A Load Analysis of the Potomac River at Chain Bridge Comparing A Dense-Sampling Direct Method with a Regression-Based Method

Submitted by

The Occoquan Watershed Monitoring Laboratory
Virginia Tech
9408 Prince William Street
Manassas, Virginia 20110

To

The Metropolitan Washington Council of Governments
777 North Capitol Street, NE, Suite 300
Washington, DC 20002

Final Version: 28 February 2018

VIRGINIA TECH

Metropolitan Washington Council of Governments

Acknowledgements

This report was prepared under a data analysis grant provided by the Metropolitan Washington Council of Governments (MWCOG) to Virginia Tech's Occoquan Watershed Monitoring Laboratory (OWML). Principal authors of the report were Saurav Kumar, Adil Godrej and Harold Post, all at OWML.

The authors would like to thank Karl Berger of MWCOG for the various reviews of the report, and the many editorial suggestions for improvement. The authors would also like to thank Ross Mandel of the Interstate Commission on the Potomac River Basin for detailed technical comments.

Table of Contents

1 SUMMARY

Since 1982, the Occoquan Watershed Monitoring Laboratory (OWML) has, on behalf of the Metropolitan Washington Council of Governments (MWCOG), operated a gaging and automated sampling station at Chain Bridge on the Potomac River. Paced by flow measurements at Little Falls Dam, storm samples are automatically composited by in-field flow weighting. Base flow sampling is typically weekly. This relatively dense sampling strategy has led to a large database upon which accurate direct load computation may be based. It can also be used to compare and contrast other methods that depend on regression. In recent years, the USGS (United States Geological Survey) has developed a modified weighted regression method–called the Weighted Regression on Time, Discharge, and Season (WRTDS)–to address some of the issues with other regression-based fluvial load estimation schemes, and to perform a better trend analysis. This report compares the long term loads and load trends estimated by WRTDS with those estimated by the dense sampling-based direct method that OWML uses.

The OWML dataset from 1983-2013 was used for the load calculations. Over 2000 samples were collected in the period, with 1151 non-storm, 620 storm, and 443 composite samples. The average censored (below detect) percentage is 16, 5, 6, 23, 20, 7, 7% for NH3N (ammonia nitrogen), OXN (oxidized nitrogen), TN (total nitrogen), SRP (soluble and reactive phosphorus), TSP (total soluble phosphorus), TP (total phosphorus), and TSS (total suspended solids), respectively. With, on an average, more than 50 non-composite discrete samples (both storm and non-storm) every year for over thirty years the dataset may be considered very rich compared to the other major Chesapeake Bay non-tidal monitoring sites, where about 20 samples are collected every year. In terms of annual load estimation, because all storms are sampled in a composite manner, the storm loads can be considered to be calculated with great accuracy. For all of the nutrient and sediment parameters sampled in the OWML program, storm loads comprise anywhere from 24% (OXN) to 41% (TSS) of the total amount of annual loads (Table 2).

The OWML Direct Method (DL: direct load) of load computation was one of the three load methods compared in this study. The other two methods were: (i) WRTDS applied to OWML data (WL-O: WRTDS load-OWML data), and (ii) WRTDS applied to USGS data (WL-U: WRTDS load-USGS data). The DL method is described in detail in another section of this report. The main characteristic of the DL method is that it computes loads directly from the observed data, with some interpolation between baseflow sampling events, and has very accurate storm loads. It does not rely on either a calibrated regression model or other statistical techniques to estimate daily concentrations to compute loads. The DL method is probably more accurate with a denser (more-frequent sample collection) dataset. The WRTDS method was developed for sparser datasets and for trends analyses.

An annual load comparison between the DL and WL-O (meaning, the same OWML dataset, but using the two different computation methods) indicated that WL-O generally overpredicted the

load more often than it underpredicted, compared to the DL method. The maximum overprediction percent difference (PD) from the DL for the entire 31-year period ranged from 21% for oxidized nitrogen (OXN) to 95% for ammonia nitrogen (NH3-N). The maximum underprediction PD ranged from 13% for total nitrogen (TN) to 48% for total soluble phosphorus (TSP). Full annual results of this comparison are shown in Table 3.

Overall statistical testing of the 31 years of annual loads suggested that the differences were statistically significant (≠0) with > 95% confidence for TSP, TP, NH3N, TN and TSS. In other words, the two methods of load computation/estimation (direct load and WRTDS) result in statistically significantly different loads. An analysis of storm and non-storm loads separately indicates that the WRTDS method does better in predicting the loads for storm days and overpredicts for non-storm days. Further analysis of the prediction equation generated and the half-widths used to assign weights for calibrating WRTDS methods is needed to ascertain the cause of this difference. For the 31-year period analyzed, the WRTDS method does a good job of predicting loads, and this is unsurprising because the method was developed for trend analysis. However, if annual loads are required, such as for setting annual load allocations and targets, the WRTDS method should be used with caution as, depending on the year and parameter, varying relative differences were observed (Table 3).

Comparison of daily load differences for storm days and non-storms days (Figure 27) show that WL-O predictions for storm days are better than for non-storm days. On non-storm days, WL-O appears to be predicting significantly higher fluxes. A higher percentage of storm samples available in the OWML database used to calibrate the WL-O prediction equation may explain some of the over estimation at low base flows. However, further investigation of the prediction equation is needed to conclusively explain the cause. It may be noted that daily loads difference are an important input to the water resources models, such as the Chesapeake Bay model, and large differences between daily loads observed between two load computation methods, and trends in the difference, need further investigation.

Comparison between WL-O and WL-U was restricted to only four parameters (SRP, TP, OXN, and TN) because these were the only parameters common to both the OWML and USGS datasets. WRTDS with two different input datasets yielded similar annual load estimates. Among the four parameters analyzed, the most variation was found in SRP. For SRP a pattern in the difference was also noticeable, from negative differences in the early 1980s to positive differences after the 2000s. A statistically significant difference between WL-O and WL-U was found only for OXN. Further investigation of the raw data and predictive equations is needed to explain the cause(s) of these observations.

It is important to note that results presented here are for the Potomac River. It is reasonable to expect that WRTDS will perform similarly in other streams draining large watersheds such as those for the Susquehanna and James Rivers. However, these results, as noticed with ESTIMATOR and other regression-based methods, may not necessarily apply to flashier streams draining smaller watersheds. Comparative studies for smaller watersheds are necessary to establish the application of WRTDS for smaller, flashier streams.

The Potomac River site at Chain Bridge is unique in the sense that it is the only site on a large tributary of the Chesapeake Bay where a single-point sampling scheme is feasible, due to the nature of the well-mixed water resulting from the restricted channel characteristics at Chain Bridge. It is also a site at which the loads computed by the direct method estimate the greatest portion of the loads arising from areas upstream of tidal waters. Thus it serves as a means to compare the loads generated by the more rural areas upstream of Chain Bridge with those loads generated downstream, which includes most of the Washington metropolitan area.

2 INTRODUCTION

Fluvial loads and observed concentrations of water quality parameters are essential to establish the state of a waterbody and its tributary watershed, and, over time, assess trends in water quality. For the Potomac River, which is a major tributary to the Chesapeake Bay, important water quality parameters of interest include:

1. Total nitrogen (TN)
2. Nitrate and nitrite nitrogen (OXN)
3. Ammonia nitrogen (NH_3-N)
4. Total phosphorus (TP)
5. Total soluble phosphorus (TSP)
6. Soluble reactive phosphorus (SRP)
7. Total suspended solids (TSS)

Load estimates of these water quality parameters are used to establish budgets for nonpoint pollution sources and design remediation/mitigation strategies in the region. There are several methods that may be used to estimate loads. These methods have their strengths and weaknesses. For example, methods that rely on extensive sampling produce reliable load estimates but are resource intensive to undertake, and methods that do not need extensive sampling often are not very accurate at shorter time scales. Given the implications of fluvial loads on watershed management strategy, it was deemed important to verify load estimates used by the Chesapeake Bay Model. This report describes one such exploration carried out in two phases to compare loads and trends estimated by two methods at Chain Bridge on the Potomac River. The first phase concentrated on comparing loads obtained from two methods (described later), and the second phase assessed the time trends observed in loads computed by the same two methods.

2.1 WHAT IS FLUVIAL LOAD?

Fluvial load may be defined as the mass of a constituent of interest transported by a stream in a given time. Most methods for estimating fluvial loads involve estimation or measurement of stream flow rates and concentrations of the constituent of interest. Loading/flux or instantaneous transport rate is the product of flow rate (Q) and concentration (c). The total fluvial load (L) over a period of time (t) may be computed as in Equation 1 [*Ferguson*, 1987].

$$L = \int_0^t Qc\,dt \approx \sum_{i=0}^{n} Q_i c_i \Delta t_i \qquad (1)$$

$where\ ,$

$$L = fluvial\ load\ over\ time\ T, kg$$
$$t = load\ computation\ time, s$$
$$Q = instantaneous\ flow\ rate, \frac{m^3}{s}$$
$$c = instantaneous\ concentration\ of\ the\ constituent\ of\ interest\ , \frac{kg}{m^3}$$

$$n = number\ of\ steps\ such\ that\ \sum_{j=0}^{n} \Delta t_j = t$$

$$Q_i = representative\ flow\ rate\ for\ interval\ \Delta t_i, \frac{m^3}{s}$$

$$c_i = representative\ concentration\ for\ interval\ \Delta t_i, \frac{kg}{m^3}$$

As Δt_i becomes smaller (or *n* becomes larger), the summation term gets closer to the integral term. It is, therefore, evident that the most accurate estimates for loads may be obtained using continuous measurements of flow rates and in-stream concentrations of the water quality parameters of interest. Near-continuous flow measurement may be done using a variety of techniques, and some water quality parameters such as temperature and dissolved oxygen may also be measured in a near-continuous fashion. However, the measurement of many water quality parameters such as those listed earlier require sample collection followed by laboratory analysis. This requirement practically rules out long-term high-frequency water quality measurements. Nevertheless, with judiciously frequent water quality measurements, excellent estimates of fluvial loads may be made. For example, the concentration for a stream with a large watershed is not expected to change much from day to day during non-storm flow periods, and very good to excellent load estimates may be obtained for non-storm periods using a weekly or bi-weekly water quality sampling scheme. During storm events, when the concentrations of the parameters of interest (and, consequently, load) may be expected to vary rapidly, a compositing method of load estimation that yields an event mean concentration (EMC) for the storm may be employed.

In resource-constrained scenarios, where frequent composite storm sampling and analysis is not feasible, regression methods are often used for estimating the concentration of the parameters of interest. Regression-based methods estimate the concentration of the constituent of interest by relating flow and other readily measurable surrogate parameters to the concentration of the constituent of interest. Typically, regression-based methods involve less sample collection and analysis when compared to direct sampling methods, as the observed concentration data are only used for calibration of a regression equation. The putative trade-off for this method is often reliability of estimated loads.

There is considerable evidence that suggests that in smaller watersheds the regression based methods do not perform adequately, and even in larger watersheds significant differences were found in long-term studies [*Robertson and Roerish*, 1999; *Toor et al.*, 2008; *Kumar et al.*, 2013]. To improve the efficiency of the regression equation-based fluvial load computation, different sampling methods, such as hydrological-based sampling, storm chasing, sampling in the rising or falling limb of the hydrograph, and adaptive cluster sampling have been used with varying degrees of success [*Robertson and Roerish*, 1999; *Horowitz*, 2003; *Sadeghi et al.*, 2008; *Arabkhedri et al.*, 2010; *Sadeghi and Saeidi*, 2010]. Data stratification techniques such as seasonal stratification and magnitudinal stratification are also reported to be successful. Using similar methods, the USGS has developed a modified weighted regression method–Weighted Regression on Time, Discharge, and Season (WRTDS)–to address some of the issues with regression-based fluvial load estimation schemes [*Hirsch et al.*, 2010]. It may be noted that WRTDS utilizes multiple regression equations (> 2000 for about 10 years of data) and internal smoothing, which is more complicated and likely to yield better estimates than the typical regression method that uses one equation for the period of interest.

Virginia Tech's Occoquan Watershed Monitoring Laboratory (OWML), on behalf of the Metropolitan Washington Council of Governments (MWCOG), has operated a gaging and automated sampling station at Chain Bridge on the Potomac since 1982. Paced by flow measurements at Little Falls Dam, storm samples are automatically composited by flow weighting. Base flow sampling is typically weekly. This relatively dense sampling strategy has led to a large database upon which accurate direct load computation may be based. It can also be used to compare the accuracy of other methods that depend on less data. This report compares the long term loads estimated by WRTDS with those estimated by the dense sampling-based direct method that OWML uses.

2.1.1 Load Trends

Load trends are used to describe changes in fluvial load over time at the same location. Though these temporal load trends are useful for assessing the health of the receiving water body, they have limited utility in evaluating watershed management plans, particularly the land-use based interventions made in the watershed. This is because naturally varying flow strongly influences the concentrations and loads of the constituents of interest, and a large portion of the variations in the observed fluvial load may be attributed to the natural variations in the flow rather than any interventions that were made as a part of a watershed management plan. Through various statistical/mathematical operations (discussed later) the effect of the natural variations in streamflow may be removed from fluvial loads to develop flow–independent (flow–adjusted or flow–normalized) fluvial load trends. By removing the effect of natural streamflow variations that confound the year-to-year fluvial load variation, these flow–independent fluvial load trends are better suited to study changes in the watershed and their impact on long-term in-stream water quality. This report only looks at flow–independent fluvial load trends. It may be noted that sometimes a distinction is made between flow–adjusted and flow–normalized fluvial load based on computation method. Nevertheless, both these load trends seek to represent the variations in the fluvial loads after accounting for variations in streamflow.

2.2 Scope of this Report

In mid-2013, MWCOG commissioned OWML to compute and study fluvial loads for the Potomac River (first phase). The objective was the comparison of fluvial loads for the seven water quality parameters listed above for the historical record (1983-2013) using two methods: the direct method (used by OWML) and the Weighted Regression on Time, Discharge, and Season (WRTDS) developed by USGS (first phase). Upon completion of first phase, a second phase was instituted in late 2014. The objective of the second phase was to extend the work done in the first phase and investigate differences in long term flow–independent trends estimated by the two methods. This report describes the data and methods used to compute these fluvial loads and flow–independent trends, tabulates the fluvial loads computed, and compares the loads and trends estimated by various applications of the two methods.

3 Study Site

The OWML monitoring station (PR01) on the Potomac River at Chain Bridge is on the Virginia-DC border. The watershed for PR01 is about 2.8 x10^4 km^2 and extends into Virginia, West Virginia, Maryland, Pennsylvania, and Washington DC (Figure 1). Due to a constriction in the Potomac River channel near Chain Bridge the water is well-mixed and suitable for representative point sampling. The site, however, is not ideally suited for stream gaging due to the variable backwater effects from the tidal fresh water portion of the estuary and the topology of the stream bed near Chain Bridge which complicates development of low and moderate flow rating tables. About 2.4 km upstream of Chain bridge is the Little Falls Dam. The wide dam with a mid-stream island is ill-suited for representative point sampling, but provides a good control to establish a stable flow rating. The USGS has monitored the stage at this site since 1930, and maintains a current rating curve. There is very little additional drainage area between Little Falls Dam and Chain Bridge. Thus the flow measurements at the gaging station may be used for fluvial load estimation with water quality measurements taken at Chain Bridge. Note that all sampling is conducted under the assumption of a well-mixed flow regime at Chain Bridge due to narrowing of the Potomac River at that point and this assumption has been verified against the USGS cross-section integration approach. OWML maintains stage measurement equipment and dataloggers independent of the USGS at the Little Falls Dam, and these measurements are used to trigger flow composite sampling at the Chain Bridge water quality monitoring station when a storm has been detected. Figure 2 shows a schematic of the connection between equipment at Little Falls Dam and Chain Bridge. Details of station operation have been reported earlier by *Post and Grizzard* [1987].

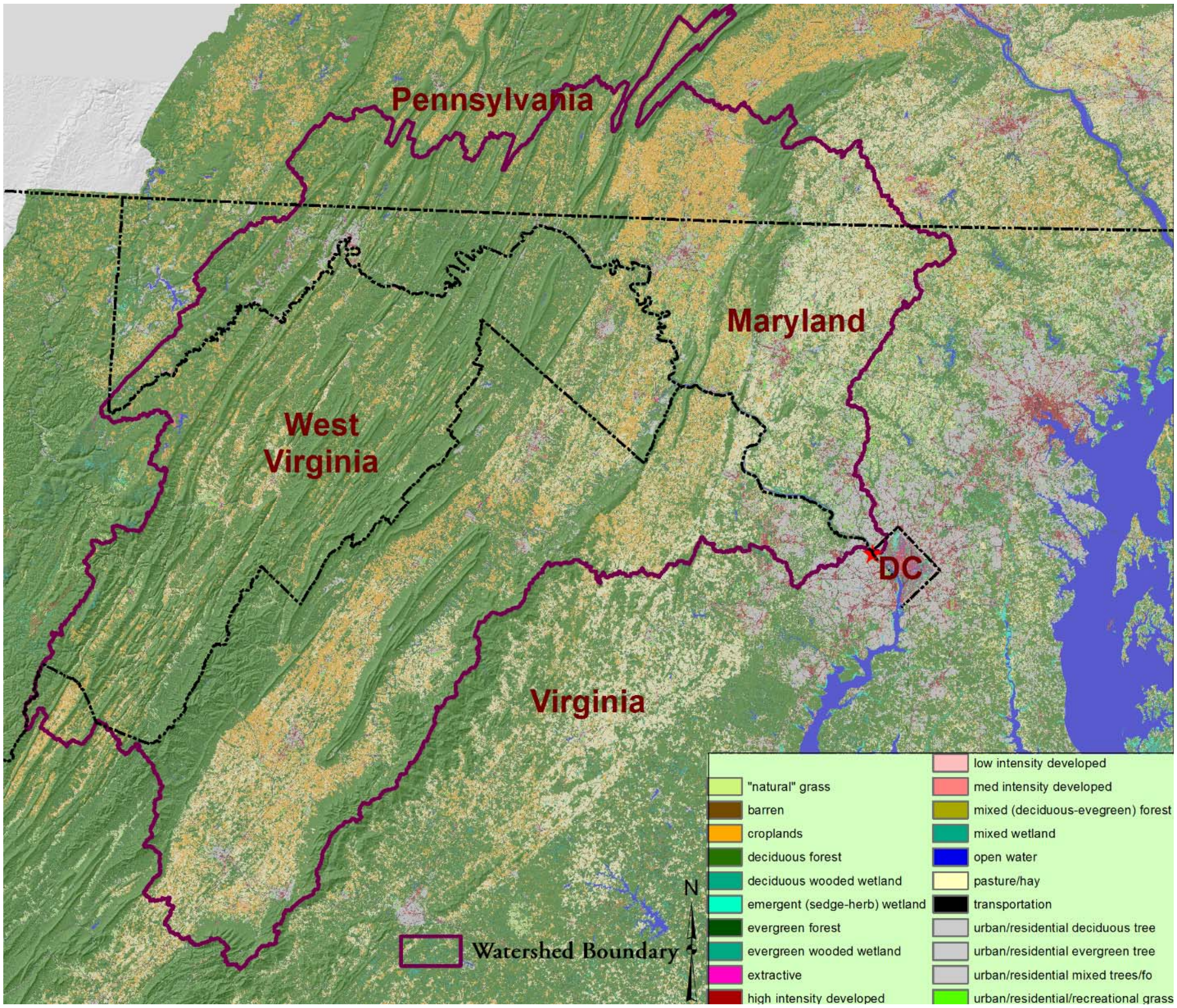


Figure 1. Potomac Watershed for PR01 stream station at Chain Bridge. The station is marked with a red star on the DC -VA border.

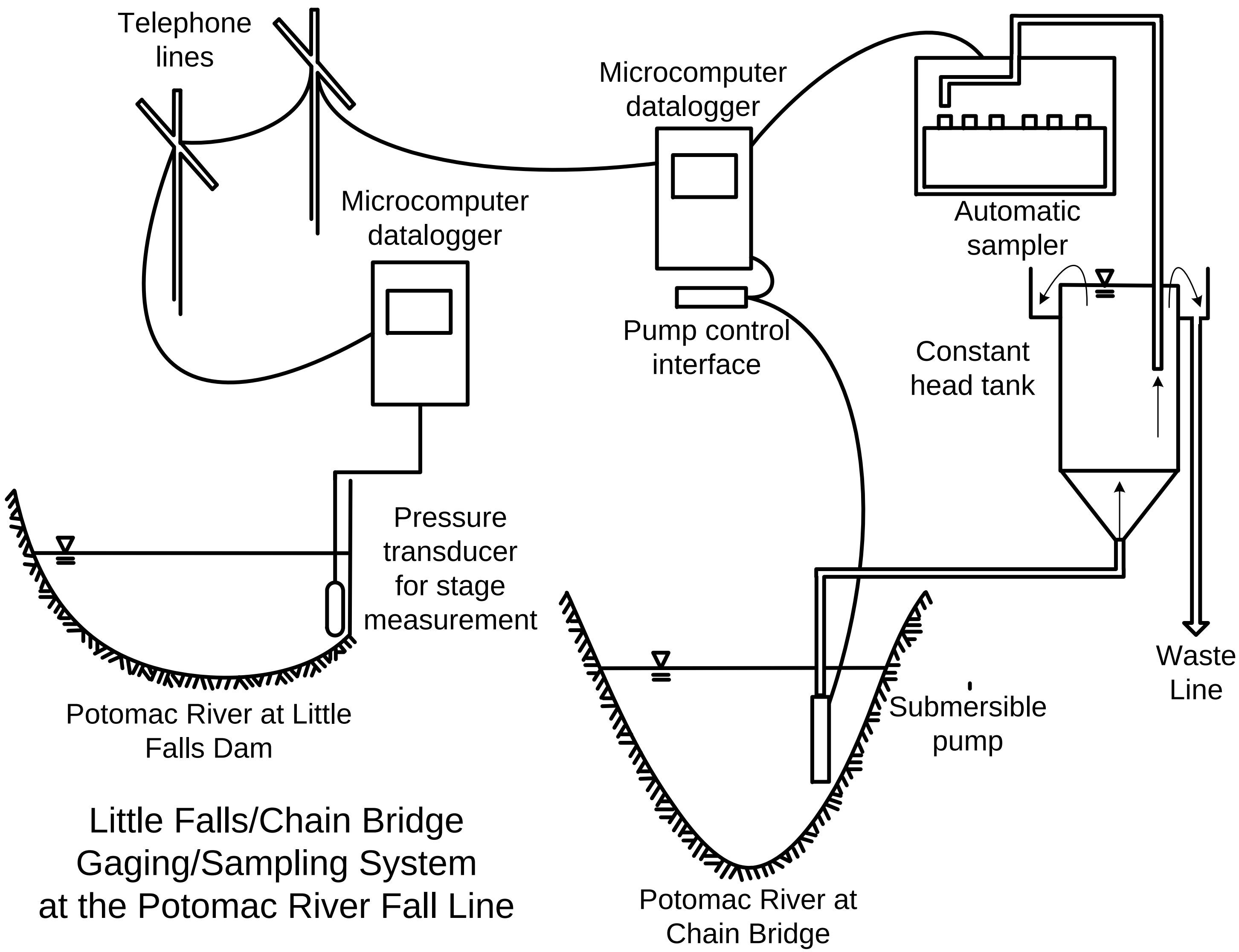


Figure 2. Schematic showing communication between flow measurement devices at Little Falls dam and water quality sampling devices at Chain Bridge about 2.4 km downstream (adapted from Post and Grizzard [1987]).

4 METHODS

4.1 DATA COLLECTION

Water quality and flow data for all stations were obtained from the OWML database of observed water quality and quantity conditions. Flows are currently measured and recorded by OWML at Little Falls using a Pressure Systems® 200S pressure transducer and Sutron® 8210 data loggers. Although equipment has been updated or changed since the start of the monitoring program in 1982 to the present, to insure data integrity appropriate QA/QC checks are performed regularly, and when new equipment is deployed on the field. Discrete manual samples are collected either on a weekly (most of the year) or biweekly (winter) cycle. In addition, flow-weighted composite samples using a Manning® auto-sampler are collected for all storm events (minus those missed due to equipment malfunction, failure of power, or other such problems). In order to facilitate comparison with various regression methods, since November 1995 discrete storm samples are also collected for up to fine storms per year, with five discrete samples per storm event.

All water quality constituents were analyzed in-house at OWML using methods in accordance with the latest edition available at the time of *Standard Methods for the Examination of Water and Wastewater* published by American Public Health Association. Because of the long term nature of the study (1983-present) laboratory analytical methods may have changed over time. Appropriate QA/QC protocols and comparability studies performed whenever methods are changed or modified insures that concentration data are accurate, and that new methods do not result in a loss of accuracy and precision.

4.1.1 Composite Sampling

OWML attempts to collect a flow-weighted composite sample (composite sample) for all storm events. The compositing is performed by an onsite auto-sampler so that only one sample may be analyzed in the laboratory to obtain the EMC. For the purposes of composite sampling, a storm event is said to begin in the Potomac River when the on-site microcomputer detects two consecutive stage increases greater than 0.01 ft and the flow is greater than the expected baseflow, Q (described later; also see equation 3), for more than two hours. Once storm sampling is initiated, the incremental volume (ΔV) is computed every minute using equation 2, and summed to compute cumulative storm flow volume.

$$\Delta \boldsymbol{V} = \frac{\boldsymbol{Q_i}+\boldsymbol{Q_{i-1}}}{\boldsymbol{2}} \times (\boldsymbol{t_i} - \boldsymbol{t_{i-1}}) \qquad (2)$$

$where$,

$$\Delta V = incremental\ flow\ volume, m^3$$
$$Q_i = discharge\ at\ present\ stage\ reading, \frac{m^3}{s}$$
$$Q_{i-1} = discharge\ at\ previous\ stage\ reading, \frac{m^3}{s}$$
$$t_i = time\ at\ present\ stage\ reading, s$$
$$t_{i-1} = time\ at\ previous\ stage\ reading, s$$

The ΔV values are summed for each increment of time to compute storm volume. Storm volume is used to trigger sample collection. Every time the total storm flow is increased by a pre-determined volume, which may be adjusted based on the length and intensity of expected event (typically $V=400\times10^6$ ft^3), that has passed the station, the sampler retrieves a fixed aliquot and adds it to a refrigerated container.

Sample collection continues in this manner until the end of the storm, which is marked by the stream flow declining to the point where the recession limb of the hydrograph equals the expected base flow computed by equation 3. In this manner, because there is an equal aliquot of sample collected for each equal increment of flow volume, a flow-weighted composite sample is automatically constructed (Figure 3). For storms that exceed refrigerated sample holding times, multiple composite samples may need to be employed, with each representing a section of the storm event (typically 1-2 days). See *Post and Grizzard* [1987] for more details on the composite sampling method employed.

$$\boldsymbol{Q = Q_b + 5 \times 10^{-4} A \cdot t} \tag{3}$$

$where$,

$Q = discharge\ value\ at\ flow\ separation\ line\ (expected\ baseflow\ value), \frac{m^3}{s}$

$Q_b = discharge\ at\ beginning\ of\ the\ storm, \frac{m^3}{s}$

$A = area\ of\ drainage\ basin, square\ miles$

$t = time\ elapsed\ from\ start\ of\ storm, s$

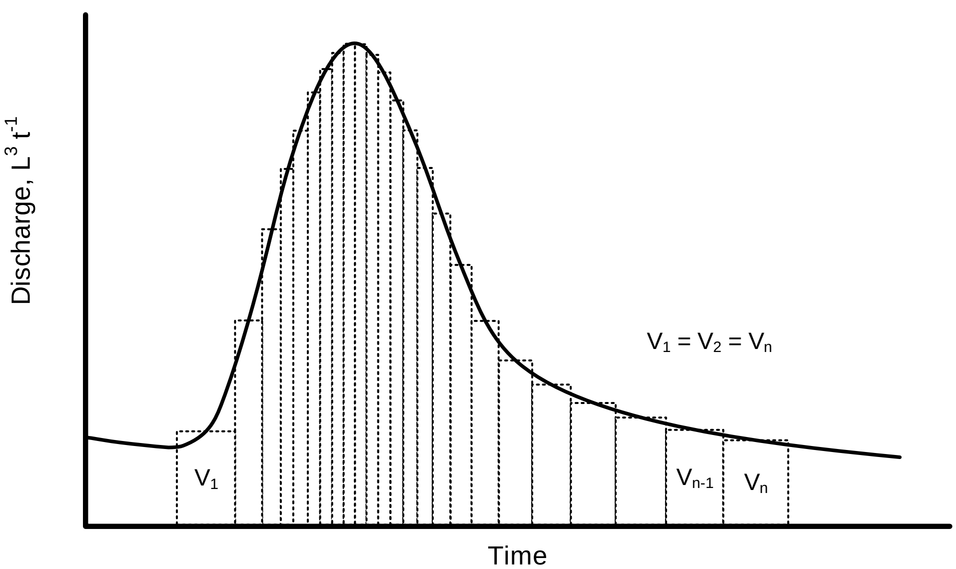


Figure 3. Flow-weighted composite sampling schematic. After every equal increment of volume V_{in} (typically 400×10^6 ft^3) has passed the sampling point, one fixed volume sample is taken and added to a refrigerated carboy, effectively compositing samples on-site by taking one sample for every rectangle of equal volume shown in the figure.

4.1.2 Dataset Used

Table 1 summarizes the available water quality dataset for the seven parameters of interest in the Potomac River including count and censoring percentage. The rows labeled 'B' indicate non-storm flow (baseflow), discrete storm flow samples are labeled 'R' (for runoff), and the label 'EMC' (event mean concentration) is used to describe the composite samples. It may be seen that over 2000 samples were collected in the study period from 1983-2013, with 1151 non-storm samples, 620 storm samples, and

443 composite. The average censored (below detect) percentage is 16, 5, 6, 23, 20, 7, 7% for NH_3N, OXN, TN, SRP, TSP, TP, and TSS respectively. With, on an average, more than 50 non-composite discrete samples (both storm and non-storm) every year for over thirty years the dataset may be considered very rich among the other major Chesapeake Bay non-tidal monitoring sites where about 20 samples are collected every year.

Table 1. Summary of dataset used in this study. Blue horizontal bars are a pictorial rendering of the data counts.

Data Count					Percentage Censored and Missing Data (%)																											
					NH_3N				OXN				TN				SRP				TSP				TP				TSS			
Year	B	EMC	R	Total	B	EMC	R	Total	B	EMC	R	Total	B	EMC	R	Total	B	EMC	R	Total	B	EMC	R	Total	B	EMC	R	Total	B	EMC	R	Total
1983	15	24	6	45	7	17	17	13	0	4	0	2	7	8	0	7	7	13	17	11	27	33	50	33	0	4	0	2	7	4	0	4
1984	33	22	23	78	30	18	0	18	0	18	0	5	0	18	0	5	21	18	4	15	0	18	4	6	0	18	4	6	0	18	0	5
1985	31	23	28	82	19	4	0	9	3	4	0	2	6	9	4	6	13	4	0	6	3	4	82	30	0	9	4	4	0	0	0	0
1986	39	7	1	47	21	14	0	19	10	14	0	11	10	14	0	11	46	14	100	43	5	14	0	6	0	14	0	2	5	14	0	6
1987	36	14	2	52	39	43	0	38	3	36	0	12	3	36	0	12	17	36	0	21	14	36	0	19	3	36	0	12	11	36	0	17
1988	49	5	4	58	33	0	25	29	6	0	0	5	12	0	0	10	24	0	0	21	20	20	0	19	12	0	0	10	12	20	25	14
1989	38	23	4	65	34	9	25	25	0	9	0	3	8	13	0	9	16	9	0	12	32	52	0	37	18	13	0	15	8	4	0	6
1990	40	14	1	55	30	14	0	25	0	14	0	4	0	14	0	4	40	14	0	33	33	43	100	36	5	21	0	9	5	7	0	5
1991	53	6	2	61	40	50	0	39	13	50	0	16	15	50	0	18	34	50	0	34	49	50	0	48	36	50	0	36	26	50	0	28
1992	42	13	2	57	26	15	0	23	5	15	0	7	5	15	0	7	26	23	0	25	40	23	50	37	21	8	0	18	12	23	0	14
1993	38	13	4	55	37	31	50	36	3	23	50	11	5	23	50	13	37	23	50	35	34	23	50	33	8	23	50	15	11	23	75	18
1994	41	9	2	52	51	0	50	42	7	0	0	6	7	0	0	6	59	0	0	46	49	11	50	42	10	0	0	8	10	0	0	8
1995	35	10	17	62	17	20	29	21	0	20	0	3	0	20	0	3	11	30	18	16	43	60	53	48	17	20	0	13	3	30	0	6
1996	27	31	61	119	44	29	3	19	0	19	0	5	7	23	0	8	48	29	7	22	52	55	41	47	22	23	0	11	0	19	0	5
1997	42	11	17	70	12	36	6	14	0	36	0	6	0	36	0	6	33	55	6	30	33	55	29	36	7	36	0	10	0	36	0	6
1998	40	15	37	92	25	33	5	18	0	33	0	5	0	33	0	5	38	33	3	23	8	33	0	9	3	33	0	7	0	33	0	5
1999	39	9	28	76	15	22	0	11	0	22	0	3	0	22	0	3	10	22	0	8	13	22	4	11	3	22	0	4	5	22	0	5
2000	38	11	19	68	13	18	0	10	3	18	0	4	3	18	0	4	29	18	0	19	8	18	0	7	0	18	0	3	5	18	0	6
2001	38	12	29	79	24	17	0	14	3	25	0	5	3	25	0	5	61	33	10	38	11	33	7	13	0	25	0	4	0	17	0	3
2002	39	12	25	76	15	17	4	12	0	8	0	1	3	8	4	4	15	8	0	9	3	8	4	4	0	8	4	3	3	8	0	3
2003	30	26	38	94	20	31	0	15	0	31	0	9	3	31	0	10	27	31	3	18	7	38	5	15	0	35	3	11	0	31	0	9
2004	30	22	58	110	13	27	10	15	0	23	0	5	0	23	2	5	33	27	5	17	0	23	3	6	0	23	0	5	0	27	3	7
2005	39	10	27	76	8	0	4	5	0	0	0	0	0	0	0	0	13	10	4	9	5	0	0	3	0	0	0	0	3	0	0	1
2006	33	14	36	83	6	14	3	6	0	7	0	1	0	7	0	1	21	7	6	12	3	7	0	2	0	7	0	1	9	7	0	5
2007	40	9	23	72	15	11	4	11	0	11	0	1	5	11	4	6	30	11	13	22	15	11	9	13	3	11	4	4	3	11	0	3
2008	38	19	29	86	13	11	3	9	5	5	0	3	5	5	0	3	18	21	28	22	11	5	7	8	3	5	0	2	13	5	0	7
2009	35	11	26	72	6	0	0	3	0	0	0	0	0	0	0	0	34	0	0	17	3	0	0	1	0	0	0	0	0	0	0	0
2010	41	9	8	58	2	11	0	3	5	11	0	5	0	11	0	2	56	33	50	52	22	11	0	17	5	11	0	5	5	11	0	5
2011	35	20	28	83	6	15	4	7	0	15	4	5	0	15	0	4	54	25	4	30	29	30	0	19	3	15	0	5	11	15	0	8
2012	41	9	14	64	17	11	0	13	2	11	0	3	0	11	0	2	37	44	0	30	20	11	0	14	7	11	0	6	12	11	0	9
2013	36	10	21	67	6	0	5	4	3	0	0	1	0	0	0	0	44	20	19	33	17	10	0	10	0	10	0	1	11	0	0	6
Total	1151	443	620	2214	21	18	5	16	3	16	0	5	4	17	1	6	31	21	7	23	20	27	13	20	7	17	1	7	7	16	1	7

4.2 LOAD COMPUTATION

The two methods of load computation utilized in this study are:

1) Direct Method, which used the OWML dataset of weekly/bi-weekly water quality data for the seven parameters of interest during non-storm flows and EMC for storm events, along with near-continuous (15 minutes to hourly) flow measurements.
2) WRTDS Method, which is based on a weighted regression technique described by *Hirsch et al.* [2010]. The WRTDS method uses the same OWML weekly/bi-weekly water quality data for parameters of interest during non-storm flows, but uses the discrete samples taken during storm events (shown as R on Table 1), along with daily average flow data. Note that in earlier years, there may be as few as one 'R' sample in a year. This is because prior to 1995-96, OWML did not collect storm discrete samples side-by-side with storm composite samples. The few discrete storm samples that were collected were when the river was in storm on a day when a baseflow sample was scheduled to be collected.

In addition to the load computations done on OWML data, comparisons were made with the WRTDS loads that the USGS had computed based on their water quality data. To satisfy ourselves that we were applying the WRTDS method to OWML data appropriately, we also computed WRTDS loads from the USGS water quality data and got similar results to those obtained by the USGS. This provides us with assurance that when we applied the WRTDS methods to OWML data, we were doing it correctly.

4.2.1 Direct Method

The direct method of fluvial load computation is an extension of the first principle of load calculation described in Equation 1. In this method a water quality concentration reading is assigned for every recorded flow reading. For non-storm flows in the Potomac River, regular periodic discrete samples measurements (weekly, bi-weekly) are used to interpolate concentrations at every data point where the flow is recorded (usually hourly during non-storm flows). For storm events, the flow-composite EMC value is assigned to all the recorded storm time flows (every fifteen minutes). Loads for a desired period of time are then computed using Equation 1. Note that in this method storm discrete samples are not used.

4.2.1.1 Data Vetting

For this study, prior to computation of fluvial loads the water chemistry and flow data sets were manually analyzed for errors. Water chemistry data are vetted as part of the standard OWML QA/QC protocol, and were presumed to be largely error-free. However, the automated high frequency flow recordings had not been manually vetted. Therefore, before load computation over 30 years of flow data were vetted, and missing or erroneous flow data were filled or replaced, respectively, using USGS fifteen minute and daily flow data from the same station. OWML has developed in-house software for the Direct Method of load computation with the required vetting and load computation operations included.

4.2.2 WRTDS Method

WRTDS was developed by USGS as a successor to the ESTIMATOR method of load estimation. USGS in their recent studies [*Moyer et al.*, 2013] have compared the WRTDS with the ESTIMATOR method for the nine non-tidal observation sites in the Chesapeake Bay and found some differences between the methods. The WRTDS uses a five parameter equation (Equation 4) to estimate concentration based of flow and time. This method is different from most other regression methods as the parameters of equation 4 ($\hat{\beta}_0$ to $\hat{\beta}_4$), instead of being fixed, are estimated for every combination of q (daily average flow) and t (time) where concentration has to be computed. The data set used to compute $\hat{\beta}_0$ *to* $\hat{\beta}_4$ is weighted based on the distance from q and t at the estimation point. Weight (w) is computed for three different distances: a) time distance, b) season distance, and c) discharge distance using the "tri-cube weight function" (Equation 5). Net-weight is taken as the product of these three weights. A 10 year half-window width is used for trend distance, 0.5 (decimal time) half-window width is used for seasonal distance, and 2.0 (ln(Q)) half window width for weight computation. These half-window widths (h) are similar to what was used and found to be optimum by USGS for non-tidal load computation in the Chesapeake Bay watershed [*Hirsch et al.*, 2010; *Moyer et al.*, 2013]. All load computations for the WRTDS methods were performed by using the R software Exploration and Graphics for RivEr Time-series (EGRET) package developed by USGS [https://github.com/USGS-R/EGRET/wiki, Access Date: 06/02/2014]. To save computational effort the EGRET package first computes a 3-dimensional matrix of expected concentration. The first dimension of this matrix is the time in years, the second dimension is the time in months, and the third dimension consists of 14 levels of discharge spanning the whole observed flow record set. This 3-dimensional matrix is then used to estimate concentration by linear regression at all daily average flow conditions. A much more detailed explanation about the WRTDS method may be obtained from the software webpage [https://github.com/USGS-R/EGRET/wiki, Access Date: 06/02/2014]. Daily fluvial loads estimates are then computed as product of daily concentration estimate and daily average flow. These daily loads may be aggregated as required to yield annual loads.

$$\ln(c) = \hat{\beta}_0 + \hat{\beta}_1 \ln(Q) + \hat{\beta}_2(t) + \hat{\beta}_3 \sin(2\pi t) + \hat{\beta}_4 \cos(2\pi t) + \varepsilon \qquad (4)$$

$c\ is\ concentration\ in\ \frac{mg}{l}$
$q\ is\ observed\ daily\ flow, in\ cfs$
$t\ is\ the\ decimal\ time\ in\ years$
$\hat{\beta}_0\ to\ \hat{\beta}_4\ are\ respective\ estimates\ for\ regression\ coefficient$
$\varepsilon\ is\ the\ unexplained\ variation$

$$w = \begin{cases} \left(1 - \left(\frac{d}{h}\right)^3\right) & if\ |d| \le h \\ 0 & if\ |d| > h \end{cases} \qquad (5)$$

$w\ is\ the\ weight$
$d\ is\ the\ distance\ from\ estimation\ point\ to\ data\ point$
$h\ is\ the\ half\ window\ width$

4.2.3 Flux Comparisons

Three load series using two methods of load computation were assembled and analyzed in this study:

1) Direct-Load (DL), the loads computed by the direct method using the OWML composite and non-storm time concentration datasets.
2) WRTDS-Load with OWML data (WL-O), the loads computed using the WRTDS method using the OWML storm and non-storm time discrete concentration datasets.
3) WRTDS-Load using USGS data (WL-U), these loads were obtained from "Water Quality Loads and Trends at Nontidal Monitoring Stations in the Chesapeake Bay Watershed" website

(http://cbrim.er.usgs.gov/; Access Date: 06/02/2014) and through email communication with staff at the USGS Virginia Water Science Center, Richmond, Virginia.

Annual loads obtained from these three methods are compared in this study. Comparisons were done using all three load series together, and separate statistical comparisons were done using a non-parametric matched-pair signed rank test comparing DL with WL-O and WL-O with WL-U. A matched-pair signed rank test was done for all constituents of interest, and if the fluvial load series were found to be substantially different (α=0.05) an unbiased magnitude-of-difference between loads was calculated with the "Hodges-Lehmann Estimator (Δ)" as suggested by *Helsel and Hirsch* [2002].

4.2.4 Flow–independent Fluvial Load Trends

As discussed earlier flow–independent fluvial load trends are developed to minimize the impact of natural variation in streamflow on fluvial loads trends. These flow–independent fluvial load trends highlight the load trends that may be attributed to changes/interventions in the watershed. WRTDS computes daily "flow–normalized" fluvial load by estimating daily flow–normalized concentration and multiplying it with the daily flow. For computation of daily flow–normalized concentration it is assumed that all flows that where seen on a particular day of year in the period of record are equally likely. By using the flows seen on a particular day of year (e.g. Jan 1) along with decimal time representing the day (e.g. Jan 1, 2001) concentrations are imputed for all flow values for a particular day. Imputation is done from the 3-dimensional matrix of expected concentrations developed earlier (see section 4.2.2). Daily flow–normalized concentration is then computed as the mean of all probable concentrations for the day (see *Hirsch et al.* [2010] for more details). For example, in a 10-year period of record (2000-2010) for Jan 1, 2001, all ten flows seen on Jan 1 will be used to impute 10 possible concentrations for Jan 1, 2001. The flow–normalized concentration for Jan 1, 2001 is then computed as the mean of 10 concentrations.

Daily flow–normalized fluvial loads were obtained for WL-O and WL-U from the EGRET program and aggregated annually (calendar year) to compute the annual flow–normalized fluvial load. Moving window averages for daily fluvial loads were computed for DL to minimize the impact of flow variance (DL-MA). Multiple moving window widths were used for computing the moving average. It was observed that within the study period the ratio between maximum to minimum averaged flow is ≈10 (one order of magnitude) when the averaging window width is 285 days (see Appendix H). Thus 285 days was used as one window width (DL-MA-285 days). In addition, a 2-year window was also tried (DL-MA-2 years). It may be noted that averaging over large periods of time (as is the case here) minimizes the daily fluctuations (see Appendix H) in flow and thus is beneficial for looking at long term trends independent of flow variations.

Another method was used to analyze flow–adjusted DL based on removal of exogenous variable for trend analysis as described by *Helsel and Hirsch* [2002]. In this method, variations in daily loads that may be explained by variations in daily flow was removed by using LOWESS regression and residuals were analyzed for trend (DL-Residuals). As with WL-O and WL-U trends, the daily loads were aggregated annually. Additionally, all annual loads trends were centered and rescaled. Note: one additional method that relied on regressing a seven parameter equation (similar to equation 4, with additional t and ln(Q) squared terms) was also tried with daily and weekly averaged flow and concentration data. The objective of this method was to study the coefficients of time terms for trends. However, statistical assumptions required to carry out such a regression could not be satisfied and the method was not used.

Akritas-Theil-Sen nonparametric line slope (slope) was also computed for change in annually averaged flow-independent daily (avg-daily) load computed through all methods. To further characterize change in avg-daily load, separate slopes were computed for years <1990, 1990-1999, and >1999.

5 Results and Discussion

5.1 Raw Data Summary

The 31-year period of record from 1983-2013 was used in this study. In the analysis period a total of 2214 water quality samples were taken, out of which 1151 where non-storm flow samples, 620 were storm flow discrete and 443 were storm composites (Table 1). Figure 4-6 shows the concentration distribution for the seven parameters of interest over the entire study period. It is clear from the plots that the median concentration for all constituents of interest were higher for storm discrete and composite samples than the non-storm samples. The fact that storm discrete sample quantile ranges and median values are fairly similar to those of the EMC distribution indicate that the discrete sampling scheme was probably not biased towards low-concentrations (typically the falling limb on a storm hydrograph). The annual concentration distribution of all seven parameters of interest are shown in Appendix A.

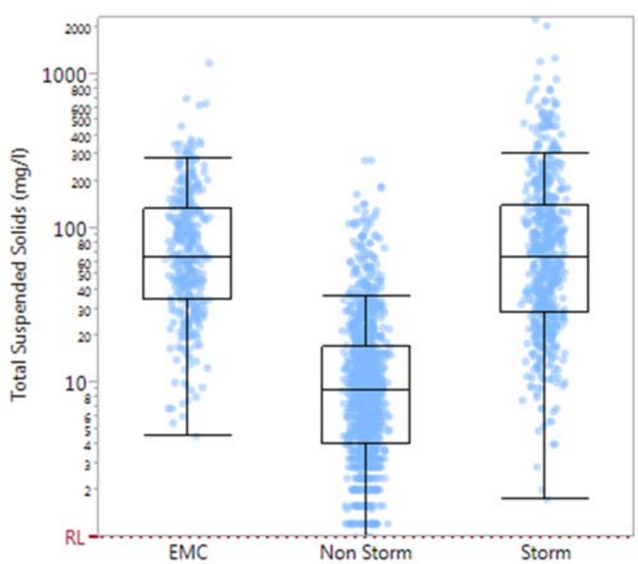


Figure 4. Concentration distribution of the total suspended solids species in the period of analysis. RL indicates the Reporting Limit for TSS.

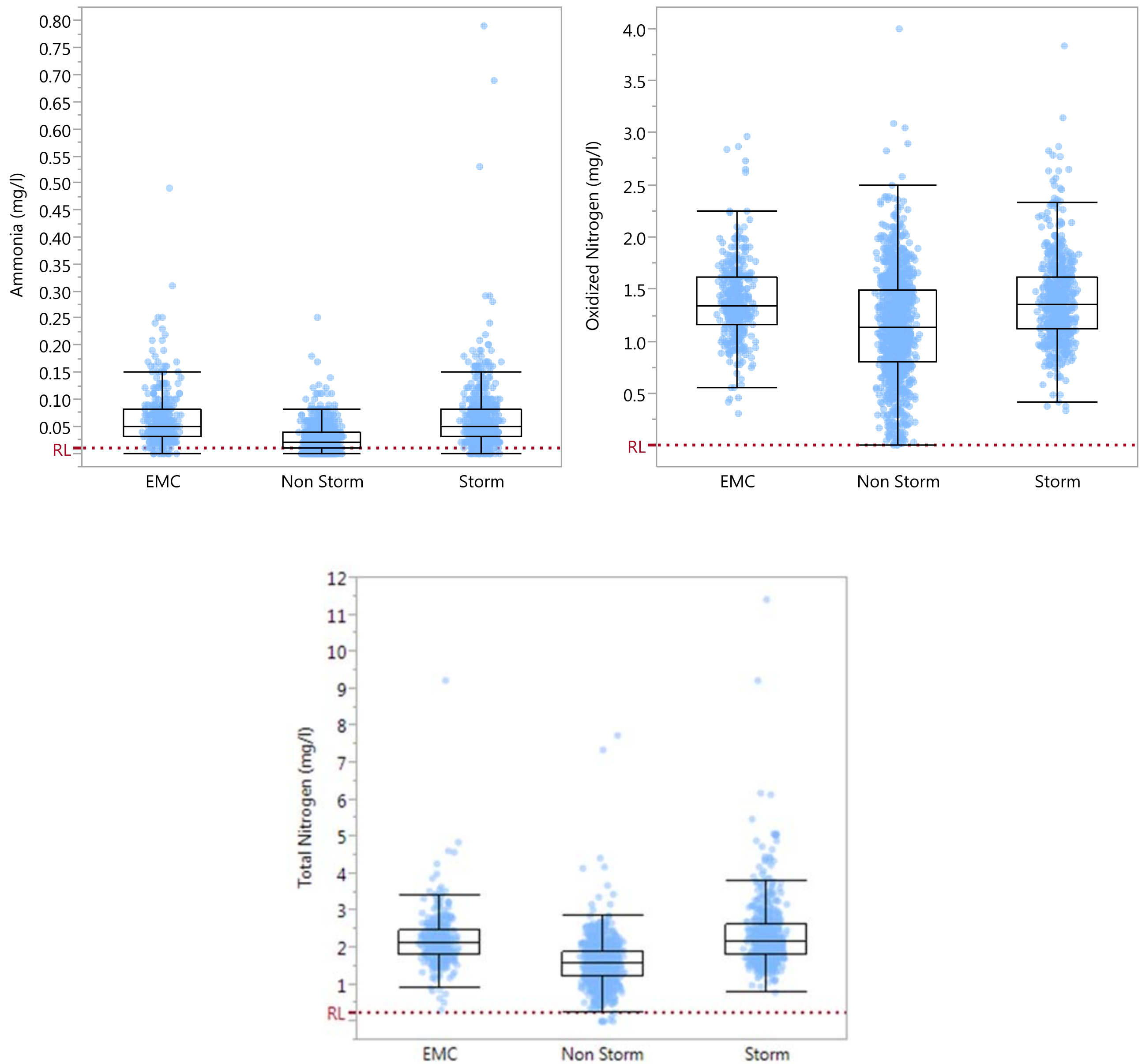


Figure 5. Concentration distribution of the Nitrogen species in the period of analysis. RL indicates a Reporting Limit.

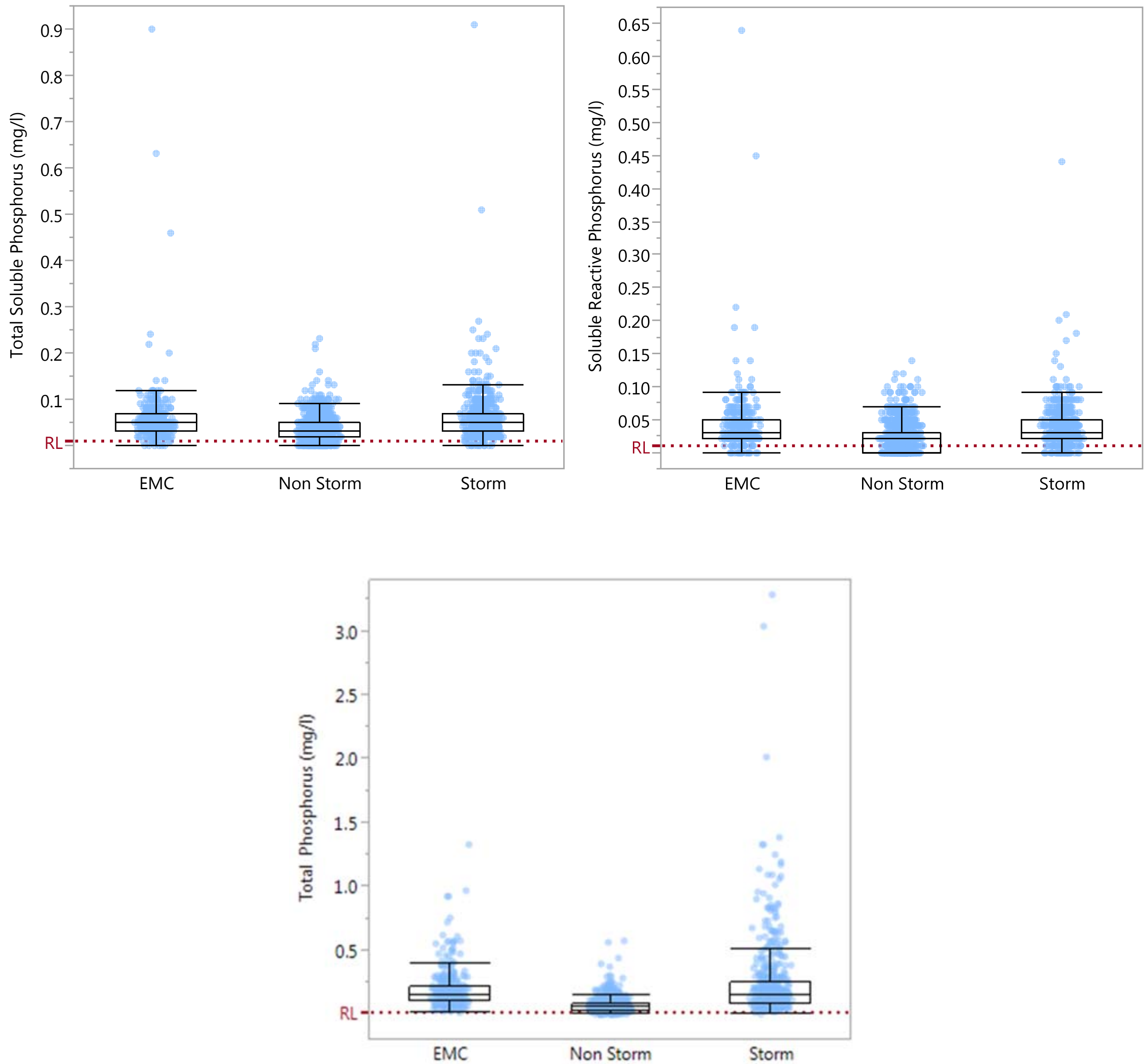


Figure 6. Concentration distribution of the Phosphorus species in the period of analysis. RL denotes the Reporting Limit.

Typically, non-storm flow is recorded every hour as long as it does not change over that period, else it is recorded more frequently, while storm flow is recorded every 15 minutes. Figure 7 shows the Storm and Non-Storm flow distribution. Note that OWML started automated flow sampling from 1985, and the data in this distribution are from 1985 onwards. Annual and monthly distribution of the flows classified as Storm and Non-Storm is presented in Appendix B. Median storm length in the Potomac River is about 110 hours (4.5 days). However, storms varying in length from 18 hours to 314 hours (~13 days) have been seen in the last 30 years (Figure 8). The longer storms pose a unique sample collection and analysis problem. As mentioned earlier, for some of these storm the onsite flow-weighted compositing process for estimating EMC is divided into multiple shorter time composites, to avoid exceeding either the recommended sample storage time before analysis or the total volume available for the compositing. In such cases multiple samples are analyzed and an effective EMC is computed and assigned to the storm event.

Analysis was done to identify the relationship, if any, between the EMC for constituents of interest and volume of storm (ft^3), average flow in storm (cfs), and storm length (hours). Statistically significant relations were found between average flow and TN, TP, and TSS, but the explanatory power of the relation (R^2) was very weak, rendering such determinations to be not very useful. Some portions of this analysis with statistically significant relationships are presented in Appendix C.

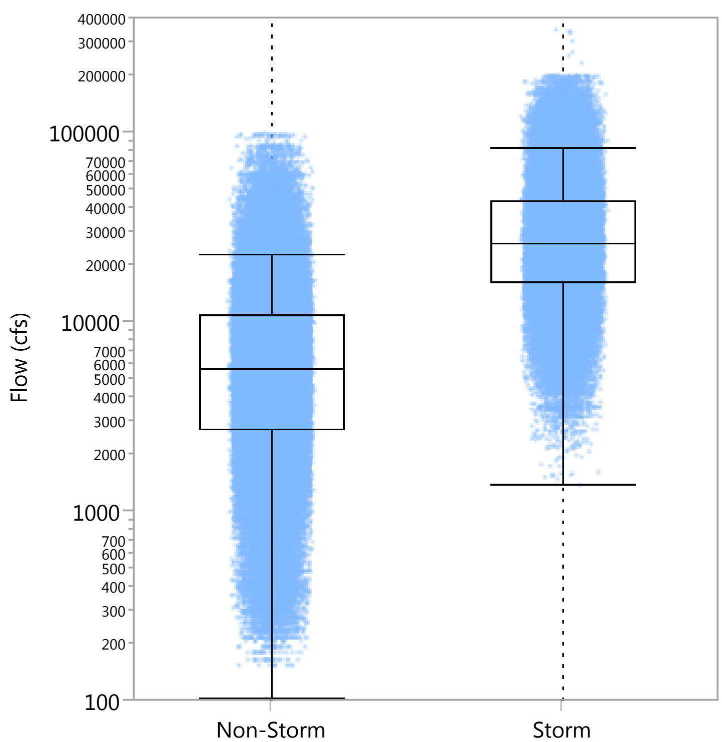


Figure 7. Distribution of flows observed in the Potomac River during the study period.

Figure 8. Study period and annual distribution of storm length. The median storm length for each year is marked in blue on the annual distribution chart on the right.

5.2 Fluvial Loads

5.2.1 Using the WRTDS software

As discussed in the Methods section, WRTDS is available from USGS as an R package. As an initial test of our abilities to utilize WRTDS, loads were computed using the data (both flow and water quality) available from the USGS for the study site. The flow data were retrieved automatically from the USGS repository using methods available in the EGRET R package (Station ID: 01646500 and Parameter: 00060 were used for retrieval). The water quality data were obtained from the USGS Virginia Water Science Center team (*personal communication*). Using these USGS datasets we were able to reproduce very closely the loads available from USGS Virginia Water Science Center, which gives us confidence that we performed the WRTDS load computation procedure correctly. Also, we found that the R package for WRTDS is very well designed and documented, making it easy for others to adopt the method.

5.2.2 Annual Load Comparisons

Annual loads (Calendar Year) computed by the direct methods (using OWML composite samples) are presented in Appendix D. These loads and several other aggregations such as USGS Water Year, Seasonal Loads, Daily Loads, and Monthly Loads for all parameters analyzed may be obtained from https://testserver.owml.vt.edu/mwcog/ (Access Date: 6/3/2014). The website also allows easy visualization of these load values. This website is a test server. Its continued existence will depend on decisions taken by COG stakeholders and staff in the near future. In the event that the decision is to not extend the website, we will create an appendix to this report with many plots considered suitable for use by others.

Simple comparison of annual load fluxes (Figures 9-15) show that there are some differences between the annual fluxes computed by the direct method (DL) and loads computed by the WRTDS method (WL) for most years. The overall trends, however, are similar. The fluxes computed using WRTDS with two different datasets (WL-O and WL-U) were also found to be much closer. Note that in Figure 15 Suspended Sediment (SS) load for WL-U is plotted along with Total Suspended Solid (TSS) loads for DL and WL-O. SS loads area presented for WL-U as TSS loads were not available. There is evidence that TSS and SS loads are not interchangeable [*Gray et al.*, 2000]. However, they may be expected to behave similarly. Further analysis of DL loads suggests that between 24-41 % of average annual load is carried on days with a storm event (Table 2). As expected, the percentage annual average load carried by storm event followed the order TSS>TP>NH_3-N>SRP>TSP>OXN>TN.

Table 2. Percent of annual load carried by storm event annually.

Year	Average Flow (cfs)	Percent load carrried by storm						
		SRP	TSP	TP	NH_3-N	OXN	TN	TSS
1983	1.48E+04	32%	31%	37%	35%	30%	32%	41%
1984	1.65E+04	37%	35%	39%	38%	30%	32%	43%
1985	1.16E+04	34%	31%	42%	44%	29%	35%	46%
1986	8.11E+03	30%	29%	38%	35%	25%	28%	43%
1987	1.12E+04	29%	29%	38%	38%	24%	28%	44%
1988	8.66E+03	28%	28%	38%	35%	18%	25%	43%
1989	1.21E+04	37%	34%	39%	41%	25%	29%	43%
1990	1.03E+04	32%	31%	39%	39%	22%	25%	43%
1991	9.18E+03	22%	20%	29%	27%	18%	20%	35%
1992	1.01E+04	22%	29%	38%	34%	20%	24%	40%
1993	1.74E+04	38%	36%	43%	42%	28%	31%	46%
1994	1.71E+04	32%	30%	36%	35%	21%	24%	40%
1995	9.38E+03	22%	23%	33%	26%	19%	22%	37%
1996	2.80E+04	39%	38%	44%	37%	29%	33%	44%
1997	1.02E+04	25%	23%	32%	30%	20%	22%	37%
1998	1.84E+04	35%	34%	38%	38%	29%	31%	41%
1999	6.55E+03	22%	18%	25%	28%	19%	19%	34%
2000	7.78E+03	18%	23%	29%	29%	17%	19%	37%
2001	7.23E+03	31%	26%	29%	31%	25%	24%	38%
2002	7.04E+03	23%	22%	31%	28%	25%	26%	37%
2003	2.61E+04	38%	38%	42%	42%	32%	35%	43%
2004	1.58E+04	33%	31%	37%	37%	24%	27%	41%
2005	1.08E+04	27%	28%	40%	29%	23%	27%	43%
2006	9.75E+03	31%	30%	35%	33%	23%	25%	41%
2007	8.63E+03	27%	27%	35%	34%	25%	27%	43%
2008	1.09E+04	25%	28%	34%	32%	26%	28%	43%
2009	1.03E+04	27%	27%	35%	31%	24%	27%	43%
2010	1.14E+04	29%	33%	40%	29%	24%	27%	45%
2011	1.67E+04	40%	39%	42%	37%	31%	33%	46%
2012	8.92E+03	32%	31%	34%	27%	18%	20%	42%
2013	1.18E+04	24%	26%	32%	23%	17%	20%	40%
Average	1.23E+04	30%	29%	36%	34%	24%	27%	41%

5.2.2.1 Comparing WRTDS load using two datasets

The WRTDS method may be used to compute both loads and flow normalized loads. The flow normalized load reduces the impact of flow variation on annual loads and may be useful for analyzing trends in flux without the impact of changes in flow conditions. Both loads and flow normalized loads for four water quality parameter SRP, TP, OXN, and TN computed using WRTDS with two different input data sets (WL-O and WL-U) were compared (Figures 16-19). It is evident from the figures that WRTDS with two different input datasets yielded similar annual load estimates. Among the four parameters analyzed, the most variation was found in SRP. For SRP a pattern in the difference was also noticeable, from negative differences in the early 1980s to positive differences after the 2000s. A statistically significant difference between WL-O and WL-U was found only for OXN. It may be seen from Figure 18 that the OXN flux estimated by WL-O is consistently higher than OXN flux computed by WL-U. Since the utility of the WRTDS method is largely in computation of long-term fluvial loads trends, the general agreement in the flow-normalized flux (Figure 16-18, Figure 28) obtained from two different data-sources (WL-U and WL-O), albeit using the same method, provides further confidence in the WRTDS methodology.

5.2.2.2 Comparing Direct Method Loads with WRTDS Load

Difference between fluxes computed by the direct method (DL) and the WRTDS method using the OWML data (WL-O) was more substantial in magnitude than between WL-O and WL-U. Nevertheless, fluxes estimated by WL-O were able to capture the annual variation (Figures 9-15). The observed percentage difference (WL-O–DL)/DL and magnitude of difference (WL-O–DL) is presented in Figures 20-26 and Table 3. It may be seen that the percentage difference varies based on the year and the parameter of interest. Annual loads prediction for TN and OXN by WL-O were closest to the observed loads estimated by the DL method; with less than ±25% difference. All other parameters have at least one year with greater than ±45% difference (Table 3).

Overall statistical testing of the 31 years of annual loads suggested that the differences were statistically significant (≠0) with > 95% confidence for TSP, TP, NH_3N, TN and TSS. The computed magnitude-of-difference, estimated using the Hodges–Lehmann estimator Δ is shown in Table 4, along with the 95% confidence bound. Using the best estimate for Δ and the median annual flux estimated within the study period it was computed that WRTDS is over predicting the TSP, TP, NH_3-N, TN and TSS fluxes by about 11%, 9%, 14%, 3%, and 12% respectively. These percentage differences are within the “typical scenario average” range computed by *Harmel et al.* [2006] using error propagation in various steps of sample collection and analyses, suggesting that the overall fluxes estimated over the 31-year study period by WL-O are reasonably similar to the DL given the uncertainty in sample collection and other analyses. However, it may be seen from Table 3 that for some years the percentage difference is much higher than the overall percentage difference observed for 31 years of data. Since annual fluxes are important for setting and monitoring water quality goals in the watershed, annual WL-O estimates should be used cautiously as they may be significantly higher or lower based on parameter and year being analyzed. Plots for percentage difference between WL-U and DL for four comparable parameters SRP, TP, OXN, and TN are shown in Appendix F and they also show yearly variation.

Further investigation is needed to understand the reasons for the observed higher percentage difference in some years. Mechanism of transport for the constituent coupled with the inadequacies in storm time loads predictions have been cited in the literature as some of the causes for poor performance of regression-based load estimation methods. However, comparison of daily flux differences for storm days and non-storms days (Figure 27) show that WL-O prediction for storm days are better than for non-storm days. On non-storm days, WL-O appears to be predicting significantly higher fluxes. A higher percentage of storm samples available in the OWML database used to calibrate the WL-O prediction equation may explain some of the over estimation at low base flows. However, further investigation of the prediction equation is needed to conclusively explain the cause.

It may be noted that daily loads difference are an important input to the water resources models, such as the Chesapeake Bay model, and large differences between daily loads observed between two load computation methods, and trends in the difference need further investigation. Appendix G presents some more daily load analysis.

5.2.2.3 Comparing different methods of flow-independent trends

Figure 28 and Table 5 summarize the trends observed from different methods. All methods used in this study to compute flow-independent trends show an essentially decreasing trend, with no statistically significant increasing trend. In the study period, statistically significant decreasing trends were observed for SRP, TSP and NH_3-N by all methods. For other parameters (TP, OXN, TN, and TSS) decreasing overall trends were observed, but trends by some methods were not statistically significant from zero (Table 5). Magnitude of trends, expressed in [kg/day]/year (Table 5) were however not similar for all five methods. Note that the magnitude of trend may be multiplied by 365 to approximate the annual change in load of the constituent in kg.

The magnitude of annual reduction is higher (though not always statistically significant) in the period before 1990, compared to any other later period for all parameters and all methods used in this study. In the decade 1990-1999, magnitudes of change in flow-independent loads are statistically insignificant (no different from zero), and a few have positive slopes, for all parameters and methods except TN measured by WL-O and WL-U, and SRP measured by WL-O. After 1999, magnitudes of the rate of change in flow-independent flux seem to be better (more decline in flux) than what was observed in 1990-1999. This study did not investigate the cause of the apparent decadal variation in detail.

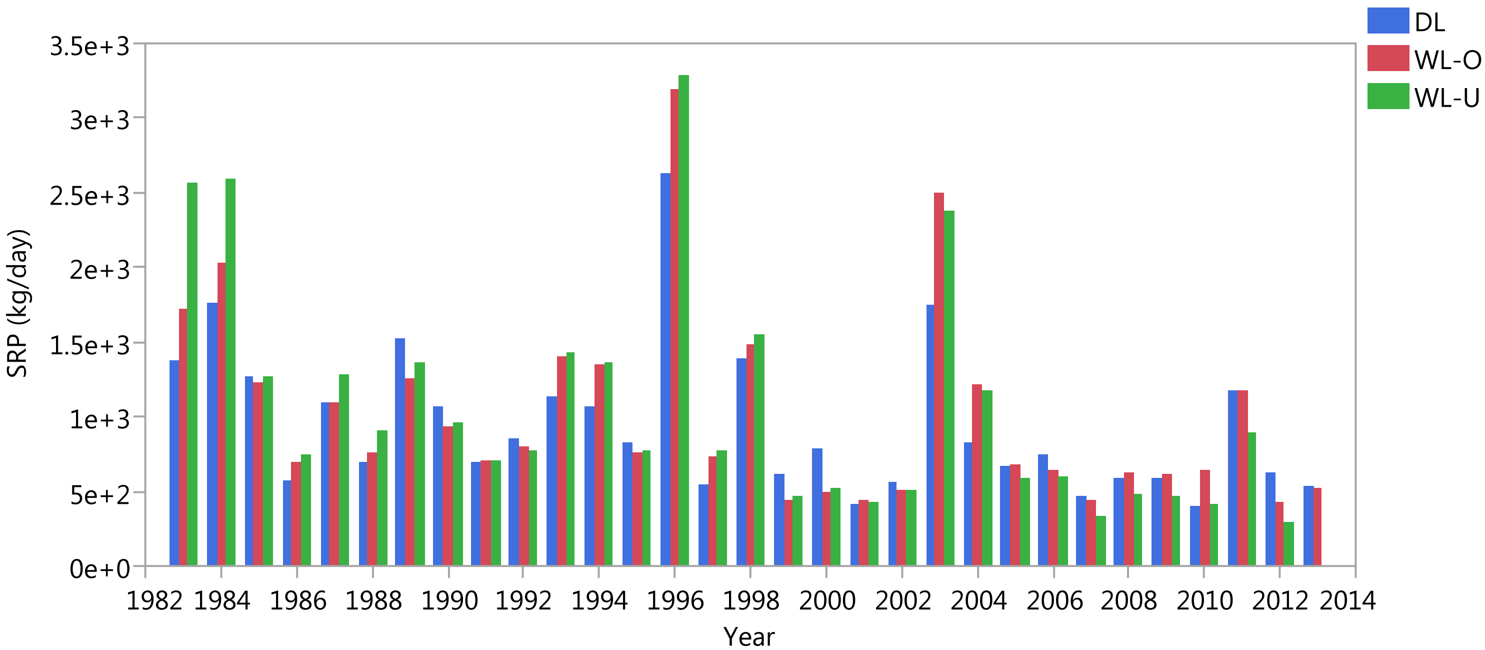


Figure 9. Annual SRP flux computed using direct method (DL) compared with SRP flux estimated by WRTDS using OWML (WL-O) and USGS (WL-U) datasets.

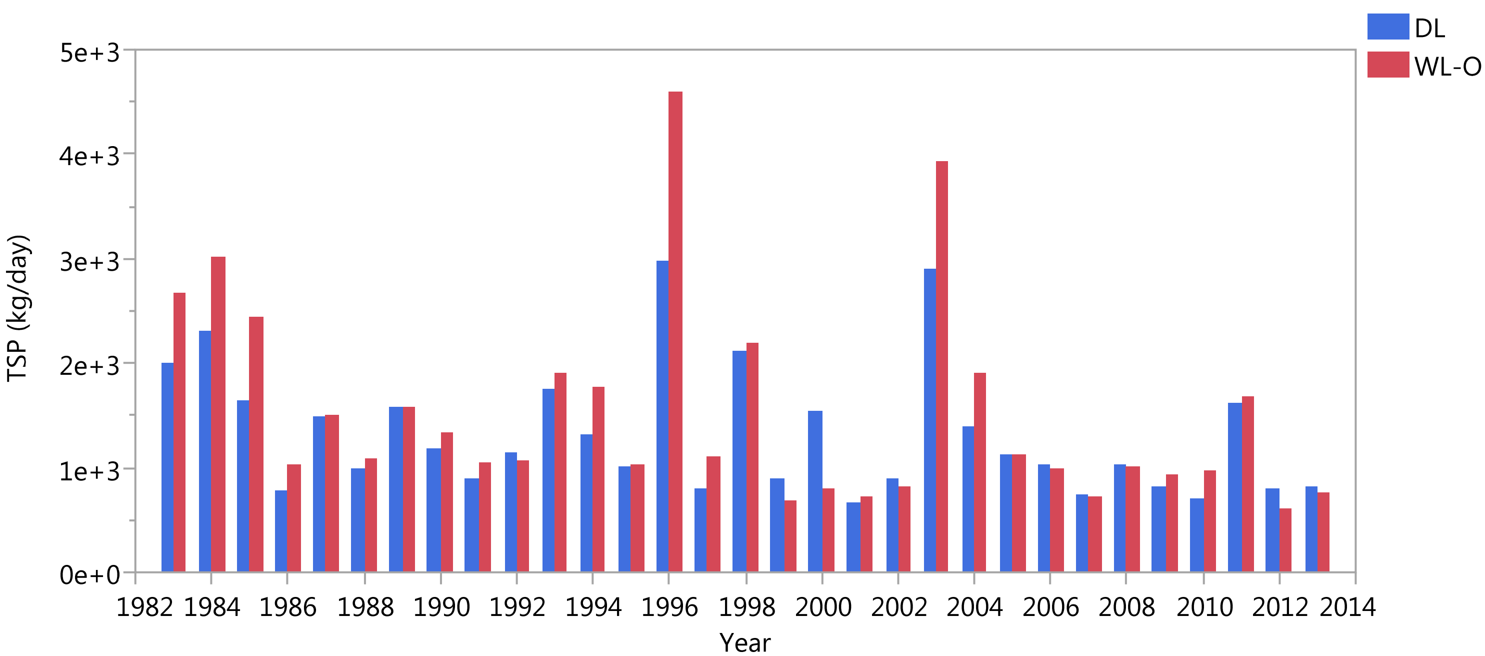


Figure 10. Annual TSP flux computed using direct method (DL) compared with TSP flux estimated by WRTDS using OWML (WL-O) datasets. Note that USGS does not record TSP at Potomac River.

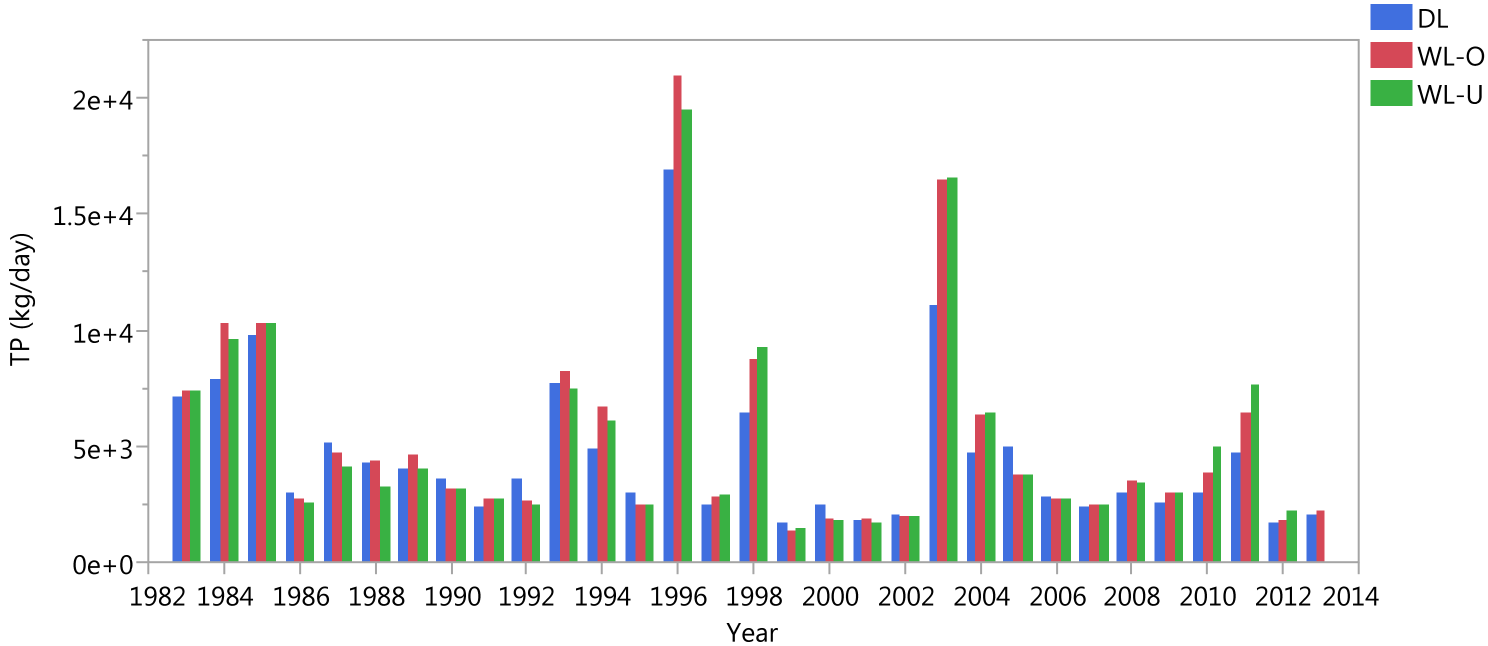


Figure 11. Annual TP flux computed using direct method (DL) compared with TP flux estimated by WRTDS using OWML (WL-O) and USGS (WL-U) datasets.

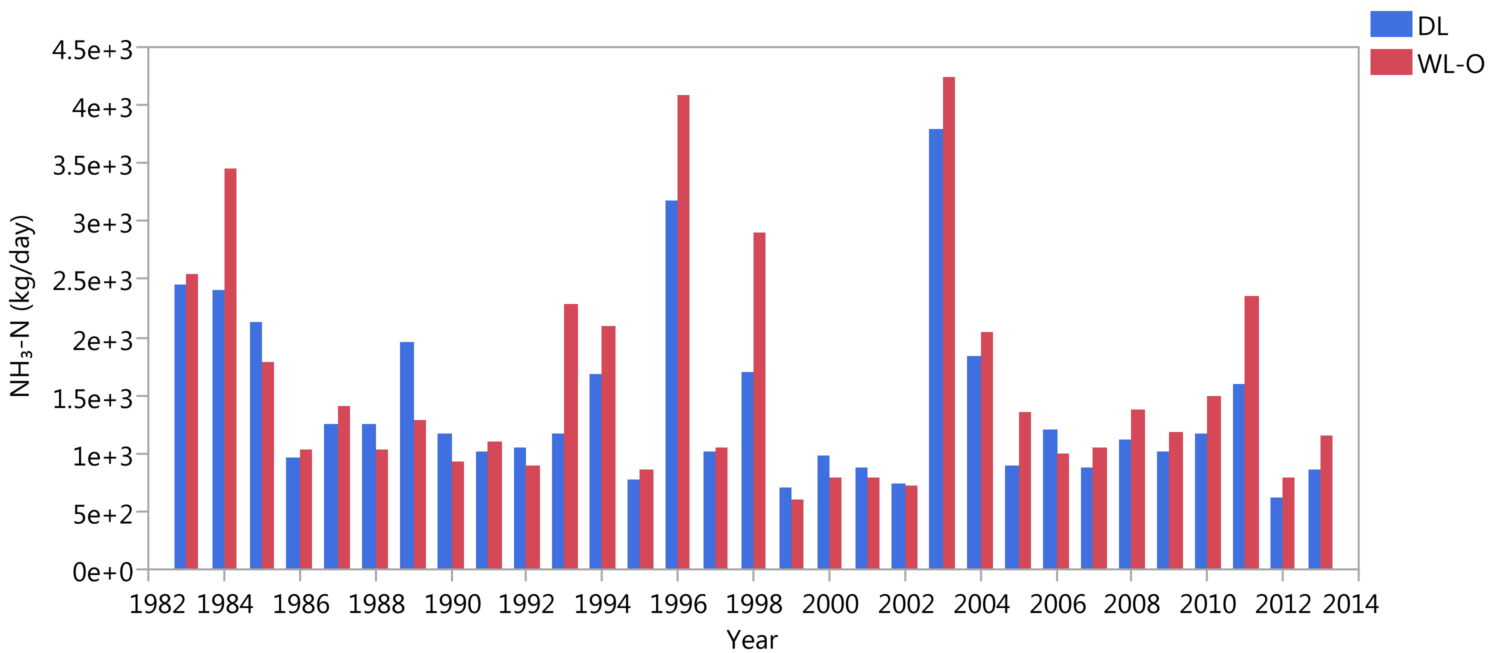


Figure 12. Annual Ammonia-N flux computed using direct method (DL) compared with Ammonia-N flux estimated by WRTDS using OWML datasets. Note that USGS does not record Ammonia-N.

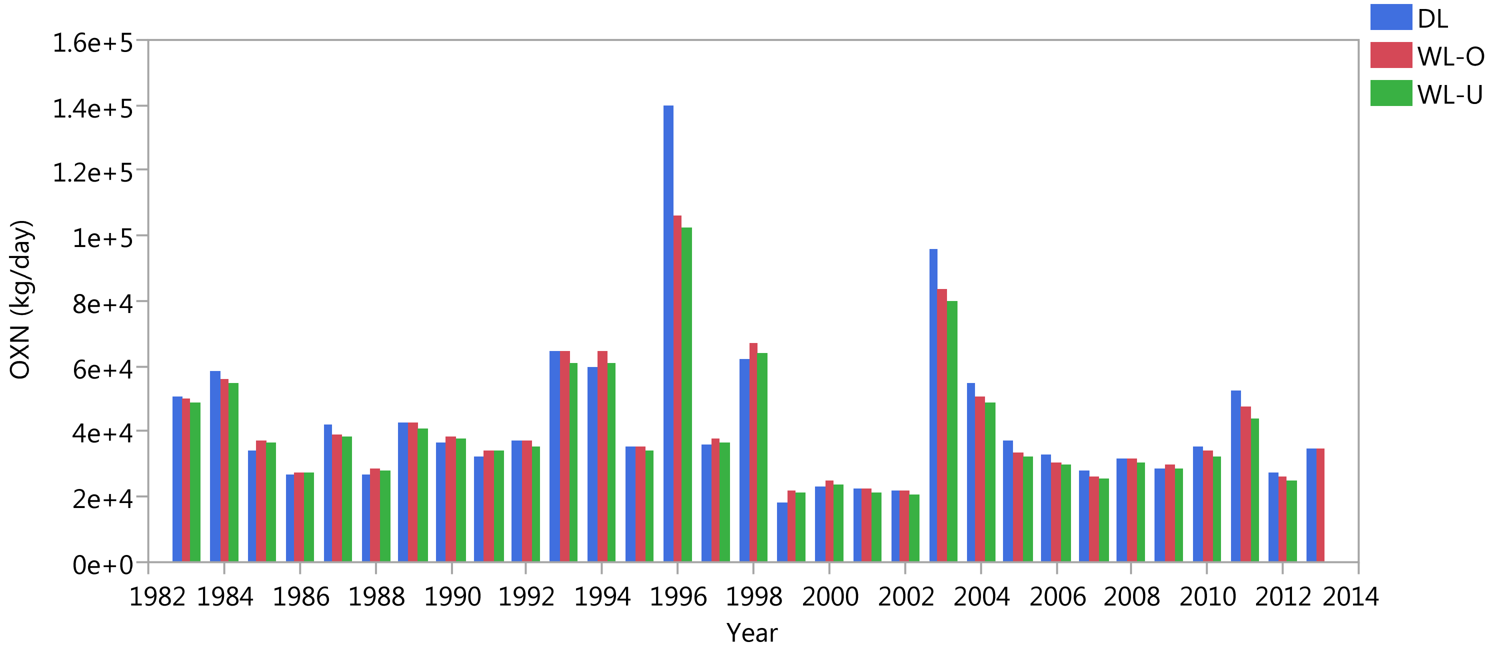


Figure 13. Annual oxidized nitrogen flux computed using direct method (DL) compared with OXN flux estimated by WRTDS using OWML (WL-O) and USGS (WL-U) datasets.

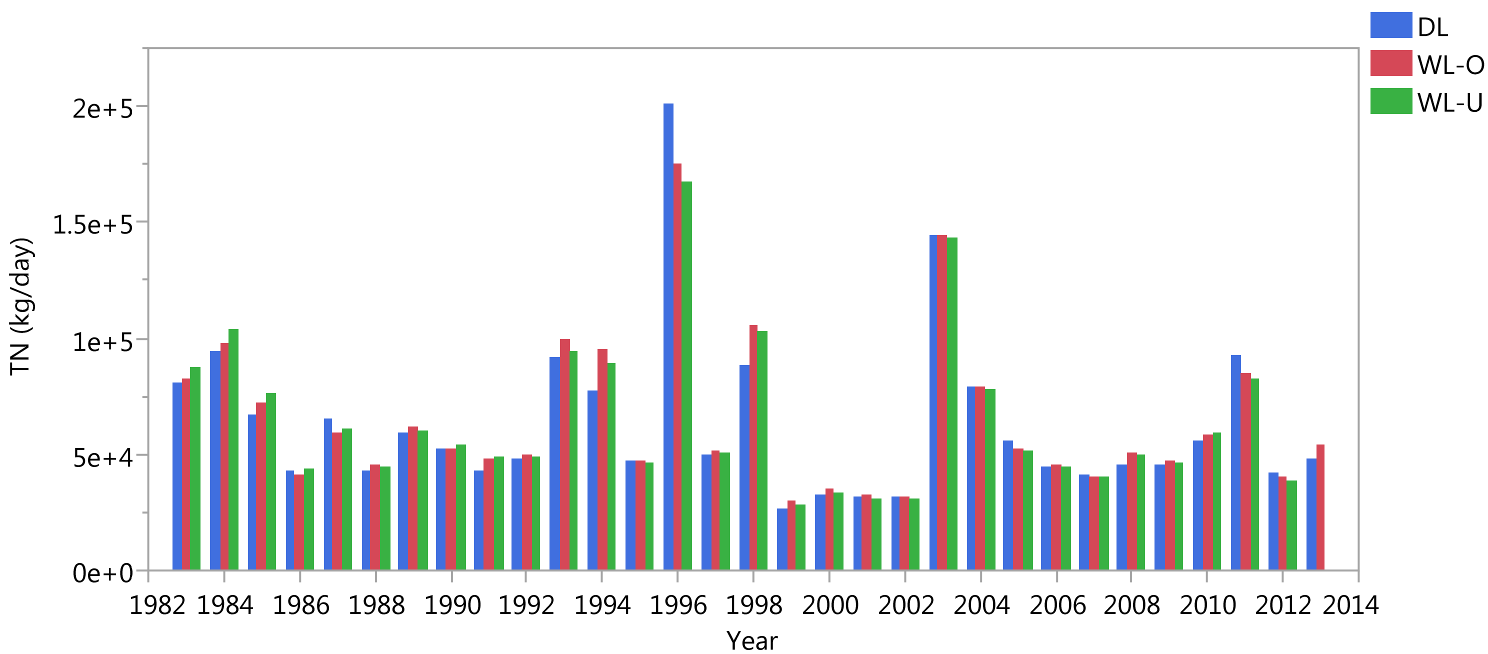


Figure 14. Annual total nitrogen flux computed using direct method (DL) compared with total nitrogen flux estimated by WRTDS using OWML (WL-O) and USGS (WL-U) datasets.

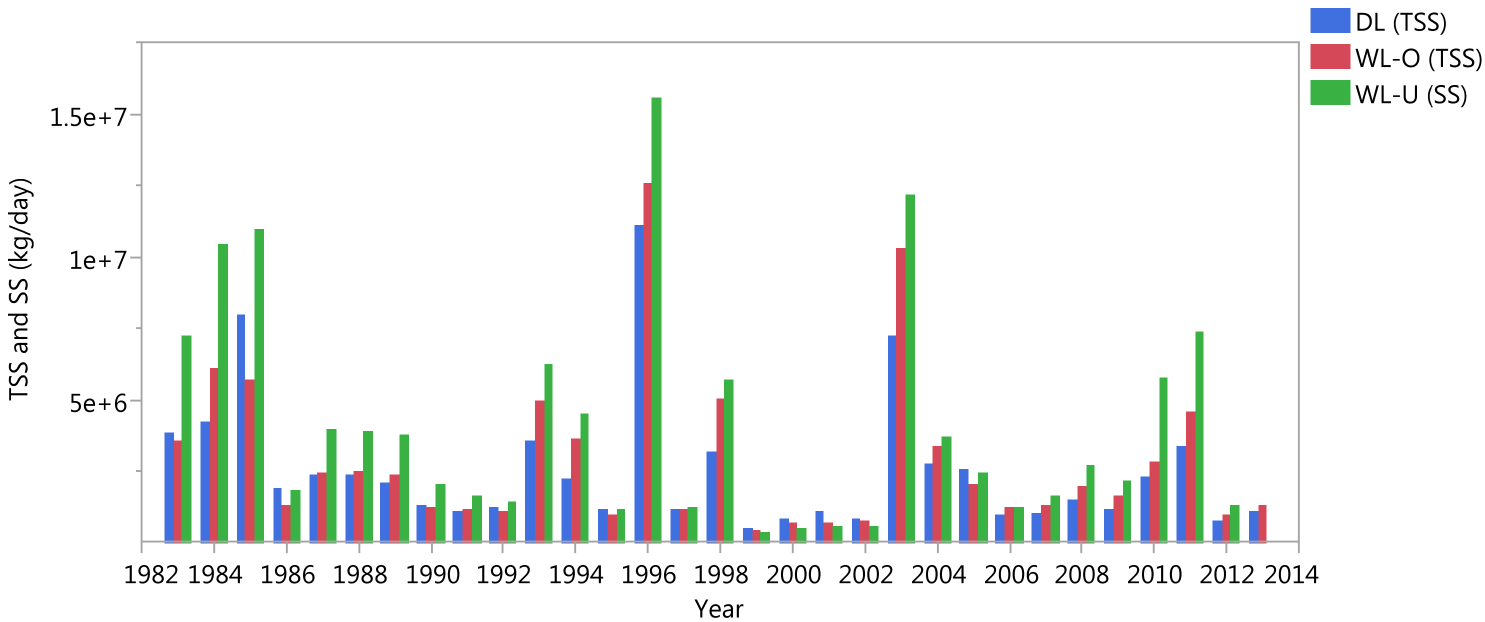


Figure 15. Annual Total Suspended Solid (TSS) flux computed using direct method (DL) along with TSS flux estimated by WRTDS using OWML (WL-O) and Suspended Sediment (SS) flux estimated using USGS (WL-U) datasets. Note that TSS and SS are not interchangeable.

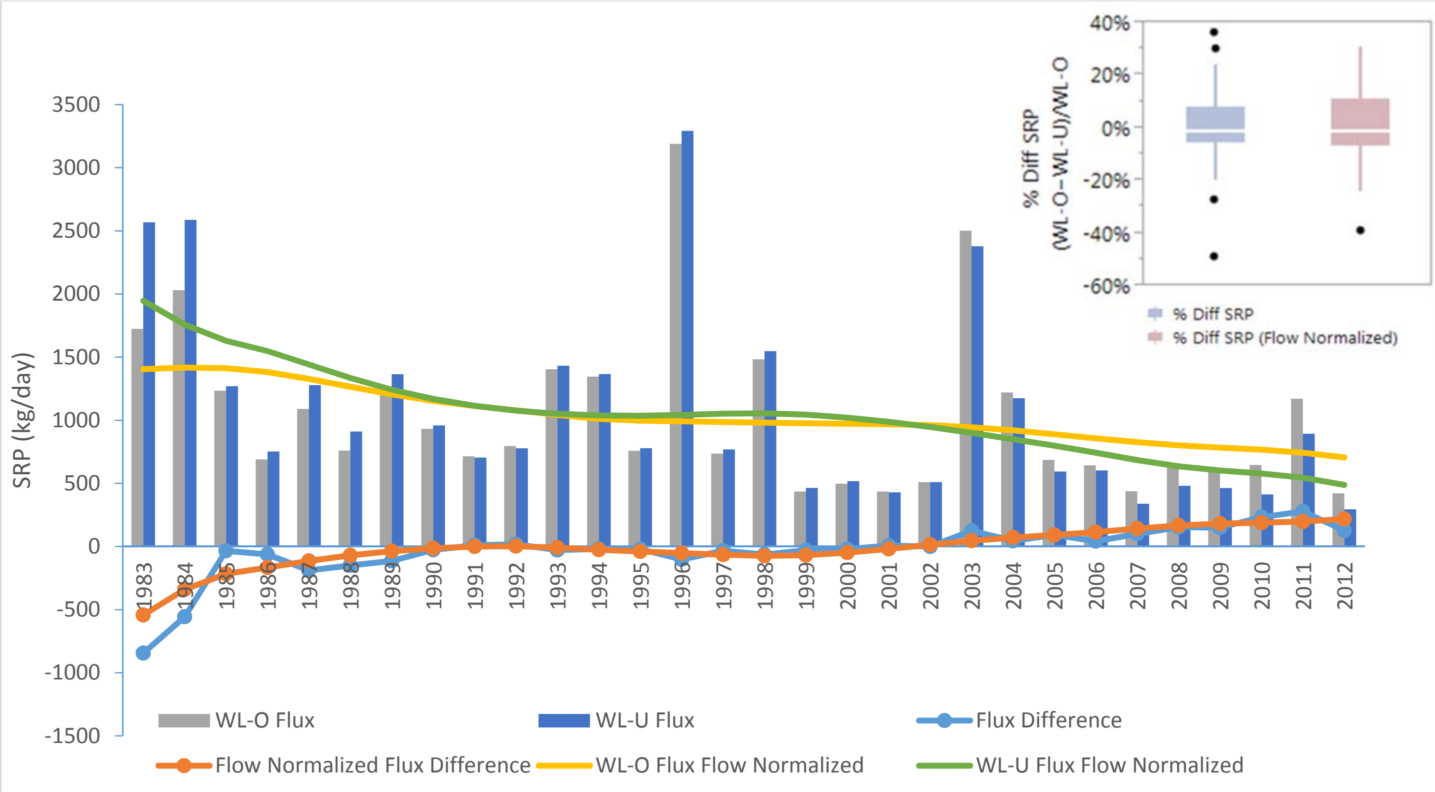


Figure 16. Comparison of SRP fluxes computed by WRTDS using OWML data (WL-O) and WRTDS using USGS data (WL-U).

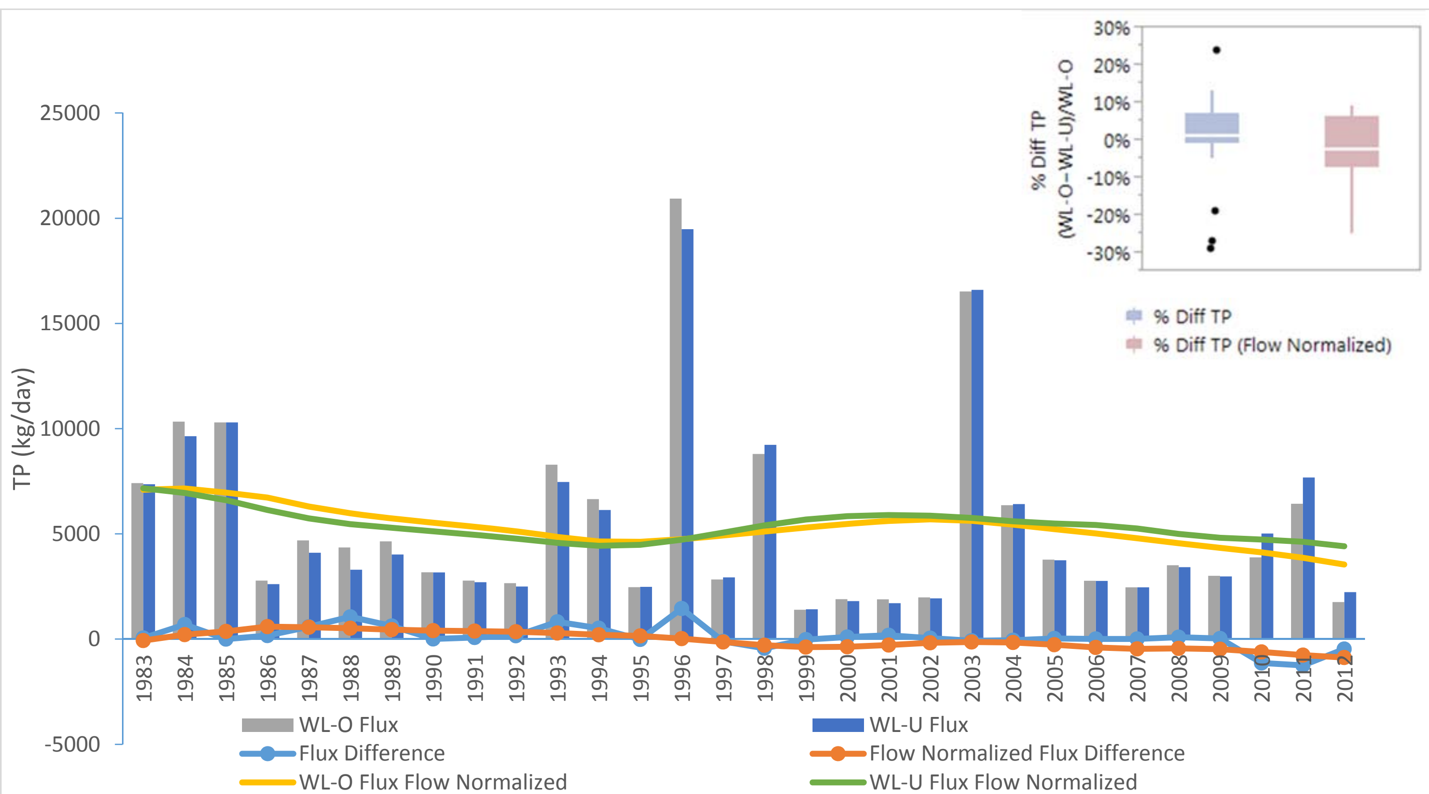


Figure 17. Comparison of TP fluxes computed by WRTDS using OWML data (WL-O) and WRTDS using USGS data (WL-U).

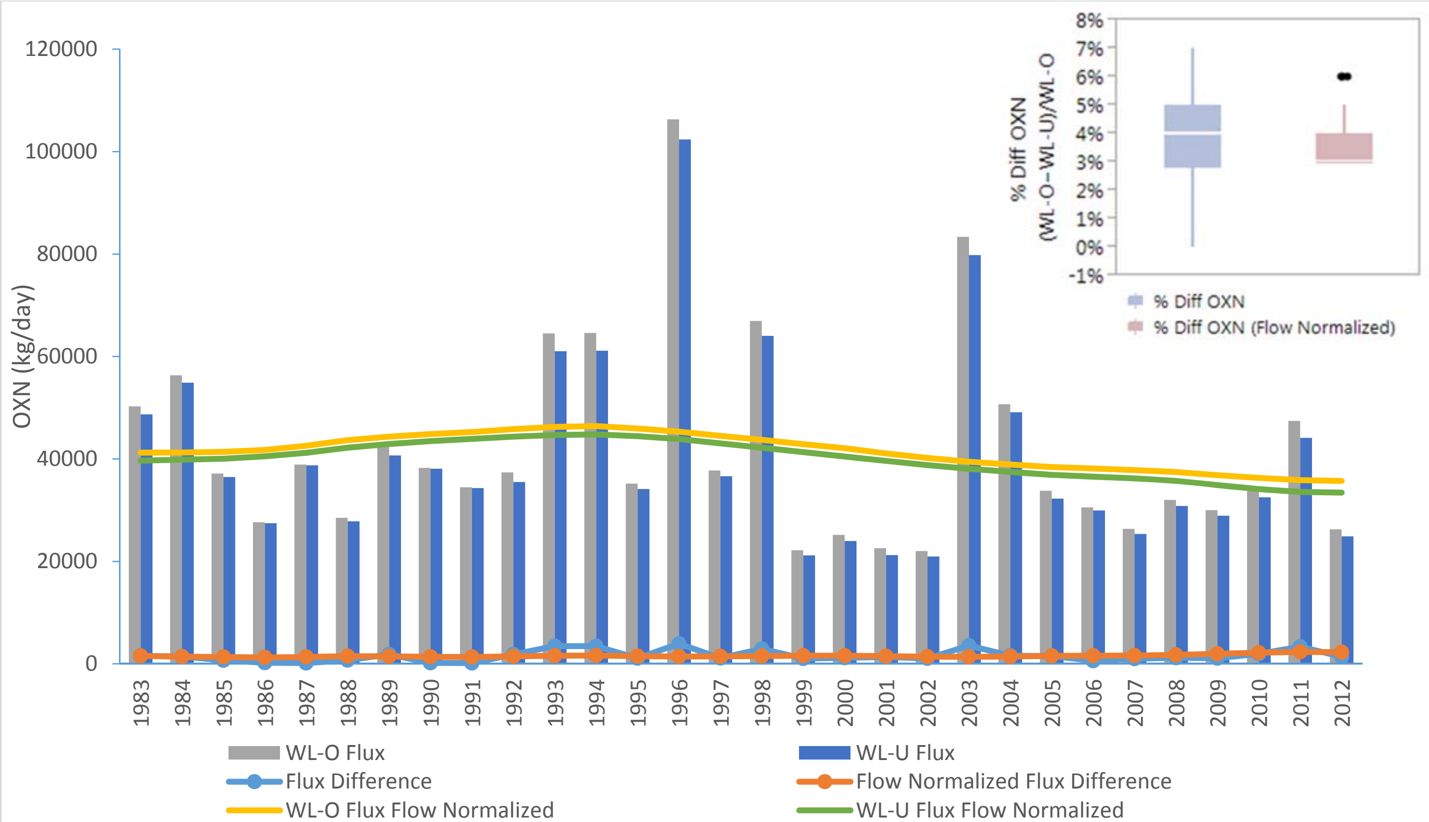


Figure 18. Comparison of oxidized nitrogen fluxes computed by WRTDS using OWML data (WL-O) and WRTDS using USGS data (WL-U).

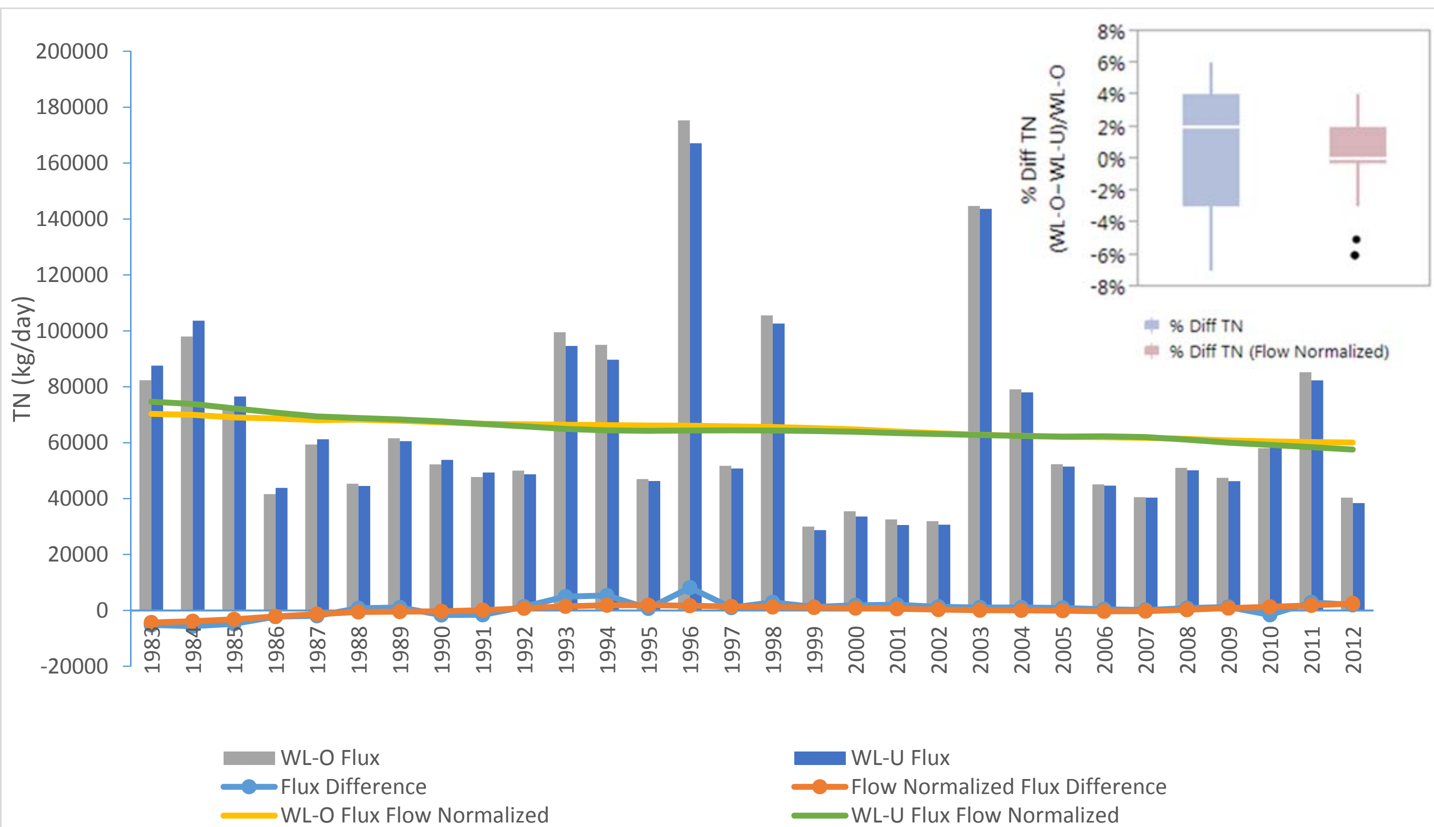


Figure 19. Comparison of TN fluxes computed by WRTDS using OWML data (WL-O) and WRTDS using USGS data (WL-U).

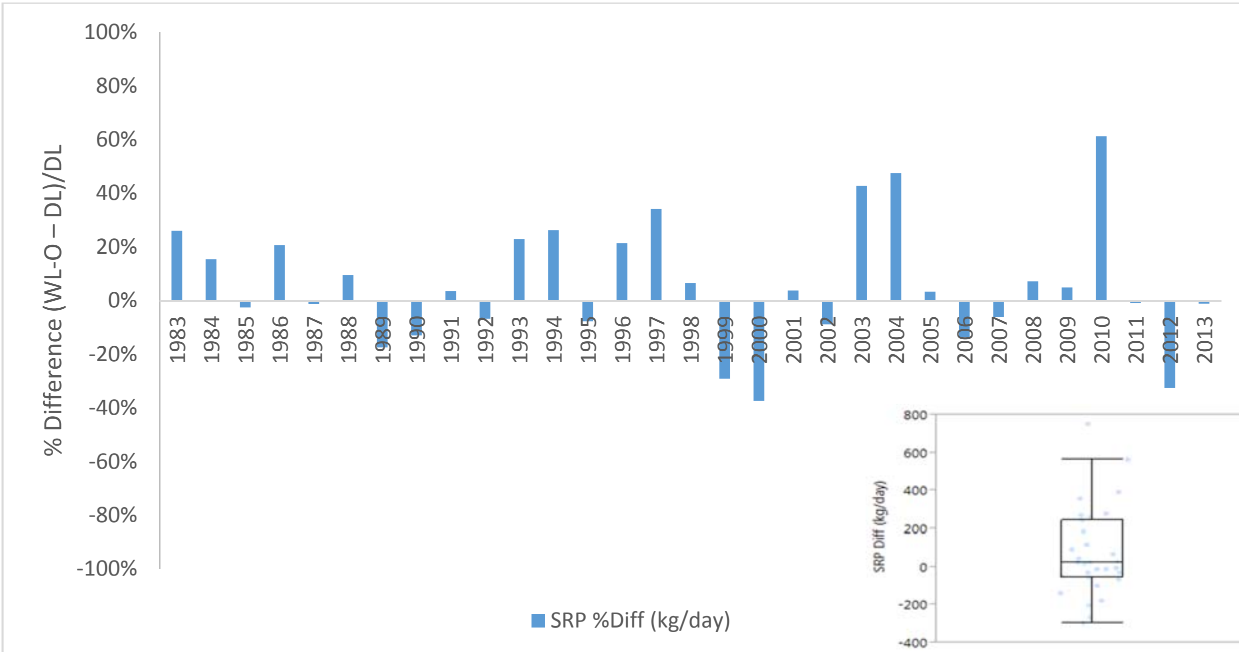


Figure 20. Difference between annual average SRP fluxes computed by direct method (DL) and WRTDS method using OWML data (WL-O).

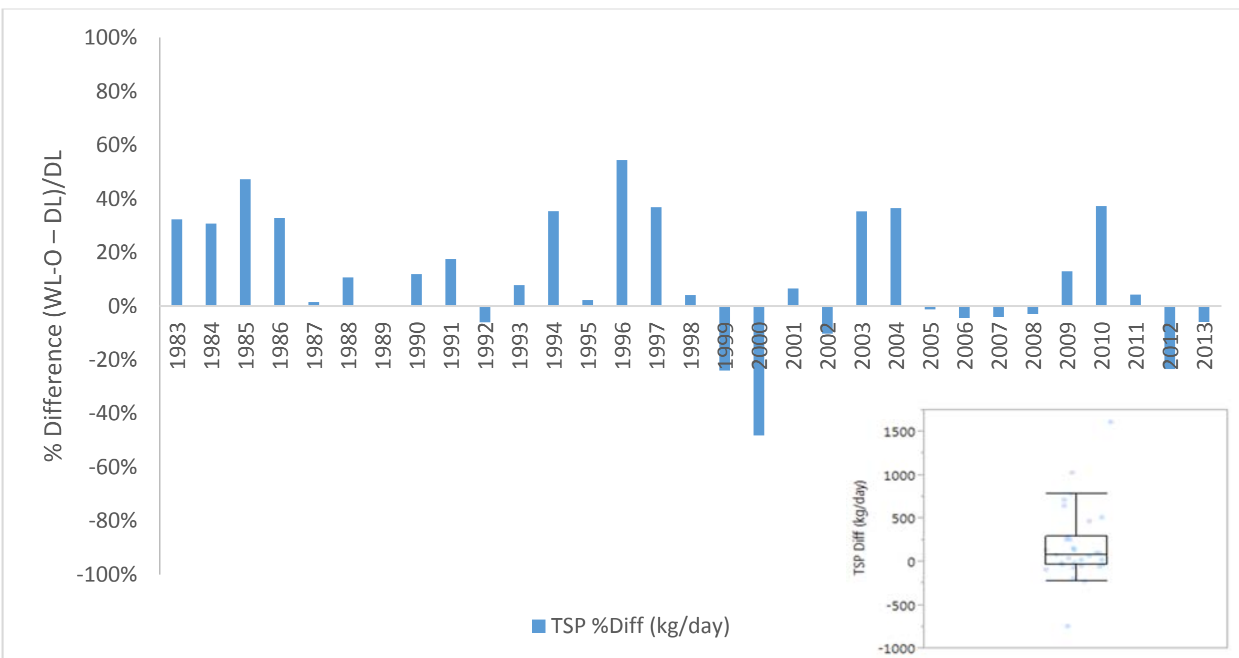


Figure 21. Difference between annual average TSP fluxes computed by direct method (DL) and WRTDS method using OWML data (WL-O).

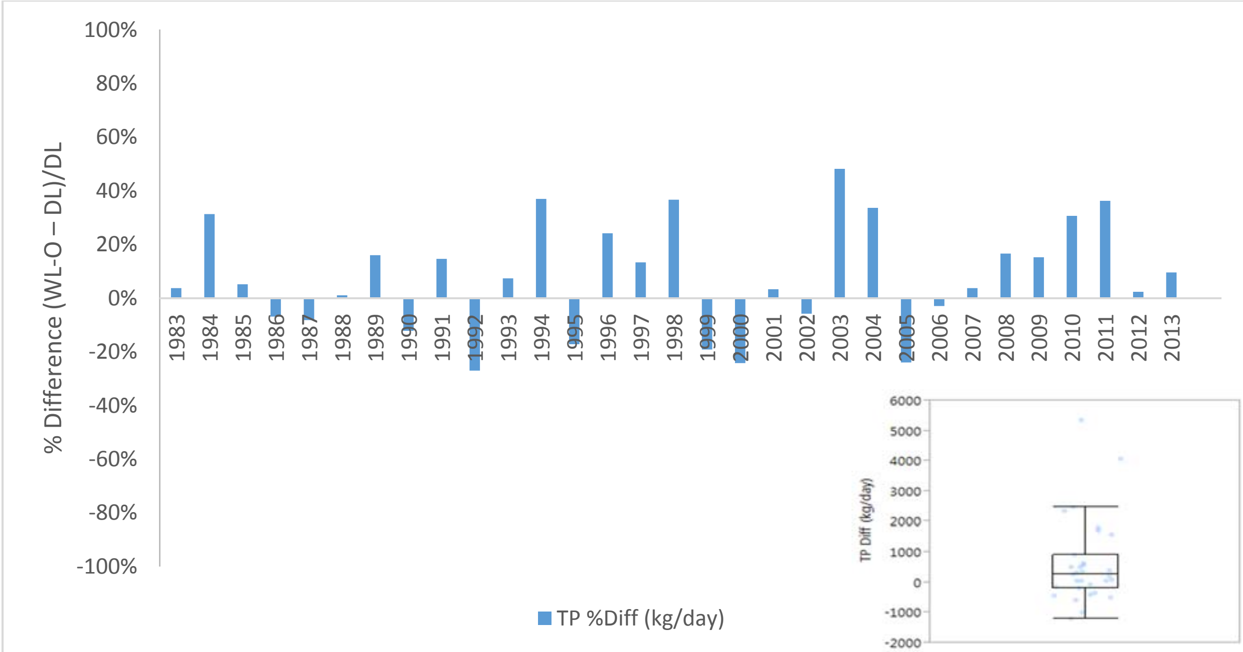


Figure 22. Difference between annual average TP fluxes computed by direct method (DL) and WRTDS method using OWML data (WL-O).

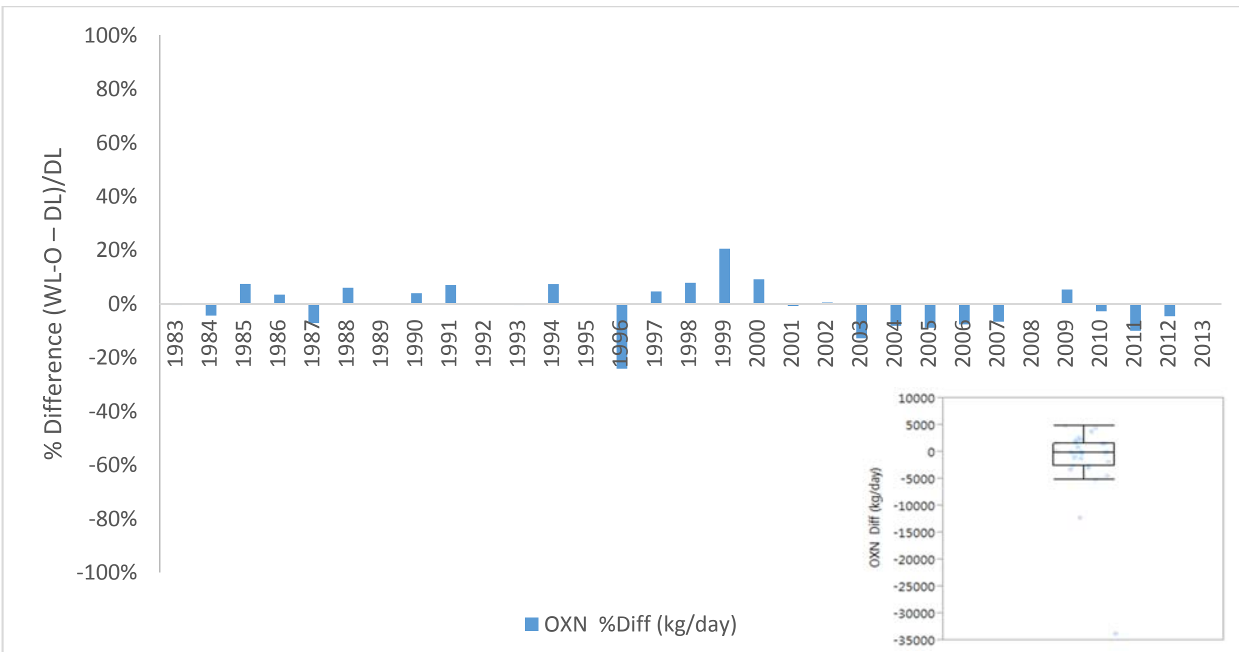


Figure 23. Difference between annual average OXN fluxes computed by direct method (DL) and WRTDS method using OWML data (WL-O).

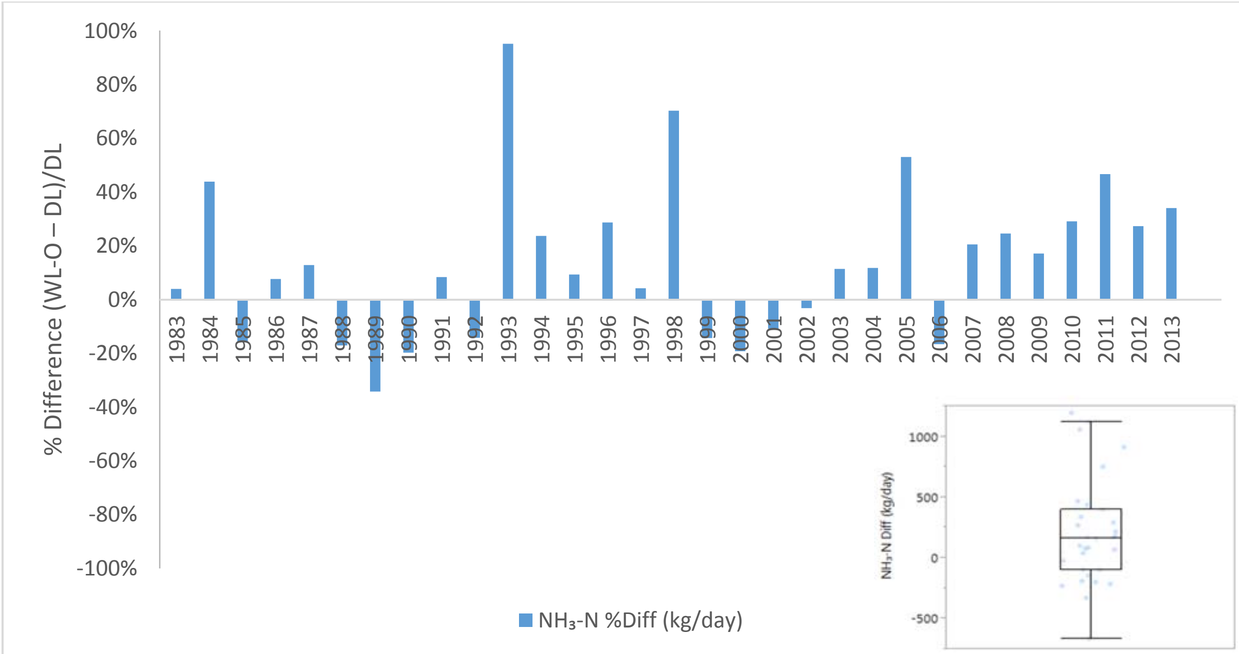


Figure 24. Difference between annual average NH_3-N fluxes computed by direct method (DL) and WRTDS method using OWML data (WL-O).

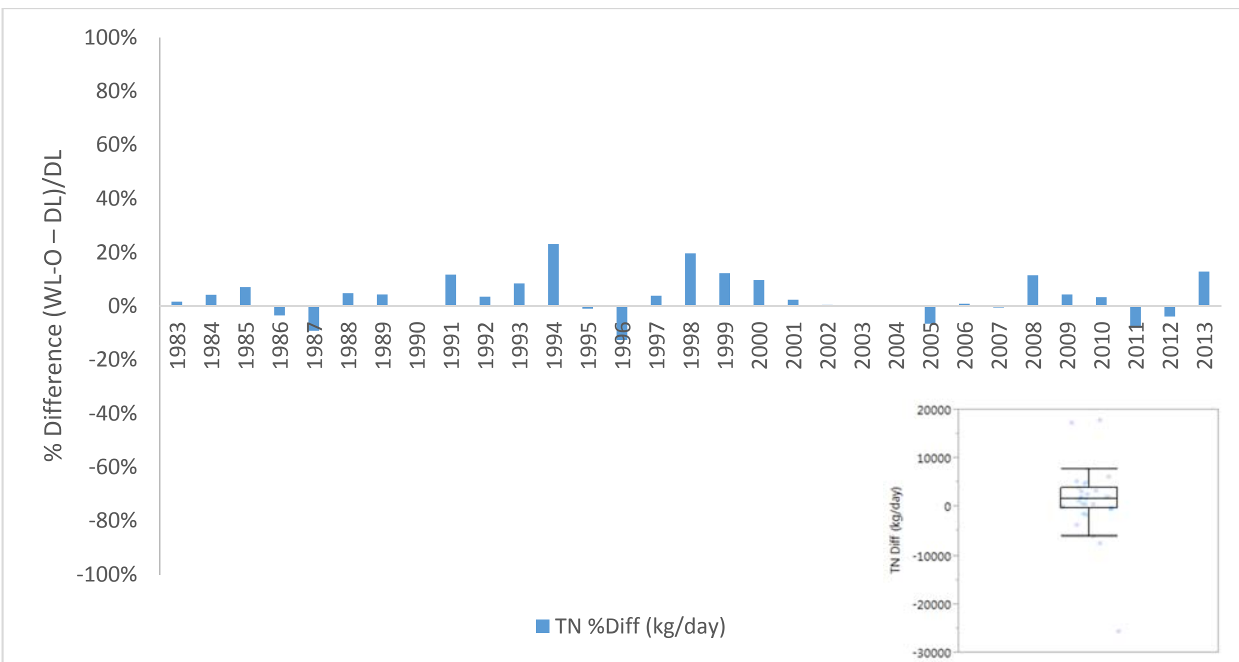


Figure 25. Difference between annual average TN fluxes computed by direct method (DL) and WRTDS method using OWML data (WL-O).

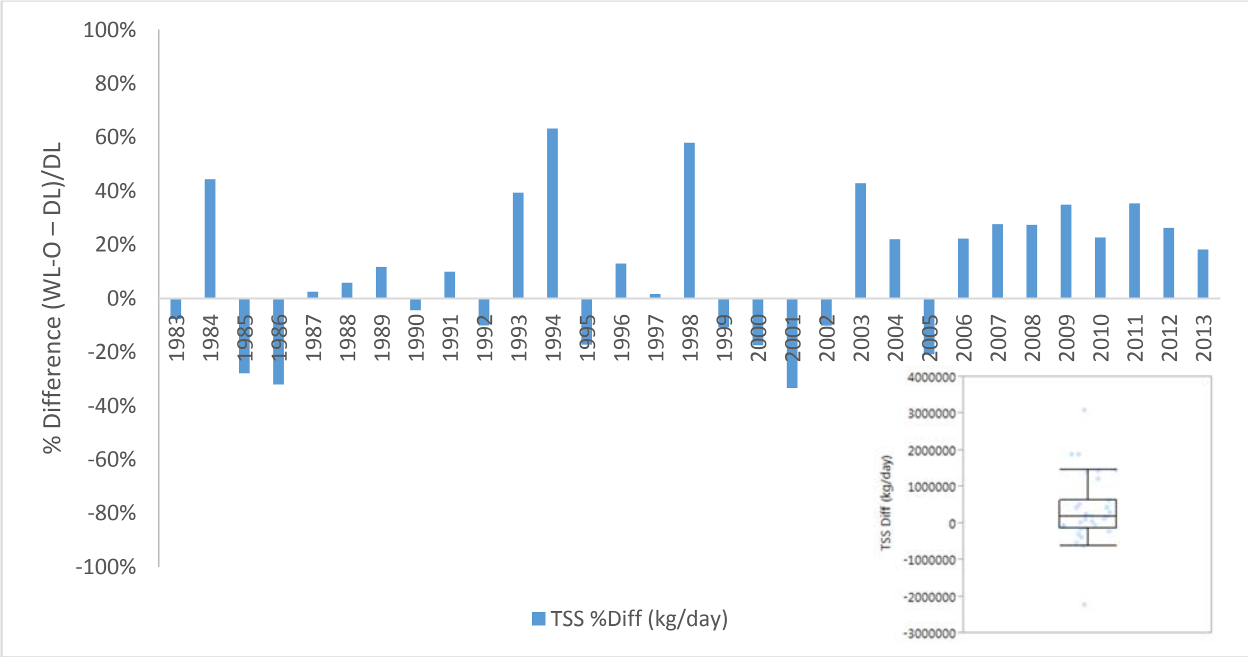


Figure 26. Difference between annual average TSS flux computed by direct method (DL) and WRTDS method using OWML data (WL-O).

Table 3. Percentage difference between WL-O and DL computed as (WL-O – DL)/DL. The colors visualize the magnitude of percentage difference. Red represent under prediction by WL-O and green is over prediction by WL-O when compared to DL. Note that the color scale intensity varies for each parameter: the highest and lowest differences in each parameter (shown at bottom of table) are assigned the same intensity.

Year	Percent difference (WL-O–DL)/DL						
	SRP	TSP	TP	NH_3-N	OXN	TN	TSS
1983	26%	32%	4%	4%	0%	2%	-8%
1984	15%	31%	31%	44%	-4%	4%	44%
1985	-3%	47%	5%	-16%	8%	7%	-28%
1986	21%	33%	-7%	8%	3%	-3%	-32%
1987	-1%	1%	-8%	13%	-7%	-9%	2%
1988	9%	11%	1%	-17%	6%	5%	6%
1989	-17%	0%	16%	-34%	0%	4%	12%
1990	-13%	12%	-12%	-20%	4%	0%	-4%
1991	3%	18%	15%	8%	7%	12%	10%
1992	-7%	-6%	-27%	-14%	0%	3%	-10%
1993	23%	8%	7%	95%	0%	8%	39%
1994	26%	35%	37%	24%	7%	23%	63%
1995	-8%	2%	-17%	9%	0%	-1%	-17%
1996	21%	54%	24%	29%	-24%	-13%	13%
1997	34%	37%	13%	4%	5%	4%	2%
1998	7%	4%	37%	70%	8%	20%	58%
1999	-29%	-24%	-19%	-14%	21%	12%	-11%
2000	-37%	-48%	-24%	-19%	9%	10%	-17%
2001	4%	7%	3%	-11%	-1%	2%	-33%
2002	-9%	-10%	-6%	-3%	1%	0%	-10%
2003	43%	35%	48%	11%	-13%	0%	43%
2004	48%	37%	34%	12%	-8%	0%	22%
2005	3%	-1%	-24%	53%	-9%	-7%	-21%
2006	-14%	-4%	-3%	-17%	-8%	1%	22%
2007	-6%	-4%	4%	21%	-7%	-1%	28%
2008	7%	-3%	17%	25%	0%	11%	27%
2009	5%	13%	15%	17%	5%	4%	35%
2010	61%	37%	31%	29%	-3%	3%	23%
2011	-1%	4%	36%	47%	-10%	-8%	35%
2012	-33%	-24%	2%	27%	-5%	-4%	26%
2013	-1%	-6%	10%	34%	0%	13%	18%
Largest overprediction	61%	54%	48%	95%	21%	23%	63%
Largest underprediction	-37%	-48%	-27%	-34%	-24%	-13%	-33%

Table 4. Magnitude of difference between DL and WL-O.

Parameter	Hodges–Lehmann Δ and 95% Confidence Interval (kg/day)
SRP	42.4 (-20, 142.45)
TSP	**121.6 (24.30, 303.05)***
TP	**336.7 (30.5, 932.5)***
NH_3-N	**156.8 (29.5, 336.5)***
OXN	-141.9 (-1420.0, 799.1)
TN	**1518.1 (105, 2986)***
TSS	**230752 (7500, 676000)***

* significant at 95%

Table 5. Annual rate of change for annually overaged daily loading ([kg/day]/year) computed using several methods.

	DL-MA-285 days	DL-MA-2 years	DL-Residual	WL-O	WL-U
			SRP		
<1990	-53	-79	-10	-38*	-103*
1990-1999	-14	-1	-18	-18*	-8
>1999	-10	-5	-18*	-25*	-41*
All	-19*	-18*	-18*	-23*	-33*
			TSP		
<1990	-126	-82*	-87	-91*	
1990-1999	27	29	6	-12	
>1999	-34	-28	-50*	-53*	
All	-17*	-15*	-21*	-35*	
			TP		
<1990	-693	-846*	-280	-279*	-542*
1990-1999	28	183	-118	-33	73
>1999	-52	-143	-143*	-215*	-132*
All	-100*	-95*	-87*	-92*	-37
			NH_3-N		
<1990	-124	-171	-37	-99*	
1990-1999	-4	21	-7	6	
>1999	-14	-18	-35*	-25*	
All	-15*	-17*	-14*	-10*	
			OXN		
<1990	-1476	-1452	124	588*	1193*
1990-1999	1902	2769	-483	-263	-273
>1999	693	434	-243	-457*	-600*
All	-132	-112	-176*	-383*	-493*
			TN		
<1990	-4926	-4432*	-1503	-427*	-1368*
1990-1999	2414	3323	-532	-192*	-81*
>1999	1292	839	-140	-359*	-597*
All	-372	-294	-315*	-344*	-320*
			TSS		
<1990	-565128	-554958	-171928	-128085*	
1990-1999	26433	155010	-32662	-3568	
>1999	-6436	-115	-32652	-75854*	
All	-49921*	-49493	-34213*	-30191*	

* significant at 95%

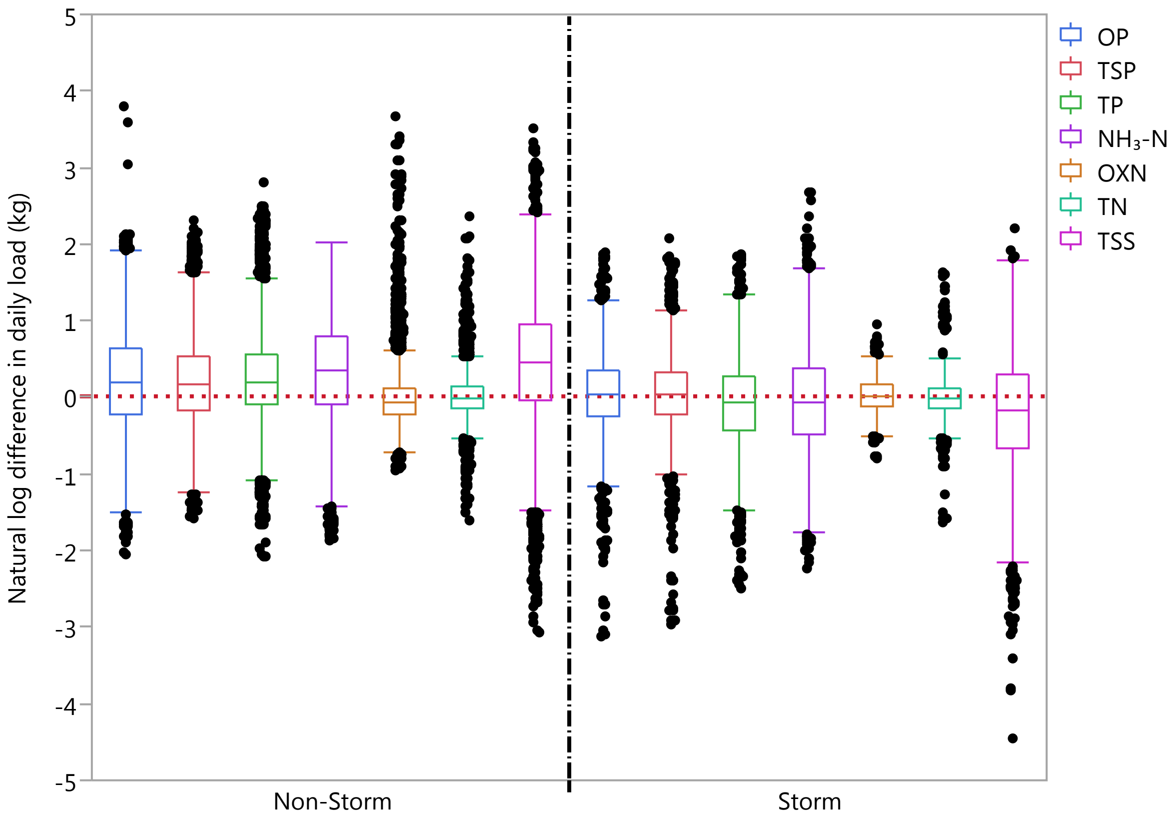


Figure 27. Difference in daily load computed as ln(WL-O)–ln(DL) between days with storm events (Storm, right) and days without storm (Non-Storm, left) for the study period.

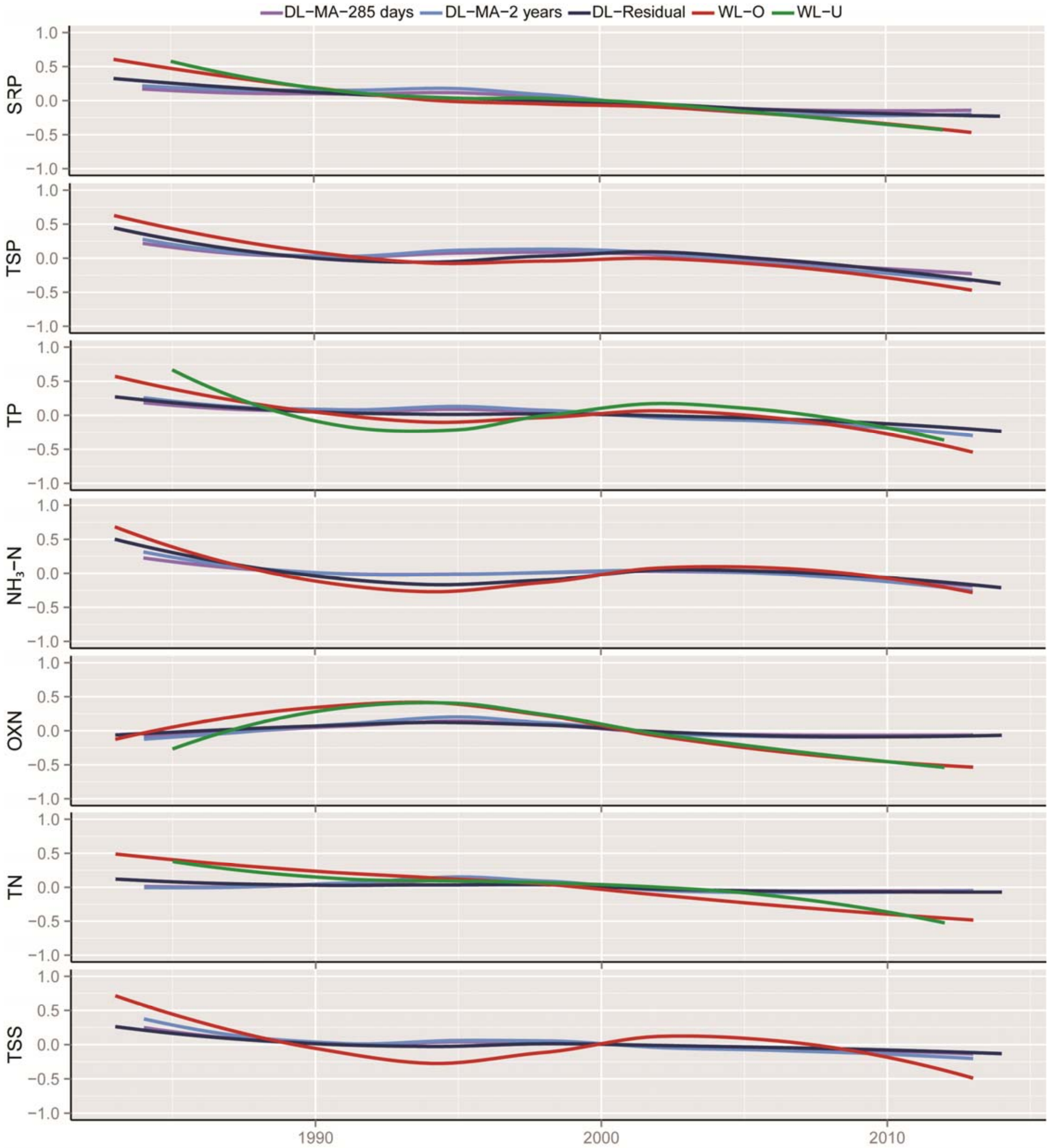


Figure 28. Flow-independent trends estimated by various methods. DL-MA-285 days is the trend line estimated for DL with a moving average window width of 285 days. DL-MA-2 years is the trend line estimated for DL with a moving average window width of 2 years. DL-Residual is the trend estimated from residuals after adjusting for streamflow variations.

6 CONCLUSIONS AND FUTURE INVESTIGATIONS

Over 30 year of water resources data, both quality and quantity, are available for the Potomac River. These data include weekly (bi-weekly in winter) non-storm sampling augmented by flow-composite and discrete storm sampling, and high-frequency flow measurement. Thus, the fluvial loads computed using these data are expected to be excellent estimates of the "true load" in the river. In addition, the raw data analysis revealed that there is significant non-storm and discrete storm data to drive the WRTDS method for load estimation and enable comparison between the two load computation/estimation methods.

Comparison of loads estimated by WRTDS using the two datasets, OWML (O) and USGS (U), were very similar for all parameters. The flow-normalized trends were also found to be similar. However, for OXN, although the flow-normalized and non-flow-normalized flux trends were similar, a consistent over-prediction by WL-O when compared with WL-U was observed. More analysis is needed to ascertain the cause of this consistent difference. For SRP, another trend was observed where in the 1980s the WL-U was overestimating flux when compared to WL-O, and starting in the 2000s WL-U started underestimating SRP flux. Thus, the difference between the two (WL-O−WL-U) gradually moved from negative in the early part of the study period to positive in the later years.

Comparison of WL-O with DL suggested that for aggregation over long periods (31 years analyzed) the WRTDS method does a good job of predicting fluxes. This makes it suitable for long-term trend analysis. However, if annual loads are required, such as for setting annual load allocations and targets, the WRTDS method should be used with caution as, depending on the year and parameter, varying relative differences were observed (see Table 3), some of which were quite significant in magnitude.

For individual parameters, TN and OXN loads were found to be better estimated by the WRTDS method than TP, SRP, TSP, NH_3N, and TSS loads. There is evidence from segregating daily loads among storm and non-storm days that the WRTDS method is doing a better job for storm-days and over predicting flux for non-storm days. This observation suggests that higher weights for the non-storm time samples may be the reason for the large over-prediction by WRTDS method. However, further analysis of the prediction equation generated and the half-widths used to assign weights for calibrating WRTDS methods is needed to ascertain the cause of this difference.

Overall, flow-independent trends computed by the five methods suggest a general decline in the flux entering Chesapeake Bay from the Potomac River. For SRP, TSP, and NH_3-N the decline is statistically significant for all methods used. For TP, OXN, TN, and TSS a majority of methods used in this study (> 3 out of 5) also show a statistically significant declining trend. There is some difference in the magnitude of decline when partitioned by decades. Before 1990 higher magnitudes of decline were observed, from 1990-1999 there was no significant decline, and after 1990 a more robust decline (compared to 1990-1999) is observed. The reasons for this decadal variation need to be investigated in detail, but it may be speculated that the point-source control measures put in place before 1990 and wide-scale adoption of non-point control measures after 1999 may be the cause of decline in those era. However, this observation does suggest that very long term trends may not be a good estimate for future reduction that may be anticipated, and perhaps should not be used for planning purposes. These results are not unexpected as it is quite likely that the magnitude of annual reduction will decline after the easier-to-control measures have been adopted and further reduction needs significantly more investment. A follow on study may investigate the nature of diminishing returns in detail along with future projections for reduction that may be anticipated.

It is important to note that results presented here are for the Potomac River. It is reasonable to expect that WRTDS will perform similarly in other streams draining large watershed such as those for the Susquehanna and James Rivers. However, these results, as noticed with ESTIMATOR and other regression-based methods, may not necessarily apply to flashier streams draining smaller watersheds. Comparative studies for smaller watersheds are necessary to establish the application of WRTDS for smaller, flashier streams.

The Potomac River site at Chain Bridge is unique in the sense that it is the only site on a large tributary of the Chesapeake Bay where a single-point sampling scheme is possible, due to the nature of the well-mixed water at the constriction at Chain Bridge. For the Metropolitan Washington area, it is also the site whereby accurate loads computed by the direct method are indicative of the non-point loads arising from areas upstream of the urban/suburban Metro region. Thus, it can serve as a means to compare the loads generated by rural areas upstream of the Metro area with those loads generated by the Metro area, and provide an insight into appropriate control strategies.

7 References

Arabkhedri, M., F. Lai, I. Noor-Akma, and M. Mohamad-Roslan (2010), An application of adaptive cluster sampling for estimating total suspended sediment load, *Hydrol. Res.*, *41*(1), 63-73.

Ferguson, R. I. (1987), Accuracy and precision of methods for estimating river loads, *Earth Surf. Processes Landforms*, *12*(1), 95-104.

Gray, J. R., G. D. Glysson, L. M. Turcios, and G. E. Schwarz (2000), Comparability of Suspended-Sediment Concentration and Total Suspended Solids Data, Water-Resources Investigation Report 00-4191, U.S. Geological Survey, Reston, VA.

Harmel, R., R. Cooper, R. Slade, R. Haney, and J. Arnold (2006), Cumulative uncertainty in measured streamflow and water quality data for small watersheds, *Trans. ASABE*, *49*(3), 13.

Helsel, D., and R. Hirsch (2002), *Statistical Methods in Water Resources*, U.S. Geological Survey.

Hirsch, R. M., D. L. Moyer, and S. A. Archfield (2010), Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs, *J. Am. Water Resour. Assoc.*, *46*(5), 857-880, doi: 10.1111/j.1752-1688.2010.00482.x.

Horowitz, A. J. (2003), An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations, *Hydrol. Processes*, *17*(17), 3387-3409.

Kumar, S., A. N. Godrej, and T. J. Grizzard (2013), Watershed size effects on applicability of regression-based methods for fluvial loads estimation, *Water Resour. Res.*, *49*(11), 7698-7710, doi: 10.1002/2013WR013704.

Moyer, D. L., R. M. Hirsch, and K. E. Hyer (2013), Comparison of Two Regression-Based Approaches for Determining Nutrient and Sediment Fluxes and Trends in the Chesapeake Bay Watershed, U.S. Geological Survey Scientific Investigations Report 2012-5244, U.S. Geological Survey Richmond, VA.

Post, H. E., and T. J. Grizzard (1987), The monitoring of stream hydrology and quality using microcomputers, in *Symposium on Monitoring, Modleing and Mediating Water Quality*, pp. 199-207, American Water Research Association, Syracuse, New York, USA.

Robertson, D. M., and E. D. Roerish (1999), Influence of Various Water Quality Sampling Strategies on Load Estimates for Small Streams, *Water Resour. Res.*, *35*(12), 3747-3759.

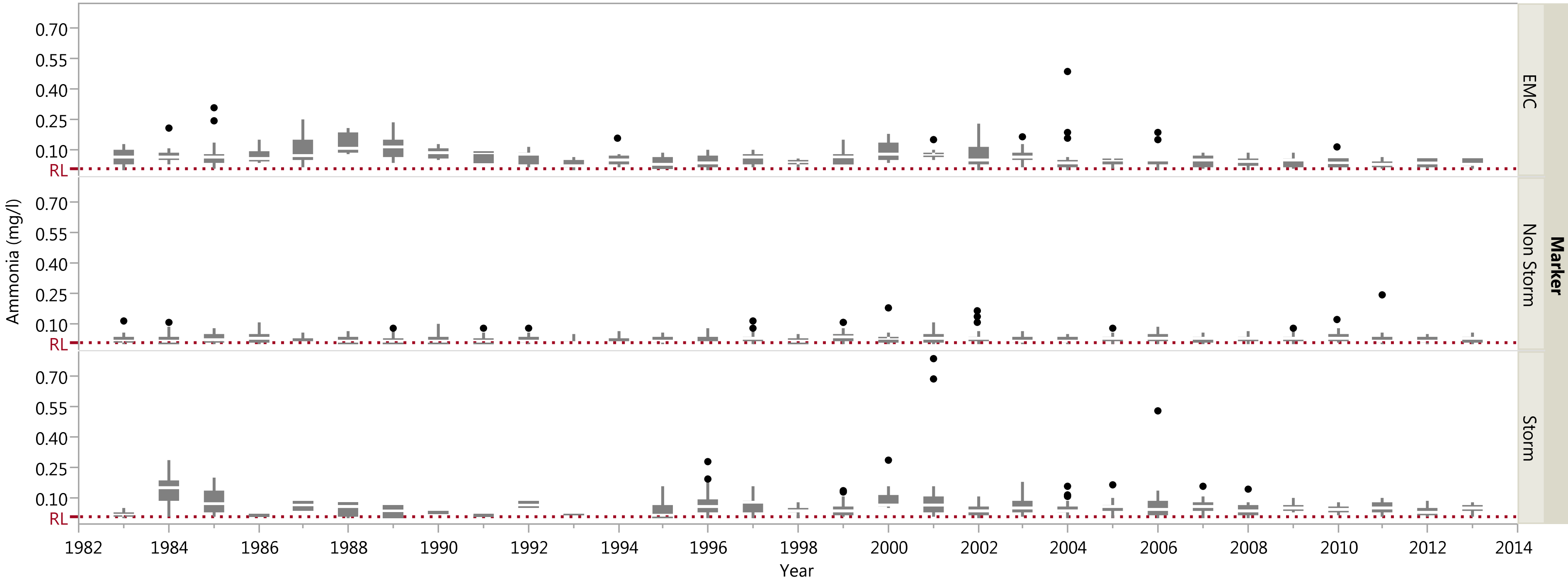
Sadeghi, S. H. R., and P. Saeidi (2010), Reliability of sediment rating curves for a deciduous forest watershed in Iran, *Hydrol. Sci. J.*, *55*(5), 821-831.

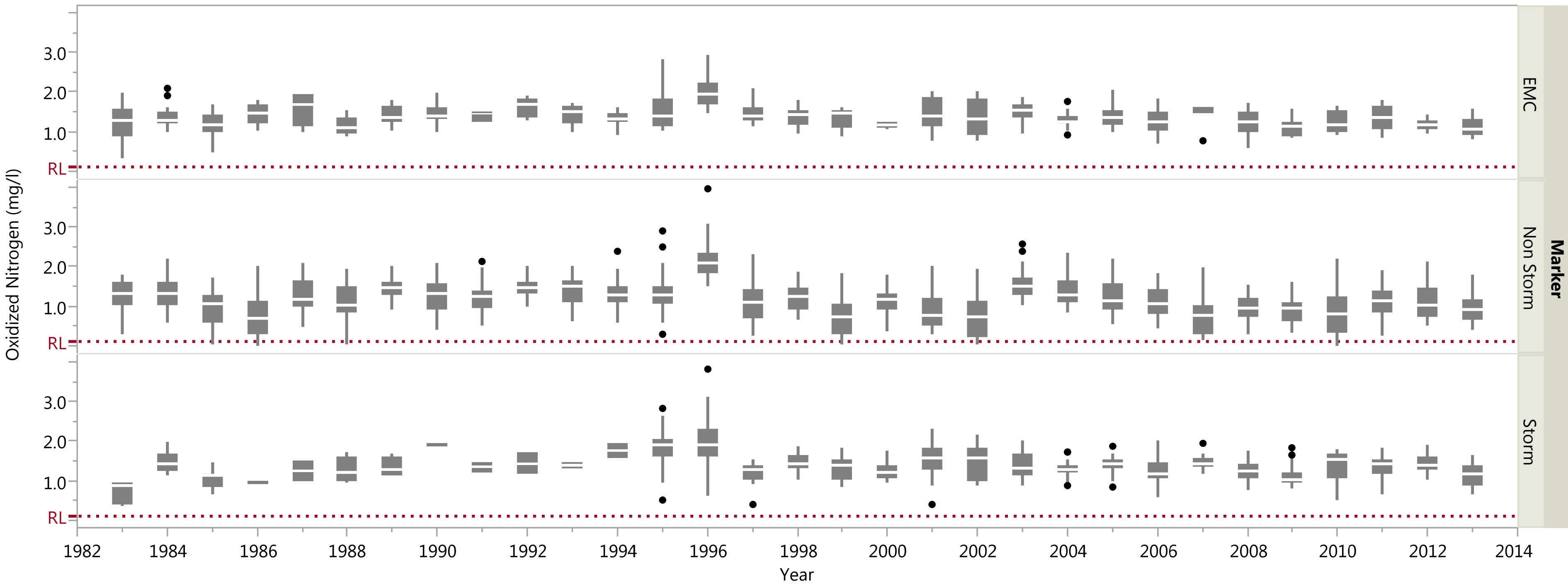
Sadeghi, S. H. R., T. Mizuyama, S. Miyata, T. Gomi, K. Kosugi, T. Fukushima, S. Mizugaki, and Y. Onda (2008), Development, evaluation and interpretation of sediment rating curves for a Japanese small mountainous reforested watershed, *Geoderma*, *144*(1–2), 198-211, doi: 10.1016/j.geoderma.2007.11.008.

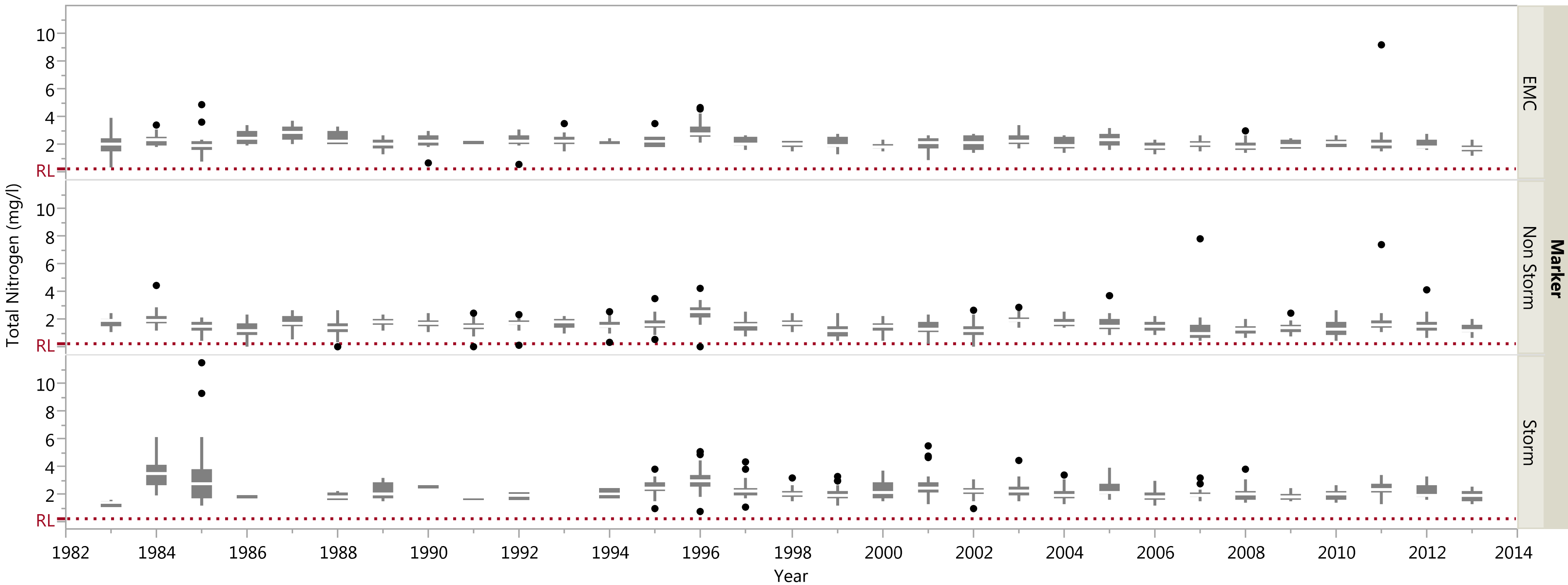
Toor, G. S., R. D. Harmel, B. E. Haggard, and G. Schmidt (2008), Evaluation of regression methodology with low-frequency water quality sampling to estimate constituent loads for ephemeral watersheds in Texas, *J. Environ. Qual.*, *37*(5), 1847-1854, doi: 10.2134/jeq2007.0232.

Appendix A

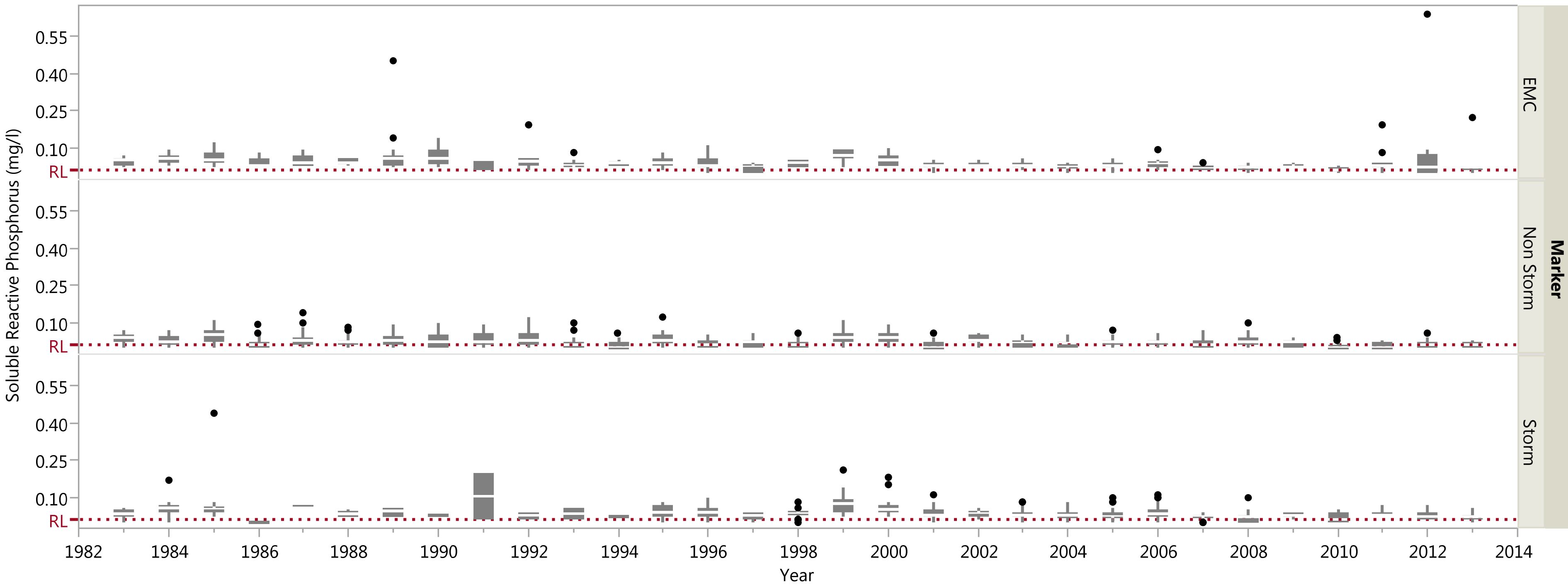
Annual distribution of the observed concentration of the seven constituents of interest. The reporting limit (RL) is plotted in red.



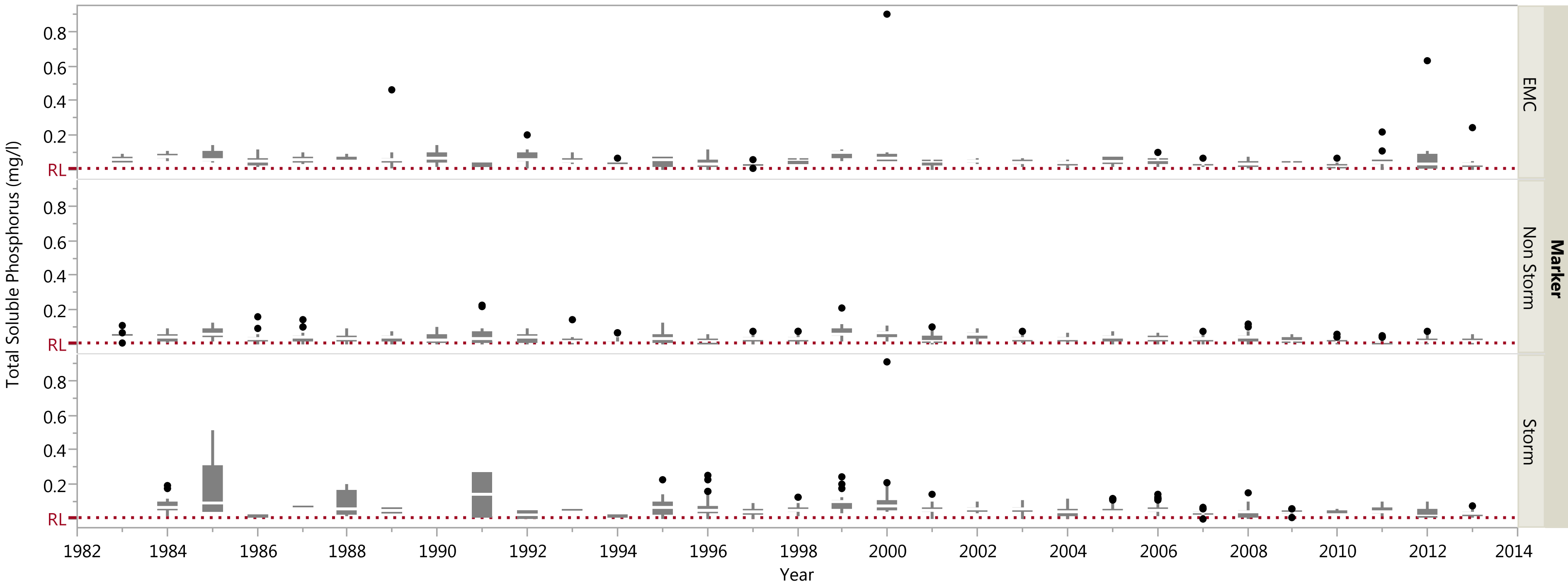

Oxidized Nitrogen (mg/l)
3.0
2.0
1.0
RL
EMC
Non Storm
Storm
Marker
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
Year

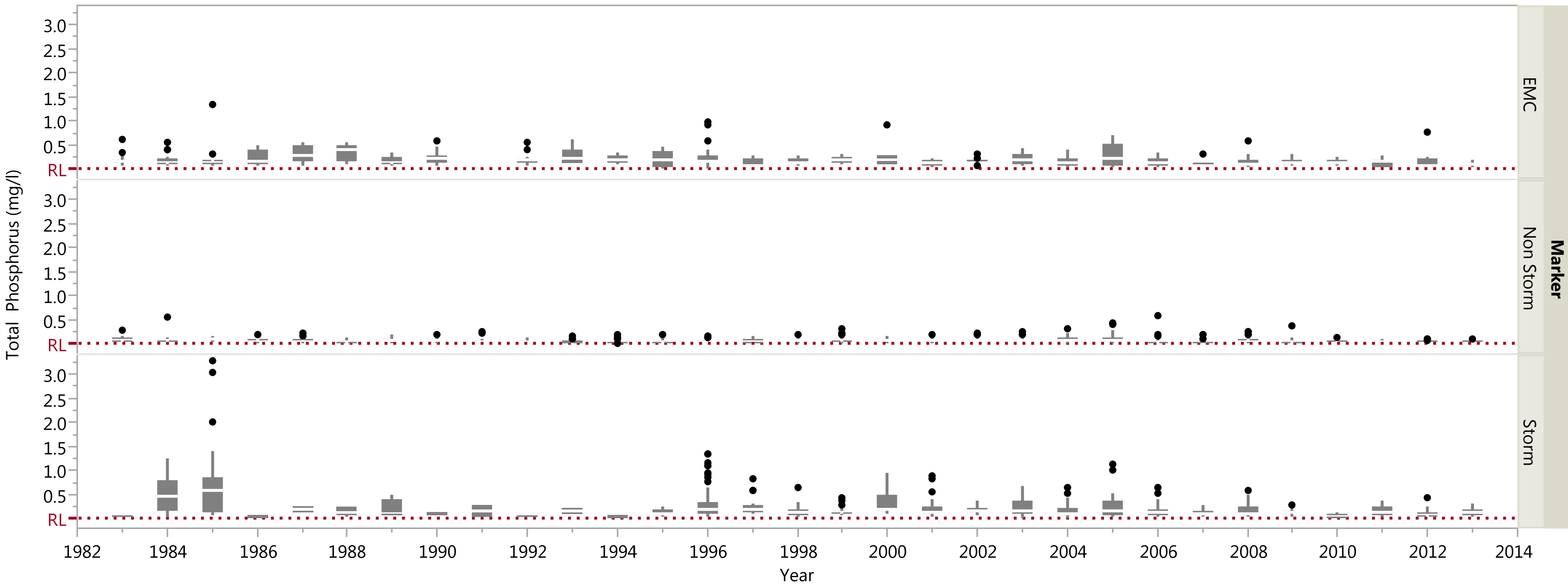


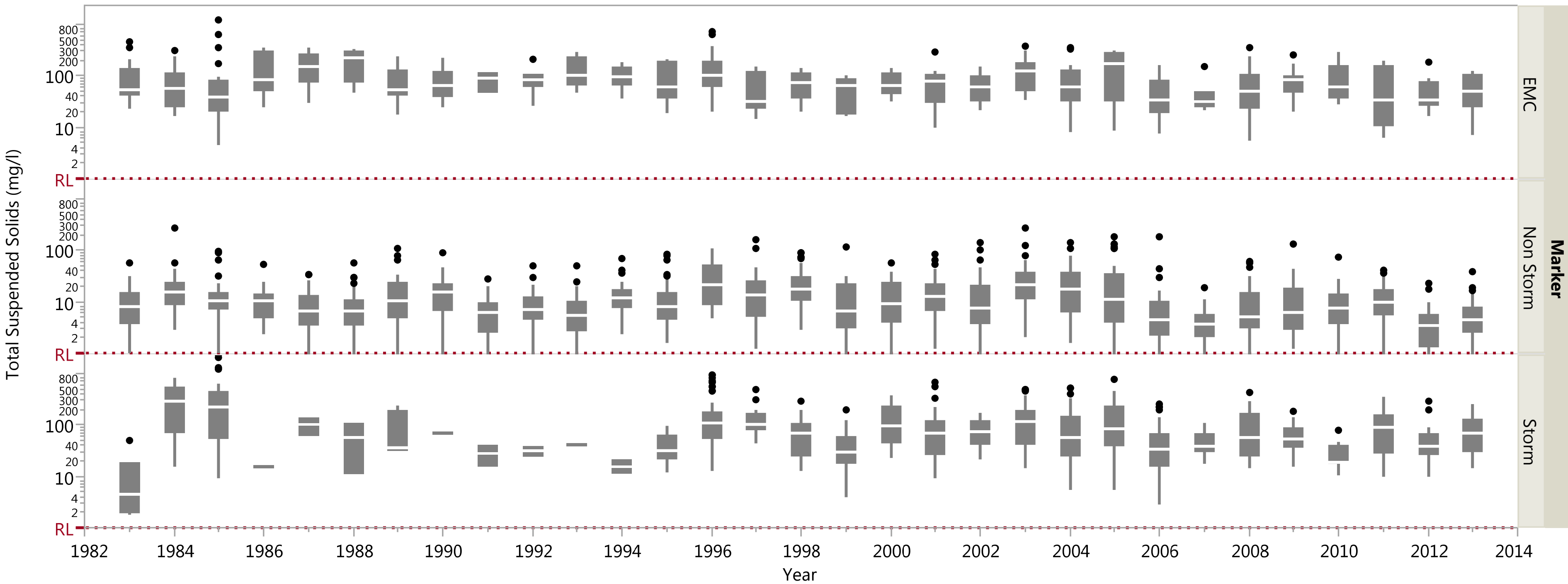
Total Nitrogen (mg/l)
10
8
6
4
2
RL
EMC
Non Storm
Storm
Marker
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
Year



Soluble Reactive Phosphorus (mg/l)
0.55
0.40
0.25
0.10
RL
EMC
Non Storm
Storm
Marker
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
Year

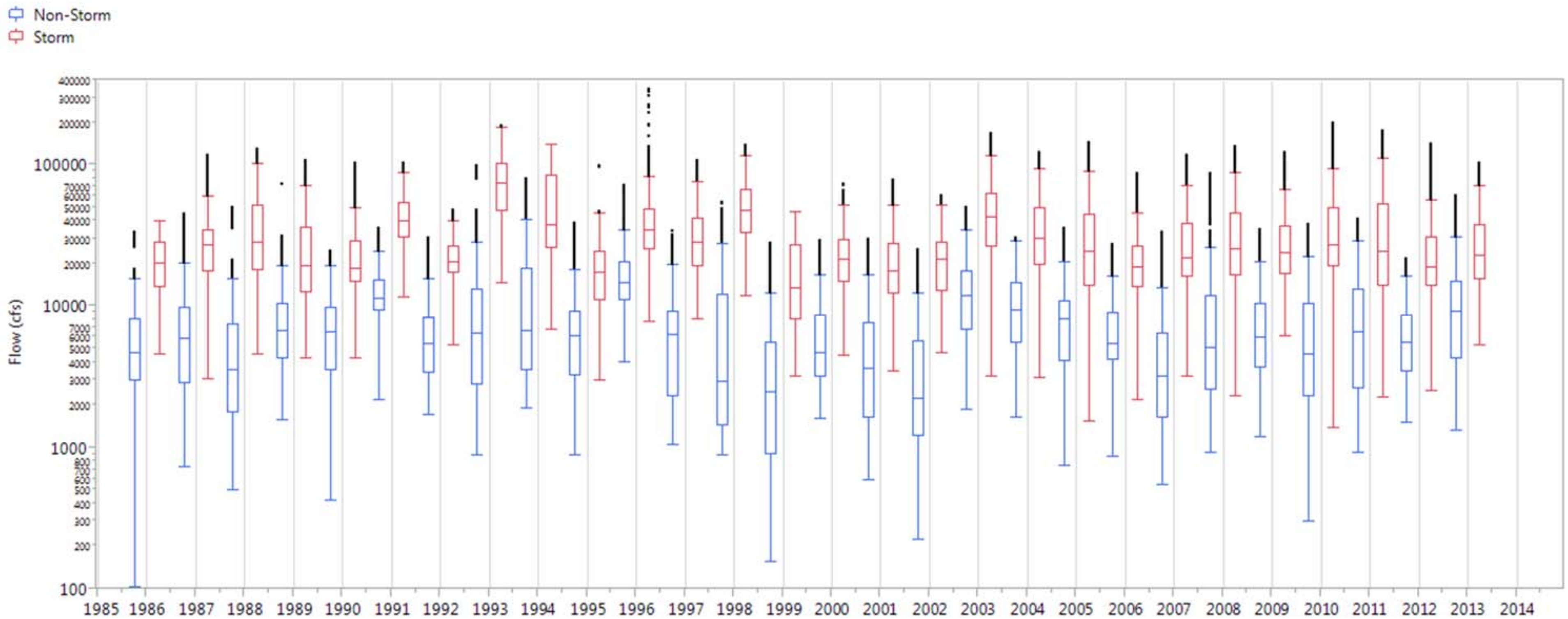

Total Soluble Phosphorus (mg/l)
0.8
0.6
0.4
0.2
RL
EMC
Non Storm
Storm
Marker
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
Year


Total Phosphorus (mg/l)
3.0
2.5
2.0
1.5
1.0
0.5
RL
EMC
Non Storm
Storm
Marker
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
Year

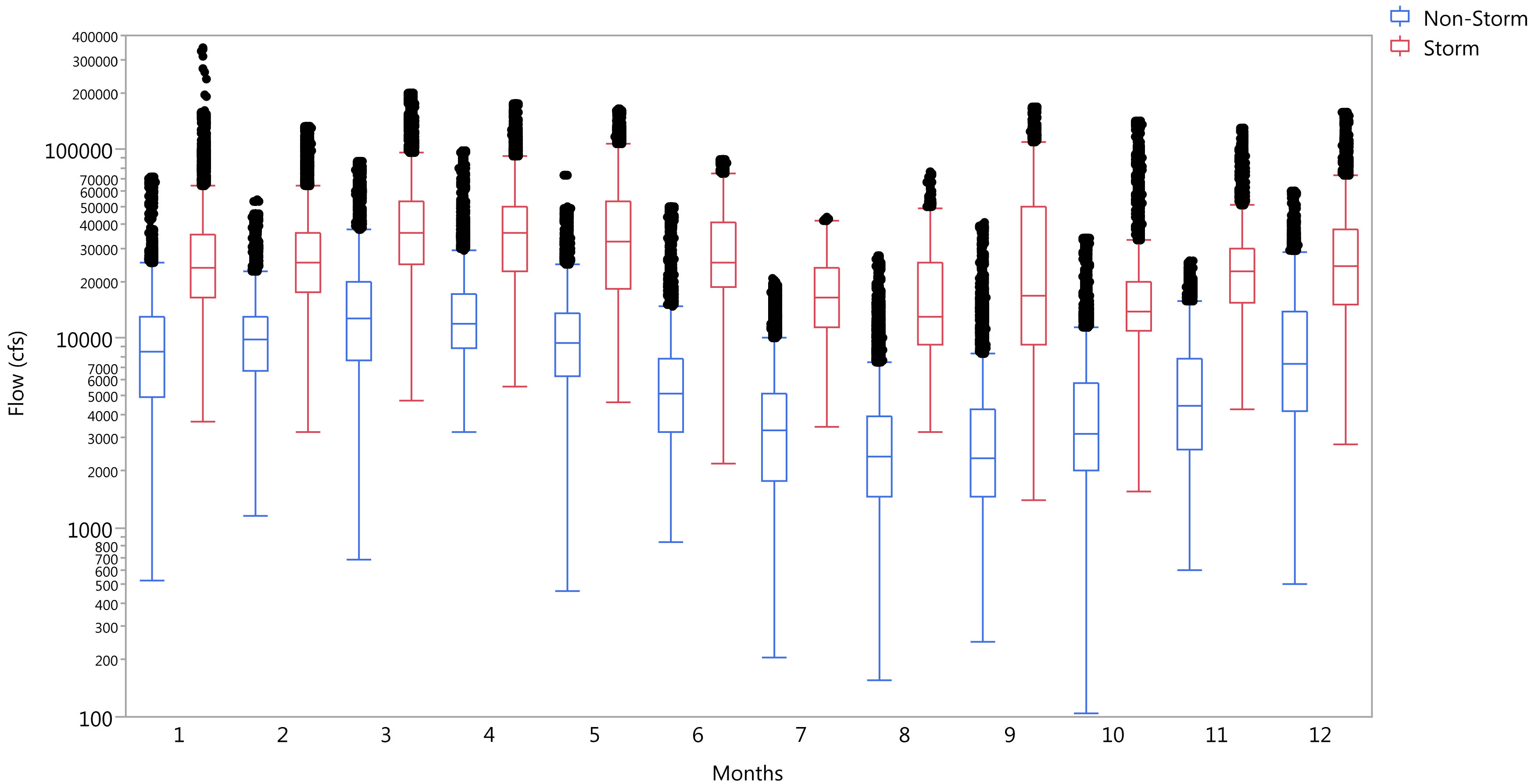

Total Suspended Solids (mg/l)
800
500
300
200
100
40
20
10
4
2
RL
EMC
Non Storm
Storm
Marker
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
Year

Appendix B

Flow classification from 1985-2013.



Annual flow distribution for storm and non-storm flows

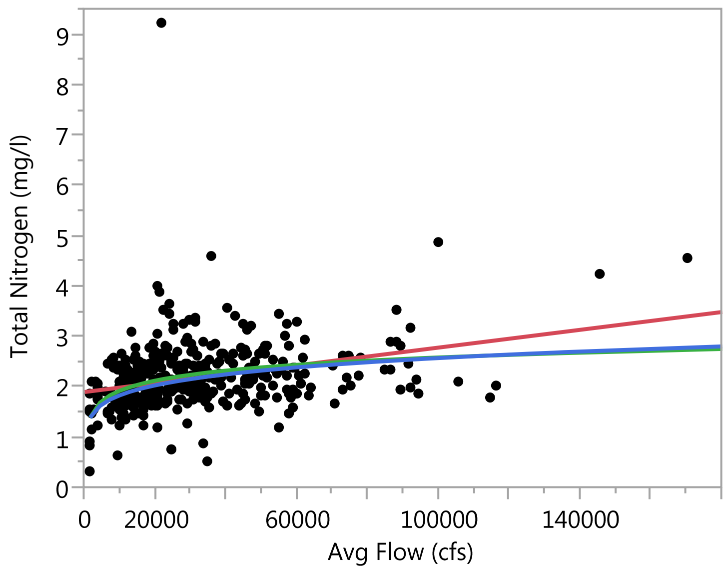


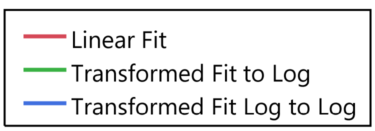
Monthly flow distribution Annual flow distribution for storm and non-storm flows (1985-2013). Note that Month 1 is January, 2 is February and so on.

Annual and monthly flow distribution for storm and non-storm flows. Note that Month 1 is January, 2 is February and so on.

Appendix C

Bivariate Fit of Total Nitrogen (mg/l) By Avg Flow (cfs)





Linear Fit

Total Nitrogen (mg/l) = 1.9106418 + 8.8315e-6*Avg Flow (cfs)

Summary of Fit

RSquare	0.091289
RSquare Adj	0.088878
Root Mean Square Error	0.654541
Mean of Response	2.18752
Observations (or Sum Wgts)	379

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	16.22577	16.2258	37.8732
Error	377	161.51570	0.4284	**Prob > F**
C. Total	378	177.74147		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1.9106418	0.056166	34.02	<.0001*
Avg Flow (cfs)	8.8315e-6	1.435e-6	6.15	<.0001*

Transformed Fit to Log

Total Nitrogen (mg/l) = -0.686436 + 0.285221*Log(Avg Flow (cfs))

Summary of Fit

RSquare	0.112531
RSquare Adj	0.110177
Root Mean Square Error	0.646845
Mean of Response	2.18752
Observations (or Sum Wgts)	379

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	20.00150	20.0015	47.8038
Error	377	157.73997	0.4184	**Prob > F**
C. Total	378	177.74147		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.686436	0.416996	-1.65	0.1006
Log(Avg Flow (cfs))	0.285221	0.041253	6.91	<.0001*

Transformed Fit Log to Log

Log(Total Nitrogen (mg/l)) = -0.727752 + 0.145704*Log(Avg Flow (cfs))

Summary of Fit

RSquare	0.156194
RSquare Adj	0.153956
Root Mean Square Error	0.273489
Mean of Response	0.740397
Observations (or Sum Wgts)	379

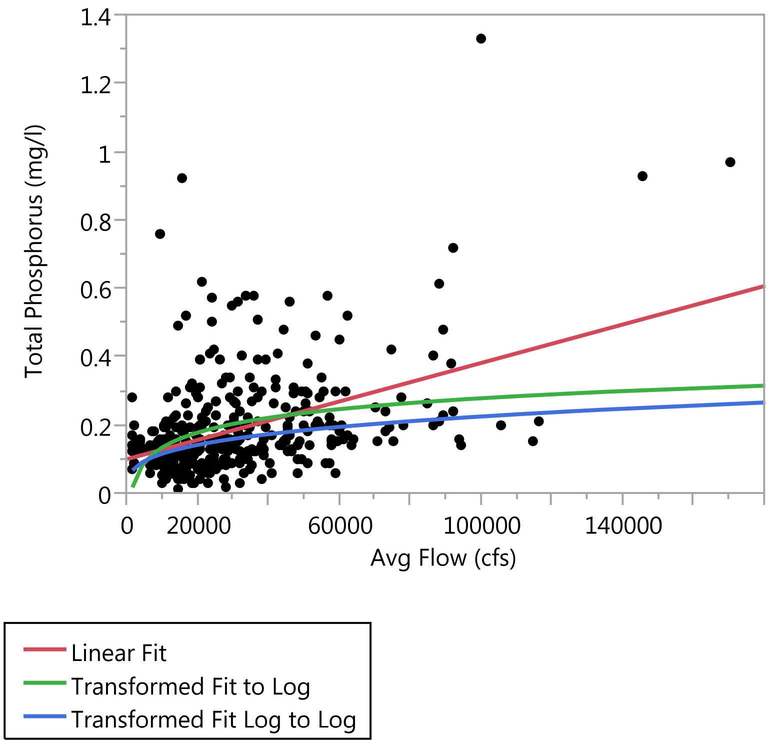
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	5.219664	5.21966	69.7851
Error	377	28.198181	0.07480	**Prob > F**
C. Total	378	33.417845		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.727752	0.176308	-4.13	<.0001*
Log(Avg Flow (cfs))	0.145704	0.017442	8.35	<.0001*

Bivariate Fit of Total Phosphorus (mg/l) By Avg Flow (cfs)



Linear Fit

Total Phosphorus (mg/l) = 0.1024119 + 2.8176e-6*Avg Flow (cfs)

Summary of Fit

RSquare	0.189413
RSquare Adj	0.187216
Root Mean Square Error	0.137856
Mean of Response	0.190916
Observations (or Sum Wgts)	371

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	1.6386629	1.63866	86.2254
Error	369	7.0126255	0.01900	**Prob > F**
C. Total	370	8.6512884		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.1024119	0.011919	8.59	<.0001*
Avg Flow (cfs)	2.8176e-6	3.034e-7	9.29	<.0001*

Transformed Fit to Log

Total Phosphorus (mg/l) = -0.436168 + 0.0622421*Log(Avg Flow (cfs))

Summary of Fit

RSquare	0.109187
RSquare Adj	0.106773
Root Mean Square Error	0.144517
Mean of Response	0.190916
Observations (or Sum Wgts)	371

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	0.9446120	0.944612	45.2286
Error	369	7.7066764	0.020885	**Prob > F**
C. Total	370	8.6512884		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-0.436168	0.093545	-4.66	<.0001*
Log(Avg Flow (cfs))	0.0622421	0.009255	6.73	<.0001*

Transformed Fit Log to Log

Log(Total Phosphorus (mg/l)) = -4.735669 + 0.2825876*Log(Avg Flow (cfs))

Summary of Fit

RSquare	0.114294
RSquare Adj	0.111894
Root Mean Square Error	0.639463
Mean of Response	-1.88862
Observations (or Sum Wgts)	371

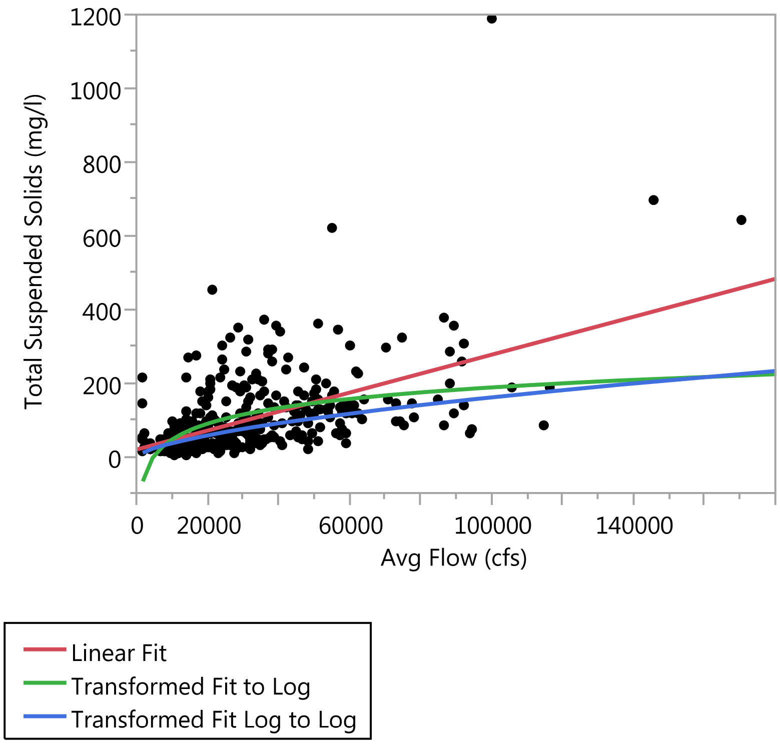
Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	19.47111	19.4711	47.6167
Error	369	150.88899	0.4089	**Prob > F**
C. Total	370	170.36010		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-4.735669	0.41392	-11.44	<.0001*
Log(Avg Flow (cfs))	0.2825876	0.040952	6.90	<.0001*

Bivariate Fit of Total Suspended Solids (mg/l) By Avg Flow (cfs)



Linear Fit

Total Suspended Solids (mg/l) = 21.417188 + 0.0025772*Avg Flow (cfs)

Summary of Fit

RSquare	0.293915
RSquare Adj	0.292017
Root Mean Square Error	94.2283
Mean of Response	102.3757
Observations (or Sum Wgts)	374

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	1374899.1	1374899	154.8489
Error	372	3302977.7	8879	**Prob > F**
C. Total	373	4677876.8		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	21.417188	8.12819	2.63	0.0088*
Avg Flow (cfs)	0.0025772	0.000207	12.44	<.0001*

Transformed Fit to Log

Total Suspended Solids (mg/l) = -511.432 + 60.912996*Log(Avg Flow (cfs))

Summary of Fit

RSquare	0.193654
RSquare Adj	0.191487
Root Mean Square Error	100.6963
Mean of Response	102.3757
Observations (or Sum Wgts)	374

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	905891.7	905892	89.3407
Error	372	3771985.1	10140	**Prob > F**
C. Total	373	4677876.8		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-511.432	65.14772	-7.85	<.0001*
Log(Avg Flow (cfs))	60.912996	6.444442	9.45	<.0001*

Transformed Fit Log to Log

Log(Total Suspended Solids (mg/l)) = -2.041148 + 0.6196945*Log(Avg Flow (cfs))

Summary of Fit

RSquare	0.283078
RSquare Adj	0.28115
Root Mean Square Error	0.798946
Mean of Response	4.203385
Observations (or Sum Wgts)	374

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	93.75871	93.7587	146.8846
Error	372	237.45331	0.6383	**Prob > F**
C. Total	373	331.21202		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-2.041148	0.516896	-3.95	<.0001*
Log(Avg Flow (cfs))	0.6196945	0.051132	12.12	<.0001*

Appendix D

Annual loads computed by different methods (all data are in **pounds for the year**)

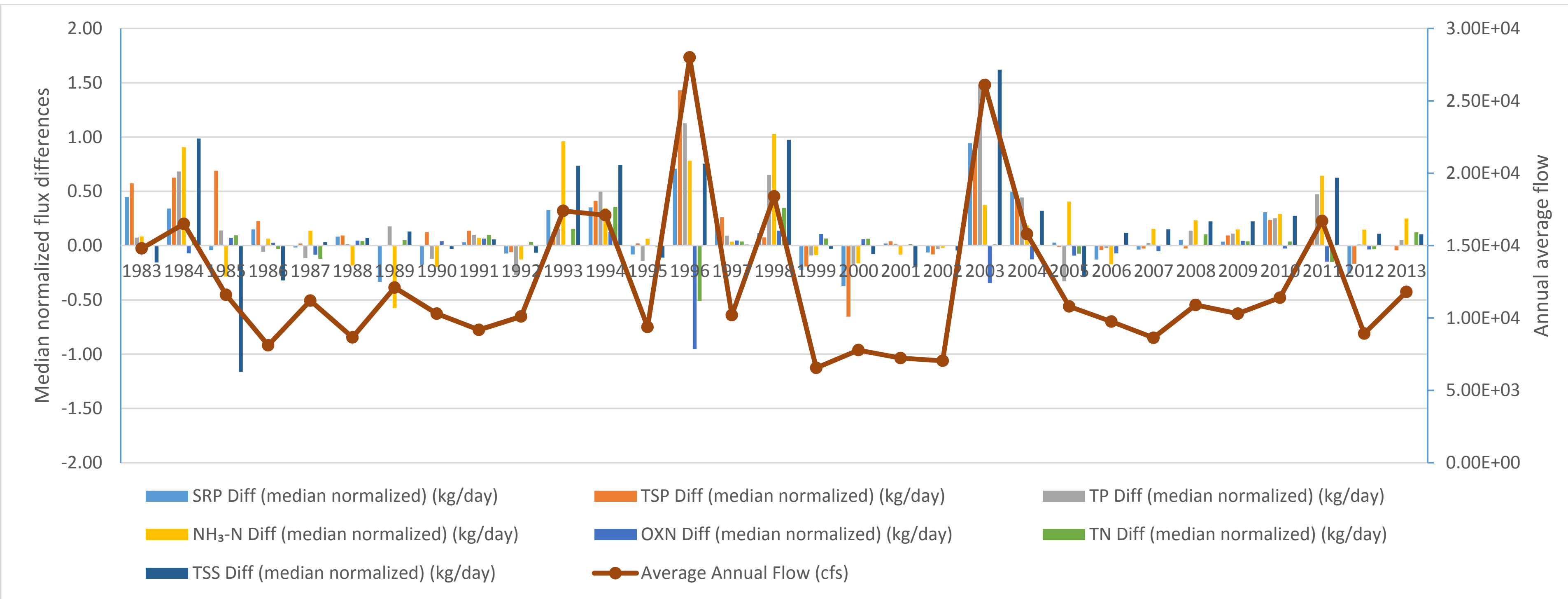
	Loads from direct method							Loads from WRTDS using OWML data							Loads from WRTDS published by USGS				
Year	**SRP**	**TSP**	**TP**	**NH_3-N**	**OXN**	**TN**	**TSS**	**SRP**	**TSP**	**TP**	**NH_3-N**	**OXN**	**TN**	**TSS**	**SRP**	**TP**	**OXN**	**TN**	**SS**
1983	1.1E+06	1.6E+06	5.8E+06	2.0E+06	4.1E+07	6.5E+07	3.1E+09	1.4E+06	2.1E+06	6.0E+06	2.0E+06	4.0E+07	6.6E+07	2.9E+09	2.1E+06	5.9E+06	3.9E+07	7.0E+07	5.8E+09
1984	1.4E+06	1.9E+06	6.4E+06	1.9E+06	4.8E+07	7.6E+07	3.4E+09	1.6E+06	2.4E+06	8.3E+06	2.8E+06	4.5E+07	7.9E+07	5.0E+09	2.1E+06	7.8E+06	4.4E+07	8.4E+07	8.4E+09
1985	1.0E+06	1.3E+06	7.9E+06	1.7E+06	2.8E+07	5.4E+07	6.4E+09	9.9E+05	2.0E+06	8.3E+06	1.4E+06	3.0E+07	5.8E+07	4.6E+09	1.0E+06	8.3E+06	2.9E+07	6.2E+07	8.9E+09
1986	4.6E+05	6.3E+05	2.4E+06	7.8E+05	2.2E+07	3.5E+07	1.5E+09	5.6E+05	8.3E+05	2.2E+06	8.4E+05	2.2E+07	3.3E+07	1.0E+09	6.0E+05	2.1E+06	2.2E+07	3.5E+07	1.5E+09
1987	8.9E+05	1.2E+06	4.1E+06	1.0E+06	3.4E+07	5.3E+07	2.0E+09	8.8E+05	1.2E+06	3.8E+06	1.1E+06	3.1E+07	4.8E+07	2.0E+09	1.0E+06	3.3E+06	3.1E+07	4.9E+07	3.2E+09
1988	5.6E+05	8.0E+05	3.5E+06	1.0E+06	2.2E+07	3.5E+07	1.9E+09	6.1E+05	8.8E+05	3.5E+06	8.4E+05	2.3E+07	3.7E+07	2.1E+09	7.3E+05	2.7E+06	2.2E+07	3.6E+07	3.2E+09
1989	1.2E+06	1.3E+06	3.2E+06	1.6E+06	3.4E+07	4.8E+07	1.7E+09	1.0E+06	1.3E+06	3.7E+06	1.0E+06	3.4E+07	5.0E+07	1.9E+09	1.1E+06	3.2E+06	3.3E+07	4.9E+07	3.0E+09
1990	8.6E+05	9.6E+05	2.9E+06	9.4E+05	3.0E+07	4.2E+07	1.1E+09	7.5E+05	1.1E+06	2.6E+06	7.5E+05	3.1E+07	4.2E+07	1.0E+09	7.7E+05	2.6E+06	3.1E+07	4.3E+07	1.7E+09
1991	5.6E+05	7.2E+05	2.0E+06	8.2E+05	2.6E+07	3.4E+07	8.9E+08	5.7E+05	8.4E+05	2.2E+06	8.9E+05	2.8E+07	3.8E+07	9.8E+08	5.7E+05	2.2E+06	2.8E+07	4.0E+07	1.3E+09
1992	6.9E+05	9.2E+05	2.9E+06	8.4E+05	3.0E+07	3.9E+07	1.0E+09	6.4E+05	8.7E+05	2.1E+06	7.2E+05	3.0E+07	4.0E+07	9.0E+08	6.3E+05	2.0E+06	2.9E+07	3.9E+07	1.2E+09
1993	9.2E+05	1.4E+06	6.2E+06	9.4E+05	5.2E+07	7.4E+07	2.9E+09	1.1E+06	1.5E+06	6.7E+06	1.8E+06	5.2E+07	8.0E+07	4.0E+09	1.2E+06	6.0E+06	4.9E+07	7.6E+07	5.0E+09
1994	8.6E+05	1.1E+06	3.9E+06	1.4E+06	4.8E+07	6.2E+07	1.8E+09	1.1E+06	1.4E+06	5.4E+06	1.7E+06	5.2E+07	7.6E+07	3.0E+09	1.1E+06	4.9E+06	4.9E+07	7.2E+07	3.6E+09
1995	6.6E+05	8.1E+05	2.4E+06	6.3E+05	2.8E+07	3.8E+07	9.9E+08	6.1E+05	8.3E+05	2.0E+06	6.9E+05	2.8E+07	3.8E+07	8.2E+08	6.3E+05	2.0E+06	2.7E+07	3.7E+07	9.6E+08
1996	2.1E+06	2.4E+06	1.4E+07	2.6E+06	1.1E+08	1.6E+08	9.0E+09	2.6E+06	3.7E+06	1.7E+07	3.3E+06	8.6E+07	1.4E+08	1.0E+10	2.7E+06	1.6E+07	8.3E+07	1.3E+08	1.3E+10
1997	4.4E+05	6.5E+05	2.0E+06	8.1E+05	2.9E+07	4.0E+07	9.5E+08	5.9E+05	8.9E+05	2.3E+06	8.4E+05	3.0E+07	4.2E+07	9.7E+08	6.2E+05	2.4E+06	2.9E+07	4.1E+07	1.0E+09
1998	1.1E+06	1.7E+06	5.2E+06	1.4E+06	5.0E+07	7.1E+07	2.6E+09	1.2E+06	1.8E+06	7.1E+06	2.3E+06	5.4E+07	8.5E+07	4.1E+09	1.2E+06	7.4E+06	5.2E+07	8.3E+07	4.6E+09
1999	4.9E+05	7.3E+05	1.4E+06	5.7E+05	1.5E+07	2.2E+07	4.2E+08	3.5E+05	5.5E+05	1.1E+06	4.9E+05	1.8E+07	2.4E+07	3.7E+08	3.7E+05	1.1E+06	1.7E+07	2.3E+07	3.0E+08
2000	6.4E+05	1.2E+06	2.0E+06	8.0E+05	1.9E+07	2.6E+07	6.9E+08	4.0E+05	6.4E+05	1.5E+06	6.4E+05	2.0E+07	2.9E+07	5.7E+08	4.2E+05	1.5E+06	1.9E+07	2.7E+07	4.4E+08
2001	3.4E+05	5.4E+05	1.5E+06	7.1E+05	1.8E+07	2.6E+07	9.2E+08	3.5E+05	5.8E+05	1.5E+06	6.3E+05	1.8E+07	2.6E+07	6.2E+08	3.4E+05	1.4E+06	1.7E+07	2.5E+07	4.8E+08
2002	4.5E+05	7.3E+05	1.7E+06	6.0E+05	1.8E+07	2.6E+07	6.9E+08	4.1E+05	6.5E+05	1.6E+06	5.8E+05	1.8E+07	2.6E+07	6.2E+08	4.1E+05	1.6E+06	1.7E+07	2.5E+07	4.9E+08
2003	1.4E+06	2.3E+06	9.0E+06	3.1E+06	7.7E+07	1.2E+08	5.8E+09	2.0E+06	3.2E+06	1.3E+07	3.4E+06	6.7E+07	1.2E+08	8.3E+09	1.9E+06	1.3E+07	6.4E+07	1.2E+08	9.8E+09
2004	6.7E+05	1.1E+06	3.8E+06	1.5E+06	4.5E+07	6.4E+07	2.3E+09	9.8E+05	1.5E+06	5.1E+06	1.7E+06	4.1E+07	6.4E+07	2.7E+09	9.5E+05	5.2E+06	4.0E+07	6.3E+07	3.0E+09
2005	5.3E+05	9.1E+05	4.0E+06	7.2E+05	3.0E+07	4.5E+07	2.1E+09	5.5E+05	9.0E+05	3.0E+06	1.1E+06	2.7E+07	4.2E+07	1.6E+09	4.8E+05	3.0E+06	2.6E+07	4.1E+07	2.0E+09
2006	6.0E+05	8.3E+05	2.3E+06	9.7E+05	2.7E+07	3.6E+07	8.2E+08	5.2E+05	8.0E+05	2.2E+06	8.0E+05	2.5E+07	3.6E+07	1.0E+09	4.8E+05	2.2E+06	2.4E+07	3.6E+07	1.0E+09
2007	3.8E+05	6.1E+05	1.9E+06	7.0E+05	2.3E+07	3.3E+07	8.4E+08	3.5E+05	5.8E+05	2.0E+06	8.5E+05	2.1E+07	3.3E+07	1.1E+09	2.7E+05	2.0E+06	2.0E+07	3.2E+07	1.4E+09
2008	4.8E+05	8.3E+05	2.4E+06	8.9E+05	2.6E+07	3.7E+07	1.3E+09	5.1E+05	8.1E+05	2.8E+06	1.1E+06	2.6E+07	4.1E+07	1.6E+09	3.9E+05	2.8E+06	2.5E+07	4.0E+07	2.2E+09
2009	4.7E+05	6.6E+05	2.1E+06	8.1E+05	2.3E+07	3.7E+07	9.9E+08	4.9E+05	7.5E+05	2.4E+06	9.5E+05	2.4E+07	3.8E+07	1.3E+09	3.7E+05	2.4E+06	2.3E+07	3.7E+07	1.8E+09
2010	3.2E+05	5.7E+05	2.4E+06	9.4E+05	2.9E+07	4.5E+07	1.9E+09	5.2E+05	7.9E+05	3.1E+06	1.2E+06	2.8E+07	4.7E+07	2.3E+09	3.3E+05	4.0E+06	2.6E+07	4.8E+07	4.7E+09
2011	9.5E+05	1.3E+06	3.8E+06	1.3E+06	4.2E+07	7.5E+07	2.7E+09	9.4E+05	1.4E+06	5.2E+06	1.9E+06	3.8E+07	6.9E+07	3.7E+09	7.2E+05	6.2E+06	3.5E+07	6.6E+07	5.9E+09
2012	5.0E+05	6.5E+05	1.4E+06	5.1E+05	2.2E+07	3.4E+07	6.4E+08	3.4E+05	4.9E+05	1.4E+06	6.4E+05	2.1E+07	3.3E+07	8.1E+08	2.4E+05	1.8E+06	2.0E+07	3.1E+07	1.1E+09
2013	4.3E+05	6.6E+05	1.6E+06	6.9E+05	2.8E+07	3.9E+07	8.9E+08	4.2E+05	6.2E+05	1.8E+06	9.2E+05	2.8E+07	4.4E+07	1.0E+09					

Appendix E

Annual differences in flux computed as WL-O−DL normalized by the median flux estimated by DL (Table 1) for various parameter of interests. Annual average flow is also plotted on secondary Y axis.

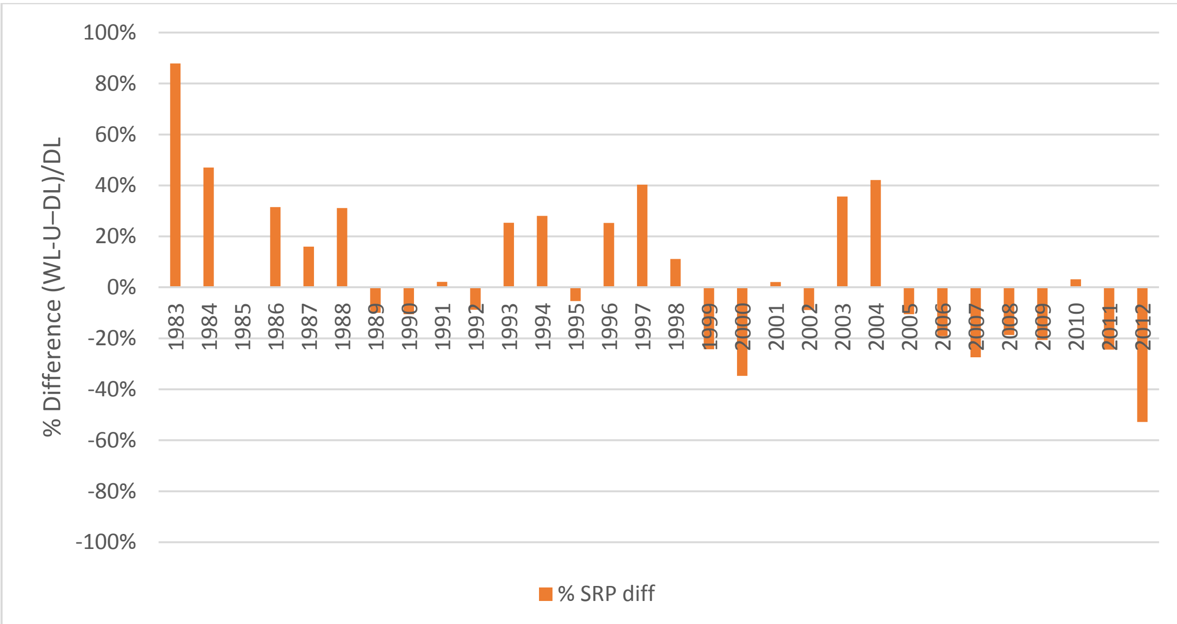
Table 1. Median flux observed in the study period.

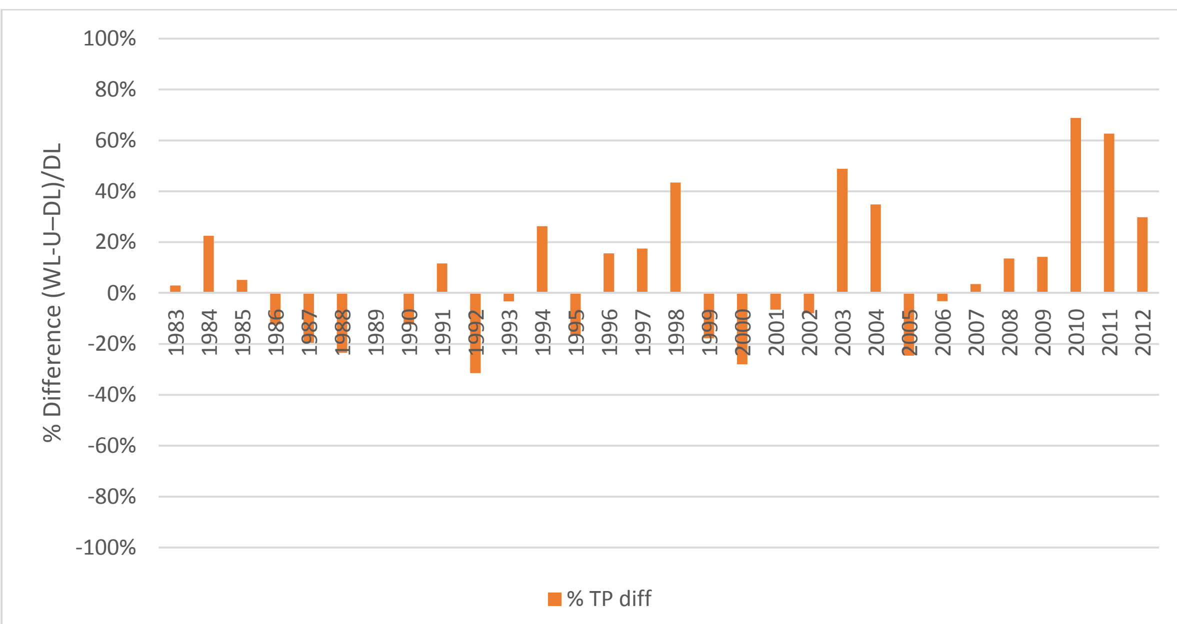
SRP (kg/day)	TSP (kg/day)	TP (kg/day)	NH_3-N (kg/day)	OXN (kg/day)	TN (kg/day)	TSS (kg/day)
7.93E+02	1.13E+03	3.62E+03	1.16E+03	3.54E+04	4.98E+04	1.91E+06

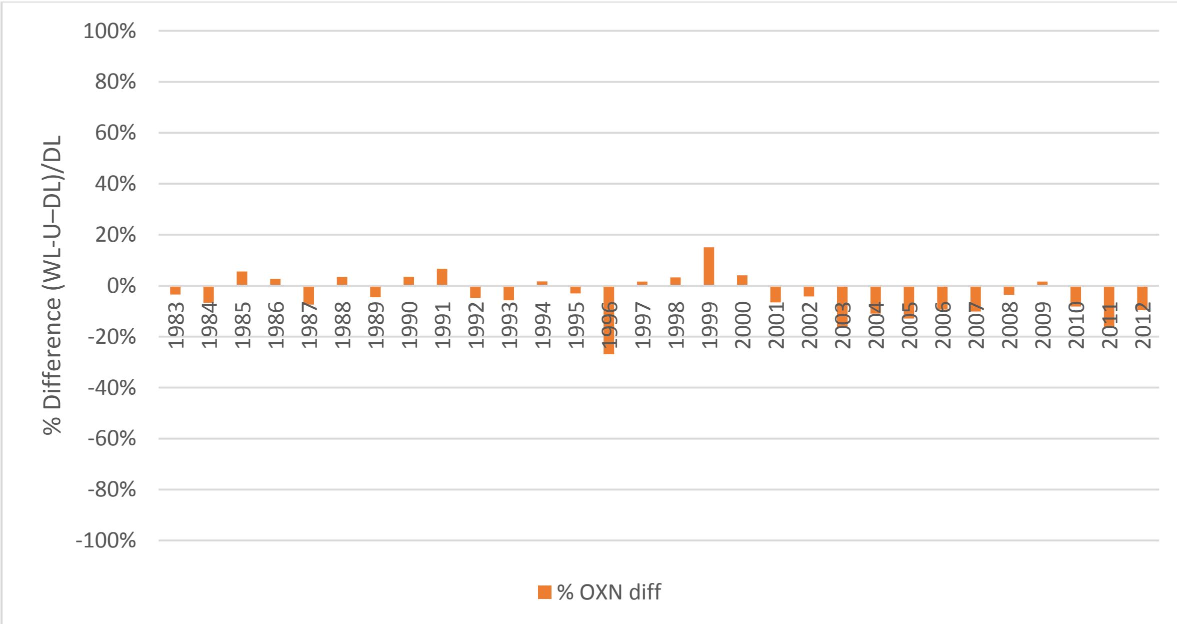


Appendix F

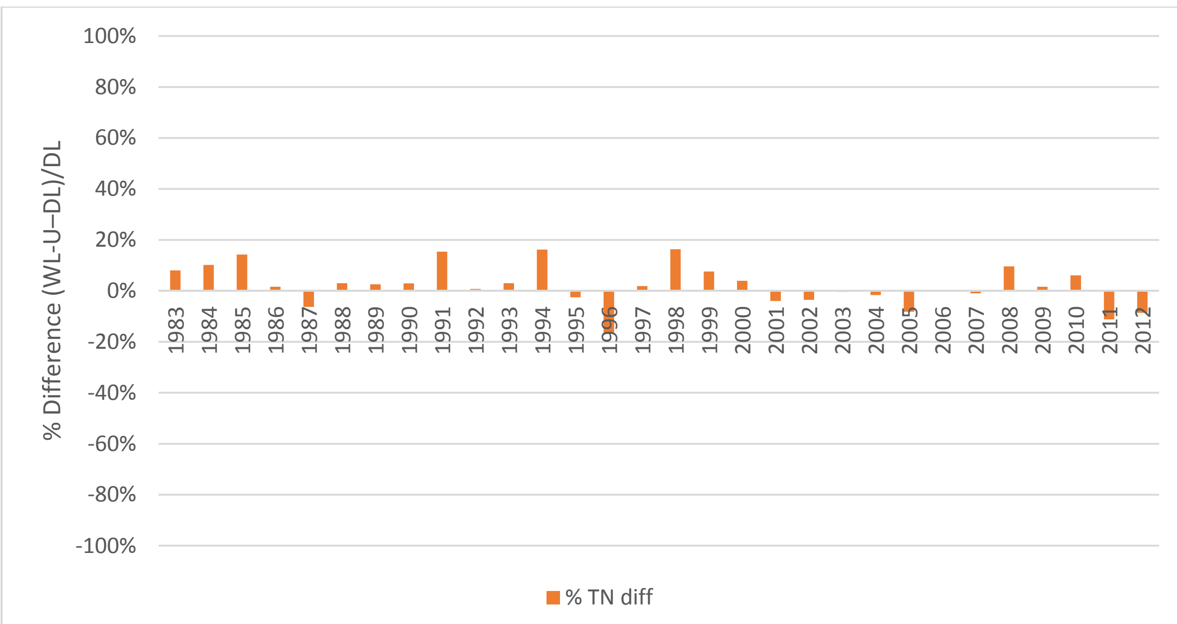
Percentage difference plots for fluvial loads comparing WL-U (WRDS loads computed using USGS data) and DL (direct loads using OWML flow composite)







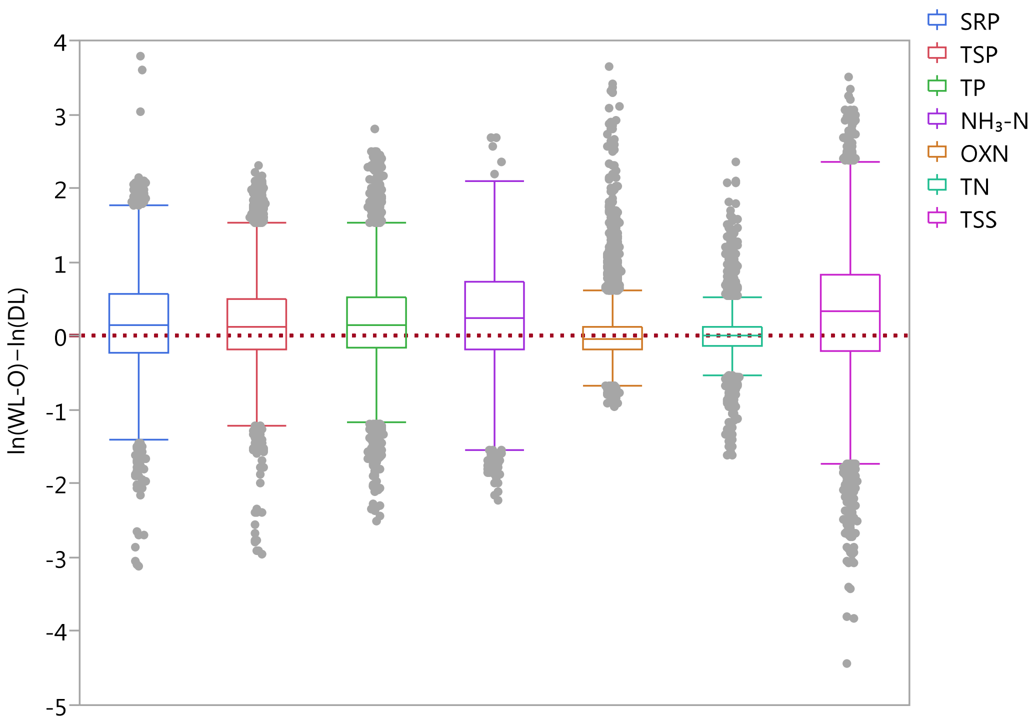
100%
80%
60%
40%
20%
0%
-20%
-40%
-60%
-80%
-100%
% Difference (WL-U–DL)/DL
1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
% OXN diff



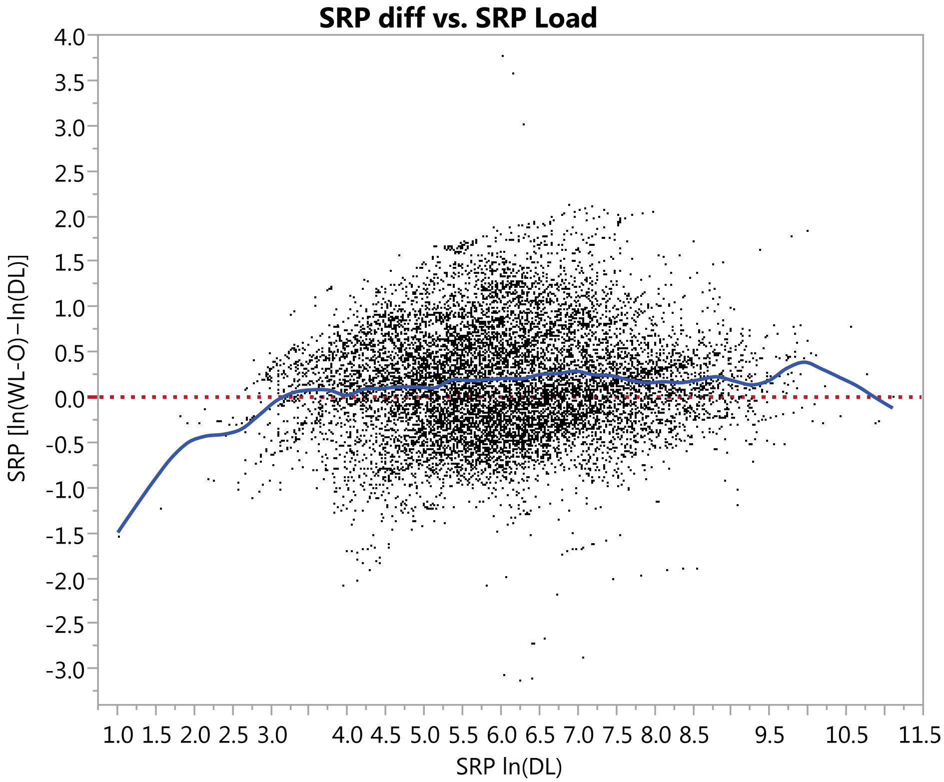
100%
80%
60%
40%
20%
0%
-20%
-40%
-60%
-80%
-100%
% Difference (WL-U–DL)/DL
1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
% TN diff

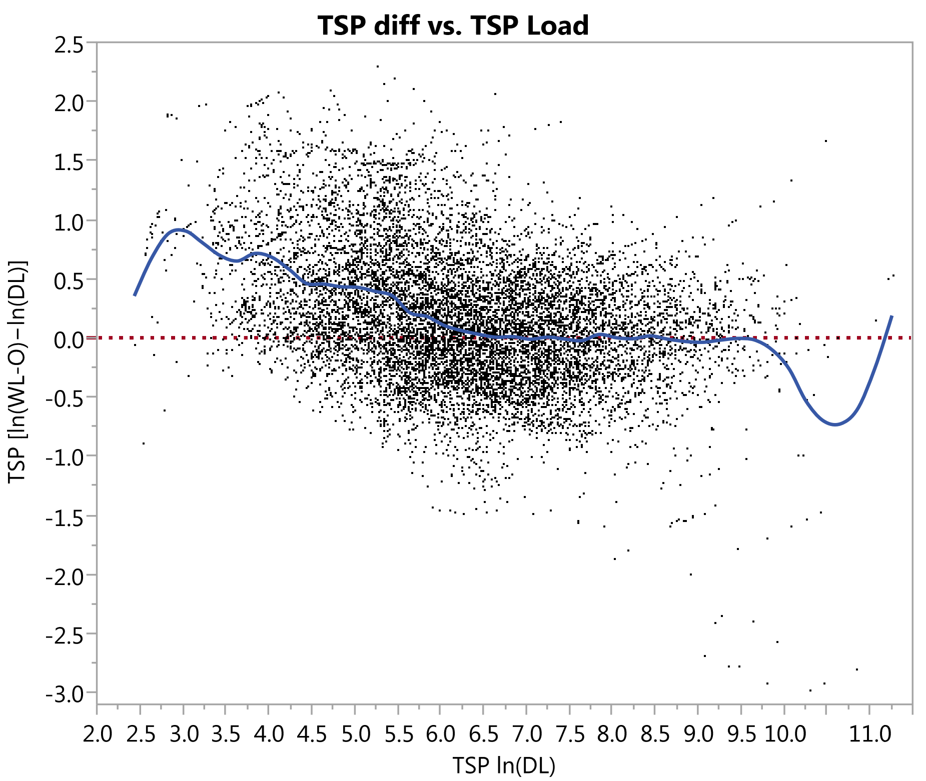
Appendix G

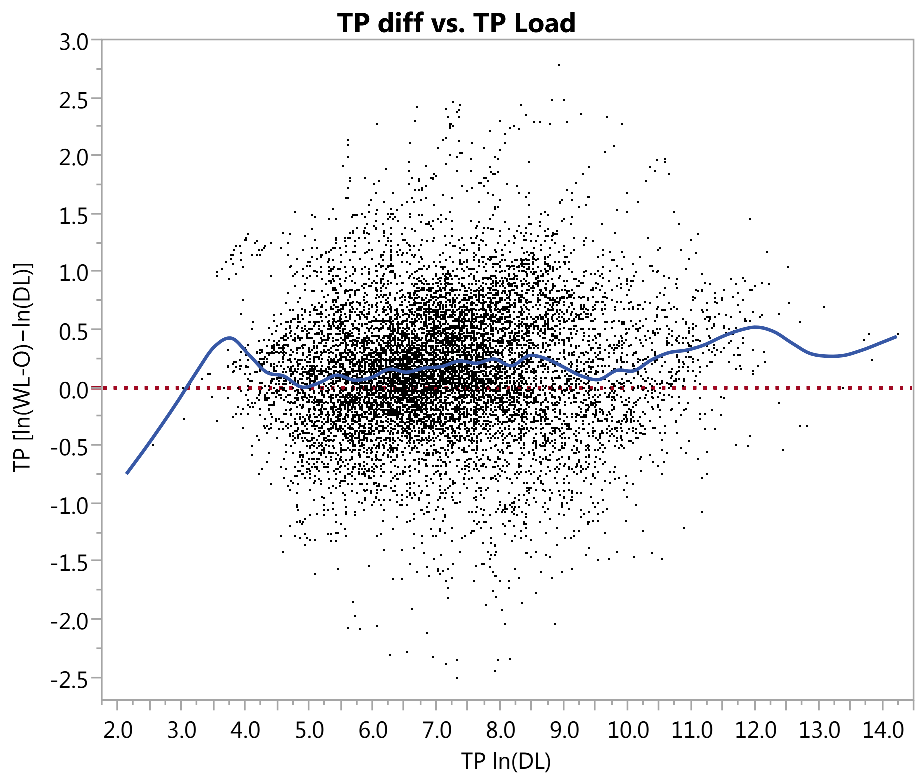
Daily Loads Difference for various parameters

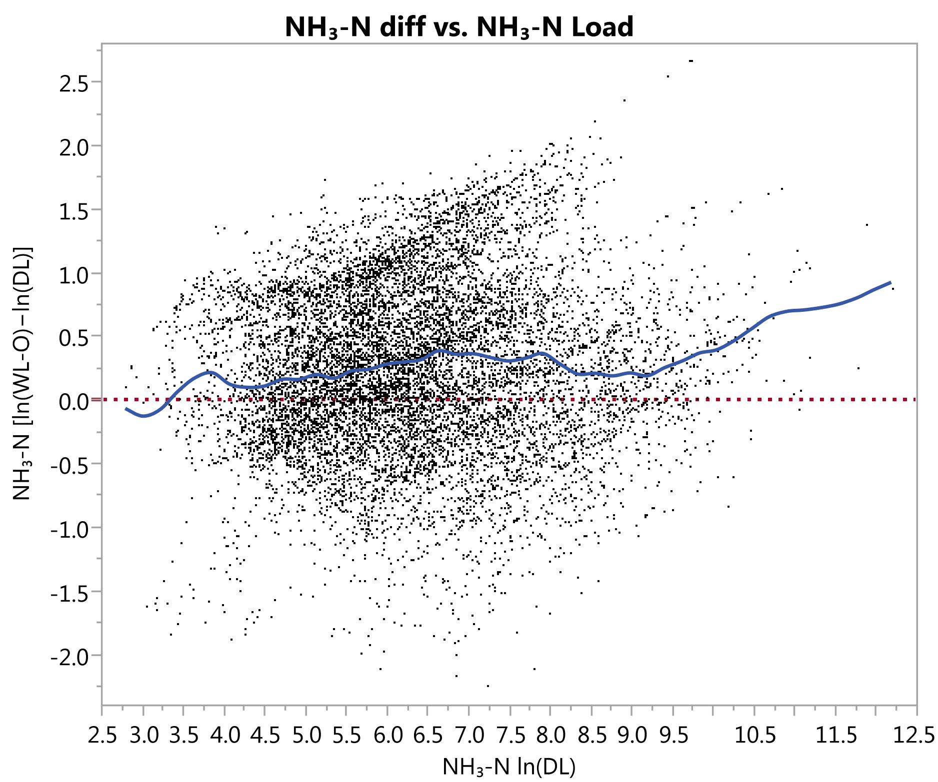


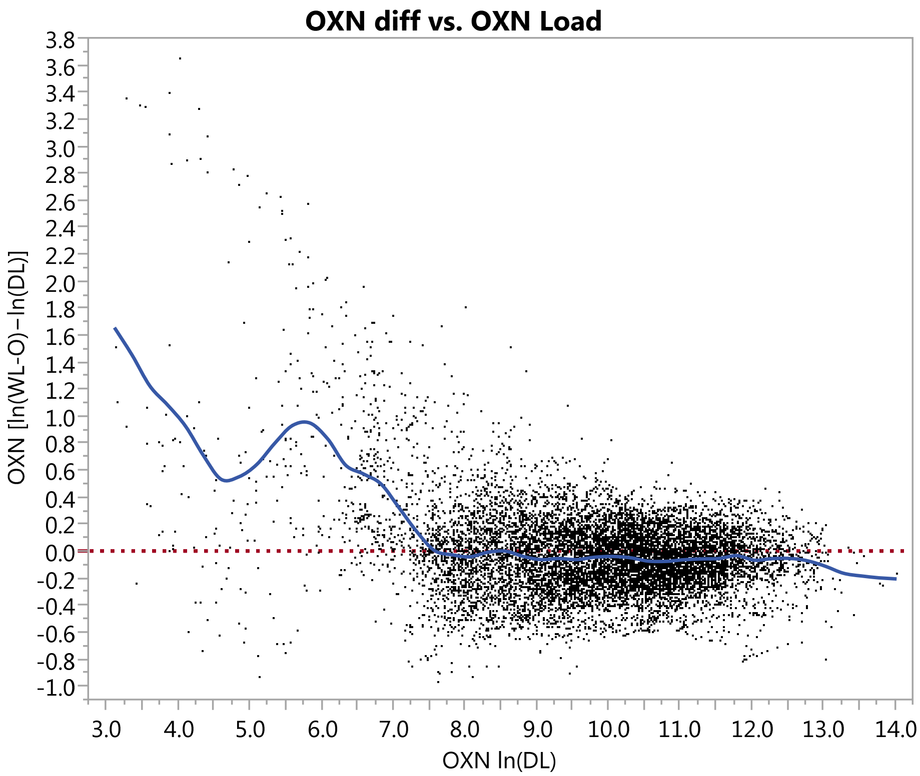
Change in daily load difference between two methods with magnitude of load

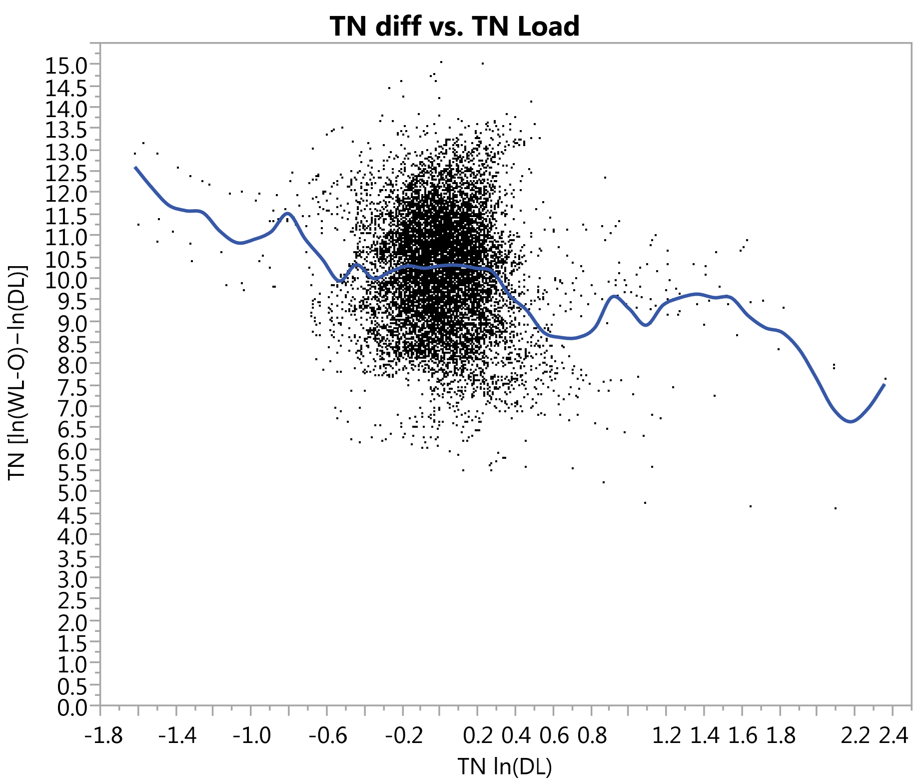


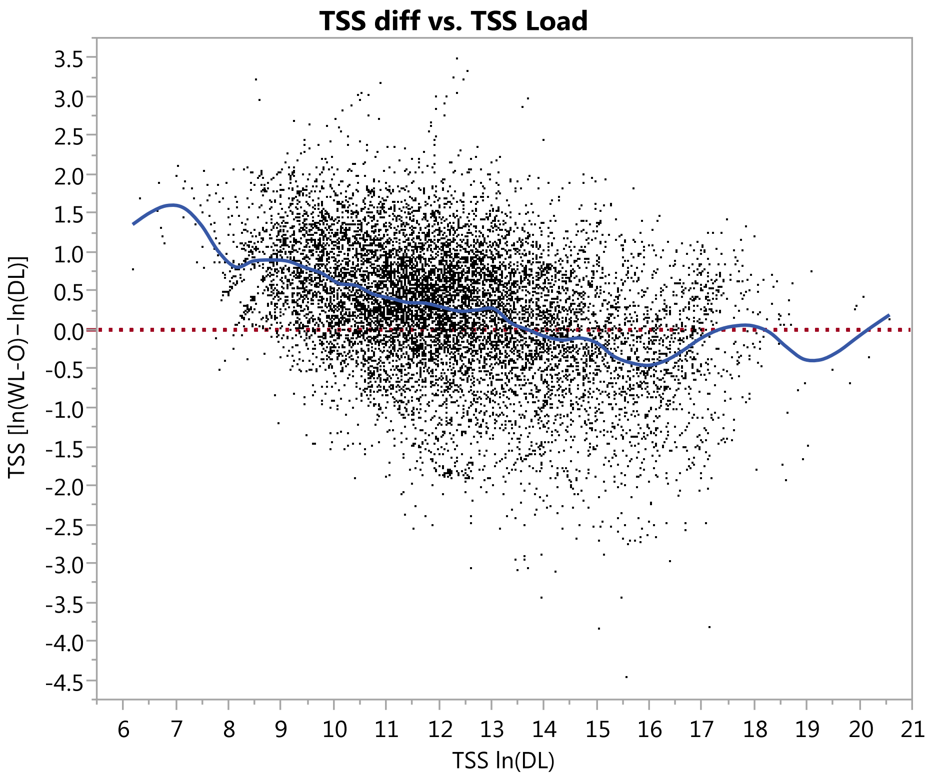



TP diff vs. TP Load
TP [ln(WL-O)−ln(DL)]
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
12.0
13.0
14.0
TP ln(DL)

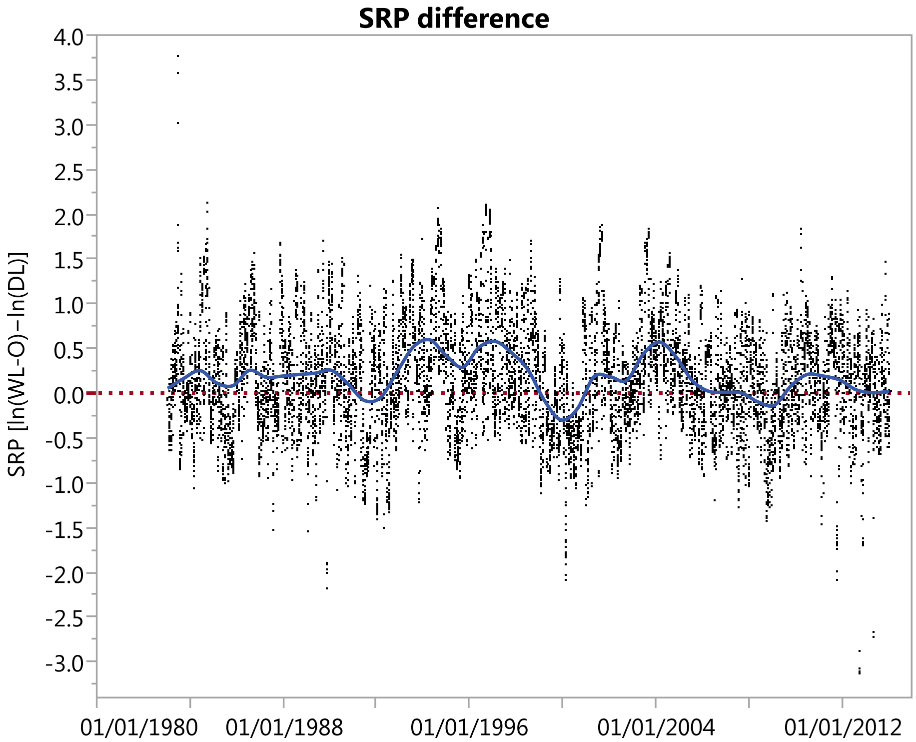

NH3-N diff vs. NH3-N Load
NH3-N [ln(WL-O)−ln(DL)]
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
9.5
10.5
11.5
12.5
NH3-N ln(DL)

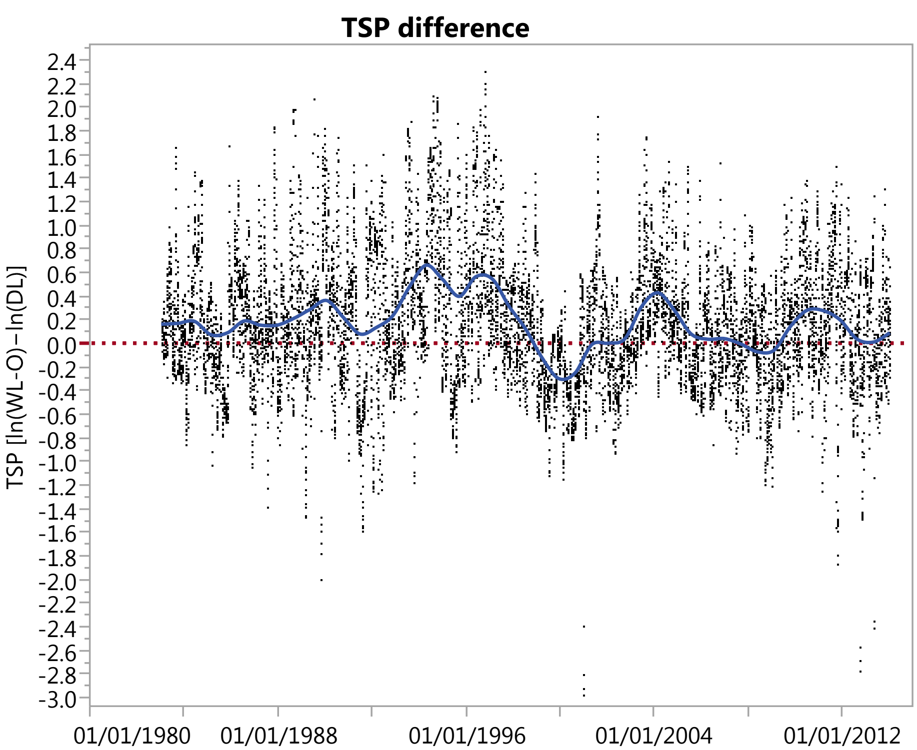

OXN diff vs. OXN Load
OXN [ln(WL-O)−ln(DL)]
OXN ln(DL)
3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0 12.0 13.0 14.0

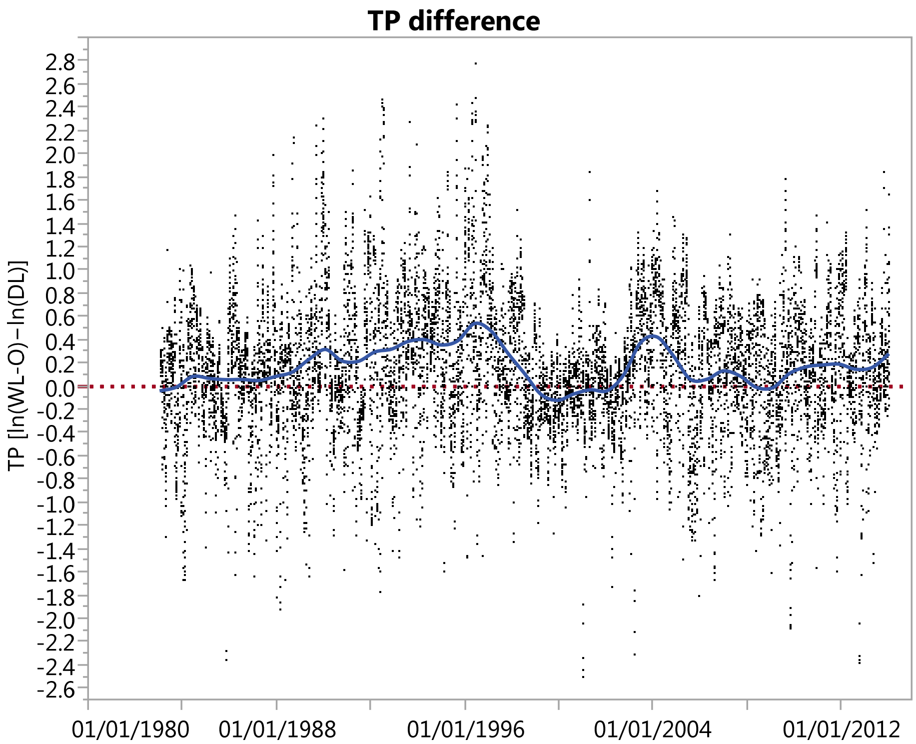

TN diff vs. TN Load
TN [ln(WL-O)−ln(DL)]
TN ln(DL)
15.0
14.5
14.0
13.5
13.0
12.5
12.0
11.5
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-1.8 -1.4 -1.0 -0.6 -0.2 0.2 0.4 0.6 0.8 1.2 1.4 1.6 1.8 2.2 2.4

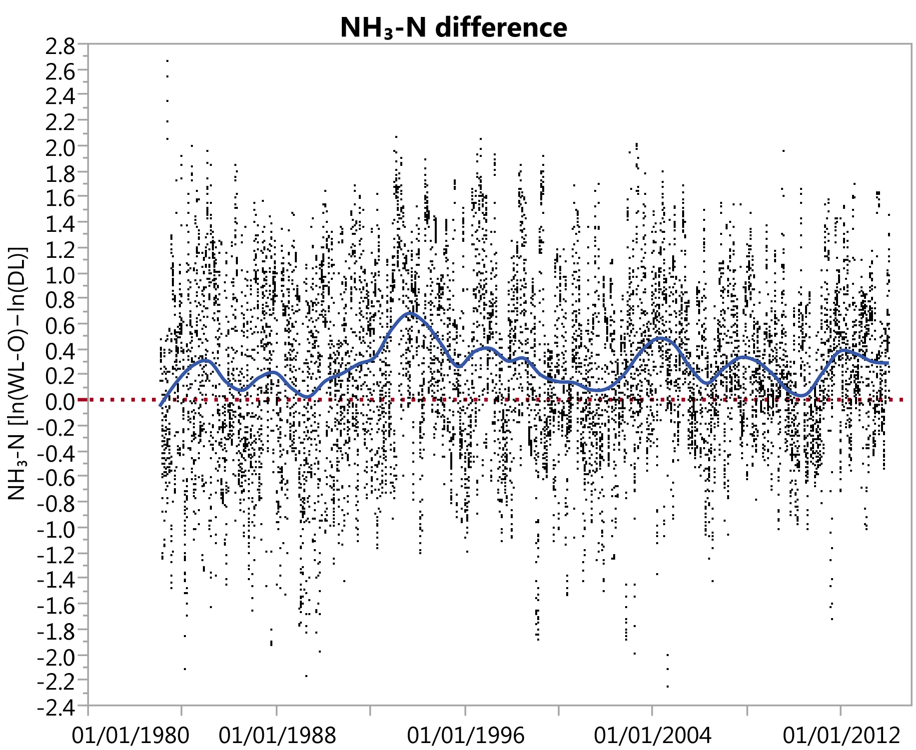

TSS diff vs. TSS Load
TSS [ln(WL-O)−ln(DL)]
TSS ln(DL)
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
-4.5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

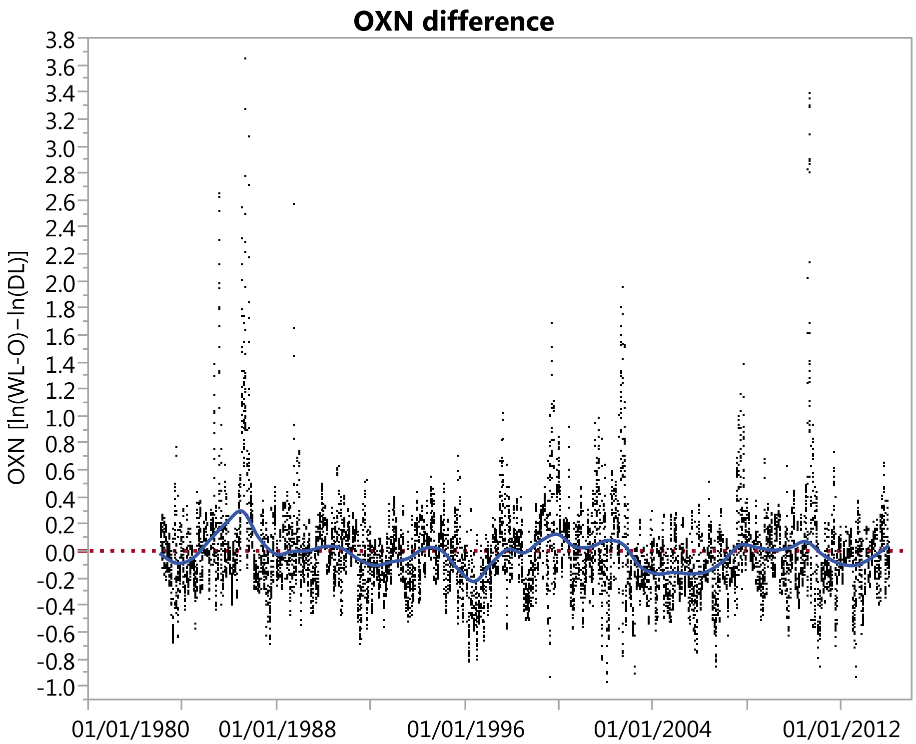
Variation in the load difference computed by two methods with time

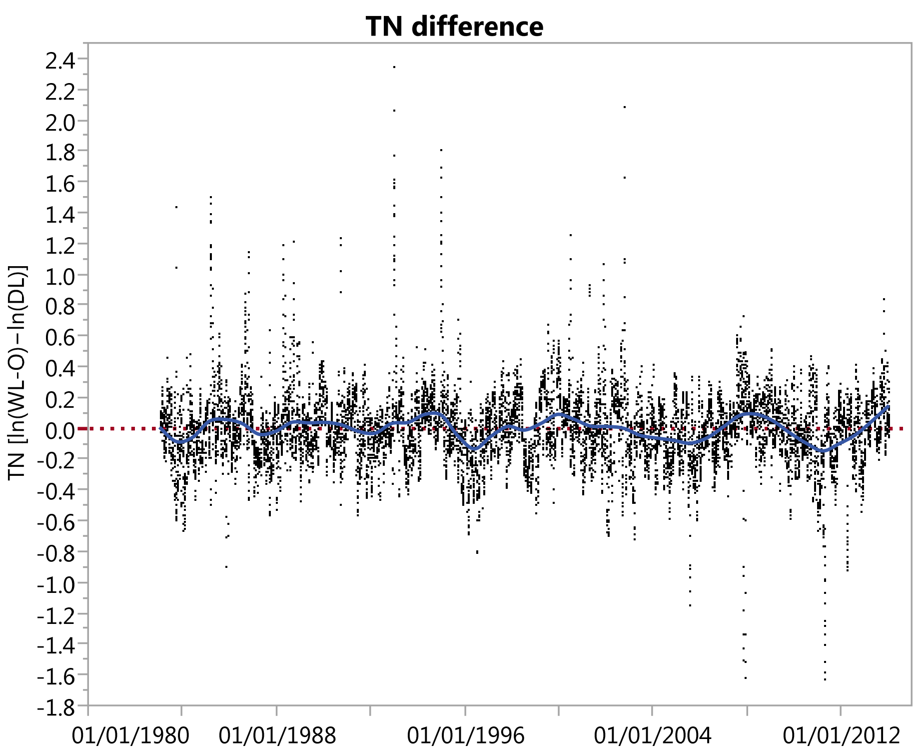


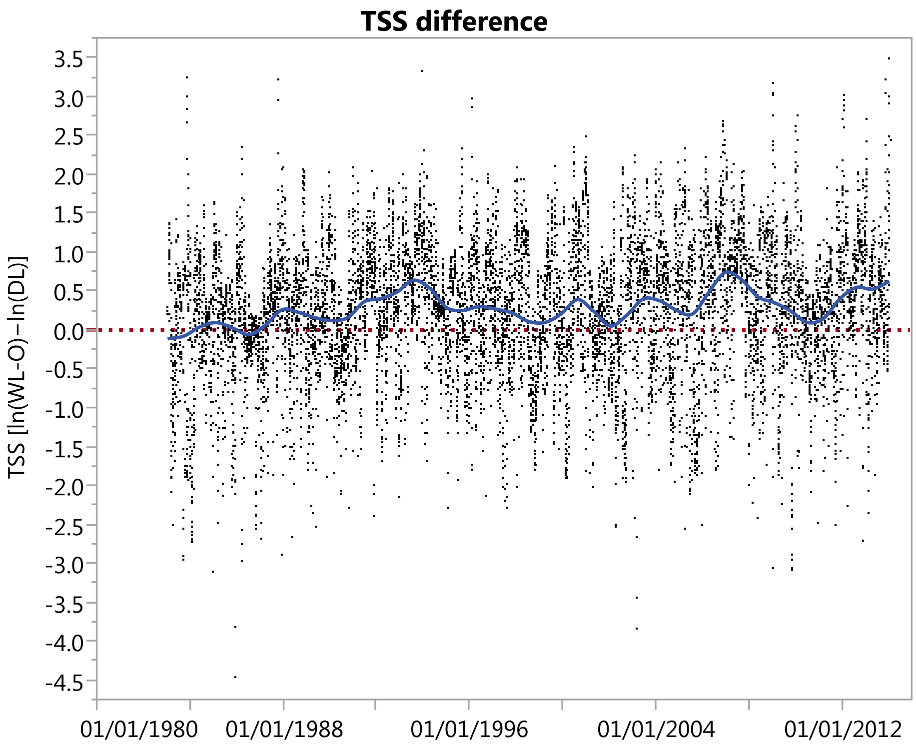



TP difference
TP [ln(WL-O)−ln(DL)]
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
-1.4
-1.6
-1.8
-2.0
-2.2
-2.4
-2.6
01/01/1980
01/01/1988
01/01/1996
01/01/2004
01/01/2012

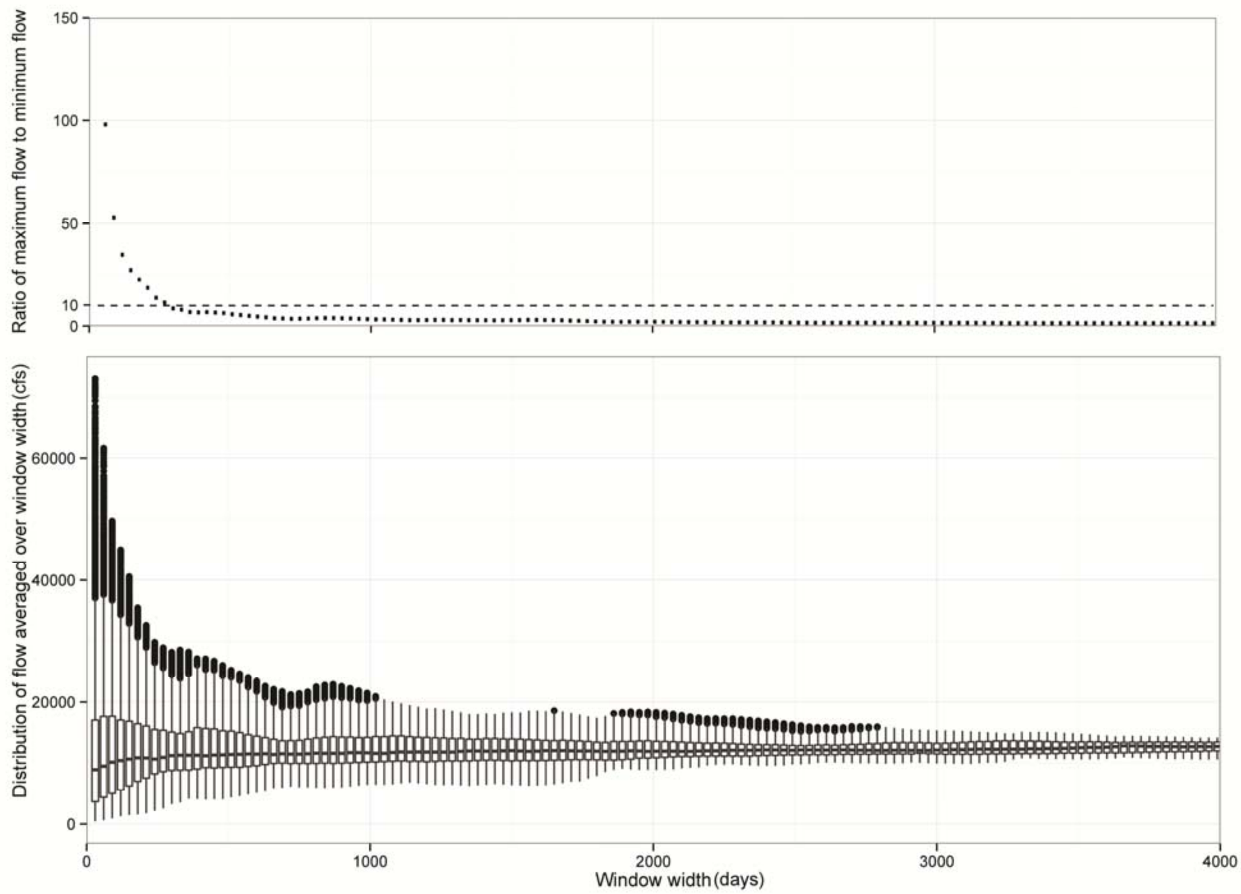

NH₃-N difference
NH₃-N [ln(WL-O)−ln(DL)]
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
-1.4
-1.6
-1.8
-2.0
-2.2
-2.4
01/01/1980
01/01/1988
01/01/1996
01/01/2004
01/01/2012


OXN difference
OXN [ln(WL-O)−ln(DL)]
3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
01/01/1980
01/01/1988
01/01/1996
01/01/2004
01/01/2012


TN difference
TN [ln(WL-O)−ln(DL)]
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
-1.4
-1.6
-1.8
01/01/1980
01/01/1988
01/01/1996
01/01/2004
01/01/2012


TSS difference
TSS [ln(WL-O)−ln(DL)]
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
-4.5
01/01/1980
01/01/1988
01/01/1996
01/01/2004
01/01/2012

Appendix H



Top: Change in ratio between the maximum and minimum average flows with increasing averaging width. The ratio is plotted for averaging window widths from 30 days to 3990 days in increments of 30 days. Around a window width of 285 days the ratio ≈10. **Bottom:** Distribution of flows in the Potomac River when averaged over the window width. Note for plotting window width increases from 30 days to 3990 days in increments of 30.