CONTINUOUS REAL-TIME WATER-QUALITY MONITORING OF KANSAS STREAMS

Victoria G. Christensen, Andrew C. Ziegler, Patrick P. Rasmussen, and Xiaodong Jian*

ABSTRACT: A continuous, real-time water-quality monitoring system was developed for 13 stream sites in Kansas that eliminates the waiting time inherent in chemical analyses reported by a laboratory and provides continuous estimates of constituent concentrations and loads. The U.S. Geological Survey (USGS) monitoring system is described, and its effectiveness in characterizing water quality is evaluated using discrete water-quality samples collected from 1995 through 2002 and continuous water-quality monitor data from the first 4 years of operation with examples from two stream sites on the Little Arkansas River in south-central Kansas. Because sensor technology currently is not available to directly measure many chemicals of interest in a stream, regression models are developed to relate constituents in laboratory-analyzed samples with in-stream continuous-sensor measurements. Concentration estimates traditionally have used continuous streamflow data only; however, most constituents are more accurately estimated with continuous specific conductance or turbidity measurements. As the hourly sensor measurements are transmitted from the stream sites to the USGS computers in Lawrence, Kansas, the models are applied, and the computed estimates displayed on a Web page at http://ks.water.usgs.gov/Kansas/rtqw/. Currently, continuous estimated concentrations and loads of sediment, major ions, select nutrients, atrazine, and indicator bacteria, and uncertainty of the estimates are displayed. Information from this system is used by water suppliers to modify treatment of water, by State and local agencies in total maximum daily load (TMDL) programs, and to alert recreational water users of potential health risks.

KEY TERMS: continuous real-time water-quality, loads, sediment, nutrients, bacteria, atrazine.

RESULTS

Between 1995 and 2002, the U.S. Geological Survey (USGS), in cooperation with local, State, and Federal agencies, collected numerous discrete water samples from 13 stream sites in Kansas. Two of those sites are the Little Arkansas River near Halstead (site 07143672) and near Sedgwick (site 07144100). Samples were analyzed for sediment, major ions, nutrients, pesticides, indicator bacteria, and other constituents. Samples were collected throughout a range of streamflow conditions and constituent concentrations as described in Wilde and Radtke (1998). Both stream sites also were equipped with water-quality monitors from 1998 through 2002 that record continuous measurements of specific conductance, pH, water temperature, dissolved oxygen, and turbidity. These monitors were serviced about twice per month according to methods described in Wagner and others (2001). Regression models were developed using S-Plus Statistical Software (MathSoft, Inc., 1999) to describe relations between the discrete samples and the continuous cross-section-averaged water-quality sensor measurements, streamflow, stage, and time (Table 1). Site-specific regression models were developed using plots of each possible explanatory variable against the response variable and visually examining the plots of the residuals for patterns. Explanatory and response variables (except time) were log transformed, if necessary, to remove curvature in the data. An overall model-building method was used (Helsel and Hirsch, 1992, p. 312-314). Generally, if there were several acceptable models (F-test p-value less than 0.05), the one with the lowest PRESS statistic was chosen. PRESS (acronym for "PRediction Error Sum of Squares") is one of the best measures of the goodness of fit of a regression model (Helsel and Hirsch, 1992, p. 248). Explanatory variables were included in a model only if there was a physical basis or explanation for their inclusion. For variables that were log transformed, retransformation of regression-estimated concentrations was necessary. However, retransformation can cause an underestimation of chemical loads when adding individual load estimates over a long period of time. A Duan's bias correction factor or smear factor (Duan, 1983) was applied to the annual load calculation to correct for this underestimation. Cohn and others (1989), and Hirsch and others (1993) provide additional information on interpreting the results of regression-based load estimates. Uncertainty of the estimates for regression models was determined using 90-percent prediction intervals (Helsel and Hirsch, 1992). Probabilities of exceeding water-quality standards, recommended criteria, or guidelines of the State of Kansas and U.S. Environmental Protection Agency also were determined. Regression methods are described in Helsel and Hirsch (1992), Hirsch and others (1993), and Christensen (2000).

* Respectively, Hydrologist, Supervisory Hydrologist, Hydrologist, Hydrologist, U.S. Geological Survey, 4821 Quail Crest Place, Lawrence, KS 66049, Phone: (785) 832-3539, Fax: (785) 832-3500, E-mail: aziegler@usgs.gov.

Table 1. Regression Models for Continuous Concentration and Load Estimates and Load Smearing Factors for Little Arkansas River at Highway 50 near Halstead (site 07143672) and near Sedgwick (site 07144100), Kansas, 1995-2002.

[n, number of samples; R^2, square of coefficient of determination; MSE, mean square error; Smear, smear correction factor; *Alk*, alkalinity in milligrams per liter as calcium carbonate; *SC*, specific conductance in microsiemens per centimeter at 25 degrees Celsius; *Q*, streamflow in cubic feet per second; HCO_3, bicarbonate in milligrams per liter; $CaCO_3$, hardness in milligrams per liter; *DS*, dissolved solids in milligrams per liter; *TSS*, total suspended solids in milligrams per liter; *Turb*, turbidity in nephelometric turbidity units; *Ca*, calcium in milligrams per liter; *Na*, sodium in milligrams per liter; *Cl*, chloride in milligrams per liter; *F*, fluoride in milligrams per liter; SO_4, sulfate in milligrams per liter; *TON*, total organic nitrogen in milligrams per liter ; *TP*, total phosphorus in milligrams per liter; *As*, arsenic in micrograms per liter; *Atr*, atrazine in micrograms per liter; *D*, day of year, *FCB*, fecal coliform bacteria in colonies per 100 milliliters; *ECB*, *Escherichia coli* bacteria in colonies per 100 milliliters]

Constituent	Regression Model	n	R^2	MSE	Smear
	Little Arkansas River at Highway 50 near Halstead (site 07143672)				
Alkalinity	$\log_{10} Alk = 0.646 \log_{10} SC - 0.102 \log_{10} Q + 0.509$	155	0.949	0.00484	1.015
Bicarbonate	$\log_{10} HCO_3 = 0.846 \log_{10} SC - 0.179$	156	.912	.00831	1.021
Hardness	$CaCO_3 = 0.273SC + 9.85$	157	.960	897	1.000
Dissolved solids	$DS = 0.551SC + 25.3$	158	.977	2,310	1.000
Total suspended solids	$\log_{10} TSS = 0.943 \log_{10} Turb + 0.0110$	48	.933	.0294	1.127
Suspended sediment concentration	$\log_{10} SSC = 0.945 \log_{10} Turb + 0.132$	43	.948	.0201	1.047
Calcium	$Ca = 0.0870SC + 2.34$	157	.955	102	1.000
Sodium	$Na = 0.107SC - 15.4$	157	.966	114	1.000
Chloride	$Cl = 0.209SC - 33.7$	156	.957	558	1.000
Sulfate	$\log_{10} SO_4 = 0.900 \log_{10} SC - 1.07$	158	.885	.0126	1.035
Total organic nitrogen	$\log_{10} TON = 0.439 \log_{10} Turb - 0.816$	35	.850	.0127	1.034
Total phosphorus	$TP = 0.000929Turb + 0.325$	36	.893	.0207	1.000
Arsenic	$\log_{10} As = -0.227 \log_{10} Q + 1.16$	33	.663	.0223	1.172
Atrazine	$\log_{10} Atr = -0.000790SC + 0.359 \sin(\frac{2\pi D}{365}) - 0.643 \cos(\frac{2\pi D}{365}) + 0.430$	78	.552	.326	1.806
Fecal coliform bacteria	$\log_{10} FCB = 1.13 \log_{10} Turb + 0.378$	23	.69	.249	2.122
Escherichia coli bacteria	$\log_{10} ECB = 1.011 \log_{10} Turb + 0.439$	23	.63	.290	2.282
	Little Arkansas River near Sedgwick (site 07144100)				
Alkalinity	$\log_{10} Alk = 0.671 \log_{10} SC - 0.121 \log_{10} Q + 0.574$	131	.936	.00588	1.014
Bicarbonate	$\log_{10} HCO_3 = 0.000462SC - 0.174 \log_{10} Q + 2.32$	131	.913	.00787	1.013
Hardness	$CaCO_3 = 0.318SC + 0.442$	131	.965	471	1.000
Dissolved solids	$DS = 0.562SC + 17.2$	125	.958	518	1.000
Total suspended solids	$\log_{10} TSS = 1.16 \log_{10} Turb + 0.000605SC - 0.831$	54	.911	.0402	1.119
Suspended sediment concentration	$\log_{10} SSC = 0.715 \log_{10} Turb + 0.188 \log_{10} Q + 0.185$	47	.908	.0419	1.141
Calcium	$Ca = 0.0979SC - 0.0587$	131	.962	48.5	1.000
Sodium	$Na = 0.101SC + 3.79 \log_{10} Q - 23.4$	131	.965	41.1	1.000
Chloride	$Cl = 0.203SC + 25.7 \log_{10} Q - 112$	131	.941	207.	1.000
Fluoride	$\log_{10} F = 0.434 \log_{10} SC - 1.75$	57	.621	.0112	1.000
Sulfate	$SO_4 = 0.0592SC + 1.00$	131	.924	35.5	1.000
Total organic nitrogen	$TON = 0.00232Turb + 0.863$	39	.743	.263	1.000
Total phosphorus	$TP = 0.000567Turb + 0.506$	36	.549	.0385	1.000
Arsenic	$\log_{10} As = -0.250 \log_{10} Q + 1.30$	34	.682	.021	1.041
Atrazine	$\log_{10} Atr = -0.000647SC + 0.504 \sin(\frac{2\pi D}{365}) - 0.745 \cos(\frac{2\pi D}{365}) + 0.237$	205	.604	.183	1.442
Fecal coliform bacteria	$\log_{10} FCB = 1.19 \log_{10} Turb + 0.198$	28	.79	.210	1.579
Escherichia coli bacteria	$\log_{10} ECB = 1.17 \log_{10} Turb + 0.111$	28	.73	.278	1.806

DISCUSSION

There are several advantages to a continuous, real-time water-quality monitoring system. The first advantage is that the information is real time. This eliminates waiting time inherent in chemical analyses reported by a laboratory and provides continuous estimates of constituent concentrations and loads. The continuous estimates can assist water managers in the operation of water-treatment facilities and allow for decisions about the sanitary quality of water with respect to water recreation activities and public safety. Concentration and load information is available on the World Wide Web (http://ks.water.usgs.gov/Kansas/rtqw/) for water managers to make decisions relative to rapid changes in water quality. An example of the information available for suspended sediment is presented in Figure 1. Knowledge of the probability that current water-quality concentrations meet or exceed water-quality standards also allows the public to make decisions about the use of a stream for recreational purposes. Additionally, continuous estimates provide information that is essential to obtain accurate constituent load information. Because the data are available on an hourly basis, rapid changes in water quality can be quantified and used to more accurately estimate the daily load during storm runoff when concentrations and loads can vary by as much as four orders of magnitude. Additionally, the continuous data result in a more accurate estimation of the loads for different time periods-- daily, weekly, monthly, or annually. The continuous data also can be used to construct frequency distribution (duration) curves to determine percentage of time that concentrations exceed water-quality standards or to assist with development of water-quality standards.

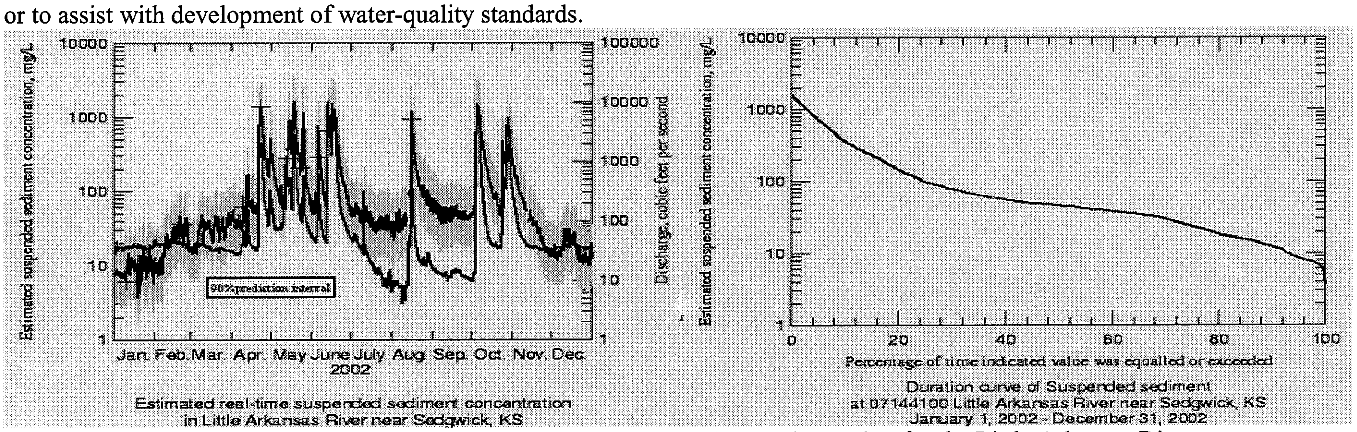

Figure 1 Hourly Estimated Suspended-Sediment Concentrations and Duration for the Little Arkansas River near Sedgwick, Kansas, January 1, 2002, Through December 31, 2002.

Continuous streamflow information is available nationally at more than 7,000 USGS stream gages, but many constituents in streams are more strongly related to specific conductance or turbidity than to streamflow; therefore, a more accurate estimation of load is possible by using the regression models with the water-quality sensor data as explanatory variables. A previous study for calendar year 2001 indicated that differences between instantaneous measured suspended-sediment load and streamflow-estimated suspended-sediment load were as great as 100 percent at some sites (Christensen and others, 2001), whereas the difference between measured instantaneous suspended-sediment load and turbidity-estimated suspended-sediment load generally was less than 10 percent. A similar comparison between multiple-regression-estimated and streamflow-estimated atrazine (Table 2) was made to determine if specific conductance and day of year provided better estimates of atrazine than streamflow. For both monitoring sites on the Little Arkansas River, the percentage difference was about twice as large for the streamflow-estimated load than for the specific-conductance and day-of-year-estimated load. Specific conductance and day of year provided a more accurate estimate of atrazine loads probably because the larger values of specific conductance are associated with base flows when ground water dominates and has small concentrations of atrazine. However, both models over estimate atrazine loads probably because of bias toward higher values and the seasonal application of atrazine. Improvements in sensor technology may provide better explanation of atrazine concentrations.

Table 2. Comparison of Measured Instantaneous Atrazine Loads to Streamflow- and Multiple-Regression-Estimated Atrazine Loads in the Little Arkansas River at Highway 50 near Halstead and near Sedgwick, Kansas, 1995-2002.

Variables	Little Arkansas River at Highway 50 near Halstead	Little Arkansas River near Sedgwick
Number of samples	78	205
Mean measured atrazine concentration (micrograms per liter)	5.19	5.29
Mean measured streamflow (cubic feet per second)	668	918
Mean measured instantaneous atrazine load (pounds per day)	23.8	28.4
Mean streamflow-estimated instantaneous atrazine load (pounds per day)	51.6	49.6
Percentage difference of streamflow-estimated load from measured load	**-120**	**-75**

Mean specific-conductance-estimated instantaneous atrazine load (pounds per day)	39.6	36.0
Percentage difference of specific-conductance-estimated load from measured load	**-66**	**-27**

The primary limitations for the continuous water-quality monitoring approach are associated with the water-quality sensors and the consistency of the relations with the sensor. Turbidity sensors have an upper limit of measurement depending on the instrument used (Sadar, 2002). Exceeding this upper limit of measurement prevents accurate estimation of constituent concentrations and loads during the greatest turbidities that generally occur during the greatest streamflows when the largest loads are likely to occur. This is an important consideration in the sediment-rich streams of Kansas, but may not be as important in streams that are less turbid. However, concentrations and loads can be described with greater accuracy for the days when the sensor is not maximized. The turbidity sensor was at maximum for 4 percent of the time during 2002 (Figure 1). As technology improves, sensor maximization potentially can be eliminated, and direct sensor measurement of constituents of interest made possible. Additionally, continued discrete water-quality samples are required to verify that the relations between the sensors and constituents estimated remain consistent.

CONCLUSIONS

Site-specific regression models coupled with continuous, in-stream water-quality data can estimate in-stream constituent concentrations and loads continuously and in real time. Estimated hourly constituent concentrations, loads, probability of exceeding water-quality standards, uncertainty of the estimates, and frequency distributions can be rapidly displayed on the World Wide Web. These estimates can provide water-quality information to resource managers and the public that is otherwise not available. In addition to the utility of the regression models for estimating concentrations, they also are useful for estimating total maximum daily loads (TMDLs) and to alert recreational water users of potential health risks. However, continued sampling is required to verify that the statistical models are consistent.

This innovative approach utilizes the existing USGS stream-gaging network. In the future, real-time estimates can be developed elsewhere in the United States using the USGS national network of more than 7,000 streamflow gages to provide continuous, real-time information on streamflow and water-quality concentrations and loads in streams that can be used to improve the treatment of drinking water and to monitor the environment. This approach offers the potential to improve our understanding of the dynamics of watersheds and nonpoint-source pollution. As new sensors become available for direct measurement of constituents, the use of the network sites to provide continuous, real-time water-quality information will improve. Additional advances in technology are hoped to directly measure chemicals of interest and decrease the data uncertainty and the operation and maintenance requirements for the sensors. The increasing public interest in TMDLs and water quality in general make this approach of regional, national, and international importance.

ACKNOWLEDGMENTS

The use of brand, trade, or firm names in this paper is for identification purposes only and does not constitute endorsement by the U.S. Geological Survey. Without the dedicated field personnel to collect samples and maintain the water-quality monitors, none of this work is possible. Their efforts are greatly appreciated.

REFERENCES

Christensen, V.G., Jian, X. and Ziegler, A.C., 2000. Regression Analysis and Real-Time Water-Quality Monitoring to Estimate Constituent Concenatrtions, Loads, and Yields in the Little Arkansas River, South-Central Kansas, 1995-99. U.S. Geological Survey Water-Resources Investigations Report 00-4126, 36 p.

Christensen, V.G., Ziegler, A.C., and Jian, Xiaodong, 2001. Continuous Turbidity Monitoring and Regression Analysis to Estimate Total Suspended Solids and Fecal Coliform Bacteria Loads in Real Time. Proceedings for Federal Interagency Sedimentation Conference, March 2001, Reno Nevada.

Cohn, T.A., DeLong, L.L., Gilroy, E.J., Hirsch, R.M., and Wells, D., 1989. Estimating Constituent Loads. Water Resources Research 25(5):937–942.

Duan, N., 1983. Smearing Estimate—A Nonparametric Retransformation Method. Journal of the American Statistical Association 78:605–610.

Helsel, D.R., and Hirsch, R.M., 1992. Statistical Methods in Water Resources. New York, Elsevier, 522 p.

Hirsch, R.M., Helsel, D.R., Cohn, T.A., and Gilroy, E.J., 1993. Statistical Analysis of Hydrologic Data, *in* Maidment, D.R., ed., Handbook of Hydrology. New York, McGraw-Hill, Inc., pp. 17.1–17.55.

MathSoft, Inc. 1999. S-Plus 2000--Modern Statistics and Advanced Graphics, Users Guide.Cambridge, Massachusetts,558 p.

Sadar, M. J., 2002. Turbidity Instrumentation--An Overview of Today's Available Technology. Accessed May 2, 2002, at URL at water.usgs.gov/osw/techniques/TSS/sadar.pdf

Wagner, R.J., Mattraw, H.C., Ritz, G.F., and Smith, B.A., 2001. Guidelines and Standard Procedures for Continuous Water-Quality Monitors-- Site Selection, Field Operation, Calibration, Record Computation, and Reporting. U.S. Geological Survey Water-Resources Investigations Report 00-4252, 53 p.

Wilde, F.D., and Radtke, D.B., eds., 1998. Field Measurements, *in* National Field Manual for the Collection of Water-Quality Data. U.S. Geological Survey Techniques of Water-Resource Investigations, book 9, chap. A6, pp. 3-20.