

STREAM DIURNAL VARIATIONS, SAMPLE BIAS AND MONITORING STRATEGY IN AN AGRICULTURAL KARST WATERSHED IN CENTRAL KENTUCKY

Alex Fogle
Kentucky Geological Survey

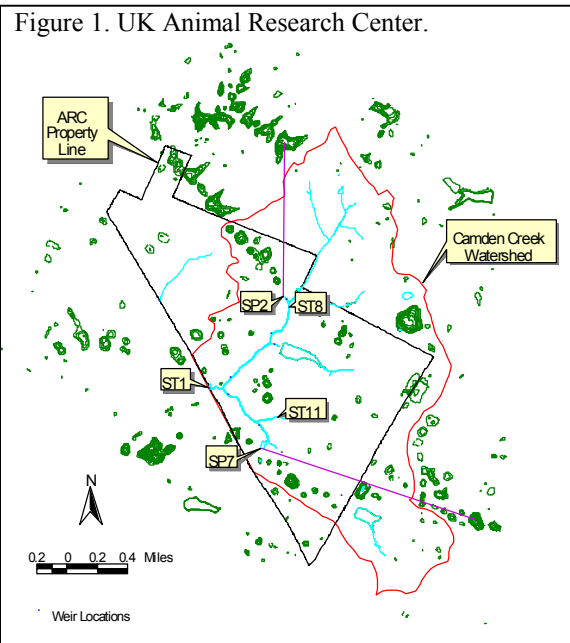
Joseph L. Taraba
Biosystems and Agricultural Engineering
Department
University of Kentucky, Lexington KY

Background

The University of Kentucky Animal Research Center (ARC; Figure 1) lies within the Inner Blue Grass physiographic region of Kentucky and is characterized by broad, shallow sinkholes with caverns and underground drainage ways, low relief, broad valleys and ridges, sparse rock outcrops and thick, fertile, limestone and shale residual soils (Keagy et al., 1993). The geology of the area is characterized as Lexington Limestone, high in phosphorus content, with most member units having minor shale bedding with one significant upper member interbedded with shale. Soils are moderate to well-drained silt loams derived from the high phosphatic limestone. Maury soils are found on broad ridgetops and cover 70% of the area. They are fertile and deep (10-15 ft. over bedrock) with a silty-clay subsoil. McAfee soils cover 15% of the area and are located on moderately steep slopes. They are shallow (< 3 ft. over bedrock) with a silty-clay subsoil. Parts of the ARC are used for precision agriculture operations and include tobacco, row crops, small grains, and animal research grazing plots. The predominant land use outside the ARC boundary is pasture for horse and cattle farms.

A water monitoring network was established at the ARC in October 1996 that has provided background data for comparing post-BMP implementation results with pre-BMP water quality values. Implementation of the nutrient management program at the ARC begins the Spring of 2002 as the Beef and Swine Research Units populate their facilities with animals. The water monitoring network, established at the ARC, is capable of providing sufficient data suitable for assessing the impact of ARC activities on water quality and quantity. One important justification for the network is that much of the water that flows through the ARC originates from land outside the ARC.

Surface and ground water quality data will have been collected for nearly four years by the end of the 2001 water year (October 1). In addition, two years of storm flow data will have been collected. This data is the background to assess the impact of the ARC animal waste nutrient management program BMPs on water resources over the next three to five years. Preliminary analysis of the water quality data has shown that sampling strategies can introduce statistical error or bias into a data set due to the existence of diurnal variations in stream water quality parameters. Implications of this analysis are that an unacceptable level of Type I and Type II statistical errors can be introduced to the watershed assessments of the nutrient management BMPs. The water quality database, containing both pre- and post- BMP implementation data, will be used to develop sampling strategies and



methods for use in karst terrains and small watersheds that reduce error and optimize the utility of data gathered for determining agricultural watershed TMDLs, nutrient loads and water concentrations.

A problem with assessing the success of a BMP in reducing pollution is the inability of controlling the practices of other landowners within a watershed. The network established at the ARC is capable of separating pollution loads that originate from sources other than the ARC. This has been demonstrated by the detection of spikes of triazines and nutrients (Taraba, Ross, and Fogle, 1999) from springs, SP2 and SP7.

Data collected from the ARC monitoring network has demonstrated that the ARC has had a negative impact on surface water quality, relative to incoming water quality, under current crop and animal production practices (Taraba, Ross, and Fogle, 1999). The introduction of BMP practices, such as injecting swine manure into ARC cropland, composting solids wastes before surface land application, and the establishment of riparian zones, is hypothesized to maintain or improve surface water quality that exits the ARC. Errors associated with the sampling protocols need to be minimized to avoid the masking of statistically significant positive or negative impacts.

A few stream water quality trends have been noted at the ARC during the pre-BMP period: 1) The concentrations of fecal coliforms at ST-1, the ARC discharge point, have been increasing since the beginning of 1998 (Taraba, Ross, and Fogle, 1999). This could be due to either additional animal numbers at the farm and increased use of riparian zones for pasturing beef cattle or it is a naturally occurring trend due to weather. 2) Seasonal trends in $\text{NO}_3\text{-N}$ have been observed. The $\text{NO}_3\text{-N}$ concentrations at ST-1 vary from nearly zero mg/l during the drier fall months to around 5 mg/l during the wet winter months, with local peaks occurring after spring fertilization. The winter levels of $\text{NO}_3\text{-N}$ can be attributed to the breakdown of organic matter and leaching during wetter weather. 3) Significantly higher P and $\text{NO}_3\text{-N}$ concentrations were found in water draining the row crop areas when compared to the water draining the pasture areas. Nitrogen isotope samples have been collected and are being analyzed to reveal the sources (inorganic fertilizer, organic matter, or manures) of the $\text{NO}_3\text{-N}$ in ARC surface waters. Collection of nitrogen isotope samples, after the nutrient management BMPs have been implemented, will reveal any changes in the source and spatial distribution of $\text{NO}_3\text{-N}$ as a result of increased use of animal manure for plant fertilization.

Sampling Strategies

The USEPA Protocol for Developing Nutrient TMDLs does not provide guidance for a stream sampling strategy to obtain stream nutrient loads. Stream sampling strategies have varied widely and each has its associated error and bias. Robertson and Roerish (1999), in a study utilizing subsampling to represent different sampling strategies, found that the most effective sampling strategy to estimate loads in small streams depended on the length of the study. Fixed-period monthly sampling with storm chasing appeared to be the most effective and resulted in the most precise annual loads but still resulted in overestimations of 25-50%. For longer studies, Robertson and Roerish found semimonthly sampling provided the least biased and most precise estimates of constituent loads. Our baseline data at ST-1 can be used as the basis for “true” nutrient flux to compare sampling strategies (subsamples created from database). Richards and Holloway (1987) found that flow stratified sampling and load calculation utilizing the Beale Ratio Estimator provided the most precise load estimations. They noted that bias was strongly related to sampling frequency and decreased fairly rapidly with increased sampling frequency. Also, Richards and Holloway stated that sampling strategies based upon sampling theory tended to overestimate the number of samples required to achieve a given precision. Currens (1997), studying a karst spring in southwestern Kentucky, recommended biweekly sampling supplemented by bihourly samples from storm flows.

Another problem that has not been adequately addressed in the literature is the effect of diurnal variations on sample error and bias. Research has gone into the study of diurnal variations in temperature (e.g. Jacobs, et al., 1997; Lowney, 2000; Younus, et al., 2000; Constanz, et al., 1994), dissolved oxygen (e.g. Guasch, et al., 1998), electroconductivity (e.g. Kobayashi, et al., 1990) and NO₃-N (e.g. Hessen, et al., 1997; Christensen, et al, 1990; and Jordan, et al., 1997). Work has also been done on the design of sampling strategies to determine mass loads (e.g. Robertson and Roerish, 1999; Richards and Holloway, 1987; Preston, et al., 1989; and Cohn, et al., 1989). However, nothing could be located in the literature that dealt with the possibility of the introduction of error and bias into a dataset due to diurnal variations.

Variable	No. of Days in Data Record	Sampling Period	mean	standard deviation	RMSE
Temp	1066	24 hour	-0.090	2.453	2.455
		8 am - 5 pm	0.949	2.406	2.586
		8 am - noon	-0.858	1.655	1.864
		noon - 5 pm	2.527	1.752	3.075
		max - min	0.326	0.412	0.525
pH	1066	24 hour	-0.004	0.294	0.294
		8 am - 5 pm	0.158	0.264	0.308
		8 am - noon	0.012	0.192	0.192
		noon - 5 pm	0.286	0.255	0.383
		max - min	0.063	0.056	0.084
NO ₃ -N	786	24 hour	0.005	0.437	0.437
		8 am - 5 pm	0.056	0.352	0.357
		8 am - noon	0.087	0.309	0.321
		noon - 5 pm	0.035	0.449	0.450
		max - min	0.050	0.272	0.276
E.C.	1003	24 hour	0.000	0.025	0.025
		8 am - 5 pm	0.000	0.024	0.024
		8 am - noon	0.009	0.020	0.022
		noon - 5 pm	-0.008	0.022	0.024
		max - min	-0.003	0.013	0.013

Methods and Results

Table 1 is a comparison of randomly selected daily water variables and their “true” daily means. The “true” daily means were determined from data collected at 10 minute intervals at ST-1 with a YSI 6000 Water Quality Datalogger by averaging the 144 data points collected each day. Random values were selected from the data record within given periods of the day, such as 8 am to 5 pm, or noon to 5 pm. Table 1 contains the means, standard deviations, and root mean square errors (RMSE) of the differences between the random sample and the daily mean. Samples collected between noon and 5 pm were the worst at estimating the daily mean for temperature, pH, and NO₃-N. The largest error was produced for E.C. by sampling between 8 am and noon. Averaging the daily minimum and maximum values produced the best estimates of the daily means. In a long term study, with a goal of estimating average annual loads or concentrations emanating from a watershed, it would seem reasonable that collecting data that was sampled as near the daily mean as possible would produce better results. Conversely, it would seem that consistently collecting data near maximum or minimum values could produce bias in the data.

Mean water variable values at each hour of the day, taken from the same data collected to produce Table 1, are plotted in Figure 2. It is plain from Figure 2 that if one researcher consistently sampled at, say, 6 am and another consistently sampled at 4 pm, they would arrive at different values for pH within the same stream. Can significantly different results be obtained by sampling at different times within the day within a given sampling strategy, such as biweekly or monthly grab sampling? The data was subsampled at biweekly and monthly intervals and the water quality values at 8 am, 10 am, noon, 2 pm, 4 pm, and a random value occurring between 8 am and 4 pm determined for each subsample. Utilizing the student’s t-test, the values obtained at each time were compared to determine if they were significantly different. Table 2 shows that significantly different pH values could be

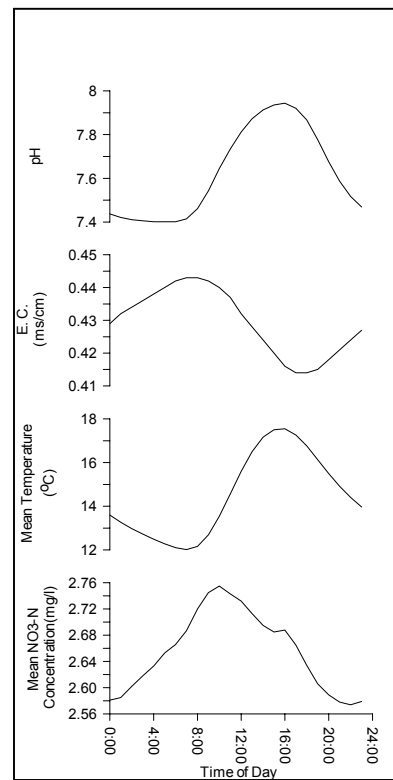


Table 2. Results of Student t-test Comparison of pH Values at Different Times of the Day within Seasons.

	8-10	8-12	8-2	8-4	10-12	10-2	10-4	12-2	12-4	2-4
winter										
mean	-0.166	-0.308	-0.371	-0.340	-0.141	-0.205	-0.174	-0.064	-0.033	0.031
std	0.136	0.229	0.294	0.338	0.102	0.178	0.235	0.094	0.158	0.077
df	15	15	15	15	15	15	15	15	15	15
t statistic	1.220	1.341	1.264	1.005	1.386	1.152	0.740	0.678	0.206	0.405
P = .9	1.753	1.753	1.753	1.753	1.753	1.753	1.753	1.753	1.753	1.753
P = .95	2.131	2.131	2.131	2.131	2.131	2.131	2.131	2.131	2.131	2.131
Sig @ .9	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Sig @ .95	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
spring										
mean	-0.365	-0.655	-0.798	-0.769	-0.290	-0.433	-0.404	-0.143	-0.114	0.029
std	0.157	0.260	0.337	0.430	0.177	0.249	0.326	0.197	0.250	0.197
df	11	11	11	11	11	11	11	11	11	11
t statistic	2.329	2.522	2.369	1.789	1.638	1.739	1.239	0.727	0.457	0.148
P = .9	1.796	1.796	1.796	1.796	1.796	1.796	1.796	1.796	1.796	1.796
P = .95	2.201	2.201	2.201	2.201	2.201	2.201	2.201	2.201	2.201	2.201
Sig @ .9	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Sig @ .95	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
summer										
mean	-0.208	-0.468	-0.614	-0.678	-0.260	-0.406	-0.469	-0.146	-0.209	-0.064
std	0.208	0.384	0.456	0.471	0.237	0.339	0.366	0.125	0.216	0.152
df	15	15	15	15	15	15	15	15	15	15
t statistic	1.001	1.219	1.345	1.439	1.095	1.197	1.281	1.169	0.970	0.419
P = .9	1.753	1.753	1.753	1.753	1.753	1.753	1.753	1.753	1.753	1.753
P = .95	2.131	2.131	2.131	2.131	2.131	2.131	2.131	2.131	2.131	2.131
Sig @ .9	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Sig @ .95	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
fall										
mean	-0.062	-0.159	-0.249	-0.296	-0.097	-0.188	-0.235	-0.091	-0.138	-0.047
std	0.032	0.083	0.170	0.223	0.062	0.153	0.205	0.097	0.151	0.062
df	16	16	16	16	16	16	16	16	16	16
t statistic	1.938	1.924	1.464	1.328	1.578	1.225	1.145	0.933	0.914	0.759
P = .9	1.746	1.746	1.746	1.746	1.746	1.746	1.746	1.746	1.746	1.746
P = .95	2.120	2.120	2.120	2.120	2.120	2.120	2.120	2.120	2.120	2.120
Sig @ .9	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
Sig @ .95	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

obtained from a biweekly sampling regimen with as little as two hours between sampling times, especially if the sampling was only conducted during a particular season. Comparison of temperatures also produced significantly different values, whereas NO₃-N and electroconductivity did not appear to vary significantly with different sampling time. This would imply that a significant difference could be found between two identical watersheds in which no change occurred when sampling at different times, a Type II error. More analysis needs to be done to determine if a sampling protocol should and can be developed that addresses the error and bias by diurnal variation.

Preliminary results from ARC baseline data, collected at ST-1 from June 1997 through 2000, indicate that biweekly sampling with a continuous flow record may provide acceptable estimations of annual load. Table 3

shows the results of comparing two sampling strategies with actual loads of NO₃-N and Total Solutes. Flow, NO₃-N and E.C. data was collected continuously (10 minute intervals). Grab samples were also collected biweekly through the period. Loads were calculated three ways: True loads utilizing the continuous flow and concentration data and estimated loads using 1) grab sample concentrations and a “grab flow” and 2) grab sample concentrations and the average flow rate for 1, 2 and 4 week periods. The average flow was determined by using the continuous flow record for 3.5 (1 week), 7 (2 week), and 14 (4 week) days, respectively before and after the grab sample collection time. Wet and dry weather periods were roughly (56% wet, 44% dry) equally represented. As can be seen in Table 3, grab sampling on a weekly or biweekly basis with collection of a continuous flow record appears to be an acceptable method of estimating of annual Nitrate-N loads, especially if weekly or biweekly sampling is done. Total Solutes appears to be better estimated by utilizing weekly grab samples and flows. More data needs to be collected to determine if this is a viable alternative to supplementing grab sampling with storm chasing or flow stratified sampling strategies, which are more costly.

The ARC stream baseline data has been used to estimate the nutrient concentrations and loads from various land uses. The USEPA Protocol for

Table 3. Comparison of Stream Load Estimation Methods for Nitrate-N and Total Solutes.

Load Estimation Method	NO ₃ -N						Total Solutes					
	Grab Sample with Grab Flow (kg/ha)			Grab Sample with Total Flow Volume (kg/ha)			Grab Sample with Grab Flow (kmol/ha)			Grab Sample with Total Flow Volume (kmol/ha)		
Period	1 week	2 week	4 week	1 week	2 week	4 week	1 week	2 week	4 week	1 week	2 week	4 week
mean	-0.041	-0.021	0.084	-0.014	-0.008	0.002	-0.048	-0.066	-0.078	-0.022	-0.041	-0.015
stdev	0.107	0.429	0.587	0.061	0.135	0.169	0.174	0.343	0.681	0.331	0.443	0.719
median	-0.001	-0.009	-0.004	0.000	0.001	-0.002	0.000	-0.004	-0.017	-0.001	-0.001	0.005
n	37	33	23	37	33	23	51	53	41	51	53	41
RMSE	0.114	0.430	0.593	0.063	0.136	0.169	0.180	0.349	0.686	0.332	0.445	0.720
% diff	-13.11%	-3.19%	13.92%	-4.46%	-1.30%	0.26%	-19.25%	-14.14%	-8.53%	-9.07%	-8.74%	-1.60%
Weeks	37	66	92	37	66	92	51	106	164	51	106	164

Developing Nutrient TMDLs recommends strategies to estimate source loads for TMDL development. Tables in the protocol present literature stream values for total dissolved inorganic N and ortho-P concentrations, and typical total P and N loading for various agricultural land uses (see Tables 4 and 5). These values are recommended when no watershed data is available. The USEPA values for pasture are 1/3 the values found for the ARC pasture areas. The ARC stream flow nutrient loads from row crop areas for Total N and ortho-P exceed the maximum USEPA guidance values. For the pasture watershed, the ortho-P nutrient load exceeds the USEPA maximum guidance value while Total N load is slightly above the USEPA median guidance value.

Table 4. ARC Stream Nutrient Concentrations (10/1996 - 6/2000 Compared to USEPA Guidance Values.

Land Use	ARC Watershed Mean (mg/l)		USEPA Central USA Mean (mg/l)	
	Total Inorganic N	Total Ortho-P	Total Inorganic N	Total Ortho-P
> 90% Agriculture with ~ 20% row crop (ST-1)	2.88 A	0.21 A		
> 90% Agriculture with ~ 0% row crop (ST-8)	2.86 A	0.21 A		
> 90% Agriculture with ~ 60% row crop (ST-11)	9.56 B	0.38 B		
> 90% Agriculture			0.77	0.085

Table 5. ARC Stream Nutrient Loads Compared to USEPA Guidance Values.

Land Use	Nutrient	ARC Watershed (kg/ha-yr)	USEPA (kg/ha-yr)		
			Min.	Max.	Median
>90% Agriculture with >20% Row Crop (ST-1) (10/1996-6/2000)	Total N	13.01			
	Ortho-P	0.85			
>90% Agriculture with >60% Row Crop (ST-11) (4/1999-6/2000)	Total N	18.06			
	Ortho-P	0.59			
> 90% Agriculture with >0% Row Crop (ST-8) (4/1999-6/2000)	Total N	4.52			
	Ortho-P	0.28			
UKARC (4/1999-6/2000)	Total N	14.18			
	Ortho-P	0.38			
Pasture	Total N		1.2	7.1	4.2
	Ortho-P		0.01	0.25	0.13

References

- Christensen, P.B., L.P. Nielsen, J. Sorensen, and N.P. Revsbech. 1990. Denitrification in nitrate-rich streams: Diurnal and seasonal variation related to benthic oxygen metabolism. *Limnol. Oceanogr.*, 35(3), 640-651.
- Cohn, T. A., L.L. DeLong, E. J. Gilroy, R. M Hirsch, and D. K. Wells. 1989. Estimating constituent loads. *Water Resources Research*, 25(5), 937-942.
- Constanz, J., C. L. Thomas, and G. Zellweger. 1994. Influence of diurnal variations in stream temperature on streamflow loss and groundwater recharge. *Water Resources Research*, 30(12), 3253-3264.
- Currens, J. C. 1997. A sampling plan for conduit-flow karst springs: Minimizing sampling cost and maximizing statistical utility. in Beck and Stephenson., eds., *The Engineering Geology and Hydrology of Karst Terranes*, Proceedings of the 6th Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impact of Karst, Springfield, MO, 6-9 April, 1997:Rotterdam, A. A. Balkema, p. 193-198.
- Guasch, H., J. Armengol, E. Marti, and S Sabater. 1998. Diurnal variation in dissolved oxygen in two low-order streams. *Water Research*, 32(4), 1067-1074.
- Hessen, D. O., A. Henriksen, and A. M. Smelhus. 1997. Seasonal fluctuations and diurnal oscillations in nitrate of a heathland brook. *Water Research*, 31(7), 1813-1817.
- Jacobs, A. F. G., T. H. Jetten, D. C. Lucassen, B. G. Heusinkveld, and J. P. Nieveen. 1997. Diurnal temperature fluctuations in a natural shallow water body. *Agri. and Forest Meteorology*, 88, 269-277.
- Jordan, T. E., D. L. Correll, and D. E. Weller. 1997. Relating nutrient discharges from watersheds to land use and streamflow variability. *Water Resources Research*, 33(11), 2579-2590.
- Keagy, D., J. S. Dinger, A. W. Fogle, and L. V. A. Sendlein. 1993. Interim report on the occurrence of pesticides, nitrate, and bacteria on groundwater quality in a karst terrain – the inner Bluegrass Region, Woodford County, Kentucky. Kentucky Geological Survey Open File Report OF-94-04.
- Kobayashi, D., K. Suzuki, and M. Nomura. 1990. Diurnal fluctuation in stream flow and in specific electric conductance during drought periods. *Journal of Hydrology*, 115, 105-114.
- Lowney, C. L. 2000. Stream temperature variation in regulated rivers: Evidence for a spatial pattern in daily minimum and maximum magnitudes. *Water Resources Research*, 36(10), 2947-2955.
- Preston, S. D., V. J. Bierman, and S. E. Silliman. 1989. An evaluation of methods for the estimation of tributary mass loads. *Water Resources Research*, 25(6), 1379-1389.
- Richards, R. P., and J. Holloway. 1987. Monte Carlo Studies of Sampling Strategies for Estimating Tributary Loads. *Water Resources Research*, 23(10), 1939-1948.
- Robertson, D. M., and E. D. Roerish. 1999. Influence of various water quality sampling strategies on load estimates for small streams. *Water Resources Research*, 35(12), 4747-3759.
- Taraba, J.L., I.J. Ross, and A.W. Fogle. 1999. Monitoring impacts of the Animal Research Center on surface and groundwater quality: 271 Phase III 1st Year Progress Report, University of Kentucky, College of Agriculture.

16. Younus, M., M. Hondzo, and B. A. Engel. 2000. Stream temperature dynamics in upland agricultural watersheds. *Journal of Environmental Engineering*, 126(6), 518-526.