Introduction

Methods

Study region

I'll probably mention anything relevant to this in the intro.

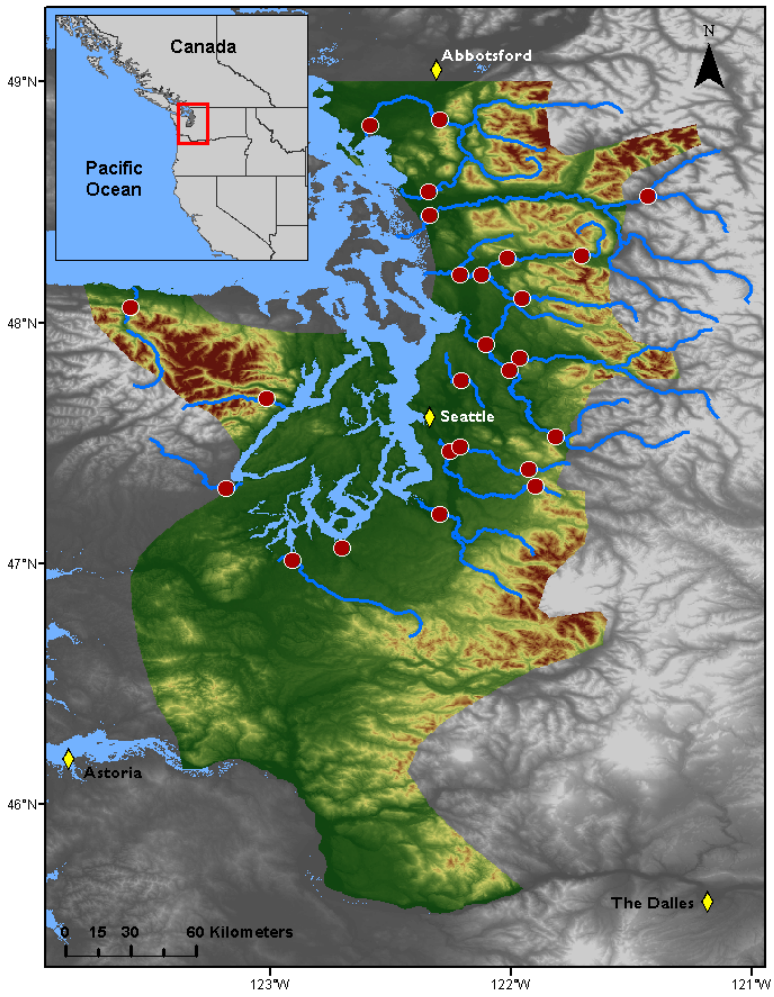

Figure 1 Site locations (red points) in relation to the region across which climate data were aggregated (colored topography).

Water and climate data

The two response variables used in our analyses were water temperature and discharge. We obtained monthly water temperature readings from 1978 through 2015 via the Washington Department of Ecology's River and Stream Water Quality Monitoring program cite . In all, 24 monitoring sites within the Puget Sound region (Fig. 1) were included, representing 19 rivers across 9 counties, and ranging from 4 to 775 m in elevation. For one site at each river, monthly discharge time series were available, either for the same location as one of the temperature monitoring sites, or for another location within 30 km on the same major reach. Discharge data were aggregated by monthly mean from the USGS Washington Water Science Center (collected daily 1978-2007) and the USGS National Water Information System (collected at 15-minute intervals 2008-2015). At least one discharge monitoring site was available for every river represented in the temperature dataset.

Potential climatic predictors of water temperature and discharge included mean and max air temperature (°C), total precipitation (cm), snowmelt (cm), and hydrological drought (Palmer Hydrological Drought Index), averaged by month across the response variable time series. All but snowmelt are available through the U.S. Climate Divisional Dataset, developed by the National Centers for Environmental Information (NCEI) cite . We acquired climatic predictor data grouped by Washington State climate division, and all but two of our sites fell within divisions 3 (Puget Sound Lowland) and 4 (East Olympic/Cascade Foothills). We therefore aggregated these data by monthly mean across the two regions (after verifying their post-standardization similarity), resulting in a single dataset of four regional, climatic predictors. A snowmelt time series was then added to this dataset, using monthly mean records 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 cite . We calculated monthly snowmelt for each site as the absolute value of negative differences in cumulative snow water equivalent from each month to the next. The snowmelt time series was assigned zeros for any positive differences (accumulations).

Time series analysis

Response time series were modeled using dynamic factor analysis (DFA; Zuur et al. 2003), a multivariate technique that can be thought of as an analog to principal component analysis in the time domain. In DFA, response time series are fit with a linear combination of shared, random-walk trends (usually many fewer than the total number of response series), predictors (which can have unique effects on each response series), and random error. We chose DFA over a traditional multivariate state space approach for two reasons. First, it provides advantages

in computational efficiency, as 1-5 shared trends often adequately capture variation across dozens of responses, and at much lower parameter cost. 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 Kristensen et al. 2015, which we called using package TMB in R R Core team 2016....

DFA takes the following form:

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t, \text{ where } \mathbf{w}_t \sim \text{MVN}(0, \mathbf{Q}) \tag{1}$$

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t, \text{ where } \mathbf{v}_t \sim \text{MVN}(0, \mathbf{R}) \tag{2}$$

$$\mathbf{x}_0 \sim \text{MVN}(0, \mathbf{\Lambda}) \tag{3}$$

At time step t, the $m \times 1$ vector of shared trends ($\mathbf{x}$) is a function of $\mathbf{x}$ in the previous step, plus normal error ($\mathbf{w}$; $m \times 1$; Eq. 1). This is the definition of a random walk. The $n \times 1$ response vector ($\mathbf{y}$) at time t is a function of the shared trends and their factor loadings ($\mathbf{Z}$; $n \times m$), covariates ($\mathbf{d}$; $q \times 1$) and their river-specific effects ($\mathbf{D}$; $n \times q$), and a second normal error term ($\mathbf{v}$; $n \times 1$; Eq. 2). $\mathbf{R}$ and $\mathbf{Q}$ are variance-covariance matrices of order m, and $\mathbf{Q}$ is set to identity for model identifiability (Harvey 1989). The initial state of the shared trend vector ($\boldsymbol{x}_0$) 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 river temperature and discharge, we used fixed factors to absorb recurring seasonal variation not related to the predictors, with one factor level for each month. These factors were incorporated into the covariate matrix ($\mathbf{d}$). Thus, the coefficient in $\mathbf{D}$ relating, say, air temperature (predictor) and water temperature (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 air temperature across the whole time series. We call this relationship "coupling." We were also interested in coupling by month for specific predictors, which required that the focal predictor in a particular model be arranged seasonally. Concretely,

$$
\mathbf{d} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \\ 12 \\ 13 \\ 14 \\ 15 \\ 16 \\ \\ 25 \end{array}
\begin{array}{c}
\begin{array}{ccccc} \text{Jan}_{1978} & \text{Feb}_{1978} & \text{Mar}_{1978} & \cdots & \text{Dec}_{2015} \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
\theta^{(1)}_{precip} & \theta^{(2)}_{precip} & \theta^{(3)}_{precip} & \cdots & \theta^{(T)}_{precip} \\
\theta^{(1)}_{air} & 0 & 0 & \cdots & 0 \\
0 & \theta^{(2)}_{air} & 0 & \cdots & 0 \\
0 & 0 & \theta^{(3)}_{air} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \theta^{(T)}_{air}
\end{pmatrix}
\end{array}
$$

is the covariate matrix structure necessary to account for exogenous seasonal variation (rows 1-12), and overall effect of precipitation (row 13), while also yielding the by-month effect of air temperature (rows 13-24) on the response ($\mathbf{y}$).

Additional, non seasonal variation due to exogenous effects loads onto the shared trends, and a portion of remaining variation is absorbed by error matrix $\mathbf{v}$. We fit models using four unique error structures ($\mathbf{R}$), to allow for different suites of unknown drivers affecting rivers. We included shared variance and zero covariance, individual variance and zero covariance, shared variance and shared covariance, and unconstrained error. Details on these structures and their implications can be found in Holmes et al. (2012). The best models for water temperature and discharge were determined with AIC.

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 (hereafter "% glaciation") and mean elevation across their watersheds. The three classes are loosely based on the classification scheme and language of the Climate Impacts Group at the University of Washington cite, and are here delineated according to Table 1.

Table 1 Watershed classification scheme

Classification	Abb.	Glaciation (%)	Mean elev. (m)
Rain-dominated	RD	< 0.7	< 600
Rain-and-snow	RS	< 0.7	≥ 600
Snow-dominated	SD	≥ 0.7	-

After model selection, climatic predictor effect sizes for each river were back-transformed to their original scales and regressed against landscape predictors in

order to identify possible watershed-scale controls on coupling. To achieve this, we amassed an additional dataset of landscape features. These were collected individually for each of the watersheds corresponding to our 24 river sites, using the EPA's StreamCat (stream-catchment) data library cite and the National Hydrography Dataset (NHDPlusV2) cite. 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 (km^2), base flow index, 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 (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 was also included. Finally, we calculated area over 1000 m (km^2) and mean slope (percent rise) by delineating watersheds from a digital elevation model in ArcMap cite.

Results

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 (Fig. 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 stats throughout this section. RD, meanwhile, 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.

Meanwhile, the amplitude of T_{air} oscillation exceeded that of any river class, dipping below T_{water} in autumn to a minimum of 3.2°C in December, and rising above RS and SD in March to an August maximum of 17.4°C. T_{air} did not overtake RD T_{water} until August, by which time the latter had begun to decline.

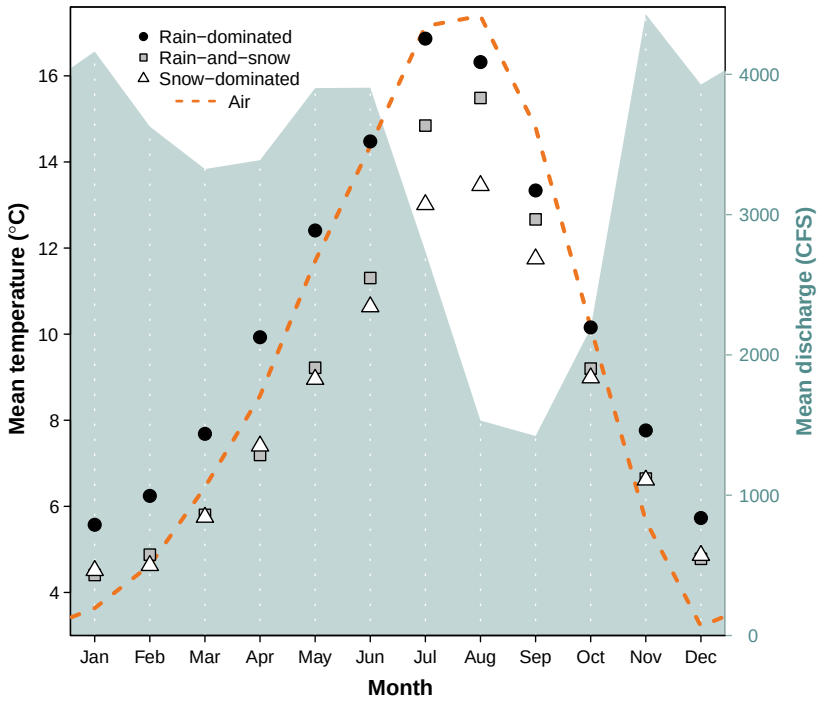

Figure 2 Monthly mean T_{water} by river class, and T_{air} and Q across classes, from 1978 to 2015. All depicted series represent discrete data. see appendix for standard deviations?

The combined hydrograph of all rivers reveals two primary peaks, one beginning in late spring and the other extending from late fall to early winter, with a prominent trough in late summer. Spring peak discharge coincides noticeably with a separation in water temperature between SD and RS, while the summer trough coincides with separation of RD and T_{air}. November marks both the autumn peak in discharge and the point at which T_{air} falls below T_{water}.

DFA results, aggregated across months and years for each site, revealed a trend toward reduced $T_{air} \rightarrow T_{water}$ coupling with greater watershed elevation ($p = 0.04, R^2 = 0.18$; Fig. 3a). On average, a 1°C change in T_{air} corresponded to a 0.53 ± 0.03°C change in T_{water} 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 $Precip \rightarrow T_{water}$ coupling ($p = 0.03, R^2 = 0.21$; Fig. 3b), where a monthly change in total precipitation of 1 cm corresponded to a 0.02 ± 0.009°C change in T_{water} for RD, -0.003 ± 0.009°C for RS, and 0.004 ± 0.02°C for SD. There was no evidence of coupling between snowmelt and T_{water} (Fig. 3c), but this predictor

was included in the most parsimonious DFA model selected via AIC and R^2 see appendix. It is of note that the strongest examples of $T_{air} \rightarrow T_{water}$ and $Precip \rightarrow T_{water}$ 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, though the Elwha's watershed is larger.

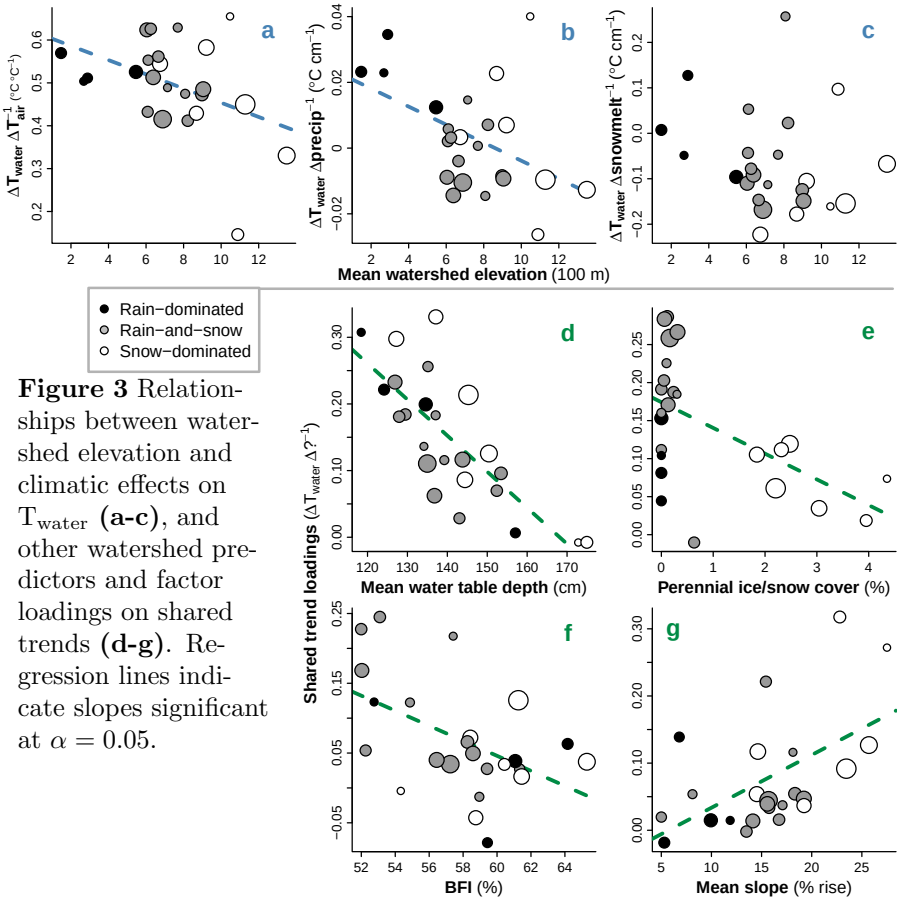

Figure 3 Relationships between watershed elevation and climatic effects on T_{water} **(a-c)**, and other watershed predictors and factor loadings on shared trends **(d-g)**. Regression lines indicate slopes significant at $\alpha = 0.05$.

In addition to the three climate predictors above, the best T_{water} model also included five shared trends. Of these, four correlated significantly with at least one known watershed predictor. Figure 3 depicts the strongest correlated variables with each trend (insets d-e). These are, in arbitrary order of relevance, mean water table depth ($p < 0.001, R^2 = 0.60$; Fig. 3d), % glaciation ($p < 0.01, R^2 = 0.30$; Fig. 3e), BFI ($p = 0.01, R^2 = 0.25$; Fig. 3f), and mean slope ($p < 0.01, R^2 = 0.29$; Fig. 3g). The fifth shared trend was not correlated with any variables in the watershed predictor dataset.

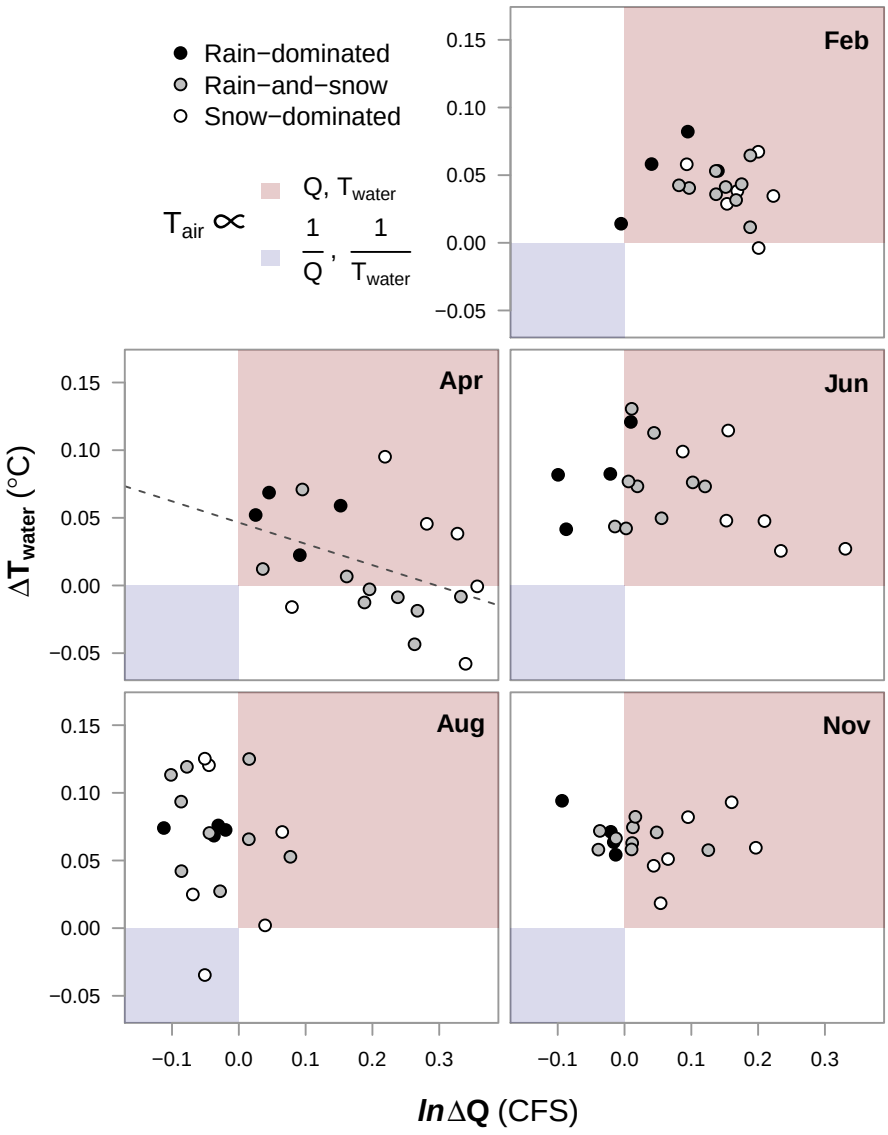

Figure 4 Relationship between $T_{air} \rightarrow T_{water}$ and $T_{air} \rightarrow Q$. Both axes are expressed per 1°C change in T_{air}. The red quadrant designates proportionality between all three variables, the blue inverse proportionality between each response and T_{air}. Regression lines indicate slopes significant at $\alpha = 0.05$.

To examine possible sub-season interactions between T_{air}, T_{water} and Q, we performed an additional DFA with Q as the response. In both models, T_{air} was allowed to have unique monthly effects. These effects, taken together, can be understood in relation to the four quadrants of the Cartesian coordinate system

(increasing clockwise from upper right; Fig. 4).

In mid-winter (exemplified by February), all river classes primarily occupy the first quadrant, signifying $T_{\text{air}} \propto T_{\text{water}}$ and $T_{\text{air}} \propto Q$, where $\propto$ denotes proportionality. RD shows the weakest Q response. By spring, many RS and SD sites develop an inverse relationship between T_{air} and T_{water}, denoted $T_{\text{air}} \propto \frac{1}{T_{\text{water}}}$, while RD sites change little from their winter state. June and August see a procession of most sites into the near fourth quadrant, with SD trailing. This signifies $T_{\text{air}} \propto \frac{1}{Q}$, though $T_{\text{air}} \propto T_{\text{water}}$ remains. One stark exception is again the Elwha river, which occupies quadrant three. By fall, RS and SD have begun progress back toward their winter states, led by SD. RD, meanwhile, remain essentially unmoved from summer.

Discussion

The effects of climate on T_{water}, determined by dynamic factor analysis, suggest that nearly all rivers included in our dataset were influenced strongly by air temperature, precipitation, and/or snowmelt across 38 years of monthly data (Fig. 3). At most monitoring sites, T_{water} closely tracked changes in T_{air}, on average responding to increases and decreases with proportional movements of up to 66% magnitude. However, some rivers only weakly track T_{air}, and several patterns in the intensity of this coupling correlate strongly with watershed features relating to ice, groundwater, and slope. Glaciation and yearly snow burden are prominent among these, and for reasons of ecological and hydrological implication, the primary focus of the following discussion.

Without any analysis, a "buffering" effect (hereafter contrasted with "coupling") of ice on river temperature can be seen in the yearly patterns of $_{\text{water}}$ relative to T_{air} (Fig. 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 rivers that receive glacial runoff (SD), this effect appears to remain, buffering them from summer temperature rise where RS rivers instead take on the character of RD (Fig. 4). In an extreme case, the Elwha River was actually cooler in August during those years in which air temperature was higher, likely 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 fall, by way of glacial runoff. For RD rivers, which receive little to no input from ice, summer temperature is entirely dictated by that of the surrounding air, and whatever rain falls through it. Though higher-elevation watersheds will always produce colder water, independent of the influence of ice, 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 streams

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 $T_{\text{air}} \rightarrow T_{\text{water}}$ and $T_{\text{air}} \rightarrow Q$ coupling between 1978 and 2015, but did not detect any widespread patterns appendix B.

In addition to the three climate predictors, five shared trends were fit by the most parsimonious DFA model. These represent additional drivers responsible for structuring water temperature across some or all of the 24 sites included in the analysis. Each monitoring sites' factor loading on a particular shared trend indicates the degree to which the trend accounted for variance in T_{water} at that site. While the precise identities of these drivers cannot be attained with certainty, they can be inferred through correlation with predictor variables. In this way, we determined the most likely landscape drivers of T_{water} to be perennial ice and snow cover, mean watershed slope, and groundwater influx. Slope's likely mechanism of influence is increased turbulence and mixing of water and air in steep, headwater streams, which allows convective warming and cooling to occur more rapidly cite.

SD RS colder throughout the year, less influenced by changes in air. RD actually remains warmer than air throughout the spring, and matches it through July. snow isn't signif because of the arrangement of points along the x axis. however, there's a negative snowmelt-¿Twater for the glac streams and a positive one for the RS streams Why might each of the landscape predictors be such? water table depth: ground water farther from surface, less responsive ground water (to whatever it is) BFI: also ground water influx ice snow: separation of RD RS?: this could be the rain-on-snow effect slope: faster and more turbulent water, greater mixing; also correlated with snow/glac what could the other factor be (trend 3)? random error, or a geological or morphological or microclimatic feature not captured by our watershed predictors fog3: step through

The relationship between river classification and factor loadings is less clear

Appendix A

DFA plots here blargenshnargen

Appendix B

over-time plots here