Arkansas Water Resources Center

INVESTIGATION OF OPTIMUM SAMPLE NUMBER AND TIMING FOR DETERMINING POLLUTION LOADS

A final report submitted to the
U.S. Geological Survey

by
T.S. Soerens, Ph.D., P.E.
Department of Civil Engineering
Marc A. Nelson, Ph.D., P.E.
Arkansas Water Resources Center
University of Arkansas
Fayetteville, Arkansas

Publication No. PUB-182

October 2000

Arkansas Water Resources Center
112 Ozark Hall
University of Arkansas
Fayetteville, Arkansas 72701

Prepared for United States Department of the Interior

Investigation of Optimum Sample Number and Timing for Determining Pollution Loads

Investigators: Thomas S. Soerens, University of Arkansas, Fayetteville, Arkansas
Marc A. Nelson, University of Arkansas, Fayetteville, Arkansas
Jean Spooner, North Carolina State University, Raleigh, NC

ABSTRACT

In order to determine the impacts of non point source (NPS) pollution and to develop Total Maximum Daily Loads (TMDLs), accurate measurements of pollution loads in streams are critical. The objectives of this study were to accurately determine pollutant loads at two sites by intensive storm sampling, to develop sub-sampling and other data analysis techniques, to determine the effect of sample interval on load calculation accuracy, and to find the minimum sample interval required to determine storm loads at a required accuracy. The two stream sites used were a 1st order and a 3rd order stream in the Illinois River basin in Arkansas. The samples were analyzed for NO_3-N, NH_4-N, TKN, Ortho-P, Total-P, and TSS. Storm loads were calculated by multiplying discharged volume by concentration for each sampling interval and summing over the storm. The loads calculated using the 30 minute interval data were termed the "best estimate" load. Loads were also calculated for 60, 120, and 240 minute sampling intervals using subsets of the data. The load estimates for the longer sampling intervals were expressed as a percentage of the best estimate load. The results showed that as sampling interval increased the error of the load estimate increased. For example, the Moores Creek data indicated that if we desire that the calculated TSS load is within 5% of the best estimate load with a 95% confidence level, we need to sample approximately every 50 minutes during a storm. This optimum sampling interval varies with the parameter measured and with the stream order. The information gained from this study should help water quality investigators develop sampling schemes to meet their goals of accuracy, precision, efficiency, and cost.

KEYWORDS

Hydrology, TMDL, water sampling

INTRODUCTION

Accurate measurements of pollution loads in streams are critical for determining the impacts of non point source (NPS) pollution and for developing TMDLs. A common sampling method for determining pollution loads is to continuously monitor flow and intermittently collect water samples. Loads are then calculated by multiplying the measured concentration by the discharge between sampling intervals or between the midpoint of sampling intervals. This method, in effect, assumes that the sample concentration represents the concentration in the stream between samples. This can lead to errors when the concentrations are not constant. Much or most of the load in a stream is transported during storms, and often the majority of the storm load is transported during the "first flush" or the rising limb of the hydrograph. This load is missed unless storms are intensively sampled.

Storm loads can be estimated from limited data using a load calculation model that incorporates the correlation between concentration and flow (Haines, 1997). However, usually the sampling data does not contain enough high flow concentration data to confidently make those correlations (Green, 1998), the correlations are not very strong, and the resulting load calculations can be inaccurate. Intensive storm sampling using frequent discrete samples or flow-weighted composites can be used to more accurately measure storm loads.

The purpose of this study was to determine the optimum number and timing of storm and baseflow water quality sampling to determine pollutant loads in streams with high precision and accuracy. The objectives of this study were to:

- Accurately determine pollutant loads at two sites by sampling storm runoff events at thirty-minute intervals.
- Develop sub-sampling and other data analysis techniques to determine the effect of sample interval on load calculation accuracy.

- Find the minimum sample interval required to determine storm loads at a required accuracy.

This study used two sites in the Illinois River basin in Arkansas. Nutrient loads in the Illinois River are of great interest to the states of Arkansas and Oklahoma. The two states have agreed to a 40% reduction in total phosphorus load entering Lake Tenkiller in Oklahoma. The baseline is the 1980-1993 data and a five-year moving average is used for assessment. By 1997, sampling data indicated a reduction of 22.9% in the main stem of the Illinois River (Maner, 1998). The Illinois River basin includes point and non-point pollution sources. Point sources, namely wastewater discharges, have been improved in the past couple of decades and the current focus is on non-point sources. Northwest Arkansas is home to many confined animal operations and a number of best management practices (BMPs) for agricultural non-point pollution control have been researched, demonstrated, and implemented in the region.

METHODOLOGY

The two sampling locations used were:

1. Moores Creek, a small 1st order stream with a drainage area of about 1000 hectares in the headwaters of the Illinois River. Moores Creek is impacted primarily by non-point source pollution from agriculture, forest, and low-density housing.

2. Illinois River near Siloam Springs Arkansas, a larger 3rd to 5th order stream with a drainage area of about 150,000 hectares. The Illinois River is impacted by urban point source and rural non-point source pollution.

Gauges at the sites continuously measured and recorded stage and calculated discharge. The Illinois River site is U.S. Geologicacl Survey (USGS) stream-gaging station numbered 07195430, a real time gaging station that transmits its data to the Arkansas District office in Little Rock through the Geostationary Operational Environmental Satellite (GOES). This station is located at the Oklahoma-Arkansas topographic map Watt Quadrangle, T.17N., R.33W., NE ¼, NE ¼, NE ¼, SE ¼, Sec. 31, latitude 3606'31", longitude 9432'00", 5.0 miles south of Siloam Springs. It was put in operation July 14, 1995 until current year in the cooperation with the Arkansas Soil and Water Conservation Commission (ASWCC).

Automatic samplers installed at the sites were triggered by the stage gauges to take storm samples at 30 minute intervals during the rising limb and 60 minute intervals during the falling limb of the storm hydrographs. All samples were collected from the sites within 24 hours and analyzed at the Arkansas Water Quality Lab using U.S. Environmental Protection Agency (EPA) approved analysis and QA/QC procedures. The samples were analyzed for NO_3-N, NH_4-N, TKN, Ortho-P, Total-P, and TSS. The data are available in Microsoft Excel spreadsheets upon request from the Arkansas Water Resources Center.

Storm loads were calculated by multiplying discharged volume by concentration for each sampling interval and summing over the storm. The loads calculated using the 30 minute interval data were termed the "best estimate" load. Loads were also calculated for 60, 120, and 240 minute sampling intervals using subsets of the data. The load estimates for the longer sampling intervals were expressed as a percentage of the best estimate load.

Optimum sampling intervals were calculated as the sampling interval that gives an estimate within 5% of the best estimate load with 95% percent confidence. An approximate 95% confidence interval for the mean load estimate error for each sampling interval was calculated as the mean plus or minus 1.96 multiplied by the standard deviation. A regression line was then fit to the upper and lower confidence levels versus sampling interval (forced to 100% at t=0). The optimum sampling interval is the greater of the intervals where the upper or lower line crosses 5% error. The 95% confindence interval and 5% error were the optimum sampling interval criteria chosen in this study. Different confidence levels and different error criteria could be chosen and would lead to different optimum sampling intervals. Appendix A, adapted from Lo (2000) provides more detail on the load and optimum sampling interval calculations.

RESULTS

Table 1 summarizes the storms analyzed at the Illinois River Hwy 59 site in 1998 and Table 2 summarizes the 1999 storms at the Illinois River Hwy 59 site. Tables 3 and 4 summarize the Moores Creek site storms for 1998 and 1999, respectively.

Table 1. Illinois Hwy 59 site 1998 storms.

Storm	Date Started	Date Ended	Time Started	Time Ended	Duration (hrs.)	Max. Stage (ft)	Total Discharge (ft^3)
1	11/29/97	11/30/97	3:35 AM	3:05 PM	35.5	4.6	51,639,011
2	12/21/97	12/22/97	10:11 AM	10:43 AM	24.5	6.2	107,524,669
3	12/24/97	12/26/97	2:43 AM	8:43 AM	54.0	9.0	429,903,575
4	1/4/98	1/6/98	7:13 AM	11:13 AM	52.0	19.3	2,264,694,658
5	1/7/98	1/9/98	1:24 PM	12:25 PM	48.0	12.2	1,103,129,415
6	1/26/98	1/27/98	2:16 AM	12:10 PM	33.9	6.1	116,938,622
7	2/26/98	2/27/98	12:30 AM	11:00 AM	34.5	7.6	220,993,756
8	3/6/98	3/7/98	1:27 AM	8:30 AM	31.0	6.3	109,712,833
9	3/7/98	3/9/98	6:23 PM	3:09 PM	44.8	10.1	488,585,793

Table 2. Illinois Hwy 59 site 1999 storms.

Storm	Date Started	Date Ended	Time Started	Time Ended	Duration (hrs.)	Max. Stage (ft)	Total Discharge (ft^3)
1	1/30/99	2/1/99	9:00 AM	8:30 AM	35.5	6.3	132,475,532
2	2/6/99	2/8/99	7:30 AM	9:00 AM	49.5	12.3	573,579,786
3	5/4/99	5/6/99	3:30 PM	9:30 AM	42.0	14.8	1,016,154,362
4	5/10/99	5/12/99	8:00 PM	8:00 AM	36.0	9.0	285,502,249
5	5/22/99	5/24/99	9:00 PM	10:00 AM	37.0	9.4	385,774,702

Table 3. Moores Creek 1998 storms.

Storm	Date Started	Date Ended	Time Started	Time Ended	Duration (hrs.)	Max. Stage (cm)	Total Discharge (ft^3)
1	12/24/97	12/24/97	12:32 AM	5:08 PM	16.5	37.3	3880527
2	1/4/98	1/5/98	8:28 AM	4:28 AM	20.0	92.6	20514420
3	1/7/98	1/8/98	1:29 PM	1:05 PM	23.5	44.9	7599363

Table 4. Moores Creek 1999 storms.

Storm	Date Started	Date Ended	Time Started	Time Ended	Duration (hrs.)	Max. Stage (in)	Total Discharge (ft^3)
1	11/30/98	12/1/98	1:23 AM	12:27 PM	35.0	38.8	3485618
2	1/30/99	2/1/99	4:23 PM	9:37 AM	41.0	29.7	2903436
3	2/6/99	2/8/99	8:15 PM	9:31 AM	37.5	46.7	5184836
4	5/22/99	5/24/99	5:21 PM	8:42 AM	39.5	31.7	2982485
5	3/19/00	3/20/00	5:22 AM	3:24 AM	23.0	47.6	7033556

Figure 1 shows a plot of the discharge and the relative TSS concentrations at the Illinois River Hwy 59 site during 1999 storm #3 starting 5/4/99. We see that concentrations are higher during the rising limb of the hydrograph, coming in several slugs. Note that after the spike on the rising limb of the hydrograph, the discharge remains high but the TSS concentration has decreased. Ammonium-N, TKN, total P, and PO_4-P concentrations followed the pattern of the TSS concentrations. Nitrate decreased with higher flows, showing a dilution effect.

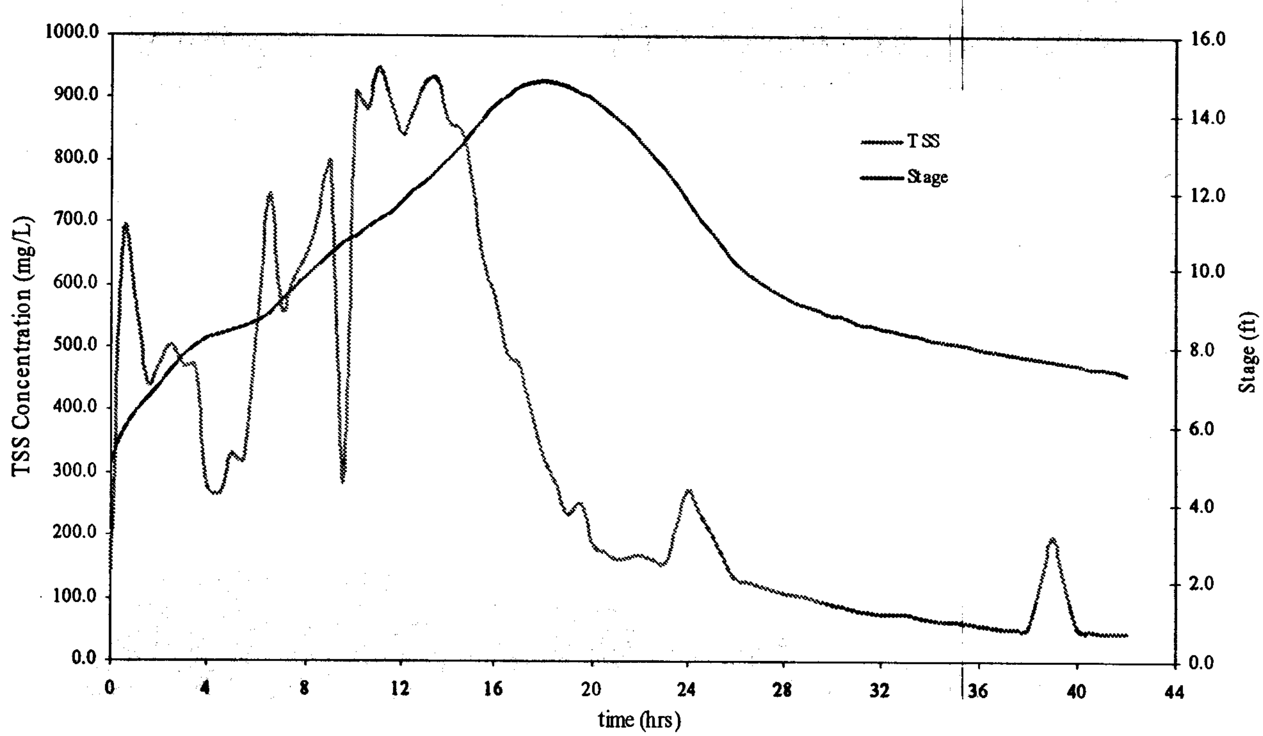

Figure 1. 5/4/99 Storm discharge and TSS concentration - Illinois River at Hwy 59

Figures 2 shows a plot of percent of storm load versus percent of volume for 1998 storm #4 at the Illinois River Hwy 59 site. We see that more than 50% of the TSS load is transported during the first 20% of the discharge volume. Because of the high concentrations during the rising limb of the hydrograph, all of the parameters except nitrate have most of their load transported during the first portion of the storm.

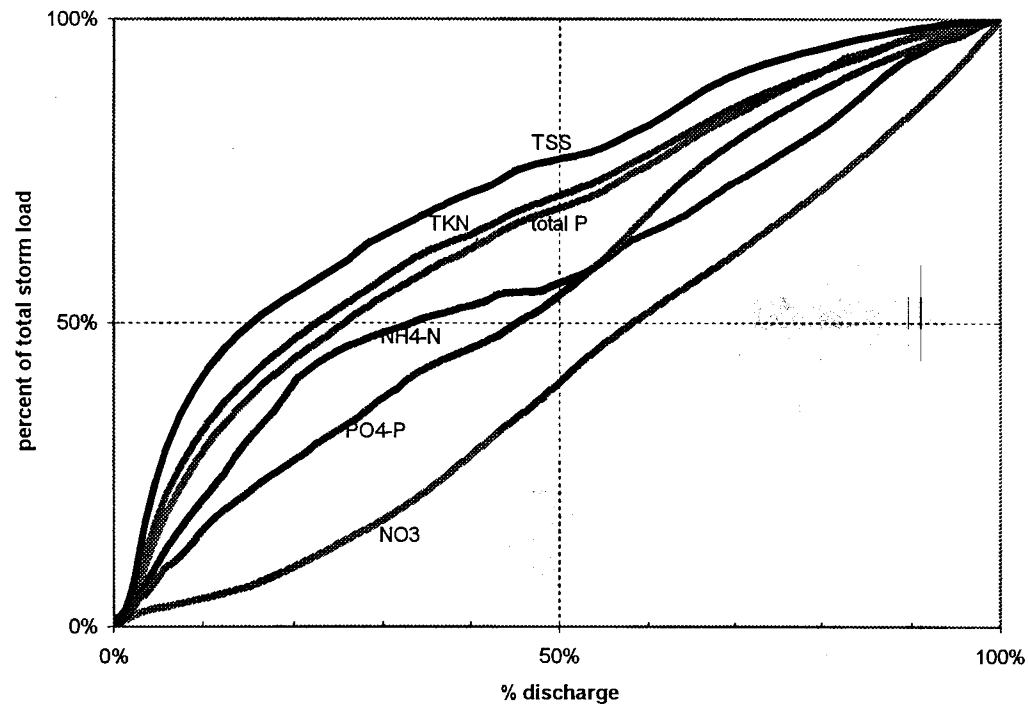

Figure 2. Percent load versus percent volume - Illinois River 1998 storm #4.

Figure 3 shows the calculated TKN loads for nine 1998 storms at the Illinois River Hwy 59 site as a percent of the best estimate load plotted versus sampling interval. We see that as sampling interval increased the error of the load estimate increased. The optimum sampling interval calculation is demonstrated on Figure 3. If we desire that the calculated load is within 5% of the best estimate load with a 95% confidence level, Figure 3 tells us that we need to sample approximately every 4.5 hours during a storm. With these criteria, the optimum sampling interval for TKN at the Illinois River Hwy 59 site is 4.5 hours.

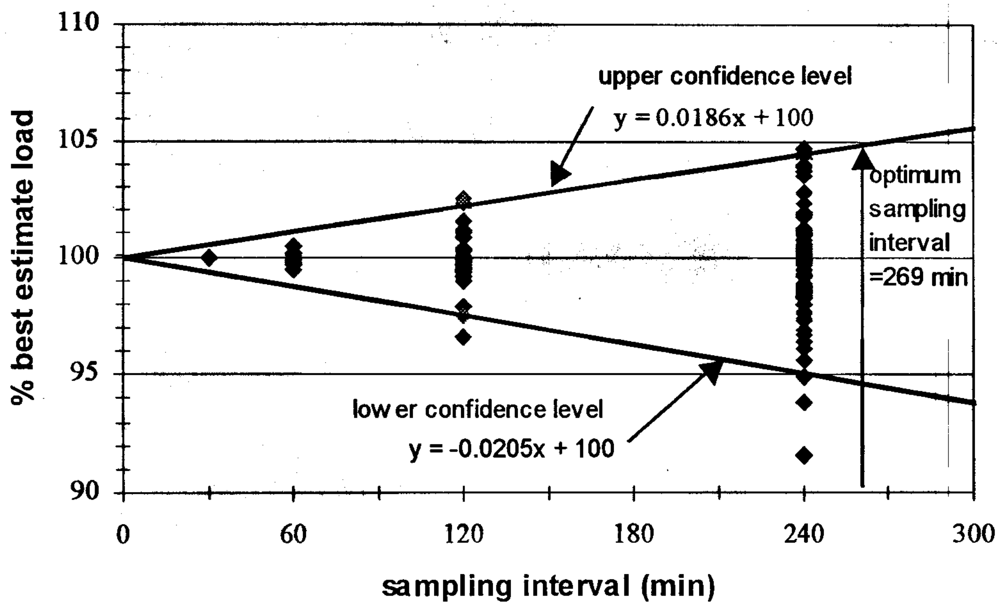

Figure 3. Optimum sampling interval calculation method

Table 5 shows the optimum sampling intervals for the Illinois River 1998 data. The "average" and the "median" columns show the mean and median of the nine optimum sampling intervals. The column named "Population" is the optimum sampling intervals calculated from the data of all nine storms put together.

Table 5 Illinois River 1998 Optimum Sampling Intervals (hr)

	Storm											
	1	2	3	4	5	6	7	8	9	Average	Median	Population
NO_3-N	30.86	20.83	23.15	14.88	46.30	55.56	23.15	21.93	19.38	28.45	23.15	22.52
Total-P	5.91	4.73	15.43	2.16	4.41	6.72	3.13	3.16	4.68	5.59	4.68	4.07
NH_4-N	1.14	0.63	2.37	1.61	5.38	0.59	1.64	0.94	0.54	1.65	1.14	0.96
TKN	4.15	4.68	12.25	3.22	11.90	2.69	2.47	9.06	4.48	6.10	4.48	4.30
PO_4-P	8.77	7.44	28.74	1.53	10.82	10.16	4.50	4.93	4.87	9.09	7.44	4.11
TSS	1.14	3.75	2.94	4.23	12.44	3.59	1.41	4.99	1.25	3.97	3.59	2.27

Table 6 shows the optimum sampling intervals for the Illinois River 1999 data. Table 7 shows the average optimum sampling intervals for both years of the Illinois River data.

Table 6. Illinois River 1999 Optimum Sampling Intervals (hr)

	Storm						
	1	2	3	4	5	average	median
NO_3-N	46.30	18.94	6.22	3.33	21.37	19.23	18.94
Total-P	4.23	3.95	2.50	1.92	5.75	3.67	3.95
NH_4-N	5.67	3.05	0.69	0.76	3.12	2.66	3.05
TKN	5.18	4.15	4.01	2.10	4.82	4.05	4.15
PO_4-P	8.42	3.31	1.06	4.50	23.81	8.22	4.5
TSS	3.84	3.75	1.62	1.40	2.85	2.69	2.85

Table 7. Illinois River optimum sampling intervals (hr) for two years – 1998 & 1999

	average	median
NO3-N	25.16	21.65
Total-P	4.91	4.32
NH3-N	2.01	1.38
TKN	5.37	4.32
TOC	8.78	6.19
TSS	3.51	3.27

Table 8 and table 9 show the optimum sampling intervals for Moores Creek for the 1998 and 1999 data, respectively. Table 10 shows the average optimum sampling intervals Moores Creek for the two years together.

Table 8. Moores Creek 1998 optimum sampling intervals (hr)

	Storm				
	1	2	3	average	median
NO3-N	1.38	2.08	9.36	4.27	2.08
Total-P	5.95	0.75	3.97	3.56	3.97
NH3-N	1.70	0.59	1.63	1.31	1.63
TKN	4.21	0.66	7.31	4.06	4.21
TOC	5.79	0.97	9.92	5.56	5.79
TSS	1.42	0.49	1.71	1.21	1.42

Table 9. Moores Creek 1999 Optimum Sampling Intervals (hr)

	1	2	3	4	5	average	median
NO3-N	5.91	5.83	2.65	1.87	11.90	5.63	5.83
Total-P	1.02	0.78	0.44	1.79	1.61	1.13	1.02
NH4-N	0.70	0.86	0.76	1.37	7.58	2.25	0.86
TKN	1.24	1.34	0.40	3.03	1.23	1.45	1.24
PO4-P	2.05	2.01	2.24	3.17	30.86	8.07	2.24
TSS	0.51	0.31	0.34	1.19	2.16	0.90	0.51

Table 10. Moores Creek Optimum Sampling Intervals

for Two Years - 1998 and 1999

	average	median
NO3-N	5.12	4.24
Total-P	2.04	1.32
NH3-N	1.90	1.12
TKN	2.43	1.29
TOC	7.13	2.70
TSS	1.02	0.85

The optimum sampling interval varies with the parameter measured and with the stream order. Figure 4 shows the two-year median optimum sampling interval for the six measured parameters for both sites. We see that the pollutants that show the most "peaking" effect during a storm require the smallest sampling interval. Also, Moores Creek, which is flashier than the main branch of the Illinois River, requires a shorter sampling interval for accurate storm load calculations.

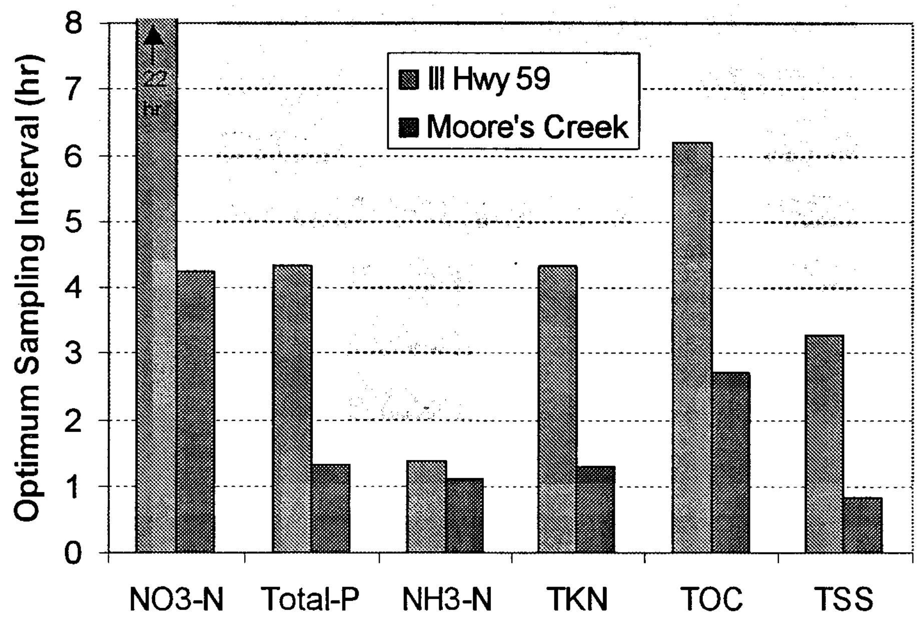

Figure 4. Optimum sampling intervals – two-year medians

DISCUSSION

The optimum sampling interval appears to be a function of the parameter measured and of the drainage basin size or stream order. It is desirable to quantify the effect of these variables on the optimum sampling interval. Figure 5 shows a possible relationship between sample interval and basin size.

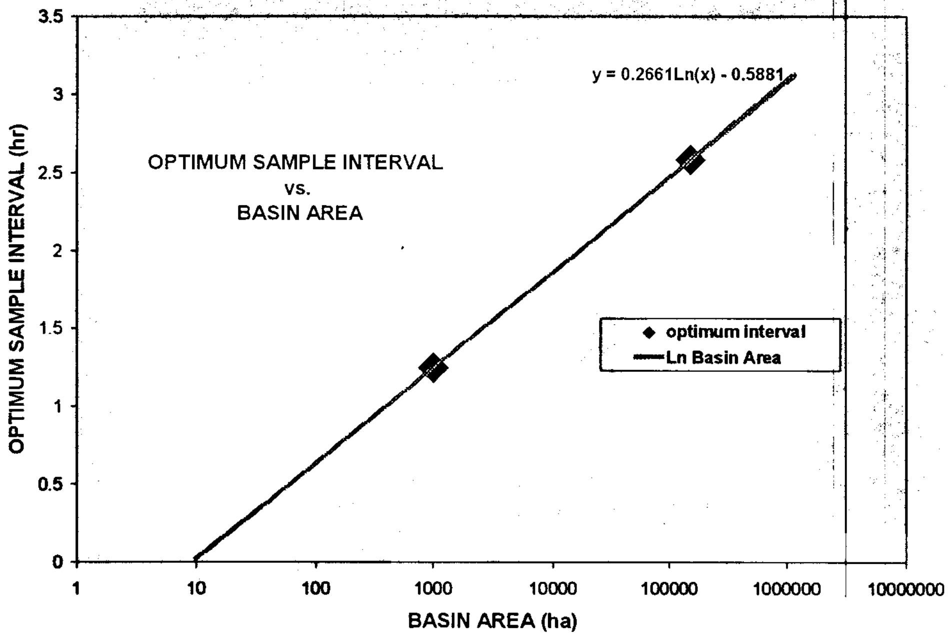

Figure 5. Possible relationship between basin size and optimum sampling intervals.

The data set that has been compiled in this project can be used to examine the effects of various variables on storm load determination. The data set and the methods developed will help in designing sampling and monitoring systems.

CONCLUSIONS

From this study we make the following conclusions:

- The sample interval affects load calculation precision and accuracy.
- An optimum sample interval can be calculated using sub-sampling techniques.
- The optimum sample interval varies by parameter measured and by drainage basin size.

REFERENCES

Haines, K., An Evaluation of Techniques used to Determine Pollution Loads, M.S. Thesis, University of Arkansas, Fayetteville, Arkansas, 1997.

Lo, Yian-Mei, Optimum Storm Sampling For The Calculation Of Nutrient Loads In The Illnois River, M.S. Project Report, University of Arkansas Department of Civil Engineering, Fayetteville, Arkansas, 2000.

Green, R., Flow Concentration Relations and Phosphorus Loading in the Illinois River, in Steele, K.F. Ed., Quality of Surface and Ground Water and Best Management Practices, Proceedings of the Arkansas Water Resources Center Annual Research Conference, April 8 and 9, 1998, AWRC Publication No. 270, September 1998, p.19.

Maner, M., Recent Total Phosphorus Loads in the Illinois River Watershed in Arkansas Compared to Loads in 1980-93, Arkansas Department of Pollution Control and Ecology, March 1998.

Appendix A

Load and Optimum Sampling Interval Calculations

DATA ANALYSIS AND LOAD CALCULATIONS

The study of the optimum sampling interval (OSI) was essentially focused on the pollutant concentration collected at an intensive sampling interval, which was the 30-minute sampling interval during the rising portion and the 60-minute sampling interval during the falling portion of a storm event. The non-fixed, infrequent sampling interval during the stream base flow of the entire hydrograph was not considered into the project's NPS pollutant storm load analysis. Hence, the selected raw data for this project that consisted of measured stages, calculated discharges, and six analyzed pollutant concentrations had to be classified according to its individual storm. There were a total of nine storms during the period from November 29, 1997 to March 9, 1998 that were analyzed in this project. For each storm event considered, the raw data was recompiled into a spreadsheet similar to Table A-1, having pollutant concentrations at 30-minute sampling intervals on the rising portion of the hydrograph and 60-minute sampling intervals on the falling portion of the hydrograph. Such data, as seen in Table A-1, was compiled for each of the nine storm events.

Table A-1. Compiled raw data of time, data, stage, discharge, and pollutant concentrations of Illinois River Hwy 59 1998 storm #1.

11/29/97				max	2.036	0.480	0.193	0.853	4.119	62.000
Time	Stage	Discharge (cfs)	Specimen No.		NO_3-N	Total-P	NH_4-N	TKN (mg/L)	PO_4-P	TSS
3:35 AM	3.99	235.85	981140		1.964	0.255	0.043	0.616	0.178	62.000
4:05 AM	4.07	261.57	981141		1.977	0.200	0.031	0.302	0.177	12.600
4:35 AM	4.17	295.60	981142		1.979	0.235	0.018	0.236	0.176	9.350
5:05 AM	4.24	320.65	981143		1.976	0.220	0.027	0.279	0.174	11.050
5:35 AM	4.29	339.18	981144		2.016	0.215	0.025	0.505	0.199	11.650
6:05 AM	4.34	358.22	981145		2.026	0.280	0.058	0.469	0.243	16.350
6:35 AM	4.41	385.76	981146		1.992	0.355	0.065	0.610	0.289	22.650
7:05 AM	4.48	414.32	981147		1.975	0.420	0.101	0.733	0.344	25.650
7:35 AM	4.54	435.35	981148		1.984	0.435	0.120	0.753	0.365	29.900
8:05 AM	4.58	452.54	981149		2.027	0.450	0.116	0.853	0.373	28.300
8:35 AM	4.6	461.27	981150		2.029	0.480	0.109	0.850	0.397	26.900
9:05 AM	4.6	461.27	981151		1.960	0.470	0.111	0.825	0.375	22.800
9:35 AM	4.59	456.90	981152		1.983	0.460	0.091	0.813	0.370	19.000
10:05 AM	4.59	456.90	981153		1.978	0.405	0.062	0.717	0.342	18.100
10:35 AM	4.58	452.54	981154		1.987	0.380	0.050	0.653	0.303	14.900
11:05 AM	4.58	452.54	981155		1.984	0.325	0.029	0.563	0.274	18.000
11:35 AM	4.59	456.90	981156		1.956	0.290	0.038	0.537	0.252	13.700
12:05 PM	4.58	452.54	981157		2.005	0.280	0.025	0.680	0.235	12.800
12:35 PM	4.59	456.90	981158		2.016	0.265	0.027	0.500	0.223	10.650
1:05 PM	4.59	456.90	981159		1.994	0.260	0.025	0.470	0.215	9.150
1:35 PM	4.58	452.54	981160		2.025	0.210	0.014	0.410	0.212	9.450
2:05 PM	4.56	443.90	981161		2.002	0.200	0.023	0.404	0.206	13.250
2:35 PM	4.56	443.90	981162		2.036	0.210	0.024	0.398	0.200	8.700
3:05 PM	4.55	439.62	981163		2.008	0.200	0.021	0.434	0.197	8.000
3:35 PM	4.54	435.35			2.019	0.193	0.013	0.386	0.197	8.025
4:05 PM	4.52	426.87	981164		2.029	0.185	0.0045	0.338	0.197	8.050
4:35 PM	4.51	422.67			1.103	0.100	0.193	0.309	4.119	4.750
5:05 PM	4.51	422.67	981165		2.020	0.195	0.048	0.421	0.187	9.500
5:35 PM	4.50	418.48			2.017	0.200	0.035	0.393	0.189	9.450
6:05 PM	4.49	418.48	981166		2.013	0.205	0.022	0.364	0.190	9.400
6:35 PM	4.48	414.32			2.017	0.200	0.023	0.343	0.195	9.575
7:05 PM	4.47	410.18	981167		2.021	0.195	0.023	0.321	0.199	9.750
7:35 PM	4.47	410.18			2.022	0.200	0.025	0.330	0.195	10.075
8:05 PM	4.46	406.06	981168		2.023	0.205	0.027	0.338	0.191	10.400
8:35 PM	4.46	406.06			2.014	0.200	0.023	0.325	0.188	11.000
9:05 PM	4.45	401.95	981169		2.004	0.195	0.018	0.311	0.185	11.600
9:35 PM	4.45	401.95			2.000	0.190	0.028	0.308	0.185	10.850
10:05 PM	4.45	401.95	981170		1.997	0.185	0.037	0.304	0.184	10.100

On the falling portion of the storm hydrograph in all nine data sets, the 30 minutes missing pollutant concentration from the 60-minute sampling interval was estimated by linear interpolation between two sampled concentration values. An example of linear interpolation applied to part of the storm hydrographic data and the calculated discharge volume are shown in Table A-2. The stage hydrograph and the pollutant concentration calculated for every 30 minutes during those nine separate storm events were plotted with time. The discharge volume for the 30-minute interval hydrograph of the storm event analyzed is calculated as follows:

$$V_t = \left(\frac{t_i - t_{(i-30)}}{2} + \frac{t_{(i+30)} - t_i}{2} \right) * Q * 86400 \qquad (1)$$

where V_t is the volume at time t (ft^3)

t_i is the time at time t (day)

$t_{(i-30)}$ is the time 30 minutes before time t_i

$t_{(i+30)}$ is the time 30 minutes after time t_i

Q is the discharge at time t (ft^3/sec)

86400 is the conversion factor of $\frac{\text{second}}{\text{day}}$

Likewise, the pollutant load for the 30-minute interval hydrograph is the product of the discharge volume calculated earlier and the pollutant concentration, at time t:

$$L_t = \frac{Ct * Vt * 62.428}{1{,}000{,}000} \qquad (2)$$

where L_t is the pollutant load at time t (lb)

C_t is the pollutant concentration at time t (mg/L)

V_t is the volume at time t (ft^3)

62.428 is the conversion factor of $\frac{L \cdot lb}{ft^3 \cdot kg}$

1,000,000 is the conversion factor of mg/kg

Table A-2. Linear interpolation of the 60-minute pollutant concentrations and the calculated discharge volume of Illinois River Hwy 59 1998 storm #1.

11/29/97

Time	Stage	Discharge (cfs)	Specimen No.	NO_3-N	Total-P	NH_4-N	TKN	PO_4-P	TSS	Volume (cf)
				(mg/L)						
3:35 AM	3.99	235.85	981140	1.964	0.255	0.043	0.616	0.178	62.000	
4:05 AM	4.07	261.57	981141	1.977	0.200	0.031	0.302	0.177	12.600	470830
4:35 AM	4.17	295.60	981142	1.979	0.235	0.018	0.236	0.176	9.350	532073
5:05 AM	4.24	320.65	981143	1.976	0.220	0.027	0.279	0.174	11.050	577175
5:35 AM	4.29	339.18	981144	2.016	0.215	0.025	0.505	0.199	11.650	610517
6:05 AM	4.34	358.22	981145	2.026	0.280	0.058	0.469	0.243	16.350	644798
6:35 AM	4.41	385.76	981146	1.992	0.355	0.065	0.610	0.289	22.650	694367
7:05 AM	4.48	414.32	981147	1.975	0.420	0.101	0.733	0.344	25.650	745777
7:35 AM	4.54	435.35	981148	1.984	0.435	0.120	0.753	0.365	29.900	783625
8:05 AM	4.58	452.54	981149	2.027	0.450	0.116	0.853	0.373	28.300	814579
8:35 AM	4.6	461.27	981150	2.029	0.480	0.109	0.850	0.397	26.900	830283
9:05 AM	4.6	461.27	981151	1.960	0.470	0.111	0.825	0.375	22.800	830283
9:35 AM	4.59	456.90	981152	1.983	0.460	0.091	0.813	0.370	19.000	822412
10:05 AM	4.59	456.90	981153	1.978	0.405	0.062	0.717	0.342	18.100	822412
10:35 AM	4.58	452.54	981154	1.987	0.380	0.050	0.653	0.303	14.900	814579
11:05 AM	4.58	452.54	981155	1.984	0.325	0.029	0.563	0.274	18.000	814579
11:35 AM	4.59	456.90	981156	1.956	0.290	0.038	0.537	0.252	13.700	822412
12:05 PM	4.58	452.54	981157	2.005	0.280	0.025	0.680	0.235	12.800	814579
12:35 PM	4.59	456.90	981158	2.016	0.265	0.027	0.500	0.223	10.650	822412
1:05 PM	4.59	456.90	981159	1.994	0.260	0.025	0.470	0.215	9.150	822412
1:35 PM	4.58	452.54	981160	2.025	0.210	0.014	0.410	0.212	9.450	814579
2:05 PM	4.56	443.90	981161	2.002	0.200	0.023	0.404	0.206	13.250	799028
2:35 PM	4.56	443.90	981162	2.036	0.210	0.024	0.398	0.200	8.700	799028
3:05 PM	4.55	439.62	981163	2.008	0.200	0.021	0.434	0.197	8.000	791307
3:35 PM	4.54	435.35		2.019	0.193	0.013	0.386	0.197	8.025	783625
4:05 PM	4.52	426.87	981164	2.029	0.185	0.0045	0.338	0.197	8.050	768374
4:35 PM	4.51	422.67		2.025	0.190	0.026	0.380	0.192	8.775	760804
5:05 PM	4.51	422.67	981165	2.020	0.195	0.048	0.421	0.187	9.500	760804
5:35 PM	4.50	418.48		2.017	0.200	0.035	0.393	0.189	9.450	753271
6:05 PM	4.49	418.48	981166	2.013	0.205	0.022	0.364	0.190	9.400	753271
6:35 PM	4.48	414.32		2.017	0.200	0.023	0.343	0.195	9.575	745777
7:05 PM	4.47	410.18	981167	2.021	0.195	0.023	0.321	0.199	9.750	738320
7:35 PM	4.47	410.18		2.022	0.200	0.025	0.330	0.195	10.075	738320
8:05 PM	4.46	406.06	981168	2.023	0.205	0.027	0.338	0.191	10.400	730901
8:35 PM	4.46	406.06		2.014	0.200	0.023	0.325	0.188	11.000	730901
9:05 PM	4.45	401.95	981169	2.004	0.195	0.018	0.311	0.185	11.600	723519

linear interpolation applied at the 60-minute sampling interval

This 30-minute load interval method is statistically correct with the assumption that the pollutant concentration and the flow discharge are representative of the flow interval (Yaksich et al., 1983). Each pollutant's calculated load at the 30-minute interval in all nine data sets was summed. Each pollutant's total load then becomes a critical data in the determination of the optimum NPS sampling strategy. The total loads from each 30-minute interval data were numerically compared to total pollutant loads form other sampling intervals. A confidence level could therefore be applied to examine the extent of statistical bias of each sampling interval.

Based on the 30-minute interval data set for each storm, each pollutant load was calculated at 60, 120 and 240-minute sampling intervals. For each of these sampling intervals, pollutant concentrations for each pollutant were arranged in a temporal sequence with respect to their sampling intervals. For example, the first sequence of the 60-minute sampling interval consisted of pollutant concentrations from the 30-minute sampling interval at time 1:00 PM, 2:00 PM, 3:00 PM, and so forth. The second sequence would be those pollutant concentrations at time 1:30 PM, 2:30 PM, 3:30 PM, and so forth. Table A-3 shows NO_3-N concentrations for part of storm #1 at 30-minute interval and their corresponding concentrations in ordered sequences for the other three intervals. Thus, there were total two sequences for 60-minute sampling interval, four sequences for 120-minute sampling interval, and eight sequences for 240-minute sampling interval.

Table A-3 NO_3-N concentrations used for different sampling intervals

11/29/97

Time	Specimen No.	NO_3-N (mg/L)	60-Minute 1	60-Minute 2	120-Minute 1	120-Minute 2	120-Minute 3	120-Minute 4	240-Minute 1	240-Minute 2	240-Minute 3	240-Minute 4	240-Minute 5	240-Minute 6	240-Minute 7	240-Minute 8
3:35 AM	981140	1.964	1.964		1.964				1.964							
4:05 AM	981141	1.977		1.977		1.977				1.977						
4:35 AM	981142	1.979	1.979				1.979				1.979					
5:05 AM	981143	1.976		1.976				1.976				1.976				
5:35 AM	981144	2.016	2.016		2.016								2.016			
6:05 AM	981145	2.026		2.026		2.026								2.026		
6:35 AM	981146	1.992	1.992				1.992								1.992	
7:05 AM	981147	1.975		1.975				1.975								1.975
7:35 AM	981148	1.984	1.984		1.984				1.984							
8:05 AM	981149	2.027		2.027		2.027				2.027						
8:35 AM	981150	2.029	2.029				2.029				2.029					
9:05 AM	981151	1.960		1.960				1.960				1.960				
9:35 AM	981152	1.983	1.983		1.983								1.983			
10:05 AM	981153	1.978		1.978		1.978								1.978		
10:35 AM	981154	1.987	1.987				1.987								1.987	
11:05 AM	981155	1.984		1.984				1.984								1.984
11:35 AM	981156	1.956	1.956		1.956				1.956							
12:05 PM	981157	2.005		2.005		2.005				2.005						
12:35 PM	981158	2.016	2.016				2.016				2.016					
1:05 PM	981159	1.994		1.994				1.994				1.994				
1:35 PM	981160	2.025	2.025		2.025								2.025			
2:05 PM	981161	2.002		2.002		2.002								2.002		
2:35 PM	981162	2.036	2.036				2.036								2.036	
3:05 PM	981163	2.008		2.008				2.008								2.008

Sampling Intervals and Their Sequences

For load calculations, each color signifies which and how the pollutant concentration is estimated:
linear interpolation between two sampled values
picked from the previous sampled value
picked from the latter sampled value

The estimation of missing 30-minute pollutant concentrations from the storm data is important in an accurate calculation of pollutant loads. For the 60-minute sampling interval, the pollutant load was computed according to Equation 2. The 30-minute sampling intervals missing pollutant concentration was estimated by linear interpolation between two sampled concentration values (Table A-3):

$$L_t = \frac{\left(\dfrac{C_{(i-30)} + C_{(i+30)}}{2}\right) * V_t * 62.428}{1{,}000{,}000} \qquad (3)$$

where L_t is the pollutant load at time t (lb)

$C_{(i-30)}$ is the pollutant concentration at 30 minutes before time t (mg/L)

$C_{(i+30)}$ is the pollutant concentration at 30 minutes after time t (mg/L)

V_t is the volume at time t (ft^3)

Likewise, Equation 3 was similarly applied to the load calculation for both 120 and 240-minute sampling intervals. The method applied to "best" estimate the missing 30-minute sampling interval concentrations was shown in Table A-3. Missing concentrations colored blue at the middle of the sampling interval were estimated by linear interpolation between two sampled values. Missing concentrations colored green before the middle concentration were substituted with the previous sampled value. Missing concentrations colored purple after the middle concentration were substituted with the latter sampled value.

For each storm event, the calculated load of each pollutant in all four sampling intervals was summed. Due to the pollutant concentrations' temporal sequence as mentioned earlier, there were two total pollutant loads for each pollutant at the 60-minute sampling interval, four total pollutant loads at the 120-minute sampling interval, and eight total pollutant loads at the 240-minute sampling interval. The total pollutant load for each sampling interval's several sequences was then compared to the corresponding total pollutant load for the 30-minute interval. The comparison yielded a statistical analysis called the Best Estimate Load (BEL), in percent. This could be used to determine the pollutant's OSI. The equation of the BEL is as follow:

$$\text{Percent of BEL (\%)} = \frac{\text{load at other}}{\text{load at 30 - minute interval}} \times 100\% \qquad (4)$$

Again, there were two BELs, in percent, for each pollutant at the 60-minute sampling interval, four at the 120-minute sampling interval, and eight at the 240-minute sampling interval. Table A-4 shows the BELs of NO_3-N for each sampling interval of storm #1.

Table A-4. NO_3-N percent of best estimate loads in percent for each sampling interval of Illinois River Hwy 59 1998 storm #1.

Sampling Interval (min)	NO_3-N Best Estimate Loads (%)							
	1	2	3	4	5	6	7	8
0								
30	100.0							
60	100.1	99.9						
90								
120	99.9	100.1	100.2	99.8				
150								
180								
210								
240	99.6	100.2	100.3	99.8	100.2	100.0	100.1	99.6
270								
300								

All the calculated percent of each pollutant BELs with sampling intervals were plotted on graphs. A 95.0 percent confidence level with ± 5.0 percent standard error was applied to the entire number of BELs for each sampling interval (Montgomery and Runger, 1994):

$$95.0 \text{ percent confidence level} = \bar{x} \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right) \quad (5)$$

where $\bar{x}$ is the mean value of pollutant BELs in percent within a group for each sampling interval

±1.96 is the upper and lower percentage point of the 95.0 percent standard normal distribution

$\sigma / \sqrt{n}$ is the standard error for samples in a group

σ is the sample's standard deviation with the formula as $\sqrt{\dfrac{n\sum x^2 - (\sum x)^2}{n(n-1)}}$

The calculated upper and lower 95.0 percent confidence limits of the pollutant BEL at each sampling interval were plotted on graphs. The intercept for both regression lines was set at the 100.0 percent total BEL at zero sampling interval. Figure A-1 shows an example.

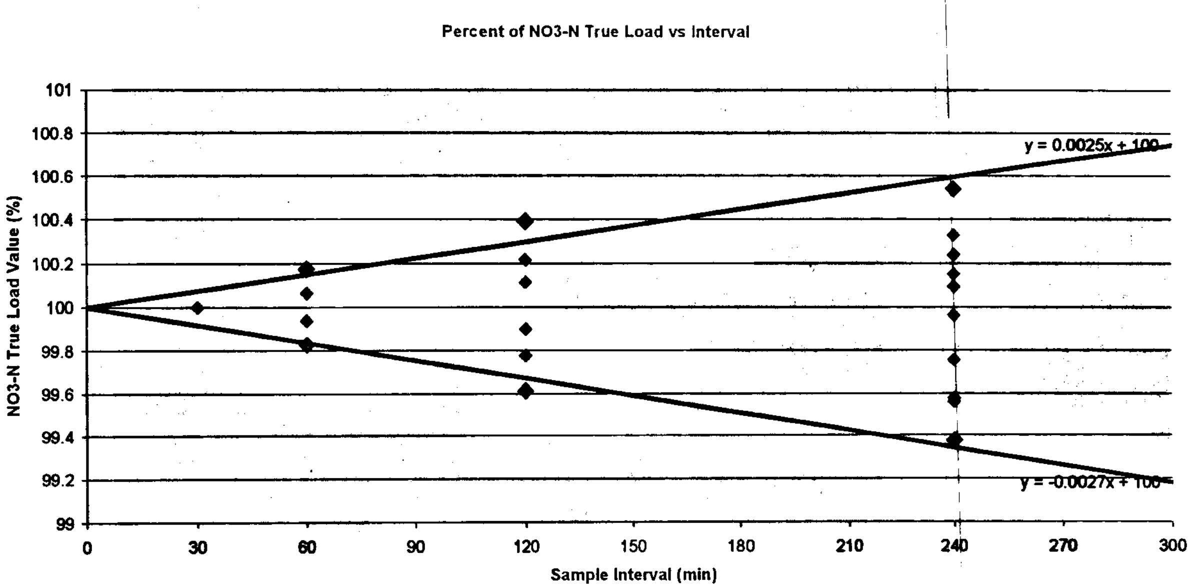

Figure A-1 NO_3-N loads with upper and lower confidence intervals

For the comparison of the OSI result, the 95.0 percent confidence level and linear regression estimators were also applied to the entire 1998 Illinois River sample population of each pollutant at each sampling interval of all nine storms. The 95.0 percent confidence level equation used was similar to that of Equation 5, only that the standard deviation for the entire population is as follows:

$$95.0\ \%\ \text{confidence level} = \bar{x} \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right) \qquad (6)$$

where $\bar{x}$ is the mean value of the calculated percent of BELs for the entire population of each pollutant

at each sampling interval

± 1.96 is the upper and lower percentage point of the 95.0 percent standard normal distribution

$\sigma / \sqrt{n}$ is the standard error for the population

σ is the standard deviation for the population with the formula

$$\sigma = \sqrt{\frac{n\sum x^2 - (\sum x)^2}{n^2}}$$

Based on the upper and lower regression lines of the 95.0 % confidence interval, the OSI for each pollutant was determined where the lines equaled ± 5.0 percent error (variance) of the total BEL. Since there were two sampling interval values determined at ± 5.0 percent error, the ultimate OSI for each pollutant was determined based on the worst case where sampling interval time was the shortest.

LITERATURE CITED

Montgomery, D. C. and G. C. Runger, 1994. Applied Statistics And Probability For Engineers. John Wiley & Sons, Inc.

Yaksich, S. M. and F. H. Verhoff, 1983. Sampling strategy for river pollutant transport. Journal of Environmental Engineering. 109(1), 219 – 31.