HYDROLOGICAL PROCESSES, VOL. 11, 1131–1144 (1997)

A TEST OF TOPMODEL'S ABILITY TO PREDICT SPATIALLY DISTRIBUTED GROUNDWATER LEVELS

JAN SEIBERT,[1*] KEVIN H. BISHOP[2] AND LARS NYBERG[1]

[1]*Institute of Earth Sciences, Department of Hydrology, Uppsala University, S-752 36 Uppsala, Sweden*
[2]*Department of Forest Ecology, Swedish University of Agricultural Sciences, S-901 83 Umeå, Sweden*

ABSTRACT

TOPMODEL was calibrated to a small catchment using precipitation and runoff data. Acceptable fits of simulated and observed runoff were obtained during both the calibration and validation periods. Predictions of groundwater levels using this calibration did not agree well with observations at the 37 points within the catchment where groundwater levels were measured, including three locations with continuous recordings. Groundwater level observations at one single point in time, however, sufficed to calibrate new topographic–soil indices that improved the prediction of the local groundwater levels at the observed tubes. This suggests that spatially distributed calibration data are necessary to exploit reliably TOPMODEL's ability to predict spatially distributed hydrology. The mean or recalibrated transmissivity values at these 37 points differed from the catchment mean as determined by the precipitation–runoff calibration. Thus, while groundwater information can help in predicting groundwater levels at specific locations, increasing the number of local groundwater level measurements is not sufficient to improve the spatially distributed representation of subsurface flow by TOPMODEL for the catchment as a whole, as long as the groundwater information is not integrated in the precipitation–runoff calibration. © 1997 by John Wiley & Sons, Ltd.

Hydrol. Process., Vol. 11, 1131–1144 (1997).

(No. of Figures: 5 No. of Tables: 0 No. of Refs: 23)

KEY WORDS TOPMODEL; groundwater levels; model validation

INTRODUCTION

There is a great interest in developing spatially distributed catchment models of hydrological processes. Fully distributed, physical models of water flow, however, are complex, and the formidable data requirements of such models limit their utility. Topographic data may simplify the problem of spatially distributed modelling because topography is such an important factor in the spatial distribution of hydrological processes. The slope component of topography can serve as a surrogate for the lateral hydraulic gradient, which is often parallel to the ground surface, and the location within the catchment provides information on the amount of lateral flow passing that specific location. As such, topography, which is easy to measure compared with many other catchment parameters, might provide enough information for modelling a satisfactory level of spatial distribution without requiring prohibitive amounts of input data.

TOPMODEL (Beven and Kirkby, 1979) is a physically based, semi-distributed catchment model of runoff generation which uses topographic information in the form of an index that describes the tendency of water to accumulate and to be moved downslope by gravitational forces. While many runoff models succeed in predicting flow without the benefit of topographic information, the TOPMODEL approach is attractive because it can be used to predict the spatial distribution of water content, groundwater levels and lateral flows at the time step of the model. This feature of TOPMODEL has been exploited together with

* Correspondence to: Jan Seibert.

Contract grant sponsor: ENCORE.
Contract grant number: STEP-CT90-0113.

CCC 0885–6087/97/091139–14$17·50

Received 5 February 1996
Accepted 6 December 1996

geochemical studies (Robson *et al.*, 1992), surface–vegetation–atmosphere models (Famiglietti and Wood, 1994) and ecological simulations (Band *et al.*, 1993), where the geographical variation in water content is a central concern.

Intriguing as this spatially distributed hydrological information is, however, little has been done to test whether the information that can be gleaned from TOPMODEL actually coincides with the spatial distribution of flow and water content. Most of the efforts to validate the spatially distributed predictions of TOPMODEL have dealt with the extent of saturated areas (Beven and Kirkby, 1979; Beven *et al.*, 1984) that can be observed at the soil surface. Indeed, the original focus of TOPMODEL was on defining the extent of saturation excess overland flow.

In Fennoscandia and many other areas, runoff is controlled more by the dynamics of the water table than by the extension of saturated areas (Bishop, 1991). Therefore, observations of the water table are more important for testing the spatially distributed aspects of TOPMODEL in these areas. Jordan (1994) found that the use of the topographic index was not satisfactory for prediction of local groundwater levels at about 15 tubes in 3·6 ha catchment. Groundwater level data, however, are only just beginning to be exploited as a source of validation information. Therefore, much remains to be done to evaluate the ability of topography to serve as the basis for spatial distribution of subsurface processes in hydrological models, and to assess the ability of the TOPMODEL approach to couple hydrology to topography.

The present study uses an unusually comprehensive set of information on runoff, groundwater levels, topography and soil properties to test the ability of TOPMODEL to predict groundwater levels at 37 points within a small till catchment in south-western Sweden. The initial prediction of groundwater levels was made from a calibration of TOPMODEL arrived at using only precipitation and runoff data, together with certain physical properties of the catchment soils. These predictions did not agree well with the observations, but the prediction of local groundwater levels was improved by calibrating a local topographic index (alternatively a local transmissivity) from a single observation of groundwater level. The potential for using such spatially distributed groundwater observations to improve the calibration of TOPMODEL at the catchment scale are also discussed.

METHODS

Study site

The study was conducted on the small (6300 m^2) catchment G1 ROOF on the west coast of Sweden, near Lake Gårdsjön (58°N, 12°E; Figure 1). The catchment is afforested with 80–100-year-old Norway spruce (*Picea abies* Karst.) and some Scots pine (*Pinus sylvestris* L.). Elevation of the catchment ranges from 123 to 143 m above sea level, with slopes varying from 0 to 40%. The topography is characterized by a central valley with steep side slopes. The 1020 mm yr^{-1} mean precipitation between 1991 and 1994 has resulted in an average runoff of 326 mm yr^{-1}. A more complete description of the catchment can be found in Andersson *et al.* (1996).

Catchment topography was measured on a 5 m × 5 m grid net. Soil depths were determined with a steel probe at 240 uniformly distributed points *c.* 5 m from one another. Soil depth varied from 0 to 1·40 m, with a mean of 0·43 m. A gneissic granodiorite bedrock underlies these soils, with outcrops at several locations. The texture of the soils is dominated by the silt and sand fractions with each comprising 38% of the soil, while clay accounts for 8% and organic material for 15% (Nyberg, 1995b). The soil porosity was high, with a volume-weighted catchment mean of 50%. The volume of each soil horizon in the catchment, together with porosity and soil moisture retention properties has been calculated (Seibert, 1993; Nyberg, 1995a; Bishop *et al.*, 1996).

Both direct measurements of saturated hydraulic conductivity (K_s) on undisturbed soil samples in the laboratory and indirect inferences from groundwater level–runoff relationships indicated that K_s decreased with depth. Using the latter method, Nyberg (1995b) calculated a larger decrease in K_s with depth than indicated by direct measurements. This decrease could be approximated with an exponential function.

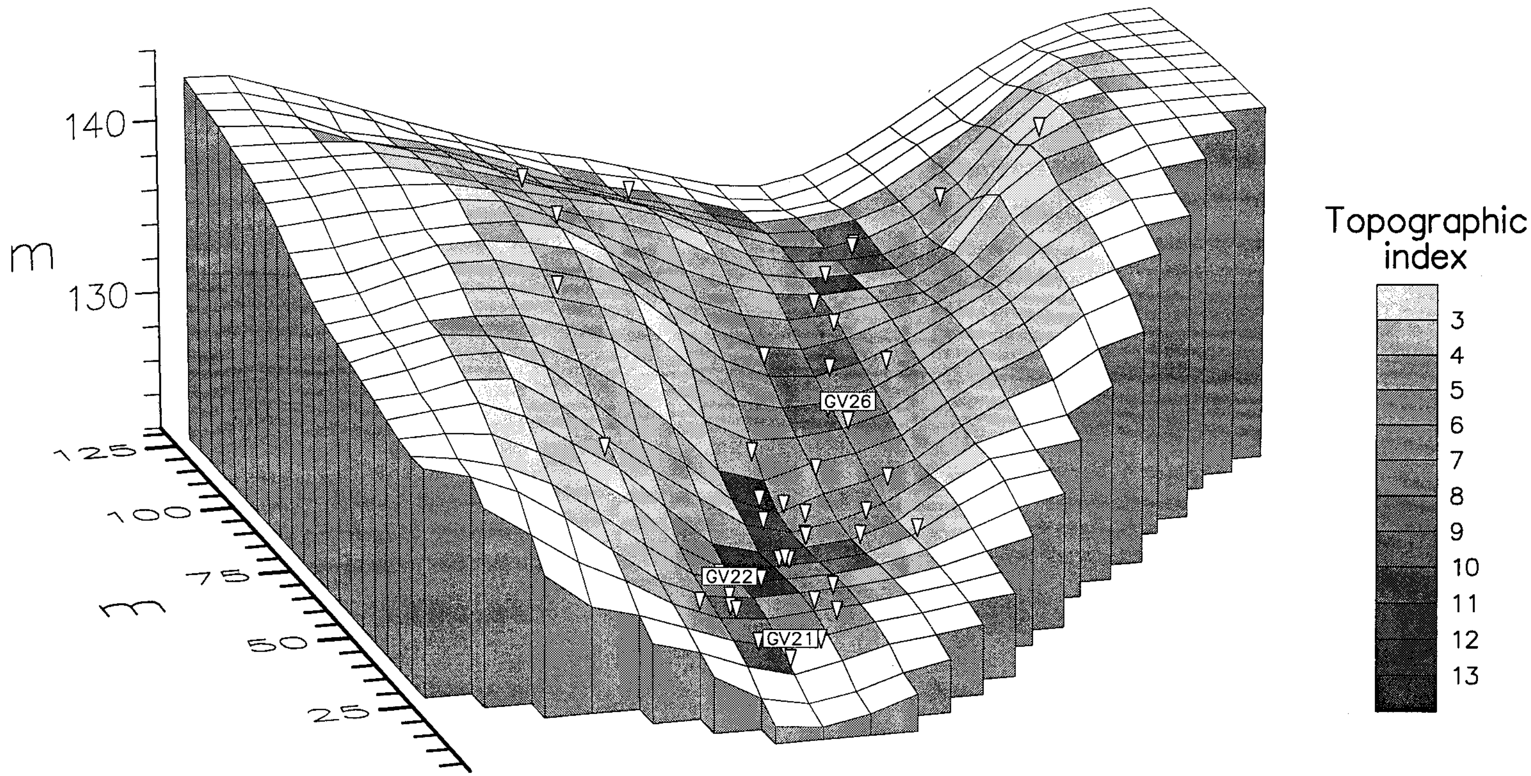

Figure 1. The 6300 m^2 G1 ROOF catchment with the location of manual and automatic (GV21, GV22 and GV26) groundwater tubes (▽). The values of the topographic index I_x on a 5 m × 5 m grid are shown across the topographic surface

In 1989 a project was started in the catchment to investigate the response of a forested ecosystem to a major reduction of acid input (Bishop and Hultberg, 1995). A transparent plastic roof was constructed below the tree crowns, at a height of 2–4 m, to intercept the acid deposition. Beneath the roof, water input to the catchment was simulated by an irrigation system. The quantity of sprinkled water corresponded to the observed throughfall, but offset in time, and with a constant intensity of 2·5 mm h^{-1}, which was the maximum output of the irrigation system. The sprinklers were installed near the soil surface (*c.* 0·5 m), so interception could be ignored.

Because there was no natural brook draining the small catchment, a pit was dug at the outlet in 1989 and dammed. The outflow from the dam was piped to a 0·7 m^3 tank where the discharge was recorded by counting when the tank filled before being automatically emptied. From April 1991 onwards, the rate of water level rise in the tank was also measured every 15 min.

Groundwater levels were measured manually in 37 groundwater tubes twice a month, and more intensively during some events, between April 1990 and September 1993. In three tubes, (GV21, GV22 and GV26; Figure 1) groundwater levels were also measured automatically with pressure transducers every 30 minutes. The slope of the groundwater table calculated from this network of groundwater tubes was similar to the topographic slope.

The maximum possible water storage in the soils with the entire catchment saturated to the surface is 300 mm. Between April 1991 and March 1994, the amount of water held within the catchment varied between 140 and 260 mm, with a mean of 206 mm (Bishop *et al.*, 1996). The hydrological response to precipitation events is rapid enough that the catchment can return to pre-event water storage and flows after two to three weeks. Once water is flowing from the catchment (when water storage is above 190 mm), the response to precipitation occurs within hours.

TOPMODEL concepts

A sufficient number of developments and variations on the basic principles of the original TOPMODEL have been made allowing to describe it now as an 'approach' rather than any specific model. In this approach there are two sources of runoff: outflow from subsurface saturated zones including return flow, q_{gw}, and overland flow caused by precipitation on saturated areas, q_{sat}. The latter term represents saturation excess overland flow (Dunne and Black, 1970). In the version of the model used in this study, q_{sat} is the entire amount of precipitation reaching the ground on the saturated areas of the catchment during the model time step. The total discharge from the catchment is the sum of these two components.

In early applications of TOPMODEL, the major role of subsurface flow was to determine the extent of saturated overland flow, and subsurface flow itself was only a minor contributor to runoff (e.g. Beven and Kirkby, 1979). In Fennoscandian catchments, and other areas where isotype hydrograph separation studies have shown large contributions of water displaced from the catchment (Rodhe, 1987), it is appropriate for the subsurface flow component to be the major source of runoff.

There are two major assumptions about lateral subsurface flow. The first is that the subsurface flow at any point in the catchment occurs in accordance with Darcy's law. Two subsidiary assumptions are made in the present application of TOPMODEL (and most others). The first is that the local topographic slope, $\tan \beta$, is equal to the local lateral hydraulic gradient. The second is that the lateral saturated hydraulic conductivity of the soil decreases exponentially with depth downwards into the soil profile. Both of these subsidiary assumptions were justified by investigators of groundwater dynamics and soil hydraulic properties on the G1 ROOF catchment (Nyberg, 1995b). The experience from other studies of runoff generation in Fennoscandia suggests that this conceptualization is also appropriate for many till catchments in the region because there is usually a strong increase in saturated hydraulic conductivity towards the soil surface (Lundin, 1982; Bishop, 1991).

These assumptions make it possible to define the transmissivity of the soil profile at some location x as a function of groundwater level, z_x [m below surface], as well as to calculate the lateral subsurface flow per unit width q_x[$m^2\ h^{-1}$], at a point from the local groundwater [Equation (1)]. In Equation (1), $T_{0,x}$ [$m^2\ h^{-1}$] is the transmissivity when the soil is saturated to the surface, and f [m^{-1}] determines the rate at which K_s decreases with depth.

$$q_x = \tan \beta_x T_{0,x} e^{-f z_x} \tag{1}$$

The second major assumption of TOPMODEL is that subsurface flow can be approximated by a succession of steady-state flow rates, and that the input to the saturated zone is uniform across the catchment.

The combined topographic effects of the upslope contributing area drained through a point per unit contour length, a_x [m], and the slope angle (β) at that point are represented by the topographic index, I_x [Equation (2)]. For steep slopes at the edge of a catchment, a is small and β is large, which yields a small value for I_x. High index values are found in areas with a large upslope area and a small slope, e.g. valley bottoms.

$$I_x = \ln\left(\frac{a_x}{\tan \beta_x}\right) \tag{2}$$

The groundwater level at any point within the catchment, z_x, can be related to the mean catchment water table $\bar{z}$ (m below surface) [Equation (3)] (Sivapalan *et al.*, 1987). In this equation, T_e is a mean value of transmissivity when the soil is saturated to the surface, and λ is the arithmetic mean value of the topographic index for the catchment. Total subsurface flow, q_{gw} (m^3 h^{-1}), from the catchment with an area A (m^2) is given using z_x as calculated by Equation (3) in Equation (1) and integrating over the line where substrate flow enters the stream channel [Equation (4)]. This defines a unique relationship between $\bar{z}$ and q_{gw}.

$$z_x = \bar{z} - \frac{1}{f}(I_x - \lambda + \ln T_e - \ln T_{0,x}) \tag{3}$$

with

$$\ln T_e = \frac{1}{A}\int_A \ln T_0 \, dA$$
$$q_{gw} = A T_e \, e^{-f\bar{z} - \lambda} \tag{4}$$

One major modification of the TOPMODEL version used in this study relative to several recently published versions (Beven *et al.*, 1995) concerning the relationship between groundwater level and catchment water storage should be noted. Usually in TOPMODEL one effective porosity is used that gives a linear groundwater level–water storage relationship. Studies of water storage in the G1 ROOF catchment (Seibert, 1993; Bishop *et al.*, 1996) found that there was an exponential relationship between groundwater storage and groundwater level owing to an increase of drainable soil volume towards the surface. Therefore, an exponential storage function derived from the water storage studies was implemented into TOPMODEL.

Application of TOPMODEL to the G1 ROOF catchment

TOPMODEL was applied to the G1 ROOF catchment with a one-hour time step using hourly precipitation and flow records. The Priestley–Taylor (1972) equation was used to calculate the daily potential evapotranspiration. The hourly evapotranspiration was interpolated from the daily Priestley–Taylor estimate using a sine-shaped distribution during daytime.

The value of I_x on a 5 m × 5 m grid (Figure 1) was calculated using the multidirectional flow algorithm proposed by Quinn *et al.* (1991).

In theory, the maximum saturated transmissivity ($T_{0,x}$) can vary spatially in TOPMODEL. Nevertheless, as in almost all applications of TOPMODEL, $T_{0,x}$ has been assumed to be spatially homogeneous (i.e. $T_{0,x}$ equals T_e at all points) for the runoff simulations, since there was no information available to justify any particular spatial distribution.

In the version of TOPMODEL used in this study there were six parameters to evaluate. Three of these were related to water storage, i.e. two parameters describing the exponential groundwater level–water storage relationship, and the maximum storage in the root zone (SRMAX, Beven *et al.*, 1995). These values were assigned using data from previous studies (Seibert, 1993; Bishop *et al.*, 1996) based on measurements of soil characteristics (Nyberg, 1995b). This left three parameters, T_e, f and $K^v_{0,s}$ to calibrate, with the latter being the vertical saturated hydraulic conductivity (m h^{-1}) at the soil surface (Beven *et al.*, 1995).

The objective function used to evaluate the goodness-of-fit provided by calibrations of these three parameters was the sum of squared differences between predicted and observed hourly runoff. The optimization was carried out manually, starting with model runs in which the parameter value combinations varied over very wide ranges, followed by examination of 'minimum error valleys' using more tightly spaced trials of parameter value combinations.

After calibration of the three model parameters using catchment precipitation and runoff data, the spatial distribution of the hydrology predicted by TOPMODEL was tested by comparing the instantaneous groundwater levels in 20–37 different groundwater tubes on 32 different occasions with the prediction of local groundwater levels using Equation (3) during the hour in which the observation was made. In these calculations, the value of I_x for the groundwater tube was set equal to I_x in the 5 m × 5 m grid cell where the tube was located. Errors in the simulation of local groundwater levels could be caused both by an erroneous relationship between mean and local groundwater level and by an inaccurately simulated mean groundwater level. To avoid the second type of error TOPMODEL was forced to simulate the observed runoff correctly. Therefore, the mean groundwater table simulated by TOPMODEL, $\bar{z}$, was calculated so that the simulated runoff from the groundwater store, q_{gw}, on each occasion equalled the observed runoff. The assumption that q_{sat} was insignificant on these occasions was reasonable because there were observations neither during irrigation events nor during extremely high flow conditions.

The observed values of z_x did not match the model predictions well. Therefore, an improvement of the predictions was attempted by changing the local values of $T_{0,x}$ and/or I_x. This was done by rearranging Equation (3) to relate I_x and $T_{0,x}$ to the difference between the z_x and $\bar{z}$ [Equation (5)] and defining a new, 'calibrated' index. Such calibration to give the best groundwater level prediction, however, decouples the index value from the catchment topography. This groundwater calibrated index is therefore referred to as the groundwater index, G_x. Assuming that $T_{0,x}$ is equal to T_e (i.e. ln $(T_e/T_{0,x}) = 0$), each observed groundwater level could be used to estimate a value of I_x for that site

$$G_x = f(\bar{z} - z_x) + \lambda = I_x + \ln \frac{T_e}{T_{0,x}} \tag{5}$$

Any single groundwater observation could be used to calculate G_x for each of the 37 groundwater tubes. These values of G_x which were thus obtained from one single occasion, were then used to simulate the groundwater levels in that tube on the other occasions when groundwater data were available (up to 31 occasions in the manual groundwater tubes). The success of the groundwater simulations yielded by this calibration was evaluated using the mean squared error (MSE) between observed and simulated groundwater levels.

RESULTS

The topographic index map agreed well with a subjective impression of the catchment (Figure 1). High values, which indicate wet areas, were found in the valley bottom and in the 'saddle' at the top of the central valley. The slopes had generally lower index values, but the steep section of the central valley through which the upper region of the catchment is drained had higher values of I_x. The mean value of I_x was 5·1.

The time-series of precipitation and runoff data used to calibrate T_e, f and $K^v_{0,s}$ extended from October 1991 to January 1992 (Figure 2). Modelled runoff was most sensitive to variations in values of T_e and f. This

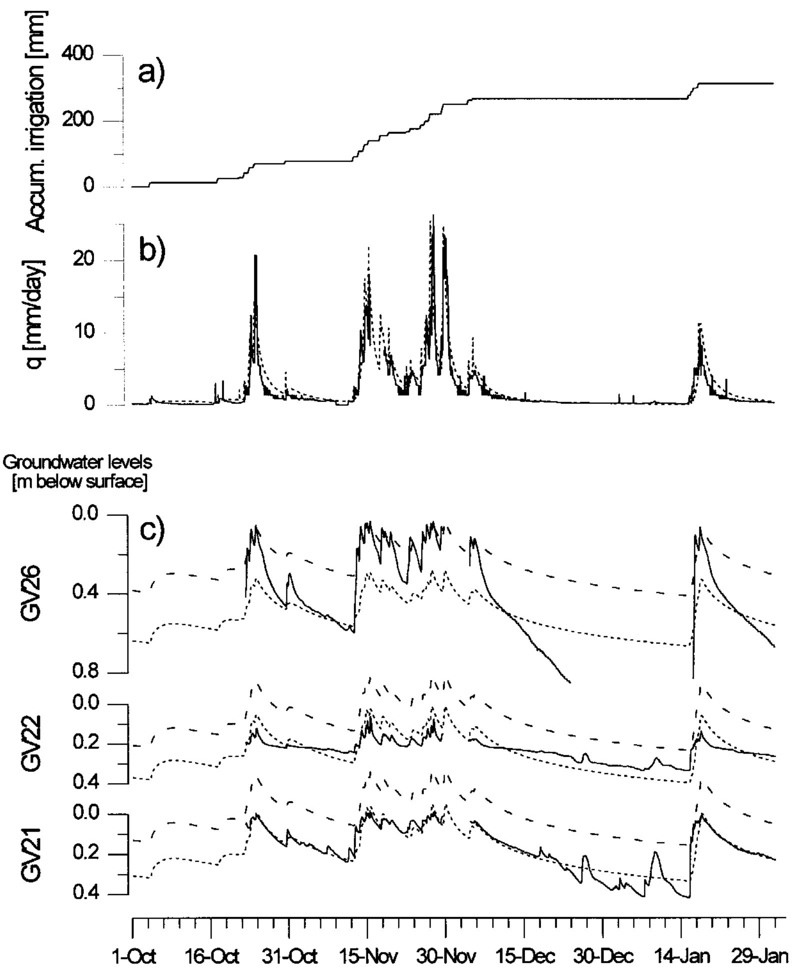

Figure 2. Time-series from the TOPMODEL calibration period (October 1991 to January 1992). (a) Accumulated sprinkling input. (b) Runoff with observed and simulated (dotted) hourly discharge. The unevenness in the hourly observed values arises from the gauging method based on filling of a tank. (c) The observed water levels at the automatically monitored groundwater tubes GV26, GV22 and GV21. The simulations based on the precipitation–runoff calibration are shown by dashed curves. Changing I_x made the simulated mean level equal to the observed mean level, but did not change the amplitude of the groundwater level fluctuations (dotted curves)

sensitivity made it easy to identify a small range of values that gave a good fit between simulated and measured runoff. The optimized value for T_e was 1·8 m^2 h^{-1}, 13 m^{-1} for f and 2·2 m h^{-1} for $K^v_{0,s}$. The decrease of conductivity with depth determined by f was much larger than measurements of saturated hydraulic conductivity (Nyberg, 1995b) suggest.

A validation run for the runoff predictions was performed using the period from October 1992 to January 1993. Predicted runoff compared well with hourly observed runoff. The Nash–Sutcliffe efficiency (Nash and Sutcliffe, 1970) was 0·77 for the calibration period and 0·69 for the validation period. Problems with the runoff data, i.e. the steps in the hydrograph caused by the gauging technique, made it difficult to achieve a better fit.

In the simulation, saturation excess overland flow, i.e. precipitation on to saturated surfaces, contributed 6% of the runoff volume. Up to 30% of the catchment was simulated to be saturated during irrigation events, which was reasonable given observed groundwater level–runoff relationships (Seibert, 1993), and observations of overland flow in the central valley (Nyberg, 1995a).

The simulation of local groundwater levels by TOPMODEL can be assessed in terms of three different characteristics: (1) the general trend (did the simulation go up and down in time with the observations); (2) the difference between the mean of the simulated and observed groundwater levels; as well as (3) the amplitude of the fluctuations in groundwater level.

The trends in the simulated hourly groundwater levels at the three automatic monitored locations corresponded to the observed trends (Figure 2). Thus the correlation between predicted and observed groundwater level was good, with values for the coefficient of determination r^2, of over 0·85 at all the automatic groundwater tubes. The simulated levels, however, were all too high, and the amplitude of simulated fluctuations in groundwater level at two of three sites differed from the observations.

The mean of the simulated levels in a groundwater tube was determined by the value of I_x (and/or $T_{0,x}$). By changing the value for I_x with one unit the local water table could be lowered/raised by about 0·08 m (using the calibrated catchment value for f, which was used to determine the difference between simulated values of $\bar{z}$ and z_x at sites with a given value of I_x). The I_x of 9·6 for the grid cell of GV22 resulted in an average groundwater level that was 0·16 m above the observed value (Figure 2). A change of I_x to 7·5 was needed to match the average simulated level to the average observed mean at that point. The situation was similar at the other automatic groundwater tubes. The I_x values for the grid cells were 10·6 for GV21 and 7·3 for GV26. Instead of these, values of 8·3 and 4·0, respectively, were needed to match the average of the groundwater level simulations to the observed averages.

The success of the simulation in reproducing the amplitude of the groundwater level fluctuation varied considerably between the three sites where high resolution time-series were available. At GV21 the simulated amplitude matched that of the observations well, while at GV26 the simulated amplitude was roughly half that of the observed amplitude, and at GV22 the simulated amplitude was approximately double that of the observations (Figure 2). Thus the overall impression of the groundwater simulation for the automatically monitored groundwater tubes is that it does not describe the fluctuation reliably beyond simply recreating the trends in groundwater levels.

The manual groundwater measurements at an additional 34 points within the catchment provided an opportunity to assess how widespread such errors were, although on fewer occasions (Figure 3). The correlation between the simulated groundwater level using the topographic index I_x and the observed groundwater level was weak for all 32 investigated occasions with 80% of the r^2 values between 0·1 and 0·35.

From Equation (3) it can be seen that the plots of I_x against the local groundwater levels should have a slope equal to $-1/f$. Therefore, the larger spatial coverage provided by the manual measurements also provided an opportunity to test the calibration of the f parameter for the catchment. The values for f calculated from groundwater level simulations using I_x ranged from 15 to 50, with a median of 22. This compares with the f value of 13 obtained from the precipitation–runoff calibration. The value of f suggested in this way tended to increase somewhat with the amount of catchment runoff on the measured occasion.

The use of observed groundwater levels to calculate the G_x for use in Equation (3) improved the prediction of the local water table significantly both in terms of MSE and r^2 (Figures 3 and 4). On average, the MSE decreased from 0·067 and 0·017 m^2 and r^2 increased from 0·23 to 0·73. In almost all cases G_x was lower than I_x, with an average difference of 1·8.

DISCUSSION

The Gårdsjön catchment is well suited for application of TOPMODEL in several important aspects. The assumption of an exponential relationship between soil depth and lateral saturated hydraulic conductivity was established experimentally, the slope of the soil surface approximates that of the groundwater surface and the dominant runoff mechanisms in the catchment are Darcian subsurface flow, complemented by saturation excess overland flow in the central valley bottom during periods of high flow. Therefore, it is appropriate that the precipitation–runoff calibration of TOPMODEL was able to predict the runoff response well. Nonetheless, the spatial distribution of groundwater levels was not well described in terms of

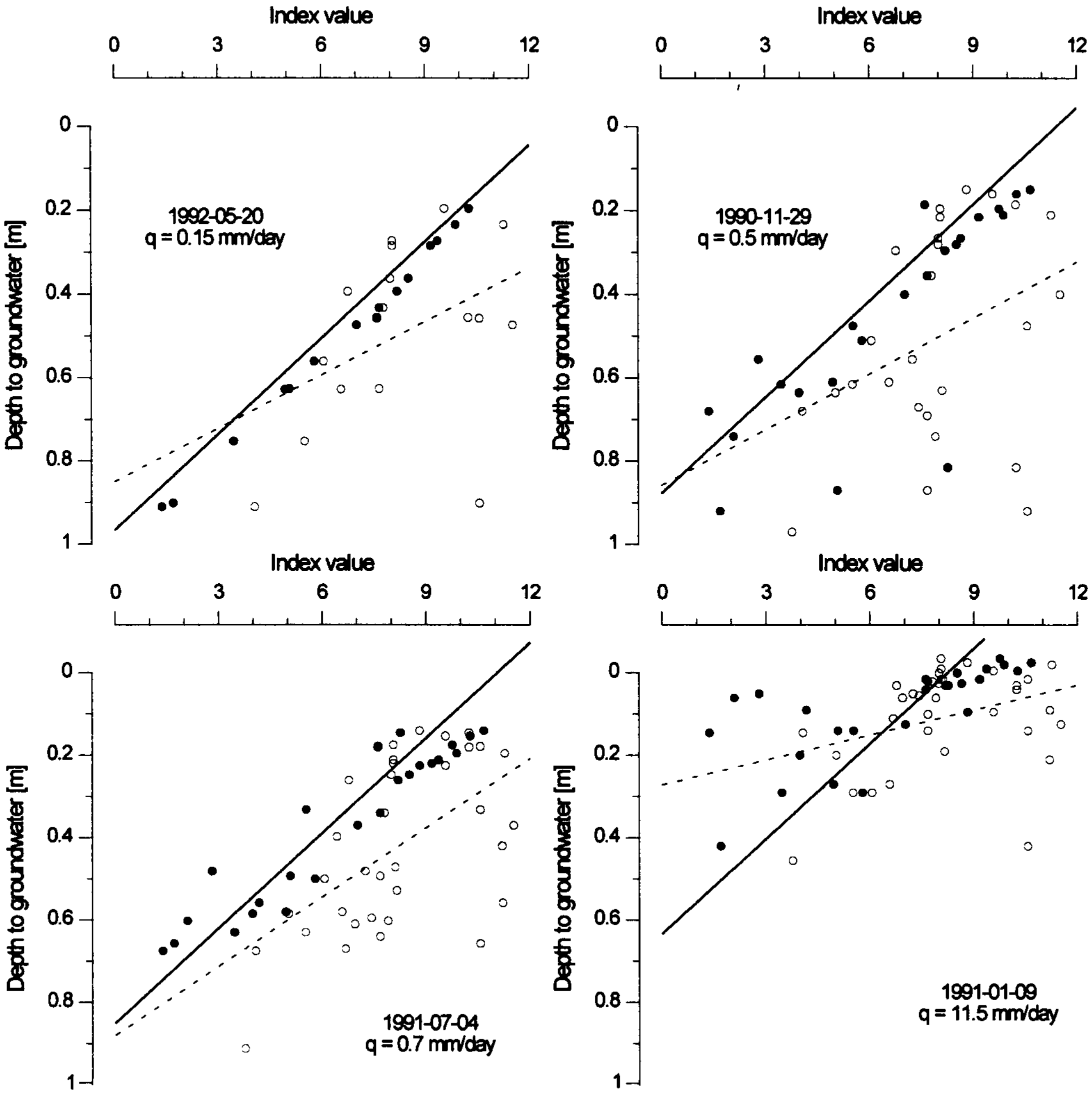

Figure 3. The observed groundwater levels plotted against the topographic index, I_x (open symbols) or groundwater index, G_x, based on the calibration using the measurement occasion on 4 April 1991 (filled symbols) which gave results typical for calibration on most of the measured occasions. The dashed line is the regression line between groundwater levels and I_x and the solid line shows the groundwater level simulated by the model at each index value. This makes the vertical distance from the solid line to each symbol equal to the error in the simulated groundwater level. Four examples of the 32 measurement occasions are shown, ranging from low flow (0·15 mm d^{-1}) to high flow (11·5 mm d^{-1}) conditions

absolute levels. The simulated and observed amplitudes of groundwater level fluctuations were also considerably different at two of the three sites where continuous observations were available.

TOPMODEL simulates the local groundwater levels from the assumption of steady-state flow rates. Therefore, the simulated levels rise and fall always in parallel over the entire catchment. In reality, groundwater levels, of course, do not behave like this and TOPMODEL can provide only an approximation of the real groundwater dynamics.

It was, however, possible to improve the prediction of the groundwater level at a particular point by using a single observation of groundwater levels to define a groundwater index, G_x, which replaces the index

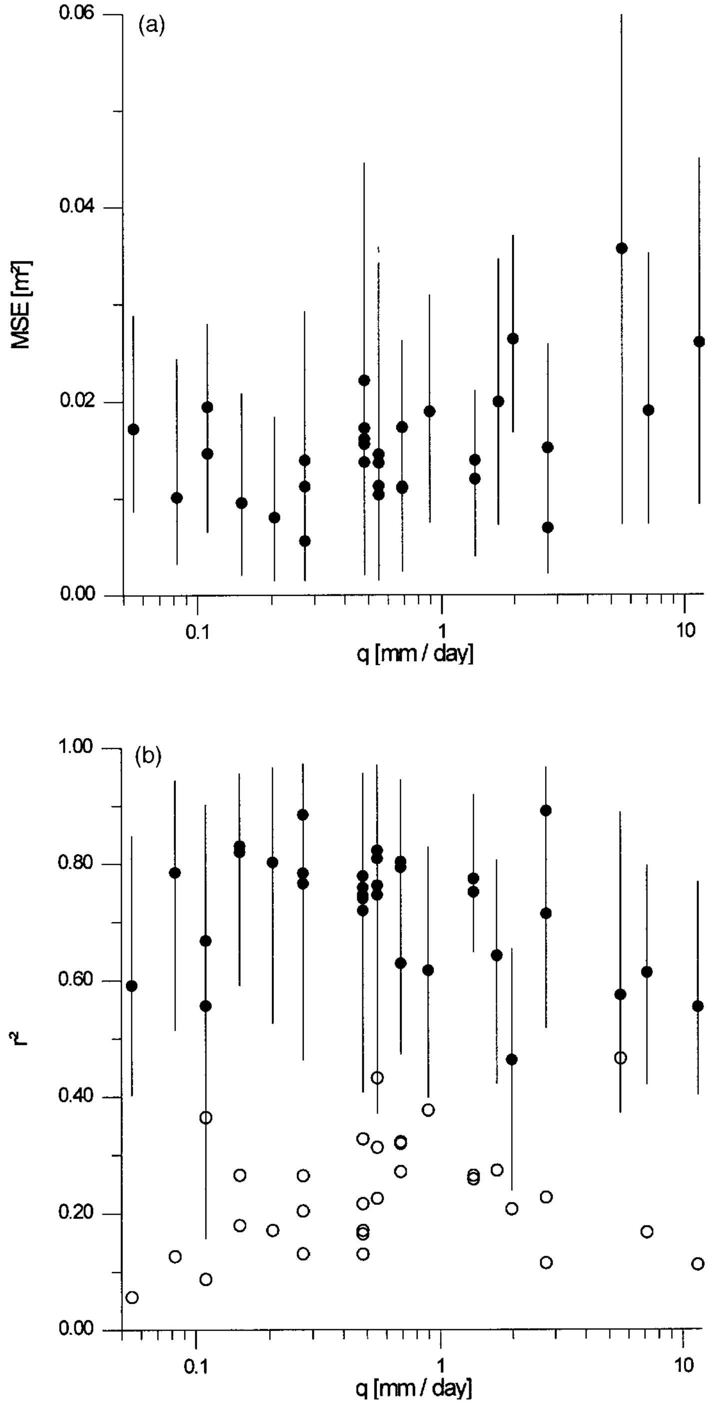

Figure 4. Goodness-of-fit. (a) The goodness-of-fit (MSE) of the simulated groundwater levels using G_x (filled symbols for the mean, together with vertical lines showing the 80% interval). The symbols represent the fit of all simulations based on the calibration done using occasions with a catchment runoff given on the horizontal axis. Using I_x the mean MSE was 0·067 m^2 with 80% between 0·047 and 0·093 m^2. (b) The goodness-of-fit (r^2) of the simulated groundwater level using I_x (open symbols) and G_x (filled symbols for the mean, together with vertical lines showing the 80% interval). The symbols represent the fit of all simulations of occasions with a catchment runoff given on the horizontal axis based on the different calibrations

calculated on the basis of topography. While this recalibration succeeds in improving the groundwater level predictions at specific points of interest within a catchment, it cannot be concluded that measurement of more groundwater levels and recalibration of local G_x values will lead to a calibration of TOPMODEL that better reflects the spatial distribution of subsurface hydrological processes in the catchment as a whole (i.e. where observations of groundwater levels are not available).

The difference between the local values of I_x and G_x can result from subgrid variability of I_x, spatial variations of transmissivity or errors in the assumptions underlying TOPMODEL, with the assumption of a uniform value of f, and a steady-state relationship between the groundwater levels within the catchment likely to be the most critical assumptions.

Subgrid variability is one contributor to the difference between G_x and I_x, since the exact value of I_x for the groundwater tube may differ from the value for the grid square; but analysis of groundwater levels from tubes in the same grid square suggested that this was a minor source of error. Different calculation methods for determining I_x can also give somewhat different values in a particular grid square. Subgrid variability or differences in calculation algorithms, however, would tend to give a random distribution of errors. At all 37 groundwater tubes, though, the calibration of G_x almost always yielded a value lower than I_x. At the automatic groundwater tubes, the value of G_x was at least one unit lower than I_x in any of the adjacent cells (equivalent to a difference in groundwater level of 0·08 m). This suggests that more than just subgrid variation or differences in the algorithm for calculating I_x was needed to explain the systematic reduction in the value of I_x needed to improve the simulation of the groundwater level at each tube.

One possibility is that not all of the upslope catchment area is contributing to the cell with a groundwater tube (i.e. a_x should be smaller). This is most likely to be the case during instances of low flow, when upslope areas are dry, but the local G_x was lower even when calibrated using high flow.

Local variation of $T_{0,x}$ is another potential contributor to the difference between I_x and G_x [Equation (5)]. The large spatial variation in soil depth at the 5 m × 5 m grid scale suggests a spatial variation in $T_{0,x}$. Even if spatial variability in $T_{0,x}$ exists, though, this does not necessarily compromise the use of TOPMODEL at the catchment scale with a uniform value of $T_{0,x}$. As long as the actual transmissivity is stochastically distributed over a certain area, a mean value of $T_{0,x}$ will give runoff results in TOPMODEL that are similar to those attained using the actual distribution of $T_{0,x}$. This is because the range of $a_x/\tan\beta_x$ is much larger than the expected range of $T_{0,x}$, i.e. the distribution of $T_{0,x}$ is of minor importance for the distribution of $\ln(a_x/T_{0,x}\tan\beta_x)$. The apparent lack of a spatial pattern in soil depth (Andersson *et al.*, 1996) makes it likely that there is also no spatial pattern in $T_{0,x}$, thus making it acceptable to use a mean value of T_0 at the catchment scale. The spatial variation in $T_{0,x}$, however, will make it more difficult to predict the groundwater level at specific points, and thus provides a rationale for the local recalibration of parameters.

Assuming that differences between G_x and I_x arise solely from local variation in $T_{0,x}$, the median value of $T_{0,x}$ calculated from each groundwater value available at a point is 3–5000 times higher than T_e (for 80% of the groundwater tubes $T_{0,x}$ values were 8–2000 times higher). This higher $T_{0,x}$ at the groundwater tubes may reflect, in part, a tendency to locate groundwater tubes where soils are deeper and $T_{0,x}$ is higher, but it still suggests that even increasing the number of points and using more representative locations for estimations of local $T_{0,x}$ values from groundwater level observations would not result in a distribution of $T_{0,x}$ that would agree with the mean value of T_e from the precipitation–runoff calibration. Thus, it is unlikely that the parameterization from the precipitation–runoff calibration can be improved simply by using local groundwater measurements to make local adjustments to parameters after having set the mean catchment values with the precipitation–runoff calibration, since there is an interplay between the different parameters and their spatial distribution.

This is particularly apparent in the case of f, which defines how rapidly the transmissivity (and lateral subsurface flow) decreases as the groundwater level falls. The value of f also determines the amplitude of the groundwater level fluctuations relative to changes in subsurface runoff. A higher value of f means that smaller changes in groundwater level result in greater changes in subsurface runoff.

The slope of the groundwater level observations against I_x in Figure 3 suggested that higher values of f were appropriate. If higher values of f had been used in the precipitation–runoff simulation, they would have reduced the amplitude in the simulation, thus bringing that aspect of the TOPMODEL predictions into better agreement with the observations at the automatic groundwater tube GV22, but into worse agreement for the other two automatic tubes.

Unlike $T_{0,x}$ the value of f cannot be defined locally without fundamental alterations to TOPMODEL's mathematical structure. This presents a dilemma when different values of f are needed to simulate the amplitude of groundwater level fluctuations at different locations, because improvements in amplitude prediction at one location may involve degradation of the simulation at other locations.

Simultaneous calibration of TOPMODEL using both groundwater level data and runoff data could help capitalize on the information in groundwater data to improve the spatial distribution of TOPMODEL predictions, by exploiting the linkages in catchment response to parameter values. Not only would this help constrain local values of T_0 and I, so that their distribution coincided with the mean T_e, but it would also make use of more of the information in groundwater data for defining an optimal f. However, this will not overcome the problems in the fundamental assumptions of TOPMODEL, such as a uniform value of f, or the assumption of a steady-state relationship between the groundwater levels in TOPMODEL.

A possible indication of a problem with the steady-state assumption is the structure of the errors in the prediction of groundwater level after calibration of G_x using a single groundwater level. There was a weak tendency that the r^2 decreased and the MSE increased as the runoff increased, i.e. as the catchment water level got closer to the soil surface at the time of the simulated event (Figure 4a), or the occasion used for calibration (Figure 4b). In other words, the quality of prediction and calibration declined at higher flow situations when the catchment was less likely to be at a steady state. Another possible explanation is increasing local variation in the soil properties close to the soil surface.

A more direct indication that the hydrological response of different areas of this very small catchment were not in a steady state with respect to water storage and discharge is the way in which the water tube could still be rising at one groundwater tube further up the central valley of the catchment at GV26, while the water table had already begun to fall near the outlet (GV21 and GV22) (Figure 5). Such non-steady state behaviour with respect to water storage and runoff is likely to be more pronounced in larger, less responsive catchments. As a consequence of the assumption of steady-state flow rates, TOPMODEL is only capable of simulating groundwater levels that rise and fall simultaneously. This is a major restriction and has been questioned before (Iorgulescu and Jordan, 1994; Barling *et al.*, 1994), because groundwater dynamics are often found to vary from place to place (e.g. Hinton *et al.*, 1993).

It will be important to assess the limitations set by the steady-state assumption underpinning TOPMODEL, as spatially distributed data is used to try to improve TOPMODEL's representation of catchment hydrology. Adjusting the time step of TOPMODEL to the response time of the catchment being modelled may be one way of mitigating any effects of the steady-state assumption on model performance. In this respect, it is worth noting that the ability of TOPMODEL to predict both runoff and groundwater level at Gårdsjön declined when TOPMODEL was run at a daily time step, since this obscured the close connection in time between precipitation, groundwater level and runoff. An optimal time step with respect to the steady-state assumptions could lie somewhere between hourly and daily time steps for the G1 ROOF catchment.

CONCLUSIONS

TOPMODEL may offer a means of providing spatially distributed hydrological information with input data requirements that are modest compared with most other models that seek to provide distributed hydrological information. The thin layer of unconsolidated deposits, the heterogeneity of soil depths and the marked small-scale topography make the G1 ROOF catchment well suited for application of TOPMODEL in some respects, but more problematic in others. The success of predicting local groundwater levels may, therefore, be different on other catchments.

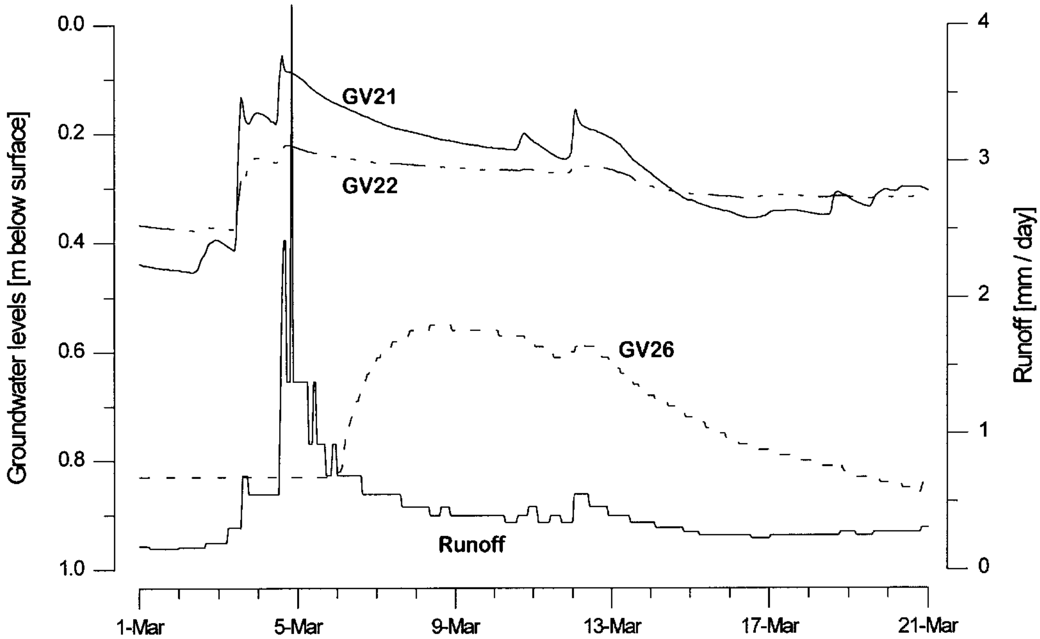

Figure 5. Conflicting trends in groundwater levels at GV21, GV22 and GV26. Time-series of groundwater levels at the automatically monitored groundwater tubes GV21 (near outlet), GV22 and GV26 at the top of the central valley for a period in March 1992

The experience from the G1 ROOF catchment, however, shows that despite a successful calibration of runoff predictions with TOPMODEL, the prediction of local groundwater levels was not satisfactory. This suggests that despite the use of topographical information, TOPMODEL may not be a reliable guide to the spatial distribution of hydrology when calibrated only from precipitation and runoff. On the other hand, a single groundwater observation was able to improve the predictions of local groundwater by substituting calibrated values of G_x for the local value for I_x. This approach may suffice for improving the average groundwater level prediction at specific points of interest, but not for the amplitude at that point, or the average level for the catchment as a whole. Better prediction of the spatial distribution of subsurface processes by TOPMODEL, however, might be achieved by integrating spatially distributed information on groundwater levels with precipitation–runoff data in the parameter calibration process. Optimal use of catchment area in the calibration process will help to show to what extent TOPMODEL can be relied on to provide the much sought after spatial distribution in hydrological simulations.

ACKNOWLEDGEMENTS

We thank the staff of IVL, and Ingvar Andersson and Filip Moldan in particular, for their assistance, as well as Allan Rodhe and two anonymous reviewers for their comments. This project was supported financially by the European Network of Catchments Organized for Research on Ecosytems (ENCORE, Contract STEP-CT90-0113), the Swedish Natural Science Research Council, the Swedish Environmental Protection Agency, National Power and PowerGen (UK), as well as the Swedish national power industry (Vattenfall and Elforsk).

REFERENCES

Andersson, I., Bishop, K. H., Borg, G. Chr., Giesler, R., Hultberg, H., Huse, M., Kjønaas, O. J., Moldan, F., Nyberg, L., Nygaard, P. Holm, Nystöm, U., and Stuanes, A. O. 1996. 'The covered catchment site: a general and specific description of physiography and vegetation of three small coniferous forest catchments at Gårdsjön, southwestern Sweden', in Skeffington, R. and Hultberg, H. (Eds), *Experimental Reversal of Acid Rain Effects: the Gårdsjön Project*. Wiley and Sons, London.

Bank, L. E., Patterson, P., Nemani, R., and Running, S. W. 1993. 'Forest ecosystem processes at the watershed scale: incorporating hillslope hydrology', *Agric. For. Meteorol.*, **63,** 93–126.
Barling, R. D., Moore, I. D., and Grayson, R. B. 1994. 'A quasi-dynamic wetness index for characterizing the spatial distribution of zones of surface saturation and soil water content', *Wat. Resour. Res.,* **30,** 1029–1044.
Beven, K. J. and Kirkby, M. J. 1979. 'A physically based variable contributing area model of basin hydrology', *Hydrol. Sci. Bull.,* **24,** 43–69.
Beven, K. J., Kirkby, M. J., Schofield, N., and Tagg, A. F. 1984. 'Testing a physically-based flood forecasting model (TOPMODEL) for three U.K. catchments', *J. Hydrol.*, **69,** 119–143.
Beven, K. J., Lamb, R., Quinn, P., Romanovicz, R., and Freer, J. 1995. 'TOPMODEL', in Singh V. P. (Ed.), *Computer Models of Watershed Hydrology*. Water Resource Publications, Colorado. Chapter 18.
Bishop, K. H. 1991. 'Episodic increases in stream acidity, catchment flow pathways and hydrograph separation', *PhD Thesis*, Cambridge University, Cambridge. 241 pp.
Bishop, K. H. and Hultberg, H. 1995. 'Reversing acidification in a forest ecosystem: the Gårdsjön covered catchment experiment', *Ambio*, **24,** 85–91.
Bishop, K. H., Nyberg, L., Seibert, J., and Rodhe, A. 1997. 'Water storage and hydrological dynamics on a till catchment near Gårdsjön, SW Sweden', *Wat. Resour. Res.*, submitted.
Dunne, T. and Black, R. D. 1970. 'Partial area contributions to storm runoff in a small New England watershed', *Wat. Resour. Res.*, **6,** 1296–1311.
Famiglietti, J. S. and Wood, E. F. 1994. 'Multiscale modeling of spatially variable water and energy processes', *Wat. Resour. Res.*, **30,** 3061–3078.
Hinton, M. J., Schiff, S. L., and English, M. C. 1993. 'Physical properties governing groundwater flow in a glacial till catchment', *J. Hydrol.*, **142,** 229–249.
Iorgulescu, I. and Jordan, J. P. 1994. 'Validation of TOPMODEL on a small Swiss catchment', *J. Hydrol.*, **159,** 255–273.
Jordan, J. P. 1994. 'Spatial and temporal variability of stormflow generation processes on a Swiss catchment', *J. Hydrol.*, **153,** 357–382.
Lundin, L. 1982. 'Soil moisture and ground water in till soil and the significance of soil type for runoff', *PhD Thesis*, Uppsala University, UNGI Report Nr. 56. 216 pp.
Nash, J. E. and Sutcliffe, J. V. 1970. 'River flow forecasting through conceptual models; part 1 — a discussion of principles', *J. Hydrol.*, **10,** 282–290.
Nyberg, L. 1995a. 'Soil- and groundwater distribution, flowpaths and transit times in a small till catchment', *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 97*, 35 pp.
Nyberg, L. 1995b. 'Water flow path interactions with hydraulic properties in till soil at Gårdsjön, Sweden', *J. Hydrol.*, **170,** 255–275.
Quinn, P., Beven, K. J., Chevalier, P., and Planchon, O. 1991. 'The prediction of hillslope flow paths for distributed hydrological modeling using terrain models', *Hydrol. Process.*, **5,** 59–79.
Robson, A., Beven, K. J., and Neal, C. 1992. 'Towards identifying sources of subsurface flow: a comparison of components identified by a physically based runoff model and those determined by chemical mixing techniques', *Hydrol. Process.*, **6,** 199–214.
Rodhe, A. 1987. 'The origin of streamwater traced by oxygen-18', *PhD Thesis*, Uppsala University, UNGI Report Series A Nr. 41, Uppsala, Sweden. 260 pp.
Seibert, J. 1993. *Water storage and flux in a micro-catchment at Gårdsjön, Sweden*. Institution of Geosciences, Hydrology Section, Uppsala University, Sweden. 46 pp.
Sivapalan, M., Beven, K. J., and Wood, E. F. 1987. 'On hydrologic similarity 2: a scaled model of storm runoff production', *Wat. Resour. Res.*, **23,** 2266–2278.