WORKSHOP

ON

MODELLING OF HYDROLOGIC SYSTEMS

4-8 September, 2000

Hydrological Modelling - General Concepts

by
A. V. Shetty

Organised by

Hard Rock Regional Centre
National Institute of Hydrology
Belgaum-590 001 (Karnataka)

Hydrological Modelling- General Concepts

A. V. Shetty
Scientist 'C'
Hard Rock Regional Centre
National Institute of Hydrology
Belgaum-590 001 (Karnataka)

1.0 Introduction

Hydrology deals principally with movement, distribution and storage of moisture. Most hydrologic problems are related to either quantity or quality of water or both. Determination of water yield, streamflow hydrograph, frequency of magnitude, volume, duration and interarrival times of flood-peaks dam breach etc. are some typical water quantity problems. These problems can be addressed in:

(a) Time domain-involving reconstruction of the past (prediction) and construction of the future (forecasting) on different scales viz. Continuous time or desecrate time such as hour or less, daily, weekly, ten daily, monthly, seasonally, annual and longer.

(b) Space domain-involving spatial variability and its sampling, regionalisation, effect of land use change etc. on different scales such as channel, field or plot, watershed river basin consisting of number of watersheds, continental or global.

(c) Frequency domain-involving determining frequency of Extremes (high as well as low), volumes, means, hydrologic space-time characteristics, etc.

In general, various approaches to the study of hydrologic problems can be grouped under two categories: (1) physical science approach - also referred to as a basic pure, causal, dynamic or theoretical approach, and (2) systems approach - also known as an operational, applied, empirical, black box or parametric approach.

Water management is the application of all available knowledge to the practical development of water resources. One of the fundamental sciences of water management is hydrology. The engineers engaged in design, construction or operation of hydraulic works must solve practical problems. These are of varied nature and in most cases hydrology is needed for their solution. These include (a) rural water management (b) river training (c) municipal water management (d)structural and hydraulic design of water control structures for different purposes. Some typical questions that the hydrologist is called upon to answer are: (i) Is the flow of a particular stream at a particular site is sufficient to meet the needs of (a) a city or industry seeking a water supply (b) an irrigation project (c) a proposed power development (d) navigation (e) recreation. (ii) Would a storage reservoir be required in connection with any of the proposed uses and if so, what should be its capacity? (iii) In the design of a flood that may be expected to occur with any specified frequency. (iv) What would be the effect of draining upland area or a swampy region upon the flow of the stream from watershed? (v) How would certain changes in land use or the removal of forests, affect the groundwater level or the stream flow from such an area?

2.0 Deterministic and Stochastic Processes

Hydrologic models are mathematical formulations to simulate natural hydrologic phenomena which are considered as 'processes or as systems'. Any phenomena which undergoes continuous changes particularly with respect to time may be called a process. As practically all hydrologic phenomena change with time, they are hydrologic processes. The hydrologic processes and their models can be divided into two broad classes:

2.1 Deterministic process

If the chance of occurrence of the variables involved in a process is ignored and the model is considered to follow a definite law of certainty but not any law of probability, the process and its model are described 'deterministic'. For example, the conventional routing of flood flow through a reservoir is a deterministic process and the mathematical formulation of unit hydrograph theory is a deterministic model.

2.2 Stochastic or probabilistic process

If the chance of occurrence of the variables is taken into consideration and the concept of probability is introduced in formulating the model or process, then process and its model are described as 'stochastic or probabilistic'. For example, the probability of the flow is taken into account in the probability routing, the process and the governing model employed to simulate the process the process are considered as stochastic or probabilistic.

3.0 Catchment as a System

Dooge defines a system as 'Any structure, device, a scheme or procedure, real or abstract, that interrelates in a given time reference, an input, cause or stimulus, of matter, energy or information. And an output, effect or response, of information, energy or matter.

The hydrological cycle of a drainage basin is a sequential, dynamic system in which water is a major throughput. Actual hydrologic system is a non-stationary stochastic process. However, since it is very complicated mathematically, the hydrologic system is generally treated as deterministic and modeled by deterministic model e.g instantaneous unit hydrograph.

In practice, the hydrologist confines his attention to individual basins or catchment areas. Thus he leaves the problems of the atmosphere to the meteorologist, those of the lithosphere to the geologist and those of the seas to the oceanographer. This narrows down his concern to the particular subsystem of the total hydrological cycle.

Though classical hydrology described the hydrological cycle in terms of surface runoff, interflow, and groundwater flow. In practice quantitative hydrology usually ignores this three fold division and considers the hydrograph being made up of a direct storm response and a base flow. Thus in the analysis of the relationship between storm rainfall and flood runoff, the system analyzed by the practical hydrologist corresponds closely to that indicated above.

The catchment system in this simplified approach consists of three subsystems.

1. The sub system involving direct storm response.
2. The sub system involving groundwater response.
3. The sub system involving the soil phase which has a feedback loop to the separation of precipitation into precipitation excess and infiltration.

There are many ways to classify systems and their models (Singh, 1988). Each classification is based upon a particular set of system characteristics. The same system can therefore fit into different categories. Thus systems could be (I) abstract or physical, (ii) natural or devised (iii) Open loop or closed loop, (iv) simple or complex (v) stable or unstable (vi) damped or undamped (vii) continuous time or discrete time (viii) causal or non causal (ix) memory or no memory (x) time invariant or time variant (xi) linear or non-linear (xii) lumped or distributed, and (xiii) deterministic or stochastic. Hydrologic systems are normally physical, sequential dynamic, natural, open loop, complex, stable, damped, causal, memory, stochastic, time variant, not-linear distributed systems.

The hydrologic behavior of watershed is a very complicated phenomenon which is controlled by an unknown large number of climatic and physiographic factors that very with both time and space. The catchment behavior is distributed, nonlinear, time variant and stochastic in nature.

The physics of separate hydrologic processes is known and the differential equations governing their deterministic behavior can be written for physically homogeneous basins. In using these equations to study natural heterogeneous systems, the system is lumped into elements which are effectively homogeneous. This method can be improved by considering the coefficients of the these equations to have a random component. However, in seeking general understanding of complex hydrologic systems there are conceptual and computational limitations in considering non-linear stochastic system. Hence deterministic linear system approach is adopted.

4.0 Hydrological Models

In simple terms a hydrological model is a simplified description of (parts of) the hydrological cycle. However, the term hydrological model is often understood to be and is used more narrowly as a computer-based mathematical model. The development and application of such models have increased tremendously during the last two or three decades, so that engineering hydrology today usually involves consideration of some kind of hydrological model. With the current rapid developments within computer technology and hydrology the application of computer-based hydrological models can only continue to increase in the near future (Storm, 1989).

A mathematical model provides a quantitative mathematical description of the processes or phenomena, i.e. a collection of mathematical equations (often partial differential equations), logical statements, boundary conditions and initial conditions, expressing relationships between inputs, variables and parameters.

The usual aim is to model the interaction of an input (e.g. rainfall) with a system (e.g. a catchment) to produce an output (e.g. the outflow hydrograph). The hydrological cycle is represented mathematically to imitate the natural system. The mathematical functions employed can be designed to simulate the natural hydrological processes as closely as present knowledge, mathematical constraints, data availability and user requirements allow. Depending on the required accuracy of results, effort to be spent in data collection, effort to be spent in modeling and available funds, the model can approximate the natural system more or less closely (storm, 1989).

4.1 Classification

Hydrological models can be classified in different ways. Two main groups of mathematical methods emerge from those which involve optimization and those which do not. Here optimization is referred to strictly in the sense of decision making rather than in the optimization of model parameters. The non-optimizing methods are generally associated with the assessment of hydrological data and are used to quantify the physical process. Methods involving optimization are concerned with the problem of selecting the "best" solution among a number of alternatives in a planning process.

Non-optimizing methods are divided into two fundamentally different approaches, the deterministic and the statistical. However, although the deterministic and the statistical methods are fundamentally different, a strong interplay between the two approaches exist, mainly because the processes involved in the hydrological cycle are partly causal and partly random. Hence, some deterministic models contain random functions to relate processes, while some statistical models contain causal or deterministic functions as part of their structure. The interplay between the two approaches also includes the subsequent analysis of the information gained from the different models (storm, 1989).

4.2 Black Box or Empirical Models

These contain no physically-based transfer function to relate input to output: n other wards no consideration of the physical process is involved. Such models usually depend upon establishing a relationship between input and output, calibrated from existing hydrometeorological records. Within the range of calibration data such models may be highly successful, often because the formal mathematical structure carries with it an implicit understanding of the underlying physical system. However, in extrapolating beyond the range of calibration, the physical link is lost and the prediction then relies on mathematical technique alone. Given the inherent linearity of many back-box models, which contrasts with the non-linearity of hydrological systems, such extrapolation is of dubious worth and is not recommended. Thus, for example, black box models cannot be used to predict the effects of a future change in land-use. Probably the best known black box models in hydrology are the unit hydrograph principles.

Black box models were developed and extensively applied before advances in computer technology made it possible to use more physically correct (and thus more complex) models. Today, .black box principles are more often used to form components of a larger model. E.g. the unit hydro graph is often used for streamflow routing in conceptual rainfall-runoff models.

4.3 Lumped Conceptual Models

These occupy an intermediate position between the fully physically-based approach and empirical black-box analysis. Such models are formulated on the basis of a relatively small number of components, each of which is a simplified representation of one process element in the system being modeled.

The mode of operation may then be characterized as a book-keeping system continuously accounting for the moisture contents in the storage. The non-linear form of such models reflects the thresholds present in hydrological systems, which cannot adequately by incorporated within a linear model. The source of this non-linearity is often the soil moisture condition, whether controlling groundwater recharge or surface/subsurface storm runoff.

4.4 Fully Distributed Physically Based Models

These are based on our understanding of the physics of the hydrological processes which control catchment response and use physically-based equations to describe these processes. From their physical basis such models can simulate the complete runoff regime, providing multiple outputs (e.g.river discharge, phreatic surface level and evaporation loss) while black box models can offer only one output. Also, almost by definition, physically-based models are spatially distributed since the equations from which they are formed generally involve one or more space coordinates. They can therefore simulate the spatial variations in hydrological conditions within a catchment as well as simple outflows and bulk storage volumes. On the other hand, such models make huge demands in terms of computational time and data requirements and are costly to develop and operate.

Unlike lumped conceptual models, physically-based distributed models do not consider the transfer of water in a catchment to take place between a few defined storage. Instead the transfers of mass momentum and energy are calculated directly from the governing partial differential equations, for example the Saint Venant equations for surface flow, the Richards equation for unsaturated zone flow and the Boussinesq equation for groundwater flow. These cannot be solved analytically for cases of practical interest and solutions must instead be obtained using approximate numerical methods, an approach which has become feasible only with the introduction of powerful computers.

Physically-based distributed models treating single components of the hydrological cycle have been developed and applied extensively over the last two decades. Almost all groundwater models, for instance, conform to this type,. However, physically-based distributed catchment models, integrating submodels of the major components of the hydrological cycle within one model, have progressed less rapidly. This is largely because of the heavy computer and data requirements of such models, although there are also numerical difficulties, such as mass balance errors, to be overcome in modeling the transfer of data between the separate submodels. Nevertheless, several physically-based distributed models have been successfully developed and tested during the past decade, although not applied operationally on a routine basis for practical projects. Prominent among these is the SHE modeling system (storm, 1989).

The above types of hydrological models can also be broadly classified in two groups: (I) Event based streamflow simulation models, and (ii) continuous stream simulation models.

In the event based streamflow simulation models, direct runoff hydrograph or its peak characteristics are modeled. Streamflow simulation for individual storm events is required for various hydraulic structures, urban and highway drainage, planning of flood control works, urban planning and development etc. A number of event based stream simulation models have been developed during last two decades. Some of the event based stream flow simulation model have been briefly described by Brown et al. (1974). The various aspects of event based stream flow simulation models including building an event based streamflow simulation model and some of the commonly used event based models have been discussed by Singh (1989).

Continuous streamflow simulation models simulate streamflow for long periods of time. This model maintain continuous accounting or water in storage and in the watershed. In these models the emphasis is a simulation of the entire aldn phase of the hydrologic cycle. In these models the hydrologic processes such as evaporation, transpiration, depression storage, infiltration, interception, sub-surface flow and base flow etc. are also taken into consideration. Some of these processes are not considered in the event based models, some are lumped and considered with approximate Larson et al. (1983) have discussed assembling these components into a continuous stream flow simulation model. Singh (1989) also briefly mentioned arrangements of the components of some of these models.

Description of (parts of) the hydrological cycle. However, the term hydrological model is often understood to be and is used more narrowly as a computer based mathematical model. The development and application of such models have increased tremendously during the two or three decades, so that engineering hydrology today usually involves consideration of some kind of hydrological model. With the current rapid development within computer technology and hydrology the application of computer based hydrological models can only continue to increase in the near future.

5.0 Selection of Appropriate Model Type

A large number of hydrological models exists. However, many of the models function in fundamentally the same way. For instance, at least 20 different rainfall-runoff models of the lumped conceptual type (like the NAM model) exist. Although these models at first sight may look very different they have fundamentally the same structure and basically function according to the same principles. Thus, differences in performance among the better half of the lumped, conceptual rainfall-runoff models are believed to be mostly dependent on the hydrologist who calibrates and operates the model, while the models themselves (apart from ease of operation, user friendliness, etc.) are basically of the same quality.

Thus the question "which model is most appropriate for my particular hydrological problem?" cannot be answered strictly by giving the name of one model. Recommendations are instead given as to which of the above-mentioned model types are most appropriate for the different kinds of hydrological problems.

For some hydrological problems the selection of model type is more or less obvious, e.g. probabilistic models for frequency analysis or stochastic time series models for generation of long (100-1000 years) synthetic streamflow series. Therefore, only the fields of applicability of the different deterministic simulation models are discussed (Storm, 1989) in the following paras:

Empirical (black box) models are mainly of interest as single event models or as subcomponents or more complicated models.

Lumped, conceptual models are especially well suited to simulation of the rainfall-runoff process when hydrological time series sufficiently long for a model calibration exist. Thus typical fields of application are:

Extension of short streamflow records based on long rainfall records.
Real-time rainfall-runoff simulation for example flood forecasting.

Other fields of possible application, to which the lumped conceptual models are not especially well suited, but where they can be used if no better model or method is available, are the following.

Prediction of runoff from ungauged catchments, i.e.i.e. catchments where calibration is not possible. In such cases the model parameters are typically estimated by calibrating against hydrologically similar, neighbouring catchments.

General water balance studies, availability of groundwater resources, rrigation needs, analyses of variations in water availability due to climatic variability, etc.

6.0 Role of Physically Based Distributed Models

Physically-based distributed models can in principle be applied to almost any kind of hydrological problem. Obviously, there are many problems for which the necessary solutions can be obtained using cheaper and less sophisticated empirical, lumped conceptual or statistical models. However, for the more complicated problems there may be little alternative, but to use a physically-based distributed model. Some examples of typical fields of application (Storm, 1989) are

6.1 Catchment Changes

These include both natural and man-made changes in land-use, such as the effects of forest fires, urbanization and forest clearance for agricultural purposes. The parameters of a physically-based, distributed model have a direct physical interpretation which means that they can be evaluated for the new state of the catchment before the change actually occurs. This enables the effects of changes to be examined in advance of such changes. In addition, the characteristically localized nature of catchment changes can easily be accounted for within the spatially distributed model structure.

6.2 Ungauged Catchments

An application in a previously ungauged catchment requires the initiation of a programme of field work to provide data and parameters for calibration. Here, the physical significance of its model parameters enables e.g. the SHE to be applied on the basis of a much shorter, and therefore more cheaply obtained, hydrometeorological records than is necessary for more conventional models. Similarly the catchment parameters can be estimated from intensive short-term field investigations.

6.3 Spatial Variability

Spatial variability in catchments inputs and outputs. Distributed models can be used to examine the effects on flood flow of different directions of storm propagation across a catchment and also the effects of localized river and groundwater abstractions and recharge. The facility is beyond the capability of lumped catchment models which can deal only with quantities averaged across the catchment.

6.4 Movement of Pollutants and Sediments

In order to model the movement of pollutants and sediments, it is first necessary to model the water flows which provide the basic dispersion mechanism. Most water quality and sediment problems are distributed in nature, so distributed models are the most suitable for supplying the basic information on water flows.

7.0 Calibration and Validation of Hydrological Models

Hydrological models are the mathematical models having some unknown co-efficient known as parameters. Model calibration means the estimation of those parameters from historical input-output records. Model validation means judging the performance of the calibrated mode over that portion of historical records which have not been used for the calibration.

For model calibration the methods, which have been commonly used, include (i) manual parameter assessment using 'Trial and Error' procedure, (ii) automatic parameter assessment using numerical optimization procedure and (iii) a combination of (I) and (ii). For the model validation, various validation criteria, developed based on the observed and computed output records, are used.

In this lecture the following aspects of the hydrological modeling have been discussed in brief:

(i) Hydrological processes considered in stream flow simulation models
(ii) Hydrological Modelling procedures
(iii) Goodness of fit and accuracy criteria
(iv) Model Calibration and validation methods
(v) Model Calibration and validation methods
(vi) Model validation including the schemes for Systematic Validation of simulation models
(vii) Sensitivity analysis; and
(viii) Extrapolation from calibration conditions etc.

7.1 Hydrological Modelling Procedures

The following procedures are usually followed for Hydrological Modeling:

- Develop a suitable model structure to simulate various component processes keeping in mind the quantity and quality of the data available and nature of the problems for which the modeling is required.
- Calibrate the developed model using the historical records.
- Validate the model using the historical records which have not been considered for calibration.
- Perform sensitivity analysis study to identity the most sensitive parameters of the model which require proper investigation before arriving at the final parameter values.
- Use the calibrated and validated model for solving the specific hydrological problem for which the development of the model is intended for.

7.2 Concept of Deterministic Mathematical Modelling And Sources of Uncertainty

Basically four sources of uncertainty occur in deterministic simulation, the disagreements between recorded and simulated output resulting from:

1. Random or systematic errors in the input data, e.g. precipitation temperature, or evapotranspiration used to represent the input conditions in time and space for the catchment.
2. Random or systematic errors in the recorded output data, e.g. water level or discharge data used for comparison with the simulation output.
3. Errors due to non-optional parameter values.
4. Errors due to incomplete or biased model structure.

Thus, during the calibration process only error source 3 is minimized, whereas the disagreement between simulated and recorded output is due to all four error sources. The measurement errors and errors source 2 serve as a 'background noise' and give a minimum level of disagreement below which further parameter or model adjustments will not improve the results. The objective of a calibration process is then to reduce the error source 3 until it is insignificant compared with the data error sources 1 and 2.

During a calibration process it is of the utmost importance to ensure that a clear distinction is drawn between the different error sources, so that it not attempted to compensate for errors for one source by adjustment within another source, e.g. compensate for a data error by parameter adjustments. Otherwise the calibration will degenerate to curve fitting, which may result in a reasonable fit within the calibration period but will inevitably give poor simulation results for other periods. In the following five examples it would be physically incorrect and fatal for future predictions to try to compensate for the following discrepancies between recorded and simulated flows using parameter adjustments:

- Both flood peak and runoff volume for a hydrograph are under predicted, owning to an underestimation of the average precipitation, Error source -1
- Discrepancies are observed between simulated and recorded flow in a period where the recorded flow is known to be very uncertain owing to problems with the rating curve. Error source 2.
- A flood peak is under predicted as a result of embankments being breached whereas the model has been developed assuming non-breaching embankments. Error source 4.
- Travel time for high flows is smaller than the travel time for low flows but the routing model is linear with the travel time independent of flow regime. Error source 4.
- The base flow in low flow periods decreased during the calibration period owing to ground water abstraction and lowering of the ground water tables but ground water abstraction cannot be accounted for directly in the applied model. Error source 4.

7.3 Model Calibration

Model calibration in general involves manipulation of a specific model to reproduce the response of the catchment under study within some range of accuracy. In a calibration procedure an estimation is made of the parameters, which cannot be assessed directly firm field data. All empirical (black box) models and all lumped, conceptual models contain parameters whose values have to be estimated through calibration. The fully distributed physically-based models contain only parameters which can be assessed from field data, so that in theory a calibration should not be necessity if sufficient data are available. However, for all practical purposes the distributed, physically-based models also require some kind of calibration, although the allowed parameter variations are restricted to relatively narrow intervals compared with those for the empirical parameters in empirical or lumped, conceptual models.

7.3.1 Calibration Methods

In principle three different calibration method can be applied:

a. 'Trial and Error', manual parameter assessment
b. Automatic, numerical parameter optimization
c. A combination of (a) and (b)

The trial and error method implies a manual parameter assessment through a number of simulation runs. This method is by far the most widely used and is the most recommended methods, especially for the more complicated models. A good graphical representation of the simulation results is a prerequisite for the trial and error method. An experienced hydrologist can usually achieve a calibration using visual hydrograph inspection within 5-15 simulation runs.

Automatic parameter optimization involves a numerical allegorical algorithm which optimizes or minimize a given numerical criterion. The objective of automatic parameter optimization is to search through the many combinations and permutation of parameter levels to achieve the set which is the optimum or 'best' in terms of satisfying the criterion of accuracy. Several optimization techniques have been used for calibration of hydrological models. A decade ago the most popular was Rosenbrocks's method)Rosenbrock, 1960).

The advantages of automatic parameter optimization over the trial and error method are:

- Automatic optimization is quick, because almost all work is carried out by computer.
- Automatic optimization is less subjective than the trial and error method, which to a large degree depends on visual hydrograph inspection and the personal judgment of the hydrologist.

Disadvantages of automatic parameter optimization include:-

- The criterion to be optimized has to be a single numerical eriterion based on a single variable; as discussed in earlier section, though, selection of an appropriate criterion under these constraints is a complicated task
- If the model contains more than a very few parameters the optimization will probably result in a local optimum instead of the global one.
- The theories behind the search algorithms assume that the model parameters are mutually independent. This assumption is usually not satisfied in practice.
- An automatic routine cannot distinguish between the different error sources mentioned earlier. Therefore, an automatic optimization algorithm will try to compensate, e.g. for data errors by parameter adjustments, with the results that the parameters values often become physically unrealistic and give poor simulation results when applied to a period different from the calibration period.

Combination of the trial and error and automatic parameter optimization method could involve, for example, initial adjustment of parameter values by trial and error to delineate rough orders of magnitude, followed by fine adjustment using automatic optimization within the delineated range of physically realistic values. The reverse procedure is also possible, first carrying out sensitivity tests by automatic optimization to identify the important parameters and then calibrating them by trial and error. The combined method can be very useful but does not yet appear to have been widely used in practice.

Finally, given the large number of parameters in a physically based distributed model like the SHE, it is not realistic to obtain an accurate calibration by gradually varying all the parameters single or in combination. A more sensible approach is to attempt a coarser simulation suing only the few parameters to which the simulation from sensitivity analysis. However, experience suggests that the soil parameters will usually require the most attention because of their role in determining the amount of precipitation which infiltrates and hence the amount which forms overland flow.

The above methods of calibration consider single objective function. In case multi objective function is required to be considered, then two types of approaches, viz. Classical approach and pareto approach may be utilised. In classical approach a combined objective function is desired assigning the weights to the various objective function depending upon the user requirement. In pareto approach a set of parameter values are determined using search algorithm in such a way that the global optima is achieved considering the multi objective function.

7.4 Model Validation

If the model contains a large number of parameters it is nearly always possible to produce a combination of parameter values which permits a good agreement between measured and simulated output data for a sort calibration period. However, this does not guarantee an adequate model structure or optimal parameter values. The calibration may have been achieved purely by numerical curve fitting without considering whether the parameter values so obtained are physically reasonable. Further, it might be possible to achieve multiple calibrations or apparently equally satisfactory calibrations based on different combinations of parameter values. In order to find out whether a calibration is satisfactory, or which of several calibrations is the most correct, the calibration should therefore be tested (validated) against data different from those used for the calibration (e.g. Stephenson and Freeze, 1974). Kelmes (1986) states that a simulation model should be tested to show how well it can perform the kind of task for which it is intended. Performance characteristics derived from the calibration data set are insufficient as evidence of satisfactory model operation. Thus the validation data must not be the same as those used for calibration but must represent a situation similar to that to shich the model is to be applied operationally.

Klemes (1986) further noted that a central question is: what are the grounds for credibility of a given hydrological simulation model? Usually they concern the goodness of fit of the model output to the historical record in a calibration period, combined with an assumption that the conditions under which the model will be used will be similar to those of the calibration period. Clearly, though, this is insufficient for a physically-based distributed model which is designed specially to simulate conditions different from those likely to be available for calibration, e.g. when simulating the impact of a future land-use change. In that case a demonstration of model transposability is required. Initially transposability referred to geographical transposability within one hydrologically homogeneous region. However, its scope has since been broaden to include transposability from one land use type to another, from one region to another and, recently, from one climate to another.

7.4.1 Schemes for Systematic Validation of Simulation Models

The hierarchical scheme proposed by Klemes (1986), should be referred to here. The scheme is briefly discussed below:

The scheme is called hierarchical because the modelling tasks are ordered according to their increasing complexity, and the demand of the test increase in the same direction. Two major categories are proposed for the process to be simulated, in particular:

1. Stationary conditions, and
2. Non stationary conditions

Each of them being sub-divided into two hierarchical sub-groups according to whether the simulation is to be done for:

a. The same station (basin) which was used for calibration or
b. A different station (basin)

Here, the term stationary is used to denote physical conditions that do not appreciably change with time:

Typical examples of modeling tasks in these four classes of increasing difficulty are as follow:

1a. Filling in a missing segment of, or extending a stream flow record
1b. Simulation of a stream flow record in an ungauged basin
2a. Simulation of streamflow record in a gauged basin for conditions after a change in land use, climate or both
2b. Simulation of a streamflow record in an ungauged basin for conditions after a change in land use, climate or both

The following tests are recommended as a minimum standard for operational testing of models for the above four levels of difficulty of the simulation task:

1a. Split sample test
1b. Proxy basin test
2a. Differential split sample test
2b. Proxy basin differential split sample test.

(1a) **Split-sample test:** The available record should be split into two segments one of which should be used for calibration and the other for validation. If the available record is sufficiently long so that one half of it may suffice for adequate calibration, it should be split into two equal parts, each of them should be used in turn for calibration and validation, and result from both arrangement compared. The model should be judged acceptable only if the two results are similar and the errors in both validation runs acceptable. If the available record is not long enough for a 50/50 splitting, it should be split in such a way that the calibration segment is long enough for a meaningful calibration, the remainder serving for validation. In such a case, the splitting should be done in two different ways, e.g. (a) the first 70% of the record for calibration and the last 30% for validation; (b) the last 70% for calibration and the first 30% for validation. The model should qualify only if validation results from both cases are acceptable and similar. If the available record cannot be meaningfully split, then only a model which has passed a higher level test should be used.

(1b) **Proxy-basin test:** This test should be required as a basic test for geographical transposability of a model, i.e. transposability within a region. If stremflow in an ungauged basin C is to be simulated, two gauged basins A and B within the region should be selected. The model should be calibrated on basin A and validated on basin b and vice versa. Only if the two validation results are acceptable and similar can the model command a basic level of credibility with regard to its ability to simulate the streamflow in basin C adequately.

This kind of test should also be required when an available streamflow record in basin C is to be extended and is not adequate for a split-sample test as described above. In other words, the inadequate record in basin C would not be used for model development and the extension would be treated as simulation in an ungauged basin (the record in C would be used only for additional validation, i.e. for comparison with a record simulated on the basis of calibrations in A and B).

Consider geographical transposability between regions I and II. If steamflow needs to be simulated in an as yet unspecified ungauged basin C (or on a number of such basins) in region II the procedure should be as follows. First, the model is calibrated on the historic record of a gauged basin D in region I. Streamflow measurements are started on at least two different subsititure basins, A and B, in region II and maintained for at least three years. Then the model is validated on these three-year records of both A and B and judged adequate for simulation in a basin C if errors in both validation runs, A and B, are acceptable and not significantly different. After longer records in A and B become available, these two basins can be used for model development and subjected to the simpler test for transposability within a region as described above, using A and B as proxy basins for C. Of course, the substitute basins a and B, would not be chosen randomly but would be selected so as to be representative of the conditions in region II, and, as far as possible, with due consideration of future stream gauging needs.

(2a) **Differential split-sample test:** This test should be required whenever a model is to be used to simulate flows in a given gauged basin under conditions different from those corresponding to the available flow record. The test may have several variants depending on the specific nature of the change for which the flow is to be simulated.

For a simulation of the effect of a change in climate, the test should have the following from. Two periods with different values of the climate parameters of interest should be identified in the historic record, e.g. one with high average precipitation, the other with low. If the model is intended to simulate streamflow for a wet climate scenario then it should be calibrated on a dry segment of the historic record and validated on a wet segment. 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.

If segments with significantly different climatic parameters cannot be identified in the given record, the model should be tested in a substitute basin in which the differential split-sample test can be done. This will always be the case when the effect of a change in land use, rather than in climate, is to be simulated. The requirement should be as follows: to find a gauged basin where a similar land-use change has taken place during the period covered by the historic record, to calibrate the model on a segment corresponding to the original land use and validate it on the segment corresponding to the changed land use.

Where the use of substitute basins is required for the testing, two substitute basins should be used, the model fitted to both and the results for the two validation runs compared. Only if the results are similar can the model be judged adequate. Note that in this case (two substitute basins) the differential split-sample test is done on each basin independently which is different from the proxy-basin test where a model is calibrated on one basin and validated on the other.

A differential split-sample test can arise by default from a simple split-sample test if the only meaningful way of splitting an available record is such that the two segments exhibit markedly different conditions.

(2b) **Proxy-basin differential split-sample test:** This test should be applied in cases where the model is supposed to be both geographically and climatically (or land-use-wise) transposable. Such universal transposability is the ultimate goal of hydrological modelling, a goal which may not be attained in decades to come. However, models with this capability are in high demand and hydrologists are being encouraged to develop them despite the fact that thus far even the much easier problem of simple geographical transposability within a region has not been satisfactorily solved.

The test to demonstrate such a general transposability may have different forms depending on the specific modelling task involved. In the simplest case of geographical and climatic transposability within a region (e.g. for a model intended for assessment of impact of climatic change in an ungauged basin C), the test should have the following form. Two gauged basins, A and B, with characteristics similar to those of basin C are selected and segments with different climatic parameters, e.g. w for wet and d for dry, are identified in the historic records of both of them. Then, for an assessment of the impact of a dry climate scenario, the model is first calibrated on Aw and validated on Bd, and then calibrated on Bw and validated on Ad. It is judged adequate if errors in both validation runs Ad and Bd are acceptable and not significantly different. By analogy, a model intended for an assessment of the impact of a wet climate scenario would have to be calibrated/validated on Ad/Bd, and on Bd/Aw, and judeged adequate if results from Bw and Aw are adequate and similar.

7.5 Sensitivity Analyses

Analysis of the sensitivity of the simulation results to changes in parameter values and analysis of parameter stability can served as model tests. Such analyses can be carried out in different ways. The influence of the length of the calibration period on parameter uncertainty as well as parameter stability with time can also be evaluated from such analysis.

7.6 Extrapolation from Calibration Conditions

If the calibration is based on a narrow range of data, the model, even of physically-based, may not be applicable outside this range. For example, if the data based contains only small floods, the model, even if properly validated in the operational sense, cannot be trusted to simulate very large floods adequately. The calibration/validation exercise should therefore be based on as wide range of conditions as possible. This approach can also be useful in eliminating incorrect calibrations in cases where it has been possible to achieve multiple calibrations based on different combinations of realistic parameter values. The incorrect calibrations are less likely to support acceptable simulations based on data outside the range used for calibration.

8.0 Remarks

In order to optimally develop and utilise India's water resources to meet the demands for various uses for our growing population, the application of hydrological modeling techniques would be very much necessary. This will be required not only for deciding about water yield or design parameters, but also for understanding and evaluating effects of developmental and other activities on hydrological regime of river basins.

Finally, the choice of model for a particular " real-world" application is likely to be heavily influenced by non-hydrological criteria such as the time, manpower and money available to support the project, availability of data, desired accuracy of results and computer resources. Selecting a model requires balancing the degree to which the model represents the hydrological system against the general difficulty in obtaining a result. If a highly complex mathematical representation of a system is used, the risk of not representing the system is minimized but the difficulty of obtaining a useful result is maximized. Many data will be required, programming effort and computer time are large, the general mathematical complexity may even render the problems formulation intractable and the resource constraints of time, money and manpower may be exceeded. Conversely, if a greatly simplified mathematical model is applied without a proper examination of its physical significance, the difficulties in obtaining a result may be reduced but the risk of not representing the physical system is increased.

The model calibration and validation are the important aspects of the hydrological modelling proper calibration and validation of the hydrological model is necessary before using the model for simulation. For the validation of the modes, the hierarchical scheme discussed in the lecture may be adopted. In order to ascertain the uncertainity in the parameters as well as parameter stability the sensitivity analysis must be carried out.

Reference:

1. Abbott, M.B., J.C. Bathurst, J.A. Cunge, P.E. O'Connell and J. Rasmussen, 1986. An Introduction to the European Hydrological System - System Hydrologique European "SHE". History and Philosophy based distributed modelling system. Jour. of Hydrology, 87:pp. 45-59.

2. Aitken, A.P., 1973. Assessing Systematic Errors in Rainfall Runoff Model. Journal of Hydrology, 20 121-136.

3. Bathurst, J.C., 1986. Physically-based distributed modelling of an upland catchment using the Systeme Hydrologique European. J Hydrol., 87(1/2), 79-102.

4. Bergston, S., 1976. Development and application of a conceptual runoff model for Scandinavian Catchments. Department of Water Resources Engineering, University of Lund. Bulletin Series A No. 52.

5. Betson. R.P., R.L. Tucker and F.M.Haller, 1969. Using Analytical Method to Develop a Surface Runoff Model. Water Resources Research, Vol. R, No. 1, pp. 103-111.

6. Carrigan, P.H., Jr., 1973. Calibration of U.S. Geological Survey Rainfall-Runoff Model for Peak Flow Synthesis-Natural Basins. Report No. U.S.G.S.-W.R.D.-73-026, U.S.Geological Survey-WRD, Reston, Va, U.S.A.

7. Crawford, N.H. and Linsley, 1966. Digital Simulation in Hydrology. The Stanford Watershed Simulation Model 'IV'. Technical Report No. 39, Department of Civil Engineering, Stanford Univ. Stanford, California.

8. DHI, 1988. Lecture notes and Exercises - Training course on Hydrological Computerized Modelling System (SHE).

9. Fleming, G., 1975. Computer Simulation Techniques in Hydrology Elsevier.

10. James, L.D., 1972. Hydrologic Modelling, Parameter Estimation and Watershed Characteristics. Journal of Hydrology, Vol. 17, pp. 283-307.

11. Klemes, V., 1986. Operational Testing of Hydrological Simulation Models. Hydrol. Sci. J., 31(1), 13-24.

12. Laurenson, E.M. and R.G. Mein, 1983. RORB-Version 3 : Runoff Routing Program User Manual. 2nd ed., Deptt. Of Civil Engg., Monash University, Australia.

13. Ministry of Water Resources, 1987. National Water Policy.

14. NIH, 1982-83. Modelling of Daily Runoff for Kasurnala Basin Using Berson and USGS Models. Report No. CS-2, National Institute of Hydrology, Roorkee.

15. Rosenbrock, K.H., 1960. An Automatic Method for finding the greatest or least value of a function. The Computer Journal, Vol. 7(3).

16. Storm, B., 1989. Introduction to Hydrological Modelling. Workshop on 'Application of SHE Model to Sub-basins of River Narmada', Bhopal, pp 4-15.

17. Singh, V. P., 1988. Hydrologic Systems-Rainfall Runoff Modelling. Vol. I, Prentice Hall, New Jersey.

18. Singh, V.P., 1989. Hydrologic Systems - Watershed Modeling. Vol. II, Prentice Hall, New Jersey.

19. Woolhiser, D.A., 1973. Hydrologic and Watershed Modelling - State of the Art. Transactions of the ASAE, 16, pp. 533-559.