10.1098/rsta.2002.1008

THE ROYAL SOCIETY

Advances in real-time flood forecasting

By Peter C. Young

Centre for Research on Environmental Systems and Statistics, University of Lancaster, Lancaster LA1 4YQ, UK and Centre for Resource and Environmental Studies, Institute of Advanced Studies, Australian National University, Canberra ACT 2020, Australia

Published online 24 May 2002

This paper discusses the modelling of rainfall-flow (rainfall-run-off) and flow-routeing processes in river systems within the context of real-time flood forecasting. It is argued that deterministic, reductionist (or 'bottom-up') models are inappropriate for real-time forecasting because of the inherent uncertainty that characterizes river-catchment dynamics and the problems of model over-parametrization. The advantages of alternative, efficiently parametrized data-based mechanistic models, identified and estimated using statistical methods, are discussed. It is shown that such models are in an ideal form for incorporation in a real-time, adaptive forecasting system based on recursive state-space estimation (an adaptive version of the stochastic Kalman filter algorithm). An illustrative example, based on the analysis of a limited set of hourly rainfall-flow data from the River Hodder in northwest England, demonstrates the utility of this methodology in difficult circumstances and illustrates the advantages of incorporating real-time state and parameter adaption.

Keywords: rainfall-flow processes; data-based mechanistic models; recursive estimation; real-time forecasting; parameter adaption; variance adaption

1. Introduction

The primary objective of this paper is to describe recent research on the design of flood-forecasting procedures—procedures that can be applied to the problem of predicting future flow volumes and, therefore, future flood events in river systems. The aim of this research is to produce an online, real-time approach to flood forecasting that is inherently stochastic and so able to predict not only the likely level of future flow, but also the uncertainty associated with this prediction. In this manner, the probability of a flood occurring in the near future is quantified and this additional information can then be used as a basis for decision making and operational management in flood-prone locations.

The paper has another, underlying objective that is of deeper philosophical and methodological significance and is, in part, a response to the recent renewed interest in the so-called (see, for example, Klemes 1983; Silvert 1993; and the references cited therein) 'top-down' (or 'holistic') approach to modelling natural systems. Interest in top-down modelling has been revived largely because the alternative 'bottom-up' or 'reductionist' philosophy, which dominated much of the research conducted during

One contribution of 18 to a Discussion Meeting 'Flood risk in a changing climate'.

the last century, has failed to solve the many problems of identifying, estimating and validating models of natural environmental systems (see, for example, the discussion on this in Oreskes *et al*. (1994), Young *et al*. (1996) and Shackley *et al*. (1998)).

The author's early papers on this topic (e.g. Young 1978) rejected the idea of 'deterministic reductionism', i.e. the view that a model can be constructed in the form of deterministic equations based on the modeller's perception of the physical system, and they presented initial thoughts on a more objective, statistical approach to modelling poorly defined systems of all kinds. The papers also adumbrated very similar anti-reductionist arguments that have appeared recently in the hydrological literature and express some of these same views within the present hydrological context (Jakeman & Hornberger 1993; Beven 2000*a*). Quite similar anti-reductionist views are also appearing in other areas of science: for instance, in a recent lecture (Lawton 2001), the current Chief Executive of the Natural Environment Research Council (NERC) recounted the virtues of the top-down approach to modelling ecological systems (although, for some reason, he did not appear to accept that such reasoning could also be applied to other natural systems, such as the physical environment).

This paper briefly outlines the general problem of modelling rainfall-flow processes and the use of such models in flow-forecasting system design. It then goes on to describe the data-based mechanistic (DBM) approach to top-down modelling within the context of real-time flow forecasting. This includes two main components. First, stochastic DBM modelling based on the statistical identification and estimation of physically meaningful, nonlinear, transfer-function models. Second, the design of an adaptive Kalman filter forecasting algorithm based on a stochastic state-space formulation of these models. This approach has the virtue of being inherently stochastic and, because it is formulated in Bayesian, recursive estimation terms (e.g. Kalman 1960; Bryson & Ho 1969; Young 1984), it provides an ideal basis for real-time implementation and the introduction of *adaptive* procedures.

Such adaptive forecasting is motivated by a view that the rainfall-flow and riverine flow processes are inherently 'non-stationary', i.e. no completely fixed model with constant parameters, estimated by reference to historic rainfall-flow data, will be able to characterize the catchment behaviour for all times into the future. As a result, the forecasting system should be based on models that are able to adjust to any, normally small, changes in the catchment behaviour not predicted accurately enough by the initially estimated model. Also, as we see later, it is advantageous to introduce an additional adaptive mechanism to handle the problem of *heteroscedasticity* in rainfall-flow data, i.e. the fact that the noise on the flow measurements has a variance that changes markedly depending on the flow volume.

It is important to emphasize that the hydrological models considered in this paper are all of a 'lumped-parameter' variety (i.e. they consist of the linear and nonlinear transfer functions that are the discrete-time equivalents of differential equations and describe the temporal behaviour only at selected spatial nodes within the catchment system). The alternative 'distributed-parameter' models, which involve spatio-temporal aspects of the catchment and are described by models such as partial differential equations in time and space (or some equivalent of these), are not considered at all.

Such distributed models are clearly attractive in these days of geographical information systems (GISs) and weather radar, since they are, potentially at least, able to exploit spatial information of this type. Within a flood-forecasting system, distributed models are of particular relevance because they can hope to predict the spatio-temporal progress of flood inundation, as in Romanowicz & Beven (1998) and Beven *et al*. (2000). However, the parenthetical comment in the title of the latter paper 'Mapping the probability of flood inundation (even in real time)' hints at the difficulties of using such computationally intensive models in real-time applications, even if the other theoretical and practical problems associated with such large (and often reductionist) models could be solved. In the meantime, there is a clear need for research on simpler approaches that involve the efficient amalgamation of distributed- and lumped-parameter concepts, e.g. the distributed models of rainfall and its distribution throughout the catchment could provide improved estimates and forecasts of the rainfall inputs for lumped-parameter models, such as those discussed in the present paper. For, as we see in the practical example described later in § 6, it is the inadequacy and inconsistencies of the rainfall inputs and forecasts that appear to most limit the accuracy of the flow and flood forecasts.

2. Rainfall-flow modelling

Characterization of the nonlinear dynamic relationship between rainfall and river flow is one of the most interesting modelling problems in hydrology. It has received considerable attention over the past 40 years, with mathematical and computer-based models ranging from simple 'black-box' representations to complex, physically based catchment models. It would be impossible to review this enormous literature here. Fortunately, however, there are many books that deal in whole, or in part, with this challenging area of science and engineering. Useful texts of this type are Anderson & Burt (1985, see particularly the chapter by Wood & O'Connell), Shaw (1994), Singh (1995) and Beven (2000*b*). The latter book, in particular, provides a clearly written review of the whole topic that not only deals critically with many recent developments but also provides an excellent introduction to the subject at the start of the 21st century. In addition, two recent reports by the UK Environment Agency (Moore & Bell 2000; Bell *et al*. 2000; see also Moore *et al*. 2000) are of considerable importance in both reviewing and comparing rainfall-flow models within the real-time forecasting context. Unfortunately, as the authors point out, only a limited subset of *transfer-function* (TF) models were considered (isolated event-mode, linear TF models), so the comparative results are not particularly relevant to the present paper, which deals with much-less-restricted and more-advanced TF models.

Wheater *et al*. (1993) categorized rainfall-flow models into four, broad types: conceptual models, physics-based models, metric models and hybrid metric–conceptual (HMC) models. Conceptual and physics-based models tend to be the slaves of deterministic reductionist thinking (see above). As a result, they are often very large and suffer from problems of over-parametrization that make rigorous statistical identification and estimation difficult or even impossible. At the other extreme, metric models—such as the neural network (e.g. Tokar & Johnson 1999) and neuro-fuzzy types (e.g. Jang *et al*. 1997)—are the epitome of 'black-box' modelling, revealing very little of their internal structure that has any physical meaning.

The HMC models are an attempt to combine the ability of metric models to efficiently characterize the observational data in statistical terms (the 'principle of parsimony' (Box & Jenkins 1970); or 'Occam's razor', its medieval equivalent), with the advantages of *simple* conceptual models that have a prescribed physical interpretation within the current scientific paradigm. In practical engineering terms, it is often an advantage if the end-user understands the nature of the forecasting algorithm, so that this physically meaningful interpretation helps to engender confidence in the nature of the resulting design. Also, it is an essential component of the DBM modelling procedures that are used in this paper. For these reasons, HMC models provide an attractive vehicle for real-time flood forecasting.

Within the category of HMC models, two main approaches to modelling can be discerned; approaches which, not surprisingly, can be related to the more general deductive and inductive approaches to scientific inference that have been identified by philosophers of science from Francis Bacon (1620) to Karl Popper (1959) and Thomas Kuhn (1962).

The *hypothetico-deductive* **approach.** Here, the *a priori* conceptual model structure is effectively a (normally simple) theory of hydrological behaviour based on the perception of the hydrologist/modeller and is strongly conditioned by assumptions that derive from current hydrological paradigms. A typical current example is the IHACRES model of Jakeman *et al*. (1990).

The *inductive* **approach.** Here, theoretical preconceptions are avoided as much as possible in the initial stages of the analysis. In particular, the model structure is not pre-specified by the modeller but, wherever possible, it is inferred directly from the observational data in relation to a more general class of models. Only then is the model interpreted in a physically meaningful manner, most often (but not always) within the context of the current hydrological paradigm. Typical examples are the DBM rainfall-flow models presented in Young (1993, 1998, 2001*a*, *b*), Young & Beven (1994) and Young *et al*. (1997).

The inductive DBM approach to modelling forms the basis for the research described in the rest of this paper. Previous publications (Young 1978, 1998; Beck 1983; Young *et al*. 1996; and the references cited therein) map the evolution of this DBM philosophy and its methodological underpinning in considerable detail. As these references demonstrate, DBM models can be of various kinds depending upon the nature of the system under study. In the present flood-forecasting context, however, they take the form of nonlinear, stochastic TF representations of the rainfall-flow processes active in the river catchment.

3. Transfer-function modelling: historical background

TF modelling originally derives from the systems and control literature, where it has been used for over half a century as a major tool in modelling and control system design for linear dynamic systems. TF models also have an obvious appeal in hydrological terms, since the unit impulse response of the TF is an amplitude-scaled equivalent of the hydrological instantaneous unit hydrograph (IUH) (see, for example, Shaw 1994). As a result, TF models were quickly assimilated into hydrological research and have figured prominently in the hydrological literature for many years. Early examples are Dooge (1959) and Nash (1959), the latter introducing the now well-known ‘Nash cascade’, which is a chain of first-order transfer functions used for flow routeing (i.e. flow–flow modelling along the river channel). Since then, there have been many references to TF models in the hydrological literature, again too numerous to review here. The present author (Young 1986) interpreted existing flow-routeing models in TF terms, showing how they could be recursively estimated and used for flow-forecasting purposes.

While useful for modelling flow processes in river channels, an early application of TF modelling to rainfall-flow data (Young 1974) demonstrated that linear TF models could only characterize rainfall-flow dynamics in the short term, as a description of the dynamics associated with *individual* storm events. However, if the input (numerator) parameters of the TF were allowed to vary, the model could then capture the effects of temporal changes in the catchment soil-water storage and modify the rainfall-run-off behaviour accordingly. When combined with methods of recursive estimation (e.g. Young 1974, 1984, 1999*b*), such time variable parameter (TVP) models could then form the basis for *parameter-adaptive* flood-forecasting procedures (see Cluckie 1993; Lees *et al*. 1994).

The TVP model in Young (1974) led quickly to the hypothetic-deductive formulation of the nonlinear 'Bedford–Ouse' model (BM) (e.g. Whitehead & Young 1975; Young 2001*b*). This consists of two components connected in series: an *effective rainfall* (sometimes erroneously referred to as 'rainfall excess') nonlinearity, which accounts for the catchment storage effects and helps to remove the requirement for the TVPs; and a constant-parameter, linear TF, which models the underlying IUH dynamics. This special type of model (known as a 'Hammerstein' model in the systems literature) is an HMC model, as discussed previously, since the nonlinearity is one particular conceptualization of the catchment storage dynamics and its effect on the rainfall-run-off process.

Using the alternative, inductive approach of DBM modelling, the author (Young 1993) showed that the variations in the input parameters of the earlier TVP TF model could be considered as being 'state dependent'; and the resulting nonlinear state-dependent parameter (SDP) model could be identified and estimated using advanced methods of recursive fixed-interval smoothing (Young 2000, 2001*a*). In particular, this SDP analysis showed that the input parameters in the TF are dependent upon the changes in flow, with the flow effectively acting as a *surrogate* measure of the catchment storage (see below, as well as the discussion in Young & Beven (1994), and the papers of Lees (2000*a*, *b*)). In the resulting SDP model, the effective rainfall nonlinearity is identified directly from the rainfall-flow data, so avoiding the intuitive conceptualization of the BM and IHACRES models. As we shall see, this model is also in a useful, minimally parametrized, form that is well suited to flood forecasting. At this point in time, therefore, it constitutes one of the most advanced TF models being used in flood forecasting and can be seen as a logical successor to previous TF models.

4. The generic catchment model based on TF concepts

Within the catchment-modelling context, TF models are of two types: the nonlinear rainfall-flow model; and the linear flow-routeing model. The complete model used in flood forecasting and warning applications is comprised of both types linked in a manner that reflects the physical nature of the catchment under study (e.g. Lees *et al*. 1994). In this paper, however, we concentrate almost completely on the rainfall-flow component. This is not because flow routeing is unimportant in real-time flood

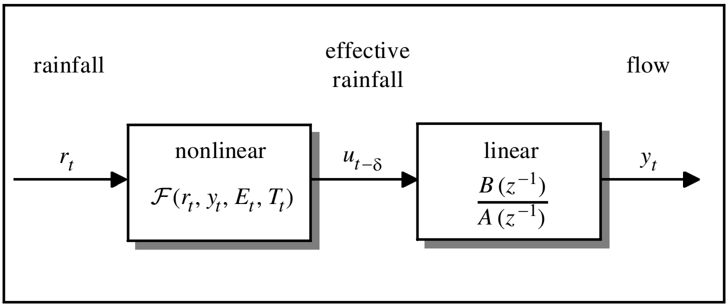

Figure 1. Block diagram of the generic TF rainfall-flow model.

forecasting; it is simply that the advances reported in this paper relate almost entirely to rainfall-flow modelling.

Previous DBM modelling of rainfall-flow data based on SDP estimation (see the references in the previous section) has confirmed many aspects of earlier hydrological research and identified the nonlinear model structure shown in figure 1.† The two components of the TF model are the linear component, which models the basic, underlying IUH behaviour; and the nonlinear component, which models the relationship between the *measured rainfall* r_t and the *effective rainfall* u_t, so controlling the magnitude of the hydrograph contribution through time.

If a constant, uniform sampling interval of Δt time units (e.g. 1 h) is used, the flow y_t at sample time t is related to past, sampled values of itself and present and past sampled values of u_t by the linear, discrete-time equation:

$$y_t = -a_1 y_{t-1} - \cdots - a_n y_{t-n} + b_0 u_{t-\delta} + b_1 u_{t-\delta-1} + \cdots + b_m u_{t-\delta-m} + \eta_t,$$

or, in TF terms,

$$y_t = \frac{B(z^{-1})}{A(z^{-1})} u_{t-\delta} + \xi_t. \tag{4.1 a}$$

In this second equation, z^{-1} is the backward shift operator, i.e. $z^{-r} y_t = y_{t-r}$, while $A(z^{-1})$ and $B(z^{-1})$ are *constant-coefficient polynomials* in z^{-1} of the following form:

$$\left.\begin{aligned} A(z^{-1}) &= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}, \\ B(z^{-1}) &= b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_m z^{-m}. \end{aligned}\right\} \tag{4.1 b}$$

The term δ is a pure time delay, measured in sampling intervals, which is introduced to allow for any temporal (advective) delay that may occur between the incidence of a change in u_t and its first effect on y_t. The noise term $\xi_t = \{1/A(z^{-1})\}\eta_t$ represents uncertainty in the relationship arising from a combination of measurement noise, the effects of other unmeasured inputs and modelling error. Sometimes, this noise variable is modelled explicitly as a coloured noise process, e.g. by an auto-regressive (AR) or auto-regressive moving-average (ARMA) model (Box & Jenkins 1970).

The structure (order) of the TF model (4.1 a) is defined by the triad $[n \quad m \quad \delta]$ and this is normally identified statistically from the data during the identification and estimation of the model, based on historical rainfall-flow data. This order is

† This model is similar in concept to the variable-gain-factor model suggested by Ahsan & O'Connor (1993), although its identification, estimation and implementation are quite different.

normally low, with $n \leqslant 2$, $m \leqslant 3$, while the value of δ is defined by the nature of the catchment and the location of the measurement devices, so its range is more difficult to define *a priori*. The general TF model form $B(z^{-1})/A(z^{-1})$ defines the input–output relationship between u_t and y_t and its unit impulse response is a scaled version of the underlying IUH. But, as we see later, it can also be decomposed into a parallel connection of lower-order processes. This decomposition not only makes the physical interpretation of the TF more transparent, it can also improve its performance in forecasting terms when implemented within a real-time flood-forecasting system (see §§ 5 and 6 below).

In general, the nonlinear component $\mathcal{F}(r_t, y_t, E_t, T_t)$ in figure 1 denotes a nonlinear functional relationship defining the unobserved catchment storage (or some surrogate for this). In addition to the rainfall r_t, it can involve other relevant measured variables, such as the temperature T_t, potential evaporation E_t and flow y_t, all of which could help to define the changes in soil moisture and storage if they are available. In the case of the standard DBM model, however, the nonlinearity is defined more simply by the equation

$$u_t = cf(y_t)\, r_t \quad \text{or} \quad u_t = cf(r_t)\, r_t. \tag{4.1 c}$$

The physical significance of these nonlinear functions is discussed below. Typically, they are initially identified through SDP estimation in non-parametric (graphical or 'look-up' table) form, without any prior assumptions about their nonlinear nature. This is then parametrized in some simple manner: for example, in Young (1993), Young & Beven (1994), Young & Tomlin (2000), Lees (2000*a*, *b*) and the present paper, $f(y_t)$ is defined as a power law $f(y_t) = y_t^{\gamma}$ with the power-law exponent γ estimated from the data. However, later research has shown that other parametric functions may be more effective and this is a suitable topic for future research (see § 7). The attraction of this SDP estimation approach is that the nonlinear function is inferred from the rainfall-flow data and not assumed *a priori*, as in HMC models such as the BM and IHACRES models.

The DBM model in the above form appears to have wide application potential. In addition to rivers in Australia (e.g. Young *et al*. 1997) and the USA (Young 2001*a*, *b*), it has been combined with an adaptive gain updating scheme in the parameter-adaptive Dumfries flood-warning system (Lees *et al*. 1994), which has been operating successfully without major modification since 1991; and it has been embedded within the Kalman filter to provide a *state-adaptive* forecasting system for the River Hodder in northwest England (see § 6 below and Young & Tomlin (2000)).

As we have stressed, an important aspect of DBM modelling is that the model can be interpreted in physically meaningful terms. In this regard, let us consider first the nonlinear effective rainfall equations in (4.1 *c*). Of course, in the case where $u_t = cf(y_t)\, r_t$, the relationship does *not* mean that the effective rainfall is physically a function of flow. Rather, the measured flow y_t is effectively acting here as an objectively identified *surrogate* for the catchment storage. This seems sensible from a hydrological standpoint, since flow is clearly a function of the catchment storage and its pattern of temporal change is likely to be similar. Moreover, it can be shown (Beven 2000*b*, p. 94; Lees 2000*a*) that there are parallels between the form of this nonlinearity and the well-known hydrologic concept of a 'dynamic contributing area'.

The effective rainfall from equation (4.1 *c*) provides the input to the linear TF model component (4.1 *a*). Often, in the case of hourly data, this TF is identified as a second order $\begin{bmatrix}2 & 2 & \delta\end{bmatrix}$ model† characterized by real eigenvalues (the roots of the $A(z^{-1})$ polynomial). In this case, the TF can be decomposed into a parallel pathway form, with first-order storage equations in each pathway (e.g. Wallis *et al.* 1989; Jakeman *et al.* 1990; Young 1992). In the case of the River Hodder example described in § 6, for instance, the partitioned flow components passing down these pathways, $x_{1,t}$ and $x_{2,t}$, are generated by the following equations.

(1) A 'quick-flow' pathway described by a first-order TF,

$$x_{1,t} = \frac{\beta_1}{1+\alpha_1 z^{-1}} u_{t-4}, \tag{4.2 a}$$

which has a partition percentage of 56%, a residence time of 5.5 h and an advective time delay of 4 h, so producing a total travel time of 9.5 h.

(2) A 'slow-flow' pathway described by a first-order TF,

$$x_{2,t} = \frac{\beta_2}{1+\alpha_2 z^{-1}} u_{t-4}, \tag{4.2 b}$$

with a partition percentage of 44%, a residence time of 84 h and the same advective time delay of 4 h, so producing a total travel time of 88 h.

Given these derived model parameters, the most obvious physical interpretation of the DBM model is that the effective rainfall affects the river flow via two main pathways. First, the initial rapid rise in the hydrograph derives from the 'quick-flow' pathway, probably as the aggregate result of the many surface processes active in the catchment. And the long, elevated tail in the recession of the hydrograph arises from the 'slow-flow' component, most likely the result of water displacement (probably of 'old water') from the storage within the groundwater system.

It must be emphasized that the estimated TF and its decomposition are stochastic objects and so the uncertainty that is inherent in their derivation needs to be taken into consideration when interpreting the model in these physically meaningful terms. In the above example, for instance, Monte Carlo simulation (MCS) analysis (e.g. Young 1999*a*) shows that, while the estimated quick-flow pathway dynamics are quite well defined, the slow-flow pathway dynamics are highly uncertain with a very skewed distribution towards larger residence times (Young 2001*c*).

5. Data assimilation: the recursive Kalman filter, state and parameter-adaptive forecasting

Most conventional methods of flow forecasting use the estimated ('calibrated') model for generating forecasts. But if we are concerned with forecasting flow several hours ahead, rather than with simply modelling the rainfall-flow data, then it cannot be assumed that the estimated model provides the optimum vehicle for generating such forecasts. The reason for this is obvious. The parameters of the model are normally estimated by minimizing some form of cost function that involves either the error between the model-generated flow and the measured flow, or the one-step-ahead

† In the case of daily data, a $\begin{bmatrix}2 & 3 & \delta\end{bmatrix}$ is more normal, so allowing u_t to have an instantaneous (within 1 day) effect on y_t.

prediction errors (as in maximum-likelihood (ML) estimation). However, the error that is relevant for multi-step-ahead forecasting purposes is not the 'fitting' or simple prediction error, it is the multi-step-ahead forecasting error based on the required forecast lead time. In other words, within the flood-forecasting and warning context, a catchment model based on rainfall-flow and flow-routeing TF models should not be considered as an end in itself. Rather, it is a major component of a *data-assimilation* system that collects data from remote sensors within the catchment and 'blends' these data with the model in a statistical manner to produce forecasts for multiple time-steps into the future. In the case of stochastic TF models such as those discussed above, an obvious statistical framework for data assimilation is the Kalman filter (KF hereafter), based on a stochastic state-space (SS) formulation of the catchment model, as described, for example, in Young & Tomlin (2000).

Formulation of the KF equations introduces additional, unknown parameters, normally termed 'hyperparameters' to differentiate them from the model parameters. In the present context, these take the form of noise variance ratio (NVR) parameters, which specify the nature of the stochastic inputs to the state equations and so define the level of uncertainty in the evolution of each state (the quick- and slow-flow states, respectively) relative to the measurement uncertainty. The presence of the NVR parameters in the KF is important because it allows for 'state adaption', i.e. the estimates of the state variables are continually adjusted to allow for the presence and effect of the unmeasured stochastic disturbances.

Clearly, the NVR hyperparameters have to be estimated in some manner on the basis of the data. One well-known approach is to exploit ML estimation based on *prediction error decomposition* (see Schweppe 1965; Young 1999*b*). Another, used later in the example of § 6, is to assume that *all* the parameters of the state-space model are unknown and re-estimate them by minimizing the variance of the multi-step-ahead forecasting errors. In effect, this optimizes the memory of the recursive estimation and forecasting algorithm (Young & Pedregal 1999) in relation to the rainfall-flow data in order to achieve optimum multi-step-ahead forecasts.

Although the parameters and hyperparameters of the KF-based forecasting system can be optimized in this manner, we cannot be sure that the system behaviour may not change sufficiently over time to require their adjustment. In addition, it is well known that the measurement noise ξ_t is quite highly heteroscedastic, i.e. its variance changes quite radically over time, with much higher variance occurring during storm events. For these reasons, it is wise to build some form of parameter and variance adaption into the forecasting algorithm, as discussed fully in Young (2001*c*) and used in the example described in § 6 below.

One limitation of using an adaptive KF as the data-assimilation engine in flow forecasting is its assumption that the stochastic processes are Gaussian. Since the KF is inherently a Bayesian recursive estimation procedure (see Bryson & Ho 1969), the most obvious way of removing this restriction is to consider extending the algorithm using Bayesian numerical methods that exploit MCS. Early use of MCS in hydrology (e.g. Whitehead & Young 1979) was inhibited by computational limitations, but, in recent years, the advances in computers have led to an explosion of research in this area. At the moment, there are a wide spectrum of such methods available ranging from Markov chain Monte Carlo (MCMC) (see, for example, Gamerman 1997), through Monte Carlo filtering algorithms (e.g. Kitagawa 1996; Thiemann *et al.* 2001), to simpler non-recursive approaches such as the generalized likelihood uncertainty estimation (GLUE) procedure (Beven & Binley 1992). Research is continuing on the best approach in the present context but simple Monte Carlo extensions to the adaptive KF described above are yielding promising results.

6. An illustrative example: adaptive flow forecasting for the River Hodder in northwest England

This example is concerned with the analysis of hourly flow, measured during 1993, at Hodder Place gauging station on the River Hodder in northwest England. This section presents an outline of the major results obtained in this analysis: they were generated using estimation procedures from the CAPTAIN toolbox developed in CRES at Lancaster for use in MATLAB™(see http://www.es.lancs.ac.uk/cres/captain/). The full results are reported in Young (2001*c*), which includes a more comprehensive discussion and can be downloaded at www.es.lancs.ac.uk/cres/papers/papers01.html.

The River Hodder has a catchment area of 261 km^2 and it forms part of the larger River Ribble catchment area of 456 km^2. DBM model identification and estimation is based on 720 h of hourly rainfall-flow data measured during January 1993, as shown in figure 2. The rainfall series r_t, measured in mm h^{-1}, is based on a Thiessen polygon average of all available rain gauges within the Hodder catchment; while the flow series y_t, measured in the same units as the rainfall (computed by dividing the gauged volumetric flow rate by the catchment area), is obtained from an Environment Agency flow gauge located at Hodder Place. The subsequent validation and forecasting analysis is based on a further 480 h of rainfall-flow data measured later, during December 1993, as shown in figure 3. Young & Tomlin (2000) have previously used this second dataset to illustrate how a second-order, nonlinear DBM model can provide the basis for KF-based flow forecasting, so the present analysis can be seen as an extension of these earlier studies.

It must be emphasized that these datasets were not chosen to produce good results: indeed, the modelling and forecasting problem they pose is quite difficult, since the estimation sample size $N = 720$ covers a very short period (just over a month) and the measured data (particularly the rainfall) are not particularly good quality. Note also that the validation dataset exhibits quite significantly larger maximum flow rates than those in the estimation dataset, so that the predictive and extrapolative ability of the nonlinear model is evaluated in the face of this larger envelope of rainfall-flow conditions.

(*a*) *Identification and estimation*

Based on the estimation dataset, the finally identified and estimated DBM model takes the form

$$y_t = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} u_{t-4} + \xi_t, \qquad u_t = \{c y_t^{\gamma}\}\, r_t. \tag{6.1}$$

The optimized parameter estimates are

$$\hat{a}_1 = -1.821(0.012), \qquad \hat{a}_2 = 0.823(0.011), \qquad \hat{b}_0 = 0.102(0.0030),$$
$$\hat{b}_1 = -0.1002(0.003), \qquad \hat{\gamma} = 0.281(0.009), \qquad \hat{c} = 1.17,$$

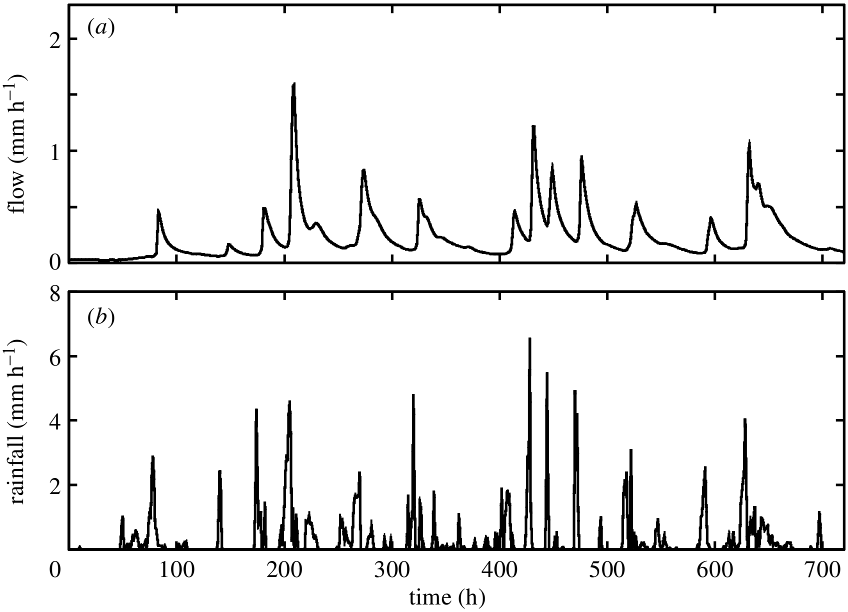

Figure 2. Hourly rainfall-flow data for the River Hodder during January 1993.

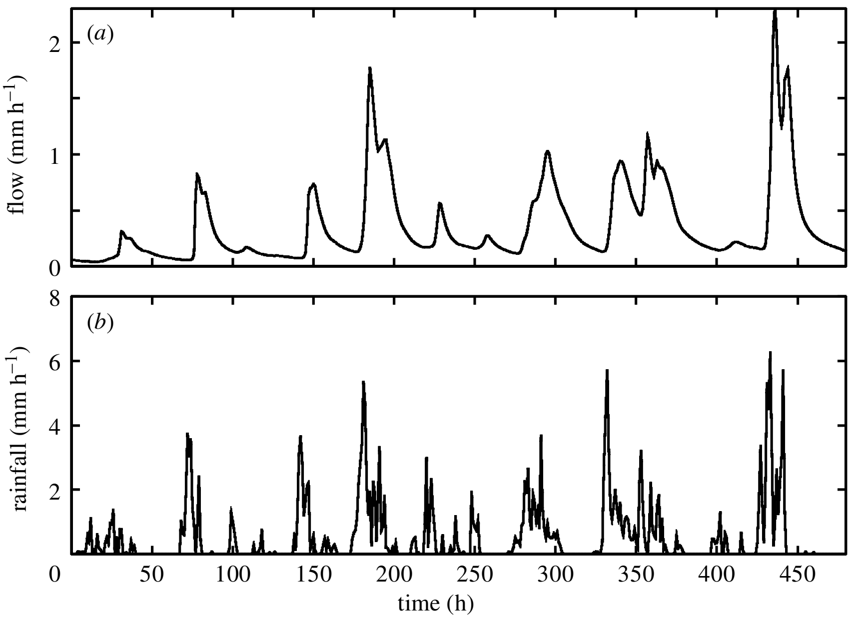

Figure 3. Hourly rainfall-flow data for the River Hodder during December 1993.

where the figures in parentheses adjacent to the estimates are the standard error bounds. The noise model is identified as a third-order AutoRegressive (AR(3)) model (e.g. Box & Jenkins 1970) and the associated parameter estimates are

$$\hat{c}_1 = -1.114(0.037), \qquad \hat{c}_2 = 0.507(0.052), \qquad \hat{c}_3 = -0.153(0.037).$$

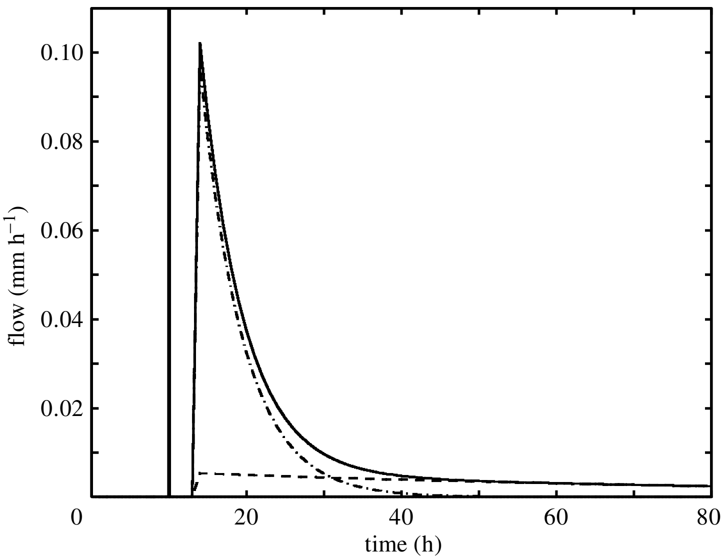

Figure 4. Estimated unit hydrograph (unit impulse response) based on the DBM model (solid line); impulsive rainfall input (vertical line at $t = 10$ h); quick-flow component (dash-dotted line); slow-flow component (dashed line).

The model (6.1) has a coefficient of determination $R_{\mathrm{T}}^2 = 0.844$ based on its response or 'simulation' error (i.e. 84% of the flow variance is explained by the model),† while the standard coefficient of determination based on the final stochastic residuals from the AR(3) noise model (i.e. one-hour-ahead prediction errors) is $R_1^2 = 0.95$. The auto (acf) and partial (pacf) autocorrelation functions of these stochastic residuals show no significant temporal correlation, although, as is normal in rainfall-flow models, they are highly heteroscedastic (see later) and correlated to a minor extent with the rainfall input.

As required by the DBM modelling strategy, the model (6.1) can be interpreted well in physically meaningful terms. Based on a partial fraction expansion of the linear TF, as shown in equations (4.2 *a*) and (4.2 *b*), it can be interpreted as a parallel connection of 'quick' and 'slow' first-order processes. In figure 4, the estimated hydrographs (impulse responses) associated with this parallel decomposition are compared with the complete hydrograph of the linear TF in the model (6.1). It is clear that much of the initial response is associated with the quick-flow pathway, while the main effect of the slow-flow pathway is to raise the longer-term tail of the hydrograph recession.

(*b*) *Adaptive forecasting*

As pointed out in § 5, the estimated or 'fitted' model (6.1) does not necessarily provide the best basis for multi-hour-ahead forecasting. In order to design the flow-forecasting system, therefore, it is necessary to reoptimize the model parameters, and any other associated hyperparameters of the KF, based on an appropriate forecasting cost function. For simplicity in this illustrative example, we use a simple

† Often referred to in the hydrological literature as the 'Nash–Sutcliffe efficiency' (Nash & Sutcliffe 1970).

least-squares cost function in the error between the specified multi-hour-ahead forecast and the measured flow over the estimation dataset. Given the 4 h advective time delay in the model (6.1), it makes sense to assume here that the major objective of forecasting in the present example is to optimize the 4 h-ahead forecasts. Of course, other cost functions could be used, such as a likelihood function based on the four-step-ahead forecasting errors, but this simple least-squares cost function will suffice for the present example and makes immediate physical sense, given the nature of the forecasting problem defined here. In this example, the length of the validation dataset is not really sufficient to consider the updating of *all* the model parameters but we are able to evaluate the effectiveness of simpler scalar gain and variance adaption procedures.

Figure 5 shows how the system performs on the validation dataset during December 1993. This figure also shows, plotted from above, the adjustments to the gauged rainfall made by the nonlinear effective rainfall coefficient cy^{γ}. It should be emphasized that the nonlinear rainfall-flow model parameters and hyperparameters used over this validation dataset are those optimized on the basis of the estimation dataset alone; *and they are maintained at these values over the whole of the validation dataset*. Given the limited size of the estimation dataset, however, it is not surprising that the optimized model is not entirely appropriate for the later December 1993 period, and the gain and variance adaption mechanisms are active in improving the forecasts. In particular, the adaptive gain reduces significantly when forecasting is applied over the December 1993 validation period (see Young 2001*c*). This indicates the value of such adaption in correcting for any deficiency in the estimated model. The effect of introducing the variance adaption is particularly noticeable in its adjustment of the standard error bounds plotted in figure 5 (dotted lines), which widen considerably over the peak-flow periods. This would not happen in the standard, non-adaptive KF algorithm, as pointed out by Lees (2000*a*).

Table 1 gives some indication of the forecasting performance achieved with and without adaption (as measured by appropriately defined coefficients of determination, R_i^2) when compared with other forecasting procedures under various settings of the forecasting system. The two other forecasting options are

(a) the 'standard' TF forecasting system in which the TF model (6.1) is used directly in its full form, without parallel decomposition or incorporation in the KF; and

(b) the naive forecasting system, in which the 4 h-ahead forecast $\hat{y}_{t+4}$ at any sampling instant t is simply set to the flow measurement y_t.

At first, these comparative results are surprising, since it is clear that the naive forecaster performs better than the standard TF forecaster based directly on the model (6.1). The main reason for this is that, as it stands, the TF model (6.1) is not good for forecasting because the numerator parameters $\hat{b}_0$ and $\hat{b}_1$ in the estimated TF model are approximately the same value and different in sign. This induces a near-differencing operation and causes 'spikes' in the forecasts that considerably degrade the forecasting performance. This problem is completely avoided, however, by the physically meaningful decomposition of the TF and its incorporation, in this decomposed form, within the KF forecasting engine.

Finally, although the forecasting system here has been designed for 4 h-ahead forecasts, it produces forecasts for any requested forecasting interval. For instance,

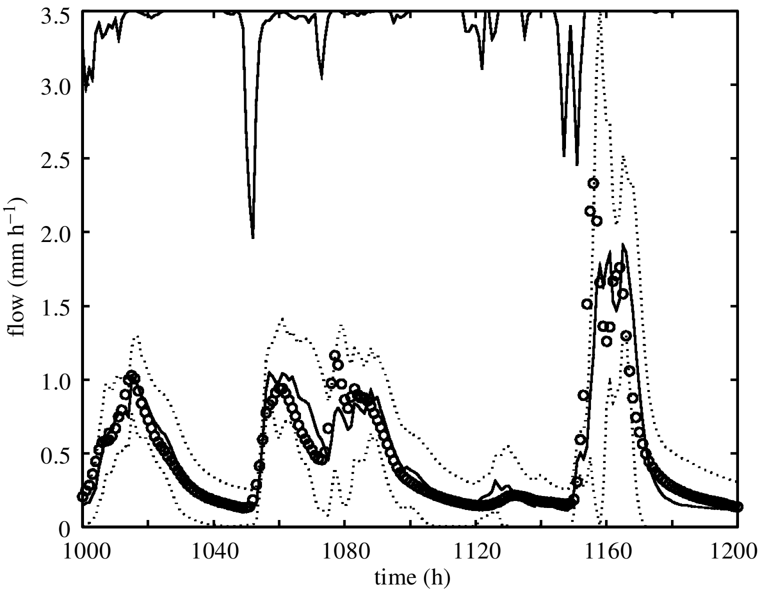

Figure 5. Comparisons of 4 h-ahead flow forecasts (solid line) and measured flow (circles) for the period between 10.00 and 12.00 during December 1993. The dotted lines show the standard error bounds. Plotted at the top is d_t, the difference between the gauged rainfall r_t and the effective rainfall u_t, subtracted from 3.5, for clarity: i.e. $d_t = 3.5 - (r_t - u_t)$.

Table 1. *Comparative forecast evaluation*

(Comparison of R_i^2 values obtained from various forecasting situations.)

level of adaption	four-step ahead	one-step ahead	four-step: no decomposition, no KF	naive forecast
only state adaption (SA)	0.820	0.876	0.313	0.461
SA + gain adaption	0.834	0.882	0.328	0.461
SA + variance adaption	0.841	0.876	0.350	0.461
SA + both	0.842	0.874	0.358	0.461

the coefficients of determination for the forecasts over all lead times from 1 to 6, R_i^2, $i = 1, 2, \ldots, 6$, are given below:

$$R_1^2 = 0.874, \qquad R_2^2 = 0.856, \qquad R_3^2 = 0.847,$$
$$R_4^2 = 0.842, \qquad R_5^2 = 0.764, \qquad R_6^2 = 0.658.$$

Of course, the forecasts for periods other than 4 h will not necessarily be optimal and may be improved by explicit optimization for the specified forecasting interval. For instance, the comparative figures obtained when the optimization is based on separate optimization at each sampling interval is as follows:

$$R_1^2 = 0.939, \qquad R_2^2 = 0.877, \qquad R_3^2 = 0.845,$$
$$R_4^2 = 0.842, \qquad R_5^2 = 0.767, \qquad R_6^2 = 0.658.$$

So we see that worthwhile advantage is obtained in the case of forecasting intervals from 1 to 3 h ahead. Although this would require three additional KF algorithms acting in parallel, the algorithms are so simple that the increase in the computational burden is quite acceptable. Also, note how the forecasts are degraded more for forecasting intervals greater than 4 h. This is because, after this interval, it is necessary to forecast the rainfall into the future, and here these forecasts are simply set to zero. Improved performance would be expected, therefore, if rainfall forecasts were available for prediction intervals greater than 4 h.

7. Conclusions

This paper describes some recent advances in stochastic modelling and forecasting that provide the basis for the implementation of real-time flow- and flood-forecasting systems. It argues that deterministic reductionist (or 'bottom-up') models are inappropriate for real-time forecasting because of the inherent uncertainty that characterizes river-catchment dynamics and the problems of model over-parametrization that are a natural consequence of the reductionist philosophy. The advantages of alternative, 'top-down', data-based mechanistic (DBM) models, statistically identified and estimated in an inductive manner directly from rainfall-flow data, are discussed. In particular, the paper shows how DBM models in the form of nonlinear, stochastic, transfer-function equations can be developed using powerful methods of recursive time-series analysis. Not only are these models able to characterize well the rainfall-flow dynamics of the catchment in a parametrically efficient manner, but, by virtue of the DBM modelling strategy, they can be interpreted in hydrologically meaningful terms. Most importantly in the forecasting context, the models are also in an ideal form for incorporation into a forecasting (or 'data-assimilation') algorithm based on a special, adaptive version of the Kalman filter algorithm.

The practical example described in the paper demonstrates how, with the minimum of rainfall-flow data and no available rainfall forecasts, the recursive estimation approach proposed here can generate useful flow forecasts for several hours ahead; forecasts that could form the basis for flood-warning-system design. This approach can be considered as a natural development of the Dumfries flood-warning system (Lees *et al*. 1994) mentioned in § 4. Both of these recursive approaches to real-time forecasting can be contrasted with more conventional, non-recursive real-time forecasting procedures proposed previously. A typical example is the adaptive scheme suggested by Brath & Rosso (1993), which addresses some of the same statistical issues raised in the present paper. However, it operates on an event basis rather than continuously; it uses repeated en bloc optimization rather than recursive estimation; it is based on a simple conceptual model with *a priori* assumed structure and parametrization; and it is computationally much more demanding.

Of course, there remain a number of methodological problems still to be solved. The DBM models discussed in the paper perform well but they cannot be considered completely satisfactory while the model residuals retain their current unsatisfactory statistical characteristics. In particular, the correlation remaining between the residuals and the rainfall input shows that the model is still not fully explaining the complete rainfall-flow process (although the remaining unexplained variance represents only a small proportion of the total variance). This limitation of the current DBM models (shared, the author believes, by all current rainfall-flow models, whatever their type) is almost certainly due to deficiencies in the effective rainfall nonlinearity and, possibly, the presence of other, smaller nonlinearities in the system, as yet unquantified. There is clear need for more research on this fascinating subject and, although such research would require the analysis of a large and comprehensive rainfall-flow database covering a wide array of different catchment behaviour, it would provide useful information for *all* existing rainfall-flow modelling studies, not just those discussed in this paper.

The author is very grateful to his colleagues Professor Keith Beven, Dr Paul McKenna and Dr Renata Romanowicz for reading and commenting on a draft of this paper. Naturally, the author is responsible for any errors or omissions.

References

Ahsan, M. & O'Connor, K. M. 1993 A simple nonlinear rainfall-runoff model with a variable gain factor. *J. Hydrol.* **155**, 151–183.

Anderson, M. G. & Burt, T. P. (eds) 1985 *Hydrological forecasting*. Wiley.

Bacon, F. 1620 *Novum organum*. In *The works*, vol. 3 (ed. and transl. B. Montague 1854), pp. 343–371. Philadelphia, PA: Parry & MacMillan.

Beck, M. B. 1983 Uncertainty, system identification and the prediction of water quality. In *Uncertainty and forecasting of water quality* (ed. M. B. Beck & G. Van Straten), pp. 3–68. Springer.

Bell, V. A., Carrington, D. S. & Moore, R. J. 2000 Comparison of rainfall-runoff models for flood forecasting. II. Calibration and evaluation of models. Environment Agency R&D Technical Report W242.

Beven, K. J. 2000*a* Uniqueness of place and process representations in hydrological modelling. *Hydrol. Earth Syst. Sci.* **4**, 203–213.

Beven, K. J. 2000*b* *Rainfall-runoff modelling: the primer*. Wiley.

Beven, K. J. & Binley, A. M. 1992 The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.* **6**, 279–298.

Beven, K. J., Romanowicz, R. & Hankin, B. 2000 Mapping the probability of flood inundation (even in real time). In *Flood forecasting: what does current research offer the practitioner?* (ed. M. J. Lees & P. Walsh), pp. 56–63. (BHS occasional paper no. 12, produced by the Centre for Ecology and Hydrology on behalf of the British Hydrological Society.)

Box, G. E. P. & Jenkins, G. M. 1970 *Time-series analysis, forecasting and control*. San Francisco, CA: Holden-Day.

Brath, A. & Rosso, R. 1993 Adaptive calibration of a conceptual model for flash flood forecasting. *Water Resources Res.* **29**, 2561–2572.

Bryson, A. E. & Ho, Y.-C. 1969 *Applied optimal control*. Waltham, MA: Blaisdell Publishing.

Cluckie, I. D. 1993 Real-time flood forecasting using weather radar. In *Concise encyclopedia of environmental systems* (ed. P. C. Young), pp. 291–298. Oxford: Pergamon.

Dooge, J. C. I. 1959 A general theory of the unit hydrograph. *J. Geophys. Res.* **64**, 241–256.

Gamerman, D. 1997 *Markov chain Monte Carlo*. London: Chapman & Hall.

Jakeman, A. J. & Hornberger, G. M. 1993 How much complexity is warranted in a rainfall-runoff model? *Water Resources Res.* **29**, 2637–2649.

Jakeman, A. J., Littlewood, I. G. & Whitehead, P. G. 1990 Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *J. Hydrol.* **117**, 275–300.

Jang, J.-S. R., Sun, C.-T. & Mizutani, E. 1997 *Neuro-fuzzy and soft computing*. Upper Daddle River, NJ: Prentice-Hall.

Kalman, R. E. 1960 A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Engng* **82**, 35–45.

Kitagawa, G. 1996 Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *J. Computat. Graph. Stat.* **5**, 1–25.

Klemes, V. 1983 Conceptualisation and scale in hydrology. *J. Hydrol.* **65**, 1–23.

Kuhn, T. 1962 *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Lawton, J. 2001 Understanding and prediction in ecology. Institute of Environmental and Natural Sciences. Lancaster University, Distinguished Scientist Lecture.

Lees, M. J. 2000*a* Advances in transfer function based flood forecasting. In *Flood forecasting: what does current research offer the practitioner?* (ed. M. J. Lees & P. Walsh), pp. 41–55. (BHS occasional paper no. 12, produced by the Centre for Ecology and Hydrology on behalf of the British Hydrological Society.)

Lees, M. J. 2000*b* Data-based mechanistic modelling and forecasting of hydrological systems. *J. Hydroinf.* **2**, 15–34.

Lees, M., Young, P. C., Beven, K. J., Ferguson, S. & Burns, J. 1994 An adaptive flood warning system for the River Nith at Dumfries. In *River flood hydraulics* (ed. W. R. White & J. Watts), pp. 65–75. Wallingford: Institute of Hydrology.

Moore, R. J. & Bell, V. A. 2000 Comparison of rainfall-runoff models for flood forecasting. I. Literature review of models. Environment Agency R&D Technical Report W241.

Moore, R. J., Bell, V. A. & Carrington, D. S. 2000 Intercomparison of rainfall-runoff models for flood forecasting. In *Flood forecasting: what does current research offer the practitioner?* (ed. M. J. Lees & P. Walsh), pp. 69–76. (BHS occasional paper no. 12, produced by the Centre for Ecology and Hydrology on behalf of the British Hydrological Society.)

Nash, J. E. 1959 Systematic determination of unit hydrograph parameters. *J. Geophys. Res.* **64**, 111–115.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models: discussion of principles. *J. Hydrol.* **10**, 282–290.

Oreskes, N., Shrader-Frechette, K. & Belitz, K. 1994 Verification, validation, and confirmation of numerical models in the Earth sciences. *Science* **263**, 641–646.

Popper, K. 1959 *The logic of scientific discovery*. London: Hutchinson.

Romanowicz, R. & Beven, K. J. 1998 Dynamic real-time prediction of flood inundation probabilities. *Hydrol. Sci. J.* **43**, 181–196.

Schweppe, F. 1965 Evaluation of likelihood functions for Gaussian signals. *IEEE Trans. Inform. Theory* **11**, 61–70.

Shackley, S., Young, P. C., Parkinson, S. D. & Wynne, B. 1998 Uncertainty, complexity and concepts of good science in climate change modelling: are GCMs the best tools? *Climatic Change* **38**, 159–205.

Shaw, E. M. 1994 *Hydrology in practice*, 3rd edn. London: Chapman & Hall.

Silvert, W. 1993 Top-down modelling in ecology. In *Concise encyclopedia of environmental systems* (ed. P. C. Young), p. 60. Oxford: Pergamon.

Singh, V. P. (ed.) 1995 *Computer models of watershed hydrology*. Highlands Ranch, CO: Water Resources Publications.

Thiemann, M., Trosset, M., Gupta, H. & Sorooshian, S. 2001 Bayesian recursive parameter estimation for hydrologic models. *Water Resources Res.* **37**, 2521–2535.

Tokar, A. S. & Johnson, P. A. 1999 Rainfall-runoff modeling using artificial neural networks. *J. Hydrol. Engng* **4**, 232–239.

Wallis, S. G., Young, P. C. & Beven, K. J. 1989 Experimental investigation of the aggregated dead zone model for longitudinal solute transport in stream channels. *Proc. Inst. Civil Engrs 2* **87**, 1–22.

Wheater, H. S., Jakeman, A. J. & Beven. K. J. 1993 Progress and directions in rainfall-run-off modelling. In *Modelling change in environmental systems* (ed. A. J. Jakeman, M. B. Beck & M. J. McAleer), pp. 101–132. Wiley.

Whitehead, P. G. & Young, P. C. 1975 A dynamic-stochastic model for water quality in part of the Bedford–Ouse River system. In *Computer simulation of water resources systems* (ed. G. C. Vansteenkiste), pp. 417–438. Amsterdam: North-Holland.

Whitehead, P. G. & Young, P. C. 1979 Water quality in river systems: Monte Carlo analysis. *Water Resources Res.* **15**, 451–459.

Young, P. C. 1974 Recursive approaches to time-series analysis. *Bull. Inst. Math. Applic.* **10**, 209–224.

Young, P. C. 1978 A general theory of modeling for badly defined dynamic systems. In *Modeling, identification and control in environmental systems* (ed. G. C. Vansteenkiste), pp. 103–135. Amsterdam: North-Holland.

Young, P. C. 1984 *Recursive estimation and time-series analysis*. Springer.

Young, P. C. 1986 Time-series methods and recursive estimation in hydrological systems analysis. In *River flow modelling and forecasting* (ed. D. A. Kraijenhoff & J. R. Moll), pp. 129–180. Dordrecht: Reidel.

Young, P. C. 1992 Parallel processes in hydrology and water quality: a unified time series approach. *J. Inst. Water Environ. Manag.* **6**, 598–612.

Young, P. C. 1993 Time variable and state dependent modelling of nonstationary and nonlinear time series. In *Developments in time series analysis* (ed. T. Subba Rao), pp. 374–413. London: Chapman & Hall.

Young, P. C. 1998 Data-based mechanistic modelling of environmental, ecological, economic and engineering systems. *Environ. Model. Softw.* **13**, 105–122.

Young, P. C. 1999*a* Data-based mechanistic modelling, generalised sensitivity and dominant mode analysis. *Comput. Phys. Commun.* **117**, 113–129.

Young, P. C. 1999*b* Nonstationary time series analysis and forecasting. *Prog. Environ. Sci.* **1**, 3–48.

Young, P. C. 2000 Stochastic, dynamic modelling and signal processing: time variable and state dependent parameter estimation. In *Nonstationary and nonlinear signal processing* (ed. W. J. Fitzgerald, A. Walden, R. Smith & P. C. Young), pp. 74–114. Cambridge University Press.

Young, P. C. 2001*a* The identification and estimation of nonlinear stochastic systems. In *Nonlinear dynamics and statistics* (ed. A. I. Mees), pp. 127–166. Birkhäuser.

Young, P. C. 2001*b* Data-based mechanistic modelling and validation of rainfall-flow processes. In *Model validation: perspectives in hydrological science* (ed. M. G. Anderson & P. D. Bates), pp. 117–161. Wiley.

Young, P. C. 2001*c* Advances in real-time forecasting. Centre for Research on Environmental Systems and Statistics, Lancaster University, Report no. TR/176.

Young, P. C. & Beven, K. J. 1994 Data-based mechanistic modelling and the rainfall-flow nonlinearity. *Environmetrics* **5**, 335–363.

Young, P. C. & Pedregal, D. 1999 Recursive and en-bloc approaches to signal extraction. *J. Appl. Stat.* **26**, 103–128.

Young, P. C. & Tomlin, C. M. 2000 Data-based mechanistic modelling and adaptive flow forecasting. In *Flood forecasting: what does current research offer the practitioner?* (ed. M. J. Lees & P. Walsh), pp. 26–40. (BHS occasional paper no. 12, produced by the Centre for Ecology and Hydrology on behalf of the British Hydrological Society.)

Young, P. C., Parkinson, S. D. & Lees, M. 1996 Simplicity out of complexity in environmental systems: Occam's razor revisited. *J. Appl. Stat.* **23**, 165–210.

Young, P. C., Jakeman, A. J. & Post, D. A. 1997 Recent advances in data-based modelling and analysis of hydrological systems. *Water. Sci. Technol.* **36**, 99–116.