Journal of Hydrology 396 (2011) 158–169

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Rating curve estimation of nutrient loads in Iowa rivers

Greg A. Stenback [a,*], William G. Crumpton [a], Keith E. Schilling [b], Matthew J. Helmers [c]

[a] Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50011, United States
[b] Iowa Geological and Water Survey, Iowa Department of Natural Resources, Iowa City, IA 52242, United States
[c] Agricultural and Biosystems Engineering, Iowa State University, Ames, IA 50011, United States

ARTICLE INFO

Article history:
Received 19 April 2010
Received in revised form 5 November 2010
Accepted 8 November 2010

This manuscript was handled by Geoff Syme, Editor-in-Chief.

Keywords:
River nutrient load
Nitrate load estimation
Total phosphorus load estimation

SUMMARY

Accurate estimation of nutrient loads in rivers and streams is critical for many applications including determination of sources of nutrient loads in watersheds, evaluating long-term trends in loads, and estimating loading to downstream waterbodies. Since in many cases nutrient concentrations are measured on a weekly or monthly frequency, there is a need to estimate concentration and loads during periods when no data is available. The objectives of this study were to: (i) document the performance of a multiple regression model to predict loads of nitrate and total phosphorus (TP) in Iowa rivers and streams; (ii) determine whether there is any systematic bias in the load prediction estimates for nitrate and TP; and (iii) evaluate streamflow and concentration factors that could affect the load prediction efficiency. A commonly cited rating curve regression is utilized to estimate riverine nitrate and TP loads for rivers in Iowa with watershed areas ranging from 17.4 to over 34,600 km^2. Forty-nine nitrate and 44 TP datasets each comprising 5–22 years of approximately weekly to monthly concentrations were examined. Three nitrate data sets had sample collection frequencies averaging about three samples per week. The accuracy and precision of annual and long term riverine load prediction was assessed by direct comparison of rating curve load predictions with observed daily loads. Significant positive bias of annual and long term nitrate loads was detected. Long term rating curve nitrate load predictions exceeded observed loads by 25% or more at 33% of the 49 measurement sites. No bias was found for TP load prediction although 15% of the 44 cases either underestimated or overestimate observed long-term loads by more than 25%. The rating curve was found to poorly characterize nitrate and phosphorus variation in some rivers.

Published by Elsevier B.V.

1. Introduction

Accurate estimation of nutrient loads in rivers and streams is critical for many applications, such as determining sources of nutrient loads in watersheds (Alexander et al., 2008; Preston et al., 1989; Robertson et al., 2009), development of waste-load allocation schemes via the Total Maximum Daily Load (TMDL) program (US Environmental Protection Agency, 2000), calibrating and validating watershed models (Ullrich and Volk, 2010; Gassman et al., 2007), evaluating long-term trends in loads (Littlewood et al., 1998; Schilling and Zhang, 2004) and estimating riverine flux to the Gulf of Mexico (Goolsby et al., 2000) and oceans (Meybeck, 1982). The instantaneous load is the product of nutrient concentration $C(t)$ and discharge $Q(t)$ at time t and the load over an extended period of time T is given by

$$L_T = \int_0^T C(t)Q(t)dt.$$

In practice, continuous measures of concentration and discharge are not available and the integration is approximated with a summation. Although daily or more frequent discharge measurements on many streams and rivers are available from the United States Geological Survey (USGS) or associated with specific projects, nutrient concentrations are often measured less frequently, such as weekly or monthly. Accordingly, to estimate nutrient loads over an extended period of time, such as monthly or annually, concentrations or loads must be estimated to fill in time periods when no data are available. It is recognized that estimation of nutrient loads is a problem subjected to many potential sources of error and uncertainty (Guo et al., 2002).

Many studies have compared various methods for load estimation (e.g., Preston et al., 1989; Robertson and Roerish, 1999; Guo et al., 2002; Aulenbach and Hooper, 2006; Moatar and Meybeck, 2005; Li et al., 2006; Zamyadi et al., 2007; Ullrich and Volk, 2010). In some cases, load uncertainty and model performance were evaluated by undersampling against a true measured load (i.e., Guo et al., 2002; Li et al., 2006) whereas in others, different algorithm methods were applied to the same dataset (i.e., Moatar and Meybeck, 2005; Zamyadi et al., 2007). Results from these studies indicate there is often wide variability in nutrient load

* Corresponding author. Tel.: +1 515 294 3983.
E-mail address: gastenba@iastate.edu (G.A. Stenback).

0022-1694/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.jhydrol.2010.11.006

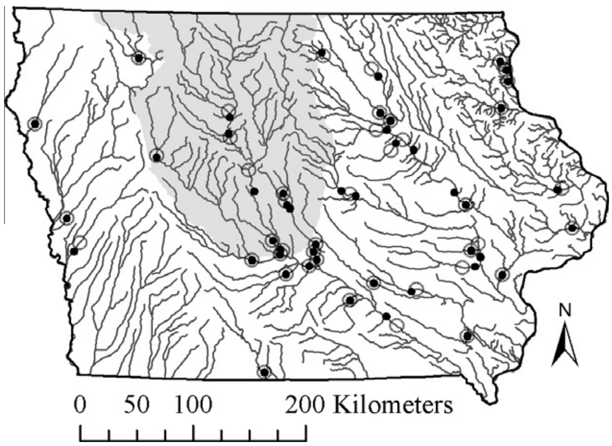

Fig. 1. Sampling sites (small circles) and corresponding river gage stations (large open circles) with Iowa's major river systems and the Des Moines Lobe.

estimates. Estimation of annual phosphorus loads have shown errors of 30% (Robertson and Roerish, 1999) and 34% (Moatar and Meybeck, 2005) and annual nitrate load estimates have differed by as much as 64% depending on sampling strategy, load estimation method and monitoring period used (Ullrich and Volk, 2010).

In this study, we focused exclusively on the accuracy of one method, the widely used multiple regression model of Cohn et al. (1992), to evaluate how the model predicted nitrate and phosphorus loads at many locations in Iowa. This model is incorporated into the USGS LOADEST (2004) computer program for load estimation, which is widely used to estimate N and P loads in rivers (e.g., Goolsby et al., 2000; Goolsby and Battaglin, 2001; Hooper et al., 2001; Aulenbach, 2006; Maret et al., 2008; Eshleman et al., 2008). The model has been utilized by the USGS to estimate nutrient flux in major rivers flowing to the Gulf of Mexico (USGS, 2009a) and to calculate "observed" loads in the USGS SPARROW model (USGS, 2009b). Both of these applications of the Cohn et al. (1992) regression model have a great deal of significance for

Table 1
Nitrate data source and sampling characteristics.

Sample site	Nitrate data source[a]	n	Non-detect count	Sample date range	Average days per sample
Beaver Creek near Cedar Falls	IA STORET	121	0	September-99 to September-08	27.1
Beaver Creek near Grimes	IA STORET	108	2	October-99 to September-08	30.1
Black Hawk Creek at Waterloo	IA STORET	92	0	October-01 to September-08	27.5
Bloody Run Creek Site #1	IA STORET	600	0	October-91 to August-08	10.3
Boyer River near Missouri Valley	IA STORET	107	0	October-99 to August-08	30.2
Cedar Creek near Bussey	IA STORET	131	28	October-99 to September-08	24.7
Cedar Creek near Oakland Mills	IA STORET	119	28	October-98 to September-08	30.4
Cedar River at Cedar Rapids	CRWD	1105	0	January-01 to May-08	2.4
Cedar River downstream of Waterloo	IA STORET	96	0	October-00 to September-08	30.1
Cedar River near Charles City	IA STORET	96	0	October-00 to September-08	30.1
Cedar River near Conesville	IA STORET	108	1	October-99 to September-08	30.1
Cedar River near Janesville	IA STORET	109	0	September-99 to September-08	30.1
Cedar River upstream of Cedar Rapids	IA STORET	96	0	October-00 to September-08	30.1
Des Moines River DMRWQN Site 1	DMRWQN	200	15	January-99 to February-08	16.6
Des Moines River downstream of Des Moines	IA STORET	96	0	October-00 to September-08	30.1
Des Moines River near Runnells	DMRWQN	537	0	October-85 to December-07	15.1
Des Moines River upstream of Fort Dodge	IA STORET	96	3	October-00 to September-08	30.0
Des Moines River upstream of Ottumwa	IA STORET	96	0	October-00 to September-08	30.1
Des Moines River, Des Moines	DMWW	1857	0	December-98 to May-08	2.2
English River at Riverside	IA STORET	222	11	January-99 to September-08	15.9
Floyd River near Sioux City	IA STORET	139	0	October-98 to August-08	25.8
Iowa River downstream of Iowa City	IA STORET	97	0	May-00 to September-08	31.4
Iowa River downstream of Marshalltown	IA STORET	96	0	October-00 to September-08	30.0
Iowa River upstream of Marshalltown	IA STORET	96	1	October-00 to September-08	30.0
Maquoketa River near Maquoketa	IA STORET	107	0	October-99 to August-08	30.1
Middle River near Indianola	IA STORET	108	23	October-99 to September-08	30.1
North Raccoon River near Sac City	IA STORET	139	1	November-98 to August-08	25.6
North Skunk River	IA STORET	108	4	October-99 to September-08	30.1
Ocheyedan River near Spencer	IA STORET	107	2	September-99 to August-08	30.5
Old Mans Creek near Iowa City	IA STORET	186	5	May-00 to September-08	16.4
Raccoon River at Van Meter	DMRWQN	198	0	January-99 to December-07	16.4
Raccoon River upstream of Des Moines	IA STORET	102	0	October-00 to September-08	28.3
Raccoon River, Des Moines	DMWW	1583	0	January-99 to May-08	2.2
Shell Rock River at Shell Rock	IA STORET	126	0	October-99 to September-08	25.7
Sny Magill	IA STORET	519	0	October-91 to September-01	7.0
Soldier River near Pisgah	IA STORET	119	0	October-98 to August-08	30.1
South Skunk River near Oskaloosa	IA STORET	102	2	October-99 to March-08	30.1
South Skunk River upstream of Ames	IA STORET	95	2	October-00 to August-08	30.1
South Skunk River, above Ames WWTP	Ames WWTP	229	11	December-02 to January-08	8.2
South Skunk River, below Ames WWTP	Ames WWTP	259	1	December-02 to January-08	7.3
Squaw Creek SQW2	Project data[b]	202	0	October-95 to September-05	18.1
Thompson Fork – Grand River at Davis City	IA STORET	108	44	October-99 to September-08	30.1
Turkey River near Garber	IA STORET	107	0	October-99 to August-08	30.1
Walnut Creek WNT1	Project data[b]	204	0	October-95 to September-05	17.9
Walnut Creek WNT2	Project data[b]	204	0	October-95 to September-05	17.9
Wapsipinicon River at De Witt	IA STORET	111	3	October-98 to August-08	32.3
West Fork Des Moines River near Humboldt	IA STORET	107	3	October-99 to August-08	30.1
Winnebago River upstream of Mason City	IA STORET	107	0	October-00 to September-08	27.0
Yellow River near Volney	IA STORET	115	0	October-04 to September-08	12.4

[a] Iowa STORET database (IA STORET), Des Moines Water Works (DMWW), Cedar Rapids Water Department (CRWD), Ames Waste Water Treatment Plant (WWTP), Des Moines River Water Quality Network (DMRWQN).
[b] Walnut Creek Watershed Monitoring Project (Schilling et al., 2006).

agricultural states such as Iowa that are major contributors to nutrient loads (Alexander et al., 2008; Goolsby et al., 1999; Robertson et al., 2009).

We evaluated the model performance by comparing model-predicted nutrient loads to actual measured loads. How well the regression model predicts measured values is assumed to be indicative of how well the model might be expected to perform for days when no samples were collected. The objectives of this study were to: (i) document the performance of the multiple regression load model to predict loads of nitrate and total phosphorus (TP) in Iowa rivers and streams; (ii) determine whether there is any systematic bias in the load prediction estimates for nitrate and TP; and (iii) evaluate streamflow and concentration factors that could affect the load prediction efficiency.

2. Methods

The load prediction model developed by Cohn et al. (1992) is

$$\ln(L) = \beta_0 + \beta_1 \ln(Q) + \beta_2[\ln(Q)]^2 + \beta_3 t + \beta_4 t^2 + \beta_5 \sin(2\pi t) + \beta_6 \cos(2\pi t) + \varepsilon \quad (1)$$

where $L = CQ$ is the load or flux, C is concentration, Q is discharge, t is time in decimal years, $\beta_0, \beta_1, \ldots, \beta_6$ are regression coefficients, and ε is assumed to be an independent and normally distributed error with zero mean and constant variance. Eq. (1) is a model of the logarithm of load and thus a back-transform with bias correction to the data scale is necessary. Both Gilroy et al. (1990) and Cohn et al. (1989) provide equations for the antilog back-transform and these are built into the *LOADEST* computer program. In addition to Eq. (1), *LOADEST* has other models the user can choose from and provides the ability to specify user defined models. *LOADEST* provides an adjusted maximum likelihood estimation (AMLE) procedure for data sets containing censored values, which are most commonly associated with concentration results determined to be below an analytical detection limit. *LOADEST* also provides a least absolute deviation (LAD) procedure for cases where the regression residuals do not appear to follow the normal distribution and constant variance assumptions, however, the LAD output is not provided if there are censored data values. Because $\ln(L) = \ln(QC) = \ln(C) + \ln(Q)$ and $\ln(Q)$ is one of the predictor variables, regressing either $\ln(C)$ or $\ln(L)$ on the terms on the right hand side of Eq. (1) will yield identical load predictions, concentration predictions, and regression residuals. The ordinary least squares regression properties and

Table 2
Total phosphorus (TP) data source and sampling characteristics.

Sample site	TP data source[a]	n	Non-detect count	Sample dates	Average days per sample
Beaver Creek near Cedar Falls	IA STORET	125	22	October-98 to April-09	30.5
Beaver Creek near Grimes	IA STORET	109	4	October-99 to April-09	31.8
Black Hawk Creek at Waterloo	IA STORET	94	5	October-01 to April-09	29.1
Bloody Run Creek Site #1	IA STORET	601	346	October-91 to April-09	10.6
Boyer River near Missouri Valley	IA STORET	108	0	October-99 to April-09	32.2
Cedar Creek near Bussey	IA STORET	132	24	October-99 to April-09	26.1
Cedar Creek near Oakland Mills	IA STORET	120	9	October-98 to April-09	31.9
Cedar River Downstream of Waterloo	IA STORET	99	2	November-99 to April-09	34.7
Cedar River near Charles City	IA STORET	97	3	October-00 to April-09	32.0
Cedar River near Conesville	IA STORET	109	3	October-99 to April-09	31.8
Cedar River near Janesville	IA STORET	113	12	October-98 to April-09	33.7
Cedar River Upstream of Cedar Rapids	IA STORET	99	6	November-99 to April-09	34.7
Des Moines River Downstream of Des Moines	IA STORET	99	1	November-99 to April-09	34.6
Des Moines River near Runnells	DMRWQN	197	1	January-99 to December-07	16.5
Des Moines River Site 1	DMRWQN	197	0	January-99 to December-07	16.5
Des Moines River Upstream of Fort Dodge	IA STORET	99	4	October-99 to April-09	34.9
Des Moines River Upstream of Ottumwa	IA STORET	99	0	November-99 to April-09	34.8
English River at Riverside	IA STORET	225	10	October-98 to April-09	17.0
Floyd River near Sioux City	IA STORET	140	0	October-98 to April-09	27.4
Iowa River Downstream of Iowa City	IA STORET	102	0	November-99 to July-09	34.3
Iowa River Downstream of Marshalltown	IA STORET	99	1	October-99 to April-09	35.0
Iowa River Upstream of Marshalltown	IA STORET	99	9	October-99 to April-09	35.0
Maquoketa River near Maquoketa	IA STORET	108	13	October-99 to April-09	32.0
Middle River near Indianola	IA STORET	113	12	October-98 to April-09	33.8
North Raccoon River near Sac City	IA STORET	65	0	July-03 to May-09	32.9
North Skunk River	IA STORET	113	10	October-98 to April-09	33.9
Ocheyedan River near Spencer	IA STORET	111	22	October-98 to April-09	34.5
Old Mans Creek near Iowa City	IA STORET	187	10	May-00 to April-09	17.4
Raccoon River at Van Meter	DMRWQN	197	0	January-99 to December-07	16.5
Raccoon River Upstream of Des Moines	IA STORET	105	3	November-99 to April-09	32.6
Shell Rock River at Shell Rock	IA STORET	127	9	October-99 to April-09	27.1
Sny Magill Creek Site #1	IA STORET	517	349	October-91 to September-01	7.0
Soldier River near Pisgah	IA STORET	120	8	October-98 to April-09	31.9
South Skunk River near Oskaloosa	IA STORET	109	3	October-99 to April-09	31.8
South Skunk River Upstream of Ames	IA STORET	99	6	October-99 to May-09	35.4
Squaw Creek W2	Project data[b]	112	3	October-00 to September-05	16.1
Thompson Fork – Grand River at Davis City	IA STORET	109	16	October-99 to April-09	31.7
Turkey River near Garber	IA STORET	108	19	October-99 to April-09	32.0
Walnut Creek T1	Project data[b]	113	8	October-00 to September-05	15.9
Walnut Creek T2	Project data[b]	113	3	October-00 to September-05	15.9
Wapsipinicon River at De Witt	IA STORET	112	8	October-98 to April-09	34.2
West Fork Des Moines River near Humboldt	IA STORET	112	6	November-98 to April-09	33.9
Winnebago River Upstream of Mason City	IA STORET	110	12	November-99 to April-09	31.2
Yellow River near Volney	IA STORET	116	1	October-04 to April-09	14.2

[a] Iowa STORET database (IA STORET), Des Moines River Water Quality Network (DMRWQN).
[b] Walnut Creek Watershed Monitoring Project (Schilling et al., 2006).

diagnostics apply to the log-space regression output. The back-transform to the data scale is separate from the regression and is conducted to transform the log-space regression output to the data scale and some regression properties will not apply to the back-transformed values. For example, the average regression ln(L) prediction for the calibration data will equal the average of the ln(L) observations, but the back-transform average predicted load will rarely, if ever, equal the average observed load and deviations should be expected whenever the underlying statistical assumptions are not valid. Although the regression R^2 for a fit to ln(L) will be greater than the R^2 for a fit to ln(C), neither are applicable to the load prediction after the back-transform from the log-space to the data scale.

The data used for this work are publically available and specific sampling details were not under our control. These are typical datasets that might be used by research or government agencies for water quality assessment and our intent is to assess the use of regression Eq. (1) using these publically available datasets. Nitrate and TP data for Iowa streams available from Iowa's STORET Database, the Des Moines River Water Quality Network (DMRWQN), and several long-term project Iowa Geological and Water Survey studies were used in the analysis. In total, data from 49 nitrate and 44 TP sampling locations sufficiently close to US Geological Survey (USGS) gage stations having discharge data were selected to evaluate the performance of the rating curve regression given by Eq. (1) (Fig. 1). Nitrate and TP sampling site, sampling frequency, and selected sampling characteristics are listed in Tables 1 and 2, respectively, and corresponding flow station information is listed in Table 3.

For each data set, a regression of the form given in Eq. (1) was run and the *LOADEST* AMLE output was used for subsequent comparisons. We compared loads predicted by the model to observed loads on days when a surface water concentration was measured. Herein we use the term "partial load" to describe the summation of daily loads over a specified time period for only those days having a measured concentration. In practice, interest may be focused on monthly, seasonal, annual, or longer term loads, but in this work, annual and period of record partial loads are used to compare predicted loads with observed loads. For comparative purposes, annualized loads were calculated as the average daily load

Table 3
Flow data sources for water sampling sites.

Sample site	USGS flow station	Flow station name	Drainage area (km^2)
Beaver Creek near Cedar Falls	05463000	Beaver Creek at New Hartford, IA	899
Beaver Creek near Grimes	05481950	Beaver Creek near Grimes, IA	927
Black Hawk Creek at Waterloo	05463500	Black Hawk Creek at Hudson, IA	785
Bloody Run Creek Site #1	05389400	Bloody Run Creek near Marquette, IA	88.40
Boyer River near Missouri Valley	06609500	Boyer River at Logan, IA	2256
Cedar Creek near Bussey	05489000	Cedar Creek near Bussey, IA	969
Cedar Creek near Oakland Mills	05473400	Cedar Creek near Oakland Mills, IA	1373
Cedar River at Cedar Rapids	05464500	Cedar River at Cedar Rapids, IA	16861
Cedar River downstream of Waterloo	05464000	Cedar River at Waterloo, IA	13328
Cedar River near Charles City	05457700	Cedar River at Charles City, IA	2730
Cedar River near Conesville	05465000	Cedar River near Conesville, IA	20168
Cedar River near Janesville	05458500	Cedar River at Janesville, IA	4302
Cedar River upstream of Cedar Rapids	05464500	Cedar River at Cedar Rapids, IA	16861
Des Moines River DMRWQN Site 1	05481300	Des Moines River near Stratford, IA	14121
Des Moines River downstream of Des Moines	05487500	Des Moines River near Runnells, IA	30186
Des Moines River near Runnells	05487500	Des Moines River near Runnells, IA	30186
Des Moines River upstream of Fort Dodge	05480500	Des Moines River at Fort Dodge, IA	10852
Des Moines River upstream of Ottumwa	05489500	Des Moines River at Ottumwa, IA	34639
Des Moines River, Des Moines	05482000	Des Moines River at 2nd Avenue, Des Moines	16175
English River at Riverside	05455500	English River at Kalona, IA	1487
Floyd River near Sioux City	06600500	Floyd River at James, IA	2295
Iowa River downstream of Iowa City	05454500	Iowa River at Iowa City, IA	8472
Iowa River downstream of Marshalltown	05451500	Iowa River at Marshalltown, IA	3968
Iowa River upstream of Marshalltown	05451500	Iowa River at Marshalltown, IA	3968
Maquoketa River near Maquoketa	05418500	Maquoketa River near Maquoketa, IA	4022
Middle River near Indianola	05486490	Middle River near Indianola, IA	1303
North Raccoon River near Sac City	05482300	North Raccoon River near Sac City, IA	1813
North Skunk River	05472500	North Skunk River near Sigourney, IA	1891
Ocheyedan River near Spencer	06605000	Ocheyedan River near Spencer, IA	1103
Old Mans Creek near Iowa City	05455100	Old Mans Creek near Iowa City, IA	521
Raccoon River at Van Meter	05484500	Raccoon River at Van Meter, IA	8912
Raccoon River upstream of Des Moines	05484500	Raccoon River at Van Meter, IA	8912
Raccoon River, Des Moines	05484900	Raccoon River at Fleur Drive at Des Moines, IA	9389
Shell Rock River at Shell Rock	05462000	Shell Rock River at Shell Rock, IA	4522
Sny Magill	05411400	Sny Magill Creek near McGregor	92.18
Soldier River near Pisgah	06608500	Soldier River at Pisgah, IA	1054
South Skunk River near Oskaloosa	05471500	South Skunk River near Oskaloosa, IA	4235
South Skunk River upstream of Ames	05470000	South Skunk River near Ames, IA	816
South Skunk River, above Ames WWTP	05471000	South Skunk River below Squaw Cr. near Ames, IA	1440
South Skunk River, below Ames WWTP	05471000	South Skunk River below Squaw Cr. near Ames, IA	1440
Squaw Creek W2	05471040	Squaw Creek near Colfax	47.03
Thompson Fork – Grand River at Davis City	06898000	Thompson River at Davis City, IA	1816
Turkey River near Garber	05412500	Turkey River at Garber, IA	4002
Walnut Creek T1	05487540	Walnut Creek near Prairie City	17.46
Walnut Creek T2	05487550	Walnut Creek near Vandalia	52.16
Wapsipinicon River at De Witt	05422000	Wapsipinicon River near De Witt, IA	6050
West Fork Des Moines River near Humboldt	05476750	Des Moines River at Humboldt, IA	5843
Winnebago River upstream of Mason City	05459500	Winnebago River at Mason City, IA	1362
Yellow River near Volney	05389000	Yellow River near Ion, IA	572

for sampled days within a given year multiplied by the number of days per year. Annualized yields were calculated as the annualized load divided by the drainage area.

Statistics examined to assess the regression model performance included the regression R^2 and mean square error (MSE), Lilliefors test for normal error distribution, Durbin-Watson's test for autocorrelation in regression residuals and observed and predicted partial loads. Positive regression residual autocorrelation would indicate potential for extended periods of over or under prediction of loads. For all measurement sites, sample size, percent of non-detects, correlation between load and discharge, sampling interval, drainage area, concentration average and variance, maximum observed and predicted concentration, and stream flashiness (Baker et al., 2004) were assessed to determine whether these factors affected model performance. The efficiency of the AMLE regression model applied to Eq. (1) was further evaluated using the Nash–Sutcliffe efficiency index (Nash and Sutcliffe, 1970) to compare predicted loads with observed loads. This efficiency index is calculated as

$$E = 1 - \frac{\sum_{i=1}^{n}(L_{o,i} - L_{p,i})^2}{\sum_{i=1}^{n}(L_{o,i} - \bar{L}_o)^2} \tag{2}$$

where $L_{p,i}$ is the predicted load for day i, $L_{o,i}$ is the observed load for day i, and $\bar{L}_o$ is the average observed load over the n days of record. Efficiency (E) values range between negative infinity and 1. An efficiency value close to one indicates a good correspondence between modeled and observed loads and a value less than zero indicates model performance worse than simply using the average load. This

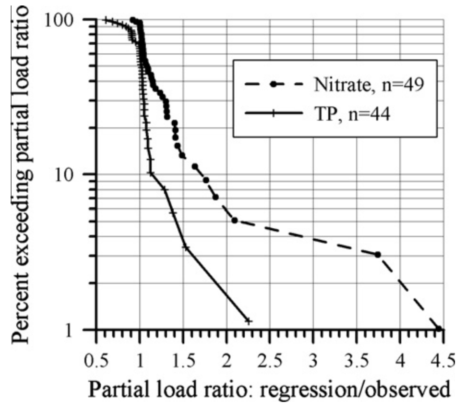

Fig. 2. Period of record prediction to observed partial load ratio frequency distribution.

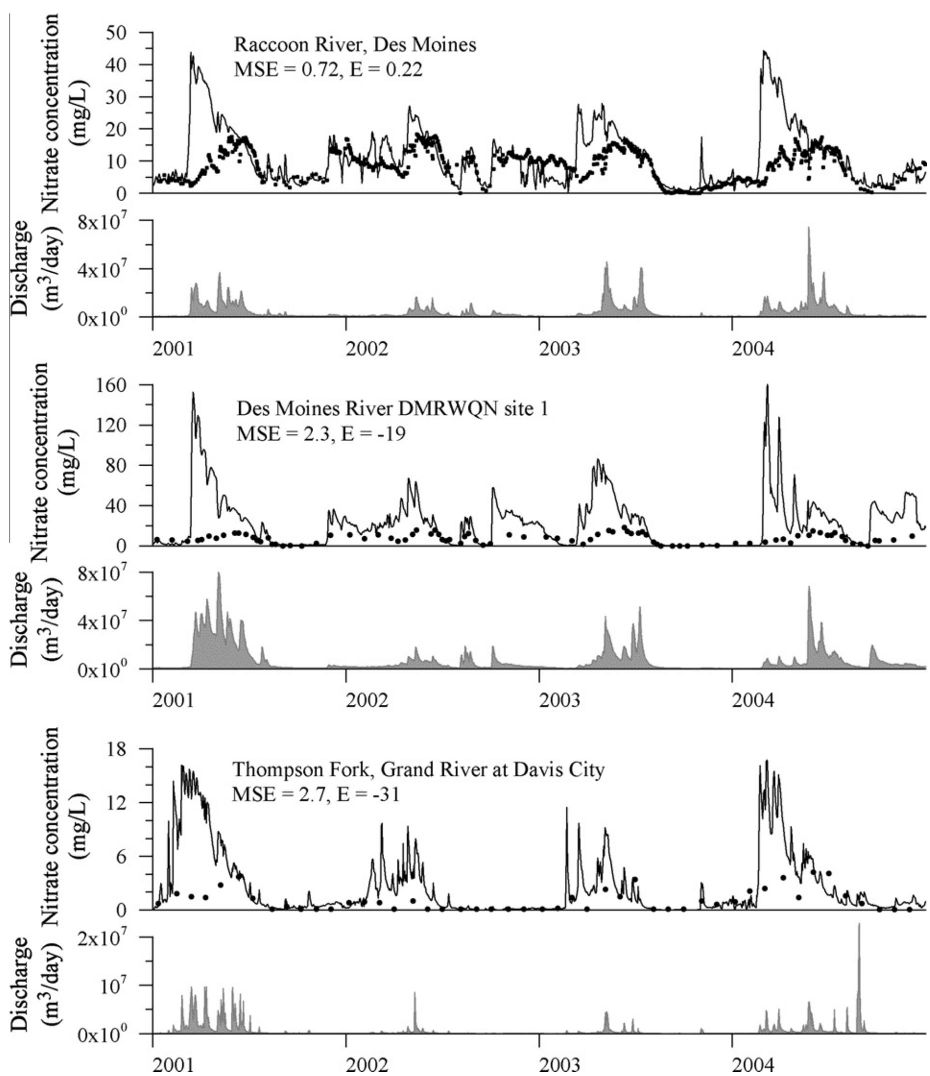

Fig. 3. Observed (solid circles) and predicted (solid lines) nitrate concentrations and river discharge for selected data sets.

efficiency calculation is similar to the regression model R^2 except that the efficiency is based on the load predictions after the back-transform from the logarithmic scale to the data scale, whereas the regression model R^2 is calculated from the regression output on the logarithmic scale. We determined efficiencies and partial loads using both the detection limit and zero for non-detect concentrations to obtain "observed" loads and found no difference between them large enough to alter any conclusion made in this work.

3. Results

3.1. Nitrate

The R^2 of the 49 nitrate data set regressions ranged from 0.800 to 0.975. However, E values ranged from −31 to 0.966, with nine cases (19%) having efficiencies less than zero. The predicted partial loads were within 10% of the observed partial loads in 51% of the cases, but overestimated the partial load by 25% or more in 33% of the cases, and overestimated the partial load by more than a factor of two in three of the 49 cases (Fig. 2). Two of 49 cases had ratios of predicted to observed partial loads less than one, but in neither case were predicted partial loads underestimated by more than 8%. The average overestimation of predicted partial loads was by a factor of 1.3 (1.18 excluding the two largest values).

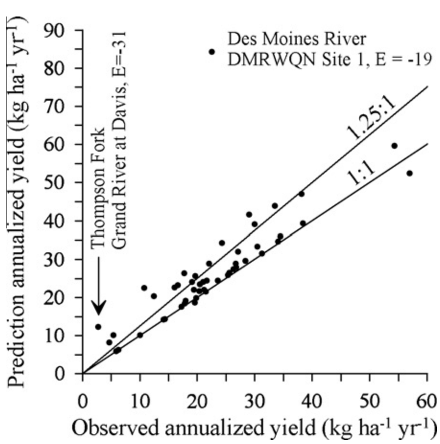

Fig. 4. Period of record regression prediction nitrate yield versus observed nitrate yield.

In some cases, the regression model (Eq. (1)) does not appear to adequately describe the way nitrate concentrations vary in streams (Fig. 3). Nitrate concentrations predicted by the regression model for some Iowa rivers often exceeded 30 mg/l while measured values rarely exceeded 10–20 mg/l. At the Des Moines River DMRWQN site 1, predicted nitrate concentrations approached an astonishingly high 160 mg/l (Fig. 3). Furthermore, these over-predictions often occur during high flow periods and thus have substantial adverse impact on load prediction.

Fig. 4 shows that nearly all the predicted nitrate yields exceed the observed yields suggesting systematic bias toward overestimation of nitrate loads with the AMLE regression Eq. (1). While the point labeled DMRWQN Site 1 in Fig. 4 appears to be a gross outlier, the Thomson Fork, Grand River point actually has a lower efficiency and is relatively further from the 1:1 line. Fig. 5 shows the ratio of predicted to observed partial loads calculated for all years with 12 or more samples plotted against drainage area, observed annualized load, and long term efficiency. The partial load ratio plotted against drainage area shows evidence of positive bias with substantial variability in annual and long-term load predictions for low to moderate drainage areas with no bias (ratios near one) for the five largest drainage areas (Fig. 5). A regression of annualized partial load ratio on the observed annualized load indicates that the nitrate partial load ratio does not depend on load magnitude but does show a positive bias of 26% ($p < 0.001$). Of the 410 annualized partial loads in Fig. 5, 28% overestimate the observed partial load by 25% or more while only 2% underestimate the observed partial load by 25% or more. Predicted to observed partial load ratios tend to increase and become more variable with decreasing efficiency, however, the relationship between these variables is weak (Fig. 5). While long term partial load ratios are near one for long term efficiencies above 0.85, the annual partial load ratios

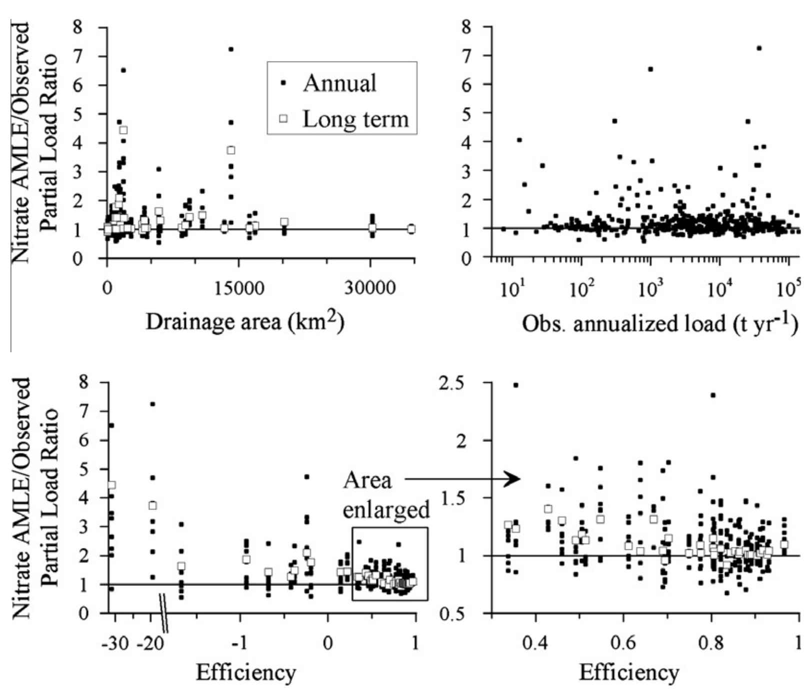

Fig. 5. Annualized nitrate AMLE to observed partial load ratio versus drainage area, annualized observed load, and long-term load prediction efficiency.

vary considerably across all efficiency levels. This is not simply a result of sparse data but rather is indicative of the performance of the model for simultaneous long term and short term load prediction. The Des Moines River, Des Moines site has an average of 13 samples per month, a long term efficiency of 0.88, a long term partial load ratio of 1.05, and annualized partial load ratios ranging from 0.88 to 1.47. The Cedar River at Cedar Rapids site has an average of 12.5 samples per month, a long term efficiency of 0.78, a long term partial load ratio of 1.09, and annualized partial load ratios ranging from 0.89 to 1.56. Accordingly, the long term efficiency and long term partial load ratio by themselves may be misleading and should be used in conjunction with relevant short term predicted versus observed partial load ratios as diagnostic measures of model performance.

Fig. 6 shows plots of nitrate load prediction efficiency and long term predicted to observed partial load ratios verses the regression MSE and R^2, the Durbin-Watson d and Lilliefors normality test p-value applied to the regression residuals, and the ratio of the maximum predicted concentration over the calibration data to the maximum observed concentration. These indicate that efficiency declines with increasing MSE, while the partial load ratio is near one only for MSE < 0.2 and increases with increasing MSE. Fig. 7 shows the regression log-scale and back-transformed load prediction versus observed values for a well performing and a poor performing AMLE regression. The well performing Black Hawk Creek at Waterloo log-scale regression model yields reasonable back-transformed data-scale load predictions while the poor performing Des Moines River DMRWQN Site 1 regression log-scale model performs poorly at low loads giving a large regression MSE resulting in back-transformed load predictions that grossly overestimate observed loads (Fig. 7). Although the efficiency tends to decline with decreasing regression R^2, there is considerable scatter and only R^2 values above 0.96 consistently show high efficiency and partial load ratios near one. Although the Durbin-Watson d indicates that positive autocorrelation is likely in 24 of the 49 cases (on the basis of the suggestion by Ott (1993) whereby $d < 1.5$ indicates positive autocorrelation) and Lilliefors test indicates non-normal residuals in 42 cases ($p < 0.05$), the model performance appears largely unrelated to the regression model assumptions of independence and residual normality. Load prediction efficiency tends to decrease and partial load ratio tends to increase as the maximum predicted to maximum observed concentration ratio increases, although these relationships are weak (Fig. 6).

It is useful to examine details associated with the poorest regression model performance, Thompson Fork – Grand River at Davis City and the Des Moines River DMRWQN site 1, with load prediction efficiencies of −31 and −19, respectively. The Thompson Fork has a watershed area of 1816 km^2, which is just below the median area of 1195 km^2 for sites used in this work, while the Des Moines River DMRWQN site 1 watershed area is 14,121 km^2, about 12 times the median. The Thompson Fork has considerably lower flow rates and greater flashiness than the Des Moines River site. The Thompson Fork site has much lower flow-weighted average (FWA) nitrate concentration but greater concentration coefficient of variation than the Des Moines River site. The regression residuals for both show marginal evidence for positive autocorrelation on the basis of the Durbin-Watson d statistic. The normal error distribution assumption for the Thompson Fork regression appears reasonable ($p > 0.1$) but the Des Moines River regression residuals appear to violate this assumption ($p < 0.01$).

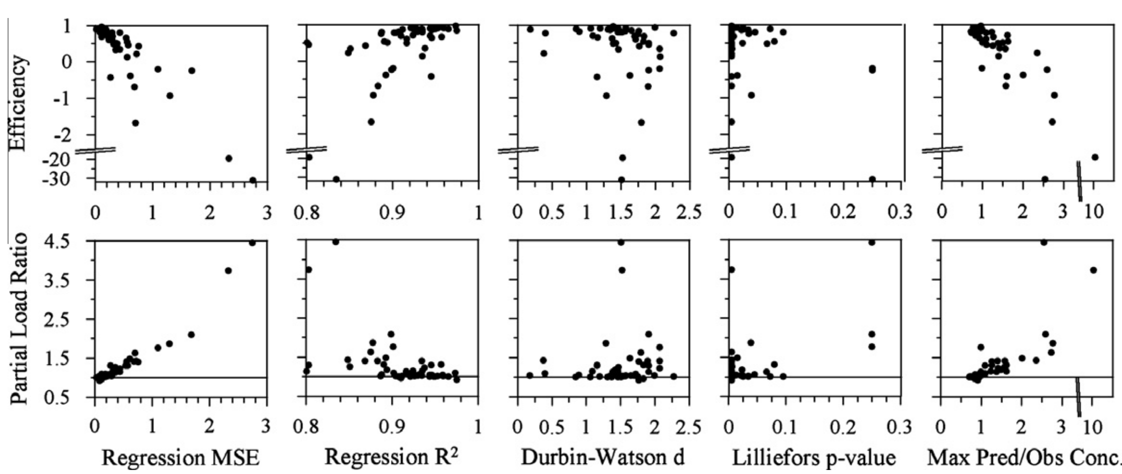

Fig. 6. Nitrate load prediction efficiency and long term AMLE to observed partial load ratio versus regression model diagnostics. The efficiency axis is split separating the bulk of the results from the two cases with $E = -19$ and -31.

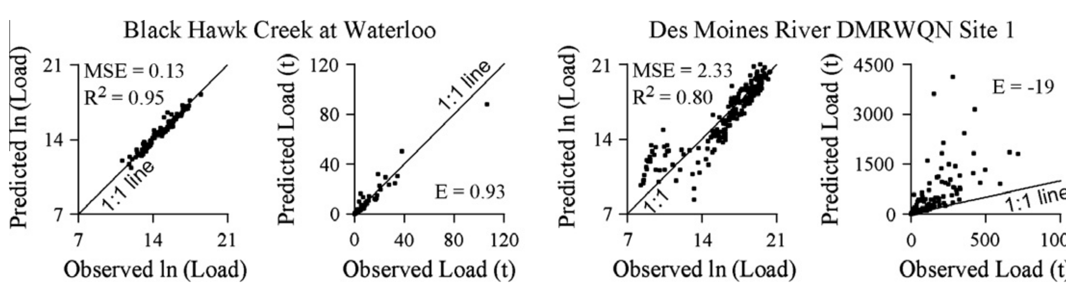

Fig. 7. Log-scale and data scale nitrate regression predicted versus observed load show good performance when the regression fit is good as at Black Hawk Creek and apparent back-transform bias when the regression fit is poor as at Des Moines River DMRWQN site 1.

The correlations between load and discharge for the Thompson Fork and Des Moines River data are 0.88 and 0.91, respectively, and regression R^2 values are 0.835 and 0.803, respectively. The regression residuals for both sites show heteroskedasticity as decreasing residual variability with respect to increasing predictor variable ln(*Q*). In both these cases, the effect of this heteroskedasticity is an overestimation of load at high discharge resulting from the back-transform Bessel function, which is a function of residual MSE (see Cohn et al., 1989). Both cases predict a concentration in excess of two times the largest observed concentration in the calibration data set. The regression diagnostics that best indicate poor load prediction performance for both these data sets are regression residual heteroskedasticity and MSE > 2.

Model performance varied considerably for sites in the Des Moines River above (DMRWQN site 1) and below (in Des Moines) Saylorville Reservoir. The Des Moines River is the largest river in central Iowa with drainage areas above and below the reservoir of 14,121 and 16,175 km^2, respectively, at the USGS gage stations utilized here. The reservoir extends about 27 km above the dam with an area of about 24 km^2 at full pool. The reservoir attenuates nitrate concentrations so that the upstream FWA-nitrate concentration is slightly greater than downstream, the upstream nitrate concentration variance is 83% greater than downstream, and the upstream data show 8% non-detects while the downstream data show no non-detects. Both of these stations have similar load-discharge correlations and stream flashiness index, the regression residuals show evidence of non-normality and the downstream station shows stronger positive autocorrelation in the regression residuals. Despite this, the regression on the upstream data performs very poorly ($E = -19$) while the downstream station regression performs relatively well ($E = 0.88$) for nitrate load prediction. The upstream regression residuals show strongly decreasing residual variance with increasing predictor variable ln(*Q*) while the downstream residuals show a similar, but much weaker, relationship. The regression diagnostics that best indicate potential poor nitrate model performance are the MSE (2.3 upstream and 0.20 downstream), heteroskedastic residual variance, and maximum predicted to maximum observed nitrate concentration ratio (10.1 upstream and 0.87 downstream).

Given variability in model performance to predict nitrate loads in Iowa streams, are there conditions under which the regression model performance would be anticipated to work better than others? Fig. 8 shows plots of nitrate load prediction efficiency versus selected sample and site characteristics. Sample size ranged from 95 to over 1800, and even a very large sample does not guarantee a good fit as indicated by the efficiency of $E = 0.22$ with 1583 observations (Fig. 8). The bulk of the sample sizes range between about 100 and 200, and yet these have efficiencies ranging from −31 to nearly 1. While the efficiency is generally low when there is a high percentage of non-detect data, the efficiency ranges from low to high with little or no non-detect data. The efficiency appears largely unrelated to the correlation between load (*L*) and discharge (*Q*), stream flashiness, and flow-weighted-average (FWA) concentration. The efficiency tends to decrease with increasing nitrate concentration coefficient of variation. Both high and low efficiencies are found with sample frequencies averaging between 2 and 30 days per sample. The nitrate load prediction efficiency is above 0.8 for the two largest drainage areas in this work, but otherwise the efficiency has no relation with drainage area.

While our focus has been on the general performance of regression Eq. (1) because of its apparent common use, other regression model forms may be useful if Eq. (1) performs poorly. Examination of the Raccoon River, Des Moines and Des Moines River DMRWQN Site 1 data show evidence for two nitrate concentration cycles per year. We replaced the $\sin(2\pi t)$ and $\cos(2\pi t)$ terms in Eq. (1) with $\sin(4\pi t)$ and $\cos(4\pi t)$ terms to allow two cycles per year and refit these Raccoon River and Des Moines River models using *LOADEST*. The AMLE did not fit well in either case, but the *LOADEST* LAD predictions did fit the Raccoon River data relatively well, yielding an efficiency of 0.75 (up from the AMLE efficiency of 0.22). A regression of the form given by Eq. (1) but with two cycles per year and fit to the untransformed concentrations on the Des Moines River DMRWQN Site 1 data yielded a load prediction efficiency of 0.85 (up from the AMLE efficiency of −19). These examples serve only to illustrate that relatively simple alternatives tailored to each specific site may be available that can significantly improve regression based load prediction in the event that the AMLE regression based on Eq. (1) does not perform well. Additionally, moving regressions based on, say, the most recent 5 years of data have been used to improve prediction and there is sufficient data in most of the data sets we examine to do this. However, the seven sites with longest records in this work include four with 10 year records and three with 11, 16 and 22 year records all have load prediction efficiencies ranging from 0.82 to 0.97 and would not likely see great improvement with a moving window regression, while the two sites with only 5 year records have efficiencies of −0.42

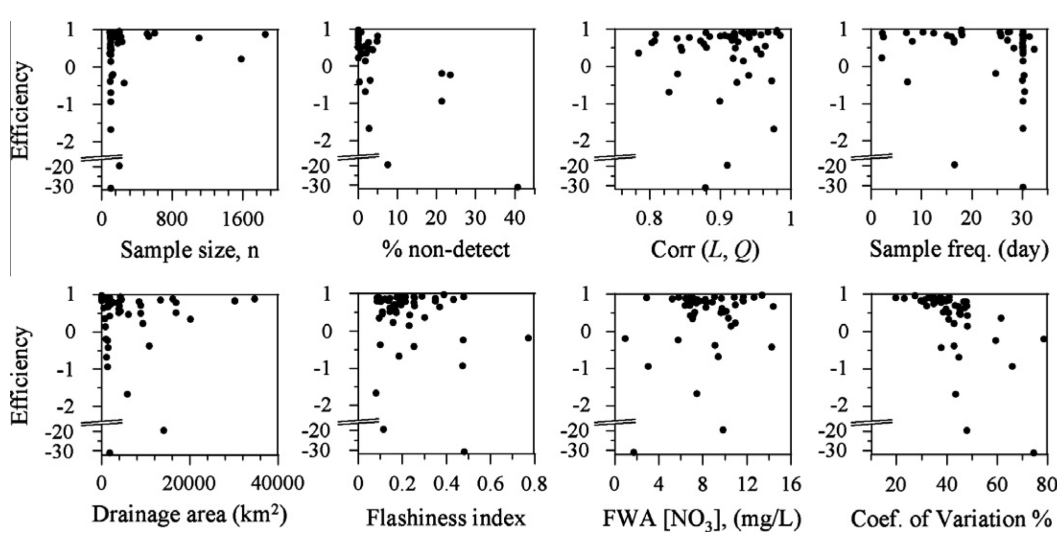

Fig. 8. Nitrate load prediction efficiency versus sample and site characteristics. The efficiency axis is split separating the bulk of the results from the two cases with $E = -19$ and −31.

and 0.67 and could not be improved with a 5-year moving regression approach.

3.2. Total phosphorus

TP regression R^2 values ranged from 0.526 to 0.950 for the 44 regressions. However, efficiency values ranged from −10 to 0.988, with five cases (11%) having efficiencies less than zero. The long-term regression predicted partial loads are within 10% of the observed loads in 73% of the cases, overestimate the long-term load by 25% or more about 8% of the cases, and overestimate the long-term load by more than a factor of two in one of the 44 cases (Fig. 2). Three (7%) period of record regression prediction partial loads underestimated observed TP loads by more than 25%. The efficiency does not show a clear pattern versus the regression R^2, and the site with poorest efficiency ($E = -10$) and greatest long term predicted to observed partial load ratio had an R^2 of 0.928 (Fig. 9). While the Durbin-Watson d indicates that positive autocorrelation is likely ($d < 1.5$) in only 3 of the 44 cases and Lilliefors test indicates non-normal residuals ($p < 0.05$) in 26 cases, the load prediction efficiency and long term predicted to observed partial load ratio appear largely unrelated to the regression model assumptions of residual normality and statistical independence (Fig. 9). The efficiency and long term partial load ratio tend to be consistently near one only for low regression MSE (Fig. 9) and the partial load ratio is consistently near one when the maximum predicted concentration divided by the maximum observed concentration is near one (Fig. 9).

The TP load prediction efficiency tends to decline with increasing percentage of non-detect data, decreasing correlation between load and discharge, and stream flashiness index, although these relationships are very weak (Fig. 10). The efficiency appears largely unrelated to sample size (n), sampling frequency (the majority are approximately bi-weekly or monthly), percent of non-detect data, stream flashiness, and flow-weighted-average (FWA) TP concentration, but does show a weak tendency to increase with increasing correlation between load and discharge and decrease with increasing concentration coefficient of variation. The prediction efficiency varies between 0.6 and 0.7 for the five largest drainage areas in this work, but otherwise varies widely for smaller drainage areas. The site with lowest efficiency, Cedar Creek near Bussey with $E = -10$, has reasonable sample size ($n = 132$) with 18% non-detect data, low drainage area, very high stream flashiness index, and high FWA TP concentration and variance, and although the

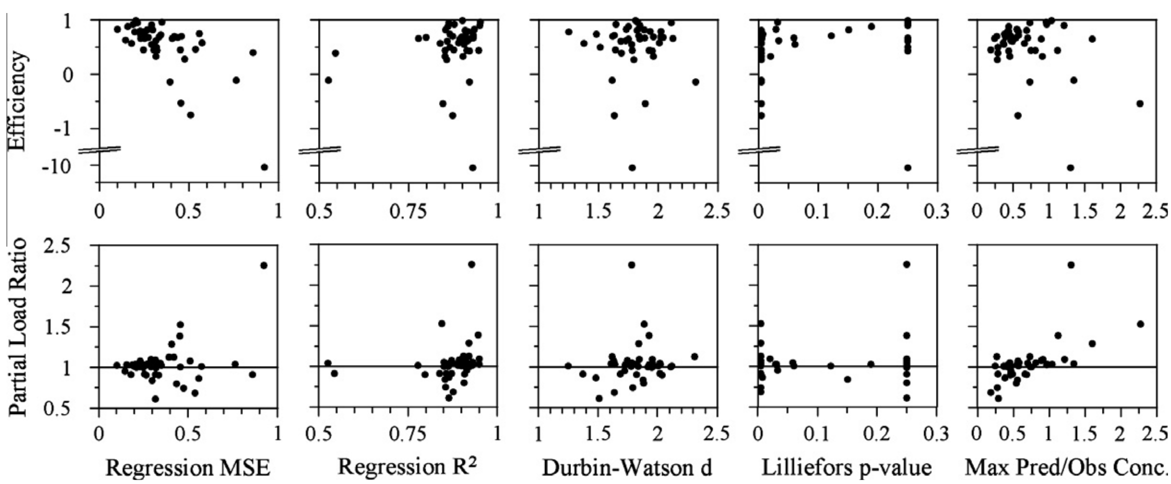

Fig. 9. TP load prediction efficiency and long term AMLE to observed partial load ratio versus regression model diagnostics. The efficiency axis is split separating the bulk of the results from the one case with $E = -10$.

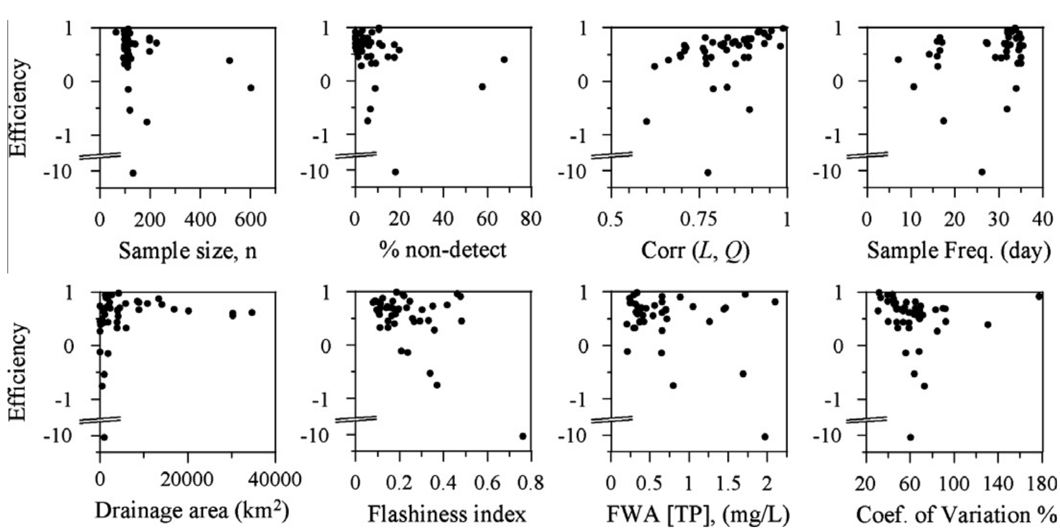

Fig. 10. TP load prediction efficiency versus sample and site characteristics. The efficiency axis is split separating the bulk of the results from the case with $E = -10$.

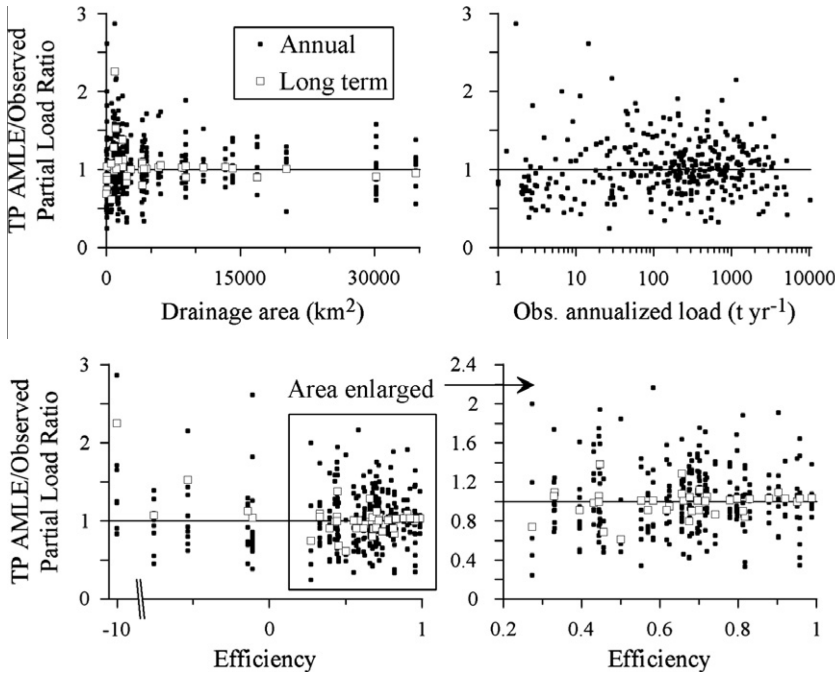

Fig. 11. Annual TP AMLE to observed partial load ratio versus drainage area, annualized observed load, and long-term load prediction efficiency.

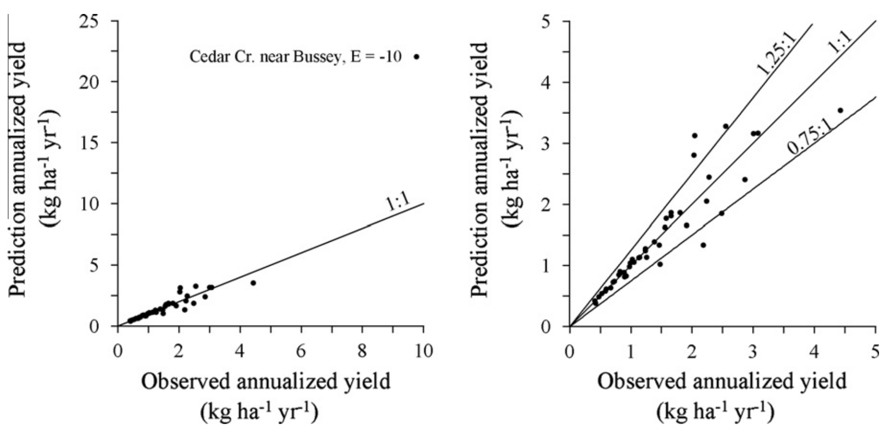

Fig. 12. Period of record TP prediction annualized yield versus observed annualized yield. The right panel plot axes are truncated to exclude the Cedar Cr. near Bussey value.

residuals appear to follow a normal distribution with no autocorrelation, the regression MSE is large. This combination of features yields poor load prediction efficiency and long term prediction to observed partial load ratio above two for this site.

Partial predicted versus observed load ratios calculated for years with 12 or more samples are plotted against drainage area, observed annualized load, and efficiency (Fig. 11). A regression of partial load ratio on observed annualized load indicates that the expected TP partial load ratio does not depend on the relative load magnitude and shows no statistically significant bias. Of the 338 partial annual loads in Fig. 11, 21% overestimate and 20% underestimate the observed partial load by 25% or more. The partial load ratio versus drainage area shows considerable annual load prediction variability at all drainage areas with (Fig. 11). While long term partial load ratios are all near one for efficiencies above 0.83, the annual partial load ratios vary considerably across all efficiency levels. As was observed with the nitrate regressions, the efficiency may be strongly affected by just one or several high loads among many poorly fit low loads resulting in highly variable annualized partial load ratios with a high long term efficiency value (see Fig. 11).

Fig. 12 shows period of record regression prediction TP yield versus observed yield. Apart from the gross outlier (Cedar Cr. near Bussey, Fig. 12), there is no evidence for systematic bias with TP load estimation. However, about 16% of the cases show greater than 25% long-term regression load prediction error. There are a few sites which have low load prediction efficiency and yet have reasonable long-term performance as judged by comparison of period of record predicted and observed partial loads. In these cases, some annual partial loads are over-predicted while other annual partial loads are under-predicted resulting in low efficiency with the appearance of reasonable long-term performance.

4. Discussion

The results of this study suggest that the multiple regression model given in Eq. (1) and available in the USGS *LOADEST* program shows systematic bias toward overestimation of nitrate loads in many Iowa rivers, shows no bias for TP loads, and shows considerable lack of precision in annual load prediction for both nitrate and TP. The terms in the regression model are insufficient to adequately characterize nitrate variation in some rivers resulting in a poor regression fit with large MSE and heteroskedastic residuals, in particular, decreasing residual variance with increasing ln(Q). In these cases, the regression and back-transform from the logarithmic scale under which the regression is performed to the data scale yields a positive load estimation bias (Fig. 7). Others have cautioned against using regression models for prediction of pollutant loads when the relationship between flow and concentration is not log-linear (Preston et al., 1989). Aulenbach and Hooper (2006) found that their regression models showed larger short term variations in storm sample residuals compared with baseflow samples. Likewise, there may be less error in regression model predicted loads or flow-weighted nitrate concentrations when comparisons are made during baseflow periods (Schilling, 2002).

While this study focused specifically on the regression model given in Eq. (1) as packaged in the *LOADEST* program, the issues raised in this study apply to all regression models used to predict non-point source dissolved groundwater constituents such as nitrate and chloride. When water quality parameters do not uniformly relate to discharge, regression model performance suffers and other load estimation schemes are recommended. Comparing a variety of load estimation procedures, Zamyadi et al. (2007) concluded simple linear interpolation between measured concentrations with a fixed sampling scheme yielded the most precise and accurate load estimations and Moatar and Meybeck (2005) concluded that both linear interpolation and a flow-weighted-average method often yielded the least biased load estimates. Li et al. (2006) compared the regression model of Cohn et al. (1992) to simple cokriging of suspended sediment and discharge and found that cokriging was preferred because it considered the temporal correlation of the data and reproduced the measured values.

The accuracy of the regression model varied considerably between streams with some long-term load predictions within 10% of the measured values while other predictions deviated more than 25% from measured values. Preston et al. (1989) observed that bias and precision of regression estimators were inconsistent among test cases they studied. In this study, the most consistent condition under which the regression model would be more accurate was very low regression MSE. The efficiency of nitrate load prediction was largely unrelated to many key site or sample specific variables and a case-by-case assessment of performance is recommended. Hence, there does not appear to be a reliable *a priori* evaluation to assess when the regression model might estimate nitrate or TP loads in Iowa streams reasonably well on the basis of variables we examined, except that sites with low concentration coefficient of variation tend to have high load prediction efficiency and sites with a high percentage of non-detect data tend to have low load prediction efficiency (Figs. 8 and 10). Furthermore, for nitrate, sites with large drainage areas performed well overall (Fig. 5). A low concentration coefficient of variation may result from either low concentration variability or an inadequate calibration sample, and thus may not be a reliable indicator of expected model performance if the calibration data do not well represent actual concentration variation.

Lack of uniformity in load prediction accuracy is particularly disconcerting when attempting to compare loading rates across regions. Nitrate loads or yields may be appear to be higher in some basins compared to others when, in actuality, the difference may be due to inability of the regression model to accurately predict nitrate loads. If the model errors were consistent across all streams, then an argument could be made that regression model estimates are comparable (though not necessary accurate). However, when regression model performance varies stream by stream with little consistency, it would not be possible to assign differences in loading patterns to actual watershed conditions without evaluating model performance on a case-by-case basis.

Results from this study have implications for calibrating watershed models, particularly for nitrate. Using a load estimation regression like Eq. (1) that may overestimate nitrate export from watersheds would steer model calibration toward incorrect adjustment of input factors. For example, perhaps model factors such as fertilizer applications could be adjusted downward to achieve model calibration if predicted nitrate yields were indeed 25% too high as shown in some Iowa cases. Furthermore, lack of consistent error in the model prediction would render comparisons of estimated loads among watershed regions difficult to assess. For example, national reconnaissance studies have used regression model results from *LOADEST* as the "observed loads" to calibrate the USGS SPARROW model (Alexander et al., 2008). If as shown for the Iowa sites, the *LOADEST* model results displayed inconsistent bias and poor precision, this could make such comparisons less reliable. Ullrich and Volk (2010) recommended the use of multiple load estimation methods to establish acceptable calibration for simulation models. We support this recommendation and would encourage further refinement of nitrate load estimation methods that would better capture the relation of nitrate and TP concentrations to discharge or other relevant factors.

5. Conclusions

The results of this study suggest that the multiple regression model given in Eq. (1) and available in the USGS *LOADEST* program shows an overestimation bias for annual and long term nitrate load prediction in some Iowa rivers, shows no bias for TP load prediction, and shows substantial lack of precision in annual and long-term load prediction for both nitrate and TP. Because of the back-transform associated with this regression model, the regression model R^2 could not be used to indicate adequate model performance at the data scale. Long term model performance was not related to regression residual distribution or autocorrelation. Good model performance was strongly related to low regression MSE and poor performance was associated with high MSE and residual heteroskedasticity. The efficiency index (Eq. (2)) and direct comparison of modeled and observed loads were found to be reasonable indicators of regression model performance.

Acknowledgements

We thank the Ames Waste Water Treatment Plant, the Des Moines Water Works, and the Cedar Rapids Water Department for their assistance with the procurement and sampling details of nitrate and phosphorus data utilized in this work. Data was also obtained from the Des Moines River Water Quality Network Database 1967–2009, which is supported by the Rock Island Corps of Engineers, with assistance from Donna S. Lutz.

References

Alexander, R.B., Smith, R.A., Schwarz, G.E., Boyer, E.W., Nolan, J.V., Brakebill, J.W., 2008. Differences in phosphorous and nitrogen delivery to the Gulf of Mexico from the Mississippi River Basin. Environ. Sci. Technol. 42, 822–830.

Aulenbach, B.T., 2006. Annual Dissolved Nitrate Plus Nitrite and Total Phosphorus Loads for the Susquehanna, St. Lawrence, Mississippi–Atchafalaya and Columbia River Basins for 1968 to 2004. Open File Report 2006-1087, US Geological Survey, Reston, VA.

Aulenbach, B.T., Hooper, R.P., 2006. The composite method: an improved method for stream-water solute load estimation. Hydrol. Process. 20, 3029–3047.

Baker, D.B., Richards, R.P., Loftus, T.T., Kramer, J.W., 2004. A New Flashiness Index: Characteristics and applications to Midwestern rivers and streams. J. Am. Water Resour. Assoc. 2004 (April), 503–522.
Cohn, T.A., DeLong, L.L., Gilroy, E.J., Hirsch, R.M., Wells, D.K., 1989. Estimating constituent loads. Water Resour. Res. 25 (5), 937–942.
Cohn, T.A., Caulder, D.L., Gilroy, E.J., Zynjuk, L.D., Summers, R.M., 1992. The validity of a simple statistical model for estimating fluvial constituent loads: an empirical study involving nutrient loads in chesapeake bay. Water Resour. Res. 28 (9), 2353–2363.
Eshleman, K.N., Kline, K.M., Morgan II, R.P., Castro, N.M., Negley, T.L., 2008. Contemporary trends in the acid–base status of two acid-sensitive streams in Western Maryland. Environ. Sci. Technol. 2008 (42), 56–61.
Gassman, P.W., Reyes, M., Green, C.H., Arnold, J.G., 2007. The soil and water assessment tool: historical development, applications, and future directions. Transactions ASABE 50, 1211–1250.
Gilroy, E.J., Hirsch, R.M., Cohn, T.A., 1990. Mean square error of regression-based constituent transport estimates. Water Resour. Res. 26 (9), 2069–2077.
Goolsby, D.A., Battaglin, W.A., 2001. Long-term changes in concentrations and flux of nitrogen in the Mississippi River Basin, USA. Hydrol. Process. 15, 1209–1226.
Goolsby, D.A., Battaglin, W.A., Lawrence, G.B., Artz, R.S., Aulenbach, B.T., Hooper, R.P., Keeney, D.R., Stensland, G.J., 1999. Flux and sources of nutrients in the Mississippi–Atchafalalya River Basin. White House Office of Science and Technology Policy Committee on Environmental and Natural Resources Hypoxia Work Group.
Goolsby, D.A., Battaglin, W.A., Aulenbach, B.T., Hooper, H.P., 2000. Nitrogen flux and sources in the Mississippi River Basin. Sci. Total Environ. 248, 75–86.
Guo, Y., Markus, M., Demissie, M., 2002. Uncertainty of nitrate-N load computations for agricultural watersheds. Water Resour. Res. 38 (10), 1185. doi:10.1029/2001WR001149.
Hooper, R.P., Aulenbach, B.T., Kelly, V.J., 2001. The national stream quality accounting network: a flux-based approach to monitoring the water quality of large rivers. Hydrol. Process. 15, 1089–1106.
Li, Z., Zhang, Y.-K., Schilling, K., Skopec, M., 2006. Cokriging estimation of suspended sediment loads. J. Hydrol. 327, 389–398.
Littlewood, I.G., Watts, C.D., Custance, J.M., 1998. Systematic application of United Kingdom river flow and quality databases for estimating annual river mass loads (1975–1994). Sci. Total Environ. 210–211, 21–40.
Maret, T.R., MacCoy, D.E., Carlisle, D.M., 2008. Long-term water quality and biological responses to multiple best management practices in rock creek, Idaho. J. Am. Water Resour. Assoc. 44 (5), 1248–1269.
Meybeck, M., 1982. Carbon, nitrogen and phosphorus transport by world rivers. Am. J. Sci. 282, 401–450.
Moatar, F., Meybeck, M., 2005. Compared performance of different algorithms for estimating annual nutrient loads discharged by the eutrophic River Loire. Hydrol. Process. 19, 429–444.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part 1: A discussion of principles. J. Hydrol. 10 (3), 282–290.
Ott, R.L., 1993. An Introduction to Statistical Methods and Data Analysis. Duxbury Press, Belmont, CA.
Preston, S.D., Bierman Jr., V.J., Sillman, S.E., 1989. An evaluation of methods for the estimation of tributary mass loads. Water Resour. Res. 25, 1379–1389.
Robertson, D.M., Roerish, E.D., 1999. Influence of various water quality sampling strategies on load estimates for small streams. Water Resour. Res. 35, 3747–3759.
Robertson, D.M., Schwarz, G.E., Saad, D.A., Alexander, R.B., 2009. Incorporating uncertainty into the ranking of SPARROW model nutrient yields from Mississippi/Atchafalya river basin watersheds. J. Am. Water Resour. Assoc. 45, 534–549.
Schilling, K.E., 2002. Chemical transport from paired agricultural and restored prairie watersheds. J. Environ. Qual. 31, 1184–1193.
Schilling, K.E., Zhang, Y.-K., 2004. Baseflow contribution to nitrate-nitrogen export from a large agricultural watershed USA. J. Hydrol. 295, 305–316.
Schilling, K.E., Hubbard, T., Luzier, J., Spooner, J. 2006. Walnut Creek watershed restoration and water quality monitoring project: Final Report. Iowa Geological Survey Technical Information Series 49, Iowa Department of Natural Resources, Iowa City, IA.
Ullrich, A., Volk, M., 2010. Influence of different nitrate-N monitoring strategies on load estimation as a base for model calibration and evaluation. Environ. Monit. Assess. doi:10.1007/s10661-009-1296-8.
US Environmental Protection Agency, 2000. Total Maximum Daily Load (TMDL) Program. Rep. EPA841-F-00-009, Washington, DC.
USGS, 2009a. USGS Open-File Report 2007-1080 – Streamflow and Nutrient Fluxes of the Mississippi–Atchafalaya River Basin and Subbasins for the Period of Record Through 2005, Methods Used to Estimate Nutrient Fluxes, <http://toxics.usgs.gov/pubs/of-2007-1080/methods.html>.
USGS, 2009b, Application of Spatially Referenced Regression Modeling for the Evaluation of Total Nitrogen Loading in the Chesapeake Bay Watershed, <http://md.water.usgs.gov/publications/wrir-99-4054/html/index.htm>.
USGS *LOADEST*, 2004. Load Estimator (*LOADEST*): A Fortran Program for Estimating Constituent Loads in Streams and Rivers. Techniques and Models Book 4, Chapter 5. US Geological Survey, Reston, VA.
Zamyadi, A., Gallichand, J., Duchemin, M., 2007. Comparison of methods for estimating sediment and nitrogen loads from a small agricultural watershed. Can. Biosyst. Eng. 49, 127–136.