Spatial Statistical Network Models for Stream and River Temperature in New England, USA

Naomi E. Detenbeck and Alisa Morrison (ORISE participant), U.S. Environmental Protection Agency, Atlantic Ecology Division, Narragansett, RI

Ralph W. Abele, U.S. Environmental Protection Agency, Region 1, Boston, MA

Darin Kopp (former student services contractor), U.S. Environmental Protection Agency, Atlantic Ecology Division, Narragansett, RI[1]

Corresponding Author: N.E. Detenbeck, U.S. EPA Atlantic Ecology Division, 27 Tarzwell Drive, Narragansett, RI, 02882. (detenbeck.naomi@epa.gov)

[1]Current address: Wrigley Institute, Arizona State University, Tempe, AZ

Key points:

- Lotic thermal regimes for May – September can be described based on the July/August median, diurnal change, and growing season maximum

- New England stream temperature models show evidence of strong spatial autocorrelation along flow pathways and between points

- Spatial statistical network models predict NE summer monthly median stream temperatures with a RMS prediction error of 1.4 – 1.5 $^{\circ}$C.

Abstract Watershed managers are challenged by the need for predictive temperature models with sufficient accuracy and geographic breadth for practical use.   We described thermal regimes of New England rivers and streams based on a reduced set of metrics for the May to September growing season (July or August median temperature, diurnal rate of change, and magnitude and timing of growing season maximum) chosen through principal component analysis of 78 candidate metrics. We then developed and assessed spatial statistical models for each of these metrics, incorporating spatial autocorrelation based on both distance along the flow network and Euclidean distance between points. Calculation of spatial autocorrelation based on travel or retention time in place of network distance yielded tighter-fitting Torgegrams with less scatter but did not improve overall model prediction accuracy. We predicted monthly median July or August stream temperatures as a function of median air temperature, estimated urban heat island effect, shaded solar radiation, main channel slope, watershed storage (percent lake and wetland area), percent coarse-grained surficial deposits, and presence or maximum depth of a lake immediately upstream, with an overall root-mean-square prediction error of 1.4 and 1.5° C, respectively. Growing season maximum water temperature varied as a function of air temperature, local channel slope, shaded August solar radiation, imperviousness, and watershed storage. Predictive models for July or August daily range, maximum daily rate of change, and timing of growing season maximum were statistically significant but explained a much lower proportion of variance than the above models (5-14% of total) .

Index terms: 1878 Water/energy interactions; 0438 Diel, seasonal, and annual cycles; 0466 Modeling; 1816 Estimation and forecasting; 1819 Geographic Information Systems (GIS)

Key words: stream, river, temperature regime, spatial statistical model, New England

1. Introduction

Regional assessments of thermal impacts to aquatic resources in response to climate change, land-use change, and water resource management are a critical component of adaptation strategies. Description of stream thermal regimes requires assessment of not only the magnitude of temperature but also frequency and duration of temperature extremes. Numerous metrics have been proposed to describe stream/river thermal regimes [*Jones and Schmidt*, 2013], and researchers are still struggling with the need to describe thermal regimes in a parsimonious fashion [*Olden and Naiman*, 2010], analogous to characterization of flow regimes [*Poff and Ward*, 1989].

Regional temperature models are needed for characterizing and mapping current stream thermal regimes [*Maheu et al.,* 2015], establishing reference condition and aquatic life use categories [*Hill et al*., 2013], assessing and prioritizing past impacts and remediation strategies [*Poole and Berman*, 2000], and predicting future impacts and identifying critical thermal refugia [*Isaak et al.*, 2010; Wiley *et al*., 2010]. In the northeastern United States (New England), states and Federal agencies have a variety of needs for comprehensive temperature data and predictive modelling results, including assessment of thermal regime classes as the basis for implementing temperature criteria, establishment or refinement of existing temperature criteria, protection of coldwater fisheries habitat, setting priorities for riparian restoration and protection, trend detection, and regulating discharges from power plants and reservoirs [*North Atlantic Landscape Conservation Cooperative*, 2012]. Early attempts to describe stream and river thermal regimes across New England were hampered by the availability of stream temperature monitoring data and, as a result, Olivero and Anderson [2008] used known temperature preferences of aquatic communities to develop predictive models of stream and river temperature classes.

Historically, when developing predictive temperature models for streams, there has been a tradeoff between accuracy of model predictions and practical spatial extent of model coverage. Mechanistically-

3

based heat budget models such as SNTEMP [*Krause et al.,* 2004] can predict stream temperature within a few tenths of a degree. However, the only mechanistic model that has been linked with a GIS interface to facilitate regional application is BASIN TEMP [*Allen* 2008], which is not commercially available. One intermediate solution could be the application of WET-Temp [*Cox and Bolte,* 2007], a spatially explicit network-based model for continuous temperature simulation. However, combined preprocessing and run times for an entire region would be prohibitive. LeBlanc et al. [1997] identified a second intermediate solution using a simulation model of the effects of urbanization on water temperature in unregulated streams. They determined that model outputs were sensitive to only four of the model inputs: vegetation transmissivity, channel width, sun angle, and groundwater discharge, thus paving the way to development of a much simpler predictive model. This approach has only been applied at the reach scale, however, and needs to be incorporated into a network model to allow examination of cumulative effects on temperature throughout a watershed.

Statistical models provide an alternative approach to mechanistic models and are generally easier to implement at a regional scale. Statistical models to predict thermal regimes generally are developed with one of two goals: 1) prediction of reference thermal regimes (e.g., fish thermal guild) [*Wehrly et al.,* 2003; *Moore,* 2006; *Isaak et al.,* 2010; *Hill et al.,* 2013, 2014; *Moore et al.,* 2013; *DeWeber and Wagner,* 2014; *Hilderbrand et al.,* 2014; *Maheu et al.,* 2015] or 2) prediction of thermal sensitivity to determine potential response to climate change-related increases in air temperature [*Kelleher et al.,* 2012; *Kanno et al.,* 2014; *Johnson et al.,* 2014; *Gu et al.,* 2015; *Segura et al.,* 2015]. Here we focus solely on models developed to predict thermal regimes.

Most regional regression models of stream temperature means or maxima using GIS-derived watershed attributes as independent variables have prediction errors of 2-3°C at best [*Wehrly et al.,* 2009], although Gallice et al. [2015] recently achieved a root mean square error (RMSE) of 1.3 °C for

4

monthly mean stream temperatures using a physics-derived statistical model. Data mining techniques

such as neural network analysis yield models that can explain daily mean temperatures with an RMSE of

0.9 – 1.5 ° C for individual streams [*Sahoo et al.,* 2009]. Due to data limitations, these models have only

been applied at the regional scale to predict stream temperature classes with thresholds differing by 2°

C, not a continuous temperature distribution [*McKenna et al.*, 2010]. (McKenna et al. based their

predictions on a large data set of instantaneous temperature measurements rather statistical

summaries of continuous time series.) Random forest models or boosted regression trees typically

perform as well or better than neural network models and should be applicable to temperature

predictions [*Hastie et al.*, 2009]. More recently, spatial statistical models have been developed to

improve regression modeling techniques by taking into account the spatial covariance structure

inherent in stream networks [*Ver Hoef and Petersen*, 2010; *Laaha et al.*, 2013]. Using Ver Hoef and

Petersen's approach, Isaak et al. [2010] have reported r^2 values of greater than 0.90 for spatial statistical

model predictions of mean or maximum stream temperatures in the Pacific Northwest, with root mean

square error (RMSE) of approximately 1° C. Unlike the earlier approaches, spatial statistical models

developed by Ver Hoef and Petersen [2010] describe spatial autocorrelation as part of the error

structure based on distance along the flow network as well as standard Euclidean distances between

observation points. The related topological kriging approach differs in that Laaha et al. [2013] describe

spatial autocorrelation as a function of shared source (watershed) area rather than based on stream

distance and do not distinguish between spatial autocorrelation in the upstream versus downstream

direction.

Given the multiple needs for a comprehensive description of thermal regimes and an understanding

of the factors affecting stream temperature sensitivity to the combined stressors of development and

interannual weather variability, our goal for this study was to develop predictive models for a

parsimonious set of metrics describing the thermal regime of streams/rivers across New England. Our

5

study objectives were to 1) develop a stream temperature monitoring database, 2) choose a suite of temperature metrics to evaluate, and 3) develop a regional temperature prediction model for New England based on Ver Hoef and Petersen's [2010] Spatial Statistical Network (SSN) approach. To accomplish these objectives, we developed the Thermal Database of New England using existing stream/river temperature time series data from state, federal, and nongovernmental organization sources. After filtering the database by applying quality assurance criteria, we derived a reduced set of metrics to describe independent aspects of the thermal regime. Finally, we developed a suite of spatial statistical temperature models to predict these metrics based on air temperature and watershed attributes. We also evaluated approaches to enhance the predictive power of Ver Hoef and Petersen's [2010] modeling approach by incorporating effects of watershed storage and retention time and water body morphometry effects on heat exchange.

2. Methods

2.1. Study Area

The study area is comprised of the New England states of Maine, New Hampshire, Vermont, Massachusetts, Connecticut, and Rhode Island in the northeastern United States (Figure 1). Stream temperature data were compiled from across the region for the years 1995 – 2010. After quality assurance checks were completed, 539 station-years of data were available for stations reporting daily water temperature statistics only and 448 station-years of data were available for stations reporting water temperatures at 15-minute to hourly intervals. Data availability varied across the region (Table 1).

We expect the thermal characteristics of streams and rivers across New England to vary between coldwater (<18.45° C), coolwater (18.45 – 22.30° C), and warmwater (> 23.0 $^\circ$ C mean July temperature) regimes [*Beauchene et al.*, 2014] based on gradients in air temperature, solar radiation, localized

6

topographic or riparian shading, ice and/or snow cover in winter, variation in surface water retention

time, groundwater inputs, impervious cover (IC), and anthropogenic water inputs and withdrawals.

Annual average air temperature generally decreases from south to north and inland from the coast,

from about 11° C in southeastern CT to 2° C in northwestern Maine, with the exception of a narrow

climatic zone along the coast with moderate temperatures. July average maximum air temperatures

show a similar trend, ranging from 23° C in southeastern CT to 17° C in northwestern Maine (climate

normals 1981 – 2010) [*Northeast Regional Climate Center*, 2016]. Between the periods of 1971-2000

and 1981-2010, climate normals have shifted over much of the region, with an increase of up to 0.33° C

in both northern New England and Cape Cod and lesser changes elsewhere along the coast. Degree of

cloudiness and unshaded solar radiation [*National Renewable Energy Laboratory*, 2015] vary little across

the region, by only 10-12 percent of possible sunshine.

Differences among hydrophysiographic regions in surficial deposits and slope can influence

groundwater recharge and discharge to streams, moderating and stabilizing water temperatures.

Hydrophysiographic regions in the Northeast [*Randall*, 2001] (Figure S1) generally follow Fenneman's

[1938] physiographic boundaries. Hydrophysiographic regions are formed by geologic and glacial

processes, causing differences in distribution of coarse-grained stratified drift relative to fine-grained

stratified drift, till, bedrock and the location of streams. Southern New England is characterized by low

to moderate relief with small, closely spaced valleys containing abundant coarse-grained stratified drift.

In a few low relief areas with sandy deltaic or fluvial outwash, there are extensive outwash-plain

aquifers. In contrast, northern and northeastern Maine, also with low to moderate relief, have relatively

little coarse-grained stratified drift, although there are numerous eskers across Maine, because the

source of meltwater for this region was a nearly stagnant residual ice cap.

7

Stream temperature extremes are particularly susceptible to low flow events. Low flows in central New England are influenced not only by the distribution of coarse surficial sediment, but also by water availability. In areas of high relief of central New England, low flow statistics are positively correlated with percent of the basin underlain by sand and gravel deposits and elevation (influencing water availability), with a smaller correlation with till and surficial fine-grained stratified drift, and negative correlation with wetland and lake area which reduces base flow due to evapotranspiration. In areas of low relief, influences on low flow statistics are similar, with the exception that mean runoff per unit area better predicts water availability than elevation [*Wandle and Randall*, 1994; *Archfield et al.*, 2010; *Morrison et al.*, 2016].

2.2 Thermal Metric Development and Selection to Describe Thermal Regimes

Temperature time series were obtained from a variety of state and Federal data sources (Table 1). Quality assurance and pre-processing steps are described in detail in Supplemental Text S1. We calculated 78 thermal metrics from raw temperature time series (Table S1). Most of the thermal metrics were based on those included in the ThermoStats version 2 program from the Ontario Ministry of the Environment [*Jones and Schmidt*, 2013], but calculated using SAS (©SAS Institute, Carey, NC) since ThermoStat version 2 does not allow batch processing. Candidate metrics included indicators of magnitude aggregated at monthly and growing season time scales (average, minimum and maximum), running averages, monthly duration curves, diurnal rate of change, timing of growing season maxima, and fraction of daily average temperatures suitable for coldwater or warmwater fish (Table S1). Some of these variables could only be generated for detailed time series data recorded at 15 minute to 1 hour time intervals. Some data sources only provide access to daily statistics (mean, minimum and maximum stream temperature values) so that metrics such as maximum daily rate of change cannot be calculated. In these cases, we assessed daily analogues (e.g., daily range). Prior to calculating metrics we eliminated

8

days with less than 95% complete readings and months with less than 95% of daily values. We filled the

resulting gaps using the SAS PROC EXPAND to minimize biases related to a potential uneven distribution

of missing values over the diurnal cycle. PROC EXPAND does a linear interpolation between

observations to fill in gaps to produce a regular time series before calculating temporal averages.

To reduce the number of candidate variables to a set of relatively independent variables capturing

most of the variation across sites, we applied principal components analysis to the suite of 78 thermal

metrics using PROC PRINCOMP in SAS and then selected individual metrics that were highly correlated

with the major axes of variation. We also applied PROC FACTOR with varimax rotation for comparison.

Prior to conducting PCAs, we normalized thermal metrics applying Box-Cox transformations in SAS, with

the exception of the fraction of days suitable for coldwater or warmwater fish, to which we applied an

arc-sine square root transformation.   Most data sources did not reflect year-round data collection, so

we focused initial analyses on the growing season (May to September). We performed separate

analyses with the data set containing daily statistics including "complete" data records for June through

September (n = 98 sets), and again with the fine resolution data set containing detailed time series with

recording intervals of less than or equal to one hour for June through September (n = 105 sets). To

assess the potential for bias in daily statistics related to missing values, we calculated some metrics after

filling gaps in time series using the PROC EXPAND function in SAS to generate complete time series at 15

minute intervals. To ensure that PCA results were robust, for both the daily statistics and detailed time

series data sets, we repeated the PCAs on subsets of the variables that were complete for a larger

number of records (up to 539 station-years for daily statistics and up to 448 station-years for detailed

data sets). In subsequent analyses for predictive modeling we used less restricted data sets for all

metrics not related to growing season maxima, which required a longer period of record. In making final

selections of thermal metrics we also considered robustness of measures (i.e., favoring median values

9

over averages because they are less likely to be skewed by outliers) and biological relevance (i.e., timing of periods of greatest stress to fish communities).

2.3. Derivation of Waterbody, Watershed and Meteorological Attributes

2.3.1. Watershed Attributes

We delineated watersheds associated with stream temperature monitoring stations using the Basin Delineator tool associated with NHDPlusV1 [*Horizon Systems, Incorporated*, 2010] for stream monitoring stations that were located on NHDPlusV1 stream reaches. For monitoring stations located on smaller reaches not included in the NHDV1 dataset or for points that failed to process using Basin Delineator, we used ArcHydro version 1.4 tools in conjunction with ESRI's ArcMap 9.3 geographic information system (©ESRI, Redland, CA). Using the ArcMap FLOWLENGTH command, we defined 120 meter stream buffers based on distance along flow paths and used these to calculate buffer zone attributes. We acquired several NHDPlus reach-scale attributes from tables included in the NHDPlus datasets, i.e., estimated mean discharge based on unit runoff models (MAFLOWU), local reach gradient (GRADIENT), area-weighted mean annual air temperature (AREAWTMAT), and reach velocity (MAVELU). Using USGS protocols available within ArcHydro tools, we calculated main channel slope between the 10th and 85th percentile of main channel. We summarized watershed attributes from watershed boundaries and grid or polygon layers acquired from state or Federal data sources (Table S2) using ArcMap 9.3 and/or ArcMap addons: Hawth's Tools version 3.27 [*Beyer*, 2004] or Geospatial Modeling Environment tools [*Beyer*, 2012]. Using a grid mosaic of either best available surficial geology coverages to estimate coarse surficial deposits or of Soil Survey Geographic Database (SSURGO) soils attributes (hydrologic soil group and drainage class) [*USDA Natural Resources Conservation Service*, 2012], we calculated three alternative indicators of groundwater inputs to streams. We extracted SSURGO attributes from individual databases using the Soil Data Viewer tool [*USDA Natural Resources Conservation Service*,

10

2011] and filled in minor gaps in coverage using State Soil Geographic (STATSGO) data [*Schwartz and Alexander,* 1995].

We evaluated the potential influence of urban heat islands on stream temperatures. Natural surfaces tend to use a large proportion of absorbed radiation in the evapotranspiration process. In contrast, urban surfaces tend to absorb radiation and limit the escape of reflected radiation due to the "urban canyon effect". In addition, urban environments produce waste heat from air conditioning, refrigeration, industrial processes and vehicles. The elevated land surface temperature in urban areas is associated with elevated air temperatures. Most weather stations are placed in rural areas to avoid confusing temporal trends related to climate change with temporal trends related to changes in development and the associated urban heat island (UHI) effect [*Kalnay et al.*, 2006]. Thus our air temperature metrics tended to underestimate the local influence of cities. To test for a heat island effect on surface water temperatures, we calculated two different indices of urban heat island effects. The first index was based on urban center population size [*Oke*, 1973] calculated from New England town/city boundaries [*CT DEP*, 2006] and 2010 population by Census Block [*US Census Bureau*, 2015] . The second index of heat island effects was based on the difference between our interpolated 2006 air temperatures and the 2006 monthly averages of remotely sensed land surface temperatures from the National Aeronautics and Space Administration (NASA) TERRA/MODIS data set [*Wan*, 1999; *NASA Earth Observations*, 2014].

Shading of solar radiation occurs at both landscape scales (topographic shading) and local riparian scales. We accounted for topographic shading by using the ArcMap solar radiation function which incorporates effects of topographic shading based on our digital elevation models (30 meter pixels) as well as changes in solar angle with latitude and over diurnal and seasonal cycles. Our approach to evaluating the effect of solar radiation on stream temperature is constrained by the availability and

11

accuracy of local estimates of stream shading combined with the need for complete coverage of shading estimates over our entire region of interest. Detailed studies of stream shading effects on temperature regimes have employed hemispherical photography [*Isaak et al.*, 2010], but most monitoring programs do not have the resources to apply these methods. The most frequent measurement of stream/river shading collected during aquatic habitat assessments in the United States involve a series of measurements of canopy closure along stream transects using a spherical densiometer [*Kaufmann et al.*, 1999]. Canopy closure is a measure of the proportion of the sky hemisphere obscured by vegetation (or other objects) when viewed from a single point [*Jennings et al.*, 1999]. Thus it reflects obstructions to radiation coming from all directions over the course of a day and season and is influenced both by channel width and riparian vegetation density and height. A less common measurement, applied by the USGS NAWQA program, is an estimate of canopy angle from the sampler's point of view, which is affected by channel width and riparian vegetation height [*Fitzpatrick et al.*, 1998].

To support temperature model development, spatially continuous estimates of shading are needed both to fill gaps in habitat assessment coverage as well as to allow extension of model predictions to unmonitored sites. The best available data available across New England consist of continuous maps of canopy cover available for the U.S. based on remote sensing imagery. Canopy cover is an estimate of obstruction of the land (or water) surface based on a vertical (not multi-directional) projection of vegetation. Although canopy cover in the riparian zone is not synonymous with canopy closure, it is possible to develop statistical models relating the two types of measurements to facilitate development of continuous estimates of canopy closure over streams and rivers [*Isaak et al*, 2010]. These statistical models also need to take the effect of stream width into account because it affects canopy angle, as well as the effect of vegetation height on the relationship between canopy cover and canopy closure [*Jennings et al.,* 1999]. Thus we developed a series of empirical relationships to predict canopy

12

openness (1 – canopy closure) as a nonlinear function of both stream width and riparian canopy cover stratified by land-cover type to reflect the influence of vegetation height.

We estimated solar radiation inputs corrected for time of year, geographic coordinates and topographic shading using the Solar Radiation Toolset in ArcMap and assuming clear skies. The Solar Radiation Toolset includes both direct and diffuse radiation in the estimate of global radiation and allows the user to specify the proportion of global normal radation flux that is diffused. We based our estimates on clear sky conditions and did not adjust settings for degree of cloudinesss because we did not have complete time series of the degree of cloud cover across the region. We corrected ArcMap solar radiation estimates for effects of riparian shading using empirical relationships. We developed a nonlinear model using SAS PROC NLIN to predict percent canopy from densiometer measurements taken during stream habitat surveys (Table S3) as a function of bankfull width and NLCD 2001 percent canopy estimates within a 120 meter radius circle surrounding densiometer measurement points:

$PtOpen_{NLCD}$ = 100 - PTCAN01v1

where $PtOpen_{NLCD}$ = percent open based on NLCD canopy cover, and

PTCAN01v1 = NLCD percent canopy (2001).

If bankfull width is less than canopy diameter then we assume that % canopy cover is equal to NLCD % canopy:

$$predOpen = PtOpen_{NLCD}$$

where predOpen = predicted percent open

If bankfull width is greater than canopy diameter (CD) then we assume direct shading for the portion with overhanging canopy is equal to NLCD percent canopy:

13

$$predOpen_{direct} = PtOpen_{NLCD} * \left(\frac{CD}{EstBankfullWidth}\right)$$

where EstBankfullWidth = bankfull width if known or estimated bankfull width if not measured

Shading for the portion without overhanging canopy follows a Michaelis-Menten type function which converges to 100% open at large widths (EstBankfullWidth) with a "half-saturation constant", HS:

$$max = \left(\left(PtOpen_{NLCD} * \frac{CD}{EstBankfullWidth}\right) + \left(100 * \frac{(EstBankfullWidth - CD)}{EstBankfullWidth}\right) - (PtOpen_{NLCD})\right)$$

$$predOpen_{indirect} = \left(\frac{max * (EstBankfullWidth - CD)}{HS + (EstBankfullWidth - CD)}\right)$$

$$predOpen = predOpen_{direct} + predOpen_{indirect}$$

Values of CD and HS were estimated by fitting the nonlinear model. Following the approach of Isaak et al. [2010], we generated separate predictive equations for riparian zones dominated by deciduous or mixed tree cover, coniferous tree cover, or more "open" NLCD cover classes such as pasture or emergent wetlands. Finally, we calculated local shaded solar radiation as the product of canopy openness and (topographically-shaded) solar radiation.

2.3.2. Waterbody Attributes

We tested both lake and stream/river morphological characteristics as potential predictors of temperatures in running waters. We tested attributes to predict the influence of upstream lakes based on variables in an empirical model developed for lake epilimnion temperature [*Kettle et al.,* 2004]: maximum lake depth, the interaction of solar radiation and maximum lake depth and log_e(maximum lake depth x smoothed air temperature). We calculated stream/river width-to-depth ratios to reflect the potential influence of river morphometry on heat exchange between air and water. Greater surface area and shallower depths facilitate heat transfer [*Poole and Berman*, 2000].

14

Where possible, we acquired waterbody attributes from stream or lake survey data. We compiled stream/river dimensions from habitat monitoring surveys conducted during base flow conditions by state or Federal agencies (Table S2) and derived lake survey data from compilations published by Hollister et al. [2011]. We estimated reach dimensions not available from habitat surveys based on empirical equations for New England channel morphometry developed by Bent [2006]. We based maximum lake depths not available from lake survey data on estimates generated by Hollister et al. [2011] using topography (slopes) adjacent to lakes. We increased predicted maximum lake depths for NHDPlus lake polygons of less than one meter to one meter based on the assumption that shallower water bodies would support growth of aquatic vegetation and would be classified as wetlands rather than as lakes in the National Hydrography Dataset [*Cowardin et al.,* 1979].

2.3.3. Meteorological Attributes

We derived air temperature metrics corresponding to water temperature metrics from several sources. Weather station data and climatological grids of monthly minima and maxima temperatures were obtained from the PRISM Climate Group at Oregon State University (*Daly et al*., 2008) and from DAYMET [*Thornton et al.,* 2014]. We acquired weather data (air temperature and precipitation) for New England from the National Oceanographic and Atmospheric Administration (NOAA) National Climate Data Center Integrated Surface Data [*NOAA NCEI*, 2011]) through web-based queries, and derived a reduced set of stations after filtering for records at least 90% complete during the growing season. We matched water temperature monitoring stations to the nearest weather station for each year of record using Thiessen polygons generated around weather station points.

We used PRISM data sets to correct air temperature time series from weather stations nearest stream temperature stations for differences in elevation and distance to coast between data pairs. PRISM grids are generated using empirical equations to predict air temperatures using weather station

15

data with corrections for effects of elevation and distance from the coast [*Daly et al.*, 2008], however the underlying equations are not published. Therefore, we back-calculated transfer functions relating air temperature metrics at stream stations to air temperature metrics at weather stations from PRISM records extracted for each of those points.

We determined that various biases were present in both the PRISM and DAYMET datasets which produced spatial patterns in residuals from our initial models. DAYMET algorithms do not assume temporal stationarity for lapse rates, however they do not correct for aspect or coastal influence [*Thornton et al.,* 1997]. PRISM algorithms include elevation as a predictor and weight nearby points based on similarity (distance, elevation, aspect, distance from coast; [*Daly et al.,* 2008]). However, PRISM models assume the lapse rate (elevation influence) is constant over time. Initial examination of data for our region showed this was not a valid assumption. Calculated lapse rates appeared to increase between 1985 and 2010 to the east and decrease over time to the west of the Green Mountains in Vermont. Thus, using the geographically-weighted regression (GWR) procedure in ArcMap 10.1 to produce 800 meter grids for local regression coefficients, we generated our own yearly 30 meter grids of July and August median air temperatures from the U.S. Historical Climatology Network (USHCN) dataset [Menne et al., 2012] for comparison with PRISM-derived values. We included terms for elevation, latitude, and a dummy variable for coastal climate zone. We chose the number of neighboring points to include by using the GWR optimization option in ArcMap that is based on Akaike Information Criteria (AIC) values. We controlled for differences in aspect across major drainage divides by developing separate prediction grids by major drainage basin. Based on examination of initial model residuals, we excluded weather stations on the shoreline or offshore islands, those adjacent to large water bodies, and those in the middle of large municipalities.

16

2.4. Development of Stream Network and Spatial Statistical Network Models

2.4.1 Temporal distribution

Most temperature monitoring stations had only one year of data available. To avoid potential problems with pseudo-replication and to match water temperature dates with canopy cover dates as closely as possible, when multiple years of water temperature monitoring data were available we used thermal metrics from the most recent year of data available within the range of 1995 – 2010. Many of our predictive variables were static through time but for those which varied from year to year we matched the year of the predictor (e.g., analogous air temperature metric) to the water temperature monitoring year or, for land cover (LC) data and canopy cover (CC), to the nearest available year of data (1992 LC, 2001 LC, 2006 LC and CC, 2010 LC and CC).

The year of collection for temperature monitoring data used in construction of our model ranged from 1995 to 2010, with a median value of 2005. Thus our model should be robust to variability in response between relatively warm and cool and wet and dry years. For example, based on geographically weighted regression results, median July and August air temperatures averaged across New England varied by 4.0 and 2.6 ° C, respectively, between 1995 and 2010. August maximum temperature averages hit record highs in five of the six New England states in 2010. Between 1995 and 2010 in the northeastern United States July values for Palmer's drought severity index varied between -3.39 (dry, 1999) and 3.71 (wet, historic high in 1996) [Northeast Regional Climate Center, 2016].

2.4.2 Development of spatial statistical network models

The stream network models developed by ver Hoef and colleagues [*Ver Hoef et al.*, 2006; *Peterson and Ver Hoef*, 2010] are based on moving average calculations used to create a wide range of autocovariance functions. Models can incorporate spatial autocorrelation based on Euclidean distance

17

as well as spatial autocorrelation restricted to distances measured along flow networks. If one can travel downstream from one point to another, those points are called "flow-connected" while if one has to traverse downstream and then back upstream in a network to travel between two points, those points are referred to as "flow-unconnected". Models based on hydrologic distance that only allow autocorrelation for "flow-connected relationships" are called "tail-up" models because the tail of the moving average function points upstream. (Variograms typically describe a wedge-shaped pattern, with variance between pairs of points increasing with interpoint distance; see Supplemental materials for examples.) At stream confluences, segment weights (typically based on flow volume or watershed area) are used to proportionally split the moving average function between upstream segments. Thus, smaller tributaries will have less influence on a downstream location. In contrast, "tail-down" models include autocorrelation along stream networks between both flow-connected and unconnected points.

The Spatial Stream Network (SSN) software [*Ver Hoef et al*., 2014] is used to parameterize mixed models of the form:

$$Y = X\beta + \partial_{eu}z_{eu} + \partial_{tu}z_{tu} + \partial_{td}z_{td} + \partial_{nug}z_{nug}$$

where Y = dependent variable

X = matrix of fixed effect independent variables

β = parameter vector for fixed effects

∂ = variance component

z = random effect

eu = Euclidean autocorrelation

18

td = taildown autocorrelation

tu = tailup autocorrelation

nug = nugget effect

Application of ver Hoef's spatial statistical network models [*ver Hoef et al.*, 2014] requires construction of a continuous stream network without divergent flow paths or complex confluences (more than three reaches converging or two reaches converging without a downstream reach). We preprocessed the NHDPlus version 1 hydrography flow lines to meet these requirements using tools available in the ArcMap FloWs version 9.3, 1.0.1 or 1.0.2 [*Theobald et al.*, 2006] and STAR extensions [*Peterson and Ver Hoef*, 2014] to remove shoreline reaches, correct or remove reaches with missing flow direction, identify and remove minor flow paths from braided channels, and to adjust reach nodes slightly to avoid the occurrence of complex confluences. Using the STAR functions with ArcMap 9.3 or 10.1, we exported the network components to a Spatial Stream Network (SSN object) for use in the R package, SSN version 1.1.2 [*Ver Hoef et al.,* 2014]. The .ssn directory contains the spatial, attribute, and topological information needed to create an SSN object in R.

We conducted subsequent model development using R 3.02 with ArcMap 10.0 [*R Core Team*, 2010, 2013]. We fit an initial set of predictive models for each of the selected thermal metrics, computing spatial covariance matrices based on both Euclidean distance and flow-connected or flow-unconnected distance along flow networks. We developed and compared alternative spatial statistical models using protocols outlined in Ver Hoef et al. [2014], and applied Aikake's Information Criteria (AIC) to evaluate relative model performance [*Burnham and Anderson*, 2002]. First, we compared a simple linear regression model without spatial autocorrelation and containing only an air temperature metric to predict the corresponding water temperature metric with an alternative model containing the air

19

temperature metric and incorporating spatial autocorrelation based on Euclidean (straight-line) distance and flow-connected or flow-unconnected flow path distance. We then compared alternative models using the full suite of model predictors and all covariance matrix types (Euclidean, tailup, taildown). Tailup covariance structures only consider spatial autocorrelation between pairs of points that are flow-connected, while taildown covariance structures consider spatial autocorrelation along the stream network regardless of flow direction. We substituted alternate indicators of gradient (local gradient, main channel slope), of potential groundwater inputs (fraction coarse surficial deposits, Hydrologic Soil Group A, Extremely well-drained soils), of air temperature metrics (PRISM- or DAYMET-based, geographically weighted regression (GWR) estimates), and of the urban heat island index (population-based, land surface temperature difference). We then applied the equivalent of a backwards stepwise regression technique to the best of these models, removing each worst-fitting nonsignificant parameter term in sequence. Finally, we optimized the fit of the model by substituting all alternate forms of autocovariance functions available in the SSN package for Euclidean distance, flow-connected distance, and flow-unconnected distance (Cauchy, Spherical, Exponential, Gaussian, Mariah, Linear with Sill). We retained the model with the combination of covariance terms providing the best fit. We removed covariance functions contributing relatively little to the overall variance. We applied model diagnostics as outlined in Ver Hoef et al. [2014] to examine the distribution of residual errors, assess potential heterogeneity of variance, identify and assess model prediction outliers, and evaluate overall model fit and variance components. We also plotted studentized residuals versus each of the predictor variables to identify potential nonlinearities and used this as the basis for subsequent testing of second- and third-order terms where necessary.

We can expect the thermal regime of streams and rivers to be affected by retention time. Stream and river temperature is not at equilibrium with air temperature or solar radiation, which change diurnally, seasonally, and along the longitudinal profile of a river. It is possible that spatial

20

autocorrelation structure for temperature is better described by interpoint "flow time" measured by traveltime plus lake retention time rather than by stream length. Using the following predictors: July median air temperature (from PRISM), main channel slope, mean July shaded solar radiation, lake and wetland storage area (%), ln maximum lake depth, we compared three alternative modeling approaches for median July temperature: 1) a model based on July median air temperature alone, with no spatial structure, 2) models based on reach length as a measure of distance, and 3) models based on retention time plus network traveltime as a measure of interpoint flow time.   We estimated traveltime as the inverse of estimated reach velocity in the NHDPlus database. We calculated lake retention time based on the sum of lake inflow or outflow discharges (NHDPLus MAFLOWU variable) and estimated lake volume based on area and maximum depth assuming a conical shape. For the reach length models, we also compared a model which included only upstream lake predictors for reaches immediately downstream of lakes and with watershed predictors for all other reaches against a model which included *both* lake and watershed predictors for reaches immediately downstream of lakes.

3.   Results

3.1 Stream/River Thermal Regimes in New England

Stream temperature data are available from across most of New England although stations are not evenly distributed, with gaps in coverage for western and southern Vermont, northern Rhode Island, west-central New Hampshire and southeastern Massachusetts (Figure 1). The 78 temperature metrics we evaluated using PCA could be represented by a reduced suite of metrics representing overall temperature magnitude, temperature variability, and timing of growing season daily maximum. The first three principal components of the full season (June – September) daily statistics data set explained a total of 90.5% of the variance, with 79.4%, 6.4%, and 4.7% of variance

21

associated with Principal Components 1 (PC1), 2 (PC2), and 3 (PC3), respectively. PC1 represented the overall magnitude of temperature values, with a significant negative correlation with percent days suitable for coldwater fish and strongest positive correlations with numerous variables including monthly average temperature for July through September ($p < 0.05$; Figures S2, S3). The only metrics not significantly correlated with PC1 were daily temperature ranges and the timing of the growing season maximum daily temperature. PC2 and PC3 both had strong significant correlations with daily temperature ranges for June – September as well as the timing of the growing season maximum daily average temperature ($p < 0.05$).

The first three principal components of the full season (May – September) fine resolution temperature logger data set explained a total of 87.9% of the variance, with 69.3%, 13.1%, and 5.4% of variance associated with PC1, PC2, and PC3, respectively. Again, PC1 represented the overall magnitude of temperature metrics, with significant correlations with all metrics except for the maximum daily positive or negative rates of change. PC2 and PC3 had strongest correlations with the maximum daily positive or negative rates of change ($p < 0.05$).

Based on the results of the principal component analyses, we chose a reduced set of temperature metrics for model development: median July (MEDIAN7) or August (MEDIAN8) temperature, daily temperature range for July (DRANG7) or August (DRANG8), maximum positive or negative diurnal rates of change (degrees/hour) for July (MROC7+, MROC7-) or August (MROC8+, MROC8-), and timing (Julian Day) of growing season daily maximum (JDGSmx). We modeled both July and August metrics because different data sources tended to have different sampling windows. We chose to examine median rather than mean temperature because it should be more robust to outliers. We evaluated the response of both daily range and maximum rate of change because the latter metrics can only be calculated for stations with fine resolution temperature logger data

22

available. In addition, we chose growing season maximum daily average temperature (GSmax) for subsequent modeling because it represents the period of maximum stress for fish populations.

For Box-Cox transformed metrics calculated from daily statistics, correlation coefficients between our selected magnitude and timing variables (r = -0.03 to 0.35), between magnitude and rate of change variables (r = -0.30 to -0.10) and between range and timing variables (r = -0.03 to 0.23) were all low, although some were statistically significant (cutoff r = 0.197 at p = 0.05 for n = 98 in full growing season set). For transformed metrics calculated from 15 minute time series, correlation coefficients between magnitude and rate of change variables were also low (r = -0.29 to 0.03) with most not statistically significant (cutoff r = 0.236 at p = 0.05 for n = 68).

3.2. Shade Prediction Models

The nonlinear models used to predict percent open sky (100 - % canopy closure) as a function of stream width and NLCD canopy cover (stratified by vegetative class) were able to reproduce the general patterns of response, with highly variable canopy openness at narrow stream widths converging to 100% open as stream width increased (Table 2; Figure 2a). However, there was a bias in the model fit, with over predictions of canopy openness (under predictions of shade) at high densiometer readings and under prediction of canopy openness (over prediction of shade) at low densiometer readings (Figure 2b).

3.3. Comparison of Spatial Modeling Approaches with Stream Distance Versus Retention and Travel Time

The initial model for July median temperature with no spatial structure performed poorly (r^2 = 0.12). Although the scatter in Torgegrams showing patterns of variance along distance gradients was reduced considerably when we used lake retention + stream travel time as the distance metric, the hybrid model with reach length distance metric performed the best based on AIC values (Table 3). Thus,

23

we developed all final models using reach length as the distance metric, and including both watershed predictors and lake predictors for points on reaches immediately downstream of lakes. (Tests in Table 3 were based on a slightly earlier version of the model for July median temperature so fit statistics differ from those in Table 4.)

3.4. *Best Predictive Models for Reduced Set of Stream/River Thermal Metrics*

Models were successful in capturing main effects related to natural variability (i.e., reference condition), as well as potential anthropogenic impacts due to land use/land cover change (percent impervious cover (IC), urban heat island index, lake + wetland storage), water use (mean discharge) or weather (air temperature metrics, mean annual air temperature as a proxy for groundwater temperatures; Table 4).

Prediction accuracy was fairly consistent across watershed size, individual networks, and years. There was a slight tendency for the occurrence of outliers (|studentized residual| > 2) for watersheds in the range of 5 to 100 square kilometers drainage area (Figure 3a). Median residuals were close to zero for all independent flow networks within New England except for cases of three or fewer observations per network (Figure 3b). Model errors did not appear to increase or decrease with time, although not surprisingly, the spread of error is greater for years with more observations (Figure 3c). Median residuals were close to zero for all years except for 1999. The year 1999 showed consistently positive residuals (underprediction) and had the lowest Palmer Drought Severity Index within the model domain.

Main effects included in the predictive models for stream thermal metrics included corresponding air temperature metrics in all but one case (timing of growing season maximum). Shaded solar radiation was also significantly associated with increased monthly median and daily range values. Other significant ($p < 0.05$) main effects included morphometry (width-to-depth ratio, maximum lake

24

depth upstream, watershed area) and landscape variables related to retention time (main channel slope, lake + wetland storage, drainage density), inputs and potential temperature of groundwater (surficial sediments with high infiltration rates, mean annual air temperature), and anthropogenic inputs (% impervious cover (IC); Table 4). Indicators of landscape retention time (higher watershed storage, upstream lake, lower slope) tended to increase median temperatures and growing season maxima but had opposite effects on temperature variability. Watershed storage tended to moderate temperature variability but high gradient systems also had lower daily ranges. Fine scale variability (maximum rate of change) was higher in systems with lower retention time (higher main channel slope) or higher percent imperviousness. The presence of a lake immediately upstream tended to accentuate maximum negative rates of change (recovery) in July, but this effect decreased as lake depth increased.

Influences on July and August median temperatures were similar, with the exception that variables related to potential groundwater inputs (coarse surficial deposits, interaction between coarse surficial deposits and main channel slope) significantly influenced August but not July median values. Predictors for July and August ranges were also similar, with the exception that variables affecting water balance (extremely well-drained soils, imperviousness) affected only July ranges, and shading affected only August ranges.

For median monthly stream temperature models, some higher order main effects were significant, indicating nonlinear responses to main channel slope and shaded solar radiation. Main channel slope tended to have the strongest impact on median temperatures at intermediate values. In contrast to the findings of Mohseni et al. [1998], we found no significant higher order effects for air temperature.

In general, main channel slope (indicator of stream flashiness) was a better predictor of thermal metrics than local reach slope. Coarse surficial deposits were usually a better predictor than soil properties (soil drainage class or hydrologic group) even though the latter data sets were available at a

25

finer spatial resolution. For models predicting July or August median temperatures, GWR estimates of air temperature outperformed interpolated air temperature values based on PRISM. Urban heat island effects predicted by urban population size performed equally well or better in models than estimates based on land surface – air temperature differences.

Our results captured an interactive effect of stream morphometry with diel air temperature range rather than with solar radiation, as well as a main effect of solar radiation and the moderating influence of discharge on diel range. In addition to previously documented factors, we also quantified the effects of runoff (imperviousness) versus base flow (extremely well-drained soils, main channel slope) and the moderating influence of watershed storage on diel temperature ranges in streams and rivers.

For July diel range, potential impacts for changes in watershed storage and imperviousness are similar in magnitude (Table 5). Potential impacts due to water withdrawals depend both on unimpacted base flows, level of withdrawal and potential changes in width to depth ratio. For August diel range, potential impacts are slightly greater for changes in watershed storage as compared to changes in solar radiation (Table 5).

Best predictive models for monthly median, monthly range, and growing season maximum included covariance components based on both flow-connected or flow-unconnected distance and Euclidean distance (Figure 4; see Supplemental Figures S6 for remaining Torgegrams). The latter covariance component dropped out of predictions for August metrics. In contrast, predictions of timing of growing season maximum and maximum August diurnal negative rates of change included only spatial covariances based on Euclidean distance. A large proportion of the explained variation in models for maximum rates of change was related to spatial structure in comparison to main factor effects (Table 4).

26

3.5. Distribution of lotic thermal regimes across New England

For the first time we are able to map the predicted distribution of lotic thermal regimes across New England, excluding watersheds in northern Maine which extend outside of the U.S. We applied our spatial statistical model to predict median July temperatures based on 2006 land-cover conditions and July median air temperatures averaged over the period 2002 to 2010. The latter corresponds to the period of overlap with studies setting thermal thresholds for fish communities. We used stream thermal thresholds for mean July temperatures delineating transitions between coldwater, transitional coolwater, and warmwater fish communities derived by Beauchene et al. [2014] to classify expected stream and river thermal regimes: < 18.45 $^\circ$ C coldwater, 18.45 – 22.3 $^\circ$ C coolwater, and > 22.3 $^\circ$ C warmwater. As expected, coldwater reaches predominate in mountainous regions of Vermont, New Hampshire, and Maine (Figure 5). Even in areas where coldwater habitat predominates, however, large mainstem rivers are classified as coolwater or warmwater. (See for example Connecticut River along the Vermont/New Hampshire border.)

In summarizing resu1ts, we have excluded predictions with independent variables that were outside of the range of the model development data set. Within the model domain and based on 2006 land-use (NLCD wetlands + open water, NLCD percent canopy), urban heat index, and 2002 – 2010 air temperatures, the model predicts that 27.1% of stream + river kilometers have regimes that should support coldwater fish communities, 62.3% of stream kilometers should support coolwater fish communities, and 10.6% of stream kilometers should support warmwater fish communities. Predictions for NHDPlus nontidal reaches outside of the model domain (15.5% of total nontidal) have a distribution of 20.6% coldwater kilometers, 64.4% coolwater kilometers, and 15.0% warmwater kilometers, suggesting that under-representation of warmwater reaches in the monitoring dataset can bias our estimate of available habitat.

27

Development of a comprehensive regional model of lotic thermal regimes allowed us to identify

gaps in representation of driving factors affecting thermal regimes in the existing monitoring network

across New England. Only 1.3% percent of all model predictions for median July temperature in nontidal

reaches fell below (1.3%) or above (0.001%) of the range of available stream measurements (11.26 –

26.79 ° C). However, 15.5% of nontidal stream reach kilometers did have input variables that were

outside of the range of observations used to calibrate the model (model domain). Cases with

independent variables outside of the model domain were associated with shaded solar radiation (8043

of 59736 reach values < model observations and 3821 reach values > model observations), main channel

slope (881 reach values > model observations) and air temperature (589 reach values < model

observations and percent storage (229 reach values > model observations) for median July temperature.

Thus high elevation, steep gradient streams, streams at high and low extremes of shaded solar radiation,

and streams in high storage watersheds were poorly represented in the monitoring database supporting

our model.

4.   *Discussion*

4.1 Anthropogenic versus Natural Controls on Stream/River Thermal Regimes in New England

We identified a parsimonious set of four types of metrics to describe the growing season thermal

regime of New England streams: magnitude, daily and subdaily variation, and timing of growing season

maxima. This subset explained 88-91% of variation in our comprehensive set of 78 thermal metrics. We

were successful in producing robust prediction models for July and August median stream temperatures

and growing season maximum, with RMSPE of 1.4 – 1.5 ∘C (median) and 1.67 ∘C, respectively. Models

28

for predicting daily range and maximum rates of change were less successful, explaining only 5-14% of total variance, or 9-25% of variance once spatial autocorrelation was excluded.

Our empirical models are more comprehensive than most currently available in the literature. Most statistical temperature models developed for individual states, provinces, regions, or the conterminous United States have focused mainly on natural controls of inter-stream variability (air temperature, precipitation, watershed area, discharge, percent forest cover, glacial coverage, lake coverage, catchment elevation, channel slope, aspect, base flow index), with lesser attention given to the effects of riparian cover and percent developed land.

Our results highlight the importance of considering multiple mechanisms of anthropogenic impact on July and August median stream temperatures, including weather variability (changes in mean air temperature), the urban heat island index, reductions in riparian zone cover, net groundwater or surface water withdrawals, imperviousness, and changes in watershed storage (loss of natural wetlands, dam removal, or gain in constructed wetlands or storm water best management practices (BMPs)). The maximum "potential impact" of variables influenced by anthropogenic activities was calculated as the product of each regression model coefficient with the range of the associated variable in the observation data set. Based on regression coefficients and the range of predictors in our data set, we predict the greatest potential impact of anthropogenic changes on July median stream temperatures to be: changes in watershed storage > urban heat island effects > loss of riparian cover with increased solar radiation exposure (Table 5). In contrast, for August median temperatures, we predict the greatest potential impact of anthropogenic changes to be loss of riparian cover, then change in watershed storage (Table 5).

We expected the urban heat index effect on median July temperatures to be positive rather than negative. Predicted UHI effect based on municipal population size was positively correlated with the

29

difference between average land surface temperatures and interpolated air temperatures, the latter based mainly on rural weather stations ($p < 0.05$). It is possible that the apparent urban heat island effect on median July stream temperatures represents an evaporative cooling effect with the *rate* of increase in stream temperature slowing at high air temperatures, similar to the nonlinear stream – air temperature relationship at high air temperatures observed by Mohseni et al. (1998). Alternatively, the negative coefficient for UHI could be an artifact related to a positive association with another unmeasured factor negatively correlated with stream temperature. Urbanization can be associated with both excess water withdrawals in some locations and local surpluses in other locations due to interbasin transfers. Effects of urbanization on base flows could be even more pronounced in August than in July, which could explain differences in sensitivity between months. The wide range of variation in the relationship between elevated land surface temperature and UHI suggests that our measure of urban heat island effects is crude and needs improvement. In addition, better data are needed concerning development effects on stream baseflows with associated consequences for stream temperature.

Previous empirical models have not identified nonlinearities in response of stream temperature to solar radiation or to main channel slope as we did, although Fullerton et al. [2015] have described a wide variety of longitudinal patterns of temperature in large rivers, with asymptotic patterns the most frequent pattern observed. Nonlinear responses to solar radiation could reflect an increase in evaporative cooling at high temperatures, similar to the nonlinearities observed by Mohseni et al. [1998] in relationships of water temperature to air temperature, or the tendency for water depth to increase (increasing inertia to temperature change) and velocity to decrease (increasing retention time and approach to equilibrium) with distance downstream. The biases in our shading model could also contribute to the nonlinearities observed. Availability of region-wide LIDAR coverage will allow us to improve shading models in the future through more detailed solar exposure modelling. The observed nonlinearities in response to main channel slope could result from a combination of the interaction of

30

slope with coarse surficial deposits as well as changes in channel complexity along the river continuum. Groundwater discharges to streams vary as a function of water availability (precipitation), surficial sediment conductivity, and gradient. High gradient streams have high channel slopes but thin soils and relatively little coarse surficial sediment available for groundwater infiltration and storage. At the other extreme of very low channel gradients, braided channels can develop where floodplains are unconstrained, with complex microtopography and flow paths and localized cold spots [*Arscott et al.*, 2001], and low gradient rivers are also more abundant along the coast with potential oceanic moderating influences.

Our analysis of factors affecting diel stream/river temperature range was more comprehensive than existing studies, capturing both broad patterns related to climatic gradients as well as the influence of finer scale controls. Controls on diurnal patterns of stream temperature have been studied relatively little compared to evaluation of factors affecting mean stream temperatures. At a continental scale, correlation between mean summer daily ranges and landscape variables is very weak, while diel variability in air temperature can explain broad patterns in diel range of stream temperatures [*Maheu et al.*, 2015]. For reaches within meso-scale rivers with limited shading and minor groundwater influence, the solar-to-stream power ratio was the only reliable predictor of river diel temperature range normalized by mean annual basin temperature [*Link et al*. 2013].

Our ability to predict rate of change variables at the scale of hours or days was relatively poor in comparison to our ability to predict median or maxima stream temperatures at the monthly scale. Unlike mechanistic models, our statistical models do not reflect antecedent conditions. Monthly thermal metrics are more likely to capture equilibrium conditions while rate of change variables are sensitive to finer scale variability in both time and space. Our initial regressions of water versus air temperature showed evidence of two to three day lag periods in response, which we did not account for

31

in our predictive models. Modelling of rate of change variables at finer temporal scales could be improved by more extensive testing to determine the optimum extent of forested buffers upstream influencing downstream water temperatures and incorporating indicators of interannual variation in baseflow, both of which could influence recovery rates.

Most studies of the effects of development on stream thermal regimes have relied on detailed mechanistic models applied to individual reaches [*Janke et al.,* 2011; *Sabouri et al.*, 2013] with relatively few statistical modeling approaches at the local [*Nelson and Palmer*, 2007], regional, or national scales [*Segura et al.*, 2015]. In headwater Maryland watersheds, daily average stream temperatures were best predicted by a combination of percent deforestation and watershed area [*Nelson and Palmer*, 2007]. Superimposed on increased daily averages were periodic temperature spikes associated with individual storm events. Nelson and Palmer were able to predict frequency of spikes as a function of percent deforestation in the buffer zone, IC, and average discharge, but were unable to predict the magnitude of storm surge temperatures. In contrast, we were able to predict changes in both maximum positive and negative rates of change in stream temperatures as a function of multiple factors, including percent imperviousness, but could only explain a small portion of variability.

Variation in response to IC across sites and precipitation events is influenced by thermal storage capacity of different impervious surfaces [*Janke et al.*, 2011] and the timing and magnitude of events [*Winston et al.*, 2011]. In addition, thermal surges in storm water runoff can be either mitigated or exacerbated by storm water infrastructure [*Winston et al.*, 2011; *Sabouri et al.*, 2013]. In New England, outflow from the larger surface water BMPs (e.g., retention ponds) can exacerbate impacts of heated urban runoff by acting as heat sinks while larger subsurface BMP systems with greater contact with groundwater (e.g., infiltration systems) exhibit greater thermal buffering, with outflows similar to groundwater temperature [*Jacobs*, 2011].

32

The low proportion of variation explained for maximum rates of change in stream temperature was likely due to the variation in IC impacts across different classes of IC [*Janke et al*., 2011], and to the stochastic nature of thermal surges [*Nelson and Palmer*, 2011]. Estimation of impacts of IC on thermal surges in cross-watershed comparisons could be improved by including frequency of surges as a response metric. Better characterization of spatial variability in impacts of storm water runoff surges across subwatersheds will require characterization of impervious cover classes and detailed information on type of gray and green storm water practices, data that are typically not available in consistent format across large regions [A. C. Morrison, N. E. Detenbeck, and R. Abele, Approaches for calculating aggregate BMP contributions at a watershed scale, submitted to Landscape and Urban Planning 2014].

4.2. Comparison of Model Performance with Those of Existing Spatial Statistical Models for Stream Temperature

Although the datasets examined and cross-validation methods vary, our prediction models for median July and August stream temperature appear to outperform regional regression models which have achieved minimum prediction errors of 2-3°C [*Wehrly et al.,* 2009], reducing the RMSE to 1.4–1.5°C. The majority of model errors (predicted minus observed) were on the order of one to two degrees C (Figure 6). If we had run equivalent models for median July and median August stream temperatures without the spatial autocovariance error structure, we would have included additional terms in our July median temperature model (fraction coarse deposits and mean discharge for median July temperature model) and dropped upstream lake effects for the median August temperature model.

We did not achieve the RMSE levels of $\leq 1^{\circ}$C reported by Isaak et al. [2010] for some western systems, but the accuracy of these models depends in part on the density of available monitoring stations. In addition, Isaak et al. included duplicate measurements at individual monitoring stations from different years by artificially shifting the location of one point within each pair by 100 meters along

33

stream lines. To minimize this potential pseudoreplication, we only retained the annual observation

closest to the year 2006 (the most recent year of land-cover data available at time of initial data

analysis). Unfortunately, the SSN package can not yet incorporate both spatial and serial

autocorrelation into the modelled error structure. Other factors potentially limiting performance of the

spatial statistical network models in the Northeast include the high density of small dams in New

England [*Martin and Apse*, 2011], the intensity of water withdrawals and recharge for municipal and

industrial purposes [*Weiskel et al.,* 2009], the counteracting fine-scale effects of urban infrastructure not

captured by our landscape variables, the stochastic nature of storm water runoff thermal surges, and

other unknown disturbances not accounted for in our dataset. Hill et al. [2013] noted a significant

degradation in model performance between predictions of thermal regime at reference versus

disturbed sites.

4.3.   Model Limitation, Future Data and Model Improvements

Data quality and availability could have limited the accuracy and precision of our model predictions.

Temperature monitoring stations are not sited randomly and there are geographic gaps in coverage and

differences in temporal resolution and sampling windows among monitoring agencies. Improved

temporal coverage would improve predictions of growing season maxima as well as allowing

researchers and managers to assess potential thermal impacts in other seasons and the influence of the

seasonal accumulation of degree days. Estimates of solar shading could be improved by improved

spatial resolution of hydrography and riparian buffer characterization. The spatial resolution of the

NHDPlus stream network (1:100,000) has an intrinsic spatial error that limits our ability to meaningfully

characterize buffer zones of less than 120 meters in width; this issue is exacerbated by the limited

resolution of regional canopy cover and land-use data (30 meter resolution).   Indicators of groundwater

input would be enhanced by the availability of regionally consistent medium resolution coverages of

glacial geology features. Although finer scale coverages of soils and soil attributes exist, the glacial

34

geology coverages tended to produce better predictors of moderating effects of groundwater, probably

because they reflect the depth of deposits as well as surficial features. For example, the median value

for percent extremely well-drained soils across our input data sets was 45% as compared to the lesser

value of 7% for coarse surficial deposits. Ideally, an assessment-level model of potential groundwater

discharge and recharge areas such as that of Baker et al. [2003] or Wehrly et al. [2003] could be updated

for use in ArcMap 10.1 and applied in the stream temperature prediction model; resources available for

this project limited our ability to do this. Baker and colleagues' analysis was based on Darcy's Law,

which predicts groundwater movement as a function of differences in groundwater head along a flow

path, area, and hydraulic conductivity of substrate. Our analysis considered effects of percent coarse

surficial sediments as an indicator of hydraulic conductivity, effects of main channel slope, and potential

interactions. Main channel slope captures the effect of vertical gradients parallel to the stream channel

but not lateral gradients adjacent to the stream. In many but not all cases, channel slope is correlated

with the width of floodplains and lateral gradients [*Rosgen and Silvey*, 1996]Consistent regional datasets

to describe or predict variation of base flow over time and space, and finer resolution data on water

withdrawals and discharges would improve our ability to model anthropogenic impacts. Finally, regional

estimates of connected (effective) IC by cover class, surrogates for underground stormwater

infrastructure, and more consistent characterization of gray and green infrastructure BMPs would

improve our ability to evaluate the impacts of development and potential mitigation strategies.

5. Research and Management Implications

While summarizing the best current knowledge of factors driving variability in thermal regimes, our

models also provide a foundation for future improvements as monitoring data and GIS coverages

improve. Our model will be useful in 1) producing regional maps of thermal regimes characterized both

by summer median temperatures and daily range [*Maheu et al.,* 2015], 2) predicting reference condition

35

in the absence of anthropogenic impacts, and 3) identifying critical thermal refugia. As evidenced by our

outlier analysis (section 2.2), these models are sufficiently accurate to allow managers to identify

aberrant temperature regimes related to discharges. We intend to use these maps in conjunction with

regional fish monitoring data to examine potential impacts of development on fish communities, as well

as the influence of moderating factors. While our current analyses focused on characterization of

reference (or impacted) condition at a static point in time, in the future we will expand our approach to

evaluate landscape factors affecting the thermal sensitivity of streams. This will allow us to model

combined scenarios of land use and climate change to evaluate and prioritize alternative mitigation

strategies for minimizing impacts.

Acknowledgments

and Jeff Hollister for providing preliminary technical reviews of this manuscript and three anonymous reviewers of the final publication.

References

Allen, D.M. (2008) Development and application of a process-based basin-scale stream temperature model. Doctoral dissertation. University of California, Berkeley, CA.

Archfield, S.A., Vogel, R.M., Steeves, P.A., Brandt, S.L., Weiskel, P.K. and Garabedian, S.P. (2010), The Massachusetts Sustainable Yield Estimator: A decision –support tool to assess water availability at ungagged stream locations in Massachusetts, U.S. Geological Survey Scientific Investigation Report 2009-5227, Reston, VA.

Arscott, D.B., Tockner, K., and Ward, J.V. (2001), Thermal heterogeneity along a braided floodplain river (Tagliamento River, northeastern Italy), *Can. J. Fish. Aquat. Sci.,***58**,2359-2373, doi 10.1139/f01-183.

Baker, M.E., M.J. Wiley, M.L. Carlson, and P.W. Seelbach (2003), A GIS model of subsurface water potential for aquatic resource inventory, assessment, and environmental management, *Env. Manage*., **32**, 706-719, doi: 10.1007/s00267-003-0018-1.

Beauchene, M., M. Becker, C.J. Bellucci, N. Hagstrom, and Y. Kanno (2014), Summer thermal thresholds of fish community transitions in Connecticut Streams, *N. Amer. J. Fish. Manage*.,**34**,119-131, doi: http://dx.doi.org/10.1080/02755947.2013.855280.

Bent, G. C. (2006) Equations for estimating bankfull channel geometry and discharge for streams in the northeastern United States [abs.]: Proceedings of the Eighth Interagency Sedimentation Conference, Reno, Nevada, April 2–6, 2006, p. 1026, accessed April 16, 2013, http://pubs.usgs.gov/misc_reports/-FISC_1947-2006/pdf/1st-7thFISCs-CD/8thFISC/Poster_Bent_AbstractOnly.pdf.

37

Beyer, H.L. (2004), Hawth's Analysis Tools for ArcGIS, Available at

http://www.spatialecology.com/htools.

Beyer, H.L. (2010), Geospatial Modelling Environment Version: 0.7.2 RC2, Available at

www.spatialecology.com/gme.

Burnham, K. P., and D. R. Anderson (1998), Model Selection and Inference: A Practical Information-

Theoretic Approach, Springer-Verlag, New York, NY, doi:10.1007/b97636.

Cowardin, L. M., V. Carter, F. C. Golet, E. T. LaRoe (1979), Classification of wetlands and deepwater

habitats of the United States. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.

C.

Cox, M. M., and J. P. Bolte (2007), A spatially explicit network-based model for estimating stream

temperature distribution, *Environ. Modell. Softw.*, **22**, 502-14, doi:10.1016/j.envsoft.2006.02.011.

CT DEP (2006), Northeastern United States Town Boundary Polygon. Connecticut Department of

Environmental Protection, Hartford, CT. Accessed online at: http://222.ct.gov/deep.

Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis and P. P. Pasteris

(2008), Physiographically sensitive mapping of climatological temperature and precipitation across the

conterminous United States, *Int. J. Climatol.*, **28**, 2031-2064, doi:10.1002/joc.1688.

DeWeber, J. T. and T. Wagner (2014), A regional neural network ensemble for predicting mean daily

river water temperature, *J. Hydrol.*, **517**, 187-200, doi:10.1016/j.jhydrol.2014.05.035.

Faustini, J. M., P.R. Kaufmann, and A. T. Herlihy (2009), Downstream variation in bankfull width of

wadeable streams across the conterminous United States, *Geomorphology*, **108**(3-4), 292-311, doi:

38

10.1016/j.geomorph.2009.02.005.

Fenneman, N.M. (1938), Physiography of Eastern United States, McGraw-Hill, New York, N.Y.

Fitzpatrick, F.A., I.R. Waite, P.J. D'Arconte, M.R. meador, M.A. Maupin, and M.E. Gurtz (1998), Revised methods for characterizing stream habitat in the National Water-Quality Assessment Program: U.S. Geological Survey Water-Resources Investigations Report 98-4052, 67 pp.

Fullerton, A.H., C.E. Torgerson, J.J. Lawler, R.N. Faux, E.A. Steel, T.J. Beechie, J.L. Ebersole, and S.G. Leibowitz (2015), Rethinking the longitudinal stream temperature paradigm: region-wide comparison of thermal infrared imagery reveals unexpected complexity of river temperature, *Hydrol. Proc.*, **29**(22), 4719-4737, DOI: 10.1002/hyp.10506.

Gallice, A., B. Schaefli, M. Lehning, M.B. Parlange, and H. Huwald (2015), Stream temperature prediction in ungauged basins: review of recent approaches and description of a new physics-derived statistical model, *Hydrol. Earth Syst. Sci.,* **19**, 3727-3753, doi: 10.5194/hess-19-3727-2015.

Gu, C., W. P. Anderson Jr, J. D. Colby, and C. L. Coffey (2015), Air-stream temperature correlation in forested and urban headwater streams in the Southern Appalachians, *Hydrol. Process.*, **29**, 1110-1118, doi:10.1002/hyp.10225.

Hastie, T., R. Tibshirani, and J. Friedman (2009), The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd edition. Springer, New York, NY.

Herb, W.R. and H. G. Stefan (2011), Modified equilibrium temperature models for cold-water streams, *Water Resour. Res.*, **47**, 1-13, doi:10.1029/2010WR009586.

39

Hilderbrand, R. H., M. T. Kashiwagi, and A. P. Prochaska (2014), Regional and local scale modeling of stream temperatures and spatio-temporal variation in thermal sensitivities, *Environ. Manage.*, **54**, 14-22, doi:10.1007/s00267-014-0272-4.

Hill, R. A., C. P. Hawkins, and D. M. Carlisle (2013), Predicting thermal reference conditions for USA streams and rivers, *Freshwater Sci*, **32**(1), 39-55, doi:http://dx.doi.org/10.1899/12-009.1

Hill, R. A., C. P. Hawkins, and J. Jin (2014), Predicting thermal vulnerability of stream and river ecosystems to climate change, *Climatic Change*, **125**, 399-412, doi:10.1007/s10584-014-1174-4.

Hollister, J. W., W. B. Milstead, and M. A. Urrutia (2011), Predicting maximum lake depth from surrounding topography, *PLoS ONE,* **6**(9), e25764, doi:10.1371/journal.pone.0025764.

Horizon Systems Corporation (2010), National Hydrography Dataset Plus: NHDPlus V1 Tools, Accessed at http://www.horizon-systems.com/NHDPlus/NHDPlusV1_tools.php

Isaak, D. J., C. H. Luce, B. E. Rieman, D. E. Nagel, E. E. Peterson, D. L. Horan, S. Parkes and G. L. Chandler (2010), Effects of climate change and wildfire on stream temperatures and salmonid thermal habitat in a mountain river network, *Ecol. Appl.*, **20**, 1350-71, http://dx.doi.org/10.1890/09-0822.1.

Jacobs, J. (2011) Examination of Thermal Impacts from Stormwater Best Management Practices, University of New Hampshire Stormwater Center, Durham, NH, USA. (http://www.unh.edu/unhsc/sites/unh.edu.unhsc/files/progress_reports/UNHSC%20EPA_Thermal_Study_Final_Report_1-28-11.pdf)

Janke, B. D., O. Mohseni, W. R. Herb and H. G. Stefan (2011), Heat release from rooftops during rainstorms in the Minneapolis/St. Paul Metropolitan Area, USA, *Hydrol. Process.*, 25, 2018-2031, doi:10.1002/hyp.7954.

40

Jennings, S.B., N.D. Brown, and D. Sheil (1999), Assessing forest canopies and understorey illumination: canopy closure, canopy cover, and other measures, *Forestry* 72:59-73.

Johnson, M. F., R. L. Wilby, and J. A. Toone (2014), Inferring air-water temperature relationships from river and catchment properties, *Hydrol. Process.*, **28**, 2912-2928, doi:10.1002/hyp.9842.

Jones, N.E. and B. Schmidt (2013), ThermoStat Version 3.1: Tools for Analyzing Thermal Regimes. Ontario Ministry of Natural Resources, Aquatic Research and Development Section. http://people.trentu.ca/nicholasjones/thermostat.htm.

Kalnay, E., M. Cai, H. Li, and J. Tobin (2006), Estimation of the impact of land-surface forcings on temperature trends in the eastern United States, *J. Geophys. Res.*, **111**, D06106, doi:10.1029/2005JD006555.

Kanno, Y., J. C. Vokoun, and B.H. Letcher (2014), Paired stream-air temperature measurements reveal fine-scale thermal heterogeneity within headwater brook trout stream networks, *River Res. Appl.*, **30**, 745-755, doi:10.1002/rra.2677.

Karl, T. R., H. F. Diaz, and G. Kukla (1988), Urbanization: Its detection and effect in the United States climate record, *J. Climate*, **1**, 1099-1123, doi:http://dx.doi.org/10.1175/1520-0442(1988)001<1099:UIDAEI>2.0.CO;2.

Kaufmann, P.R., P. Levine, E.G. Robison, C. Seeliger, and D.V. Peck (1999), Surface waters: Quantifying physical habitat in wadeable streams, Washington (DC): U.S. Environmental Protection Agency. Report EPA/620/R-99/003. 158 p.

41

Kelleher, C., T. Wagener, M. Gooseff, B. McGlynn, K. McGuire, and L. Marshall (2012), Investigating controls on the thermal sensitivity of Pennsylvania streams, *Hydrol. Process.*, **26**, 771-785, DOI:10.1002/hyp.8186.

Kettle, H., R. Thompson, N. J. Anderson, and D. M. Livingstone (2004), Empirical modeling of summer lake surface temperatures in southwest Greenland, *Limnol. Oceanogr.*, **49**, 271-282, doi:10.4319/lo.2004.49.1.0271.

Krause, C. W., B. Lockard, T. J. Newcomb, D. Kibler, V. Lohani, and D. J. Orth (2004), Predicting influences of urban development on thermal habitat in a warm water stream, *JAWRA*, **40**, 1645-8, doi:10.1111/j.1752-1688.2004.tb01612.x.

Laaha G., J. O. Skøien, F. Nobilis and G. Blöschl (2013), Spatial prediction of stream temperatures using top-kriging with an external drift, *Environ. Model. Assess.*, **18**, 671–683, DOI 10.1007/s10666-013-9373-3.

LeBlanc, R. T., R. D. Brown, and J. E. Fitzgibbon (1997), Modeling the effects of land use change on the water temperature in unregulated urban streams, *J. Env. Manage.*, **49**, 445-69, doi:10.1006/jema.1996.0106.

Link, O., A. Huerta, A. Stehr, A. Monsalve, C. Meier, and M. Aguayo (2013), The solar-to-stream power ratio: A dimensionless number explaining diel fluctuations of temperature in mesoscale rivers, *River Res. Appl.*, **29**, 792-803, doi:10.1002/rra.2579.

Long, D.L. and R. L. Dymond (2014), Thermal pollution mitigation in cold water stream watersheds using bioretention*, JAWRA*, **50**, 977-987, doi:10.1111/jawr.12152.

Maheu, A., N. L. Poff, and A. St-Hilaire (2015), A classification of stream water temperature regimes in the conterminous USA, *River. Res. Appl.*, Published online doi:10.1002/rra.2906.

Markle, J. M. and R. A. Schincariol (2007), Thermal plume transport from sand and gravel pits – Potential thermal impacts on cool water streams, *J. Hydrol.*, **338**, 174-195, doi:10.1016/j.jhydrol.2007.02.031.

Martin, E. H. and C. D. Apse (2011), Northeast Aquatic Connectivity: An Assessment of Dams on Northeastern Rivers. The Nature Conservancy, Eastern Freshwater Program. Accessed 04/23/14 from http://rcngrants.org/sites/default/files/final_reports/NEAquaticConnectivity_Report.pdf.

Maxwell, J. R., Edwards, C. J., M. Jensen, S. Paustian, H. Parrott, and D. Hill (1995), A hierarchical framework of aquatic ecological units in North America (Nearctic Zone), General Technical Report *NC-176*, North Central Forest Experiment Station, USDA Forest Service, St. Paul, MN.

McKenna, J. E., Jr., R. S. Butryn, and R. P. McDonald (2010), Summer stream water temperature models for Great Lakes streams, *Trans. Am. Fish. Soc.*, **139**, 1399-1414, doi:10.1577/T09-153.1.

Menne, M.J., I. Durre, R.S. Vose, B.E. Gleason, and T.G. Houston (2012), An overview of the Global Historical Climatology Network-Daily Database, *J. Atmos. Ocean. Technol.*, **29**, 897-910.

Mohseni, O., H. G. Stefan, and T. R. Erikson (1998), A nonlinear regression model for weekly stream temperatures, *Water Resour. Res.*, **34**, 2685-2692, doi:10.1029/98WR01877.

Moore, R. D. (2006), Stream temperature patterns in British Columbia, Canada, based on routine spot measurements, *Can. Water Resour. J.*, **31**, 41-56, doi:10.4296/cwrj3101041, 41-56.

Moore, R. D., M. Nelitz and E. Parkinson (2013), Empirical modelling of maximum weekly average stream temperature, *Can. Water Res. J.*, **28**, 135-147, doi:10.1080/07011784.2013.794992.

43

Morrison, A.C., A.J. Gold, and M.C. Pelletier (2016), Evaluating key watershed components of low flow regimes in New England streams, *J. Environ. Qual.* (accepted) doi:10.2134/jeq2015.08.0434.

NASA Earth Observations (NEO) (2015), Land Surface Temperature, Accessed at http://neo.sci.gsfc.nasa.gov/view.php?datasetId=MOD11C1_M_LSTDA.

National Renewable Energy Laboratory (2015) Dynamic Maps, GIS Data, & Analysis Tools. December 23, 2015, Accessed at http://www.nrel.gov/gis/data_solar.html.

Nelson, K. C. and M. A. Palmer (2007), Stream temperature surges under urbanization and climate change: Data, models, and responses, *JAWRA*, **43**, 440-452, doi:10.1111/j.1752-1688.2007.00034.x.

NOAA NCEI (2011), Integrated Surface Dataset, National Oceanic and Atmospheric Administration, National Center for Environmental Information (formerly National Climatic Data Center), Accessed at http://www.ncdc.noaa.gov/

North Atlantic Landscape Conservation Cooperative (2012), Stream Temperature and Data Modeling Meeting: May 3, 2012. Accessed at http://northatlanticlcc.org/projects/impacts-of-climate-change-on-stream-temperature/temperature-data-and-modeling-meeting.

Northeast Regional Climate Center (2016), Summary tables, Accessed at http://www.nrcc.cornell.edu/regional/tables/tables.html.

Olden, J. D. and R.J. Naiman (2010), Incorporating thermal regimes into environmental flows assessments: modifying dam operations to restore freshwater ecosystem integrity, *Freshwater Biol.*, **55**, 86–107, doi:10.1111/j.1365-2427.2009.02179.x.

44

Olivero, A.P. and M.G. Anderson (2008), Northeast Aquatic Habitat Classification System, Nature

Conservancy, Eastern Regional Office, Boston, MA. Accessed 7/21/2009 from

http://rcngrants.org/content/northeastern-aquatic-habitat-classification-project)

Oke, T.R. (1973), City size and the urban heat island, *Atmos. Env.*, **7**, 769-779, doi:10.1016/0004-6981(73)90140-6.

Peterson, E. E. and J. M. Ver Hoef (2014), STARS: An ArcGIS toolset used to calculate the spatial

information needed to fit spatial statistical models to stream network data, *J. Stat. Softw.*,

56(2), doi:10.1.1.368.2858

Poff, N. L. and J. V. Ward (1989), Implications of streamflow variability and predictability for lotic

community structure: a regional analysis of streamflow patterns, *Can. J. Fish. Aquat. Sci*., **46**, 1805–1817,

doi:10.1139/f89-228.

Poole, G.C. and C.H. Berman (2000), Pathways of Human Influence on Water Temperature Dynamics

in Stream Channels. U.S. Environmental Protection Agency, Region 10, Seattle, WA.

www.krisweb.com/biblio/gen_usepa_pooleetal_2000_pathways.pdf.

R Development Core Team (2010), R: A language and environment for statistical computing. R

Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

R Core Team (2013), R: A language and environment for statistical computing. R Foundation for

Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Randall, A. D. (2001), Hydrogeologic framework of stratified-drift aquifers in the glaciated northeastern United States, U.S. Geological Survey Professional Paper *1415-B*, United States Geological Survey, Reston, VA.

Rosgen, D.L. and H.L. Silvey (1996) Applied River Morphology. Wildland Hydrology Books, Fort Collins, CO.

Sabouri, F., B. Gharabaghi, A. A. Mahboubi, and E. A. McBean (2013), Impervious surfaces and sewer pipe effects on storm water runoff temperature, *J. Hydrol.*, **502**, 10-17, doi:10.1016/j.jhydrol.2013.08.016

Sahoo, G. B., S. G. Schladow, and J. E. Reuter (2009), Forecasting stream water temperature using regression analysis, artificial neural network, and chaotic non-linear dynamic models, *J. Hydrol.*, **378**, 325-42. doi:10.1016/j.jhydrol.2009.09.037.

Schwartz, G.E. and R.B. Alexander (1995), Soils data for the Conterminous United States Derived from the NRCS State Soil Geographic (STATSGO) Data Base. Edition 1.1. U.S. Geological Survey Open-File Report 95-449. Accessed at http://ww.water.usgs.gov/lookup/getspatial?ussoils.

Segura, C., P. Caldwell, G. Sun, S. McNulty, and Y. Zhang (2015), A model to predict stream water temperatures across the conterminous USA, *Hydrol. Process.*, **29**, 2178-2195, doi:10.1002/hyp.10357.

Theobald, D. M., J. B. Norman, E. Peterson, S. Ferraz, A. Wade, and M. R. Sherburne (2006), Functional Linkage of Water basins and Streams (FLoWS) v1 User's Guide: ArcGIS tools for Network-based analysis of freshwater ecosystems, Natural Resource Ecology Lab, Colorado State University, Fort Collins, CO.

46

Thornton, P. E., S. W. Running, and M. A. White (1997), Generating surfaces of daily meteorological variables over large regions of complex terrain, *J. Hydrol.*, **190**, 214-51, doi:http://dx.doi.org/10.1016/S0022-1694(96)03128-9.

US Census Bureau (2014), Geography, Accessed at http://www.census.gov/geo/www/tiger.

USDA NRCS (2011) Soil Data Viewer 6.0 User Guide, United States Department of Agriculture, Natural Resources Conservation Service, Accessed at http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_052432.pdf.

Ver Hoef, J.M. and E.E. Petersen (2010) A moving average approach for spatial statistical models of stream networks, *J. Am. Stat. Assn.,* **105**(489):6-18, DOI: 10.1198/jasa.2009.ap08248.

Ver Hoef, J. M., E. E. Peterson, D. Clifford, and R. Shah (2014), SSN: An R package for spatial statistical modeling on stream networks, *J. Stat. Softw.*, **56**(3), 1-45.

Wandle, S.W., Jr. and A. D. Randall (1994), Effects of surficial geology, lakes and swamps, and annual water availability on low flows of streams in central New England, and their use in low-flow estimation, U.S.G.S. Water-Resources Investigations Report *93-4092*, U.S. Geological Survey, Marlborough, MA.

Wann, Z. (1999), MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD) Version 3.3. (Accessed online http://modis.gsfc.nasa.gov/data/atbd/atbd_mod11.pdf)

Wehrly, K.E., M.J. Wiley, and P.W. Seelbach (2003) Classifying regional variation in thermal regime based on stream fish community patterns, *Trans. Am. Fish. Soc.*, **132**, 18-38, DOI: 10.1577/1548-8659(2003)132<0018:CRVITR>2.0.CO;2.

47

Wehrly, K. E., T. O. Brenden, and L. Wang (2009), A comparison of statistical approaches for predicting stream temperatures across heterogeneous landscapes, *JAWRA*, **45**(4), 986-997, doi:10.1111/j.1752-1688.2009.00341.x.

Weiskel, P. K., S. L. Brandt, L. A. DeSimone, L. J. Ostiguy, and S. A. Archfield (2010), Indicators of streamflow alteration, habitat fragmentation, impervious cover, and water quality for Massachusetts stream basins (Ver 1.8, July 2, 2013), U.S. Geological Survey Scientific Investigations Report *2009–5272*, 70 p., plus CD–ROM. (Also available at http://pubs.usgs.gov/sir/2009/5272/).

Wiley, M.J., D.W. Hyndman, B.C. Pijanowski, A.D. Kendall, C. Riseng, E.S. Rutherford, S.T. Cheng, M.L. Carlson, J.A. Tyler, R.J. Stevenson, P.J. Steen, P.L. Richards, P.W. Seelbach, J.M. Koches, and R.R. Rediske (2010) A multi-modeling approach to evaluating climate and land use change impacts in a Great Lakes River Basin, *Hydrobiologia*, **657**, 243-262, DOI 10.1007/s10750-010-0239-2.

Winston, R. J., W. F. Hunt, and W. G. Lord (2011), Thermal mitigation of urban storm water by level spreader-vegetative filter strips, *J. Environ. Eng.*, **137**, 707-716, doi:http://dx.doi.org/10.1061/(ASCE)EE.1943-7870.0000367.

Table 1. Data sources for temperature time series used in development of New England temperature model.

Geographic extent	Agency[a]	Web site	Acquisition Date	Contact name	Stations
New England	USGS	http://waterdata.usgs.gov	6/23/2010		14
New England	USGS NAWQA	http://infotrek.er.usgs.gov/apex/f?p=NAWQA:HOME:2755121354639815	9/30/2010		9
Connecticut	DEP		2/20/2011	Mike Beauchene	219
Maine	DEP		8/18/2009	Susanne Meidel	234
Maine	University of Maine; ME DMR ME Atlantic Salmon Commission	No longer available online	7/13/2010	Richard Dill	100
Massachusetts	DEP		9/1/2010	Richard Chase	96
New Hampshire	DFW		7/7/2010	Matt Carpenter	118
Vermont	DFW		7/21/2010	Rod Wentworth	96
Charles River watershed, MA	Charles River watershed association		8/7/2010	Julie Wood	9
Wood-Pawcatuck	Wood-Pawcatuck Watershed Association	www.wpwa.org	6/21/2010	Denise Poyer	13

[a] USGS = United States Geological Survey, US EPA = United States Environmental Protection Agency, NAWQA = National Water Quality Assessment, DEP = Department of Environmental Protection, DMR = Department of Marine Resources, DFW = Department of Fisheries and Wildlife.

49

Geographic extent	Agency[a]	Web site	Acquisition Date	Contact name	Stations
watershed, RI					

Table 2. Coefficients of equations predicting percent open sky as a function of percent NLCD canopy in buffer zone and estimated bankfull width based on nonlinear model fits. CD (canopy diameter) is the estimate of the bankfull width beyond which stream width increases stream exposure and HS is the half-saturation constant, the channel width (beyond CD) at which shading is reduced by 50%. Open covers include all NLCD land-use classes other than deciduous forest (including woody wetlands) and coniferous forest.

Dominant cover	CD	HS
Deciduous forest	5.6	8.1
Coniferous forest	6.0	4.8
Open	10.8	7.3

51

Table 3. Comparison of alternate models for July median temperature, comparing effect of incorporating different distance metrics and strategies for describing upstream lake effects. Model in italics represents the best solution based on Aikake's Information Criterion.[a]

Upstream Lake Effect Model Variables	AIC	RMS Prediction Error (deg C)	r^2	n parameters
No network distance metric (nonspatial)				
None	3736		0.119	2
Length distance metric (km)				
None	3057.7	1.641	0.333	9
Separate	3114.6	1.674	0.279	13
Hybrid	*3045.8*	*1.627*	*0.340*	*13*
Retention time + traveltime along network (days)				
None	3125.4	1.704	0.289	13
Hybrid	3099.7	1.683	0.284	13

[a] Separate = separate lake effect and watershed effect variables; reaches with upstream lakes have watershed variables set to zero; other reaches have lake variables set to zero. Hybrid = both lake effect and watershed effects included for reaches with upstream lakes; other reaches have lake variables set to zero.

52

Table 4. Final spatial statistical network model structure to predict New England stream and river thermal metrics. Dependent variables are stream/river water temperature metrics. Independent variables are measured at the watershed scale with the exception of point values at monitoring locations (air temperature metrics, mean discharge, width-to-depth ratios) or NHDPlus reach scale estimates (local channel gradient, shaded solar radiation). Metrics for maximum positive (+) or negative (-) rates of change (MROC) were subjected to a BoxCox transformation before analysis with an index (lambda) of +MROC+: -0.5, +MROC8: -0.3; -MROC7: -0.5; -MROC8: -0.4. AIC = Aikake's Information Criterion, RMSPE = root mean square prediction error, NRMSPE = normalized RMSPE. MEDIANn = median temperature for month n, DRANGn = daily range for month n, GSmax = growing season maximum daily average, JDGSmax = Julian day of growing season maximum. Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$. Spatial autocovariance model forms: C = Cauchy, G = Gaussian, LS = Linear with sill, M = Mariah, S = Spherical.

	MEDIAN7	MEDIAN8	DRANG7	DRANG8	GSmax	JDGSmx	+MROC7	+MROC8	-MROC7	-MROC8
					Dependent variable					
Sample size	798	825	794	776	197	248	469	491	469	491
y-intercept	7.25***	9.75***	3.2***	1.8***	15***	214***	-0.10**	0.95***	1.02***	1.05***
					Independent Variable Coefficients (significance)					
Air temp metric	MEDIAN7	MEDIAN8	DRANG7	DRANG8	GSmax		+MROC7	+MROC8	-MROC7	-MROC8
	0.65***	0.51***		0.08***	0.19***					
Urban heat index-pop	-0.33***									
Watershed area			0.001*	0.0018**	-0.009***	-0.05*				
Local channel gradient					-0.6***					
Main channel slope	-153.8***	-202.3***	-15.6***	-12***						3.05***
Squared main channel slope	2573***	3758***								
Cubed main channel	-13430***	-20090***								

53

slope										
July shaded solar radn	0.15***									
Aug shaded solar radn		0.21***			0.05**					
Extremely well-drained soils			0.7*			35**				
Coarse surficial deposits		-3.94*					0.15**	0.12*		
Imperviousness			.016*				-0.01***	-0.008***	-0.013***	-0.0083***
Drainage density						-35***				
Lake + wetland storage	0.068***	0.030*	-0.02*	-.022**	0.10*					
Mean discharge			-.0013*	-.0027**		0.07*				
Mean annual air temp										
WidthtoDepth (WtoD)			-0.0005*		-0.023***					
WtoD x Air Temp			1.2e-4***		0.0013***					
Ln(max Lake Depth)		0.52**							-0.11*	
Lake Upstream (0/1)	1.18***	0.34(*)							0.017ns	
Covariance components (fraction variance explained)										
Tail-up	LS	LS	S	LS	LS				S	M
	0.26	0.20	0.10	0.14	0.02				0.21	0.17
Tail-down	LS				S		LS	S		
	0.08				0.14		0.23	0.26		
Euclidean	G	C	S	S	G	G	G		G	G
	0.13	0.17	0.34	0.39	0.25	0.38	0.31		0.38	0.14
Nugget	0.12	0.20	0.42	0.38	0.17	0.47	0.36	0.48	0.4	0.72
Model performance										
Generalized r^2	0.41	0.43	0.14	0.10	0.41	0.14	0.10	0.05	0.05	0.07
fixed effects/fixed + nugget	0.77	0.69	0.25	0.20	0.72	0.23	0.22	0.10	0.11	0.09
AIC	2777	2728	2375	2095	754.8	2238	556.2	712.7	771.7	872.5
RMSPE	1.49	1.42	1.71	0.97	1.63	21	0.11	0.15	0.28	0.24
NRMSPE (%)	9.6	8.9	17.0	11.7	11.0	18.4	3.8	15.0	15.7	14.2

Model performance with no spatial autocovariance included										
AIC	3011	2907	2592	2298	806.3	2303	651.6	763.6	839.0	917.8
RMSPE	1.85	1.68	2.09	1.19	2.02	26	0.50	0.55	0.62	0.65

Table 5. Predicted effects of anthropogenic factors over range of predictor variables.

Temperature metric/ predictors	Median (range)	Coeff	Min	Max	Potential impact
MEDIAN7	21.6 (16.7-26.9)				
Urban heat index		-0.33	0.0	8.9	-2.95
July shaded solar		0.15	1.0	17.4	2.46
Lake + wetland		0.068	0.0	45.5	3.10
MEDIAN8	21.0 (16.0-26.9)				
Aug shaded solar		0.21	0.8	14.5	2.88
Lake + wetland		0.03	0.0	45.5	1.37
DRANG7	2.9 (0.4-10.5)				
Urban heat index		-0.28	0.0	8.9	-2.50
Imperviousness		0.016	0.0	48.7	0.78
Lake + wetland		-0.019	0.0	45.5	-0.86
DRANG8	2.6 (0.2-15.4)				
Aug shaded solar		0.05	0.8	14.5	0.69
Lake + wetland		-0.022	0.0	45.5	-1.00
GSmax	23.2 (9.6-30.1)				
Lake + wetland		0.1	0.0	45.5	4.55
July MROC+	1.1 0 - 26.6				
Imperviousness		0.15	0.0	48.7	7.30
Aug MROC+	1.3 (0.3-38.7)				
Imperviousness		0.084	0.0	48.7	4.09
July MROC-	-0.8 -39.3 - -0.3				
Imperviousness		-0.15	0.0	48.7	-7.30
Aug MROC-	-0.8 -57.6 - -0.17				
Imperviousness		-0.022	0.0	48.7	-1.07

Figures

Figure 1. Location of stream/river temperature monitoring stations used in model development for different stream/river thermal metrics (growing season maximum, monthly maximum rate of change (degrees C/hour), monthly medians).

Figure 2. (a) Percent open sky versus estimated bankfull width, comparing observed open sky from densiometer measurements (diamond) with model predictions (square). b) Model predicted open sky as compared to open sky measurements from densiometer readings. Predictive model was a nonlinear function setting % open equal to (1 - % NLCD canopy cover) for stream widths less than estimated tree canopy diameter and increasing to a maximum of 100% following a Michaelis-Menten function for the portion of stream width greater than estimated tree canopy diameter. - - - 1:1 line.

Figure 3. Variation in model error (studentized residual) for median July temperature predictions as a function of a) drainage area, b) year, and c) individual stream/river network. Asterisks at base of graph c) indicate networks with only one to three points.

Figure 2. Torgegram for spatial statistical network model for median July stream temperatures showing spatial autocorrelation (increasing variance with interpoint distance) for both flow-connected and flow-unconnected points. Size of points indicates relative number of points contributing to each estimate.
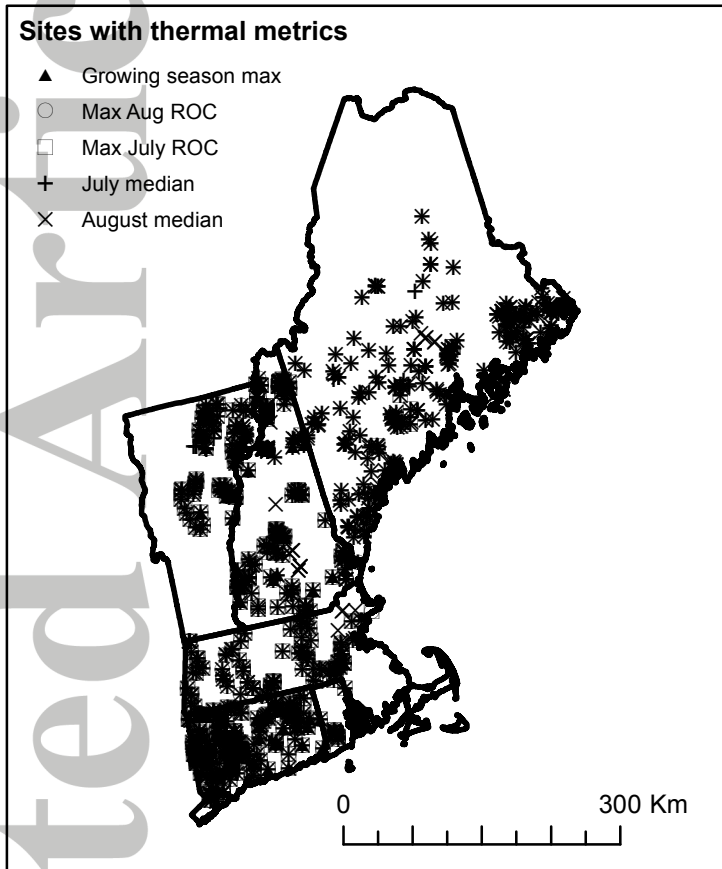
Figure 5. Predicted thermal regime of New England streams and rivers in temperature model spatial domain based on 2006 land condition and median July air temperature averaged over the period 2002-2010.

57

Figure 6. Predicted versus observed for a) median July stream temperature and b) median August stream temperature for final spatial statistical network models based on leave-one-out cross validation (LOOCV) process.
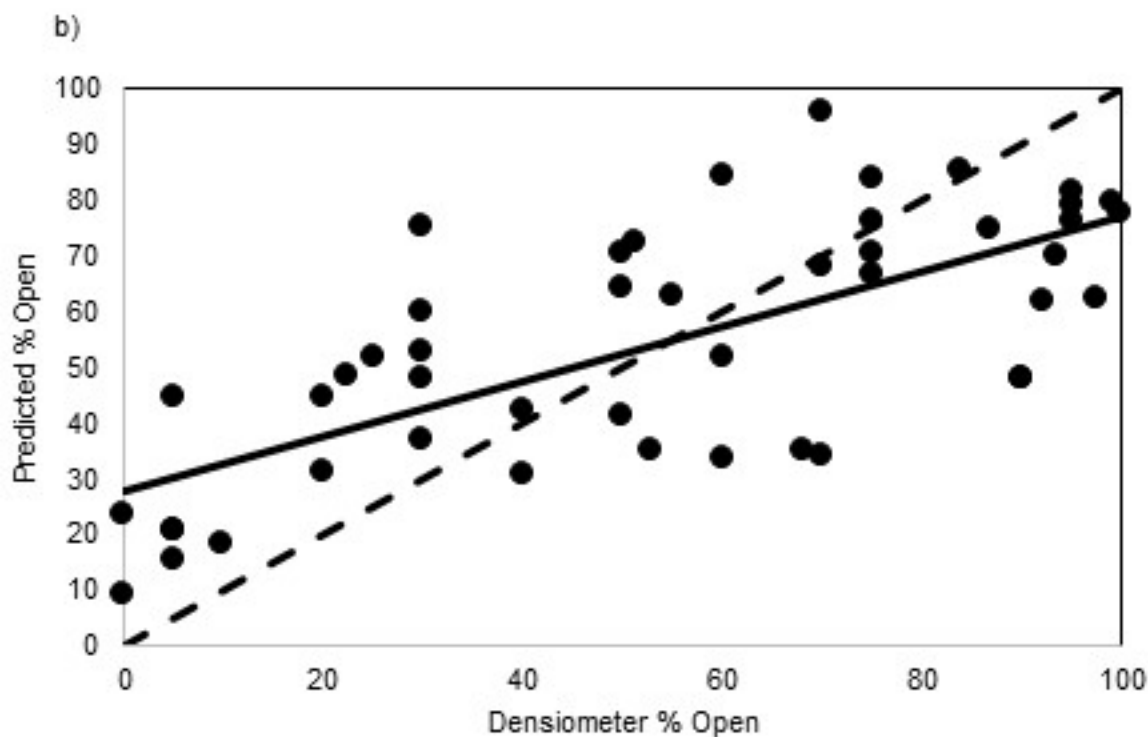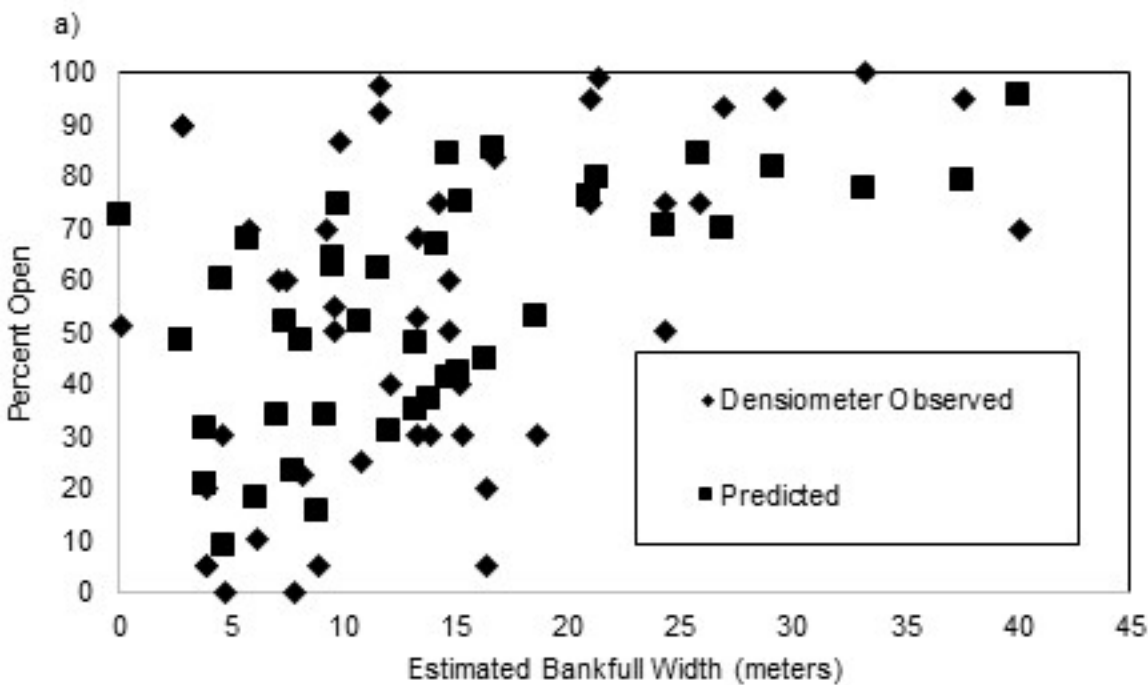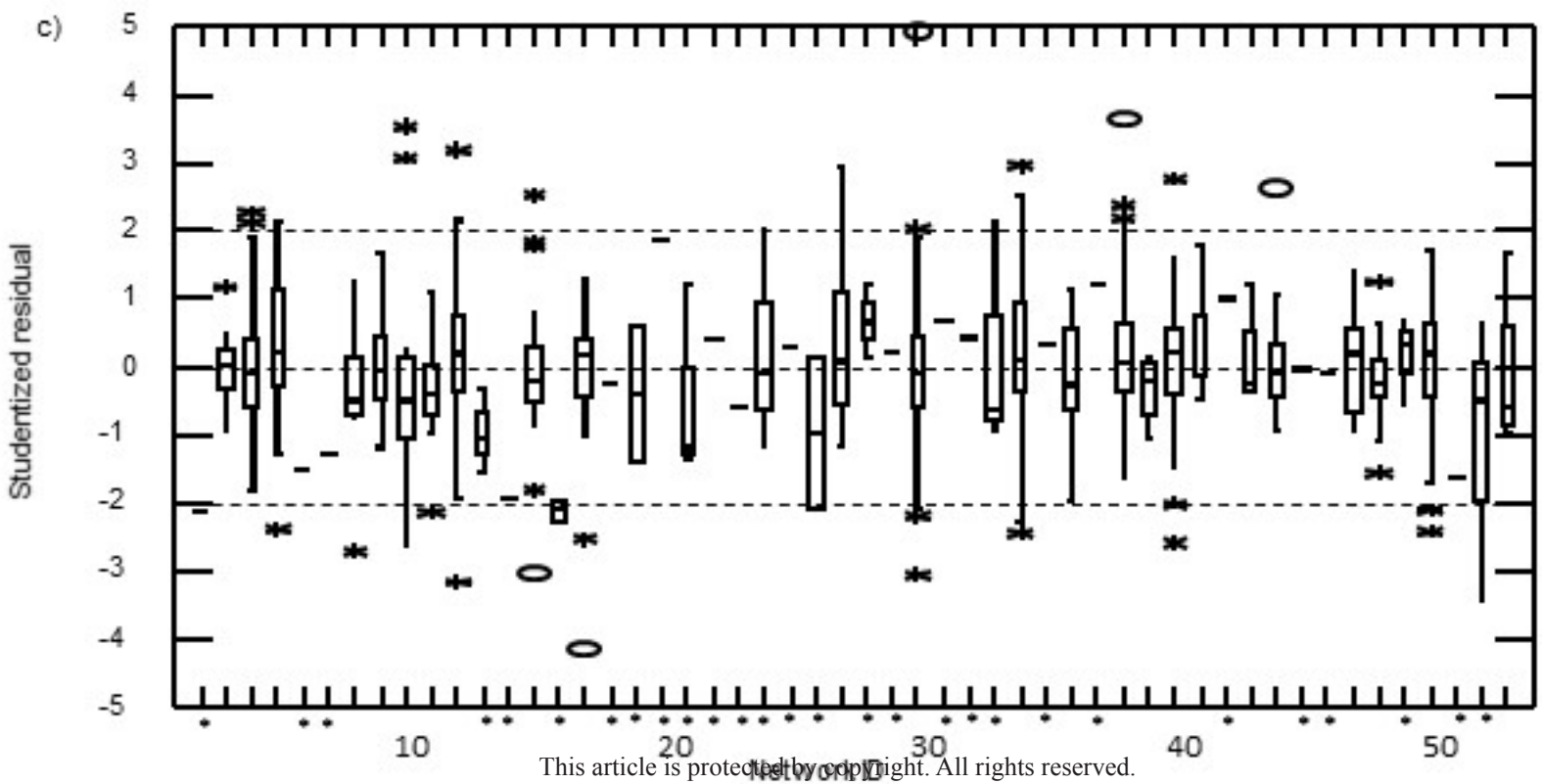
58

Figure 1. Figure

Sites with thermal metrics

▲ Growing season max
○ Max Aug ROC
□ Max July ROC
+ July median
✕ August median

0   300 Km

Figure 2. Figure

a)

b)

Figure 3. Figure

a)

b)

c)
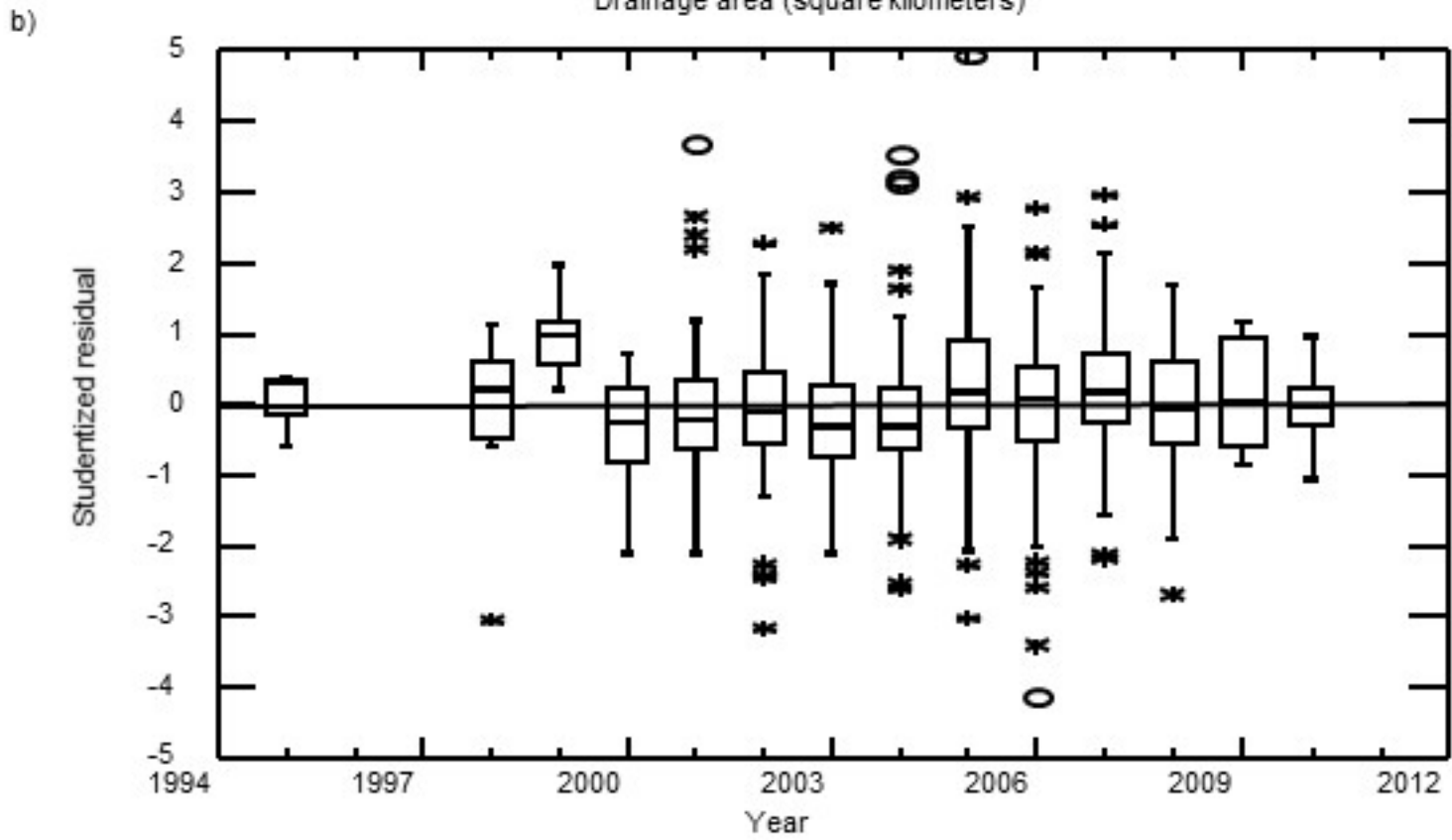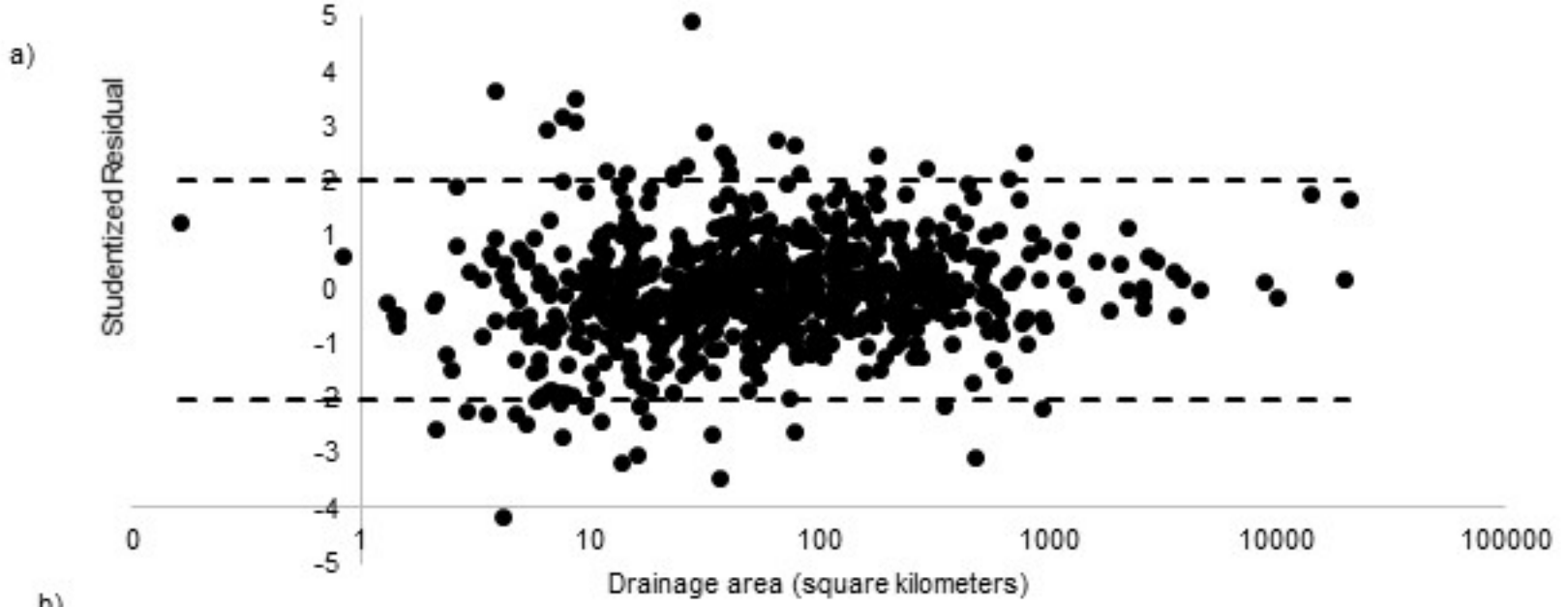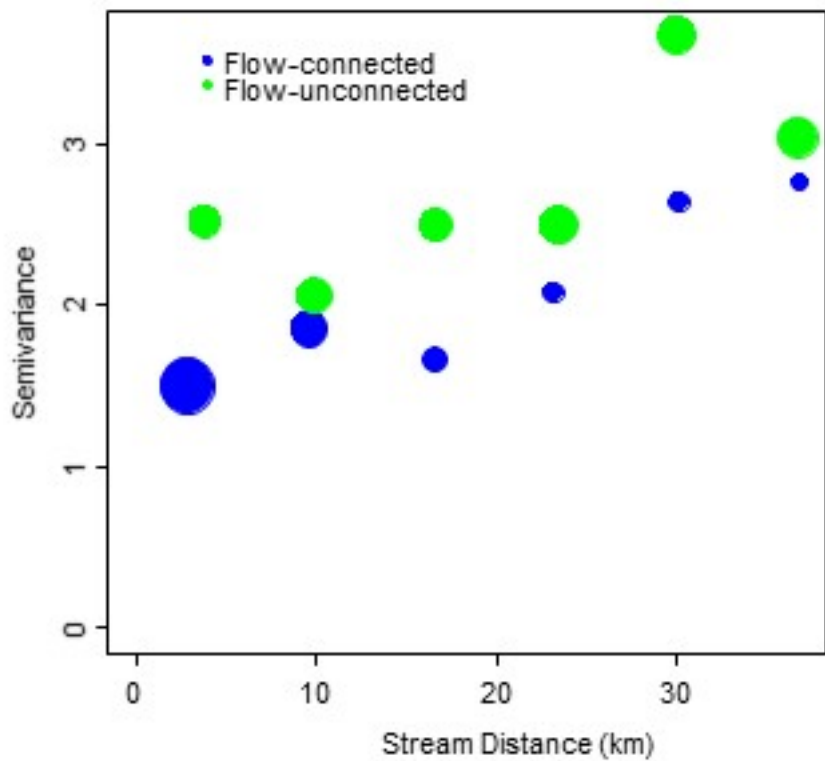
Figure 4. Figure

Figure 5. Figure

Legend

Flowlines

Thermal Regime

—— Cold
—— Cool
—— Tidal
—— Warm

Elevation (m)

Value

High : 1916

Low : 0

M E

V T

N H

M A

R I

C T

| 0 | 45 | 90 | 180 Kilometers |

figure 6. Figure