Assessment of models for estimation of land-derived nitrogen loads to shallow estuaries

Ivan Valiela*, Jennifer L. Bowen, and Kevin D. Kroeger

Boston University Marine Program

Marine Biological Laboratory

Woods Hole, MA 02543

Corresponding Author. Fax.: 508-289-7949.

Email address: valiela@bio.bu.edu

Abstract

We compared performance of several models used to estimate land-derived nitrogen loads to shallow receiving estuaries. Models included in our comparison differed in complexity and approach, and predicted either loads or concentrations in estuary water. In all cases, we compared model predictions to measured loads or concentrations, as appropriate. Measured nitrogen loads to nine estuaries on Cape Cod, MA, were obtained as the product of mean concentrations in groundwater about to seep into estuaries multiplied by the annual recharge of groundwater. Measured annual mean nitrogen concentrations in estuaries were obtained by extensive sampling surveys. The validity of this procedure to measure loads was verified by comparison against seepage meter data.

Responsiveness of model predictions was generally good: predictions increased significantly as measured values increased in 8 of the 10 models evaluated. Precision of predictions was significant for all models. Three models provided highly accurate predictions; we calculated correction terms that could be applied to predictions from the other models to improve accuracy. Four of the models provided reasonable predictive ability. We ran simulations with somewhat different versions of two of the models; in both cases, the modified versions yielded improved predictions.

The more complex models tended to be more responsive and precise, but not necessarily more accurate or predictive. Simpler models are attractive because they demand less information for use, but models with more comprehensive formulations, and emphasis on processes tended to perform better.

Different models predicted widely different partitioning of land-derived nitrogen loads from wastewater, fertilizers, and atmospheric deposition. This is of concern, because mitigation options would be based on such partitioning of predictions. Choice of model to be used in management decisions or for research purposes therefore is not a trivial decision.

Keywords: seepage meters, groundwater flow, model complexity, nitrogen loads, model comparisons

1. Introduction

Awareness of eutrophication in coastal waters has prompted interest in the importance of land-derived nitrogen enrichment of these systems (GESAMP, 1990; National Research Council, 1994; Goldberg, 1995). Perception of the issue of eutrophication has entered into play in many decisions about management, permitting, and planning of the development of the coastal landscape in many parts of the world. On the coast of the northeastern U. S., for instance, many municipalities and states have insisted that the impact on nitrogen loads from land be evaluated before projects are approved. This need for information has prompted the development of models with which to assess nitrogen loads from land to water.

Comprehensive reviews of information on nitrogen loads from land have been done (Howarth et al.; 1996); the loading data included in these reviews were largely obtained by multiplying nitrogen concentrations in the few major rivers of the world, by annual river flow. Most management questions, however, deal with decisions involving smaller parcels of land. In this paper we therefore deal with models that predict contributions from watersheds that are smaller in area and empty into smaller estuaries, where concentration and river flow data might not be available. Many models have been broadly applied to estimate land-derived nutrient loads to such estuarine settings, but they have seldom been verified versus measured nitrogen loads and have not been inter-compared.

Models that estimate nitrogen loads to estuaries are so diverse as to resist ready classification. The greatest difference among the models is the complexity of their formulations. Models used in estimation of nitrogen loads may be relatively simple, requiring only estimates of such variables as number of people in a watershed and certain loss coefficients (Gaines, 1986; Cole et al., 1993; Meeuwig, 1999), or have a complex structure with many components (Koppelman, 1978; Frimpter, 1990; Kellogg et al., 1996; Valiela et al., 1997; Birkinshaw and Ewen 2000). Ideally, we want models that are sufficiently complex to reflect the nature of the complex systems they represent, and sufficiently adapted to local conditions to produce accurate predictions, yet simple enough to be generally applicable to many different types of situations. Much has been written about the matter of sufficient model complexity (Jakeman and Hornberger, 1993; Rastetter et al., 1992; Hakanson, 1999; Hakanson 2000, for example), addressing the incompatible requirements for accuracy, realism, and generality in models (Levins, 1966). The model comparisons we discuss below afford a chance to examine whether performance by models of widely different complexity differs in a systematic fashion.

Nitrogen loading models also differ substantially in the terms and processes included. Discussion of the many differences among models would be too long to include here. Suffice it to say that the models differ even in something as basic as whether or not to include inputs by wastewater, fertilizers, and atmospheric deposition, the three major sources of land-derived nitrogen. Even when these terms are included the models differ in many details as to how the inputs are estimated. The diversity of approaches and formulations prevents crafting a systematic classification of these models, and makes selection of representative models to compare a challenge.

To keep the number of comparisons within feasible limits, we selected models that spanned the wide range of approaches and model structures. We included largely mechanistic models based on processes rather than those that apply empirical fitted regression equations such as those of Omernik (1976) or Meeuwig (1999). We selected seven models that predicted land-derived loads, and three models that predicted concentrations in the water of estuaries.

We included models that ranged from simple (Gaines, 1986; Cole et al., 1993; Caraco and Cole 1999) to complex structure (Kellogg et al., 1996; Valiela et al., 1997). Comparing these models provided some notion of the degree of complexity that might be warranted in models designed to assess nutrient loads. We also included models designed for relatively small geographical parcels (Eichner and Cambareri, 1992; Valiela et al., 1997; Costa et al., 1999), as well as models developed for larger watersheds (Cole et al., 1993; Caraco and Cole, 1999), to see if differences in spatial scales mattered. We constrained the number of models to be tested by including only recent versions of models. For example, models have been developed since the 1970's to estimate nitrogen loads from coastal watersheds where groundwater transport is the major pathway of nutrient input and on-site septic systems are the predominant form of wastewater treatment; in this case, we omitted earlier models (Koppelman, 1978; Lee and Olson, 1985; Lamb et al., 1987; Valiela and Costa, 1988; Nelson et al., 1988; Frimpter, 1990) in favor of later developments (Eichner and Cambareri, 1992; Valiela et al., 1997; Costa et al., 1999). We also wanted to include models that estimated wastewater nitrogen inputs using a nitrogen per capita contribution approach (Gaines, 1986; Johnes, 1996; Valiela et al., 1997; Costa et al., 1999), and the water use per person method (Eichner and Cambareri, 1992).

Few published models or models used by management agencies have received critical verification versus measured values. Such a comparison seems a useful requirement before models are applied to answer questions that will influence management decisions and involve considerable investment of public funds. In some cases, authors calibrate their models by simulations in which the model predictions are compared to observations, and the values of the terms are changed to best fit the data. We avoided this practice both in the development of our own models, and in the treatment of the other models selected for the comparisons. Particularly with complex models, there are so many terms that can be "tuned" that almost any measured data set can be made to fit. This is contrary to the intent of this paper, which rather focuses on the relative performance of models as given by the authors. This is also important in prospective use of the models for management: users will simply take the model and apply it to their circumstance without the extensive sampling that is necessary for calibration.

We compared predictions from the models we selected versus empirical data on nitrogen loads and nitrogen concentrations from a series of estuaries on Cape Cod, Massachusetts that we have studied in some detail (Foreman et al., submitted; Kroeger et al., 1999; Valiela et al., 1997; Valiela et al., 2000). As a group, these estuaries are subject to a range of land-derived nitrogen loads that encompass 75% of the estuaries of the world (Nixon, 1992), and hence challenges the models across a broad set of conditions.

Below we first ascertain that, indeed, we have reasonable measured nitrogen loads to the estuaries. Second, we compare predictions by the various models to empirically measured nitrogen loads from the Cape Cod estuaries by means of regression statistics. We conclude with an examination of the relative ability of the various models to address management and remediation issues.

2. Materials and Methods

Below we provide capsule descriptions of each of the models used in this comparison. These descriptions merely describe our understanding of the procedures; we may or may not agree with the rationale used in structuring the different models. We made every effort to apply the models as closely as possible to the author's intent. We used the authors' plan as far as possible, and to make fair comparisons to measured estimates we adapted model inputs to apply to conditions in the Cape Cod watersheds. In all cases, we specify just what local adaptations were applied. For convenience we assigned a brief acronym to each model.

2.1. Description of models

2.1.1. NLM

This Nitrogen Loading Model (Valiela et al., 1997; Valiela et al., 2000) predicts total dissolved nitrogen loads to shallow estuaries from rural to suburban watersheds, where the watersheds are underlain by unconsolidated sands, and hence groundwater flow is the dominant transport vehicle. NLM considers nitrogen inputs from wastewater (via septic systems, using values for per capita contributions of N), fertilizer use on turf and agriculture (determined from local data), and atmospheric deposition (estimated from local data).

NLM allows nitrogen from the three major sources to be lost during passage through the array of different land cover types on a watershed (residential, turf, impervious surfaces, etc.), followed by losses during travel through the soils, vadose zone, and aquifer under the land cover mosaic. Loss rates in each component of watersheds were determined from empirical measurements and literature values. NLM then adds the surviving nitrogen from each source (wastewater, fertilizer, and atmospheric deposition) to produce an estimate of the total nitrogen that will enter the receiving estuary.

In some Cape Cod watersheds there are large freshwater ponds, and portions of the watersheds lie up-gradient from the ponds. Valiela et al. (1997) developed loss coefficients to describe the interception of nitrogen as the groundwater passes through these large ponds. Similar results were reported in reviews by Kelly et al. (1987) and by Howarth et al. (1996). NLM considers these losses by first calculating the total nitrogen load from the part of the watershed up-gradient from the ponds, and then applying the loss coefficient. None of the other models in this comparison consider such losses, but, to make the predictions most applicable to the watersheds we used in our comparisons, we added the same loss terms for ponds to each model's predictions. In all cases, the changes in estimated N loads resulting from up-gradient ponds were small.

2.1.2. MANAGE

This model (Method for Assessment, Nutrient-loading, And Geographic Evaluation of non-point pollution; Kellogg et al., 1996) estimates total dissolved nitrogen loadings from surface runoff and nitrate-nitrogen from groundwater. MANAGE has two major components designed to separately deal with nitrogen loads transported via groundwater and via runoff. The amount of available rainfall after evapotranspiration that contributes to either surface runoff or groundwater recharge is partitioned using runoff coefficients based on land cover and soil hydrologic groups.

The groundwater component calculates NO_3-N loads to groundwater from wastewater, lawn fertilizers, agricultural fertilizers, pets, and atmospheric deposition (as leaching from unfertilized areas) based on published values from local and national research. To assess nitrogen contributions by wastewater from septic systems, the groundwater component of MANAGE assumes that 80% of nitrogen exiting from septic systems leaves the tank and enters the aquifer. Losses of NO_3 during transport are not estimated. Nitrogen derived from fertilizer use is calculated as turf or agricultural area multiplied by loss coefficients of 6 or 30% for lawn and agricultural use, respectively. The groundwater component of MANAGE considers that 0 or 1.3 kg N ha of watershed^{-1} y^{-1} of atmospheric-derived nitrogen leach from forest, pasture and other unfertilized areas, representing atmospheric deposition from these areas. The atmospheric contribution is included within leaching estimates from fertilized areas and in surface pollutant runoff coefficients. Direct atmospheric deposition to surface waters is calculated as 9.8 kg ha^{-1} by the surface runoff component of the model.

The surface water component of MANAGE uses published export coefficients to estimate nitrogen and phosphorus loads from 21 land use types. The coefficients represent generalized sources of nitrogen from fertilizers, atmospheric deposition, and so on, in runoff from different land use types. To estimate the amount of wastewater effluent that might be entering the surface waters, MANAGE assumes that a certain percentage of septic systems in areas with a high water table and slowly permeable soils fail (with the % failure based on soil type), and assumes that 80% (beyond the riparian zone) and 100% (within the riparian zone) of the N in wastewater effluent from failed systems travels via surface runoff or shallow groundwater to the receiving water body.

We used only the groundwater component of MANAGE in our comparisons to measured data, because on Cape Cod precipitation quickly percolates into the sandy glacial soils and hence surface runoff is negligible (Oldale, 1992). Although MANAGE provides a default number of houses (based on the residential land use density) it is possible to make the predictions more locally relevant by using local parcel records or Census block-group data. We entered the actual number of houses for the land parcels in the watershed of our Cape Cod estuaries, as well as a locally specific housing occupancy rate.

2.1.3. BBP

This model is used by the Buzzards Bay Project (Costa et al., 1999), a joint effort of the Commonwealth of Massachusetts and the U. S. Environmental Protection Agency, to estimate loads to estuaries, considering total dissolved nitrogen contributed by wastewater inputs, fertilizers, livestock, and atmospheric deposition to pervious and impervious surfaces. Nitrogen loads from all sources are calculated from land use-specific nitrogen load coefficients that pool nitrogen loads from each source. The land use data to which the loading coefficients are applied are derived from literature and from information specifically collected by the Buzzards Bay Project.

In our treatment of the other model comparisons, we applied the models to the Cape Cod watersheds and estuaries. In this case, however, the values we report were computed by the Buzzards Bay Project staff, using their compilation of land use data and watershed delineations.

2.1.4. CCC

This model was developed by the staff of the Cape Cod Commission, the county planning agency for Cape Cod, Massachusetts, to calculate inputs of nitrogen from septic systems, fertilizer use, and atmospheric deposition for specific residential or commercial land use parcels (Eichner and Cambareri, 1992). This protocol has been broadly used in many decisions about land use and management in the Cape Cod region (for example, Eichner et al., 1998; Eichner et al., 1998a; Eichner et al., 2000).

The protocol used by CCC involves two somewhat different calculations of nitrogen loads, which are then averaged. First, an estimate of N loads is obtained by calculations based on water used, using information from the Commonwealth of Massachusetts Sanitary Code (110 gallons of water used per bedroom d^{-1}, 2 people per bedroom in each residential building, 3 bedrooms per housing unit) multiplied by an average nitrate concentration (0.35 mg NO_3/L) in wastewater leaving septic systems. These values were selected as conservative overestimates. Second, the CCC protocol calculates the amount of wastewater nitrate loaded to estuaries. Wastewater flow is obtained using a local estimate of people per house from 1990 census data, multiplied by a mean flow of 55 gallons per person per day, and this is then multiplied by 0.35 mg NO_3/L in effluent, taken as a representation of the concentrations leaving the septic system. Then, the CCC protocol averages the two values obtained, and uses the average as the predicted estimate of wastewater nitrate loading.

Fertilizer inputs are calculated in CCC based on a fertilizer application rate multiplied by a loss coefficient, both compiled from the literature. Atmospheric nitrogen inputs are considered in two ways. First, there is a contribution via a "background" concentration of 0.05 mg NO_3/L in groundwater. This "background" was estimated as the most frequently occurring nitrate concentration in samples of groundwater collected from drinking and monitoring wells on Cape Cod. An additional atmospheric contribution is allowed through deposition on impervious surfaces. CCC considers that precipitation (plus some undefined contribution by vehicles, pets, and wildlife) deposits nitrogen onto paved areas and roofs, and that 1.5 or 0.75 mg NO_3/L for pavement or roofs, respectively, finds its way to groundwater and estuaries. CCC assumes a standard lawn, roof and driveway area per house, based on surveys from Koppelman (1978).

The issue of what CCC predicts seems unsettled. The text of Eichner and Cambareri (1992), and the practice of the staff of the Cape Cod Commission (T. Cambareri, pers. comm.) is to consider that the CCC calculation predicts total dissolved nitrogen loads. On the other hand, the calculations, as noted above, are done in terms of nitrate (except for fertilizer inputs, which are given in the text as total N), so it seems reasonable to think that the prediction refers to nitrate loads. To be consistent with the author's interpretation, we compared CCC predictions to measured TDN. To be consistent with the calculation procedure, however, we also ran a comparison of the CCC estimates to measured nitrate loads. We were also unsure about the purpose or effect of the averaging step with the likely overestimate from the State Statutory Code. To evaluate the effect of averaging with the overestimate, we ran CCC omitting the averaging step.

2.1.5 PJM

This model (Johnes, 1996) estimates total nitrogen exported annually to a water body from land-derived sources using an input-and-loss-coefficient approach similar to that of Omernik (1976), Beaulac and Reckow (1982), and Soranno et al. (1996). PJM requires input data on the amount of nitrogen from human and livestock wastes, fertilizer use, nitrogen fixation in different land cover types, and atmospheric deposition. Then, the nitrogen from each of these inputs is multiplied by "export coefficients" (values for losses within the watershed before the nutrients reach the receiving water body) derived from the literature.

We applied PJM to the Cape Cod estuaries by entering local data for human populations, land uses, and atmospheric deposition derived for the specific Cape Cod watersheds. We ran three different simulations with PJM.

First, we ran PJM in a configuration as close as possible to the original version in Johnes (1996). Land use data for the Cape Cod watersheds did not exactly match the classifications used in PJM, but we re-sorted the Cape Cod land use data to approximate the categories required by PJM as much as possible. We then applied the nitrogen loss coefficients given in Table 2 of Johnes (1996) (4-17% for wastes from animals, 36% for waste from humans, 5-30% for fertilizers, and 37% for atmospheric deposition, respectively) and applied them to the re-sorted Cape Cod watershed data.

Second, we ran PJM with loss coefficients derived for a Cape Cod groundwater-based system. PJM uses export coefficients that include losses that are not defined in detail, but they most likely pertain to nutrient losses during surface runoff-dominated transport. To evaluate just how much difference such locally-relevant loss terms may make, we ran PJM with groundwater-based export coefficients (35%, 21%, and 11% for wastewater, fertilizer, and atmospheric nitrogen, respectively, from Valiela et al., 1997). By comparing the first to the second simulations, we ask whether the differences in loss coefficients, and formulation of transport via surface or subsurface flow, alter the resulting estimates produced by PJM.

Third, we ran PJM with the original loss coefficients but without terrestrial nitrogen fixation terms. PJM is unusual among nitrogen loading models in that it considers nitrogen fixation in soils within the watersheds. PJM assigns rates of nitrogen fixation of 20 kg N ha^{-1} yr^{-1} to natural vegetation land cover. These rates were chosen from British sources, but are higher than the 1-5 kg N ha^{-1} yr^{-1} reported as an average for U.S. temperate forest soils (Schlesinger, 1991). For agricultural and grazing land covers, PJM uses N fixation rates of 50 and 10 kg N ha^{-1} yr^{-1}. These are also higher than those reported for US conditions, and would overestimate nitrogen loads. To assess the importance of the fixation term, we ran PJM omitting nitrogen fixation.

2.1.6. OSF

This model, based on on-site septic system and fertilizer inputs, was applied for land management purposes in Martha's Vineyard, Massachusetts (Gaines, 1986). Its interest is to provide first-order estimates in circumstances where funds and information were limited. OSF has a rather simple structure: it adds annual load of septic N per house (6.8 kg N y^{-1}) to an annual load from lawn fertilizers (4.8 kg N y^{-1}), to predict N load. Atmospheric loads are assumed to be taken up within the watershed. We assumed that the loads refer to TDN, though the form of nitrogen is not defined.

To apply this model to Cape Cod estuaries, we multiplied these coefficients by the number of houses, and the area of lawns in each watershed (0.05 ha of lawn per house, from Valiela et al., 1997).

2.1.7. C&C

This is another model with a relatively simple structure. Land-derived nitrate loads are calculated as the sum of wastewater, fertilizer, and atmospheric deposition (Caraco and Cole, 1999). Nitrate from human sewage is estimated as contributions from the urban population multiplied by a 60% reduction in N in sewage treatment plants. Rural areas are assumed to contribute insignificantly to wastewater nitrate loads. Loads from fertilizers are calculated as the area in agriculture multiplied by a mean fertilizer use per unit area. Atmospheric deposition was defined from meteorological data, and multiplied by 2 to account for dry deposition. To allow for losses of fertilizer or atmospheric nitrate within the watershed, Caraco and Cole (1999) developed a nitrate export coefficient that was a function of the rate of water flow from the watershed. The percent of nitrogen exported ranged from 6 to 40%, depending on the annual throughput of freshwater.

To adapt C&C to Cape Cod estuaries, we substituted recharge for runoff, and used the local recharge rate of 0.52 m y^{-1} (equivalent to 45% of annual precipitation). We then inserted that recharge value into a function derived from the NO_3 export and watershed runoff volume data given in Caraco and Cole (1999) to estimate a watershed nitrate export coefficient applicable to Cape Cod estuaries. For Cape Cod estuaries, this export coefficient was 23% of the nitrate delivered to the watershed surface by atmospheric deposition and fertilizer use. We used data on total number of people in the watershed, and allowed for the 60% loss of wastewater nitrate. We used local data on fertilizer use as given in Valiela et al. (1997), and local agricultural acreage. We further assumed that the total nitrogen was converted to nitrate by the time the nitrogen arrived at the estuary. We used data from Lajtha et al. (1995) on NO_3 deposition to Cape Cod watersheds for regionally appropriate deposition data.

2.1.8. CPCP

This is the simplest model we test in our comparisons: it merely asks whether the mean nitrate content can be adequately predicted by knowledge of the human population on a watershed. CPCP, taken from Cole et al. (1993), consists of a simple log-log regression that relates the number of people on a watershed to the mean annual concentration of nitrate in the receiving river water. No distinction is made about surface or subsurface water flow. The assumption is that the mean concentration in the receiving river is a direct product of the land-derived load.

To apply the CPCP model to Cape Cod estuaries, we used the number of houses found in each watershed, multiplied by occupancy of 1.79 people per dwelling (Valiela et al., 1997). With these data, we used the Cole et al. (1993) regression to predict the mean annual concentration of nitrate in the groundwater about to enter into each estuary. We then compared these predicted values versus corresponding measured values of nitrate concentrations in groundwater.

2.1.9. ELM

The Estuarine Loading Model (Valiela et al., submitted) predicts mean annual concentrations of dissolved inorganic nitrogen in the water within estuaries, and requires inputs of land-derived nitrogen loads from an external source (NLM in our case). ELM considers nitrogen losses within the upper reaches of estuaries and in fringing wetlands, nitrogen fixation and denitrification within the estuaries, regeneration in the benthos, direct atmospheric deposition on the estuary surface, and the effects of estuarine water residence times and tidal exchange. From these inputs and transformations, ELM estimates the mean annual concentration of dissolved inorganic nitrogen (DIN) in the water within the receiving estuaries.

2.1.10. DVM

This model incorporates nitrogen loads and water residence times to predict the concentration of nitrogen in estuaries (Dettmann, in press). The concepts underlying this model parallel the well-known Vollenweider (1976) loading approach, which has been widely applied in lakes. Although Hakanson (2000) is not convinced that the Vollenweider approach provides the most balanced or useful models for eutrophication studies, we had interest in including a Vollenweider-like model to evaluate how applicable that approach might be to estuaries. DVM assumes that the rate of internal N losses is proportional to the content of N in the water column, and that net N export over the seaward boundary is a function of freshwater residence time.

This model predicts mean annual TN concentration in estuaries based on an externally-provided land-derived nitrogen load from the watershed, the residence time of water in the estuary, and inputs of N across the seaward boundary as

$$[N]_{\mathrm{ave}} = \{L_l\,\tau_{FW}/V + [N_s]\}*(1/1+\alpha\tau_{fw}),$$

where $[N]$=the mean annual concentration of total nitrogen, L_l=the annual N input from land, V=estuary volume, τ_{FW}=freshwater residence time, N_S= the mean concentration of N entering across the seaward boundary, and α is a 1st order rate coefficient (α=0.01 d^{-1} for estuaries).

To adapt DVM to Waquoit estuaries, we used NLM estimates of loads of TDN, and residence time estimates for each estuary obtained by a simplification by J. Kremer (University of Connecticut, Avery Point, CT), based on a 2-dimensional hydrodynamic model created by T. Isaji (ASA Associates Narragansett, RI), as reported in Valiela et al. (submitted). DVM requires data in TN entering estuaries from deeper water. These data are seldom available for most estuaries. To supply DVM with data for inputs from deeper water, we used concentrations of total nitrogen (nitrate, ammonium, dissolved organic nitrogen, plus particulate nitrogen) recorded at the saltiest end of each estuary. This was the best approximation available to us for concentrations in flooding tidal waters in Cape Cod estuaries. The requirement to add PN unfortunately meant that we only had measured values for three estuaries for this comparison.

2.2 Measured estimates of N loads and concentrations

We first validated the use of the concentration times recharge method of estimating measured N loads. This method involves multiplication of nitrogen concentrations in groundwater samples taken from the seepage face (McBride and Pfannkuch, 1975; Lee, 1977; Bokuniewicz, 1980) and multiplying these by the annual flow of groundwater through the seepage face. We estimated annual groundwater flow from annual precipitation minus evapo-transpiratory loss.

We needed to ascertain that this approach adequately represented the total amount of nitrogen entering the estuary, because it is possible that within the glacial outwash sediments in the Cape Cod area there are glacial melt drainage channels buried within the aquifer, where coarser sediments might allow groundwater to flow into the estuary much beyond the shore (Cambareri and Eichner, 1998), rather than through the seepage face. Elsewhere (Valiela et al., 2000) we reported how an intensive sampling of groundwater nutrient concentrations is required to aptly define nutrient content of recharge to estuaries, because of the large variation in nutrient concentrations in groundwater parcels. There is no turbulent mixing pressure in groundwater flow to mix and hence smooth out concentrations in the aquifer.

If we show that the calculation of water flow based on precipitation and evapo-transpiration on the watershed is similar to the flow obtained at the seepage face, we would have confidence that our measured estimates (groundwater concentrations times recharge volume) reasonably represent the measured quantity of nitrogen entering the estuary.

2.2.1. Sampling of groundwater nutrients

Groundwater about to enter the estuary was sampled at the seepage face by driving piezometers below the water table at sites just above the high tide mark, and pumping water out of the aquifer. We repeated this sampling at stations all around the periphery of the estuaries so that no significant groundwater flows went unsampled. The number of samples varied depending on the dimensions of the shoreline of the estuary. The spacing between sampling stations varied between 1 m and 50 m; examination of the data did not show evident differences in the distribution of concentrations taken from the different spatial distributions, so we treated measurements from all the piezometer samples alike.

To prevent differences in groundwater flow, number of stations, or other differences from one place to another from biasing our estimates of nitrogen loads, we subdivided the watersheds into recharge areas (Fig. 1 in Valiela et al., 2000). We assumed that, within a given recharge area, groundwater flow along the seepage face was the same at all locations, and we averaged the concentrations from all groundwater samples taken within that specific recharge area. We delineated the recharge areas based on land surface features and groundwater flow lines obtained from hydrological modeling, using the MODFLOW (McDonald & Harbaugh, 1988) groundwater transport model (Fig. 1 in Valiela et al., 2000). Hydraulic conductivities, porosities, and watertable contours of this part of Cape Cod are well defined (LeBlanc et al., 1991; Solomon et al., 1995; Portniaguine & Solomon, 1998), so that we could run MODFLOW to delineate particle tracks and the watersheds (Fig. 1 in Valiela et al., 2000).

We measured nitrate (NO_3), ammonium (NH_4), and dissolved organic nitrogen (DON) concentrations in samples of groundwater. We did not consider nitrite (NO_2) independently because, as is common in most natural waters, concentrations of nitrite were one to two orders of magnitude lower than those of the other nitrogen species. In addition, the analytical method we used records NO_3 + NO_2 combined. Concentrations of NO_3 were determined colorimetrically on a Lachat Autoanalyzer using cadmium column reduction of NO_3 to NO_2 (QuickChem® method 31-107-04-1-C). DON was measured as nitrate after persulfate digestion (modified from D'Elia et al. 1977). Concentrations of NH_4 were measured colorimetrically on a Lachat Autoanalyzer using a standard alkaline phenol method (QuickChem® method 31-107-06-1-C).

Details on the collection of estuarine water column samples, including station locations, replication, and methods of analysis, are documented elsewhere (Valiela et al., 1992; Valiela et al., 1997; Valiela et al., 2000; Foreman et al. submitted). Briefly, the water column samples were taken monthly for 4 years from a series of 5 stations (near-surface and near-bottom) in each estuary. Nutrient concentrations were analyzed using the same methods described for groundwater samples.

2.2.2. Estimation of groundwater flow

This section of the work consisted of two parts. First, we used data on annual precipitation from published papers (Valiela et al., 1978; Lajtha et al., 1995), plus estimates of annual evapotranspiration for our latitude (Thornthwaite and Mather, 1957; Running et al., 1988; Eichner and Cambareri, 1992), and areas of the watersheds of Cape Cod estuaries (Valiela et al., 1992) to estimate annual recharge to the groundwater below the watersheds.

Second, to test whether most of the annual discharge from the aquifer does indeed seep near the shore (where we sampled the concentrations of nutrients in groundwater), we ran a series of seepage meter measurements (Valiela et al., 1990). The seepage meters consisted of plastic cylindrical chambers about 0.4 m in diameter, 0.3 m tall, and with no bottom. Thin, wetted plastic bags were connected to ports on top of the chambers to vent pressure within the headspace (the water filled space above the sediment), and hence, allow free flow of ground- and seawater through the sediments within the chambers. The chambers were pressed into the sediments so that a headspace of less than 0.1 m remained. Water samples were taken through sampling ports set into the upper surface of the chambers to determine time courses of changes in chlorinity, and concentrations of dissolved nutrients during several hours of deployment in situ. Chloride concentrations were measured with a chloridometer, and nutrients were measured as described for the piezometer samples.

We deployed sets of 9 seepage meters, set into the sediments at intervals of 3 m in a 3x3 grid across the intertidal range in the shore of three of the Waquoit Bay estuaries (Childs River, Quashnet River, and Sage Lot Pond, Fig. 1 in Valiela et al., 2000). We repeated the deployment during June-July and Oct-Nov, for a total of 4 deployments per site. Deployments were done from high to low tide in each case. In the deployment of each seepage meter, we measured volume within the seepage meter (from the height of the head space), and salinity of the headspace water within the seepage meter at the start and at the end of the measurement interval. We also measured the volume of water and the chlorinity that accumulated in each pressure-venting plastic bag during the measurement period. From these volume and salinity measurements we calculated the volume of fresh groundwater that had flowed into the seepage meters over the measurement interval (ΔV_{fw}) as

$$\Delta V_{fw} = [V_H(‰Cl_2 - ‰Cl_1) + (V_B(‰Cl_2 - ‰Cl_{SW})] / (‰Cl_{FW} - ‰Cl_{SW}),$$

where volume of the head space is V_H, $‰Cl_2$ and $‰Cl_1$ are the concentrations of chloride in samples taken at the start and end of the sampling interval (times 1 and 2), V_B is the volume of water that flowed into the venting plastic bag, and $‰Cl_{FW}$ and $‰Cl_{SW}$ are the chloride concentrations in samples of freshwater and seawater end members of this land/sea mixing interface, both of which we also measured.

Once we verified that estimates of annual groundwater flows calculated from recharge rates were comparable to groundwater flows measured in the seepage meters, we felt justified in using the recharge-based estimates to multiply by mean nutrient concentrations measured in groundwater samples. In turn, these measured nutrient fluxes could then be used to verify the nutrient loads predicted by the various models.

2.3. Quantification of model performance

Our approach to quantify model performance was to compare measured vs. predicted data using four regression-related statistical features, which we define below and interpret as suggestive of responsiveness, precision, accuracy, and utility.

First, we assessed responsiveness of a model by use of F_{reg}, which in this context evaluates the degree to which model predictions track changes in actual measured values. Second, we calculated r for the plots of predicted vs. measured values and determined the statistical significance of r as in Motulsky (1995); the values of r and their statistical likelihood furnished an idea of scatter around the regression line, a measure of the precision (consistency of repeated estimations) of model predictions. Third, we assessed accuracy (approximation to actual values) of model predictions by a t-test between the slope of the regression line and the slope of a 1:1 line (Steel and Torrie, 1960) that would indicate a perfect fit of predicted to actual measured values. Fourth, we evaluated the predictive ability of the models using the criterion developed by Prairie (1996), and used by Hakanson (1999), based on R^2. Values of R^2 lower than 0.65 are unable to differentiate between more than two categories, hence have low predictive ability. We further assessed the relative utility for application of the various models by two means: the nature and variety of data inputs required, and the ability of the models to predict terms useful to managers.

3. Results

3.1. Measured estimates of nitrogen loads

3.1.1. Groundwater flows

Flow rates calculated from the seepage meter deployments showed no seasonal differences (Table 1, compare summer to fall means). This seems reasonable, because flow through the seepage face integrates passage of water parcels that arrived on the watershed at different times (up to 80 years ago, given the watershed dimensions), and hence within-year differences might be not obvious. There were also similar flows of fresh and sea water through the seepage meters (Table 1, compare means for fresh vs. seawater flows). Through tidal forcing, volumes of seawater equivalent to the volumes of fresh groundwater course through the near-shore sediments. This is significant because it suggests the potential for considerable biogeochemical activity within the shallow sediments at the seepage face: reactions such as denitrification could be stimulated by the alternation of nitrate-laden freshwater and organic-matter bearing seawater.

From measurements in Table 1, plus data on dimensions of the seepage face, we estimated flow of groundwater from areas up-gradient of the shoreline from which we sampled groundwater nutrients (Table 2). The dimensions of the seepage face (length and width) were determined from the following lines of evidence. To calculate the length of the seepage face we examined aerial photos of the watersheds, and measured the periphery of the presumed seepage face, based on previous hydrological determinations (Valiela et al., 1990). To calculate the width of the seepage face, we plotted data on chlorinity of piezometer samples (taken adjacent to seepage meter while deployed) versus the distance from the shore line (Fig. 1). The piezometer samples were taken from depths of 30 to 150 cm in sediments of Childs River, Quashnet River, and Sage Lot Pond (Fig. 1 in Valiela et al., 2000). These data show substantial variation in chlorinity, but the relevant feature for present purposes is that water coursing through the seepage face is fresher (lower chlorinity) nearer shore, and becomes saltier farther offshore. By 8 m from shore, the influence of fresh groundwater water became minor. We conclude that in these Waquoit Bay sites the seepage face is largely constrained to within 8 m from the shoreline.

The second way we calculated flows of fresh groundwater was to use data on area of recharge, multiplied by the annual freshwater recharge (45% of annual precipitation, Lajtha et al., 1995), to get the values shown in the 4^{th} row of Table 2. Estimates of groundwater flow obtained by the recharge calculation are remarkably similar to the values obtained using the seepage meter results (Fig. 2). The best fit line for the points for the three Waquoit estuaries (Childs River, Quashnet River, and Sage Lot Pond) lies quite close to the 1:1 line. From the gap between the fitted regression line and the 1:1 line in Fig. 2, we calculated that flow of groundwater through the seepage face makes up, on average, 91% of the total flow of groundwater expected to have recharged to the catchment area feeding the seepage face. This further means that if there is some spring-borne flow of groundwater passing under the seepage face, it should be, on average, only about 9% of the total flow.

The resemblance between the seepage meter-derived flows and the recharge-derived flows simultaneously supports the use of recharge-based flows, and strengthens the justification for the comparative approach we apply for verification of the predictions of the various models. These results agree with similar comparisons of seepage meter data and recharge estimates made by Giblin and Gaines (1990) in Town Cove, another Cape Cod estuary. In general, these shallow estuaries are fed by groundwater flowing largely through a narrow seepage face, and flow through deeper channels within submerged sediments may be of a smaller magnitude.

3.1.2. Groundwater nutrient concentrations

Nutrient concentrations in groundwater about to enter the several estuaries varied greatly in the samples used for these comparisons; more details of these analyses are given in Valiela et al. (2000), but the useful feature was that we had data for all relevant nitrogen species. For the comparisons in this paper, we used concentrations of those nitrogen species specified by each model. In some comparisons, we used total nitrogen (the sum of DIN and DON), and in others we used nitrate alone. The DVM model requires the use of TN (the sum of DIN, DON, and particulate N). The different watersheds provided a considerable range of different mean concentrations of nitrogen, which, once multiplied by the annual recharge estimates, provided a wide range of loading rates (Valiela et al., 1992; Valiela et al.; 2000; Valiela et al., 2000a). This made it possible to compare model predictions over a significant range of measured values.

3.2. Predictions from models

3.2.1. Specific evaluations of the models

We compared predictions obtained by each of the models to actual field measurements by plotting predicted vs. measured values (Figs. 3, 4, and 5). We plotted predicted values on the x axis because this made it possible to meet the assumption of no error term for x: the models all gave single explicit values. As already noted, based on our definition, we assessed responsiveness of model predictions by examining F_{reg} for the regressions, precision by study of r values, accuracy by the results of t-tests between the regression and the 1:1 line of perfect fit, and predictive ability by examining the R^2 values for the plots of measured vs. predicted. Statistical tests for the plots of Figs. 3, 4, and 5 are reported in Table 3, and the statistics were calculated as specified in Sokal and Rohlf (1995), Motulsky (1995), Steel and Torrie (1960), and Prairie (1996). We should add that perhaps had we had more estuaries to work with, the results would have differed. In general, though, it is not common to find that the multiplicity of data required to test the models is available for many estuaries.

The responsiveness of model predictions to increases in actual loads or concentrations was generally good. Regressions from all models but CPCP and DVM yielded significant F_{reg} values (Table 3). We could roughly classify the models, on the basis of the F_{reg} values, into three categories (Table 4): high (NLM, BBP, CCC, PJM, OSF, ELM), intermediate (MANAGE, C&C), and low responsiveness (CPCP, DVM).

Precision of model predictions was variable, ranging from 0.41 to 0.98 (Table 3). The r values were highly significant for some models (NLM, BBP, CCC, PJM, OSF, C&C, and ELM), significant for one (MANAGE), and not significant for two models (CPCP and DVM). Most models therefore were reasonably precise in repeated application to the various estuaries in the comparisons; in Table 4 we qualitatively sorted the models as to level of precision, based on values of r from Table 3.

Few of the models represented actual values accurately (Figs. 3, 4, and 5, and Table 3). Three models (NLM and ELM, and the version of CCC without averaging and assuming it estimates TDN) provided slopes of the measured vs. predicted regression that were not significantly different from the slopes of the 1:1 lines that showed a perfect fit between predicted and actual values. Predictions from MAN, BBP, other versions of CCC, versions of PJM, OSF, and C&C significantly overestimated the measured values. In most cases this could result because the loadings are source estimates that do not explicitly account for losses that occur during travel to the estuary. F_{reg} were not significant for CPCP and DVM (Table 3) so we did not run t-tests vs. the 1:1 lines for predictions from these models. From the slopes of the regression equations fitted to measured vs. predicted points (Table 3) we calculated correction terms to improve accuracy of predictions by these models (Table 3, 6th column of numbers). From the results of the t-tests and the magnitude of the correction terms we assigned the various models into two groups, in terms of accuracy (Table 4). The most accurate predictions (those with no significant difference from the 1:1 line, and no need for a correction term) were produced by NLM, ELM, and the version of CCC with no averaging and assuming it predicts TDN. All the other models made predictions that required some corrections for accuracy (Table 3).

In terms of the predictive ability of the models, we found (Table 3, last column of numbers) a group of models with values well above (>0.85) the Prairie threshold of 0.65 for R^2 (NLM, CCC with no averaging and assuming estimation of TDN, PJM without N_2 fixation, OSF, and ELM). A second group of models furnished lower values of R^2, but still above the 0.65 threshold (BBP, CCC assuming prediction of NO_3, PJM, C&C, and DVM). A third group of models provided estimates below the Prairie threshold for useful predictive ability (MANAGE, CPCP) (Table 4).

Three models (NLM, ELM, and CCC with no averaging and assuming prediction of TDN) produced predictions that were statistically indistinguishable from measured values. The case of CCC is perplexing, because to accept that version of the model, we need the supposition that although the calculations are done on the basis of nitrate, the prediction refers to TDN. This requires the assumption that all the forms of nitrogen become converted to nitrate before entering the estuaries. We do know that substantial amounts of ammonium and dissolved organic N in groundwater seep into estuaries of Cape Cod: nitrate contributes 6-70% of the nitrogen in groundwater about to seep into Waquoit Bay estuaries (Table 5). This being so, it is biogeochemically unclear how the CCC version with no averaging, and assuming prediction of TDN, performed as well as it did. Were we to use the CCC model, we would feel more comfortable avoiding the leap of faith that a nitrate-based calculation predicts a total dissolved nitrogen load. We would use the output of the version of CCC with no averaging, accepting, that, as the calculations suggest, CCC predicts nitrate loads, and then applying a suitable correction term to improve accuracy of the values (Table 3). This seems a more biogeochemically convincing use of this model.

Most of the other models in our comparisons overestimated nitrogen loads and concentrations; DVM underestimated values. Losses of nitrogen within watersheds were consistently under-evaluated by the formulations of most models. It has been argued that for management purposes it may be desirable to work with models that overestimate loads. We fail to see the advantage in using biased estimates for any purpose. In sum (Table 4), NLM, ELM, and CCC (with no averaging and assuming TDN prediction), provided statistically adequate precision and accuracy; BBP, OSF, and PJM provided sufficient precision but needed corrections to add accuracy.

We should add that we have throughout ignored the substantial propagated uncertainty in any model prediction. For NLM, for example, the coefficient of variation in estimates of loads was 12% based on standard errors, and 39% based on standard deviations (Valiela et al., 1997). Similarly, measured estimates also inevitably include considerable variation: field estimates of mean annual concentrations, for instance, must bear associated errors of 20-30%. Changing coefficients to change the fit of model prediction to measured value to within 2%, as done in Johnes (1996), may overstep the limits of the data. Thus, interpretation of various statistical tests that we have carried out (Figs. 3, 4, and 5, Table 3) must be tempered by the realization that there is some considerable uncertainty in all model predictions. We could not calculate propagated uncertainty of other model predictions, but differences among models, and between model predictions and actual values must be weighted with this uncertainty in mind, at least qualitatively. Differences between loading estimates of less than 10-20% of the means should be viewed as within the margin of uncertainty.

3.2.2. Aggregate model features

To evaluate how model complexity might affect the various indicators of model performance, we plotted the various statistics calculated in Table 3 versus numbers of components in the model. Responsiveness of model predictions (Fig. 6 top left) increased significantly as model complexity increased, although the scatter was considerable. Model accuracy (Fig. 6 top right) was not clearly related to number of components in the models, although the lowest values of the *t*-tests were obtained for the two models with the largest number of terms; perhaps a large number (ca. 90) of terms improves accuracy. Precision of model predictions (Fig. 6 bottom left) increased as number of terms increased, but relatively high precision was achieved by some simple as well as complex models (NLM, OSF, CCC both with and without averaging, and assuming TDN prediction), PJM (using local loss coefficients and Not including fixation), and ELM). Some simple models (CPCP, DVM, MANAGE) may yield imprecise predictions at least for the estuaries that we tested. Predictive ability of models (Fig. 6 bottom left) was achieved from some simple or complex models; some simple models (CPCP, DVM) produced predictions with low predictive ability. Simplicity by itself, therefore, is not a defining aspect; appropriate formulation of processes in models seems to play an even more significant role in model success (Hakanson, 1999; Hakanson, 2000).

At least for nitrogen loading models relatively high complexity seems to confer good performance. Simpler models are appealing because of ease of application and presumed generality. In our comparisons, some simpler models, such as OSF, were reasonably precise (though to an extent inaccurate), but others, such as C&C and CPCP, did not cope well with the task of predicting concentrations or loads to the estuaries. In any case, for prediction at the scale of the watershed, the attraction of simplicity needs to be balanced with concern for lower precision and accuracy. Perhaps they are better suited to larger scale predictions that encompass a wider variety of climates, land uses and topography.

Complexity of models does not by itself guarantee precision or accuracy, as made evident by the comparisons with MANAGE and several versions of the CCC models. These complex formulations did not uniformly provide good matches to measured loads. On the other hand, the fairly complex BBP and PJM models yielded reasonably precise estimates, which, with adjustments to increase accuracy, could produce good estimates of loads.

Models that predicted nitrate loads or concentrations (MANAGE, CCC (assuming NO_3 prediction), CPCP, DVM) tended to perform less well than models that dealt with total nitrogen loads (NLM, CCC (assuming TDN prediction), OSF, BBP, PJM). We suspect that leaving out forms of nitrogen other than nitrate leads to omission of important terms in the model, contributing to the lack of fit. Models apparently need to capture at least some of the biogeochemical details to function adequately. The unprepossessing performance of the Vollenweider approach model (DVM) was surprising in view of the success of this approach in freshwater. It may be that here also we have evidence that key processes—in this case about dynamics of water flow and exchange in estuaries—require more detailed formulation than was included in DVM.

Models designed to deal primarily with surface delivery of nutrients, such as C&C and PJM, performed as well as many models explicitly including subsurface freshwater flow (Table 3 and Fig. 3). The case of PJM is especially instructive. The results of runs of PJM with its own loss coefficients (obtained from surface runoff data) and with a different set of loss coefficients estimated by NLM (obtained from groundwater flow data) were remarkably similar (compare rows 5 and 6 in Table 3). This suggests that the PJM model is relatively robust, and that load estimates are not overly sensitive to small differences in loss terms, regardless of whether the terms were derived from surface or subsurface data. The accuracy of PJM predictions, however, did improve greatly when we omitted terrestrial N_2 fixation (R^2 increased, and the t relative to the 1:1 line decreased, Table 3). The improvement seems a reasonable result since, as noted earlier, the nitrogen fixation rates used in PJM appeared too high, at least for our sites.

Explicit development for local application also may not confer ability to make reliable predictions. For instance, the BBP and the CCC models, which furnished quite different results, were both designed for application in Cape Cod. The MANAGE nitrogen loading model, designed as one component of a broader-risk based approach for use in Rhode Island, did not seem readily applicable to our test sites, while the PJM model, a British creation, did better.

These considerations suggest that details of the various terms in these models carry less weight than we might have expected in influencing the fit of model predictions. For example, formulation of the model to describe transport of nitrogen via surface runoff or via groundwater did not seem to make a huge difference. To balance against this conclusion, we need to recall that, in general, the models with more terms did better than models with fewer terms (Fig. 6).

The results of these comparisons demonstrate that choice of model for use in research or management is not a trivial decision. Models differ in complexity and performance, and the choice of model (rather than the data themselves) can point to one or another management option. Hakanson (1999, 2000) argues that it is not enough to make a model as simple as possible and include input variables and terms whose values might be readily accessible for users. Rather, model predictions will be more successful if models 1) use mechanistic formulations that include realistic depictions of the most fundamental ecological and biogeochemical processes; and 2) are based on carefully evaluated major driving variables defined with minimal uncertainties. Such features characterize NLM and ELM, models designed as analogs of the major paths and fates of nitrogen in coastal environments. The formulation of these models focused on the processes and terms that are likely to be quantitatively of major magnitude.

3.3. Practical applications

Model simplicity and accessible data inputs might be insufficient to guarantee model accuracy, but these features are important for applications of models. The application of the various nitrogen loading models would be easiest if the number of needed inputs were low, if the information needed as inputs were readily accessible through GIS and other sources, and if the output from the models could be easily translated into variables that furnished practical information.

The models differed greatly in number of items needed as inputs (Table 4), from 2 to 25. Simpler models of course required fewer inputs, but there was no evident relationship of any of the statistics we used to evaluate the models and the number of inputs (data not shown). Model performance was hence independent of number of inputs required by the models. The kind of data inputs were largely number of buildings and land use types, which ought to be readily accessible. Some models asked for inputs on atmospheric (C&C, PJM) and fertilizer inputs (C&C). ELM and DPM required an estimate of water residence times in estuaries, which is more demanding.

The relative utility of the output from the models is a difficult issue to evaluate; as a proxy for the many possible uses for the outputs, we chose to see if the models could separately estimate the proportion of total nitrogen load that is delivered through wastewater disposal, use of fertilizers, and by atmospheric deposition (Table 5). We selected this criterion because before management actions can be pondered, the models first need to quantify the magnitude of the three major sources of nitrogen. It is only after we have this information in hand that we can prioritize different options for management and restoration of nitrogen loads. Depending on the relative proportions of the three major sources, managers might opt for quite different strategies for restoration or maintenance of loading rates to estuaries. In addition, models that can separate these three sources are also more likely to be able to be used in simulations where we can evaluate the effects of different remedial actions. Thus, we argue that ability to sort out the contributions from these three sources is a reasonable proxy for usefulness of the models. To be fair, some of these models were designed just to estimate total loads, rather than to examine details of loading sources; nonetheless, we included them here so as to examine model utility.

To examine the way in which different models apportion the contributions of nitrogen loads, we calculated the partition of land-derived nitrogen loads to the Waquoit Bay estuarine system (Valiela et al. 1997). Based on the ease of partitioning of source contributions as the proxy for utility, the models fall into three groups.

Two models allowed easy separation of the three major N sources to watersheds (Table 5). NLM and PJM readily furnish estimates of wastewater, fertilizer, and atmospheric contributions (Valiela et al., 1997; Johnes, 1996). These models appear to be widely applicable to many different areas, and useful as points of departure for management evaluations.

Three more models furnish an indirect but calculable separation of sources. CCC allows a direct estimation of wastewater contributions, and of fertilizer nitrogen loads from fertilizer applied to lawns, but not from fertilizer use on agricultural parcels. For atmospheric deposition, CCC assumes that atmospheric nitrogen falling on natural vegetation is intercepted before it reaches the estuary; a part of atmospheric nitrogen that falls on impervious surfaces is allowed to reach the estuary. C&C combines locally appropriate inputs of atmospheric and fertilizer nitrogen as input terms in the model, and then calculates a final load from each source after considering the losses that occurred in transit. It also provides estimates of nitrogen derived from wastewater. MANAGE, in its surface runoff component uses loss coefficients that include presumed losses for the sum of wastewater, fertilizer and atmospheric nitrogen for each type of land cover. In most coastal watersheds with sandy soils, the surface water runoff component is relatively small, typically less than 20% of the inputs. The subsurface component of MANAGE allows estimates of wastewater, lawn fertilizer, agricultural fertilizer, and pet waste contributions directly. MANAGE implicitly calculates atmospheric contributions by assigning a nitrogen contribution rate to areas of land covers that are not fertilized. It is impossible to differentiate between atmospheric and fertilizer nitrogen on lawns and agricultural fields that receive both inputs.

Two other models do not provide for ready separation of atmospheric and fertilizer nitrogen sources (Table 5). BBP includes atmospheric deposition of nitrogen to undeveloped land, but, like MANAGE does not explicitly include nitrogen from deposition on agricultural or turf parcels. Both BBP and MANAGE provide estimates of nitrogen derived from wastewater. OSF does not separate fertilizer from atmospheric nitrogen, and does not consider agricultural land; it just yields load estimates for watersheds whose nitrogen is primarily delivered from septic systems. OSF was developed for use on Martha’s Vineyard, and hews to the island’s special conditions. OSF would be unsuited for application to areas with large portions of the watersheds devoted to agriculture. ELM, DPM, and CPCP do not differentiate contributions from separate sources, hence were not included in Table 5.

There is an additional aspect of the results shown in Table 5 that should be of concern to managers. The models furnish widely different values for the percentages of the nitrogen loads deriving from the three major sources. Wastewater contributions, for example, vary from 13 to 83 percent of the load, depending on the model used. Clearly, managers that find that wastewater adds only 13% of the load to their estuaries will focus on other sources. On the other, managers that calculate that 83% of the load comes from wastewater will want to mitigate this input. Thus, choice of model, rather than actual conditions, can set the priority given to managing options. We could, in addition, wonder whether the breakdown of sources might affect the fit of the overall predictions. Might the very good fit of the CCC (the version that predicts TDN and omits the averaging step; Table 3) depend on an underestimation of atmospheric contributions and an overestimation of wastewater inputs (Table 5)? It is troubling that the different models yield such different assessments of the major sources of nitrogen, because choice of management options or research conclusions should respond to actual facts, rather than choice of model. This is just one reason that justifies careful verification of model predictions, as done here, and use of models that not only prove most accurate, but also reflect the ecological and biogeochemical processes and structure of the systems under examination.

Acknowledgments

This research was funded through grants from the U.S. Geological Survey's National Institute for Water Resources. Support also came from Victor Serveiss at the US Environmental Protection Agency's Office of Risk Assessment, a NOAA Estuarine Research Reserves Fellowship and a Palmer/McLeod Fellowship to K. D. Kroeger. The research that furnished the basis for this work was part of the Waquoit Bay LMER project, funded by NSF, NOAA, and EPA.

We would like to thank Maggie Geist for her initial work on this project. We thank Lorraine Joubert, Edward Dettmann, Nina Caraco, Arthur Gaines, Penny Johnes, and Tom Cambareri, authors of some of the models we compared in this paper, for their helpful comments on a previous version of this manuscript.

References

Beaulac, M. N., Reckow, K. H., 1982. Examination of land-use nutrient export relationships. Wat. Resour. Res. 18, 1013-1024.

Bokuniewicz, H. J., 1980. Groundwater seepage into Great South Bay, NY. Estuar., Coast., Shelf Sci. 10, 437-444.

Birkinshaw, S. J., Ewen, J., 2000. Nitrogen transformation component for SHETRAN catchment nitrate transport modeling. J. Hydrol. 230, 1-17.

Cambereri, T. C., Eichner, E. M., 1998. Watershed delineation and groundwater discharge to a coastal embayment. Ground Water 36, 626-634.

Caraco, N. F., Cole, J. J., 1999. Human impact on nitrate export: An analysis using major world rivers. Ambio 28, 167-170.

Cole, J. J., Peierls, B. L., Caraco, N. F., Pace M. L., 1993. Nitrogen loading of rivers as a human-driven process. In: McDonnell, M. J., Pickett, S. T. A. (Eds.), Humans as Components of Ecosystems: The Ecology of Subtle Human Effects and Populated Areas. Springer-Verlag, New York, pp. 138-154.

Costa, J. E., Howes, B. L., Janik, D., Aubrey, D., Gunn, E., Giblin, A. E., 1999. Managing anthropogenic nitrogen inputs to coastal embayments: Technical basis of a management strategy adopted for Buzzards Bay. Buzzards Bay Proj. Tech. Rep.

Dettmann, E., In Press. Effects of water residence time on annual export and denitrification of nitrogen in estuaries: A model analysis. Estuaries.

D'Elia, C. F., Steudler, P. A., Corwin, N., 1977. Determination of total nitrogen in aqueous samples using persulfate digestion. Limnol. Oceanogr. 22, 760-764.

Eichner, E. M., Cambareri, T. C., 1992. Nitrogen Loading. Cape Cod Comm. Tech. Bull. 91-001. Barnstable, Massachusetts, USA.

Eichner, E. M., Cambareri, T. C., Livingston, K., Lawrence, C., Smith, B., Prahm, G., 1998. The Cape Cod Coastal Embayment Project, Cape Cod Commission Water Resources Office, Barnstable, Massachusetts, USA.

Eichner, E. M., Livingston, K., Smith, B., Morrill, V., 1998. Pleasant Bay Nitrogen Loading Analysis. Cape Cod Commission Water Resources Office, Barnstable, Massachusetts, USA.

Eichner, E. M., Cambareri, T. C., Morrill, V., Smith, B., Prahm, G., 2000. Coastal Nitrogen Loading Project, Interim Report, Centerville River/East Bay, Nauset Marsh, Town Cove, and Herring River, Harwich. Cape Cod Commission Water Resources Office, Barnstable, Massachusetts, USA.

Foreman, K., Tomasky, G., Soucy, L. A., Valiela, I., Submitted. Responses of phytoplankton production to land-derived nitrogen loads in Waquoit Bay, Massachusetts, USA. Aquat. Ecol.

Frimpter, M. H., Donahue, J. J., Rapacz, M., 1990. A mass-balance model for predicting nitrate in groundwater. New Eng. Wat. Works Assoc. 104, 219-232.

Frink, C. R., 1991. Estimating nutrient exports to estuaries. J. Environ. Qual. 20, 717-724.

Gaines, A. G., 1986. Lagoon Pond Study: An Assessment of Environmental Issues and Observations on the Estuarine System. Final Report for the Town of Oak Bluffs. Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA.

GESAMP, 1990. The State of the Marine Environment. Joint Group of Experts on the Scientific Aspects of Marine Pollution. Rep. and Stud. 39. United Nations Environmental Programme.

Giblin, A. E., Gaines, A. G., 1990. Nitrogen inputs to a marine embayment: The importance of groundwater. Biogeochemistry 10, 309-328.

Goldberg, E. D., 1995. Emerging problems in the coastal zone for the twenty-first century. Mar. Poll. Bull. 31,152-158.

Hakanson, L., 1999. On the principles and factors determining the predictive success of ecosystem models, with a focus on lake eutrophication models. Ecol. Mod. 121, 139-160.

Hakanson, L., 2000. The role of characteristic coefficients of variation in uncertainty and sensitivity analyses, with examples related to the structuring of lake eutrophication models. Ecol. Mod. 131, 11-20.

Howarth, R. W., Billen, G., Swaney, D., Townsend, A., Jaworski, N. A., Lajtha, K., Downing, J. A., Elmgren, R., Caraco, N., Jordan, T., Berendse, F., Freney, J., Kudeyarov, V., Murdoc, P., Zhao-liang, Z., 1996. Regional nitrogen budgets and riverine N and P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. Biogeochemistry 35, 75-139.

Jakeman, A. J., Hornberger, G. M., 1993. How much complexity is warranted in a rainfall-runoff model? Wat. Resour. Res. 29, 2637-2649.

Johnes, P. J., 1996. Evaluation and management of the impact of land use changes on the nitrogen and phosphorus load delivered to surface waters: The export coefficient modeling approach. J. Hydrol. 183, 323-349.

Kelly, C. A., Rudd, J. W. M., Hesslein, R. H., Schindler, D. W., Dillon, P. J., Driscoll, C. T., Gherini, S. A., Hecky, R. E., 1987. Prediction of biological acid neutralization in acid-sensitive lakes. Biogeochemistry 3, 129-141.

Kellogg, D. Q., Joubert, L., Gold., A., 1996. Technical Documentation, Nutrient Loading Component of the MANAGE Geographic Information System-based Risk Assessment Method. University of Rhode Island Cooperative Extension, Department of Natural Resources Science, Woodward Hall, Kingston RI (http://www.edc.uri.edu/cewq/manage.html).

Koppelman, L. E., 1978. The Long Island comprehensive waste treatment management plan. Volume II: Summary Documentation. Nassau-Suffolk Regional Planning Board. Hauppauge, NY, USA.

Kroeger, K. D., Bowen, J. L., Corcoran, D., Moorman, J., Michalowski, J., Rose, C., Valiela, I., Nitrogen loading to Green Pond, Falmouth, MA: Sources and evaluation of management options. Environ. Cape Cod 2, 15-26.

Lajtha, K., Seely, B., Valiela, I., 1995. Retention and leaching losses of atmospherically derived nitrogen in the aggrading coastal watershed of Waquoit Bay, MA. Biogeochemistry 28, 33-54.

Lamb, B., Gold, A. J., Loomis, G., and McKey, C., 1987. Evaluation of nitrogen removal systems from on-site sewage disposal. In: On-site wastewater treatment. Proceedings of the 5th

National Symposium on Individual and Small Community Sewage Systems. Chicago, IL, pp 151-160.

LeBlanc, D. R., Garabedian, S. P., Hess, K. M., Gelhar, L. W., Quadri, T. D., Stollenwerk, K. G., Wood, W. W., 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. Wat. Resour. Res. 27, 895-910.

Lee, D. R., 1977. A device for measuring seepage in lakes and estuaries. Limnol. Oceanogr. 24, 1146-1154.

Lee, V., Olson, S., 1985. Eutrophication and management initiatives for the control of nutrient inputs to the Rhode Island coastal Lagoons. Estuaries 8, 191-202.

Levins, R., 1966. The strategy of model building in population biology. Amer. Sci. 54, 421-431.

McBride, M. S., Pfannkuch H. O., 1975. The distribution of seepage within lake beds. J. Res. US Geol. Surv. 3, 505-512.

McDonald, M. G., Harbaugh, A. W., 1988. A modular three-dimensional finite-difference groundwater flow model. US Geol. Surv. Open File Rep. 83-875.

Meeuwig, J. J., 1999. Predicting coastal eutrophication from land use: an empirical approach to small non-stratified estuaries. Mar. Ecol. Prog. Ser. 176, 231-241.

Motulsky, H., 1995. Intuitive Biostatistics. Oxford University Press. New York.

National Research Council, 1994. Priorities for Coastal Science. National Academy Press.

Nelson, M. E., Horsley, S. W., Cambareri, T. C., Giggey, M. D., Pinnette J. R., 1988. Predicting nitrogen concentrations in groundwater—An analytical model. In: Conference on Eastern Regional Groundwater Issues. National Well Water Association. Dublin, OH, USA, pp. 179-202.

Nixon, S. W., 1992. Quantifying the relationship between nitrogen input and the productivity of marine ecosystems. Proc. Adv. Mar. Technol. Conf. 5, 57-83.

Oldale, R. N., 1992. Cape Cod and the Islands: The Geologic Story. Parnassus Imprints, East Orleans, Massachusetts, USA.

Omernik, J. M., 1976. The influence of land use on stream nutrient levels. EPA Report 600/3-76-014.

Prairie, Y. T., 1996. Evaluating the predictive power of regression models. Can. J. Fish. Aquat. Sci. 53, 490-492.

Portniaguine, O., Solomon, D. K., 1998. Parameter estimation using groundwater age and head data, Cape Cod, Massachusetts. Wat. Resour. Res. 34, 637-645.

Rastetter, E. B., King, A. W., Cosby, B. J., Hornberger, G. M., O'Neill, R. V., Hobbie J. E., 1992. Aggregating fine-scale ecological knowledge to model coarser scale attributes of ecosystems. Ecol. Appl. 2, 55-70.

Running, S. W., Nemani, R. R., Hungerford, R. D., 1988. Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evaoptranspiration and photosynthesis. Can. J. For. Res. 17, 472-483.

Schlesinger, W. H., 1991. Biogeochemistry: An Analysis of Global Change. Academic Press, San Diego, CA.

Sokal, R. R., Rohlf, F. J., 1995. Biometry. W. H. Freeman, New York.

Solomon, D.K., Poreda, R. J., Cook, P. G., Hunt, A., 1995. Site Characterization using $^3H/^3He$ groundwater ages, Cape Cod, Massachusetts. Ground Water 33, 988-996.

Soranno, P. A., Hubler, S. L., Carpenter, S. R., Lathrop, R. C., 1996. Phosphorus loads to surface waters: A simple model to account for spatial pattern of land use. Ecol. Appl. 6, 865-878.

Steel, R. G. D., Torrie, J. H., 1960. Principles and Procedures of Statistics. McGraw-Hill Book Co., Inc., New York.

Thornthwaite, C. W., Mather, J. R., 1957. Instructions and tables for computing potential evapotranspiration and the water balance. Publications in Climatology. Drexel Institute of Technology 10, 185-311.

Valiela, I., Teal, J. M., Volkmann, S., Shafer, D., Carpenter, E. J., 1978. Nutrient and particulate fluxes in a salt marsh ecosystem: Tidal exchanges and inputs by precipitation and groundwater. Limnol. Oceanogr. 23, 798-812.

Valiela, I., Costa, J., 1988. Eutrophication of Buttermilk Bay, a Cape Cod coastal embayment: Concentrations of nutrients and watershed nutrient budgets. Environ. Manag. 12, 539-553.

Valiela, I., Costa, J., Foreman, K., Teal, J. M., Howes, B., Aubrey, D., 1990. Transport of groundwater-borne nutrients from watersheds and their effects on coastal waters. Biogeochemistry 10, 177-197.

Valiela, I., Foreman, K., LaMontagne, M., Hersh, D., Costa, J., Peckol, P., DeMeo-Anderson, B., D'Avanzo, C., Babione, M., Sham, C. H., Brawley, J., Lajtha, K., 1992. Couplings of watersheds and coastal waters: Sources and consequences of nutrient enrichment in Waquoit Bay, Massachusetts. Estuaries 15, 443-457.

Valiela, I., Collins, G., Kremer, J., Lajtha, K., Geist, M., Seely, B., Brawley, J., Sham, C. H., 1997. Nitrogen loading from coastal watersheds to receiving estuaries: New method and application. Ecol. Appl. 7, 358-380.

Valiela, I., Geist, M., McClelland, J., Tomasky, G., 2000. Nitrogen loading from watersheds to estuaries: Verification of Waquoit Bay Nitrogen Loading Model. Biogeochemistry 49, 277-293.

Valiela, I., Tomasky, G., Hauxwell, J., Cole, M. L., Cebrian, J., Kroeger, K., 2000a. Operationalizing sustainability: Management and risk assessment of land-derived nitrogen loads to estuaries. Ecol. Appl. 10, 1006-1023.

Valiela, I., Bowen, J. L., Stieve, E. L., Cole, M. L., Kroeger, K. D., Tomasky, G., Geist, M., Submitted. ELM, an estuarine nitrogen loading model: Verification of predicted nitrogen concentrations and links to producers. Ecol. Mod.

Vollenweider, R. A., 1976. Advances in defining critical loading levels for phosphorus in lake eutrophication. Mem. Ist. Ital. Idrobiol. 33, 53-83.

Figure legends

Figure 1. Chlorinity of pore water at different distances away from shore. Measurements were taken along transects perpendicular to shore in the same areas where we measured groundwater flows with seepage meters.

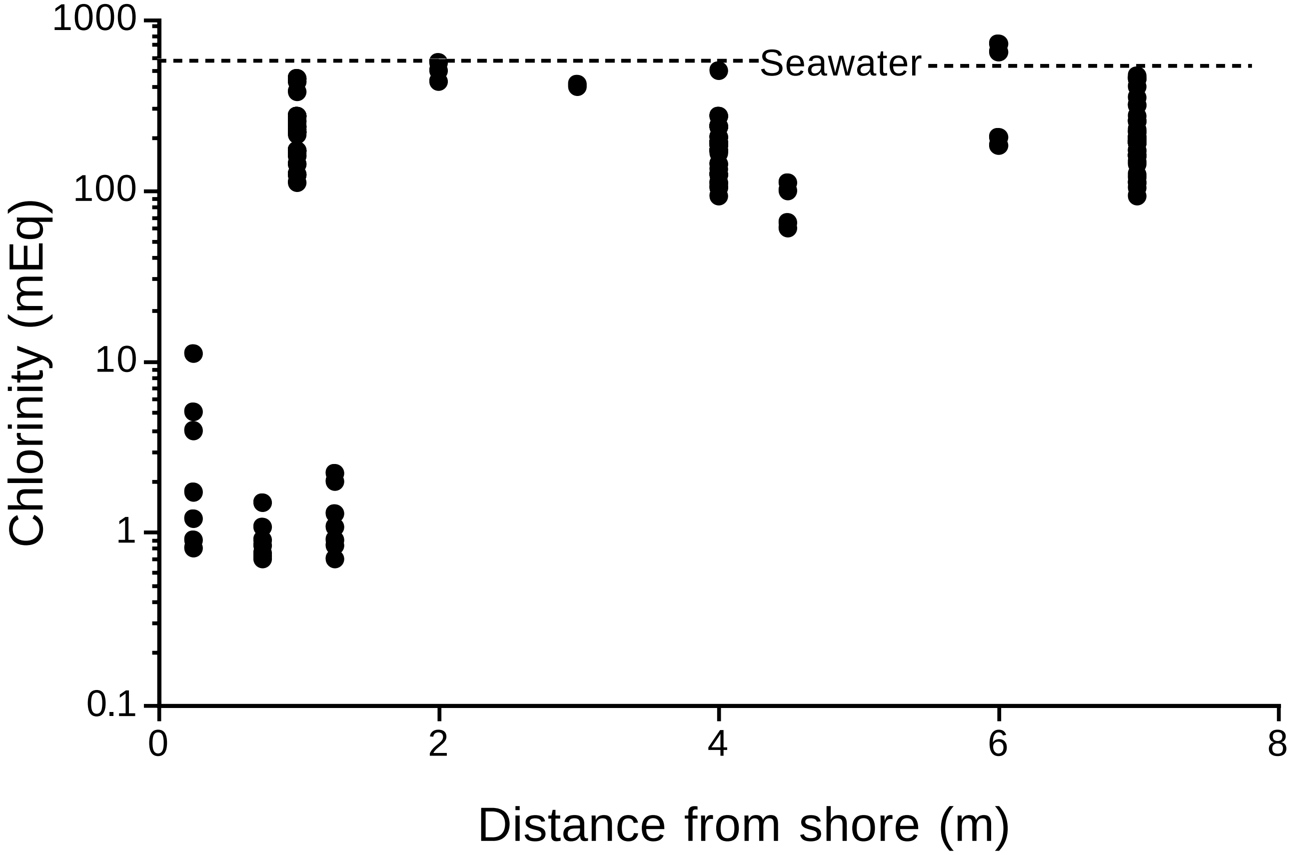
Figure 2. Comparison of annual flows of groundwater measured using seepage meters versus estimates obtained using the recharge values corrected for evapo-transpiration. Dashed line shows the 1:1 line of perfect fit.

Figure 3. Comparisons of land-derived nitrogen loads predicted by different models versus measured nitrogen loads. Acronyms as described in text, dashed line shows 1:1 line of perfect fit. Three simulations of the CCC model (center left panel) are included [using loss coefficients and the averaging procedure as described in Eichner and Cambareri (1992, squares), not using the averaging procedure described in Eichner and Cambareri (1992, circles), and using loss coefficients for each local watershed and not averaging the two methods described in text (triangles)]. Three simulations with the PJM are included [in center right panel, with original loss coefficients (circles), with local loss coefficients (triangles), and in bottom left panel, omitting terrestrial N_2 fixation]. Measured data from Valiela et al. (2000), and Kroeger et al. (1999).

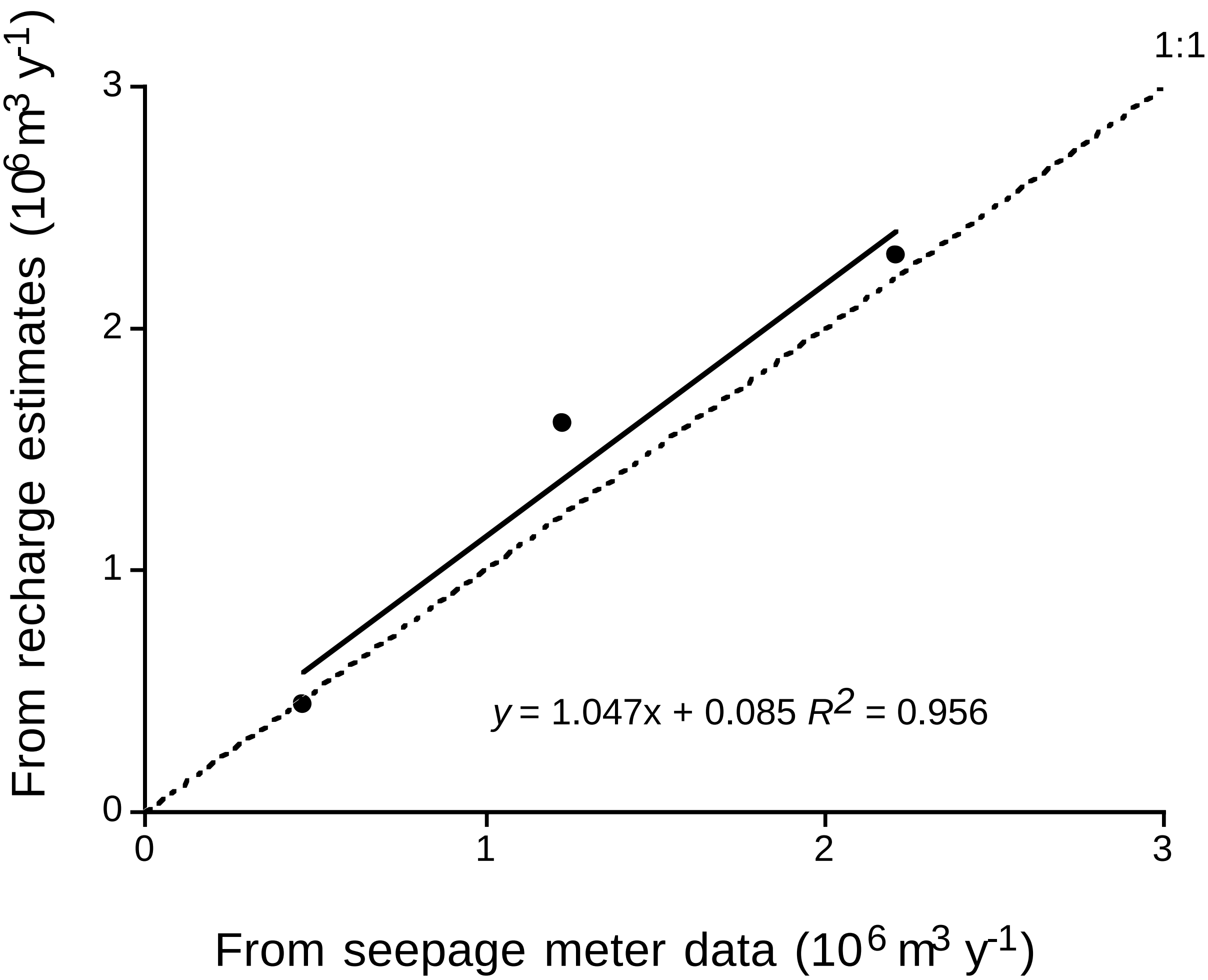
Figure 4. Comparisons of land-derived nitrate loads predicted by three models versus measured nitrate loads. Acronyms as described in text, dashed line shows 1:1 line of perfect fit. The top panel includes three simulations of the CCC model [using original loss coefficients and the averaging procedure (squares), not using the averaging procedure described in Eichner and Cambareri (1992, circles) and using only local loss coefficients and not averaging the two methods (triangles)]. Measured data from Valiela et al. (2000), and Kroeger et al. (1999).

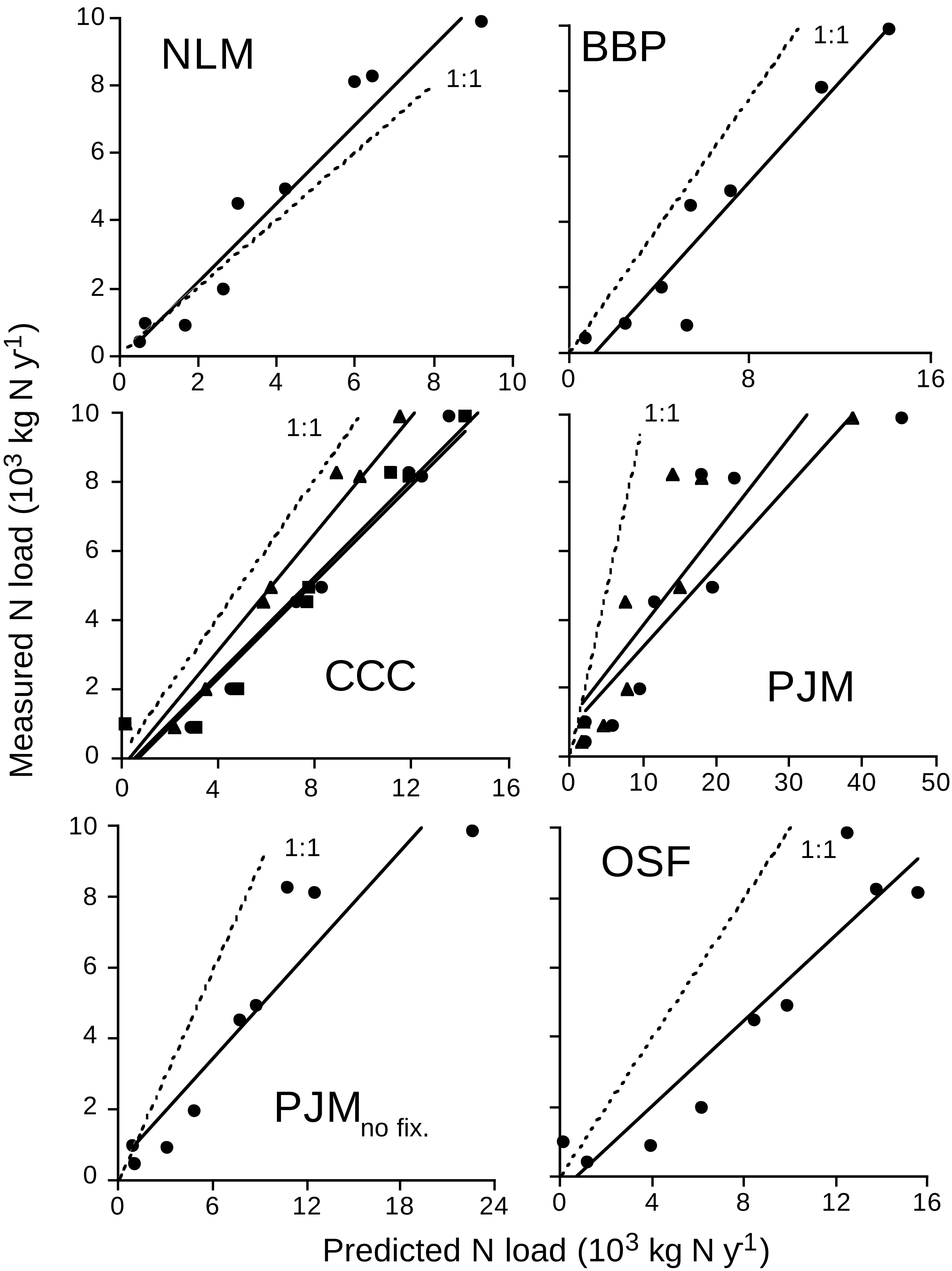
Figure 5. Comparisons of different nitrogen concentrations predicted by three models versus measured nitrogen concentrations. Acronyms as described in text, dashed line shows 1:1 line of perfect fit. Measured data from Valiela et al. (2000), and Kroeger et al. (1999).

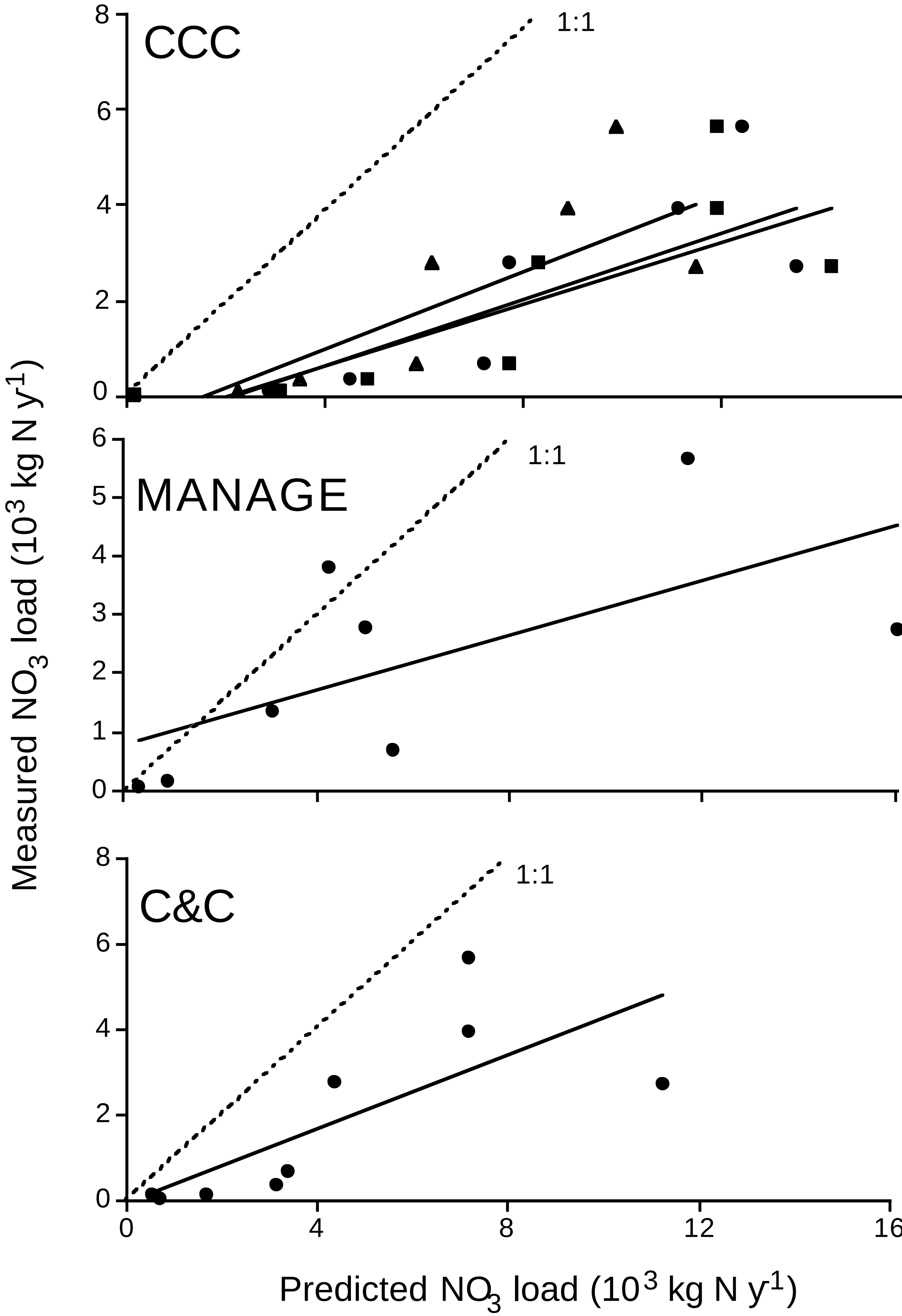
Figure 6. Comparison of model performance in terms of responsiveness, precision, accuracy, and predictive ability. See text for definitions of statistical usage. Calculations based on black circles, which show points based on the original description of each model. White symbols show predictions of models in our altered versions of the CCC (circles) and PJM (triangles) models.

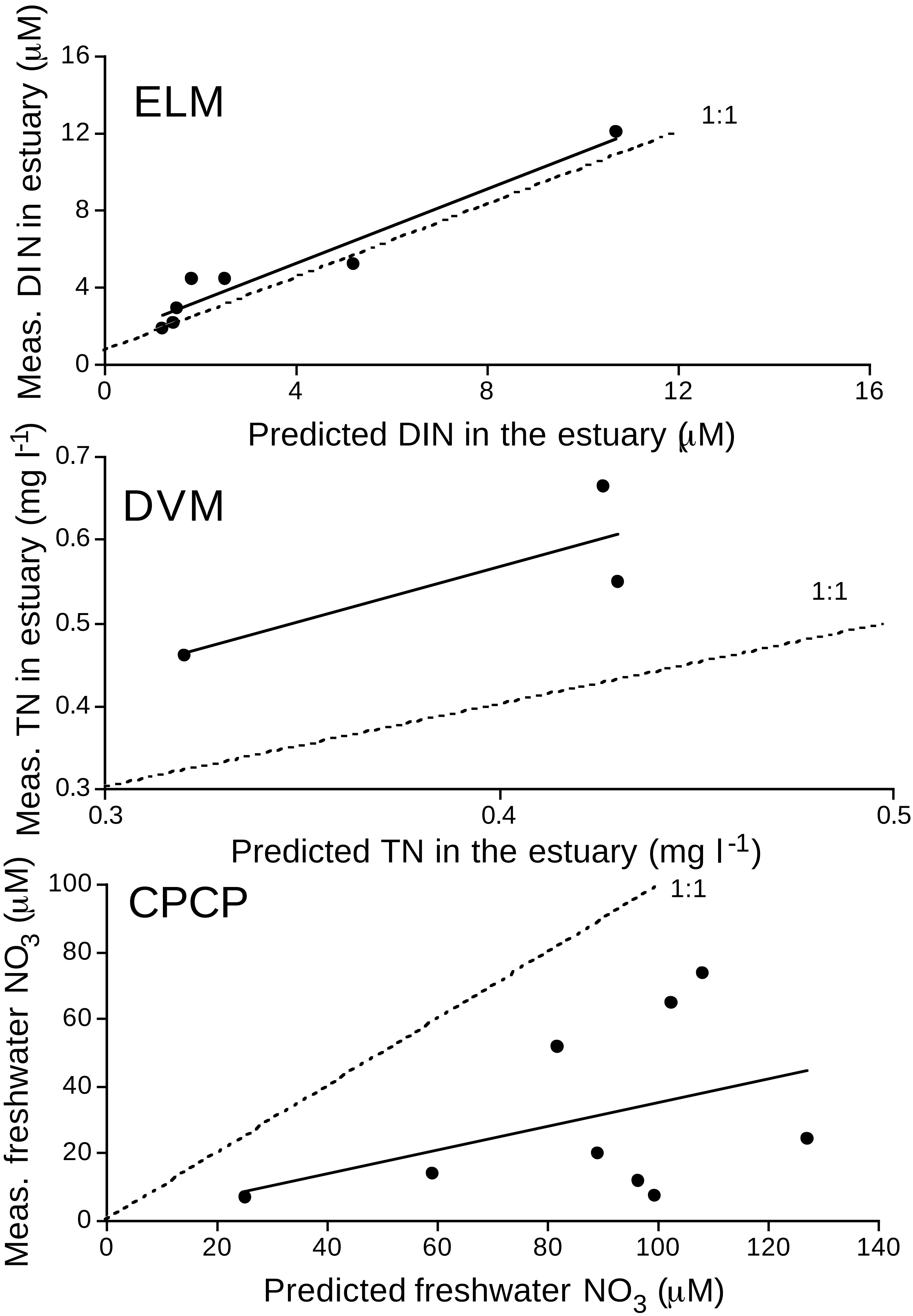


Chlorinity (mEq)
1000
100
10
1
0.1
Seawater
0
2
4
6
8
Distance from shore (m)


1:1
3
2
1
0
0
1
2
3
y = 1.047x + 0.085 R2 = 0.956
From recharge estimates (106 m3 y-1)
From seepage meter data (106 m3 y-1)


NLM
1:1
BBP
1:1
1:1
CCC
1:1
PJM
1:1
PJMno fix.
OSF
1:1
Measured N load (10^3 kg N y^1)
Predicted N load (10^3 kg N y^1)
0
2
4
6
8
10
0
8
16
0
4
8
12
16
0
10
20
30
40
50
0
6
12
18
24


CCC
MANAGE
C&C
1:1
1:1
1:1
0
2
4
6
8
0
1
2
3
4
5
6
0
2
4
6
8
0
4
8
12
16
Measured NO3 load (10^3 kg N y^-1)
Predicted NO3 load (10^3 kg N y^-1)



ELM
1:1
Meas. DIN in estuary (μM)
16
12
8
4
0
0
4
8
12
16
Predicted DIN in the estuary (μM)
DVM
1:1
Meas. TN in estuary (mg l-1)
0.7
0.6
0.5
0.4
0.3
0.3
0.4
0.5
Predicted TN in the estuary (mg l-1)
CPCP
1:1
Meas. freshwater NO3 (μM)
100
80
60
40
20
0
0
20
40
60
80
100
120
140
Predicted freshwater NO3 (μM)

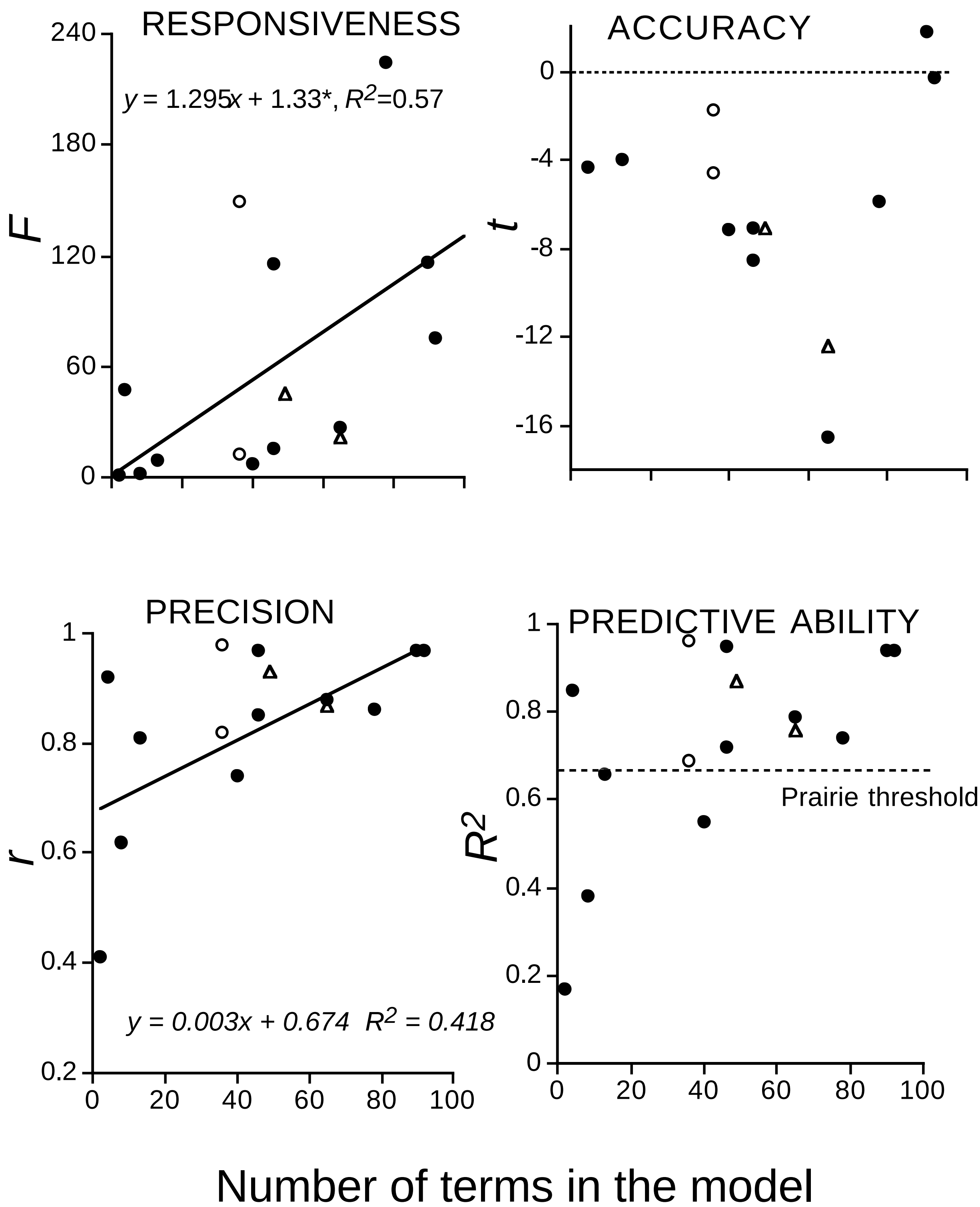

RESPONSIVENESS
y = 1.295x + 1.33*, R2=0.57
F
240
180
120
60
0
ACCURACY
t
0
-4
-8
-12
-16
PRECISION
r
1
0.8
0.6
0.4
0.2
y = 0.003x + 0.674 R2 = 0.418
0
20
40
60
80
100
PREDICTIVE ABILITY
R2
1
0.8
0.6
0.4
0.2
0
Prairie threshold
Number of terms in the model

Table 1. Calculated rates (x ± s.d.) of flow of fresh- and seawater into sets of 9 seepage meters deployed in the summer and fall in three different estuaries of Waquoit Bay.

		Flow rates (L m^{-2} hr^{-1})		
		Childs River	Quashnet River	Sage Lot Pond
Freshwater flow	Jun-Jul	5.6 ± 6.4	3.3	1.8 ± 1.2
	Oct-Nov	3.4 ± 3.5	6.6 ± 11.9	3.6 ± 0.95
	Mean	4.5	5.0	2.7
Seawater flow	Jun-Jul	4.6 ± 3.5	5.2 ± 5.4	1.8 ± 1.8
	Oct-Nov	4.4 ± 3.7	2.5 ± 3.4	1.0 ± 1.4
	Mean	4.5	3.9	1.4

Table 2. Comparison of the annual flows of fresh groundwater into sections of estuaries of Waquoit Bay, as obtained from measured flows using seepage meters, and by calculations based on annual freshwater recharge.

	Childs River	Quashnet River	Sage Lot Pond
Calculated seepage face dimensions (m^2)[a]	31,152	50,976	19,824
Mean flow of freshwater ($l\ m^{-2}\ yr^{-1}$) calculated from seepage meter data	4.5	5.0	2.7
Annual flow of groundwater ($m^3\ yr^{-1}$) estimated from seepage meter data	$1.2x10^6$	$2.2x10^6$	$0.46x10^6$
Annual flow of groundwater ($m^3\ yr^{-1}$) estimated from recharge data[b]	$1.6x10^6$	$2.2x10^6$	$0.44x10^6$
% of annual recharge volume accounted for by flow through the seepage face	77	96	105

[a]These numbers were calculated from values for the length of the periphery of the seepage face around each estuary, obtained from aerial photos (3,894; 6,372; and 2478 meters, respectively for CR, QR, and SLP). Periphery was multiplied by 8 m for height of seepage face, from Fig. 2.

[b]Estimated from area of catchment feeding the portion of the shores within CR, QR, and SLP where we deployed the seepage meters, multiplied by total precipitation x 45% (correction for evapotranspiration) to get annual recharge (Lajtha et al. 1995).

Table 3. Statistical evaluation of the predictions from the models included in this paper.

Models	Measured N load or conc. =	F_{reg}	r	R^2	$t_{vs.\ 1:1}$	Correction needed to improve accuracy
Predicting loads						
NLM	1.19 x-152.5	116.6**	0.97**	0.94	1.79 ns	—
MANAGE	0.29 x – 22.4	7.2*	0.74*	0.55	-6.62 **	71%
BBP	0.66 x + 327.2	224.6**	0.86**	0.74	-5.9 **	44%
CCC vs. TDN						
Original version	0.74x-431.8	115**	0.98**	0.95	-7.1 **	26%
No averaging	0.72x-444.4	124**	0.97**	0.95	-4.20 *	28%
No averaging, local coeff.	0.88x-330.8	150**	0.98**	0.96	-3.5 ns	—
CCC vs. NO_3						
Original version	0.35x-712.7	15.1**	0.85**	0.72	-8.5 **	65%
No averaging	0.37x-677.6	14.5**	0.84**	0.71	-6.5 **	63%
No averaging, local coeff.	0.43x-541.9	12.2*	0.82**	0.67	-5.4 **	57%
PJM						
Original	0.24x + 825.9	26.7**	0.88**	0.79	-16.5 **	76%
Local coeff.	0.27x+1,113.1	21.9**	0.87**	0.76	-12.4 **	73%
Local coeff., no N_2 fixation	0.49x + 505.3	45.6**	0.93**	0.87	-7.03 **	51%
C&C	0.59x-476.6	8.99**	0.81**	0.66	-3.96 **	41%
OSF	0.61x - 440.0	47.4**	0.92**	0.85	-4.3 **	43%
Predicting concentrations						
CPCP	0.35x-0.45	1.4 n.s.	0.41 n.s.	0.17	—	65%
DVM	1.32x+0.04	1.9 n.s.	0.81 n.s.	0.66	—	-32%
ELM	0.97x+1.43	75.6**	0.97**	0.94	-0.29 ns	—

Table 4. Number of term and inputs needed by the models included in this paper, with a summary of the assessments of performance. We assess performance based on the statistics of Table 3, as "+++" to indicate higher, "++", intermediate, and "+", lower performance.

Models	Number of terms	Number of inputs	Summary assessment of model performance			
			Responsiveness	Precision	Accuracy	Predictive ability
Predicting loads						
NLM	90	16	+++	+++	+++	+++
MANAGE*	40	19	++	+	+	++
BBP	78	25	+++	+++	+	++
CCC vs. TDN						
Original	46	11	+++	+++	+	+++
No averaging	39	10	+++	+++	++	+++
No averaging, local coeff.	39	10	+++	+++	+++	+++
CCC vs. NO_3						
Original	46	11	+++	+++	+	+++
No averaging	39	10	+++	+++	+	++
No averaging, local coeff.	39	10	+++	+++	+	++
PJM						
Original	65	12	+++	+++	+	++
Local coeff.	65	12	+++	+++	+	++
Local coeff., no N_2 fixation	49	12	+++	+++	+	+++
C&C	13	4	++	++	+	++
OSF	4	1	+++	+++	+	+++
Predicting concentrations						
CPCP	2	2	+	+	+	+

DVM	8	6	+	+	+	++
ELM	92	16	+++	+++	+++	+++

* Includes only the subsurface transport component of MANAGE.

Table 5. Percentage of total dissolved nitrogen in groundwater input to different estuaries of Waquoit Bay, MA. From Valiela et al. (2000).

	% of TDN		
Estuary	NO_3	NH_4	DON
Childs River	70	10	20
Hamblin Pond	34	8	59
Quashnet River	28	14	79
Jehu Pond	19	13	68
Eel Pond	15	24	61
Sage Lot Pond	6	15	79

Table 6. The percentage of nitrogen contributed by the three major sources predicted by the different models. ELM, DPD, and CPCP are not included because they predict N concentrations to the Waquoit Bay estuarine system rather than land-derived loads.

Model	Atmospheric N	Fertilizer N	Wastewater N
NLM	32	21	47
CCC			
Original	1	16	83
No averaging	2	17	81
No averaging, our coeff.	2	21	77
PJM			
Original	30	51	18
Our coeff.	11	76	13
Our coeff. No N_2 fixation	57	8	35
C&C	27	24	49
MANAGE*	44		56
BBP*	48		52
MVM*	42		58

* Model does not explicitly differentiate between loads from fertilizer application and from atmospheric deposition, and so we include only the combined term.