International Journal of Remote Sensing

May, 2005

An integrated modeling system for management of the Patuxent River estuary and basin, Maryland, USA.

M. R. Williams[1], T. R. Fisher[1], W. R. Boynton[3], C. F. Cerco[4], M. W. Kemp[1], K. N. Eshleman[2], S.-C. Kim[4], R. R. Hood[1], D. A. Fiscus[2], and G. R. Radcliffe[1]

[1]University of Maryland, Center for Environmental Sciences, Horn Point Laboratory, 2020 Horn Point Rd, Cambridge, MD 21613 USA, email: williams@hpl.umces.edu, fax: (410) 221-8390

[2]University of Maryland, Center for Environmental Sciences, Appalachian Laboratory, 301 Braddock Rd, Frostburg, MD 21532 USA

[3]University of Maryland, Center for Environmental Sciences, Chesapeake Biological Laboratory, One Williams Street, Solomons, MD 20688 USA

[4]Engineering Research and Development Center, Waterways Experiment Station, Halls Ferry Road, Vicksburg, MS 93180 USA

Abstract

The Patuxent River watershed is a heavily impacted basin (2290 km^2) and estuarine tributary (120 km^2) of the Chesapeake Bay, USA. To assist management of the basin, we are testing a coupled modeling system composed of a watershed model (HSPF), an estuarine circulation model (CH3D), and an estuarine water quality model (CE-QUAL-ICM). The modeling system is being tested to guide the development of Total Maximum Daily Loads (TMDLs), and therefore errors in the models must be carefully evaluated. A comparison of daily total nitrogen (TN) concentrations simulated in HSPF with observations indicated that there was no significant bias, with rms error = 37%. In contrast, modeled total phosphorus (TP) and total suspended solids (TSS) had significant bias with larger rms errors (65% and 259%, respectively). In the estuary, CH3D accurately simulated tides, temperature, and salinity. CE-QUAL-ICM overestimated nitrogen (N) and phosphorus (P) in the upper estuary and underestimated in the lower estuary, primarily because intertidal marshes are not currently a model component. Model errors declined from short ($\leq$ 1 day) to long (multi-year) time scales as under- and overestimations cumulatively canceled. Watershed model errors propagate into the estuarine models, interacting with each subsequent model's errors, which limits the effectiveness of this TMDL management tool at short time scales.

Introduction

Increasing nutrient inputs to receiving waters have been associated with rising human population densities, changes in land use, and intensification of agricultural practices in many watersheds (Cole et al. 1993, Howarth et al. 1996, Galloway 1998, Boyer et al. 2002). Problems associated with water quality degradation are increasingly a threat to aquatic systems worldwide, particularly in estuaries (Valiela et al. 1992, Howarth et al. 2000). For example, nutrient enrichment from urban wastewater and agricultural runoff is responsible for excessive phytoplankton production, the decline of submerged aquatic vegetation, increasing abundance of nuisance algae blooms, and the increasing extent and duration of hypoxic and anoxic waters in the United States (Turner and Rabalais 1991, Vitousek et al. 1997, Boesch et al. 2001).

The Chesapeake Bay and its tributaries are part of this group of impacted estuaries. The Patuxent River, a tributary of the Chesapeake Bay lying entirely within Maryland, is being used as a model for testing environmental and management strategies without the sociopolitical complications associated with multi-state jurisdictional conflicts (Boynton et al. 1995). Extensive development and population increases in the Patuxent basin in the 1960s and 1970s increased wastewater discharges to the river, and the resulting pollution and ecological degradation of the Patuxent estuary resulted in agreements to regulate development and to remove nutrients from wastewater (D'Elia et al. 2003). For example, in the 1980s, municipal, state and federal agencies agreed to achieve a 40% reduction in nutrient loading to the Chesapeake Bay (including the Patuxent) by the year 2000. Upgrading wastewater plants resulted in large reductions in nutrient discharges to the aquatic environment in the upper Patuxent watershed, but these have not yet produced dramatic improvements in water quality in most of the estuary because of continuing and substantial diffuse source loads. There apparently needs to be further reductions in both point and nonpoint (NPS) nutrient discharges (D'Elia et al. 2003, Jordan et al. 2003, Weller et al. 2003) to improve water quality in the estuary.

Continuing population increases in the Patuxent basin will make this goal harder to achieve. There is continued development in the upper basin as residential communities in the Baltimore-Washington corridor as well as in the lower basin for tourism on the Bay. Development in both parts of the basin present challenges to water resource managers attempting to improve water quality in the Patuxent estuary. Managers will need a clear understanding of 1) nutrient sources, 2) quantitative predictions of the effects of increasing population and changing land use on nutrient loads, and 3) how these loads will be processed in the estuary to control water quality conditions.

Remote sensing is an important tool commonly used to accurately determine the land use/land cover (LULC) of watersheds (Weller et al. 2003, Williams et al. 2004). A practical application of remote sensing is to determine LULCs and include these as essential parameters in hydrochemical models. Conceptually, watershed hydrochemical models describe fluxes of water and associated sediment and nutrients from various land use types to rivers and estuaries (Krisanova et al. 1999), whereas estuarine models predict internal water fluxes, and nutrient transformations and transport. Because of the complexity of factors involved in water quality problems, simulation models have been widely used as tools to predict the known hydrological and biogeochemical processes controlling sediment and nutrient transport and dynamics in rivers and estuaries. Such models allow alternative land use or management scenarios to be evaluated, based on our current understanding of watershed and estuarine biogeochemistry.

The objective of this paper is to describe a linked modeling system for the Patuxent watershed and estuary. The modeling system includes a watershed model (HSPF), a 3-dimensional estuarine circulation model (CH3D), and a water quality model (CE-QUAL-ICM), which will be used to develop Total Maximum Daily Loads (TMDLs) for nutrient management of the Patuxent basin. If successful, these integrated models are potentially transferable to other subsystems of the Chesapeake watershed. Here we present initial analyses of model errors and describe several changes that have been incorporated into the estuarine models that improve the predictive capability of the integrated modeling system.

Study Site

The Patuxent River drains a 2290 km^2 watershed and is the sixth largest tributary of the Chesapeake Bay (figure 1). Estimates of year 2000 land cover in the watershed were determined from a 30 m x 30 m grid based on ETM+ imagery obtained from the Chesapeake Bay RESAC (S. Prince, PI, Geography Dept. University of MD). Land cover within the basin is 16% urban (part of the Baltimore-Washington metropolitan corridor), 20% agricultural, 64% forest, and 0.4% wetlands. The watershed has two reservoirs above the fall line that separates the Piedmont from the Coastal Plain (see figure 1), and these reservoirs supply drinking water for local municipalities, many of which are located outside of the basin (interbasin transfer). About 27% of the watershed is in the Piedmont above the fall line, and the remainder (73%) lies on the coastal plain. Nutrient inputs are from a variety of sources, including agriculture, industrial wastes, and over 120 000 m^3 d^{-1} of effluent from nine major sewage treatment facilities. Although nitrogen (N) inputs from sewage effluent increased about 10 times between 1963 and 1989, the removal of phosphorus (P) from effluent, due to the ban on detergent P in 1986 and by managed P removal, has reduced P concentrations in effluent by about a third (Magnien et al. 1992). Similarly, application of BNR technology (biological N reduction, which employs successive oxic and anoxic conditions to induce denitrification) has significantly reduced summer N concentrations in effluent by 75% (Boynton et al. 1995); however, increasing wastewater volume and high winter N concentrations continue to load the estuary with both N and P.

Methods

Model descriptions

Watershed hydrochemical model

Hydrological Simulation Program in Fortran (HSPF) is a comprehensive model that simulates hydrologic and biogeochemical processes in watersheds with pervious and impervious land surfaces, as well as within streams and well-mixed impoundments (Bicknell et al. 1997). The model can simulate urban and agricultural land-use, surface and subsurface processes, runoff, sediment export, and the fate and transport of nutrients, pesticides and other constituents. The model was developed over several decades by the Environmental Protection Agency (EPA) and is commonly used to assess the effects of land-use changes, flow diversions and point or non-point source treatment alternatives on the hydrology and water quality of watersheds. The strength of the model lies in its ability to continuously simulate the comprehensive range of hydrological and associated water quality processes in watersheds with complex land-use. The weakness of the model is that it is heavily parameterized, requiring considerable and often unknown values of inputs or parameters.

Structurally, HSPF is a lumped parameter model which is divided into three blocks. Each block simulates processes occurring in: 1) pervious land (e.g. forest, agriculture, etc.), 2) impervious land (e.g. high density urban), and 3) streams, lakes and reservoirs. HSPF is a complex, highly parameterized model, particularly for nitrogen processes. For example, in the pervious land block, there is a module that simulates the behaviour of nitrate, ammonium and organic N in four soil layers, and another module that simulates N in its various forms in the stream reaches. Transformations of N in pervious lands include plant uptake of inorganic forms, fixation, return to soil from plant tissues, immobilization, mineralization, nitrification, denitrification, adsorption/desorption of ammonium, and partitioning of organic N into dissolved and particulate forms, either as labile or refractory species. The N transformations are simulated individually in each of the four soil layers. Nitrogen that collects on impervious lands (e.g. parking lots, streets) from atmospheric deposition and urban activities is advected by overland flow to the aquatic module without transformation.

C, N, P and eroded soil are transported in the model by rain and groundwater to the aquatic module. Within the aquatic system, advected materials undergo further processing and transport in stream reaches. Several different routines are used to simulate inorganic N, P and total suspended solids (TSS) in the reaches. These include advection, exchanges between the sediments and the overlying waters, nitrification and denitrification, adsorption and desorption, and ionization and volatilization of ammonia. Additional sinks and sources of N and P are simulated for plankton and benthic populations and associated reactions. In all three modules, biochemical reactions are modeled with either first-order or Michaelis-Menten kinetics.

In our study, the model was run for 11 to 14-year periods. For TN, TP and TSS, we used model output and observations for 1984-1994, whereas simulated and observed flow were available for 1984-1997. The model runs in hourly time steps with inputs of meteorological data, such as precipitation, air temperature, dew-point temperature, solar radiation, and wind speed. Other input data for the model included atmospheric deposition loads, septic system loads, and fertilizer application rates. These data were obtained by the modeling subcommittee of the Chesapeake Bay Program and detailed information on model parameters and data sources is available elsewhere (www.chesapeakebay.net/model.htm).

Estuarine hydrodynamic model

The numerical hydrodynamic model (CH3D - Curvilinear Hydrodynamics in 3 Dimensions) transports salt and water within a three-dimensional system of volume elements (voxels, figure 1). CH3D solves conservation equations for water mass, momentum, salinity, and heat on a boundary-fitted grid of voxels in the horizontal and vertical planes at five-minute time steps. A finite difference solution scheme is used to solve vertically-averaged equations in order to yield the water surface elevations on tidal time scales.

CH3D is validated by comparing modeled output to observed data over tidal to seasonal periods (e.g. tidal elevations, time series of temperature and salinity at Chesapeake Bay Program (CBP) monitoring stations). In our study, the model was applied to simulate estuarine hydrodynamics for the 1984-1994 period, and the output from CH3D for each of 488 870 voxels was then used to drive the three-dimensional water quality model (Johnson 2001).

Estuarine water quality model

The central computations of the water quality model (CE-QUAL-ICM) are algal biomass, dissolved oxygen, and water clarity. Through primary production expressed in units of carbon (C), algae provide the energy required by the ecosystem to function, although excessive primary production leads to large vertical fluxes of organic matter, decomposition in bottom waters, and deficits of oxygen. In order to compute algae and dissolved oxygen concentrations, the model uses twenty-four state variables, including dissolved, particulate, organic and inorganic forms of C, N and P. CE-QUAL-ICM treats each voxel as a control volume that exchanges material with adjacent voxels, as described by CH3D. The model solves, for each voxel and for each state variable, a three-dimensional conservation of mass equation, details of which are described in Cerco and Cole (1994). Inputs to a benthic submodel include sinking particulate (organic and inorganic) material and dissolved oxygen. The sediment organic matter decays to inorganic nutrients, which are recycled in benthic decomposition processes and diffuse to the overlying water along a concentration gradient. Benthic macrofauna, which are supported by inputs of particulate organic matter, can substantially enhance cycling of nutrients and consumption of dissolved oxygen. Key processes and phenomena relevant to the water quality model simulation include 1) oxygen consumption and diffusion, 2) the spring phytoplankton bloom, 3) nutrient limitation, 4) sediment-water interaction, and 5) nitrogen and phosphorus budgets.

Nitrogen budget

An empirical N budget was developed for the Patuxent to provide validation for the modeling system. Values in the nitrogen budget were updated from those used in the Boynton et al. (1995) estimates in order to coincide with the period used in the current modeling effort. New information included here are: 1) a time series of data from 1985-2000, 2) measurements of long-term nutrient burial (based on ^{210}Pb profiling), 3) an evaluation of N and P losses due to burial and denitrification in the tidal marshes (Merrill

1999), 4) a box model of the tidal Patuxent that enables estimates of nutrient transport at key portions of the system (Hagy et al. 2000), 5) availability of several watershed models estimating diffuse source nutrient inputs, and 6) a system-wide 'experiment' of attempted N reductions, including the completion of biological nitrogen reduction (BNR) technology at all of the major sewage treatment plants in the basin.

Bathymetry, voxel grid, and marsh areas

An improvement to the estuarine models was the development and incorporation of a refined bathymetric grid. Using data collected by the Maryland Department of the Environment in August, 2001, as well as NOAA point soundings, we created a point coverage and grid file in ArcGIS. The point coverage was the actual sounding data with spatial coordinates, whereas the grid file was extrapolated from the point coverage to provide a spatially continuous and smoothed depth field for the development of the new estuarine voxel grid.

The refined bathymetric grid was complemented by the development of a more detailed voxel grid (figure 1). We attempted to match the model voxels with the true bathymetry and estuarine shoreline as much as possible. The original voxel grid in use for the Patuxent was developed from the perspective of Chesapeake Bay as a whole by the Chesapeake Bay Program, and had relatively low resolution in the Patuxent (136 by 96 by 19). Consequently, the spatial fidelity relative to shorelines and bathymetry was low when viewed from within the Patuxent, particularly in the upper estuary. The bathymetry data described above and a shoreline arc file created from USGS topographic maps were used to develop a higher resolution voxel grid (152 by 166 by 19) which preserved more of the true shoreline and bottom bathymetry, and which also incorporated more of the tributary structure of the Patuxent, including Western Branch in the upper estuary.

Intertidal estuarine marsh areas were estimated from digital, georeferenced topographic maps compiled by USGS. These raster maps (projection UTM, datum NAD 27) display marsh areas visible from aerial photographs with a spatial resolution of ~2 m and marsh polygons were digitized as polygons on-screen using ArcGIS v8.2. Marsh polygons adjacent to an estuarine voxel or a HUC14 watershed were given appropriate attributes to provide linkages between land and water.

Equations

Statistics for model errors were calculated as:

$$
\begin{aligned}
&ME = \sum(O - P)/N \\
&AME = \sum|O - P|/N \\
&RE = \sum|O - P|/\sum O \\
&RMS = 100 * (\mathrm{sqrt}(\mathrm{sum}(O - P)^2) / \mathrm{sum}(O))
\end{aligned}
\qquad (1)
$$

where O and P are observed and predicted values, respectively. N denotes the number of observations, ME is the mean error, AME is the absolute mean error, RE is the relative error, and RMS is the root-mean-square (rms) error.

Results

Watershed hydrochemical model – calibration errors

We determined the model bias for the calibrated output of HSPF generated by the CBP phase 4.3 model. We compared model output at several time scales with the data collected by the United States Geological Survey (USGS) for Bowie, Maryland (segment 340 of the Patuxent basin – see figure 1). We used the model output that was generated over the calibration period of 1984 – 1995 for total nitrogen (TN), total phosphorus (TP) and total suspended solids (TSS), and output over 1984 – 1997 for flow (river discharge). A scatter plot of daily model output for flow versus observed flow data indicates that the relationship is unbiased with a slope of 0.80 (not significantly different from 1, $p>0.05$) and an r^2 of 0.71 (figure 2a). The rms error (root-mean-square difference between predicted and observed, relative to observed) is a measure of precision (± 93%) and reflects the scatter about the 1:1 line. The rms error indicates that on a daily basis most of the predicted flow values are within a factor of two of the observed flows (figure 3a), and the average absolute error indicates an average daily bias of +56%.

Daily flow data were also aggregated to annual discharge for 1984-1997. This comparison showed good agreement between the observed and predicted annual discharge for 1984-1997 (r^2=0.83; figure 4). Annual rms errors indicated a precision (rms error) of ± 33%, a considerable improvement compared to daily rms errors of 93% because daily model under- and overestimates tend to cancel at longer time scales. The comparison of annual runoff values suggested that there is moderate model bias for the Bowie station (approximately equal scatter about the 1:1 line with 33% rms errors). For the entire 14-year period, flow had cumulative model errors that declined again to +25%, indicating a long-term model bias by overpredicting flow by 25% even at the decadal time scale. The dependence of model errors on time scale is a common aggregational effect (e.g. Lee et al. 2000). Hence, it is harder to predict a short-term hydrologic response to a rain event than the long-term hydrologic behaviour of a watershed.

HSPF also produced reasonably accurate simulations of TN. Observed TN data collected on a specific date and time were compared with modeled TN averaged for the entire day. The log-log plot of these data in figure 2b shows that there is substantial scatter about the 1:1 line but fairly good symmetry on either side of this line, indicating little apparent bias. The slope was 0.68 (not significantly different from 1, $p>0.05$), rms errors were 37%, and average absolute errors were 27%. As for flow, the frequency distribution of the % difference between predicted and observed TN values relative to observed TN concentrations indicate that errors are approximately log normally distributed and range from −70% (low predicted values) to +140% (high predicted values) (figure 3b). The median error is -17% and the average absolute error is +27%, indicating a slight tendency of the model to overestimate TN at the daily time scale. Most predicted values fall within ±50% of the observed values.

Model estimation of TP and TSS estimates were weaker than for TN and flow. For instance, the comparison of modeled and observed TP (figure 2c) shows clear evidence of model bias (slope of 0.32, significantly <1, $p<0.05$) and considerable scatter (rms errors of 65%). Moreover, it is apparent from the log-log plot that the model is over-predicting the lower concentrations of TP observed at this station. As for TN, the % errors for TP were log normally distributed and ranged from −90% to +400%, with most falling between ±100% (figure 3c). The average absolute error was 44%, and the median error was –8%.

The model had even more difficulty accurately predicting concentrations of TSS. There was more scatter (rms errors of 259%) and asymmetry in the relationship of the modeled output of TSS versus observed TSS data for the Bowie station than that of TP (figure 2d). Moreover, there was evidence of model bias (slope of 0.49, significantly <1, $p<0.05$), and rms errors were large (259%), with a relatively poor coefficient of determination (r^2=0.22). The average absolute error (accuracy) was 121%. In a large fraction of the paired predicted and observed values, there was a large range of observed TSS values (2-500 mg L^{-1}) for which modeled values varied over the more restricted range of only 7-30 mg L^{-1}.

The time-scale dependence of HSPF model errors is summarized in table 1. For flow, TN, TP, and TSS, as we increased the time scale (aggregated the observed data and model output), rms and average absolute model errors decreased. Flow aggregation was a simple summing of daily values, whereas TN, TP, and TSS were aggregated as export fluxes (concentration x flow). Despite any compounding of modeled concentration and flow errors, there was better agreement between the observed and modeled annual and decadal fluxes than observed and modeled concentrations at the daily time scale. For instance, the precision of TN predictions improved from the daily time scale (average absolute errors of 27%, n=363) to the decadal time scale (cumulative error + 6%, n = 1). This analysis (table 1) indicates that at longer time scales, the model accurately predicts N (cumulative error of +6%, probably similar to observational errors) and fairly accurately predicts fluxes of P and water (+16% to +25% cumulative errors, although TSS errors are considerably higher (+122% cumulative errors).

A more robust test of the watershed model's predictive capabilities was done by evaluating validation errors associated with the model output. All errors described above are calibration errors, or those which could not be further minimized by model parameter adjustment. A true test of a model's ability is to apply a calibrated model to a basin with independent observations (not used in calibration) that can be used to estimate validation errors. An example of validation error analysis for flow was done for a sub-basin of the watershed at Laurel, Maryland (i.e. segment 330, in figure 1), which accounts for about 55% of the Piedmont area of the entire watershed and includes two reservoirs. The Laurel model output for flow agreed less well with observations, and there was more scatter in the log-log plot (figure 5). For the station at Laurel, the average observed and predicted annual runoff values were 8 and 14 cm y^{-1}, respectively, with annual rms errors of 166% and a cumulative error of +67% over years 1984 to 1997. Hence, there was a positive bias with larger rms errors at the Laurel site compared to the Bowie site (table 1).

Unbiased predictions by the watershed model are essential for the model system to function effectively as a management tool. Errors associated with watershed model output are propagated through the estuarine models, further compounding errors which may be associated with these models. Even if HSPF produces unbiased estimates of export of water, N, P, and sediments on a time scale relevant for management (e.g. seasonal or annual time scales), at shorter time scales with insufficient averaging to cancel errors, deviations of model predictions from real conditions will propagate through the estuarine models. Only if the estuarine models respond slowly to the short term HSPF errors will an unstable error cascade be avoided.

Estuarine circulation model

CH3D simulations of tides, temperature and salinity had reasonable precision and accuracy. For instance, statistics of salinity at CBP monitoring stations along the Patuxent River channel show that there was relatively low bias and good precision overall (table 2). For salinity precision, the rms error increased slightly from the mouth of the estuary (i.e. at river km 0, rms error $\simeq$ 0.2 ppt) to the middle of the transect (at km 45, rms errors $\simeq$ 0.3 ppt for the bottom), where the horizontal salinity gradients are steep. Further upstream in low salinity regions, errors decreased to very low values (rms error $\simeq$ 0.01). Model bias (mean error) behaved similar to rms error in surface waters (i.e. -0.2 ppt at river km 0, -1 ppt at river km 34, and -0.6 ppt at river km 45), indicating a slight tendency to underestimate salinity throughout most of the estuary. In general, CH3D reproduces salinity in the estuary to within 0.5 units.

Estuarine water quality model

Preliminary simulations of water quality in the Patuxent estuary indicated large discrepancies in the upper estuary. The water quality model tended to overestimate N and P in the upper reaches of the estuary compared to observed surface water concentrations of nutrients, and there was a large positive bias for average values of predicted-observed TN and nitrate concentrations as a function of position along the axis of the estuary (figure 6a). Likewise, there were large average negative biases in modeled-observed P concentrations, particularly in the middle estuary (figure 6b). Interestingly, the distributions of model errors in figures 6a and b are similar to the distribution of intertidal marshes along the Patuxent estuary (figure 6c). Intertidal marsh area increases abruptly downstream of river km 90 (the approximate head of tide) and then gradually decreases toward the mouth of the Patuxent estuary. The discrepancies (modeled output minus observed N and P concentrations) are generally well correlated with the area of intertidal marshes (table 3, figure 7), particularly in the region from river km 90 to 20, where marsh area exceeds estuarine water surface area (see figure 1). The correlations in table 3 and figure 7 strongly suggest that the intertidal marshes of the middle estuary are interacting with estuarine waters, a role confirmed below by an alternative approach using a box model. Marshes are not currently a component of the estuarine water quality model, but this comparison of modeled and observed concentrations and marsh location and area suggests that the intertidal marshes play a critical role in the processing of estuarine N and P and should be included for more accurate model predictions.

TN budget

Average annual TN inputs from all sources were determined for the period 1985-2000. The fall line load was compiled by USGS measurements of flow and nutrient concentrations. This approach uses statistical modeling and includes point, diffuse and direct atmospheric deposition of TN to the river surface. The loads to the middle estuary included direct atmospheric deposition to estuarine surface waters, estimated septic inputs, point source inputs, and all other diffuse source inputs as estimated from the Chesapeake Bay Program watershed hydrochemical model (HSPF). Inputs to the lower basin were computed in the same fashion as for the middle basin except that there were no significant point source inputs. Transport from the middle to the lower basin and from (flow x concentration) are within 6% of those calculated with the empirical data (table 1). Hence, the watershed model is doing a good job estimating the loading of TN, probably because highly soluble nitrate is the dominant component of the TN but, more importantly, the positive bias of flow compensates for the negative bias of TN concentrations in TN loading (figures 9a and b). Moreover, there is a large negative bias associated with the model output for TP and TSS concentrations (figures 9c and d), although only the first 5 years of the calibration period for TSS is negative (figure 9d) whereas the P bias is consistently negative (figure 9c). The abrupt change in the trajectory of TSS in 1989 could be related to some environmental forcing affecting HSPF, such as the wetter-than-average precipitation that occurred in 1989 (see figure 9a); however, the other constituents do not exhibit similar changes in their trajectories. For example, cumulative modeled and observed TN and TP concentrations diverge from one another throughout the calibration period, whereas there are essentially 3 instances where the model bias is quite large for predicted flow, and these result in a divergence or convergence of the cumulative trends.

The ranges of watershed model predications for TP and TSS are much less than the ranges of observations (figure 2c, d). Because TP and TSS concentrations are commonly higher during stormflow events (Jordan et al. 2003), model biases for these two constituents may be indicative of mismatches in the timing of discharge events (e.g. modeled water discharge lags or precedes observed discharge). Unfortunately, as far as we are aware, there appear to be no instances where comprehensive measurements of TN, TP or TSS were done over the rising and falling limbs of any stormflow hydrographs at the Bowie station. However, we were able to evaluate whether there were obvious mismatches in the timing and magnitude of stormflow peaks in a comparison of daily model output of flow and flow measurements from the Bowie station.

In a sample of 106 major stormflow hydrographs taken from throughout the calibration period (1984-1997), we made qualitative estimates of the timing and magnitude of modeled and observed stormflow peaks. In terms of the relative magnitude of stormflow peaks, HSPF, as currently calibrated, was relatively unbiased in that about 54% of the integrated flows associated with hydrographs were underestimated (by about a factor of 2), 32% were overestimated (by a factor of 2), and about 14% were similar in magnitude. However, in terms of the relative timing of stormflow peaks, the model had a pronounced premature peak bias (i.e. the modeled hydrograph preceded the observed stormflow hydrograph by 1 day in 76% of the storms). Alternatively, only 2% of the modeled hydrographs were delayed (by 1 day), and 22% had observed peaks on the same day as those of the modeled output.

Evidently, this premature stormflow bias is a result of the lumped parameter nature of the HSPF model. Water is modeled to move from the entire basin to the downstream node instantaneously, thereby circumventing the process whereby water gradually flows through a watershed after a storm event, eventually making its way to the outflow. At the Bowie station, observed flows peak 1-2 days following a storm event, and this is presumably the cause of the timing mismatch. If this model behaviour is scale-dependent, the premature prediction of hydrograph peaks might increase in larger watersheds, as real travel times increase. The implications of the mismatch in the timing of stormflow peaks are that the chemical concentrations of stream water during the period prior to the storm event are likely to be different than those during the event. Consequently, because stormflow is commonly responsible for a large amount of particulate transport, and TP is commonly correlated with this transport, there would be a tendency to over- or underestimate loading of these two constituents. It is intriguing, however, that the model tends to overestimate the annual loading of all the chemical constituents we analyzed in this study primarily because modeled flows exceed observed flows (figure 9a). One might expect, due to the premature stormflow peak bias, and the generally lower concentrations of TN, TP and TSS (figures 9b - d), that the model would underestimate these constituents. However, the consistently higher modeled flows (figure 9a) appear to dominate the loading calculation.

We have identified several biases and sources of error in HSPF as applied to the Patuxent. These are the lack of basin retention of water following storm events, and insufficient model flexibility or fidelity for processes influencing the movement of TP and TSS. While the model is capable of estimating fluxes of water, TP, and TSS, more attention should be given to these three parts of the model to improve model performance. It is interesting to note that the model's greatest complexity and parameterization of constituents transported by water lies in the N processes, and the model's N performance was the best of the three constituents at the annual time scale, although P was reasonably good at decadal time scales. With the exception of TSS, the errors summarized in table 1 suggest that model runs for testing management scenarios are likely to have cumulative model errors of only 6-25% if run for a decade or more. This will provide an accurate, climatologically average watershed response to a management action. What is currently unknown, however, is the potential cascading effect of short-term errors as watershed model outputs are fed into the estuarine models. This is an area that we are currently exploring.

Evidence of a marshland nutrient sink

Nutrient budgets done in successive sectors of a watershed are an effective means by which empirical data can be used to determine source or sink areas in an aquatic system. Previous efforts to calculate nutrient budgets in the Patuxent watershed (Boynton et al. 1995) have shown that 1) there is moderate loading of both N and P relative to other estuarine systems, 2) there are important point and diffuse sources responsible for nutrient loading, 3) there are substantial losses of N from burial and denitrification, and 4) export to Chesapeake Bay was small for N and almost zero for P, implying retention within the estuary. As population pressure increases and remedial efforts are implemented, changes are likely to occur in the spatial and temporal dynamics of nutrient processing in the estuary. Accordingly, the TN budget for the watershed has been updated to help us identify possible new source and sink areas and the relative magnitudes of these areas in different years.

There was just over a factor of two difference between the lowest and highest load years on record (figure 8). This represents a substantial difference, considering that nutrient management goals for Chesapeake systems generally aim for smaller reductions than indicated here for inter-annual variability. The loads were all higher during the wet year of 1996, particularly in the middle basin. Even though the Patuxent basin is typically thought to be point source dominated, increased loads occurring after full implementation of BNR at the major sewage treatment plants indicate that diffuse sources are also important. This finding is supported by the observation that the lowest loads occurred in 1991, prior to full implementation of BNR technology.

Losses within the estuary were significant in the revised budget (figure 8). During the dry year about 34% of all N entering the estuary upstream of the mesohaline zone was removed, while in the wet year 45% of all inputs upstream of the mesohaline zone were removed. These losses appear to be related to denitrification and long-term burial of N in both sub-tidal and tidal marsh sediments. Direct measurements of these losses have yet to be systematically measured, but preliminary estimates suggest that measured rates are sufficient to account for these large estimated losses in the budget.

Intertidal and non-tidal marshes were not considered as landscape sinks for TSS, N, and P in either the watershed or estuarine modeling. However, because marshland areas are commonly regarded as important nutrient processing centres, their omission from the estuarine water quality model would likely result in a significant overestimation of simulated N and P concentrations compared to areas of the upper estuary where intertidal marshes are abundant. Further evidence of the importance of intertidal marshes as a nutrient sink is provided in our preliminary CE-QUAL-ICM simulations. For instance, in figure 6, the difference between the modeled output and observed concentrations of nutrients is what the estuarine water quality model cannot explain based on watershed model loads, observed estuarine concentrations, and modeled water column processes. One weakness of this approach is that some process other than marsh effects could be causing the difference. However, the significant correlation of, for instance, nutrient concentrations and the area of marshland per unit of estuary length (table 3, figure 7) is strong evidence in support of an intertidal marsh effect on the system. We note, however, that this water quality model run purposefully eliminated shoreline interactions to enhance our chance of detecting a marsh effect, and that in other model runs the correlations of marsh area with model errors is weaker. Although we have not yet quantified the magnitude of any mechanism for marsh nutrient retention, we are currently investigating seasonal plant uptake, denitrification, and long-term sediment accumulation as possible mechanisms of N and P removal by marshes. Consequently, one of the major improvements of the system of integrated models presented here is the inclusion of a detailed polygon coverage of intertidal marshes (figure 1) that can be utilized as N, P, and TSS sinks by the estuarine models.

Incorporating marshland effects into the estuarine models

As indicated above, preliminary simulations of the estuarine water quality model suggest that the lack of intertidal marshes in the model is responsible for the significant overestimation of simulated N and P concentrations in the upper estuary. In order to include these important landscape components in the coupled Patuxent modeling system, we created a detailed polygon coverage of wetlands from USGS topographic maps (figure 1). Polygons defined included 1) those which exchange laterally with the estuarine voxels, 2) those which occur at the downstream end of a creek and exchange with an estuarine voxel, and 3) those which are surrounded by land and interact with the land only. In addition to the attributes normally assigned by ArcGIS, each intertidal marsh polygon has attributes consisting of the adjacent estuarine voxel cell with which it exchanges water and materials as well as the number of the nearest estuarine km from the estuarine transect line (figure 1). These marsh polygons will also have attributes such as denitrification rates, CNP burial rates, and biomass. These modifications to the estuarine models are likely to improve the predictive capabilities of the system of integrated models.

Nevertheless, one of the key issues related to the preparation of a TMDL management tool concerns accurate estimation of nutrient loading rates. For some of the sources typically affecting estuaries, estimates are either quite accurate (e.g. point source discharges which are required to be monitored by National Pollution Discharge System permits) or are relatively small (e.g. direct atmospheric deposition to surface waters); errors in these sources are, therefore, relatively unimportant. In cases where most of the drainage basin is above the head of tide, excellent estimates can be developed from high frequency flow and concentration measurements (e.g. USGS data sets). However, about 52% of the drainage basin of the Patuxent is located below the head of tide, and the application of a watershed hydrochemical model such as HSPF is likely to have a great deal of uncertainty in this area. Because our calibration and validation error analysis of the watershed model indicates that there are considerable biases even at the gauged outflow to the head of the estuary, the biases associated with the application of the watershed hydrochemical model in the watershed area below the gauged station are likely to be even larger. Therefore, improvements in the model calibration are necessary before the system of integrated models can be used accurately as a TMDL management tool for the Patuxent Estuary.

Acknowledgements

We thank Gary Shenk of the Chesapeake Bay Program for phase 4.3 model output data. Funding was provided by the Maryland Department of the Environment and the University of Maryland, Center for Environmental Sciences IAN program. We would also like to thank UM RESAC and Steve Prince of the Geography Department, University of Maryland, for land use/land cover interpretation of the Patuxent watershed.

References

BICKNELL, B. R., IMHOFF, J. C., KIDDLE, J. L., JR., DONIGIAN, A. S., and JOHANSON, R. C., 1997, Hydrologic Simulation Program – FORTRAN, User's manual for release 11. EPA/600/R-97/080. U.S. Environmental Protection Agency, National Research Laboratory, Athens, Georgia, USA.

BOYER, E. W., GOODALE, C. L., JAWORSKI, N. A., and HOWARTH, R. W., 2002, Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern U.S.A. Biogeochemistry, 57/58, 137-169.

BOYNTON, W.R., GARBER, J. H., SUMMERS, R., and KEMP, W. M., 1995, Inputs, transformations, and transport of nitrogen and phosphorus in Chesapeake Bay and selected tributaries. Estuaries, 18, 285-314.

BOESCH, D. F., BRINSFIELD, R. B., and MAGNIEN, R. E., 2001, Chesapeake Bay eutrophication: Scientific understanding, ecosystem restoration, and challenges for agriculture. Journal of Environmental Quality, 30, 303-320.

CERCO, C. F., and COLE, T. M., 1994, Three-dimensional eutrophication model of Chesapeake Bay. Volume 1: Main Report. U.S. Army Corps of Engineers Waterways Experiment Station, Vicksburg, MS. Technical Report EL-94-4.

COLE, J. J., PEIERLS, B. L., CARACO, N. F., and PACE, M. L., 1993, Nitrogen loading of rivers as a human-driven process, 141-157 pp. In McDoneil, M.J. and Pickett, S.T.A. (eds), Human as components of ecosystems: The ecology of subtle human effects and populated areas. Springer-Verlag, New York, New York, USA.

D'ELIA, C. F., BOYNTON, W. R., and SANDERS, J. G., 2003, A watershed perspective on nutrient enrichment, science, and policy in the Patuxent River, Maryland: 1960-2000. Estuaries, 26, 171-185.

GALLOWAY, J. N., 1998, The global nitrogen cycle: changes and consequences. Environmental Pollution, 102, 15-24.

HAGY, J. D., BOYNTON, W. R., and SANFORD, L. P., 2000, Estimation of net physical transport and hydraulic residence times for a coastal plain estuary using box models. Estuaries, 23, 328-340.

HOWARTH, R. W., ANDERSON, D., CHURCH, T., GREENING, H., HOPKINSON, C., JUBER, W., MARCUS, N., NAIMAN, R., SEGERSON, K., SHARPLEY, A., and WISEMAN, W., 2000, Clean Coastal Waters – Understanding and reducing the effects of nutrient pollution. Committee on the Causes and Management of Coastal Euthropication. Ocean Studies Board and Water Science and Technology Board, Commission on Geosciences, Environment and Resources, National Research Council. National Academy of Sciences, Washington, D.C., USA.

HOWARTH, R. W., BILLEN, G., SWANEY, D., TOWNSEND, A., JAWORSKI, N., LAJTHA, K., DOWNING, J. A., ELMGREN, R., CARACO, N., JORDAN, T., BERENDSE, F., FRENEY, J., KUDEYAROV, V., MURDOCH, P., and ZHAO-LIANG, Z., 1996, Regional nitrogen budgets and riverine N and P fluxes for the drainages in the North Atlantic Ocean: Natural and human influences. Biogeochemistry, 22, 1-65.

JOHNSON, B. K., 2001, 10-year (1985-94) simulation with refined three-dimensional numerical hydrodynamic, salinity, and temperature model of the Chesapeake Bay and its tributaries. U.S. EPA Chesapeake Bay program, Annapolis, MD. www.chesapeakebay.net/modsc.htm

JORDAN, T. E., WELLER, D. E., and CORRELL, D. L., 1999, Sources of nutrient inputs to the Patuxent River Estuary. Estuaries, 26, 226-243.

KRISANOVA, V., and BECKER, A., 1999, Integrated modelling of hydrological processes and nutrient dynamics at the river basin scale. Hydrobiologia, 410, 131-138.

LEE, K-Y., FISHER, T. R., JORDAN, T. E., CORRELL, D. L., and WELLER, D. E., 2000, Modeling the hydrochemistry of the Choptank River basin using GWLF and Arc/Info: 1. Model calibration and validation. Biogeochemistry, 49, 143-173.

MAGNIEN, R. E., SUMMERS, R. M., and SELLNER, K. G., 1992, External nutrient sources, internal nutrient pools, and phytoplankton production in Chesapeake Bay. Estuaries, 15, 497-516.

MERRILL, J. Z., 1999, Tidal freshwater marshes as nutrient sinks: particulate nutrient burial and denitrification. Ph.D. dissertation, Univ. of Maryland, UMI 9981789, 342 pgs.

TURNER, R.E., and RABALAIS, N. N., 1991, Changes in Mississippi River water quality this century. BioScience, 41, 140-147.

VALIELA I., COLLINS, G., KREMER, J., LAJTHA, K., GEIST, M., SEELY, B., BRAWLEY, J., and SHAM, C. H., 1997, Nitrogen loading from coastal watersheds to receiving estuaries: New method and application. Ecological Applications, 7, 358-380.

VITOUSEK, P. M., ABER, J. D., HOWARTH, R. W., LIKENS, G. E., MATSON, P. A., SCHINDLER, D. W., SCHLESINGER, W. H., and TILMAN, D. G., 1997, Human alteration of the global nitrogen cycle: sources and consequences. Ecological Applications, 7, 737-750.

WELLER, D. E., JORDAN, T. E., and CORRELL, D. L., 2003, Effects of land-use change on nutrient discharges from the Patuxent River watershed. Estuaries, 26, 244-266.

WILLIAMS, M. R., FILOSO, S., and LEFEBVRE, P., 2004, Effects of land-use change on solute fluxes to floodplain lakes of the central Amazon. Biogeochemistry, 68, 259-275.

Table 1. Summary of model biases for daily predictions and the average annual export errors (modeled vs observed; flow-weighted for TN, TP and TSS at the USGS gauging station at Bowie, Maryland). The model overestimated average annual export for all constituents.

Constituent	time scale	units	n	rms error	ave. absolute error	cumulative error
Flow	daily	$m^3\ d^{-1}$	5114	93%	56%	
	annual	$cm\ y^{-1}$	14	33%	26%	
	decadal	$m\ decade^{-1}$	1			+25%
TN	daily conc.	$mg\ N\ L^{-1}$	363	37%	27%	
	annual flux	$kg\ yr^{-1}$	12	20%	15%	
	decadal flux	$Mg\ decade^{-1}$	1			+6%
TP	daily conc.	$mg\ P\ L^{-1}$	411	65%	44%	
	annual flux	$kg\ yr^{-1}$	12	59%	40%	
	decadal flux	$Mg\ decade^{-1}$	1			+16%
TSS	daily conc.	$mg\ L^{-1}$	438	259%	121%	
	annual flux	$kg\ yr^{-1}$	12	194%	130%	
	decadal flux	$Mg\ decade^{-1}$	1			+122%

Table 2. Statistics for salinity at selected stations in the Patuxent estuary and river. Here, ME is mean error, AME is absolute mean error, RE is relative error, and RMS is root-mean-square (rms) error. ME is a measure of bias. Both AME and RMS are indicators of precision.

	Surface				Bottom			
River km	ME (ppt)	AME (ppt)	RMS (ppt)	RE	ME (ppt)	AME (ppt)	RMS (ppt)	RE
0	-0.24	1.11	0.16	0.0038	-0.16	1.22	0.18	0.0032
9	-0.24	1.21	0.17	0.0044	-0.08	1.14	0.16	0.0031
15	-0.31	1.17	0.16	0.0043	0.05	1.11	0.15	0.0030
24	-0.51	1.32	0.17	0.0053	0.25	1.17	0.16	0.0034
34	-1.00	1.55	0.20	0.0052	-0.69	1.30	0.18	0.0033
45	-0.65	1.55	0.20	0.0000	-1.96	2.40	0.33	0.0057
55	0.22	0.60	0.11	0.0014				
64	0.26	0.27	0.07	0.0006				
72	0.03	0.03	0.02	0.0000				
78	0.00	0.00	0.00	0.0000				
99	-0.01	0.03	0.01	0.0000				

Table 3. Coefficients of determination between marsh areas and errors of N and P concentrations generated by the estuarine water quality model. r^2 was calculated for all stations (top line), and excluding the three lower stations below km 20 with little marsh area. The lower 30 km reach near the mouth of the Patuxent estuary is more influenced by fluxes of water and nutrients from the main-stem of Chesapeake Bay than is the upper estuary. The asterisk indicates that the relationship is not significant ($p>0.05$).

statistic	stations	Nitrate	TN	Phosphate	TP
r^2	all	0.54	0.82	0.37	0.002*
r^2	1-6	0.18*	0.84	0.90	0.54

Table 4. Estimated N budget for the Patuxent River and estuary (figure 8). Average annual rates of N inputs and transport are given as kg N d^{-1}, and rates are expressed as a % of total inputs (4). Losses of N to tidal marshes in the middle estuary (7) and burial + denitrification in lower estuary were estimated by difference.

		1991 (dry)		1996 (wet)	
	process	kg N d^{-1}	%	kg N d^{-1}	%
1	input to river at head of tide	1857	45	2840	33
2	input to middle estuary	1604	39	4265	50
3	input to lower estuary	683	17	1470	17
4	total terrestrial inputs	4144	100	8575	100
5	inputs to middle estuary (1+2)	3461	84	7105	83
6	transport to lower estuary	2300	56	3900	45
7	estimated loss to tidal marshes	1161	28	3205	37
8	inputs to lower estuary (6+3)	2983	72	5370	63
9	export to Bay	950	23	1800	21
10	estimated burial and denitrification in lower estuary	2033	49	3570	42

Figures

Figure 1. Map of the Patuxent River basin indicating its location in the Chesapeake Bay watershed. Modeling segments 330 (Laurel) and 340 (Bowie) are located above and below the fall line, respectively. Examples of marshland coverage and voxel grid improvements to the estuarine models are provided.

Figure 2. Log-log plots of daily measurements vs modeled output of flow, TN, TP, and TSS for Segment 340 (Bowie). Plots indicate the slope of the relationship, average absolute error, root-mean-square (rms) error, and sample size (n=x) for each constituent.

Figure 3. Frequency distributions of the errors ((predicted – observed)/observed) for daily measurements of all constituents for Segment 340. Distributions are log-normal. Various statistics are provided, such as the median error that is a measure of accuracy. The mode and median, average absolute, and rms errors are percentages.

Figure 4. Daily flow data aggregated to annual discharge for 1984-1997 at the Bowie (Segment 340) and Laurel (Segment 330) stations.

Figure 5. Log-log plot of daily measurements vs modeled output of flow for Segment 330 (Laurel).

Figure 6. Modeled output (CE-QUAL-ICM) minus observed data errors of N and P concentrations for the Patuxent estuary. Marsh area is summed over 5 km intervals and divided by 5 to represent the average area of intertidal marshes in km^2 per km length along the Patuxent estuary. This is essentially equivalent to the average width of marshes in km along the length of the estuary.

Figure 7. The relationship of phosphate output from CE-QUAL-ICM with marsh area along a reach of the Patuxent estuary. Coefficients of determination of the modeled output errors with marsh area are generally significant for inorganic and total N and P, and using only the stations from river km 80 to 30 km where larger marsh areas occur (i.e., 1 to 6 from the headwater of the estuary) tends to improve the relationship (r^2 for NO_3 = 0.99 for stations 3 to 6).

Figure 8. A summary of total nitrogen loads from all sources to the Patuxent River estuary for the lowest (1991) and highest (1996) loading years on record (1985-2000).

Figure 9. Daily cumulative modeled output and observed data for the calibration period for all constituents.

figure 1

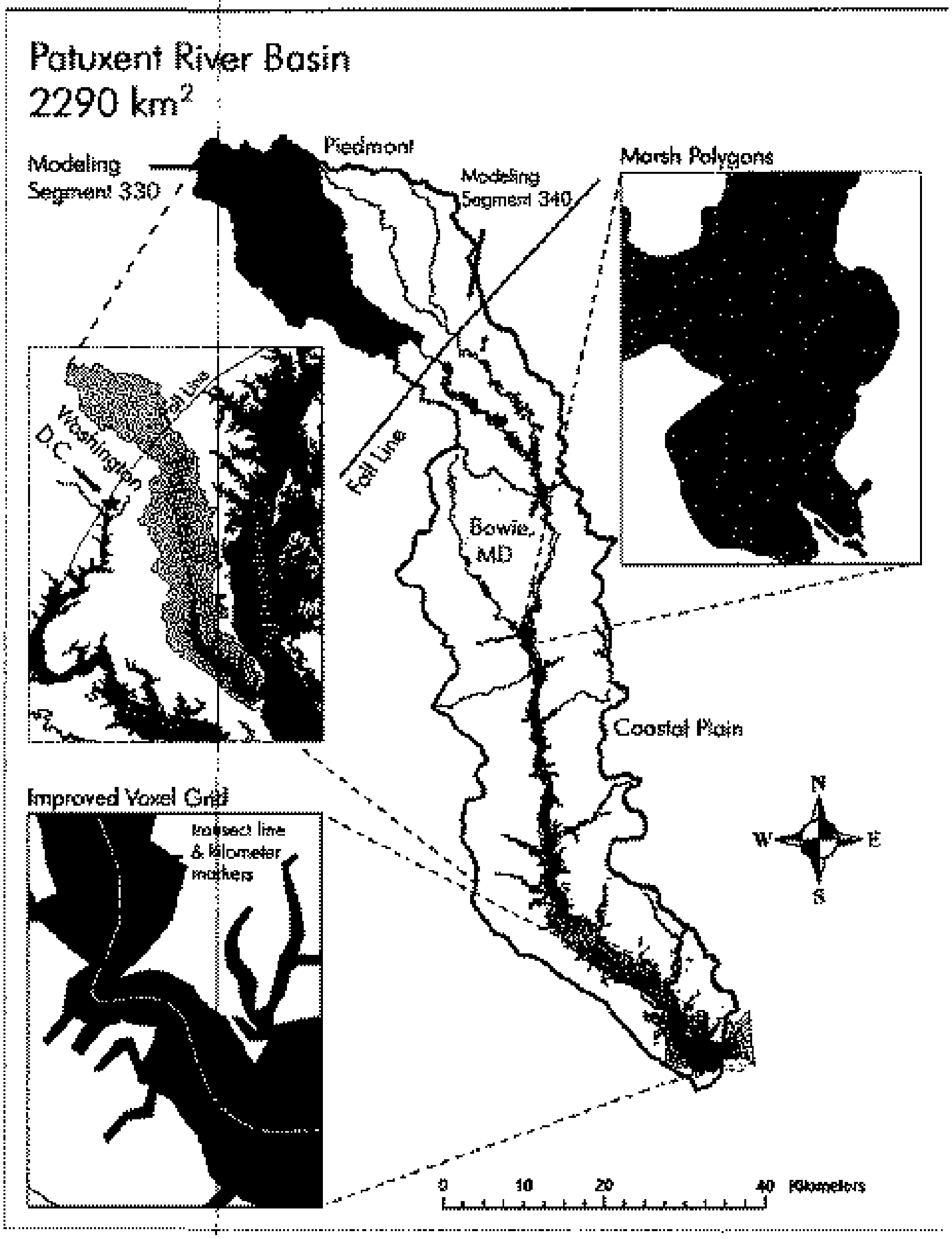

Patuxent River Basin
2290 km²
Modeling Segment 330
Piedmont
Modeling Segment 340
Marsh Polygons
Fall Line
Washington D.C.
Fall Line
Bowie, MD
Coastal Plain
Improved Voxel Grid
transect line & kilometer markers
N
W
E
S
0
10
20
40
Kilometers

figure 2

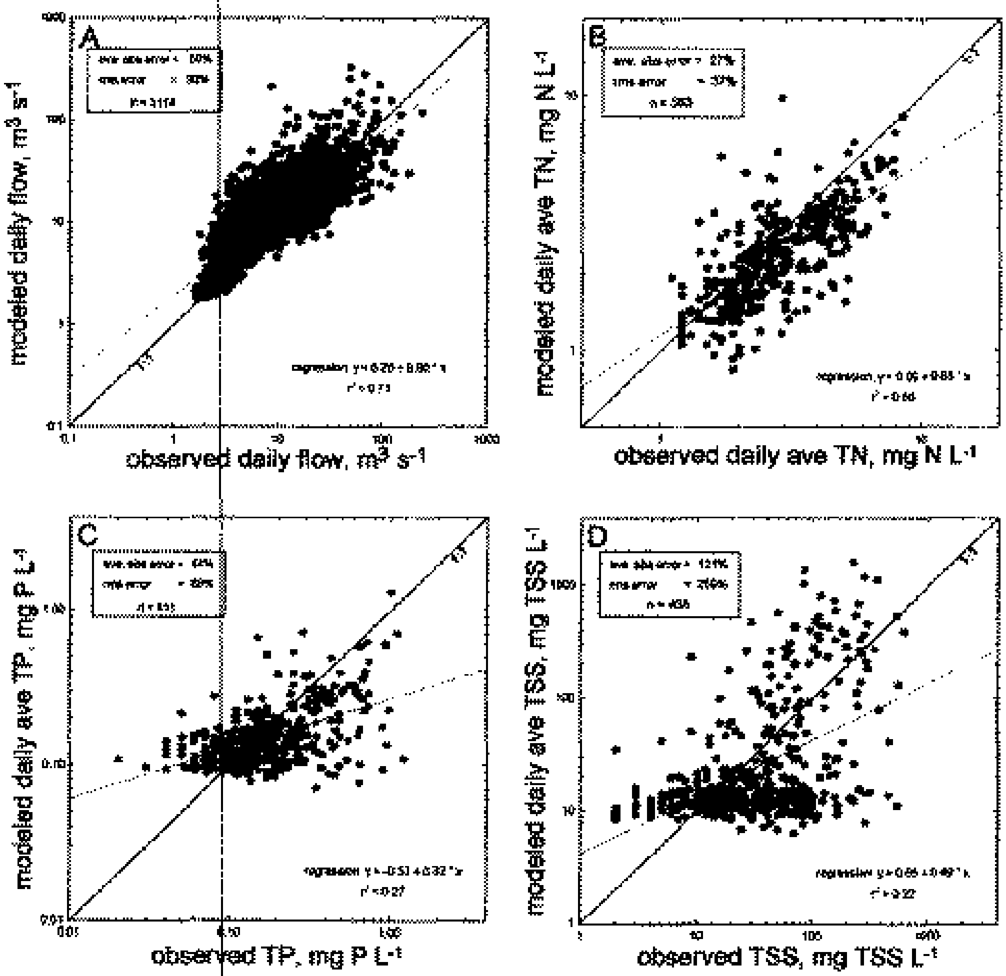

A
B
C
D
modeled daily flow, m^3 s^{-1}
observed daily flow, m^3 s^{-1}
modeled daily ave TN, mg N L^{-1}
observed daily ave TN, mg N L^{-1}
modeled daily ave TP, mg P L^{-1}
observed TP, mg P L^{-1}
modeled daily ave TSS, mg TSS L^{-1}
observed TSS, mg TSS L^{-1}

figure 3

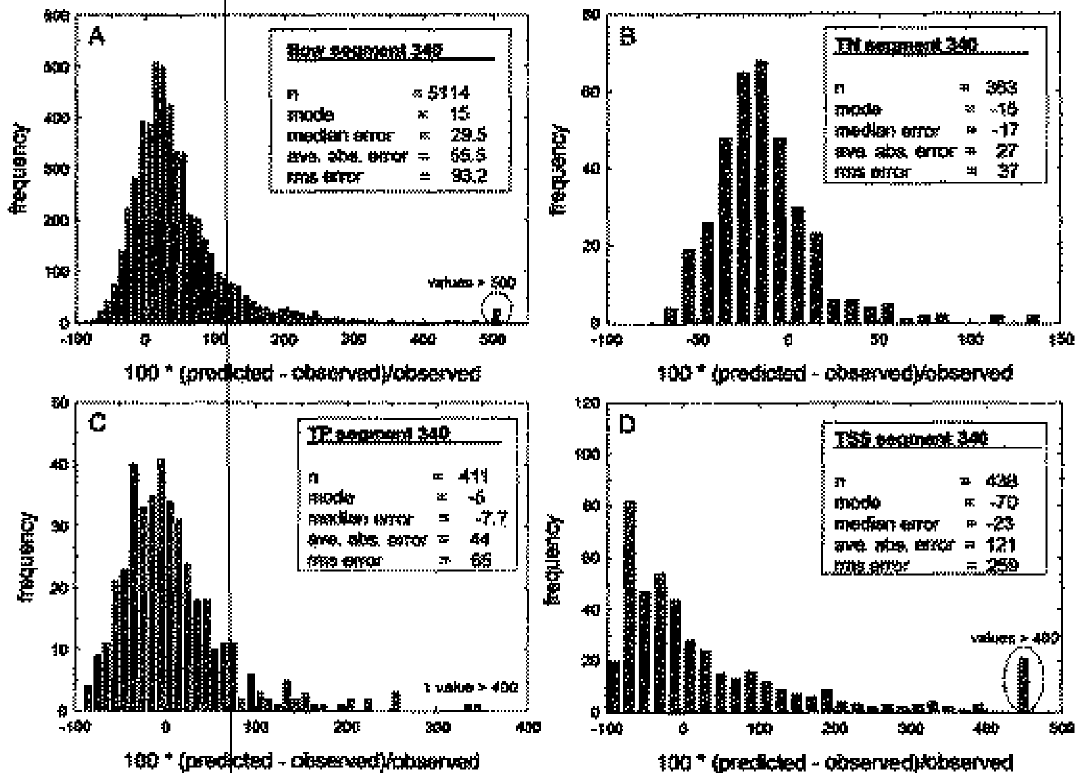

A
flow segment 340
n = 5114
mode = 15
median error = 29.5
ave. abs. error = 55.5
rms error = 93.2
values > 500
frequency
100 * (predicted - observed)/observed
B
TN segment 340
n = 353
mode = -15
median error = -17
ave. abs. error = 27
rms error = 37
frequency
100 * (predicted - observed)/observed
C
TP segment 340
n = 411
mode = -5
median error = -7.7
ave. abs. error = 44
rms error = 65
1 value > 400
frequency
100 * (predicted - observed)/observed
D
TSS segment 340
n = 438
mode = -70
median error = -23
ave. abs. error = 121
rms error = 259
values > 400
frequency
100 * (predicted - observed)/observed

figure 4

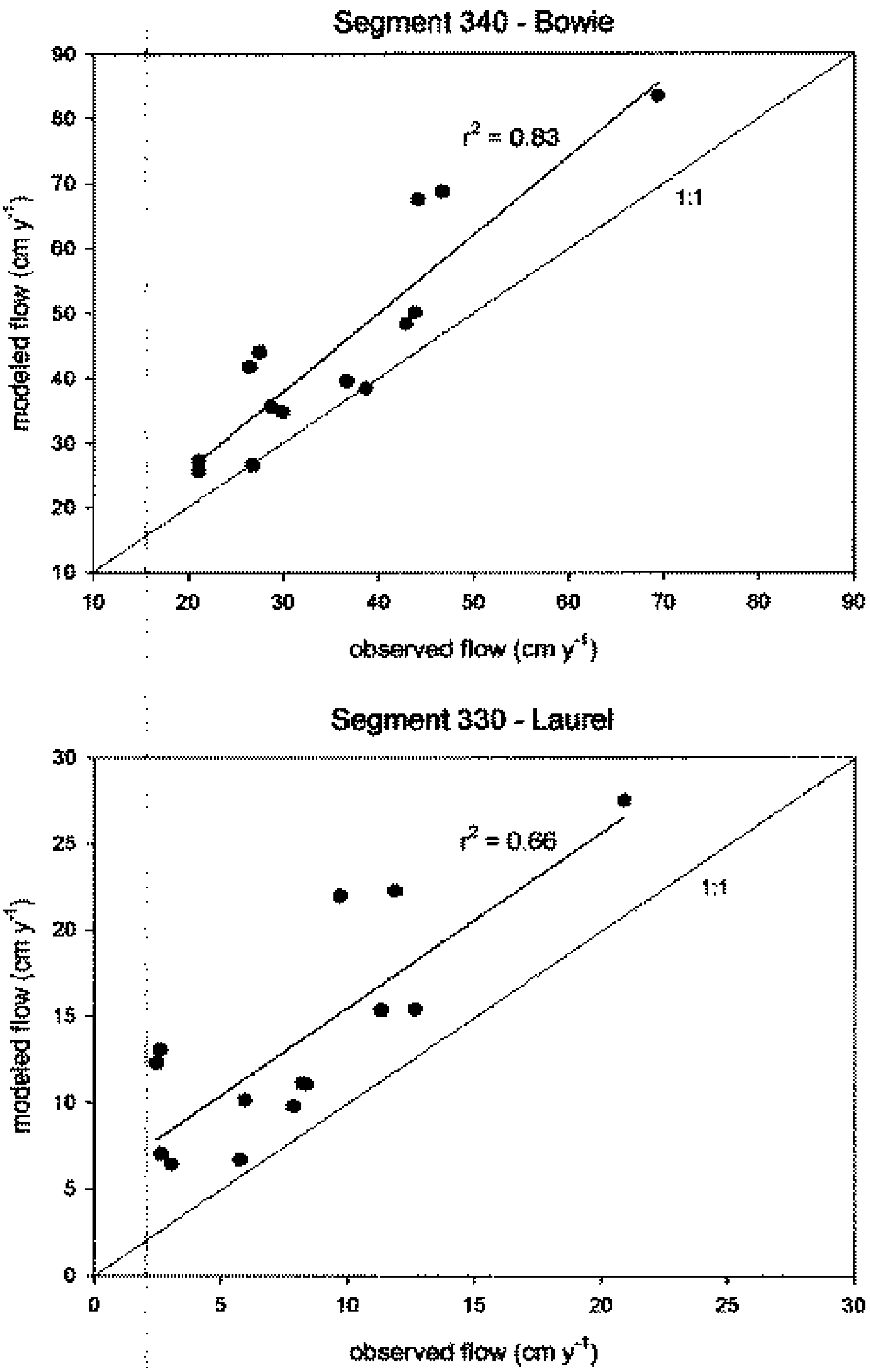

Segment 340 - Bowie
r^2 = 0.83
1:1
modeled flow (cm y⁻¹)
observed flow (cm y⁻¹)
Segment 330 - Laurel
r^2 = 0.66
1:1
modeled flow (cm y⁻¹)
observed flow (cm y⁻¹)

figure 5

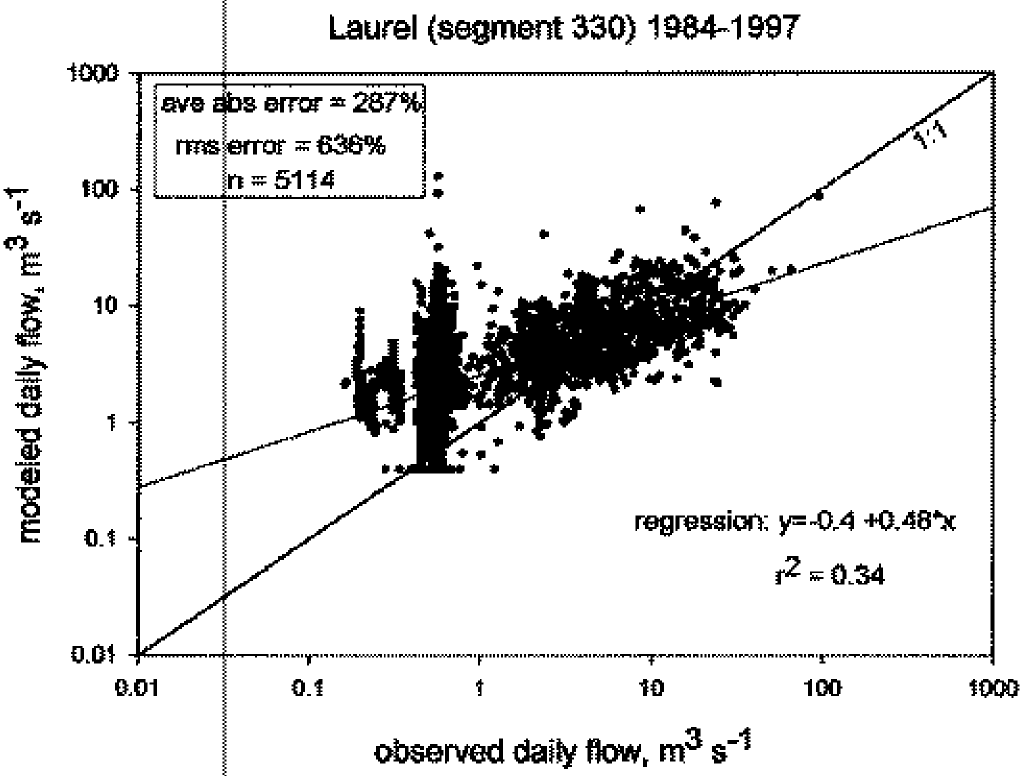

Laurel (segment 330) 1984-1997
ave abs error = 287%
rms error = 636%
n = 5114
1:1
regression: y=-0.4 +0.48*x
r2 = 0.34
modeled daily flow, m3 s-1
observed daily flow, m3 s-1
1000
100
10
1
0.1
0.01
0.01
0.1
1
10
100
1000

figure 6

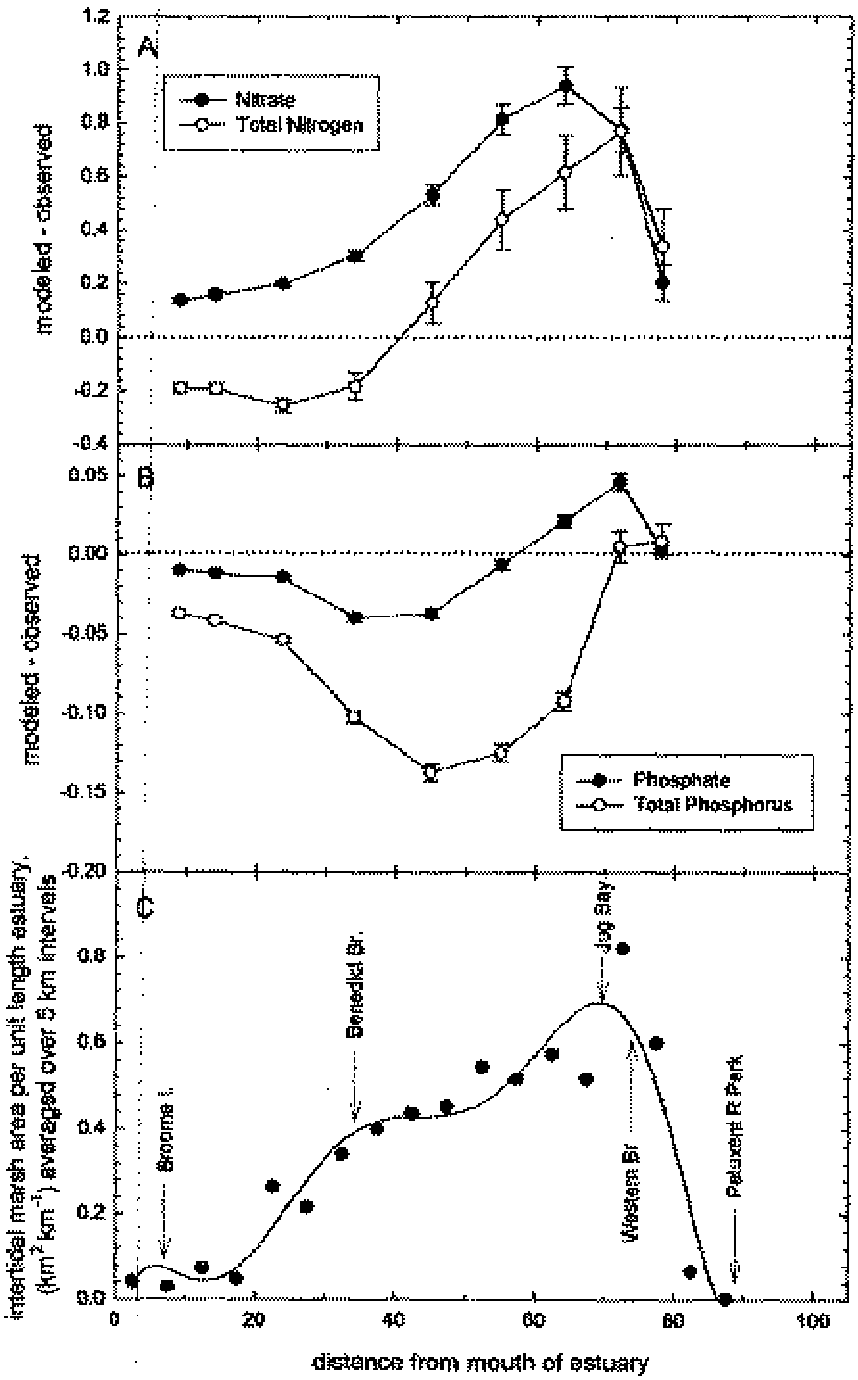

A
Nitrate
Total Nitrogen
modeled - observed
B
modeled - observed
Phosphate
Total Phosphorus
C
intertidal marsh area per unit length estuary, (km² km⁻¹) averaged over 5 km intervals
Broome Is
Benedict Br
Jug Bay
Western Br
Patuxent R Park
distance from mouth of estuary

figure 7

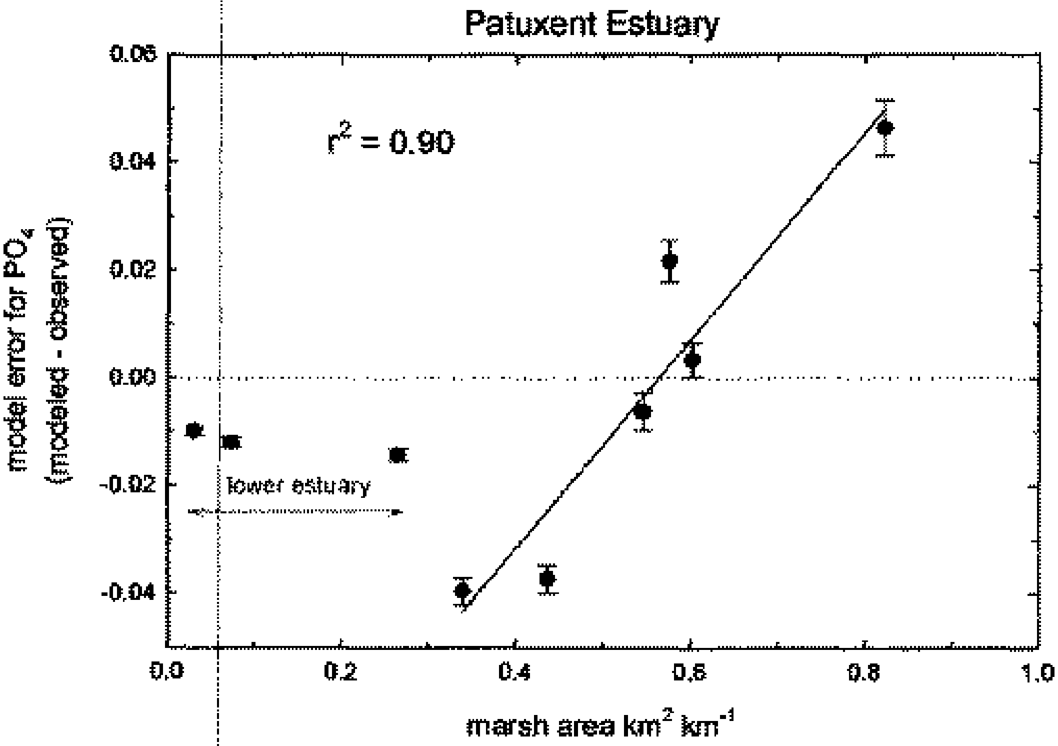
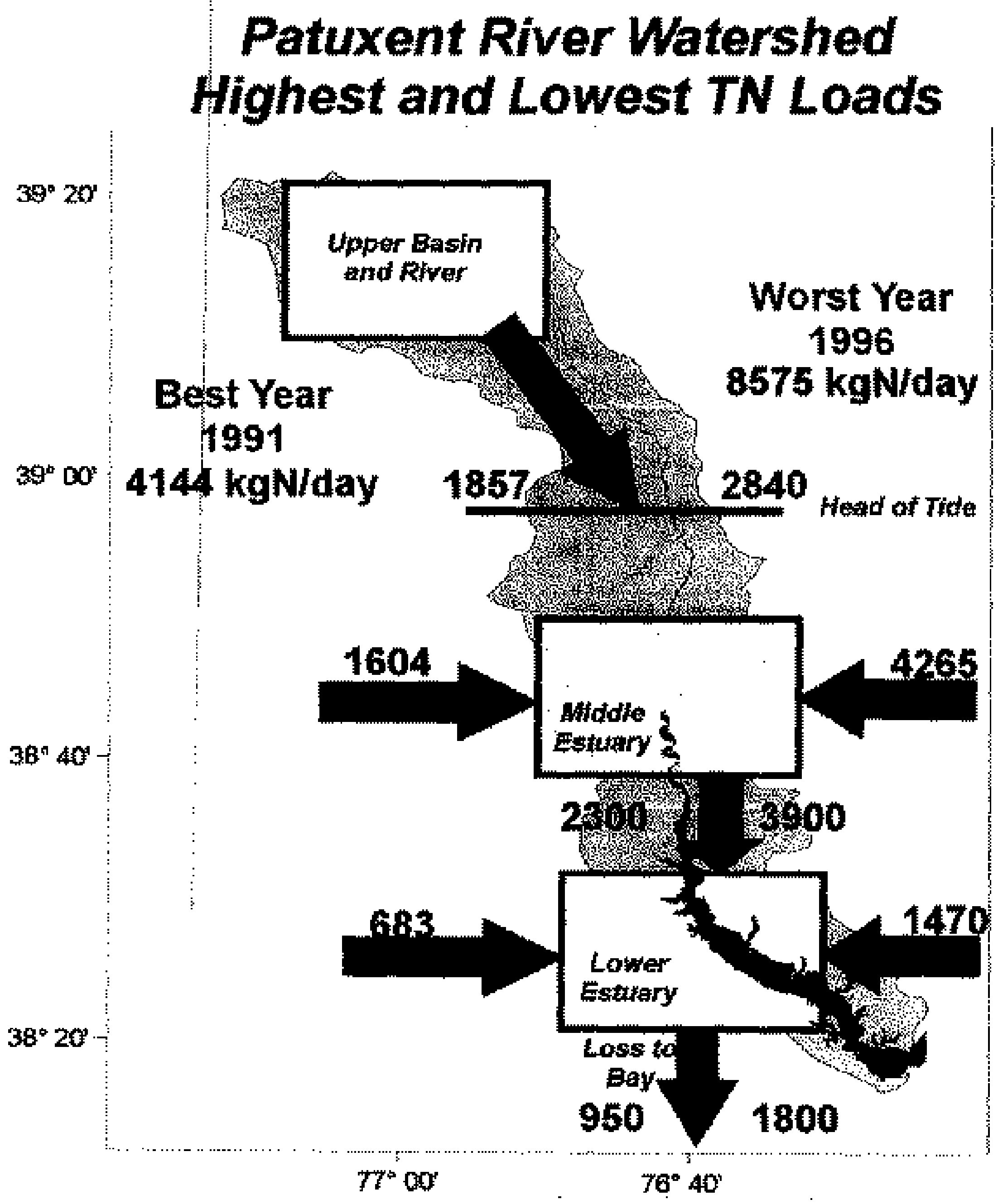

Patuxent Estuary
$r^2 = 0.90$
model error for PO_4 (modeled - observed)
lower estuary
marsh area $km^2\ km^{-1}$
0.06
0.04
0.02
0.00
-0.02
-0.04
0.0
0.2
0.4
0.6
0.8
1.0

figure 8


Patuxent River Watershed
Highest and Lowest TN Loads
39° 20'
39° 00'
38° 40'
38° 20'
77° 00'
76° 40'
Upper Basin and River
Worst Year
1996
8575 kgN/day
Best Year
1991
4144 kgN/day
1857
2840
Head of Tide
1604
4265
Middle Estuary
2300
3900
683
1470
Lower Estuary
Loss to Bay
950
1800