



Surface water flow in landscape models: 2. Patuxent watershed case study

Alexey A. Voinov *, Helena Voinov, Robert Costanza

Institute for Ecological Economics (IEE), University of Maryland, P.O. Box 38, Solomons, MD 20688-0038, USA

Received 27 July 1998; received in revised form 8 January 1999; accepted 25 February 1999

Abstract

We developed a system for modeling regional, spatially explicit hydrology in a way that allows integration with broader plant, nutrient, and socio-economic model components. The relatively coarse spatial and temporal resolution of these broader components relative to existing hydrologic models requires a different approach to hydrologic modeling. We developed a quasi-empirical model which still retains as much of the underlying process dynamics as possible. We tested the approach in the Patuxent watershed, Maryland, USA where we needed a time step of 1 day (compared to 1 h in most hydrologic models) and a spatial resolution of 1 km². The model was developed, calibrated and tested at several spatial and temporal scales and extensive testing shows that it performs at least as well as more complex hydrologic models, even using the 1 day and 1 km² resolution. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Landscape spatial modeling; Hydrology; Temporal and spatial resolution; DEM; Overland flow

1. Introduction

This is the second paper in the series on quasi-empirical modeling of surface runoff in complex landscape models (see Voinov et al. (1998)). Large landscape models are designed for analysis and management over areas that are orders of magni-

tude greater than those traditionally encountered in hydrological applications (Beven and Kirkby, 1979; Beasley and Huggins, 1980; Grayson et al., 1992). Besides, hydrology plays an auxiliary role in these applications with much further complexity embodied in the other modules involved. In this context the issues of scale and resolution become especially important and certain trade-offs in hydrologic complexity are required.

Model complexity can be measured in three dimensions: temporal, spatial and structural, the latter standing for the level of detail in describing

* Corresponding author. Tel.: +1-410-3267207; fax: +1-410-3267354.

E-mail address: voinov@cbl.umces.edu (A.A. Voinov)

the processes, variables and interactions in the system. For the model to be feasible we need to compromise between the three dimensions in order to keep the overall complexity within reasonable bounds. There are numerous examples of fairly complex models (Park et al., 1979; Parton et al., 1988 etc.), which are spatially confined to one point or several boxes linked with simplified spatial fluxing. On the other hand there are examples of models run as part of geographic information systems (GIS) (ESRI, 1992; Watkins et al., 1996; Ye et al., 1996) at a very high level of spatial resolution. These models are structurally simplified, performing analysis of only some aspects of landscape ecology. Their temporal resolution also tends to be more aggregated. This occurs because even with the supercomputing power at hand, the modeling process implies tedious analysis of model performance and calibration, which is possible only up to a certain level of complexity. As the modeling system becomes overcomplicated in all the three dimensions, its behavior gets as hard to predict and understand as the operation of the real-life landscape.

The Patuxent watershed covers an area of 2356.2 km². It stretches for about 150 km from the Piedmont area of the Appalachians to the Chesapeake bay (Fig. 1). The Patuxent landscape model (PLM) built for this watershed is an integrated ecological economic model and therefore needs to be fairly complex in order to cover a variety of issues of landuse development and landuse change in addition to interaction with ecological processes occurring at the landscape scale (Bockstael et al., 1995). Therefore it is especially important to find the right balance between the levels of complexity in each of the model components and in its spatial and temporal resolution.

In this paper we focus on the hydrologic module of the model, because hydrology is probably one of the most important transport mechanisms, both delivering the essential elements to the biota and removing the constituents that are in excess. It is therefore the component that links other modules spatially. Our previous experience with the Everglades landscape model (ELM) (Voinov et al., 1998), turned out to be insufficient for the area characterized by steeper hillslopes and almost no standing water. The flow patterns in this

watershed were quite different, requiring a revision of both the unit model and its spatial implementation.

Our major goal was to build a terrestrial analogue to the hydrologic module previously developed and calibrated for wetlands. In addition to that we analyzed the model performance at a variety of scales to see in what way they were different and what adjustments were required when switching from one resolution to another. The model has been constructed for two spatial resolutions and tested for several areas. Calibration and preliminary analysis was performed for a small subwatershed (22.5 km²) at a 200 m resolution. A larger subwatershed, that covered approximately half of the whole watershed area was then simulated at the same resolution. It became clear that this resolution was too fine and too time-consuming to suit the needs of the full integrated model on the whole watershed scale. Therefore we upscaled to a 1 km² resolution and analyzed the associated differences in model performance.

We explore an approach similar to the one used in other landscape models (Sklar et al., 1985), ELM in particular (Voinov et al., 1998; Fitz et al., 1999). The watershed is represented as a uniform grid of cells (Fig. 2), in each cell local vertical hydrological dynamics is modeled. The cells are then connected by a model of lateral horizontal flows of surface and ground water. The two sub-models are updated independently and run in sequence. This approach seems to be most appropriate for the integrated type of analysis that PLM is pursuing, when ecological and economic modules operating at a variety of spatial and temporal scales, need to be brought together and run in concert.

2. Hydrologic unit model

Based on the hydrologic module of the General ecosystem model (GEM) (Fitz et al., 1996) the unit model in PLM simulates vertical fluxes for a locality that is assumed horizontally homogeneous. GEM takes into account a variety of hydrologic processes and parameters including the following:

1. Transpiration associated with plant growth, physiology and relative humidity.
 2. Evaporation using pan evaporation estimates and pan coefficients.
 3. Rainfall based on precipitation data interpolated over seven stations.
 4. Seepage of water from that stored above the sediment/soil surface into that stored in sediment pore space (either in unsaturated or saturated storage).
 5. Roughness coefficient that depends on dynamic simulation of plant biomass, numeric density, and plant morphology.
- The traditional scheme of vertical water movement (Novotny and Olem, 1994), also implemented in GEM, assumes that water is fluxed

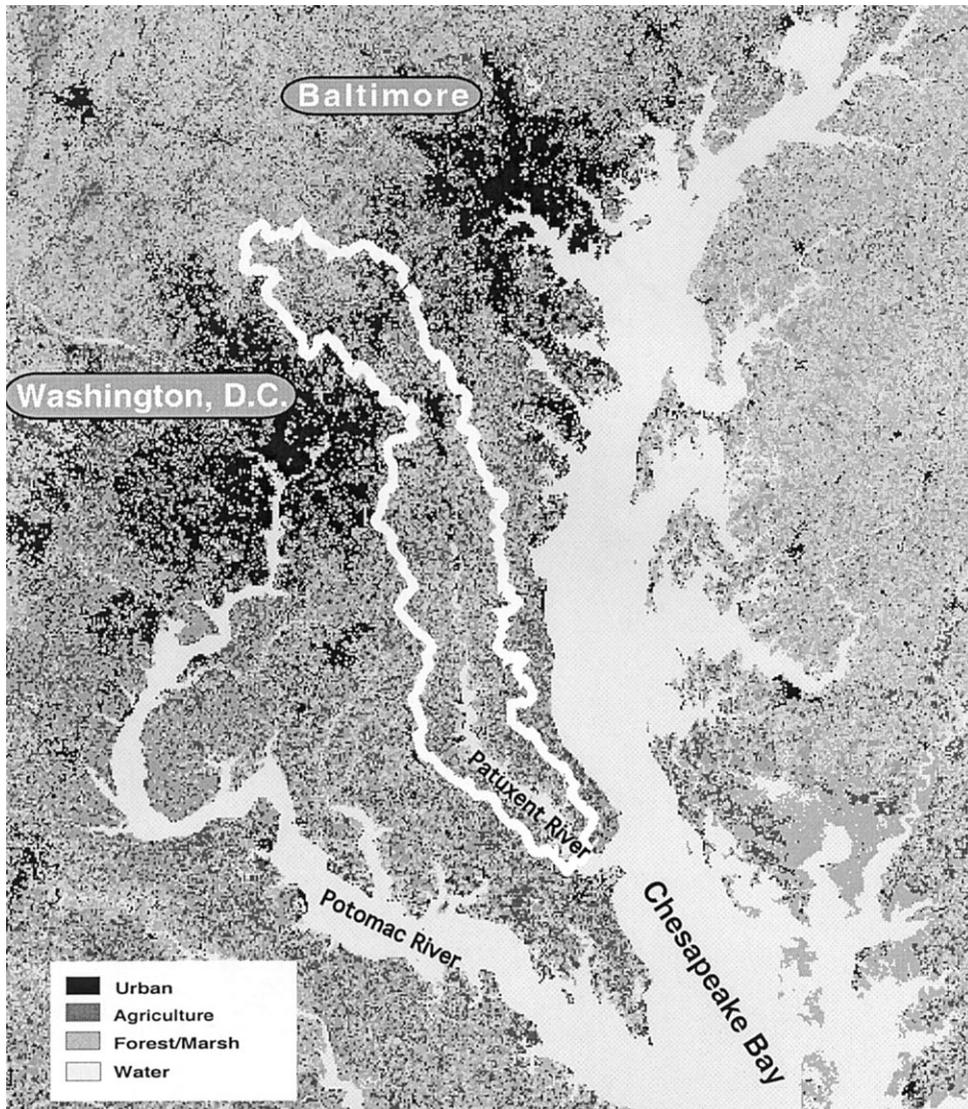


Fig. 1. Patuxent watershed location. The Patuxent River is one of the major tributaries of the Chesapeake Bay. Its drainage basin (approx. 2356 km²) lies between the Baltimore and Washington DC metro areas. The background map is based on 30 m cell resolution Landsat Thematic Mapper scenes for Chesapeake Bay Land Cover Classification 1988/1989.

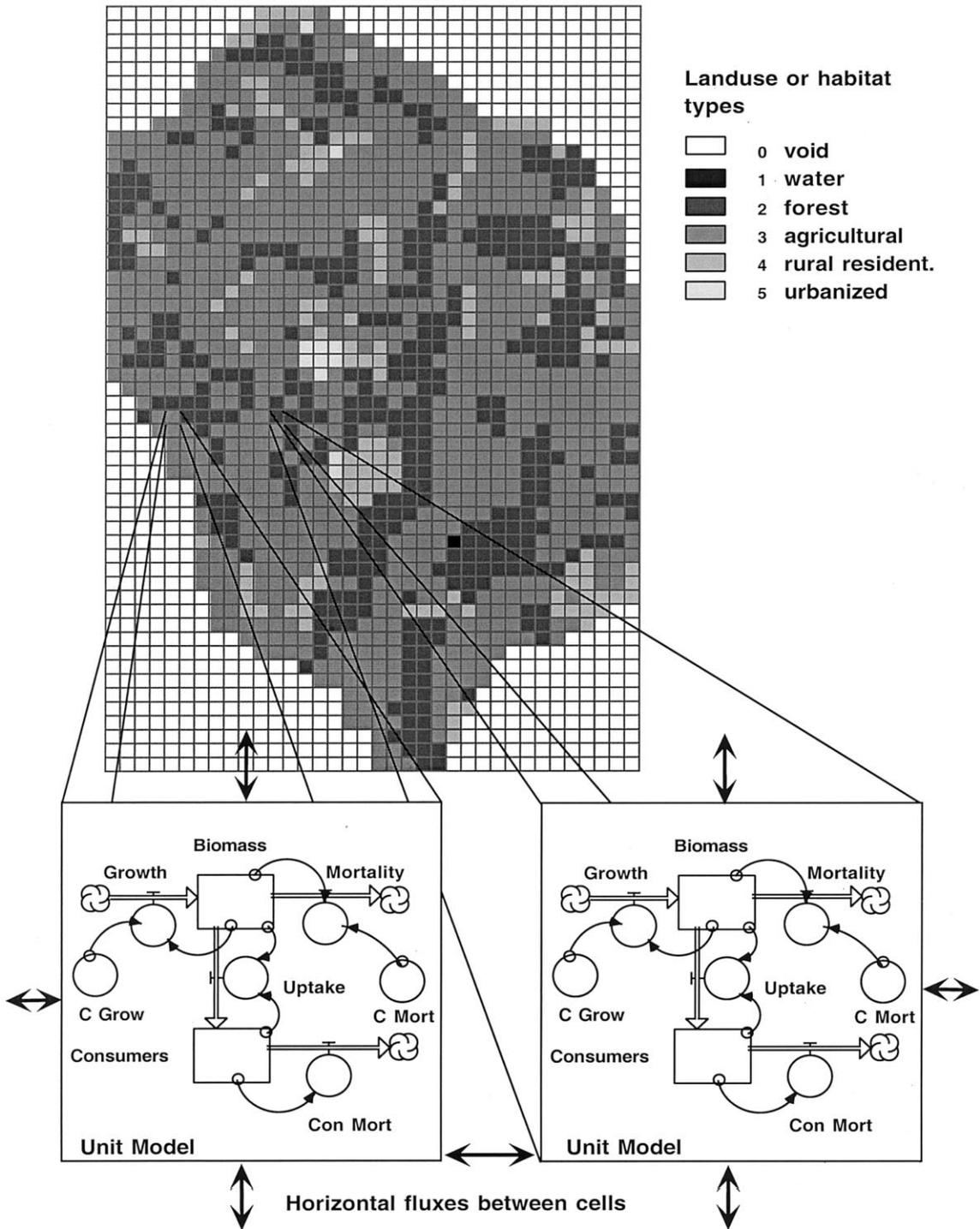


Fig. 2. Spatial representation of the landscape as a uniform grid of cells with local hydrological dynamics linked by horizontal flows. A unit model that simulates vertical hydrological flows is replicated in each cell.

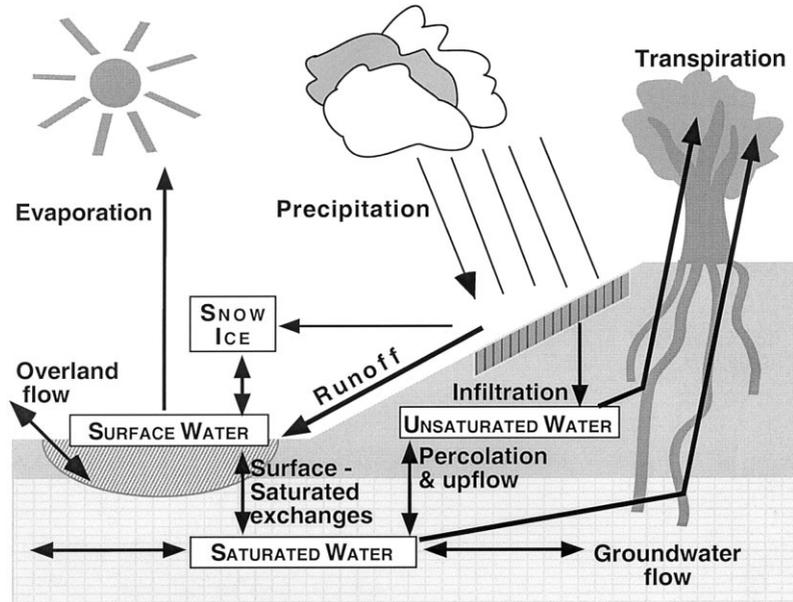


Fig. 3. The unit hydrological model. The state variables are *surface water*, *snow/ice*, *unsaturated water*, and *saturated water*. The major processes are precipitation, evapotranspiration, infiltration and percolation.

along the following pathway: rainfall → surface water → water in the unsaturated layer → water in the saturated zone. In each of the stages some portions of water are diverted due to physical (evaporation, runoff) and biological (transpiration) processes, but in the vertical dimension the flow is controlled by the exchange between these four major phases. This same conceptual model is also the basis of more detailed descriptions of infiltration and percolation (Rawls et al., 1992).

However for the temporal and spatial scales of a watershed landscape model this scheme needs to be reconsidered. First, at a time step of a day or more in steady state in terrestrial landscapes with a pronounced gradient, we do not observe surface water over unsaturated layers. During a rapid rainfall event surface water may accumulate in quantities that are not immediately infiltrated, but over the course of the day this water will either infiltrate or will be removed by horizontal runoff. There is a great variability in the infiltration rates, according to estimates of Skaggs and Khaleel (1982), the infiltration can vary between $0.048\text{--}216\text{ m day}^{-1}$, depending on the soil type and rainfall pattern. For the soil types on the Patuxent

watershed infiltration varies from $0.15\text{ to }6.2\text{ m day}^{-1}$ (Maryland Department of State Planning, 1973). In most cases this can certainly accommodate most of the water remaining from a rainfall occurring over a daily time period.

Second, climatic data are rarely available for better than daily resolutions. Moreover if the model is intended to be run over large areas for many years, the diel rainfall data become inappropriate and impossible to project for scenario runs. Therefore in the chosen modeling scale it should be realized that a certain amount of detail will be lost in any case: a flash rainfall event with a downpour of several centimeters occurring over a short time and therefore most probably flushed away with surface runoff becomes indistinguishable from a drizzle that delivers the same amount of water over 24 h, entirely infiltrating into the unsaturated layer.

With this in mind, we use a simplified unit hydrologic model (Fig. 3) based on the following set of hypotheses.

1. It is assumed that there can be no surface water on top of the unsaturated layer. Therefore the rainfall (or snowmelt) is channeled

directly into the unsaturated layer. Surface water appears only if the unsaturated layer becomes saturated, or when the infiltration rate is exceeded by the rainfall amount.

2. The surface water that may be present in a cell is that in rivers, creeks and depressions. There is no sheet surface flow, that is gradually infiltrated. All surface water is removed by horizontal runoff or by replenishing the saturated water in case that storage has been drained by lateral flow.
3. The surface water variable also takes into account the shallow subsurface flow that may occur during rainfall. Following the approach used in the CNS model (Haith et al., 1984) we assume that the surface water variable includes the water in the upper active 10 cm of the soil.

Conceptually the choice and meaning of the *surface water* and *saturated water* variables in PLM are close to the slow and quick flow separation (Jakeman and Hornberger, 1993; Post and Jakeman, 1996) assumed in empirical models of runoff. In this case the *surface water* variable accounts for the quick runoff, while the *saturated water* storage performs as the slow runoff, defining the baseflow rate between rainfall events.

We thus assume that the spatial and temporal scale of the type of a landscape model that we are building, allows us to simplify the structure of the unit model assumed in GEM, eliminating two otherwise important state variables (water in river and water in subsurface flow) and the corresponding processes. The precipitation (rainfall + snowmelt) (P) is now directed into the unsaturated zone at a rate that is determined by the infiltration function. Formalization of the infiltration process needs to match the chosen spatial and temporal resolution. This essentially continuous process needs to be defined as a discrete event based process and the continuous distribution of water in the unsaturated layer that should control ponding, infiltration and vertical percolation should be aggregated into one spatially averaged and uniform value. As in the ELM case study (Voinov et al., 1998) again we allow a good deal of empiricism to be included into the process based modeling paradigm that we employ.

We define the amount of water infiltrated as:

$$F = \min (F_i, F_s, F_\sigma, F_h, P).$$

where F_i is the infiltration rate; F_σ is the slope parameter for the infiltration (the steeper the slope, the smaller the infiltration); F_s , the soil parameter for the infiltration; and F_h is the habitat parameter for the infiltration. These three parameters are based on the slope, soil and land use maps, respectively. Besides, for each time step the infiltration capacity I_c is calculated as $I_c = U_h \rho$, where U_h is the unsaturated depth and ρ is porosity. If $F > I_c$ then the unsaturated zone is totally eliminated and the incoming precipitation is added directly to the saturated water. Whatever precipitation is in excess of either the infiltration rate ($P - F$) or the capacity of the unsaturated zone ($P - I_c$), is added to the surface water, which is then subject to runoff.

In addition to the spatial features of the infiltration process, there are also certain temporal patterns that need be accommodated in the model. Inspecting the flow data measured at gaging stations and comparing it to the patterns of precipitation, one can notice that in most cases after a long dry period the runoff is considerably lower than when it is caused by a series of rainfall events. This makes sense and can be explained by the patterns of vertical distribution of water in the unsaturated layer. If we want to avoid the spatially heterogeneous description in the vertical dimension we have to adopt certain empirical rules that would modify the infiltration rate according to the observed patterns.

We do this in the model with the help of a counter $n \geq 1$ that is incremented by d^+ every time when $P < 0.001$ m and decreased by d^- whenever $P > 0.001$ m. The infiltration rate is then adjusted: $F^\circ = nF$.

By setting $d^+ < d^-$ we make sure that the infiltration gradually grows during the dry period and then rapidly declines to the original value once a rainfall series starts.

Another observation that follows from the rainfall versus runoff data analysis is that exceptionally high precipitation usually results in relatively higher runoff. If data were available and analyzed at a finer resolution this could probably be explained by the temporal patterns of the peak

rainfalls that are occurring in this area. The largest amount of rainfall tends to be delivered in shorter time periods resulting in rapid accumulation of water in the immediate subsurface zone, that blocks further infiltration and increases runoff. To account for this empirical generalization, we further modify the infiltration rate: $\Phi = fF^\circ$ where f is a hyperbolic function

$$f = \begin{cases} 1, & \text{if } P < P_{\max} \\ (P_{\max}/P)^4, & \text{otherwise} \end{cases}$$

P_{\max} is the threshold precipitation value, which causes higher runoff when exceeded.

By assuming these empirical extensions for the infiltration formalization we gain the flexibility of accounting for spatial and temporal variations in the infiltration process, while reducing the resolution and heterogeneity in the model variables that would otherwise be necessary.

3. Spatial hydrologic model

In the spatial implementation the major hypothesis that we are testing is that overland and channel flow can be modeled similarly. Traditionally, in most models of overland flow the surface water is moved according to two separate algorithms: one for the 2-dimensional flux across the landscape and the other for the 1-dimensional channel flow. This approach we find in some of the classical spatial hydrologic models such as ANSWERS (Beasley and Huggins, 1980) or SHE (Abbott et al., 1986). However, once again bearing in mind the spatial and temporal scale of the Patuxent model, as well as its overall complexity, we intend to use the same algorithm for both types of overland flow.

Taking into account the temporal resolution of the model, we may assume that whatever surface water is present in a cell, is already in the creeks, streams or rivers. Taking into account the spatial resolution we may further assume that in every cell there will be a stream or river present, where all the surface water can accumulate. Therefore it makes sense to consider the whole area as a linked network of channels, where each cell presents a channel reach which discharges into a single next

channel reach. Based on the elevation map and given a series of cells marking the mouth of the river, which is the outlet for water drained from the watershed, we can create a link map that specifies into which of the eight directions (N, NW, W, SW, S, SE, E, NE) a cell drains its water (Fig. 4). This link map can be either created by an ARC/INFO procedure, or it may be created by an internal algorithm, which is a part of the SME package (Spatial Modeling Environment, (Maxwell, 1995)).

After the water head in each raster cell is modified due to vertical fluxes, the surface water movement between the raster cells (and associated transport of constituents) is calculated. The hydrology of the whole watershed is modeled as a merge of two algorithms. One of them operates in areas with a pronounced elevation gradient. It is a simplified runoff algorithm that takes a certain portion of water from a cell and adds it to the next cell downstream, according to the link map (Fig. 5A).

The other algorithm is implemented in areas with standing water. These areas are marked as open water on the landuse map and they are much more similar to the conditions that we have encountered in our previous exercises with the Everglades Landscape model (Voinov et al., 1998). In this case we may look at the flow between two adjacent cells i and $i+1$ as flow in an open channel and use the so called *slope-area method* (Boyer, 1964), which is a kinematic wave approximation of St. Venant's momentum equation. The flux ($\text{m}^3 \text{d}^{-1}$) in this case is described by the empirical Manning's equation for overland flow. The equation is further modified to ensure that there is no flux after the two cells are equilibrated and then it is accelerated by the multi-cell dispersion algorithm. This method has been elaborated in our previous paper and we will not discuss it further here.

The simplified runoff algorithm that we have applied for the steeper areas on the watershed required some investigation and adjustment. First it became clear that the runoff operation needs to be performed at a different time scale than the 1 day time step adopted in the rest of the ecological

components of the model. Over 1 day water could travel considerably further than over one cell, especially when the spatial resolution was fine enough. The simplest way to treat this inconsistency was to decrease the time step of this part of the hydrologic module by several (10–20) times. The size of the time step and the number of iterations needed for the hydrologic module could be calibrated so that water was moved sufficiently fast.

As an alternative, we have tested another method which instead of reiterating the whole fluxing algorithm N times, identified the path of the water on the link map over N cells and then dumped the water from the current cell directly into that cell N steps downstream (Fig. 5B). The results of both algorithms were remarkably close

and we switched to the second one since it performed much faster. However, as we will see below in the calibration section, this approach was still very much dependent on the spatial scale of the model: as we switched to larger areas it occurred that the delivery rate was not high enough and the number N had to be increased.

In an effort to endogenize this process and make the model robust to rescaling we have adopted yet another method of calculating N . It was assumed that the length of the pathway that water was to travel over one time step was a function of the water head in the current cell (Fig. 5C). A stepwise function worked best for this purpose:

$$N = mW^s / (K^s + W^s)$$

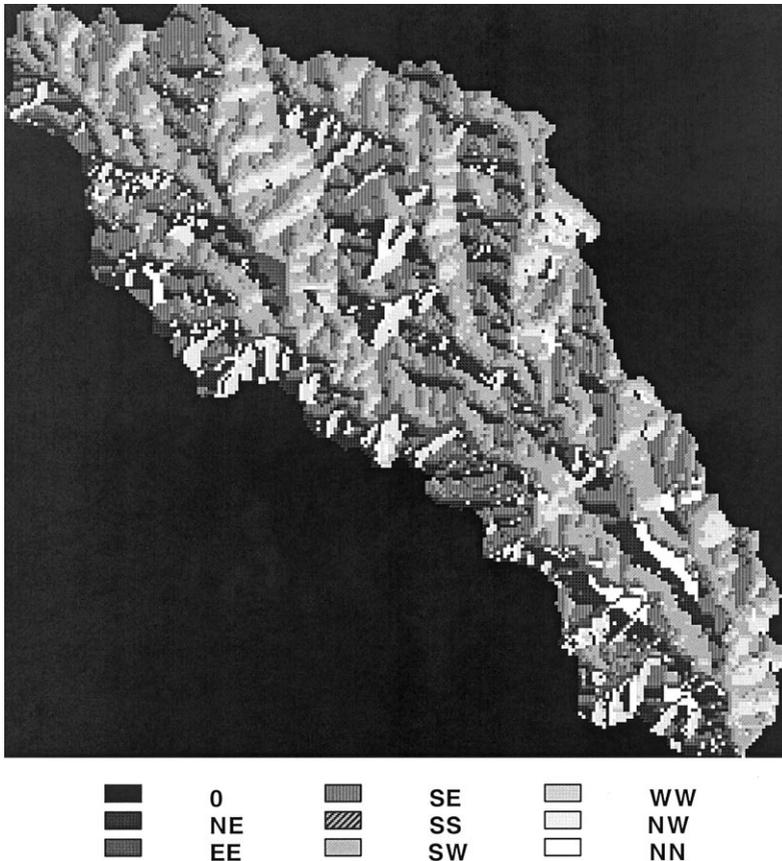


Fig. 4. Link map. Each cell is marked by a number that indicates the direction of the next cell along the path of surface water flow: 2 = NN, 3 = NW, 4 = WW, 5 = SW, 6 = SS, 7 = SE, 8 = EE, 9 = NE.

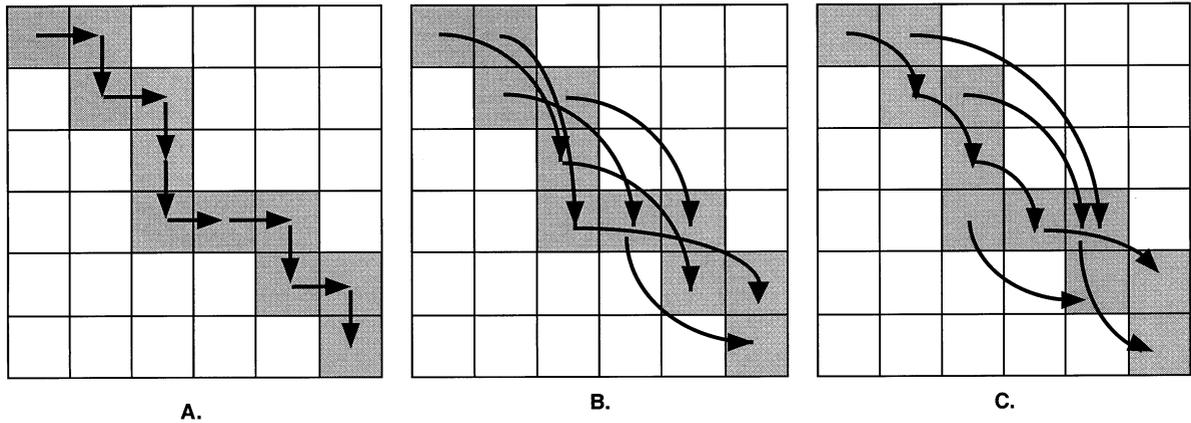


Fig. 5. Calculation of pathways on the Link Map. The water may flow over more than one cell in a daily time step. (A) Defining the water transport by running the whole hydrologic model for N iterations. (B) Defining the recipient cell as the N th cell down the link map and fluxing the water directly there. (C) Defining the pathway as function of the water head in current donor cell.

where N is the cell distance; m is the maximum allowed cell distance; W is the water head in the current cell; K is the threshold head at which the cell distance is half the maximum and s is a parameter that controls the steepness of the step function.

For saturated water fluxing a modified Darcy equation was employed. For each cell ij the flux to an adjacent cell ω was determined as a function of horizontal conductivity and the saturated water head difference between the current cell and the average of the nine cells that included the current one and the eight neighboring ones:

$$f_{ij,\omega} = (h_{ij} \cdot P_{\omega} - G_{\omega}) \cdot C_{\omega},$$

where h_{ij} is the targeted head; $h_{ij} = \sum_{k=i,j,\omega} P_k G_k / \sum_{k=i,j,\omega} P_k C_k$; $\omega = \{i \pm 1, j \pm 1\}$ is the set of neighboring cells for cell i, j ; P_{ω} is porosity in cell ω ; G_{ω} is the amount of water in the saturated storage in cell ω ; C , is the horizontal conductivity in cell C_{ω} .

4. Data sets

Spatial hydrologic modeling requires extensive data sets. In Fig. 6 we present the basic spatial coverages that have been employed in our modeling effort and some of the derived layers that were also essential for the hydrologic module. Spatial

fluxes of surface water in watershed models are predominantly driven by the elevation gradient. The best elevation data in a digital format can be obtained by direct digitizing of the topographic maps of the area modeled. However this is a time consuming, tedious and expensive procedure. As an alternative we have relied on the United States Geological Survey's (USGS) Digital Elevation Model (DEM) data that are available for downloading from the Internet (USGS, 1995).

USGS offers elevation data in 1 degree grid coverages for the four quadrangles covering the Patuxent watershed. DEM grids are based on 1:250 000 USGS maps with 3-arc second grid spacing. Grids constructed from USGS 1-degree DEMs are not immediately suitable for the analysis of such topographic features as volume, slope, or accurate visibility, because they measure the x , y (planar) locations as latitude and longitude, while the z value (elevation height) is measured in meters. Consequently, the actual distance on the ground represented by one ground unit is not constant, and the ground distance units and the surface elevation units are not the same. To make this surface model compatible with other layers of information and suitable for analysis, the ground units in the 1° USGS DEM have been projected into nonangular units of measure such as the UTM coordinates. After reprojection the grid has been rescaled to the 200 m resolution, which is the

highest resolution currently used in the PLM. The vertical resolution of the DEM maps is 1 m. This resolution is not fine enough to show the actual shoreline in the estuarine part of the watershed that lies between 0 and 1 m of elevation. The shoreline was derived from the river network map (TIGER/LINE-US Bureau of Census (USCB, 1996)) and then combined with the estuarine bathymetry data (NOAA/National Ocean Service, 1997) and elevation map to make the land-water boundary visible.

Another inconsistency is observed when the elevation layer is combined with the river network and ‘depressions’ show along the courses of streams. This is caused by widely known imperfections in DEM data and determined primarily by the resolution (distance between sampling points). Another common cause of these ‘depressions’ is storing the elevation data as an integer number and hence rounding off the real data. The additional aggregation to the 200 m and 1 km resolutions performed for our modeling purposes adds to the problem. This prevented the proper flow of water downstream in the hydrologic mod-

ule and resulted in an erroneous link map. Corrections may be performed using several ARC/INFO hydrologic functions that allow one to determine and fill depressions. This was also directly incorporated into the PLM hydrologic module, that used the river network map as a reference and smoothed out all the depressions along the river beds.

Using a GIS the DEM data have been preprocessed to create several other raster maps needed for the hydrologic model. Watershed Boundary (Studyarea map), Slopes and Aspects layers have been calculated by the Watershed Basin Analysis Program in GRASS-Geographic Resources Analysis Support System (USACERL, 1993). The best results were achieved with a threshold of 500 cell units used as the minimum size of an exterior watershed for the 200 m resolution.

The River Network coverage has been acquired from the TIGER/LINE (USCB, 1996) in a vector format. The database contained numerous errors: streams that were not continuous, missing channels (improperly digitized or missing on the original maps or photos because they may have been

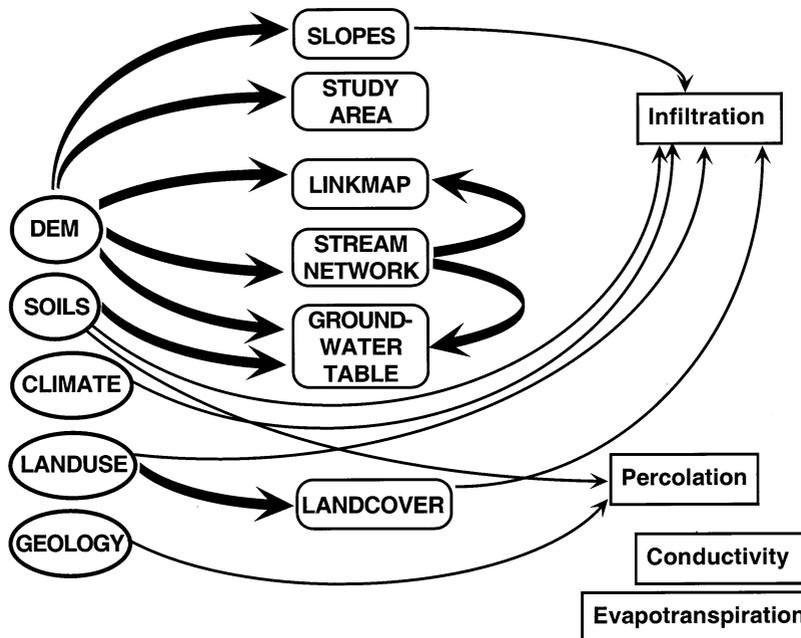


Fig. 6. Spatial GIS layers and maps used in the model. There are six basic coverages. Additional maps are created during preprocessing and model initialization. Other spatial parameters and variables are calculated and updated during model runs.

dried out at the time the photos/maps were interpreted). The hydrologic analysis tools in the ARC/INFO GRID module (ESRI, 1994) were applied to correct the digitized stream network. Using the digital elevation model as an input we delineated the drainage system and then quantified its characteristics. For any location in the grid, those tools also gave us the upslope area contributing to that point and the downslope path water would follow. A 'hydrologically proper' surface, without any artificial pits or hills, was produced and flow directions and flow accumulations were determined. Water channels were identified for different threshold amounts of water accumulation (product of the number of cells draining into a target cell and the size of the precipitation event). These water channels were used as a background coverage to manually correct stream discontinuities for the digitized River Network. The corrected River Network was converted into a raster (cell-based) format in order to comply with other data layers. This River Network map produced from the elevation data turned out to be more consistent, than the original vector map.

The Soils layer was originally imported from the State Soil Geographic (STATSGO) data base (NRCS USDA, 1997) which has been compiled using a USGS 1:250 000 scale, 1 by 2° quadrangle series as a map base. The STATSGO Data Base was downloaded in GRASS format and reprojected from the Albers Conical Projection to the needed UTM projection. Every map unit on a STATSGO coverage contains up to 21 components (segments) for which there are attribute data. One of the disadvantages of this dataset is that these components cannot be spatially identified, which reduces the STATSGO application to the coarse regional scale.

After analyzing the tabular information it was clear that aggregation criteria did not include hydrological properties, because one map unit could contain soils from very different hydrological groups. Therefore we could use only some general hydrological parameters from STATSGO but most of the spatially explicit soil data was taken from the Patuxent Watershed Counties Soils map available at the Maryland Office of

Planning (MOP) (Maryland Department of State Planning, 1973).

The Groundwater Table Map, required as an initial condition for the model, was approximated from a series of spatial and point data sets using the GRASS overlay and interpolation techniques. The reference points were taken from:

- MOP Soils map and the unsaturated depth data that was provided by the Maryland Department of State Planning;
- The elevation and river network coverages, along which the ground water table was assumed to reach the surface;
- 15 well measurements of the groundwater level over the watershed area (James et al., 1990).

The groundwater depth data were interpolated over the whole watershed with these data sets as reference points. After that the model was run for 100 days, the Groundwater Table Map was regenerated, saved and then fed back into the model for subsequent runs as the initial condition for the depth of the water table. This improved the performance of the Hydrological Module by significantly decreasing the initial adjustment period in the model runs.

Land Use coverages have been acquired from Maryland Office of Planning in a vector format and then rasterized for the required cell resolutions. The climatic data series were taken from the Earthinfo Inc. NCDC Summary of the Day Database (EARTHINFO, 1993). The point time series for precipitation, temperature, humidity and wind were then interpolated across the study area to create spatial climatic coverages. The calibration procedures were mostly based on USGS gaging data also available for downloading from the Web (USGS, 1995a). There are 13 gaging stations in the area that have data for the time period that matches the one defined by the climatic data series, that is 1980–94.

It should be noted that the PLM hydrologic module was entirely based on data available either from the Web or from the State and Federal Agencies and Institutions. There were no special data collection efforts undertaken for the sole purpose of this project. This fact makes us some-

what optimistic about the possibilities of applying the PLM methods elsewhere in other watersheds and landscapes.

5. Calibration and scaling

Calibrating and running a model of this level of complexity and resolution is quite tedious and requires careful planning and decomposition. There is little chance to get the model running as a whole immediately and a stage-by-stage approach is most fruitful.

Therefore we first identified two resolutions to run the model—the 200×200 m and the 1×1 km. On the one hand, the 200 m resolution was more appropriate to capture some of the intrinsic processes associated with landuse change. On the other hand it was certainly too detailed and required too much CPU time to perform the numerous model runs required for calibration and to run the full model over decades which is essential for scenario analysis. At the 1 km resolution the 58 905 cells of the 200 m version were replaced by 2352 cells.

Second, we identified a hierarchy of subwatersheds. The whole Patuxent area has been split into a hierarchy of nested subwatersheds to perform analysis at smaller scales. Three levels have been identified (Fig. 7). A small subwatershed (Cattail creek) in the North of the Patuxent drainage basin has been chosen as a starting point for our spatial analysis. From there we moved on to the upper portion of the Patuxent watershed, that included approximately half of the area and ended at Bowie. And finally we looked at the whole Patuxent watershed. We thus started with running the model over 566 cells, then jumped to 23 484 cells and finally moved to all the 58 905 cells in the study area of the Patuxent watershed at the 200 m resolution.

We staged a set of experiments with the small Cattail creek subwatershed to test sensitivity of the hydrologic module. It was established from this exercise that there are three crucial parameters that control the surface water flow in the model. These were the infiltration rate, the horizontal conductivity, controlling the baseflow rate

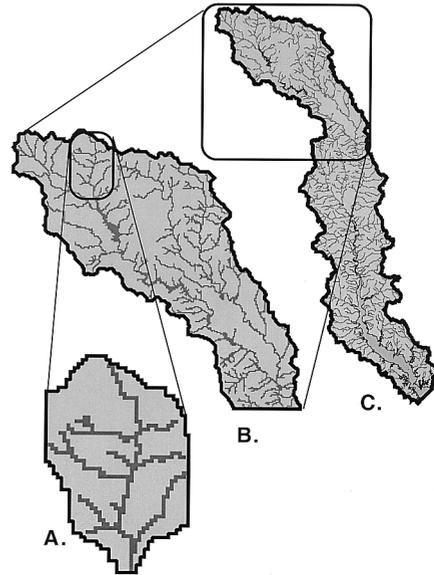


Fig. 7. Hierarchy of subwatersheds on the Patuxent drainage basin used to calibrate and analyze the model. (A) 22.5 km² Cattail Creek; (B) 939.9 km² Upper Patuxent draining at Bowie; (C) Full study area.

and the number of iterations, or the path length of water transport in the hydrologic module. The infiltration rate effectively controlled the height of peaks in the river water flow. The conductivity determined the amount of flow in the low period and by changing the number of iterations we could modify the length of the peaks and the delivery rate downstream.

Calibration of the hydrologic module was conducted against USGS data from several gaging stations over the watershed. First the model was calibrated for the 1990 data for the small Cattail subwatershed. Then the model was run for 5 consecutive years (1986–1990). Fig. 8 displays the annual dynamics of rainfall for 1984–1992, which show that the 1986–1990 period gives a good sample of various rainfall conditions that may be observed in the watershed, 1989 being a wet year and 1986 representing dry year conditions. The results displayed in Fig. 9 are in fairly good agreement with the data and may be considered as a partial model verification, because none of the parameters have been changed after the initial calibration using 1990 data. Some of the flow

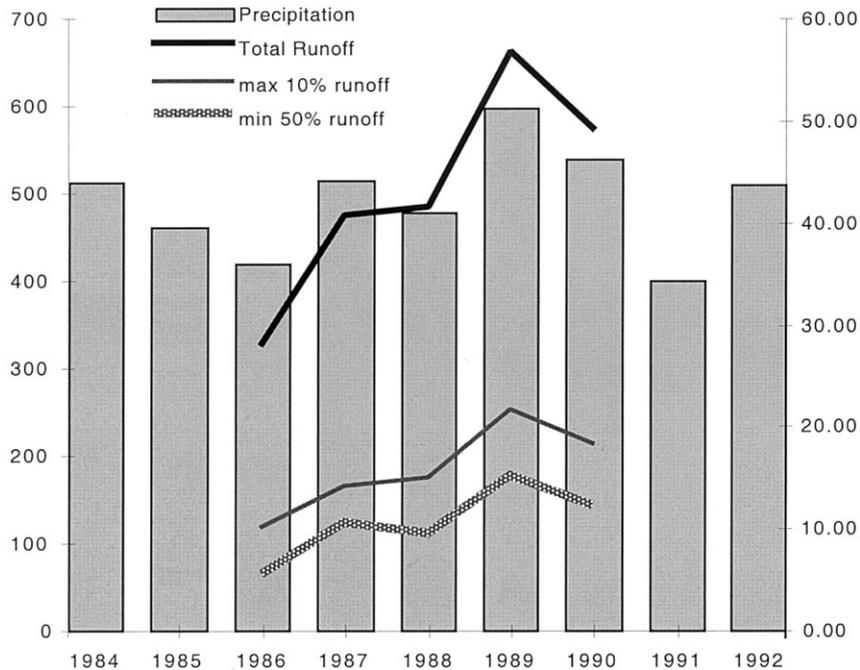


Fig. 8. Annual rainfall (inches) data for 1984–1992 and runoff data for the 1986–1990 period over which the model was analyzed.

statistics are presented in Table 1, were we have also included calibration results from the Hydrologic Simulation Program-Fortran (HSPF) (Donigian et al., 1984), that has been previously applied to the Patuxent watershed (AQUA TERRA, 1994). Surprisingly we attained a considerably better fit to data with our model, than with HSPF. HSPF is a more complex model that uses much more elaborate data sets for its performance (e.g. hourly rainfall data), although its spatial resolution is not as fine as in our model (aggregated subwatersheds).

We did not have any reliable data to calibrate the spatial dynamics of ground water. However we tracked the total amount of water in saturated and unsaturated storage to make sure that the model is in quasi steady state with respect to groundwater. The dynamics of these integrated values (Fig. 10) were in good agreement with the total amount of rainfall received by the watershed (Fig. 7), responding with a lower level of the groundwater table in dry years and a rising water table during wet periods.

The fit was less accurate once we started running the same model with the same parameters over the half-watershed area. One parameter that needed to be adjusted was the number of iterations N in the hydrologic module. Now that a larger watershed was involved it turned out that a better fit could be obtained if the number of iterations was further increased. Apparently this

Table 1
Model verification for Cattail Creek subwatershed and comparison to the HSPF model statistics

	Data	Model	% error	HSPF % error
Total flow	2510.41	2527.58	0.68	8.2
Max 10%	930.2	925.79	-0.48	4.9
Min 50%	587.3	596.25	1.50	-14.7
Total 1986	326.16	282.24	-15.66	
Total 1987	472.83	469.25	-0.76	
Total 1988	482.01	414.22	-16.37	-0.7-
Total 1989	660.62	748.29	11.72	+18.1
Total 1990	568.78	611.31	6.96	

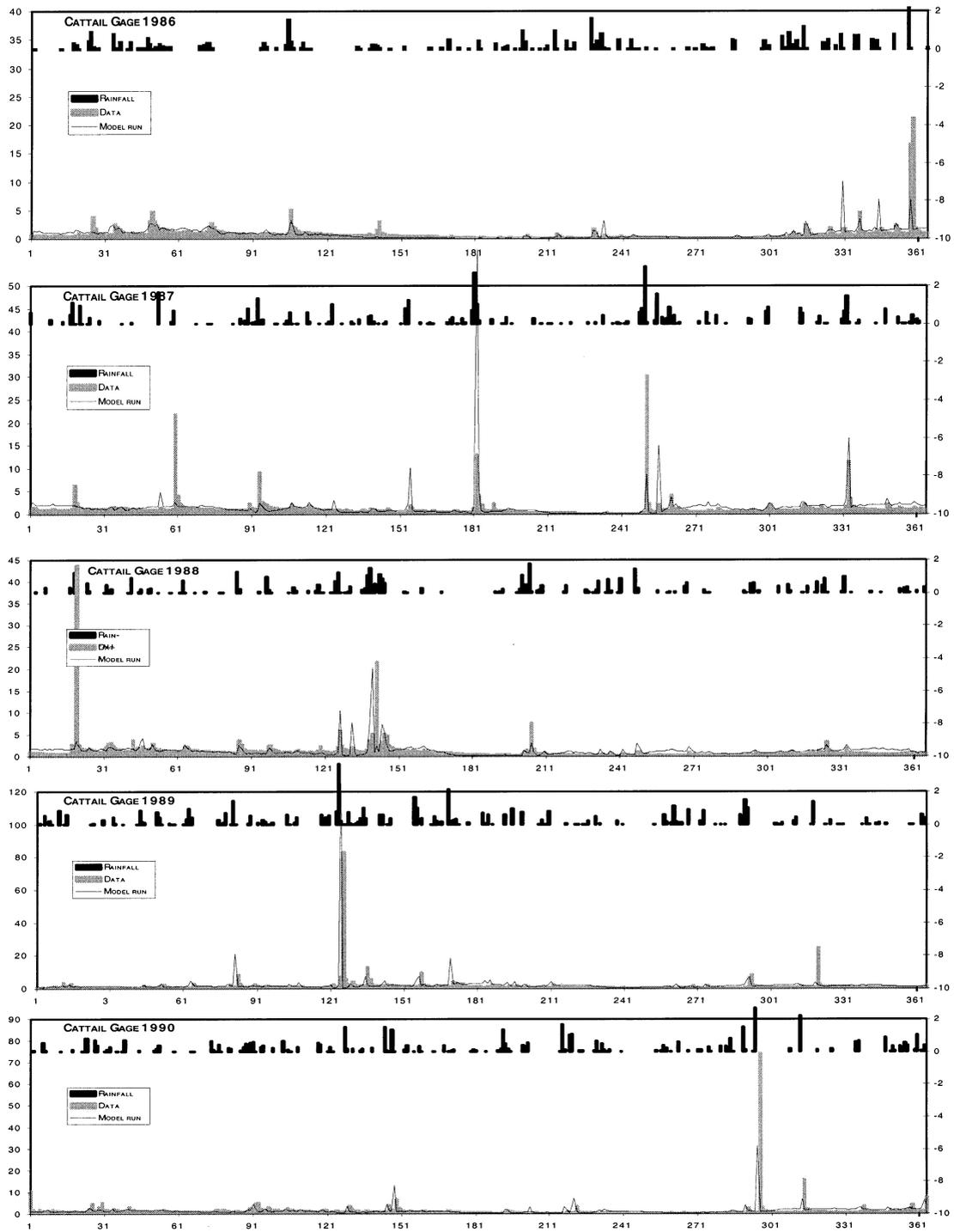


Fig. 9. Calibration and verification results for Cattail Creek (1986–1990).

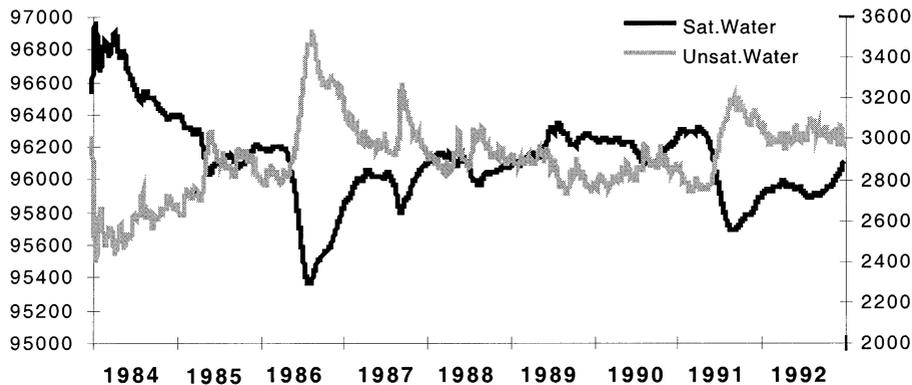


Fig. 10. Dynamics of groundwater in the model. As the groundwater table drops during dry years (compare to Fig. 7) the amount of water in unsaturated storage increases and vice versa.

was because at this larger scale we needed to move water further and faster to better simulate the short-term high peaks that were observed in the data. This was a clear illustration of the fact that different scales present new emerging behavior of the system, and that rescaling is always a delicate process that cannot be done mechanically. The best fit to data was obtained when running the model with the selfadjusting method for N with the maximum number of iterations $m = 100$ (Fig. 11). Interestingly, the Cattail subwatershed was still performing as well as before with this value of m . This could be expected since the previous analysis showed that there was no sensitivity of subwatersheds to increases in N beyond 20 ($m = 20$).

The calibration statistics for the half watershed area are summarized in Table 2. This time the fit was not as good as for only the Cattail subwatershed, clearly indicating that the increased spatial extent of the model presented more heterogeneity that was not accounted for in the Cattail subwatershed calibration. One potential source of error is the groundwater dynamics. Within the subwatershed we assumed that the groundwater dynamics closely follows the surface water flows and confined the groundwater to the subwatershed area. This is probably not accurate for the Cattail Creek and at larger scales the groundwater patterns are apparently even more different and result in added error in the simulation.

One other possible source of error was the

spatial representation of climatic data, which we had measured at several points and then interpolated over the study area. The sensitivity analysis showed that the overall annual flows are highly sensitive to particular climatic timeseries and to the spatial patterns of climatic data. The Cattail hydrology was driven by two climatic stations. The half-watershed model incorporated data from eight stations. In reality over this larger area there is certainly more variability in climatic factors that potentially causes some of the flash runoff that we see in the data, but which we cannot mimic with the model based on only the few meteorologic stations we have smoothed over the whole area. Nevertheless the general hydrologic trends seem to be well captured by the model, and the total flows were still better simulated than with the HSPF model.

Our next experiment was related to the resolution of the cells in the model. Using GIS operations we have rescaled the input maps, switching from a 200 m to a 1 km cell resolution. Model runs for the 1 km resolution were remarkably close to the results we were getting from the 200 m model (Fig. 12). This finding was especially promising for analysis of the other modules of the full ecological economic model. Since most of the horizontal spatial dynamics is stipulated by the hydrologic fluxes, we can expect that the coarser 1 km resolution would be sufficient for the spatial analysis of the integrated model of the watershed.

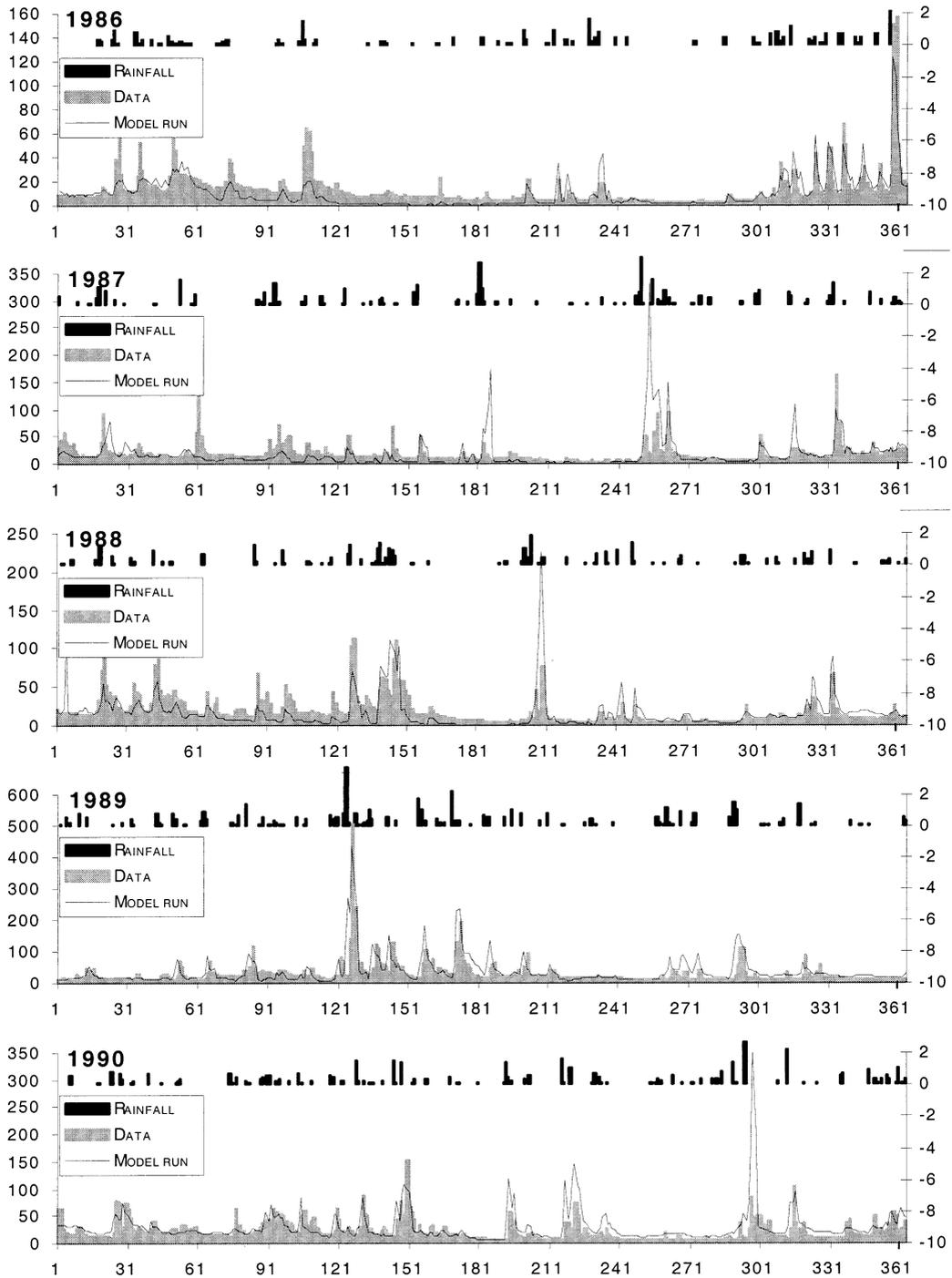


Fig. 11. Calibration for Upper Patuxent at Bowie (1986–1990).

The correlation between the model results at two resolutions turned out to be even higher than the correlation between some of the data sets involved. For example in Table 3 we present the comparison of the model correlations with the corresponding correlations between the elevation and slope maps considered for the 200 m and 1 km resolutions. Two methods of aggregation from 200 m to 1 km were explored. One is the average method (A), when the aggregate value in the 1 km grid is defined by the average of the 25 individual values in the 200 m cells. The other is the majority method (M), that takes the most frequently occurring value as the aggregate value. Both aggregation methods produce certain error. The A method works slightly better for the Elevation coverage, while the M method gives less error for the Slope coverage. However for both methods the slope layers were very poorly correlated. Naturally the model performed differently on the 200 m and the aggregated 1 km data sets. But for both aggregation methods the model performance on the aggregated 1 km maps correlated even better to the 200 m model, than the 1 km spatial data set correlated to the original 200 m data.

The comparisons of flows at gaging stations is a convenient way to look at the model output, calibrate and analyze the model performance. It integrates a lot of 2-dimensional spatial information in a normalized 1 dimensional picture. Another way to view the output of a spatial model, which is especially important to localize potential

accumulations of water and other spatial inconsistencies, is to output the model variables as series of maps, which can be then compiled into graphic animations. While the format of a Journal publication is not suitable for this kind of output, it may be well displayed on the Internet. We therefore refer the reader to our WWW page at <http://iee.umces.edu/PLM>, to view further model output.

6. Conclusions

Most of the traditional hydrologic models used in ecological applications seem to borrow much from engineering hydrology that usually enjoyed the luxury of an abundance of well performed and replicated measurements for particular locations (weirs, culverts) or relatively small areas and for quite short-termed events (rainfalls). The temporal and spatial resolution in those models could be quite fine, the processes could be considered in much detail and, with all the data available, model calibration was still not a problem. As the modeled systems become more integrated and as more variables and processes in addition to the hydrologic ones are to be accounted for by the model, we can afford less detail for each of the individual processes and formalizations, if we are to keep the model at an overall complexity level feasible for analysis and understanding. As a result more empiricism must be added to the model

Table 2

Model verification and comparison to the HSPF model statistics for the half subwatershed draining at Bowie and the Unity subwatershed

	Unity				Bowie			
	Data	Model	% Error	HSPF % error	Data	Model	% Error	HSPF % error
Total flow	3950.54	3981.31	0.8	-2.1	36617.43	37978.78	3.6	9.7
Max 10%	1410.15	2148.13	41.5	2.3	12497.58	16546.70	27.9	15.1
Min 50%	826.76	626.78	-27.5	-12.1	7917.98	6582.62	-18.4	9.0
Total 1986	484.52	446.30	-8.2		4752.94	4352.84	-8.8	
Total 1987	816.48	942.00	14.3		6446.08	7041.22	8.8	
Total 1988	819.30	792.10	-3.4	-11.6 - +8.0	6751.99	5841.62	-14.5	-2.6 - +25.3
Total 1989	960.30	949.45	-1.1		10507.98	11881.88	12.3	
Total 1990	869.94	851.47	-2.1		8158.45	8861.23	8.3	

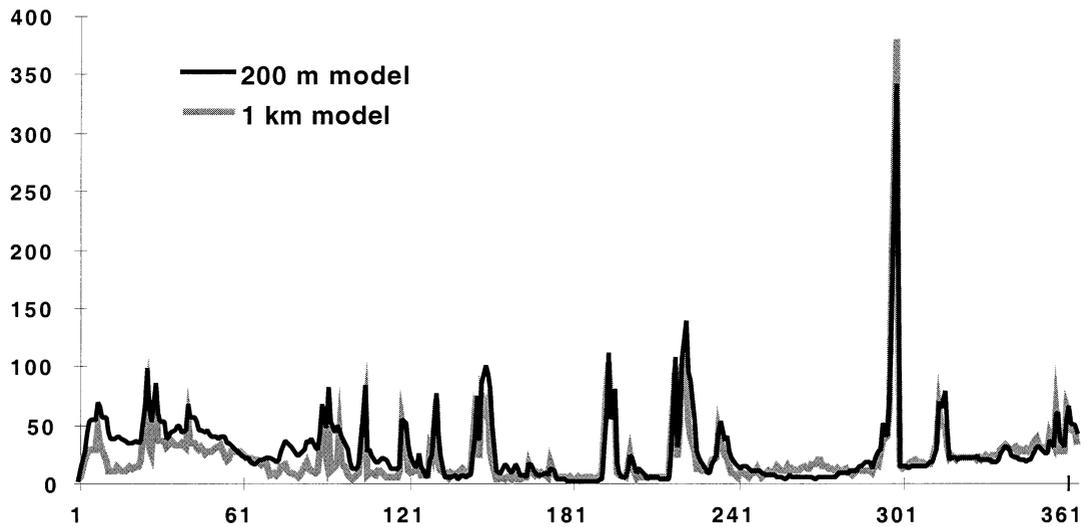


Fig. 12. Comparisons of model output generated at 1 km resolution and 200 m resolution.

structure, certain processes become lumped together, the models become more aggregated in space and in time, spatially distributed parameters become averaged over space and assumed homogeneous. In this case direct comparisons with other model structures are hardly meaningful, because of the obvious inconsistencies and different levels of empiricism. The only criterion that can be still used to evaluate the model performance is the verifications against existing measurements and cross calibration with other models.

In the hydrologic module of PLM we went somewhat beyond the traditional amount of empiricism seen in spatial hydrologic models of watersheds. We have included some new empirical functions to describe infiltration and the horizontal propagation of surface water across the landscape. The trade-offs that we gained were less variables, simpler model structure, less data needed for calibration, and faster performance. On the other hand this added empiricism may significantly impede model applicability to other watersheds. However, this was not the case in our applications to the several subwatersheds on the Patuxent drainage basin. We have succeeded in using the same model with the same set of parameters for several watersheds, varying in spatial resolution and scale. In most cases the model

performance was comparable to that of HSPF—the more complex and detailed model widely used in various EPA programs (e.g. Chesapeake Bay Program (CBP, 1998)).

The model is also limited by the data used and the spatial, temporal and structural scales assumed. In our case, the error can be explained by the following factors:

- coarse 1 day temporal scale that cannot distinguish between the types of rainfall events; spatial and temporal interpolation of climatic data;
- vertical aggregation of variables and processes; discrete rather than continuous formalization of processes;
- horizontal aggregation within cells, neglect of spatial units that are smaller than the cell size (40 000 m² or 1 km²).

Even with these assumptions we were able to produce a fairly good fit to the available gaging stations data and managed to simulate the long-term aggregated runoff dynamics that was essential for the overall Patuxent Landscape Model. It was somewhat unexpected that the coarser 1 km resolution model could mimic data almost as well as the 200 m model. Apparently the temporal (1 day) and structural (4 hydrologic variables) aggrega-

Table 3

Correlations between flows, elevation and slopes at 200 m and 1 km resolutions (M, majority method; A, average method)

Map 1	Map 2	Ag. Method	Maps Correl.	Model Correl.
Elevation 200m	Elevation 1 km	M	0.921	0.922
Slope 200 m	Slope 1 km	M	0.25	
Elevation 200 m	Elevation 1 km	A	0.926	0.935
Slope 200 m	Slope 1 km	A	0.23	
Elevation 1 km (A)	Elevation 1 km (M)		0.996	
Slope 1 km (A)	Slope 1 km (M)		0.725	

gation in the model is better matched by the 1 km spatial aggregation. Therefore the added detail that comes with the 200 m fine resolution is unessential for the overall model performance. However additional studies are required to support this statement and find some guidelines for the overall balanced model scaling in space, time and structural complexity. At this time we could only see that the spatial resolution was less important than the size and characteristics of the watershed. Larger spatial extent resulted in larger simulation errors.

Acknowledgements

Funding for this research has been provided by the US EPA/NSF Water and Watersheds Program through the US EPA Office of Research and Development (Grant no. R82-4766-010). Our sincere thanks are due to our project collaborators: Tom Maxwell, Lisa Wainger, Ferdinando Villa, Roel Boumans, and Elena Irwin.

References

- Abbott, M.B., J.C. Bathurst, J.A. Cunge, P.E. O'Connell, J. Rasmussen, 1986. An introduction to the European hydrological system—systeme hydrologique Europeen, 'SHE' 2: Structure of a physically-based, distributed modelling system. *J. Hydrol.* 87: 61–77.
- AQUA TERRA, M., USGS, 1994. Patuxent River Basin Watershed Model, 94.
- Beasley, D.B., Huggins L.F., 1980. ANSWERS (Areal Non-point Source Watershed Environment Response Simulation), User's Manual. West Lafayette, Purdue University.
- Beven, K.J., Kirkby, M.J., 1979. A physically-based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* 24 (1), 43–69.
- Bockstael, N., Costanza R., Strand I., Boynton W., Bell K., Wainger L., 1995. Ecological economic modeling and valuation of ecosystems. *Ecol. Econ.* 14, 143–159.
- Boyer, M.C., 1964. Streamflow measurements. In: Chow, V.T. (Ed.), *Handbook of Applied Hydrology*. McGraw-Hill, New York, pp. 15.1–15.42.
- CBP, 1998. The Chesapeake Bay Program, <http://www.chesapeakebay.net/Vbayprogram/>.
- Donigan, A.S., Imhoff J.C., Bicknell B.R., Kittle J.L., 1984. Application Guide for Hydrological Simulation Program-FORTRAN (HSPF). Athens, GA, US EPA, Environmental Research Laboratory.
- EARTHINFO, I., 1993. SD on CD. NCDC Summary of the Day. East. Boulder, CO, EARTHINFO, Inc.
- ESRI, 1992. Cell-based Modeling with GRID 6.1. Supplement-Hydrologic and distance modeling tools. Redlands, CA, ESRI.
- ESRI, 1994. ARC/INFO. GRID-CELL Modeling. Redlands, CA, Environmental Systems Research Institutes, Inc.
- Fitz, H.C., Costanza R., Voinov A., 1999. A dynamic spatial model as a tool for integrated assessment of the Everglades, USA. In: R. Costanza. (Ed.), *Ecological Economics for Integrated Modeling and Assessment* (in press).
- Fitz, H.C., DeBellevue E., Costanza R., Boumans R., Maxwell T., Wainger L., Sklar F., 1996. Development of a general ecosystem model for a range of scales and ecosystems. *Ecol. Model.* 88; (113): 263–295.
- Grayson, R.B., Moore, I.D., McMahon, T.A., 1992. Physically based hydrologic modeling: 1. A Terrain-based model for investigative purposes. *Water Res. Res.* 28 (10), 2639–2658.
- Haith, D.A., Tubbs L.J., Pickering N.B., 1984. Simulation of pollution by soil erosion and soil nutrient loss. Wageningen, Pudoc.
- Jakeman, A.J., Hornberger, G.M., 1993. How much complexity is warranted in a rainfall-runoff model? *Water Resources Res.* 29 (8), 2637–2649.
- James, R.W.J., Hornlein J.F., Strain B.F., Smigaj M.J., 1990. Water Resource Data. Maryland and Delaware. Water Year 1990. Vol. 1. Atlantic Slope Basins, Delaware River through Patuxent River., US Geological Survey Water Data Report.

- Maryland Department of State Planning, 1973. Natural Soil Groups of Maryland. Technical Report, MD Dept. of State Planning.
- Maxwell, T., 1995. Distributed Modular Spatio-Temporal Simulation. <http://kahir.cbl.cees.edu/SME3>.
- NOAA/National Ocean Service, 1997. MapFinder Home Page., <http://mapindex.nos.noaa.gov/>.
- Novotny, V., Olem, H., 1994. Water Quality. Prevention, Identification, and Management of Diffuse Pollution. Van Nostrand Reinhold, New York.
- NRCS USDA (1997). State Soil Geographic (STATSGO) Data Base, <http://www.nrcs.usda.gov/statsgo.html>.
- Park, R.A., Groden, T.W., Desormeau, C.J., 1979. Modifications to the model CLEANER requiring further research. In: Scavia, A.R.D. (Ed.), Perspectives on Aquatic Ecosystem Modeling. Ann Arbor, MI, pp. 87–108.
- Parton, W.J., Stewart J.W.B., Cole C.V., 1988. Dynamics of C, N, P, and S in grassland soils: A model. Biogeochemistry 5: 109–131.
- Post, D.A., Jakeman, A.J., 1996. Relationship between catchment attributes and hydrological response characteristics in small Australian Mountain Ash catchments. Hydrol. Process. 10, 877–892.
- Rawls, W.J.L.R., Ahuja, D.L., Brakensiek, S.A., 1992. Infiltration and soil water movement. In: Maidment, D.R. (Ed.), Handbook of Hydrology. McGraw-Hill, New York, pp. 5.1–5.51.
- Skaggs, R.W., Khaleel, R., 1982. Infiltration. In: C.T. Haan. (Ed.), Hydrologic Modeling of Small Watersheds. St. Joseph, Mich. American Society of Agricultural Engineers 5, 4–166.
- Sklar, F.H., Costanza R., Day J.W., Jr., 1985. Dynamic spatial simulation modeling of coastal wetland habitat succession. Ecol. Model. 29: 261–281.
- USACERL, 1993. GRASS Version 4.1. User's reference manual., Open GRASS Foundation Center for Remote Sensing, Boston University.
- USCB, 1996. TIGER: The Coast to Coast Digital Maps Database, US Census Bureau, <http://www.census.gov/geo/www/tiger/>.
- USGS, 1995. 1-Degree USGS Digital Elevation Model, US Geological Survey, http://edcwww.cr.usgs.gov/glis/hyper/guide/1_dgr_dem.
- USGS, 1995a. Maryland NWIS-W Data Retrieval, <http://waterdata.usgs.gov/nwis-w/MD/>.
- Voinov, A., Fitz, C., Costanza, R., 1998. Surface water flow in landscape models: 1. Everglades case study. Ecol. Model. 108 (1-3), 131–144.
- Watkins, D.W., Jr., McKinney D.C., Maidment D.R., Lin M.D., 1996. GIS and Groundwater Modeling. J. Water Resour. Plan. Manag., ASCE 122 (2), 88–96.
- Ye, Z., Maidment D.R., McKinney D.C., 1996. Map-based surface and subsurface flow simulation models: An object-oriented and GIS approach, http://www.ce.utexas.edu/prof/maidment/gishyd97/library/ye/rep96_5.htm.