Correcting for bias in freshwater total nitrogen concentrations obtained with a modified standard (NS4743) method

Norwegian Institute for Water Research

REPORT

Title		Serial number	Date
Correcting for bias in freshwater total nitrogen concentrations obtained with a modified standard (NS4743) method		7538-2020	05.10.2020
Author(s)		Topic group	Distribution
Jan-Erik Thrane, Heleen de Wit, Tomas A. Blakseth, Liv Bente Skancke and Øyvind Garmo.		Chemical analysis	Open
		Geographical area	Pages
		Norway	22

Client(s)	Client's reference
NIVA	
	Printed NIVA
	Project number

Concentrations of total nitrogen (TotN) in freshwater monitoring programs were unreliable for a limited period, probably because of inadequate sample digestion when using a modified version of method NS 4743 ("method A") for TotN analysis. The method resulted in underestimation of TotN, especially for humus-rich surface waters. Where re-analysis of samples with the original NS 4743 method ("method B") is not possible, TotN may be corrected using a regression model fitted to ca. 1000 freshwater samples with analytical results from both method A (likely resulting in biased TotN concentrations; $TotN_{old}$) and method B (resulting in presumably correct values; $TotN_{new}$). Here we present the dataset, the correction model, and a validation of the model using independent datasets. The regression model used $TotN_{old}$ and TOC and their interaction as predictors, and had an R^2 of 0.956 and a root mean squared error (RMSE) of 31.87 µg N/L. The model predicted $TotN_{new}$ well when tested on independent datasets from lakes and headwater streams. We recommend that the function is used to estimate TotN in long-term monitoring programs of lakes and headwater streams (1-3 order streams) where TotN was analysed with the modified standard method (method A), but not for rivers (>6-order streams) since the difference between TotN from the two methods was not significant for these water types.

Four keywords	Fire emneord
1. Total nitrogen	1. Total nitrogen
2. Analytical methods	2. Analysemetoder
3. Monitoring	3. Overvåking
4. Regression	4. Regresjon

This report is quality assured in accordance with NIVA's quality system and approved by:

Øyvind Garmo
Project Manager

Thorjørn Larssen
Research Manager

Correcting for bias in freshwater total nitrogen concentrations obtained with a modified standard (NS4743) method

Preface

In March 2020, NIVA researchers discovered suspicious values for total nitrogen concentration collected over a period spanning more than one year. Some of the samples were available for reanalysis, others were not. Available samples were reanalysed and data from these samples were used to find a correction function for the samples that could not be reanalysed. Here we present the results of our investigation into the matter and describe how erroneous values can be corrected. On behalf of the authors

Hamar, 25.09.2020

Øyvind Garmo

Table of contents

Summary

Evaluation of long-term changes in the carbon to nitrogen ratio (C:N ratio) in monitoring data showed that the C:N-ratio increased considerably in 2019, and expert judgement suggested that values were outside the natural range for C:N. The increase in C:N coincided with a change of method for total N (TotN) analysis to a modified version on NS 4743 ("method A"), which occurred in September 2018. Suspicions were therefore raised about the reliability of the analytical results, especially because the C:N ratios were most deviating at high total organic carbon (TOC) concentrations, which usually are associated with high organic N concentrations. It was hypothesized that the organic N content was underestimated, and a quality check of the method was initiated.

The investigation revealed that insufficient sample digestion due to the use of other chemicals and conditions during digestion, appeared to cause underestimated concentrations of TotN, especially for samples with a high fraction of organically bound N. The oxidation power of the oxidizing agent was found to be too low, and the time of digestion was also slightly too short according to NS 4743 standard. Therefore, the method for analysis of TotN was adapted to follow the NS 4743 standards ("method B", or the "standard method"), and all samples from a large regional lake survey in 2019 ("1000 lakes") were rerun using method B.

Because of the methodological issues, several monitoring programs had received presumably underestimated TotN concentrations between September 2018 and 2019. Most of those samples were, however, no longer available for re-analysis. To avoid either underestimated or void values for TotN for the stations and periods concerned, a function for correcting the underestimated TotN concentrations was needed.

Since the samples from the "1000 lakes"-program were analyzed with both methods for TotN, we used this dataset to fit a regression model for correcting the old (presumably biased) concentrations resulting from method A. The best regression model included the old TotN concentrations and TOC as predictors, as well as an interaction term between the two variables. This indicated that the difference between the old (method A) and new (method B) TotN concentrations was dependent on TOC. The model had an R^2 of 0.956 and a root mean squared error (RMSE) of 31.87 µg N/L. We validated the model on data from independent monitoring datasets containing data from lakes, headwater streams and larger rivers, and found that the model predicted the new TotN concentrations with adequate accuracy for lakes and headwater streams (1-3 order streams). We thus recommend that the function is used to correct the old TotN concentrations (resulting from method A) from lakes and headwater streams. TotN concentrations from rivers (>6-order streams), however, does not seem to need a correction, as the difference between TotN from the two methods was not significant. We speculate that this is due to differences in quality (e.g. lability or origin) of the organic matter between headwaters and rivers further down in catchments.

Since the modified standard method (method A) likely caused biased TotN concentrations, we recommend that freshwater monitoring programs use the NS 4743 standard method (method B) and not method A for analysis of total nitrogen. The recovery of native N present in headwater samples containing high concentrations of refractory natural organic matter must be documented before new or adjusted methods are implemented.

Sammendrag

En evaluering av langtidstrender i karbon til nitrogen-ratioen (C:N ratioen) fra overvåkingsdata viste at C:N-ratioen hadde økt betydelig i 2019, og ekspertvurderinger indikerte at verdiene var utenfor den naturlige variasjonsbredden for C:N. Økningen i C:N-ratio sammenfalt med en endring av metode for analyse av total nitrogen (TotN) til en modifisert versjon av NS 4743 («metode A»), som ble gjort i september 2018. Det ble derfor reist mistanke om påliteligheten til de analytiske resultatene, spesielt siden C:N-ratioene var mest avvikende ved høy konsentrasjon av totalt organisk karbon (TOC), som vanligvis er assosiert med høy konsentrasjon av organisk N. Det ble framsatt en hypotese om at innholdet av organisk N var underestimert, og en det ble igangsatt en kvalitetssjekk av metoden.

Undersøkelsen viste at utilstrekkelig oppslutning av prøvene, trolig på grunn bruk av andre kjemikalier og forhold under selve oppslutningen, så ut til å være årsaken til underestimerte konsentrasjoner av TotN, spesielt for prøver med høy andel organisk bundet N. Oksidasjonskraften til den oksiderende agensen ble funnet å være for lav, og oppslutningstiden var også litt for kort med hensyn til NS 4743-standarden. Metoden for analyse av TotN ble derfor tilpasset til NS 4743 standarden («metode B»), og alle prøvene fra et stort nasjonalt innsjøovervåkingsprogram («1000 sjøer»-prosjektet) ble re-analysert ved bruk av metode B.

På grunn av de metodiske utfordringene har det blitt rapportert antatt underestimerte konsentrasjoner av TotN for flere overvåkingsprogrammer mellom september 2018 og 2019. De fleste av disse prøvene er ikke lengre tilgjengelig for re-analyse, og for å unngå enten underestimerte eller tapte verdier for TotN fra de berørte stasjonene og periodene, var det behov for en funksjon for å korrigere de underestimerte verdiene.

Siden prøvene fra «1000 sjøer»-undersøkelsen ble analysert med begge metoder for TotN, benyttet vi dette datasettet til å tilpasse en regresjonsmodell for å korrigere de gamle (antatt underestimerte) konsentrasjonene som ble målt med metode A. Den beste regresjonsmodellen inkluderte de gamle TotN-konsentrasjonene og TOC som prediktorer, samt en interaksjon mellom de to variablene. Dette indikerer at forskjellen mellom de gamle (metode A) og nye (metode B) TotN-konsentrasjonene avhenger av TOC. Modellen hadde en R^2 på 0.956 og et gjennomsnittlig avvik (*root mean squared error;* RMSE) på 31,87 µg N/l. Vi validerte modellen på uavhengige datasett med data fra innsjøer, elver/bekker høyt oppe i vassdragene og større elver, og fant at modellen predikerte de nye TotN-konsentrasjonene med adekvat nøyaktighet for innsjøer og bekker/elver høyt oppe i vassdragene (1-3 ordens bekker/elver). TotN-konsentrasjoner fra større elver (>6-orden), ser derimot ikke ut til å trenge korreksjon, ettersom forskjellen mellom TotN fra de to metodene ikke var signifikant for denne typen elver. Vi spekulerer i om dette kan ha en sammenheng med forskjeller i kvaliteten på det organiske materialet (graden av labilitet eller opphav) mellom og bekker/elver høyt oppe og lengre ned i vassdragene.

Siden bruk av den modifiserte standardmetoden (metode A) trolig medførte underestimert TotN anbefaler vi at overvåkingsprogrammer for ferskvann bruker NS 4743-standarden (metode B) og ikke metode A for analyse av TotN. Gjenfinning (recovery) av ikke-tilsatt nitrogen i vannprøver med høy konsentrasjon av refraktorisk organisk materiale må dokumenteres før nye eller justerte metoder tas i bruk.

1   Introduction

During an evaluation of long-term change in the carbon to nitrogen ratio (C:N ratio) of dissolved organic matter (DOM), a considerable increase was found for the Langtjern catchment in 2019. Expert judgement suggested that the values were outside the normal range for C:N ratios found in natural waters, and it was suggested that organic nitrogen was underestimated. Organic nitrogen is estimated as the difference between total nitrogen and the sum of nitrate and ammonia. The change in C:N ratio coincided with a shift in method for routine analysis of total N, which was done in the autumn of 2018. However, this shift was accompanied by a quality check and it was initially considered that the changes were within tolerable limits of uncertainty for the analysis.

A regional survey for Norwegian lakes was undertaken in 2019 (Hindar et al. 2020). The water samples were among others analyzed for concentrations of total nitrogen. An exploratory data analysis indicated that the C:N ratios of DOM had increased profoundly since 1995, and again suspicions were raised with regard to the reliability of the analytical results, especially because the C:N ratios were most deviating at high TOC concentrations which usually associated with high organic N concentrations. Hypotheses put forward was that incomplete digestion of organic nitrogen to nitrate, prior to determination of nitrogen, gave too low measurements.

A quality check of the method (referred to in this report as "method A" or "the modified standard method") was decided on. The investigation of the analytical procedures revealed that insufficient digestion due to use of other chemicals and terms of condition during digestion indeed appeared to be the cause of the deviating TotN values. The result seemed to be a significant underestimation of TotN that was well outside the acceptable uncertainty interval. This was true especially for the samples with a high level of organic bound nitrogen. The oxidation power of the oxidizing agent was found to be too low, and the time of digestion was also slightly too short according to NS 4743 standard. Therefore, the method for analysis of total nitrogen was adapted to follow the NS 4743 standards (referred to in this report as "method B" or "the standard method"), and all samples for the regional lake survey were rerun using this method. All other freshwater monitoring programs also adopted this approach.

Several monitoring programs thus had received presumably biased TotN concentrations (hereby referred to as $TotN_{old}$) since the fall of 2018, until the new method (method B) again was established. TotN values from the method B is hereby referred to as $TotN_{new}$. Most of those samples were, however, no longer available for re-analysis. To avoid either underestimated or void values for TotN for the stations and periods concerned, a correction function was needed. This function should result in unbiased TotN estimates, with acceptable uncertainty.

Here we present a TotN correction function based on a comparison of $TotN_{old}$ (resulting from method A) and $TotN_{new}$ (resulting from method B) from the regional lake survey (ca 1000 lakes, predominantly headwaters). We also present data from other monitoring programs spanning small headwater catchments (1-3 order streams and lakes within the 1-3 order stream network) to large river basins (over 6-order streams). Each sample had measurements of $TotN_{old}$ and $TotN_{new}$, in addition to TOC and major cations and anions. The regional survey includes large gradients in TOC, TotN and organic N, and should therefore serve as a good basis for estimating a correction function. We present results from a model selection procedure, tests of model fit, and a validation of the final model on independent datasets. The dataset used for the estimations is presented in this report and will be available in a data repository.

2 Methods

2.1 Analytical methods

The NS 4743 standard method (Standard Norge, 1993; "method B") as well as the modified standard method ("method A") are based on oxidation of organic nitrogen and inorganic nitrogen species to nitrate with peroxydisulfate under alkaline conditions. Nitrate is subsequently reduced to nitrite by a copper-coated cadmium reductant at a pH of 8.0-8.5 in solution. In a reaction with sulfanilamide and N-(1-naphtyl) ethylenediamine, nitrite forms an azo-compound. The compound's absorption of light at 540 nm is measured. The most important differences between the NS 4743 standard (Method B) and the modified standard (Method A) is presented in the following.

1. **Oxidizing agent**
   The amount of potassium peroxydisulfate added to the sample is the same for the two methods, but the amount of sodium hydroxide is 2,5 times higher following method B than in method A. Method A instead uses boric acid in combination with sodium hydroxide. This may cause a difference in digestion effect for some samples.

2. **Temperature and time during autoclaving**
   Method B uses a temperature of 120 degrees Celsius over a period of 30 minutes. Method A applies a temperature of 121 degrees Celsius, but only for a period of 20 minutes. This difference in digestion time can potentially cause differences for the samples containing a high level of total organic carbon (TOC) and organic bound nitrogen (TON)

3. **Reagent for buffering of pH during analysis**
   This is a critical parameter to ensure sufficient reduction of the nitrate to nitrite in the sample when the sample passes through the cadmium reductant. The two methods use the same chemical and concentration for this buffer. However, method B includes a step checking the pH, while this is not done using method A.

4. **Control of cadmium reductant**
   The function of the cadmium reductant is essential for analysis of total nitrogen and nitrate. If the efficiency drops below a certain level (normally 90%), the reductant should be replaced. If the nitrate in the sample is not being sufficiently reduced to nitrite, this nitrogen will not be measured. In method B the effect of the cadmium reductant is controlled in the start of each series of analysis. For method A this was not done, and hence it cannot be excluded that some samples have been underestimated because of this.

2.2 Monitoring data

Monitoring data from the "1000-lake survey" (Hindar et al. 2020) was used to fit regression models and estimate parameters for correcting the presumably biased concentrations resulting from the modified standard method. All samples from the 1000-lakes survey was measured with both the modified standard method and the standard method for TotN. Samples were also analyzed for a large suite of other water chemical parameters. The 1000-lake survey was done the autumn of 2019 and is a representative lake survey for all of Norway. The lakes are mostly headwaters with little agricultural activity in their catchments, and hence relatively low TotN concentrations.

Monitoring data from three other monitoring programs ("Økofersk sur", "Reference rivers", and "The Norwegian river monitoring programme"; see **Table 1**) were used to validate the model fitted to the 1000-lakes dataset. A subset of samples from each program was analyzed both with the modified standard method and the standard method, allowing for independent tests of model fit. "Økofersk sur" is a program for monitoring temporal trends that started in 1986. It mainly comprises circa 100 acid-sensitive headwater lakes that are sampled once a year, and 6 small calibrated catchments ("Feltforskningsstasjoner") that are sampled weekly (or bi-weekly) (Garmo et al. 2019). The "Reference rivers" monitoring program comprises pristine (low human impact) rivers and streams with varying catchment sizes (Thrane et al. 2019). The Norwegian River Monitoring Programme monitors rivers with relatively large catchments and with a varying amount human impact (Gundersen et al. 2019).

Table 1. Overview of monitoring programs and data used for model fitting and model validation.

Monitoring program	Nr of sites	Period of presumably biased TotN samples	Period where TotN was measured with both methods
"1000-lakes", regional lake survey	Ca. 1000 lakes	2019 (ca. 1000 samples)	2019 (all samples were re-analyzed)
Økofersk-sur	Ca. 100 lakes, 2 rivers and 6 calibrated catchments (so-called "Feltforsknings-stasjoner")	Sept. 2018- Jan. 2019 (ca. 600 samples)	Sep 2019-Jan 2020 (n = 94 lake samples and 43 samples from calibrated catchments)
Reference rivers	38 rivers and streams	Sept. 2018- Jan. 2019 (ca. 550 samples)	Jan 2020 (n = 24)
Norwegian river monitoring programme	20 rivers	Sept. 2018- Jan. 2019 (ca. 350 samples)	Jan 2020 (n = 33)

3   Results and discussion

Exploratory data analysis of 1000-lake dataset.
Figure 1 shows that the TotN values resulting from the standard method (TotN$_{new}$) are systematically higher than the values resulting from the modified standard method (TotN$_{old}$) in the 1000-lake dataset. Exploratory data analysis (**Figure 2**) indicated that the bias between TotN from the two methods depends on TOC and in the C:N ratio of the dissolved organic matter (DOM). A low C:N ratio indicates N-rich organic matter. The intercept and slope values for the regressions of TotN$_{new}$ vs. TotN$_{old}$ were higher at low C:N ratio and high TOC, and we interpret this as an indication that the underestimation was most severe when most N was present in the form of organic N. This indicates that it was incomplete oxidation of organic matter that was causing the biased in Tot N concentrations measured with the modified standard method, i.e. points 1 and 2 in the chapter 2.1.
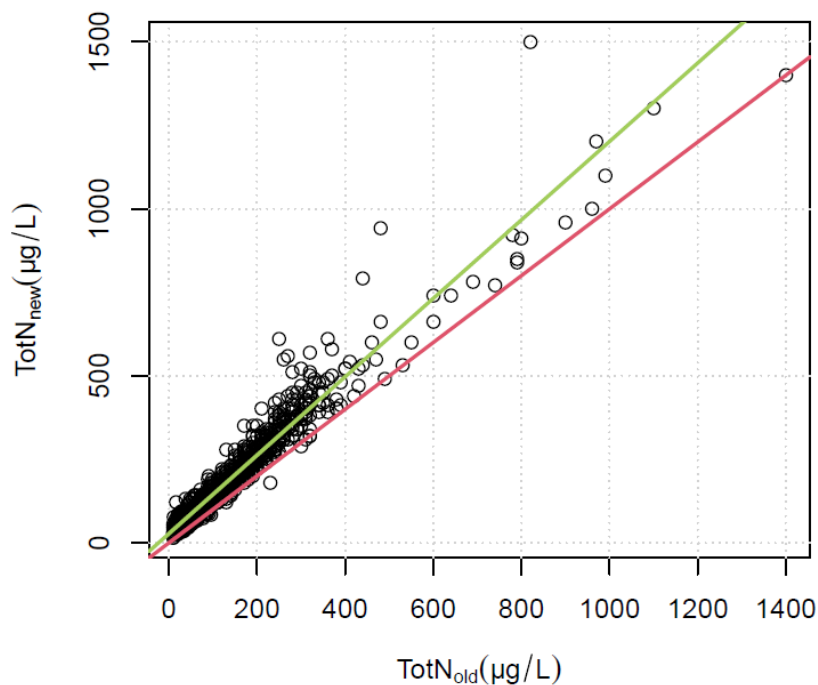

Figure 1. TotN values from the standard method (TotN$_{new}$; µg/L) plotted as a function of the TotN values from the modified standard method (TotN$_{old}$; µg/L) from the 1000-lakes dataset. The red line is the 1:1 line, the green line is the linear regression line (intercept = 27.7 µg/L; slope = 1.17; r2=0.936).
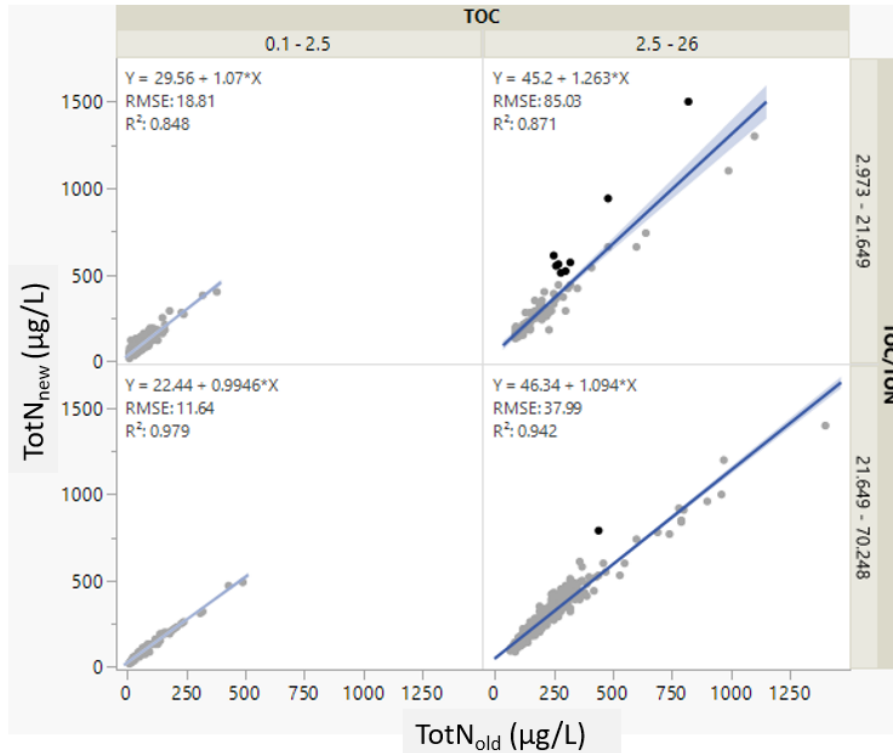
Figure 2. TotN from the standard method (TotN$_{new}$; μg/L) plotted against TotN values from the modified standard method (TotN$_{old}$; μg/L) for two classes of TOC concentrations (0.1-2.5 mg/L [below median TOC] and 2.5-26 mg/L [above median TOC]) and two classes of C:N ratios (below and above median C:N) . Results from linear regressions of TotN$_{new}$ vs TotN$_{old}$ are given in the top left of each panel.

3.1   Model selection

To find the best regression equation for correcting the TotN values from the modified standard method (TotN$_{old}$), we first did a model selection based on different regression models fitted to the 1000-lakes dataset. We tested various models where TotN from the standard method (TotN$_{new}$) was modelled as a function of TotN$_{old}$ with additive or interactive effects of potentially important predictor variables. These were total organic carbon (TOC) concentration, total organic nitrogen (TON = TotN$_{old}$ – [nitrate + ammonium]) concentration, total inorganic nitrogen (TIN = nitrate + ammonium) concentration, and TOC:TON ratio (i.e, C:N ratio of the organic matter). We tested models fitted to both untransformed and log-transformed data. To evaluate model fit, we used model R^2, root mean squared error (RMSE; the average difference between observed and predicted TotN$_{new}$) and AIC-value (Table 2). We also checked diagnostic plots (residual vs. fitted and observed vs. predicted; chapter 3.1.2) and did a cross validation of the model (chapter 3.1.3). The final choice of model was done after validating the set of best fitting models on four independent datasets, where TotN had been analysed with both methods (see chapter 2.2 and 3.2).

Table 2. Results from the different linear regression models fitted to predict $TotN_{new}$ in the 1000-lakes dataset. R^2 = fraction of variance explained by the model; RMSE = root mean squared error [the average difference between observed and predicted $TotN_{new}$]; AIC = Akaike Information Criterion, where a low value indicates a better fit. AIC is not comparable among untransformed and log-transformed models. The RMSE values for log-transformed models are back-calculated to untransformed scale and are therefore comparable among all models. All variables in all models were highly significant ($p < 10^{-16}$) except from the ones marked with an asterisk (p > 0.05). The best five models based on the above criteria are marked with numbers in parenthesis in column one.

Model #	Predictior variables & model structure	R^2	RMSE	AIC
m1	$TotN_{old}$	0.9419	36.65	9914.15
m2 (3/4)	$TotN_{old}$ + TON	0.9606	30.21	9533.77
m3 (1)	$TotN_{old}$ * TON	0.9622	29.59	9493.46
m4	$TotN_{old}$ + TOC	0.9559	31.95	9644.50
m5 (5)	$TotN_{old}$ * TOC	0.9564	31.78	9634.58
m6 (3/4)	$TotN_{old}$ + TIN	0.9606	30.21	9533.77
m7 (2)	$TotN_{old}$ * TIN	0.9608	30.12	9528.76
m8	$TotN_{old}$ + TOC:TON-ratio*	0.9419	36.66	9915.56
m9	$TotN_{old}$ * TOC:TON-ratio	0.9438	36.09	9885.77
m10	$\log(TotN_{old})$	0.9256	45.83	-364
m11	$\log(TotN_{old})$ + $\log$(TON)	0.9283	47.49	-408.41
m12	$\log(TotN_{old})$ * $\log$(TON)	0.9431	35.88	-633.92
m13	$\log(TotN_{old})$ + $\log$(TOC)	0.9347	49.68	-492
m14	$\log(TotN_{old})$ * $\log$(TOC)	0.9431	35.95	-627
m15	$\log(TotN_{old})$ + $\log$(TOC:TON-ratio)	0.9277	45.46	- 400.65
m16	$\log(TotN_{old})$ * $\log$(TOC:TON-raio)*	0.9278	45.41	-398.85
m17	$\log(TotN_{old})$ + $\log$(TIN)	0.9276	47.01	-389.83
m18	$\log(TotN_{old})$ * $\log$(TIN)	0.9338	43.4141	-475.81

3.1.1 Results from model selection

A model with only $TotN_{old}$ as predictor (i.e., a constant correction factor) explained 94% of the variation in $TotN_{new}$ with a root mean squared error (RMSE) of 36.65 µg/L (m1 in **Table 2**). Adding either TON, TOC or TIN (= [TotN – TON]) as predictors increased the R^2 and reduced the RMSE by 5-7 µg/L (models m2-m7 in **Table 2**). This indicates that the correction can be improved by including a variable representing the amount of organic material in the sample. The best model, judging only from the results in **Table 2**, was a model with $TOTN_{old}$ and TON as predictors (m3). The difference in fit between this model and models with TOC (m4 and m5) or TIN (m6 and m7), however, was minimal and probably of little practical importance. The similarity in fit between models using TOC and TON is due to the high correlation between TOC and TON (R^2 = 0.83 with both variables log-transformed). Including both TOC and TON in the model did therefore not improve model fit much. Adding an additive effect of TOC:TON-ratio (the C:N ratio of the organic matter) to the model if TON or TOC was already in, did not improve the fit significantly.

We validated the five best models from **Table 2** on independent datasets (see chapter 3.2), and found that the model with TOC and an interaction with $TotN_{old}$ (m5 in **Table 2**) had slightly better predictive ability than the models with TON and TIN as predictors. We therefore decided on m5 as the overall best model:

$$TotN_{new} = a + b1 \times TotN_{old} + b_2 \times TOC + b_3 \times (TotN_{old} \times TOC)$$

Estimated coefficients and their confidence intervals are shown in Table 3. Note that both $TotN_{new}$, $TotN_{old}$ and TOC were in units of µg/L when estimating coefficients. Therefore, *when using the model*

for prediction, all variables must be in units of µg/L. The model had an R^2 of 0,9564 and an RMSE of 31.78 on 983 degrees of freedom. An RMSE on 31,78 means that the average deviation between observed and predicted TotN$_{new}$ was 31,78 µg/L.

Table 3: Estimated regression coefficients, standard error (SE), 95% confidence interval, and *p*-value for the coefficients.

Coefficient	Estimate	SE	2.5 %	97.5%	*p*-value
a	26.58	2.008	22.644	30.525	$< 10^{-16}$
b_1	1.028	0.0142	1.000	1.056	$< 10^{-16}$
b_2	0.004046	0.000503	0.00306	0.00503	$< 10^{-16}$
b_3	0.00000656	0.00000190	0.00000284	0.0000103	$< 10^{-3}$

3.1.2 Model fit on the 1000-lakes dataset

Predicted and observed concentrations were close to unity (values fell along the 1:1 line in **Figure 3**A) and there were no obvious deviations from linearity (**Figure 3**A and B). The model thus seemed to precisely predict TotN$_{new}$ along the whole gradient of TotN concentrations. Predicted and observed TotN$_{new}$ also showed a linear relationship when plotted for different classes of TOC and C:N-ratio (Figure 4). We therefore generally expect low bias in the predictions for water samples with different levels of these parameters. There was, however, more variance in samples with high TOC and low C:N, especially in samples with TotN > 500 µg/L, where the model also tended to underestimate TotN slightly (Figure 4B). This was also evident from the residual plots (Figure 5). Overall, this indicates a reasonably good fit to the data, but with higher uncertainty for samples with the combination of high TOC and low C:N ratio.

Absolute prediction error (the absolute value of observed minus predicted TotN$_{new}$) increased with TotN concentration (Figure 6A). The relative prediction error (in % of TotN concentration), however, was highest for samples with low TotN (Figure 6B). The relative prediction error was nevertheless quite low. When dividing the TotN concentrations into classes (0-50 µg/L; 50-100 µg/L; 100-200 µg/L; 200-400 µg/L; 400-800 µg/L and 800-1500 µg/L), median relative prediction error was < 10% for all classes except 0-50 µg/L and 50-100 µg/L, where the it was 27% and 13%, respectively (Figure 6B).
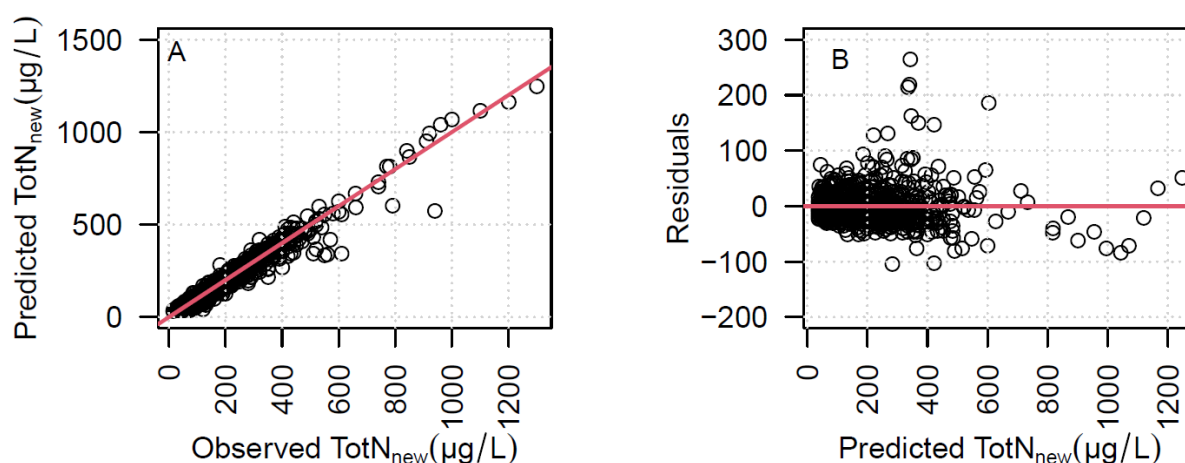

Figure 3: A) Predicted TotN$_{new}$ (µg/L) as a function of observed TotN$_{new}$ (µg/L) in the 1000-lakes dataset. The red line shows the 1:1 line. B) Model residuals as a function of predicted TotN$_{new}$.
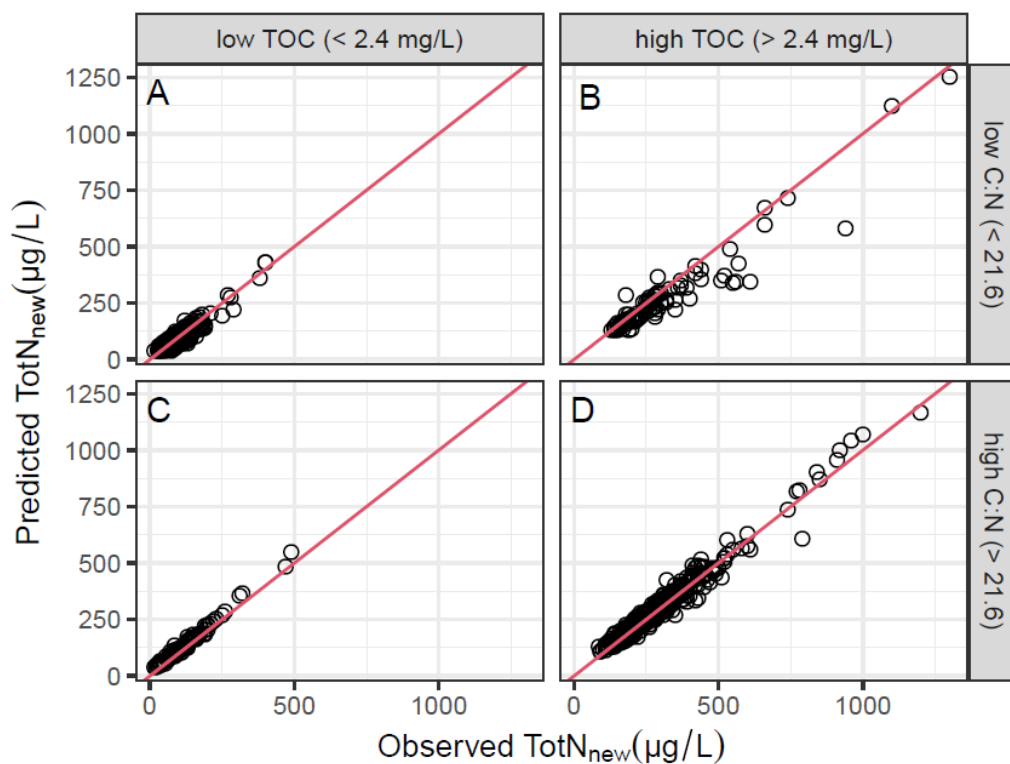
Figure 4. Predicted TotN$_{new}$ (µg/L) as a function of observed TotN$_{new}$ (µg/L) plotted for different combinations of "low" (< median) and "high" (> median) levels of TOC and TOC:TON ratio. Red diagonal line is the 1:1 line.
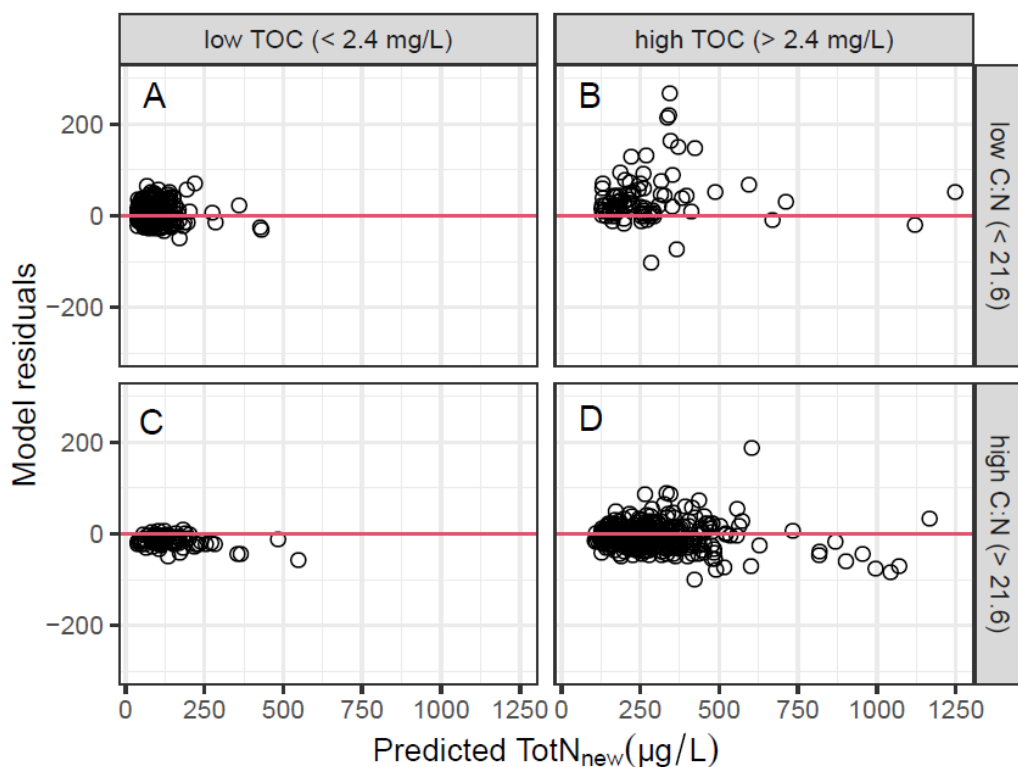

Figure 5. Model residuals as a function of predicted TotN$_{new}$ (µg/L) plotted for different combinations of "low" (< median) and "high" (> median) levels of TOC and TOC:TON ratio.
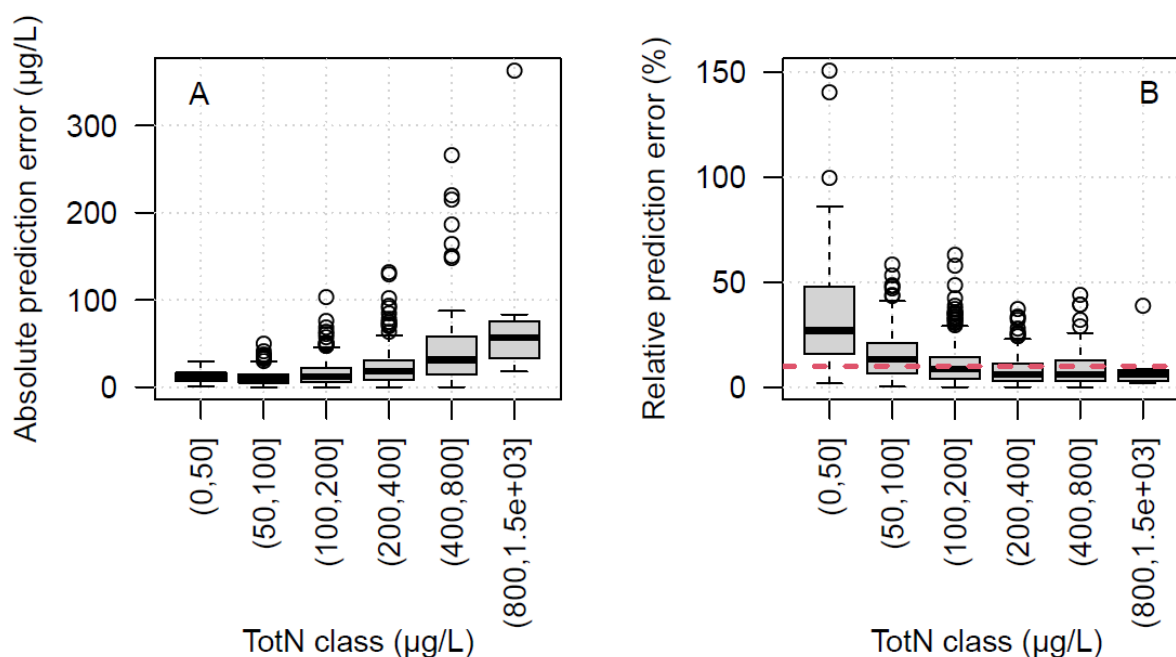
Figure 6. Absolute (A) and relative (B) prediction error for TotN$_{new}$. Absolute prediction error was calculated as the absolute value of [observed TotN$_{new}$ – precicted TotN$_{new}$], while relative prediction error was calculated as absolute prediction error divided by the observed TotN$_{new}$ concentration. TotN$_{new}$ was divided into classes along the x-axis to show the distribution of prediction errors for different concentration classes. The horizontal red line in B) show the 10 % prediction error.

3.1.3 Cross validation of the model

To test how well the model predicted "unseen data", we ran a Monte-Carlo type cross validation using the 1000-lakes dataset. In this procedure, we iteratively created "training datasets" by randomly sampling 100 lakes from the full dataset (n = 983). For each iteration, the regression model (m5: TotN$_{new}$ ~ TotN$_{old}$ * TOC) was fitted to the training dataset. The estimated coefficients were then used to predict TotN$_{new}$ in a unique "test dataset" created by randomly sampling 100 other lakes (including none of the lakes from the training dataset) from the full dataset. For every iteration, we calculated the RMSE and R^2 as measures of fit the between observed and predicted TotN$_{new}$.

Results from the cross validation showed that the model predicted TotN$_{new}$ with high accuracy also when tested on "unseen" data points (Table 4). Median RMSE and R^2 was 32.39 µg/L and 0.9560, respectively.

Table 4. Summary of results from 1000 iterations of a cross validation of the regression model. RMSE = Root mean square error = mean difference between observed and predicted TotN$_{new}$ when tested on the "unseen" data in the test dataset. R^2 = the fraction of variance in observed TotN$_{new}$ explained by the predicted TotN$_{new}$.

	Min	Q1	Median	Q3	Max
RMSE	17.83	26.86	32.39	38.8	98.29
R^2	0.8722	0.9388	0.9560	0.9691	0.9903

3.2   Test of model fit on independent datasets

An independent validation of the model (m5 in **Table 2**) was done by testing it on data from three different monitoring projects (see **Table 1Table 3**). In these projects, several samples had been analysed with both methods so that data on both $TotN_{old}$ and $TotN_{new}$ was available. The model was tested by predicting concentrations of $TotN_{new}$ (µg/L) using the estimated model parameters (**Table 3**) and concentrations of $TotN_{old}$ (µg/L) and TOC (µg/L), and evaluating the predicted concentrations against the observed concentrations.

3.2.1 Difference between the $TotN_{new}$ and $TotN_{old}$ in in the independent datasets

As in the 1000-lakes dataset, the independent datasets generally showed that TotN was underestimated by the modified standard method relative to the standard method. For all the data combined (n= 194), $TotN_{new}$ was higher than $TotN_{old}$ in in 77 % of the samples (Figure 7A & B). The difference was highly significant ($p < 10^{-16}$ in a one sample t-test of log($TotN_{new}$: $TotN_{old}$) = 0).

Notably, the degree of underestimation by the modified standard method differed between projects, likely because the samples from the different projects varied from headwater lakes and streams (Økofersk sur) to rivers relatively far down in catchments (Norwegian river monitoring programme and Reference rivers). Underestimation by the modified method was most prominent in the samples from the Økofersk sur lakes (Figure 7C & D) and Feltforskning stations (six streams included in the Økofersk sur program; Figure 7E & F), with an average underestimation of 32% and 27%, respectively ($p < 10^{-16}$ in t-test of log($TotN_{new}$: $TotN_{old}$) = 0). In the data from the Norwegian river monitoring programme (Figure 7G & H) and Reference rivers (Figure 7I & J), there was no significant difference between the TotN values from the two methods ($p > 0.05$ in t-test of log($TotN_{new}$: $TotN_{old}$) = 0).

The difference in the effect of method may be related to variation in the fraction of organic N or C:N ratio between the samples from the different projects. For example, the samples from the Reference rivers and the Norwegian river monitoring programme had significantly lower fraction of organic N and lower C:N ratio than the samples from the two other projects (Figure 8).
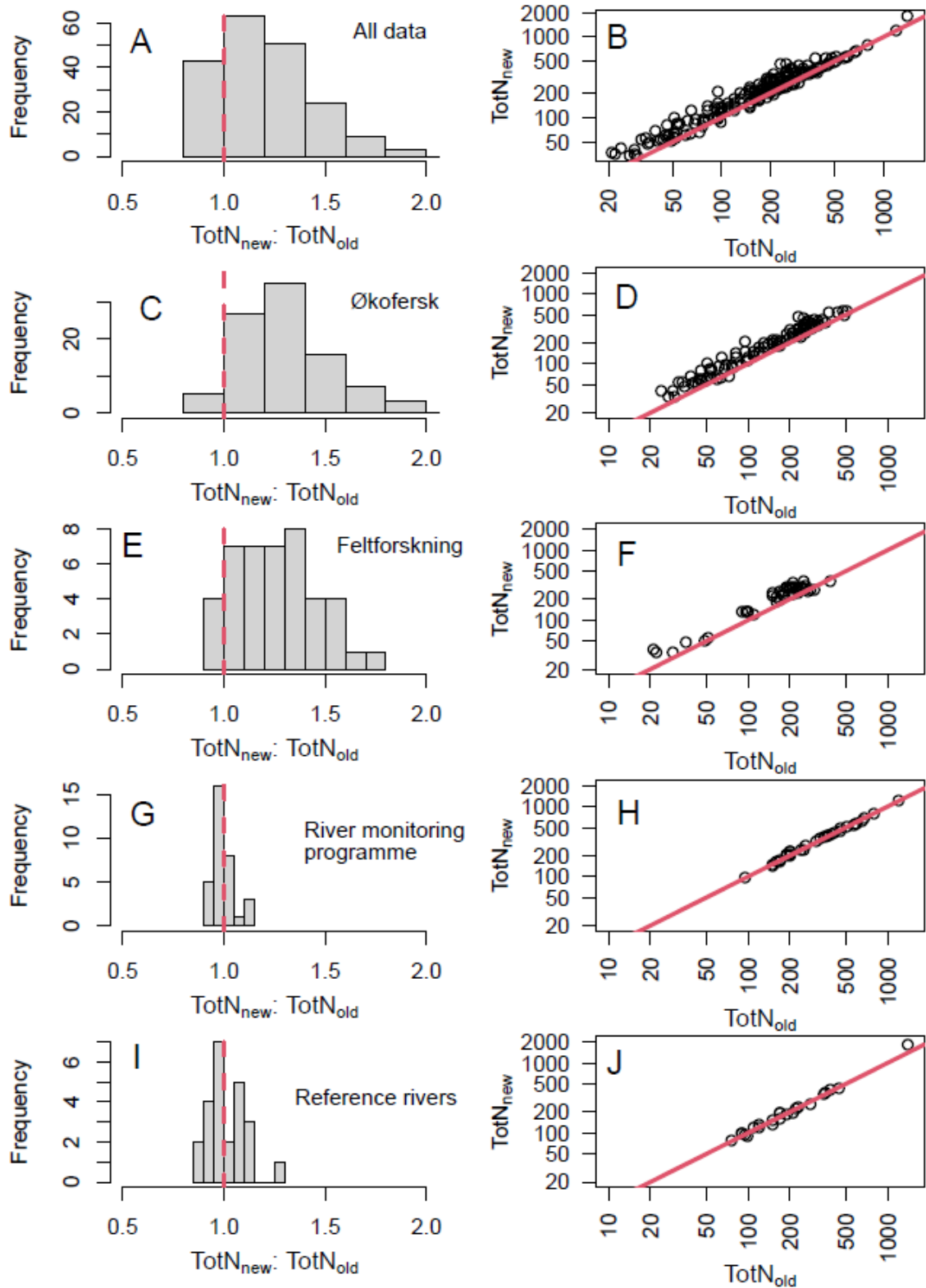
Figure 7. Left column (A-E): Ratio of $TotN_{new}$: $TotN_{old}$ in the independent datasets where TotN had been analysed by both methods. "Økofersk" (C & D) are lakes from the Økofersk sur program, while "Feltforskning" (E & F) are six headwater streams included in the same monitoring project. Vertical lines show the 1:1 ratio. Right column (F-J): Plots of $TotN_{new}$ vs. $TotN_{old}$. Diagonal red line equals the 1:1 line. Note logarithmic axes in the scatterplots.
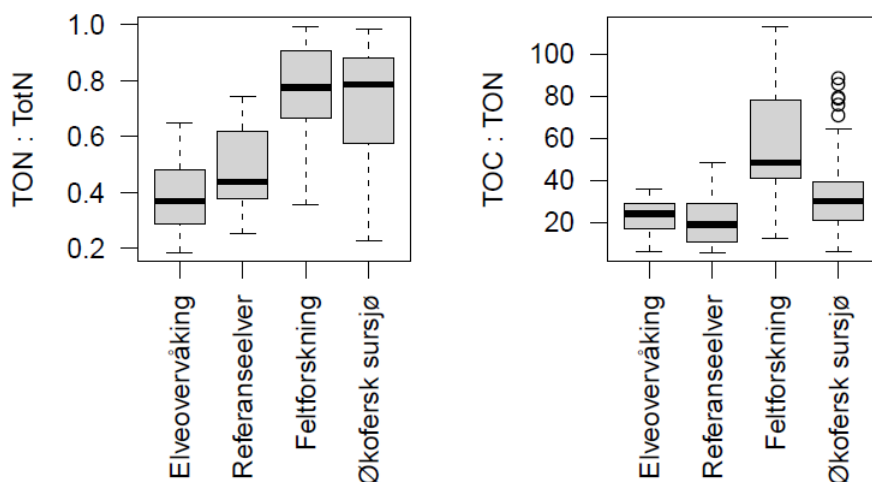
Figure 8. A) Fraction of total organic nitrogen (TON) relative to total N in the samples from the four different monitoring programs. B) Ratio of TOC to TON (C:N ratio of the organic matter).

3.2.2 Performance of the regression model on the independent dataset

The observed and the predicted $TotN_{new}$ was highly correlated ($R^2 = 0.958$) and with relatively low error (RMSE = 33 µg/L) in the pooled dataset from all four monitoring programs (Figure 9A-B). The slope and intercept were estimated as 1.06 (95% CI: 1.03-1.09) and 13.3 (95% CI: 4.2-22.4), respectively, indicating that the model overestimated TotN slightly when considering data from all four monitoring programs pooled. The model fit, however, differed between projects.

Considering the data from lakes in Økofersk sur (headwater lakes with negligible local sources of pollution), the model estimated $TotN_{new}$ with high precision (Figure 9C-D; $R^2 = 0.973$, RMSE = 23 µg/L; one outlier removed). The slope of observed vs. predicted $TotN_{new}$ was not significantly different from 1 (95% CI: 0.94 - 1.01) with an intercept of 16.0 (95% CI: 7.6 – 24.4). This indicates that the model works well for correcting TotN values from lakes with similar characteristics as the ones in the Økofersk program.

Observed and predicted TotN from the Feltforskning project (small catchments with negligible local sources of pollution) had a lower correlation ($R^2 = 0,83$, RMSE = 46.8 µg/L) and a slope of 1.13 (95% CI: 0.97-1.29). The deviation from 1:1 was, however, mainly caused by to six data points sampled between 16/9 and 29/9-2019 (marked with red circles in Figure 9F). Removing these six points from the dataset, the R^2 increased to 0.969 and RMSE reduced to 16.3 µg/L. The slope then also changed closer to unity (0.97, 95% CI: 0.90-1.02), indicating a reasonably good fit to these types of waters.

The TotN values from the Norwegian river monitoring programme (rivers with varying human impact) and Reference rivers (rivers with low human impact) did not differ between the two methods to begin with (**Figure 7**G-J). Applying the correction model on the data from these two programs thus caused a slight overestimation of TotN compared to the observed concentrations (Figure 9G-J). Due to the agreement between $TotN_{old}$ and $TotN_{new}$ in these programs to begin with, no correction seems to be needed.
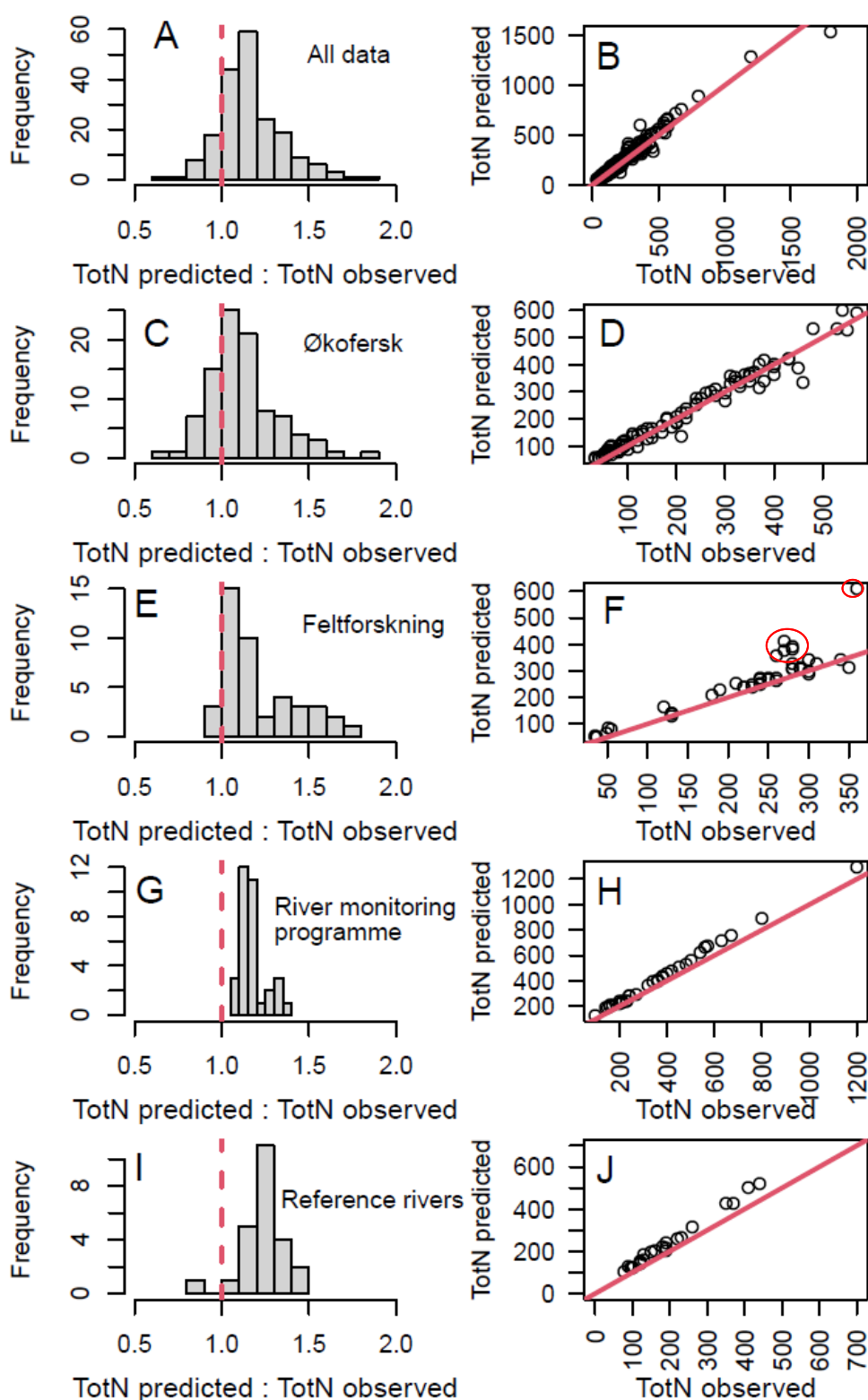
Figure 9. Left panels: Ratio of predicted TotN$_{new}$ and observed TotN$_{new}$ in the independent datasets where TotN had been analysed with both methods. "Økofersk" (C & D) are lakes from the Økofersk sur program, while "Feltforskning" (E & F) are six headwater streams included in the same monitoring project. Vertical lines show the 1:1 ratio. Right panels: Plots of predicted TotN$_{new}$ as a function of observer TotN$_{new}$. Red line shows the 1:1 line.

4   Conclusion and recommendations

The aim of this report was to present a correction function for TotN, as the use of a modified standard method (method A) likely led to underestimation of TotN for several long-term monitoring programs that was well outside the acceptable uncertainty interval. To avoid underestimated and void values for TotN for the stations and periods concerned, a correction function was needed. This function should result in unbiased TotN estimates, with acceptable uncertainty.

This report shows that relatively unbiased estimates for $TotN_{new}$ can be found using $TotN_{old}$, TOC and its interaction, based on a regression model fitted to the regional 1000-lake survey that had values for $TotN_{old}$ and $TotN_{new}$. The model had an R^2 of 0.9564 and an RMSE of 31,78 µg/L, and fitted the data well based on diagnostic plots. The function was tested and validated on independent datasets that also had values for TotN from both methods, and this resulted in acceptable values of $TotN_{new}$ for headwater streams and lakes. For rivers, however, $TotN_{old}$ and $TotN_{new}$ did not show a significant difference to begin with. We speculate that this may be because aquatic processing has resulted in either i) removal of the fraction of organic N that is difficult to digest, and/or, ii) made organic N more easily available and thus easier to digest. The origin of the organic matter in the rivers may also play a role (perhaps a higher proportion of the DOM originates from lowland soils that contain more labile organic matter and nitrogen). However, one of the river programs was aimed at natural systems without agricultural influences in the catchment and these rivers had the same relationships between old and new TotN as the rivers that also drain developed and agriculturally impacted catchments.

The methodological bias was probably caused by incomplete oxidation of organic nitrogen species to nitrate under the conditions applied in the modified standard method (method A). We recommend that freshwater monitoring programs use the NS 4743 standard method (method B) and not the modified method for analysis of total nitrogen. The recovery of native N present in headwater samples containing high concentrations of refractory natural organic matter must be documented before new or adjusted methods are implemented. We also recommend that TotN values obtained from the modified standard method are replaced with corrected values, especially samples from headwater lakes and streams. Our analysis indicates, however, that samples from the Norwegian river monitoring programme and Reference rivers does not need to be corrected, as the concentrations obtained by the two methods were not significantly different. Corrected values should flagged with a reference to this report in the publicly available databases.

The recommended model is:

$$TotN_{new} = a + b1 \times TotN_{old} + b_2 \times TOC + b_3 \times (TotN_{old} \times TOC)$$

Where estimated coefficients are: a: 26.58; b_1: 1.028; b_2: 0.004046; b_3: 0.00000656

When using the model for prediction, all variables must be in units of µg/L.

5 References

Garmo, Ø.A., Skancke, L.B., 2019. Overvåking av langtransportert forurenset luft og nedbør. Årsrapport – vannkjemiske effekter 2018. Miljødirektoratet No. M1466/2019.

Gundersen, C.B., Kaste, Ø., Sample, J.E., Braaten, H.F.V., Selvik, J.R., Hjermann, D.Ø., Norling, M.D., Guerrero, J.-L., 2019. The Norwegian river monitoring programme – water quality status and trends in 2018. NIVA-rapport 7441.

Hindar A., Garmo Ø., Austnes K., Sample J. E. 2020. Nasjonal innshøundersøkelse 2019. NIVA-rapport 7530.

Standard Norway 1993. Norwegian Standard NS 4743. Water analysis - Determination of total nitrogen after oxidation by peroxodisulphate.

Thrane, J.E., Persson, J., Røst Kile, M., Bækkelie, K.A., Myrvold, K.M., Garmo, Ø.A., M. Grung, Calidonio, J.L.G., de Wit, H., Moe, T.F., 2020. Overvåking av referanseelver 2020 - Basisovervåking i henhold til vannforskriften. NIVA-rapport 7485.

NIVA: Norway´s leading centre of competence in aquatic environmentes

NIVA provides government, business and the public with a basis for preferred water management through its contracted research, reports and development work. A characteristic of NIVA is its broad scope of professional disciplines and extensive contact network in Norway and abroad. Our solid professionalism, interdisciplinary working methods and holistic approach are key elements that make us an excellent advisor for government and society.

NIVA

Norwegian Institute for Water Research