

todo:  
include snowmelt in model run. see if it's worth including with precip and temp as a climate predictor  
introduce terminology: responses = water temp, Q climatic predictors = air temp, precip, drought landscape predictors = many

## Introduction

## Methods

### Study region

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

### Data collection

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, precipitation, and hydrologic drought, averaged by month across the response variable time series. These data 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.

For post-hoc analyses of potential sub-watershed-scale drivers of response-predictor coupling, we amassed an additional set of landscape predictor data. These were collected individually for each of the watersheds that correspond to our 24 river sites, using the EPA's StreamCat (stream-catchment) data collection [cite](#) and the National Hydrography Dataset (NHDPlusV2) [cite](#). Land-

scape predictor categories include lithology, land use, population and road density, and soil type, as well as other categories summarized in [Table A1](#).

## 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), covariates (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, it allows greater inferential parsimony. Being a multivariate technique, it also provides an advantage over univariate alternatives in that covariance structure among responses can be specified and compared.

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

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

$$\mathbf{x}_0 \sim \text{MVN}(\boldsymbol{\pi}, \boldsymbol{\Lambda}) \quad (3)$$

## Appendix A

Table A1: ...at the scale of individual stream reaches (segments bounded by sources, confluences, or mouths) and their corresponding watersheds. Watersheds are calculated as land contributing flow to a reach, and have been determined for 2.6 million reaches within the conterminous United States.