hydrologic measurements, like temperature, humidity, solar radiation, evapotranspiration, and antecedent soil moisture, may be needed as well as precipitation for accurate water balance calculations or accurate crop yield estimates.

For details concerning point precipitation measurement, calculation of areal precipitation, etc., refer to the course "*Catchment Hydrology*".

3.2 SEPARATION OF SNOWFALL FROM PRECIPITATION

Precipitation includes rainfall, snowfall, and other processes by which water falls to the land surface, such as hail and sleet. The first two forms constitute the major part of precipitation and are of importance in hydrologic models.

Snow and snowfall play a significant part in the hydrologic regime in many parts of the world. Snow has received attention as a water resource, primarily in the northern part of North America, Europe, and Asia. All of Sweden receives snow in hydrologically significant amounts. For hydrologic purposes, the water content is more important than depth, unless one is interested in the insulating properties of the snow as in soil freezing studies. Snow water equivalent and snow density are more useful for hydrologic modelling.

Whether precipitation falls as rain or snow can have a very significant influence on the estimation of runoff, especially for the spring flows. Model performance is therefore sensitive to decisions made concerning the form of precipitation. The problem with determination of the form of precipitation is usually solved in a rather simple manner in most modelling processes. The air temperature is accepted as a determining factor

meaning that snow accumulation starts as soon as the temperature is lower than a certain threshold value. It may be noted that some models use a fixed value of the threshold temperature, whereas others treat it as a calibration parameter. This method has been used by the U.S. Corps of Engineers (1956), Anderson (1973) and HBV model among others. According to an investigation made by the U.S. Corps of Engineers, the threshold value may vary between -1.7°C and +4.4°C when studying hourly values. An investigation of daily values made at the Lilla Tivsjön climate station in Sweden is shown in Fig.3.1 (Bergström, 1975). This investigation shows that the threshold value may vary between -2.5°C and +4.0°C when studying daily values.

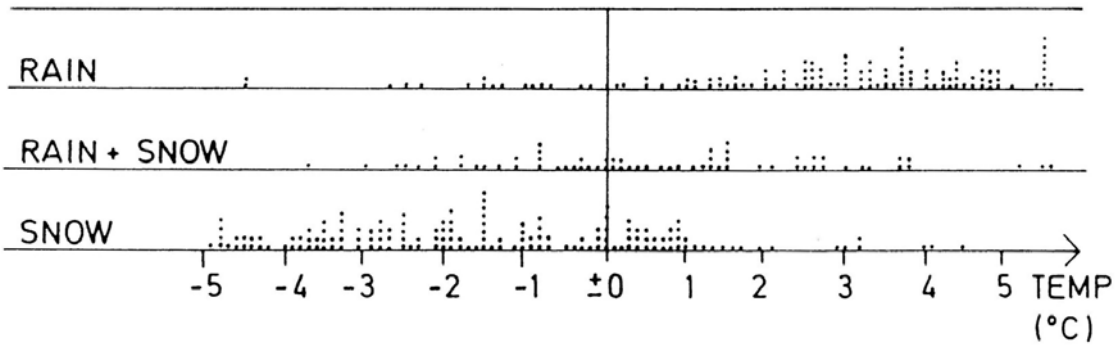


Fig.3.1. The observer's note on the form of precipitation related to mean daily temperature. Each point represents one day with precipitation.

Methods used in conceptual hydrological models for distinguishing the rainfall and snowfall are quite simple. Some examples of such methods are discussed as follows:

(1) In the HBV model and many others,

$$P_s = 0,\ P_r = P_t \quad \text{when } T_a \geq T_o \qquad (3.1)$$
$$P_r = 0,\ P_s = P_t \quad \text{when } T_a < T_o$$

Where

P_r = amount of precipitation in the form of rain (mm)
P_s = amount of precipitation in the form of snow (mm)
P_t = total precipitation (mm)
T_a = mean daily air temperature (°C)
T_o = Threshold temperature (°C)

(2) In the Hydrocomp (1969) model the division is based on the expression shown in equation (3.2).

$$P_s = 0,\ P_r = P_t \quad \text{when } T_o \geq (33^oF \approx 0.6\,^oC) \qquad (3.2)$$
$$P_r = 0,\ P_s = P_t \quad \text{when } T_o < (33^oF \approx 0.6\,^oC)$$

And T_o is calculated by

$$T_o = T_a - (T_a - T_d)(0.12 + 0.008T_a) \qquad (3.3)$$

where

T_d = dew-point temperature, and other notations have the same meaning as in equation 3.1.

(3) Willen (1971) and Moussavi (1988) used the following equation for estimating the form of precipitation in their runoff models:

$$\begin{aligned} &P_s = 0,\ P_r = P_t && \text{when } T_{min} \geq T_o \\ &P_r = 0,\ P_s = P_t && \text{when } T_{max} < T_o \\ &\text{percent rain} = [(t_{max} - t_o)/(t_{max} - t_{min})]\times 100 && \text{when } T_{min} \leq T_o \leq T_{max} \end{aligned} \qquad (3.4)$$

where: t_{max} is the daily maximum air temperature, t_{min} is the daily minimum air temperature, and other notations have the same meaning as before.

(4) Shih et al (1972) specify the division by the functions shown in equation (3.5)

$$\begin{aligned} &P_s = 0,\ P_r = P_t && \text{when } T_a \geq T_r \\ &P_r = 0,\ P_s = P_t && \text{when } T_a \leq T_s \\ &P_r = P_t\left[\frac{T_a - T_s}{T_r - T_s}\right] && \text{when } T_s \leq T_a \leq T_r \end{aligned} \qquad (3.5)$$

where

T_r = limiting temperature above which precipitation will be rain, e.g., 38^{o}F (3.3^{o}C)

T_s = limiting temperature below which precipitation will be snow, e.g., 30^{o}F (-1.1^{o}C). Other notations have the same meaning as before.

(5) Xu et al. (1996) used the following equation in the monthly snow and water balance model.

$$\begin{aligned} &P_S = p_t\left\{1 - \exp\left[-(T_a - T_r)/(T_r - T_s)\right]^2\right\}^+ && T_r \geq T_s \\ &P_r = P_t - P_s \end{aligned} \qquad (3.6)$$

where

T_r = threshold temperature above which precipitation will be rain (2^{o}C).

T_s = threshold temperature above which snowmelt process begins (-2^{o}C).

Other notations have the same meaning as before.

3.3 MODELLING OF SNOWMELT

Problems of snowmelt runoff modelling associated with the climatic and physiographic conditions of these regions are functions of data availability, regional characteristics, modelling approach, and model application. Many of these problems are common to all models and regions, whereas others are unique to specific models or regions. The more universal problems are generally associated with data constraints, whereas the more unique problems are associated with model formulation and the climatic and physiographic conditions of a region.

Most models of snowmelt use variations of the energy balance method pioneered by Wilson (1941) in which he outlined the sources of energy that cause snowmelt. In this section, the use of the energy balance method and its simplifications are first outlined, and secondly, the application of various techniques of snowmelt calculations as incorporated into currently used models is described.

3.3.1 Energy balance approach:

The energy balance approach uses a form of the energy balance equation for a snowpack that can be written as (US Army, 1956):

$$H = H_{sn} + H_{\ln} + H_c + H_e + H_g + H_p + H_q \qquad (3.7)$$

where

H = energy available for snowmelt (net heat transfer to snowpack from its environment).
H_{sn} = net shortwave radiation
H_{ln} = net longwave radiation
H_c = convective heat flux
H_e = latent heat flux
H_g = conduction of heat from the ground
H_p = heat content of rain drops
H_q = change in energy content of the snowpack.

If H is the total net change in energy, the melt M, is calculated as (Haan et al., 1982):

$$M = H/L_f \qquad (3.8)$$

where L_f is the latent heat of fusion of ice.

The use of energy balance technique results in a model which may be very close to being correct, but which may be unwieldy to use, except in very specialised, highly instrumented situations (Kuzmin, 1973; Haan et al., 1982). Among the variables necessary for a complete heat budget computation according to equation (3.7), can be mentioned:

-total solar radiation,
-albedo,
-longwave radiation balance (effective radiation)
-air temperature,

-air humidity,
-wind speed,
-temperature gradients in the soil and in snow,
-precipitation.

In addition to these variables some physical parameters governing heat exchange with the atmosphere, heat transfer within the snowpack, liquid water content in the snow and drainage of the snowpack, would have to be estimated. Limits on the availability of some of these data and on techniques to extrapolate point measurements to areal mean values have restricted most applications of equation (3.7) to snowmelt studies at a point or on small plots (Leavesley, 1989). A few basin scale models that use equation (3.7) are currently being developed and tested; these models include the Institute of Hydrology Model, IHDM (Morris, 1980) and the SHE model (Jonch-Clausen, 1979).

To work with limits imposed by data availability on the energy balance approach, many investigators have studied the relative importance of the various energy balance components, this greatly aids in simplifying the computations when the situation justifies it or more detailed data are not available. Various modified versions of equation (3.7) have been used. In most of these models, H_{sn}, H_{ln}, H_p and $\mathrm{H_q}$ are computed using measured data and the remaining components are parameterized and fitted, or are assumed to be negligible. Examples include the Precipitation and Runoff Modelling System, PRMS (Leavesley et al., 1983) and the Snowmelt Model, MELTMOD (Leaf & Brink, 1973).

3.3.2 Simplifications

Zuzel and Cox (1975) measured daily values of wind, air temperature, vapor pressure, net radiation, and melt at a point. They found that for an area with continuous snow cover, vapour pressure, net radiation, and wind run explained 78% of the variations in melt, whereas air temperature and net radiation explained 60%. Temperature alone had a coefficient of determination of 0.51 and net radiation was 0.40.

Raffelson (1974) investigated the energy balance of isolated snowdrifts in Wyoming during melt. He found the sensible and latent heat components were about the same size, and both substantially larger than the radiation component. O'Neill (1972) and Gray and O'Neill (1974) found that net radiation was the predominant energy source for snowmelt for the Canadian Prairies when the snow cover was continuous, supplying 93% of the melt energy. For non-continuous cover, advection of heat from bare ground to isolated drifts caused 44% of melt energy to be supplied by sensible heat transfer and 56% by net radiation. For an isolated drift, Cox and Zuzel (1976) found that 69% of the energy available for melt and evaporation came from sensible heat input. The Crops of Engineers (1960) assigned a constant value to shortwave radiation during rain periods. King and Molnau (1976) noted that temperature index methods seem to work well for calculating snowmelt during overcast periods, indicating that radiation was relatively unimportant during those periods. Kuzmin (1973) explored five different simplifications of the basic energy balance method. He found that the use of temperature was possible for plains when mean daily temperature was greater than 2°C.

In conceptual hydrologic models, emphasis has been put on determining snowmelt by use of air temperature or a temperature index because of the ease of obtaining air temperatures and because temperature is the most easily extrapolated meteorological variable.

3.3.3 Temperature index approach (degree-day method):

Usually, an air temperature index approach has the form of:

$$M = C_m(T_a - T_b) \tag{3.9}$$

where

M = snow melt (mm/day);

C_m = degree day coefficient (mm/day/^oC);

T_a = air temperature (^oC); and

T_b = a base temperature (^oC).

For most cases, T_b is assumed to be constant. It can either be determined by experience, e.g. Granger and Male (1977) used $T_b = 0$, or be estimated by model calibration.

C_m and T_a are assumed to integrate the effects of several of the individual energy-budget components in equation (3.7). This is a very broad assumption and is a source of error for a variety of conditions. To minimize this error, most temperature index models apply a number of adjustments to C_m and T_a. C_m is adjusted to incorporate knowledge of the relations between it and measurable spatial and temporal variations in basin and climate characteristics. Anderson (1973) allows C_m to vary from a minimum on December 21 to a maximum on June 21, using a sine curve. For an Iowa watershed with no forest cover, the C_m ranged from 7.3 to 3.6 mm/°C/day. McKay (1968) presents curves of degree day factors for a shallow prairie snowpack. Gartska (1944) noted a strong correspondence between cumulative runoff and cumulative degree-hours above 0 °C. This relationship seemed consistent within a storm, but varied between storms. King (1976) used the degree-day method on small watersheds in the Palouse Prairie. He used C_m as a function of cumulative degree-hours with good success. However, he found that different functions may be needed for basins with different aspects because of the rolling topography.

Bengtsson (1976) developed the idea of an equilibrium temperature to use in place of the base temperature. This is the temperature at which no net transfer of heat between the air and snow takes place. By equating the energy balance approach with degree-day factor, he found that C_m could be determined as a function of wind speed for a forested watershed and a function of solar radiation for nonforested areas

HBV Model (Bergström, 1976, 1995) and the Snowmelt Runoff Model, SRM (Martinec et al., 1983) use a different value of C_m in each basin zone depending on the vegetation characteristics of the zone. The HBV model holds each value of C_m constant for the entire melt season while SRM varies C_m as a function of snowpack density. In distributed models, different values of C_m for each basin zone or grid cell are used, they also vary the magnitude of C_m through the melt season to account for the effects of seasonal variation in day length on C_m. The Streamflow Synthesis and Reservoir Regulation Model, SSARR (US Army, 1975) uses an antecedent temperature index to adjust C_m seasonally.

In the literature, there are many different versions of equation (3.9) in which T_a is replaced by, e.g., maximum daily temperature, minimum daily temperature, and combinations of these variables. Examples are:

1) $M = C_m(T_{max} - T_b)$ (Martinec & Range, 1986)
2) $M = C_m T_{max}$ (Power, 1986)
3) $M = C_m (T_{max})^2$ (Woo, 1972)
4) $M = C_m (T_{max} - T_b) + b$ (Lang, 1984)
5) $M = C_m (T_{max} - T_{min}) + b$ (Moussavi, 1988)

where b is a model parameter.

A different temperature index equation was used by Xu et al. (1996) in the snow and water balance model, where snowmelt, M_t is considered as a function of temperature and the snow storage, sp_t.

$$M_t = sp_t \left\{1 - \exp[(T_a - T_s)/(T_r - T_s)]^2\right\}^+ \quad (3.10)$$

where T_r and T_s have the same meaning as in equ (3.6).

3.4 SNOWMELT IN HYDROLOGIC MODELS

Many hydrologic models include routines which will compute the amount of snowmelt by any one or combination of methods mentioned in previous sections. Very few models have been designed primarily as snowmelt models; normally, the snowmelt routine is added to the precipitation section where the water input to the main part of the hydrologic model is determined. Few examples of such snowmelt routine are discussed in this section.

Utah Water Research Laboratory Model

The flow chart for this hybrid model (Riley et al., 1969) is shown in Fig.3.2. This is a routine in a hybrid computer model and illustrates some of necessary steps in a mass budget of snow on the ground. This model has been used successfully in mountain snowpack situations, but there is nothing in its development which suggests it would not work on agricultural catchments.

Ohio State University Model

The Ohio State University Model (OSUM) (Fig.3.3) is derived from the Kentucky Watershed Model (KWM) (Ricca, 1972). The OSUM includes a snowmelt routine developed specially for agricultural watersheds and was tested on the Coshocton Watersheds. The model includes simplified versions of each of the energy balance terms and requires daily average dewpoint, wind run, solar radiation, maximum and minimum temperature.

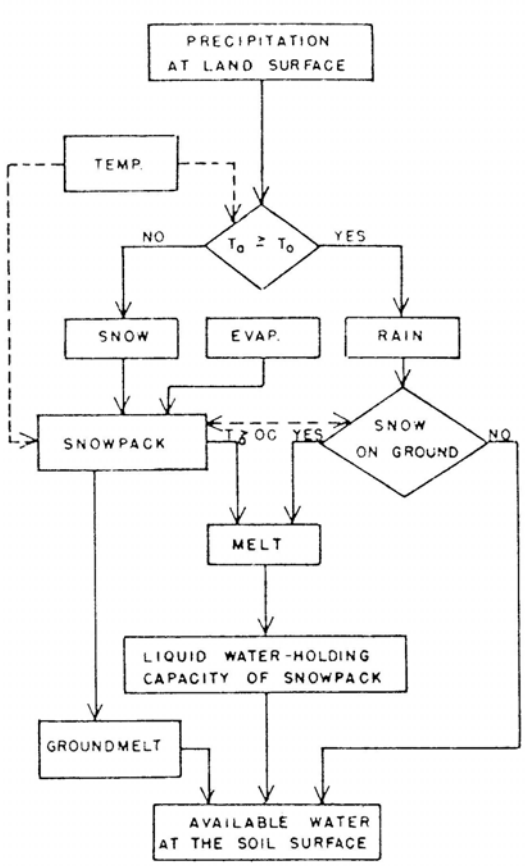


Fig.3.2 Snow accumulation and ablation (from Riley et al., 1969)

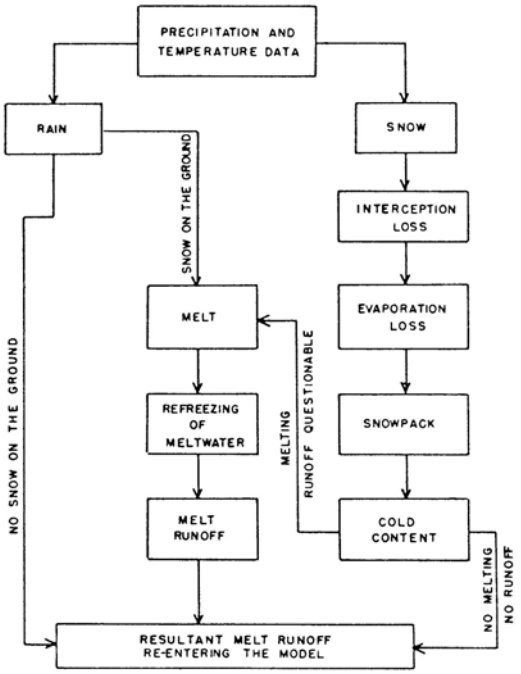


Fig.3.3 Block diagram of snowmelt for the Ohio State University Model (from Ricca, 1972)

REFERENCES OF CHAPTER 3

Anderson, E.A., 1973. National weather service river forecasting system- Snow accumulation and ablation model. NOAA Technical Memorandum NWS HYDRO-17, Silver Spring, MD.

Bengtsson, L., 1976. Snowmelt estimated from energy budget studies. Nordic Hydrology 7: 3-18.

Bergström, S., 1975. The development of a snow routine for the HBV-2 model. Nordic Hydrology, 6(2).

Bergström, S., 1976. Development and application of a conceptual runoff model for Scandinavian catchments. SMHI report, Nr. RHO 7.

Bergström, S., 1995. The HBV model. In: Singh, V.P. (Ed.) Computer Models of Watershed Hydrology. Water Resources Publication, Highlands, Ranch, CO, USA, pp 443-476.

Cox, L.M. and J.F. Zuzel, 1976. A method for determining sensible heat transfer to late-lying snowdrifts. Western Snow Conf. Proc. 44: 23-28.

Crops of Engineers, 1960. Runoff from snowmelt. US Army Crops of Engineers.

Gartska, W.U., 1944. Hydrology of small watersheds under winter conditions of snow cover and frozen soil. Geophys. Union Trans., Part 6: 838-871.

Granger, R.J. D.H. Male, 1977. Melting of a prairie snowpack. American Meteorological Society, Proc. 2nd Conf. On Hydrometeorol. pp 261-268.

Gray, D.M. and A.D.J. O'Neill, 1974. Application of the energy budget for predicting snowmelt runoff. In: Henry S. Stanford and James L. Smith (ed.), Advanced Concepts and Techniques in the Study of Snow and Ice Resources, National Academy of Sciences.

Haan, C.T., et al., 1982. Hydrologic Modeling of small Watersheds, Americam Society of Agricultural Engineers.

Hydrocomp, Inc. 1969. Hydrocomp simulation programming operation manual (2nd edition). Hydrocomp, Inc., Palo Alto, CA.

Jonch-Clausen, T., 1979. Systeme Hydrologique Europeen: a short description. SHE Report 1, Danish Hydraulics Institute, Horsholm, Denmark.

King, D.L., 1976. Simulation of the snow hydrology of the Paluse Prairie. University of Idaho, Master of Science Thesis.

Kuzmin, P.P. 1973. Melting of snow cover. Israel Program of Scientific Translation, published for National Science Foundation, Translation TT 71-50095.

Lang, H., 1984. Forecasting meltwater runoff from snow-covered areas and from glacier basins. In: J.R. Moll (Ed), Real-time river flow forecasting. Landbhouwhoge school, Wageningen, The Netherlands, Report 6: 113-146.

Leaf, C.F. and Brink, G.F., 1973. Computer simulation of snowmelt within a Colorado subalpine watershed. US dept of Agriculture, Forest Service, Research Paper RM-99.

Leavesley, 1989

Leavesley, et al., 1983

Martinec, et al., 1983

Martinec, J. and A. Rango, 1986. Parameter values for snowmelt runoff modelling. J. Hydrol., 84:197-219.

Mckay, G.A., 1968. Problems of measuring and evaluating snow-cover. In: snow Hydrology, Proc. Of Workshop Seminar, Canadian Nat. Comm. For the Int. Hydrol. Decade, pp49-65.

Molnau, M., 1971. Comparison of runoff from a catchment snow pillow and a small forested watershed, Western Snow Conf. Proc. 39: 39-43.

Morris, E.M., 1980. Modelling the flow of mass and energy within a snowpack for hydrological forecasting. Ann. Glaciol., 4:198-203.

Moussavi, M.,1988. Hydrologic Systems Modelling of Mountainous Watersheds in Semi-arid Regions. PhD thesis, Katholieke Universiteit Leuven, Belgium.

O'neill, A.D.J., 1972. The energetics of shallow prairie snowpacks. University of Saskatchewan, PhD Dissertation.

Power, J.M., 1986. Canada case study: water supply. In: G. Young (Ed). Techniques for prediction of runoff from glacierized areas. IAHS publ. No. 149. Institute of Hydrology, Wallingford, Oxfordshire, UK.

Raffelson, C.N., 1974. Evaporation from snowdrifts under oasis conditions. University of Wyoming, Master of Science Thesis.

Ricca, V.T., 1972. The Ohio State University version of the Stanford streamflow simulation model: Part 1 – Technical aspects. Ohio State University, Water Resources Center.

Riley, J.P. et al., 1969. Snowmelt simulation. Western Snow Conf. Proc. 37: 49-56.

Shih, G.B., et al., 1972. Computer modelling of a coniferous forest watershed. In: Age of Changing Priorities for land and water, ASCE.

US Army, 1975

US Crops of Engineers, 1956. Snow Hydrology. US Army, Crops of Engineers, North Pacific Division, Portland.

Vandewiele, G.L. and Ni-Lar-Win, 1993. Monthly water and snow balance models on basin scale, *In: Runoff and Sediment Yield Modelling,* K. Banasik and A. zbikowski (eds), Warsaw, 83-88.

Willen, D.W., et al., 1971. Simulation of daily snow water equivalent and melt. Western Snow Conference Proceedings, Billings, Mont.:1-8.

Wilson, W.T., 1941. An outline of the thermodynamics of snowmelt. Am. Geophys. Union Trans., Part 1: 182-195.

WMO, 1986.

Woo, M.K., 1972. A numerical simulation model for snow storage in small coastal basins, Southern British Columbia. IAHS Publ. no. 107, 2:992-1014.

Xu, C-Y, S. Halldin and J. Seibert, 1996. Regional water balance modelling in the NOPEX area: development and application of monthly water balance models. J. Hydrol. 180: 211-236.

Zuzel, J.F. and L.M. Cox, 1975. Relative importance of meteorological variables in snowmelt. Water Resour. Res. 11(1): 174-176.

CHAPTER 4
EVAPOTRANSPIRATION IN HYDROLOGIC MODELS

4.1 GENERAL

Three terms are used in this course. (1) The term evaporation, ET_0 is used for open/free water evaporation, i.e. the physical process involving a phase change from liquid to vapor by which water is returned to the atmosphere from lakes and reservoirs and, in some cases, from river channels in a river catchment. (2) The term actual evapotranspiration, AET describes all the processes by which liquid water at or near the land surface becomes atmospheric water vapor. Looking at a global average, two-thirds of the precipitation that falls on the continents is evapotranspired. Of this amount, 97% is from land surfaces and 3% is open-water evaporation. (3) The term potential evapotranspiration, ET is the maximum rate of evapotranspiration from a vegetated catchment under conditions of unlimited moisture supply. Of the three terms, the potential evapotranspiration or free water evaporation is usually used as an input together with precipitation to many hydrologic models. The term actual evapotranspiration is an important output for most hydrologic models.

Accurate spatial and temporal estimations of evapotranspiration are required for hydrologic models. Many methods of estimating evapotranspiration, whether for hydrologic models or irrigation scheduling, have been developed. In general, the procedure is to first estimate a potential evapotranspiration based on meteorological factors, then compute the amount of that potential that is utilized by the actual evapotranspiration processes, given the current status of the plant- and soil-moisture-related characteristics.

4.2 ESTIMATION OF FREE WATER EVAPORATION AND POTENTIAL EVAPOTRANSPIRATION

The potential for evapotranspiration is usually defined as an atmospheric determined quantity. There exist a multitude of methods for measurement and estimation of ET. Certain of these methods are accurate and reliable; others provide only a rough approximation. Most of the methods were developed for use in specific studies and are most appropriate for use in climates similar to where they were developed. It is not uncommon to use an equation for determination of evaporation from open water that was actually developed for determination of potential evapotranspiration from vegetated lands, and vice versa (see also Winter et al. 1995) although they are not the same as defined in previous section.

In general, techniques for estimating potential ET or ET_0 are based on one or more atmospheric variables, like solar or net radiation and air temperature and humidity, or some measurement related to these variables, like pan evaporation. Because climatic variables usually do not vary significantly over small areas, ET estimates can often be transferred some distance with minimal error. For most hydrologic applications, this is necessary because data are rarely available on the area where needed.

In the sections that follow many of the most commonly used techniques for estimating evaporation and potential evapotranspiration are described.

4.2.1 Climatological Methods

4.2.1.1 Air temperature-based methods

In certain regions of the world, meteorological and climatological data may be quite limited. Models based almost solely on air temperature may be used in such cases to provide estimates of ET. The temperature methods are some of the earliest methods for estimating ET (Jensen et al., 1990). If estimates are made for periods of several weeks or a month, reasonable approximations are possible. Some of the more common temperature-based models are described below. Most temperature-based equations take the form:

$$ET = cT^a \tag{4.1}$$

or

$$ET = c_1 d_l T(c_2 - c_3 h) \tag{4.2}$$

in which ET is potential evapotranspiration, T is air temperature, h is a humidity term, c_1, c_2, c_3 are constants, d_l is day-length. Many temperature-based equations have been developed and used. The following seven temperature-based equations each representing a special form of the equations (4.1) or (4.2) are discussed, namely: Thornthwaite (1948), Linacre (1977), Blaney-Criddle (1950), Hargreaves (1985), Kharrufa (1985), Hamon (1961), and Remanenko (1961) methods.

(1) Thornthwaite method:

A widely used method for estimating potential evapotranspiration was derived by Thornthwaite (1948) who correlated mean monthly temperature with evapotranspiration as determined from water balance for valleys where sufficient moisture water was available to maintain active transpiration. In order to clarify the existing method, the computational steps of Thornthwaite equation are discussed as follows:

Step 1: The annual value of the heat index I is calculated by summing monthly indices over a 12-month period. The monthly indices are obtained from the equation

$$i = \left(\frac{T_a}{5}\right)^{1.51} \tag{4.3a}$$

and

$$I = \sum_{j=1}^{12} i_j \tag{4.3b}$$

in which I = annual heat index; i = monthly heat index for the month j, (which is zero when the mean monthly temperature is 0 °C or less); T_a = mean monthly air temperature (degree Celsius); and j = number of months (1 - 12).

Step 2: The Thornthwaite general equation, Eq. 4a, calculates unadjusted monthly values of potential evapotranspiration, ET' (in mm), based on a standard month of 30 days, 12 hr of sunlight/day.

$$ET' = C(\frac{10T_a}{I})^a \quad (4.4a)$$

in which C = 16 (a constant); and $a = 67.5\times10^{-8}I^3 - 77.1\times10^{-6}I^2 + .0179I + .492$.

The value of the exponent a in the preceding equation varies from zero to 4.25 (e.g. Jain and Sinai, 1985), the annual heat index varies from zero to 160, and ET' is zero for temperature below zero degree Celsius.

Step 3: The unadjusted monthly evapotranspiration values ET' are adjusted depending on the number of days N in a month $(1 \le N \le 31)$ and the duration of average monthly or daily daylight d (in hr) which is a function of season and latitude.

$$ET = ET'(\frac{d}{12})(\frac{N}{30}) \quad (4.4b)$$

in which ET = adjusted monthly potential evapotranspiration (mm); d = duration of average monthly daylight (hr); and N = number of days in a given month, 1 - 31 (days).

Thornthwaite's equation has been widely criticized for its empirical nature but is widely used. Because Thornthwaite's method of estimating ET can be computed using only temperature, it has been one of the most misused empirical equations in arid and semi-arid irrigated areas where the requirement has not been maintained (Thornthwaite and Mather, 1955).

(2) Linacre Method

For the case of well-watered vegetation with an albedo of about 0.25, Linacre (1977) simplified Penman formula to give the following expression for the evaporate rate:

$$ET = \frac{500T_m/(100-A) + 15(T_a - T_d)}{(80 - T_a)} \quad (4.5)$$

where ET = Linacre potential evapotranspiration in mm/d, $T_m = T+0.006h$, h is the elevation (meters). A is the latitude (degrees) and T_d is the mean dew-point temperature. T_a, T_m and T_d are in °C. This formula requires only geographical data (A and h), the mean and the dew-point temperature.

(3) Blaney-Criddle method

The Blaney-Criddle (1950) procedure for estimating ET is well known in the western USA and has been used extensively elsewhere also (Singh, 1989). The usual form of the Blaney-Criddle equation converted to metric units is written as:

$$ET = kp(0.46T_a + 8.13) \quad (4.6)$$

where ET is evapotranspiration from reference crop, in mm, for the period in which p is expressed. T_a is mean temperature in °C, p is percentage of total daytime hours for the

used period (daily or monthly) out of total daytime hours of the year (365×12), and k is monthly consumptive use coefficient, depending on vegetation type, location and season. According to Blaney-Criddle, for the growing season (May to October) k varies from 0.5 for orange tree to 1.2 for dense natural vegetation.

(4) Kharrufa method

Kharrufa (1985) derived an equation through correlation of ET/p and T in the form of:

$$ET = 0.34\, p\, T_a^{\,1.3} \tag{4.7}$$

where ET = Kharrufa potential evapotranspiration in mm/month, T_a and p have the same definitions as given in equ (4.6).

(5) Hargreaves method

Hargreaves and Samani (1982, 1985) proposed several improvements for the Hargreaves (1975) equation for estimating grass-related reference ET. Because solar radiation data frequently are not available, Hargreaves and Samani (1982, 1985) recommended estimating R_s from extraterrestrial radiation, R_A, and the difference between mean monthly maximum and minimum temperatures, TD in °C. The resulting form of the equation is:

$$ET = 0.0023\, R_A\, TD^{1/2}\, (T_a+17.8) \tag{4.8}$$

The extraterrestrial radiation, R_A, is expressed in equivalent evaporation units. For a given latitude and day R_A is obtained from tables or may be calculated using a set of equations (see Jensen et al., 1990, page 179). The only variable for a given location and time period is air temperature. Therefore, the Hargreaves method has become a temperature-based method.

(6) Hamon method

Hamon (1961) derived a potential evapotranspiration method based on the mean air temperature and is expressed as

$$ET = 0.55\, D^2\, Pt \tag{4.9}$$

where ET is potential evapotranspiration in inch/day, D is the hours of daylight for a given day in units of 12 hr, and Pt is a saturated water vapour density term calculated by:

$$Pt = \frac{4.95 e^{(0.062 T_a)}}{100} \tag{4.10}$$

where T_a is daily mean air temperature in °C.

(7) Remanenko method

Remanenko (1961) derived an evaporation equation based on the relationship using mean temperature and relative humidity:

$$ET = 0.0018\, (25+T_a)^2\, (100-Rh) \tag{4.11}$$

where T_a is the mean air temperature in °C, Rh is the mean monthly relative humidity, which is calculated by:

$$Rh = \frac{e^{\circ}(T_d)}{e^{\circ}(T_a)} \tag{4.12}$$

in which e°(T) is the saturated vapour pressure calculated by (see Bosen, 1960):

$$e^{\circ}(T) = 33.8679\left[(.00738T + .8072)^8 - .000019\left|1.8T + 48\right| + .001316\right] \tag{4.13}$$

A comparative study of the above discussed temperature-based methods was done by Xu and Singh (2001).

4.2.1.2 Solar radiation-based methods

The radiation-based approach has had wide application in estimation of potential evapotranspiration (ET) of land areas. Many empirical formulae have been derived based on this approach (Jensen et al., 1990; Xu and Singh, 1999). Certain methods based on solar radiation also involve a temperature term.

Empirical radiation-based equations for estimating potential evaporation generally are based on the energy balance (Jensen et al., 1990). Most radiation-based equations take the form:

$$\lambda ET = C_r(wR_s) \quad \text{or} \quad \lambda ET = C_r(wR_n) \tag{4.14}$$

where λ is the latent heat of vaporisation (in calories per gram), ET is the potential evapotranspiration (in mm per day), R_s is the total solar radiation (in calories per cm^2 per day), R_n is the net radiation (in calories per cm^2 per day), w is the temperature and altitude-dependent weighting factor, and C_r is a coefficient depending on the relative humidity and wind speed. Eight popular radiation-based equations were evaluated and compared in this study: Turc (1961), Makkink (1957), Jensen and Haise (1963), Hargreaves (1975), Doorenbos and Pruitt (1977), McGuinness and Bordne (1972), Abtew (1996), and Priestley and Taylor (1972) equations. For the sake of completeness, these equations are briefly summarised in what follows. For more complete discussion, the reader is referred to the cited literature.

(1) Turc method

Under general climatic conditions of western Europe, Turc (1961) computed ET in millimetres per day for 10-day periods as

$$E_t = 0.013\frac{T}{T+15}(R_s + 50) \qquad \text{for RH} \geq 50 \tag{4.15}$$

$$E_t = 0.013\frac{T}{T+15}(R_s + 50)(1 + \frac{50 - RH}{70}) \qquad \text{for RH} < 50 \tag{4.16}$$

where T is the air temperature in ºC, R_s is the total solar radiation in cal/cm^2/day, and RH is the relative humidity in percentage.

(2) Makkink Method

Makkink (1957) estimated ET in millimetres per day over 10-day periods for grassed lands under cool climatic conditions of the Netherlands as:

$$ET = 0.61\frac{\Delta}{\Delta+\gamma}\frac{R_s}{58.5} - 0.012 \tag{4.17}$$

where Δ is the slope of saturation vapour pressure curve (in mb/ºC), γ (in mb/ºC) is the psychromatic constant. These quantities are calculated as (see also Singh, 1989):

$$\Delta = 33.8639[0.05904(0.00738T+0.8072)^7 - 0.0000342] \tag{4.18}$$

$$\gamma(mb/°C) = \frac{c_p P}{0.622\lambda} \tag{4.19}$$

$$\lambda \text{ (cal/g)} = 595 - 0.51T \tag{4.20}$$

$$P = 1013 - 0.1055\ EL \tag{4.21}$$

where EL is elevation (in metres), λ (in calories per gram) is latent heat, and P (in millibar) is atmospheric pressure. The specific heat of air c_p (in cal/g/ºC) varies slightly with atmospheric pressure and humidity, ranging from 0.2397 to 0.260. An average value of 0.242 is reasonable.

On the basis of later investigation in the Netherlands and at Tåstrup, Hansen (1984) proposed the following form of the Makkink equation

$$ET = 0.7\frac{\Delta}{\Delta+\gamma}\frac{R_s}{\lambda} \tag{22}$$

where all the notations have the same meaning and units as in (4.17).

(3) Jensen-Haise method

Jensen and Haise (1963) evaluated 3000 observations of ET as determined by soil sampling procedures over a 35-year period, and developed the following relation:

$$\lambda ET = C_t(T - T_x)R_S \tag{4.23}$$

where λ and R_s have the same meaning and units as before, ET is in mm/day, C_T (temperature constant) = 0.025, and T_x = -3 when T is in degree Celsius. These coefficients were considered to be constant for a given area.

(4) Hargreaves method

Hargreaves (1975) and Hargreaves and Samani (1982, 1985) proposed several equations for calculating potential evapotranspiration, ET (in mm/day). One of the equations is written as

$$\lambda ET = 0.0135(T + 17.8)R_S \tag{4.24}$$

All variables have the same meaning and units as before. The Hargreaves method was derived from eight years of cool season Alta fescue grass lysimeter data from Davis, California.

(5) Doorenbos and Pruitt Method

Doorenbos and Pruitt (1977) presented a radiation method for estimating ET using solar radiation. The method is an adaptation of the Makkink (1957) method and was recommended over the Penman method when measured wind and humidity data were not available or could not be estimated with reasonable confidence.

$$ET = a\left[\frac{\Delta}{\Delta + \gamma} R_S\right] + b \tag{4.25}$$

where R_s is solar radiation in mm/day, b = -0.3 mm/day and a is an adjustment factor that varies with mean relative humidity and daytime wind speed. The adjustment factor a was presented in graphic and tabular forms, and can also be calculated from

$$\begin{aligned} a = {} & 1.066 - 0.13\times10^{-2} RH + 0.045U_d - 0.20\times10^{-3} RH \times U_d \\ & - 0.315\times10^{-4} RH^2 - 0.11\times10^{-2} U_d^2 \end{aligned} \tag{4.26}$$

where RH is the mean relative humidity in percentage and U_d is the mean daytime wind speed in m/s.

(6) McGuinness and Bordne method

McGuinness and Bordne (1972) proposed a method for calculating potential evapotranspiration based on an analysis of a lysimeter data in Florida.

$$ET = \{(0.0082T - 0.19)(R_S/1500)\}2.54 \tag{4.27}$$

where ET is in cm/day for a monthly period, T is in degrees Fahrenheit, and R_s is in cal/cm^2/day.

(7) Abtew method

Abtew (1996) used a simple model that estimates ET from solar radiation as follows

$$ET = K\frac{R_S}{\lambda} \tag{4.28}$$

where ET is in mm/day, R_s is in $MJm^{-2}d^{-1}$, λ is in MJ kg^{-1}, and K is a dimensionless coefficient.

(8) Priestley and Taylor method

Priestley and Taylor (1972) proposed a simplified version of the combination equation (Penman, 1948) for use when surface areas generally were wet, which is a condition required for potential evaporation, ET. The aerodynamic component was deleted and the energy component was multiplied by a coefficient, $\alpha = 1.26$, when the general surrounding areas were wet or under humid conditions.

$$ET = \alpha \frac{\Delta}{\Delta + \gamma} \frac{R_n}{\lambda} \tag{4.29}$$

where R_n is the net radiation (cal $cm^{-2}d^{-1}$), and other notations have the same meaning and units as in equation (4.17).

A comparative study of the radiation-based methods was done by Xu and Singh (2000).

4.2.1.3 The Penman combination method

Penman (1948) was among the first to develop a method considering the factors of both energy supply and turbulent transport of water vapour away from an evaporating surface. The physical principles combine the two approaches, i.e. the mass-transfer and the energy balance. The basic equations are later modified and rearranged to use meteorological constants and measurements of variables made regularly at climatological stations. Following Shaw (1989), the Penman equation (4.34) may be derived as follows:

In a simplified energy balance equation:

$$H = E_o + Q \tag{4.30}$$

where H is the available heat energy, E_o is energy for evaporation (latent heat flux) and Q is energy for heating the air (sensible heat flux).

The values of E_o and Q can be defined by the aerodynamic equations:

$$E_o = f(u)(e_s - e_d) \tag{4.31}$$

and

$$Q = \gamma f_1(u)(T_s - T_a) \tag{4.32}$$

γ is the hygrometric constant (0.27 mm of mercury/ºF) to keep units consistent. It is generally assumed that $f(u) = f_1(u)$. If the aerodynamic equation (4.31) is based on the air humidity using the air temperature T_a, then:

$$E_a = f(u)(e_a - e_d) \tag{4.33}$$

where e_a is the saturated vapor pressure at air temperature T_a, and thus ($e_a - e_d$) is the saturation deficit (e_d, the vapor pressure of the air, is the saturated vapor pressure at the dew point T_d). The temperature, T_a, is easily measured, whence e_a is easily obtained, whereas e_s in equation (4.31) is difficult to evaluate.

If Δ represents the slope of the curve of saturated vapor pressure plotted against temperature, then:

$$\Delta = \frac{de}{dT} \approx \frac{e_s - e_d}{T_s - T_d} \approx \frac{e_a - e_d}{T_a - T_d} \quad \text{(if gradients are small)}$$

then from equation (4.32):

$$\begin{aligned} Q &= \gamma f(u)\left[(T_s - T_d) - (T_a - T_d)\right] \\ &= \gamma f(u)\left[\frac{(e_s - e_d)}{\Delta} - \frac{(e_a - e_d)}{\Delta}\right] \\ &= \frac{\gamma E_o}{\Delta} - \frac{\gamma E_a}{\Delta} \end{aligned}$$

substituting for Q in the energy balance equation (equation 4.30):

$$E_o = H - \frac{\gamma E_o}{\Delta} + \frac{\gamma E_a}{\Delta}$$

$$\Delta E_o + \gamma E_o = \Delta H + \gamma E_a$$

$$E_o = ET = \frac{\Delta}{\Delta + \gamma} H + \frac{\gamma}{\Delta + \gamma} E_a \qquad (4.34)$$

This final equation is the basic Penman formula for open water evaporation. It requires values of H and E_a as well as Δ for its application.

If net radiation measurements are available, then H, the available heat may be obtained directly. More often, H is calculated from incoming (R_I) and outgoing (R_O) radiation determined from sunshine records, temperature and humidity, using:

$$H = R_I(1 - r) - R_O \qquad (4.35)$$

where r is the albedo and equals 0.05 for water. R_I is a function of R_a, the theoretical radiation (fixed by latitude and season) modulated by a function of the ratio, n/N, of measured to maximum possible sunshine duration. Using r = 0.05 givens:

$$R_I(1 - r) = 0.95 R_a(0.18 + 0.55 n/N) \qquad (4.36)$$

the term R_O in equation (4.35) is given by:

$$R_O = \sigma T_a^4 \ (0.56 - 0.09\sqrt{e_d}\ )(0.10 + 0.90 n/N) \qquad (4.37)$$

Where σT_a^4 is the theoretical black body radiation at T_a which is then modified by functions of the humidity of the air (e_d) and the cloudiness (n/N).

Thus H in equation (4.34) is obtained from values found via equations (4.36) and (4.37) inserted into equation (4.35).

Next, E_a in equation (4.34) is found using the coefficients derived by experiment for open water:

$$E_a = 0.35(0.5 + u_2/100)(e_a - e_d) \tag{4.38}$$

Finally, a value of Δ is found from the curve of saturated vapor pressure against temperature corresponding to the air temperature, T_a.

The equations given are those originally published by Penman. The four measurements required to calculate the open water evaporation are thus:

T_a mean air temperature for a week, 10 days or a month, ºF or ºC
e_d mean vapor pressure for the same period, mm of mercury
n bright sunshine over the same period, h day^{-1}
u_2 mean wind speed at 2 m above the surface, miles day^{-1}

With meteorological observations made in various units and the tendency to work now in SI units, care is needed in converting measurements into the appropriate units for the formula. The evaporation ET is finally in mm/day.

4.2.1.4 Penman-Monteith method

The Penman combination method (equ (4.34)) was further developed by many researchers, an excellent work was done by Monteith (1963, 1964) who introduced resistance terms and arrived at the following equation for ET from surfaces with either optimal or limited water supply:

$$ET = \frac{(R_n - G)\Delta + \dfrac{\rho c_p (e_s - e_a)}{r_a}}{\lambda\left[\Delta + \gamma\left(1 + \dfrac{r_s}{r_a}\right)\right]} \tag{4.39}$$

where: ET = evapotranspiration; R_n = net radiation; Δ = rate of increase with temperature of the saturation vapor pressure of water at air temperature; ρ = density of air; c_p = specific heat of air at constant pressure; (e_s-e_a) = vapor pressure deficit of air; r_a = aerodynamic resistance to water vapor transport; λ = latent heat of vaporization of water; γ = psychometric constant; r_s = bulk (canopy) surface resistance to water transport. G = soil heat flux.

The Penman-Monteith approach as formulated above includes all parameters that govern energy exchange and corresponding latent heat flux (evapotranspiration) from uniform expanses of vegetation. This model requires data on r_a and r_s which are not readily available. The FAO (Allen et al. 1998) recommended equations for computing r_a and r_s and substituted them into equation (4.39). From the original Penman-Monteith

equation (equation 4.39) and the equations of the aerodynamic and surface resistance, the FAO Penman-Monteith method for calculating reference (potential) evapotranspiration ET can be expressed as (Allen et al. 1998):

$$ET = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_a + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \qquad (4.39a)$$

where: ET = reference evapotranspiration [mm day^{-1}], R_n = net radiation at the crop surface [MJ m^{-2} day^{-1}], G = soil heat flux density [MJ m^{-2} day^{-1}], T = mean daily air temperature at 2 m height [°C], u_2 = wind speed at 2 m height [m s^{-1}], e_s = saturation vapour pressure [kPa], e_a = actual vapour pressure [kPa], e_s - e_a = saturation vapour pressure deficit [kPa], Δ = slope vapour pressure curve [kPa $°C^{-1}$], γ = psychrometric constant [kPa $°C^{-1}$].

Apart from the site location, the FAO Penman-Monteith equation requires air temperature, humidity, radiation and wind speed data for daily, weekly, ten-day or monthly calculations.

The procedure for using equ (4.39a) for computing reference evapotranspiration has been given in Chapter 3 of the FAO paper 56 (Allen et al., 1998), which is briefly summarized in what follows. It is important to verify the units in which the weather data are reported.

Latent Heat of Vaporization (λ)

$\lambda = 2.501 - (2.361 \times 10^{-3})\ T_a$

where: λ = latent heat of vaporization [MJ kg^{-1}], T_a = air temperature [°C].

Atmospheric Pressure (P)

$$P = 101.3\left(\frac{293 - 0.0065z}{293}\right)^{5.26}$$

where: P = atmospheric pressure [kPa] at elevation z [m].

Saturation Vapour Pressure (e_s)

$$e_s(T_a) = 0.611 \exp\left(\frac{17.27T_a}{T_a + 237.3}\right)$$

where: $e_s(T_a)$ = saturation vapour pressure function [kPa] and T_a = air temperature [°C].

Actual vapour pressure (e_a)

$$e_a(T_d) = 0.611 \exp\left(\frac{17.27T_d}{T_d + 237.3}\right)$$

where: $e_a(T_d)$ = actual vapor pressure function [kPa] and T_d = dew point temperature [°C].

Slope Vapour Pressure Curve (Δ)

$$\Delta = \frac{4098 e_s(T_a)}{(T_a + 237.3)^2} = \frac{2504 \exp\left(\frac{17.27 T_a}{T_a + 237.3}\right)}{(T_a + 237.3)^2}$$

where: Δ = slope vapour pressure curve [kPa C^{-1}] and T_a = air temperature [°C].

Psychrometric Constant (γ)

$$\gamma = \frac{C_p P}{\varepsilon \lambda} \times 10^{-3} = 0.00163 \frac{P}{\lambda}$$

where: γ = psychrometric constant [kPa C^{-1}], c_p = specific heat of moist air = 1.013 [kJ kg^{-1} $°C^{-1}$], P = atmospheric pressure [kPa], ε = ratio molecular weight of water vapour/dry air = 0.622 and λ = latent heat of vaporization [MJ kg^{-1}].

Short Wave Radiation on a Clear-Sky Day (R_{so})

The calculation of R_{so} is required for computing net long wave radiation. A good approximation for R_{so} according to FAO (Allen et al., 1998) for daily and hourly periods is:

$R_{so} = (0.75 + 2 \times 10^{-5} z) R_a$

where: R_{so} = short wave radiation on a clear sky day [MJ m^{-2} d^{-1}], z = station elevation [m], R_a = extraterrestrial radiation [MJ m^{-2} d^{-1}].

Extraterrestrial radiation for daily periods (R_a)

The extraterrestrial radiation, R_a, for each day of the year and for different latitudes is estimated from the solar constant, the solar declination and the time of the year by:

$$R_a = \frac{24(60)}{\pi} G_{sc} d_r \left[\omega_s \sin(\varphi)\sin(\delta) + \cos(\varphi)\cos(\delta)\sin(\omega_s)\right]$$

where: R_a = extraterrestrial radiation [MJ m^{-2} day^{-1}], G_{sc} = solar constant = 0.0820 MJ m^{-2} min^{-1}, d_r = inverse relative distance Earth-Sun, ω_s = sunset hour angle, φ = latitude [rad] and δ = solar decimation.

The equations for calculating d_r, ω_s, φ and δ are given in chapter 3 of FAO paper 56 (Allen et al., 1998).

Net solar or net shortwave radiation (R_{ns})

The net shortwave radiation resulting from the balance between incoming and reflected solar radiation is given by:

$R_{ns} = (1-\alpha)R_s$

Where: R_{ns} = net solar or shortwave radiation [MJ m^{-2} day^{-1}], α = albedo or canopy reflection coefficient, which is 0.23 for the hypothetical grass reference crop [dimensionless] and R_s = the incoming solar radiation [MJ m^{-2} day^{-1}].

Net longwave radiation (R_{nl})

The net outgoing longwave radiation is calculated by

$$R_{nl} = \sigma \left[\frac{T_{\max,K}^4 + T_{\min,K}^4}{2} \right] \left(0.34 - 0.14\sqrt{e_a}\right) \left(1.35 \frac{R_s}{R_{so}} - 0.35\right)$$

where: R_{nl} = net outgoing longwave radiation [MJ m^{-2} day^{-1}], σ = Stefan-Boltzmann constant [4.903 10^{-9} MJ K^{-4} m^{-2} day^{-1}], $T_{max, K}$ = maximum absolute temperature during the 24-hour period [K = °C + 273.16], $T_{min, K}$ = minimum absolute temperature during the 24-hour period [K = °C + 273.16], e_a = actual vapour pressure [kPa], R_s/R_{so} = relative shortwave radiation (limited to ≤ 1.0), R_s = measured solar radiation [MJ m^{-2} day^{-1}] and R_{so} = calculated clear-sky radiation [MJ m^{-2} day^{-1}].

Net radiation (R_n)

The net radiation (R_n) is the difference between the incoming net shortwave radiation (R_{ns}) and the outgoing net longwave radiation (R_{nl}):

$R_n = R_{ns} - R_{nl}$

Soil heat flux (G)

For vegetation covered surface and calculation time steps are 24 hours or longer, a calculation procedure proposed by FAO (Allen et al., 1998), based on the idea that the soil temperature follows air temperature is as follows,

$$G = c_s \frac{T_i - T_{i-1}}{\Delta t} \Delta z$$

where: G = soil heat flux [MJ m^{-2} day^{-1}], c_s = soil heat capacity [MJ m^{-3} $°C^{-1}$], T_i = air temperature at time i [°C], T_{i-1} = air temperature at time i-1 [°C], Δt = length of time interval [day], Δz = effective soil depth [m], which for a time interval of one or few days is about 0.10 – 0.20 m. Different equations are proposed by Allen et al. (1998) in calculating G depending on the computation time periods.

4.2.2 Micrometeorological Methods

4.2.2.1 The mass-transfer-based methods

The mass-transfer method is one of the oldest methods (Dalton, 1802; Meyer, 1915; Penman, 1948) and is still an attractive method in estimating free water surface evaporation, ET_0, because of its simplicity and reasonable accuracy. The mass-transfer methods are based on the Dalton equation which for free water surface can be written as:

$$ET_0 = C(e_s - e_a) \tag{4.40}$$

where ET_0 is free water-surface evaporation, e_s is the saturation vapor pressure at the temperature of the water surface, e_a is the actual vapor pressure in the air, and C is an empirically determined constant involving some function of windiness.
Therefore equation (1) is expressed as:

$$ET_0 = f(u)(e_s - e_a) \tag{4.41}$$

where *f(u)* is the wind function. This function depends, among other factors, on the observational heights of the wind speed and vapor pressure measurements. Although the two heights need not be the same, the same experimental layout must be used for a particular value of the function. The mass-transfer method has had wide application in the estimation of lake evaporation and many empirical formulae have been derived based on this approach (Singh, 1989). Examples of empirical equations of this type are included in Table 4.1.

An inspection of the above mass-transfer-based equations reveals that three major meteorological factors considered to affect evaporation are (1) vapor pressure gradient, (2) wind speed, and (3) temperature. The air pressure, fluid density, and water surface elevation for a given location may not greatly affect the rate of evaporation. Table 1 also shows that specific formulas have resulted from the analysis of limited and site specific meteorological data. The data collection procedures are not only varied but are frequently inconsistent. Usually, such inconsistencies are a major source of site specific modifications and adaptations of these types of equations. Specifically, the elevations at which temperature and vapor pressure are measured vary widely. As a result, estimates of moisture gradient and wind velocity are affected.

An evaluation and comparison of mass-transfer methods was performed by Singh and Xu (1997a). More recently, a cross-comparison of mass-transfer, radiation and temperature based evaporation models was done by Xu and Singh (2002).

Table 4.1 Some mass-transfer-based evaporation equations for estimation of evaporation.

No.	Author	Equation	remarks
1)	Dalton (1802)	ET_0 (in./mo)=$a(e_s-e_a)$	a=15 for small, shallow water, and a=11 for large deep water
2)	Fitzgerald (1886)	ET_0 (in./mo)=$(.4+.199u)(e_s-e_a)$	
3)	Meyer (1915)	ET_0 (in./mo)=$11(1+.1u)(e_s-e_a)$	e_a is measured at 30 ft above the surface
4)	Horton (1917)	ET_0 (in./mo)=$.4[(2-\exp(-2u))(e_s-e_a)]$	
5)	Rohwer (1931)	ET_0 (in./da)=$.77(1.465-.0186p_b)\cdot(.44+.118u)(e_s-e_a)$	p_b= barometric pressure in in. of Hg.
6)	Penman (1948)	ET_0 (in./da)=$.35(1+.24u_2)(e_s-e_a)$	
7)	Harbeck et al (1954)	ET_0 (in./da)=$.0578u_8(e_s-e_a)$ ET_0 (in./da)=$.0728u_4(e_s-e_a)$	
8)	Kuzmin (1957)	ET_0 (in./mo)=$6.0(1+.21u_8)(e_s-e_a)$	
9)	Harbeck et al (1958)	ET_0 (in./da)=$.001813u(e_s-e_a)(1-.03(T_a-T_w))$	T_a = average air temperature ^oC +1.9^oC; T_w = average water surface temperature ^oC.
10)	Konstantinov (1968)	ET_0 (in./da)=$.024(t_w-t_2)/u_1+.166u_1)(e_s-e_a)$	
11)	Remanenko (1961)	ET_0 (cm/mo)=$.0018(T_a+25)^2(100-hn)$	hn =relative humidity
12)	Sverdrup (1946)	$ET_0(in./h)=\frac{.623\rho K_o^2(u_8-u_2)(e_2-e_8)}{p[\ln(800/200)]^2}$	K_o=von Karman's const ρ = density of air p=atmospheric pressure
13)	Thornthwaite & Holzman (1939)	$ET_0(in./h)=\frac{.623\rho K_o^2(u_8-u_2)(e_2-e_8)}{p[\ln(800/200)]^2}$	

The wind speed (monthly mean) u is measured in miles per hour and vapor pressure e, in inches of Hg. The subscripts attached to u refer to height in meters at which the measurements are taken; no subscript refers to measurements near the ground or water surface.

4.2.2.2 Aerodynamic method

Theories, principles, and procedures involved in the aerodynamic methods are discussed in the course "Process Hydrology". One example of such methods is briefly shown here. Thornthwaite and Holzman (1942) were among the first modern micrometeorologists to apply the aerodynamic approach to measurement of ET. They proposed a relationship involving the gradients of specific humidity q and the logarithmic wind profile. Their expression, given here without derivation, is

$$ET = \rho_a k^2 \frac{(q_2 - q_1)(U_2 - U_1)}{\ln(z_2 / z_1)^2} \tag{4.42}$$

where ρ_a = density of moist air, k = von Karman's constant. Over a rough cropped surface z - d is substituted for z. An error analysis of this method is given by Thompson and Pinker (1981).

Following Thornthwaite and Holzman's work, many others (e.g., Pasquil, 1950; Pruitt, 1963; Dyer, 1974) have proposed stability-corrected aerodynamic methods for estimating the flux of vapor. Aerodynamic methods require stringently accurate observations of wind speed and specific humidity or vapour pressure at a number of heights above the surface, as well as temperature measurement to permit stability corrections to be made. Because of its origins in classical fluid dynamics theory, aerodynamic methods have been popular with scientists. However, the methods have not reached a degree of development that makes them applicable for routine use, for example, in hydrological modeling.

4.2.2.3 Bowen ration-energy balance method

Bowen (1926) introduced a relationship between latent heat flux, λE and sensible heat flux, H known as the Bowen ration β. This is defined by

$$\beta = \frac{H}{\lambda E} = \frac{PC_p}{\lambda}\left(\frac{M_a}{M_w}\right)\left(\frac{K_h}{K_w}\right)\frac{\partial T / \partial z}{\partial e / \partial z} = \gamma \frac{K_h}{K_w}\frac{\partial T / \partial z}{\partial e / \partial z} \tag{4.43}$$

where M_w and M_a are the molecular weights of water vapor and air, K_h and K_w are the turbulent exchange coefficients for sensible heat and water vapor. Other notations are previously defined.

This relationship is generally simplified by assuming that the turbulent exchange coefficient for heat transport K_h = the exchange coefficient for water vapor transport K_w and that $(\partial T / \partial z)/(\partial e / \partial z) \approx \Delta T / \Delta e$ where $\Delta T = T_2 - T_1$, and $\Delta e = e_2 - e_1$. Equation (4.43) then becomes

$$\beta \approx \gamma \frac{\Delta T}{\Delta e} \tag{4.44}$$

a simplified form of energy balance equation at the earth's surface can be written as:

$$R_n + S + \lambda E + H = 0 \tag{4.45}$$

From (4.43), H = βλE. Substitution into (4.45) and solution for λE yields

$$\lambda E = -\frac{R_n + S}{1+\beta} = -\left[\frac{R_n + S}{1+\gamma \Delta T / \Delta e}\right] \tag{4.46}$$

Equation (4.46) is the so called "Bowen ratio-energy balance" (BREB) method of estimating λE.

4.2.3 The pan Method

Measured evaporation from a shallow pan of water is one of the oldest and common methods for estimating ET_0. It is an indirect integration of the principal atmospheric variables related to ET_0. Pans are inexpensive, relatively easy to maintain and simple to operate. In humid regions, pans may also give realistic estimates of potential evapotranspiration, ET. However, care must be taken in relating evaporation from pans to ET in arid climates (Rosenberg et al., 1983, page 262). Given some standardization of pan shape, environmental setting, and operation, good correlations have been developed between pan evaporation, E_p, and potential evaporation, ET, by a simple relation

$$ET = C_{ET} E_p \tag{4.47}$$

where C_{ET} is a coefficient.

Pan-to-ET coefficients (C_{ET}) are necessary because evaporation for a pan is generally more than for a well-wetted vegetated surface, or even a pond, due to the pan's excessive exposure and lower reflectance of solar radiation. The values of C_{ET} vary normally from 0.5 to 1.0. The actual value depends, among other factors, on the type of pan, the location of the measurement, and the season. Although specific coefficient values for application to any given situation or pan may have to be found by calibration, mean monthly values are usually shown in a table or graphically shown in a map for some major meteorological stations or regions.

4.2.4 Relationship between ET_o and ET

Most subsequent refinements of Penman's formula for ET_0 have been concerned with adapting it to calculate potential evapotranspiration, ET. Penman himself began with a purely empirical approach, comparing calculated ET_0 with ET losses from well-watered plots covered in base soil and short-cropped grass at Rothamsted Research Station, near London. Collecting data from similar plots in a wide variety of climates from Europe to the humid tropics, Penman produced the empirical formula

$$ET = f\,ET_0$$

Where f is a seasonal correction factor which covers the effects of differing insolation intensity, day length, stomatal response and geometry. He concluded from the experiment that:

(1) The evaporation rate (as measured by *f* in the equation) for continuously wet bare soil is 0.9 times that an open water surfaces exposed to the same weather conditions in all seasons.
(2) The corresponding relative evaporation rate from turf with a plentiful water supply varies with season of the year. Provisional value of *f* for southern England are:

Midwinter (Nov – Feb)	0.6
Spring and autumn (Mar, Apr, Sep, Oct)	0.7
Midsummer (May – Aug)	0.8
Whole year	0.75

He still recommended local empirical confirmation wherever possible, because of the large variation that factors such as age and species can cause.

4.3 ESTIMATION OF ACTUAL ET IN HYDROLOGIC MODELS

A large number of methods have been developed in recent years for actual ET predictions each has its own requirements and emphases. The available methods range from quite simple to very complex. The most complex and physical realistic method used for actual evapotranspiration calculation used in the physically-based models is the Penman-Monteith equation (Equation 4.39a) as discussed in the previous section. It is seen that this method requires many variables that might not be available. For most conceptual hydrological models, this method is too sensitive to data requirement.

In conceptual catchment models, the most investigators have found it necessary to derive "actual" evapotranspiration as a function of potential evapotranspiration and the dryness of the soil (Palmer, 1965; Saxton and McGuinness, 1982; Dyck, 1983). As the model storage ratio (actual soil moisture storage divided by the maximum storage) is representative of the 'wetness' of the soil, it would be conceptually acceptable to extract moisture at the potential rate when the storage was full, that is at field capacity, and reduce the extraction to zero when the storage was empty (when the soil moisture deficit had reached its maximum). However, the nature of the function that estimates actual evapotranspiration for conditions between these limits is not known. A number of functions operating between the limits of potential rate and zero have been tried by a number of modellers. A general form of such equations can be shown as

$$AET = ET \cdot f(SMT / SMC) \tag{4.48}$$

where SMT is the actual soil moisture storage, and SMC is the soil moisture storage at field capacity.

Some examples for $f(SMT / SMC)$ and other functions are given in Table 4.2. The ratio of actual evapotranspiration (AET) to potential evapotranspiration (ET) varies with drying of soil and that the shape of this curve differs, amount other factors, with the type of soil, as illustrated in Figure 4.1.

Table 4.2 Some examples for function $f(SMT/SMC)$.

Reference	$f(SMT/SMC) =$	
DAILY VALUES		
Minhas et al. (1974)	$\frac{1-\exp(-\gamma SMT)}{1-2\exp(-\gamma SMC)+\exp(-\gamma SMT)}$	(4.49)
Norero(1969)	$\left[1+(SMT/SMC)^{b\cdot k}\right]^{-1}$	(4.50)
	$k = 2.69\,exp(-0.09\,PET)^{-0.62}$	(4.51)
Baier & Robertson (1966)	$\sum_{j=1}^{n} k_j (SMT_{j,i-1}/SMC_j) Z_j$	(4.52)
Koitzsch & Golf (1983)	$\frac{1}{1-0.533M_i}\frac{SMT_{i-1}}{SMC}$	(4.53)
HBV & many others	$\frac{SMT}{LP}$	(4.54
	(in the following equations $RAT = SMT/SMC$)	
Roberts (1978)	$(\mathrm{RAT})^{0.5}$	(4.55)
	$RAT^2/((RAT^2)+(1-RAT)^2)$	(4.56)
	$2\times RAT^2\times(1/(1+RAT)^{RAT})$	(4.57)
	$2\times RAT\times(1/(1+RAT)^{RAT})$	(4.58)
	$RAT^{1/2}+(RAT^{1/2}-RAT)$	(4.59)
	RAT^2	(4.60)
	RAT	(4.61)
5-DAY VALUES		
Renger et al. (1974)	$0.2+2.0\times RAT-1.2\times RAT^2$	(4.62)
MONTHLY VALUES		
Budyko & Zubenok (1961)	RAT	(4.63)
Xu et al. (1996)	$1-a_1^{((SMT_{i-1}+P_i)/ET)}$	(4.64)

Where: SMT = actual soil moisture; SMC = soil moisture at field capacity; $SMT_{j,i-1}$ = actual soil moisture in the j-th zone at the end of the previous day (i-1); Z_j = fraction of available soil moisture at which AET<ET and plant stress sets in; K_j = fraction of soil moisture extraction at that zone; γ = free parameter; b = soil specific constant; M = vegetation canopy density.

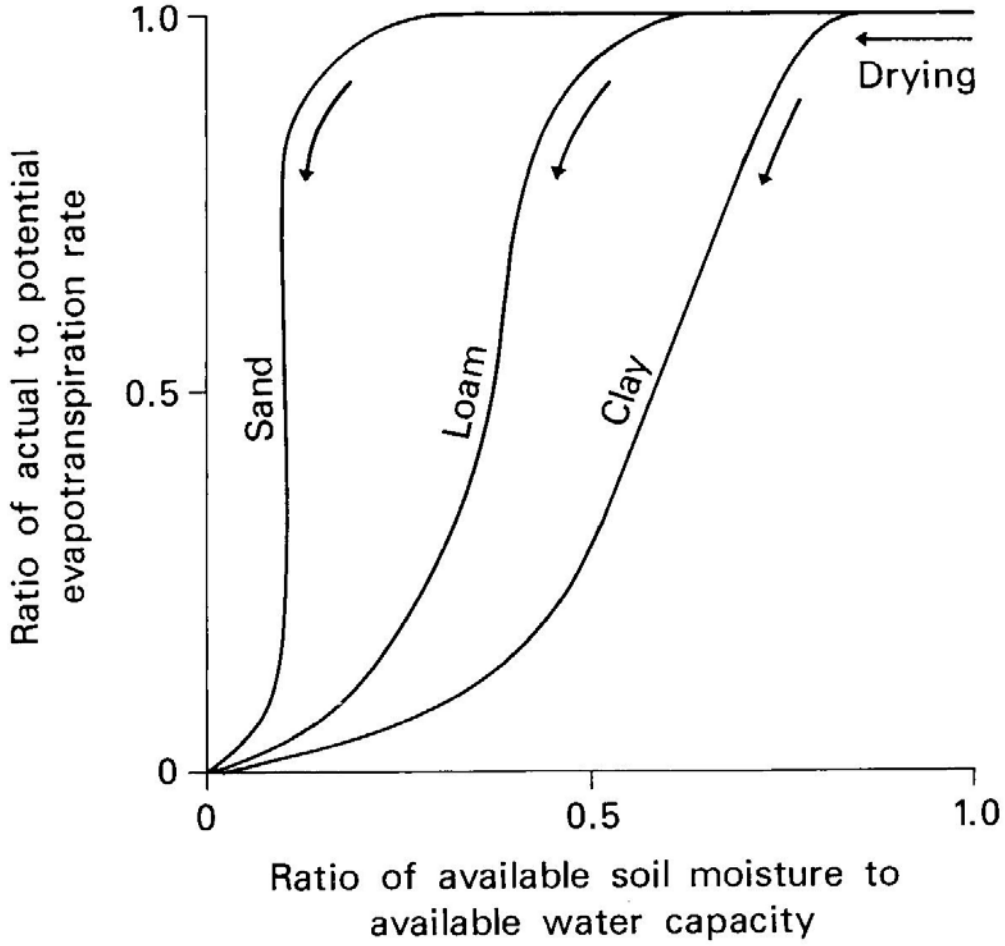


Fig.4.1 The graph compares the rates of reduction in evapotranspiration as the soil dries out (right to left) for sand, loam and clay, quantified as a proportion of the total water-holding capacity of the root zone, the part of the soil that contains water available to plants ("the available water capacity"). In sand, plants can extract water at full potential rates until near their wilting point, but in clay the supply is restricted by smaller soil pores so that uptake and transpiration rates can rapidly fall below potential much earlier. Based on Dunne and Leopold (1978), see also Jones (1997).

REFERENCES OF CHAPTER 4

Abtew, W., 1996. Evapotranspiration measurement and modeling for three wetland systems in South Florida. *Water Resources Bulletin*, 32, 465-473.

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M., 1998. Crop evapotranspiration – guidelines for computing crop water requirements – FAO Irrigation & Drainge Paper 56. FAO, 1998. ISBN 92-5-104219-5.

Baier, W. and Robertson, G.W., 1966. A new veraltile soil moisture budget. *Can. J. Plant Sci*., 46.

Blaney, H.F. and Criddle, W.D., 1950. Determining water requirements in irrigated area from climatological irrigation data. US Department of Agriculture, Soil Conservation Service Tech. Pap. No. 96, 48pp.

Bosen, J.F., 1960. A formula for approximation of saturation vapor pressure over water. Monthly Weather review 88(8):275-76.

Bowen, I.S., 1926. The ratio of heat losses by conduction and by evaporation from any water surface. Phys. Rev. 27: 779-787.

Brutsaert, W.H., 1982. Evaporation into the Atmosphere: Theory, History, and applications. Boston: D. Reidel Publishing Company, Massachusetts.

Budyko, M. and Zubenok, L.I., 1961. Opredelenie isparenija poverchnosti sushi. *AN SSSR Ser. Geogr*., 6.

Dalton, J., 1802. Experimental essays on the constitution of mixed gases: on the force of steam or vapor from water or other liquids in different temperatures, both in a Torricelli vacuum and in air; on evaporation; and on expansion of gases by heat. *Manchester Literary Philosophical Society Mem. Proceedings 5*: 536-602.

Doorenbos, J. and Pruitt, W.O., 1977. Crop water requirements. FAO, Irrigation and Drainage Paper 24, 144pp.

Dyck, S., 1983. 'Overview on the present status of the concepts of water balance models'. *New Approaches in Water Balance Computations (Proceedings of the Hamburg Workshop)*, Van der Beken, A. and Herrmann, A. (Eds), IAHS Publ. No.148: p3-19.

Dyer, A.J., 1874. A review of flux-profile relationship. Boundary-layer Meteorol. 7: 363-372.

Fitzgerald, D., 1886. Evaporation. *Transactions of the American Society of Civil Engineers 98(HY12)*: 2073-85.

Gurney, R.J. and Camillo, P.J., 1984. Modelling daily evapotranspiration using remotely sensed data. *J. hydrol.*, 69:305-324.

Hamon, W.R., 1961. Estimating potential evapotranspiration. *J. Hydraul. Div., Proc. Am. Soc. Civil Eng*., 87: 107-120.

Harbeck, G.E., 1962. A practical field technique for measuring reservoir evaporation utilizing mass-transfer theory. *Geological Survey Professional Paper 272-E*, 101-5, Washington, D.C.: U.S. Government Printing Office.

Harbeck, G.E., et al. 1954. Water loss investigations, Vol.1, Lake Hefner studies. *U.S. Geological Survey Paper* 269, Washington, D.C.

Harbeck, G.E., et al. 1958. Water loss investigations, Lake Mead studies. *U.S. Geological Survey Professional Paper* 298, Washington, D.C.

Hargreaves, G.H. and Samni, Z.A., 1982. Reference crop evapotranspiration from temperature. *TRANSACTION of the ASAE.*

Hargreaves, G.H. and Samni, Z.A., 1985. Reference crop evapotranspiration from temperature. *TRANSACTION of the ASAE*

Hargreaves, G.H., 1975. Moisture availability and crop production. *TRANSACTION of the ASAE* 18, 980-984.

Haude, W., 1955. Zur Bestimmung der Verdunstung auf möglichst einfacher Weise. Mitteilungen des Deutschen Wetterdienstes, Nr.11, Band 2, Bad Kissingen.

Helfrich, K.R., Adams, E.E., Godbey, G.E. and Harleman, D.R.F., 1982. Evaluation of the models for predicting evaporative water loss in cooling impoundments. Palo Alto, Calif.: Electric Power Res. Inst. (Interim report no. CS-2325).

Horton, R.E., 1917. Rainfall interception. *Monthly Weather Review 47(9)*:603-23.

Hounam, C.E., 1971. Problems of evaporation assessment in the water balance. Report No.3 on WMO/IHD Projects, Geneva.

Jain, P.K. and Sinai, G., 1985. Evapotranspiration model for semiarid regions. *Journal of Irrigation and drainage Engineering* 111(4): 369-379.

Jensen, M.E. and Haise, H.R., 1963. Estimation of evapotranspiration from solar radiation. *Journal of Irrigation and Drainage Division, Proceedings of the American Society of Civil Engineers 89*, 15-41.

Jensen, M.E., 1973. Consumptive use of water and irrigation requirements. ASAE, New York.

Jensen, M.E., Burman, R.D. and Allen, R.G., 1990. *Evapotranspiration and Irrigation Water Requirements*. American Society of Civil Engineers, New York, 1990.

Jones, D.M.A., 1966. Variability of evapotranspiration in Illinois. Circular 89, Illinois State Water Survey, Urbana, Ill.

Jones, J.A.A., 1997. Global Hydrology – Processes, resources and environmental management. Pearson Education Limited, Harlow, England, 399pp.

Kharrufa, N.S., 1985. Simplified equation for evapotranspiration in arid regions. *Beiträge zur Hydrologie*, Sonderheft 5.1, 39-47.

Klemes, V., 1986. Operational testing of hydrological simulation models, *Hydrol. Sci., J.,* 31(1):13-24.

Kohler, M.A., 1957. Meteorological aspects of evaporation phenomena, 421-436. IUGG Meeting held September in Toronto, Canada.

Kohler, M.A., Nordenson, T.J. and Fox, W.E., 1955. Evaporation from pans and lakes. Weather Bureau Research Paper 38, U.S. Department of Commerce, Washington, D.C.

Koitzsch, R. and Golf, W., 1983. Algorithmus zur Berechnung der realen verdunstung.

Konstantinov, A.R., 1968. Evaporation in nature, Leningrad.

Kuzmin, P.O., 1957. Hydrophysical investigations of land waters. *International Association of Scientific Hydrology Publication 3*:468-78.

Linacre, E.T., 1977. A simple formula for estimating evaporation rates in various climates, using temperature data alone. Agricultural Meteorology, 18, 409-424.

Linsley, R.K., Jr., M.A. Kohler and J.L.H. Paulhus, *Hydrology for Engineers*, McGraw-Hill, New York, 1982.

Makkink, G.F., 1957. Testing the Penman formula by means of lysimeters. *Journal of the Institution of Water Engineers*, 11, 277-288.

McCulloch, J.S.G., 1965. Tables for the rapid computation of the Penman estimate of evaporation. East African Agricultural and Forestry Journal XXX(3):286-95.

McGuinness, J.L. and Bordne, E.F., 1972. A compariosn of lysimeter-derived potential evapotranspiration with computed values. Tech. Bull. 1452, 71pp. Agric. Res. Serv., U.S. Dept. of Agric., Washington, D.C., 1972.

Meyer, A.F., 1915. Computing runoff from rainfall and other physical data. *Transactions, American Society of Civil Engineers 79*:1055-155.

Meyer, A.F., 1944. Evaporation from lakes and reservoirs. Minnesota Resources Commission, St. Paul, Minn.

Minhas, B.S., et al., 1974. Toward the structure of a production function for wheat yields with dated input of irrigation water. *Wat. Resour. Res.* 10(3):383-393.

Monteith, J.L., 1963. Gas exchange in plant communities. Environmental Control of Plant Growth (L.T. Evans, ed.) Academic Press, New Rork.

Monteith, J.L., 1964. Evaporation and evironment. The state and movement of water in living organisms. 19th Symp. Soc. Exp. Biol., Academic Press, New Rork.

Monteith, J.L., 1965. Evaporation and environment, *Proc. Symp. Exptl. Biol*., 19: 205-234.

Morton, F.I., 1968. Evaporation and climate: a study in cause and effect. *Scientific Series no.4*, Inland Water Branch, Department of Energy, Mines and Resources, Ottawa, Canada.

Morton, F.I., 1990. Studies in evaporation and Their lessons for the Environmental Sciences. *Canadian Water Resources Journal* 15(3):261-285.

Morton, F.I., 1994. Evaporation research - A critical review and its lessons for the Environmental Sciences. *Critical Reviews in Environmental Science and Technology*, 24(3):237-280.

Nash, J.E. and Sutcliffe, J., 1970. River flow forecasting through conceptual models Part I. A discussion of principles. *J. Hydrol.*, 10: 282-290.

Norero, A.L., 1969. A formula to express evapotranspiration as a function of soil moisture and evaporative demands of the atmosphere. Utah State University thesis.

Palmer, W.C., 1965. *Meteorologic drought.* Res. Pap. US Weather Bur., 45, p58.

Pasquill, F., 1950. Some further considerations of the measurements and indirect evaluation of natural evaporation. Q.J.Roy. Meteorol. Soc. 76: 287-301.

Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. *Proceedings, Royal Society, London* 193:120-45.

Priestley, C.H.B. and Taylor, R.J., 1972. On the assessment of the surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review*, 100, 81-92.

Pruitt, W.O., 1963. Application of several energy balance and aerodynamic evaporation equations under a wide range of stability. Final report to USAEPG, Univ. of California-Davis. pp107-124.

Renger, M., et al. 1974. Beurteilung bodenkundlicher, kulturtechnischer und hydrologischer Fragen mit Hilfe von klimatischer Wasserbilanz und bodenphysikalischen Kennwerten. *Z. Kulturtechnik u. Flurbereinigung* 15, 148-160.

Roberts, P.J.T., 1978. *A Comparison of the Performance of Selected Conceptual Models of the Rainfall-runoff Process in Semi-arid Catchments Near Grahamstown*, Report 1/78, Hydrological Research Unit, Rhodes University, Grahamstown.

Rohwer, C., 1931. Evaporation from free water surfaces. U.S. Department of Agriculture, Technical Bulletin 271, Washington, D.C.

Remanenko, V.A., 1961. Computation of the autumn soil moisture using a universal relationship for a large area. *Proc. of the Ukrainian Hydrometeorological Research Institute,* No.3, Kiev.

Rosenberg,N.J., et al., 1983. *Microclimate: The Biological Environmental*. Wiley, New York, 495p.

Saxton, K.E., and McGuinness, J.L. ,1982. 'Evapotranspiration'. In: *Hydrologic Modeling of Small Watersheds*, Haan, C.T., Johnson, H.P. and Brakensiek, D.L.,

Saxton, K.E., H.P. Johnson and R.H. Shaw, 1974. Modeling evapotranspiration and soil moisture. TRANSACTIONS of the ASAE 17(4): 668-672.

Schulz, E.F., 1962. A graphical procedure to estimate potential evapotranspiration by the Penman method. Report CER 62EFS49, Department of Civil Engineering, Colorado State University, Fort Collins, Colo.

Sill, B.L., 1981. Free and forced convection effects on evaporation. *Journal of Hydraulic Engineering 109(9):*1216-31.

Singh, V.P. and Xu, C-Y, 1997a. Evaluation and generalization of 13 equations for determining free water evaporation. *Hydrological Processes,* 11, 311-323.

Singh, V.P. and Xu, C-Y, 1997b. Sensitivity of mass transfer-based evaporation equations to errors in daily and monthly input data. *Hydrological Processes,* 11, 1465-1473.

Singh, V.P., 1989. Hydrologic Systems Vol.II Watershed Modelling, Prentice-Hall, Inc.

Sledgley, R.A., 1977. A simple evapotranspiration model for use in assessing hydrologic changes in catchments response to changing land use. *Proceedings, Hydrology and Water Resources Symposium, Institution of Engineers (Australia)*, Perth, Australia, 123-7.

Sutton, O.G., 1949. The application to Micrometeorology of the theory of turbulent flow over rough surfaces. *Royal Meteorological Society*, Quart. J., 75, No.236.

Sverdrup, H.U., 1946. The humidity gradient over the sea surface. *Journal of Meteorology* 3:1-8.

Tanner, C.B., 1966. Comparison of energy balance and mass transport methods for measuring evaporation. *Proceedings, ASAE Conference on Evapotranspiration and Its Role in Water Resources Management,* 45-48, Held December 5-6, in Chicago, Ill.

Thompson, O.E. and R.T. Pinker, 1981. An error analysis of the Thornthwaite-Holzman equations for estimating sensible and latent heat fluxes over crop and forest canopies. J. Appl. Meteorol. 20: 250-254.

Thornthwaite, C.W. and Holzman, B., 1939. The determination of land and water surfaces. *Monthly Weather Review 67*:4-11.

Thornthwaite, C.W. and Holzman, B., 1942. Measurement of evaporation from land and water surface. USDA Tech. Bull. 817: 1-75.

Thornthwaite, C.W. and Mather, J.R., 1955. The water balance. *Publ. Climatol. Lab. Climatol. Dresel Inst. Technol.*, 8(1):1-104.

Thornthwaite, C.W., 1944. A contribution to the report of the Cttee. on tarnspiration and evaporation, *Trans. AGU*, 25:686-693.

Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. Geogr. Rev., 38(1): 55-94.

Turc, L., 1961. Estimation of irrigation water requirements, potential evapotranspiration: a simple climatic formula evolved up to date. *Annals of Agronomy*, 12: 13-49.

Van Bavel, C.H.M., 1968. Discussion of "Climate and Evaporation from Crops." *Journal of the Irrigation and Drainage Division, Proceedings of the American Society of Civil Engineers* 68:533-5.

Visentini, M., 1936. Association international d'hydrologie scientifique. Bulletin No.22: 119-137.

Weisman, R.N., 1975. Comparison of warm water evaporation equations. *Journal of Hydraulics Division, Proceedings of the American Society of Civil Engineers 101(HY10)*:1303-13.

Winter, T.C., Rosenberry, D.O. and Sturrock, A.M., 1995. Evaluation of 11 equations for determining evaporation for a small lake in the north central United States. *Wat. Resour. Res.* 31(4):983-993.

Xu, C-Y and Singh, V.P., 1998. Dependence of evaporation on meteorological variables at different time-scales and intercomparison of estimation methods. *Hydrological Processes*, 12, 429-442.

Xu, C-Y and Singh, V.P., 2000. Evaluation and generalization of radiation-based methods for calculating evaporation. *Hydrological Processes* 14, 339-349.

Xu, C-Y and Singh, V.P., 2001. Evaluation and generalization of temperature-based methods for calculating evaporation. *Hydrological Processes* 15, 305-319.

Xu, C-Y and Singh, V.P., 2002. Cross-comparison of mass-transfer, radiation and temperature based evaporation models. *Water Resources Management* 16, 197-219.

Xu, C-Y, J. Seibert and S. Halldin, 1996. Regional water balance modelling in the NOPEX area: development and application of monthly water balance models, J. Hydrol. 180: 211-236.

Young, A.A., 1947. Evaporation from water surface in California: Summary of pan records and coefficients, 1881-1946. Bulletin 54, Public Works Department, Sacramento, California.

CHAPTER 5
RUNOFF IN HYDROLOGIC MODELS

5.1 INTRODUCTION

Runoff, which may be variously referred to as *streamflow, stream* or *river discharge*, or *catchment* or *watershed yield*, is normally expressed as a volume per unit of time. The m^3/s, i.e., one cubic meter per second. Runoff may also be expressed as a depth equivalent over a catchment, i.e., millimeters per day or month or year. This is a particularly useful unit for comparing precipitation and runoff rates and totals since precipitation is almost invariably expressed in this way.

Runoff or streamflow comprises the gravity movement of water in channels which may vary in size from the one containing the smallest ill-defined trickle to the ones containing the large river. In a general sense, this water representing the excess of rainfall over evapotranspiration, when allowance is made for storage on and under the ground surface.

Models of runoff processes have been developed for a wide variety of purposes, from the 'one-off' design of engineering structures and water supply systems to modern real-time models used continuously in river regulation schemes, they are also providing valuable for studying the potential impacts of changes in landuse or climate. Model outputs vary from predictions of peak discharges or total volumes of flood flow to the complete specification of the distribution of flow over time, either for individual storm events in event models or for continuous sequences of flows in continuous or sequential models. In this chapter, event-based models and continuous models of runoff processes are discussed separately and more details are placed on the former. This emphasis has been chosen because event-oriented models provide a direct means of continuous simulation. Continuous simulations can be obtained by specifying a continuous precipitation input.

5.2. SOURCES AND COMPONENTS OF RUNOFF

In order to select a method of modelling, it is essential to recognize the different runoff components and their regime. The different runoff terminologies used in the literature have resulted in much confusion and ambiguity about the sources and components. Ward (1972) provided a consistent and unambiguous terminology which has been adopted in this chapter. The total runoff from a typically heterogeneous catchment area may be conveniently divided into four component parts: channel precipitation, overland flow, interflow, and groundwater flow (see figure 5.1).

Channel precipitation

Direct precipitation onto the water surfaces of streams, lakes, and reservoirs makes an immediate contribution to streamflow. In relation to other components, however, this amount is normally small in view of the small percentage of catchment area normally covered by water surfaces. In catchments containing a large area of lakes or swamps channel precipitation may make a substantial contribution to streamflow.

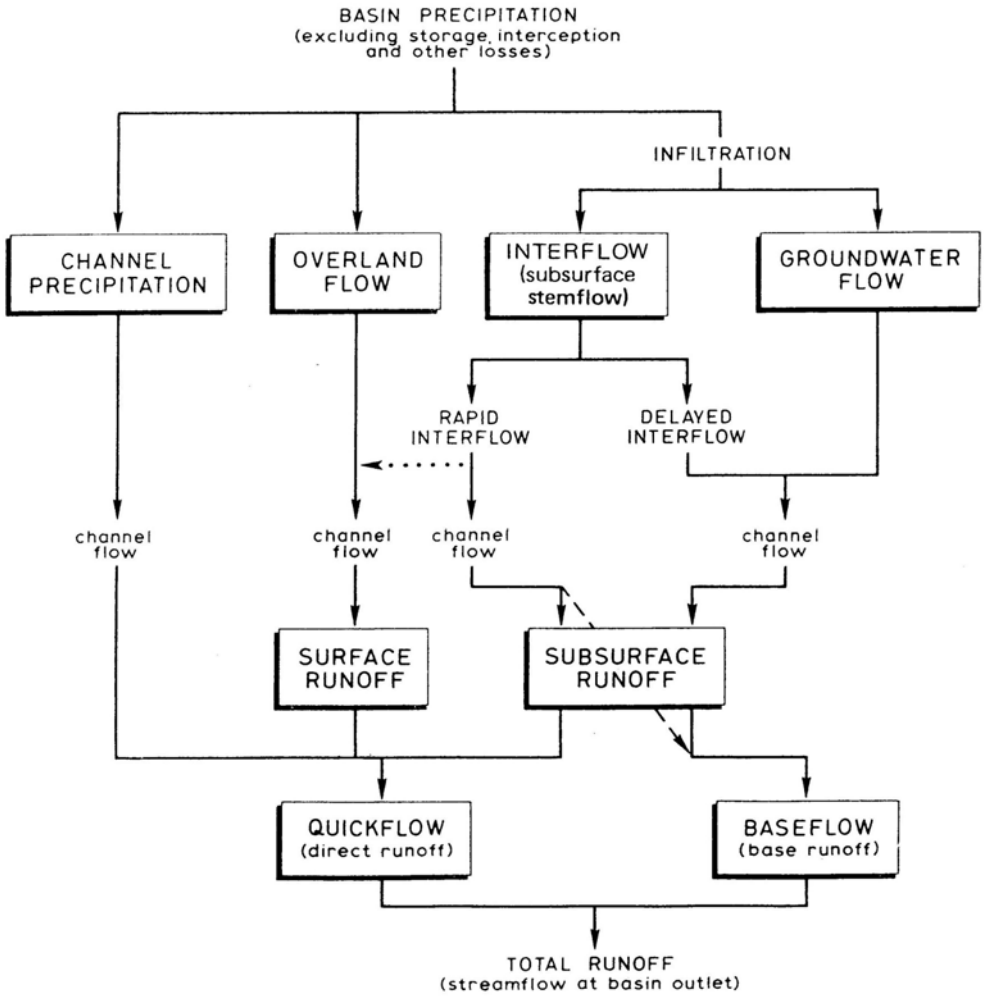


Fig.5.1 Diagrammatic representation of the runoff process (Ward, 1972)

Overland flow

Overland flow comprises the water which, falling to infiltrate the surface, travels over the ground surface towards a stream channel either as quasi-laminar sheet flow or, more usually, as flow anastomosing in small trickles and minor rivulets. The main cause of overland flow is the inability of water to infiltrate the surface and in view of the high value of infiltration characteristic of most vegetation-covered surfaces it is not surprising that overland flow is a rarely observed phenomenon (except on laboratory models!). Conditions in which it assumes considerable importance include the saturation of the ground surface, the hydrophobic nature of some very dry soils, the deleterious effects of many agricultural practices on infiltration capacity, and freezing of the ground surface. Surface runoff may then be defined as that part of the total runoff which travels over the ground surface to reach a stream channel and thence through the channel to reach the drainage basin outlet.

Interflow

Water which infiltrates the soil surface and then moves laterally through the upper soil horizons towards the stream channels, either as unsaturated flow or, more usually, as shallow perched saturated flow above the main groundwater level is known as interflow. Alternative terms found in the literature include subsurface storm flow, storm-seepage, and secondary base flow. The general condition favouring the

generation of interflow is one in which lateral hydraulic conductivity in the surface horizons of the soil is substantially greater than the overall vertical hydraulic conductivity through the soil profile. Then during prolonged or heavy rainfall water will enter the upper part of the profile more rapidly then it can pass vertically through the lower part, thus forming a perched saturated layer from which water will 'escape' laterally, i.e., in the direction of greater hydraulic conductivity. They may be several levels of interflow below the surface corresponding to textural changes between horizons and to the junction between weathered mantle and bedrock. In addition, some hydrologists argue that water may travel downslope through old root holes and animal burrows and other subsurface pipes. In view of the variety of possible interflow routes it is to be expected that some will result in more rapid movement of water to the stream channels than will others, so that it is sometimes helpful to distinguish between rapid and delayed interflow (see Fig.5.1). The experimental evidence has long indicated that the interflow may account for up to 85 per cent of total runoff (Ward, 1972).

Groundwater flow

Most of the rainfall which percolates through the soil layer to the underlying groundwater will eventually reach the main stream channels as groundwater flow through the zone of saturation. Since water can move only very slowly through the ground, the outflow of groundwater into the stream channels may lag behind the occurrence of precipitation by several days, weeks, or often years. Groundwater flow also tends to be very regular, representing as it does, the overflow from the slowly changing reservoir of moisture in the soil and rock layers. In general, groundwater flow represents the main long-term component of total runoff and is particularly important during dry spells when surface runoff is absent.

Snowmelt

In some areas, particularly at high altitudes or in high latitudes, a large proportion of streamflow may be derived from the melting of snows and glaciers. Although, in terms of the phase relationship between precipitation and runoff, snow accumulation and melt pose particular problems, in terms of the present discussion snowmelt does not represent a special case or merit consideration as a fifth component of runoff. Snow falling directly on to the stream surface has already been discussed under the heading of channel precipitation, while water generated by the process of snowmelt will either flow over the ground surface as overland flow or will infiltrate to become interflow and groundwater flow depending on whether the sub-snowpack surface is saturated and/or frozen.

Surface and subsurface runoff

The foregoing discussion should have clarified the definition and role of four other runoff terms, illustrated in Fig.5.1, which are used somewhat indiscriminately in the literature, i.e., surface and subsurface runoff and quickflow and baseflow. Surface runoff, as has been shown, is that part of total runoff which reaches the drainage basin outlet via overland flow and the stream channels, although it may in some circumstances also include interflow which has discharged at the ground surface at some distance from the stream channel. Subsurface runoff is the sum of interflow and groundwater flow and is normally equal to the total flow of water arriving at the stream as saturated flow into the stream bed itself, and as percolation from the seepage faces on the stream bank.

Quickflow and baseflow

Quickflow, or direct runoff, is the sum of channel precipitation, surface runoff and rapid interflow and will clearly represent the major runoff contribution during storm periods and is also the major contributor to most floods. It will be observed that quickflow and surface runoff as defined above cannot be used synonymously. Baseflow or base runoff may be defined as the sustained or fair-weather runoff and is the sum of groundwater runoff and delayed interflow, although some hydrologists prefer to include the total interflow as illustrated by the broken line in Fig.5.1. Again it will be observed that baseflow and groundwater flow, as defined above cannot be used synonymously; indeed Hewlett (1961, 1963) demonstrated that baseflow from steep mountain drainage basins may consist almost entirely of unsaturated lateral flow from the soil profile. Hewlett and Nutter (1969) suggested that in upland forested catchments about 85 per cent of total runoff might consist of baseflow.

The separations of runoff components can be done by using of hydrograph analysis, to this point refer to the course of "*Catchment Hydrology*".

5.3 APPROXIMATE INFILTRATION MODELS

Before modelling of various runoff components is discussed, modelling of infiltration process is described. Infiltration is the process of entry of water into a soil through the soil surface. For example, it is the infiltration capacity of the soil that determines for a given storm, the amount and time distribution of rainfall excess that is available for runoff and surface storage. Hence, an understanding of infiltration and factors affecting it is important to the determination of surface runoff as well as the subsurface movement and storage of water within the catchment. Philip (1969) and Morel-Seytoux (1973) have presented excellent reviews of the infiltration processes. Infiltration can be characterized for most initial and boundary conditions of interest by solving the governing differential equations using numerical methods. Although these solutions provide a physically consistent means of quantifying infiltration in terms of soil properties governing movement of water and air, such elaborate procedures are rarely used in practice. One reason is that numerical solutions are usually expensive due to computational requirements. A more severe limitation is the difficulty of obtaining necessary soil property data. Attempts to characterize infiltration for field applications have usually involved simplified concepts which permit the infiltration rate or cumulative infiltration volume to be expressed algebraically in terms of time and certain soil parameters. Some of the approximate models have been developed by applying principles governing soil water movement for simplified boundary and initial conditions. The parameters in such models can be determined from soil water properties, when they are available. Other models are strictly empirical and the parameters must be obtained from measured infiltration data or estimated using more approximate procedures. The most obvious characteristics of the infiltration process are that for ponded surfaces the rate decreases rapidly with time during the early part of an infiltration event. Although attributed to different physical phenomena, this characteristic is reflected by all of the approximate infiltration equations.

Considerable literature exists on approximate infiltration models. This section presents, as an example, one empirical and one conceptual infiltration models. Physically based models employing the Richards equation will be discussed separately in Chapter 8 together with the SHE model.

We define a number of terms that are to be used in this section and give their notations and dimensions here.

- The *infiltration rate*, *f* is the rate at which water enters into the soil surface. It is expressed as volume per unit area per unit time and has the dimensions of length per unit of time.
- Cumulative *infiltration, F* denotes the volume of infiltration from the beginning of time t, or the rainfall event. It also is called infiltration volume or accumulated infiltration and is measured in centimetres.
- *Infiltration capacity,* f_p is the maximum rate at which soil can absorb water through its surface and has the dimensions of length per unit of time. A distinction should be made between f and f_p; $0 \le f \le f_p$.

Horton Model

The Horton model (Horton, 1939, 1940) is one of the best-known infiltration models in hydrology. Horton recognised that the infiltration capacity decreased with time until it approached a more-or-less constant rate. He expressed the view that the decrease in infiltration capacity was controlled probably more by the factors operating at the soil surface than by the flow process in the body of the soil.

Horton (1940) hypothesized that infiltration is similar to exhaustion process according to which the rate of performing work is proportional to the amount of work remaining to be performed. In case of infiltration, the work remaining to be performed at any time t is equal to that of changing the infiltration rate f to its ultimate constant value f_c. The rate of performing work is *df/dt*. The amount of work remaining to be performed is $(f - f_c)$. Since f decreases with t,

$$\frac{df}{dt} = -k(f - f_c) \tag{5.1}$$

where k is a proportionality factor dependent on soil type and initial moisture content. The initial condition is:

When t = 0, f = f_0,

Equation (5.1) can be integrated to yield

$$f = f_c + (f_o - f_c)e^{-kt} \tag{5.2}$$

and

$$\frac{dF}{dt} = f_c + (f_o - f_c)e^{-kt} \qquad \text{where } f = dF/dt \tag{5.3}$$

The initial condition is: When t = 0, F = 0. Then

$$\begin{aligned} F &= f_c t + \frac{1}{k}(f_o - f_c)(1 - e^{-kt}) \\ &= f_c t + \frac{1}{k} f_o - \frac{1}{k} f_c - \frac{1}{k} f_o e^{-kt} + \frac{1}{k} f_c e^{-kt} \\ &= f_c t + \frac{1}{k}(f_o - f_c) - \frac{1}{k}(f_o - f_c)e^{-kt} \end{aligned} \tag{5.4}$$

This model is simple and fits well to experimental data. The principal weakness of this model lies in determination of its parameters f_o, f_c, and k. These have to be determined by data fitting. The parameters f_o, f_c, and k can be estimated in two ways (see Singh, 1989).

- First taking the logarithm of equation (5.2)

$$\ln(f - f_c) = \ln(f_o - f_c) - kt \tag{5.5}$$

Equation (5.5) represents a straight line on a semilog plot whose slop, -k, and intercept, $\ln(f_o - f_c)$, can readily be determined (Toebes, 1963). For given infiltration data, f_c is taken to be the lowest value of f where it tends to become constant. The value of $(f - f_c)$ at t = 0 is $(f_o - f_c)$.

- Second, a least squares method can be used (Blake, et al., 1968) directly to estimate parameters of equation (5.3), which is of the form

$$F = a_0 + a_1 t - a_0 e^{-kt}, \qquad a_0 = \frac{1}{k}(f_o - f_c), \qquad a_1 = f_c$$

This equation can be fitted to the experimental data. During the last 5 min or so (assuming that the experiment is continued until the runoff becomes constant), the exponential term becomes small. Therefore,

$$F = a_0 + a_1 t$$

the values of a_o and a_1 can be estimated using the data for this time interval.

To determine k, we can write

$$-kt = \ln \frac{F - a_o - a_1 t}{-a_o}$$

By applying the linear regression analysis to the remainder of the data, excluding those for which

$$F < (a_0 + a_1 t)$$

the parameter k can be obtained.

The Green-Ampt method:

A simple conceptual model, based on Darcy's law, was proposed by Green and Ampt (1911). For details about the assumptions, deriving procedure, etc., refer to the course "*Process hydrology*". The section presents only the calculation procedure of the method (see also Chow et al., 1988):

Following the flow chart (Fig.5.2):

Step 1: Calculate the current potential infiltration rate, f_t from the known value of cumulative infiltration F_t. For the Green-Ampt method, one uses

$$f_t = K\left(\frac{\Psi\Delta\theta}{F_t}+1\right) \tag{5.6}$$

where:
θ is the moisture content, i.e. the ratio of the volume of water to the total volume,
$\Delta\theta = (1-s_e)\theta_e$ is the change of moisture content,
s_e is the effective saturation,
θ_e is the effective porosity
Ψ is suction head,
K is hydraulic conductivity

Step 2: Calculation of Cumulative infiltration, F_t

Case 1: If $f_t \le i_t$ then (ponding occurs throughout the interval), and

$$F_{t+\Delta t} = F_t + \Psi\Delta\theta \ln\left[\frac{F_{t+\Delta t}+\Psi\Delta\theta}{F_t+\Psi\Delta\theta}\right] = K\Delta t \tag{5.7}$$

Case 2: if $f_t > i_t$ and no ponding at the beginning of the interval. Assume that this remains so throughout the interval, then the infiltration rate is i_t and a tentative value for cumulative infiltration at the end of the time interval is

$$F'_{t+\Delta t} = F_t + i_t\Delta t \tag{5.8}$$

Step 3: Calculate a corresponding infiltration rate $f'_{t+\Delta t}$ from $F'_{t+\Delta t}$.

If $f'_{t+\Delta t}$ is greater than i_t, then $F_{t+\Delta t} = F'_{t+\Delta t}$, no ponding throughout the interval.

If $f'_{t+\Delta t}$ is less than or equal to i_t, ponding occurs during the interval. The cumulative infiltration F_p at ponding time is found by setting $f_t = i_t$ and $F_t = F_p$ in equation (5.6) and solving for F_p to give, for the Green-Ampt equation,

$$F_p = \frac{K\Psi\Delta\theta}{i_t - K} \tag{5.9}$$

The ponding time is then $t + \Delta t'$, where

$$\Delta t' = \frac{F_p - F_t}{i_t} \tag{5.10}$$

and the cumulative infiltration $F_{t+\Delta t}$ is found by substituting $F_t = F_p$ and $\Delta t = \Delta t - \Delta t'$ in equation (5.7).

The excess rainfall values are calculated by subtracting cumulative infiltration and other losses from cumulative rainfall and then taking successive differences of the resulting values. Direct runoff can then be calculated from excess rainfall.

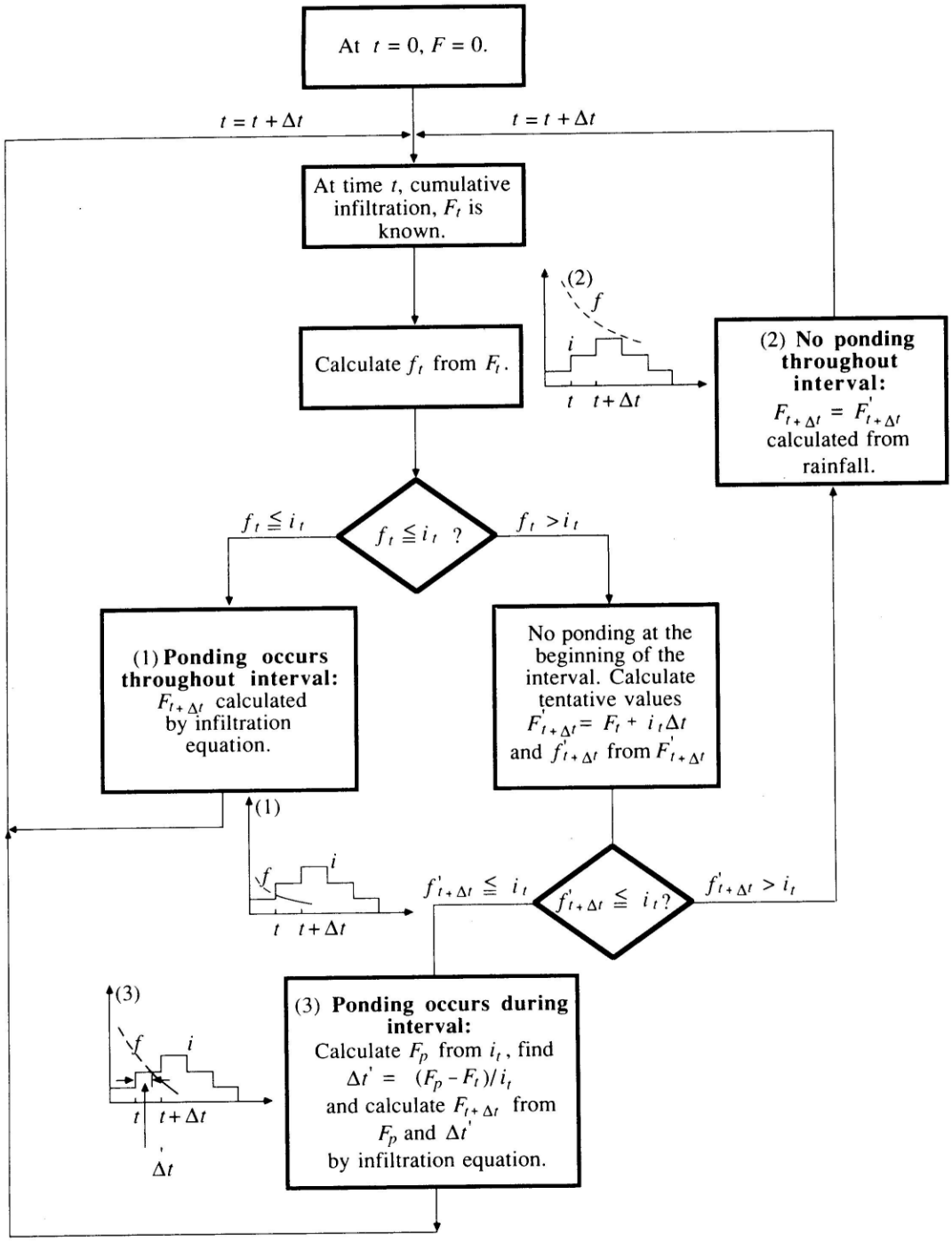


Fig.5.2 Flow chart for determining infiltration and ponding time under variable rainfall intensity (after Chow et al., 1988).

5.4 SIMULATING RUNOFF WITH LUMPED MODELS

The number of runoff components to be analysed depends on the characteristics of the basin and the objective of the separation including the time base to be considered (Figure 5.1). In most conceptual lumped catchment models, the following four components (left) or even the right two components may be identified and modelled explicitly:

surface flow fast interflow	fast components
slow interflow base flow	slow components

A lumped approach to the modelling of runoff considers the catchment as a spatially singular entity which transforms rainfall excess into an outflow hydrograph. The approach ranges from the use of a mathematical transfer function, or "black-box" approach, to a modelling of the detailed interrelationship of processes for hydrologically significant phenomena.

5.4.1 Event-based models

Streamflow simulation for individual storms is needed to solve a wide variety of water resources problems, including design of hydraulic structures such as dams, culverts, bridges, spillways; urban and highway drainage; planning of flood-control works; source pollution; disposal of waste material; evaluation of environmental impacts of land use and management practices; and planning of soil conservation works.

Many event-based streamflow simulation models have been developed (Renard et al., 1982; Singh, 1988). Some of the models are summarised in Table 5.1. The objective of this section is to present the general concepts of event-based streamflow simulation and then make some comments about the models. No attempt is made to describe all the models. Complete details of the models can be found in the cited references.

5.4.1.1 Determination of effective/excess rainfall hyetograph (ERH)

Excess rainfall, or effective rainfall, is that rainfall which is neither retained on the land surface nor infiltrated into the soil. The graph of excess rainfall vs. time, or excess rainfall hyetograph (ERH), is a key component of the study of rainfall runoff relationships. The excess rainfall hyetograph (ERH) may be determined from rainfall (hyetograph) data in one of two ways, depending on whether streamflow data are available for the storm or not.

In case the rainfall and streamflow data area available, the ERH can be determined by using the Φ–index method which consists of the following steps:

- Estimate the baseflow by applying a base flow separation method.
- Calculate the direct runoff which equals to the observed streamflow minus baseflow.
- Compute the volume and depth of direct runoff, V_d and r_d. And

$$V_d = \sum_{i=1}^{1} Q_{d,i}\Delta t \qquad r_d = \frac{V_d}{catchment\ area} \tag{5.11}$$

where Q_d is the ordinate of the direct runoff.

- Estimate the rainfall loss rate Φ, and M, the number of nonzero pulses of excess rainfall, by solving the following equation with a trial and error method (an example can be sound on page 138 of Chow et al, 1988).

$$r_d = \sum_{m=1}^{M} (R_m - \Phi\Delta t) \tag{5.12}$$

where R_m is the rainfall depth in pulse m.

- Calculate the excess rainfall hyetograph (ERH). The ordinates of ERH are found by subtracting $\Phi\Delta t$ from the ordinates of the observed rainfall hyetograph.

In case, the streamflow data are not available, the ERH can be calculated from either the SCS curve-number method or infiltration abstraction method (Chow et al., 1988).

5.4.1.2 Computation of direct runoff hydrograph (DRH)

At this stage a decision has to be made regarding the type of the model to be used in light of available data. This approach ranges from the use of a mathematical transfer function, or "block-box" approach, such as the time-area method, and unit hydrograph method (For details refer to the course of “Process Hydrology”), to conceptual method, such as the linear-channel and linear-reservoir (Nash, 1957).

Suppose, a linear conceptual model is chosen for this purpose. Then the IUH (instantaneous unit hydrograph) for this model must be derived as well as a method for estimating the IUH parameters. Thereafter, the IUH is convoluted with the ERH estimated already to obtain the DRH. To summarise, three tasks are performed for this choice: (a) computation of the IUH, (b) determination of the IUH parameters, and (C) convolution of the IUH with the ERH.

5.4.1.3 The Instantaneous unit hydrograph

If the excess rainfall is of unit amount and its duration is infinitesimally small, the resulting hydrograph is an impulse response function called the instantaneous unit hydrograph (IUH). For an IUH, the excess rainfall is applied to the drainage area in zero time. Of course, this is only a theoretical concept and cannot be realized in actual catchments, but it is useful because the IUH characterizes the catchment’s response to rainfall without reference to the rainfall duration. Therefore, the IUH can be related to catchment geomorphology.

The convolution integral is

$$Q(t) = \int_0^t u(t-\tau)I(\tau)d\tau \tag{5.13}$$

If the quantities I(τ) and Q(t) have the same dimensions, the ordinate of the IUH must have dimensions $[T^{-1}]$. The properties of the IUH are as follows, with $l = t - \tau$.

$0 \le u(l) \le$ some positive peak value for $l > 0$
$u(l) = 0$ for $l \le 0$
$u(l) \to 0$ as $l \to \infty$
$\int_0^\infty u(l)dl = 1$ and $\int_0^\infty u(l)ldl = t_L$

The quantity t_L is the lag time of the IUH. It can be shown that t_L gives the time interval between the centroid of an excess rainfall hyetograph and that of the corresponding direct runoff hydrograph.

The IUH can be determined by various methods of mathematical inversion, using, for example, orthogonal functions such as Fourier series (O'Donnell, 1960) or Laguerre functions (Dooge, 1973); integral transforms such as the Laplace transform (Chow, 1964), the Fourier transform (Blank et al., 1971), and the Z transform (Bree, 1978). The complexity of the methods has deterred practising engineers from applying them in day-to-day problems. Techniques incorporating catchment parameters have had their attractions. These methods, based on unit hydrograph theory, are more akin to mathematical models, and will be discussed hereafter.

5.4.1.4 The Nash linear conceptual model

In view of the difficulty of directly deriving instantaneous unit hydrographs (IUHs), the influence of a catchment in transferring rainfall excess into direct runoff was viewed conceptually as equivalent to transferral through a series of reservoirs linked by channels (Nash, 1957).

- **A linear reservoir** is a conceptual reservoir in which the storage, S, is directly proportional to the outflow, Q, or

$$S = KQ \tag{5.14}$$

The proportionality constant K, is known as the storage coefficient. The difference between inflow (I) and outflow (Q), is the time rate of change in storage, i.e., by continuity,

$$I - Q = dS/dt \tag{5.15}$$

Substituting equation (5.14) in equation (5.15),

$$I - Q = KdQ/dt \tag{5.16}$$

or

$$dQ/dt + Q/K = I/K \tag{5.17}$$

the solution for which is:

$$Q = I(1-e^{-t/K}) \tag{5.18}$$

If the inflow stops at time t = t_0 at which time the outflow Q = Q_0, then from equation (5.16) with I = 0 and $\tau = t - t_0$,

$$dQ / d\tau + Q / K = 0 \tag{5.19}$$

for which the solution is:

$$Q = Q_o e^{-\tau/K} \tag{5.20}$$

For an instantaneous inflow which fills a reservoir of storage S_0 in time $t_0 = 0$,

$$Q_0 = S_0/K \tag{5.21}$$

combining equations (5.20)and (5.21), and since $\tau = t$

$$Q = (S_o / K) e^{-t/K} \tag{5.22}$$

The IUH for a linear reservoir, in which S_0=1 and inflow is instantaneous, is

$$u(t) = (1/K) e^{-t/K} \qquad \text{for } t \geq 0 \tag{5.23}$$
$$u(t) = 0 \qquad \text{fot } t < 0$$

as shown in Fig.5.3 (left). If the inflow were a unit pulse of duration D, the UH u(D, t), as shown in Fig.5.3(right), and the direct runoff Q(t) (ignoring the dimensions of u(D, t) and Q(t)) would be

$$Q(t) = u(D;t) = I\left[\ 1 - e^{-t/k}\ \right], \qquad t \leq D \tag{5.24}$$

and

$$Q(t) = u(D;t) = Q_p e^{-(t-D)/k} \qquad t > D \tag{5.25}$$

where Q_p represents the hydrograph peak and is given by

$$Q_p = I\left[\ 1 - e^{-D/k}\ \right] \tag{5.26}$$

Here I = 1/D. the hydrograph peak will occur at the end of the duration of unit pulse.

As $t \to \infty$, $Q(t) = I(t)$. This implies an equilibrium conditions: outflow becoming equal to inflow. As $t \to 0$, Q = 0. As I(t) terminates at t = D, the recession starts immediately. For an instantaneous inflow, which fills the reservoir of storage S in t = 0, Q_p = S/k. The equation of outflow is simply

$$Q(t) = \frac{S}{k} e^{-t/k} \tag{5.27}$$

For a unit inflow, $S = 1$ and $I(t) = \delta(t)$. Consequently $Q(t) \to u(t)$, the IUH is the same as given by equation 5.23.

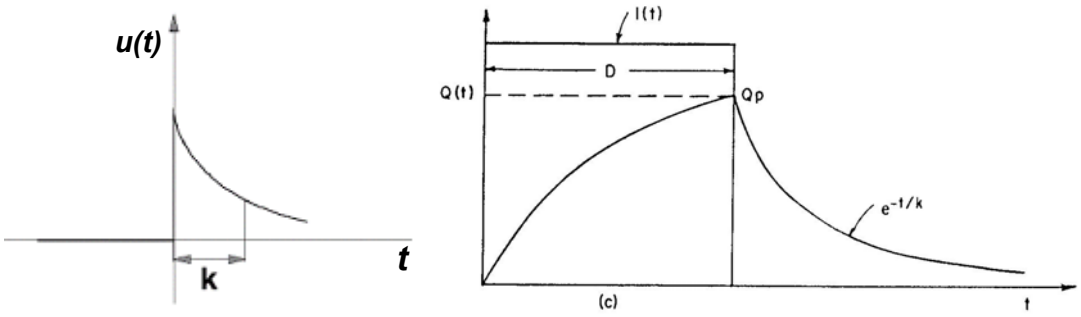


Fig.5.3 A linear reservoir: hydrograph due to a pulse of instantaneous input (left); hydrograph due to a pulse of duration D hours duration (right).

- **A linear channel** is a fictitious channel in which the cross-sectional flow area at a section is proportional to the discharge, or

$$A = CQ \tag{5.28}$$

The proportionality constant, C, is known as the translation coefficient. The velocity at a section of the channel therefore remains constant, but may vary from section to section. An inflow hydrograph or excess rainfall hyetograph routed through a linear channel remains unchanged in shape and is merely translated in time, i.e., y(t) = x(t-T) and this is shown in figure 5.4.

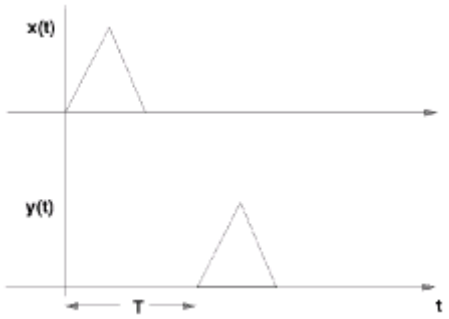


Figure 5.4. A linear channel which is just a translation without changing of the hydrograph.

- **Cascaded reservoirs**: A conceptual model of a catchment having the same hydrologic response as a series of n linear reservoirs, each having the same storage coefficient, K, was formulated by Nash (1957). This model has proved to be a simple but effective method for deriving catchment IUHs.

As shown diagrammatically in Fig.5.5, a hydrograph-shaped outflow curve is developed and modified by successively routing the outflow from one reservoir as inflow to the next lower reservoir.

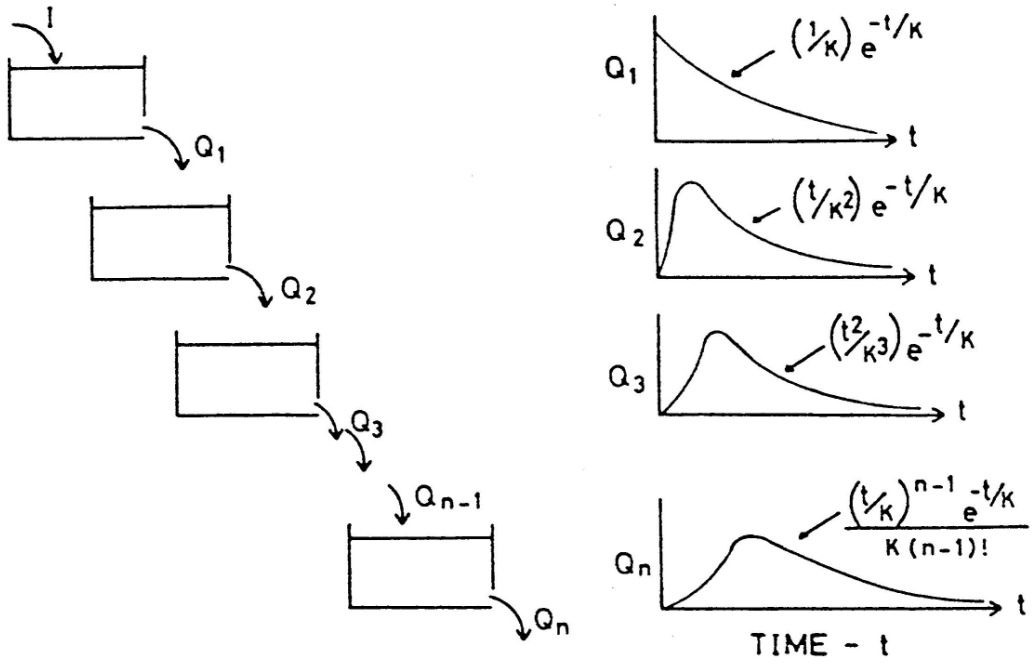


Fig.5.5 Cascaded reservoirs of the Nash model. Catchments having considerable impervious area, such as small urban catchments, commonly exhibit IUHs of the one or two reservoir shape due to the rapid response. Conversely, large flat agricultural catchments having little channel formation exhibit the considerable lag of a large number of reservoir routings.

Assuming instantaneous unit input into the initial reservoir, the outflow from this reservoir (previously developed as equation (5.23) may in turn be considered as input to the second reservoir. Using τ as the variable in the convolution integral, the outflow from the second reservoir may be obtained as

$$Q_2 = \int_0^t I(\tau)u(t-\tau)d\tau = \int \left|(1/K)e^{-\tau/K}\right\|(1/K)e^{-(t-\tau)/K}\left|d\tau\right.$$
$$= (t/K^2)e^{-t/K} \tag{5.29}$$

For n repetitions of the above convolution, the generalised formula for the IUH of the conceptualized drainage basin may be derived as

$$Q_n = u(t) = \frac{1}{K(n-1)!}(t/K)^{n-1}e^{-t/K} \tag{5.30}$$

in which the value *n* is not necessarily an integer. When n is not an integer, (n-1)! is replaced by $\Gamma(n)$ in equation (5.30). $\Gamma(n)$ can be interpolated from tables of the gamma function. This equation expresses the instantaneous unit hydrograph of the proposed model; mathematically, it is a gamma probability distribution function. The integral of the right side of the equation over t from zero to infinity is equal to 1.

The two parameters, *K*, the storage constant for each of the reservoirs and *n*, the number of reservoirs, may be simply evaluated by taking incremental moments of the

excess rainfall hyetograph (ERH) and the direct runoff hydrograph (DRH), and substituting in the formulas

$$M_{D1} - M_{E1} = nK \tag{5.31}$$

and

$$M_{D2} - M_{E2} = n(n+1)K^2 + 2nKM_{E1} \tag{5.32}$$

in which:

M_{D1} = first moment of the DRH about the time origin divided by the total direct runoff
M_{D2} = second moment of the DRH about the time origin divided by the total direct runoff
M_{E1} = first moment of the ERH about the time origin divided by the total effective rainfall
M_{E2} = second moment of the ERH about the time origin divided by the total effective rainfall

The value nK, as indicated by equation (5.31), represents the time lag between centroids of the rainfall and runoff curves.

An example: Given the ERH and the DRH shown in Fig.5.6, determine n and K for the IUH.
Solution: Determine the moments of the excess rainfall hyetograph and the direct runoff hydrograph. Each block in the ERH and DRH has duration 6 h = 6×3600 s = 21600 s. The rainfall has been converted to units of m^3/s by multiplying by the watershed area to be dimensionally consistent with the runoff. The sum of the ordinates in the ERH and in the DRH is 700 m^3/s, so the area under each graph = 700 × 6 =4200 (m^3/s) h.

$$M_{E1} = \sum\left[\frac{\text{incremental area} \times \text{moment arm}}{\text{total area}}\right]$$

$$= \frac{6}{4200}\left[100\times 3 + 300\times 9 + 200\times 15 + 100\times 21\right]$$

$$= 11.57 \text{ h}$$

The second moment of area is calculated using the parallel axis theorem.

M_{E2} = {∑[incremental area ×(moment arm)^2]

+ ∑[second moment about centroid of each increment]}/total area

$$= \frac{6}{4200}\{\left[100\times 3^2 + 300\times 9^2 + 200\times 15^2 + 100\times 21^2\right]$$

$$+ \frac{1}{12}6^3\left[100 + 300 + 200 + 100\right]\}$$

$$= 166.3 \text{ h}^2$$

By a similar calculation for the direct runoff hydrograph

$M_{D1} = 28.25$ h
$M_{D2} = 882.8$ h^2

Solve for nK using (5.31):

$$\begin{aligned} nK &= MD1 - ME1 \\ &= 28.25 - 11.57 \\ &= 16.68 \end{aligned}$$

Solve for n and K using (5.32):

$$\begin{aligned} M_{D2} - M_{E2} &= n(n+1)K^2 - 2nKM_{E1} \\ &= n^2K^2 + nK \times K + 2nKM_{E1} \end{aligned}$$

Hence

$$882.8 - 166.3 = (16.68)2+16.68K+2\times16.68\times11.57$$

and solving yields K = 3.14 h

Thus

$$\begin{aligned} n &= 16.68/K \\ &= 16.68/3.14 \\ &= 5.31 \end{aligned}$$

These values of n and K can be substituted into Eq. (5.30) to determine the IUH of this watershed. By using the methods described in Process Hydrology Course, the corresponding unit hydrograph can be determined for a specific rainfall duration. Surface (direct) runoff hydrograph of any rainfall event can then be calculated.

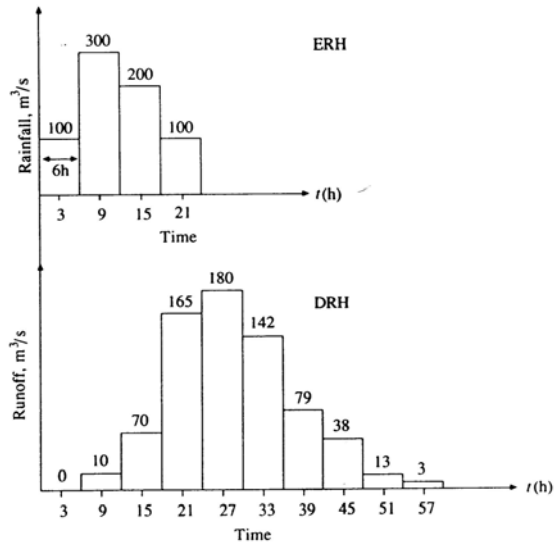


Fig.5.6 Excess rainfall hyetograph (ERH) and direct runoff hydrograph (DRH) for calculation of n and K in a linear reservoir model (from Chow et al., 1988).

5.4.2 Flow routing

Depending upon the drainage pattern and the existence of dams or reservoirs, bridges, and the like, within the basin, the calculated direct runoff hydrograph, DRH of each sub-basin is to be routed during its journey to the catchment outlet. Here again, a decision has to be made about the routing method. Suppose that the Muskingum method is chosen for flow routing through channels, and Level-pool method is used for routing through reservoirs. Then the parameters of the Muskingum method (weighting factor, X, and lag time, K) are estimated first for each channel reach. Because this model is linear, it will be sufficient to obtain its IUH and then convolute it with the reach inflow hydrograph, which is usually the DRH, to compute the reach-outflow hydrograph. For reservoir flow routing, storage-elevation and outflow-elevation relationships must be established for each reservoir. These relationships are then combined to form a storage-outflow graph, which is then used to perform flow routing through the reservoir. These and other reservoir flow routing and channel flow routing methods are discussed in the course of "*Catchment hydrology*".

5.4.3 Continuous streamflow simulation models

Continuous streamflow simulation (CSS) has many applications, such as (1) extending streamflow records, (2) flow forecasting, (3) evaluating the effect of land-use practices on catchment response, (4) design urban drainage, highway culverts, reservoirs, and the like, water-quality modelling, (5) irrigation planning and management. An extensive listing of applications of the CSS models is given by James et al (1982).

5.4.3.1 Comparison with event based models

As the name suggests, CSS models allow simulation of streamflow for long periods of time and thus more fully utilise capability of the digital computer. These models maintain a more or less continuous accounting of the water in storage in the catchment. Because of long periods of time, such hydrologic processes as evaporation and transpiration, infiltration, interception, depression storage, subsurface flow, and baseflow assume added significance. These processes are calculated separately in most conceptual models. In event-based streamflow simulation some of these processes are neglected, some are lumped, and some are considered with considerable approximation, for the period of simulation is usually as long as the duration of the DRH. The emphasis in CSS is on simulation of the entire land phase of the hydrologic cycle, whereas the emphasis in EBSS is on modelling the DRH or its peak characteristics. Thus, CSS models are models of the hydrologic cycle, whereas EBSS models are models of rainfall-runoff cycle. It is logical to say that CSS models are more general and encompass EBSS models as their special cases.

Naturally then, discussion of CSS models involves what has already been presented about EBSS models plus discussion of components not included in EBSS models, such as, evapotranspiration (chapter 4), snowmelt (chapter 3), subsurface runoff (groundwater runoff plus delayed interflow, to be discussed in the section), soil moisture storage (to be discussed in this section).

Table 5.1

Some event-based streamflow simulation models

Model			Model Components						
Name	*Author(s)*	*Year*	*Baseflow Separation*	*DR Volume*	*Infiltration and Loss*	*DR Hydrograph*	*Channel Routing*	*Reservoir Routing*	*Parameter Optimization*
HEC-1	Hydrologic Engineering Center	1981, 1982	Yes	SCS curve number and two other methods	Variable loss rate method	Clark's and Snyder's unit hydrograph methods	Muskingum method and five other methods	Storage-indication method	Automatic calibration capability
TR-20	Soil Conservation Service	1973	Constant rate method	SCS curve number method	SCS curve number method	Unit hydrograph method	Convex method	Storage-indication method	No
USGS	Dawdy et al.	1972	Constant rate method	Soil moisture accounting	Philip equation	Clark's unit hydrograph	Translation method	No	Rosenbrock's method
HYMO	Williams and Hann	1973	No	SCS curve number method	SCS curve number method	Nash model	Variable storage coefficient method	Storage-indication method	No
SWMM	Metcalf and Eddy, Inc. et al.	1971	No	Loss accounting	Horton's equation	Hydraulic method	Hydraulic routing method	No	No
WAHS	Singh	1983	Recession equation	SCS curve number method	Philip's equation	Geomorphological unit hydrograph method	Linear reservoir	No	Rosenbrock-Palmer method
RORB	Laurenson and Mein	1983	Two options	No	Constant and variable loss rate methods	Nonlinear storage routing	Nonlinear storage routing	Yes	No
WBNM	Boyd et al.	1979a, 1979b	No	Yes	Φ-index	Linear as well as storage elements for routing	Storage routing	No	Yes
FHSM	Foroud and Broughton	1981	Yes	Yes	Modified Horton's equation	Time area curve + a linear reservoir	No	No	Nonlinear least square curve fitting
XJM	Zhao et al.	1980	Yes	Yes	Storage capacity curve	Unit hydrograph method	Muskingum method	No	No
GAWSER	Ghate and Whiteley	1977, 1982	Yes	Yes	Holtan's equation	Time area curve + convolution	HYMO method	No	No
MIT	Maddaus and Eagleson	1969	No	No	Any suitable model	Linear channel and reservoir	Linear	No	Optimization
HM	Huggins and Monke	1968	No	Yes	Holtan's equation	Kinematic wave method	No	No	No
Kansas	Smith and Lumb	1966	Yes	Yes	Soil moisture accounting	Lag and route method	No	No	No
IHM	Morris	1980	Yes	Yes	Richards equation	St. Venant equation	St. Venant equation	No	No

5.4.3.2 Simulation of subsurface runoff

In general, quantitative simulation of interflow and baseflow is much less advanced than is the simulation of either infiltration or movement of subsurface water within its own domain. Further, except in few physically-based models, almost no one has tried to include physically-based simulation of subsurface flow in general watershed models.

Fleming (1975) briefly discussed subsurface processes and how hydrologic models commonly handle them. The usual method is to consider subsurface water as resident in one or more storages or reservoirs. The following storage concepts might be applied:

single linear reservoir	$S = K'Q$	$\rightarrow$	$Q = K \cdot S$
single logarithmic reservoir	$S = K'\ln Q$	$\rightarrow$	$Q = K \cdot e^{S}$
single nonlinear reservoir	$S = K'Q^{m}$	$\rightarrow$	$Q = K \cdot S^{1/m}$

where S = storage; Q = reservoir outflow (discharge); K and K′ = storage constants; and m = exponent.

The storage is usually updated by a balance equation, which is usually a simple accounting of inflows and outflows,

$$S_t = S_{t-1} + I\Delta t - O\Delta t \tag{5.33}$$

where

S_t = total water in storage at time t

I = a summation of such inflow rates as infiltration or inflowing seepage.

O = a summation of such outflow as evapotranspiration, outflowing seepage (baseflow, interflow, etc.)

Subsurface outflow rates are usually expressed as functions of the amount of subsurface water remaining in storage.

For example, in the HBV model, the interflow and groundwater flow are calculated by the following equations, respectively:

$$Q_1 = K_1 S_{uz} \tag{5.34}$$
$$Q_2 = K_2 S_{lz} \tag{5.35}$$

where Q_1, Q_2 = Runoff components, K_1 and K_2 = recession coefficients (parameters), S_{uz} and S_{uz} are the storages at upper zone and lower zone, respectively.

where

The Stanford Watershed model computes the interflow on a 15-min time interval according to equation:

$$q_i = \left\{1.0 - (IRC)^{1/96}\right\} \cdot SRGX \tag{5.36}$$

where

q_i = interflow volume entering the channel during a 15-min time interval

IRC = daily recession rates of interflow (1/96 converts to 15-min interval)

SRGX = volume of interflow storage.

The ground water discharge for each 15-min time interval is computed by:

$$G_g = (1-(KK24)^{1/96}(1+KU \cdot S)S_{gw} \quad (5.37)$$

where

KK24 = minimum observed daily ground water recession constant
KU = variable ground water recession parameter
S = ground water slope
S_{gw} = ground water storage

More equations used in other models will be discussed in Chapter 8 while discussing the particular models.

5.4.3.3 Building a continuous streamflow simulation model

From the preceding discussion, it is clear that building a catchment model involves modelling the various components of the hydrologic cycle and maintaining a continuous water balance involving these components. Some of these components are interactive and involve iterative calculation. Larson et al. (1982) presented a good discussion on assembling these components into a catchment model. Fig.5.7 shows a general conceptual framework for building a catchment model. Many of the models, as summarized in Table 5.2, posses similar arrangements of components.

5.5 DISTRIBUTED MODELING OF RUNOFF PROCESSES

A truly distributed hydrologic model would require the development and solution of a comprehensive set of partial differential hydrodynamic and porous media flow equations. The solution of such equations is highly boundary value dependent. A detailed description of the infinite variety of boundary conditions present in a natural watershed is not currently feasible. Therefore, those models that are currently classified as distributed parameter models only approximate this approach.

Models can be classified as distributed when they utilize data concerning the spatial distribution of controlling parameter variations in conjunction with computational algorithms to evaluate the influence of this distribution on simulated behaviour. Such models attempt to increase the accuracy of the simulation by preserving and utilizing information concerning the areal distribution of all spatial non-uniform processes characterized by the model. This increased accuracy usually comes at the expense of increased computational and data preparation effort. The ready availability on Internet and CD-ROM of data describing the land surface, especially digital elevation data for land surface terrain, has made it practical for the first time to delineate catchments in a few minutes in an automated way, and to compute the hydrologic properties of those catchments.

More details about distributed models will be discussed in chapter 8 using the SHE model as an example.

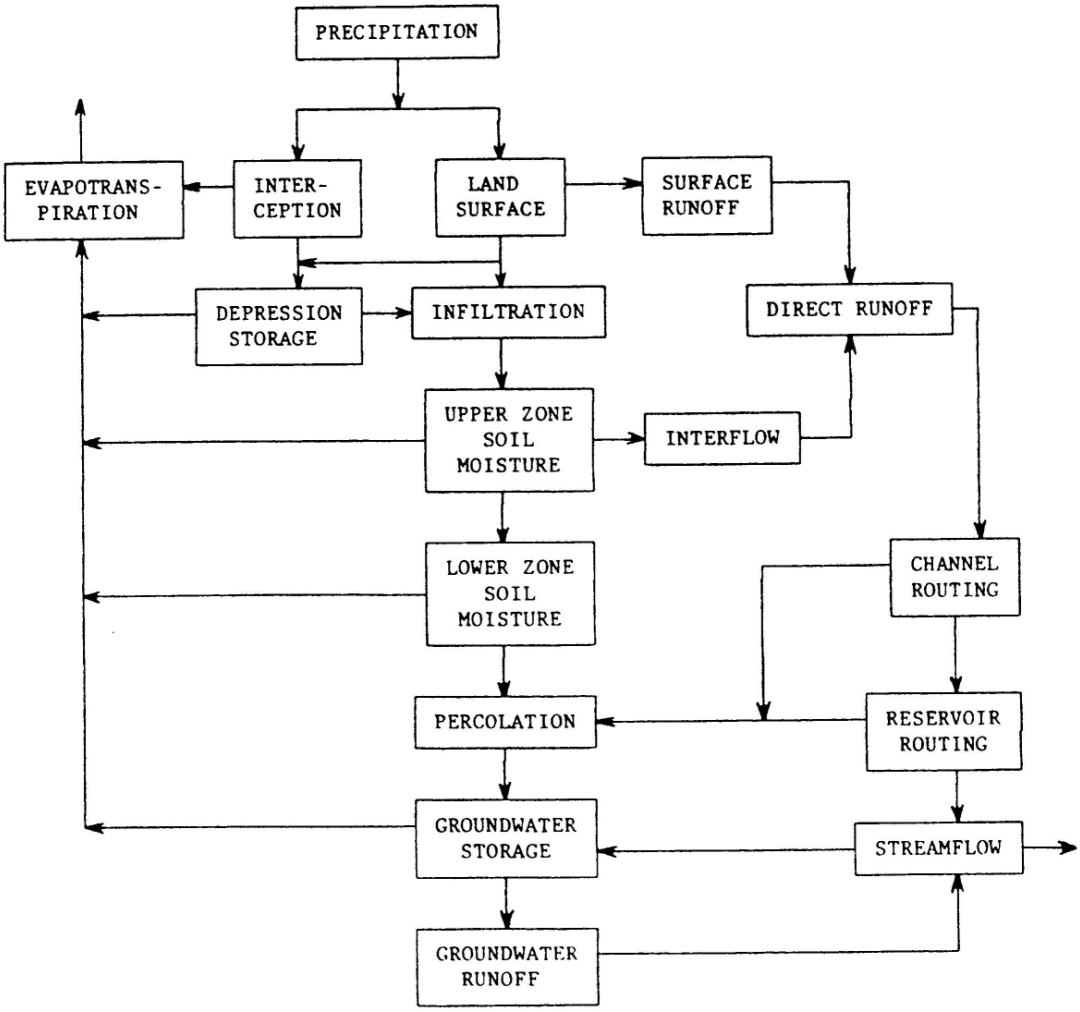


Fig.5.7 Components of a continuous catchment model (from Singh, 1988)

Table 5.2 Hydrologic components considered in some catchment models

Model		Model Components										
Name	*Author*	*Inter-ception*	*Infiltra-tion*	*Soil Moisture Storage*	*Evapotrans-piration*	*Surface Runoff*	*Snowmelt Runoff*	*Inter-flow*	*Groundwater Runoff*	*Channel Routing*	*Reservoir Routing*	*Parameter Optimiza-tion*
SWM IV	Crawford & Linsley (1966)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
KWM	Liou (1970)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
OPSET	James (1970, 1972)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
OSUM	Ricca (1972)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No
NWSRFS	Hydrology Research Laboratory (1972)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
SSARR	U.S. Army Engineer Division, North Pacific (1975)	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
API	Sittner et al. (1969)	No	No	Yes	No	Yes	No	No	Yes	No	No	No
USDA	Holtan et al. (1975)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
TWM	Claborn and Moore (1970)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
TANK	Sugawara et al. (1984)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
HBV	Bergstrom (1976)	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
SHE	Abbott et al. (1986a, 1986b)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
CEQUEAU	Charbonneau et al. (1977)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
MC	Deschenes et al. (1985a, 1985b)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
SCM	Refsgard (1981)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
SRBM	Bultot and Dupriez (1986a, 1986b)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes
UBCWM	Quick and Pipes (1977)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
MRM	Porter and McMahon (1971, 1975)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes
HYSIM	Manly (1978)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
ARBM	Chapman (1968)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes
BM	Boughton (1966)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes
HHM	Ando et al. (1983)	No	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes
TVA	Tennessee Valley Authority (1972)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes
USUWSM	Andrews et al. (1978)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes

5.6 A REVIEW OF MODELS USED IN WATER RESOURCES ASSESSMENT

5.6.1 Long-term water balance models

A traditional way of water resources assessment was based on the long-term average water balance equation over a basin in a form of

$$P = AE + Q \tag{5.38}$$

where *P, AE* and *Q* are the long-term average annual precipitation, evapotranspiration and streamflow, respectively. To solve equation (5.38) and get available water resources, Q, two terms, P and AE, must be known. Areal precipitation P is usually computed from point measurement. The key element in the long-term water balance of a large catchment or a region is the value of the actual long-term evapotranspiration (AE).

The first attempts at linking actual evapotranspiration to precipitation and potential evapotranspiration were made in the early years of last century on the basis of available measurements of catchment rainfall and runoff. The fundamental assumption in the formulae commonly suggested for the long-term water balance of catchments is that the ratio of actual to potential evapotranspiration may be expressed as a function of the ratio of precipitation to potential evapotranspiration. Well-known formulae of this type include Schreiber (1904); see also Dooge (1992),

$$\frac{AE}{PE} = \frac{P}{PE}\left[1 - \exp\left(-\frac{PE}{P}\right)\right] \tag{5.39}$$

It was formulated on the basis of measured precipitation and runoff in a number of catchments in Europe. A few years later, Ol'dekop (1911) suggested

$$\frac{AE}{PE} = \tanh\left(\frac{P}{PE}\right) \tag{5.40}$$

based on measurements in Russia.

Budyko and Zubenok (1961) examined the long-term water balance data for 1200 regions throughout the U.S.S.R. and found that these data fell within the limits of the formulae proposed by Schreiber and Ol'dekop. Turc (1954) proposed a simpler formula based on measurements from African catchments, which later was somewhat modified by Pike (1964) on the basis of further measurements:

$$\frac{AE}{PE} = \frac{P/PE}{\sqrt{1 + \left(P/PE\right)^2}} \tag{5.41}$$

The long-term water-balance method for estimating renewable water resources from meteorological data, though being very simple, has a number of essential disadvantages. First, in arid and semiarid regions, river runoff is very small by the absolute value and close to the error of determination of evaporation and precipitation. Second, it is impossible to estimate water resources for seasons and months. These data are crucial

for modern planning of water management. Third, this technique is inapplicable to estimate water resources of the countries and regions located in the basins of international rivers. In this case a larger volume of river runoff comes from outside rather than is formed on the territory at issue.

5.6.2 Monthly water balance models

Simple water balance models that simulate hydrographs of streamflow on the basis of available meteorological data and few physically relevant parameters, have been used by hydrologists and agricultural engineers in the assessment of regional water resources. Such models were first developed in the 1940s by Thornthwaite (1948) and have since been adopted, modified, and applied to a wide spectrum of hydrological problems (e.g. Alley, 1984; Schaake and Liu, 1989; Xu, 1999). A general review on monthly water balance models being used all over the world is made by Xu and Singh (1998).

The general structure of all water balance models is similar and building such a model involves writing equations that relate the rates of change of water properties within the control volume to flow of those properties across the control surface. For example, a simple soil water balance model for a control volume drawn around a block of topsoil is:

$$S(t+1) = S(t) + P(t) - AE(t) - Q(t) \tag{5.42}$$

In which S(t) represents the amount of soil moisture stored at the time t, i.e. at the beginning of a month, S(t+1) the storage at the later time t+1, i.e. at the beginning of next month, and the flow across the control surface during the interval [t, t+1], i.e. during the month considered, consists of precipitation P(t), actual evapotranspiration, AE(t), and soil moisture surplus, Q(t), which supplies streamflow and groundwater recharge. Solving this equation requires dealing with time series of the four variables: S, P, AE, Q, and possibly of other variables related to them. The water balance models differ in how AE and Q are conceptually considered and mathematically represented. One of the limitations of monthly water balance models is its inability to adequately account for possible changes in individual storm runoff characteristics at the time steps they are applied.

5.6.3 Conceptual lumped-parameter models

Many conceptual lumped-parameter models have been developed since the 1960s with the primary objective of flood forecasting of river basins. They are since then also used for simulation purpose for hydrologic design and water resources assessment at different scales. A few representative models will be briefly mentioned herewith as examples and for details the cited references should be consulted.

The remarkable Stanford Watershed Model IV by Crawford and Linsley (1964) represents the first great success in combining all the main hydrological processes within a computer model. This model is widely known and has been applied to many catchments throughout the world. Several models have followed, developing the concept further. A frequently used model in this group is the Sacramento Soil Moisture Accounting Model (Burnash et al., 1973). This model has been used by many researchers as one of the standard tools in the United States in flood forecasting, water resources assessment and studies on the impact of climate change. The HBV model (Bergström, 1976) is widely used in the Nordic countries as a standard tool to forecast

stream floods, to assess surface water resources and to simulate climate change effects. Applications of the HBV models have been made in some 30 countries (Bergström, 1992). In China and other Asian countries, the Xinanjiang model is used as a standard tool for a number of hydrologic simulation purposes. The model was developed in 1973 and published in 1980 (Zhao et al., 1980; Zhao, 1992). It has also been tested in the United States, Germany, Belgium, France, and Sweden.

Many other models having a similar structure but with different process conceptualisations, have been used in many regions of the globe (See also Leaveley, 1994). Among others, the Institute of Royal Meteorology Belgium model (Bultot and Dupriez, 1976) has been applied to basins in Belgium (Bultot et al., 1988) and Switzerland (Bultot et al., 1992). The HYDROLOG model (Porter and McMahon, 1971) was applied to two basins in South Australia (Nathan et al., 1988). The Hydrologic Simulation Program - FORTRAN (HSPF) model (U.S.E.P.A., 1984) has been applied to a basin in Newfoundland, Canada (Ng and Marsalek, 1992).

Compared with monthly water balance models, conceptual lumped-parameter models enable a more detailed assessment of the magnitude and timing of process response to climate change. However, these capabilities are accompanied by an increase in the number of process parameters and in the amount and types of input data needed to run the simulations.

5.6.4 Macroscale hydrologic models – a GIS supported modelling system

According to Maidment (1996), a spatial hydrologic model is one which simulates the water flow and transport in a specified region of the earth using GIS data structures. There are at least four primary motivations for the development of such a system.

- First, for a variety of operational and planning purposes, water resource managers responsible for large regions need to estimate the spatial variability of resources over large areas, at a spatial resolution finer than can be provided by observed data alone.
- Second, hydrologists and water managers are interested in the effects of land-use and climate variability and change over a large geographic domain.
- Third, there is an increasing need for using hydrologic models as a base to estimate point and non-point sources of pollution loading to streams.
- Fourth, hydrologists and atmospheric modellers have perceived weaknesses in the representation of hydrological processes in regional and global atmospheric models.

Leading models in this category include the one developed by Vörösmarty et al. (1989), the VIC model (Wood et al., 1992) and the Macro-PDM (Arnell, 1999). These spatial hydrologic models, in the literature named macro-scale hydrologic models (MHM), are conceptual water balance accounting models, which can be applied repeatedly over a large geographic domain on a regular grid without the need for calibration at the catchment scale. This is because what is feasible on the catchment scale, where parameters may be derived from careful observations or be calibrated using observed data is not feasible over a large area. Compared with the conceptual lumped-parameter models the macro-scale models have considerably fewer parameters. Their parameters can furthermore be estimated from spatial data sets, covering attributes as diverse as land cover, soil type and climate. Macro-scale hydrologic models are state-of-the-art tools in assessing regional and continental scale water resources.

5.7 WHEN CAN LUMPED MODELS BE USED, AND WHEN MUST DISTRIBUTED MODELS BE USED?

The distinction between lumped and distributed models is not only one of lesser or greater sophistication, but also intimately bound up with the purposes for which such models are to be used; answers to certain questions can only be attempted by using models with spatially distributed parameters. Blackie and Eeles (1985) give a good account of the uses of lumped models, under five headings:

(i) quality control and infilling of missing data;
(ii) extensions of historic flow records;
(iii) generation of synthetic data runs for civil engineering design work and other applications;
(iv) water resources assessment;
(v) water resources management including real-time forecasting.

In a paper which appeared in the same collection as Blackie and Eeles (1985), Beven (1985) quotes Beven and O'Connel (1982) as defining the role of distributed models in hydrology. The four areas offering the greatest potential are given as:

(i) forecasting the effects of land-use changes;
(ii) forecasting the effects of spatially variable inputs and outputs;
(iii) forecasting the movements of pollutants and sediments;
(iv) forecasting the hydrological response of ungauged catchments where no data are available for calibration of a lumped model.

Distributed-parameter models remain the most objective approach to answering questions such as: how the pattern of runoff will be changed in the area above 500 m in a catchment is planted to forest; or, if an accidental spillage of toxic chemical occurs at the bridge crossing this stream, how far the river channel network will be affected, how long its effects will last, and to what extent groundwater will be affected. Lumped-parameter models have little to offer for such questions.

In recent years, however, there has been a stocktaking of what is being, and can be, achieved by the use of distributed models, even with the most powerful computer support. Grayson et al. (1992) expressed the doubts as follows, in a paper discussing future directions for physically-based, distributed-parameter models (that is, those in which parameters describing processes are allowed to vary spatially over the river basin):

> The attraction of these models is their potential to provide information about the flow characteristics at points within the catchments, but current representations in process-based models are often too crude to enable accurate, a priori application to predictive problems. The difficulties relate to both the perception of model capabilities and the fundamental assumptions and algorithms used in the models. In addition, the scale of measurement for many parameters is often not compatible with their use in hydrologic models. The most appropriate uses of process-based, distributed-parameter model are to assist in the analysis of data, to test hypotheses in conjunction with field studies, to improve our understanding of processes and their interactions and to identify areas of poor understanding in our process descriptions. The misperception that model complexity is positively correlated

with confidence in the results is exacerbated by the lack of full and frank discussion of a model's capability/limitations and reticence to publish poor results... Model development is often not carried out in conjunction with field programs designed to test complex models, so the link with reality is lost.

5.8 A CALL FOR NEW GENERATION DISTRIBUTED MODELS

5.8.1 The present situation of hydrological modelling

As is shown in figure 5.8, successful analyses have been performed using physically-based distributed models with fine resolution data and using conceptual hydrologic models with coarse-scale data. One of the challenging fields in the hydrological modelling exercises is the application of physically-based distributed models to the meso or regional scale. There are at least four motivations for developing new generation distributed models that can be used in meso or regional scale (in the literature they are also referring as macro-scale hydrological modeling and spatial hydrology modelling):

- Water balance assessment of ungauged sites and large geographical regions,
- Assessment of the effects of land-use and climate variability,
- Estimating point and non-point sources of pollution loading to streams, and
- Improve the representation of hydrological processes in regional and global atmospheric models.

The traditional physically-based distributed models cannot be directly applied to large scales with coarse resolutions because: (1) The physically-based partial differential equations used in most distributed models are defined based on the fine resolution data and may not valid in large scale with coarse resolution due to the spatial heterogeneity and nonlinear nature of soil-vegetation-atmosphere transfer processes. (2) Large amount of data used to estimate the parameters of those models may not be available in large geographical regions. (3) Calibration of the large number of parameters of such models might not be feasible in large geographical regions.

The new generation distributed models should, therefore, have the following key characteristics:

- The model should be transferable from one geographical location to another. Model parameters should therefore be physically relevant.
- The model should be applied either to every sub-basin in the spatial domain or on a regular grid.
- Runoff must be routed from the point of generation (the fundamental unit) through the spatial domain along the river network.

As to the first characteristics, it requires, on the one hand the equations and parameters should be physically relevant, and on the other hands, the models should not be too specific with respect to local conditions. It requires some kind of generality and averaging. The number of model parameters should be fewer than the traditional distributed models. Refer to the second characteristics; the question arises as to how these area elements can be defined. One option is to subdivide the catchment into so-called "hydrological response units (HRUs)" which are similar with regard to selected characteristics and which are modelled separately. Another option is to subdivide the catchment into equally-spaced square grid elements. The problems related to the first subdivision method include which characteristics should be considered relevant to the

hydrological processes. If too many, the partitioning will be very detailed resulting in too much data need to be handled. If too few, we neglect the heterogeneity of the others. The problems related to the second subdivision method include that the physical characteristics within each grid cell that considered being homogeneous may be highly heterogeneous, reducing the size of grid cells reduces the heterogeneity, but increases computation time. Considering the coupling with atmospheric models (GCMs), the second sub-division method, i.e. regular grid cells are used. Finally, considering the large scale that the new generation distribution models are applied, flow routing both within the grid cells and between the cells constitute the important part of the models.
The general structure of a new generation distribution model includes three parts:

- Runoff generation at each grid cell,
- Routing within the cells using Time-area method,
- Routing between cells using river flow routing methods.

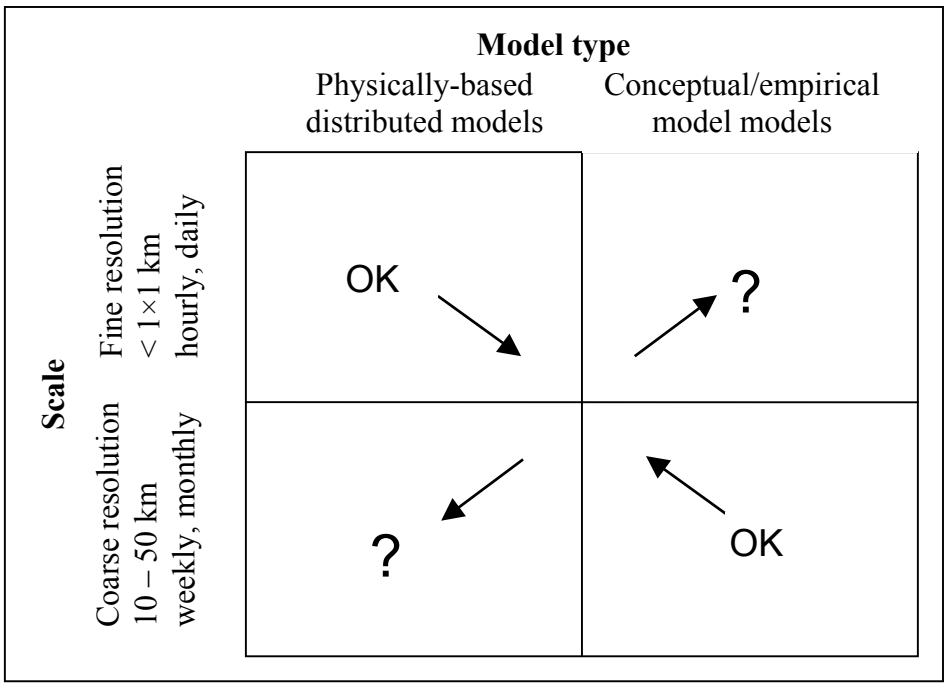


Fig.5.8 Conceptual matrix for hydrologic modelling and scale. Successful analyses have been performed using physically-based distributed models with fine resolution data and using conceptual hydrologic models with coarse-scale data. (Modified from Vörösmarty et al., 1993).

5.8.2 State-of-the-art of the new generation models

State-of-the-art new distributed model is an intergraded modeling system that combines SVAT model, groundwater model, snow model and hydrodynamic routing model and couples with GIS, DEM/DTM and GCMs.

What is a GCM?
GCM is a General Circulation Model of global atmospheric process. The outputs of such models include wind direction/speed, temperature, humidity, air pressure, precipitation, evaporation, streamflow, etc. The main inputs include CO2 and other gases. Typical spatial resolution of a GCM is 300×300 km.

With the original purpose of weather forecasting, the GCMs nowadays are used in predicting future climate scenarios (Loaiciga et al., 1996; Xu, 1999c).

What is a RCM?

RCM is a Regional scale atmospheric model, being similar to GCM in principle but with finer spatial resolution of 30×30 km. In the literature it is also referred as limited-area atmospheric model and regional climate model. It is the result of dynamic downscaling of a GCM (e.g. Giorgi et al., 1990).

What is a SVAT?

SVAT is a Soil-Vegetation-Atmosphere-Transfer scheme/model with the aim of simulating at higher performance in terms of hydrological, biogeochemical and vegetation dynamic processes. The general structure of a SVAT is shown in figure 5.9.

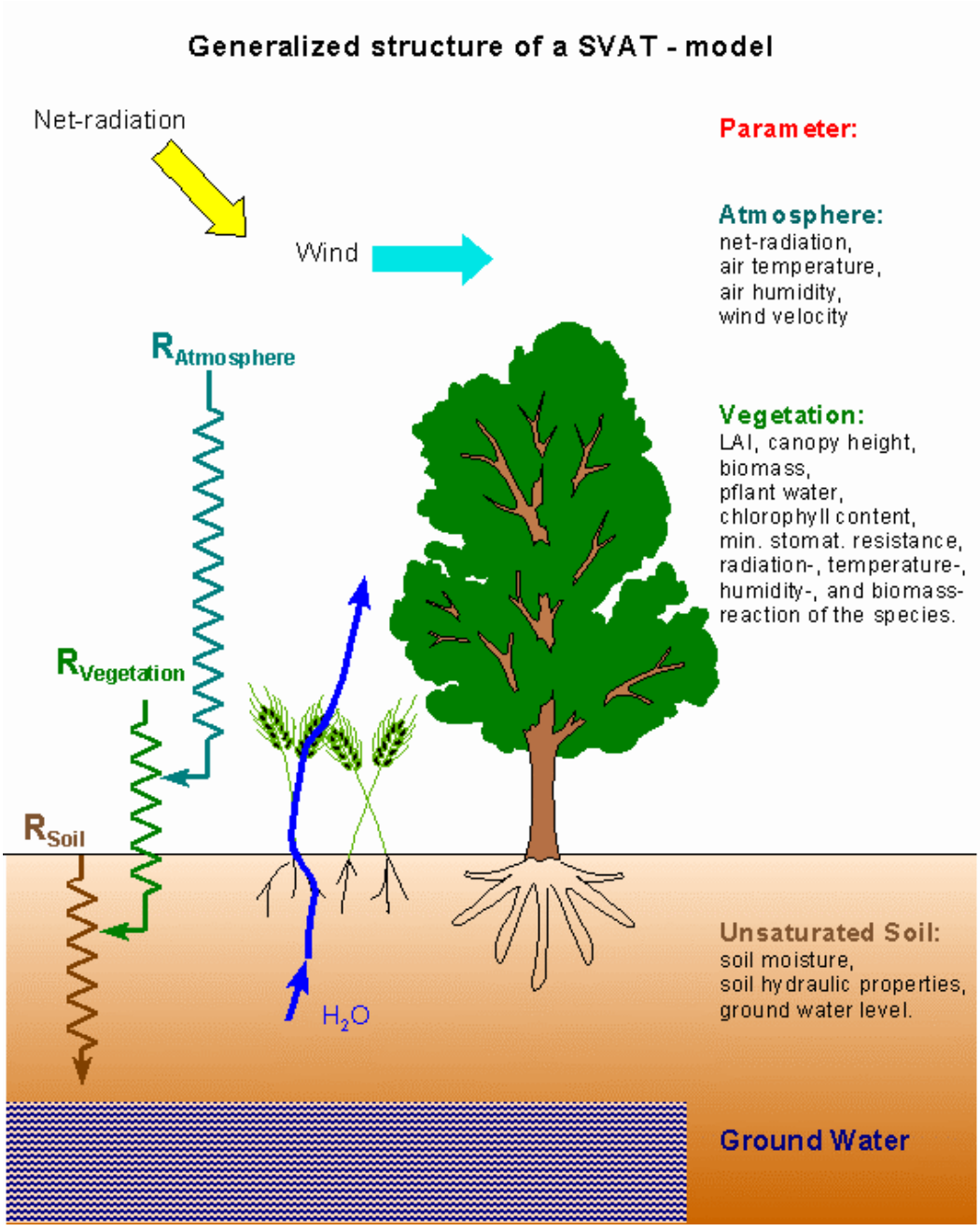


Figure 5.9 Generalized structure of a SVAT model.

There are two directions for further developments of SVATs:

- To design more comprehensive ecohydrological SVAT models capable of describing the various complex interrelationships and interdependencies between the different process variables, parameters and influencing characteristics.
- To derive simplified SVAT models, especially for use in larger scales, which are capable of simulating main processes based on a small set of key parameters, to be linked up to quantities used in the description of land-surface processes at the larger scales.

5.8.3 Approaches in developing the new generation distributed models

In principle, there are two approaches:

- "Bottom-up" approach: identifies representative hydrological areas and applies highly-detailed physically-based hydrological models, then aggregates upwards to all catchments or fundamental units in a large area (e.g. the Institute of Hydrology macromodel, Arnell, 1993; Kite et al., 1994).
- "Top-down" approach: treats each of the fundamental units as a single lumped catchment, and applies to each of them a simple conceptual hydrological model (Vörösmarty et al., 1989; Liston et al., 1994; Wood et al., 1992).

In practice, there are also two approaches:

- One is improving the energy balance processes within an existing hydrological model and enabling it to couple with an atmospheric model (e.g. Xu et al., 1994).
- Another approach is the improvement of the hydrological processes in land surface models developed for atmospheric models (Kim et al, 2001).

5.8.4 Approaches of coupling hydrologic models with atmospheric models (GCMs)

As mentioned before, one of the important motivations of the development of the new generation distributed models is to improve the representation of hydrological processes in regional and global atmospheric models. Coupling the hydrological models with the GCMs is perhaps the best way of doing so.

The traditional hydrological models cannot be coupled directly with the GCMs because of a number of gaps between them (see Table 5.3). The approaches that have been used to link GCM and hydrological models are presented in figure 5.10.

Global atmospheric GCMs have been used directly to simulate streamflow under present climate and to predict the impact of future climatic change in macroscale catchments. The analysis of GCM-predicted runoff showed that a simplistic representation of the hydrologic cycle within a global model of general atmospheric circulation leads to poor hydrologic predictive skill (e.g. Kuhl and Miller, 1992). The problem from a hydrologic point of view is that the most GCMs contain no lateral transfer of water within the land phase.

The results of literature survey showed that coupling the macroscale hydrological model with the GCM produces a better representation of the recorded flow regime than GCMs predictions of runoff for very large river basins. However, GCMs cannot 'see' smaller scale river basins because of their coarse grid resolution. Subgrid-scale hydrologic models and nesting schemes are needed to resolve the large-scale GCM predictions and predict smaller scale hydrologic phenomena (Hostetler and Giorgi, 1993).

Table 5.3: Some existing gaps between GCMs' ability and hydrology need

	Better simulated	**Less-well simulated**	**Not well simulated**
Spatial scales Mismatch	Global 300×300 km	Regional 30×30 km	Local 0 – 50 km
Temporal scales Mismatch	Mean annual & seasonal	Mean monthly	Mean daily
Vertical scale Mismatch	500 hPa	800 hPa	Earth surface
Working variables Mismatch	Wind Temperature Air pressure	Cloudiness Precipitation Humidity	Evapotranspiration Runoff Soil moisture
GCMs' ability declines →			
Hydrological importance increases →			

To circumvent these problems, 'downscaling' techniques have subsequently emerged as a mean of relating large-scale atmospheric models to regional or even basin scale hydrological models (Xu, 1999b). Two broad classes of downscaling approaches exist: dynamic methods, involving the explicit solving of the process-based physical dynamics of the system (e.g. Giorgi and Mearns, 1991); and statistical methods that use identified system relationships derived from observed data (e.g. Wigley et al., 1990; Xu, 1999c).

Due to the reason that the present GCMs and/or RCMs give different values of some climate variables and often do not provide a single reliable climate that could be advanced as a deterministic forecast for hydrological planning. Accordingly, methods of simple alteration of the present conditions are widely used by hydrologists (approach 5 in figure 5.10). Various hypothetical climate change scenarios have been adopted and climate predictions for 'double CO2' conditions have become a standard in such sensitivity studies (e.g. Loaiciga et al., 1996). The general procedure for estimating the impacts of hypothetical climate change on hydrological behavior has the following stages:

- Determine the parameters of a CHM model by calibration,
- Perturb the historical time series of observed climatic data according to some climate change scenarios by:
    - estimating average annual changes (%) in precipitation and temperature using
        - GCM result or
        - historical measurements of change or
        - personal estimates

- adjusting historic time series by multiplying the historic precipitation by a percentage change and adding an absolute change to the historic temperature.

- Simulate the hydrological characteristics of the catchment under the perturbed climate using the calibrated hydrological model.
- Compare the model simulations of the current and possible future hydrological characteristics.

In the recent studies, approaches 3 and 4 have received more efforts than others. This is because, statistical downscaling approaches linking GCMs to meteorologic and hydrologic models resolved at finer scales provide the possibility of bridging the gaps between the coarse-resolution GCMs and hydroclimatic modelling at the river-basin scale, while dynamic downscaling approach and nesting schemes provide the possibility of double-way coupling between atmospheric and hydrologic modeling. This double-way coupling is important not only because it provides better simulations of regional hydrologic scenarios, but also because it provides a feed back to meteorological modelers which in its turn is important for improving the skills of GCMs and RCMs.

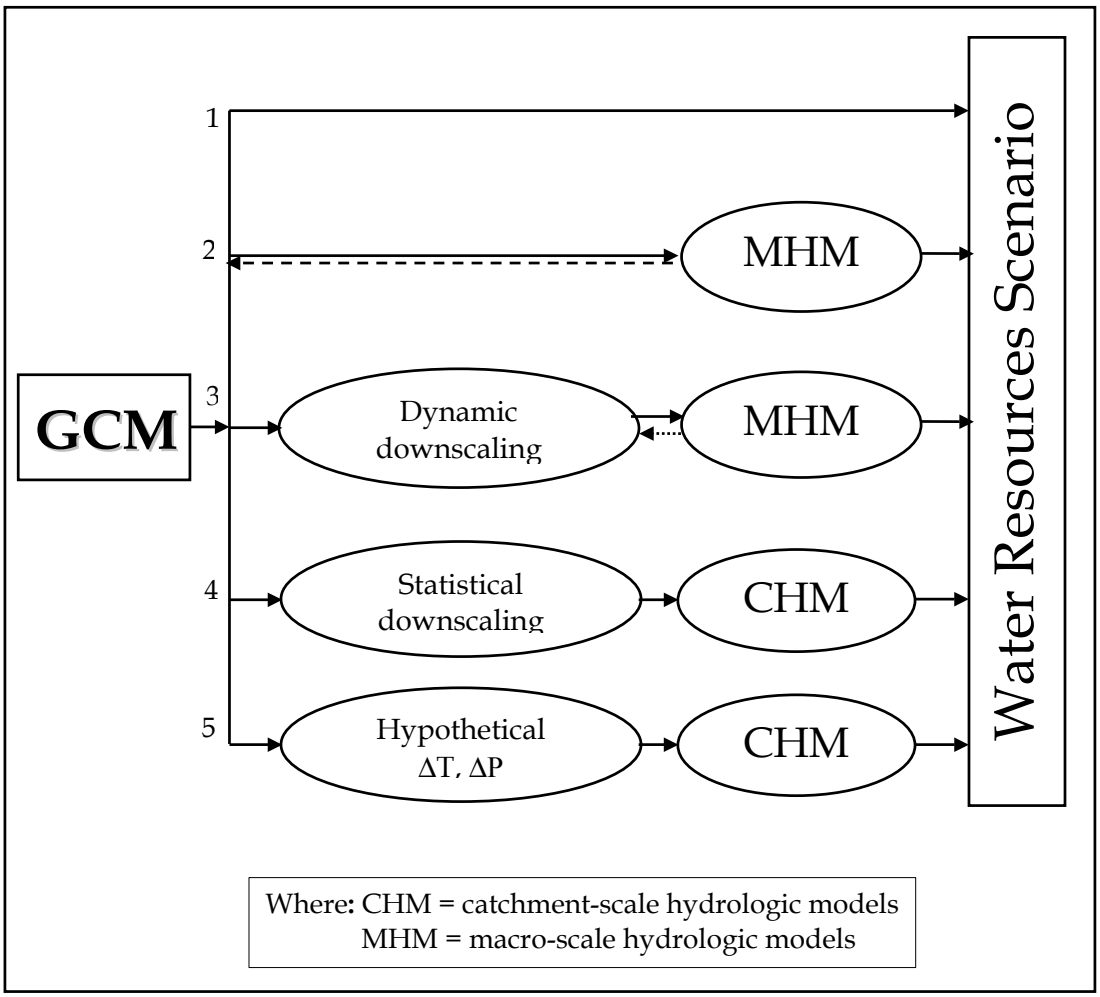


Figure 5.10 Schematic representation of the methods for assessing water resources under changing climate

REFERENCES OF CHAPTER 5

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J., 1986. An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 1: History and philosophy of physically-based, distributed modeling system. *J. Hydrol.*, 87: 45-59.

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmossen, J., 1986. An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modeling system. *J. Hydrol.*, 87: 61-77.

Alley, W.M., 1984: On the treatment of evapotranspiration, soil moisture accounting and aquifer recharge in monthly water balance models. *Water Resources Research* 20(8), 1137-1149.

Arnell, N.W., 1993. Data requirements for macroscale modeling of the hydrosphere, In: *Macroscale Modeling of the Hydrosphere*, IAHS publ. No. 214, pp. 139-149.

Arnell, N.W., 1999: A simple water balance model for the simulation of streamflow over a large geographic domain, *Journal of Hydrology* 217 (3-4), 314-335.

Bergström, S., 1976: Development and application of a conceptual runoff model for Scandinavian catchments. Department of Water Resources Engineering, Lund Institute of Technology, Bulletin Series A-52, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden.

Bergström, S., 1992: The HBV model – its structure and applications. SMHI RH report No 4. Meteorological and Hydrological Institute, Norrköping, Sweden.

Blake, G.J., et al., 1968. Infiltration in the Puketurua experimental basin. J. Hydrol. 7(1), 38-46.

Blank, D., J.W. Delleur, and A. Giorgini, 1971. Oscillatory kernel functions in linear hydrologic models. *Water Resources Research* 7(5): 1102-1117.

Bree, T., 1978. The stability of parameter estimation in the general linear model, Journal of Hydrology 37(1/2): 47-66.

Budyko, M-I. and Zubenok, L.I., 1961: On the determination of evaporation from the land surface. Izv. Akad. Nauk SSSR Ser. Geogr. 6, 3-17.

Bultot, F. and Dupriez, G.L., 1976: Conceptual hydrologic model for an average-sized catchment area: Concepts and relationships. *Journal of Hydrology* 29, 251-272.

Bultot, F., Coppens, A., Dupriez, G.L., Gellens, D. and Meulenberghs, F., 1988: Repercussions of a CO2-doubling on the water cycle and on the water balance: a case study for Belgium. *Journal of Hydrology* 99, 319-347.

Bultot, F., Gellens, D., Spreafico, M., and Schädler, B., 1992: Repercussions of CO2 doubling on the water balance - a case study in Switzerland. *Journal of Hydrology* 137, 199-208.

Burnash, R.J.C., Ferral, R.L. and McGuire, R.A., 1973: A generalized streamflow simulation system, conceptual modeling for digital computer. U.S. Department of Commerce, National Weather Service and State of California, Department of Water Resources, Sacramento, CA.

Chow, V.T., 1964. Runoff, in Handbook of Applied Hydrology, sec 14. McGraw-Hill, New York.

Chow, V.T., D.R. Maidment and L.W. Mays, 1988. Applied Hydrology, McGraw-Hill Book Company.

Crawford, N.H. and Linsley, R.K., 1964: A conceptual model of hydrologic cycle. IAHS Publ. 63, 573-587.

Dooge, J.C.I., 1973. Linear theory of hydrologic system, Tech. Bull. No. 1468, Agric. Res. Serv., pp117-124. October, U.S. Dept of Agriculture, Washington, D.C.

Dooge, J.C.I., 1992: Hydrologic models and climate change, *Journal of Geophysical Research* D3, 97, 2677-2686.

Fleming, G., 1975. Computer simulation techniques in hydrology. Elsevier Publishing Co., New York.

Giorgi, F., 1990. Simulation of regional climate using a limited area model nested in a general circulation model. *Journal of Climate* 3, 941-963.

Giorgi, F. and Mearns, L.O., 1991. Approaches to the simulation of regional climate change: a review. *Reviews of Geophysics* 29, 191-216.

Green, W.H. and Ampt, C.A., 1911. Studies on soil physics, I. Flow of air and water through soils. J. of Agricultural Sciences, 4: 1.24.

Hewlett, J.D., 1961. Watershed management, in *USFS Southeast, Forest Expt. Sta. Rept.* pp61-66.

Hewlett, J.D., and A.R. Hibbert, 1963. Moisture and energy conditions within a sloping soil mass during drainage, *JGR*, 68:1081-1087.

Hewlett, J.D., and W.L. Nutter, 1969. An outline of forest hydrology, Univ. of Georgia Press, Athens.

Horton, R.E., 1919. Rainfall Interception. Monthly Weather Review 47(9): 603-623.

Horton, R.E., 1933. The role of infiltration in the hydrologic cycle. Transactions of the American Geophysical Union 14: 446-460.

Horton, R.E., 1939. Analysis of runoff – plat experiment with varying infiltration capacity. Transactions of the American Geophysical Union 20:693-711.

Horton, R.E., 1940. An approach toward a physical interpretation of infiltration capacity. Soil Science Society of American Proceedings 5: 399-417.

Hostetler, S.W. and Giorgi, F., 1993. Use of output from high-resolution atmospheric models in landscape-scale hydrologic models: an assessment. *Water Resources Research* 29, 1685-1695.

James, L.D., et al., 1982. A taxonomy for evaluating surface water quantity model reliability. In Applied modelling in Catchment Hydrology, (ed.) V.P. Singh, 189-228. Littleton, Colo: Water Resources Publication.

Kite, G.W., Dalton, A. and Dion, K., 1994. Simulation of streamflow in a macroscale watershed using general circulation model data. *Water Resources Research* 30, 1547-1599.

Kuhl, S.C. and Miller, J.R., 1992. Seasonal river runoff calculated from a global atmospheric model. *Water Resources Research* 28, 2029-2039.

Larson, C.L. et al., 1982. Some particular Watershed Models. Chapter 10 in Hydrologic Modeling of Agricultural Watersheds, (ed.) H.T. Haan, 409-434. ASAE Monograph No.5 St. Joseph, Mich.: American Society of Agricultural Watersheds.

Lawler, 1964

Leavesley, G.H., 1994: Modelling the effects of climate change on water resources – A review. *Climatic Change* 28, 159-177.

Liston, G.E., Sud, Y.C. and Wood, E.F., 1994. Evaluating GCM land surface hydrology parameterisations by computing river discharges using a runoff routing model: application to the Mississippi Basin. *Journal of applied meteorology* 33, 394-404.

Loaiciga, H.A., Valdes, J.B., Vogel, R., Garvey, J., and Schwarz, H. (1996) Global warming and the hydrologic cycle. *Journal of Hydrology 174*, 83-127.

Maidment, D.R., 1996: GIS and hydrologic modelling – an assessment of progress. Presented at The Third International Conference on GIS and Environmental Modeling, January 22-26, Santa Fe, New Mexico.

Meriam, R.A., 1960. A note on the interception loss equation. J. of Geophysical Research 65(11): 3850-3851.

Meyer, A.F., 1942. Introduction to Runoff. Chapter 11A in Meinzer, O.E. (Ed.), Hydrology, McGraw-Hill, New York, 1942.

Morel-Seytoux, H.J., 1973. Two-Phase flows in porous media. In: Advances in Hydroscience, vol 9, edited by V.T. Chow, New York: Academic Press, 119-201.

Nash, J.E., 1957. The form of the instantaneous unit hydrograph. *International Association of Scientific Hydrology Publication* 45(3):114-121.

Nathan, R.J., McMahon, T.A. and Finlayson, B.L., 1988: The impact of the Greenhouse effect on catchment hydrology and storage-yield relationships in both winter and summer rainfall zones. In Pearman, G.I. (ed.), *Greenhouse, Planning for Climate Change,* Division of Atmospheric Research, CSIRO, East Melbourne, Australia.

Ng, H.Y.F., and Marsalek, J., 1992: Sensitivity of streamflow simulation to changes in climatic inputs. *Nordic Hydrology* 23, 257-272.

O'Donnell, T., 1960. Instantaneous unit hydrograph by harmonic analysis, Proc. Gen. Ass. Helsinki, IASH, publ.51, pp546-557.

Ol'dekop, E.M., 1911: On evaporation from the surface of river basins. Tr. Meteorol. Observ. Iur'evskogo Univ. Tartu., 4.

Philip, J.R., 1969. Theory of infiltration. In: Advances in Hydroscience, vol 5, edited by V.T. Chow, New York: Academic Press, 215-296.

Pike, J.G., 1964: The estimation of annual runoff from meteorological data in a tropical climate. *Journal of Hydrology* 2, 116-123.

Porter, J.W. and McMahon, T.A., 1971: A model for the simulation of streamflow data from climatic records. *Journal of Hydrology* 13, 297-324.

Renard, K.G., W.J. Rawls, and M.M. Fogel. 1982. "Currently available models." Chapter 13 in Hydrologic Modeling of Agricultural Watersheds, (ed.) H.T. Haan, 507-522. ASAE Monograph No.5 St. Joseph, Mich.: American Society of Agricultural Watersheds.

Rovey, C.E.K. and E.V. Richardson, 1975. Mathematical model of flow in a stream-aquifer system. *Proc. 2nd Annual Symp. on Model Technology, San Francisco, CA. (Published by ASCE, New York)*, pp439-457.

Schaake, J.C. and Liu, C., 1989: Development and application of simple water balance models to understand the relationship between climate and water resources. New Directions for Surface Water Modeling (Proceedings of the Baltimore Symposium, May 1989), IAHS Publ. no.181.

Schreiber, P., 1904: On the relationship between precipitation and the runoff of rivers in Central Europe, *Z. Meteorol*. 21, 441-452.

Singh, V.P., 1988. Hydrologic Systems. Volume 1: Rainfall-runoff modelling. Prentice Hall, New Jersey.

Singh, V.P., 1989. Hydrologic Systems, II, Watershed Modeling, Prentice Hall, New Jersey.

Thornthwaite, C.W., 1948: An approach toward a rational classification of climate. *Geogr. Rev.* 38(1), 55-94.

Toebes, C., 1963. Applied Hydrology 2: 37-59. Technical Correspondence School, New Zealand Dept of Education.

Turc, L., 1954: The water balance of soils. Relation between precipitation evaporation and flow. *Ann. Agron*. 5, 491-569.

U.S. Environmental Protection Agency (U.S.E.P.A.), 1984: Users manual for hydrological simulation program-FORTRAN (HSPF), EPA-600/3-84-066, Environmental Research Laboratory, Athens, GA.

Ullah, W. and W.T. Dickinson, 1979. Quantitative description of depression storage model using digital surface model: part I and part II, J. Hydrology 42: 63-99.

Vörösmarty, C.J., Moore, B., Gildea, M.P., Peterson, B., Melillo, J., Kicklighter, D., Raich, J., Rastetter, E. and Steudler P., 1989: A continental-scale model of water balance and fluvial transport: Application to South America. *Global Biogeochemical Cycles* 3, 241-65.

Vörösmarty, C.J., Gutowski, W.J., Person, M., Chen, T-C and Case, D. Linked atmosphere-hydrology models at the macroscale. In: Macroscale Modelling of the Hydrosphere, IAHS Publ No 214, 3-30.

Ward, R.C., 1972. Estimating streamflow using Thornthwaite's climatic water-balance, *Weather*, Feb. pp-73-84.

Wigley, T.M.L., Jones, P.D., Briffa, K.R. and Smith, G., 1990. Obtaining sub-gridscale information from coarse-resolution general circulation model output. *Journal of Geophysical Research* 95, 1943-53.

Wood, E.F., Lettenmaier, D.P. and Zartarian, V.G., 1992: A land-surface hydrology parameterisation with subgrid variability for general circulation models. *Journal of Geophysical Research* 97, D3, 2717-2728.

Xu, C-Y, 1999a: Estimation of parameters of a conceptual water balance model for ungauaged catchments. *Water Resources Management* 13(5), 353-368.

Xu, C-Y, 1999b. Climate change and hydrologic models: a review of existing gaps and recent research developments. *Water Resources Management* 13(5): 369-382.

Xu, C-Y, 1999c. From GCMs to river flow: a review of downscaling techniques and hydrologic modeling approaches. *Progress in Physical Geography* 23(2): 229-249.

Xu C-Y and Singh, V.P., 1998: A review on monthly water balance models for water resources investigations. *Water Resources Management* 12, 31-50.

Xu, L., Lettenmaier, D.P., Wood, E.F. and Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J. Geophys. Res. 99(D7), 14415-14428.

Zhao, R.J., 1992: The Xinanjiang model applied in China. *Journal of Hydrology* 135, 371-381.

Zhao, R.J., Zhang, Y.L., Fang, L.R., Liu, X.R. and Zhang, Q.S., 1980: The Xinanjiang model. Hydrological forecasting proceedings, Oxford Symposium, IAHS Publ. 129, 351-356.

CHAPTER 6
THE METHODOLOGY OF MODEL EVALUATION

6.1 PHASES OF MODEL EVALUATION

In general several levels of evaluation are necessary before the model should be applied to estimate the output from a catchment (See Fig.6.1). These are:

(i) model selection – choice of working hypotheses
(ii) model calibration - estimation of the parameter values
(iii) model validation - testing the fitted model to verify its accuracy; and
(iv) estimation of its range of applicability

Conceptually, these evaluations are distinct and follow in sequence. In practice, the boundaries for many types of models are often blurred. Of the four types of evaluation, estimation of the parameter values generally receives most attention. Nevertheless, it is important to recognise that all four evaluations are of equal fundamental importance, and neglect of any one can lead to serious error.

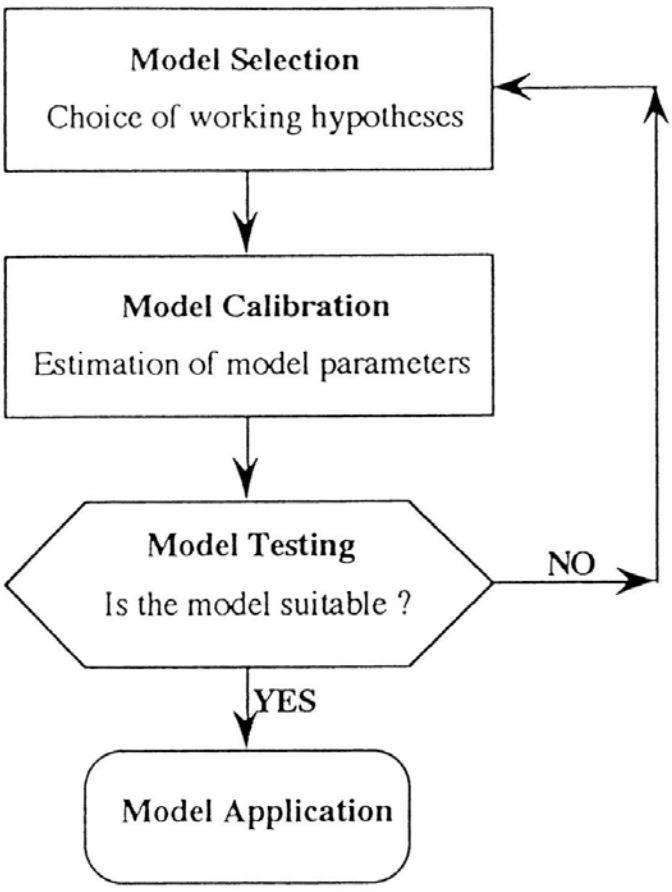


Fig. 6.1 Phases of model analysis

6.2 MODEL SELECTION

6.2.1 Problems to be considered

Hydrological practice would be improved if models were objectively chosen on the basis of making the best use of the information available and following some systematic procedure of selection and verification (Dooge, 1984). The choice of the best model depends to a large extent on the problem. Generally speaking, items that should be considered in the selection process include (Haan et al. 1982):

(a) The nature of the physical processes involved,
(b) The use to be made of the model,
(c) The quality of the data available and
(d) The decisions that rest on the outcome of the model's use.

In examining the nature of the physical processes involved, one should ask and attempt to answer such questions as: what are the processes that interact to produce the phenomenon under investigation? Are they amenable to solution by stochastic processes? Are they independent processes? Are they independent of time? Are values of the parameters likely to change with time, i.e., are they seasonal? Is the process stationary? Must future man-induced changes to be represented? If so, how?

In studying the use to be made of the model, one needs to answer: How much information is needed concerning the process being modelled? Do the data need to be presented in short time intervals or is monthly or annual data sufficient?

The quality of hydrological data describing a phenomenon affects the problem of fitting useful information from complex processes that produced the phenomenon. Several models may be capable of describing the same process, and, to a great extent, selection of the one to be used depends on a comparison of sampled data and model output.

Finally, in model selection, decisions that may rest upon the outcome of the model's use must be considered. To a great extent, these decisions will dictate the criteria that should be used to judge the quality of the model's performance. As an example, suppose that streamflow sequences will be used to determine the size of a dam to be used for water supply. In this case, the model is selected and its parameters estimated in such a way as to minimise the costs of uncertainty inherent in decisions regarding the size of the dam. Alternatively, suppose aerial rainfall data were used to study the spatial variability of soil moisture in assessing crop conditions. In this case, the model and its parameters must be selected to minimize the costs inherent in either overirrigation or losses in productivity brought on by drought induced growth stress. These are rather simplistic examples, but they serve to show the needs of the decision-maker, who may not know how to judge the quality of a model's response.

6.2.2 Criteria of selection

Thus far the problems to be considered in choosing a suitable model in general have been discussed. In most situations, however, absolute objective methods of choosing the best model for a particular problem have not yet been developed, so this choice remains a part of the art of hydrological modelling. Dawdy and Lichty (1968) suggested four criteria that can be used to choose between alternative models:

1). Accuracy of prediction
2). Simplicity of the model

3). Consistency of parameter estimates
4). Sensitivity of results to changes in parameter values

Accuracy of prediction of system output is obviously very important; it is desired when all other factors being equal, the model with minimum error variance would be superior. Simplicity refers to the number of parameters that must be estimated and the ease with which the model can be explained to clients or public bodies. When all other factors are being equal, one should choose the simplest model. Consistency of parameter estimation is an important consideration in developing hydrological models using parameters estimated by optimization techniques. If the optimum values of the parameters are very sensitive to the particular period of the record used, or if they vary widely between similar catchments, the model will probably be unreliable. Finally, models should not be extremely sensitive to input variables that are difficult to measure.

6.3 ISSUES IN MODEL CALIBRATION - PARAMETER ESTIMATION

6.3.1. Introduction to model calibration

Whatever the model form is chosen, there are some unknown constants used to represent the physical process. These so called parameters of the model must be assigned fixed numerical values before the model may be used to predict the runoff, in other words one needs to estimate these parameters such that the best agreement between modelled and observed runoff can be obtained. The process by which the parameters are selected is called model "calibration". The emphasis here is directed towards the calibration of "conceptual" hydrologic model of streamflow.

6.3.1.1 Model parameters

Many hydrologic models are based on conceptual representations of the physical processes that govern the flow of water through and over the soil. Such models usually have two types of parameters: "physical" parameters and "process" parameters (Sorooshian and Gupta, 1995).

(a) Physical parameters: physical parameters represent physically measurable properties of the watershed. Examples are: the area of the watershed, the fraction of the watershed area that is impervious, the surface area of the streams and open water bodies, surface slopes, and so on.
(b) Process parameters: process parameters represent watershed properties that are not directly measurable. Examples include: the average or "effective" depth of surface soil moisture storage, the effective lateral interflow rate, the coefficient of nonlinearity controlling rate of percolation to the groundwater storage, and so on.

6.3.1.2 Methods of parameter determination

There are two parts of parameter determination process: parameter specification and parameter estimation.

(a) **Parameter specification**: Here, we use prior knowledge about the watershed properties and behaviour to specify initial estimates for the parameters of the

model. For "physical" parameters, estimates are made using measurements obtained from maps in the field. The parameters are then typically fixed at these measured values and not adjusted further unless determined to be in error. For "process" parameters, estimates of the range (minimum and maximum values) of possible values for these parameters are determined based on judgement and understanding of the hydrology of the watershed. This uncertainty in the parameter estimates is then reduced by the process of parameter estimation described below.

(b) **Parameter estimation**: Here, we use various techniques designed to reduce the uncertainty in the estimates of the process parameters. A typical approach is to first select an initial estimate for the parameters, somewhere inside the ranges previously specified. The parameter values are then adjusted to more closely match the model behaviour to that of the watershed. The process of adjustment can be done "manually" or using computer-based "automatic" methods, as discussed below.

6.3.2 Manual calibration

To calibrate a model, we must select some aspect of watershed behaviour to which the model is to be matched: typically, we might select the streamflow hydrograph at one or more locations on the river. We then adjust the model parameters to get the simulated streamflow hydrograph to resemble the observed hydrograph for some historical data period. In manual calibration, we use a trial-and-error process of parameter adjustment; after each parameter adjustment is made, the simulated and observed hydrographs are visually compared to see if the match is improved.

With training and a good deal of experience, it is possible to obtain very good calibration using the manual approach. However, for the inexperienced and untrained person, manual calibration can be a rather frustrating and time-consuming exercise. This is mainly because the logic by which the parameters should be adjusted to improve the match is difficult to determine (due to the compensating effects which the model parameters usually have on the model output). Recent developments in computer graphics have made the process of manual calibration somewhat less tedious by enabling the effects of a parameter adjustment to be rapidly observed and compared to previous parameter trials (Brazil, 1988).

The main weakness of manual calibration is that the absence of generally accepted objective measures of comparison makes it difficult to know when the process should be terminated – i.e., whether the "best" possible fit has been obtained. Because manual calibration involves a great deal of subjective judgement, different persons may obtain very different parameter values for the same watershed. This makes it difficult to explicitly assign measures of confidence to the calibrated model and to its simulations and predictions.

6.3.3 Automatic calibration

6.3.3.1 Introduction

The development of computer-based methods for "automatic" calibration of watershed models has been motivated by:

(a) the need to speed up the process of calibration;

(b) the fact that there are few model calibration experts available for each watershed model; and
(c) the need to assign some measure of objectivity and confidence to model predictions.

Early attempts to develop automatic calibration methods were reported by Dawdy and O'Donnell (1965), Nash and Sutcliffe (1970), and Ibbitt (1970), among others. These researchers set the stage by bringing the vast body of research on statistical regression and model fitting techniques to bear on the calibration problem.

Since these beginnings, a great deal of progress has been made. However, it is important to clearly state that *automatic calibration methods have not yet matured to the point that they can entirely replace manual methods.* Although quick to provide "solutions", automatic methods still require user expertise and are typically used in conjunction with a manual procedure.

Automatic optimisation procedures are mathematical search algorithms that seek to minimize differences between selected features of modelled and observed streamflows by systematic trial alterations in the values of the model parameters. These trial alterations are called "iterations". The objective function, i.e., the quantitative measure of the fit of modelled runoff to the observed runoff, is calculated after each parameter iteration. Successful iterations are those which cause a reduction in the value of the objective function (for direct search method). During the search only the parameter set associated with the current least objective function value is retained, which, at the end of a search, is regarded as the optimal parameter set.

To illustrate the concept, a one-parameter model and a two-parameter model are used as examples. A *response surface* is formed when the objective function is plotted against the parameters. Typical *response surfaces* for one and two parameter models are shown in Fig6.2 for illustrative purpose. This concept can be extended for a model with n parameters to a *response surface* in (n+1) dimensional space and obviously it cannot be represented visually. The optimal parameter set is defined by the lowest point on the surface in the case of minimization of the objective function. This lowest point is known as the *global* optimum and discovery of the optimum is known as *convergence.* There are may be other points on the surface which are lower than all others in their immediate vicinity, but not lower than the global optimum. Such points are known as *local* optima, as shown in Fig.6.2.

A typical automatic parameter estimation procedure consists of four major elements:

(1) objective function,
(2) optimisation algorithm (to be discussed in Chapter 7),
(3) termination criteria, and
(4) calibration data, as discussed below.

In addition, processes of verification and sensitivity analysis (section 6.4) are necessary to establish confidence in the results.

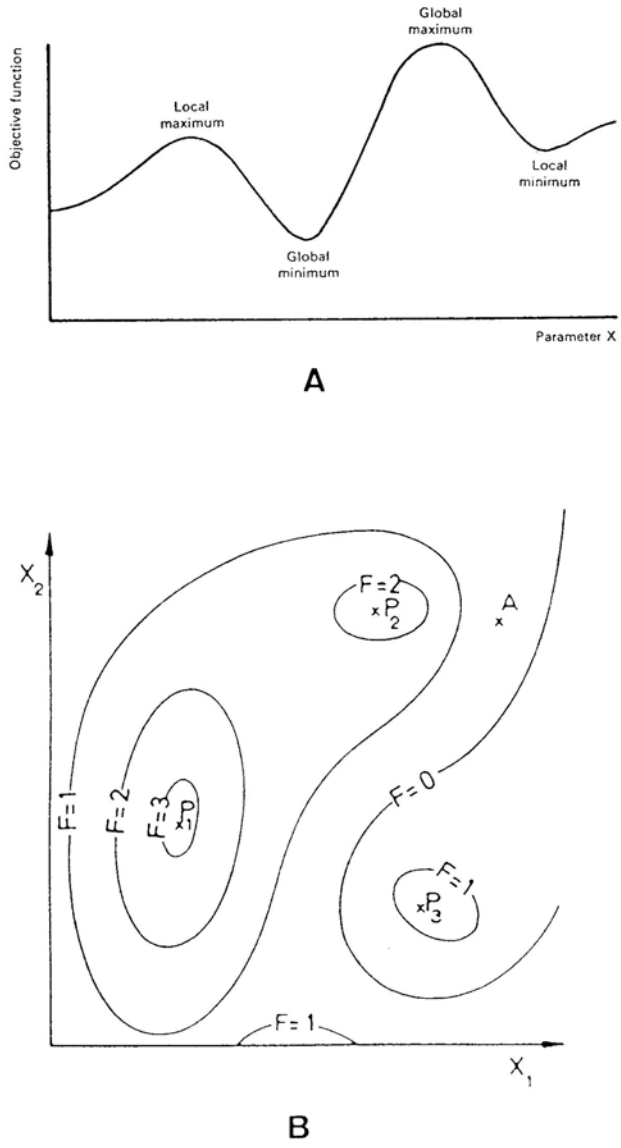


Fig.6.2 Response surface. A: one parameter model; B: two-parameter model

6.3.3.2a Objective functions

An objective function is an equation that is used to compute a numerical measure of the difference between the model-simulated output (typically the streamflow hydrograph) and the observed (measured) watershed output. The purpose of automatic model calibration is, therefore: "*to find those values of the model parameters that optimise (minimize or maximize, as appropriate) the numerical value of the objective function*".

(A) Least squares methods: Drawing from statistical regression and model-fitting theory, the most commonly used objective function has been some form of the Weighted Least Squares (WLS) function:

$$F(\theta) = \sum_{t=1}^{n} w_t \left[q_t^{obs} - q_t^{sim}(\theta) \right]^2 \tag{6.1}$$

where:

q_t^{obs} = observed (measured) streamflow value at time t;
$q_t^{sim}(\theta)$ = model simulated streamflow value at time t;
θ = vector of model parameters;
w_t = weight at time t; and
n = the number of data points to be matched.

The weights w_t indicate the importance to be given to fitting a particular hydrograph value. If the weights are all set equal to 1.0, the WLS function reduces to the familiar Simple Least Squares (SLS) function. Notice that the minimum value of the objective function F(θ) that can be attained is 0.0 (zero) if the model is able to perfectly reproduce the observed streamflow hydrograph. In general, however, a zero value is not attainable, and the purpose of automatic calibration is to find the value for θ which minimizes the value of the function.

For a proper evaluation of the model calibration, it is necessary to translate the overall calibration objective into more operational terms. The following objectives are usually considered:

1) A good agreement between the average of simulated and observed catchment runoff volume (i.e. a good water balance).
2) A good overall agreement of the shape of the hydrograph.
3) A good agreement of the peak flows with respect to timing, rate and volume.
4) A good agreement for low flows.

In this respect, it is important to note that, in general, trade-offs exist between the different objectives. For instance, one may find a set of parameters that provide a very good simulation of peak flows but a poor simulation of low flows, and vice versa. The following numerical performance statistics that are defined here measure the different calibration objectives stated above:

1) Overall volume error

$$F_1(\theta) = \left| \frac{\sum_{t=1}^{N} w_t \left[q_t^{obs} - q_t^{sim}(\theta) \right]}{\sum_{t=1}^{N} w_t} \right| \tag{6.2}$$

2) Overall root mean square error (RMSE)

$$F_2(\theta) = \left[\frac{\sum_{t=1}^{N} w_t^2 \left[q_t^{obs} - q_t^{sim}(\theta) \right]^2}{\sum_{t=1}^{N} w_t^2} \right]^{1/2} \tag{6.3}$$

3) Average RMSE of peak flow events

$$F_3(\theta) = \frac{1}{M_p} \sum_{j=1}^{M_p} \left[\frac{\sum_{t=1}^{n_j} w_t^2 \left[q_t^{obs} - q_t^{sim}(\theta) \right]^2}{\sum_{t=1}^{n_j} w_t^2} \right]^{1/2} \tag{6.4}$$

4) Average RMSE of low flow events

$$F_4(\theta) = \frac{1}{M_l} \sum_{j=1}^{M_l} \left[\frac{\sum_{t=1}^{n_j} w_t^2 \left[q_t^{obs} - q_t^{sim}(\theta) \right]^2}{\sum_{t=1}^{n_j} w_t^2} \right]^{1/2} \tag{6.5}$$

In equations (6.2) – (6.5), N is the total number of time steps in the calibration period, M_p the number of peak flow events, M_l the number of low flow events, n_j is the number of time steps in peak/low flow event no. j, θ is the set of model parameters to be calibrated, and w_t is a weighting function. Peak flow events are defined as periods where the observed discharge is above a given threshold level. Similarly, low flow events are defined as periods where the observed discharge is below a given threshold level.

Many other objective functions have been proposed or used in the literature; few of them are as follows:

$$\sum (q_{t,obs} - q_{t,sim})^r \qquad \text{r > 2} \tag{6.6}$$

$$\sum (1/q_{t,obs} - 1/q_{t,sim})^2 \tag{6.7}$$

(for use when emphasis must be placed upon low flows)

$$\sum (\log(q_{t,obs}) - \log(q_{t,sim}))^2 \tag{6.8}$$

$$\sum (\sqrt{q_{t,obs}} - \sqrt{q_{t,sim}})^2 \tag{6.9}$$

The coefficient of determination or the Nash-Sutcliffe coefficient (Nash and Sutcliffe, 1970) which is commonly adopted for evaluating the goodness-of-fit of the simulated hydrograph is a transformed and normalised measure of the overall RMSE (normalised with respect to the variance of the observed hydrograph)

$$R^2 = 1 - \frac{\sum_{t=1}^{N} w_t^2 \left[q_t^{obs} - q_t^{sim} \right]^2}{\sum_{t=1}^{N} w_t^2 \left[q_t^{obs} - \overline{q}_{obs} \right]^2} \tag{6.10}$$

where $\overline{q}_{obs}$ is the average observed discharge. In many applications, the weight, w_t is set to 1.

Use of the SLS function (6.1) and its modified forms is equivalent to making the following assumptions concerning the probability distribution of the residuals $\varepsilon_t = q_t^{obs} - q_t^{sim}$ (Clarke, 1973):

(a) that the ε_t have zero mean and constant variance σ_ε^2 (i.e., E(ε_t) = 0, E(ε_t^2) = σ_ε^2);

(b) that the ε_t are mutually uncorrelated ($E(\varepsilon_t\ \varepsilon_{t-k}) = 0$ for all k ≠ 0).

If it were known that either assumption (a) or (b), or both, were invalid, then eqn. (6.1) and it modified forms would not be the most sensible objective function; estimates of model parameters would, of course, still be obtained by minimising eqn.(6.1), but their interpretation would be fallacious.

Clarke (1973) also stated that if approximate confidence intervals are to be given for the estimated model parameters, a further assumption must be made about the probability distribution of the residuals, that is:

(c) that the ε_t are distributed normally.

The above assumptions need to be tested. The success or otherwise of the fitted model as a description of the relation between rainfall and streamflow from the catchment is illustrated by the model residuals, which also give evidence of the validity or invalidity of the assumptions (such as (a), (b) and (c) above) made in the model formulation. The procedure for testing the above assumptions is exemplified in a recently study of Xu (2001).

(B) Maximum likelihood methods: The method of maximum likelihood was developed by R.A. Fisher (1922). He reasoned that the best value of a parameter of probability distribution should be that value which maximizes the likelihood or joint probability of occurrence of the observed sample. Suppose that the sample space is divided into intervals of length *dx* and that a sample of independent and identically distributed observations $x_1, x_2, \ldots, x_n$ is taken. The value of the probability density for X = x_i is $f(x_i)$, and the probability that the random variable will occur in the interval including x_i is $f(x)dx$. Since the observations are independent, their joint probability of occurrence is given as the product

$$f(x_1)dxf(x_2)dx...f(x_n)dx = \left[\prod_{i=1}^{n} f(x_i)\right]dx^n$$

and since the interval size *dx* is fixed, maximizing the joint probability of the observed sample is equivalent to maximizing the likelihood function

$$L = \prod_{i=1}^{n} f(x_i) \qquad (6.11)$$

Because many probability density functions are exponential, it is sometimes more convenient to work with the log-likelihood function

$$\ln L = \sum_{i=1}^{n} \ln[f(x_i)] \tag{6.12}$$

Since the logarithmic function is nonotonic, the values of the θ's that maximize the logarithm of the likelihood function also maximize the likelihood function.

Example: Find the maximum likelihood estimator for the parameter λ of the distribution $f(x) = \lambda e^{-\lambda x}$ for X > 0.

Solution: For a given value x_i, the exponential probability density is

$$f(x_i) = \lambda e^{-\lambda x_i}$$

so, from Eq.(6.12), the log-likelihood function is

$$\ln L = \sum_{i=1}^{n} \ln[f(x_i)]$$

$$= \sum_{i=1}^{n} \ln(\lambda e^{-\lambda x_i})$$

$$= \sum_{i=1}^{n} (\ln \lambda - \lambda x_i)$$

$$= n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

The maximum value of $\ln L$ occurs when $\partial(\ln L)/\partial\lambda = 0;$ that is, when

$$\frac{\partial(\ln L)}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$$

so

$$\frac{1}{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$\lambda = \frac{1}{\bar{x}}$$

i.e., the parameter λ is equal to one over the sample average.

The method of maximum likelihood is the most theoretically correct method of fitting probability distributions to date in the sense that it produces the most efficient parameter estimates – those which estimate the population parameters with the least average error. But for many probability distributions, there is no analytical solution for all the parameters in terms of sample statistics; and the log-likelihood function must then be numerically maximized using an iterative procedure, which may be quite difficult.

Sorooshian and Dracup (1980) developed Maximum Likelihood based objective functions to properly account for the presence of either autocorrelation (nonindependence) or heteroscedasticity (changing variance) of the streamflow data errors. The most successful form of the Maximum Likelihood criteria has been one called HMLE (heteroscedastic Maximum Likelihood Estimator) that accounts for nonstationary variance in the streamflow measurement errors (Sorooshian, 1978, 1981; Sorooshian and Dracup, 1980; Sorooshian and Gupta, 1995). This estimator in simplified form is:

$$\min \text{HMLE}(\theta,\lambda) = \left\{\sum_{t=1}^{n} w_t \varepsilon_t^2\right\} \left\{ n \left[\prod_{t=1}^{n} w_t\right]^{1/n} \right\}^{-1} \tag{6.13}$$

where

$$w_t = f_t^{2(\lambda-1)} \tag{6.14}$$

and λ, the unknown transformation parameter which stabilizes the variance, is estimated by solving the implicit equation

$$\left[\sum_{t=1}^{n} \ln(f_t)\right] \bullet \left[\sum_{t=1}^{n} w_t \varepsilon_t^2\right] - n \bullet \left[\sum_{t=1}^{n} w_t \ln(f_t) \bullet \varepsilon_t^2\right] = 0 \tag{6.15}$$

where $f_t = q_{t,obs}$. Details of the derivation of the latter criterion and the two-stage optimisation procedure for its implementation are given by Sorooshian (1978, 1981).

6.3.3.2b Multiple objectives

Calibration based on a single performance measure is often inadequate to measure properly the simulation of all the important characteristics of the system that are reflected in the observations. This aspect is basically what causes certain scepticism in the hydrological profession for applying automatic calibration. Automatic routines that use a general multi-objective formulation of the calibration problem have been applied in rainfall-runoff modelling (Lindström, 1997; Liong et al., 1996, 1998; Gupta et al., 1998; Yapo et al., 1998; Madsen, 2000).

When using multiple objectives, the calibration problem can be stated as follows:

$$Min\{F_1(\theta), F_2(\theta), ..., F_p(\theta)\}, \qquad \theta \in \Theta \tag{6.16}$$

The optimisation problem is said to be constrained in the sense that θ is restricted to the feasible parameter space Θ. The parameter space is usually defined as a hypercube by specifying lower and upper limits on each parameter. These limits are chosen according to physical and mathematical constraints in the model and/or from modelling experiences (prior knowledge).

The solution of eq.(6.16) will not, in general, be a single unique set of parameters but will consist of the so-called Pareto set of solutions (non-dominated solutions),

according to various trade-offs between the different objectives. Formally, any member θ_j of the Pareto set has the properties (Gupta et al., 1998):

1) For all non-members θ_j there exists at least one member θ_i where $F_k(\theta_i) < F_k(\theta_j)$ for all k = 1,2,…,p.
2) It is not possible to find θ_j within the Pareto set such that $F_k(\theta_j) < F_k(\theta_i)$ for all k = 1,2,…p.

Concerning (1), the parameter space Θ can be divided into "good" (Pareto optimal) and "bad" solutions, and concerning (2) none of the "good" solutions can be said to be "better" than any of the other "good" solutions. A member of the Pareto set will be better than any other member with respect to some of the objectives, but because of the trade-off between the different objectives it will not be better with respect to other objectives.

When solving the multi-objective calibration problem, the problem is usually transformed into a single-objective optimisation problem by defining a scalar that aggregates the various objective functions. One such aggregate measure is the Euclidean distance

$$F_{agg}(\theta) = \left[(F_1(\theta) + A_1)^2 + (F_2(\theta) + A_2)^2 + \cdots + (F_p(\theta) + A_p)^2\right]^{1/2} \quad (6.17)$$

where A_i are transformation constants assigned to the different objectives, which allows the user to select relative priorities to certain objectives. The selection of transformation constants, however, is not straightforward, since the priority also depends on the value of F_i itself. For instance, if all A_i are set to zero, implicitly larger weights are given to objectives with larger F-values. For investigating the entire Pareto front, the aggregated distance measure can be adopted by performing several optimisation runs using different values of A_i.

In practical applications, the entire Pareto set may be too expensive to calculate, and one is only interested in part of the Pareto optimal solutions. In this case, it is proposed to use an aggregated objective function that puts equal weights on the different objectives. A balanced measure can be defined by assigning transformation constants in eq.(6.17) such that all $(F_i + A_i)$ have about the same distance to the origin. When using a population-based (global search method) optimisation algorithm, an initial population within the feasible region is evaluated. The minimum values of F_i $(F_{i,\min})$ are estimated from this initial population, and each of the objective functions is transformed to having the same distance to the origin as the objective function with the largest minimum value of F_i, i.e.

$$A_i = Max\{F_{j,\min}, j = 1,2,..., p\} - F_{i,\min,}\ j = 1,2,...p \quad (6.18)$$

6.3.3.3 Optimisation algorithms (to be discussed in Chapter 7)

6.3.3.4 Termination criteria

The optimisation strategies are all iterative procedures which search for the optimal parameter values by means of incremental improvement steps. Therefore, criteria are needed to determine when to stop the search. In principle, the solution exists at that point in the parameter space where the slope of the function response surface is zero and the function value is a minimum. In practice, it is virtually impossible to know when this point has been reached; hence, the criteria discussed below are more commonly used.

6.3.3.4a Function convergence

One simple way to terminate the search is to stop when the algorithm is unable to appreciably improve the value of the function over one or more iterations. While this can indicate arrival at the location of an optimum, it could also mean only that a very flat region of the response surface has been reached. If precise detection of an optimum is not considered important, then function convergence can be a very useful stopping criterion. One typical implementation of this criterion is to stop when:

$$(f_{i-1} - f_i) / f_i <= \varepsilon_f \tag{6.19}$$

where f_{i-1} and f_i are the function values at the (i-1)th and ith steps, respectively, and ε_f is the function convergence criterion (for example $\varepsilon_f = 10^{-3}$).

6.3.3.4b Parameter convergence

Another way to terminate the search is to stop when the algorithm is unable to appreciably change the parameter values and simultaneously improve the function value over one or more iterations. While this can indicate arrival at an optimum, it could also mean only that a region of high parameter interaction (long narrow valley) on the response surface has been reached. One typical implementation of this criterion is to stop when:

$$(\theta(j)_{i-1} - \theta(j)_i) / (\theta(j)_{\max} - \theta(j)_{\min}) <= \varepsilon_\theta \quad \text{for each } \theta(j) \tag{6.20}$$

where $\theta(j)_{i-1}$ and $\theta(j)_i$ are the values of the jth parameter at the (i-1)th and ith steps, respectively, and ε_θ is the parameter convergence criterion (for example $\varepsilon_\theta = 10^{-3}$).

6.3.3.4c Maximum iterations

If computer time is limited, and to ensure that the algorithm does not somehow enter an infinite loop, it is normal to terminate the search if a prespecified maximum number of iterations is exceeded, unless the parameter or function convergence criteria are met first. For random search methods, this is the normal way to terminate the search. It is not really possible to give guidelines on the value for this criterion, because it is both algorithm- and problem – dependent. The maximum iterations criterion is used as a

backup to prevent waste of computer time; if the algorithm does not terminate within a "reasonable" number of iterations, the computer code may need to be examined for "bugs".

6.3.3.4d Limitations

None of these termination criteria guaranty that the search arrival at the global optimum, except in the most trivial cases where the function is convex and well behaved. These criteria can be used in the same program, so the search will terminate when the first criterion is reached.

6.3.3.5 Calibration data

It is generally agreed that proper choice of the calibration data can do much to reduce the difficulties encountered during calibration of a hydrologic model. However, little is known objectively about what constitutes "good" calibration data. The criteria issues here are how much data are necessary and sufficient for calibration and what kind of data will give the best results (most precisely specified parameter estimates).

6.3.3.5a Quantity of data

It has been a common practice to use as much data as were available for the calibration, after setting aside part of the data set for verification (see section 6.4) of the results. However, studies by Sorooshian et al (1983) and Xu and Vandewiele (1994) indicated that the use of longer data sets than what is necessary served only to marginally improve the parameter estimates. In general, from a statistical point of view, the data set used should be at least of length 20 times the number of parameters to be estimated (for example, if there are 10 parameters, then at least 200 streamflow data points should be used for computing the function). This is of course an approximate rule of thumb. Gupta and Sorooshian (1985) showed that the standard error (j) of the estimate of parameter (j) decreases with sample size n approximately according to the formula:

$$\sigma(j) \propto \frac{1}{\sqrt{n}} \tag{6.21}$$

Because the marginal improvement in 1/n becomes small after 500 to 1000 data points, this suggests that two to three years of calibration data should be sufficient for a daily model with not more than 10 parameters, provided the data are of the *right kind*. This brings us to a discussion of data quality.

6.3.3.5b Quality of data

From the viewpoint of model calibration, the quality of the data is dependent on the information (about the parameters) contained in the data and the noise (errors) in the data. Clearly, we wish the information content to be as large as possible and the noise to be as small as possible.

Informativeness

Qualitatively, an informative data set is one which contains or represents enough variability in watershed behaviour that the different modes of operation of the hydrologic processes are properly represented. For example, if the data selected are from a relatively dry year, certain runoff processes may not be activated, therefore, the model response will be insensitive to some of the model parameters that determine the partitioning of moisture between the various subsurface and overland flow components. However, if the data selected are from a year that is so wet that the watershed remains saturated most of the time, the model response may be insensitive to other subsurface flow controlling parameters. The best choice seems to be a data set that contains a lot of "hydrologic variability". Thus, the more often the hydrologic regime switches from between dry, medium and wet modes, the more informative the data are likely to be.

Data errors

The presence of measurement and logging errors in the data causes the quality to deteriorate, thereby resulting in less confidence in the parameter estimates. In selection data for model calibration, it is desirable that the data be carefully examined for various errors.

6.4 MODEL VERIFICATION (TEST)

6.4.1. Introduction

Testing or verification or validation of a model after the parameter values are estimated is the third level of model analysis. As no model is perfect, verification requires both subjective and objective judgements on many aspects to determine whether the results provide adequate information for answering the question facing the decision-makers, and all models can be expected to fail at least on some occasions. Faulty results may stem from a variety of causes, common problems may be:

(1) Errors in the data used in calibration. Both the data used as input to the model and the data used to check model output should be checked very carefully. Data with large errors should not be used for calibration. In addition, the data used to calibrate or to test are always only a sample of the possible population of values. Relative to values predicted by a model, the actual values contain a probabilistic or stochastic component as a result of physical factors that are not modelled at all.

(2) Use of a period of record that does not contain enough events of the physical processes needed to calibrate key parameters.

(3) Inadequate or miss-representation by the model of hydrological processes found in the catchment. Model results should be compared visually with the recorded data series to look for consistent variations.

Therefore the model test in some cases is also called "diagnostic checking". If the model is good enough, then we can pass on to application. If not, one begins the whole process all over again by changing one or more working hypotheses and checking the data used. In which direction the model has to be changed mostly appears during model test.

6.4.2 Methods of model validation

Klemes (1986) proposed a model validation framework for testing the conceptual hydrologic models according the modelling tasks. Four major categories, corresponding modelling tasks and test methods are summarised in Table 6.1.

Simple split-sample testing involves dividing the available measured time-series data for the test catchment into two sets, each of them should be used in turn for calibration and validation, and results from both arrangements compared. For differential split-sample testing, the same approach is followed, but the data are divided according to rainfall rate or some other variable in an attempt to show that the model has general validity in that it can predict the values of the output variables for conditions different from those for which it was calibrated. For example, if the model is intended to simulate streamflow for a wet climate scenario then it should be calibrated on a dry set of the historic record and validated on a wet set. If it is intended to simulate flows for a dry climate scenario, the opposite should be done. In general, the model should demonstrate its ability to perform under the transition required: from drier to wetter conditions or the opposite. Proxy-catchment tests use data for two catchments. These tests can be used to show the model has even greater general validity as they involve calibrating the model against data for one catchment and then running a validation test using data for the other catchment. For differential proxy-catchment testing, the available measured time-series data for each catchment are divided into two sets according to rainfall rate or some other variable. The model is then calibrated against one of the sets (e.g. the dry period data for the first catchment) and a validation test run using a contrasting set (e.g. the wet period data for the second catchment). Calibration is required in all the four validation methods discussed above.

Beven et al. (1984) and Loague (1990) used another type of test in which the model is not calibrated, and predictions are simply compared against measurements. Recently, Ewen and Parkin (1996) proposed a method, namely a ‘blind’ approach. The central feature of this method is that it involves making predictions for a test catchment as if it were a hypothetical catchment. The modeller is, therefore, not allowed sight of the output data for the test catchment (i.e. the method involved ‘blind’ testing), and, as a result, cannot calibrate the model for the test catchment.

Table 6.1 Hierarchical scheme for operational testing of hydrologic simulation models

	Stationary conditions		Transient conditions	
	Basin A	Basin B	Basin A	Basin B
Basin A	Split-sample test	Proxy-basin test	Differential split-sample test	Proxy-basin differential split-sample test
Basin B	Proxy-basin test	Split-sample test	Proxy-basin differential split-sample test	Differential split-sample test

6.4.3. Items to be tested in model validation

The models may be tested on the following aspects.

Parameter analysis:

The evaluation of the parameter values during the optimization. The stabilization of the parameter values can be studied on the graphs of the parameter values versus number of iterations. An example of such graphs is given in Fig.6.3

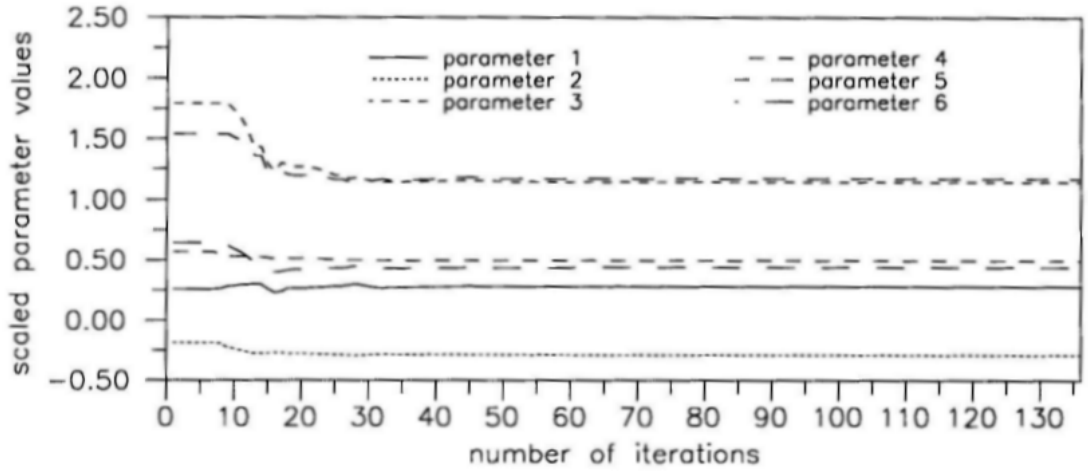


Fig.6.3 Stabilization of the parameter values during iteration process

Detailed analysis of the variance-covariance matrix. The correlation matrix of the parameters has to be checked. If the correlation coefficient between two parameters is very near to +1 or -1, this means that perhaps a model can be found with a smaller number of parameters and with the same explanatory power, or that perhaps the parameters have to be built into the model in a different way, so that their explanatory effects are more dissociated, and optimization is easier.

For answering the question whether all parameters are really necessary, one can test the hypotheses that parameters a_1, a_2, ... are significantly different from zero. This can be done by checking whether the zero value belongs to the 95% confidence interval (Xu, 2001).

Residuals analysis:

The basic issue in model testing is to determine if the hydrologic estimates (residual error) achieved by the calibration are acceptable. Residual analysis is checking whether the residuals u_t behave as is required by the model hypotheses, especially whether they are *independent, homoscedastic and normally distributed with zero expectation.*

Check on independence

The hypothesis that the residuals are mutually uncorrelated can be checked by computing the autocorrelations of the residuals, ρ_k, with time lag k and the corresponding confidence interval. In general, the autocorrelations ρ_k with time lag k is

$$\rho_k = E[(x_t - \mu)(x_{t+k} - \mu)] / \sigma^2 \tag{6.22}$$

where μ and σ^2 are the mean and variance of the residuals, respectively.
An estimate of ρ_k is

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} x_t x_{t+k} - \frac{1}{n-k}(\sum_{t=1}^{n-k} x_t)(\sum_{t=k+1}^{n} x_t)}{\left[\sum_{t=1}^{n-k} x_t^2 - \frac{1}{n-k}(\sum_{t=1}^{n-k} x_t)^2\right]^{1/2}\left[\sum_{t=k+1}^{n} x_t^2 - \frac{1}{n-k}(\sum_{t=k+1}^{n} x_t)^2\right]^{1/2}} \tag{6.23a}$$

when n is large and k is small, n/(n-k) → 1, a simpler estimator of autocorrelation coefficient is

$$r_k = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} \tag{6.23b}$$

The confidence interval for the autocorrelation coefficient of an independent series is given by the limits (Haan, 1977)

$$r_k(95\%) = \frac{1}{n-k}\left[-1 \pm 1.96\sqrt{(n-k-1)}\right] \tag{6.24}$$

If the calculated r_k falls outside these confidence limits, the hypothesis that ρ_k is zero (H_o: $\rho_k = 0$ versus H_a: $\rho_k \neq 0$) is rejected. Examples of the tests are shown in Figure 6.4.

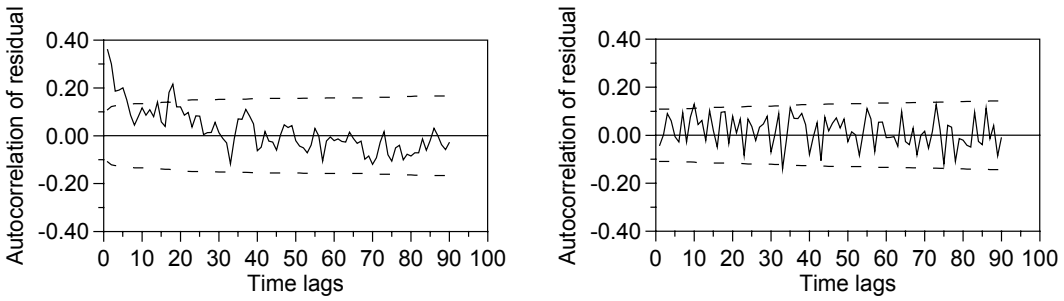


Figure 6.4. Autocorrelation of residuals. (Left) Correlated case. (Right) Uncorrelated case.

Check on trend and homoscedasticity

The residual's homoscedasticity can be checked using both the graphic method and the Kruskal-Wallis test method. In the graphic method, the general behaviour of the residuals is judged by graphs of the residuals versus important variables such as time itself, the input variables precipitation p_t and evaporation e_t, and computed runoff d_t.

The residuals versus time diagram is used for checking the absence of trend and also homoscedasticity. It could also reflect what kinds of violations are involved in the residual time series. The hypotheses of homoscedasticity and uncorrelatedness then result in a diagram such as in Fig.6.5(a). The other diagrams in Fig6.5 show violations of these hypotheses. Combinations of these violations are possible, and other possibilities exist.

The residual versus expected response diagram. Conditional expectations of the residuals have to be zero for all d_t, and conditional variances have to be equal. This appears not to be the case in Fig 6.6 where the expected residual appears to increase with d_t and the variance decreases. Both aspects lead to serious doubt on the validity of the model-hypotheses.

The residual versus observed factor diagrams. Considering all residuals at the same time leads to a marginal diagram. When only residuals are considered where all or part of the other observed factors are held constant approximately, a conditional diagram is obtained. Again violations are at the origin of reformulating the model. Fig6.7 and Fig6.8 are examples of marginal scatter-grams, which appear to satisfy model-hypotheses.

The Kruskal-Wallis test, or H test enables us to test the null hypothesis that k independent random samples come from identical populations. It is a nonparametric test. The method assumes that the variable has a continuous distribution, but nothing is said about the form of the population distribution or distributions from which the samples were drawn. The test is based on the statistic

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1) \qquad (6.25)$$

In the test, all observations are ranked jointly, and R_i is the sum of the ranks occupied by the n_i observations of the i^{th} sample, and $n_1 + n_2 + ... + n_k = n$. When $n_i > 5$ for all i and the null hypothesis is true, the sampling distribution of the H statistic is well approximated by the chi-square distribution with k-1 degrees of freedom. The null hypothesis of homoscedasticity will be rejected for a given significance level, α, if computed H is bigger than $\chi^2_{1-\alpha,k-1}$. Detailed examples of the test can be found in Xu (2001).

Check on the normality

The hypothesis that residuals are distributed normally is needed if the estimated confidence regions for the parameters are required. The normality can be tested using different methods. The Kolmogorov-Smirnov test method is discussed here. The test has several advantages. It is easy to use and the procedure is graphic; a large number of samples can be tested on the same plot; the test is nonparametric and is not subject to the very small sample limitation. The test is conducted as follows:

1) Let F(x) be the completely specified theoretical cumulative distribution function under the null hypothesis.
2) Let $F^e(x)$ be the sample cumulative density function based on n observations. For any observed x, $F^e(x)$ = k/n where k is the number of observations less than or equal to x.
3) Determine the maximum deviation, D, defined by

$$D = \max\left|F(x) - F^e(x)\right| \tag{6.26}$$

4) If, for the chosen significance level, the observed value of D is greater than or equal to the critical tabulated value of the Kolmogorov-Smirnov statistic, the hypothesis is rejected. The deviation between F(x) and $F^e(x)$ is graphically shown in Fig.6.9.

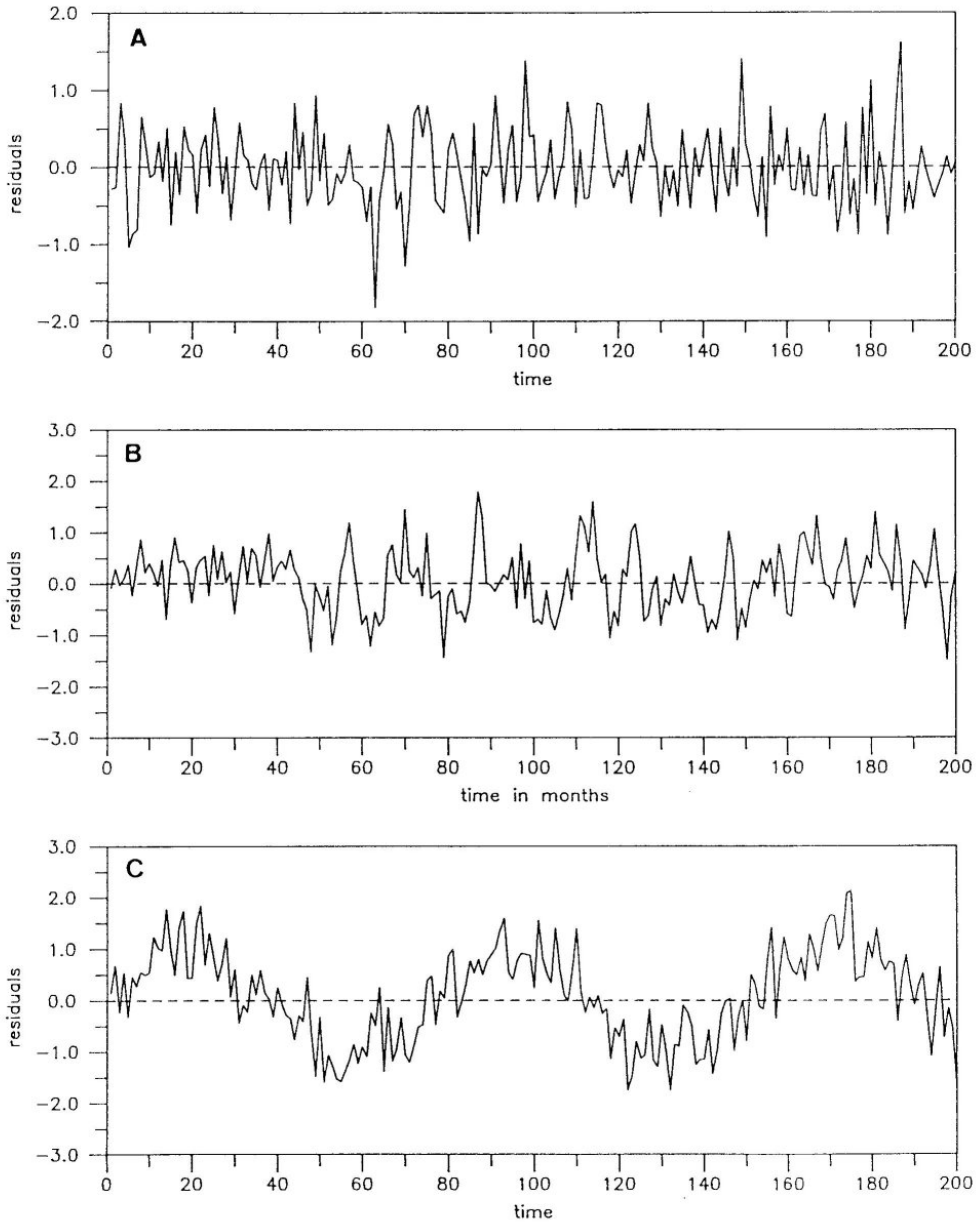


Fig.6.5 Residuals versus time diagrams. A: shows the residuals are homoscedastic and uncorrelated; B: shows the residuals are autocorrelated; C: shows a periodic component in the residuals.

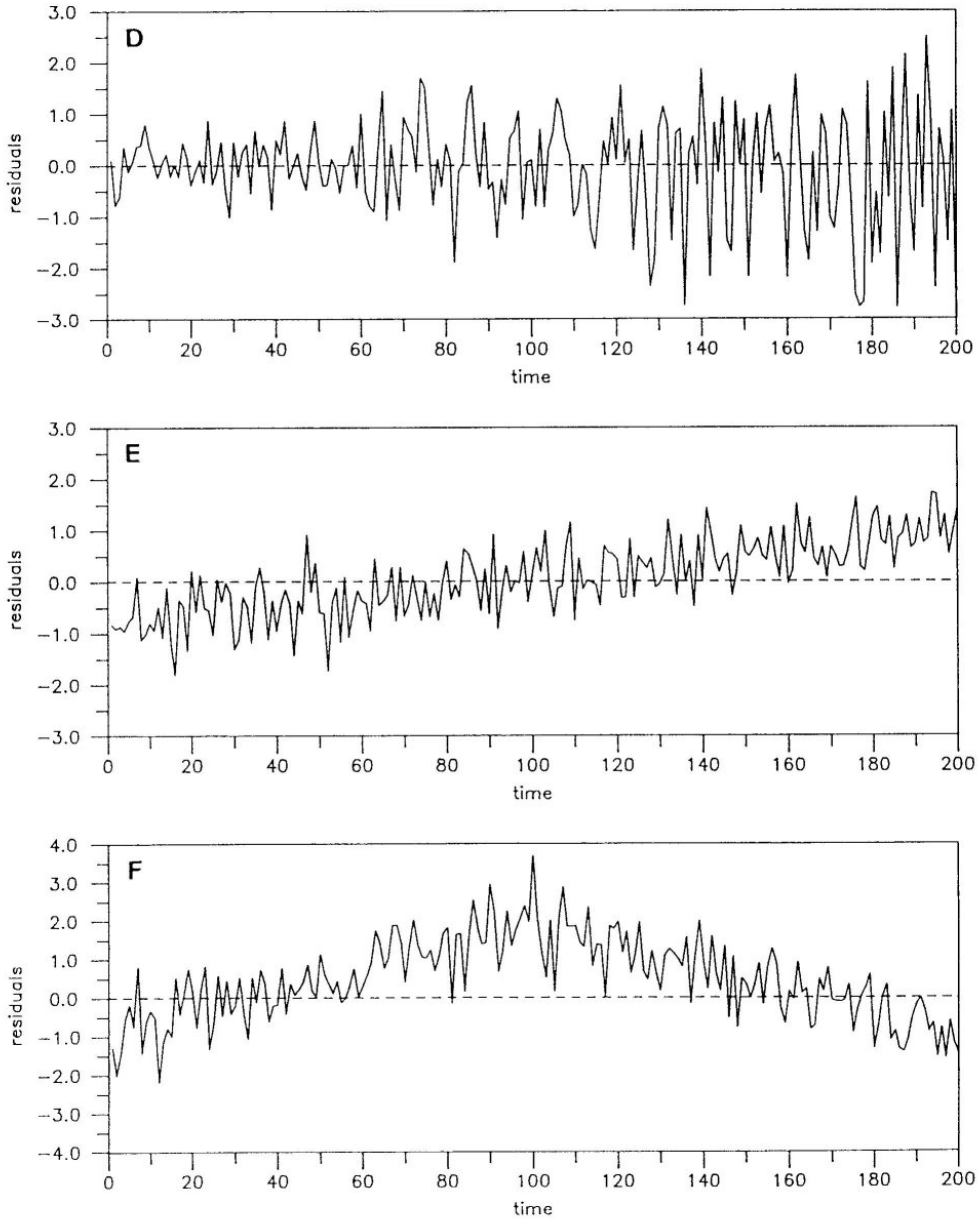


Fig.6.5 Residuals versus time diagram. D: shows the variance is time-dependent (heteroscedasticity); E and F: show the mean is time-dependent (trend)

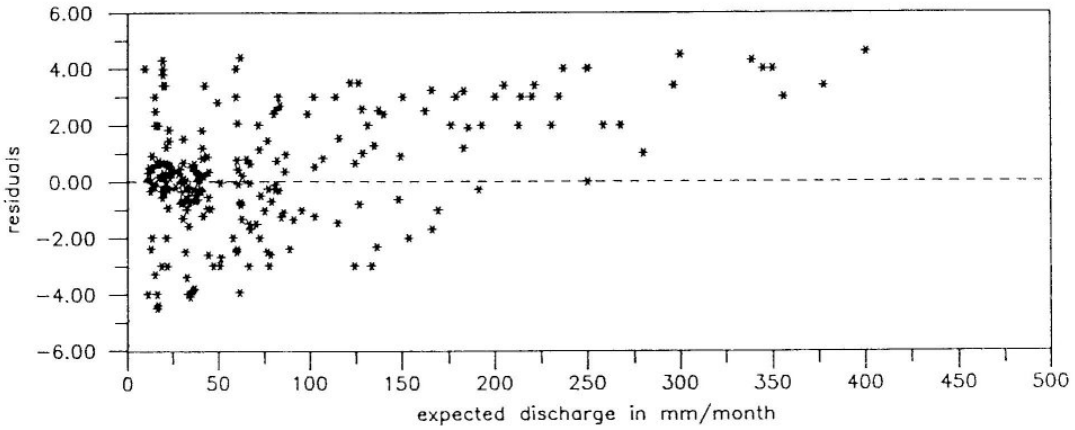


Fig.6.6 Residual versus expected response diagram. (The expected residual appears to increase with d_t and the variance decreases)

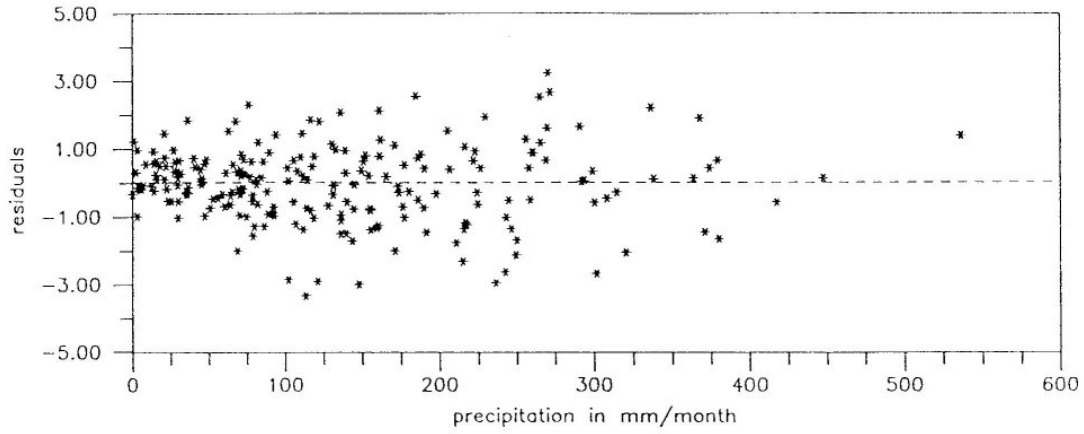


Fig.6.7 Residuals versus precipitation.

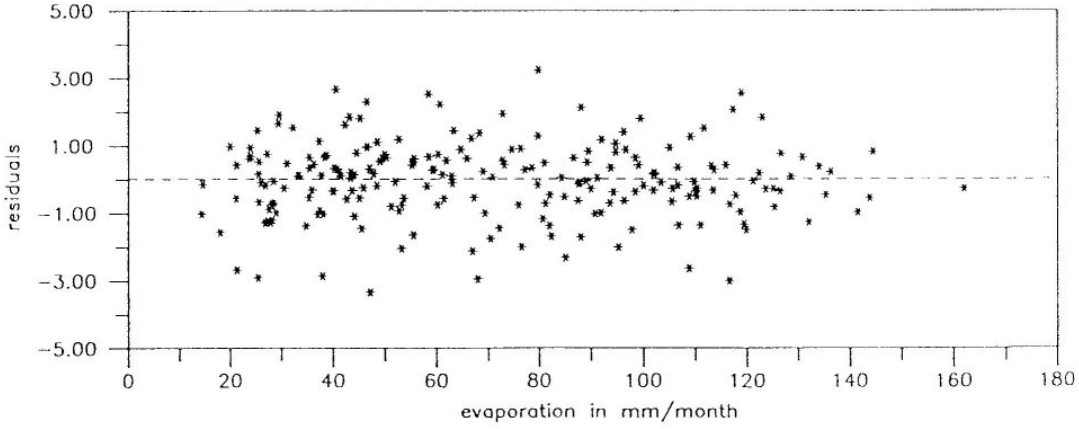


Fig.6.8 Residuals versus evaporation.

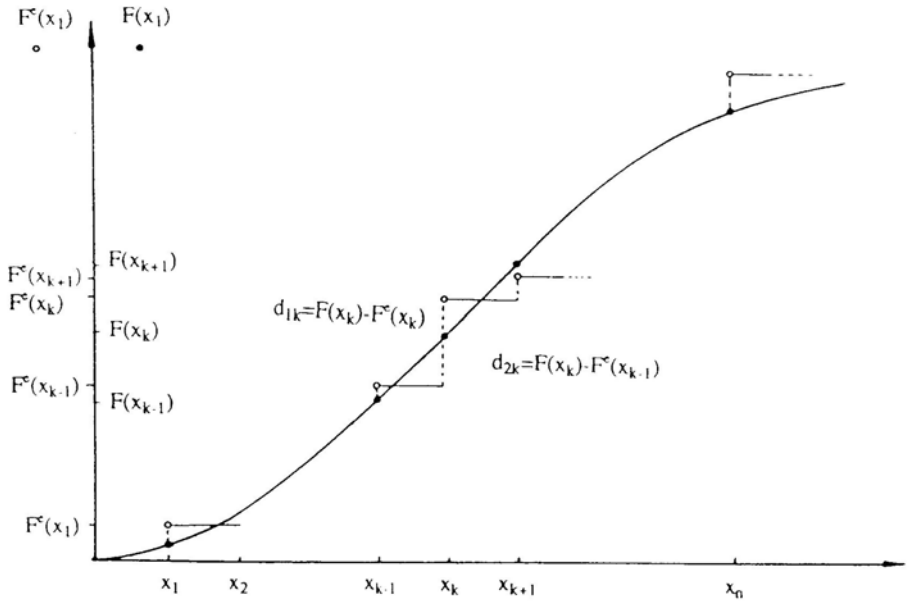


Fig.6.9 Deviation between experimental $F^e(x)$ and the theoretical F(x) distribution function values

6.4.4 Comparing modelled and observed runoff time series

Since the main aim of a rainfall-runoff model is to simulate the runoff series on the basis of rainfall record, perhaps the first and most striking comparison that can be made on the performance of a model can be seen from the plot of the observed and computed runoff series.

6.5 REGIONAL PARAMETERIZATION OF HYDROLOGICAL MODELS

6.5.1 The aims and principles of regionalization

The use of hydrological models in ungauged sites and in large geographical regions becomes a more and more important issue in hydrological study. The aim of regionalization study is to estimate parameter values of the hydrological models for any/every grid cell, sub-catchment or large geographic region without a need of calibration or "tune" the model to get the best fit. Regionalization methods aim to relate the model parameters to catchment characteristics and/or geographical location.

To be successful in the regionalisation study, the following principle is important:

- The parameter classes (soil types, vegetation types, climatological zones, geological layers, etc.) should be selected so that it becomes easy, in an objective way, to associate parameter values.
- It should explicitly be evaluated which parameters can be assessed from field data alone and which need some kind of calibration.
- The number of real calibration parameters should be kept low, both from practical and methodological points of view.

6.5.2 The methods of regional parameterization

A number of regionalization methods have been reported in the hydrological literature, which might be classified into two categories: point estimation methods and interval estimation methods.

6.5.2.1 Point estimation methods

The point estimation methods intend to provide unique value of each parameter for the ungauged catchment in case of lumped models or for each regular grid cell in case of distributed models. The point estimation methods usually do not take into consideration of parameter uncertainty.

It is noted here that in order to have better chance of success in the regionalization study it is important to list some basic requirements on the model, the model parameters and the catchments.

- First, to have meaningful statistical regression analysis, the number of gauged catchments used to optimized model parameters and establish regression equations should be more than 20, in any case not less than 10.
- Second, the number of parameters that needs to be regionalized should be kept to minima, i.e. the principle of parsimony is important in the analysis.
- Third, the automatic optimization technique should be used in order to get unique and repeatable value for each parameter in each gauged catchment. With manual calibration, every person who calibrates the model will get different values for the same parameter on the same catchment and one never knows which value should be used in the regression analysis.
- Four, most regionalisation methods assume that model parameters are independent and identically distributed for all catchments. Methods of statistical analysis of parameter values, as discussed in section 6.4.3 and more details in Xu (2001) should be performed in order to test the hypothesis, i.e., whether they are uncorrelated, identically distributed and statistically significant.

The proxy basin method

The proxy basin method for testing the geographic transferability of the hydrological models is used for any model that is assumed to be geographically transferable within a region hydrologically and climatically homogeneous. If the goal is to simulate streamflow for an ungauged basin C, then the model to be used should be calibrated on basin A and validated on basin B and vice versa. Only if both proxy-basin tests are acceptable should one consider the model as geographically transferable (Klemes, 1986). The proxy basin test has been the most common regionalization method. The main problems of the method include: (1) it is not possible to have any idea on the error of estimation on both parameter values and streamflow simulations, (2) it is not easy if not impossible to determine the degree of similarity between the ungauged catchment and the reference catchments, (3) last but not least, if there do not exist gauged catchments at the region the method is not useful. Examples of such tests include Xu (1999), Motovilov et al., (1999) and Refsgaard (1997).

Spatial interpolation methods

Spatial interpolation of model parameters is also a common method for hydrological regionalization. In order to get regional parameter values, the model parameters are firstly calibrated on all gauged catchments at the region, parameter values for the

ungauged catchments are interpolated by considering the soil type, vegetation distribution map, etc. using one of the following two techniques:

1) Direct interpolation: with the help of GIS and other computer tools, parameter values for the interested site or the whole region are interpolated. Examples of such a study include: Vandewiele and Elias (1995) interpolated parameters of a monthly water balance model to ungauged catchments using parameter values from neighbouring catchments, 75 in all, in Belgium. Bergström (1990) used the technique to make the map of parameter values of the HBV model for Sweden. Abdulla and Lettenmaier (1997) compared the interpolation technique with the multiple regression method. Guo (2001) used the interpolation technique to get the parameter values for the grid cells in ungauged sub-catchments and applied the macro-scale water balance model in a number of big catchments in China.
2) Geostatistical method – Kriging: Vandewiele and Elias (1995) compared the kriging method with the interpolation method in their geographical regionalization study and concluded that kriging method was better.

Multiple regression methods

Many attempts have been tried in the hydrological regionalization study to relate optimized parameter values (dependent variables) to catchment characteristics (independent variables) using the multiple regression technique (e.g., Jaboe and Haan, 1974; Magette et al., 1976; Andersen et al., 1983; Abdulla and Lettenmaier, 1997; Xu, 1999). Different forms of regression equations have been tried, among others, the common equations include:

- The linear method
  $Y = \beta_1 X_1 + \beta_2 X_2 + ...$
- The square root method 1
  $\sqrt{Y} = \beta_1 \sqrt{X_1} + \beta_2 \sqrt{X_2} + ...$
- The square root method 2
  $Y = \beta_1 \sqrt{X_1} + \beta_2 \sqrt{X_2} + ...$
- The logarithmic method 1
  $Log(Y) = \beta_o + \beta_1 log(X_1) + \beta_2 log(X_2) + ...$
- The logarithmic method 2
  $Y = \beta_o + \beta_1 log(X_1) + \beta_2 log(X_2) + ...$

Where Y is a dependent variable (model parameters in our case), X_1, X_2, ... are independent variables (catchment attributes), and β_1, β_2, ... are unknown regression parameters.

The dependent and independent variables might be transformed to account for nonlinear relationships as showed in some of the above equations. A multiple regression analysis is performed for each model parameter and which catchment characteristics to include for each parameter is an essential point. Sefton and Howarth (1998) presented a guide of how to carry out the analysis including exploratory correlation analysis, principal component analysis, stepwise regression and multiple regression.

An alternative to the univariate multiple regression is the multivariate regression (see Tung et al., 1997 and Engeland, 2002), which is not discussed in details in this section.

Regional calibration method

The above discussed regionalisation methods require at-site calibrations on each gauged site. This traditional at-site approach treats each site independently in an effort to obtain the best possible calibration at each site. The regional calibration approach, as described in Fernandez et al. (2000) attempts to get the best possible calibration at each site while simultaneously obtaining the best possible regional relationships between model parameters and basin characteristics. In this case the objective is to:

$$Maximiza\left[\frac{1}{m}\sum_{i=1}^{m}R_i^2\right]+\left[\frac{R_a^2+R_b^2+R_c^2+...}{n}\right] \tag{6.27}$$

where there are m = number of sites in the region, R_i^2 represents the coefficient of determination for site *i* which measures the goodness of fit of the logarithms of the modelled flows at site *i* and R_a^2, R_b^2, R_c^2, … represent the coefficient of determination associated with each of the regression models for the model parameters a, b, c, …, on basin characteristics, respectively. For example, R_a^2 is the coefficient of determination for parameter *a* regressed on basin characteristics. The idea of the objective function in equation (6.27) is to maximize the average goodness of fit of the model across all sites as well as to maximize the average goodness of fit of the regression equations that relate model parameters to basin characteristics. The results of the study showed that using the regional calibration approach improved the regional relationships between model parameters and basin characteristics significantly as compared with the traditional two-step regionalization approach, i.e. the multiple regression approach.

6.5.2.2 Interval estimation methods

Unlike in the point estimation methods, the interval estimation method intends to provide not a unique value for each parameter, they rather provide for each parameter with the most possible (maximum likelihood) value together with a probability distribution of the parameter values on the catchments. In other words, the methods provide a description of parameter uncertainty (Engeland, 2002).

The Bayesian method

The methodology, procedure and results of application are described in details in Engeland et al. (2001).

REFERENCES OF CHAPTER 6

Abdulla, F. And Lettenmaier, D.P., 1997a. Development of regional parameter estimation equations for a macroscale hydrologic model, Journal of Hydrology 197, 230-257.

Abdulla, F. And Lettenmaier, D.P., 1997a. Application of regional parameter estimation schemes to simulate the water balance of a large continent river, Journal of Hydrology 197, 258-285.

Andersen et al., 1983. Hydrologisk model for flomberegninger (in Norwegian), Hydrology department report no 2-83, NVE, Oslo, Norway, 40pp.

Bergström, S., 1990. Parametervärden för HBV-modellen I Sverige – Erfarenheter från modellkalibreringar under perioden 1975-1989 (in Swedish), SMHI Report Hydrology No 28, Norrköping, Sweden, 35pp.

Beven, K.J., Kirkby, M.J., Schofield, N. and Tagg, A.F., 1984. Testing a physically-based flood forecasting model (TOPMODEL) for three U.K. catchments, *J. Hydrol*., 69: 119-143.

Brazil, L.E., 1985. Multilevel calibration strategy for complex hydrologic simulation models. PhD Dissertation, Dept. of Civil Engineering, Colorado State University, Fort Collins, Colorado.

Dawdy, D.R. and R.W. Lichty, 1968. Methodology of hydrologic model building. In The use of analog and digital computers in hydrology. Int. Assoc. Sci. Hydrol. Symp. Proc., Tucson, AZ, 2(81):347-355.

Dooge, J.C.I., 1984. The water of the earth. *Hydrol. Sci. J.,* 29(2):149-176.

Duan, Q., 1991. A global optimisation strategy for efficient and effective calibration of hydrologic models, PhD Dissertation, Dept. of hydrology and Water Resources, The University of Arizona, Tucson, Arizona.

Engeland, K., 2002. Parameter Estimation in Regional Hydrological Models, PhD thesis, Department of Geophysics, University of Oslo.

Engeland, K., Gottschalk, L. and Tallaksen, L., 2001. Estimation of regional parameters in a macroscale hydrological model, Nordic Hydrology 32, 161-180.

Ewen, J. and Parkin, G., 1996. Validation of catchment models for predicting land-use and climate change impacts. 1. Method, *J. Hydrol*. 175: 583-594.

Fernandez, W., Vogel, R.M. and Sankarasubramanian, A., 2000. Regional calibration of a watershed model, Hydrological Sciences Journal 45(5), 689, 707.

Fulton, J., 1982. Discussion, modification, and extension of some maximum likelihood techniques for model calibration with application to rainfall-runoff models. M.S. Thesis, Case Western Reserve University.

Guo, S.L., Wang, J.X. and Yang, J., 2001. A semi-distributed hydrological model and its application in a macroscale basin in China, IAHS Publ No 270, 167-174.

Gupta, V.K. and S. Sorooshian, 1985. The relationship between data and the precision of estimated parameters, J. Hydrol. 81: 57-77.

Gupta, V.K. et al., 1998. Toward improved calibration of hydrological models: multiple and noncommensurable measures of information. Water Resour. Res. 34(4), 751-763.

Gupta, V.K., 1984. The identification of conceptual watershed models, PhD Dissertation, Case Western Reserve University, Cleveland, Ohio.

Haan, C.T., et al., 1982. Hydrologic modelling of small watersheds. American Society of Agricultural Engineers, St. Joseph, Michigan.

Ibbitt, R.P. and O'Donnell, T., 1971. Designing conceptual catchment models for automatic fitting methods. *Mathematical Models in Hydrology (Proc. of the International Conf.*, Sept. 24-28, Wageningen, p21-28.

Ibbitt, R.P., 1970. Systematic parameter fitting for conceptual models of catchment hydrology, PhD Dissertation, Imperial College of Science and Technology, University of London, London.

Jarboe, J.E. and Haan, C.T.: 1974. Calibrating a water yield model for small ungaged watersheds, *Wat. Resour. Res*. 10(2): 256-262.

Klemes, V., 1986. Operational testing of hydrological simulation models, Hydrol. Sc. J., 31: 13-24.

Lindström, G., 1997. A simple automatic calibration routine for the HBV model. Nordic Hydrology 28(3), 153-168.

Liong, S.Y. et al., 1996. Construction of multi-objective function response surface with genetic algorithm and neural network. In: Proceedings of the International Conference of Water resources and Environmental Research, 29-31 October, Kyoto, Japan, Vol.II, pp31-38.

Liong, S.Y. et al., 1998. Derivation of Pareto front with accelerated convergence genetic algorithm, ACGA. In: Babovic, V. and Larsen, L.C. (eds.). Hydroinformatics'98, Balkema, Rotterdam, The Netherlands, pp889-896.

Loague, K.M., 1990. R-5 revisited 2. Reevaluation of a quasi-physically based rainfall-runoff model with supplemental information, *Water Res. Resour*., 21: 973-987.

Madsen, H. 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, J. Hydrol. 235: 276-288.

Magette, W.L., Shanholtz, V.O. and Carr, J.C., 1976. Estimating selected parameters for the Kentucky watershed model from watershed characteristics, *Wat. Resour. Res*. 12(3): 472-476.

Motovilov, Y.G., Gottschalk, L, Engeland, K and Rodhe, A., 1999. Validation of a distributed model against spatial observations, Agricultural and Forest Meteorology 98-99, 257-277.

Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I - a discussion of principles. *J. Hydrol.*, 10: 282-290.

Refsgaard, J.C., 1997. Parameterization, calibration and validation of distributed hydrological models, Journal of Hydrology 198, 69-97.

Sefton, C.E.M. and Howarth, S.M., 1998. Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales, Journal of Hydrology 211, 1-16.

Sorooshian, S. 1981. Parameter estimation of Rainfall-runoff models with heteroscedastic streamflow errors: noninformative data case, In: Y. Haimes and J. Knolles (eds.), Water and Related Land Resources Systems, IFAC Symposium, Cleveland, Ohio, May 28-31. Pergamon Press, New York, N.Y., pp477-485.

Sorooshian, S. and J.A. Dracup, 1980. Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: correlated and heteroscedastic error cases, Water Resour. Res. 16(2): 430-442.

Sorooshian, S. and V.K. Gupta, 1995. Chapter 2: Model calibration. In: V.P. Singh (ed.) Computer Models of Watershed Hydrology, Water Resources Publications, Colorado, USA.

Sorooshian, S., 1978. Considerations of Stochastic Properties in Parameter Estimation of Hydrologic Rainfall-Runoff Models, PhD Dissertation, University of California-Los Angeles, 1978.

Sorooshian, S., V.K. Gupta and J.L. Fulton, 1983. Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runoff models – influence of calibration data variability and length on model credibility, *Water Resour. Res*. 19(1): 251-259.

Tung, Y.K., Yeh, K.C. and Yang, J.C., 1997. Regionalization of unit hydrograph parameters: a. Comparison of regression analysis techniques, Stochastic Hydrology and Hydraulics 11, 145-171.

Vandewiele, G.L. and Elias, A., 1995. Monthly water balance of ungauged catchments obtained by geographical regionalization, Journal of Hydrology 170, 277-291.

Xu, C-Y, 1999. Estimation of parameters of a conceptual water balance model for ungauged catchments, *Water Resources Management* 13(5): 353-368.

Xu, C-Y, 2001. Statistical analysis of a conceptual water balance model, methodology and case study, *Water Resources Management* 15: 75-92.

Xu, C-Y, 2001. Statistical analysis of parameters and residuals of a conceptual water balance model – methodology and case study, *Water Resources Management* 15: 75-92.

Xu, C-Y and G.L. Vandewiele, 1994. Sensitivity of Monthly Rainfall-Runoff Models to Input Errors and Data Length, *Hydrological Science Journal*, 39(2): 157-176.

Yapo, P.O. et al., 1998. Multi-objective global optimisation for hydrological models. J. Hydrol. 204, 83-97.

Yevjevich, V.M., 1972. Probability and statistics in hydrology. *Water Resources Publication, Fort Collins, Co.*

CHAPTER 7
SOME TOPICS IN OPTIMIZATION

7.1 GENERALITIES

In operational research one mostly reduces a decision problem to the choice of the values of real variables, the so-called decision variables. Therefore for all possible values one assesses the cost or the expected cost, or the profit or another value-criterion. One thus constructs a so-called *objective function* or *criterion function* or *economic function*. As a consequence the resulting problem is the optimization of a real function of a number of real variables, of which all real values are not necessarily admissible for all kinds of reasons. One thus imposes constraints in advance.

Also in statistics optimization problems appear when unknown model parameters have to be estimated. For example for finding the maximum likelihood estimates one has to maximize the likelihood function with respect to unknown parameters, of which the set of admissible values is given in advance (for example one knows that an unknown standard deviation is positive). Another example is the minimization of the sum of squared deviations (the so-called least squares method).

Also in other branches optimization problems appear.

These are examples of so-called mathematical programs.

A mathematical program is a problem of optimization of a real function $f(X)$, the objective/criterion function, with respect to a number of real variables $x_1, x_2, ..., x_n$, put together in the vector

$$X = (x_1, x_2, ..., x_n)$$

and where constraints on X are imposed in advance. These constraints imply that the vector X only takes values in a subset χ *of* R^n, the set of admissible values of X. In short a mathematical program has the form

$$\begin{array}{|c|}\hline \text{optimum } f(X) \\ X \in \chi \\ \hline \end{array}$$

An optimization can be a maximization or a minimization. If we have a method for maximizing $f(X)$, then we have at the same time a method for minimizing it by maximizing - $f(X)$.

In this context we are talking about global optimization, as opposed to the search for local optima, which are only in a neighbourhood, see Fig.7.1.

In some cases we have that

$\chi = R^n$,

and then the problem is said to be a problem of unconstrained optimization, but mostly χ is a proper subset of R^n, and then it is said to be a problem of constrained optimization.

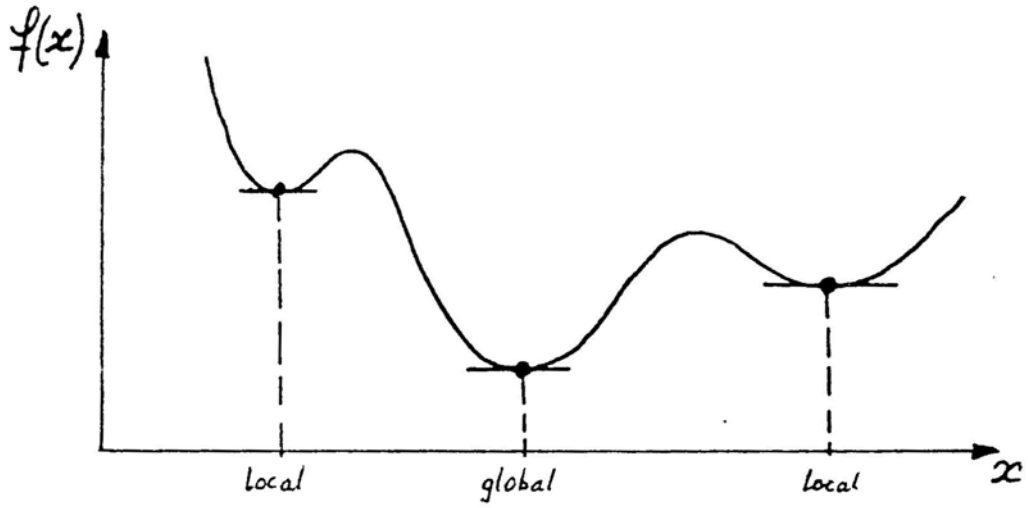


Fig.7.1. Global and local minima of a function f(x) of a scalar variable x

Nevertheless the case of unconstrained optimization is very important, either because the constraints are inactive practically or because the algorithms for constrained optimization often are reduced to a series of unconstrained optimizations.

The algorithm strongly depends on the form of f(X) and of χ .

7.2 OPTIMIZATION ALGORITHMS

7.2.1 Local search methods

Many different optimisation algorithms and computer codes are available, which can be categorized as “local search” methods and “global search” methods. Local search methods are designed to efficiently find the minimum of unimodal functions – functions for which any strategy that seeks to continuously proceed downhill (a direction of improving function value) must eventually arrive at the location of the function minimum, irrespective of where in the parameter space the search procedure is started. To understand this process, it is helpful to imagine that you are a blind person standing on the side of a mountain and must find a way from your present location to the lowest point of the valley. You would have to:

a) select a direction in which to move;
b) move some distance in that direction; and then
c) repeat the procedure over and over again until you are satisfied that no further improvement can be achieved.

Therefore, the procedure involves three main decisions:

(a) which direction to move,
(b) how far to move in that direction, and
(c) how to decide that no further improvement is possible.

Different local search strategies differ in the methods by which these decisions are made. Based on such differences, local search methods can be further classified as

"direct search" methods and "gradient" methods. We discuss these methods in more details below.

7.2.1.1 Direct search methods

A direct search optimisation strategy uses only (objective) function value information in the decision process described above. A typical method is the following. Starting at the initial point, the strategy selects some direction and step size and evaluates the function at the new point. It may do this for more than one new point. Then, based on the differences in function values between the initial and new point(s), a prediction is made of which is the best direction to move to improve the function and how large a step should be taken in that direction. Each direct search strategy has its own "idea" about how this should be done in order to quickly solve the problem. A step is taken in the trial direction, and if the new point has a lower function value than the previous point, it replaces the old one, and the procedure is repeated. If the new point turns out to be worse than the previous point, the step size is reduced and another try is made at a new location. The search stops when the strategy is unable to find a direction in which improvement is possible. In this section we discuss a number of simple methods, most of which are used as a base for more complicated methods, which try to compensate the disadvantages.

7.2.1.1a Linear search algorithms

With many search algorithms the following problem arises: starting in a point X_0 and with a given straight line through this point, find a new point on this line with smaller f(X) (supposing we have a minimization problem). A method solving this problem is called a linear search algorithm.

Let D be a set of direction numbers of the given line through X_0; the vector D thus has n components: all points X on the line then have the form

$$X = X_0 + \lambda D$$

where λ is a real number. We are looking for a value of λ such that

$$f(X_0 + \lambda D) < f(X_0)$$

we even can try to find a value of λ such that $f(X_0 + \lambda D)$ is as small as possible. This is a problem of so-called "linear minimization".

There are many linear search algorithms. We discuss one of them, the "doubling and halving algorithm" for illustrating purposes; it is not necessarily an efficient algorithm. For n = 2 it is illustrated in Figs.7.2 to 7.4.

Suppose it is known in which of both directions starting from X_0 the criterion function decreases; e.g. for $\lambda > 0$. Starting with a value $\lambda_0 (> 0)$ *of* λ, we compare

$$f(X_0) \text{ and } f(X_0 + \lambda_0 D)$$

If

$$f(X_0 + \lambda_0 D) < f(X_0)$$

we try the point $X_0 + 2\lambda_0 D$

If

$$f(X_0 + 2\lambda_0 D) < f(X_0 + \lambda_0 D)$$

we redouble the value of λ and compare $f(X_0 + 2\lambda_0 D)$ and $f(X_0 + 4\lambda_0 D)$. We continue redoubling until the new function value is greater than the preceding. The new (better) point is the penultimate value of λ. If on the contrary (9) is not true:

$$f(X_0 + \lambda_0 D) > f(X_0)$$

then we try $X_0 + \frac{\lambda_0}{2} D$

If

$$f(X_0 + \frac{\lambda_0}{2} D) \geq f(X_0),$$

we halve λ again, until the new function value is smaller than $f(X_0)$. The new (better) point then is the last value of λ.

This algorithm computes function values excluding computations of the derivatives. This algorithm can be ameliorated easily by quadratic or higher order interpolations. We do not insist further on this point.

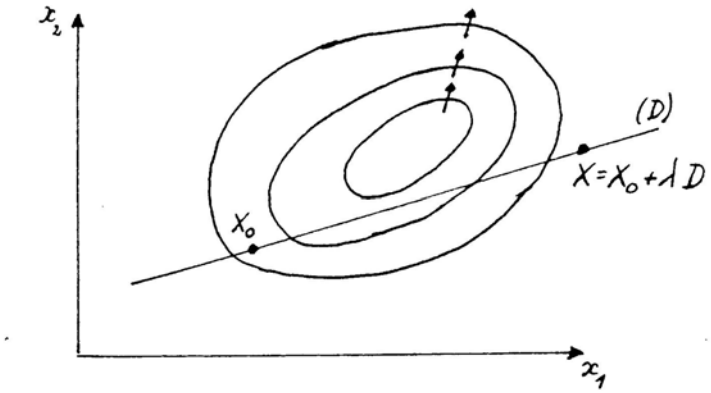


Fig.7.2. The linear minimization problem for n=2. The curves are contours of f(X).

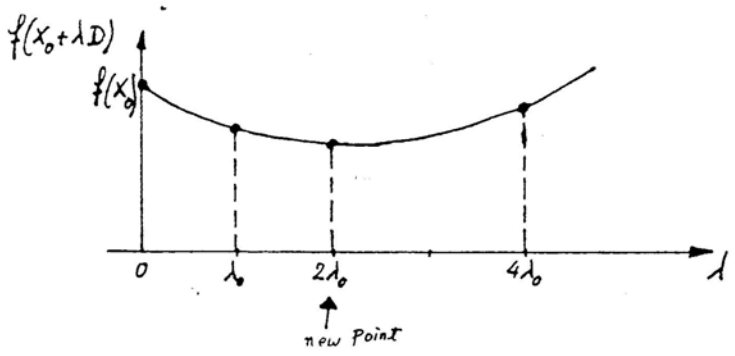


Fig.7.3. The algorithm of the text with $f(X_0 + \lambda D) < f(X_0)$

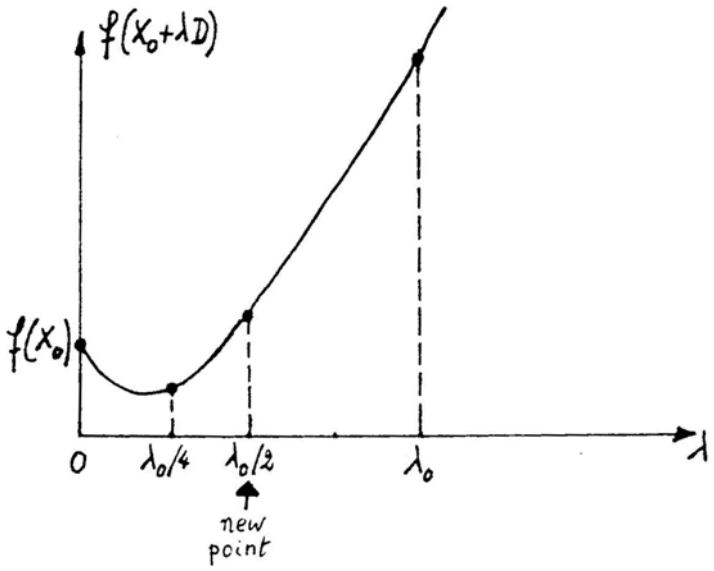


Fig.7.4. The algorithm of the text with $f(X_0 + \lambda_0 D) > f(X_0)$

7.2.1.1b Optimization along the directions of the axes

We now return to our original optimization problem. Starting in a first point X_0 we apply a linear search algorithm along the straight line through X_0 and parallel to the x_1-axis. In this way we find a second point X_1, where we apply a linear search algorithm along the straight line through X_1 and parallel to the x_2-axis, etc. This is illustrated in Fig.7.5 for n = 2.

For most functions f(X) an infinite number of steps are necessary for reaching the optimum. Therefore we have to use a rule for stopping the search; e.g. the search is stopped when the difference of two consecutive function values is smaller than a given quantity.

This method, like all methods discussed hereafter, leads to a local minimum, which is not necessarily the required global minimum. Therefore the choice of the initial point X_0 is important; it has to be chosen in the neighbourhood of the global minimum. This can be done taking account of the interpretation of the variables. It is recommended to repeat the optimization procedure for several initial points X_0; see Fig.7.6. Certainty is seldom achieved.

When a suitable linear search algorithm is used, the whole algorithm needs not to compute derivatives of f(X).

This algorithm makes very slow progress in narrow valleys of the f-hypersurface.

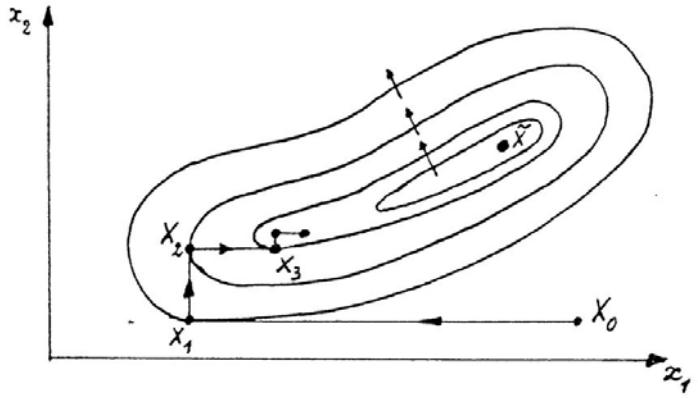


Fig.7.5. Minimization along the directions of the axes. The curves are contours of f(X).

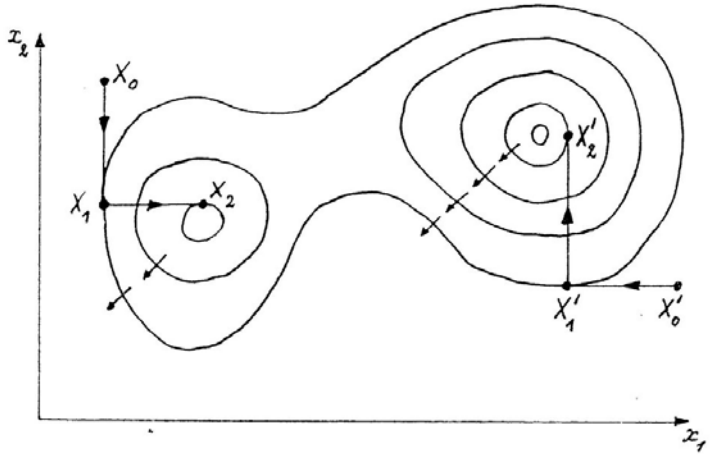


Fig.7.6. Criterion function with several minima.

7.2.1.1c An algorithm with "conjugated" directions

Methods based on the principle of conjugated directions try to bypass low progress in narrow valleys as mentioned in sections 7.2.1.1a and 7.2.1.1b.

The procedure of the algorithm with conjugated directions is illustrated schematically in Fig. 7.7 and with contours in Fig.7.8.

For n = 2 the simplest such method goes as follows.

- Along two neighbouring parallel lines with arbitrary direction D we search for the optima X_1^1 and X_1^2.
- Along the line $X_1^1X_1^2$ we search for the optimum X_1^3. Let D1 be the direction of this line:

  $D_1 = X_1^2 - X_1^1$

- Repeat step one and two. In general the i-th iteration optimizes along neighbouring lines with direction D, of which one line goes through X_{i-1}^3. These two optima define a line with direction

$D_i = X_i^2 - X_i^1$

along which we find by linear optimization the initial point X_i^3 of the (i+1)-th iteration.

The conjugated directions algorithm can be generalized to n variables.

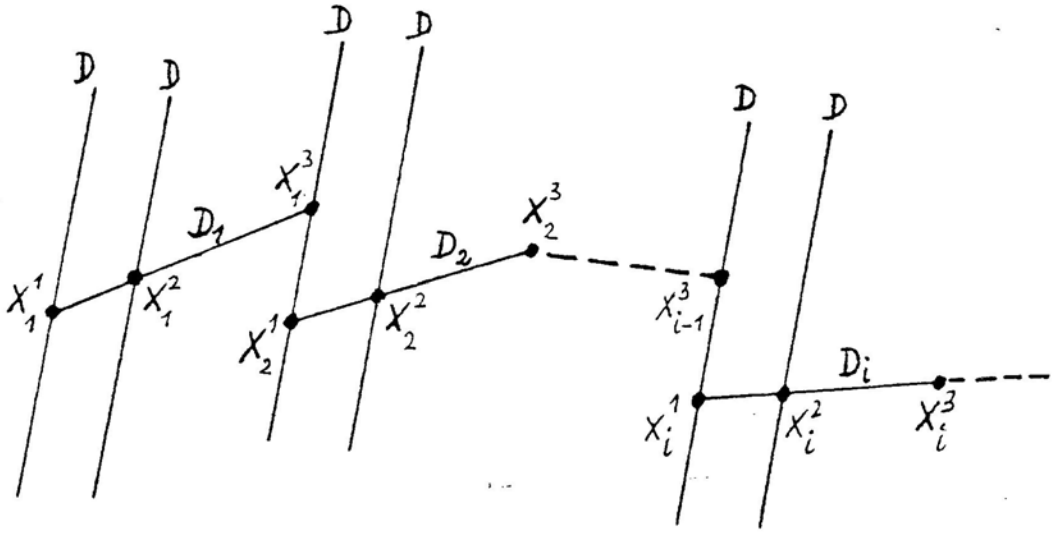


Fig.7.7. Scheme of a simple algorithm with conjugated directions.

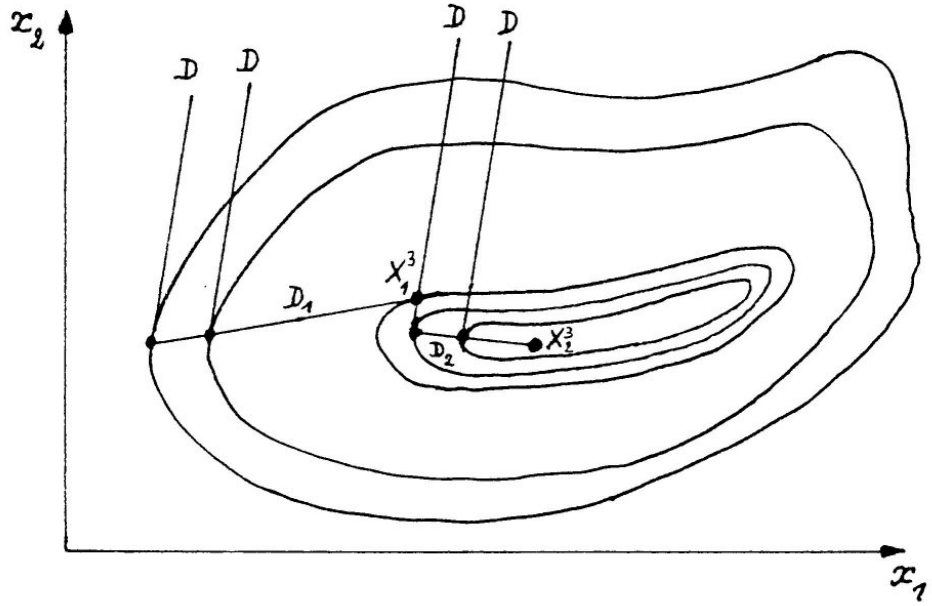
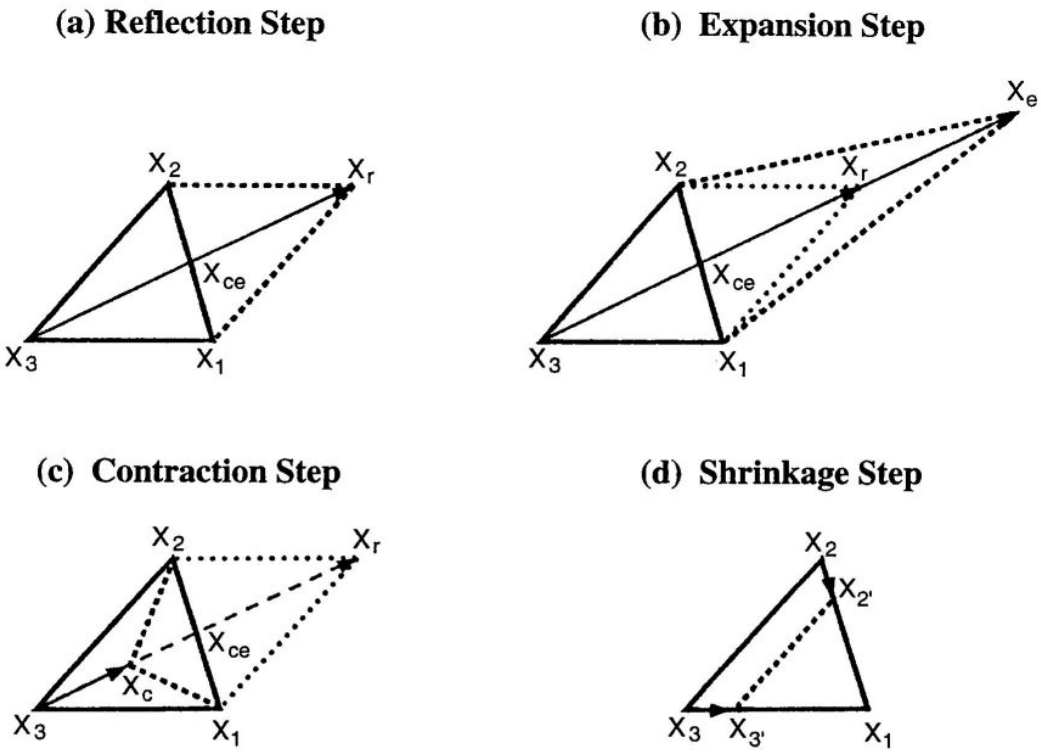


Fig.7.8. A simple algorithm with conjugated directions.

7.2.1.1d The Simplex search method

The Simplex search strategy is presented below and is illustrated in Fig.7.9 (after Sorooshian and Gupta, 1995).

(0) Select n+1 points (n = dimension) in the feasible parameter space and compute the function value at each point. This set of n+1 points is called a "simplex" (Fig.7.9).
(1) Identify the point with the worst (largest) function value.
(2) Compute the centroid of the best n points of the simplex (i.e., exclude the worst point).
(3) Locate a new point by reflecting the worst point through the centroid (see Fig.7.9a). If the function value at the reflection point is better than the worst point, go to Step 4; otherwise, go to Step 5.
(4) Locate a new point by expanding the reflection step by a factor of 2 (see Fig.7.9b). If the function value at the expansion point is better than the reflection point, replace the worst point by the expansion point. If not, replace the worst point by the reflection point. Go to step 7.
(5) Locate a new (contraction) point halfway between the worst point and the centroid (see Fig.7.9c). If the function value at the contraction point is better than the worst point, replace the worst point by the contraction point, and go to step 7. If not, go to Step 6.
(6) Shrink the simplex by moving each point (except the best point) to a location halfway between its current location and the best point (Fig.7.9d).
(7) Repeat Steps 1-6 until the size of the simplex becomes smaller than some convergence criterion. The point with the best function value is selected as an estimate of the optimum.



X_1 = the point with the lowest function value
X_2 = the point with the second lowest function value
X_3 = the point with the highest function value
X_{ce} = the centroid of the two best points X_1 and X_2
X_r = the reflection point
X_e = the expansion point
X_c = the contraction point
$X_{1'}$ and $X_{2'}$ = the shrinked points corresponding to X_1 and X_2

Fig.7.9 The Simplex direct search algorithm

7.2.1.2 Gradient methods

A gradient search optimisation strategy uses information about both the function value and the function gradient in the decision processes listed above. Most gradient methods are based on the following equation:

$$\Theta_{I+1} = \Theta_I - \rho \cdot A \cdot \nabla\Theta_I$$

where Θ_I is the present (initial) point (parameter vector), $\nabla\Theta_I$ is the function gradient matrix at the present point, ρ is a step size parameter, A is a specially chosen square matrix, and Θ_{I+1} is the new point. It can be mathematically shown that, if the matrix A is any positive definite matrix, the vector in the direction from Θ_I to Θ_{I+1} will be a

local improvement direction. Each gradient method uses different methods for selecting ρ and A. For example, the "steepest descent" approach uses the identity matrix (all diagonal values are one and off-diagonal values are zero). In general, most gradient methods use some approximation to the "Hessian" matrix (matrix of second partial derivatives of the function with respect to the parameters). If the exact Hessian matrix were to be used, the optimisation method would be the well-know "Newton method" which is the fastest way to solve a quadratic problem.

As with the direct search methods, if the new point has a lower function value than the previous point, the new point replaces the old one, and the procedure is repeated. If the new point turns out to be worse than the previous point, then the step size is reduced and another try is made at a new location. The search stops when the strategy is unable to find a direction in which improvement is possible. At this point, the gradient value will be numerically very close to zero.

7.2.1.2a Method of steepest descent or ascent

Here successive linear minimizations are applied along straight lines in the direction opposite to the local gradient (see Fig.7.10)

For n = 2 we know that G is the direction of the steepest line in the tangent plane of the f-surface. Therefore this algorithm is called the method of steepest descent (minimization problem) or of steepest ascent (maximization problem); it is called also the Cauchy-method. The efficiency of the method depends strongly on scale-transformations on the axes, since the concept of right angle is Euclidic. The method progresses very slowly in narrow valleys of the f-hypersurface.

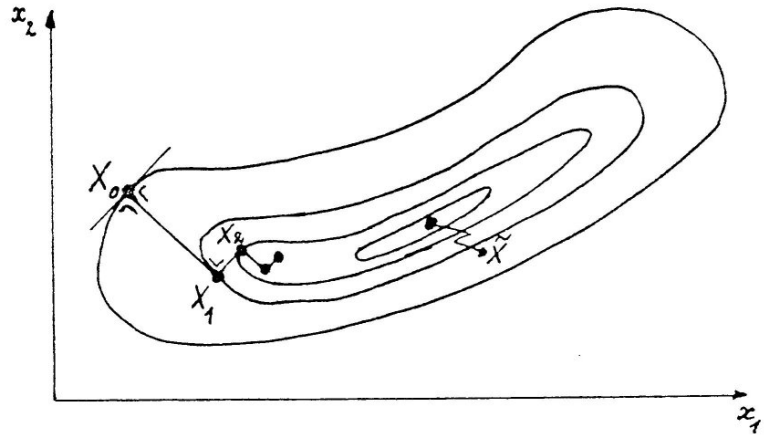


Fig.7.10. Method of steepest descent or ascent

7.2.1.2b Newton-Raphson algorithm

General aspect:

If the criterion function f(X) is differentiable, the vector of first derivatives is called the gradient

$$G(X) = \left(\frac{\partial f(X)}{\partial x_1}, \frac{\partial f(X)}{\partial x_2}, \ldots, \frac{\partial f(X)}{\partial x_n}\right)$$

If the criterion function is differentiable twice, the matrix of second derivatives is called the Hessian

$$H(X)=\begin{vmatrix} \frac{\partial^2 f(X)}{\partial x_1^2} & \frac{\partial^2 f(X)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(X)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(X)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(X)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(X)}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f(X)}{\partial x_n \partial x_1} & \frac{\partial^2 f(X)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(X)}{\partial x_n^2} \end{vmatrix}$$

Under very general conditions the Hessian is symmetrical:

$$\frac{\partial^2 f(X)}{\partial x_i \partial x_j}=\frac{\partial^2 f(X)}{\partial x_j \partial x_i}$$

If the Taylor expansion of f(X) exists for a value X_0 of X, it is written

$$f(X)=f(X_0)+G(X_0)\cdot(X-X_0)^T+\frac{1}{2}(X-X_0)\cdot H(X_0)\cdot(X-X_0)^T+\ldots$$

or for short

$$f(X)=f_0+G_0\cdot(X-X_0)^T+\frac{1}{2}(X-X_0)\cdot H_0\cdot(X-X_0)^T+\ldots$$

where T indicates transposition and where

$$f_0=f(X_0),\quad G_0=G(X_0),\quad H_0=H(X_0)$$

The Taylor expansion of the gradient is

$$G(X)=G(X_0)+(X-X_0)\cdot H(X_0)+\ldots$$

or

$$G(X)=G_0+(X-X_0)\cdot H_0+\ldots$$

If the second derivatives of f(X) are continuous for all X, then a necessary condition for $\widetilde{X}$ being a local or global minimum is that $G(\widetilde{X})=0$, and that $H(\widetilde{X})$ is positive semidefinite. A sufficient condition is that $G(\widetilde{X})=0$, and that $H(\widetilde{X})$ is positive definite.

An obvious solution method of the optimization problem is to solve the vector equation $G(\widetilde{X})=0$.

Newton-Raphson method:

When f(X) is a quadratic function, its Taylor expansion is a quadratic polynome in X. The Hessian H(X) does not depend then on X (second derivatives of a quadratic function!). This quadratic function has a minimum $\widetilde{X}$ if H(X) is positive definite, and we have

$$G(\widetilde{X}) = 0$$

The Taylor expansion of the gradient then is

$$G(\widetilde{X}) = G(X) + (\widetilde{X} - X) \cdot H(X) = 0$$

or

$$\widetilde{X} = X - G(X) \cdot H(X)^{-1}$$

As a consequence, starting with an arbitrary point X, we can find the minimum $\widetilde{X}$ by computing gradient G(X) and H(X).

The Newton algorithm (also called the Newton-Raphson algorithm) is based on the idea that f(X) can be approximated by its Taylor expansion truncated after the quadratic terms in X; consequently f(X) would be a quadratic function approximately. With an arbitrary point X we then associate the direction

$$-G(X) \cdot H(X)^{-1}$$

along which we perform a linear optimization. In the new point obtained, we repeat this operation, etc., until a stopping rule stops the algorithm.

This method requires the computation of first and second derivatives of f and also function values of *f* in view of the line minimizations. An important disadvantage in many circumstances is to have to compute second derivatives and to inverse the Hessians. The Hessian can even be singular, so that iteration is impossible. Locally, far from the minimum searched for, the Hessian can even be non-positive definite, which renders the methods much less attractive.

7.2.1.3 Choice of local search methods

In case the derivatives of f(X) are not available (e.g. f(X) is not differentiable), one has to use direct search algorithms. A first criterion thus is the availability of the elements, necessary for the algorithm.

Remark that gradient methods can be transformed into direct search algorithms, by approximating differentials by differences.

A second criterion is the amount of computation, which is necessary for sufficiently approaching the optimum. Computer time depends on the number of function values, first and second derivatives to be computed and on other computations (inversion of matrices, etc.) It is often difficult to estimate the convergence speed of an algorithm, especially the number of iterations necessary. Therefore the algorithms are tried out on so-called "test-functions". A theoretical criterion is the number of linear optimizations taken for reaching the optimum of a quadratic function.

The main weak-point of the local search methods is that most practical problems involving calibration of nonlinear hydrologic models have response surfaces that are

multi-modal – that is, there are several locations of the parameter space where the value of the function is a "local minimum". In such cases, the point where a local search algorithm terminates *will depend on the location where it is started.* As a result, it is difficult, if not impossible, to know with certainty if the procedure has located the actual "global" minimum of the function.

7.2.2 Global search methods

Global search strategies are designed to efficiently discover the minimum of multi-modal functions. Such strategies fall into three categories: deterministic, stochastic, or a combination of deterministic and stochastic. Deterministic strategies require that certain criteria related to the continuity of the function and its derivatives be satisfied to guarantee convergence to the global solution. These conditions are usually not met in the calibration of nonlinear hydrologic models. Only stochastic (random) and combination methods have been applied to the calibration of hydrologic models.

7.2.2.1 Random search methods

Random (stochastic) search methods use the random number generators built into modern digital computers to randomly sample the parameter space in search of points with improved function values. The samples are generated according to some probability distribution applied to feasible parameter space. In "pure" random search, the sampling is done using a uniform distribution. This assumes no prior knowledge of where in the feasible space the best parameter set exists. However, because pure random search does not make use of the function value information obtained during the search to guide the search, it is not very efficient, especially for problems with many parameters. Other random search methods have been developed which adaptively adjust the probability distribution used for sampling based on the function value information obtained during the search. One such method is the adaptive Random Search (ARS) proposed by Masri et al. (1978) and modified by Pronzato et al. (1984).

The ARS strategy is as follows (see also Sorooshian and Gupta, 1995):

(0) Choose a focal point (for example, this can be the "best" point obtained in the preliminary process of defining the parameter space, or it can be some arbitrary point such as the centroid of the feasible space);
(1) Generate a set of N points randomly distributed in the entire feasible space (for example, according to a uniform or normal distribution) and centered on the focal point. Store the location of the point with the best function value;
(2) Repeat Step (1) a pre-specified number of times, on the i^{th} time using the initial parameter range divided by 10^i and centered on the focal point (Fig.7.11) to restrict the search space. Each time, store the location of the point with the best function value;
(3) Compare all the stored points and determine the point with the best function value. Re-define this point to be the new focal point. Record in which range level this point was found; and
(4) Repeat Steps (1 – 3) until the best point is found in the smallest range level a user-specified successive number of times (say three). This point is chosen as the optimal parameter set.

Reports in the optimisation literature indicate that the ARS strategy works well in practice. However, Duan et al. (1992) found that the algorithm was neither effective nor efficient on a simple hydrologic model calibration problem. The best result they were able to obtain was a 30% success rate (30 out of 100 trials located the known region of the global optimum) after sampling the space 250 000 times.

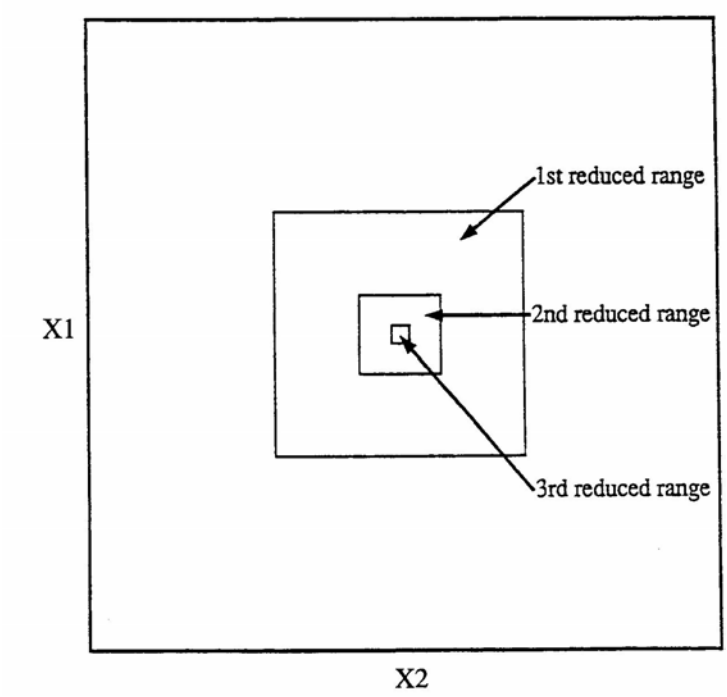


Fig.7.11 The adaptive random search algorithm

7.2.2.2 Multi-start algorithms

A simple *combination* method for dealing with multiple optima has been suggested in the hydrologic literature (e.g., Johnston and Pilgrim, 1976). In this method, one runs several trials of a local search optimisation method from randomly selected starting points in the feasible space. The validity of this "multi-start" approach can be demonstrated by the following arguments. To have confidence in the results of any stochastic optimisation procedure, we require that it has a relatively small failure probability on the problem of interest. Let us suppose that the failure probability of a local search optimisation method is F (i.e., if we were to run 100 independent randomly initiated tests of the method, we would find that $100 \times$ F of them would fail to locate the global optimum). Then, if we run the procedure r times from r independent randomly selected locations, the overall failure probability will decrease according to the equation $F(r) = F(1)^r$ and tend to zero as r becomes large. For example, if F is 0.65 (65 failures out of 100), then r equal to 12 will give a failure rate of less than 1 in 100.

The efficiency of any multi-start procedure varies nonlinearly with F, so that the number of restarts r required to achieve an overall failure probability of F(r) is given by $r = lh(F(r))/\ln(F(1))$. the curve of r versus F(1) is plotted in Fig.7.12 for the case of F(r) equal to 0.01 (1 failure in 100) and 0.05 (5 failure in 100). Clearly, for single-start failure probabilities F(1) that are less than 0.8, we do not require a very large number of restarts. However, as F(1) increases above 0.8 towards 1.0, the number of restarts required rapidly increases towards infinity, making the procedure impractical.

In Duan et al. (1992), it was demonstrated that a multi-start procedure based on the nonlinear simplex method (described earlier) worked well on simple hydrologic watershed model.

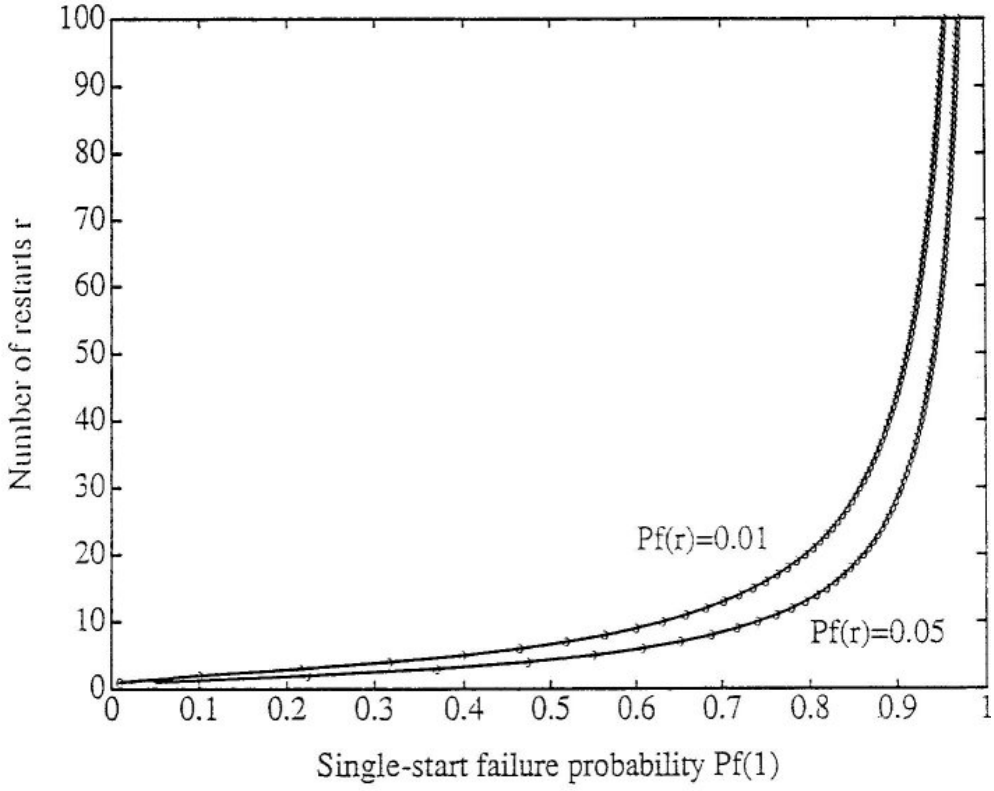


Fig.7.12 Theoretical performance of any multi-start algorithm

7.2.2.3 Shuffled complex algorithms

The Shuffled Complex Evaluation (SCE-UA) method, developed by Duan et al. (1993), is a global optimisation strategy designed to be effective and efficient for a broad class of problems. The SCE-UA strategy is presented below (see also Sorooshian and Gupta, 1995):

(0) Initialize: Select $p \geq 1$ and $m \geq n+1$, where p = number of complexes, m = number of points in each complex, and n = dimension of the problem. Compute the sample size $s = p \times m$.

(1) Generate sample: Sample s points $x_1,\ldots,x_s$ in the feasible space $\Omega = \mathbf{R}^n$. Compute the function value f_i at each point x_i. In the absence of prior information, use a uniform sampling distribution.

(2) Rank points: Sort the s points in order of increasing function value. Store them in an array $D = \{x_i, f_i,\ i = 1,...,s\}$, so that i = 1 represents the point with the smallest function value.

(3) Partition into complexes: Partition D into p complexes $A^1,...,A^p$, each containing m points, such that

$$A^k = \{x_j^k, f_j^k \,|\, x_j^k = x_{k+p(j-1)}, f_j^k = f_{k+p(j-1)}, j = 1,...,m\}.$$

(4) Evolve each complex: Evolve each complex $A^k, k = 1,..., p,$ according to the Competitive Complex Evolution (CCE) algorithm outlined separately.

(5) Shuffle complexes: Replace $A^1,..., A^p$ into D, such that D = $\{A^k, k = 1,..., p\}$. Sort D in order of increasing function value.

(6) Check convergence: If the convergence criteria are satisfied, stop; otherwise, return to Step (3).

The Competitive Complex Evolution (CCE) algorithm required for the evolution of each complex in Step (4) of the Shuffled Complex Evolution method is represented below:

(0) **Initialize:** Select q, α and β where $2 \le q \le m,\ \alpha \ge 1,$ and $\beta \ge 1.$

(1) **Assign weights:** Assign a trapezoidal probability distribution to A^k, i.e.,

$$\rho_i = \frac{2(m+1-i)}{m(m+1)},\ i = 1,...,m$$

The point x^k has the highest probability $\rho_1 = 2/(m+1).$ The point x_m^k has the lowest probability $\rho_m = 2/m(m+1).$

(2) **Select parents:** Randomly choose q distinct points $u_1,\ldots,u_q$ from A^k according to the probability distribution specified above (the q points define a "subcomplex"). Store them in array $B = \{u_i, v_i,\ i = 1,..., q\},$ where v_i is the function value associated with point u_i. Store in L the locations of A^k which are used to construct B.

(3) **Generate offspring:**

a. Sort B and L so that the q points are arranged in order of increasing function value. Compute the centroid g using the expression:

$$g = \frac{1}{q-1}\sum_{j=1}^{q-1} u_j$$

b. Compute the new point r = 2g-u_q (reflection step).

c. If r is within the feasible space **Ω**, compute the function value f_r, and go to step (d); otherwise, compute the smallest hypercube $H \subset R^n$ that contains A^k, randomly generate a point z within H, compute f_z, set r = z and set $f_r = f_z$ (mutation step).

d. If $f_r < f_q$, replace u_q by r, go to step (f); otherwise, compute c = (g + u_q)/2 and f_c (contraction step).

e. If $f_c < f_q$, replace u_q by c, go to step (f); otherwise, randomly generate a point z within H and compute f_z (mutation step). Replace u_q by z.

f. Repeat Steps (a) through (e) α times, where $\alpha \ge 1$ is a user-specified parameter.

(4) **Replace parents by offspring**: Replace B into A^k using the original locations stored in L. Sort A^k in order of increasing function value.

(5) **Iterate**: Repeat Steps (1) through (4) β times, where $\beta \geq 1$ is a user-specified parameter which determines how many offspring should be generated (how far each complex should evolve).

The version of SCE-UA used by Duan et al. (1993) used the values m = (2n+1), q = (n+1), $\alpha = 1$, and $\beta = (2n+1)$. Hence, the only variable to be specified by the user is the number of complexes p. In Duan et al. (1992) and Sorooshian et al. (1993), it was demonstrated that the performance of the SCE-UA method is far superior to that of the multi-start simplex (MSX) procedure. He claimed that the SCE-UA method appears to be the best method currently available for parameter estimation of conceptual watershed models.

7.3 DIFFICULTIES IN OPTIMIZATION

The successful application of a hydrologic watershed model depends on how well the model is calibrated. In recent years, automated approaches to calibration, such as those discussed in this chapter, have received much attention, and several difficulties in the application of such methods are (see also Ibbitt, 1970; Johnston and Pilgrim, 1976; Picup, 1977; Sorooshian and Gupta, 1983; Gan and Burges, 1990a,b):

(i) Interdependence between model parameters, by which a large number of combinations of parameter values will give similarly low values of the objective function - a change in the value of one parameter may be compensated by changes in one or more of the other parameters. For a two parameters model, a long flat-bottomed valley results in the response surface, as shown in Fig 7.13. A large number of combinations of parameter values will give similar low values of the objective function. Optimization methods make little or no progress along the floor of such valley towards its lowest point. It could be argued, of course, that this interdependence is not a problem, since any of the pairs of values in the valley is almost an optimum and the resulting output sequence is none the worse for the interdependence. However, if any meaning is to be attached to individual parameter values - if, say, parameter values are to be correlated with catchment characteristics - the values obtained from such an optimization would be meaningless (Gorgens, 1983a,b).

(ii) Indifference of the objective function to the value of a parameter, the calculated model output, and thus the value of the objective function, is not affected by changes in the value of a parameter because either the parameter is redundant or it is not active in the particular set of input data and even worse in a particular region of parameter space. Indifference causes zero gradients in some areas of the response surface, and optimization methods are not able to make further progress from such areas, as shown in Fig.7.13 for high values of X_1.

(iii) Discontinuities, or points on the response surface at which the objective function, while still continuous, is non-differentiable.

(iv) Local optima, these are points on the response surface that have lower values of the objective function than any surrounding point but have greater values than another point or points in another region of the response surface. As is shown in Fig.6.2 for a two dimensional response surface on which there is more than one closed contour for a given value of F. The 'peak' P_1 inside the contour F = 3 is the global optimum. If a local search method starts its search at point A in Fig.6.2 it would, in all probability, find a nearby local optimum such as P_2. Once at the local

optimum, the technique would be able to satisfy its built-in tests for convergence. For example, it would find that for small perturbations about P_2 only worse points could be found. Search methods have no tactics for moving to a higher peak from a lower one since they assume that only one peak exists. Local optima have all the properties of the global optimum except the value of the objective criterion F.

(v) Scaling of parameters. Different scaling of parameters changes the configuration of the response surface, affecting the difficulty of optimization, as is illustrated in Fig.7.14. Progress may be greatly improved by rescaling of parameters with the aim of producing near-circular contours of the objective function. However, the form of the response surface, particularly in multidimensional space, and the best selection of transformations are not known. Experimentation with scaling parameters is therefore desirable and may lead to more rapid progress.

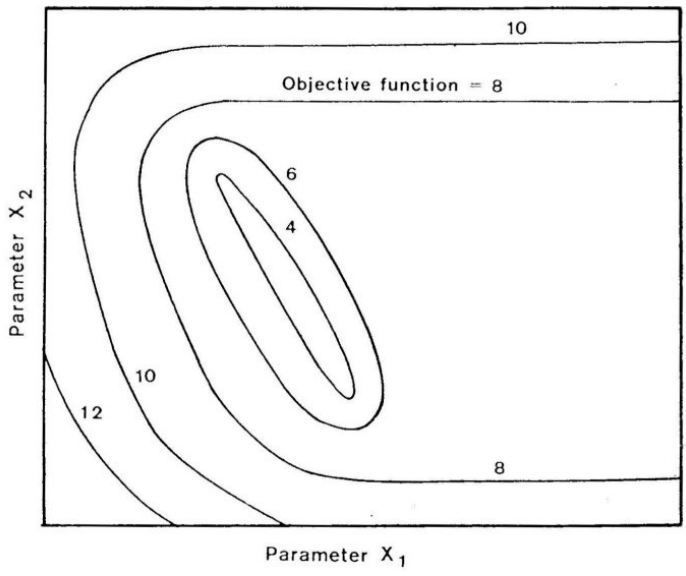


Fig.7.13 A two-dimensional response surface (hypothetical)

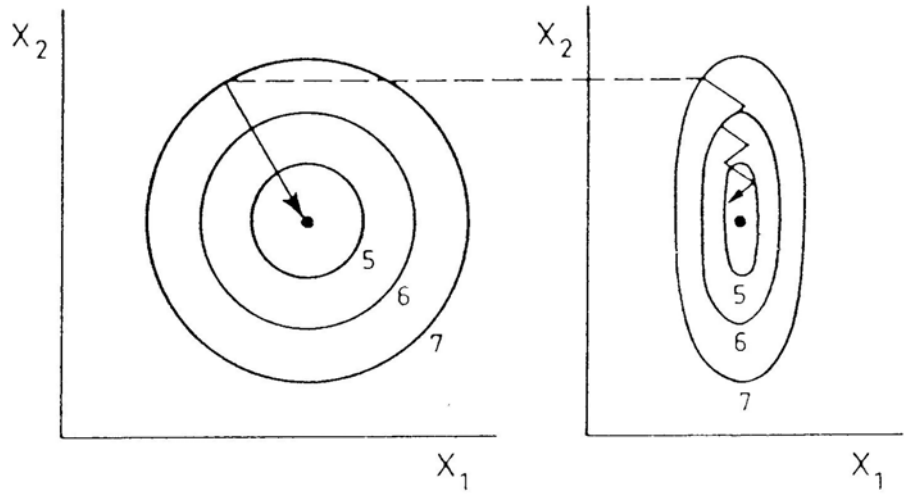


Fig.7.14 The influence of scaling on steepest descent searches

Moreover, deficiencies in model structure can also cause problems in optimization, since even the most complex models are imperfect representations of physical processes and this imperfection may lead to difficulties in optimization.

The choice and the role of the objective function are aspects of optimization that also offer serious difficulties to the modeller. It is axiomatic that the optimal set of parameters arrived at by optimization is in fact optimal only in the context of the objective function used during the process. A substantially different objective function may converge on substantially different optimum parameter set, though other conditions of optimization remain unchanged (Diskin and Simon, 1977; Pilgrim, 1975).

Studies of some effects of these difficulties have been reported by O'Connell, Nash and Farrell (1970), Mandeville, et al. (1970), Ibbitt and O'Donnell (1971), Plinston (1971), Johnson and Pilgrim (1973, 1976); Pilgrim (1975) and Gorgens (1983a). Strategies for overcoming some of these problems are suggested, although not all the problems can be solved. Possible solutions to these problems include the following measures (Gorgens, 1983a). Problem (i) can be partially redressed by optimizing interdependent parameters individually in separate searches. Problem (ii) can be avoided by setting indifferent parameters to constant values. Problem (iii) affects only steepest descent algorithms and cannot be solved except by multiple searches from different points on the response surface. Problem (iv) can often be overcome by changing initial points. Point (v) is less a problem in direct search than in steepest descent methods and can be redressed by either scaling parameters to the same order of magnitude or weighting the search steps for individual parameters according to parameter scale.

A measure that is often used to cope with more than one of the above difficulties is to constrain the values of certain parameter to a "likely range" during optimization, i.e., to prevent 'impossible' values from being chosen by the search routine or for the routine to wander into one of the difficulty-prone areas of the response surface. However, Pilgrim (1975) argues that this procedure is unjustifiable because a parameter value might pass through an impossible region during the search but then return to a realistic level. Imposition of limits also implies that the model structure and the parameters are indeed physically realistic and that the data contain no serious errors. Chapman (1975) argues conversely, i.e., that modellers should consciously strive to make their models more physically-based; then crucial parameters must be constrained to known physical limits commensurate with each catchment situation.

REFERENCES OF CHAPTER 7

Chapman, T.G., 1975. Trends in catchment modelling, *Prediction in Catchment Hydrology,* Chapman, T.G. and Dunin, F.X. (Eds.), Australian Academy of Science, Canberra, p459-482.

Diskin, M. and Simon, E., 1977. A procedure for the selection of objective functions for hydrologic simulation models. *Mathematics and Computers in Simulation XXIV,* North-Holland Publishing Company, 122-130.

Duan, Q., Gupta, V.K. and Sorooshian, S., 1993. A Shuffled complex evolution approach for effective and efficient global minimization, J. of Optimisation Theory and Applications 76(3): 501-521.

Duan, Q., Sorooshian, S. and Gupta, V.K., 1992. Effective and efficient global optimisation for conceptual rainfall-runoff models, Water resour. Res. 28(4): 1015-1031.

Gan, T.Y. and S.J. Burges, 1990a,b. An assessment of a conceptual rainfall-runoff model's ability to represent the dynamics of small hypothetical catchments 1&2. Water Resour. Res. 26(7).

Gorgens, A.H.M., 1983a. Conceptual modelling of the rainfall-runoff process in semiarid catchments. Report 1/83, Hydrological Research Unit, Rhodes University, Grahamstown, 347p.

Gorgens, A.H.M., 1983b. Reliability of calibration of a monthly rainfall-runoff model: the semiarid case. *Hydrol. Sci. J.*, 28(4):485-498.

Ibbitt, R.P. and O'Donnell, T., 1971. Designing conceptual catchment models for automatic fitting methods. *Mathematical Models in Hydrology* (Proceedings of Warsaw Symposium), IAHS Publ. No.101: p461-475.

Johnson, P.R. and Pilgrim, D.H., 1973. A study of parameter optimisation for a rainfall-runoff model, Report No.131, Water Resources Laboratory, University of New South Wales, Austrlia.

Johnson, P.R. and Pilgrim, D.H., 1976. Parameter optimization for watershed models. *Wat. Resour. Res.*, 10(2):256-262.

Mandeville, A.N., et al., 1970. River flow forecasting through conceptual models. Part III - the Ray catchment at Grendon Underwood, *J. Hydrol.*, 11: 109-128.

Masri, S.F., et al., 1978. An adaptive random search method for identification of large scale nonlinear systems, Proceedings of 4th IFAC Symposium on identification and system parameter estimation 1976, North-Holland.

O'Connell, P.E., et al., 1970. River flow forecasting through conceptual models. Part II - the Brosna catchment at Ferbane, *J. Hydrol.*, 10: 317-329.

Picup, G., 1977. Testing the efficiencies of algorithms and strategies for automatic calibration of rainfall-runoff models. Hydrological Science Bulltin 22(2): 257-274.

Pilgrim, D.H., 1975. Model evaluation, testing and parameter estimation in hydrology. *Prediction in Catchment Hydrology,* Chapman, T.G. and Dunin, F.X. (Eds.), Australian Academy of Science, Canberra, p305-333.

Plinston, D.T., 1971. Parameter sensitivity and interdependence in hydrological models. *British Ecological Society 12th Symposium,* Grange-over-Sands, p237-247.

Pronzato, L.E. et al., 1984. A general purpose global optimiser: implementation and applications, *Mathematics and computers in Simulation*, Vol.26: 412-422.

Sorooshian, S. and V.K. Gupta, 1983. Automatic calibration of conceptual rainfall-runoff models: the question of parameter observability and uniqueness, Water Resour. Res. 19(1): 251-259.

Sorooshian, S. and V.K. Gupta, 1995. Model calibration, In: V.P. Singh (ed.) Computer Models of Watershed Hydrology, Chapter 2.

CHAPTER 8
SOME PARTICULAR CATCHMENT MODELS

The following hydrological models will be discussed in details during the lecture. Documents about the listed models will be distributed to the students before the course starts. Other hydrological models, which are not listed below might also be discussed.

8.1 WASMOD – A conceptual, stochastic, lumped water balance model.
Reference: Xu, C-Y, 2002. WASMOD – The Water And Snow Balance Modeling system. In: Mathematical Models of Small Watershed Hydrology and Applications (Chapter 17), V.P. Singh and D.K. Frevert (eds), Water Resources Publications, LLC, Highlands Ranch, Colorado, USA.

8.2 HBV-model – A conceptual, deterministic, lumped (semi-distributed) daily rainfall-runoff model.
Reference: Bergström, S., 1995. The HBV Model. In: Computer Models of Watershed Hydrology (Chapter 13), V.P. Singh (ed.), Water Resources Publications, Highlands Ranch, Colorado, USA.

8.3 TOPMODEL – A physically-based, semi-distributed model.
Reference: Beven, K. et al., 1995. TOPMODEL. In: Computer Models of Watershed Hydrology (Chapter 18), V.P. Singh (ed.), Water Resources Publications, Highlands Ranch, Colorado, USA.

8.4 SHE model – A physically-based, deterministic, distributed model.
Reference: Abbott, M.B., et al., 1986. An introduction to the European Hydrological System – Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system. Journal of Hydrology 87: 61-77.