Principles and Confidence in Hydrological Modelling

Sten Bergström

Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

General principles in development and application of hydrological models are discussed and related to the confidence in the results. The presentation is mainly based on the experience from the work with the HBV and PULSE models at the Swedish Meteorological and Hydrological Institute between 1971 and 1990 but has also been influenced by other modelling work. It covers a discussion on the optimal complexity of models, use of observations, calibration, control and sensitivity analysis. Special attention is given to the uncertainties encountered when using hydrological models for the simulation of extreme floods and long-term scenario simulations. Finally a few ethical problems in modelling are mentioned.

Introduction

During the last few decades mathematical models have become generally accepted tools in hydrology. From the straightforward and rather simple modelling of runoff from the late sixties and early seventies the number of applications has grown and models are now used to solve a large span of problems in hydrology and water resources. Models are also used to extrapolate our knowledge of the hydrological system into the extreme, and attempts are made to model long-term effects of man-made changes of the environment. This means that the simulations are frequently made into the unknown, where no reliable data exist for verification of the performance. Along with this development the models are gradually becoming more and more complex and are often used to integrate scientific findings from different
disciplines. The necessary interdisciplinary cooperation in model development has created a situation where no one really has the overview of all components of the model and where the user of the results has to rely upon judgements from the modellers as concerns confidence in the results. Consequently it is of outmost importance to maintain a critical and constructive attitude to all phases of the modelling process. Inappropriate model structures or combinations of substructures, incomplete model calibration and control and, not the least, poor modelling ethics can easily ruin the confidence in models as productive tools. Some recent papers in scientific journals reflect a growing concern about this risk (see, for example, Linsley 1986; Klemeš 1986, 1988; Beven, 1989; and Hauhs 1990), and an increasing number of modelling exercises carried out by relatively inexperienced modellers, with an academic degree as their primary objective, calls for a continuing discussion.

It is the objective of this paper to discuss some general principles which, in my opinion, should guide any modeller of hydrological and water resources systems. The presentation is based on experience from the development of the HBV and PULSE models which started at the Swedish Meteorological and Hydrological Institute in 1971 and is still in progress, but has also been influenced by other modelling work. The paper represents my personal views. It is my hope that it will encourage a continuing discussion on general principles in modelling. Why not in this journal?

**Choice of Model Structure**

**Level of Complexity**

The initial choice of complexity level of the model structure is a very important decision in the modelling process. It will affect the need of computer capacity and the data requirement and thus the applicability of the model. Optimal complexity is the point we should try to reach for each specific problem. It is not difficult to find examples of too complex models which have not contributed significantly to the solution of the problem. WMO has organized a few very well controlled modelling intercomparison projects related to operational runoff modelling (WMO 1975 and 1986) which are highly relevant for the discussion on model complexity. The last of the two intercomparisons concerned snowmelt modelling and covered a span from simple degree-day methods to complete energy balance models. One of the conclusions is worth quoting: »On the basis of available information, it was not possible to rank the tested models or classes of models in order of performance. The complexity of the structure of the models could not be related to the quality of the simulation results.«

Another illustrative example was given by Brandt (1990a) when applying a relatively complex nitrogen turnover model, developed for small homogeneous re-
search basins, to a larger basin of mixed land use. It was shown that equal model performance could be obtained by a simple runoff model and standard concentrations related to different land use. Thus the data demand could be kept low and the risk for overparameterization could be avoided.

The situation was the opposite when a model was sought for use in the new guidelines for spillway design in Sweden (Swedish Committee on Spillway Design 1990). It was found that increasing modelling complexity was needed when going from the simple single reservoir basin to total river systems. Otherwise the guidelines would have become inconsistent with costly rehabilitation of the wrong dams as a result.

Going from complex to simpler model structures requires an open mind, because it is frustrating to have to abandon seemingly elegant concepts and theories. It is normally much more stimulating, from an academic point of view, to show significant improvement of the model performance by increasing complexity. Nash and Sutcliffe (1970) presented a strategy for model development from the simple to the complex which may help to avoid this frustration. The idea is that only modifications that significantly improve the results are accepted. Biswas (1976) argued that the simplest model possible for the system should be sought and that, generalized, all purpose models should be avoided. Similar thoughts had earlier been presented by von Bertalanffy (1968) when stating that »oversimplifications progressively corrected in subsequent development are the most potent and indeed the only means towards conceptual mastery of nature«. Many more examples along the same lines can easily be found.

We have found that the strategy of going from the simple to the complex is still valid and very useful, in particular for the development of operational basin-oriented models. These thoughts have guided our work on the HBV and PULSE models (Bergström 1976) and are responsible for the relative simplicity of their structures. The point of diminishing returns (no model improvement) was reached surprisingly soon when increasing model complexities.

Physically Based or Conceptual

There should not be a conflict between the use of physically based process-oriented models and the simpler, more empirical conceptual ones. We just have to realize that they are developed for different purposes and should be applied accordingly. Physically based models are normally more feasible as research tools for process studies in the small scale where physical parameters are well under control, and their variability is small. So is, for example, the Swedish SOIL model used for dynamic simulations of moisture, temperatures or turnover of nutrients in well controlled soil profiles (see, for example, Jansson and Halldin 1979, and Johansson et al. 1987). Physically based hydraulic modelling can also be found feasible for the computation of flood-waves and inundated land, provided proper mapping of levels and hydraulic properties are available (see, for example, Fread 1984).
Conceptual models are more basin-oriented than physically based models. The basic assumption is that we have accepted that the great areal and vertical variability of physical processes is little known, a fact which justifies a more crude, almost statistical approach. The parameters of a conceptual model thus represent an average over a large area and often integrate several processes and their variability. This means, for example, that the soil moisture conditions computed by a conceptual model is a model-specific variable and should be considered more as an index than as the true value. The physical interpretation of the parameters of the model is consequently normally very vague and should be regarded with a sound scepticism.

Conceptual models show their strength in limited data demand, and thus great applicability in operational hydrology. They are also generally quite easily understood by a non-expert. Examples of applications are extension of short runoff records, flood forecasting and reservoir operation, design floods and estimation of hydrochemical transport.

Physically based models are often said to be superior to conceptual models as they demand less calibration or tuning of the parameters (see, for example, Refsgaard et al. 1989). This may be true to some extent, but there is a limit in performance that can not be passed even by the most complete physical description. This limit is set by the accuracy and representativeness of input data. We know, for example, that the most important driving variables, precipitation and evapotranspiration, are subject to large systematic errors which we cannot correct satisfactorily before using them as input to the model. We also know that geology, topography and land use are highly variable in most basins and even varies a lot within the square grid used as areal unit in many distributed models (Beven 1989). There is consequently a tendency to use the calibration option even in very complex model formulations like, for example, the SHE model (Bathurst 1986). As the physically based model is gradually becoming more and more conceptual the more the calibration option is accepted, the statement that a complex, physically based model is more feasible for studies of effects of land use, scenario simulations, or where input data are lacking, can certainly be challenged.

The Cult of Success

Many modellers have experienced the frustration of reaching the point of diminishing returns after long periods of hard work. They simply find themselves with results which they feel are not worth publishing. Normally they are wrong! Although reports on lack of success are rare in the scientific literature we all know that a negative advice can be invaluable. An honest statement like: »It is no use trying, I did and failed«, can help others to avoid expensive disappointments and to proceed along a better path.
Lack of success in modelling is probably more common than is success, but this fact is not reflected in the conclusions of published papers. If published at all, the real truth may be hidden behind words like »encouraging indications«, »deserves further study« or similar. Honest presentations of scientific disappointments are important contributions which make the journals more interesting to the reader. A change in attitude among scientists and publishers is requested.

Success or lack of success is often judged differently by the modeller and the user of the results. The user is interested in the applicability of the model and costs of operation. It may therefore be a relief for him that the suggested modification of the model did not improve the performance. The modified model may have entailed costly changes of an existing operational modelling system and more detailed data collection.

**Use observations!**

**Confidence in the Model**

Even if a model is based on the best available knowledge, proper control of the results against observations is still the only way of approaching confidence in the formulation of the processes. On the other hand good model performance does not guarantee that we have described the system correctly. The model must »work well for the right reason« (Klemes 1986). As many modellers have experienced, almost identical results can often be obtained with a number of model formulations.

There is no ultimate solution to the problem of model verification. There is no point where we can say that the model is verified. It is rather a process of repeated tests of hypotheses and growing confidence. In my opinion the use of control data can never be replaced in this process.

Growing confidence are thus key words in hydrological modelling. This can be obtained by applying the model under a span of different geographical, climatological and geological conditions. A complex process-oriented model should also be controlled against a variety of internal data representing more than one single process. This procedure has, for example, been followed during various stages of the development of the BIRKENES model which is used for studies of the effects of acid rain on water resources. The use of internal data forced the modellers to reevaluate parts of the model structure (Lundquist et al. 1990).

**The HBV/PULSE – Experience**

When developing the conceptual HBV and PULSE models we have tried to use more and more internal data for control and improvement of the different components of the models. In the early stage only runoff was simulated (Bergström 1976) and this was the only type of application for many years. Later on simulations of groundwater dynamics led to a revision of the response function of the model.
(Bergström and Sandberg 1983). The soil moisture dynamics of the model were controlled by the work of Andersson (1988) which threw light on problems with the seasonality of the evapotranspiration routine. Examples of simulations of short-term dynamics of hydrochemical variables can be found in the work by Bergström et al. (1985) (pH and alkalinity) and Brandt (1990a) (nitrogen). The attempts to simulate nitrogen transports in mixed basins taught us that we had made the model too complex and that we had to return to simplicity again.

Modelling of pathways and transit time distributions were performed by a modified PULSE model and controlled by the use of data on the stable isotope $^{18}O$ in precipitation and runoff (Lindström and Rodhe 1987). This study revealed that the PULSE model, like any model from the conceptual runoff family, grossly underestimates transit times unless it is modified to account for large volumes of additional subsurface tension water.

The limitations of our modelling approach were finally identified when trying to model climate-induced variations in groundwater chemistry (Bergström et al. 1990). It simply was not possible to describe short-term hydrochemical variations in groundwater by a conceptual model of this simplicity except for in a few special cases.

Although a multitude of data of different origin has been used in the long process of development of the HBV and PULSE models, we can not honestly say that their structures are finally tested and verified. But the large number of applications have gradually built up our confidence in the use of these models to a degree where we can continue our operational applications and accept the models as the foundation for further model development. We know that some basic assumptions concerning snow accumulation and melt, soil moisture accounting, recharge and groundwater dynamics and transit times are reasonable and agree relatively well with field observations. We also identified some important limitations of our approach.

**Data Availability**

Data availability is a problem in many modelling studies. The best situation is a project where data collection and modelling can go on in parallel, so that the field work can be adjusted to the need of the modellers. This lucky situation is not too common. Field work is expensive and it takes time to collect sufficient data for the models. But there are other ways. In my opinion there is a tendency among modellers to restrict themselves to data that they have collected themselves or are familiar with for other reasons. Normally there is an abundance of data available. These may have been generated by other research projects or can be found in the data bases of hydrological services or other authorities. The use of these data requires that we change our attitude as concerns ownership of data. We have to realize, and take advantage of, the fact that some scientists show their skill in the field or in laboratories and others in front of a computer. Together they can form successful teams.
Principles and Confidence in Hydrological Modelling

Calibration, Control and Presentation of Results

Calibration is the process whereby some coefficients of the model are adjusted to make the model output match the observations. This can be made either automatically or by a manual process based on visual inspection of the results. Normally some statistical criterion of the agreement between simulations and observations is used to describe the performance, for example the explained variance, $R^2$, according to Nash and Sutcliffe (1970). This criterion, like any other criterion, is a crude estimate of model performance which is site specific, specific for the time period modelled and depending on the quality of the observed data. It is therefore of limited value for a reader to restrict the presentation to the criterion of agreement. As the human eye can identify much more, the best way of presenting results is simply to show time series of the model output and the observations in the same graph and thus leave the judgement of the agreement to the reader. This standpoint was also taken by the participants in the WMO intercomparison of snowmelt models (WMO 1986).

There are also problems when visualizing results in graphical form. If, for example, simulated and observed transports of substances are compared, the domination of river flow in the two graphs can give a false impression of ability of the hydrochemical model to describe concentrations. Other well known ways of hiding uncertainties are by the presentation of accumulative plots or logarithms.

We all know that we can make many mathematical expressions match a time series of observations provided there are enough degrees of freedom. It is therefore wise to save part of the database for an independent test of the model and to concentrate on this period when discussing the results. In any case it shall be clearly stated in any presentation of results whether they refer to the calibration or to the independent test periods. Otherwise there is a risk that we are trying to explain results that do not come from the model at all.

If the model performance is significantly lower for the independent period used for validation than it was for the calibration period, the modeller should seriously consider if there are problems of overparameterization. The model may simply have too many degrees of freedom for the information contained in the observed records. This leaves the modeller with two options, model simplification or improvement of the control database. There is no general rule as concerns the necessary length of a calibration period. The data shall simply cover a sufficient range of significant events and have a resolution in time that makes it possible to find stable values of the parameters of the model.

Generalization of Model Parameters

What shall we do if we don’t have measurements of runoff and are unable to model the ungauged system with physically based models? One way, which is growing in...
popularity, is the use of generalized model parameters and empirical conceptual
models. The idea is that we calibrate our model to many basins in a region and
develop a standard parameter set for use in ungauged basins. The method has
proved to be useful to provide environmental control programmes with runoff data
from a large number of basins in Sweden (Johansson 1986). It requires, however,
that the user of the data is aware of the fact that this inexpensive, and therefore
attractive, method yields results of a lower quality than a properly calibrated model
and still lower quality than a record based on traditional measurements.

Generalization of the parameters of a conceptual model may be an alternative to
physically based modelling when it comes to studies of the effects of changing land
use in a basin. The prerequisite is that we have well defined homogeneous basins
where we can derive parameter sets for each specific land use. This method was,
for example, used by Brandt et al. (1988) in a study of the effects of partial
clearcutting of a large basin.

It is more difficult to generalize results from combined hydrological and hydro-
chemical models. The reason is that it is difficult to avoid strong interaction be-
tween the water balance components and the chemical ones. A change in the value
of one parameter of the hydrological subroutines may have negligible effects on the
water balance but recalibration of the chemical subroutines may be required (see,
for example, Bergström et al. 1987).

Into the Unknown

There are situations where no data are at hand for control of the models. Examples
of this are the estimations of extreme floods and simulations of long-term environ-
mental effects.

Extreme Floods

Design floods for dams and spillways are very much in focus all over the world.
New guidelines have been adopted or are under way in many countries. The issue
was given special attention during the conference of the International Commission
on Large Dams (ICOLD) in 1988. A key question is the estimation of extreme
precipitation and its transformation into a design flood with very low annual proba-
bility of exceedence (in the order of one in 10,000 or less). A discussion on this
specific problem can be found in the work by Bouvard (1988) or in NRC (1985).

When using a hydrological model for computation of very extreme floods there is
a combination of uncertainties which has to be considered. First of all, the design
precipitation and the hydrological conditions represent an extreme situation which
often is the result of modelling. Secondly the hydrological model will be run with
data that are outside the range of those used for calibration, and the magnitude of
the flow will be much higher than those during the period of calibration. This
requires special attention to the representation of peak flows in the calibration period. Flood-wave simulations to study the consequences of a hypothetical dam break suffer from similar extrapolation problems as those related to design floods (Kung 1989).

**Environmental Effects**

The modelling of long-term effects of acid precipitation and other types of atmospheric deposition requires combined hydrological and hydrochemical models which easily become very complex and require multidisciplinary approaches. Best known are model names like ILWAS (Gherini *et al.* 1985), MAGIC (Cosby *et al.* 1985) and BIRKENES (Christophersen *et al.* 1982). The problem of verification is obvious in this type of environmental simulations as we have very few control experiments to rely upon. The situation is similar for those scientists who are engaged in the modelling of leakage of radioactivity from nuclear waste deposits which are expected to remain for thousands of years.

Modelling the effects on water resources of a global climatological change means that we have to deal with a mix of uncertainties. While climatologists keep warning for too detailed regional interpretations of the results from the general circulation models (GCM), the hydrologists have already started to use the results in their models (see, for example, Gleick 1989, or Lemmelä *et al.* 1990). We have to recognise that reliable forecasts of the effects on water resources require much more detailed information from the climatologists than is available today. So is, for example, the flood risks in many parts of the world more related to the timing of the extreme precipitation and the soil moisture deficit than to a general increase in precipitation (Fig. 1). The timing of air temperature and the seasonality of precipitation are also critical for the retention of nutrients in lakes and reservoirs and thus for the transport of these matters to the coastal waters. While estimates of regional average precipitation changes by GCMs are very uncertain, it is still more difficult to say anything about the seasonality or magnitude and timing of extreme values.

Modelling effects of global change also addresses the question of model stationarity. A changing climate will, of course, gradually effect the vegetation and thus the evapotranspiration in a basin (see, for example, Boer *et al.* 1990). The conventional hydrological models in use today do not consider this process, which means that further model development is needed.

**Attitudes to Uncertainty**

Modelling into the unknown, as exemplified above, requires a much more humble attitude by the modeller than traditional applications of models to observed time series. At the same time the effects of the modelling may be more far-reaching. Spillway design studies are the foundation for expensive investments in our river systems. Environmental model simulations have socio-economic and political implications. Again confidence in the models is the crucial point. It is a minimum
Fig. 1. Maximum daily areal precipitation and model simulations of maximum snow melt, soil moisture deficit and maximum effective precipitation for the 3,446 km² Blankaström basin in river Emån in southeast Sweden over the period 1944-1985. The figure illustrates the importance of timing of the hydrological processes for flood generation (From Brandt 1990b). Note that all scales are identical.
requirement that the models are properly controlled and found reliable when applied to all possible observations that are at hand today and that the known physical limits, which may be reached in the scenario, are accounted for. There is no reason to believe that a model that is unable to describe the dynamics observed in our present records will be of any value in long-term scenario simulations.

**Sensitivity Analysis**

The analysis of model sensitivity is a very important component of a model development process. It helps keeping the model simple because it reveals model parameters with insignificant effect on the results. It is also a tool to identify interaction between model components and parameters. Before carrying out a model simulation into the unknown, as exemplified above, a proper sensitivity analysis will help us to identify the stability of the results in relation to the uncertainty of our assumptions.

The simplest form of sensitivity analysis is by visual comparison of graphs computed by different assumptions. These can be initial conditions, model parameters or systematic adjustments of input data. The method is straightforward and surprisingly powerful in many applications. It is also easy to understand by anyone.

More complex is the mapping of the error function topography on the basis of a criterion of agreement between computations and observations. With this method the interaction between two, or even three, model parameters can be analyzed. The interpretation requires some insight into the model, and an important limitation is the rigidity of the chosen criterion of model performance as discussed above. There are also more theoretical methods of sensitivity analyses at hand, for example statistical analysis of the response surface shape and Monte Carlo simulations of parameter values (e.g. Mein and Brown 1978, Gardner et al. 1980).

No matter what method is being used it is important that the modeller bases the analysis on realistic assumptions of the uncertainties in the model conditions. Properly made, the sensitivity analysis is a very valuable tool for anyone who wants to build up confidence in a model structure.

**Ethics in Modelling**

Is it really relevant to speak about ethics in modelling? The modeller is primarily a scientist and can therefore not be held responsible for the use or abuse of the results. The question is a classical one. In my opinion there is no answer but yes.

The complexity of hydrological models and many other models in geoscience has now reached a point where no one but the modeller has a chance to have the total overview and to judge their applicability, weaknesses and uncertainties. At the
same time society is confronting a large number of difficult environmental and technical problems, the solution of which will be very costly. Reliable model simulations are urgently needed to help us to make the right decision. Models are also used to guide politicians in sensitive negotiations on reduction of emissions into the atmosphere and the whole issue of global warming is to a high degree based upon model simulations. There are also examples where model simulations are used to solve conflicts in court.

Environmental problems and an increasing demand for water in developing countries has opened a new market for hydrological models. It is important that these models are well tested and appropriate for the technical level and data availability of the area. It can not be justified to apply more advanced models in developing countries than we use to solve similar problems at home.

The modellers must consequently take their responsibility and base their work on a scientifically sound and stable foundation. All assumptions, results and limitations must be clearly presented and interpreted to the decision makers in an understandable way. Multi-colour graphical presentations are very useful for illustrative purposes but they should not be used to impress or convince where the scientific foundation is weak. There must also be a point where the modeller has the courage to say: »No sir, this problem is too difficult to model. You have to base your decision on something else«.

Acknowledgements

This paper is based on experience in model development and application that started in 1971. The work was initially financed by the Swedish Meteorological and Hydrological Institute (SMHI) but has later also been supported by other organizations of which the Swedish Association of River Regulation Enterprises (VASO), the Swedish Natural Science Research Council and the Swedish Environmental Protection Board are the most significant. I am particularly grateful to the staff of the Hydrological and Oceanographical Research Section of SMHI for their support and criticism during preparation of the manuscript.

References

Principles and Confidence in Hydrological Modelling


Swedish Committee on Spillway Design (1990) Riktlinjer för bestämning av dimensionerande flöden för dammanläggningar. (Guidelines for the calculation of design floods for dams, in Swedish), Swedish State Power Board, Swedish Power Association, Swedish Meteorological and Hydrological Institute, Norrköping.


**Address:**

Swedish Meteorological and Hydrological Institute, S-601 76 Norrköping, Sweden.

Received: 17 October, 1990

Accepted: 30 January, 1991