

A Stream Network Model for Integrated Watershed Modeling

Zhi-Jun Liu · Donald E. Weller

Received: 8 March 2006 / Accepted: 8 January 2007 / Published online: 13 February 2007
© Springer Science + Business Media B.V. 2007

Abstract We describe a flexible, computationally efficient stream network model, which forms the core of a simulation framework that spatially integrates the contributions from point and nonpoint sources in a watershed. The model uses the map and stream topology information in the US Environmental Protection Agency's Reach File 3 to generate a spatially explicit network of stream reaches. Water and materials are routed through the stream network to the watershed outlet, and the routing process accounts for transit times and for possible nutrient losses in streams. This model can be applied wherever Reach File maps or maps from the newer National Hydrography Dataset are available, and it can be combined with models of other watershed processes to create a complete watershed simulation system. We present an application of the stream network model to two watersheds of different sizes in the Patuxent River watershed of Maryland, USA. Simulated predictions of streamflow and nitrate concentrations are either very good or good according to standards developed for evaluating the widely used Hydrologic Simulation Program – Fortran (HSPF) watershed model.

Keywords reach file · point and nonpoint sources · stream network traversal and routing · watershed models

1 Introduction

Pollution of streams and rivers in the USA and throughout the world remains of great concern, especially nonpoint source pollution by nutrients and sediment [15, 20]. There is a growing need to assess water quality and biological conditions in streams from a watershed perspective. In the USA, the imposition of total maximum daily loads as legal limits on the watershed-wide pollutant influx to a water body has created strong demand for new assessment tools [38].

Watershed models have emerged as important research and management tools, particularly in efforts to understand and control water pollution [11, 12, 16, 31, 33]. A great variety of watershed models can be found in the literature. Some, such as SHE [1, 2] and TOPMODEL [13, 14], simulate only hydrologic processes in a watershed, whereas others also include biogeochemical processes that influence water quality [5]. Some watershed models are designed for single storm events (ANSWERS [8] and AGNPS [42]); others are continuous-time models that simulate both storm and nonstorm conditions (HSPF [17, 21], SWRRB [6], and SWAT [5]).

Watershed models can be grouped into two categories of spatial representation: spatially lumped and spatially distributed. The former does not account for spatial variability of processes within a watershed, whereas the latter considers the within-watershed variations in the factors controlling the movements of water, sediment, and other materials. Spatially distributed models are a conceptual improvement over their lumped counterparts, and they can provide more accurate results when detailed data are available for parameterization and simulation. Many use grid cells to represent spatial variability of hydrologic processes and watershed characteristics. Such models are complex and data demanding, and the required data on

Z.-J. Liu (✉)
Department of Geography, University of North Carolina
at Greensboro,
Greensboro, NC 27402-6170, USA
e-mail: z_liu@uncg.edu

D. E. Weller
Smithsonian Environmental Research Center,
Edgewater, MD 21037-0028, USA

inputs and parameters are often unavailable. High uncertainties in predictions resulting from data uncertainties can limit model utility [10, 31]. There is an increasing need for spatially explicit models with less-intensive data requirements than grid cells.

White et al. [41] presented a conceptual spatial model that partitions a landscape into drainage subbasin units. Point and nonpoint sources of pollution are assigned within each subbasin, and subbasins are anchored to stream networks. The drainage subbasins and stream networks together constitute the hydrologic system through which point and nonpoint sources of pollution are aggregated. This conceptualization is the core of SPARROW, a statistical model that relates water quality to spatially referenced basin characteristics [32]. The works of White et al. [41] and Smith et al. [32] exemplify a shift in watershed modeling away from spatial representation on grid cells.

The development of a comprehensive watershed model that simulates all the important processes in a single program demands a great deal of time, manpower, and other resources. The application of such a model can be complex and computationally demanding, and often does not produce satisfactory predictions because data are not available to calibrate its large number of parameters. Such difficulties have helped drive a shift in watershed model development away from complex comprehensive simulation programs toward modular systems, such as Modular Modeling System [24, 25], which integrates simpler, exchangeable models for subsets of processes. For many practical problems, simpler models can give more precise predictions because they have fewer parameters and lower associated parameter uncertainties. Relatively simple, user-friendly modular models that can be applied to broad landscapes may also provide better assessment tools for watershed managers and decision makers.

We developed a watershed model composed of three simple modules to predict water and nutrient discharges from watersheds that can be divided into many subwatersheds. One of the modules is an existing rainfall–runoff model called TOPMODEL [13], and we created the other two modules for estimation of nutrient concentrations in runoff and stream network transport, respectively. This paper focuses on the stream network model, which was designed to be the core of our watershed simulation. We present it separately from the rest of our full watershed model because it is very generic and could be integrated into other modular watershed simulation systems. We applied the model to predict water discharge and nitrate concentrations from two watersheds of different sizes in the Patuxent River drainage of Maryland, USA, and we compared the simulated predictions to the measured streamflows and nitrate concentrations [22]. We discuss the quality of predictions and the advantages and applicability of our approach.

2 Model Design and Structure

We consider a generic watershed made up of many subwatersheds connected to a stream network (Fig. 1). The subwatersheds are sources of runoff and nonpoint sources of water pollutants, and they may also contain point sources of pollution. The stream network model uses the stream map and flow topology in the US Environmental Protection Agency's Reach File 3 to generate a stream network [36, 37], and it spatially integrates contributions from both point and nonpoint sources in the watershed (Fig. 2). The inputs to the stream network can be outputs from other models or actual measurements. Water and materials are routed through the stream network to the watershed outlet. The integration of point and nonpoint sources is accomplished in the routing process.

The following sections discuss the functions of four major components of the stream network model: the reach module, the binary search tree module, the first in–first out (FIFO) queue module, and the stream network module (Fig. 3).

2.1 The Reach File

A stream reach is a segment of a stream, and the stream network in a watershed is composed of all the stream

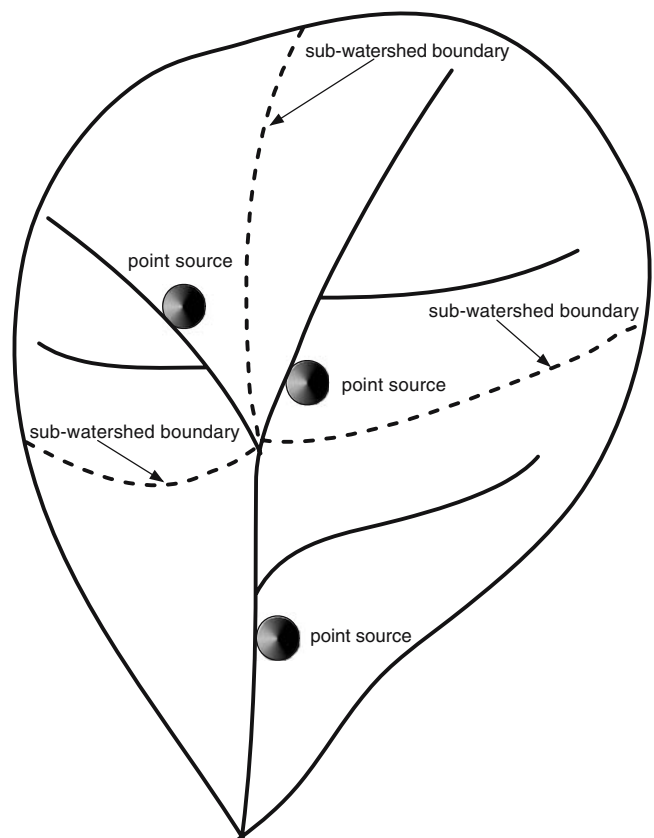


Fig. 1 A generic watershed with a stream network, subwatersheds, and point sources

Stream Network Model

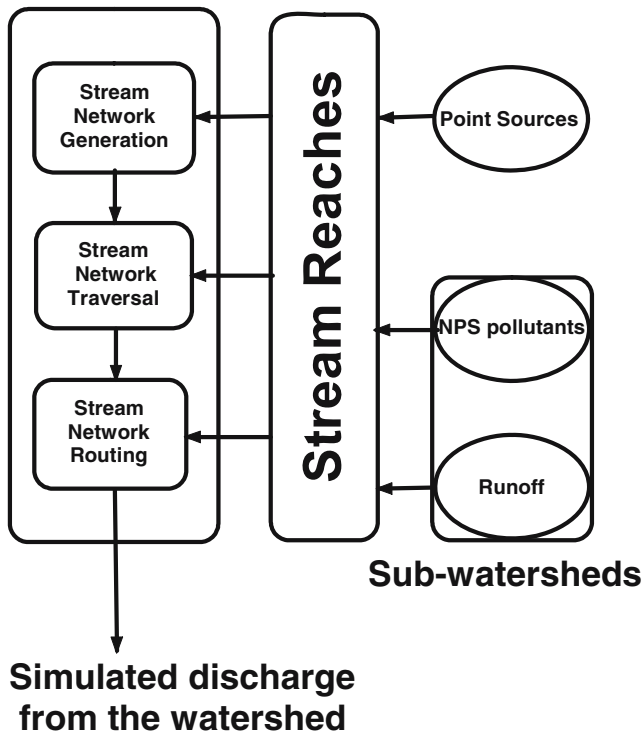


Fig. 2 A stream network-based conceptual watershed model. Inputs to stream reaches from point and nonpoint sources can be field measurements or outputs from component submodels. Point and nonpoint sources in the watershed are integrated in the stream network model

reaches in the watershed that are interconnected in an orderly fashion. Water flows downstream to the outlet of the watershed. Except for headwater streams, a typical stream reach receives water at its upstream end as well as runoff from the subwatershed directly draining to it, and it discharges water at its downstream end. Sediment and other materials are carried in water downstream to the watershed outlet in the same way.

Reach File 3 was developed by the US EPA as a network-oriented, spatially referenced geographic database of surface water features of the USA [36, 37] and was superseded by the US Geological Survey’s (USGS) National Hydrography Dataset (NHD) in 2000 [39]. Both Reach File 3 and NHD have been used in water-quality modeling of streams and rivers [3, 19, 32].

The basic unit of Reach File 3 is the stream reach, which has many attributes that define the connected stream network. The most important attributes are the reach code, which uniquely identifies each reach, and the navigation attributes, which specify the connectivity between reaches [36]. The reach code has three parts: the catalog unit (CU), the segment number, and the marker index. The CU is a unique eight-digit code assigned to a watershed and defined

as a Federal Information Processing Standard number. The CU for the Patuxent River watershed in Maryland is 02060006. The segment number is a unique four-digit number assigned to a surface water feature within a CU. When an existing segment is subsequently divided by a new tributary, the two pieces of the original segment are each assigned a different marker index value [36]. The navigation attributes of a stream reach are the identifying codes of its associated reaches. Any given ‘instant’ reach has up to five associated ones: the upstream left, upstream right, downstream, divergent, and complement reaches (Fig. 4). Besides these attributes from Reach File 3, we defined some additional attributes, such as sediment and nutrient concentrations, for simulating water and pollutant transport through the stream network.

2.2 Data Storage and Search: Binary Search Tree

To create a model that can be applied to watersheds of different sizes, we used a binary search tree as the data

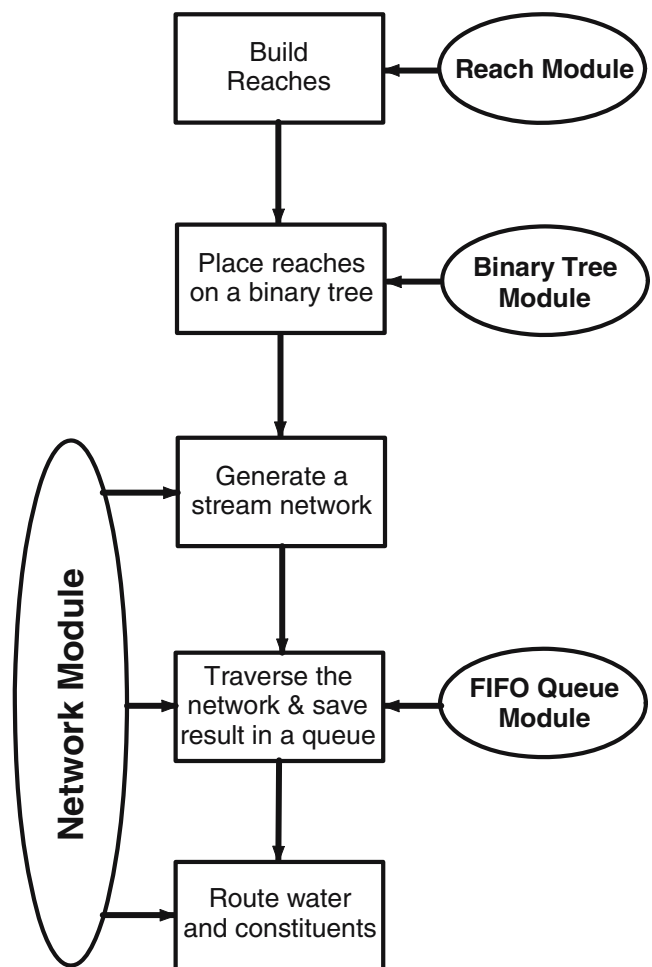


Fig. 3 The stream network model is composed of four basic modules and an application module

structure for storing and searching stream reach data [23, 34]. The tree contains one node for each reach, the location of which is determined by its unique three-part code. The binary search tree dynamically allocates storage for each node and places the node into its proper position so that the set of reaches read in from a file is sorted at all times. As a result, the code in any given node is lexicographically greater than any codes in the node's left subtree, but less than any codes in the right subtree (Fig. 5). This data structure makes it possible to accommodate varying sizes of stream networks in different watersheds.

To connect a stream reach with its associated ones, up to five searches are needed to locate them using the navigation attributes (Fig. 4). Search time and efficiency become significant issues as the size of a stream network becomes large. With the reach data stored in a binary search tree, searching is conducted with an efficient binary search algorithm. For any search, the algorithm starts at the root of the tree and compares the search code with the code stored at the root. If the two codes match, the search is finished. If the search code is less than the code at the root, searching goes to the left subtree, otherwise to the right subtree. In either case, the next step is to compare the search code with that stored in a child node (left or right). This recursive process of dividing the tree in two continues until a stream reach with the matching code is found or the entire tree has been searched. In the latter case, the reach being searched for is not in the tree [23, 34].

2.3 Stream Network Generation and Traversal

A stream network is generated by connecting each stream reach with its associated ones (Fig. 4). In the simulation program, a pointer is specified for each of the associated reaches. Once one is located using the binary search algorithm, it is connected to the instant reach by setting its designated pointer to its address. When every stream reach in the watershed is connected to its associated ones, stream network generation is completed.

Fig. 4 Reach connectivity in Reach File 3. For each instant reach (the reach under consideration), there are up to five associated reaches

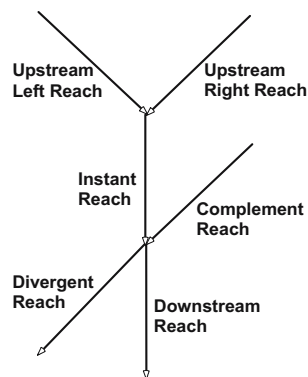
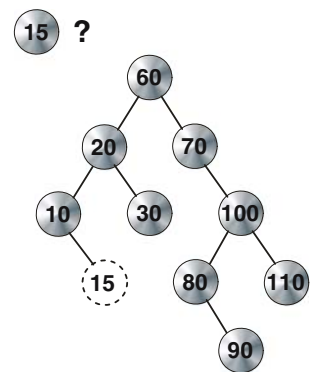
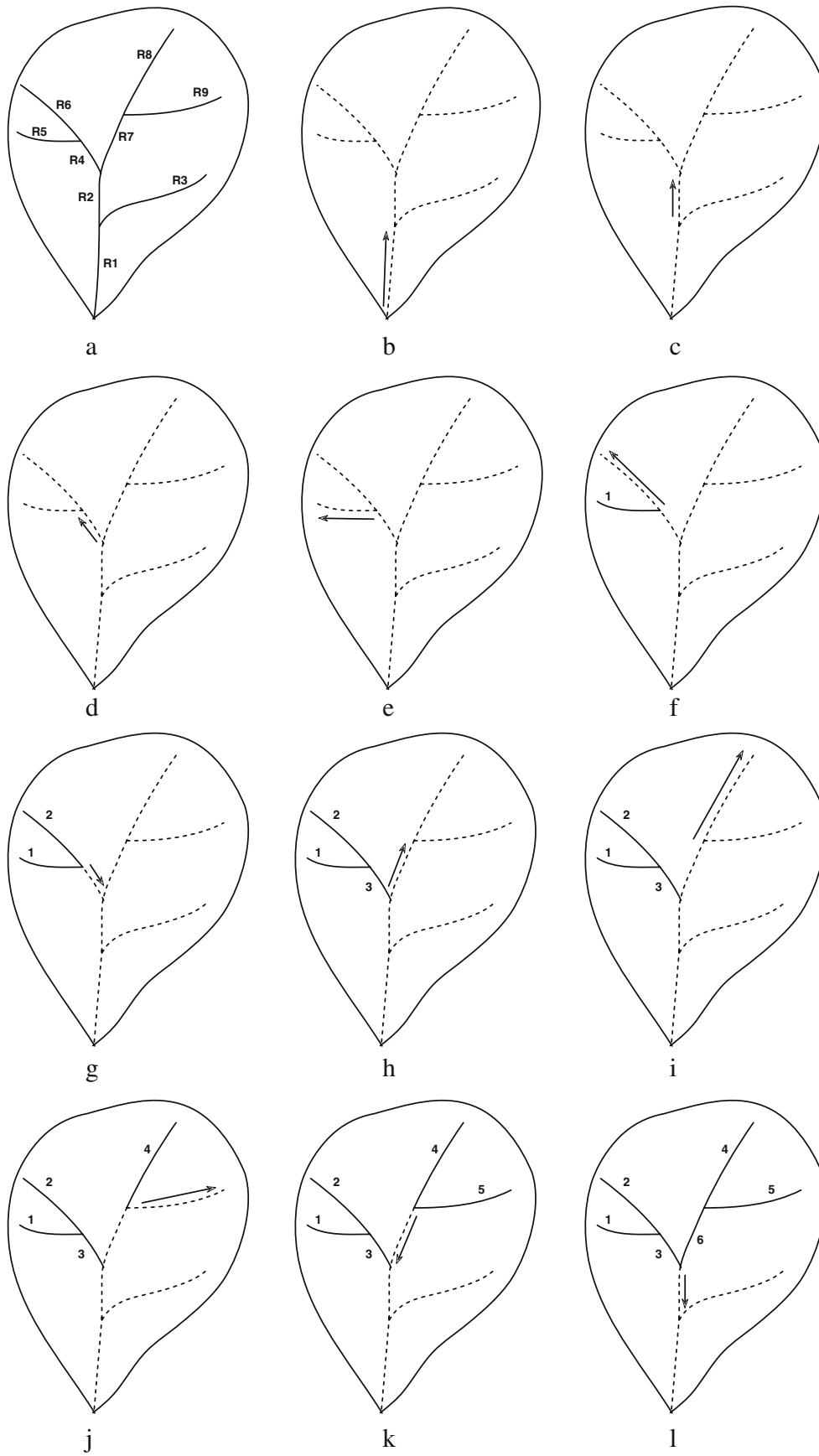


Fig. 5 An example of a binary search tree. The identification number in a node determines its position in the tree. For any node, all the nodes in its left subtree have smaller identification numbers, and all the nodes in its right subtree have larger numbers. In our model, the identification number is the reach code, which is a combination of three numbers: the catalog unit, the segment number, and the marker index



The model then traverses the stream network to establish a routing sequence that correctly moves water and materials from upstream to downstream reaches. The traversal algorithm begins with the stream network outlet and moves upstream. An example of the traversal process is given in Fig. 6, which contains nine stream segments labeled R1 through R9 (Fig. 6a). These labels are for identification only and are not stream orders. The traversal algorithm is a recursive process applied to each reach in the stream network in the same way. The algorithm can be simply described as first left, then right, and finally the current one. That is, for any given reach, the sequence of traversal is always first to its upper left, then to its upper right, and finally itself. If the upper left one is not a headwater stream, the algorithm continues to move up to the left until a headwater stream is found. When that occurs, the position of the stream reach in the routing sequence (shown by a number in Fig. 6) is saved in a queue (described below). The algorithm then checks the upper right one. If it is a headwater stream, it is placed in the queue. Then the immediate downstream reach is saved in the queue. For any given reach, only after all the reaches upstream of it are already in the queue will the one in question be placed in the queue. The stream reach at the outlet of the network will be the last one to be saved in the queue (the last one in the routing sequence). The numbers in Fig. 6 indicate the order in which these reaches are placed in a queue and in which they will be processed as water and materials are routed through the stream network. A downstream reach always has a larger number than all the ones upstream of it. In the example in Fig. 6, the outlet reach R1 has the largest number (nine) among the nine reaches, which means that it is the last one in the routing sequence.

Fig. 6 An example stream network with stream reaches identified as R1, R2,..., R9 (a). The traversal of the stream network is shown from (b) to (o), which determines the order of routing (indicated by numbers 1 to 9, where a larger number means later routing). A downstream reach always has a larger number than all of its upstream reaches, and the outlet reach has the largest number (o). The routing order is saved in a queue (p)



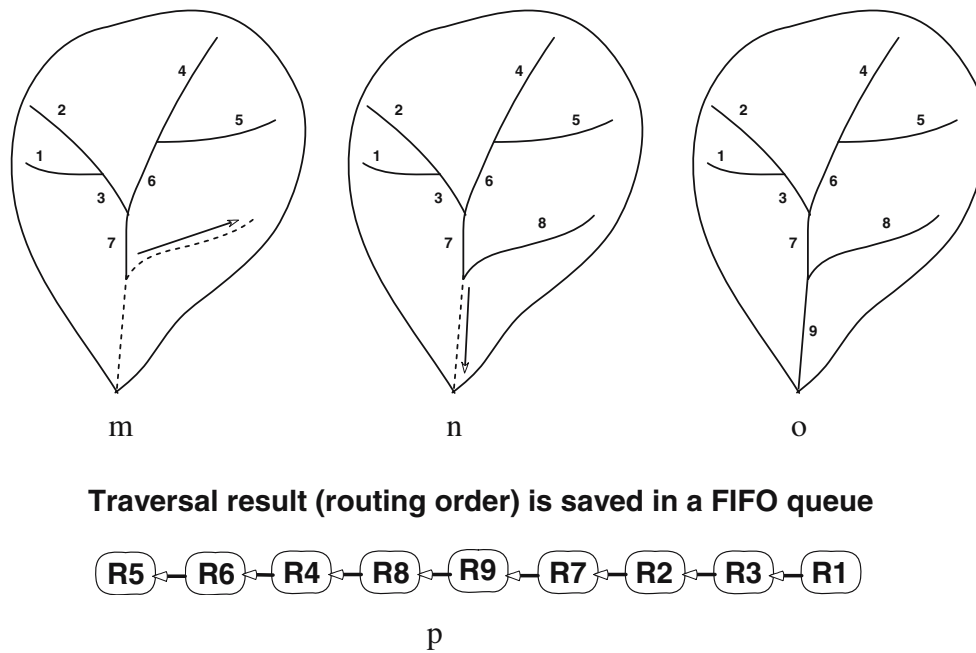


Fig. 6 (continued)

Similar methods of traversing a stream network have been used in other hydrological studies [7]. The purpose of traversing the network in our model is to order the stream reaches so that each one will be placed after all of its upstream reaches in the simulation sequence. There can be many different sequences that meet this need, and the sequence resulting from our traversal procedure is just one of the correct sequences. For example, the four headwater reaches (R5, R6, R8, and R9) in Fig. 6 can be arranged in 24 different ways that all keep the downstream ones in the same order, such as (R5, R6, R8, R9, R4, R7, R2, R3, R1), (R6, R5, R8, R9, R4, R7, R2, R3, R1), (R5, R6, R9, R8, R4, R7, R2, R3, R1), and (R8, R6, R5, R9, R4, R7, R2, R3, R1). Any method of establishing such a sequence can be used in the model. We use the recursive traversal mainly because of its simple implementation and computational efficiency.

2.4 Saving the Routing Sequence: The FIFO Queue

The structure of a stream network does not change, so it only needs to be traversed once to determine the routing sequence, which is saved for repeated use in a data structure called a first-in–first-out (FIFO) queue. Such a queue holds an ordered sequence of items. The two basic operations of a FIFO queue are enqueue and dequeue [34]. The enqueue operation adds a new item at the end of the queue, whereas the dequeue operation removes an item from the front of the queue.

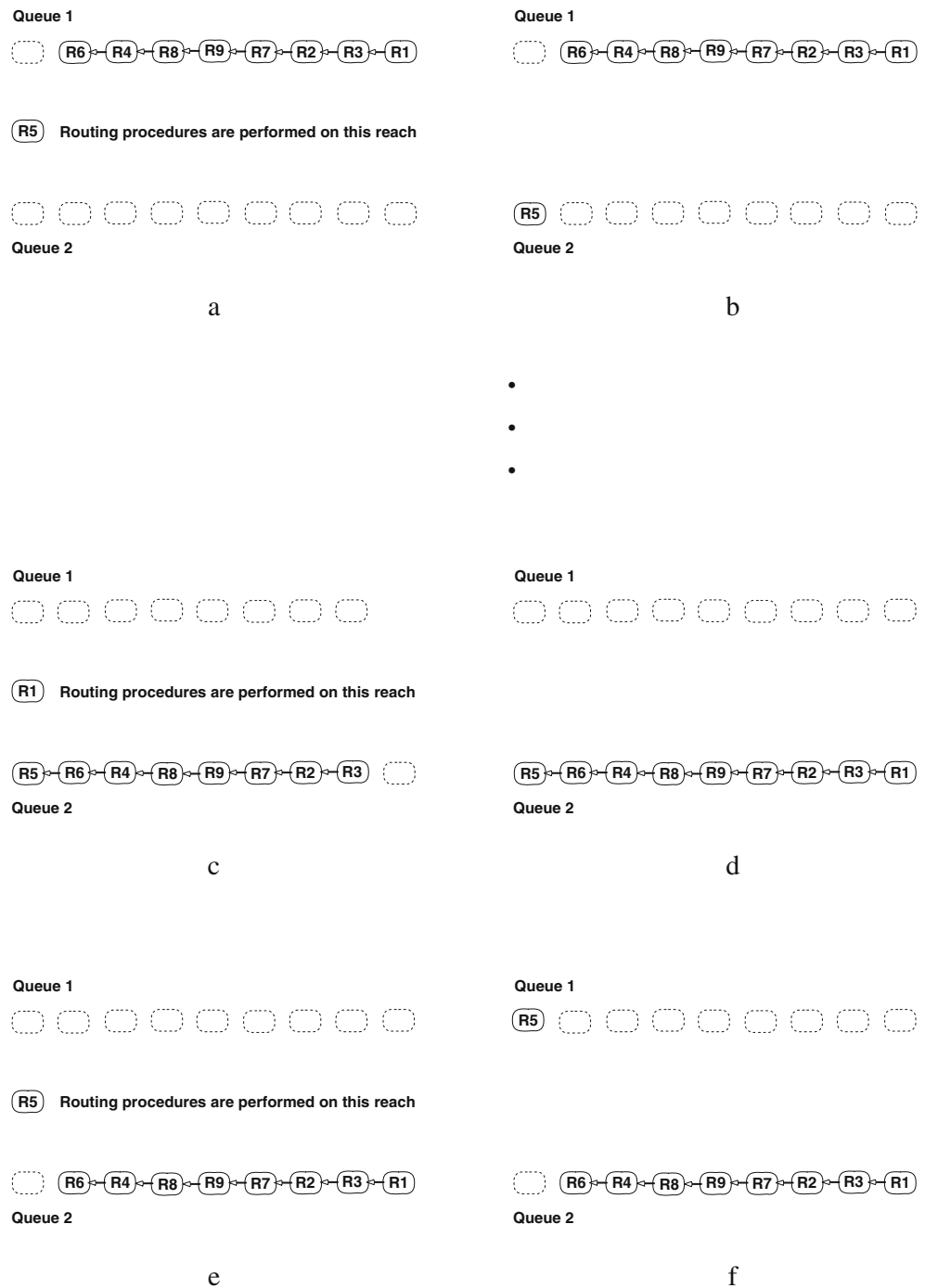
To keep the ordered sequence unchanged throughout a simulation, two FIFO queues are used (Fig. 7). At the beginning of each round of routing, one queue is filled with

reaches, and the other is empty. Routing occurs one reach at a time: A reach is taken out of one queue (dequeued); routing is performed on it; and then it is put into the other queue (enqueued). This process continues until all the reaches have been taken out of one queue and put into the other (Fig. 7), finishing one round of routing. The two queues reverse their roles when the next round of routing begins. The queue previously filled with reaches is now empty; the other now contains all the reaches in the same order. The process is repeated as many times as needed to complete a simulation (Fig. 7). To our knowledge, this queue structure has not been used in any other hydrological models.

2.5 Integrating Point and Nonpoint Sources

The stream network model simulates the routing of streamflow and selected water pollutants, and the contributions of point and nonpoint sources are integrated by the routing process. Each point source is anchored to a particular reach and treated as one of its inputs, but the contribution of nonpoint source inputs depends on how finely a watershed is partitioned. In theory, the contributing area of a stream reach can be defined as a subwatershed draining directly to it; that is, it has its own subwatershed. In practice, a watershed is often partitioned into larger units so that each subwatershed contains more than one stream reach. In this case, the nonpoint source contribution is estimated using the ratio of the reach length to the total length of all reaches in the subwatershed.

Fig. 7 Two first in–first out (FIFO) queues are used to preserve the routing order. The elements in these example queues are the reaches (R1, R2,..., R9) shown in Fig. 6. At the beginning of each round of routing, one queue is full of reaches and the other is empty. One reach at a time is taken from one queue and put into the other after routing procedures are performed on reaches (a) to (d). The two queues reverse their roles when the next round of routing begins. The one previously filled with reaches is now empty; the other now contains the same reaches in the same order (d). The same process is repeated as many times as needed to complete a simulation (e, f)



The output from a reach as a result of the routing process provides one of the inputs to its immediate downstream reach. Thus, contributions from point and nonpoint sources in the study watershed are integrated, on a reach-by-reach basis, in the process of routing through the stream network. In the simulation program, the routed water, sediment, and other water pollutants are defined as additional attributes in the reach data structure and updated in each round of routing.

3 Stream Network Routing

During a dynamic simulation of stream transport, water and materials must be moved (“routed”) through the stream network at each time step. Except for headwater streams, the stream network model calculates outflow at the downstream end of a stream reach from three inputs: the inflow at its upstream end, nonpoint source runoff from the subwatershed draining to it, and effluent from any point sources. The

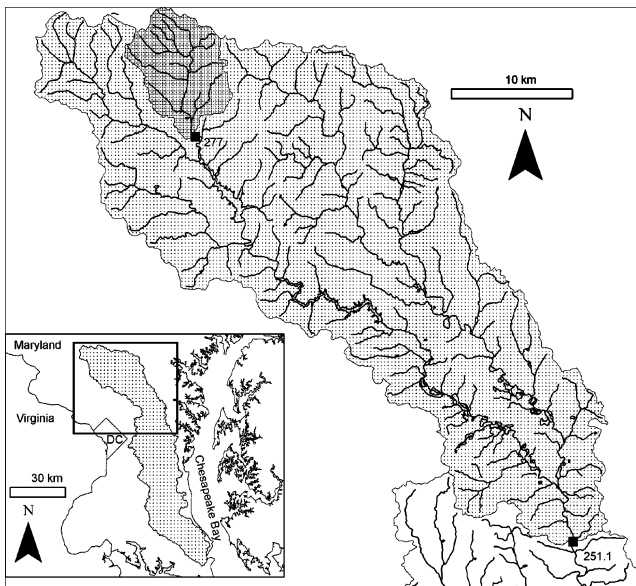


Fig. 8 The location of the Patuxent River watershed. The two subwatersheds draining to sampling stations 277 and 251.1 are outlined and stippled

algorithm also accounts for time lags due to channel storage, which may delay discharge until a later time. The chain of calculations for each time step ends at the outlet of the stream network, where the calculated outputs are the outputs from the entire stream network and its entire watershed.

The flux of water and materials from a reach can be described in the following mass balance equation:

$$O = I_{NPS} + I_{PS} + F - R \quad (1)$$

where O is the output, I_{NPS} is the nonpoint source input from the contributing subwatershed, I_{PS} is the input from point sources, F is the flux from the immediate upstream reaches, and R is the removal or loss due to in-stream processes.

When the loss term (R) in the equation is negligible, the output from a stream reach equals its total input. If the reach is long, resulting in a significant time lag between input and output, the output at any time can be described by the output and input a time interval earlier. When the loss is not negligible, the proportion of input that is removed by in-stream processes can be described by a decay function. We implement Eq. 1 in two computational steps. First, a routing procedure is applied without considering in-stream loss. Then, the output from the routing procedure supplies the input to a decay function. The result from the second step is the output from the stream reach in question.

3.1 Streamflow Routing

For streamflow, the in-stream loss, R in Eq. 1, is considered insignificant, so streamflow routing is simulated with a routing procedure only. There are many published algo-

rithms for routing water through streams, rivers, or reservoirs [27, 30, 35]. In our model, we implemented a simple algorithm called the convex method [30], which estimates the outflow leaving a reach at anytime from the outflow and inflow one time interval earlier, i.e.,

$$Q_t = cI_{t-\Delta t} + (1 - c)Q_{t-\Delta t} \quad (2)$$

where Q_t is the outflow at time t ; c is a constant, $0 \leq c \leq 1$; and the inflow and outflow one time interval earlier are $I_{t-\Delta t}$ and $Q_{t-\Delta t}$, respectively. Thus, the outflow at any time is the weighted average of the inflow and the outflow one time interval earlier, and the outflow must be between $I_{t-\Delta t}$ and $Q_{t-\Delta t}$ [30]. The constant c depends on the time interval Δt and stream reach length. The number of times routing is performed per day is the reciprocal of the time interval. For example, if the time interval is 1/4 day, routing will be performed four times per day.

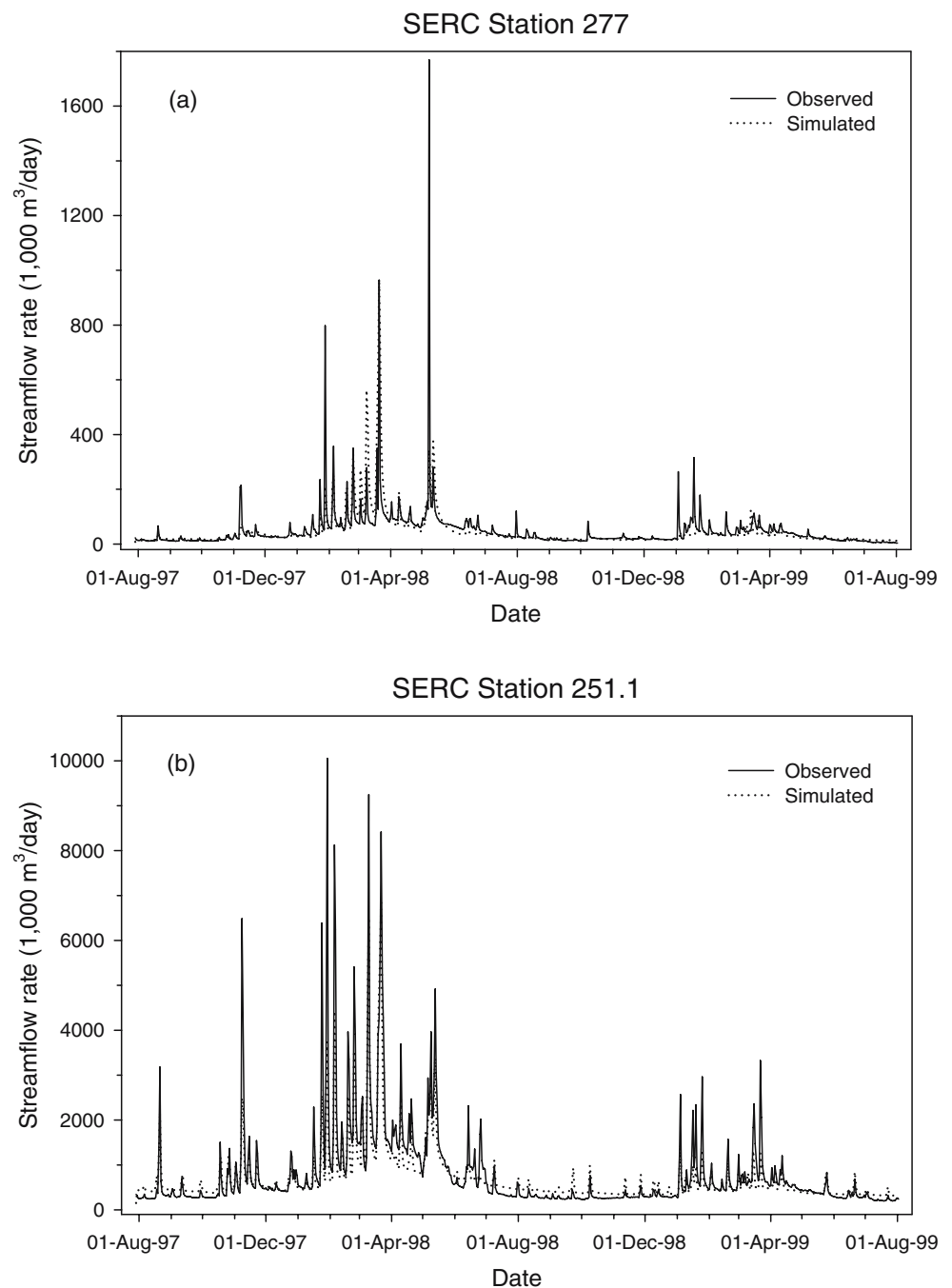
It is worth noting that this model was not designed to test or implement state-of-the-art mechanistic representations of stream channel transport. Instead, we sought to develop a flexible modeling framework with simple algorithms that can be applied with appropriate detail for entire watersheds. The convex algorithm is presented here as an example to illustrate the routing process. The simulation program could be easily modified to implement a different routing algorithm or offer a selection of several different algorithms.

3.2 Material Routing

In most practical problems, the in-stream loss of material, i.e., R in Eq. 1, is not negligible. The proportion of material entering a stream reach that is removed by in-stream processes is a function of many factors and varies among constituents. For plant nutrients, such as nitrogen, the proportion of input removed can be described as a function of water depth and travel time (residence time). The fraction of nutrient removed is inversely related to stream channel depth and positively related to water residence time [19]. Data on depth and residence time are not included in Reach File 3; however, depth can be described as a function of flow rate [3], and residence time can be described as the ratio of reach length to flow velocity. Data on reach length and flow rate are available, but velocity data are not. If velocity does not vary considerably, the fraction of nutrient removed by in-stream processes can be expressed as a function of reach length and flow rate. Different forms of decay functions can be found in the literature [41, 29, 28], and we implemented an exponential decay function as follows:

$$O_t = I_t \exp(-kL/Q_t) \quad (3)$$

Fig. 9 Simulated and observed daily streamflows at SERC sampling stations 277 (a) and 251.1 (b)



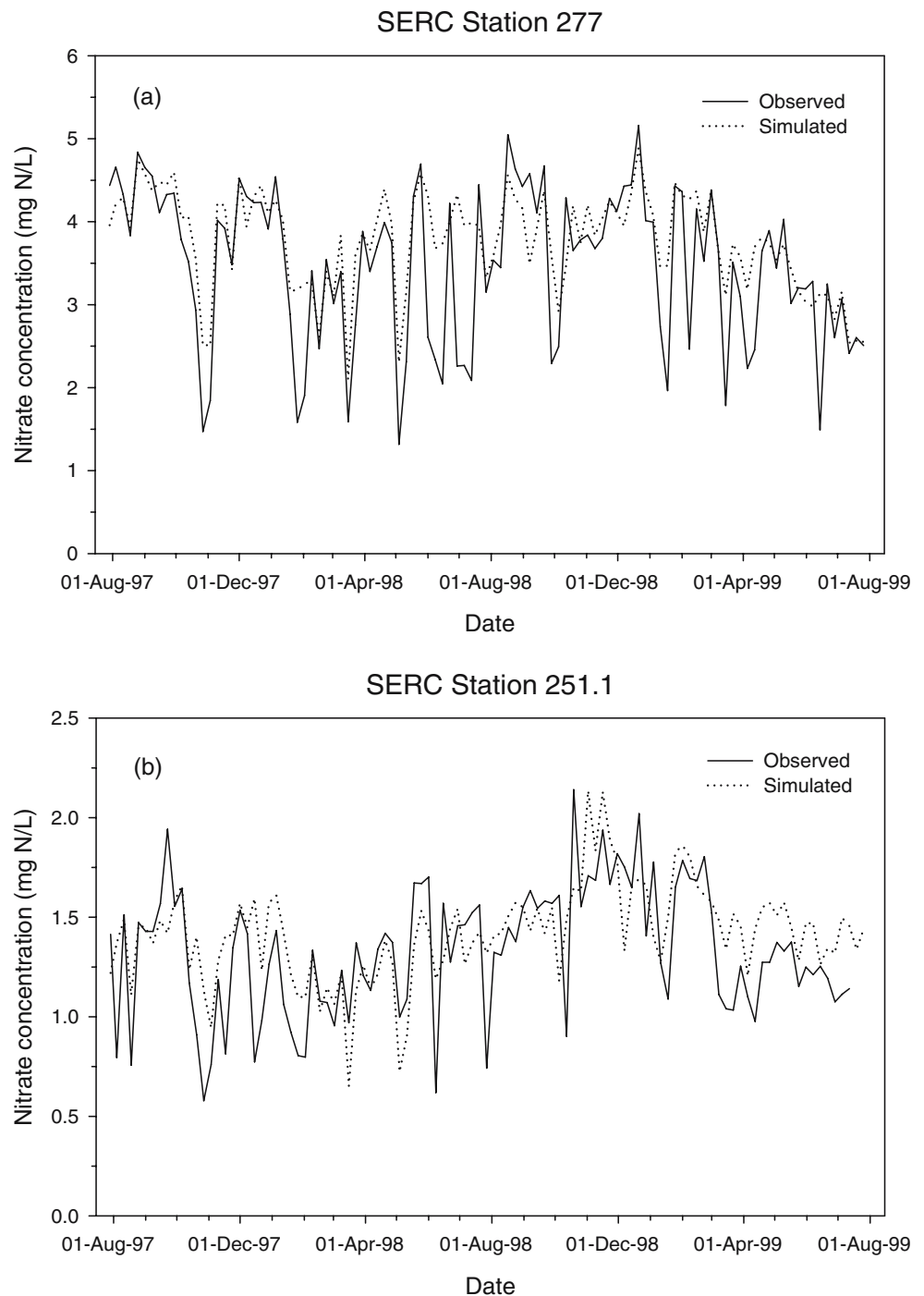
where O_t is the flux of material from the reach at time t , I_t is the input of material at time t , k is the decay parameter, L is the reach length, and Q_t is the flow rate (discharge) at time t .

Material routing is estimated in two steps. For each stream reach, routing is performed with the convex method; then output from the routing procedure supplies the input to the decay function, i.e., I_t in Eq. 3. The difference between influx (I_t) and flux (O_t) in Eq. 3 is the in-stream loss term R in Eq. 1. The end result from these two procedures, O_t in Eq. 3, is the flux of material from the reach in question, i.e., the output (O) in Eq. 1.

4 Model Application

The stream network model was applied to the watershed of the Patuxent River, a tributary to the Chesapeake Bay near the cities of Washington, DC, and Baltimore, MD (Fig. 8). The Smithsonian Environmental Research Center (SERC) monitored the discharges of water, sediment, and plant nutrients from study watersheds in the Patuxent drainage for 2 years from 1997 to 1999 [22, 40]. The monitoring data from SERC and USGS were used to validate this model. The validation process consists of two steps. The

Fig. 10 Simulated and observed weekly nitrate concentrations at SERC sampling stations 277 (a) and 251.1 (b)



first step verifies the topology of the generated stream network and the connectivity of subwatersheds to the network, and the second step compares simulated flow and nutrient concentrations with observed data.

4.1 Topology of the Generated Stream Networks

To verify the stream network topology, model-generated stream networks were compared with the Reach File maps

using a geographic information system (GIS). The routing order of each generated stream network, as stored in a queue in the simulation, was examined in comparison with the network map in the GIS. The connections of subwatersheds to their corresponding anchor reaches were also verified. In addition, for each simulated watershed, we compared the total water inputs from all the point and nonpoint sources in the watershed with the simulated streamflow at the watershed outlet. For this purpose, the

parameter c in the routing function (Eq. 2) was set at 1, so there was no delay in flow routing; the parameter k in the decay function (Eq. 3) was set at 0, so there was no loss of water. With these parameter values, the principle of conservation of mass requires that the daily streamflow at the outlet of a watershed equal the daily sum of the inputs from all the point and nonpoint sources in the watershed. Failure to achieve mass balance would indicate errors in the topology of the generated stream network or in connecting subwatersheds to their anchor reaches. The model was successfully verified using these criteria.

4.2 Simulated vs. Observed Data

We further tested the model by comparing simulated predictions of streamflow and nitrate concentrations with observed data from SERC study watersheds and USGS gauging stations [22]. Nonpoint source inputs from subwatersheds to the stream network were simulated using two other models. Runoff was predicted by TOPMODEL [13, 14], and the nitrate concentration was estimated from precipitation, streamflow, geology, land use, and seasonality using an empirical regression model [40]. The Maryland Department of the Environment provided monitoring data on point source inputs of water and nitrate.

The stream network model was run to predict streamflow and nitrate concentrations at the outlets of watersheds draining to two SERC sampling stations numbered 277 and 251.1 (Fig. 8). The former is a small, low-order watershed (58 km²), and the latter is a large mainstem watershed (908 km²) discharging to the head of the Patuxent estuary. Watershed 251.1 contains 79 subwatersheds and 15 point sources, whereas watershed 277 has 19 subwatersheds with no point sources. The runoff and nutrient inputs from subwatersheds and the point source inputs are saved in separate files and retrieved one record at a time during the simulation. A USGS gauging station is located near each of the two SERC stations (SERC 277 is USGS 01591400 and SERC 251.1 is USGS 01594440). Daily simulated predictions were directly compared to USGS measured daily

flows (Fig. 9). Because nutrient concentrations were measured weekly, weekly average simulated nitrate concentrations were first calculated from daily predictions and then compared to nitrate concentrations measured by SERC (Fig. 10). Overall, the simulated flows and nitrate concentrations were reasonably close to their observed values. The correlations of the observed and predicted flows were 0.93 and 0.63 for stations 251.1 and 277, respectively. The correlations of observed and predicted nitrate concentrations were 0.61 and 0.76 for the two stations. The main discrepancies between predicted and observed time series data are in the peak values and the lowest values. Predicted peak flows are lower and base flows are higher than their observed counterparts. A similar observation can be made for the nitrate concentrations. This pattern of discrepancy between the predictions of the stream network model and observed data appears inherited from its two feeding models, TOPMODEL and the nutrient model, because similar patterns of difference in predicted and observed high and low values were also observed for them even before their outputs were processed by the stream network model.

In addition to time series data, we also compared the total observed and simulated flow volumes and nitrate fluxes at both stations for the whole simulation period (Table 1). We used the general guidelines for evaluating simulation results of the widely used Hydrologic Simulation Program – Fortran model (HSPF) [9, 18] to assess the performance of our model. According to those guidelines, our predictions of streamflow and nitrate for the two watersheds are either very good (less than 10% difference between observed and simulated values) or good (less than 15% difference).

5 Discussion

The stream network model was developed to be the core of a flexible, efficient, modular watershed simulation in which different sets of processes occurring in a watershed can be simulated separately with different models. Our watershed simulation framework is similar to that of White et al. [41] and Smith et al. [32]. However, White et al. did not include the idea of using different models for different sets of processes, and SPARROW predicts total nitrogen or phosphorus fluxes over annual or greater time intervals rather than simulating fine-scale temporal dynamics.

Simulation accuracy depends on the quality of the input data as well as the model itself. The accuracy of our flow predictions at the stream network outlet is mainly determined by the accuracy of the TOPMODEL estimates of runoff from subwatersheds to the stream network. The

Table 1 Modeled and measured 2-year water and nitrate discharges from two study watersheds

Watershed	Data type	Streamflow (10 ⁶ m ³)	Nitrate flux (10 ³ kg)
SERC 277	Observed	35.7	123.5
	Simulated	33.6	114.6
	Difference (%) ^a	5.9	7.2
SERC 251.1	Observed	586.8	692.2
	Simulated	521.8	640.6
	Difference (%) ^a	11.1	7.5

^a Difference (%) = 100 × (observed – simulated) / observed

accuracy of predicted nutrient concentrations depends on the nutrient inputs to the stream network, which are predicted by the nutrient regression model, and on the in-stream loss, which is simulated in the stream network model. When the rate of in-stream loss is small compared to the subwatershed inputs, the proportion of prediction errors that can be attributed to the stream network model is small. On the other hand, when the rate of in-stream loss is high, the contribution of this model to the prediction errors could be substantial if it were not appropriately calibrated.

The stream network model is simple and computationally efficient, and can be applied wherever Reach File maps are available with no constraint on the size of stream network; as such, it provides the kind of modeling tool needed for environmental assessment of broad landscapes. At the time of model development, the Reach File Version 3 (RF3) was the most advanced digital hydrographic map that provided topological information for stream routing. Since then, the newer NHD has integrated the information from RF3 with other USGS maps [39] and introduced a slightly different representation of topology. With minor changes, the stream network simulation program can be easily adapted to use that new topological information. The availability of RF3 and NHD has also fostered the development of GIS data models especially designed for surface water resources, such as Arc Hydro for Environmental Systems Research Institute's ArcGIS [26]. The stream network model could be made more broadly applicable by integrating it with a GIS that supports Arc Hydro or other similar data structures.

In general, a watershed can be compartmentalized spatially or topologically. Depending on the purpose of research, the compartmentalization can be accomplished in different ways, and spatial processes in different compartments can be simulated with different component models. A watershed simulation system designed this way facilitates model development and customization because components are usually loosely coupled with one another and can be added or removed without much effort. Needed components may already exist or can be developed efficiently, which are usually simple and designed to simulate particular sets of watershed processes, and often perform as well as or better than more complex comprehensive models.

When a different compartmentalization of watershed is required, different components can be used and coupled in different order. For example, to study the effects of riparian buffers on water quality at watershed scale, a model that simulates riparian processes, such as the Riparian Ecosystem Management Model (REMM) [4], can be added to our conceptual watershed simulation framework. The outputs of runoff and nutrients from subwatersheds would then become inputs to REMM, instead of the streams, and the outputs from the riparian model would become the inputs to

the streams. Because of their simplicity and flexibility, watershed models composed of simple component modules will see more development in the future than the traditional complex comprehensive watershed simulation systems.

Acknowledgements This paper results from our participation in the COASTES project [Complexity and Stressors in Estuarine Systems, funded by the National Oceanic and Atmospheric Administration (NOAA) Coastal Ocean Program (COP), award number NA66RG012]. Additional funding was provided by NOAA-COP (award number NA03NOS4780008). Michelle Coffee, Carolyn Lieberman, Jennifer Bruggink, Kieren Tinning, Marcia Snyder, Nathan Bowden, and Nancy Lee provided valuable assistance with the analysis of watershed geography. The authors would like to thank an anonymous reviewer, whose comments and suggestions helped us improve the final version of this manuscript.

References

- Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986). An introduction to the European Hydrological System – Systeme Hydrologique European, "SHE", 1: history and philosophy of a physically-based, distributed modeling system. *Journal of Hydrology*, 87, 45–59.
- Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., & Rasmussen, J. (1986). An introduction to the European Hydrological System – Systeme Hydrologique European, "SHE", 2: structure of a physically-based, distributed modelling system. *Journal of Hydrology*, 87, 61–77.
- Alexander, R. B., Smith, R. A., & Schwarz, G. E. (2000). Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. *Nature*, 403, 758–761.
- Altier, L. S., Lowrance, R., Williams, R. G., Inamdar, S. P., Bosch, D. D., Sheridan, J. M., et al. (2002). *Riparian Ecosystem Management Model: Simulator for Ecological Processes in Riparian Zones* (USDA, Agricultural Research Service, Conservation Research Report 46) 216 pp.
- Arnold, J. G., Srinivasan, R., Mutiah, R. S., & Williams, J. R. (1998). Large area hydrologic modeling and assessment part I: model development. *Journal of the American Water Resources Association*, 34, 73–89.
- Arnold, J. G., Williams, J. R., Nicks, A. D., & Sammons, N. B. (1990). *SWRRB: A basin scale simulation model for soil and water resources management*. College Station, Texas: Texas A & M University Press.
- Band, L. E. (1989). A terrain-based watershed information system. *Hydrological Processes*, 3, 151–162.
- Beasley, D. B., Huggins, L. F., & Monke, E. J. (1980). ANSWERS: A model for watershed planning. *Transaction of the ASAE*, 23, 938–944.
- Bergman, M. J., Green, W., & Donnangelo, L. J. (2002). Calibration of storm loads in the South Prong watershed, Florida, using BASINS/HSPF. *Journal of the American Water Resources Association*, 38, 1423–1436.
- Beven, K. J. (1989). Changing ideas in hydrology—the case of physically-based models. *Journal of Hydrology*, 105, 157–172.
- Beven, K. J. (2001). *Rainfall-runoff modelling: The primer*. Chichester, England: John Wiley (360 pp.).
- Beven, K. J. (2001). On modelling as collective intelligence. *Hydrological Processes*, 15, 2205–2207.

13. Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24, 43–69.
14. Beven, K. J., Lamb, R., Quinn, P., Romanowicz, R., & Freer, J. (1995). TOPMODEL. In V. J. Singh (Ed.), *Computer models of watershed hydrology* (pp. 627–668). Highlands Ranch, Colorado: Water Resources Publications.
15. Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, W. R., Sharpley, A. N., & Smith, V. H. (1998). Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecological Application*, 8, 559–568.
16. Diplis, P. (2002). Integrated decision making for watershed management: introduction. *Journal of the American Water Resources Association*, 38, 337–340.
17. Donigian, A. S., Bicknell, B. R., & Imhoff, J. C. (1995). Hydrological simulation program – FORTRAN (HSPF). In V. P. Singh (Ed.), *Computer models of watershed hydrology* (pp. 395–442). Highlands Ranch, CO: Water Resources Publications.
18. Donigian, A. S., Imhoff, J. C., Bicknell, B. R., & Kittle, J. L. (1984). *Application guide for the hydrological simulation program – FORTRAN*. EPA 600/3-84-066, Environmental Research Laboratory, U.S. EPA, Athens, GA.
19. Horn, C. B., & Grayman, W. M. (1993). Water-quality modeling with EPA Reach File system. *Journal of Water Resources Planning and Management*, 119, 262–274.
20. Howarth, W. A., Anderson, D., Cloern, J., Elfring, C., Hopkinson, C., Lapoint, B., et al. (2000). Nutrient pollution of coastal rivers, bays, and seas. *Issues in Ecology*, 14, 433–437.
21. Johansen, N. B., Imhoff, J. C., Kittle, J. L., & Donigian, A. S. (1984). *Hydrological simulation program – FORTRAN (HSPF): User's manual for release 8*. EPA-600/3-84-066. U.S. Environmental Protection Agency, Athens, GA.
22. Jordan, T. E., Weller, D. E., & Correll, D. L. (2003). Sources of nutrient inputs to the Patuxent River estuary. *Estuaries*, 26, 226–243.
23. Kernighan, B. W., & Ritchie, D. M. (1988). *The C programming language*, 2nd ed. Englewood, NJ: Prentice Hall (272 pp.).
24. Leavesley, G. J., Restrepo, P. J., Markstrom, S. L., Dixon, M., & Stannard, L. G. (1996). *The modular modeling system – MMS: User's manual*. U.S. Geological Survey Open File Report 96-151) 200 pp.
25. Leavesley, G. J., Restrepo, P. J., Stannard, L. G., Frankoski, L. A., & Sautins, A. M. (1996). The modular modeling system (MMS) – A modeling framework for multidisciplinary research and operations. In M. Goodchild, L. Steyaert, B. Parks, M. Crane, M. Johnston, D. Maidment, & S. Glendinning (Eds.), *GIS and environmental modeling*. (pp. 155–158). Ft. Collins, CO: GIS World Books.
26. Maidment, D. R. (2002). *Arc hydro: GIS for water resources*. Redland, CA: ESRI (203 pp.).
27. McCuen, R. H. (1998). *Hydrologic analysis and design*, 2nd ed. Englewood, NJ: Prentice Hall (814 pp.).
28. Santhi, C., Arnold, J. G., Williams, J. R., Dugas, W. A., Srinivasan, R., & Hauck, L. M. (2001). Validation of the SWAT model on a large river basin with point and nonpoint sources. *Journal of the American Water Resources Association*, 37, 1169–1188.
29. Seitzinger, S. P., Styles, R. V., Boyer, E. W., Alexander, R. B., Billen, G., Howarth, R. W., et al. (2002). Nitrogen retention in rivers: model development and application to watersheds in the northeastern U.S.A. *Biogeochemistry*, 57/58, 199–237.
30. Singh, V. P. (1992). *Elementary hydrology*. Upper Saddle River, NJ: Prentice Hall (973 pp.).
31. Singh, V. P. (ed.) (1995). *Computer models of watershed hydrology*. Highlands Ranch, Colorado: Water Resources Publications (1130 pp.).
32. Smith, R. A., Schwarz, G. E., & Alexander, R. B. (1997). Regional interpretation of water-quality monitoring data. *Water Resources Research*, 33, 2781–2798.
33. Srinivasan, R., Ramanarayanan, T. S., Arnold, J. G., & Bednarz, S. T. (1998). Large area hydrologic modeling and assessment, part II: model application. *Journal of the American Water Resources Association*, 34, 91–101.
34. Stubbs, D. F., & Webre, N. W. (1989). *Data structures with abstract data types and pascal*. Pacific Grove, CA: Brooks/Cole Publishing Co (471 pp.).
35. Thomann, R. V., & Mueller, J. A. (1987). *Principles of surface water quality modeling and control*. New York: Harper & Row (644 pp.).
36. USEPA (U.S. Environmental Protection Agency) (1994). *EPA reach file version 3.0, alpha release (RF3-Alpha), technical reference*. Washington, DC: USEPA (23 pp.).
37. USEPA (U.S. Environmental Protection Agency) (1995) *EPA River Reach File Version 3.0, RF3-Final Database Design*. Washington, DC: USEPA (41 pp.).
38. USEPA (U.S. Environmental Protection Agency) (2000). *Revisions to the Water Quality Planning and Management Regulation and Revisions to the National Pollutant Discharge Elimination System Program in Support of Revisions to the Water Quality Planning and Management Regulation*. Federal Register 65 No. 135.
39. USGS (U.S. Geological Survey), National Hydrography Dataset. <http://nhd.usgs.gov>.
40. Weller, D. E., Jordan, T. E., Correll, D. L., & Liu, Z. J. (2003). Effects of land-use change on nutrient discharges from the Patuxent River watershed. *Estuaries*, 26, 244–266.
41. White, D. A., Smith, R. A., Price, C. V., Alexander, R. B., & Robinson, K. W. (1992). Spatial model to aggregate point-source and nonpoint-source water-quality data for large areas. *Computers and Geosciences*, 18, 1055–1073.
42. Young, R. A., Onstad, C. A., Bosch, D. D., & Anderson, W. P. (1989). AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. *Journal of Soil and Water Conservation*, 44, 168–173.