# Daily model of stream temperature for regional predictions

**Daniel J. Hocking, Ben Letcher, and Kyle O'Neil**

*Daniel J. Hocking (dhocking@usgs.gov), US Geological Survey, Conte Anadromous Fish Research Center, Turners Falls, MA, USA

## Abstract

Set up the problem. Explain how you solve it. Tell what you find. Explain why it's the best thing ever.

## Introduction

**Options:** Water Research, **Water Resources Research**, Freshwater Biology, Journal of Hydrology, Ecohydrology, Journal of Environmental Quality, Hydrobiologia, JAWRA

Temperature is a critical factor in regulating the physical, chemical, and biological properties of streams. Warming stream temperatures decrease dissolved oxygen, decrease water density, and alter the circulation and stratification patterns of streams (refs). Biogeochemical processes such as nitrogen and carbon cycling are also temperature dependent and affect primary production, decomposition, and eutrophication (refs). Both physical properties and biogeochemical processes influence the suitability for organisms living in and using the stream habitat beyond just primary producers. Additionally, temperature can have direct effects on the biota, especially poikliotherms such as invertebrates, amphibians, and fish [*Xu et al.*, 2010b, 2010a; *Al-Chokhachy et al.*, 2013; e.g., *Kanno et al.*, 2013]. Given commercial and recreational interests, there is a large body of literature describing the effects of temperature on fish, particularly the negative effects of warming temperatures on cool-water fishes such as salmonids . Finally, stream temperature can even affect electricity, drinking water, and recreation (see van Vliet et al 2011). Therefore, understanding and predicting stream temperatures are important for a multitude of stakeholders.

Stream temperature models can be used for explanatory purposes (understanding factors and mechanisms affecting temperature) and for prediction. Predictions can be spatial and temporal including forecasting and hindcasting. Predictions across space are especially valuable because there is often a need for information at locations with little or no observed temperature data. For example, many states have regulations related to the management of streams classified as cold, cool, and warm waters (refs), but because of the tremendous number of headwater streams it is impossible classify most streams based on observed data. Therefore, modeled stream temperature is needed to classify most streams for regulatory purposes. Forecasting can provide immediate information such as the expected temperature the next hour, day, or week as well as long-term information about expected temperatures months, years, and decades in the future. Hindcasting can be used to examine temperature variability and trends over time and for model validation. Both forecasting and hindcasting are useful for understanding climate change effects on stream temperature regimes.

Given the importance of temperature in aquatic systems, it is not surprising that there are a variety of models and approaches to understanding and predicting stream temperature. Stream temperature models are generally divided into three categories: deterministic (also called process-based or mechanistic), stochastic, and statistical [*Caissie*, 2006; *Benyahya et al.*, 2007; *Chang and Psaris*, 2013]. Deterministic models are based on heat transfer and are often modeled using energy budgets [*Caissie*, 2006; *Benyahya et al.*, 2007]. The models require large amounts of detailed information on the physical properties of the stream and adjacent landscape as well as hydrology and meteorology. These models are useful for detailed re assessments and scenario testing. However, the data requirements preclude the models from being applied over large spatial extents.

Stochastic models attempt to combine pattern (seasonal and spatial trends) with the random deviations to describe and predict environmental data [*Kiraly and Janosi*, 2002; *Sura et al.*, 2006; *Chang and Psaris*, 2013]. Stochastic models of stream temperature generally rely on relationships between air and water temperature then with random noise and an autoregressive correlation, often decomposed by seasonal and annual components. These models are mostly commonly used to model daily temperature fluctuations because of their ability to address autocorrelation and approximate the near-random variability in environmental data [*Caissie et al.*, 2001; *Kiraly and Janosi*, 2002; *Ahmadi-Nedushan et al.*, 2007]. A limitation is that the physical processes driving temperature fluctuations are not elucidated with these models. They are generally used to describe characteristics and patterns in a system and to forecast these patterns in the future [*Kiraly and Janosi*, 2002]. Additionally, stochastic models rely on continuous, often long, time series from a single or a few locations. Inference cannot be made to other locations without assuming that the patterns and random deviations are identical at those locations.

As with stochastic models, statistical models generally rely on correlative relationships between air and water temperatures, but also typically include a variety of other predictor variables such as basin, landscape, and land-use characteristics. Statistical models are often linear with normally distributed error and therefore used at weekly or monthly time steps to avoid problems with temporal autocorrelation at shorter time steps (e.g. daily, hourly, sub-hourly). Parametric, nonlinear regression models have been developed to provide more information regarding mechanisms than traditional statistical models without the detail of physical deterministic models [*Mohseni et al.*, 1998]. Researchers have also developed geospatial regression models that account for spatial autocorrelation within dendritic stream networks [*Isaak et al.*, 2010; *Peterson et al.*, 2010, 2013]. However, due to the complexity of the covariance structure of network geostatistical models, they are best used for modeling single temperature values across space (e.g. summer maximum, July mean, etc.) rather than daily temperatures [*Peterson et al.*, 2007, 2010; *Ver Hoef and Peterson*, 2010]. Additionally, statistical machine learning techniques such as artificial neural networks have been used to model stream temperatures when unclear interactions, nonlinearities, and spatial relationships are of particular concern [*Sivri et al.*, 2007, 2009; *DeWeber and Wagner*, 2014b].

In contrast with deterministic approaches, statistical models require less detailed site-level data and therefore can be applied over greater spatial extents than process-based models. They also can describe the relationships between additional covariates and stream temperature, which is a limitation of stochastic models. These relationships can be used to understand and predict anthropogenic effects on stream temperature such as timber harvest, impervious development, and water control and release [*Webb et al.*, 2008]. Quantifying the relationship between anthropogenic effects, landscape characteristics, meteorological patterns, and stream temperature allows for prediction to new sites and times using statistical models. This is advantageous for forecasting and hindcasting to predict and understand climate change effects on stream temperatures. This is critical because not all streams respond identically to air temperature changes and the idiosyncratic responses may be predicted based interactions of known factors such as flow, precipitation, forest cover, basin topology, impervious surfaces, soil characteristics, geology, and impoundments [*Webb et al.*, 2008].

Letcher et al. [2015] outline six general challenges of statistical stream temperature models including accounting for 1) the non-linear relationship between air and water temperature at high and low air temperatures, 2) different relationships between air and water temperature in the spring and fall (hysteresis), 3) thermal inertia resulting in lagged responses of water temperature to changes in air temperature, 4) incomplete time series data and locations with large differences in the amount of available data, 5) spatial and temporal autocorrelation, and 6) important predictors of stream water temperature other than air temperature. They developed a statistical model that addresses aspects of non-linear relationships, hysteresis, thermal inertia, and spatial and temporal autocorrelation but their analysis was limited to a single small network of streams with long time series [*Letcher et al.*, 2015].

We describe a novel statistical model of daily stream temperature that incorporates features of stochastic models and extends the Letcher et al. [2015] framework to large geographic areas. This model handles time series data of widely varying duration from many sites using a hierarchical mixed model approach to account for autocorrelation at specific locations within watersheds. It incorporates catchment, landscape, and meteorological covariates for explanatory and predictive purposes. It includes an autoregressive function to account for temporal autocorrelation in the time series, a challenge with other statistical models at fine temporal resolution. Additionally, our hierarchical Bayesian approach readily allows for complete accounting of uncertainty. We use the model to predict daily stream temperature across the northeastern United States over a 34-year time record.

# Methods

## Study area

**Map of data locations: size = amount of data, color/shape = training-validation - Kyle, Ana, or Matt make? See deWeber 2014 for example**

## Water temperature data

We gathered stream temperature data from state and federal agencies, individual academic researchers, and non-governmental organizations (NGOs). The data were collected using automated temperature loggers. The temporal frequency of recording ranged from every 5 minutes to once per hour. This data was consolidated in a PostgreSQL database linked to a web service at http://www.db.ecosheds.org. Data collectors can upload data at this website and choose whether to make the data publicly available or not. The raw data is stored in the database and users can flag problem values and time series. For our analysis, we performed some automated and visual QAQC on the sub-daily values, summarized to mean daily temperatures and performed additional QAQC on the daily values. The QAQC was intended to flag and remove values associated with logger malfunctions, out-of-water events (including first and last days when loggers were recording but not yet in streams), and days with incomplete data which would alter the daily mean. We developed an R (ref) package for analyzing stream temperature data from our database, including the QAQC functions which can be found at https://github.com/Conte-Ecology/conteStreamTemperature. The R scripts using these functions for our analysis are available at https://github.com/Conte-Ecology/conteStreamTemperature_northeast.

Stream reach (stream section between any two confluences) was our finest spatial resolution for the analysis. In the rare case where we had multiple logger locations within the same reach recording at the same time, we used the mean value from the loggers for a given day. In the future, with sufficient within reach data, it would be possible to use our modeling framework to also estimate variability within reach.

*Stream network delineation*

## Meteorological (, Climatic,) and landscape data - separate landscape if use climate data for future projections

**Table of Variables** - include part of the model they're in (fixed, site, huc, year)

## Statistical model

Statistical models of stream temperature often rely on the close relationship between air temperature and water temperature. However, this relationship breaks down during the winter in temperature zones, particularly as streams freeze, thereby changing their thermal and properties. Many researchers and managers are interested in the non-winter effects of temperature. The winter period, when phase change and ice cover alter the air-water relationship, differs in both time (annually) and space. We developed an index of air-water synchrony ($Index_{sync}$) so we can model the portion of the year that it not affected by freezing properties. The index is the difference between air and observed water temperatures divided by the water temperature plus 0.000001 to avoid division by zero.

We calculate the $Index_{sync}$ for each day of the year at each reach for each year with observed data. We then calculate the 99.9% confidence interval of $Index_{sync}$ for days between the 125 and 275 days of the year (05 May and 02 October). Then moving from the middle of the year (day 180) to the beginning of the year, we searched for the first time when 10 consecutive days were not within the 99.9% CI. This was selected as the spring breakpoint. Similarly moving from the middle to the end of the year, the first event with fewer than 16 consecutive days within the 99.9% CI was assigned as the autumn breakpoint. Independent breakpoints were estimated for each reach-year combination. For reach-years with insufficient data to generate continuous trends and confidence intervals, we used the mean break points across years for that reach. If there was not sufficient local reach information, we used the mean breakpoints from the smallest hydrologic unit the reach is nested in (i.e. check for mean from HUC12, then HUC10, HUC8, etc.). More details regarding the identification of the synchronized period can be found in Letcher et al. (*in review*). The portion of the year between the spring and autumn breakpoints was used for modeling the non-winter, approximately ice-free stream temperatures.

We used a generalized linear mixed model to account for correlation in space (stream reach nested within HUC8). This allowed us to incorporate short time series as well as long time series from different reaches and disjunct time series from the same reaches without risk of pseudoreplication (ref: Hurlbert). By limited stream drainage area to <200 $km^2$ and only modeling the synchronized period of the year, we were able to use a linear model, avoiding the non-linearities that occur at very high temperatures due to evaporative cooling and near 0 C due to phase change (ref: mohseni).

We assumed stream temperature measurements were normally distributed following,

$$t_{h,r,y,d} \sim \mathcal{N}(\mu_{h,r,y,d}, \sigma)$$

where $t_{h,r,y,d}$ is the observed stream water temperature at the reach ($r$) within the sub-basin identified by the 8-digit Hydrologic Unit Code (HUC8; $h$) for each day ($d$) in each year ($y$). We describe the normal distribution based on the mean ($mu_{h,r,y,d}$) and standard deviation ($\sigma$) and assign a vague prior of $\sigma = 100$. The mean temperature is modeled to follow a linear trend

$$\omega_{h,r,y,d} = X_0 B_0 + X_{h,r} B_{h,r} + X_h B_h + X_y B_y$$

but the expected mean temperature ($\mu_{h,r,y,d}$) was also adjusted based on the residual error from the previous day

$$\mu_{h,r,y,d} = \begin{cases} \omega_{h,r,y,d} + \delta(t_{h,r,y,d-1} - \omega_{h,r,y,d-1}) & \text{for } t_{h,r,y,d-1} \text{ is real} \\ \omega_{h,r,y,d} & \text{for } t_{h,r,y,d-1} \text{ is not real} \end{cases}$$

where $\delta$ is an autoregressive [AR(1)] coefficient and $\omega_{h,r,y,d}$ is the expected temperature before accounting for temporal autocorrelation in the error structure.

$X_0$ is the $n \times K_0$ matrix of predictor values. $B_0$ is the vector of $K_0$ coefficients, where $K_0$ is the number of fixed effects parameters including the overall intercept. We used **???XX???** fixed effect parameters including the overall intercept. These include **??latitude, longitude, upstream drainage area, percent forest cover, elevation, surficial coarseness clas-**

**sification, percent wetland area, upstream impounded area, and an interaction of drainage area and air temperature??**. We assumed the following distributions and vague priors for the fixed effects coefficients

$$B_0 \sim \mathcal{N}(0, \sigma_{k_0}), \text{for } k_0 = 1, ..., K_0,$$

$$B_0 = \beta_0^1, ..., \beta_0^{K_0} \sim \mathcal{N}(0, 100)$$

$$\sigma_{k_0} = 100$$

**??The effects of air temperature on the day of observation ($d$) and mean air temperature over the previous 7 days varied randomly with reach nested within HUC8, as did precipitation, the previous 30-day precipitation mean, and the interactions of air temperature and precipitation (all 4 combinations).??**

$B_{h,r}$ is the $R \times K_R$ matrix of regression coefficients where $R$ is the number of unique reaches and $K_R$ is the number of regression coefficients that vary randomly by reach within HUC8. We assumed prior distributions of

$$B_{h,r} \sim \mathcal{N}(0, \sigma_{k_r}), \text{for } k_r = 1, ..., K_R,$$

$$\sigma_{r_0} = 100$$

$X_h$ is the matrix of parameters that vary by HUC8. We allowed for correlation among the effects of these HUC8 coefficients as described by Gelman and Hill [**???**].

$B_h$ is the $H \times K_H$ matrix of coefficients where $H$ is the number of HUC8 groups and $K_H$ is the number of parameters that vary by HUC8 including a constant term. In our model, $K_H = K_R$ and we assumed priors distributions of

$$B_h \sim \mathcal{N}(M_h, \Sigma_{B_h}), \text{for } h = 1, ..., H$$

where $M_h$ is a vector of length $K_H$ and $\Sigma_{B_h}$ is the $K_H \times K_H$ covariance matrix.

$$M_h \sim MVN(\mu^h_{1:K_h}, \sigma^h_{1:K_h})$$

$$\mu^h_1 = 0; \mu^h_{2:K_h} \sim \mathcal{N}(0, 100)$$

$$\Sigma_{B_h} \sim \text{Inv-Wishart}(diag(K_h), K_h + 1)$$

Similarly, we allowed the some effects of some parameters ($X_y$) to vary randomly by year with potential correlation among the coefficients. The intercept, day of the year ($day$), $day^2$, and $day^3$ all varied randomly with year such that $K_y = 4$. We assumed prior distributions of

$$B_y \sim \mathcal{N}(M_y, \Sigma_{B_y}), \text{for } y = 1, ..., Y$$

where $M_y$ is a vector of length $K_Y$ and $\Sigma_{B_y}$ represents the $K_Y \times K_Y$ covariance matrix.

$$M_y \sim MVN(\mu^y_{1:K_y}, \sigma^y_{1:K_y})$$

$$\mu_1^y = 0; \mu_{2:K_y}^y \sim \mathcal{N}(0, 100)$$

$$\Sigma_{B_y} \sim \text{Inv-Wishart}(diag(K_y), K_y + 1)$$

To estimate all the parameters and their uncertainties, we used a Bayesian analysis with a Gibbs sampler implemented in JAGS (ref) through R (ref) using the rjags package (ref). This approach was beneficial for hierarchical model flexibility and tractability for large datasets. We used vague priors for all parameters so all inferences would be based on the data.

**info on length of burn-in and sampling iterations and thinning**

## Model validation

To validate our model, we held out 10% stream reaches at random. We also held out 10% of remaining reach-year combinations with observed temperature data at random. Additionally, we excluded all 2010 data because it was an especially warm summer across the northeastern U.S. based on the mean summer daymet air temperatures. This approach was also used by [*DeWeber and Wagner*, 2014a] and helps to test the model's predictive ability under future warming conditions. This included reaches with no data located within subbasins with and without data and how well the model predicts in warm years without data, which will be important if using this model with future climate predictions. The most challenging validation scenario was at reaches within HUC8s without any data in a year without any data. In total, 26.4% of observations and 33.3% of reaches were held out for validation.

## Derived metrics

We use the meteorological data from daymet to predict daily temperatures for all stream reaches (<200 km$^2$) in the region for the synchronized period of the year from 1980-2013. The predictions are conditional on the specific random effects where available and receive the mean effect for reaches, HUC8s, and years when no data was collected. From these daily predictions, we derive a variety of metrics to characterize the stream thermal regime. These include mean (over the 34 years) July temperature, mean summer temperature, mean number of days per year above a thermal threshold (18, 20, 22 C used by default), frequency of years that the mean daily temperature exceeds each of these thresholds, and the maximum 7-day and 30-day moving means for each year and across all years. We also provide predictions of cold, cool, and warm waters specific to states with regulations related to these classifications.

## Climate change projections (future paper?)

# Results

To fit the model, we used 129,026 daily temperature observations from 627 stream reaches representing 1,051 reach-year combinations within 44 HUC8 subbasins between 1995 and 2013, excluding all records from 2010.

*Evaluation of MCMC convergence (visual and R-hat)* The iterations of the three MCMC chains converged on a single area of high posterior probability while exhibiting minimal autocorrelation, based on visual inspection of the iteration traceplots, partial vs. full density plots, autocorrelation plots. The potential scale reduction factors (PSRF, $\hat{R}$) for all parameters and the multivariate PSRF were $< 1.1$, further indicating good convergence of the MCMC chains [*Brooks and Gelman*, 1998].

*Coefficient estimates from the model*

Most variables and their interactions were significant with 95% Credible Intervals that did not overlap zero (Table 1). The only non-significant parameters were the interactions of air temperature and forest cover and air temperature and Impounded Area. Drainage area alone was not significant but it was significant in its interactions with all combinations of air temperature and precipitation (Table 1). Air temperature (1-day and 7-day) was the primary predictor of daily water temperature. The effect of air temperature was dampened by interactions with precipitation and drainage area (negative 3-way interactions; Table 1). There was also a large autocorrelation coefficient (AR1 = 0.77), indicating that if the other deterministic parameters in the model predicted temperature to be over- or under-estimated by 1 C yesterday, they will be similarly over- or under-estimated by 0.77 C today.

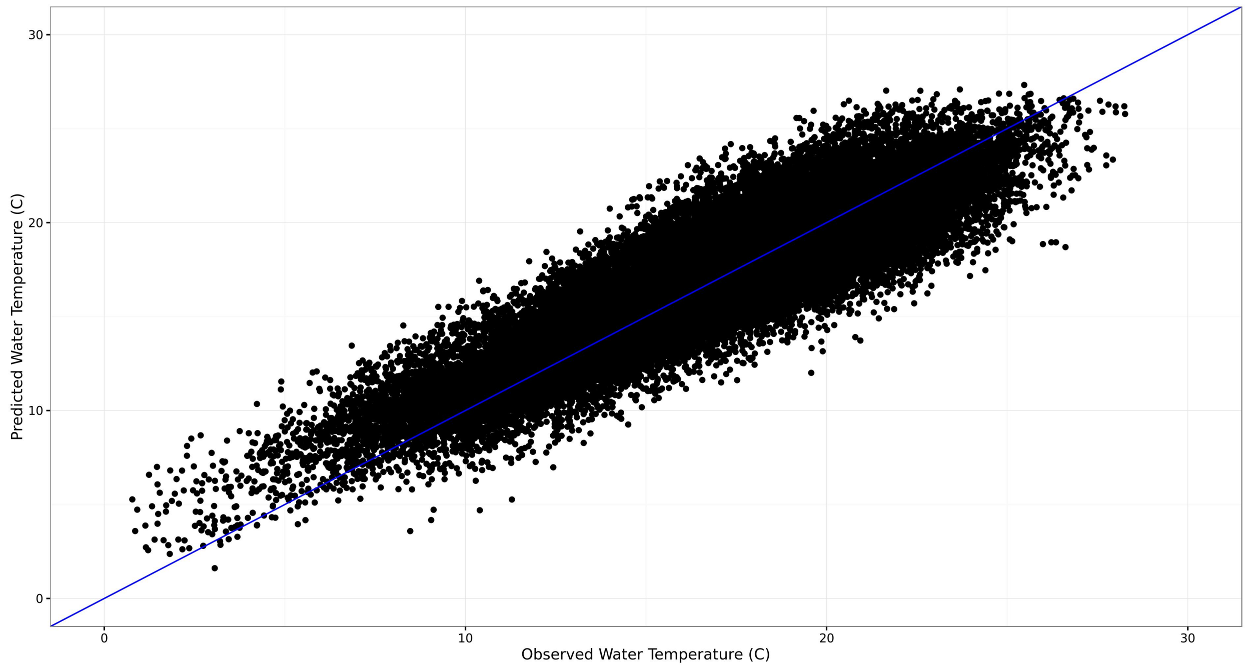
*Variability at the reach and huc scales*

There was much more unexplained random variation among sites than among HUC8, but the effects of air temperature on water temperature were only slightly more variable among sites compared with HUC8. There was very little random variability among years not explained by other parameters (Table 1).

*Evaluation of model fit and predictive power*

The overall Root Mean Squared Error (RMSE) was 0.58 C and the residuals were normally distributed and unbiased (exhibiting no visual heterogeneity), indicating that the model was a good estimate of the process generating the data. These predicted values are adjusted for residual error, but to understand how well the model predicts temperatures when the previous day's observed temperature is unknown it is better to use the predictions prior to adjusting with the residual AR1 term. The RMSE for the fitted data using unadjusted predictions was 0.89 C. All additional predictions and summaries use the unadjusted predictions to better understand the predictive abilities of the model.

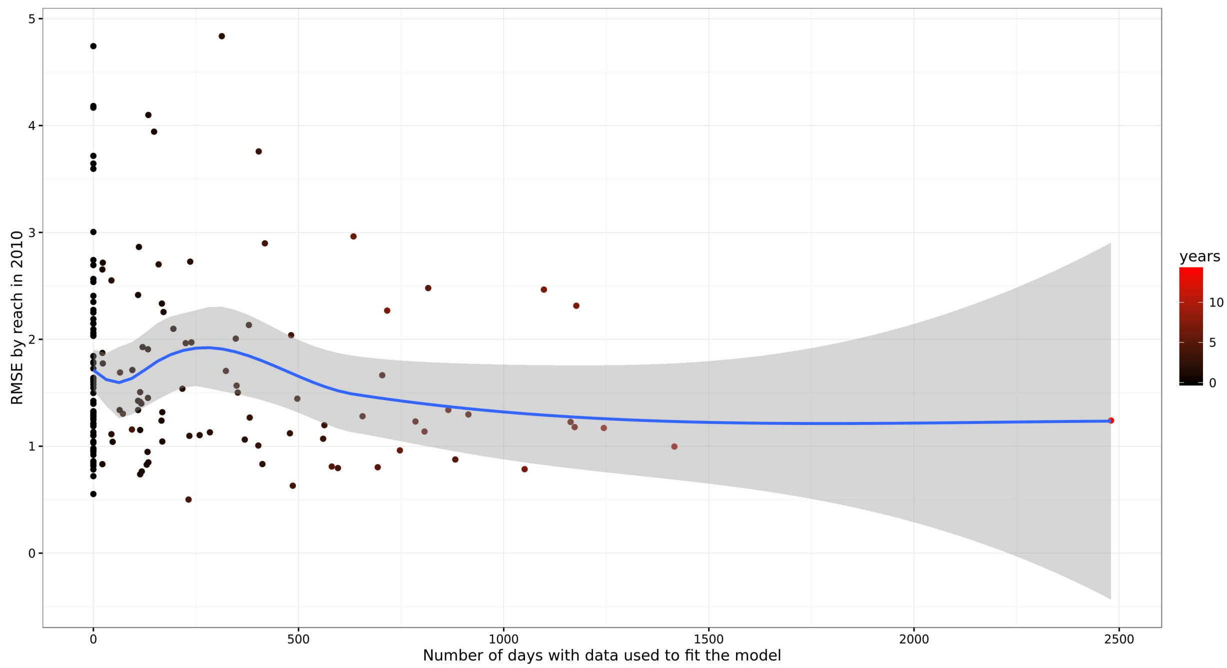
Specifically, to evaluate the spatial and temporal predictive power of our model, we used independent validation data consisting of 46,290 daily temperature observations from 313 stream reaches representing 383 reach-year combinations within 36 HUC8 subbasins between 1998 and 2013. The overall unadjusted RMSE for all validation data was 1.81 C. Similar to the fitted data, there was no bias in the predictions of the validation data, with the potential exception of slight over-prediction at very low temperatures and possible slight under-prediction at very high temperatures (figure - appendix?)




.

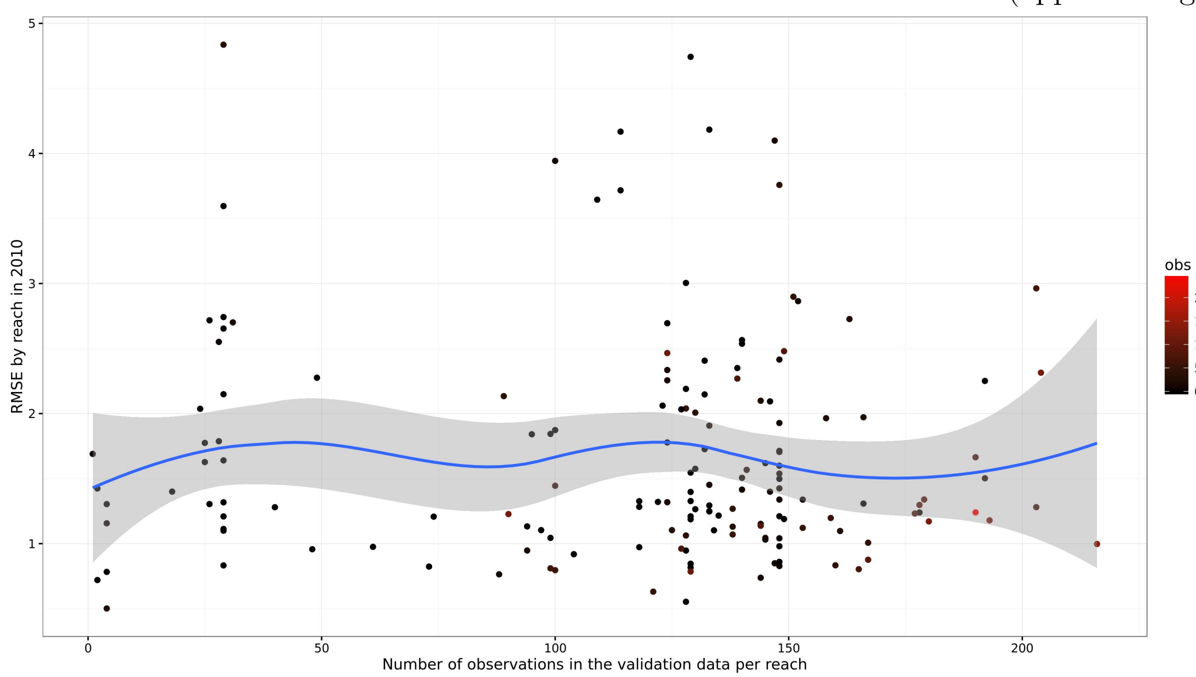
To assess predictive accuracy in warm years without data, we calculated the RMSE for all reaches in 2010 (excluded from model fitting) to be 1.85 C. The RMSE in 2010 for reaches that had data in other years used in the modeling fitting was 1.77 C, whereas reaches that had no data in other years had an overall RMSE of 1.95 C in 2010 (no information about the specific reach or year in a warm year). We performed 10,000 bootstrap samples on the 2010 RMSE by reach and calculated the 95% confidence interval to be 0.73 - 4.17 C. Inference from this interval assumes the reaches sampled in 2010 represent a random sample of the reaches of interest.

Interestingly, there appears to be only a slight improvement in RMSE with increases in the amount of data used in the model fitting or years of observed data (appendix figure)



.

This suggests that the predictive ability is largely based on reach-level effects not accounted for in the model. Similarly, there is no affect of the amount of data in the validation data for a reach on the RMSE estimate of that reach (appendix figure)



.

*Maps of each derived metric in appendix or special version of ICE*

# Discussion

Most aquatic organisms inhabiting streams are ectothermic and are therefore sensitive to changes in stream temperatures. Although air temperature can be used as a proxy for water temperature in small streams, there is considerable variability in the relationship between air and water temperatures. Additionally, land-use change (e.g. forest cover, impervious surfaces) and modifications to the stream network (e.g. undersized culverts, dams) influence water temperature differently than air temperature. It is also impossible to monitor water temperature across all streams; therefore, regional models are needed to predict stream temperatures across time and space accounting for differences in the landscape and land-use. Many fish biologists have focused on weekly, monthly, or summer-only models of stream temperature to relate warm conditions to trout distributions (refs). However, daily temperatures are useful because they can be used in observation processes when activity or detection is dependent on the current thermal conditions (refs) and they can be summarized into any derived metrics of interest. Depending on the species, life-stage, or management options, decision makers and biologists might be interested in different metrics such as degree days since an event (e.g. oviposition, hatching), frequency of thermal excursions, magnitude of excursions, mean summer temperature, or variability in temperature of different time frames, all of which can be derived from daily temperature predictions. Daily temperatures can also relate more closely to state agency regulations such as the frequency of daily temperatures over a threshold when classifying cold, cool, and warm streams for legal protection (MA Department of Environmental Protection, CALM Regulations, Gerry Szal *personal communication* - should probably find a real reference for this). Without knowing in advance all the potential uses of predicted stream temperatures, a daily model provides the flexibility to derive the values needed for particular decisions.

To accommodate these flexible needs, we developed a daily stream temperature model that takes advantage of diverse data sources to make predictions across a large region. Our model fit the data well as indicated by the RMSE < 1 C and had a good ability to predict daily stream temperatures across space and time. With regards to predicting temperatures in warm years without fitted data, such as 2010, the model predicted temperatures well even in reaches with no other data (RMSE = 1.95 C). The predictions were even better at reaches with data from other years (RMSE = 1.77 C), indicating that reach-specific data can improve predictions in future years but this improvement is not dramatic. The lack of dramatic improvement is likely due to multiple factors.

Some of the reach-level variability is probably accounted for by other nearby reaches within the same HUC8 (influence of HUC8 random effects). We did not have sufficient data from combinations of reaches, HUC8, and years to compare the RMSE for HUC8 with single versus multiple observed reaches, but based on similar levels of variability explained at the reach and HUC8 levels it is likely that having data from other reaches in a HUC8 improves the predictions for unmonitored reaches in the same HUC8. Therefore, on average, predictions will be worse at reaches within HUC8 with no data. There are also local conditions that vary in time to influence stream temperatures beyond what is included in the model. If the effect of these unmodeled covariates were constant in time, we would expect more of the variation to be captured by the random reach effects and therefore a larger difference in the RMSE in 2010 between reaches with other years of data and reaches with no observed data. Tim-varying ground-surface water interactions are likely a major source of the unexplained uncertainty in model predictions. Ground-surface water interactions are particularly complex in the northeastern U.S. and depend on dynamics of precipitation, temperature, snowmelt, local geology, land-use, and landscape physiognomy (refs - I'm just making this up based on physics and basic ecosystem processes). The amount of groundwater entering streams depends on these time-varying conditions but the temperature of the groundwater is also variable, depending on the residence time, depth, and past weather conditions (refs). How much the ground water affects the temperature of the stream water depends of the volume and temperature of each source of water. We do not currently have any landscape or environmental conditions that can predict these ground-surface water interactions over broad space in the northeastern U.S. However, work towards this is in progress and has been applied to other areas (refs: than and others), and any appropriate predictors could be added to our model without needed to change the overall structure of the model.

*interpretation of parameter estimates*

Of the parameters currently modeled, the current day's air temperature and the mean air temperature over the previous 7 days had the largest effect on daily stream water temperature. This is not surprising as we limited our analysis to small streams and to the synchronized period of the year when air and water temperature are most correlated. Past studies of small streams have also found air temperature to be the main predictor of stream temperature (refs) –compare specific coefficients and TS to other papers?–

*partitioning of variability*

However, the effects of air temperature and 7-day air temperature were not identical across space. These effects varied moderately across sites and HUC8 (Table 1), with similar variance for both temperature effects although the daily air temperature had a slightly larger mean effect (Table 1). Additionally, air temperature had significant 3-way interactions with precipitation and drainage area. We used 2-day precipitation x drainage area as in index of flow associated with storms and 30-day precipitation x drainage area as an index of baseflow in these small headwater streams (A. Rosner *personal communication*). Therefore, the negative 3-way interactions with air temperature are what we would expect, indicating that at high flows the effect of air temperature on water temperature is dampened. The effect size of these interactions are extremely small, likely in part because of the coarseness of using precipitation x drainage area as an index of flow and not accounting for local ground-surface water interactions.

Air temperature did not interact significantly with percent forest cover or impounded stream area. Alone forest cover had a significant, but small, negative effect on stream temperature during the synchronized period, whereas impounded area had a significant, moderately large positive effect on temperature (Table 1).

We did not include other predictors previously found to be important in statistical models because of correlation with existing covariates or a lack of variability in the potential predictor across the study area. For example, elevation can be a useful predictor of stream temperature (refs) but it lacks a specific mechanistic relationship and covaries strongly with air temperature across the region. Similarly, human development and impervious surfaces can affect stream temperature but in the northeastern U.S. these exhibited high negative correlation with forest cover and both variables could not be included in the model. As more data become available through our data portal http://db.ecosheds.org, it may be possible to separate the effects of forest cover and human development variables. Likewise, agricultural land-use can influence stream temperature or the effect of air temperature on stream temperature [**???**], but there were insufficient observations over a range of agriculture in our data to include it in the current model. Agriculture can be added to a future version of the model as we expand coverage to the mid-Atlantic region of the U.S. and as more data in added to our database. Shading can also influence local stream conditions but is challenging to quantify over large regions. As a step in this direction it would be possible to replace forest cover at the catchment or watershed scale with canopy cover within a riparian buffer area.

*derived metrics*

We used the daymet air temperature and precipitation to predict daily stream temperatures in each reach then calculated derived metrics of potential interest to biologists, managers, and policy makers.

*Disagreement (conflicting evidence? confused terminology) regarding the drivers of stream temperature*

*Benefits of our approach*

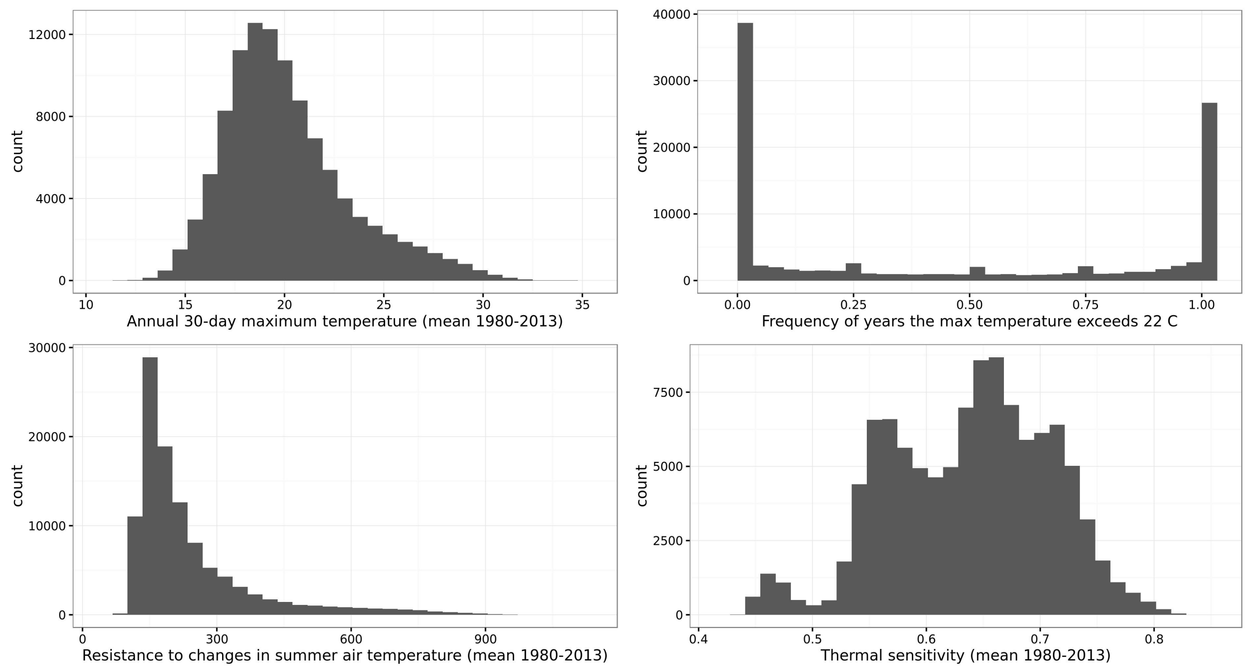



Figure 1: Figure #

relate it to the 6 challenges of statistical models the ben describes

model separates uncertainty in estimates and predictions from variability across space and time. The random reach, HUC, and year effects explicitly address spatial and temporal variability, allowing for more proper accounting of uncertainty.

lots of sensors because relatively cheap and easy to collect, but varying lengths of time at different reaches. Our model incorporates reaches with any length of time (a few days to decades). reaches will little data contribute less to the model but do provide some local and spatial information. The more data a location has the more informative so there is less shrinkage to the mean values. reaches with no data can be predicted based on covariate values and HUC-level random effects but do not get reach-specific coefficient effects.

*limitations*

ground-surface water interactions not included. However, if remotely sensed predictors could be developed, or exist in a particular region, they could easily be included as site-level predictors.

*future developments*

- groundwater
- within reach variability
- autoregressive temperature not just residuals
- detailed effects of impoundments (exponential decay with distance)
- spatial autocorrelation
- expand to larger spatial extent
- nonlinear relationships
- model winter
- adjust breakpoint sync function to adjust with different stream conditions, elevations, and locations
- dynamic model (effect of air temperature varies over time)


## Acknowledgments

Thanks to A. Rosner for thoughtful discussions related to the analysis and inference.

Groups who provided data


## Tables

Table 1. Regression summary table with coefficient estimates including the mean, standard deviation (SD), and 95% credible intervals (LCRI = 2.5%, UCRI = 97.5%).

| Parameter | Mean | SD | LCRI | UCRI |
|---|---|---|---|---|
| Intercept | 16.69 | 0.135 | 16.4182 | 16.949 |
| AirT | 1.91 | 0.022 | 1.8620 | 1.950 |

| Parameter | Mean | SD | LCRI | UCRI |
|---|---|---|---|---|
| 7-day AirT | 1.36 | 0.029 | 1.3015 | 1.417 |
| 2-day Precip | 0.06 | 0.002 | 0.0546 | 0.063 |
| 30-day Precip | 0.01 | 0.006 | 0.0005 | 0.026 |
| Drainage Area | 0.04 | 0.096 | -0.1452 | 0.232 |
| Impounded Area | 0.50 | 0.095 | 0.3181 | 0.691 |
| Forest Cover | -0.15 | 0.047 | -0.2455 | -0.059 |
| AirT x 2-day Precip | 0.02 | 0.002 | 0.0195 | 0.028 |
| AirT x 30-day Precip | -0.01 | 0.004 | -0.0224 | -0.007 |
| AirT x Drainage | -0.06 | 0.029 | -0.1170 | -0.006 |
| AirT x Impounded Area | 0.02 | 0.029 | -0.0345 | 0.077 |
| AirT x Forest | -0.02 | 0.015 | -0.0508 | 0.009 |
| 2-day Precip x Drainage | -0.04 | 0.002 | -0.0424 | -0.034 |
| 30-day Precip x Drainage | -0.06 | 0.006 | -0.0709 | -0.046 |
| AirT x 2-day Precip x Drainage | -0.01 | 0.002 | -0.0156 | -0.008 |
| AirT x 30-day Precip x Drainage | -0.01 | 0.004 | -0.0193 | -0.004 |
| AR1 | 0.77 | 0.002 | 0.7681 | 0.776 |

431 **Random effects:**

| Group | Coef | SD | Variance |
|---|---|---|---|
| Site | Intercept | 1.03 | 1.060 |
| | AirT | 0.29 | 0.083 |
| | 7-day AirT | 0.35 | 0.120 |
| HUC8 | Intercept | 0.59 | 0.345 |
| | AirT | 0.27 | 0.072 |
| | 7-day AirT | 0.26 | 0.066 |

| Group | Coef | SD | Variance |
|---|---|---|---|
| Year | Intercept | 0.28 | 0.076 |

**HUC8 coefficient correlations:**

| | Intercept | AirT | 7-day AirT |
|---|---|---|---|
| Intercept | | | |
| AirT | 0.64 | | |
| 7-day AirT | 0.338 | 0.234 | |

# Figures (do this as a separate file then merge the PDF)

Figure 1. Example of adding a figure.

# Literature Cited

Ahmadi-Nedushan, B., A. St-Hilaire, T. B. M. J. Ouarda, L. Bilodeau, E. Robichaud, N. Thiemonge, and B. Bobee (2007), Predicting river water temperatures using stochastic models : case study of the Moisie River (Quebec , Canada), *Hydrological Processes*, *34*, 21–34, doi:10.1002/hyp.

Al-Chokhachy, R., J. Alder, S. Hostetler, R. Gresswell, and B. Shepard (2013), Thermal controls of Yellowstone cutthroat trout and invasive fishes under climate change, *Global change biology*, *19*(10), 3069–81, doi:10.1111/gcb.12262.

Benyahya, L. et al. (2007), A review of statistical water temperature models, *Canadian Water Resources Journal*, *32*(3), 179–192.

Brooks, S. P. B., and A. G. Gelman (1998), General methods for monitoring convergence

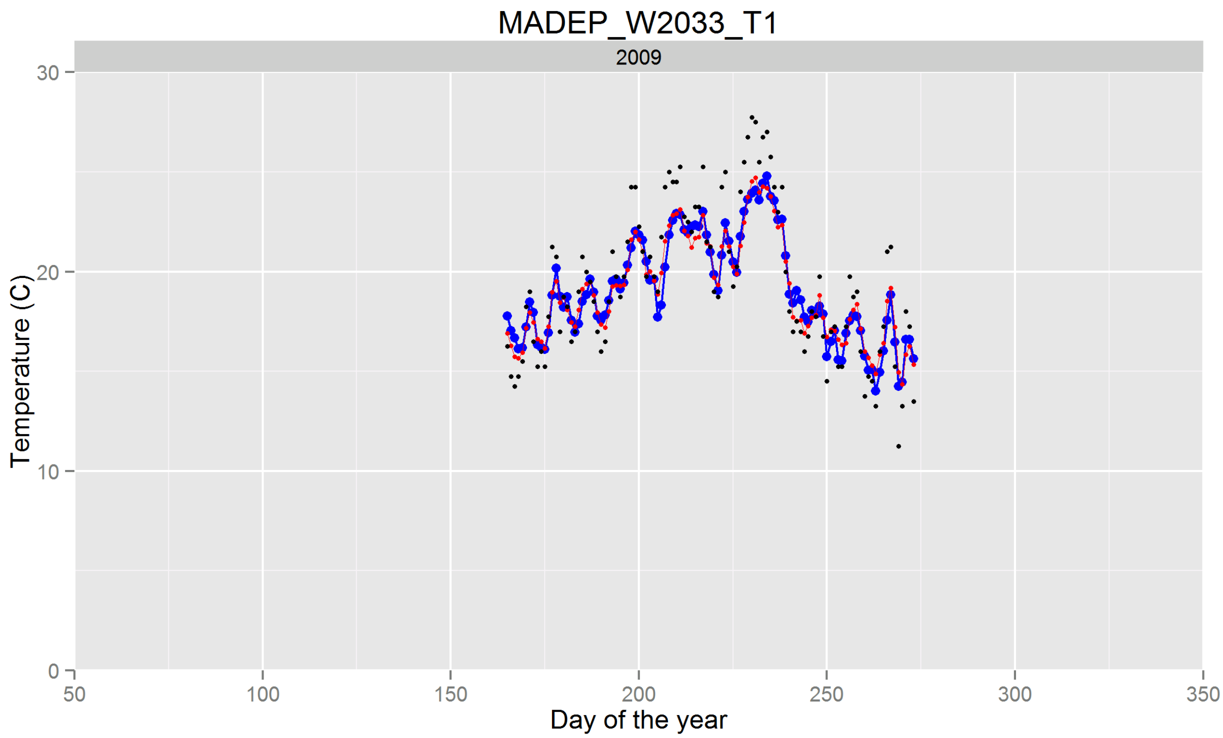



Figure 2:

of iterative simulations, *Journal of computational and graphical statistics*, *7*(4), 434–455, doi:10.2307/1390675.

Caissie, D. (2006), The thermal regime of rivers: a review, *Freshwater Biology*, *51*(8), 1389–1406, doi:10.1111/j.1365-2427.2006.01597.x.

Caissie, D., N. El-jabi, and M. G. Satish (2001), Modelling of maximum daily water temperatures in a small stream, *Journal of Hydrology*, *251*(2001), 14–28.

Chang, H., and M. Psaris (2013), Local landscape predictors of maximum stream temperature and thermal sensitivity in the Columbia River Basin, USA., *The Science of the total environment*, *461-462*, 587–600, doi:10.1016/j.scitotenv.2013.05.033.

DeWeber, J. T., and T. Wagner (2014a), A regional neural network ensemble for predicting mean daily river water temperature, *Journal of Hydrology*, *517*, 187–200,

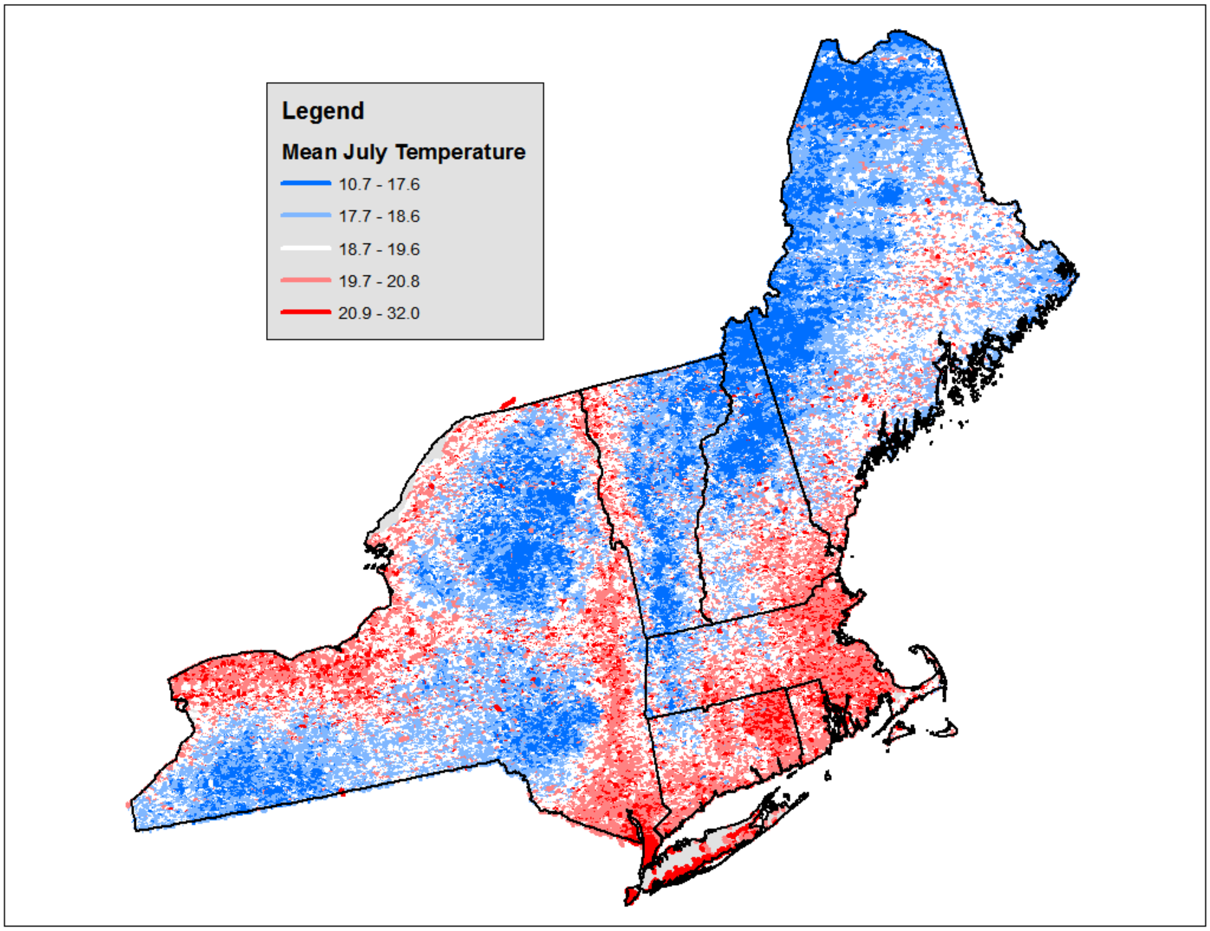



Figure 3: Predicted mean July stream temperatures from the period 1980-2013.

doi:10.1016/j.jhydrol.2014.05.035.

DeWeber, J. T., and T. Wagner (2014b), Predicting Brook Trout Occurrence in Stream Reaches throughout their Native Range in the Eastern United States, *Transactions of the American Fisheries Society*, *144*(1), 11–24, doi:10.1080/00028487.2014.963256.

Isaak, D. J. D., C. H. Luce, B. E. B. 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., *Ecological applications : a publication of the Ecological Society of America*, *20*(5), 1350–1371, doi:papers2://publication/uuid/8973E71F-5D23-47C7-A085-2AB46FFD8BF0.

Kanno, Y., J. Vokoun, and B. Letcher (2013), Paired stream-air temperature measurements reveal fine-scale thermal heterogeneity within headwater Brook Trout stream networks, *River Research and Applications*, *30*(6), 745–755, doi:10.1002/rra.

Kiraly, A., and I. Janosi (2002), Stochastic modeling of daily temperature fluctuations, *Physical Review E*, *65*(5), 1–6, doi:10.1103/PhysRevE.65.051102.

Letcher, B. H., D. J. Hocking, K. O. Neill, A. R. Whiteley, K. H. Nislow, and M. J. O. Donnell (2015), A robust hierarchical model of daily stream temperature using air-water temperature synchronization , autocorrelation , and, *Peer J Preprints*, *1578v1*, 1–39, doi:http://doi.org/10.7287/peerj.preprints.1578v1.

Mohseni, O., H. G. Stefan, and T. R. Erickson (1998), A nonlinear regression model for weekay stream temperatures, *Water Resources Research*, *34*(10), 2685–2692.

Peterson, E. E., D. M. Theobald, and J. M. Ver Hoef (2007), Geostatistical modelling on stream networks: developing valid covariance matrices based on hydrologic distance and stream flow, *Freshwater Biology*, *52*(2), 267–279, doi:10.1111/j.1365-2427.2006.01686.x.

Peterson, E. E., J. M. V. Hoef, and M. Jay (2010), A mixed-model moving-average approach to geostatistical modeling in stream networks, *Ecology*, *91*(3), 644–651.

Peterson, E. E. et al. (2013), Modelling dendritic ecological networks in space: an integrated network perspective., *Ecology letters*, *16*(5), 707–19, doi:10.1111/ele.12084.

Sivri, N., N. Kilic, and O. N. Ucan (2007), Estimation of stream temperature in Firtina Creek ( Rize-Turkiye ) using artificial neural network model, *Journal of Environmental Biology*, *28*(1), 67–72.

Sivri, N., H. K. Ozcan, O. N. Ucan, and O. Akincilar (2009), Estimation of Stream Temperature in Degirmendere River (Trabzon- Turkey) Using Artificial Neural Network Model, *Turkish Journal of Fisheries and Aquatic Sciences*, *9*, 145–150, doi:10.4194/trjfas.2009.0204.

Sura, P., M. Newman, and M. A. Alexander (2006), Daily to Decadal Sea Surface Temperature Variability Driven by State-Dependent Stochastic Heat Fluxes, *Journal of Physical Oceanography*, *36*, 1940–1958.

Ver Hoef, J. M., and E. E. Peterson (2010), A Moving Average Approach for Spatial Statistical Models of Stream Networks, *Journal of the American Statistical Association*, *105*(489), 6–18, doi:10.1198/jasa.2009.ap08248.

Webb, B., D. Hannah, R. D. Moore, L. E. Brown, and F. Nobilis (2008), Recent advances in stream and river temperature research, *Hydrological Processes*, *918*, 902–918, doi:10.1002/hyp.

Xu, C., B. H. Letcher, and K. H. Nislow (2010a), Context-specific influence of water temperature on brook trout growth rates in the field, *Freshwater Biology*, *55*(11), 2253–2264, doi:10.1111/j.1365-2427.2010.02430.x.

Xu, C. L., B. H. Letcher, and K. H. Nislow (2010b), Size-dependent survival of brook trout Salvelinus fontinalis in summer: effects of water temperature and stream flow, *Journal of Fish Biology*, *76*(10), 2342–2369, doi:10.1111/j.1095-8649.2010.02619.x.