Methods

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 so we can model the portion of the year that it not affected by freezing properties.

We used a generalized linear mixed model to….

correlation in space

incorporate short time series as well as long time series from different sites

incorporate disjunct time series from sites

We assumed stream temperature measurements were normally distributed following,

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

where $t_{s,h,d,y}$ is the observed stream water temperature at the site (s) 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 with the standard deviation (σ). The expected temperature follows a linear trend

$$\omega_{s,h,d,y} = X^0 B^0 + X_h^{huc} B_h^{huc} + X_{s,h}^{site} B_{s,h}^{site} + X_y^{year} B_y^{year}$$

but the expected temperature ($\mu_{s,h,d,y}$) is adjusted based on the residual error from the previous day

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

where δ_s is an autoregressive [AR(1)] coefficient that varies randomly by site and $\omega_{s,h,d,y}$ is the expected temperature before accounting for temporal autocorrelation in the error structure.

X_d^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. B^0 is distributed as

$$B^0 \sim \mathcal{N}(0, \sigma_{k_0}), \text{for } k_0 = 1, \ldots, K_0,$$

We used 10 fixed effect parameters including the overall intercept. These include latitude, longitude, upstream drainage area, percent forest cover, elevation, surficial coarseness classification, percent wetland area, upstream impounded area, and an interaction of drainage area and air temperature. We describe how we derived each of the parameters below.

The effects of air temperature on the day of observation (d) and mean air temperature over the previous 5 days varied randomly with site nested within HUC8, as did precipitation, the previous 30-day precipitation mean, and the interactions of air temperature and preciptation (all 4 combinations).

$B_{s,h}^{site}$ is the $S \times K_S$ matrix of regression coefficients where S is the number of unique sites and K_S is the number of regression coeffcients that vary randomly by site within HUC8.

X_h^{huc} is the matrix of parameters that vary by HUC8. We allowed for correlation among the effects of these HUC8 coefficients such that

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

as described by Gelman and Hill [-@Gelman2007], where H is the number of HUC8 groups and K_h is the number of paramaters that vary by HUC8 including a constant term. B_h^{huc} is the $H \times K_H$ matrix of coefficients and M_h is a vector of length K_h. In our model, $K_h = K_S$. \Sigma*{B*{h}} is the $K_h \times K_h$ covariance matrix.

Similarly, we allowed the some effects of some parameters (X_y^{year}) to vary randomly by year with potential correlation among the coefficients

$$B_y^{year} \sim \mathcal{N}(M_h, \Sigma_{B_y}), \text{for } y = 1, \ldots, Y$$

The intercept, day of the year (day), day^2, and day^3 all varied randomly with year so that $K_y = 4$. \Sigma*{B*{y}} represents the $K_y \times K_y$ covariance matrix.

priors