USGS
science for a changing world

Continuous Turbidity Data Used to Compute Constituent Concentrations in the South Loup River, Nebraska, 2017–18

Scientific Investigations Report 2021–5120

U.S. Department of the Interior
U.S. Geological Survey

Continuous Turbidity Data Used to Compute Constituent Concentrations in the South Loup River, Nebraska, 2017–18

By David L. Rus and Brenda K. Densmore

Scientific Investigations Report 2021–5120

U.S. Department of the Interior
U.S. Geological Survey

U.S. Geological Survey, Reston, Virginia: 2021

Acknowledgments

Contents

Figures

Tables

Conversion Factors

U.S. customary units to International System of Units

Multiply	By	To obtain
Area		
square mile (mi^2)	2.590	square kilometer (km^2)
Flow rate		
foot per second (ft/s)	0.3048	meter per second (m/s)
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

Supplemental Information

Constituent concentrations are given in milligrams per liter (mg/L), in formazin nephelometric units (FNU), or in most probable number of colonies per 100 milliliters (mpn/100 mL).

Abbreviations

AMP	Ambient Monitoring Program
CV	coefficient of variation
E. coli	*Escherichia coli*
LLNRD	Lower Loup Natural Resources District
NDEE	Nebraska Department of Environment and Energy
NO3	nitrate plus nitrite
NWIS	National Water Information System
RMSE	root mean squared error
R^2	coefficient of determination
SLR	single-variable ordinary least squares regression
SLRSS	South Loup River Special Study
TKN	total Kjeldahl nitrogen
TN	total nitrogen
TP	total phosphorus
TSS	total suspended solids
USGS	U.S. Geological Survey

Continuous Turbidity Data Used to Compute Constituent Concentrations in the South Loup River, Nebraska, 2017–18

By David L. Rus and Brenda K. Densmore

Abstract

The South Loup River in central Nebraska has been impaired by bacteria since at least 2004, which has resulted in the river not meeting its intended use as a recreational waterway. As part of a strategy for reducing the bacterial load in the river, the U.S. Geological Survey, in cooperation with the Lower Loup Natural Resources District, made continuous estimates of *Escherichia coli* (*E. coli*) and nutrient concentrations during seasonal monitoring at the South Loup River at Saint Michael, Nebraska, during 2017–18. Continuous turbidity data were collected from mid-April through October in 2017 and 2018 and were paired with 35 co-occurring discrete water samples that were analyzed for *E. coli*, nutrients, and suspended solids. Surrogate models relating the discrete concentrations to the continuous turbidity data were developed using ordinary-least-squares regression and were evaluated for model performance and uncertainty. Although the model assumptions were met for *E. coli*, the imprecision of the *E. coli* model was considerably higher than the other constituents, probably because of measurement imprecision and greater sensitivity to environmental factors. Once the models were developed, the turbidity data were used to predict continuous constituent concentrations and corresponding prediction intervals, which were made available online as part of the U.S. Geological Survey National Water Information System database. It is expected that results from these models will provide stakeholders with an understanding of constituent concentrations during the 2017–18 monitoring period and the results will also provide a good reference point for any future comparisons.

Introduction

The South Loup River in central Nebraska (fig. 1) has been impaired by bacteria since at least 2004, which has resulted in the river not meeting its intended use as a recreational waterway (Nebraska Department of Environment and Energy [NDEE], 2018). In response, a total maximum daily load restriction has been in place since December 2005 (NDEE, 2013), and more recently the Lower Loup Natural Resources District (LLNRD) completed a management plan that includes a strategy for reducing the bacterial load in the South Loup River. A part of that strategy includes the need for monitoring to better understand the spatial, source-based, and temporal characteristics of *Escherichia coli* (*E. coli*) and nutrient levels in the river. In response, the LLNRD, the NDEE, and the U.S. Geological Survey (USGS) completed the South Loup River Special Study (SLRSS). The USGS, in cooperation with the LLNRD, was responsible for the part of the SLRSS that included computing continuous estimates of *E. coli* and nutrient concentrations in the river. Such estimates can help resource managers identify periods associated with higher concentrations and can provide a baseline before mitigation activities. However, these concentration estimates can be quite uncertain when based solely on infrequent (weekly or less frequent) sample data (Garrett, 2019). Including continuous water-quality information, such as turbidity, as a surrogate for those concentrations can improve the estimates by reducing the uncertainty associated with unsampled periods.

Purpose and Scope

The purpose of this report is to describe procedures for computing a time series of concentrations and loads of selected constituents, including total suspended solids (TSS), *E. coli*, total Kjeldahl nitrogen (TKN), total nitrogen (TN), and total phosphorus (TP) using 2017–18 continuous turbidity data from the South Loup River. This report describes the techniques and guidelines followed to develop unbiased concentration and load time series from linear regression models using continuous turbidity measurements (see the "Methods" section). Linear regression models relating continuous turbidity data and discrete sample results are described, with detail on model selection and diagnostics (see the "Surrogate Models using Continuous Turbidity Data to Compute Constituent Concentrations" section). Methods used to fill gaps in the continuous turbidity data are also described. Continuous turbidity data and computed time-series concentrations and loads are available through the USGS National Water Information System (NWIS) database (USGS, 2020).

Study Area

The monitoring site on the South Loup River is associated with USGS streamflow-gaging station 06784000 (fig. 1). The site is near the unincorporated community of Saint Michael, Nebraska (not shown in fig. 1) and is commonly referred to, including in this report, as the Saint Michael site. About 15 percent of the 2,320-square-mile South Loup River Basin lies within the Nebraska Sand Hills (Soenksen and others, 1999; Hobza and Schepers, 2018), an area of grass-covered sand dunes dominated by rangeland (Bleed and Flowerday, 1989), and much of the remaining basin drains dissected loess plains (Hobza and Schepers, 2018). Most (71 percent) of the basin is undeveloped or used for rangeland, and 21 percent of the basin is irrigated cropland (Hobza and Schepers, 2018).

In 2017, annual mean streamflow at the Saint Michael site was 169 cubic feet per second (ft³/s), which was lower than the mean annual of 234 ft³/s for the entire period of record (1944–2020) for this site (USGS, 2020). A large runoff event associated primarily with rainfall in the western part of the basin led to a peak streamflow of 1,870 ft³/s on August 16,

2017. Annual mean streamflow in 2018 was 219 ft³/s and was characterized by several runoff events between May and September, although the peak streamflow of 1,460 ft³/s for 2018 was lower than that of 2017.

Methods

Continuous water-quality data and discrete samples were collected seasonally (April through October) at the Saint Michael site during 2017 and 2018. Those data were then used to develop statistical models that estimate continuous constituent concentrations for the monitoring period.

Continuous Data

A multiparameter continuous water-quality monitor was deployed seasonally at the Saint Michael site from 2017 to 2018. Methods used to operate, quality control, and service the continuous water-quality monitor, as well as procedures used
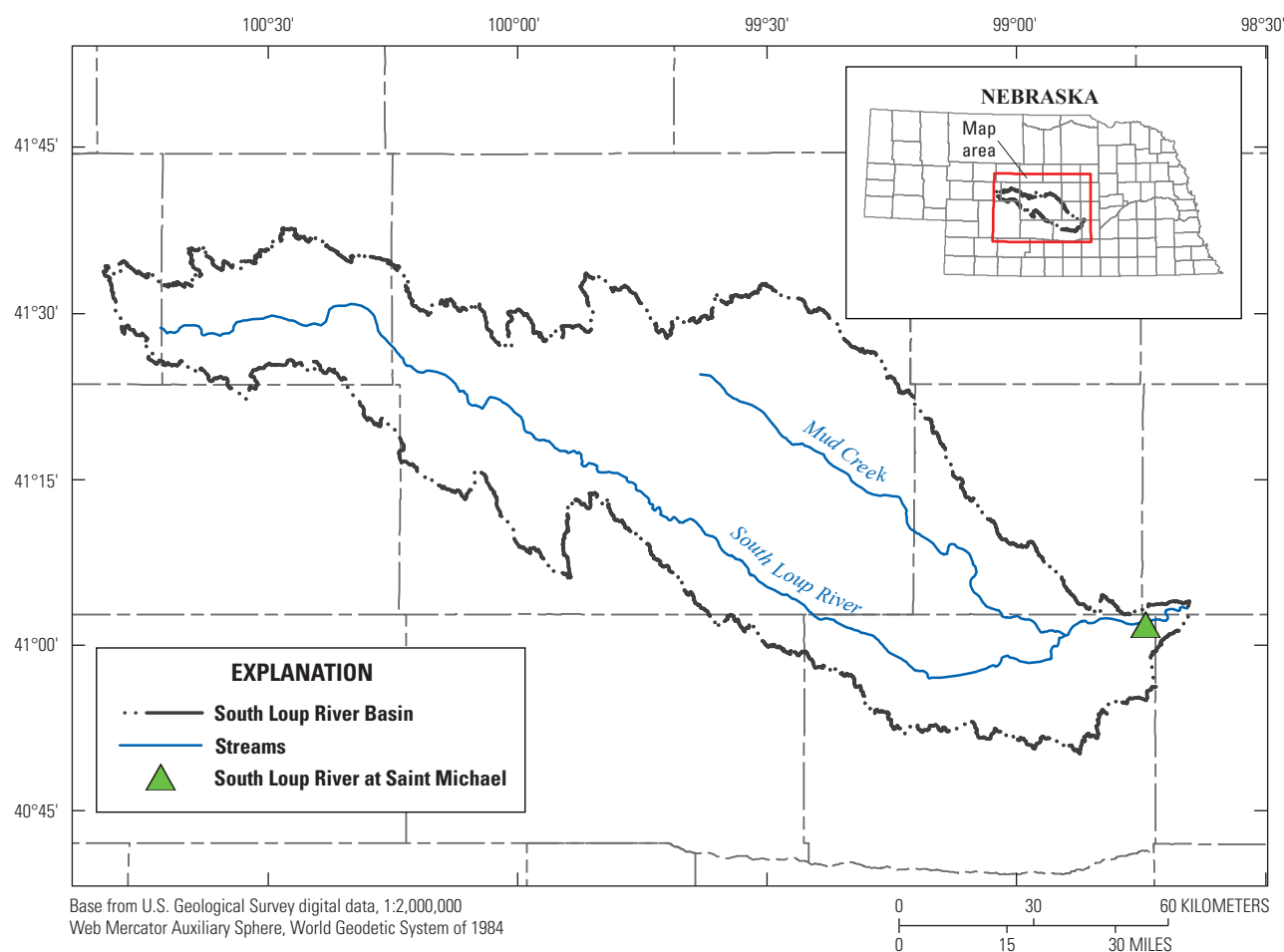

Figure 1.   Map of the South Loup River Basin and the monitoring site on the South Loup River at Saint Michael, Nebraska (U.S. Geological Survey station 06784000).

for data review and approval, followed Wagner and others (2006). The multiparameter continuous water-quality monitor collected water temperature and turbidity data every 15 minutes and transmitted the data to the USGS NWIS database (USGS, 2020). Cross sections of turbidity data collected during maintenance trips indicated that the channel was well mixed and that the placement of the continuous water-quality monitor at the monitoring location was representative of the channel. In addition to the seasonal continuous water-quality data, continuous streamflow data were also collected at the site for the full duration of the study as part of ongoing monitoring at the site. The monitoring seasons in each year were intended to encompass the Nebraska recreation season of May through September, and generally ran from mid-April through October of each year.

Continuous data were published through the USGS NWIS database (USGS, 2020). These data have periodic gaps in their continuous record for various reasons, such as equipment malfunctions, effects from debris on the sensors, or instrument fouling. As a result, measured and approved continuous turbidity data were available for 68 percent of the 2017 seasonal period and 84 percent of the 2018 seasonal period. For this study, gap-filling procedures were applied to short-duration gaps (6 hours or less). For these short-duration gaps, the missing data were filled by linear interpolation between the surrounding values. If the data gap was greater than 6 hours, then data were not filled, and the gap was retained. After completing these gap-filling procedures, measured or estimated continuous turbidity data were available for 92 and 96 percent of the 2017 and 2018 seasonal period, respectively.

Sampling Data

Discrete water-quality samples were collected by cooperating agencies, including the LLNRD as part of the SLRSS and the NDEE as part of its Ambient Monitoring Program (AMP). This resulted in 35 samples co-occurring with the continuous data that were collected between April and October in 2017 and 2018. Sampling data were reviewed by those agencies and published in the Water Quality Portal (National Water Quality Monitoring Council, 2020), from which they were compiled for this report.

The sampling procedures were intentionally similar between the two agencies and consisted of grab samples collected near the water's edge in flowing water. However, the sampling schedule for the AMP was fixed on a monthly basis, whereas the sampling schedule for the SLRSS had more flexibility. This flexibility allowed for targeting runoff events for sampling, thereby characterizing a greater range of streamflows and turbidities. A turbidity-duration curve (Rasmussen and others, 2009) was developed from the continuous turbidity data at the site to explore the representativeness of the sample data. Sampling events were overlain onto that duration curve and demonstrate that discrete samples were collected over 97 percent of the range of turbidities observed at the site during the study period (fig. 2).
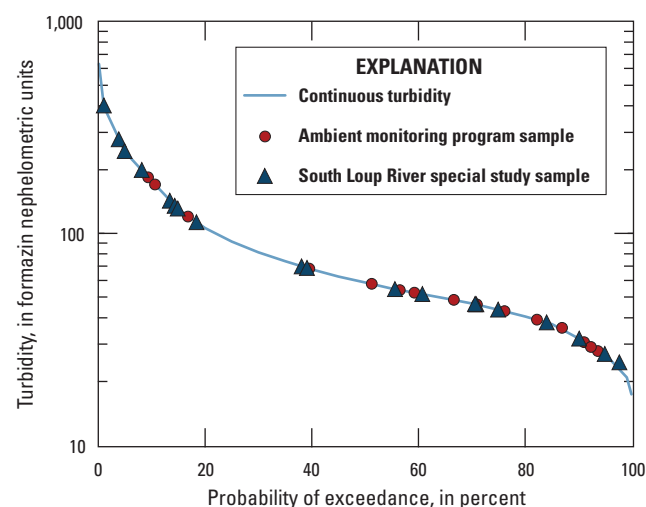
Figure 2. Duration curve for continuous turbidity and corresponding turbidity values associated with samples collected at the South Loup River near Saint Michael, Nebraska, 2017–18 by the Ambient Monitoring Program and the South Loup River Special Study (U.S. Geological Survey, 2020; National Water Quality Monitoring Council, 2020).

The samples were analyzed for bacteria, nutrients, and sediment. *E. coli* was enumerated using the Quantitray method (IDEXX, 2013) from either the NDEE laboratory (for samples collected as part of the AMP) or the LLNRD laboratory (for samples collected as part of the SLRSS). All samples were analyzed for nitrate plus nitrite (NO3) as nitrogen, TKN, TP, and TSS according to standard methods (American Public Health Association and others, 2005) at the Nebraska Public Health Laboratory. TN was estimated for each sample as the sum of NO3 and TKN analyses.

Quality-assurance assessments of the sampling data were carried out by NDEE and LLNRD personnel as part of their data publication process. These assessments relied on 23 replicate samples and 23 blank samples that were collected as part of the SLRSS to help define precision metrics for the data. Data for the 23 replicates were obtained from the NDEE in 2019 (app. 1, table 1.1). Coefficients of variation (CV; Helsel and others, 2020) were computed for each replicate pair to represent the precision of the analyses. The CVs were then averaged across all of the replicate pairs and converted to a percentage (table 1). The CVs indicate *E. coli* data were considerably less precise than the other constituents.

In addition, the published data were assessed for this study to evaluate outliers and address censored data. As a result, four *E. coli* values were subsequently removed from the study data. Two of those omissions were associated with data that were high censored, or greater than the reporting limit. High-censored data are problematic for developing statistical relations because they are not constrained. The other two *E. coli* omissions were associated with abnormally low values (both had concentrations of 3 most probable number of colonies per 100 milliliters). Conversations

Table 1.   Water-quality constituent data from discrete samples collected concurrently with continuous turbidity at the South Loup River at Saint Michael, Nebraska, 2017–18.

[Data are from the National Water Quality Monitoring Council (2020). *E. coli*, *Escherichia coli*, in most probable number of colonies per 100 milliliters; TSS, total suspended solids, in milligrams per liter; TKN, total Kjeldahl nitrogen, in milligrams per liter; NO3, nitrate plus nitrite, in milligrams per liter; N, nitrogen; <, less than; TN, total nitrogen, in milligrams per liter; TP, total phosphorus, in milligrams per liter]

Constituent	Number of samples	Minimum	Maximum	Median	Average coefficient of variation, in percent
E. coli	31	63.0	19,900[a]	488	31.3
TSS	35	40.5	1,420	158	4.42
TKN	35	0.681	4.00	1.40	3.00
NO3 as N	35	<0.05	1.56	0.422	1.69
TN	35	0.913	5.10	1.81	2.41
TP	35	0.167	1.63	0.377	2.39

[a]A high-censored *E. coli* value exceeding the reporting level of 24,196 most probable number of colonies per 100 milliliters was excluded from the statistical analysis.

with NDEE personnel about those low values did not reveal specific collection or analytical issues but noted that similar analyses at other sites on those same days also were abnormally low. Though not omitted from the study, 14 NO3 results were low censored, or lower than the reporting limit of 0.05 milligram per liter (mg/L) as nitrogen. These low-censored data were included in the study by substitution of one-half the reporting limit, or 0.025 mg/L as nitrogen. No other sample data in the dataset were considered outliers or were censored. A summary of the sample data is given in table 1.

Surrogate Model Development

The continuous data were time paired with the discrete sampling data for use in developing surrogate models relating the two datasets. Model development followed the procedures of Rasmussen and others (2009) and was intentionally parsimonious as part of the study design. Only single-variable ordinary least squares regression models (SLRs) with turbidity as an explanatory variable were considered.

Despite constraining the models to turbidity-based SLRs, some model exploration was still needed to ensure that the models met the assumptions of ordinary least squares regressions in which model residuals were homoscedastic (meaning that variance in those residuals was uniformly distributed across the range of observations). Various data transformations were explored during the regression process to assess homoscedasticity using plots of fitted and residual data, quantile-quantile plots, and striving for a probability plot correlation coefficient approaching a value of one (Helsel and others, 2020). In all cases, this exploration indicated that base-10 logarithmic transformations of the turbidity and response analyte were most appropriate for SLR of these data.

Model performance was evaluated for goodness of fit by plotting the data and through several statistical metrics, including the statistical significance (*p*-value) of the explanatory variable (in this case, turbidity), the root mean squared error (RMSE), the model standard percentage error, and the adjusted coefficient of determination (R^2). *P*-values less than a threshold level of 0.05 were considered significant. The RMSE was given in log units and characterized the variance between fitted and observed values. The RMSE was about equal to one standard deviation around a fitted value (Rasmussen and others, 2009). The model standard percentage error is simply a re-expression of the RMSE into a standard error represented as a percentage greater than or less than the fitted value. The R^2 is adjusted for the number of explanatory variables in the model and indicates the fraction of variability in the response variable that is explained by the model.

The model evaluation also considered the inherent measurement imprecision in the sampled constituents using the CV (table 1). Because the CV associated with *E. coli* sample data was an order of magnitude higher than the other constituents, poorer performance of the *E. coli* model was allowed if the model produced homoscedastic residuals and seemed to have a reasonable model fit. But uncertainty metrics associated with the *E. coli* model should be considered in its application.

One consequence of the logarithmic transformations of the data is that retransformation is needed to compute concentrations in the original units. This retransformation introduces a bias in those computations (Koch and Smillie, 1986) unless the data are perfectly correlated. To correct for retransformation bias, Duan (1983) introduced a nonparametric bias-correction factor that is calculated from the mean of the residual values. The retransformed values are then multiplied by the bias correction factor to obtain the concentration in their original units (Rasmussen and others, 2009).

Prediction intervals at the 90-percent confidence level were computed to define the range of values around the model estimate that correspond to the true stream value with 90-percent confidence. Models with greater uncertainty will have a wider prediction interval range. Prediction intervals were computed using the log-transformed data following the approach of Rasmussen and others (2009) and then retransformed back to their original units.

Surrogate Models Using Continuous Turbidity Data to Compute Constituent Concentrations

Surrogate models were developed to relate concentrations for TSS, *E. coli*, TKN, TN, and TP to continuous turbidity data collected during 2017–18 at the South Loup River at Saint Michael, Nebr. The models were developed from 31 *E. coli* and 35 nutrient discrete water-quality samples collected by the NDEE and LLNRD that were paired with continuous turbidity

data collected by the USGS. The regression models relating continuous turbidity to each constituent are shown in table 2 with coefficients and model quality indicators.

Model development was also explored for NO3 using a similar approach as was used for the constituents shown in table 2. However, an evaluation of the model indicated poor model performance and that the assumptions of SLR were not met. As a result, a final model was not developed for NO3. Given that NO3 readily dissolves into water and is associated with surface water and groundwater transport (Dubrovsky and others, 2010), it is not surprising that there is a poor relation of NO3 to turbidity, which is a measure of particulates in the water.

In all cases, the best model performance was achieved when considering logarithm-transformed values of turbidity with logarithm-transformed constituent concentrations. The models explained 51–90 percent of the variability of the response variables as shown by the adjusted R^2 values in table 2. The turbidity coefficient was significant in all models (table 2). Model residuals were normally distributed for each model based on an examination of probability plots and their

Table 2. Surrogate models for selected water-quality constituents at the South Loup River at Saint Michael, Nebraska, 2017–18.

[log, base 10 logarithmic; *EC*, *Escherichia coli*, in most probable number of colonies per 100 milliliters; *Turb*, continuous turbidity, in formazin nephelometric units; *TSS*, total suspended solids, in milligrams per liter; *TKN*, total Kjeldahl nitrogen, in milligrams per liter; *TN*, total nitrogen, in milligrams per liter; *TP*, total phosphorus, in milligrams per liter; *n*, sample size; <, less than; LMSPE, lower model standard percentage error; UMSPE, upper model standard percentage error; R^2, coefficient of determination; RMSE, root mean squared error ; PPCC, probability plot correlation coefficient]

Constituent	*Escherichia coli*	Total suspended solids	Total Kjeldahl nitrogen	Total nitrogen	Total phosphorus
Model equation	log(*EC*)=1.470log (*Turb*)+0.105	log(*TSS*)=1.060log (*Turb*)+0.285	log(*TKN*)=0.496 log(*Turb*)−0.732	log(*TN*)=0.546log (*Turb*)−0.710	log(*TP*)=0.741log (*Turb*)−1.721
Range of variable values used in model	*Turb*: 24.6 to 401 *EC*: 63 to 19,683	*Turb*: 24.6 to 401 *TSS*: 40.5 to 1,420	*Turb*: 24.6 to 401 *TKN*: 0.681 to 4.00	*Turb*: 24.6 to 401 *TN*: 0.913 to 5.1	*Turb*: 24.6 to 401 *TP*: 0.167 to 1.63
n	31	35	35	35	35
Statistical significance (*p*-value) of *Turb* coefficient	<0.001	<0.001	<0.001	<0.001	<0.001
LMSPE, in percent	63.1	29.8	19.4	16.0	16.6
UMSPE, in percent	171	42.5	24.0	19.0	19.9
Adjusted R^2 (dimensionless)	0.51	0.83	0.75	0.85	0.90
RMSE	0.43	0.15	0.093	0.075	0.079
PPCC	0.99	0.99	0.98	0.99	0.99
PPCC *p*-value	0.96	0.90	0.13	0.88	0.53
Bias correction factor (Duan, 1983)	1.59	1.06	1.02	1.01	1.02
Range of variable values used in prediction	*Turb*: 15.4 to 709 *EC*: 113 to 31,461	*Turb*: 15.4 to 709 *TSS*: 37 to 2,149	*Turb*: 15.4 to 709 *TKN*: 0.74 to 4.91	*Turb*: 15.4 to 709 *TN*: 0.88 to 7.11	*Turb*: 15.4 to 709 *TP*: 0.15 to 2.51
Range of 90-percent prediction interval	20 to 170,857	20 to 3,915	0.51 to 7.07	0.66 to 9.55	0.18 to 3.41

probability plot correlation coefficients (table 2). The RMSE in logarithm units is small (0.075–0.15) for all models except for the *E. coli* model (table 2). The lower and upper model standard percentage errors were within 20 percent for only the TN and the TP models and just slightly high for TKN (table 2). Exploratory data analyses indicated that a multiple linear regression model with the addition of streamflow might improve the prediction of TSS but would not improve the *E. coli* model.

The *E. coli* model included considerably more uncertainty than the other models, as indicated by the error indices (upper and lower model standard percentage errors, adjusted R^2, and RMSE, table 2). Given the imprecision of *E. coli* measurements as quantified by CV and discussed earlier (table 1), this uncertainty was expected. Furthermore, although *E. coli* may be introduced to a stream via the same runoff processes that carry sediment, those *E. coli* cells are sensitive to environmental factors, such as sunlight, that may lead to mortality. These complicating factors lead to greater uncertainty and should be considered when using the model.

Predicted Concentrations

Surrogate models (table 2) were used with 15-minute continuous turbidity data to predict 15-minute values of *E. coli* (fig. 3), TSS, TKN, TN, and TP from April 14 through October 24, 2017, and April 2 through October 16, 2018. In addition to predicted values, 90-percent prediction intervals were also calculated to represent the range above and below the predicted value where there is 90-percent confidence that the true concentration is in this range. The predicted values and the 90-percent prediction intervals are published in the NWIS database as part of the continuous data at the site (USGS, 2020), although gap-filling procedures used in this report are not similarly applied in the NWIS database. Continuous turbidity data ranged from 15.4 to 709 formazin nephelometric units as compared to 24.6 to 401 formazin nephelometric units in the calibration dataset (table 2). The largest prediction intervals are associated with the *E. coli* model that predicts concentrations ranging from 113 to 31,461 most probable number of colonies per 100 milliliters and prediction intervals ranging from 20 to 170,857 most probable number of colonies per 100 milliliters (fig. 3 and table 2).

A comparison of predicted *E. coli* values to the recreation season (May 1 to September 30) water-quality criterion of 126 most probable number of colonies per 100 milliliters (Nebraska Department of Environmental Quality, 2019) shows that predicted *E. coli* concentrations were above the recreation season criterion for most of the monitoring period (fig. 3). Similar criteria are not in place for the other modeled constituents (table 2).

These results are important for setting a baseline for future comparisons. However, it should also be noted that these results are only representative of the 2017–18 monitoring period, and more specifically to the relation of the constituents to turbidity during that period. The hydrologic conditions of the South Loup during 2017–18 would need to be considered before drawing conclusions to future periods that may or may not be comparable. Comparisons are similarly challenged by landscape changes that may affect one model variable more than the other. For example, land-management practices designed to reduce the application of fertilizers may reduce nutrients in the stream without affecting the turbidity levels. As a result, collection of additional data during a future reference period would be helpful to redefine the relation of turbidity (or perhaps some other surrogate measurement) to bacterial counts and stream nutrients. However, these challenges are tractable, and these models are expected to not only provide stakeholders with an understanding of constituent concentrations during the 2017–18 monitoring period but also provide a good reference point for any future comparisons.
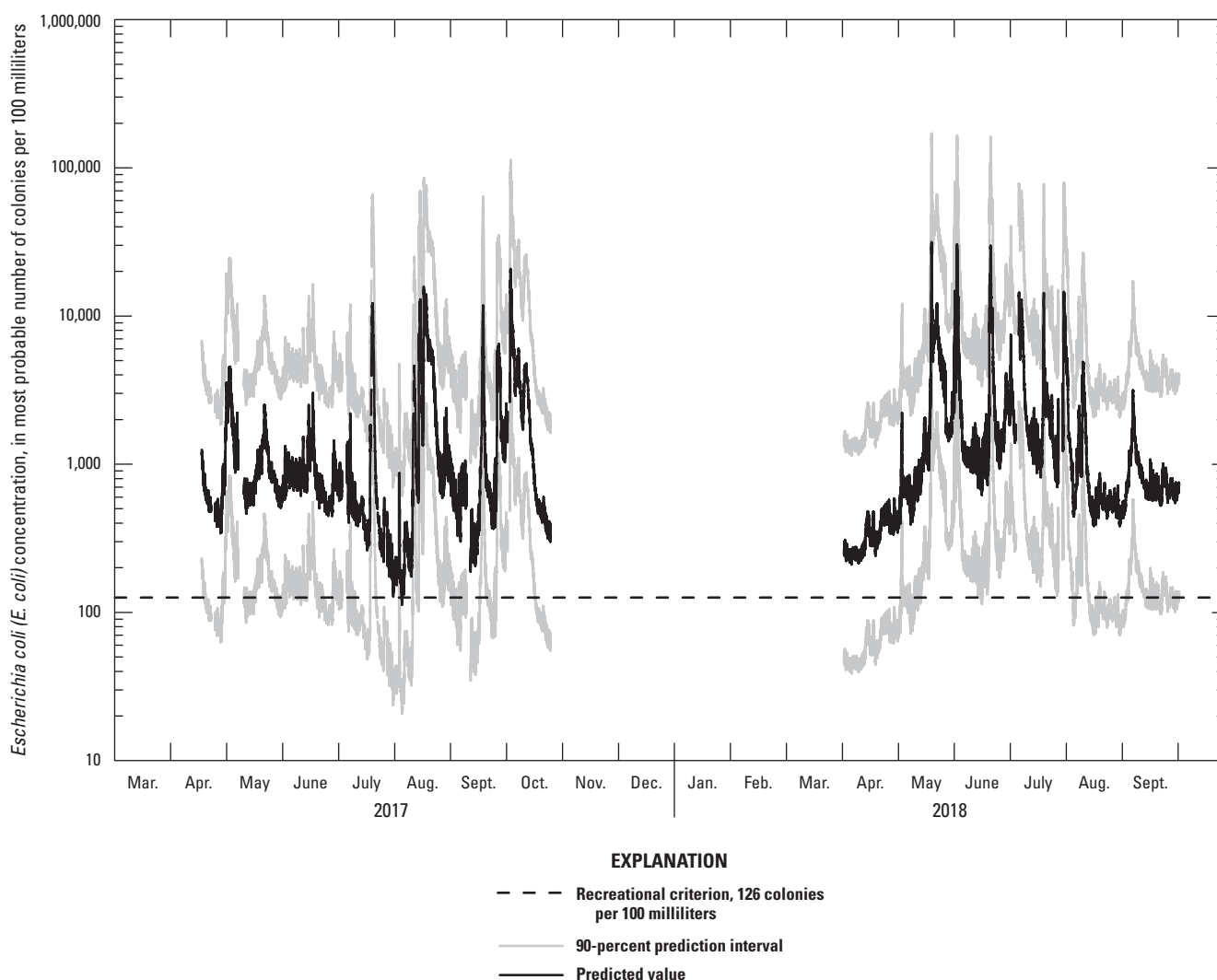
Figure 3. *Escherichia coli* concentrations predicted from continuous turbidity values at the South Loup River near Saint Michael, Nebraska, 2017–18 (U.S. Geological Survey, 2020).

Summary

The South Loup River has been impaired by bacteria since at least 2004, which has resulted in the river not meeting its intended use as a recreational waterway. As part of a strategy for reducing the bacterial load in the river, the U.S. Geological Survey (USGS), in cooperation with the Lower Loup Natural Resources District, made continuous estimates of *Escherichia coli* (*E. coli*) and nutrient concentrations in the river to help resource managers identify periods associated with higher concentrations and to provide a baseline before mitigation activities. The continuous estimates were developed from seasonal monitoring data collected at the South Loup River at Saint Michael, Nebraska during, 2017–18.

Continuous water-quality data were represented by turbidity and water temperature measurements made every 15 minutes from mid-April through October in 2017 and 2018.

The data were collected according to standard methods and are available online at the USGS National Water Information System database. Periodic gaps in the data, typically of 6 hours or less, were filled by linear interpolation for the purposes of this study.

During the same monitoring period, 35 discrete water samples were collected at the site and were analyzed for *E. coli*, nitrate plus nitrite (NO3), total Kjehldahl nitrogen (TKN), total phosphorus (TP), and total suspended solids (TSS) using standard methods. Total nitrogen (TN) was estimated for each sample as the sum of NO3 and TKN analyses. Quality assurance assessments of the data indicated that the *E. coli* data were considerably less precise than the other constituents.

The continuous data were time-paired with the discrete sampling data for use in developing surrogate models relating the two datasets. Model development followed standard procedures and were intentionally parsimonious as part of the study design. Only single-variable ordinary least squares regression

models were considered and only turbidity was considered as an explanatory variable. Different data transformations were explored to ensure that the assumptions of the regression were being met. In all cases, base-10 logarithmic transformations of the turbidity and response analyte were found to be most appropriate. Model performance was evaluated using several goodness of fit metrics and the inherent measurement imprecision of the data from which the models were based upon.

Using this approach, surrogate models were developed that related turbidity to *E. coli*, TKN, TN, TP, and TSS concentrations at the Saint Michael site in 2017–18. Model imprecision, as evaluated by the root mean squared error, was small for all models except for the *E. coli* model. Greater uncertainty was expected for *E. coli* given the measurement imprecision of *E. coli* and sensitivity to environmental factors that could lead to mortality. The data did not support the development of a similar model for NO3.

Once developed, the models were used in conjunction with the continuous turbidity data to predict constituent concentrations at a similar continuous time step. Prediction intervals at the 90-percent confidence level were also computed to define the range of values around the model estimate that correspond to the true constituent value with 90-percent confidence. The predictions and prediction intervals are available online as part of the USGS National Water Information System database. As with the model performance, the largest prediction intervals were associated with the *E. coli* estimates. Nonetheless, a comparison of predicted *E. coli* values to water-quality criteria indicated that levels in the stream were above the recreational threshold for most of the monitoring period.

The results in this report are important for setting a baseline for any future comparisons. It should be noted that these results are representative of the relation of turbidity to the modeled constituents during the 2017–18 monitoring period. Because of hydrologic variation and the potential for landscape changes to affect one model variable more than another, collection of additional data during a future reference period would be helpful to redefine the relation of turbidity (or perhaps some other surrogate measurement) to bacteria counts and stream nutrients. However, these challenges are tractable, and these models are expected to not only provide stakeholders with an understanding of constituent concentrations during the 2017–18 monitoring period but also provide a good reference point for any future comparisons.

References Cited

American Public Health Association, American Water Works Association, and Water Environment Federation, 2005, Standard methods for the examination of water and wastewater (19th ed.): Washington, D.C., American Public Health Association, 905 p.

Bleed, A.S., and Flowerday, C.A., eds., 1989, An atlas of the Sand Hills (3d ed.): Lincoln, University of Nebraska, Conservation and Survey Division, Resource Atlas no. 5b, 260 p.

Duan, N., 1983, Smearing estimate—A nonparametric retransformation method: Journal of the American Statistical Association, v. 78, no. 383, p. 605–610. [Also available at https://doi.org/10.1080/01621459.1983.10478017.]

Dubrovsky, N.M., Burow, K.R., Clark, G.M., Gronberg, J.M., Hamilton, P.A., Hitt, K.J., Mueller, D.K., Munn, M.D., Nolan, B.T., Puckett, L.J., Rupert, M.G., Short, T.M., Spahr, N.E., Sprague, L.A., and Wilber, W.G., 2010, The quality of our Nation's waters—Nutrients in the Nation's streams and groundwater, 1992–2004: U.S. Geological Survey Circular 1350, 174 p. [Also available at https://doi.org/10.3133/cir1350.]

Garrett, J.D., 2019, The use of continuous water-quality time-series data to compute nutrient loadings for selected Iowa streams, 2008–17: U.S. Geological Survey Scientific Investigations Report 2019–5054, 31 p., accessed January 5, 2020, at https://doi.org/10.3133/sir20195054.

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chap. A3, 458 p. [Also available at https://doi.org/10.3133/tm4A3.] [Supersedes U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. A3, ver. 1.1.]

Hobza, C.M., and Schepers, A.R., 2018, Groundwater discharge characteristics for selected streams within the Loup River Basin, Nebraska, 2014–16: U.S. Geological Survey Scientific Investigations Report 2018–5093, 50 p., accessed January 5, 2020, at https://doi.org/10.3133/sir20185093.

IDEXX, 2013, IDEXX Quanti-Tray/2000 insert and most probable number (MPN) table: Westbrook, Maine, IDEXX, accessed November 2020 at https://www.idexx.com/files/quanti-tray-2000-procedure-en.pdf.

Koch, R.W., and Smillie, G.M., 1986, Bias in hydrologic prediction using log-transformed regression models—American Water Resources Association: Journal of the American Water Resources Association, v. 22, no. 5, p. 717–723. [Also available at https://doi.org/10.1111/j.1752-1688.1986.tb00744.x.]

National Water Quality Monitoring Council, 2020, Water quality portal: National Water Quality Monitoring Council website and digital data, accessed January 5, 2020, at https://www.waterqualitydata.us.

Nebraska Department of Environment and Energy [NDEE], 2013, A list of TMDLs in Nebraska: Nebraska Department of Environment and Energy digital data, accessed October 1, 2020, at http://deq.ne.gov/NDEQProg.nsf/OnWeb/TMDLlist.

Nebraska Department of Environment and Energy [NDEE], 2018, 2018 water quality integrated report: Nebraska Department of Environment and Energy, 376 p., accessed January 5, 2020, at http://dee.ne.gov/NDEQProg.nsf/PubsForm.xsp?databaseName=CN=DEQSER6/O=NDEQ!!publica.nsf&documentId=89721A3F201CE5348625827A006BF7D4&action=editDocument.

Nebraska Department of Environmental Quality, 2019, Nebraska surface water quality standards: Nebraska Administrative Code, title 117, 235 p., accessed January 5, 2020, at http://deq.ne.gov/RuleAndR.nsf/pages/PDF/%24FILE/Title117_2019.pdf.

Rasmussen, P.P., Gray, J.R., Glysson, G.D., and Ziegler, A.C., 2009, Guidelines and procedures for computing time-series suspended-sediment concentrations and loads from in-stream turbidity-sensor and streamflow data: U.S. Geological Survey Techniques and Methods, book 3, chap. C4, 52 p. [Also available at https://doi.org/10.3133/tm3C4.]

Soenksen, P.J., Miller, L.D., Sharpe, J.B., and Watton, J.R., 1999, Peak-flow frequency relations and evaluation of the peak-flow gaging network in Nebraska: U.S. Geological Survey Water-Resources Investigations Report 99–4032, 110 p. [Also available at https://doi.org/10.3133/wri994032.]

U.S. Geological Survey [USGS], 2020, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed November 2, 2020, at https://doi.org/10.5066/F7P55KJN.

Wagner, R.J., Boulger, R.W., Jr., Oblinger, C.J., and Smith, B.A., 2006, Guidelines and standard procedures for continuous water-quality monitors—Station operation, record computation, and data reporting: U.S. Geological Survey Techniques and Methods, book 1, chap. D3, 51 p., 8 attachments, accessed January 5, 2020, at https://doi.org/10.3133/tm1D3.

Appendix 1.   Paired Replicate Sampling Data

Quality-assurance assessments of the sampling data were carried out by Nebraska Department of Environment and Energy and Lower Loup Natural Resources District personnel as part of their data publication process. These assessments included 23 replicate samples that were collected as part of the South Loup River Special Study to help define precision metrics for the data. Data for the 23 replicates were obtained from the Nebraska Department of Environment and Energy in 2019 and are provided here (table 1.1).

Table 1.1.   Paired replicate sampling data collected at the South Loup River near Saint Michael, Nebraska, 2017–18.

[Data provided by Nebraska Department of Environment and Energy. Dates are given in month/day/year. *E. coli*, *Escherichia coli*, in most probable number of colonies per 100 milliliters; TSS, total suspended solids, in milligrams per liter; TKN, total Kjeldahl nitrogen, in milligrams per liter; NO3, nitrate plus nitrite, in milligrams per liter; N, nitrogen; TN, total nitrogen, in milligrams per liter; TP, total phosphorus, in milligrams per liter; <, less than; >, greater than]

Sample date	Routine samples						Replicate samples					
	E. coli	TSS	TKN	NO3 as N	TN	TP	*E. coli*	TSS	TKN	NO3 as N	TN	TP
5/1/2017	6,867	504	1.43	0.848	2.278	0.569	5,172	484	1.43	0.839	2.269	0.545
5/15/2017	276	208	1.24	0.422	1.662	0.377	308	235	1.26	0.429	1.689	0.391
5/19/2017	4,106	296	1.27	0.827	2.097	0.484	2,987	326	1.34	0.852	2.192	0.522
6/12/2017	137	168	1.78	<0.05	1.805	0.356	276	168	1.68	<0.05	1.705	0.358
6/26/2017	613	89	1.7	<0.05	1.725	0.31	240	92	1.7	<0.05	1.725	0.31
7/10/2017	63	64	1.21	<0.05	1.235	0.244	387	59	1.2	<0.05	1.225	0.242
7/24/2017	435	58	1.14	<0.05	1.165	0.269	345	60	1.06	<0.05	1.085	0.247
8/3/2017	1,046	57	1.09	<0.05	1.115	0.205	1,046	58	1.05	<0.05	1.075	0.212
8/7/2017	3	58	0.982	<0.05	1.007	0.181	110	64	0.931	<0.05	0.956	0.175
8/11/2017	1,553	258	2.75	0.279	3.029	0.739	1,414	238	2.66	0.277	2.937	0.726
8/16/2017	19,863	864	4	1.1	5.1	1.4	15,531	968	4.02	1.11	5.13	1.42
8/21/2017	1,046	346	2.26	0.819	3.079	0.912	579	296	2.36	0.812	3.172	0.918
9/5/2017	261	55	0.958	<0.05	0.983	0.27	74	52	0.96	<0.05	0.985	0.265
9/18/2017	6,131	220	2.19	0.176	2.366	0.795	8,664	229	1.83	0.177	2.007	0.773
10/2/2017	1,421	322	1.49	0.747	2.237	0.485	2,420	308	1.55	0.753	2.303	0.514
5/7/2018	73	138	1.37	<0.05	1.395	0.313	63	140	1.38	<0.05	1.405	0.308
5/11/2018	261	168	1.7	<0.05	1.725	0.41	260	163	1.73	<0.05	1.755	0.424
5/16/2018	387	113	1.41	<0.05	1.435	0.299	517	122	1.39	<0.05	1.415	0.303
5/21/2018	>24,196	660	3.79	0.904	4.694	1.37	>24,196	610	3.88	0.91	4.79	1.35
6/20/2018	4,884	1,420	3.9	0.452	4.352	1.63	6,488	1,440	3.86	0.444	4.304	1.62
6/26/2018	1,300	232	1.4	1.02	2.42	0.605	1,553	218	1.4	1.01	2.41	0.615
8/20/2018	548	125	1.4	<0.05	1.425	0.302	517	115	1.27	<0.05	1.295	0.297
9/17/2018	345	128	1.05	1.06	2.11	0.273	365	120	1.29	0.693	1.983	0.332

USGS

Rus and Densmore—**Continuous Turbidity Data to Compute Constituent Concentrations, South Loup River, Nebraska, 2017–18**—Scientific Investigations Report 2021-5120