How Much Complexity Is Warranted in a Rainfall-Runoff Model?

A. J. JAKEMAN

Centre for Resource and Environmental Studies, Australian National University, Canberra, Australian Capital Territory

G. M. HORNBERGER

Department of Environmental Sciences, University of Virginia, Charlottesville

Development of mathematical models relating the precipitation incident upon a catchment to the streamflow emanating from the catchment has been a major focus of surface water hydrology for decades. Generally, values for parameters in such models must be selected so that runoff calculated from the model "matches" recorded runoff from some historical period. Despite the fact that the physics governing the path of a drop of water through a catchment to the stream involves complex relationships, evidence indicates that the information content in a rainfall-runoff record is sufficient to support models of only very limited complexity. This begs the question of what limits the observed data place on the allowable complexity of rainfall-runoff models. Time series techniques are applied for estimating transfer functions to determine how many parameters are appropriate to describe the relationship between precipitation and streamflow in the case where data on only precipitation, air temperature, and streamflow are available. Statistics from an "information matrix" provide the clues necessary for determining allowable model complexity. Time series models are developed for seven catchments with widely varying physical characteristics in different temperate climatic regimes to demonstrate the method. It is found that after modulating the measured rainfall using a nonlinear loss function, the rainfall-runoff response of all catchments is well represented using a linear model. Also, for all catchments a two-component linear model with four parameters is the model of choice. The two components can be interpreted as defining a "quick flow" and "slow flow" response of the given catchment. The method therefore provides a statistically rigorous way to separate hydrographs and parameterize their response behavior. The ability to construct reliable transfer function models for describing the rainfall-runoff process offers a new approach to investigate empirically the controls of physical catchment descriptors, land use change, climate change, etc., on the dynamic response of catchments through the extensive analysis of historical data sets.

1. INTRODUCTION

The construction and application of watershed models describing precipitation to streamflow processes has been a prime focus of hydrological research and investigations for many decades. Both the amount of effort and the complexity of models seem to have increased continually with the expansion in available computing power. Most attention has been given to catchments subject to basically temperate climatology where hydrological responses tend to be simpler or involve a subset of the processes which occur in other climatic regimes. Despite the activity in modeling the rainfall-runoff process and the concentration on temperate catchments, hydrologists have noted the lack of real progress being made in watershed modeling generally and the problems of developing the process knowledge derived at small scales for use at larger scales [e.g., Beven, 1987]. Philip [1975, p. 23] saw the need "to identify and to recognise frankly the limits of what natural science and the 'scientific method' can bring to the tasks of catchment prediction." More specifically and recently, van Genuchten [1991, p. 190] summarized that future research in catchment modeling must address the problem of permissible system and model complexity, the scales over which model components are valid, and the integration of model components into an overall balanced framework.

One of the major problems in rainfall-runoff modeling is dealing with overparameterization. Loague and Freeze [1985, p. 245] applied several models of varying complexity to a number of catchments and concluded that the "fact that simpler, less data intensive models ... provided as good or better predictions than a more physically based model is food for thought." Hornberger et al. [1985] incorporated parameterizations into a version of TOPMODEL for processes observed to be occurring in the field but found that the 13 parameters could not be reliably estimated using rainfall-runoff data; they found that four parameters seemed to suffice to represent the transformation of rainfall to streamflow. Hooper et al. [1988] examined a very simple hydrological model with six parameters and still found it to be overparameterized. Beven [1989, p. 159] comments, There is a great danger of overparameterization if it is attempted to simulate all hydrological processes thought to be relevant and fit those parameters by optimisation against an observed discharge record. ... It appears that 3 to 5 parameters should be sufficient to reproduce most of the information in a hydrological record.

Hydrologists are faced with something of a dilemma. The most frequent application of rainfall-runoff models is in cases where the only data available are precipitation, temperature, and streamflow. Models that seek to incorporate processes known to be important hydrologically (at small scales) are likely to contain a rather large number of parameters, many of which will be correlated with other parameters [e.g., Clarke, 1973]. How are these observations to be reconciled with the contentions that only a model with a few parameters
can be supported by rainfall-runoff data? It appears that there are challenges that face those concerned with rainfall-runoff modeling. How much information is contained in records of precipitation and streamflow? How complex a parameterization is warranted in a rainfall-runoff model? Does adding spatially distributed data on physical catchment descriptors, such as on terrain, hydrologic soil, and vegetation properties, permit a more detailed parameterization?

A response to these challenges is attempted here first by presenting a framework to answer the question of what reliable information may reside in concurrent precipitation-streamflow measurements for assessing the dynamic characteristics of catchment response and for prediction of streamflow. In particular, the paper addresses the limitations of precipitation-streamflow modeling when measurements of other dynamic flow or concentration variables are not used as additional prior information in model construction. The framework allows inference of the number of streamflow components that can be identified from given precipitation and stream discharge observations. Its use is illustrated for catchments spanning a range of scales and basically temperate hydroclimatological regimes. Mainly daily data are used, and the specific results are applicable to analyses involving observational time series whose length is of the order of 100 times the quick flow response time constant. In the case of daily data this is of the order of 1 year.

The outcome of applying the framework is a hypothesis that, after allowing for antecedent conditions, the response of a catchment is predominantly linear over a wide range of temperate climatological regimes and down to small catchment scale. In response to van Genuchten [1991], the “permissible model complexity” seems to be generally low, containing around half a dozen parameters, and “the scales over which the model components are valid” are very wide. With much longer observational time series, it may be possible to identify the values of additional parameters. However, as pointed out by Sorooshian et al. [1983], rather than length, it is the quality of information contained in the data, which is important; data sequences which contain larger hydrologic variability are more likely to result in reliable parameter estimates.

2. METHODS

2.1. The Model

There are many formal ways to assess the information content in data with respect to some model M. In a stochastic setting the model M and its parameters are most completely specified by its probability distribution p(M). Given an evenly spaced time sequence of N rainfall-runoff samples

\[ \{r_N, q_N\}, (r_N = r_1, r_2, \ldots, r_N; q_N = q_1, q_2, \ldots, q_N) \]

where \( r_k \) is observed rainfall and \( q_k \) is observed streamflow at time step \( k \), the conditional distribution \( p(M|\{r_N, q_N\}) \) can be called the information about M contained in \( \{r_N, q_N\} \). In the next section, it will be argued how the covariance matrix of our model parameters can be invoked to determine the information about M in the time series samples.

The model used here to extract the information in rainfall-streamflow time series data consists of one nonlinear and one linear module. The nonlinear or rainfall loss module represents the transformation of rainfall \( r_N \) to “excess” rainfall \( u_N \). At each time step \( k \) a catchment wetness index, \( s_k \), or antecedent precipitation index is calculated by a weighting of the rainfall time series, the weights decaying exponentially backward in time from step \( k \), namely,

\[ s_k = c r_k + (1 - \tau_w^{-1}) s_{k-1} \]

\[ = c [r_k + (1 - \tau_w^{-1}) r_{k-1} + (1 - \tau_w^{-1})^2 r_{k-2} + \cdots] \]  

(1)

The parameter \( \tau_w \) is approximately the time constant, or inversely, the rate at which the catchment wetness declines in the absence of rainfall. Hence a larger value of \( \tau_w \) gives more weight to the effect of antecedent rainfall on catchment wetness than a smaller one. The excess or effective rainfall is computed using

\[ u_k = r_k s_k \]  

(2)

The parameter \( c \) in (1) is chosen so that the volume of excess rainfall is equal to the total streamflow volume over the calibration period, after adjustment for change in catchment storage between the beginning and end of the period. It is the increase in storage index per unit rainfall in the absence of evapotranspiration. It is not really a free parameter but merely a normalizing one.

To account for fluctuations in evapotranspiration, a simple function of temperature can be used to modulate the rate at which the catchment dries out. Then \( \tau_w \) in (1) is replaced with the function

\[ \tau_w(t_k) = \tau_w \exp [(20 - t_k)f] \]  

(3)

where \( t_k \) is the temperature in degrees Celsius at time step \( k \). In this way, \( \tau_w \) is inversely related to the rate at which catchment wetness declines at 20°C. The parameter \( f \) is a temperature modulation factor which determines how \( \tau_w(t_k) \) changes with temperature.

The general conclusions of this paper are independent of this nonlinear loss module. The structure of the accompanying linear module identified from any \( \{u_N, q_N\} \) is independent of the values of the parameters \( \tau_w, c, \) and \( f \), whereas the parameter values in the identified linear components are dependent. An explanation for this is that the nonlinearity between rainfall and streamflow has not been strong enough to impede the identification.

The linear module of the model converts excess rainfall \( u_k \) at time step \( k \) to streamflow \( q_k \). It is a transfer function of the form

\[ x_k = -a_1 x_{k-1} - \cdots - a_n x_{k-n} + b_0 u_k + \cdots + b_m u_{k-m} \]

\[ q_k = x_k + \xi_k \]  

(4)  

(5)

where \( \xi_k \) represents the addition of all data and model errors and \( x_k \) is a hypothetical error-free streamflow variable. The \( n + m + 1 \) elements of the vector \( a = (a_1, \cdots, a_n, b_0, b_1, \cdots, b_m)^T \) are the parameters to be optimized in the linear module.

There are several reasons that this model (1)–(5) is a natural one to use for extracting the information in time series \( \{u_N, q_N\} \):

1. The model and a similar version with a simpler nonlinear loss module are good predictors of streamflow and
with an appropriate parameter estimation tech-
ique, and parameter estimation.

The associated approximation of the unit hydro-
graph is an efficient one parametrically, corresponding to a
rational function representation of a polynomial function
input of excess rainfall. The time constant can be defined as
the time taken for the peak in output of a storage to recess to
exp (−1) of that peak value. In terms of an equivalent
time formulation, the characteristic response of each
storage can be considered to be defined by a unit
hydrograph (component) of form \( I_i \exp (-t/\tau_i) \), where \( I_i \) is
the relative peak of the hydrograph response and \( v_i \)
the integral of or area under this hydrograph response compo-
nent. We define \( \tau = (\tau_1, \tau_2, \cdots, \tau_n, v_1, \cdots, v_n)^T \) and the parameters
\( \tau_w, f, \) and \( c \) of the nonlinear module as the
dynamic response characteristics (DRCs) of the catchment.
The quantities \( I_i \) can also be regarded as DRCs, but the elements of the vector \( \tau \) are sufficient to completely define
the exponential response.

When \( n = 2 \) and \( m = 1 \) in (1), excess rainfall can be
considered to travel through a configuration of two parallel
storages as illustrated in Figure 1. In this case, \( I_1 + I_2 = v_1 + v_2 = 1 \), and the storage with smaller time constant
represents the quick flow component while that with the larger
slow flow component. In this case, \( \tau \) can be
written as \( (\tau_q, \tau_s, v_q, v_s)^T \). When \( n = 2 \) and \( m = 0 \), the
storages in Figure 1 are in series. The simple relationships
between the parameters in \( \tau \) and \( a \) for the two parallel
storage configuration in Figure 1 are

\[
\begin{align*}
\tau_q &= -\Delta_1 \ln (-\alpha_q) \\
\tau_s &= -\Delta_1 \ln (-\alpha_s) \\
v_q &= \beta_q/(1 + \alpha_q)
\end{align*}
\]

Fig. 1. Systems diagram of the most common model configuration identified.
where $\Delta$ is the sampling interval for the precipitation and streamflow time series and the $\alpha$ and $\beta$ parameters are obtained from the decomposition of the polynomial transfer function in the backward shift operator $B$ ($Bu_k = u_{k-1}$) according to

$$b_0 + b_1B \over 1 + a_1B + a_2B^2 = \beta q + \beta I$$

The sampling interval selected or available for the time series data clearly has an effect on the information that can be extracted about a model $M$. Too coarse a sampling interval will result in a loss of information about response dynamics. Too fine an interval can result in numerical instabilities [e.g., Jakeman and Young, 1980]. For our model, the appropriate sampling interval to select is one that is of the order of, but preferably less than, the time constant of the quickest identifiable response. This selection can make identification of slower components numerically difficult. A particular algorithm was applied by Jakeman et al. [1990] to obviate this problem and is again used here to extract the different components.

Note that with a two-parallel storage configuration, our overall model has seven parameters, six of which must be estimated. Use of the parameter $c$ in (1) constrains the volumetric gain between excess rainfall and streamflow, $(b_0 + b_1)/(1 + a_1 + a_2)$, to be unity, so that only three of the linear module parameters need be independently estimated.

2.2. Parameter Estimation and the Form of the Covariance Matrix

The main interest of the paper is on the complexity of the linear part of the model. Estimation is therefore focused on the mean and covariance of the parameters in the linear module given values of the parameters $\tau$, $f$, and $c$ in the nonlinear module. The parameter $c$ is obtained simply from the data and $\tau$ and $f$ using (2). The latter parameters can be optimized by trial and error using a simple search technique. The search can be applied to select those values which, when used in conjunction with an automatic algorithm for estimating the parameters $a$, satisfy an objective function in fitting streamflow.

There are many algorithms available for estimating the parameters in transfer function models of the form (4). Instrumental variable techniques are preferred here because they lead to simpler algorithms with good properties if one is mainly interested in the system dynamics and one cannot model or is not interested in the precise nature of the errors. They yield consistent estimates provided the errors $\xi_k$ are uncorrelated with the input $u_k$. Therefore they do not require the errors to be Gaussian or even independent and identically distributed. They can yield asymptotically efficient estimates if the errors are stationary. A covariance matrix is a by-product of the algorithms. The algorithmic details will not be reported here. Jakeman et al. [1989, 1990] summarize the properties of various instrumental variable algorithms and cite the major literature. The simple refined version is the one preferred because it performs well in the most difficult cases, in particular the case of estimating two hydrological stores or components where one component is small in volume and decays very slowly relative to the other component.

An information matrix $I$ for any input-output time series $\{u_N, q_N\}$ and transfer function model of the form (4)-(5) with parameters $a$ can be estimated [e.g., Pierce, 1972] as

$$I = (N/\delta)^2E[\xi_k^T \xi_k]$$

where $\delta^2 = \sum_{k=0}^{N-1} u_k^2$. $\xi_k$ is an estimate of the noise-free streamflow $x_k$, and $\delta^2 = \text{Var}(\xi_k) = \text{Var}(\hat{x}_k) = \text{Var}([q_k - \hat{x}_k])$ is the model error or residual variance. As described below, the information matrix and the residual variance allow one to determine how many model parameters are supported by the data. The specific approach was first used by Young et al. [1980] and is in the spirit of the philosophy of model parsimony espoused most notably by Box and Jenkins [1976].

If a simple refined instrumental variable (SRIV) algorithm is used to estimate $a$ then an analogous matrix $I^*$ is the information matrix. This corresponds to $I$ above except that each element is replaced by an analogous asterisked variable which is filtered according to

$$\hat{x}_k = -\hat{a}_1 \hat{x}_{k-1} - \cdots - \hat{a}_n \hat{x}_{k-n} + \hat{\xi}_k$$

$$u_k = -\hat{a}_1 u_{k-1} - \cdots - \hat{a}_n u_{k-n} + u_k$$

Here $(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_n)$ are the SRIV estimates of the first $n$ elements of $a$.

The covariance matrix $P$ of the parameters in $a$ is estimated as

$$P = (I^*)^{-1}$$

Under certain conditions [e.g., Pierce, 1972], the parameters estimated in $a$ are asymptotically normally distributed with covariance $P$. Thus for any model $M$, if and $P$ are total extracts of the information in the data $\{u_N, q_N\}$. They contain the probability distribution of the linear response characteristics of the associated catchment. In the context of the model (4)-(5), either the information or covariance matrix permits specification of which configuration (or model orders $(n, m)$) can be identified unambiguously (or with specified uncertainty) from the data. Over specification of either model order as $n'$ or $m'$ will result in an $I$ matrix which is not well conditioned because of a lack of cross correlation between $\hat{x}_{k-n}$ and lagged values of $u_k$ or between $u_{k-m}$ and lagged values of $\hat{x}_k$. These cross correlations form some of the off-diagonal elements in (10). The greater the overspecification of $n$ and $m$, the worse the conditioning of the information matrix or tendency to singularity, and hence the larger the elements in the covariance matrix. Underspecification of either model order will result in a substantially higher value of $\delta^2$ than for any model orders $(n' \geq n, m' \geq m)$. In practice, the values of $\delta^2$ (or the coefficient of determination $D = 1 - \delta^2/\text{Var}(q_k)$) can be expected to decline (rise) to a plateau as the model order increases while standard measurement norms of the covariance matrix become unacceptably large.
square of its estimated mean value. It has been used by Jakeman et al. [1989, 1990], for example.

The covariance matrix in (11) is dependent on (1) the input sequence of excess rainfall $u_k$, (2) the underlying response (model) parameters (because $x_k$ is the output of the model and $x_k$ and $u_k$ are transformed from $x_k$ and $u_k$ using $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_n$). (3) the variance of the combined data and model errors, $\sigma^2$, and (4) the sample size, $N$. For any temporal pattern of excess rainfall, response parameter values, and error level (or alternatively sample size), $P$ can be evaluated using (10) and (11) to determine the absolute minimum sample size (error variance) required to achieve some predetermined accuracy in the parameter estimates $a$ and hence in the linear dynamic response characteristics $\tau$.

2.3. Example Catchments

To answer our fundamental question, How complex a parameterization is warranted in a rainfall-runoff model?, the modeling framework was applied to a selection of catchments covering a range of scales and climatic conditions (Table 1). Except for the experimental Hydrohill catchment, pairs of catchments were chosen in close proximity but with differing responses so that model performance could be examined under similar climatic conditions (almost identical calibration periods) but different catchment descriptors.

The largest pair of catchments selected for analysis consisted of the upland subcatchments of the Murrumbidgee and Cotter rivers in the Australian Capital Territory about 50 km southwest of Canberra. Their centers are about 10 km apart. The stream gauge on the larger Orroral Valley catchment is at 1090 m. Land falls steeply in both catchments from high ridges. Both contain a large diversity of vegetation, mainly native eucalypts, but the cover at Licking Hole was completely burned by brushfire just prior to the period of analysis. Soils are deep in the two, but quick flow above a semipermeable layer (R. Knee, personal communication, 1992).

<table>
<thead>
<tr>
<th>Catchment</th>
<th>$A$, km$^2$</th>
<th>Precipitation, mm/yr</th>
<th>Average Daily Temperature, °C</th>
<th>Annual Yield, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orroral Valley (Australian Capital Territory)</td>
<td>89.6</td>
<td>1101</td>
<td>18.9</td>
<td>27</td>
</tr>
<tr>
<td>Licking Hole (Australian Capital Territory)</td>
<td>20.6</td>
<td>1426</td>
<td>18.9</td>
<td>53</td>
</tr>
<tr>
<td>Monachyle (Balquhidder, Scotland)</td>
<td>7.7</td>
<td>2953</td>
<td>5.4</td>
<td>93</td>
</tr>
<tr>
<td>Kirkton (Balquhidder, Scotland)</td>
<td>6.9</td>
<td>2795</td>
<td>5.4</td>
<td>88</td>
</tr>
<tr>
<td>Watershed 36 (Coweeta, North Carolina)</td>
<td>0.49</td>
<td>2012</td>
<td>19.6</td>
<td>64</td>
</tr>
<tr>
<td>Watershed 34 (Coweeta, North Carolina)</td>
<td>0.33</td>
<td>2012</td>
<td>19.6</td>
<td>46</td>
</tr>
<tr>
<td>Hydrohill (Nanjing)</td>
<td>0.00049</td>
<td>950</td>
<td>19.5</td>
<td>na</td>
</tr>
</tbody>
</table>

Here na denotes not available.

The Kirkton has forest and grass vegetation, while the Monachyle is covered with heather and grass. Jakeman et al. [1993b] have analyzed precipitation-streamflow data from these catchments before and after experimental land use changes were effected to examine the associated changes in the hydrological response. Detailed descriptions of the catchments and the experimental Balquhidder program are given by Blackie [1987] and Johnson [1988].

The smallest pair of catchments selected was from the Coweeta Hydrological Laboratory in the United States. Coweeta is located in the Nantahala Mountains of western North Carolina. Watershed 36 is a high-elevation, steeply sloping catchment with shallow soils, and a high annual yield and a large proportion of quick flow [Swift et al., 1988]. Watershed 34 is a mid-elevation catchment with somewhat deeper soils and, consequently, substantially more delayed flow. Details of the physical characteristics of the Coweeta catchments are given by Swank and Crossley [1988].

Hydrohill is a small experimental catchment of 490 m$^2$ near Nanjing, China. It has been used to investigate isotopic heterogeneity in subsurface waters by Kendall and Gu [1991]. According to them, the catchment was constructed with a concrete aquiclude consisting of two intersecting slopes with 14° gradients overlaying bedrock. Impermeable walls enclose the catchment on the top and sides. The aquiclude was covered with 1 m of a silty loam that was free of concretions. The bulk density was adjusted to approximate the natural soil profile. Grass was then planted over the surface. After three years of settling, a drainage trench was dug at the intersection of the two slopes and the water-sampling instrumentation was installed.

Five troughs, each 40 m long and constructed of fibreglass, were installed longitudinally in the trench. These troughs were stacked on top of each other to create a set of long zero-tension lysimeters. Each trough has a 20 cm aluminum lip that extends horizontally into the soil layer to prevent leakage between layers. Waters collected in each trough pass through V-notch weirs where discharge is continuously monitored and recorded ... the uppermost trough collects rain; the next lower trough collects surface runoff. The next three troughs collect subsurface flow from soil layers spanning the depths of, respectively, 0–30 cm, 30–60 cm, and 60–100 cm.

For each catchment pair, that is, excepting Hydrohill, one year of excess rainfall and streamflow values was used to estimate model structure and the associated parameter val-
values. For Hydrohill, analysis was performed on a storm covering a period of about 34 hours with data spaced at 6 min. For all catchments analyzed, the corresponding sample sizes were sufficient to identify the appropriate configuration of linear storages and their approximate parameter values.

3. Results

For a broad range of catchments we have found, as we have for the seven catchments in this paper, that the most commonly identified configuration is the one in Figure 1 of two storages in parallel driven by excess rainfall. This configuration of the linear module identified is the same for any values of the nonlinear module parameters, $T_f$ and $f$, which yield reasonable fits to streamflow. Indeed it is the same if no nonlinearities are assumed and rainfall is treated as excess rainfall. However, only one storage may be identified if either base flow is absent or data $\{q_N, q_m\}$ are sampled over a coarse enough time interval. From monthly data, only one storage was identified by Littlewood and Jakeman [1992] for the Thames Basin at Kingston, and the match between observed and model flow was excellent.

3.1. The Paired Catchments

Identification results. A general pattern emerges from the results of the modeling approach applied to a year of daily data from each of the six catchments (Tables 2-4). In terms of $D$, the single-storage and two-series storage configurations lead to lower values than the two-parallel storage configuration (with the exception of the Monachyle catchment). The former also have fewer parameters, two and three, respectively, compared to four for the latter. In the case of the Monachyle, the slow flow volume fraction is so small ($<0.1$ as will be seen in Table 5) that the former configurations suffer no reduction in $D$ value by not fitting the low recessions. The slow flow volume fraction at Kirkton is also small enough (estimated as 0.15) to yield only small differences in $D$ among the different configurations.

With all six catchments, more complex configurations than two parallel storages yield no substantial improvement in $D$ values, less than 1.3%, and they may yield a substantial decrease. The average relative parameter error (ARPE) of the single-storage and two-parallel storage configurations is orders of magnitude lower than for the other configurations. Therefore the results show that a rainfall-runoff model configuration with more than two storages is not warranted by the data. If larger configurations are fit, the uncertainty or ambiguity in parameter estimates is exceedingly high. If configurations involving fewer parameters are estimated, the fit to streamflow for all the catchments is visually inferior, especially during long recessions, and is generally manifest as lower $D$ values.

Qualitative relation of DRC values to physical catchment descriptors. The dynamic response characteristics of the linear module for each of the six catchments, derived directly from the parameter estimates $a$, show a variability reflecting our wide choice of catchment types (Table 5). A comprehensive interpretation of the estimated DRC values is beyond the scope of this paper, but some comments are warranted. Note that size of the catchment bears little relation to the time constants of the quick and slow components. Orroral Valley, the largest catchment by far, has a much smaller $T_f$ (faster quick response) than the next three largest catchments. Licking Hole, which is much larger than both Coweeta catchments, has a smaller $T_f$ (faster slow response) than these. If DRCs are related to physical catchment descriptors, as suggested by Jakeman et al. [1992], more than catchment size is involved.

Qualitatively, some of the controlling differences between catchments in each pair can be argued. This is because the climatic forcing variables are basically the same and some of

<table>
<thead>
<tr>
<th>Configuration and Model Order $(n, m)$</th>
<th>Orroral Valley Catchment</th>
<th>Licking Hole Catchment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D$</td>
<td>ARPE, %</td>
</tr>
<tr>
<td>One storage (1, 0)</td>
<td>0.761</td>
<td>0.087</td>
</tr>
<tr>
<td>Two serial (2, 0)</td>
<td>0.768</td>
<td>3.366</td>
</tr>
<tr>
<td>Two parallel (2, 1)</td>
<td>0.787</td>
<td>0.108</td>
</tr>
<tr>
<td>Three serial (3, 0)</td>
<td>0.788</td>
<td>1.251</td>
</tr>
<tr>
<td>One serial, two parallel (3, 1)</td>
<td>0.787</td>
<td>945.474</td>
</tr>
<tr>
<td>Three parallel (3, 2)</td>
<td>0.785</td>
<td>365.680</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Configuration and Model Order $(n, m)$</th>
<th>Monachyle Catchment</th>
<th>Kirton Catchment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$D$</td>
<td>ARPE, %</td>
</tr>
<tr>
<td>One storage (1, 0)</td>
<td>0.686</td>
<td>0.492</td>
</tr>
<tr>
<td>Two serial (2, 0)</td>
<td>0.686</td>
<td>450.004</td>
</tr>
<tr>
<td>Two parallel (2, 1)</td>
<td>0.685</td>
<td>1.555</td>
</tr>
<tr>
<td>Three serial (3, 0)</td>
<td>0.690</td>
<td>29.589</td>
</tr>
<tr>
<td>One serial, two parallel (3, 1)</td>
<td>0.694</td>
<td>34.147</td>
</tr>
<tr>
<td>Three parallel (3, 2)</td>
<td>0.695</td>
<td>89.647</td>
</tr>
</tbody>
</table>
the physical catchment descriptors are similar as described in section 2.3. For the largest pair (the Australian Capital Territory catchments), terrain is of the same type, with the soils and vegetation being the obvious difference. The deep soils and reduced vegetation at Licking Hole provide relatively larger increases in storage (see separation in Figure 3 versus Figure 2) immediately following rainfall, slow flow contributing more to the hydrograph peaks, whereas an underlying less permeable soil layer at Orroral Valley provides a much flashier quick flow response ($\tau_q = 1.61$ compared to 3.79 days) and a slow flow storage which is less responsive to rainfall and recedes more than twice as slowly ($\tau_s = 97$ compared to 36 days). In the case of the Balquhidder catchments, the topography is similar, the main differences also being the vegetation and soils. The larger areas of upland peat may contribute to the greater proportion of quick flow volume in the Monachyle ($\nu_q = 0.93$ compared to 0.85 and see separation in Figure 5 versus Figure 4). For the Coweeta pair the differences in soils and topography are substantial. The steep slopes and thin soils at watershed 36 yield a much smaller proportion of slow flow volume than do the deep soils at watershed 34 ($\nu_s = 0.58$ compared to 0.80 and see separation in Figure 6 versus Figure 7).

### Correspondence between DRC values and stream hydrograph patterns

There is also a very reasonable correspondence between the patterns in the stream hydrographs in Figures 2-7 and the DRC values of the linear module, especially if one does the comparison within each pair where the climatic forcing variables are basically the same. The low value of $\tau_q$ for the Orroral Valley stream response accords well with its qualitative flashiness (Figure 2). Perhaps not so easy to appreciate from the figures are the similar relative volumes of quick (and hence slow) flow throughput for Orroral Valley and Licking Hole (Figures 2 and 3). For Orroral, the long slow flow time constant $\tau_s$ of almost 100 days and the flashiness of quick flow evidently ensure a strong contribution of slow flow to total stream volume despite the relatively low magnitude of slow flow during rainfall events. For the Balquhidder catchments, Monachyle shows a flashier quick response and a smaller slow flow volume than Kirkton (Figures 4 and 5). The DRC values (Table 5) confirm this view. But it is more difficult from inspection of the stream hydrographs to appreciate the identified longer response times of quick (and slow) components for Coweeta watershed 34 versus 36 (Table 5). This is partly due to the obviously larger slow flow volume for

### TABLE 5. Dynamic Response Characteristics of the Linear Module and Minimum 90% Confidence Intervals for the Seven Catchments

<table>
<thead>
<tr>
<th>Catchment</th>
<th>$\tau_q$, days</th>
<th>$\tau_s$, days</th>
<th>$\nu_q$, fraction</th>
<th>$\nu_s$, fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orroral Valley</td>
<td>1.61 (1.45, 1.78)</td>
<td>97 (-308, 458)</td>
<td>0.62 (0.34, 0.96)</td>
<td>0.38 (0.02, 0.66)</td>
</tr>
<tr>
<td>Licking Hole</td>
<td>3.79 (3.44, 4.18)</td>
<td>36 (25, 64)</td>
<td>0.63 (0.56, 0.69)</td>
<td>0.37 (0.30, 0.43)</td>
</tr>
<tr>
<td>Monachyle</td>
<td>0.85 (0.72, 1.00)</td>
<td>39 (-102, 86)</td>
<td>0.93 (0.81, 1.07)</td>
<td>0.07 (-0.07, 0.20)</td>
</tr>
<tr>
<td>Kirkton</td>
<td>1.35 (1.18, 1.54)</td>
<td>46 (-176, 251)</td>
<td>0.85 (0.73, 0.99)</td>
<td>0.15 (0.01, 0.28)</td>
</tr>
<tr>
<td>Coweeta 36</td>
<td>2.32 (2.07, 2.64)</td>
<td>50 (42, 61)</td>
<td>0.42 (0.38, 0.46)</td>
<td>0.58 (0.54, 0.62)</td>
</tr>
<tr>
<td>Coweeta 34</td>
<td>3.45 (2.97, 4.12)</td>
<td>69 (64, 75)</td>
<td>0.20 (0.18, 0.22)</td>
<td>0.80 (0.78, 0.83)</td>
</tr>
<tr>
<td>Hydrohill</td>
<td>0.0054 (0.2417)</td>
<td>0.77 (0.77)</td>
<td>0.23 (0.23)</td>
<td>0.83 (0.83)</td>
</tr>
</tbody>
</table>
Fig. 2. Rainfall (millimeters), model fit, residuals and total/slow separation for daily streamflow (cubic meters per second) at Orroral Valley in 1983–1984.

Fig. 3. Rainfall (millimeters), model fit, residuals and total/slow separation for daily streamflow (cubic meters per second) at Licking Hole in 1983–1984.

Fig. 4. Rainfall (millimeters), model fit, residuals and total/slow separation for daily streamflow (millimeters) at Monachyle in 1985–1986.

Fig. 5. Rainfall (millimeters), model fit, residuals and total/slow separation for daily streamflow (millimeters) at Kirkton in 1985–1986.
watershed 34 obscuring direct appreciation from the stream hydrograph of the rate of quick response.

Uncertainties in DRC values. A measure of the uncertainties for estimated mean values of the DRCs for the linear module is also provided in Table 5. These were computed by Monte Carlo sampling of the estimated covariance matrix of $a$, calculating the ensemble of $\tau$ values using (6)–(9) and rejecting the largest and smallest 5% of values to obtain 90% confidence intervals. The table illustrates that the underlying nature of a catchment’s response has a strong effect on the accuracy with which the properties of that response can be calculated. Notice the large uncertainties on $T_s$, $v_q$, and $v_s$ for Kirkton, Monachyle, and Orroral. These large uncertainties are principally due to the low magnitude of slow flow at many time steps. Absolute errors in time series calibration data and model have a larger relative effect on slow flow magnitude, swamping the underlying slow flow signal in the streamflow series to a greater degree.

There is also an effect which compounds these uncertainties. The absolute data errors are probably larger in these three catchments than in the others examined in the paper. For the Balquhidder pair, some precipitation in the winter months occurs as snowfall, and its delayed transfer to excess rainfall as it melted was not attempted (see the model’s premature response to snowfall in early January in Figures 4 and 5). In Orroral Valley the rain gauge, being situated in the largest by far of our seven catchments, will not yield as representative a measure of incident rainfall over the catchment.

3.2. The Hydrohill Catchment, China

Rainfall time series data $r_N$ from a storm on July 5, 1989, were analyzed by C. Kendall et al. (manuscript in preparation, 1993) using model (4) separately with discharge measurements $q_t^{(0)}$, $q_t^{(30)}$, $q_t^{(60)}$ and $q_t^{(100)}$ and various additions of these (e.g., $q_t^{(0+30)} = q_t^{(0)} + q_t^{(30)}$ for each $k$), employing an obvious notation to represent discharge collected at each trough. The time step selected was 0.1 hours, and analysis was undertaken from 8.7 hours after the storm commenced, still leaving almost 25 hours of record ($N = 250$) before surface discharge ceased. Deletion of the first 8.7 hours of the storm data allows omission from the analysis of some missing discharge measurements in that early period. As will be seen, rainfall after this period, from time steps 88 to 337, can be treated as excess rainfall. That is, $u_k = r_k$ is set for this catchment, allowing omission of the nonlinear module.

Transfer function model identification was applied (C. Kendall et al. (manuscript in preparation, 1993) give detailed results and compare the quick flow–slow flow separations with chemical separations) to the following discharge time series: $q_t^{(0)}$, $q_t^{(30)}$, $q_t^{(60)}$, $q_t^{(100)}$, $q_t^{(0+30+60+100)}$, $q_t^{(0+30+60+100)}$, and $q_t^{(0+30+60+100)}$, $N = 250$. The resulting model fits were found to be credible in all cases. Figure 8 shows the fit to total discharge, for example, including simulation of rainfall through the model for the first 87 time steps not used for model calibration. The identified configurations of the different models (omitting the subscript $N$) were a single storage for $u$ with $q^{(0)}$, $q^{(30)}$, $q^{(60)}$, and $q^{(0+30)}$.
ity in rainfall-runoff models parameterized using only data on

needed to fit streamflow well and to separate a slow flow
directly, only a small number (four) of parameters are
even in cases where several runoff components are observed
configurations also carry high parameter variances. Thus
accounting for the variance of observed discharge. These
in

D

fied lead to little improvement (of the order of 1% in
parameters in higher-order configurations than those identi-
tification of more than a small number of them. Estimation of
numerous components. Data and model errors prevent iden-
tification

and two storages in parallel for \( q^{(100)} \), \( q^{(60+100)} \), \( q^{(30+60+100)} \),
and \( q^{(0+30+60+100)} \).

A conclusion to be drawn from the Hydrohill results is that
higher-order configurations of linear storages can often be
well approximated by lower-order linear ones. The sum of
individual exponential decay responses (to a pulse of rain) in
different parts of the catchment can be well approximated by
the linear combination of a smaller number of exponential
terms. Using rainfall and total discharge time series data \( u, q^{(0+30+60+100)} \),
only two components, compared to five
separately identified components from measurements of four
individual troughs, are necessary to approximate 89% of the
variance of the total discharge at the outlet (Figure 8).

Of course, the response at any level is composed of
numerous components. Data and model errors prevent identi-
fication of more than a small number of them. Estimation of
parameters in higher-order configurations than those identi-
ﬁd lead to little improvement (of the order of 1% in \( D \)) in
accounting for the variance of observed discharge. These
conﬁgurations also carry high parameter variances. Thus even
in cases where several runoff components are observed
directly, only a small number (four) of parameters are
needed to fit streamflow well and to separate a slow flow
component.

4. DISCUSSION AND CONCLUSIONS

The main issue addressed in this work is that of complex-
ity in rainfall-runoff models parameterized using only data on
precipitation, air temperature, and streamflow. In this re-
spect, the actual model used is important only insofar as it
allows a statistical evaluation of the number of parameters
(taken by us to be an index of model complexity) that can be
logically supported by the data. The major result of our
study is that, for catchments in temperate climates but over
a tremendously wide range of scales, only a handful of
parameters can be reliably estimated from rainfall-runoff
data. This result confirms suggestions made by others using
conceptual or more physically based models [e.g., Beven,
1989].

Before proceeding with a discussion of the main infer-
ences drawn from the modeling work, it should be empha-
sized what our work does not do. First, there are many
reasons for building rainfall-runoff models. From a scientific
perspective, for example, insights derived from physically
based models, whether or not such models can be rigorously
parameterized using statistical methods, can be extraordi-
narily useful. Our work should not be taken as a recommen-
dation to replace in toto physically based models with
transfer function models such as used here.

Better deﬁnition of the spatial distributions of catchment
characteristics, moisture status, precipitation, and so forth is
likely to be achieved in the future as a result of promising
modern techniques. Such additional information may result
in an improved ability to identify some of the complex
hydrological mechanisms that drive catchment response.
For example, it has been argued that a knowledge of the
distribution of precipitation over a catchment would greatly
enhance our ability to simulate hydrographs [e.g., Wilson et
al., 1979]. This knowledge may become routinely available
from data obtained from sophisticated weather radar units
that are being installed around the world. The work pre-
ferred here should not be taken as an implied limitation on
possibilities for the future.

Our modeling approach may be utilitarian for such pur-
poses as flow forecasting. However, repeated applications in
this style will not, in and of themselves, lead to scientific
advances. We believe that there are two main paths to the
development of intuition and advances in the sciences that
deal with the natural environment. The first, building up
from a physically based understanding of local processes to
the large scales of natural catchments, is one with which
most scientists are comfortable. The second approach,
studying relationships at the larger scales with the aim of
discovering patterns that may subsequently be explained
using conventional scientific wisdom, is also valuable. It is
through the second approach that our modeling results
ultimately may be applied in a scientific sense. We do not
claim that we have arrived at this point. The aim of this
paper is to investigate the number of parameters that are
supported by rainfall-runoff data sets. Nevertheless, we
indulge later in this discussion in some speculation as to how
our approach might prove to be useful in empirical scientific
exploration.

4. Complexity

Complexity has been analyzed within a statistical frame-
work using a specific family of models which allows an
optional nonlinear rainfall loss model and any parallel and/or
series arrangement of linear reservoirs. Strictly, the results
are valid only under these conditions. However, we will
proffer some implications later for rainfall-runoff modeling in general.

For a broad range of catchments, we have found that the most commonly identified configuration for a rainfall-runoff model is two storages in parallel driven by excess rainfall as illustrated in Figure 1. (If base flow is negligible or the data sampling interval is too coarse, then only one storage may be identified). This four-parameter (two in the case of one identified storage) linear model obviously may require supplementation to allow for antecedent precipitation conditions and fluctuations in evapotranspiration. In humid catchments, this need add only a few more parameters.

It appears to be a robust conclusion that only a small number of conceptual storages is warranted. Orroral Valley with an area of about 90 km² is the largest catchment reported here, but similar results can be obtained on much larger catchments especially when one has a good spatial distribution of rain gauges to estimate-areal rainfall. For the 894-km² Teifi basin in Wales and the 767-km² French Broad River basin in North Carolina, for example, Jakeman et al. [1990] used hourly rainfall and streamflow for two small (0.72 and 0.34 km²) moorland catchments above Llyn Brianne, Wales, and found the two-parallel storage configuration. The Hydrohill results also show how individual discharge responses at different depths become identifiable as either only a quick component or a quick and a slow component when discharge data are aggregated in space at the “stream” outlet.

4.2. Linearity of Response

We have observed a predominant linearity in the response of watersheds over a large range of catchment scales, even if only a simple adjustment is made for antecedent rainfall conditions. The linearity assumption of unit hydrograph theory therefore seems applicable in temperate catchments and works just as well for slow flow as for quick flow. The major evidence for this is twofold. First, there is the ability of the exponential-like response of the transfer function approximation to the convolution integral to fit stream hydrograph recessions generally quite well, indicating that nonlinearities can be described by the transformation of rainfall to excess rainfall. Second, the Hydrohill results reinforce our observations of a predominantly linear response in catchments and add considerable justification for the modeling approach used in the paper to extract information. Despite considerable spatial heterogeneity in the surface wetting of the experimental catchment [see Kendall and Gu, 1991], the discharge response (without any nonlinear adjustment between rainfall and excess rainfall) is quite linear; and as found by C. Kendall et al. (manuscript in preparation, 1993), it is a linear response at all four troughs. Such linearity is perhaps surprising on so small a scale: The catchment is only 490 m² in area. Chapman [1992] has also observed linearity from analysis of event data using a nonparametric unit hydrograph method. Caroni [1986] used a unit hydrograph approach to consider variation of parameters with flow and, despite findings of nonlinearity, acknowledged that linear models may very well be good approximations.

4.3. Implications for Current Modeling Practice

The amount of information that can be gleaned from climatic time series of rainfall and temperature and from streamflow seems a great deal smaller than much of the current modeling practice, largely performed in temperate catchments, would indicate. With or without adjustment of rainfall for antecedent conditions, almost always, one quick flow component and one slow flow component are all that can be identified. Experiments by the authors with synthetic data under ideal conditions (no error in excess rainfall and Gaussian noise in streamflow) show that, for a large range of systems, three true components can be identified only for impractically low noise levels; more than two components can definitely be identified only with no noise on the streamflow series. It may be possible to identify three components in some catchments when the nature of the rainfall, dynamic response, and quality of records allow this. To find more than three components without knowledge of internal states, such as levels of hydraulically connected groundwater, would seem to be a rare achievement.

The ability to identify more than one response component with reasonable accuracy requires data sets of reasonable length and quality, as well as an algorithm that makes few assumptions about the nature of errors and that is numerically stable enough to extract the slow flow component. Even when just two components are identified, high parameter variance (see Table 5) and covariation exist for estimation from daily data of 1 year in length. Such uncertainty is exacerbated when the difference in relative volume between quick and slow flow response is large and absolute errors in rainfall or excess rainfall are high. For a wide variety of catchment types, the calibration of relatively complex models (conceptual, physically based or otherwise), solely from climatic and streamflow data, in the hope of understanding processes or inferring the compartmentalization of storage, is likely to be largely unprofitable (see also Beven [1989] and Hooper et al. [1988]).

The inclusion of spatial data on physical catchment descriptors (PCDs) would not appear to resolve the identifiability problem substantially. While the model (1)–(5) does not explicitly allow the incorporation of additional data on PCDs, the framework used permits speculation as to what additional information such data provide for model calibration. With physically based models there tend to be spatial elements (areas or volumes), representative of a level of physical homogeneity, upon which incident precipitation falls and/or to which flow from other elements may travel. Each element requires some minimal level of definition by parameters describing, for example, the rate at which water is transmitted and the potential throughput or storage. One can therefore regard such models as being nonlinear versions of our conceptual model. Indeed, precipitation enters the surface component of elements in parallel, and flow out of these storages may pass to other storages of the same element or storages of neighboring elements. The PCD data permit an attempt to specify the number and configuration of storages. However, the precipitation and streamflow data are still used in physically based models to estimate those key parameters of each storage related to the transport of
water. The results obtained in this paper help argue that,
based only independent of the scale of representative el-
ments, information on the flow needs to be obtained from

One way around this problem would be to assume
that characteristic hydrological properties are related to suitable
PCDs in a parametrically efficient way. For example, it may
be possible to obtain a simple relation between them wherein
each of the parameters associated with an element is de-
designed to have the same calibrated value over all elements or
a large subset of element types.

4.4. Opportunities

An ability to represent succinctly the response of a catch-
ment to precipitation and other climatic inputs proffers many
opportunities for enhancing our knowledge of hydrological
phenomena on a local, regional, and global basis. The transfer
function–unit hydrograph separation approach al-

An obvious research avenue is exami-

Acknowledgments. This work was undertaken while both au-
phases such as our wetness declination time constant \( \tau_w \),
temperature modulation factor \( f \), change in storage index
per unit rainfall \( c \), and the time constants and fractional
volumetric throughputs associated with each linear storage.
Such a parsimonious, yet effective and physically plausible
parameterization at catchment scale, may provide a common
basis for accumulating knowledge through collective appli-
cation of the associated identification procedure to rainfall-
streamflow data sets. An obvious research avenue is exami-
ination of the relationships between estimated DRC values
and physical catchment descriptors. Any such procedure
will need to account for systematic and random uncertainties
derived from rain gauge coverage errors and possible drifts
and shifts in stream stage-height rating curves. While the
amount and quality of time series data required to generate
sufficiently reliable DRC values for any catchment require
investigation, it is likely that there are a useful number of
catchments worldwide with adequate data.

REFERENCES

Beven, K., Towards a new paradigm in hydrology, in Water for the
Future: Hydrology in Perspective, IAHS Publ., 164, 393–403,
1987.

Blackie, J. R., The Balquhidder catchments, Scotland: The first four

Box, G. E. P., and G. M. Jenkins, Time Series Analysis, Forecast-

Caroni, E., Non-linearity and time-variance of the hydrologic re-
sponse of a small mountain catchment, in Scale Problems in
Hydrology, edited by V. K. Gupta et al., pp. 19–37, D. Reidel,

Chapman, T. G., Identification of a common unit stream and rainfall
loss patterns for a set of runoff events, Working Pap. 1992/4,
Centre for Resour. and Environ. Studies, Aust. Natl. Univ.,
Canberra, 1992.

Clarke, R. T., A review of some mathematical models used in
hydrology, with observations on their calibration and use, J.

Hooper, R. P., A. Stone, N. Christophersen, E. de Grobois, and
H. M. Seip, Assessing the Birkenes model of stream acidification
using a multisignal calibration methodology, Water Resour.

Hornberger, G. M., K. J. Beven, B. J. Cosby, and D. E. Sapping-
ton, Shenandoah watershed study: Calibration of a topography-
based, variable contributing area hydrological model to a small

Jakeman, A. J., and P. C. Young, Towards optimal modelling of
translocation data from tracer studies, paper presented at Simu-
lation Society of Australia Conference, Univ. of Queensland,
Brisbane, Aug. 27–29, 1980.

in a stream aquifer system, J. Application of catchment identifica-
tion to the Murray River, Water Resour. Res., 25, 2177–2185,
1989.

Jakeman, A. J., I. G. Littlewood, and P. G. Whitehead, Computa-
tion of the instantaneous unit hydrograph and identifiable compo-
nent flows with application to two small undrained catchments, J.

Jakeman, A. J., I. G. Littlewood, and H. D. Symons, Features and
applications of IHACRES: A PC program for identification of unit
hydrographs and component flows from rainfall, evapotranspira-
tion and streamflow data, in Proceedings of the 13th IMACS
World Congress on Computation and Applied Mathematics, vol. 4,
edited by R. Vichnevetsky and J. J. H. Miller, pp. 1963–1967,

Jakeman, A. J., G. M. Hornberger, I. G. Littlewood, P. G.
Whitehead, J. W. Harvey, and K. E. Bencala, A systematic
approach to modelling the dynamic linkage of climate, physical
catchment descriptors and hydrologic response components,

Littlewood, and P. G. Whitehead, Assessing uncertainties in
hydrologic response to climate at large scale, IAHS Publ., in

Jakeman, A. J., I. G. Littlewood, and P. G. Whitehead, An
assessment of the dynamic response characteristics of streamflow

Johnson, R. C., Changes in the sediment output of two upland
drainage basins during forestry land use changes, IAHS Publ.,

Kendall, C., and W. Gu, Development of isotopically heterogeneous
infiltration waters in an artificial catchment in Chuzhou, China,
paper presented at International Symposium on the Use of
Isotope Techniques in Water Resources Development, Int. At.

Littlewood, I. G., and A. J. Jakeman, Hydrograph separation into
dominant quick and slow flow components, paper presented at

modelling techniques on small undrained catchments, Water Resour.

Newsam, G. N., Measures of information in linear ill-posed prob-
Canberra, 1984.

Philip, J. R., Some remarks on science and catchment prediction, in
Prediction in Catchment Hydrology, edited by T. G. Chapman
and F. X. Dunin, pp. 23–30, Australian Academy of Science,
Canberra, 1975.

Pierce, D. A., Least squares estimation in dynamic disturbance
G. M. Hornberger, Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22903.
A. J. Jakeman, Centre for Resource and Environmental Studies, Australian National University, GPO Box 4, Canberra, Australian Capital Territory 2601, Australia.

(Received June 15, 1992; revised March 23, 1993; accepted April 1, 1993.)