

<sup>1</sup> **Journal Options:** Water Research, **Water Resources Research**, Freshwater Biology,  
<sup>2</sup> Journal of Hydrology, Ecohydrology, Journal of Environmental Quality, Hydrobiologia,  
<sup>3</sup> JAWRA

## <sup>4</sup> **A hierarchical model of daily stream temperature for 5 regional predictions**

<sup>6</sup> **Daniel J. Hocking, Ben Letcher, and Kyle O'Neil**

<sup>7</sup> \*Daniel J. Hocking (dhocking@usgs.gov), US Geological Survey, Conte Anadromous Fish  
<sup>8</sup> Research Center, Turners Falls, MA, USA

### <sup>9</sup> **Abstract**

<sup>10</sup> Set up the problem. Explain how you solve it. Tell what you find. Explain why it's the best  
<sup>11</sup> thing ever.

### <sup>12</sup> **Introduction**

<sup>13</sup> Temperature is a critical factor in regulating the physical, chemical, and biological properties  
<sup>14</sup> of streams. Warming stream temperatures decrease dissolved oxygen, decrease water den-  
<sup>15</sup> sity, and alter the circulation and stratification patterns of streams (refs). Biogeochemical  
<sup>16</sup> processes such as nitrogen and carbon cycling are also temperature dependent and affect  
<sup>17</sup> primary production, decomposition, and eutrophication (refs). Both physical properties and  
<sup>18</sup> biogeochemical processes influence the suitability for organisms living in and using the stream  
<sup>19</sup> habitat beyond just primary producers. Additionally, temperature can have direct effects  
<sup>20</sup> on the biota, especially poikilotherms such as invertebrates, amphibians, and fish [Xu *et al.*,

21 2010b, 2010a; *Al-Chokhachy et al.*, 2013; e.g., *Kanno et al.*, 2013]. Given commercial and  
22 recreational interests, there is a large body of literature describing the effects of tempera-  
23 ture on fish, particularly the negative effects of warming temperatures on cool-water fishes  
24 such as salmonids . Finally, stream temperature can even affect electricity, drinking water,  
25 and recreation (see van Vliet et al 2011). Therefore, understanding and predicting stream  
26 temperatures are important for a multitude of stakeholders.

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

40 Given the importance of temperature in aquatic systems, it is not surprising that there are  
41 a variety of models and approaches to understanding and predicting stream temperature.  
42 Stream temperature models are generally divided into three categories: deterministic (also  
43 called process-based or mechanistic), stochastic, and statistical [*Caissie*, 2006; *Benyahya et*  
44 *al.*, 2007; *Chang and Psaris*, 2013]. Deterministic models are based on heat transfer and  
45 are often modeled using energy budgets [*Caissie*, 2006; *Benyahya et al.*, 2007]. The models  
46 require large amounts of detailed information on the physical properties of the stream and

47 adjacent landscape as well as hydrology and meteorology. These models are useful for detailed  
48 re assessments and scenario testing. However, the data requirements preclude the models  
49 from being applied over large spatial extents.

50 Stochastic models attempt to combine pattern (seasonal and spatial trends) with the random  
51 deviations to describe and predict environmental data [Kiralay and Janosi, 2002; Sura *et al.*,  
52 2006; Chang and Psaris, 2013]. Stochastic models of stream temperature generally rely on  
53 relationships between air and water temperature then with random noise and an autoregressive  
54 correlation, often decomposed by seasonal and annual components. These models are mostly  
55 commonly used to model daily temperature fluctuations because of their ability to address  
56 autocorrelation and approximate the near-random variability in environmental data [Caissie  
57 *et al.*, 2001; Kiraly and Janosi, 2002; Ahmadi-Nedushan *et al.*, 2007]. A limitation is that  
58 the physical processes driving temperature fluctuations are not elucidated with these models.  
59 They are generally used to describe characteristics and patterns in a system and to forecast  
60 these patterns in the future [Kiralay and Janosi, 2002]. Additionally, stochastic models rely on  
61 continuous, often long, time series from a single or a few locations. Inference cannot be made  
62 to other locations without assuming that the patterns and random deviations are identical at  
63 those locations.

64 As with stochastic models, statistical models generally rely on correlative relationships  
65 between air and water temperatures, but also typically include a variety of other predictor  
66 variables such as basin, landscape, and land-use characteristics. Statistical models are often  
67 linear with normally distributed error and therefore used at weekly or monthly time steps  
68 to avoid problems with temporal autocorrelation at shorter time steps (e.g. daily, hourly,  
69 sub-hourly). Parametric, nonlinear regression models have been developed to provide more  
70 information regarding mechanisms than traditional statistical models without the detail  
71 of physical deterministic models [Mohseni *et al.*, 1998]. Researchers have also developed  
72 geospatial regression models that account for spatial autocorrelation within dendritic stream

73 networks [Isaak *et al.*, 2010; Peterson *et al.*, 2010, 2013]. However, due to the complexity of  
74 the covariance structure of network geostatistical models, they are best used for modeling  
75 single temperature values across space (e.g. summer maximum, July mean, etc.) rather than  
76 daily temperatures [Peterson *et al.*, 2007, 2010; Ver Hoef and Peterson, 2010]. Additionally,  
77 statistical machine learning techniques such as artificial neural networks have been used to  
78 model stream temperatures when unclear interactions, nonlinearities, and spatial relationships  
79 are of particular concern [Sivri *et al.*, 2007, 2009; DeWeber and Wagner, 2014b].

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

93 Letcher et al. [2015] outline six general challenges of statistical stream temperature models  
94 including accounting for 1) the non-linear relationship between air and water temperature at  
95 high and low air temperatures, 2) different relationships between air and water temperature  
96 in the spring and fall (hysteresis), 3) thermal inertia resulting in lagged responses of water  
97 temperature to changes in air temperature, 4) incomplete time series data and locations with  
98 large differences in the amount of available data, 5) spatial and temporal autocorrelation,

<sup>99</sup> and 6) important predictors of stream water temperature other than air temperature. They  
<sup>100</sup> developed a statistical model that addresses aspects of non-linear relationships, hysteresis,  
<sup>101</sup> thermal inertia, and spatial and temporal autocorrelation but their analysis was limited to a  
<sup>102</sup> single small network of streams with long time series [Letcher et al., 2015].

<sup>103</sup> We describe a novel statistical model of daily stream temperature that incorporates features  
<sup>104</sup> of stochastic models and extends the Letcher et al. [2015] framework to large geographic  
<sup>105</sup> areas. This model handles time series data of widely varying duration from many sites using  
<sup>106</sup> a hierarchical mixed model approach to account for autocorrelation at specific locations  
<sup>107</sup> within watersheds. It incorporates catchment, landscape, and meteorological covariates for  
<sup>108</sup> explanatory and predictive purposes. It includes an autoregressive function to account for  
<sup>109</sup> temporal autocorrelation in the time series, a challenge with other statistical models at fine  
<sup>110</sup> temporal resolution. Additionally, our hierarchical Bayesian approach readily allows for  
<sup>111</sup> complete accounting of uncertainty. We use the model to predict daily stream temperature  
<sup>112</sup> across the northeastern United States over a 34-year time record.

## <sup>113</sup> Methods

### <sup>114</sup> Study area

<sup>115</sup> Map of data locations: size = amount of data, color/shape = training-validation  
<sup>116</sup> - Kyle make? See deWeber 2014 for example - I will send Kyle list of series\_id (or  
<sup>117</sup> just lat,lon) of training data and of validation data

### <sup>118</sup> Water temperature data

<sup>119</sup> We gathered stream temperature data from state and federal agencies, individual academic  
<sup>120</sup> researchers, and non-governmental organizations (NGOs). The data were collected using  
<sup>121</sup> automated temperature loggers. The temporal frequency of recording ranged from every

122 5 minutes to once per hour. This data was consolidated in a PostgreSQL database linked  
123 to a web service at <http://www.db.ecosheds.org>. Data collectors can upload data at this  
124 website and choose whether to make the data publicly available or not. The raw data is  
125 stored in the database and users can flag problem values and time series. For our analysis, we  
126 performed some automated and visual QAQC on the sub-daily values, summarized to mean  
127 daily temperatures and performed additional QAQC on the daily values. The QAQC was  
128 intended to flag and remove values associated with logger malfunctions, out-of-water events  
129 (including first and last days when loggers were recording but not yet in streams), and days  
130 with incomplete data which would alter the daily mean. We developed an R (ref) package for  
131 analyzing stream temperature data from our database, including the QAQC functions which  
132 can be found at <https://github.com/Conte-Ecology/conteStreamTemperature>. The R scripts  
133 using these functions for our analysis are available at [https://github.com/Conte-Ecology/conteStreamTemperature\\_northeast](https://github.com/Conte-Ecology/conteStreamTemperature_northeast).

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

140 *Stream network delineation and landscape data*

141 Meteorological (, Climatic,) and landscape data - separate landscape if use cli-  
142 mate data for future projections

143 **Kyle** - Where the data came from with any necessary citations and any processing done  
144 (minimally descriptive) along with links to webpages, repos, README, and packages as  
145 appropriate.

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

<sup>147</sup> **Statistical model**

<sup>148</sup> Statistical models of stream temperature often rely on the close relationship between air  
<sup>149</sup> temperature and water temperature. However, this relationship breaks down during the  
<sup>150</sup> winter in temperature zones, particularly as streams freeze, thereby changing their thermal  
<sup>151</sup> and properties. Many researchers and managers are interested in the non-winter effects  
<sup>152</sup> of temperature. The winter period, when phase change and ice cover alter the air-water  
<sup>153</sup> relationship, differs in both time (annually) and space. We developed an index of air-water  
<sup>154</sup> synchrony ( $Index_{sync}$ ) so we can model the portion of the year that it not affected by freezing  
<sup>155</sup> properties. The index is the difference between air and observed water temperatures divided  
<sup>156</sup> by the water temperature plus 0.000001 to avoid division by zero.

<sup>157</sup> We calculate the  $Index_{sync}$  for each day of the year at each reach for each year with observed  
<sup>158</sup> data. We then calculate the 99.9% confidence interval of  $Index_{sync}$  for days between the 125  
<sup>159</sup> and 275 days of the year (05 May and 02 October). Then moving from the middle of the year  
<sup>160</sup> (day 180) to the beginning of the year, we searched for the first time when 10 consecutive days  
<sup>161</sup> were not within the 99.9% CI. This was selected as the spring breakpoint. Similarly moving  
<sup>162</sup> from the middle to the end of the year, the first event with fewer than 16 consecutive days  
<sup>163</sup> within the 99.9% CI was assigned as the autumn breakpoint. Independent breakpoints were  
<sup>164</sup> estimated for each reach-year combination. For reach-years with insufficient data to generate  
<sup>165</sup> continuous trends and confidence intervals, we used the mean break points across years for  
<sup>166</sup> that reach. If there was not sufficient local reach information, we used the mean breakpoints  
<sup>167</sup> from the smallest hydrologic unit the reach is nested in (i.e. check for mean from HUC12,  
<sup>168</sup> then HUC10, HUC8, etc.). More details regarding the identification of the synchronized  
<sup>169</sup> period can be found in Letcher et al. (*in review*). The portion of the year between the spring  
<sup>170</sup> and autumn breakpoints was used for modeling the non-winter, approximately ice-free stream  
<sup>171</sup> temperatures.

<sup>172</sup> We used a generalized linear mixed model to account for correlation in space (stream reach

173 nested within HUC8). This allowed us to incorporate short time series as well as long time  
174 series from different reaches and disjunct time series from the same reaches without risk of  
175 pseudoreplication (ref: Hurlbert). By limited stream drainage area to <200 km<sup>2</sup> and only  
176 modeling the synchronized period of the year, we were able to use a linear model, avoiding  
177 the non-linearities that occur at very high temperatures due to evaporative cooling and near  
178 0 C due to phase change (ref: mohseni).

179 We assumed stream temperature measurements were normally distributed following,

180 **need to decide how to handle naming subscripts vs. indexing subscripts and**  
181 **superscripts**

- 182 • maybe do naming as subscripts and indexing in bracketted subscripts
- 183 • drawback would be random vs. fixed subscripts still
- 184 • another alternative is to have different variable names for everything so don't reuse X,  
185 and B, mu, beta, or sigma
- 186 • This might be easier when I reduce the complexity of the year random effects

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

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

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

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

193 from the previous day

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

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

196  $X_0$  is the  $n \times K_0$  matrix of predictor values.  $B_0$  is the vector of  $K_0$  coefficients, where  $K_0$   
197 is the number of fixed effects parameters including the overall intercept. We used 15 fixed  
198 effect parameters including the overall intercept and interactions. These were 2-day total  
199 precipitation, 30-day cumulative precipitation, drainage area, upstream impounded area,  
200 percent forest cover within the catchment and upstream catchments and various two- and  
201 three-way interactions (Table 1?). We assumed the following distributions and vague priors  
202 for the fixed effects coefficients

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

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

$$\sigma_{k_0} = 100$$

203  $B_{h,r}$  is the  $R \times K_R$  matrix of regression coefficients where  $R$  is the number of unique reaches  
204 and  $K_R$  is the number of regression coefficients that vary randomly by reach within HUC8.  
205 The effects of daily air temperature and mean air temperature over the previous 7 days varied  
206 randomly with reach and HUC8 (Table 1). We assumed prior distributions of

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

$$\sigma_{r_0} = 100$$

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

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

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

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

$$M_h \sim MVN(\mu_{1:K_h}^h, \sigma_{1:K_h}^h)$$

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

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

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

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

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

$$M_y \sim MVN(\mu_{1:K_y}^y, \sigma_{1:K_y}^y)$$

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

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

217 To estimate all the parameters and their uncertainties, we used a Bayesian analysis with a  
218 Gibbs sampler implemented in JAGS (ref) through R (ref) using the rjags package (ref). This  
219 approach was beneficial for hierarchical model flexibility and tractability for large datasets.  
220 We used vague priors for all parameters so all inferences would be based on the data. We  
221 ran 13,000 iterations on each of three chains with independent random starting values. We  
222 discarded the first 10,000 iterations, then thinned; saving every third iteration for a total of  
223 3,000 iterations across 3 chains to use for inference.

## 224 Model validation

225 To validate our model, we held out 10% stream reaches at random. We also held out 10% of  
226 remaining reach-year combinations with observed temperature data at random. Additionally,  
227 we excluded all 2010 data because it was an especially warm summer across the northeastern  
228 U.S. based on the mean summer daymet air temperatures. This approach was also used by  
229 [DeWeber and Wagner, 2014a] and helps to assess the model's predictive ability under future

<sup>230</sup> warming conditions. This included reaches with no data located within subbasins with and  
<sup>231</sup> without data, which will be important if using this model with future climate predictions.  
<sup>232</sup> The most challenging validation scenario was at reaches within HUC8s without any data in a  
<sup>233</sup> year without any data. In total, 26.4% of observations and 33.3% of reaches were held out  
<sup>234</sup> for validation.

## <sup>235</sup> Derived metrics

<sup>236</sup> We use the meteorological data from daymet to predict daily temperatures for all stream  
<sup>237</sup> reaches ( $<200 \text{ km}^2$ ) in the region for the synchronized period of the year from 1980-2013.  
<sup>238</sup> The predictions are conditional on the specific random effects where available and receive  
<sup>239</sup> the mean effect for reaches, HUC8s, and years when no data was collected. From these  
<sup>240</sup> daily predictions, we derive a variety of metrics to characterize the stream thermal regime.  
<sup>241</sup> These include mean (over the 34 years) July temperature, mean summer temperature, mean  
<sup>242</sup> number of days per year above a thermal threshold (18, 20, 22 C used by default), frequency  
<sup>243</sup> of years that the mean daily temperature exceeds each of these thresholds, and the maximum  
<sup>244</sup> 7-day and 30-day moving means for each year and across all years. We also calculated the  
<sup>245</sup> resistance of water temperature to changes in air temperature during peak air temperature  
<sup>246</sup> (summer) based on the cumulative difference between the daily temperatures. Finally, we  
<sup>247</sup> assess the thermal sensitivity for each stream reach as the change in daily stream temperature  
<sup>248</sup> per 1 C change in daily air temperature. This is essentially the reach-specific air temperature  
<sup>249</sup> coefficient converted back to the original scale from the standardized scale.

250 Climate change projections (future paper?)

## 251 Results

252 To fit the model, we used 129,026 daily temperature observations from 627 stream reaches  
253 representing 1,051 reach-year combinations within 44 HUC8 subbasins between 1995 and  
254 2013, excluding all records from 2010.

255 *Evaluation of MCMC convergence (visual and R-hat)*

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

261 *Coefficient estimates from the model*

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

272 *Variability at the reach and huc scales*

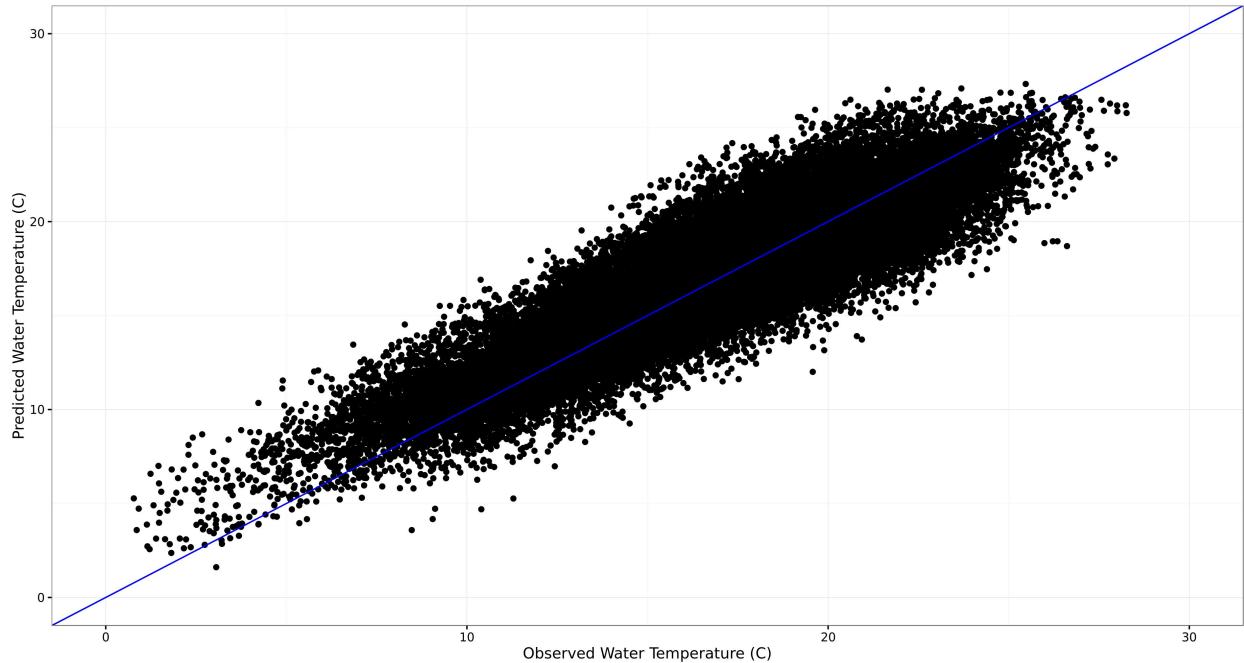
273 There was much more unexplained random variation among sites than among HUC8, but the

<sup>274</sup> effects of air temperature on water temperature were only slightly more variable among sites  
<sup>275</sup> compared with HUC8. There was very little random variability among years not explained  
<sup>276</sup> by other parameters (Table 1).

<sup>277</sup> *Evaluation of model fit and predictive power*

<sup>278</sup> The overall Root Mean Squared Error (RMSE) was 0.58 C and the residuals were normally  
<sup>279</sup> distributed and unbiased (exhibiting no visual heterogeneity), indicating that the model was  
<sup>280</sup> a good approximation of the process generating the data. These predicted values are adjusted  
<sup>281</sup> for residual error, but to understand how well the model predicts temperatures when the  
<sup>282</sup> previous day's observed temperature is unknown it is better to use the predictions prior  
<sup>283</sup> to adjusting with the residual AR1 term. The RMSE for the fitted data using unadjusted  
<sup>284</sup> predictions was 0.89 C. All additional predictions and summaries use the unadjusted values  
<sup>285</sup> to better understand the predictive abilities of the model.

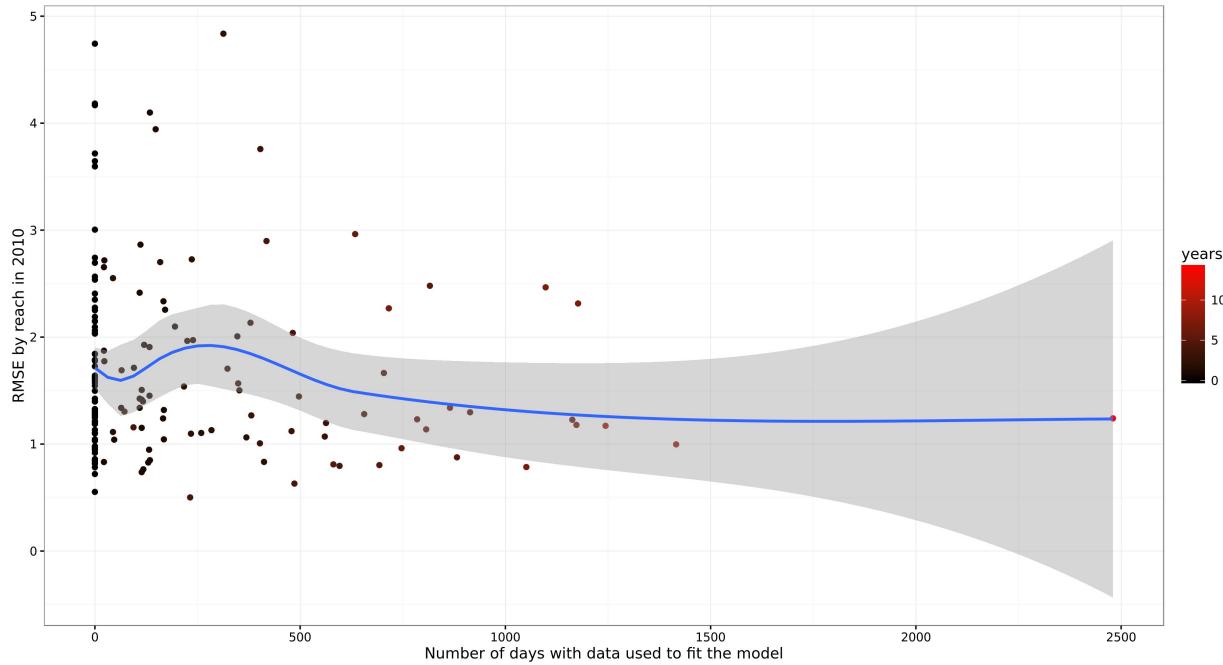
<sup>286</sup> Specifically, to evaluate the spatial and temporal predictive power of our model, we  
<sup>287</sup> used independent validation data consisting of 46,290 daily temperature observations  
<sup>288</sup> from 313 stream reaches representing 383 reach-year combinations within 36 HUC8  
<sup>289</sup> subbasins between 1998 and 2013. The overall unadjusted RMSE for all validation  
<sup>290</sup> data was 1.81 C. Similar to the fitted data, there was no bias in the predictions of the  
<sup>291</sup> validation data, with the potential exception of slight over-prediction at very low tempera-  
<sup>292</sup> tures and possible slight under-prediction at very high temperatures (figure - appendix?).



293

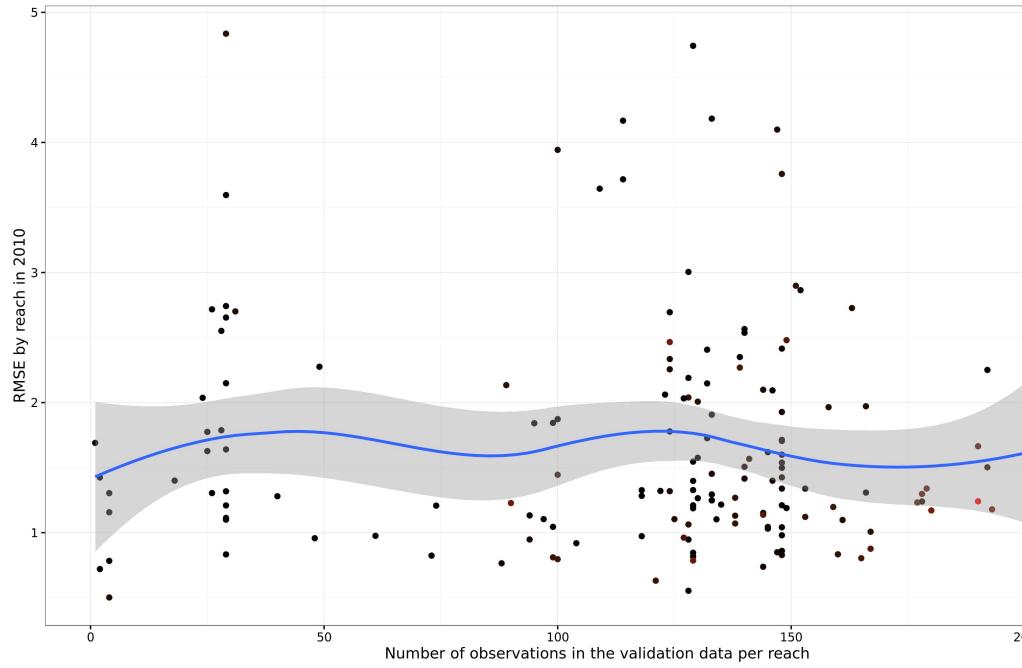
294 To assess predictive accuracy in warm years without data, we calculated the RMSE for all  
 295 reaches in 2010 (excluded from model fitting) to be 1.85 C. The RMSE in 2010 for reaches  
 296 that had data in other years used in the modeling fitting was 1.77 C, whereas reaches that  
 297 had no data in other years had an overall RMSE of 1.95 C in 2010 (no information about the  
 298 specific reach or year in a warm year).

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



301

302 Similarly, there is no affect of the amount of validation data for a reach on the RMSE estimate



303 of that reach (appendix figure).

## 304 Discussion

305 Most aquatic organisms inhabiting streams are ectothermic and are therefore sensitive to  
 306 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

<sup>334</sup> with data from other years ( $\text{RMSE} = 1.77 \text{ C}$ ), indicating that reach-specific data can improve  
<sup>335</sup> predictions in future years but this improvement is not dramatic. The lack of dramatic  
<sup>336</sup> improvement is likely due to multiple factors.

<sup>337</sup> Some of the reach-level variability is probably accounted for by other nearby reaches within  
<sup>338</sup> the same HUC8 (influence of HUC8 random effects). We did not have sufficient data from  
<sup>339</sup> combinations of reaches, HUC8, and years to compare the RMSE for HUC8 with single versus  
<sup>340</sup> multiple observed reaches, but based on similar levels of variability explained at the reach  
<sup>341</sup> and HUC8 levels it is likely that having data from other reaches in a HUC8 improves the  
<sup>342</sup> predictions for unmonitored reaches in the same HUC8. Therefore, on average, predictions  
<sup>343</sup> will be worse at reaches within HUC8 with no data. There are also local conditions that  
<sup>344</sup> vary in time to influence stream temperatures beyond what is included in the model. If the  
<sup>345</sup> effect of these unmodeled covariates were constant in time, we would expect more of the  
<sup>346</sup> variation to be captured by the random reach effects and therefore a larger difference in the  
<sup>347</sup> RMSE in 2010 between reaches with other years of data and reaches with no observed data.

<sup>348</sup> Tim-varying ground-surface water interactions are likely a major source of the unexplained  
<sup>349</sup> uncertainty in model predictions. Ground-surface water interactions are particularly complex  
<sup>350</sup> in the northeastern U.S. and depend on dynamics of precipitation, temperature, snowmelt,  
<sup>351</sup> local geology, land-use, and landscape physiognomy (refs - I'm just making this up based  
<sup>352</sup> on physics and basic ecosystem processes). The amount of groundwater entering streams  
<sup>353</sup> depends on these time-varying conditions but the temperature of the groundwater is also  
<sup>354</sup> variable, depending on the residence time, depth, and past weather conditions (refs). How  
<sup>355</sup> much the ground water affects the temperature of the stream water depends of the volume and  
<sup>356</sup> temperature of each source of water. We do not currently have any landscape or environmental  
<sup>357</sup> conditions that can predict these ground-surface water interactions over broad space in the  
<sup>358</sup> northeastern U.S. However, work towards this is in progress and has been applied to other  
<sup>359</sup> areas (refs: than and others), and any appropriate predictors could be added to our model  
<sup>360</sup> without needed to change the overall structure of the model.

361 *interpretation of parameter estimates*

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

368 *partitioning of variability*

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

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

385 We did not include other predictors previously found to be important in statistical models  
386 because of correlation with existing covariates or a lack of variability in the potential predictor

387 across the study area. For example, elevation can be a useful predictor of stream temperature  
388 (refs) but it lacks a specific mechanistic relationship and covaries strongly with air temperature  
389 across the region. Similarly, human development and impervious surfaces can affect stream  
390 temperature but in the northeastern U.S. these exhibited high negative correlation with  
391 forest cover and both variables could not be included in the model. As more data become  
392 available through our data portal <http://db.ecosheds.org>, it may be possible to separate the  
393 effects of forest cover and human development variables. Likewise, agricultural land-use can  
394 influence stream temperature or the effect of air temperature on stream temperature [???],  
395 but there were insufficient observations over a range of agriculture in our data to include  
396 it in the current model. Agriculture can be added to a future version of the model as we  
397 expand coverage to the mid-Atlantic region of the U.S. and as more data are added to our  
398 database. Shading can also influence local stream conditions but is challenging to quantify  
399 over large regions. As a step in this direction it would be possible to replace forest cover  
400 at the catchment or watershed scale with canopy cover within a riparian buffer area. Both  
401 riparian and drainage-level forest cover could be included in the model if there were sufficient  
402 data and they were not overly correlated.

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

405 *Benefits of our approach*

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

407 *Letcher et al. [2015] outline six general challenges of statistical stream temperature models  
408 including accounting for 1) the non-linear relationship between air and water temperature at  
409 high and low air temperatures, 2) different relationships between air and water temperature  
410 in the spring and fall (hysteresis), 3) thermal inertia resulting in lagged responses of water  
411 temperature to changes in air temperature, 4) incomplete time series data and locations with  
412 large differences in the amount of available data, 5) spatial and temporal autocorrelation, and*

413 6) important predictors of stream water temperature other than air temperature.

414 Our model addresses a number

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

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

424 *limitations*

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

428 *future developments*

- 429 • groundwater
- 430 • within reach variability
- 431 • autoregressive temperature not just residuals
- 432 • detailed effects of impoundments (exponential decay with distance)
- 433 • spatial autocorrelation
- 434 • expand to larger spatial extent
- 435 • nonlinear relationships
- 436 • model winter
- 437 • adjust breakpoint sync function to adjust with different stream conditions, elevations,

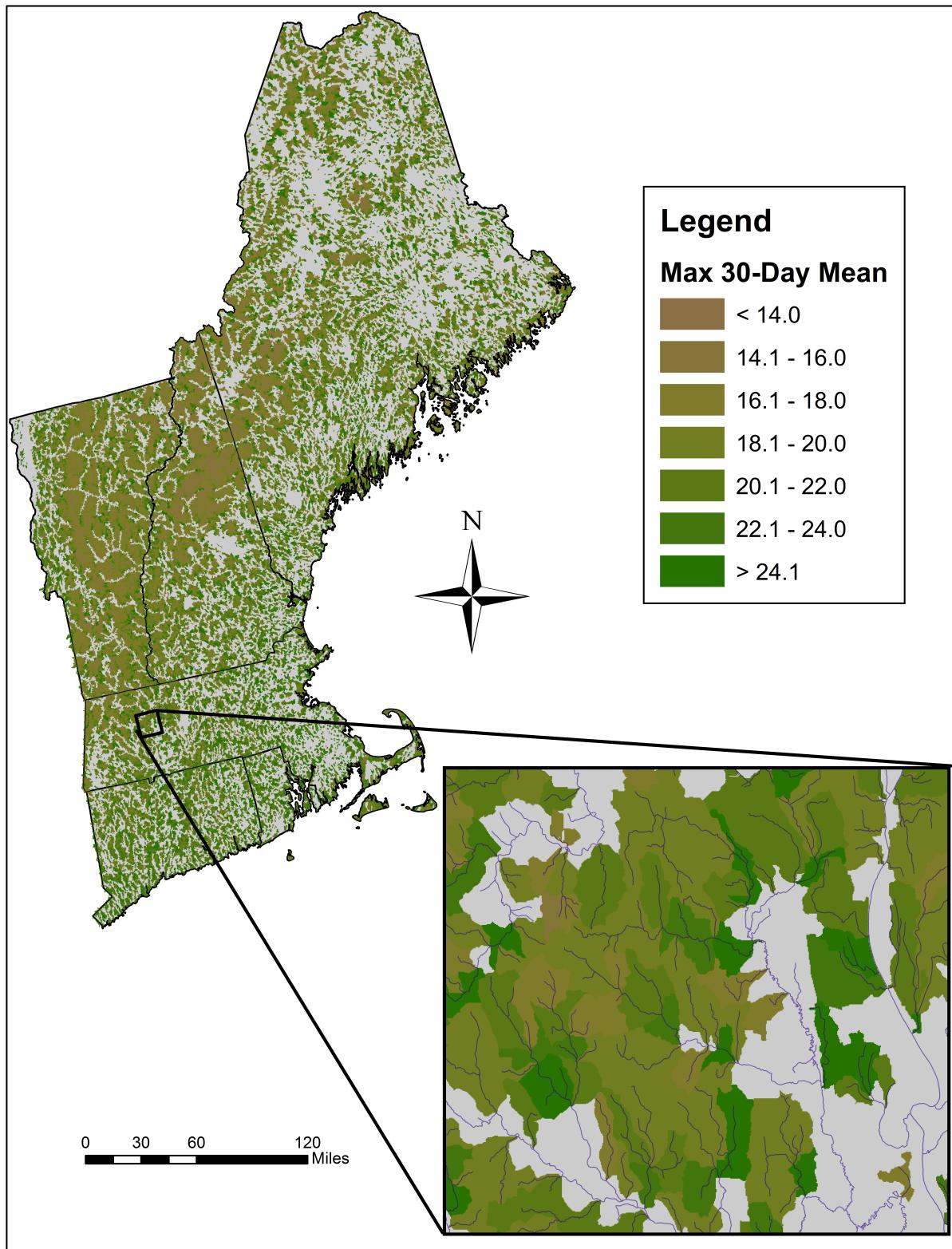
438 and locations

- 439 • dynamic model (effect of air temperature varies over time)

440 *derived metrics*

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

444 We generated maps of mean derived metrics from temperatures predicted over the daymet  
445 record (1980-2013). When scaled to view the entire region the patterns generally follow  
446 air temperature patterns with cooler temperatures at higher elevations and latitudes  
447 and warmer temperatures in urban, coastal, and lowland areas. An example of this can  
448 be seen on the annual 30-day maximum of the mean daily stream temperature map.  
449 However, when zoomed in to view individual catchments on the HUC8 or HUC10 scale,  
450 it is clear that there is considerable local variation in water temperatures (Figure #)

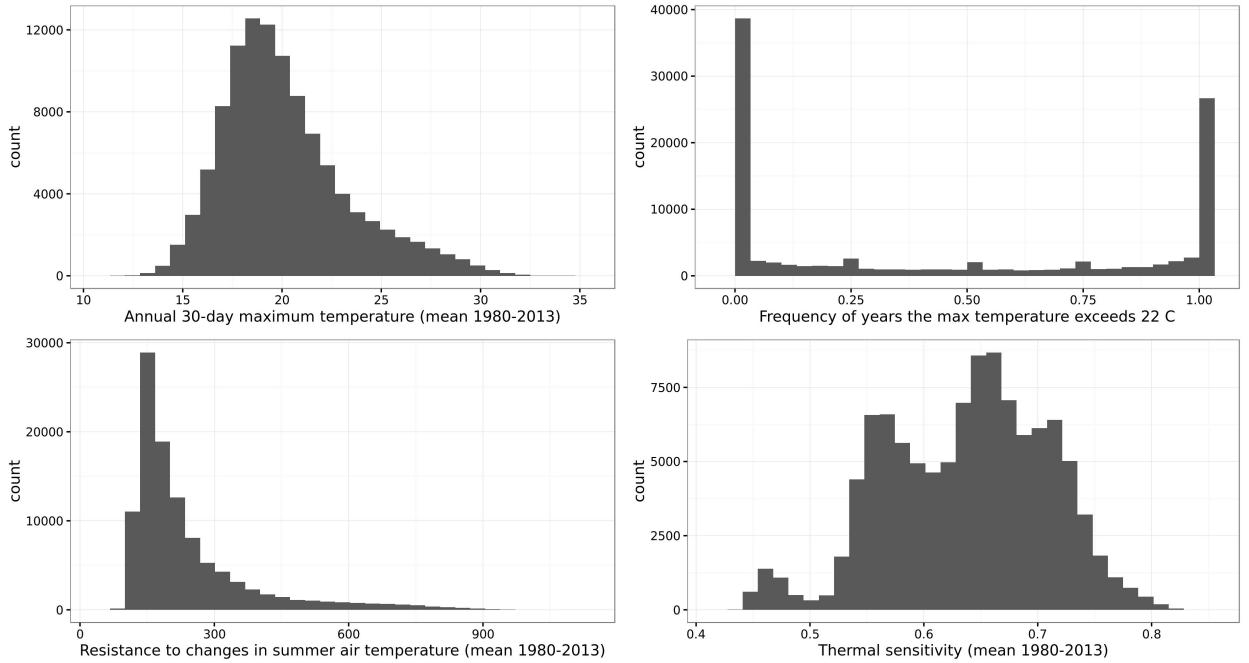


451

452 based on forest cover, drainage area, and local reach effects (unaccounted for local conditions

453 including ground-surface water interactions), as expected based on the model coefficients and  
454 past research [Kanno *et al.*, 2013]. In lieu of presenting small static maps, many of which  
455 would look somewhat similar at the regional scale, we added maps of the derived metrics to  
456 our web application which can be found at <http://ice.ecosheds.org/> *add special manuscript*  
457 *ice link*. Users can zoom to specific areas and view information about individual stream  
458 reaches and associated catchments. There is also the ability to filter to locate areas with  
459 specific conditions. Our main Interactive Catchment Explorer (ICE) for the northeastern and  
460 mid-Atlantic regions of the U.S. with information about the landscape conditions and Brook  
461 Trout occupancy in addition to stream temperatures can be found at <http://ice.ecosheds.org/>  
462 and will be regularly updated as new data become available. This is part of our web platform  
463 for Spatial Hydro-Ecological Decision Systems (SHEDS; <http://ecosheds.org/>) where we  
464 present visualizations linking datasets, statistical models, and decision support tools to help  
465 improve natural resource management decisions. Interested users can contribute, view, and  
466 download (if user-designated as publicly available) data at <http://db.ecosheds.org/>. As noted  
467 above, these data will be used to further improve model estimates and predictions, which  
468 will be presented in ICE.

469 Although many of the derived metrics relating to peak temperatures have relatively  
470 similar broad-scale spatial patterns, there are some metrics that quantify other aspects  
471 of the thermal regime. For example, we calculated the resistance of water temperature  
472 to changes in air temperature during peak air temperature (summer) based on the  
473 cumulative difference between the daily temperatures. The distribution of resistance  
474 values was much more right-skewed than the annual 30-day maximum temperature (Figure



475 #).

476 This metric is intended as a potential index of ground water influence on stream temperature.  
 477 Streams with larger resistance values would be expected to have higher ground water  
 478 influence because they would essentially be buffered from changes in air temperature during  
 479 the warmest part of the year (*could make figure to depict this for two extreme cases*). This  
 480 value could be adjusted for drainage area or flow since it is possible that larger streams  
 481 always fluctuate less and it could be divided by mean water temperature during the summer  
 482 to make it reflect the relative resistance. We anticipate future efforts to quantify the influence  
 483 of ground water in summer stream temperature and explore how well this metric is able to  
 484 predict those values. Similarly, thermal sensitivity (Figure # - histograms above) or the  
 485 size of the specific reach random effect could serve as indicators of ground water influence.  
 486 In particular, the specific reach slope of air temperature suggests that reaches with larger  
 487 coefficients are highly responsive to changes in air temperature (little ground water buffering)  
 488 and reaches with small coefficients are insensitive to changes in air temperature and therefore  
 489 likely to have significant ground water influence. These metrics are hypothesized to indicate  
 490 ground water influence but remain to be tested. Given the differences in the distributions of  
 491 these metrics (Figure # histograms), it is likely that some will be considerably more effective

492 as ground water indices than other metrics. A similar effort has recently shown promise in  
493 creating a ground water influence index from stream temperature data (ref: snyder, than and  
494 colleagues). These indices would currently only apply to reaches with observed data, so the  
495 next step would be to find landscape and geological parameters that could predict the best  
496 ground water index across the region.

## 497 Acknowledgments

498 Thanks to A. Rosner for thoughtful discussions related to the analysis and inference.  
499 J. Walker for database management and discussions.  
500 Groups who provided data

## 501 Tables

502 Table 1. Regression summary table with coefficient estimates including the mean, standard  
503 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
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

Parameter	Mean	SD	LCRI	UCRI
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

504 **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
Year	Intercept	0.28	0.076

505 **HUC8 coefficient correlations:**

	Intercept	AirT	7-day AirT
Intercept			
AirT		0.64	

	Intercept	AirT	7-day AirT
7-day AirT	0.338	0.234	

506 **Figures (do this as a separate file then merge the PDF because  
507 otherwise getting mixed in with citations during pandoc)**

508 Figure #. Map of the mean annual maximum 30-day mean stream temperature (mean  
509 temperature during the warmest 30-day period each year). The inset shows how much  
510 local variation there is that is not clearly visible on the regional map. Gray areas have no  
511 predictions, usually because they are in larger streams, outside the bounds of the data used  
512 in the model ( $>200 \text{ km}^2$  drainage area). Results are presented as catchments delineated  
513 based on the stream reaches because at this scale stream lines would blend together and not  
514 provide a smooth visual map surface - *not sure if I need to include this, maybe wait to see if*  
515 *reviewers say anything*

516 Figure 1. Example of adding a figure.

## 517 Literature Cited

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

522 Al-Chokhachy, R., J. Alder, S. Hostetler, R. Gresswell, and B. Shepard (2013), Thermal  
523 controls of Yellowstone cutthroat trout and invasive fishes under climate change, *Global*

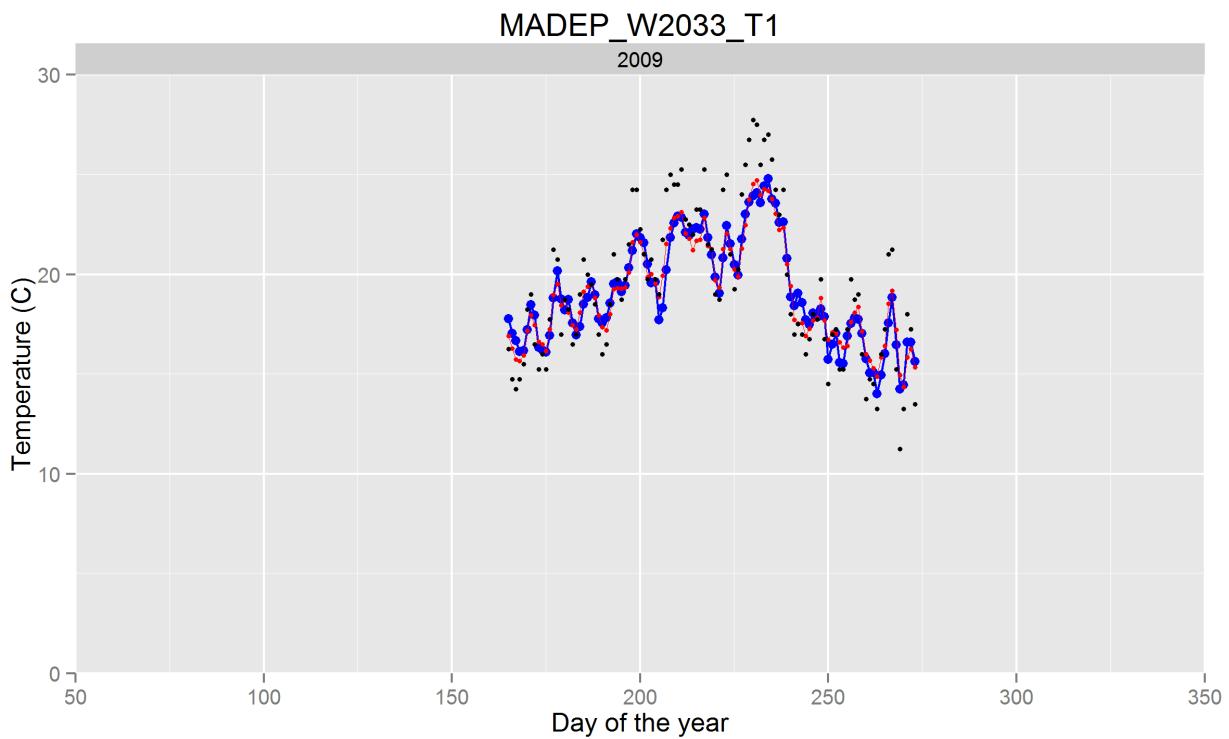


Figure 1:

<sup>524</sup> *change biology*, 19(10), 3069–81, doi:10.1111/gcb.12262.

<sup>525</sup> Benyahya, L. et al. (2007), A review of statistical water temperature models, *Canadian*  
<sup>526</sup> *Water Resources Journal*, 32(3), 179–192.

<sup>527</sup> Brooks, S. P. B., and A. G. Gelman (1998), General methods for monitoring convergence  
<sup>528</sup> of iterative simulations, *Journal of computational and graphical statistics*, 7(4), 434–455,  
<sup>529</sup> doi:10.2307/1390675.

<sup>530</sup> Caissie, D. (2006), The thermal regime of rivers: a review, *Freshwater Biology*, 51(8),  
<sup>531</sup> 1389–1406, doi:10.1111/j.1365-2427.2006.01597.x.

<sup>532</sup> Caissie, D., N. El-jabi, and M. G. Satish (2001), Modelling of maximum daily water tempera-  
<sup>533</sup> tures in a small stream, *Journal of Hydrology*, 251(2001), 14–28.

<sup>534</sup> Chang, H., and M. Psaris (2013), Local landscape predictors of maximum stream tempera-

- ture 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, 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. Hovan, 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 weekday stream temperatures, *Water Resources Research*, 34(10), 2685–2692.
- Peterson, E. E., D. M. Theobald, and J. M. Ver Hoef (2007), Geostatistical modelling on

- 560 stream networks: developing valid covariance matrices based on hydrologic distance and  
561 stream flow, *Freshwater Biology*, 52(2), 267–279, doi:10.1111/j.1365-2427.2006.01686.x.
- 562 Peterson, E. E., J. M. V. Hoef, and M. Jay (2010), A mixed-model moving-average approach  
563 to geostatistical modeling in stream networks, *Ecology*, 91(3), 644–651.
- 564 Peterson, E. E. et al. (2013), Modelling dendritic ecological networks in space: an integrated  
565 network perspective., *Ecology letters*, 16(5), 707–19, doi:10.1111/ele.12084.
- 566 Sivri, N., N. Kilic, and O. N. Ucan (2007), Estimation of stream temperature in Firtina Creek  
567 ( Rize-Turkiye ) using artificial neural network model, *Journal of Environmental Biology*,  
568 28(1), 67–72.
- 569 Sivri, N., H. K. Ozcan, O. N. Ucan, and O. Akincilar (2009), Estimation of Stream Tem-  
570 perature in Degirmendere River (Trabzon- Turkey) Using Artificial Neural Network Model,  
571 *Turkish Journal of Fisheries and Aquatic Sciences*, 9, 145–150, doi:10.4194/trjfas.2009.0204.
- 572 Sura, P., M. Newman, and M. A. Alexander (2006), Daily to Decadal Sea Surface Temper-  
573 ature Variability Driven by State-Dependent Stochastic Heat Fluxes, *Journal of Physical*  
574 *Oceanography*, 36, 1940–1958.
- 575 Ver Hoef, J. M., and E. E. Peterson (2010), A Moving Average Approach for Spatial Statistical  
576 Models of Stream Networks, *Journal of the American Statistical Association*, 105(489), 6–18,  
577 doi:10.1198/jasa.2009.ap08248.
- 578 Webb, B., D. Hannah, R. D. Moore, L. E. Brown, and F. Nobilis (2008), Recent advances in  
579 stream and river temperature research, *Hydrological Processes*, 918, 902–918, doi:10.1002/hyp.
- 580 Xu, C., B. H. Letcher, and K. H. Nislow (2010a), Context-specific influence of water tem-  
581 perature on brook trout growth rates in the field, *Freshwater Biology*, 55(11), 2253–2264,  
582 doi:10.1111/j.1365-2427.2010.02430.x.
- 583 Xu, C. L., B. H. Letcher, and K. H. Nislow (2010b), Size-dependent survival of brook trout

- <sup>584</sup> Salvelinus fontinalis in summer: effects of water temperature and stream flow, *Journal of*  
<sup>585</sup> *Fish Biology*, 76(10), 2342–2369, doi:10.1111/j.1095-8649.2010.02619.x.