A COMPARISON OF STATISTICAL APPROACHES FOR PREDICTING STREAM TEMPERATURES ACROSS HETEROGENEOUS LANDSCAPES[1]

Kevin E. Wehrly, Travis O. Brenden, and Lizhu Wang[2]

ABSTRACT: Estimating stream temperatures across broad spatial extents is important for regional conservation of running waters. Although statistical models can be useful in this endeavor, little information exists to aid in the selection of a particular statistical approach. Our objective was to compare the accuracy of ordinary least-squares multiple linear regression, generalized additive modeling, ordinary kriging, and linear mixed modeling (LMM) using July mean stream temperatures in Michigan and Wisconsin. Although LMM using low-rank thin-plate smoothing splines to measure the spatial autocorrelation in stream temperatures was the most accurate modeling approach; overall, there were only slight differences in prediction accuracy among the evaluated approaches. This suggests that managers and researchers can select a stream temperature modeling approach that meets their level of expertise without sacrificing substantial amounts of prediction accuracy. The most accurate models for Michigan and Wisconsin had root mean square errors of 2.0-2.3°C, suggesting that only relatively coarse predictions can be produced from landscape-based statistical models at regional scales. Explaining substantially more variability in stream temperatures likely will require the collection of finer-scale hydrologic and physiographic data, which may be cost prohibitive for monitoring and assessing stream temperatures at regional scales.

(KEY TERMS: rivers/streams; temperature; mixed models; geospatial analysis; landscape features; watershed management.)

INTRODUCTION

Water temperature is an important characteristic influencing stream ecosystems (Allan and Castillo, 2007). Water temperature regulates biological processes of aquatic ectotherms (Brett, 1971; Hokanson *et al.*, 1977; Taniguchi *et al.*, 1998; Gillooly *et al.*, 2002) and plays a major role in regulating biological communities (Vannote and Sweeney, 1980; Rahel and Hubert, 1991; Lyons, 1996; Hawkins *et al.*, 1997; Sponseller *et al.*, 2001; Steen *et al.*, 2008). Thus, water temperature is an essential habitat feature that should be considered in stream conservation efforts. At regional scales, however, stream temperature measurements are typically available for relatively few locations and models capable of predicting temperature at unsampled locations can be of great

management value. For example, broad-scale temperature predictions have been used to map species distributions to identify gaps in conservation efforts (McKenna *et al.*, 2006; Steen *et al.*, 2008), assess human impacts (Eaton and Scheller, 1996; Rahel *et al.*, 1996; Nelitz *et al.*, 2007), and classify stream reaches based on their potential to support different species assemblages to guide fisheries management (Seelbach *et al.*, 2006; Brenden *et al.*, 2008).

Models for predicting stream temperature have been developed using both deterministic and statistical approaches. Deterministic or physical-process models are based on energy budget equations and use inputs such as solar radiation, evaporation, and conduction to calculate thermal exchange between the atmosphere and streams. Deterministic models are typically developed for individual stream reaches and require extensive data inputs (e.g., stream geometry, state hydrology, and meteorology) to calibrate. Although deterministic models can provide very accurate estimates of stream temperature (Sinokrot and Stefan, 1993; Kim and Chapra, 1997; Gu *et al.*, 1998; Younus *et al.*, 2000; Horne *et al.*, 2004; Caissie *et al.*, 2007), the reach-specific data inputs and predictions make them impractical for use at regional scales (Mohseni *et al.*, 1998; Risley *et al.*, 2003), particularly when the goal is to estimate water temperature at hundreds of unsampled locations. In contrast, statistical models do not directly account for sources of heat flux but instead rely on variables that are correlated with water temperature such as air temperature and landscape characteristics (Caissie, 2006; Webb *et al.*, 2008). Statistical models based on air temperature explain much of the variation in water temperature when developed for specific locations or for a limited number of locations having similar characteristics (Crisp and Howson, 1982; Mackey and Berrie, 1991; Stephan and Preud'homme, 1993; Mohseni *et al.*, 1998). However, as more types of streams are considered in model development, the strength of the air-water temperature relationship diminishes (Stephan and Preud'homme, 1993; Caissie, 2006). Statistical temperature models that include multiple stream types typically incorporate geomorphic, riparian, and catchment characteristics (Hawkins *et al.*, 1997; Isaak and Hubert, 2001; Sponseller *et al.*, 2001; Abell and Allan, 2002; Scott *et al.*, 2002; Risley *et al.*, 2003; Wehrly *et al.*, 2006; Nelitz *et al.*, 2007).

Most statistical temperature models have used relatively simple methods, such as simple and multiple linear regression (MLR) and nonlinear regression techniques (reviewed in Caissie, 2006 and Webb *et al.*, 2008). Recently, a number of new modeling approaches have been used to predict water temperature such as kriging (KRIG) (Gardner *et al.*, 2003; Gardner and Sullivan, 2004), wavelet analysis (Steel

and Lange, 2007), evolutionary polynomial regression (Giustolisi *et al.*, 2007), and artificial neural networks (Risley *et al.*, 2003; Roehl *et al.*, 2006; Stewart *et al.*, 2006; Karacor *et al.*, 2007; Sivri *et al.*, 2007). Additional statistical approaches such as generalized additive modeling (GAM) and linear mixed modeling (LMM) may improve temperature prediction accuracy by accounting for nonlinear relationships between response and predictor variables and by accounting for spatial autocorrelation among sampling sites. Despite a growing list of analytical techniques, little information exists to aid in the selection of a particular modeling approach because few comparative studies have been published (Caissie, 2006; Webb *et al.*, 2008). The purpose of this study was to compare the accuracy of MLR, GAM, KRIG, and LMM for predicting July mean stream temperatures for streams located across Michigan and Wisconsin. Such an assessment of model performance is aimed to provide knowledge of model accuracy, bias, and model selection to aquatic resource managers and researchers who are interested in developing their own statistical temperature models.

MATERIALS AND METHODS

Data Collection

We based our evaluation on data from 820 wadeable stream sites in Michigan and 311 wadeable stream sites in Wisconsin (Figure 1). Streams in these states primarily drain low-elevation landscapes of glacial origin. Climate, topography, and land use are relatively similar between the states. Additionally, ground-water loading plays an important role in shaping stream temperature and flow regimes in both states (Wehrly *et al.*, 2006). Summer water temperatures for the study streams were measured using continuous-recording temperature loggers that were deployed between 1989 and 2005 by the Michigan and Wisconsin Departments of Natural Resources. July mean water temperature for each stream site was computed from the logged temperatures. July mean temperature is a useful predictor of fish assemblage structure (Wehrly *et al.*, 2003; Steen *et al.*, 2008) and July is the time in northern latitudes when temperature differences among streams are most pronounced (Caissie *et al.*, 2006; Kevin E. Wehrly, unpublished data). Of the 820 sites in Michigan where stream temperature measurements were collected, 64 sites had more than one year of temperature measurements. Similarly, of the 311 sites in Wisconsin, 31 sites had more than one year of
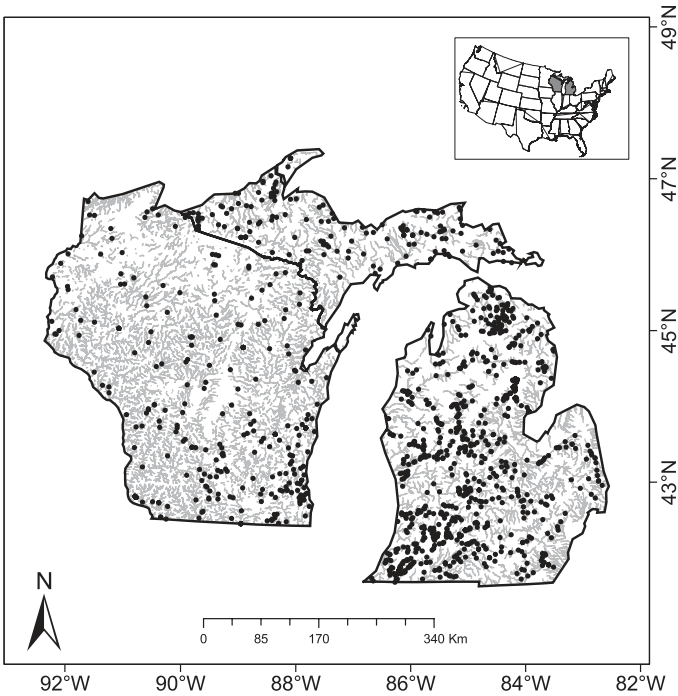
FIGURE 1. Michigan and Wisconsin
Stream Temperature Sampling Sites.

temperature measurements available. Because we were interested in characterizing the "typical" temperature regime experienced by biota at each site, we averaged temperatures across years so that only one temperature measurement per stream occurred within the datasets. Alternatively, we could have randomly selected a single year of measurement for the sites that had multiple years of measurement to be used in our evaluations. Ultimately, we found that the relative performance of the modeling approaches that we evaluated in this study was the same regardless of whether we used average temperatures or randomly selected a single year of temperature measurements. Consequently, we only report the results in which we averaged temperatures for sites with multiple years of data. The study streams were a diverse set of stream types encompassing a range of thermal conditions (Figure 2).

Model Development

Stream temperatures were modeled as a spatial random field that could be represented with the equation

$$Z(s) = X(s)\beta + e(s), \qquad (1)$$

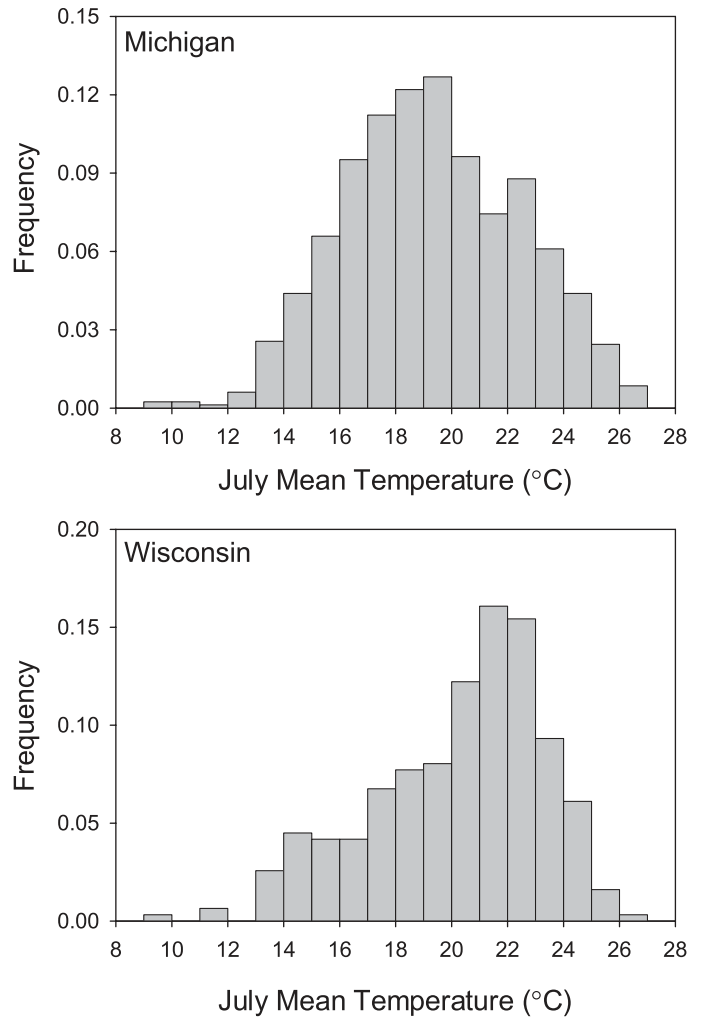where $Z(s)$ is a vector of stream temperatures (response variable) observed at locations $s_1, s_2, \ldots s_n$;

FIGURE 2. Histograms of Observed July Mean Water
Temperature in Michigan and Wisconsin Streams.

X is a matrix of model covariates (predictor variables) measured at the locations; and $e(s)$ is a vector of random model errors (Schabenberger and Gotway, 2005). For generality, it was only assumed that $e(s)$ had a mean of 0 and a variance-covariance matrix denoted by Σ.

The first approach that we used to predict stream temperature was ordinary least-squares MLR. With the MLR approach, it is assumed that stream temperatures at different locations have the same variance and are uncorrelated. By uncorrelated, we mean that temperature at one stream conveys no information about temperatures at other nearby streams. In other words, Σ is assumed to equal $\sigma^2 I$, where I is an $n \times n$ identity matrix. With this model structure, all variability in stream temperatures, apart from random white-noise error, is associated with changes in the mean function (Schabenberger and Gotway, 2005). We modeled the mean function as a linear combina-

TABLE 1. Descriptive Statistics for Predictor Variables Used to Develop Temperature Models for Michigan and Wisconsin Streams.

Predictor Variable	Michigan			Wisconsin		
	Minimum	Median	Maximum	Minimum	Median	Maximum
AREA	0.0	4.7	9.6	−0.1	4.1	10.2
FOREST	0.0	37.6	97.9	0.0	22.4	97.0
JULAIR	14.6	20.8	25.4	17.0	21.9	25.8
PERM	73.6	636.6	1,300.0	63.0	263.3	1,242.1
SLOPE	0.0	1.2	7.4	1.0	5.3	16.1
WATER	0.0	0.9	16.8	0.0	0.4	30.4

Notes: See Brenden *et al.* (2006) for a description of how the stream attributes were calculated. [AREA = $\log_e$ network catchment surface area (km^2), FOREST = percent forest land use/type in local catchments, JULAIR = July mean air temperature (°C), PERM = mean soil permeability in network catchments (cm/100 h), SLOPE = mean network catchment slope (°), and WATER = percent water in network catchments.]

tion of six landscape characteristics that were summarized for the study stream sites: percent forest land use in local catchments, percent water land type in network catchments, July mean air temperature (°C), $\log_e$ transformed network catchment surface area (km^2), mean network catchment slope (%), and mean soil permeability in network catchments (cm/100 h) (Table 1). These predictor variables were chosen because preliminary analyses revealed they were correlated with water temperature. Network and local catchments differ with respect to drainage area boundaries. We defined local catchments as only those areas that drain directly to a particular stream reach and we defined network catchments as all areas that drain to a stream reach either by overland or waterway routes (Figure 3). Predictor variables for the study streams were summarized using a previ-

ously developed stream attribute database for upper Midwest United States (U.S.) states (Brenden *et al.*, 2006). Land use/type for study streams were summarized using 30-m resolution land cover/use Geographic Information System (GIS) themes for the states (MCGI, 2004; WDNR, 2004). Air temperature was summarized from Oregon State University/Spatial Climate Analysis Service datasets (OSU/SCAS, 2004). Mean network catchment slope for stream reaches was calculated using the U.S. Geological Survey National Elevation Dataset (USGS, 2004a). Mean soil permeability was calculated from U.S. Geological soil survey database (USGS, 2004b). The MLR models were fit in SAS (SAS Institute, 2004) using the GLIMMIX procedure.

Our second approach for predicting stream temperatures was GAM. GAM is a semiparametric modeling

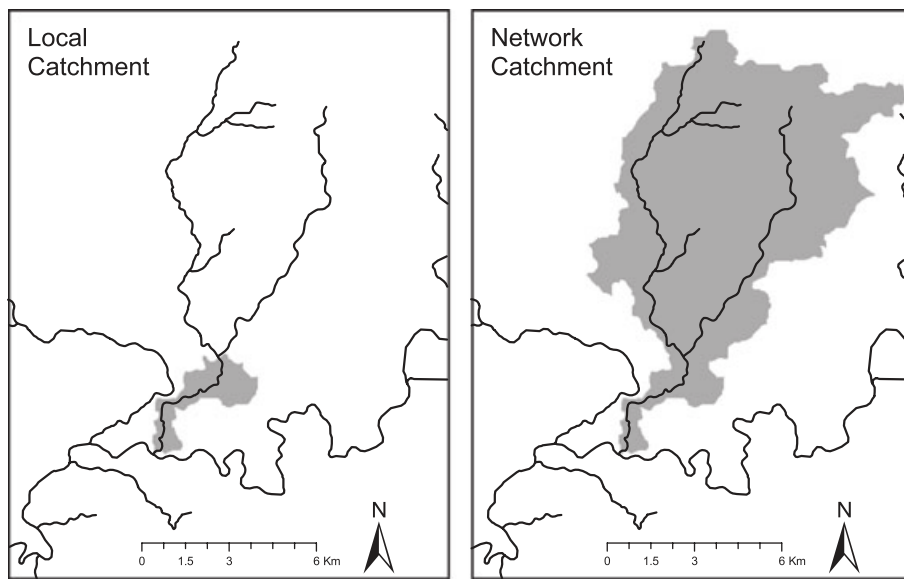FIGURE 3. Illustrated Differences Between Local and Network Catchments, Which Were Spatial Scales Used to Summarize Land Use/Type for Stream Reaches to Predict Stream Temperatures. Local catchments (left panel) included only those upstream areas that drained directly to a particular stream reach. Network catchments (right panel) included all upstream areas that drained to a stream reach either by overland or waterway routes.

approach for identifying nonlinear regression effects between response and explanatory variables (Hastie and Tibshirani, 1990). Whereas the MLR approach assumed that there was a linear relationship between stream temperature and model covariates, the GAM approach only assumed that there were smooth relationships between stream temperature and the model covariates; the exact nature of these smoothed relationships were estimated as part of the model fitting process. To our knowledge, GAM has not previously been used to predict stream temperatures, but may be advantageous for doing so because of its ability to account for nonlinearities between response and predictor variables without requiring explicit specification of model equations. Like the MLR approach, the GAM approach assumed that streams had the same variance and were uncorrelated (i.e., $\Sigma = \sigma^2\mathbf{I}$), and that all variability in stream temperatures (apart from random, white-noise error) was due to changes associated with the mean function. As with the MLR approach, we modeled the mean function for the GAM approach as a linear combination of the six landscape characteristics listed in Table 1. We fit the GAM models in SAS using the GAM procedure. We used cubic smoothing splines to estimate the relationships between stream temperature and the model covariates. The amount of smoothing for the model covariates was determined through generalized cross validation.

Our third approach for predicting stream temperature was ordinary KRIG. KRIG is an interpolation technique that predicts information at unsampled locations from measurements at sampled locations. From a prediction standpoint, KRIG can be represented with the equation

$$\mathbf{Z}(\mathbf{s}) = \mu + \mathbf{e}(\mathbf{s}), \tag{2}$$

where μ indicates that there is a constant (and unknown) mean at all locations. Whereas with the MLR and GAM approaches it was assumed that most variability in stream temperature was due to changes in the mean function, with KRIG it is assumed that most variability in stream temperatures is associated with spatial dependency among streams. Thus, with the KRIG approach, it is no longer assumed that $\Sigma = \sigma^2\mathbf{I}$. Ordinary KRIG predictions of stream temperature at unobserved locations are calculated using KRIG weights, which determine the influence of neighboring observations on predicted values. KRIG weights are affected by Σ and the covariances between stream temperatures at predicted and observed locations (Schabenberger and Gotway, 2005). It is typical with the KRIG approach to use some parametric model to capture the spatial

dependence among observations (Schabenberger and Gotway, 2005). Based on the empirical semivariograms that were calculated from the Michigan and Wisconsin datasets, we chose to use a spherical semivariogram model to model the spatial dependence among streams. Although Gardner *et al.* (2003) found that KRIG using the shortest stream length between sampling locations yielded the most accurate stream temperatures for a river network, for many watersheds in our dataset only a few stream temperature measurements occurred within the same upstream and downstream network. For example, in a hypothetical stream network, temperature measurements may be available for a site in the headwaters and for a site located tens of kilometers downstream. Because thermal regimes typically vary as a function of network position (Caissie, 2006), the temperature of the headwater site in our hypothetical stream would be more similar to other headwater streams in adjacent river networks than to the downstream site in the same network. We thus felt that the sparse temperature data available within each network were inadequate to use in-water path distance as the basis for assessing spatial dependence among observations. Instead, we used Euclidean distance, or the shortest distance between two sampling locations regardless of whether the path was overland or instream, as the basis for assessing spatial dependence among observations. We used the MIXED procedure in SAS and restricted maximum likelihood to estimate the parameters of the spherical semivariogram model. Ordinary KRIG of the stream temperatures for the Michigan and Wisconsin datasets was conducted using the KRIGE2D procedure in SAS.

Our fourth and fifth approaches for predicting stream temperatures were based on LMM. The LMM approaches differed from the other approaches used in this study in that the spatial structure in stream temperatures was incorporated as part of both the mean function and error process, rather than as part of the mean function or error process separately. Based on the relative performance of the GAM and MLR approaches, we chose only to consider linear relationships between stream temperature and the model covariates, although, with LMM, nonlinearities between response and predictor variables could also have been considered. For the LMM approaches to predicting stream temperature, the random error [$\mathbf{e}(\mathbf{s})$] in Equation (1) gets separated into two components, a component that includes smooth-scale and micro-scale variation in stream temperatures and a white-noise measurement error component (Schabenberger and Gotway, 2005). Thus, Equation (1) for the LMM approach can be decomposed to

$$\mathbf{Z}(\mathbf{s}) = \mathbf{X}(\mathbf{s})\beta + \upsilon(\mathbf{s}) + \varepsilon(\mathbf{s}), \qquad (3)$$

where $\upsilon(\mathbf{s})$ represents the smooth-scale and micro-scale variation component and $\varepsilon(\mathbf{s})$ represents the white-noise measurement error component [the KRIG approach assumes a similar decomposition of $\mathbf{e}(\mathbf{s})$]. With Equation (3), it is assumed that $\upsilon(\mathbf{s})$ has a mean of $\mathbf{0}$ and a variance-covariance matrix of $\sum_S$, and that $\varepsilon(\mathbf{s})$ has a mean of $\mathbf{0}$ and a variance-covariance matrix of $\sigma_\varepsilon^2 \mathbf{I}$. From a mixed models standpoint, $\upsilon(\mathbf{s})$ forms the random effects component of the model (Schabenberger and Gotway, 2005). As with the MLR and GAM approaches, we modeled the mean function in Equation (3) as a linear combination of the six landscape characteristics listed in Table 1. The fourth and fifth approaches differed in respect to how the spatial dependence among streams was modeled. For the fourth approach (hereafter referred to as LMM-Spherical), we predicted stream temperature by assuming $\upsilon(\mathbf{s})$ was a zero mean random field with spherical covariance structure. Thus, the fourth approach combined elements of both the MLR and KRIG approaches. For the fifth approach (hereafter referred to as LMM-Smooth), we used low-rank radial smoothing splines to model $\upsilon(\mathbf{s})$. This approach entailed placing "knots" at various locations throughout the study area. Smoothing splines were then used to model the spatial autocorrelations in stream temperatures based on distances of streams to knots (Ruppert *et al.*, 2003; Schabenberger and Gotway, 2005). This approach is computationally efficient as only stream-to-knot rather than stream-to-stream distances need to be calculated (Schabenberger and Gotway, 2005). We used the MIXED procedure in SAS to fit the models for the LMM-Spherical approach. We used the GLIMMIX procedure for the LMM-Smooth approach. Restricted maximum likelihood estimation was used to fit both the LMM-Spherical and LMM-Smooth approaches. The placement of knots for the LMM-Smooth approach was based on a space filling design, which involves placing knots at locations within the study area such that the sum of the minimum distances from each knot to the locations of the original data are minimized (Ruppert *et al.*, 2003). We used the OPTEX procedure in SAS to determine exact knot placement. A total of 60 knots, which is intermediate to the minimum and maximum numbers of knots recommended by Ruppert *et al.* (2003), were used as part of the LMM-Smooth approach.

Model Evaluation

We evaluated the accuracy of the five modeling approaches for predicting stream temperatures by fitting the models to a random selection of 75% (Michigan: $n = 615$ and Wisconsin: $n = 233$) of the study stream sites and then calculating the accuracy in stream temperature predictions for the remaining (Michigan: $n = 205$ and Wisconsin: $n = 78$) study streams (validation dataset). Accuracy was determined by calculating the mean absolute error (MAE) and root mean square error (RMSE) in temperature predictions for the validation dataset (Power, 1993). We chose to use both metrics to assist in identifying variability in model errors for the modeling approaches. We repeated the process of randomly selecting streams, fitting the models, and evaluating accuracy of the temperature predictions 500 times. This method for validating models has been given several names, including repeated random subsampling validation and delete-d validation (Good, 2006), and is beneficial for ensuring that results are not affected by a single aberrant selection. We averaged the MAE and RMSE across all 500 model runs to evaluate overall accuracy of the different modeling approaches.

Prediction of Statewide July Mean Temperatures

To evaluate the utility of landscape-based models for predicting stream temperatures across large regions, we used the best performing modeling approach to predict statewide stream temperatures for Michigan and Wisconsin using a previously developed stream attribute database for the states (Brenden *et al.*, 2006). In fitting these models, we used the full set of stream temperatures measurements for the states. To determine how well these models predicted stream temperature, we regressed predicted *vs.* observed stream temperatures using simple linear regression and tested whether the slopes and intercepts for the regression models equaled 1 and 0, respectively. The results of such tests can gage whether the prediction model is consistent and unbiased (Smith and Rose, 1995; Piñeiro *et al.*, 2008). We also calculated Theil's partial inequality coefficients for the Michigan and Wisconsin stream temperature datasets, which separates total error of model predictions into three components: U_{bias}, U_{slope}, and U_{error} (Smith and Rose, 1995; Piñeiro *et al.*, 2008). U_{bias} represents the proportion of total error associated with mean differences between observed and predicted values. U_{slope} represents the proportion of total error associated with deviance of the slope from the 1:1 line. U_{error} represents the proportion of total error associated with the unexplained variance (Smith and Rose, 1995; Piñeiro *et al.*, 2008).

RESULTS

Comparison of Modeling Approaches

For Michigan streams, the mean MAE for the five modeling approaches ranged from 1.55°C to 2.07°C, while the mean RMSE for the approaches ranged from 2.00°C to 2.64°C (Figure 4). The similarity in RMSE and MAE for the modeling approaches suggested that each avoided large outliers in predicted temperatures (Figure 4). The most accurate approach for predicting stream temperatures was the LMM-Smooth approach. The mean MAE for this approach was 1.55°C, while the mean RMSE was 2.00°C. The second best performing modeling approach was the LMM-Spherical approach, although it performed only slightly worse than the LMM-Smooth approach. The mean MAE and RMSE for the LMM-Spherical approach equaled 2.03°C and 1.56°C, respectively. The next best performing modeling approaches were the MLR and GAM approaches. The mean MAE and

RMSE for the MLR approach equaled 1.66°C and 2.11°C, respectively; the mean MAE and RMSE for the GAM approach equaled 1.68°C and 2.15°C. The KRIG approach was the least accurate modeling approach. The mean MAE and RMSE for the KRIG approach was 2.07°C and 2.63°C, respectively.

For Wisconsin streams, the mean MAE for the modeling approaches ranged from 1.83°C to 2.36°C, while the mean RMSE ranged from 2.32°C to 2.95°C (Figure 5). As with Michigan, the similarity in RMSE and MAE for the modeling approaches suggested that each avoided large outliers in predicted temperatures (Figure 5). The performance of the modeling approaches for Wisconsin streams were similar to what was found for Michigan streams; the LMM-Smooth approach resulted in the most accurate stream temperature predictions (mean MAE = 1.83°C and mean RMSE = 2.32°C), but was followed closely by the LMM-Spherical (mean MAE = 1.89°C and mean RMSE = 2.40°C), MLR (mean MAE = 1.92°C and mean RMSE = 2.45°C), and GAM (mean MAE = 1.96°C and mean RMSE = 2.48°C) approaches. The
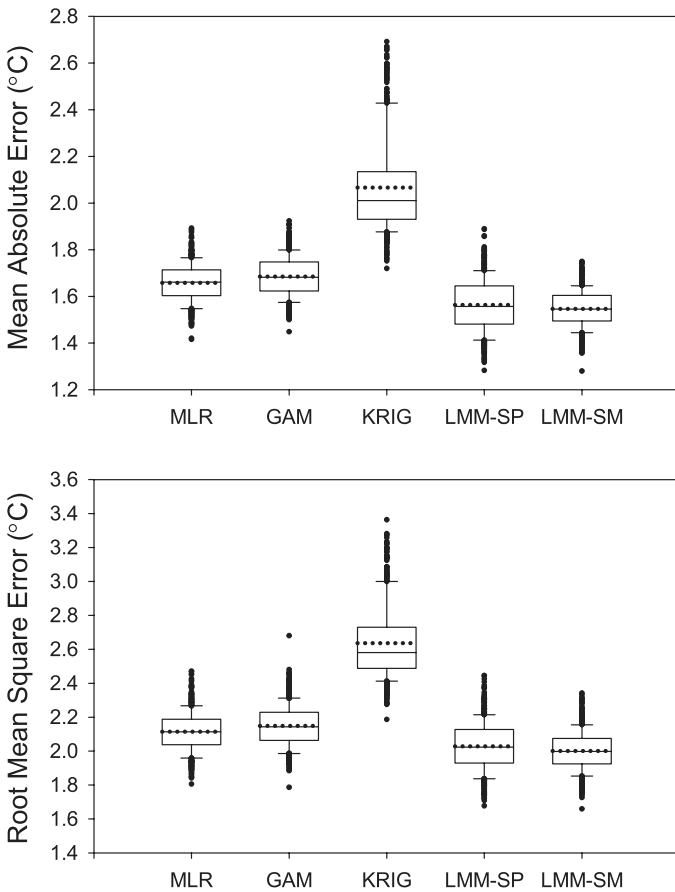

FIGURE 4. Prediction Accuracy Based on Mean Absolute Error and Root Mean Square Error for Different Modeling Approaches Used to Predict July Mean Water Temperature in Michigan Streams.
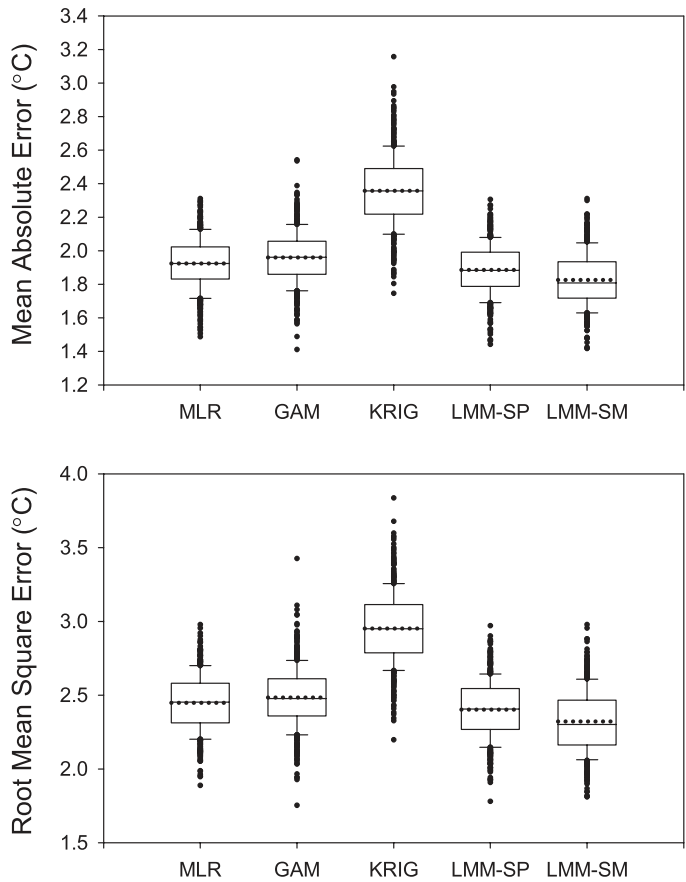

FIGURE 5. Prediction Accuracy Based on Mean Absolute Error and Root Mean Square Error for Different Modeling Approaches Used to Predict July Mean Water Temperature in Wisconsin Streams.

TABLE 2. Standardized Coefficients for Fixed Effects for the LMM-Smooth Models That Were Fit to the Full Set of Michigan and Wisconsin Stream Temperature Measurements.

	Michigan		Wisconsin	
Variable	Estimate	SE	Estimate	SE
AREA	**45.414**	2.198	**32.644**	2.493
FOREST	**−10.469**	3.117	−1.941	3.636
JULAIR	**24.265**	2.924	**17.741**	3.214
PERM	**−8.667**	3.024	**−9.377**	3.146
SLOPE	**−9.393**	2.719	−5.466	4.728
WATER	**9.766**	2.155	**8.057**	2.485

Notes: Coefficient estimates in bold font were significantly different from zero ($p < 0.0001$). [AREA = $\log_e$ network catchment surface area (km^2), FOREST = percent forest land use/type in local catchments, JULAIR = July mean air temperature (°C), PERM = mean soil permeability in network catchments (cm/100 h), SLOPE = mean network catchment slope (°), and WATER = percent water in network catchments.] LMM, linear mixed modeling.

TABLE 3. Regression Parameter Estimates and Significance Testing Results for the Regression of Observed *vs*. Predicted Stream Temperatures for the Michigan and Wisconsin Stream Temperature Datasets.

	Michigan	Wisconsin
Intercept		
Estimate	−0.38	−1.48
t	0.71	1.45
p	0.48	0.15
Slope		
Estimate	1.02	1.07
t	0.72	1.44
p	0.47	0.15

Note: The null hypotheses that were tested were intercept = 0 and slope = 1, the results of which can be used to evaluate prediction consistency and bias.

KRIG approach yielded the least accurate temperature predictions (MAE = 2.36°C and RMSE = 2.95°C).

Statewide Temperature Predictions

Based on the prediction accuracies for the different modeling approaches, we chose to use the LMM-Smooth approach to predict statewide stream temperatures for Michigan and Wisconsin. The fixed parameter coefficients for the LMM-Smooth models that were fit to the full set of Michigan and Wisconsin stream temperature datasets were similar in direction and magnitude (Table 2). For Michigan, the coefficient estimates for all variables used to model the mean function were significantly different from zero ($p < 0.004$). For Wisconsin, the coefficient estimates for percent forest land use in local catchments

and mean network catchment slope were not significantly different from zero ($p \geq 0.05$); the coefficient estimates for all other variables were significantly different from zero ($p < 0.003$). Predicted temperatures were slightly more accurate when all variables were included in the model. For this reason, we kept percent forest land use and catchment slope in the Wisconsin model.

The statewide July mean stream temperature predictions ranged from 5.2°C to 26.9°C for Michigan and from 7.6°C to 29.4°C for Wisconsin. The intercept and slope for the regression of observed *vs*. predicted stream temperatures were not significantly different from 0 and 1, respectively (Table 3). The intercept and slope for the regression of predicted *vs*. observed stream temperatures for the Wisconsin dataset also were not significantly different from 0 and 1 (Table 3). Thus, for both datasets, model predictions and observed values were linearly related and the models were unbiased. These findings were supported by the Theil partial inequality coefficients that were calculated for both datasets. In both cases, U_{error} was greater than 0.99, meaning that almost all of the predictive error resulted from unexplainable variance rather than from predictive error associated with model bias (U_{bias}) or deviation from the 1:1 line (U_{slope}). Both the Michigan ($R^2 = 63\%$) and Wisconsin ($R^2 = 60\%$) stream temperature models explained similar levels of variation in the observed stream temperature data (Figure 6).

DISCUSSION

The modeling approaches evaluated in this study varied considerably. Some methods incorporated spatial structure in stream temperatures as part of the mean function (MLR and GAM approaches) and error process (KRIG approach) separately, while others incorporated spatial structure in stream temperatures as part of both the mean function and error process (LMM-Spherical and LMM-Smooth approaches). With the exception of the KRIG approach, differences in performance among modeling approaches were relatively minor. The LMM-Spherical and LMM-Smooth approaches resulted in slight, albeit consistent, improvements in predictions for both states, suggesting that it may be somewhat advantageous to incorporate spatial structure in both the mean function and error process when trying to predict stream temperatures. However, the question remains whether additional model complexity is worth only marginal improvement in stream temperature predictions. The LMM-Smooth approach
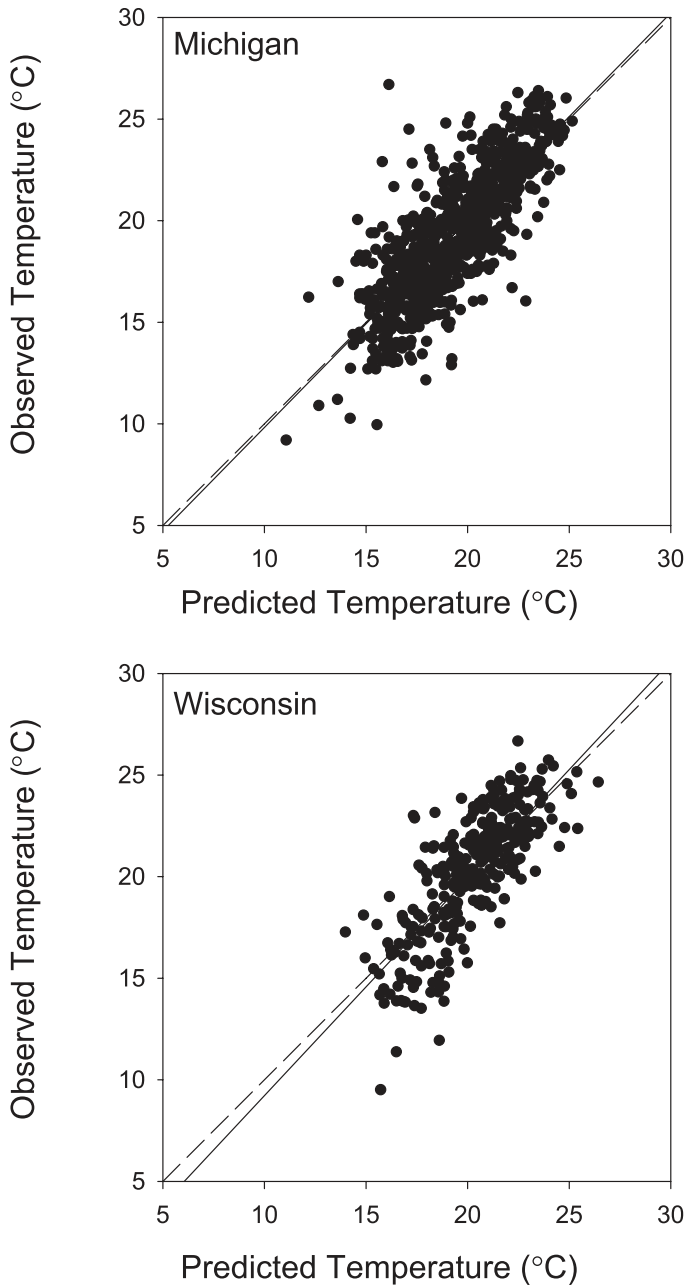
FIGURE 6. Scatterplots of Predicted *vs.* Observed July Mean Stream Temperatures for the LMM-Smooth Models Developed for Michigan and Wisconsin. The dashed line on the plots indicates 1:1 agreement between predicted and observed temperature. The solid line on the plots indicates the fitted values from a regression of observed *vs.* predicted temperatures.

performed slightly better than the LMM-Spherical approach, which likely stemmed from the spherical covariance structure not accurately accounting for the spatial dependency in temperatures among streams. The LMM-Smooth approach allows for the incorporation of spatial dependency in temperature among streams, without requiring explicit specification as to how streams are interrelated. The LMM-

Smooth has the additional advantage of being computationally simpler than the LMM-Spherical approach because fewer observations are used to model the spatial correlation among streams (Ruppert *et al.*, 2003).

Our finding that the KRIG approach resulted in the least accurate temperature predictions may have stemmed, in part, from our use of Euclidean distance for assessing spatial dependence among streams. Euclidean distance is based on the shortest distance between locations and, as a result, does not account for the spatial variability in temperature due to flow path and position in the stream network. Gardner *et al.* (2003) found that using network distance and position resulted in more accurate KRIG predictions than using Euclidean distance in a study of the Beaverkill watershed in New York. The lack of adequate data coverage in our datasets precluded the use of network path distances for the KRIG approach. In some cases temperature would have been interpolated from fewer than 10 measurements for watersheds with more than 1,000 confluence-to-confluence stream reaches at a spatial scale of 1:100,000. With more observations, it is likely that the accuracy of the stream temperature predictions from the KRIG would have improved. It remains uncertain, however, whether using KRIG alone would be more beneficial than accounting for spatial variability as part of both the mean function and error. The ability to use model covariates such as large ground-water input at particular locations may result in more accurate temperature predictions even when there is sufficient data to use network distances as the KRIG basis.

Our finding that the LMM-Smooth and LMM-Spherical approaches consistently resulted in the most accurate stream temperature predictions suggests that there may be important landscape or climate variables that were not included in the prediction model (Peterson *et al.*, 2007). These variables could be related to regional patterns in land use, surficial geology, ground-water loading, or air temperature variability. Identifying additional variables that would improve model prediction likely would be a very time consuming process. As a result, accounting for spatial autocorrelations in stream temperatures may be an efficient alternative to searching among a large number of stream attributes. When attempting to model stream temperatures based on regional landscape variables, we encourage analysts to consider incorporating spatial dependence among streams as a way of improving temperature predictions regardless of whether dependency is modeled using Euclidean or network path distances.

Although LMM using low-rank thin-plate smoothing splines to measure the spatial autocorrelation in stream temperatures was the most accurate modeling
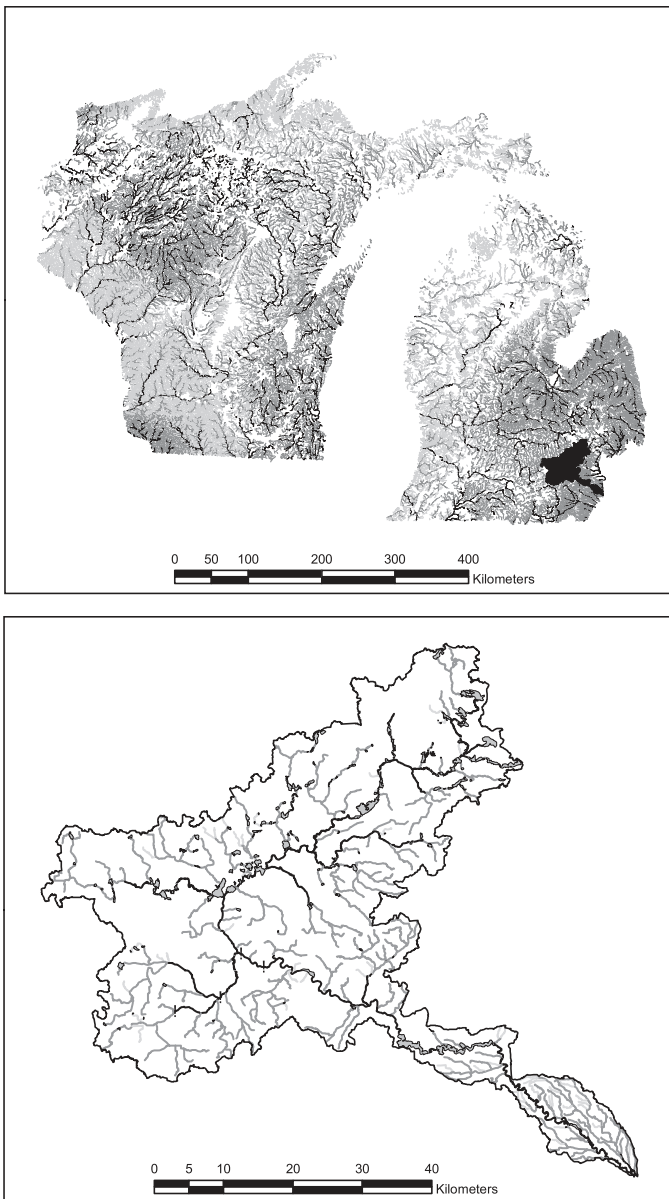
FIGURE 7. Predicted Distribution of Cold Water (light gray), Cool Water (dark gray), and Warm Water (black lines) Streams Across Michigan and Wisconsin (upper panel) and in the Huron River Watershed, Michigan (lower panel; shown in black in upper panel).

approach; overall, there were only slight differences in prediction accuracy among the evaluated approaches. One implication of this is that it permits substantial flexibility in choosing among modeling approaches without sacrificing a substantial amount of accuracy. Thus, researchers and managers can choose a method with which they are already experienced. A second implication is that in order to develop more accurate stream temperature predictions, it will likely require finer-scale hydrologic (e.g., volume of water in the channel and travel time) and physiographic features (e.g., riparian shading, location and volume of

springs, and substrate type) (Risley *et al.*, 2003; Johnson, 2004). Because obtaining measurements of these finer-scale factors will require substantially more effort, researchers and managers will need to consider the tradeoffs between increased model accuracy and the costs associated with obtaining this finer-scale information.

The scale at which conservation planning and resource assessment can be carried out is ultimately limited by data availability. Statistical temperature models based on landscape features provide a cost-effective means to generate critical habitat data at sites across large regions, such as several watersheds or even across an entire state. Although our temperature predictions are relatively coarse, they do provide the best available information to policy makers and managers who often must develop and implement strategies in the absence of field observations (Sowa *et al.*, 2007). In this study, the number of inter-confluence reaches that had temperature measures represented less than 2% of the total number of reaches across Wisconsin and Michigan, and in the Huron River watershed, MI, measured temperature data were available at only 5 out of 547 inter-confluence reaches (Figure 7). By linking temperature predictions to inter-confluence stream reaches, we were able to predict water temperature for the majority of stream reaches across Wisconsin and Michigan (Figure 7). The ability to identify and assess stream ecosystems across this region is greatly enhanced by having comprehensive, albeit coarse, temperature information.

LITERATURE CITED

Abell, R. and J.D. Allan, 2002. Riparian Shade and Stream Temperatures in and Agricultural Catchment, Michigan, USA. Proceedings of the International Association of Theoretical and Applied Limnology 28:232-237.

Allan, J.D. and M.M. Castillo, 2007. Stream Ecology: Structure and Function of Running Waters (Second Edition). Springer, Dordrecht, the Netherlands.

Brenden, T.O., R.D. Clark, Jr, A.R. Cooper, P.W. Seelbach, L. Wang, S.S. Aichele, E.G. Bissell, and J.S. Stewart, 2006. A GIS Framework for Collecting, Managing, and Analyzing Multi-Scale Landscape Variables Across Large Regions for River Conservation and Management. *In:* Influences of Landscape on Stream Habitats and Biological Assemblages, R.M. Hughes, L. Wang, and P.W. Seelbach (Editors). American Fisheries Society Special Publication 48, Bethesda, Maryland, pp. 49-73.

Brenden, T., L. Wang, and P.W. Seelbach, 2008. A Landscape-Based River Classification System for Michigan Rivers and Streams for Fisheries and Environmental Management. Transactions of American Fisheries Society 137:1621-1636.

Brett, J.R., 1971. Energetic Responses of Salmon to Temperature. A Study of Some Thermal Relations in the Physiology and Freshwater Ecology of Sockeye Salmon (Oncorhynchus Nerka). American Zoologist 11:99-113.

Caissie, D., 2006. The Thermal Regime of Rivers: A Review. Freshwater Biology 51:1389-1406.

Caissie, D., M.G. Satish, and N. El-Jabi, 2006. Predicting River Water Temperatures Using the Equilibrium Temperature Concept With Application on Miramichi River Catchments (New Brunswick, Canada). Hydrological Processes 19:2137-2159.

Caissie, D., M.G. Satish, and N. El-Jabi, 2007. Predicting Water Temperatures Using a Deterministic Model: Application on Miramichi River Catchments (New Brunswick, Canada). Journal of Hydrology 336:303-315.

Crisp, D.T. and G. Howson, 1982. Effect of Air Temperature Upon Mean Water Temperature in Streams in the North Pennines and English Lake District. Freshwater Biology 12:359-367.

Eaton, J.G. and R.M. Scheller, 1996. Effects of Climate Warming on Fish Thermal Habitat in Streams of the United States. Limnology and Oceanography 41:1109-1115.

Gardner, B. and P.J. Sullivan, 2004. Spatial and Temporal Stream Temperature Prediction: Modeling Nonstationary Temporal Covariance Structures. Water Resources Research 40:1-9.

Gardner, B., P.J. Sullivan, and A.J. Lembo, Jr., 2003. Predicting Stream Temperatures: Geostatisical Model Comparison Using Alternative Distance Metrics. Canadian Journal of Fisheries and Aquatic Science 60:344-351.

Gillooly, J.F., E.L. Charnov, G.B. West, V.M. Savage, and J.H. Brown, 2002. Effects of Size and Temperature on Developmental Time. Nature 417:70-73.

Giustolisi, O., A. Doglioni, D.A. Savic, and B.W. Webb, 2007. A Multi-Model Approach to Analysis of Environmental Phenomena. Environmental Modelling and Software 22:674-682.

Good, P.I., 2006. Resampling Methods (Third Edition). Birkhäuser, Boston, Massachusetts.

Gu, R.R., S. Montgomery, and T. Austin, 1998. Quantifying the Effects of Stream Discharge on Summer River Temperature. Hydrological Science Journal 43:885-894.

Hastie, T. and R. Tibshirani, 1990. Generalized Additive Models. Chapman and Hall, London, United Kingdom.

Hawkins, C.P., J.N. Hogue, L.M. Decker, and J.W. Feminella, 1997. Channel Morphology, Water Temperature, and Assemblage Structure of Stream Insects. Journal of the North American Benthological Society 16:728-749.

Hokanson, K.E.F., C.E. Kleines, and T.W. Thorslund, 1977. Effects of Constant Temperatures and Diel Temperature Fluctuations on Specific Growth and Mortality Rates of and Yield of Juvenile Rainbow Trout, *Salmo Gairdneri*. Journal of the Fisheries Research Board of Canada 34:639-648.

Horne, B.D., E.S. Rutherford, and K.E. Wehrly, 2004. Simulating Effects of Hydro-Dam Alteration on Thermal Regime and Wild Steelhead Recruitment in a Stable-Flow Lake Michigan Tributary. River Research and Applications 20:185-203.

Isaak, D.J. and W.A. Hubert, 2001. A Hypothesis About Factors That Affect Maximum Summer Stream Temperatures Across Montane Landscapes. Journal of the American Water Resources Association 37:351-366.

Johnson, S.L., 2004. Factors Influencing Stream Temperatures in Small Streams: Substrate Effects and a Shading Experiment. Canadian Journal of Fisheries and Aquatic Science 61:913-923.

Karacor, A.G., N. Sivri, and O.N. Ucan, 2007. Maximum Stream Temperature Estimation of Degirmendere River Using Artificial Neural Network. Journal of Scientific and Industrial Research 66:363-366.

Kim, K.S. and S.C. Chapra, 1997. Temperature Model for Highly Transient Shallow Streams. ASCE, Journal of Hydraulic Engineering 123:30-40.

Lyons, J.L., 1996. Patterns in the Species Composition of Fish Assemblages Among Wisconsin Streams. Environmental Biology of Fishes 45:329-341.

Mackey, A.P. and A.D. Berrie, 1991. The Prediction of Water Temperatures in Chalk Streams From Air Temperature. Hydrobiologia, 210:183-189.

MCGI (Michigan Center for Geographic Information), 2004. Land Cover 2001 Geographic Theme: Land Cover/Use. http://www.mcgi.state.mi.us/mgdl/?rel=thext&action=thmname&cid=5&cat=Land+Cover+2001, *accessed* August 2004.

McKenna, J.E., Jr, R.P. McDonald, C. Castiglione, S. Morrison, K. Kowalski, and D. Passino-Reader, 2006. A Process to Model Habitats and Fish-Environment Relationships in Support of Aquatic GAP Analysis in the Great Lakes. *In:* Influences of Landscape on Stream Habitats and Biological Assemblages, R.M. Hughes, L. Wang, and P.W. Seelbach (Editors). American Fisheries Society Special Publication 48, Bethesda, Maryland, pp. 555-575.

Mohseni, A., H.G. Stefan, and T.R. Erickson, 1998. A Nonlinear Regression Model for Weekly Stream Temperatures. Water Resources Research 34:2685-2692.

Nelitz, M.A., E.A. MacIsaac, and R.M. Peterman, 2007. A Science-Based Approach for Identifying Temperature-Sensitive Streams for Rainbow Trout. North American Journal of Fisheries Management 27:405-424.

OSU/SCAS (Oregon State University/Spatial Climate Analysis Service), 2004. Mean, Monthly, and Annual Maximum, Minimum, and Mean Temperature: Conterminous United States. http://www.climatesource.com/us/fact_sheets/fact_tmean_us.html, *accessed* August 2004.

Peterson, E.E., D.M. Theobald, and J.M. Ver Hoef, 2007. Geostatistical Modeling on Stream Networks: Developing Valid Covariance Matrices Based on Hydrologic Distance and Stream Flow. Freshwater Biology 52:267-279.

Piñeiro, G., S. Perelman, J.P. Guerschman, and J.M. Paruelo, 2008. How to Evaluate Models: Observed vs. Predicted or Predicted vs. Observed? Ecological Modelling 216:316-322.

Power, M., 1993. The Predictive Validation of Ecological and Environmental Models. Ecological Modelling 68:33-50.

Rahel, F.J. and W.A. Hubert, 1991. Fish Assemblages and Habitat Gradients in a Rocky Mountain-Great Plains Stream: Biotic Zonation and Additive Patterns of Community Change. Transactions of the American Fisheries Society 120:319-332.

Rahel, F.J., C.J. Keleher, and J.L. Anderson, 1996. Potential Habitat Loss and Population Fragmentation for Cold Water Fish in the North Platte River Drainage of the Rocky Mountains: Response to Climate Warming. Limnology and Oceanography 41:1116-1123.

Risley, J.C., E.A. Roehl, and P.A. Conrads, 2003. Estimating Water Temperatures in Small Streams in Western Oregon Using Neu-

ral Network Models. U.S. Geological Survey, Water-Resources Investigations Report 02-4218, Portland, Oregon.

Roehl, E., J. Risley, J. Stewart, and M. Mitro, 2006. Numerically Optimized Empirical Modeling of Highly Dynamic, Spatially Expansive, and Behaviorally Heterogeneous Hydrologic Systems – Part 1. *In:* Proceedings for the Environmental Modeling and Software Society Conference, Burlington, Vermont.

Ruppert, D., M.P. Wand, and R.J. Carroll, 2003. Semiparametric Regression. Cambridge University Press, Cambridge, United Kingdom.

SAS Institute, 2004. SAS/STAT 9.1 User's Guide. SAS Institute Inc., Cary, North Carolina.

Schabenberger, O. and C.A. Gotway, 2005. Statistical Methods for Spatial Data Analysis. Chapman & Hall, Boca Raton, Florida.

Scott, M.C., G.S. Helfman, M.E. McTammany, E.F. Benfield, and P.V. Bolstad, 2002. Multiscale Influences on Physical and Chemical Stream Conditions Across Blue Ridge Landscapes. Journal of the American Water Resources Association 38:1379-1392.

Seelbach, P.W., M.J. Wiley, M.E. Baker, and K.E. Wehrly, 2006. Initial Classification of River Valley Segments Across Michigan's Lower Peninsula. *In:* Influences of Landscape on Stream Habitats and Biological Assemblages, R.M. Hughes, L. Wang, and P.W. Seelbach (Editors). American Fisheries Society Special Publication 48, Bethesda, Maryland, pp. 25-47.

Sinokrot, B.A. and H.G. Stefan, 1993. Stream Temperature Dynamics: Measurement and Modeling. Water Resources Research 29:2299-2312.

Sivri, N., N. Killie, and O.N. Ucan, 2007. Estimation of Stream Temperature in Firtina Creek (Rize-Turkiye) Using Artificial Neural Network Model. Journal of Environmental Biology 28:67-72.

Smith, E.P. and K.A. Rose, 1995. Model Goodness-of-fit Analysis Using Regression and Related Techniques. Ecological Modeling 77:49-64.

Sowa, S.P., G. Annis, M.E. Morey, and D.D. Diamond, 2007. A GAP Analysis and Comprehensive Conservation Strategy for Riverine Ecosystems of Missouri. Ecological Monographs 77:301-334.

Sponseller, R.A., E.F. Benfield, and H.M. Valett, 2001. Relationships Between Land Use, Spatial Scale and Stream Macroinvertebrate Communities. Freshwater Biology 46:1409-1424.

Steel, E.A., and I.A. Lange, 2007. Using Wavelet Analysis to Detect Changes in Water Temperature Regimes at Multiple Spatial Scales: Effects of Multi-Purpose Dams in the Willamette River Basin. River Research and Applications 23:351-359.

Steen, P.J., T.G. Zorn, P.W. Seelbach, and J.S. Schaeffer, 2008. Classification Tree Models for Predicting Distributions of Michigan Stream Fish From Landscape Variables. Transactions of the American Fisheries Society 137:976-996.

Stephan, H.G. and E.B. Preud'homme, 1993. Stream Temperature Estimation From Air Temperature. Water Resources Research 29:27-45.

Stewart, J., M. Mitro, E.A. Roehl Jr., and J. Risley, 2006. Numerically Optimized Empirical Modeling of Highly Dynamic, Spatially Expansive, and Behaviorally Heterogeneous Hydrologic Systems – Part 2. *In*: Proceedings of the 7th International Conference on Hydroinformatics, Nice, France.

Taniguchi, Y., F.J. Rahel, D.C. Novinger, and K.G. Gerow, 1998. Temperature Mediation of Competitive Interactions Among Three Fish Species That Replace Each Other Along Longitudinal Stream Gradients. Canadian Journal of Fisheries and Aquatic Sciences 55:1894-1901.

USGS (U.S. Geological Survey), 2004a. National Elevation Dataset. http://ned.usgs.gov/, *accessed* August 2004.

USGS (U.S. Geological Survey), 2004b. Soils Data for the Conterminous United States Derived From the NRCS State Soil Geo-graphic (STATSGO) Data Base. http://water.usgs.gov/GIS/metadata/usgswrd/XML/ussoils.xml, *accessed* August 2004.

Vannote, R.L. and B.W. Sweeney, 1980. Geographic Analysis of Thermal Equilibria: A Conceptual Model for Evaluating the Effect of Natural and Modified Thermal Regimes on Aquatic Insect Communities. American Naturalist 115:667-695.

WDNR (Wisconsin Department of Natural Resources), 2004. Land-cover Dataset (WISCLAND). http://www.dnr.state.wi.us/maps/gis/datalandcover.html, *accessed* August 2004.

Webb, B.W., D.M. Hannah, R.D. Moore, L.E. Brown, and F. Nobilis, 2008. Recent Advances in Stream and River Temperature Research. Hydrological Processes 22:202-918.

Wehrly, K.E., M.J. Wiley, and P.W. Seelbach, 2003. Classifying Regional Variation in Thermal Regime Based on Stream Fish Community Patterns. Transactions of the American Fisheries Society 132:18-38.

Wehrly, K.E., M.J. Wiley, and P.W. Seelbach, 2006. Influence of Landscape Features on Summer Water Temperatures in Lower Michigan Streams. *In:* Influences of Landscape on Stream Habitats and Biological Assemblages, R.M. Hughes, L. Wang, and P.W. Seelbach (Editors). American Fisheries Society Special Publication 48, Bethesda, Maryland, pp. 113-127.

Younus, M., M. Hondzo, and M. Newton, 2000. Stream Temperature Dynamics in Upland Agricultural Watersheds. ASCE, Journal of Environmental Engineering 126:518-526.