Title: Multi-scale controls on river temperature: a multivariate, autoregressive time series approach

Authors: Michael Vlah\* and Gordon Holtgrieve

Affiliation: School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98105

\*Corresponding author: vlahm13@gmail.com, 412-735-3706

**Abstract**

Temperature is among the most important determinants of riverine biodiversity and function. It is therefore a primary freshwater management concern, particularly where temperature-sensitive aquatic organisms including fish are of high ecological, recreational, or commercial value. River temperature in the Puget Sound watershed of the Northwestern U.S.A. is affected by snowmelt, hydropower, and a variety of other drivers at both regional and local scales, but little is known of their interactions. We used dynamic factor analysis, a multivariate time-series technique incorporating dimensionality reduction, to quantify the influence of these drivers on riverine thermal regimes, synthesizing long-term climate and fine-scale land cover data. We found that primarily rain-fed rivers undergo large seasonal temperature fluctuations that closely track air temperature – high coupling – while snow-fed rivers tend to be more weakly, and in some cases inversely, coupled with air temperature fluctuation. However, variation in coupling among snow-fed rivers is high and disproportionately influenced by artificial reservoirs, which appear to magnify the buffering effect of melting snow and glacial ice on thermal regimes in summer. Our results suggest snow-influenced rivers stand to see the largest changes in temperature regime under future climate warming.

**Introduction**

The biodiversity of a stream or river and the goods and services it provides are strongly influenced by the timing and magnitude of seasonal changes in water temperature (caissie2006thermal, olden\_hypoCold). Temperature structures the distribution of ectothermic taxa within river networks (vinson1998biodiversity, huryn2000life), including fish, which thermoregulate by occupying cool seeps and confluences (berman1991behavioural) or small, warm tributaries (peterson1996natural) depending on conditions needed for optimal feeding and growth. Modifications to temperature regimes, such as through dam construction, can radically alter species and functional diversity, sometimes yielding near or complete extirpation of native invertebrates and fish (lehmkuhl1974thermal, zhong1996environmental).

Temperature is also a chief consideration in the management of fisheries, as it affects species distribution (Boisneau2008), growth and reproduction (mccullough1999review), and migration timing (boscarino2007effects). In particular, In the Puget Sound watershed of the American Pacific Northwest, several salmonid species spawn, migrate, or emerge only within the bounds of a few degrees Celsius, and reproduce optimally under even greater temperature constraints (carter2005effects). Colder water is associated with earlier spawn timing in salmonids (beechie2008hierarchical), while warmer water is associated with a greater incidence of infectious disease (sanders1978relation).

Alongside water temperature, discharge (flow) regime is a major factor in shaping biotic composition, both directly (rorslett1989effects , munn1991benthic, bunn2002basic) and by altering the heat capacity of water; i.e. greater volume absorbs less heat per unit time (smith1972river). Conversely, temperature can affect discharge by determining the timing and magnitude of flow resulting from melting snow and ice (wulf2016differentiating). This interdependence implies a complex relationship between thermal regime and climate, which affects both temperature and discharge in rivers and streams.

River networks tend to be fractal in structure, and so are naturally governed by environmental processes at multiple spatial and temporal scales. Seasonal and inter-annual variation in water temperature and discharge are, in part, functions of regional climate, including solar radiation and precipitation (eldridge1967water). These large-scale drivers may in turn be mediated or supplemented by several aspects of watershed morphology at finer scales, including slope, elevation, and geology (poole2001ecological,lisi2013association). Adding to this picture, flow regimes are highly variable across streams and rivers, with timing of peak annual discharge sometimes varying by many months even within relatively small geographic areas (reidy2012hydrogeomorphic).

Taken together, variation and hierarchy complicate our ability to predict impacts, such as those from changing regional and global climate, on riverine systems. Lisi et al. (2015) found that low-elevation watersheds in southwestern Alaska are 5-8 times more sensitive to variation in summer air temperature than high-elevation streams, due to the cooling effect of melting snow. This implies that the temperature regime of any one river sits on a continuum of relative influence by regional versus within-watershed drivers. Effective management plans must therefore integrate a diversity of factors across space and time in order to determine which rivers and watersheds are likely to see consequential changes under projected climate and land use scenarios for the Pacific Northwest (mote2010future,radeloff2012economic), and identify potential mitigating actions. However, the understanding required to do so is limited by knowledge of relationships among temperature drivers at multiple scales.

We sought to determine whether rivers in the Puget Sound region vary in sensitivity to climatic forces, including air temperature (a proxy for the effect of solar radiation), precipitation, and snowmelt. We examined climatic effects on both water temperature and discharge throughout the year. Our second aim was to identify watershed features that correlate with variation in sensitivity, i.e. *coupling* between climatic forces and river temperature/discharge, and thus to provide a more nuanced basis for predicting impacts on aquatic biodiversity and fishery health. In the Puget Sound, watersheds vary with latitude and elevation (reidy2012hydrogeomorphic, mauger2015CIG), and can be classified broadly into three categories by flow source and hydrograph shape. Rain-dominated (RD) rivers receive little or no input from snowmelt, and thus peak in discharge (Q) during the rainy season, usually between October and February. Snow-dominated (SD) rivers instead see peak flow during spring snowmelt, often in April, May, or June. Between these extremes lies a third class of rain-and-snow-driven (RS) rivers, which have appreciable peaks at both times. We hypothesized that water temperature (Twater) would be most closely coupled with air temperature (Tair) in RD rivers (ward1985thermal, garner2014river, lisi2015watershed). We expected deviations from this relationship to correlate with cold-water influx from snow and ice melt (lisi2015watershed) and with factors affecting thermal sensitivity of water, including Q and watershed slope (van2013global). Finally, we expected Q to be most strongly coupled with Tair in the SD rivers, potentially serving to lessen the thermal sensitivity of water to changes in Tair.

**Methods**

Water and climate data

We investigated climate and landscape controls on Twater and Q, as separate response variables, from January 1978 to December 2015. Time series of water temperature by month were obtained for 24 river sites via the Washington Department of Ecology's River and Stream Water Quality Monitoring program (DoEwaterData). These sites represent 19 separate watersheds across 9 counties, and range from 4 to 775 m in elevation. River order (Strahler number) at mouth ranges from 5 to 8. For at least one site in each watershed, monthly discharge time series were available through the USGS National Water Information System database (USGSdischarge), either at the same location as one of the temperature monitoring sites, or within 30 km on the same major reach.

C:\Users\Mike\git\stream_nuts_DFA\manuscript\figures\map\puget_sound_labeled.tif

Figure 1. Site locations in relation to combined Washington State climate divisions 3 and 4 (colored topography), the region across which climate data were aggregated. See Appendix C for site information.

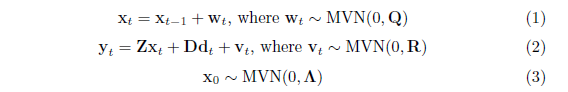
Potential climatic predictors of Twater and Q included mean and max Tair (°C), total precipitation (cm), and snowmelt (cm). All were incorporated as time series matching the span and temporal grain of the Twater and Q data. Tair and precipitation were available through the U.S. Climate Divisional Dataset, developed by the National Centers for Environmental Information (NCEI; climateData), and were aggregated across two Washington State climate divisions (Puget Sound Lowland and East Olympic/Cascade Foothills; colored topography in Figure 1). Snowmelt was derived from snow water equivalent (SWE) data from six SNOTEL sites (Bumping Ridge, Elbow Lake, Mount Crag, Park Creek Ridge, Stevens Pass, White Pass) listed by the USDA's Natural Resources Conservation Service (snowData). We calculated snowmelt for each site as the absolute value of negative differences in cumulative SWE from each time point to the next. The snowmelt time series was assigned zeros for any positive differences (accumulations).

Time series analysis

Response time series (Twater and Q) were modeled using dynamic factor analysis (DFA; (zuur2003estimating)), a multivariate technique partly analogous to principal component analysis in the time domain. In DFA, response time series are fit with a linear combination of predictors (which can exert unique effects on each response series), model-derived latent trends (which represent additional, unspecified but shared sources of variation), and random error. Each latent trend is generated as a random walk, and may explain variation in multiple response series, such that the number of latent, “shared” trends required to model each response is much smaller than the number of responses. In effect this reduces the dimensionality of predictor data required to model a multivariate system.

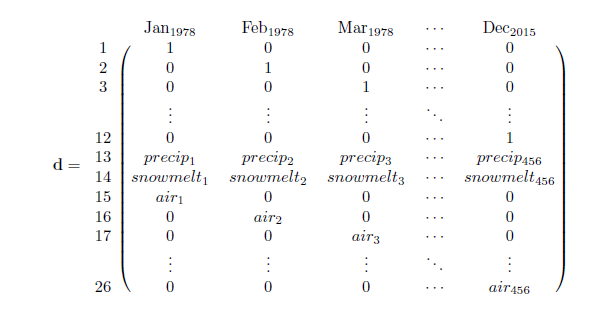
We chose DFA over a basic state space approach for two reasons. First, it provides advantages in computational efficiency, as a small number of shared trends often adequately capture variation across dozens of responses, and at much lower parameter cost (zuur2003dynamic). Second, in terms of identifying what drives the shared trends, having fewer of them allows for greater inferential parsimony. Being a multivariate technique, DFA also provides an advantage over univariate alternatives in that covariance structure among responses can be specified and compared. All models were fit using maximum likelihood estimation by automatic differentiation, with Template Model Builder software (kristensen2015tmb), which we called using package TMB in R (Rmanual,tmbPackage).

DFA takes the following form:



At time step *t*, the m\*1 vector of shared trends (**x**) is a function of **x** in the previous step, plus normal error (**w**; m\*1; Eq. 1). This is the definition of a random walk. The n\* 1 response vector (**y**) at time *t* is a function of the shared trends and their factor loadings (**Z**; n\* m), covariates (**d**; q\* 1) and their river-specific effects (**D**; n\* q), and a second normal error term (**v**; n\* 1; Eq. 2). **R** and **Q** are variance-covariance matrices of order m, and **Q** is set to identity for model identifiability (harvey1990forecasting). The initial state of the shared trend vector (**x0**) is multivariate-normally distributed with a mean of zero and a diagonal variance-covariance matrix with large variance (e.g. 5; Eq. 3). Response and predictor data were standardized to facilitate comparison of effect sizes and avoid error inflation.

Because we were interested in isolating the effects of climatic predictors on Twater and Q, we used a fixed factor to account for recurring seasonal variation not related to the predictors, with one factor level for each month. This factor was incorporated into the covariate matrix (**d**). Thus, the coefficient in **D** relating, say, precipitation (predictor) and Twater (response), represents the effect size of the former on the latter. In other words, it is the change in water temperature accompanying a unit change in precipitation across the whole time series. We call this relationship “coupling.” We were also interested in coupling by month for Tair, which required that it be arranged as twelve separate, monthly time series. Concretely,



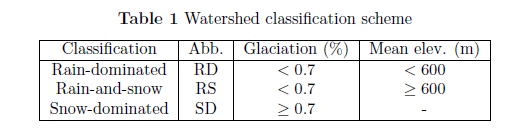
is the covariate matrix structure necessary to account for seasonal variation of unknown origin (rows 1-12), and the effects of precipitation (row 13) and snowmelt (row 14), while also yielding the effect of Tair by month (rows 15-26) on the response (**y**; Eq. 2). This is the covariate structure of the Twater model we used for subsequent analyses, not including those described in Figure 5d-e, and Appendix B. The same form was used for the Q model. We refer to this as the “full” model form, versus the “reduced” form, described in the following section.

Additional variation due to unknown factors manifests in the shared trends, which represent patterns in the response data that are shared among sites but are not attributable to the specified predictors (i.e. Tair, precipitation, and snowmelt). Finally, leftover residual variation is absorbed by error matrix **v**.

We used AIC to find the most parsimonious model while varying the number of shared trends and the error structure among responses. We included four alternative error structures (**R** in equation 2) to allow for multiple suites of unknown drivers affecting rivers, namely shared variance with zero covariance, individual variance with zero covariance, shared variance with shared covariance, and individual variance with individual covariance. Details on these structures and their implications can be found in Holmes et al. (holmes2012marss). Under AIC, negligible likelihood improvements can be inflated when multiplied by thousands of data points, undermining common rules of thumb for admitting additional parameters (burnham2003model). Thus, we had reason to doubt that the “most parsimonious” model according to AIC alone was any better than a much simpler alternative. To manage this, we required that each additional trend, covariate, or seasonal structure improve the median coefficient of determination (R2) by at least 1% in order to justify accepting its attendant complexity.

Landscape predictors and post-hoc regression

For post-hoc analyses, monitoring sites were separated into three classes based on relative areal coverage of perennial ice and snow (hereinafter “% glaciation”) and mean elevation across their watersheds. The three classes are loosely based on the classification schemes and language of Reidy Liermann et al. (reidy2012hydrogeomorphic) and the Climate Impacts Group at the University of Washington (mauger2015CIG), and are here delineated according to Table 1.



After model selection, climatic predictor effect sizes (**D**; Eq. 2) for each river were back-transformed to their original scales and regressed against multiple landscape predictors (Appendix C) in order to identify watershed-scale controls on coupling. Landscape predictors were compiled individually for each of the watersheds corresponding to our 24 river sites, using the EPA's StreamCat (stream-catchment) data library (hill2016stream) and the National Hydrography Dataset (NHDPlusV2; (mckay2012nhdplus)). Each site was mapped to an individual river reach, defined as a segment bounded on each end by a stream or river source, confluence, or mouth. The region contributing flow to this reach (its watershed) was then fetched, along with selected areal data, from the NHDPlusV2 database. Landscape attributes used as predictors were aggregated by watershed mean where applicable, and include elevation (m), total area (km2), soil permeability (cm hr-1), water table depth (cm), bedrock depth (cm), Base Flow Index (BFI; %), runoff (mm mo-1), percent perennial ice and snow coverage (National Land Cover Database [NLDC] 2006 and 2011 average), riparian population density (people km-2 within 100m of streams; 2010 census), riparian road density (km km-2; 2010 census), and percent riparian urban land (NLCD 2011). Monitoring site elevation (m) and presence of upstream dams (as full/partial/no damming of upstream mainstem and major tributaries) were also included. Finally, we calculated area above 1000 m (as % watershed area), mean slope (% rise), and mean aspect (degree from true north) by delineating and summarizing watersheds from a digital elevation model in ArcMap v. 10.4 (arcviewenvironmental).

An additional set of post-hoc regressions was performed using factor loadings on shared trends (**Z**; Eq.2) as dependent variables, with landscape predictors again as independent variables. Loadings represent the degree to which each river's temperature fluctuates with an unknown force driving the corresponding shared trend. A landscape feature that varies in proportion to these loadings may therefore be a mediator or correlate of the unknown driver, or the driver itself. To facilitate inference by way of the shared trends we used a “reduced” model form, based on three simplifications to the full model. We removed the monthly factor and the snowmelt predictor from the covariate matrix (**d**, rows 1-12 and 14) so that the trends would be free to express seasonal and elevational variation. Then, we limited the number of trends to between one and three, to avoid “trend specialization.” In other words, we optimized the trends for flexibility while concentrating their explanatory power. Additionally, we ordinated the landscape predictors via principal coordinates analysis (PCoA) in order to reduce dimensionality. Data constrained to irregular, restricted ranges were scaled to [0-1] and arcsine-squareroot transformed, along with all proportional data (The logit transform was avoided to prevent generation of infinite values.). All continuous data were then centered and scaled to unit variance before PCoA was performed. We used the Gower dissimilarity coefficient (Gower's distance) to account for association among both continuous and nominal variables (gowerDist).

**Results**

The response of river and stream water temperature to seasonal air temperature changes varied depending on whether rain, snow, or rain and snow dominated hydrology. Mean monthly temperature trends for the three river classes, aggregated across all 38 years of data, deviated by a minimum of 1.0°C in December, and a maximum of 3.9°C in July (Figure 2). SD rivers remained approximately two degrees colder than their RS counterparts through mid-late summer, and 3-4 degrees colder than RD throughout spring and summer. RD rivers were consistently warmest throughout the year. In January, RS reached a minimum of 4.4°C, and did not significantly differ from SD (Student's t: p<0.01, F=11.9). RD only attained a minimum of 5.6°C. RS reached a peak summer temperature of 16.9°C in July, while RS and SD followed in August with peak temperatures of 15.5 and 13.5°C, respectively.

Seasonal oscillations in Twater were smaller than those in Tair for all river classes. Tair dipped below Twater in autumn to a minimum of 3.2°C in December, and rose above RS and SD in March to an August maximum of 17.4°C. Tair did not overtake RD Twater until August, by which time the latter had begun to decline.

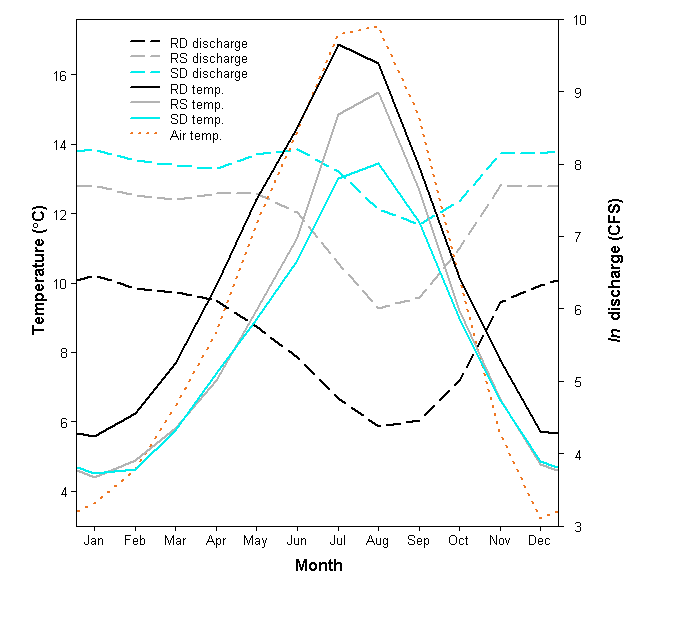
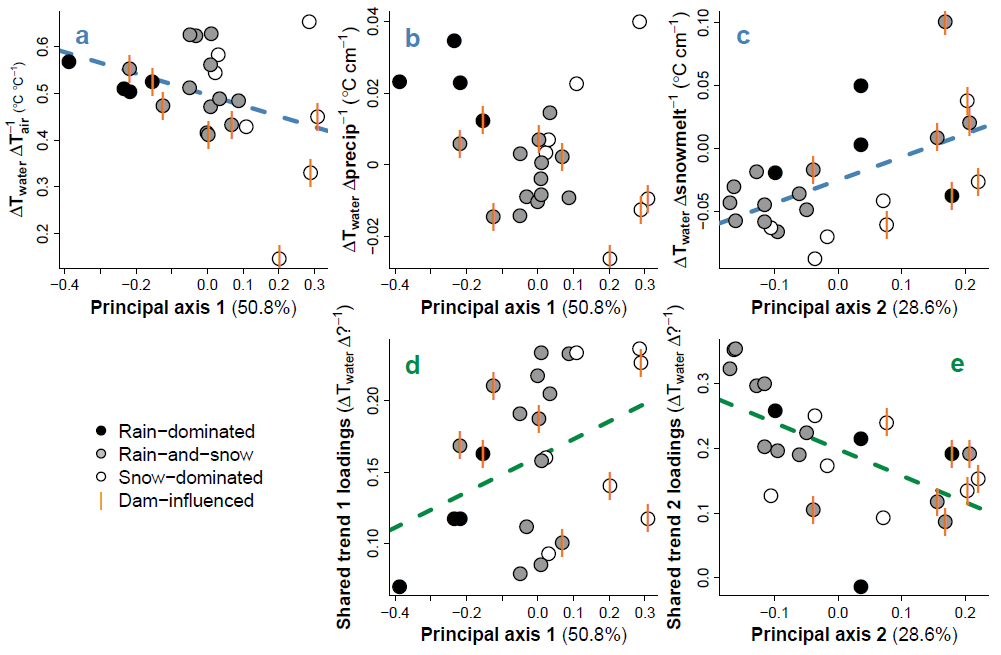


Figure 2. Monthly mean Twater and Q by river class, and regional Tair, from 1978 to 2015.

Mean absolute discharge for RD, RS, and SD across all 38 years of the time series was 583.7, 2766.9, and 5256.5 CFS, respectively. The discharge peak resulting from melting snow and ice, spanning approximately April to August, was most prominent and persisted longest in the SD rivers.



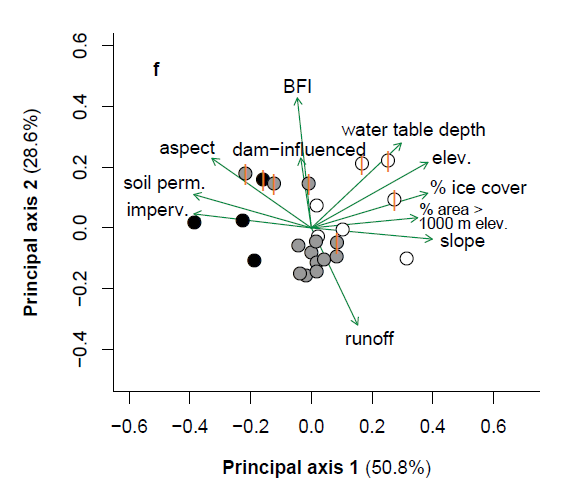


Figure 3 (a-c) Coupling between watershed elevation and climatic effects on Twater, obtained from full model fit. Coupling = effect size, based on proportionality at each time point. (d-e) Relationships between watershed features and factor loadings on shared trends, from constrained model fit. Regression lines indicate slopes significant at \alpha=0.1. (f) Ordination of landscape predictors by principal coordinates analysis. Length and direction of arrows are proportional to loading of landscape predictors onto each principal axis of their variation. Vertical bars indicate sites with upstream dams.

The full DFA fit, comprised of the effects of all three climate covariates and five shared trends, accounted for 96% of variation among response time series (mean R2 = 0.964, SD = 0.036). The effects of shared trends alone accounted for 10% of variation (mean R2 = 0.103, SD = 0.042). DFA results, aggregated across months and years for each site, revealed a trend toward reduced Tair → Twater coupling with increasing loading onto principal axis 1 (marginal significance: p = 0.08, mult. R2=0.13; Figure 3a). On average, a 1°C change in Tair corresponded to a 0.53 ± 0.03°C change in Twater at RD, a 0.51 ± 0.08°C change at RS, and a 0.45 ± 0.17°C change at SD sites. A similar trend was observed with respect to precip → Twater coupling when SD sites were excluded (p < 0.01, mult. R2 = 0.39), but not when they were included (p = 0.13, mult. R2 = 0.10; Figure 3b). Coupling between snowmelt and Twater increased with greater loading onto principal axis 2 (p < 0.01, mult. R2 = 0.30; Figure 3c). The most parsimonious model chosen via AIC and R2 is detailed in Appendix A.

The strongest examples of Tair → Twater and precip → Twater coupling were observed in the Duckabush River, while the weakest examples are from the Elwha River. Both rivers drain glaciers of the Olympic Mountain Range, and both are SD. Among SD rivers, those influenced by dams appear to couple less strongly with Tair and precip, but more so with snowmelt.

The reduced DFA fit accounted for 93% of variation among response time series (mean R2 = 0.934, SD = 0.053). The effects of shared trends alone accounted for 20% of variation (mean R2 = 0.199, SD = 0.059). Factor loadings from the reduced, two-trend model each correlated with one of the two principal axes of variation across landscape predictors, determined by PCoA (Figure 3f). The first principal axis was associated mostly with mean watershed slope, snow (% area > 1000 m) and ice, soil permeability, and other features that vary along elevational gradients, as well as mean elevation itself. Watershed scores along this axis correlated with loadings from one trend, with marginal significance (p=0.07, mult. R2=0.14; Figure 3d). The second principal axis was driven by runoff, base flow, and upstream dams, and correlated with the other trend's loadings (p<0.01, mult. R2=0.35; Figure 3e). Combined, the first two principal axes accounted for 79.4% of variation across landscape predictors.

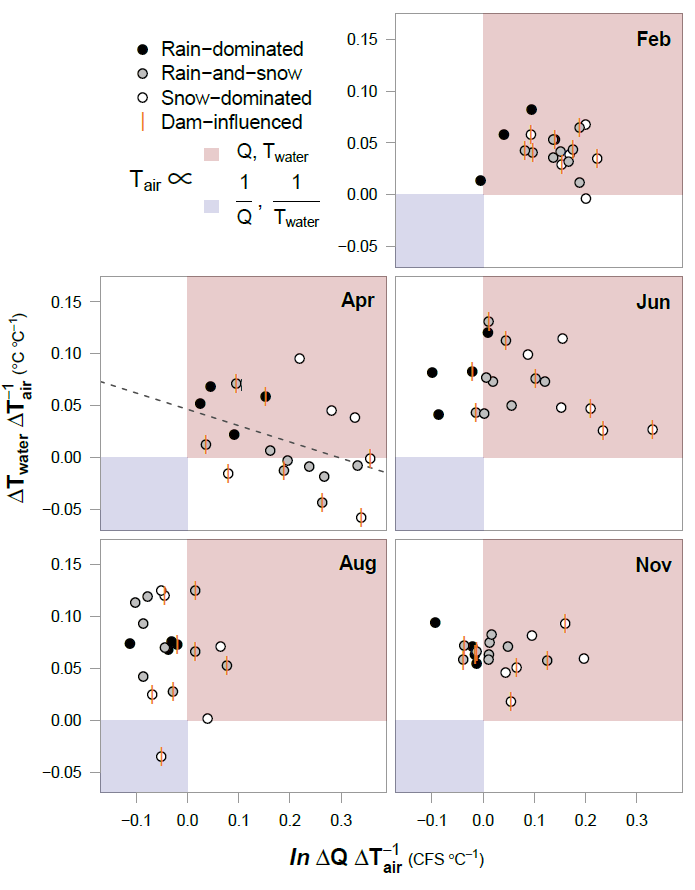


Figure 4. Relationships between Tair → Twater and Tair → Q coupling per month. The red quadrant designates proportionality between all three variables, the blue inverse proportionality between each response and Tair. Coupling = effect size, based on proportionality at each time point. Regression line indicates slope significant at alpha=0.05.

To examine possible sub-season interactions between Tair, Twater and Q, we performed an additional DFA with Q as the response. Tair was allowed to have unique monthly effects with respect to each response variable. These effects, taken together, can be conceptualized in relation to the four quadrants of the Cartesian coordinate system (increasing clockwise from upper right; Figure 4).

In mid-winter (exemplified by February), all river classes primarily occupy the first quadrant, signifying that Twater and Q both vary in proportion to Tair. In other words, warmer Februaries yield warmer water and higher discharge. RD shows the weakest Q response. By April, many RS and SD sites develop an inverse relationship between Tair and Twater, while RD sites change little from their winter state. There is an overall inverse relationship between Tair → Q and Tair → Twater coupling, with sites showing strong positive discharge responses to higher Tair showing correspondingly weak Twater responses. Early and late summer months (June and August) see a procession of most sites into the near fourth quadrant, with SD trailing. This quadrant signifies a general pattern of lower discharge and higher water T with increased air T. One stark exception is again the Elwha river, which occupies quadrant three. By autumn, RS and SD have begun progress back toward their winter states, led by SD. RD, meanwhile, remain essentially unmoved from summer.

Rivers influenced by dams do not appear to deviate appreciably from undammed systems in February, August, or November. However, SD rivers in April divide across the x-axis according to whether they are dammed. Those with dams exhibit inverse Tair → Twater coupling, while Tair fluctuates proportionately to Twater for those without. Similarly, in June, dammed SD rivers display stronger coupling between Tair and Q than those without dams.

**Discussion**

We did not detect any long-term trends in coupling between Tair and Twater. In other words, the sensitivity of Twater to changes in Tair did not vary consistently across 38 years of data (Appendix B) despite trends in Tair, precipitation, and several rivers’ individual water temperatures over the same period (Appendix D). Nonetheless, dynamic factor analysis did reveal variation among rivers and river classes in terms of Twater and Q sensitivity to climatic variation.

The effects of climate on Twater suggest that nearly all rivers included in our dataset were influenced strongly by air temperature, precipitation, and/or snowmelt across seasons (Figure 3a-c). At most monitoring sites, Twater closely tracked changes in Tair, on average responding to increases and decreases with proportional movements of 66% magnitude. However, some rivers only weakly tracked Tair, and patterns in the intensity of this coupling relate primarily to changing landscape features along an elevational gradient (Figure 3f).

A “buffering” effect (the inverse of coupling) of ice on river temperature can be seen in the yearly patterns of Twater relative to Tair (Figure 2). The aggregate hydrograph peaks due to snowmelt from April to June, at the same time that the trajectories of RS and SD (snow-influenced rivers) start to drop off relative to RD. After snowmelt begins to subside, RS and SD recover with a noticeable jump. For SD rivers, this buffering effect appears to remain into the summer months, buffering them from temperature rise when RS rivers instead approach the temperature of RD (Figure 4). In an extreme case, the Elwha River was actually cooler in August during those years in which air temperature was higher, probably due to increased runoff from Carrie and Eel glaciers. The buffering effect of ice on river temperature is therefore two-fold, acting first on all snowmelt-influenced rivers through a cold-water pulse in spring, and then on a subset of those rivers throughout summer and autumn, by way of glacial runoff. For RD rivers, which receive little to no input from snow or ice, summer temperature is entirely dictated by that of the surrounding air, and any rain falling through it.

Temperature buffering during warmer parts of the year by snow and ice appears to be enhanced by the action of dams. Eight sites on five rivers included in this study are (or were until 2014, in the case of the Elwha River) interrupted by dams or embankments, all of which release stored water from the bases of their reservoirs (hypolimnetic release). At 33 m, even the shallowest of these reservoirs is deep enough to stratify in summer, meaning released water would be delivered from a relatively cold hypolimnion (olden\_hypoCold). This certainly would have affected temperature readings for the Green, Elwha, Cedar and upper Skagit River sites, whose mainstems are or were dammed upstream of the sample location. The impact of damming on temperature at the Skokomish and the lower Skagit River sites should be lesser, as major, unobstructed river forks intercede between sample location and dam, resetting or partially resetting natural conditions (stanford2001revisiting). These sites are RS and SD, respectively, and both fall almost exactly on the regression line in Figure 4a. The upper Skagit site therefore occupies a middling space of Tair → Twater coupling between “fully” obstructed and unobstructed SD sites.

The role of reservoirs in restructuring natural temperature coupling relationships is complex (damsComplex,WarmingPostDam), and here confounded with many additional variables. Omitting all obstructed sites from Figure 3a, it would appear that no trend exists, yet we believe such omission is unwarranted. If the presence of reservoirs negated the influence of other factors, there would be no separation in coupling between obstructed sites of different river classes. Furthermore, though cold, hypolimnetic outflow should be expected to buffer Twater in summer, it alone cannot explain an *inverse* relationship between Tair and Twater. Instead, reservoirs may serve to enhance the decoupling of Tair → Twater and precip. → Twater brought on by snowmelt and glacial runoff, by selectively withholding warm water in their epilimnia and admitting cold water through their hypolimnia. Evidence for this phenomenon can be seen in the coupling of snowmelt and Twater, which is generally greater in RS and SD sites downstream of obstructions (Figure 3c). The Elwha River, which was cleared of its two dams between 2011 and 2014, will provide an excellent opportunity to compare each form of coupling with and without reservoirs, using the same dataset, once enough time has passed for signals to overcome inter-annual variability.

Water temperature in the unobstructed SD sites appears to couple more strongly with both air temperature and precipitation than in the sites with upstream dams (Figure 3a, b). In particular, the Duckabush and Puyallup Rivers (uppermost white circles in Figures 3a, 3b, and 4-Apr.) show stronger coupling relationships even than many of the RD rivers. Compared to all RS and RD rivers, and many SD, these stand out in terms of mean water table depth (Appendix C), suggesting they receive little flow from groundwater influx, which would otherwise serve to decouple Tair and Twater. They also occupy smaller watersheds than most of the other SD rivers, which yield lower overall discharge and heat capacity, and thus greater susceptibility to temperature change (caissie2006thermal). There may be additional factors at work in the SD rivers that account for the surprisingly high coupling seen in some unobstructed SD rivers. Another potential candidate is watershed slope, which increases with elevation and affects Twater by influencing residence time and evaporative cooling (via turbulence; caissie2006thermal). High slope and elevation are also associated with lower-order tributaries, and thus lower heat capacity.

Though higher-elevation watersheds will always produce colder water, independent of the influence of ice and snow, it can be expected that RS and SD rivers will grow more similar to RD as regional temperatures warm and glaciers decline. That is to say, formerly reliably cold-water rivers and associated habitats may see increases in both summer and winter average temperatures, as well as higher variability from year to year. The Elwha in particular may slip from its current state of high resistance to seasonal climatic changes. We tested for changes in mean and variance of Tair → Twater and Tair → Q coupling between 1978 and 2015, but did not detect any regular patterns (Appendix B).

In addition to the most parsimonious DFA, we fit a reduced model, designed to focus on whatever variation in Twater could be explained by landscape predictors. The two trends of this model represent additional drivers responsible for structuring water temperature across some or all of the 24 sites included in the analysis (Figure A6). We found elevation to be one of the dominant determinants of Tair → Twater and precip. → Twater coupling. Higher loading on shared trend 1 is associated with systems with higher elevation, snow and ice melt, and slope, as well as lower permeability of soils (Figure 3f). Dams (reservoirs) and BFI (essentially groundwater and reservoir contribution) were also major components of variation in coupling, along with water table depth. Groundwater, being insulated from the air, maintains relatively constant temperatures throughout the year, particular if it is deep underground. These additional sources of flow are expressed by shared trend 2. The shapes….

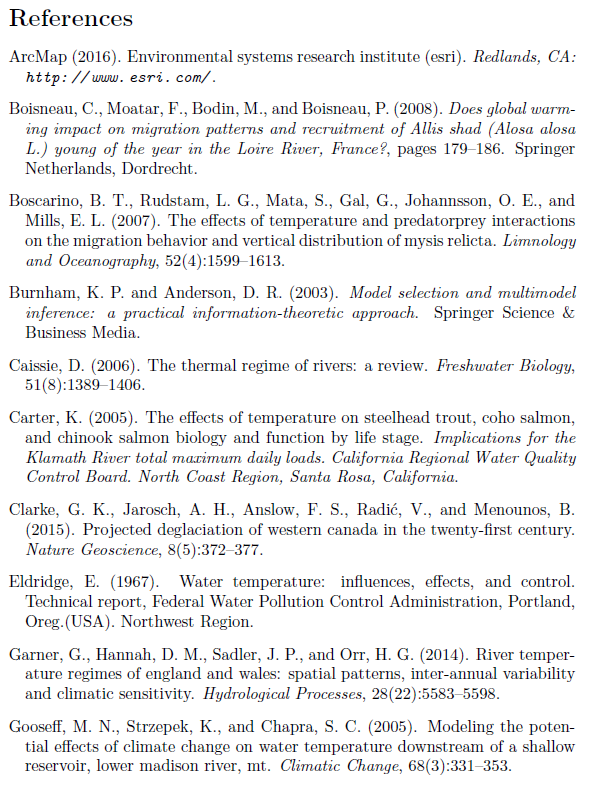
The relationship between climate and river temperature is further influenced by the interaction of discharge, and the fates of rivers in the Puget Sound watershed can be best understood by examining these factors in combination (Figure 4). Whether rain-, both-, or snow-dominated, all rivers appear to take on RD characteristics in winter, when the effects of ice lay dormant. As a result, warmer winters should on average yield warmer rivers and higher flow (less precipitation bound in ice). The critical differences between river classes play out in spring and summer, and it's during these months that future perturbations due to changing climate may be felt most acutely. For example, warmer Aprils on average produced colder water at 9 out of 15 RS and SD sites. Projected reductions in snowpack for the Puget Sound region (stewart2005changes,hamlet2013overview) can therefore be expected to fundamentally alter the responses of currently snow-influenced rivers to yearly variation in spring temperature over the long term. In the next 100 years, changes can be expected for rivers that now receive the temperature-buffering effect of glacial runoff. Glaciers continue to decline across North America, with glacial ice across Western Canada projected to decline by 70% from 2005 to 2100 (clarke2015projected), and only 3 of 13 examined North Cascades glaciers expected to survive the current climate (pelto2010forecasting).

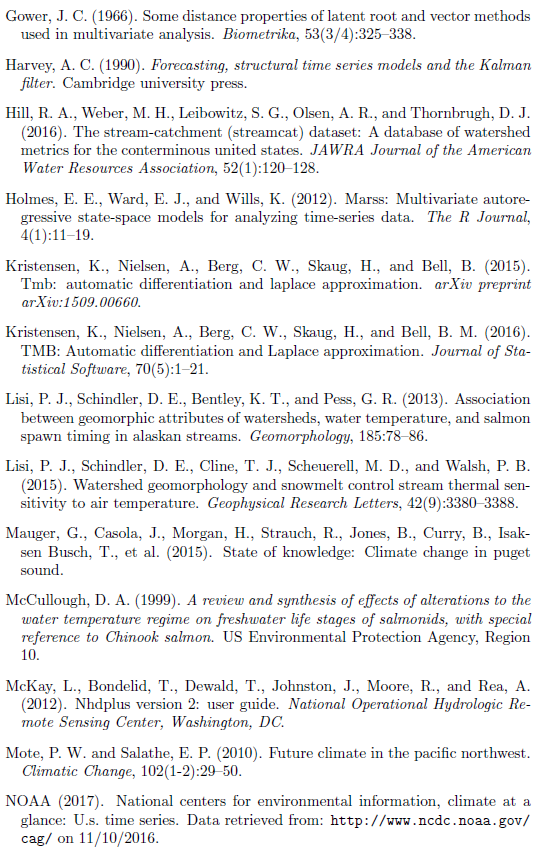
**Conclusion**

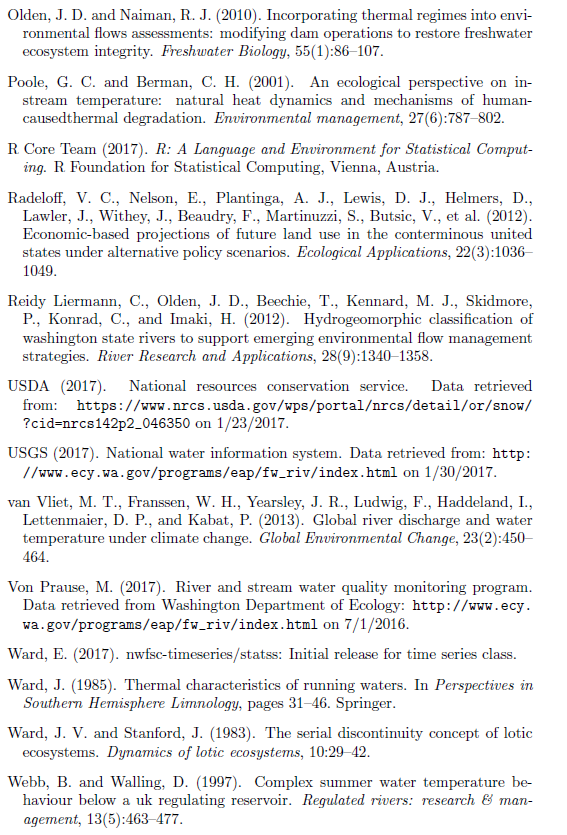
Temperature regimes across the rivers of the Puget Sound watershed are structured by a combination of climatic drivers at the regional scale, and geophysical drivers at watershed scales. In the absence of snow and ice, river temperature is closely coupled to that of the surrounding air, while discharge contributions from snowmelt and glacial runoff can dampen or even reverse this coupling in spring and summer, particularly where hypolimnetic-release reservoirs augment downstream cooling. In some cases, icemelt-influenced rivers exhibit stronger positive responses to climate patterns than their rain-driven counterparts. Our results suggest elevational variations in groundwater influx and total discharge may account for these patterns. However, while these factors and artificial reservoirs may influence the degree of coupling between climatic drivers and water temperature, only snow and ice can reverse it. Since 1978, such reversals have been widespread and commonplace, particularly during spring melt. Though we did not detect changes in this effect across historical observations, future reductions in snowpack and glacial mass are projected. Consequently, many rivers that now undergo the mildest seasonal temperature changes may be impacted most strongly.

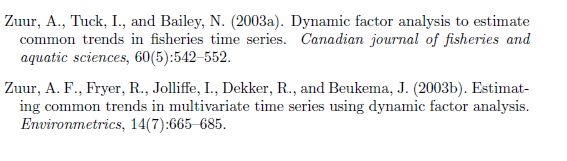
**Acknowledgements**

We thank Timothy Cline for the use of his TMB script, and Drs. Mark Scheuerell, Eric Ward, Eli Holmes, and Adrianne Smits for technical advice. Drs. Daniel Schindler and Michael Brett provided additional suggestions and guidance.









**TWO ADDITIONAL REFERENCES**:

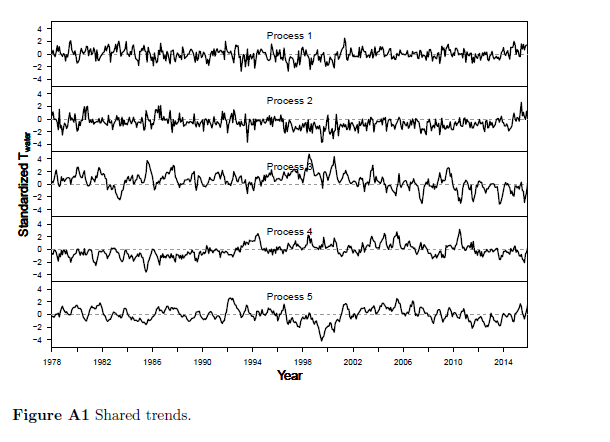
Beechie, T. J., Hamish Moir, and G. E. O. R. G. E. Pess. "Hierarchical physical controls on salmonid spawning location and timing." *Salmonid spawning habitat in rivers: physical controls, biological responses, and approaches to remediation. American Fisheries Society, Symposium*. Vol. 65. 2008.

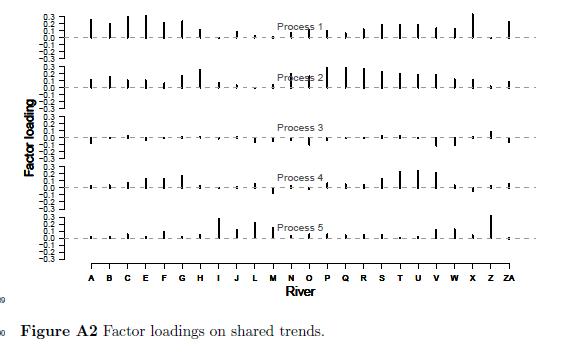
Sanders, J. E., K. S. Pilcher, and J. L. Fryer. "Relation of water temperature to bacterial kidney disease in coho salmon (Oncorhynchus kisutch), sockeye salmon (O. nerka), and steelhead trout (Salmo gairdneri)." *Journal of the Fisheries Board of Canada* 35.1 (1978): 8-11.

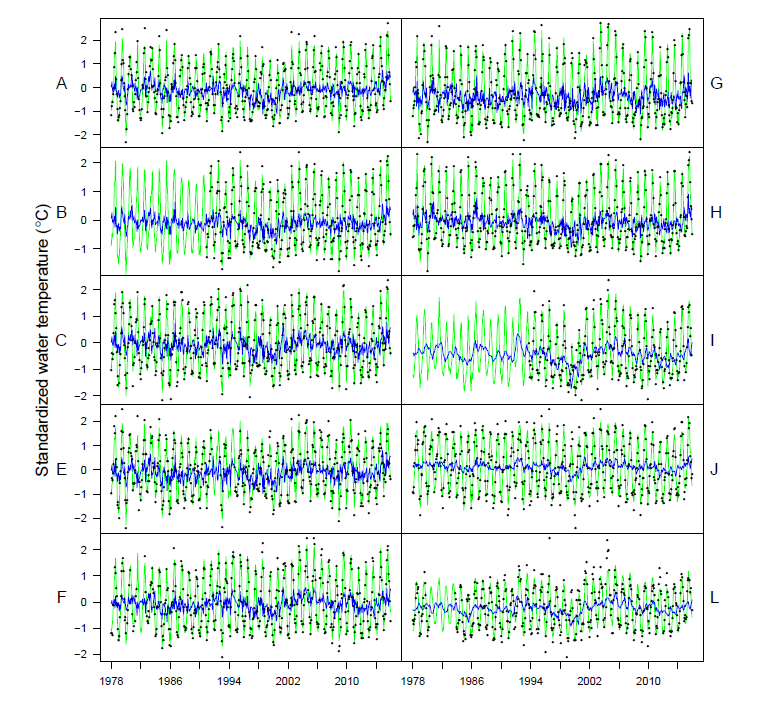
**Appendix A**

Temperature DFA output and diagnostics

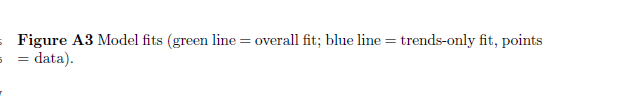
The process of selecting the best Twater and discharge models involved four climate covariates (air temperature, precipitation, snowmelt, and hydrological drought), between 1 and 15 shared trends, four within-and-among-site error structures (see methods), and two expressions of unknown seasonal variation (fixed monthly factors and Fourier series). The most parsimonious models were selected using the Akaike Information Criterion (AIC) in tandem with R2 (required increase of 1% for each additional parameter), and in each case included air temperature, precipitation, and snowmelt as covariates. The Twater model also included five shared trends and an independent and unequally distributed error structure among rivers (i.e. diagonal and unequal variance-covariance matrix). The discharge model (not shown) included six trends. All subsequent plots relate to the Twater model, and alphabetic names correspond to sampling sites (Figure 1).

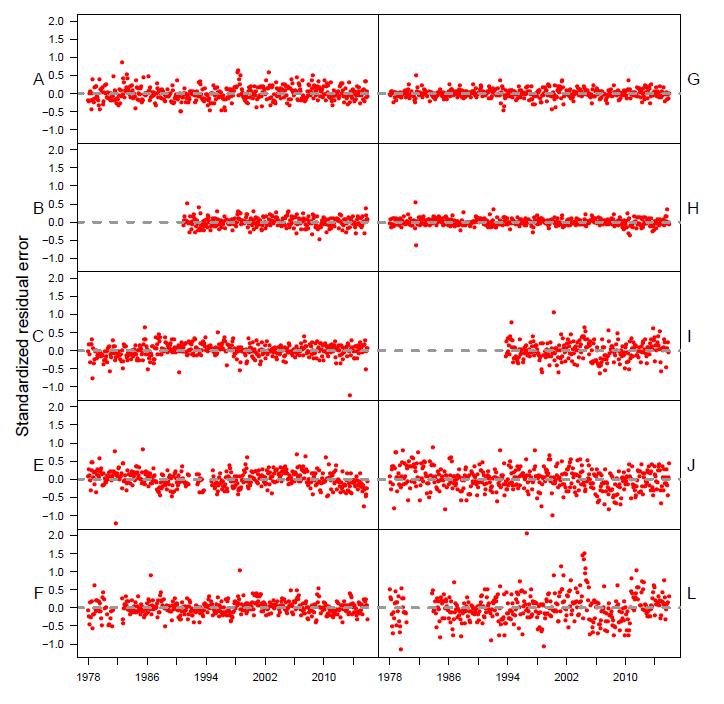






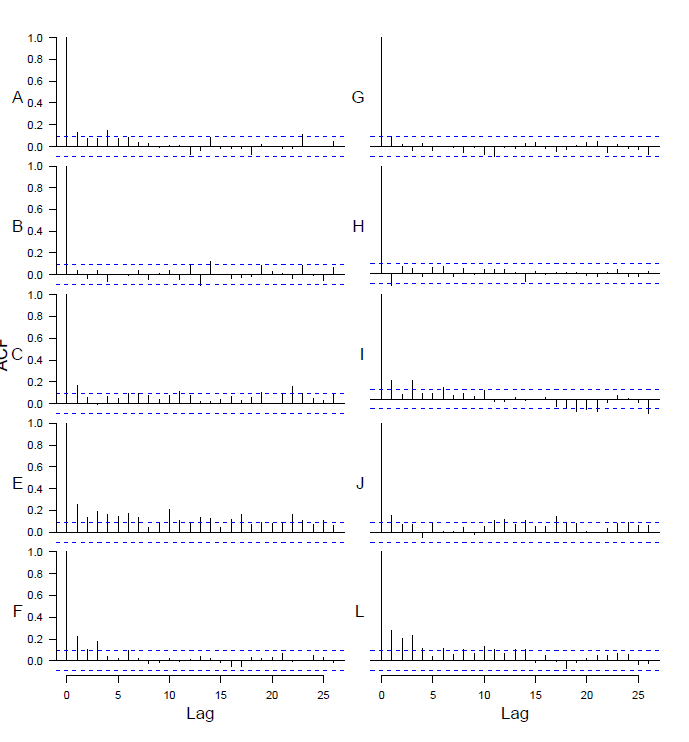
Etc…



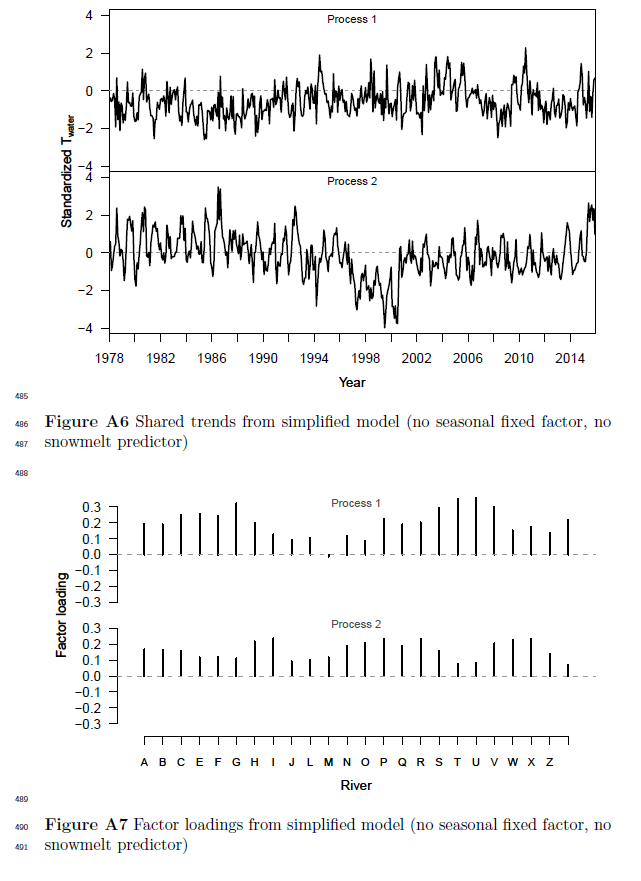


Etc…

Figure A4: Residuals.



Etc…



**Appendix B**

Testing for change in coupling over time

We used an additional DFA model to test for changes in Tair → Twater coupling over time, by dividing the 1978-2015 time series into 5 intervals and comparing central tendency and variance of effect sizes for each interval. Figures B1-B3 show mean Tair → Twater coupling for each river.

To approximate estimates of variability over time, we performed the same analyses within a Bayesian framework, and obtained uncertainty estimates from the credible intervals of the effect size (i.e. degree of coupling) posteriors. This approach yielded no trends in variation over time, and is not visualized here. For Bayesian analyses, we used R package “statss” (eric\_ward\_2017\_375646).

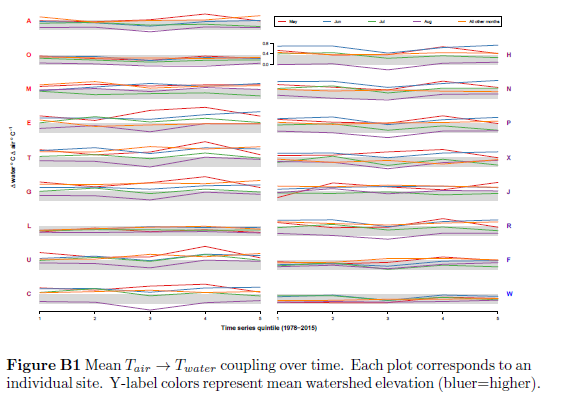
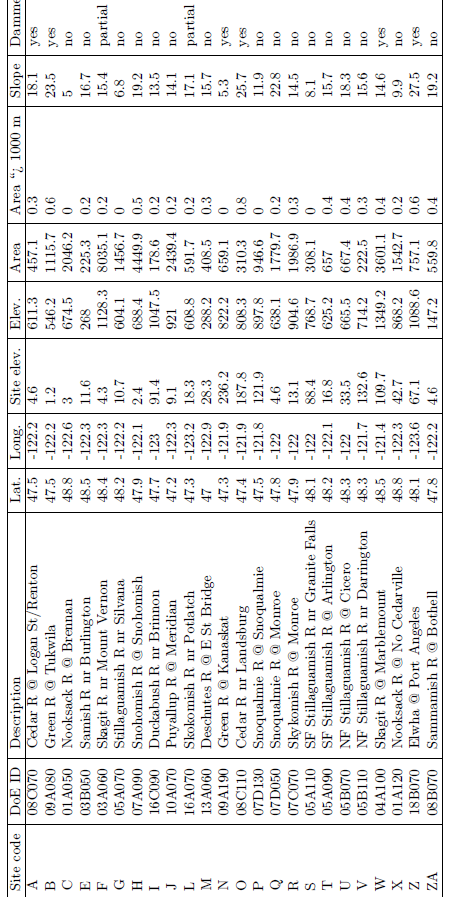
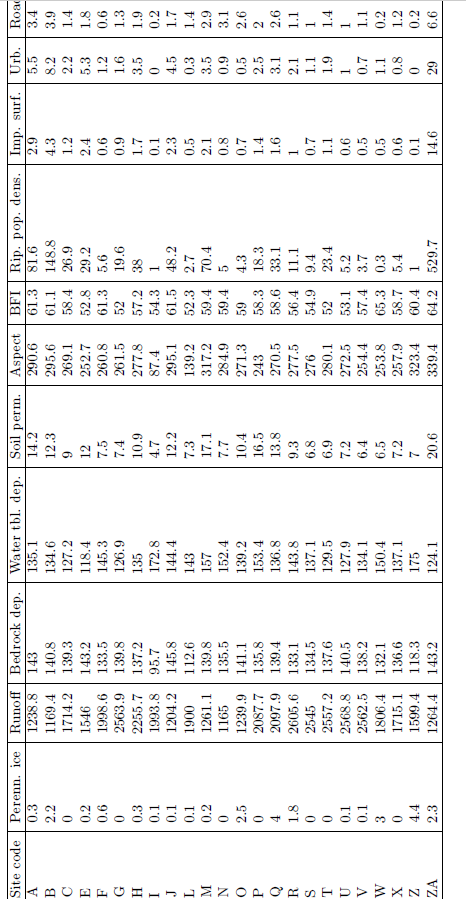


Figure B2…

Figure B3…

**Appendix C**



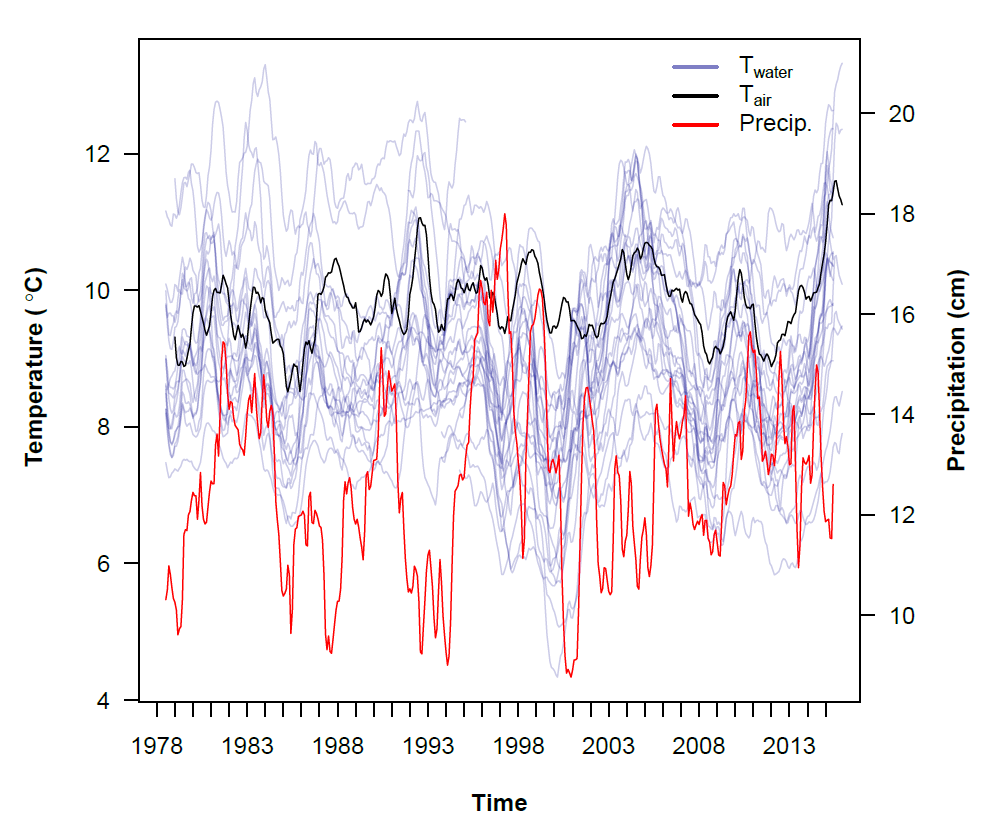


That last column is “Road density”.

**Appendix D**

Testing for trends in absolute Twater, Tair, precipitation, and snowmelt over the entire time series

We performed Kendall’s Test for Trend to determine whether T­­water or any of the climate predictors showed significant, monotonic positive or negative slopes over the 38-year interval. The Seasonal Kendall Test was avoided because of its assumption that any monotonic trends present are the same across all seasons (months, in our case). Instead, each time series was decomposed into trend, seasonal, and noise components, and only the trends (Figure D1) were used in the analyses, with six data points on either end removed during the decomposition process.



**Figure D1** Time series of Twater across all sites, and regional Tair and precipitation, with seasonality and random noise removed via decomposition.

Kendall’s test identified significant positive, monotonic trends in T­air, precipitation, and 10 out of 24 T­water time series (Tables D1, D2). The same number of Twater series was determined to be monotonically decreasing over the 37-year interval from July 1978 to June 2015.

**Table D1** Results of Kendall’s Test for Trend (with continuity correction) on climate predictor time series with seasonality and random noise removed via decomposition. Slope estimated via Thiel/Sen Estimator; intercept via Conover’s Estimator. Confidence intervals (upper and lower 95%) determined via Gilbert’s modification to Thiel/Sen method.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Predictor | Kendall’s Tau | Slope | Lower 95 | Upper 95 | Intercept | z | *p* |
| Tair | 0.128 | 0.001 | 0 | 0.001 | 9.613 | 4.015 | 0 |
| Precip | 0.101 | 0.002 | 0.001 | 0.004 | 12.136 | 3.167 | 0.002 |
| Snowmelt | 0.043 | 0 | 0 | 0.001 | 2.658 | 1.355 | 0.175 |

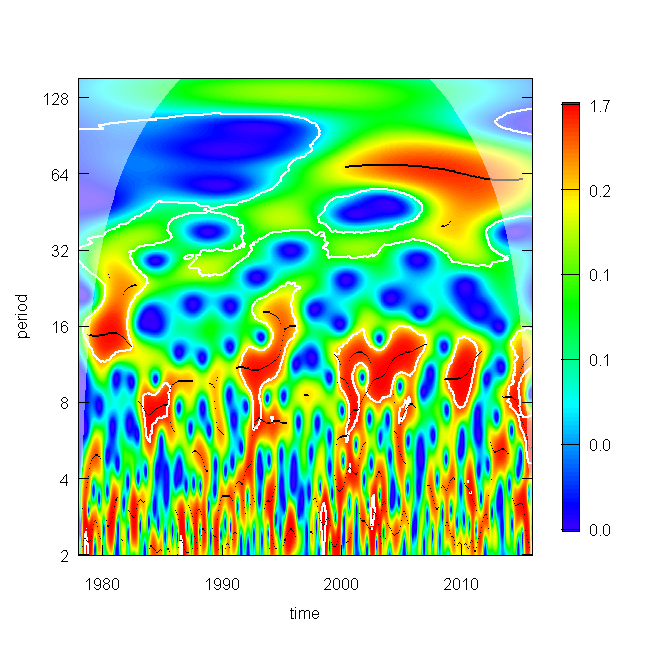
**Table D2** Results of Kendall’s Test for Trend (with continuity correction) on Twater time series with seasonality and random noise removed via decomposition. Specifications same as above. Slopes significant at α = 0.05 are bolded. Significant negative slopes are italicized.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Site | Kendall’s Tau | Slope | Lower 95 | Upper 95 | Intercept | z | *p* |
| **A** | **0.150** | **0.001** | **0.001** | **0.002** | **9.870** | **4.721** | **0.000** |
| *B* | *-0.293* | *-0.002* | *-0.002* | *-0.001* | *11.649* | *-9.270* | *0.000* |
| C | 0.003 | 0.000 | -0.001 | 0.001 | 9.177 | 0.084 | 0.933 |
| E | 0.011 | 0.000 | -0.001 | 0.001 | 9.325 | 0.358 | 0.720 |
| **F** | **0.088** | **0.001** | **0.000** | **0.001** | **8.675** | **2.780** | **0.005** |
| **G** | **0.094** | **0.001** | **0.000** | **0.002** | **8.966** | **2.975** | **0.003** |
| *H* | *-0.208* | *-0.002* | *-0.002* | *-0.001* | *10.090* | *-6.564* | *0.000* |
| *I* | *-0.326* | *-0.002* | *-0.002* | *-0.001* | *7.763* | *-10.321* | *0.000* |
| *J* | *-0.221* | *-0.002* | *-0.003* | *-0.002* | *9.836* | *-6.945* | *0.000* |
| **L** | **0.224** | **0.001** | **0.001** | **0.002** | **8.077** | **7.060** | **0.000** |
| *M* | *-0.147* | *-0.001* | *-0.001* | *-0.001* | *11.055* | *-4.635* | *0.000* |
| N | -0.061 | -0.000 | -0.001 | 0.000 | 8.572 | -1.928 | 0.054 |
| **O** | **0.096** | **0.001** | **0.000** | **0.001** | **8.297** | **3.029** | **0.002** |
| P | -0.060 | -0.001 | -0.001 | 0.000 | 8.328 | -1.877 | 0.061 |
| *Q* | *-0.416* | *-0.004* | *-0.005* | *-0.004* | *11.417* | *-13.195* | *0.000* |
| *R* | *-0.127* | *-0.001* | *-0.002* | *-0.001* | *9.085* | *-3.991* | *0.000* |
| **S** | **0.235** | **0.001** | **0.000** | **0.002** | **7.365** | **7.413** | **0.000** |
| **T** | **0.183** | **0.002** | **0.001** | **0.003** | **8.436** | **5.759** | **0.000** |
| **U** | **0.175** | **0.002** | **0.001** | **0.003** | **7.846** | **5.505** | **0.000** |
| *V* | *-0.107* | *-0.000* | *-0.000* | *0.000* | *7.942* | *-3.392* | *0.001* |
| **W** | **0.098** | **0.001** | **0.000** | **0.001** | **7.333** | **3.100** | **0.002** |
| *X* | *-0.242* | *-0.002* | *-0.002* | *-0.002* | *8.550* | *-7.615* | *0.000* |
| **Z** | **0.071** | **0.000** | **-0.000** | **0.000** | **8.167** | **2.288** | **0.022** |
| *ZA* | *-0.466* | *-0.002* | *-0.003* | *-0.002* | *11.624* | *-14.788* | *0.000* |

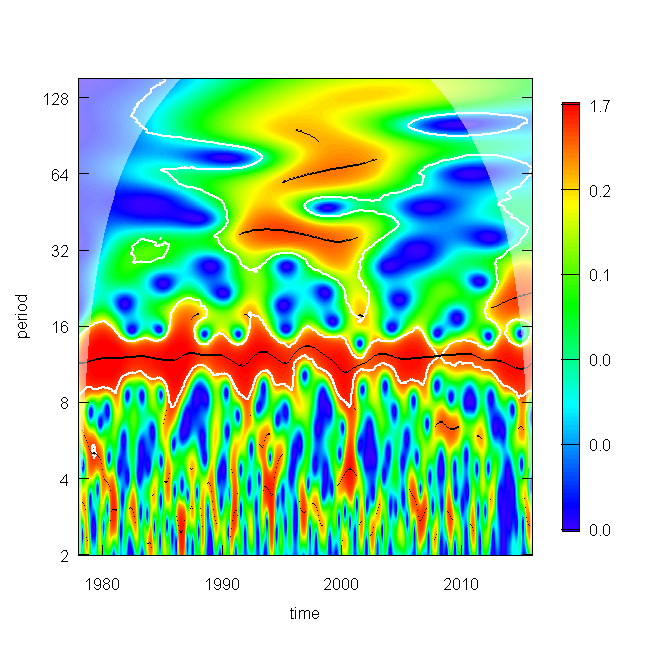
**Appendix E**

Wavelet analysis of shared trends

To identify potential sub-seasonal periodic structure in the two shared trends from the reduced fit model, we conducted a wavelet power spectrum analysis using the Morlet wavelet (RoeschWavelet). Shared trends were detrended using a Loess span of 0.75. Power spectra revealed a continuous ridge of strong 12-month periodicity in shared trend 2 (Figure E2).

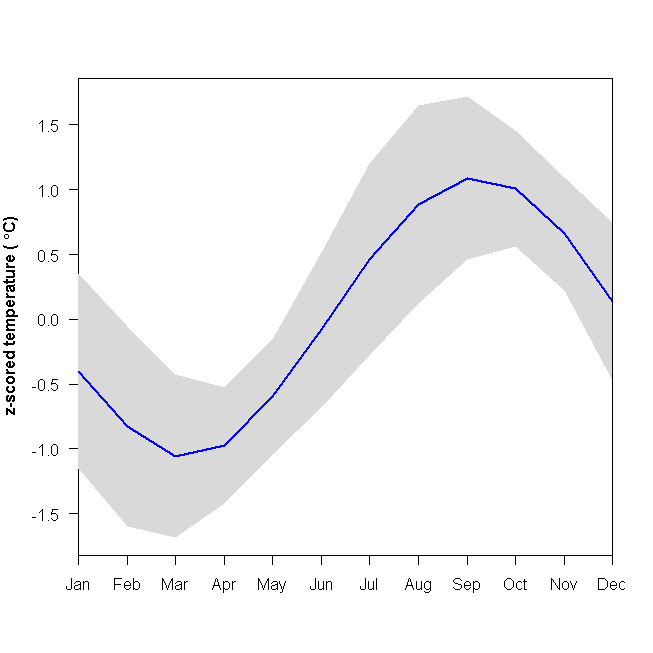


**Figure E1** Wavelet power spectrum for shared trend 1. Period refers to Fourier period. Warmer colors denote greater strength of periodicity. White polygons encompass areas of significant periodicity at α = 0.10, based on comparison with a white noise surrogate time series. Black lines correspond to “ridges” of highest power within significant regions.



**Figure E2** Wavelet power spectrum for shared trend 2. See Figure E1 for details.

Trend 2 was cleaned to remove all but the time series components contributing to the ridge at period=12. The original trend was then reconstructed using only those components, revealing a yearly recurring sinusoid patter of low values in spring and high values in late summer/early fall.



**Figure E3** Mean (blue) and standard deviation (gray) of the seasonal component of shared trend 2, averaged across the 38 years between 1978 and 2015.