THESIS

PREDICTING FLOW DURATION AND ASSESSING ITS DRIVERS IN NORTH-CENTRAL COLORADO USING CROWDSOURCED DATA

Submitted by

David Peterson

Department of Ecosystem Science & Sustainability

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2022

Master's Committee:

Advisor: Stephanie K. Kampf

Matt Ross
Sean Gallen

ABSTRACT

PREDICTING FLOW DURATION AND ASSESSING ITS DRIVERS
IN NORTH-CENTRAL COLORADO USING CROWDSOURCED DATA

Headwater streams are globally important both ecologically and for human resource needs. These streams represent the majority of stream network length, but their flow regimes are often unknown. Streams can be classified by flow regime as perennial, intermittent, or ephemeral. These classifications are used in forest land management decisions and may affect Clean Water Act jurisdiction; however, the National Hydrography Dataset (NHD) often misclassifies headwater streams. The goal of this study is to model flow duration across the stream networks of eight subbasins in north-central Colorado. We used crowdsourced flow presence/absence data from 82 sites in the Stream Tracker program and eight flow sensors to train random forest regression models; these models predicted the fraction of time a stream flows from April-September for both the average from 2016-2020 (dubbed mean annual) and yearly averages (annual). Model predictor variables included climatic, physiographic, and land cover attributes of the study area. Models were developed using a sample of the sites for training and leaving the remaining sites for model testing. The resulting mean annual model's Nash-Sutcliffe efficiency (NSE) was 0.88 for test data, and the annual model's test data had an NSE value of 0.81. We found climate variables such as snow persistence, precipitation, and potential evapotranspiration most influential in predicting flow fraction based on the random forest-ranked variable importance. Forested and herbaceous land cover as well as depth to bedrock, available water storage, hydraulic conductivity, hydrologic soil group, drainage area, and watershed curvature were also identified as important drivers. We developed maps of predicted flow fractions and compared them to NHD flow classifications. In the Cache La Poudre subbasin, the mean annual model predicted perennial flow in 10% of streams and intermittent or ephemeral flow in 90% of streams. Our model predicted nonperennial flow for 76% of the streams that were mapped as perennial in the medium-resolution NHD. Based on these findings, the NHD over-represented perennial streams, classifying them three times more than our model, and under-represented intermittent and ephemeral streams by 32% in our study area. The annual model captured interannual variability in flow fraction and highlighted isolated areas of high variability in flow fraction between years in mid-to-low elevations. The models we developed using crowdsourced data can improve flow classifications of headwater streams and inform resource management decisions in northern Colorado. Crowdsourced streamflow data can be used in streamflow predictions anywhere that nonperennial flow is common.

ACKNOWLEDGEMENTS

I would like to sincerely thank my advisor, Stephanie Kampf, for her guidance and encouragement as I learned and grew through this process. I couldn't have done this without her skill in managing crowdsourced data and her expertise when analyzing results. I'm grateful for our collaboration with John Hammond, whose planning, knowledge, and data contributions were invaluable to this effort. I also want to thank Sam Zipper, who provided the random forest code structure used in this study and graciously guided me through the ins-and-outs of that code. Thanks to Kira Puntenney-Desmond for her work organizing and promoting the Stream Tracker project. Thank you to Abby Eurich for her help with initial data processing. Feedback from my committee members, Matt Ross and Sean Gallen, has helped strengthen this work.

Funding was provided by the Arapaho & Roosevelt National Forests and the U.S. Forest Service Citizen Science Program. Thank you to Matt Fairchild for the support and thoughts on how to make this work applicable to many facets of natural resource management.

I want to thank Katie Willi, Bryce Pulver, Chelsea Wilmer, Marielle Sidell, Quinn Miller, Leonie Kiewiet, and Clara Mosso for their work in data gathering and maintaining our sensor array. Thanks to all other Stream Tracker volunteers for making this work possible. And finally, I want to take the opportunity to thank my friends and family for their support in my pursuits of a graduate degree. The list of people that have guided me along the natural resource path is long and full of wonderful teachers, mentors, and coworkers.

TABLE OF CONTENTS

1. INTRODUCTION

Headwater streams constitute an estimated 70-80% of stream length in the United States (U.S.; Downing et al., 2012). These streams are often classified by flow regime as perennial, intermittent, or ephemeral. Perennial streams have year-round flow, and nonperennial streams can be either intermittent, with seasonal flow, or ephemeral, with flow only briefly after rain or snowmelt events. Nonperennial streams encompass an estimated 51-60% of global stream length (Messenger et al., 2021) and nearly 70% of stream length in arid regions of the U.S. (Goodrich et al., 2018). The duration that a stream flows affects biogeochemical cycling of nutrients and carbon (von Schiller et al., 2017; Benstead and Leigh, 2012; Shumilova et al., 2019), availability and composition of aquatic and riparian habitat (Schilling et al., 2021; Kingsford et al., 2006), overall biodiversity (Vander Vorste and Datry, 2020), and water supply for human use (Palmer et al., 2008; Larned et al., 2010; Koundouri et al., 2017). However, the extent of these streams is predicted to increase due to climate change and land use alterations; as these systems experience increased drying, their ecosystem services will be disrupted (Palmer et al., 2008; Larned et al., 2010; Jaeger et al., 2014; Datry et al., 2014; Datry et al., 2018; Ward et al., 2020). This trend is already in effect in the southern half of the country, where the degree of intermittency in most streams has been increasing over the past 30 years (Zipper et al., 2021).

In some cases, flow classifications carry policy and management implications, and the regulatory status of nonperennial streams has been an ongoing focus of debate in the U.S. (Liebowitz et al., 2008; Acuña et al., 2014). However, it is difficult to implement regulations based on flow regimes because of a lack of available gaging data on small headwater streams (Jaeger et al. 2019; Busch et al., 2020). Furthermore, the nation's most comprehensive digital source of drainage network data, the National Hydrography Dataset (NHD; Nadeau and Rains 2007), has misclassified headwater stream types by as much as 50% in the eastern U.S. (Fritz et al. 2013).

To improve maps of stream types, researchers have developed several strategies to classify flow regimes based on climatic, physiographic, and land cover drivers. Sando and Blasch (2015) used a random forest classification model to predict streamflow class (perennial or intermittent). Jaeger et al. (2019) expanded on this approach to predict the probability of streamflow permanence in each year. Other studies focused on no-flow duration (Hammond et al., 2021; Price et al., 2021; Zipper et al., 2021; Kaplin, Blume, and Weiler, 2020) and the probability of flow at the end of the summer (Moidu et al., 2021).

Models from Hammond et al. (2021), Price et al. (2021), and Zipper et al. (2021) were developed from existing stream gauge networks, which are not evenly distributed across regions and stream types, and they have not been used to estimate flow durations throughout stream networks. Models that have been applied to entire stream networks have relied on extensive field surveys (Moidu et al., 2021; Kaplin, Blume, and Weiler, 2020), which limit the size of a study area and time period of data collection. Further, Kaplin, Blume, and Weiler (2020) did not include climatic predictors in their models, omitting an important factor for understanding nonperennial flow (Costigan et al., 2015). The PRobability Of Streamflow PERmanence (PROSPER) study is the one example in which models have been applied to estimate flow conditions throughout a large regional stream network. To do this, researchers compiled information on streamflow permanence from a wider range of stream types by combining both gage data and indirect sources such as field surveys (Jaeger et al., 2019). They used the flow status (flow/no flow) of streams at the end of the summer to classify streams as either perennial or nonperennial, then developed a random forest model to predict the probability of streamflow permanence (perennial flow) for full stream networks in the Pacific Northwest.

While the indirect data sources used to develop PROSPER can help separate perennial and nonperennial streams, many applications could also benefit from information on the flow duration and times of year most likely to have streamflow. Because of the limited gaging data on nonperennial headwater streams, few datasets exist on flow duration across a broad range of stream sizes and types. These streams are also often difficult to monitor with aerial imagery due to their small size, which makes them more likely to be obscured by vegetation cover (Fritz et al., 2013). Some studies have expanded measurements of flow duration using electrical resistance (ER) sensors, which allow detecting flow or no flow conditions at low sensor cost (Blasch et al. 2002; Chapin et al., 2014); however, there are logistical challenges to implementing and maintaining a wide array of sensors. Another useful strategy for monitoring nonperennial streams is crowdsourcing, where volunteer observers can document flow status using mobile phone apps (Kampf et al. 2018; Jaeger et al., 2021).

One of these crowdsourcing projects, Stream Tracker, has led to a large dataset of flow observations in the Colorado Front Range from 2016-2020. In this study, we examine the utility of this crowdsourced dataset for predicting streamflow duration across the region. Our study objectives are to:

1. Quantify the duration of streamflow for headwater tributaries across the elevation gradient of the Cache la Poudre watershed.
2. Examine how climate, physiographic, and land cover drivers relate to flow duration.
3. Predict mean annual and annual streamflow duration across the region.

2. METHODS

2.1 Study Area

This study focuses on the upper Cache la Poudre (CLP) watershed in the Front Range of northern Colorado, and our final models were applied to neighboring watersheds with similar characteristics, including the North Platte Headwaters, Upper Laramie, Big Thompson, St. Vrain, Clear, and Upper South Platte (Figure 1). The CLP basin is well suited for this study because it covers a wide range of climatic, land cover, topographic, and substrate characteristics. The CLP basin is 4875 km^2, although our focus at this stage is on the upper basin, which encompasses roughly 2730 km^2. This part of the basin is less influenced by land use changes and water storage or diversions, allowing us to better isolate the climatic, geologic, and vegetative influences on flow.

Headwaters of the CLP river originate in Rocky Mountain National Park, and the river runs for 203 km, eventually draining into the South Platte River. One hundred and three km of the upper CLP river system has been designated a National Wild and Scenic River. The cities of Fort Collins and Greeley, CO draw their primary drinking water from the CLP. The basin is characterized by a steep elevation gradient that impacts the climate and hydrology of the region. Elevation ranges from 1590 m to 4125 m, averaging 2560 m, and thus the CLP basin spans the Mountain and Plains hydrologic regions defined by the United States Geological Survey (Capesius & Stevens 2009). Our focus on undeveloped areas means the majority of our analysis applies to the Mountain region.

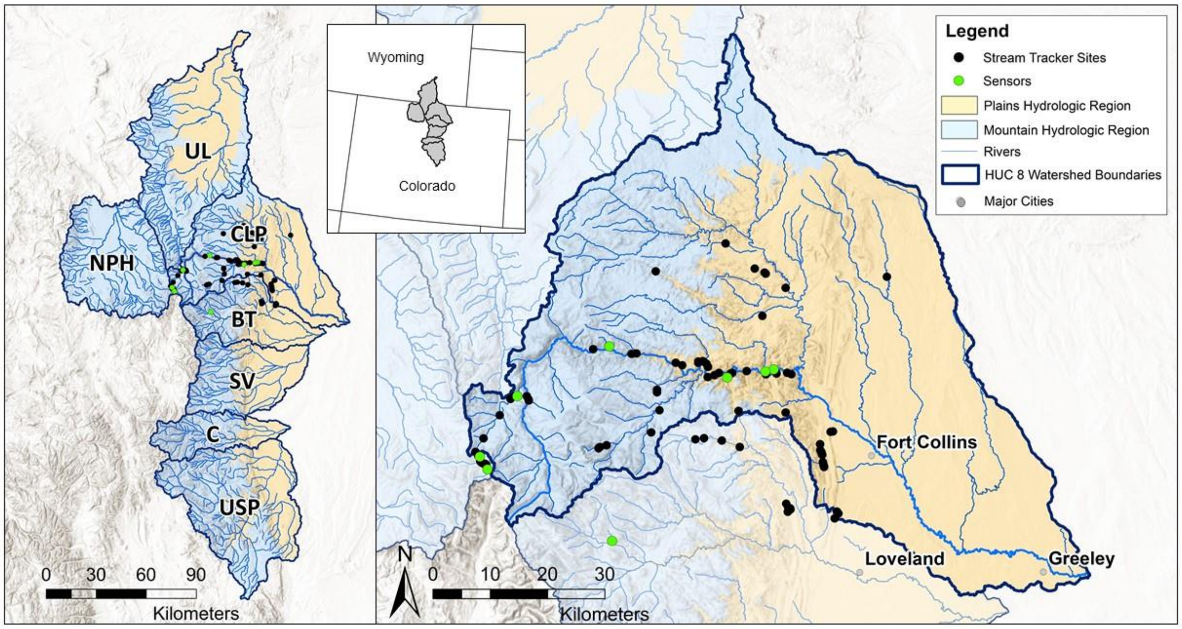

Figure 1. Map of the study area, which contains eight subbasins of northern Colorado and into southern Wyoming: the North Platte Headwaters (NPH), Upper Laramie (UL), Cache La Poudre (CLP), Big Thompson (BT), St. Vrain (SV), Clear (C), and Upper South Platte (USP). The CLP watershed is enlarged on the right. The Mountain and Plains hydrologic regions are differentiated by a 2280 m contour line. Stream Tracker and flow sensor locations used in this study are shown along with a coarse selection of NHDPlus V2 flowlines.

Precipitation of the Mountain region is snowfall dominated and averages over 1300 mm annually, while the Plains region is rainfall dominated, averaging as little as 400 mm annually (Capesius and Stevens, 2009; Kampf and Lefsky, 2016). At higher elevations, peak discharge typically occurs in early June, following snowmelt, and baseflows are sustained through the summer, supplemented by rain events. Peak discharge at lower elevations typically occurs in early April, and with peaks primarily activated by rain events (Kampf and Lefsky, 2016). Stream reaches in the upper CLP basin have slopes ranging from less than 1% (0.6°) to over 150% (56°). The mean estimated depth to bedrock from a gridded global dataset is 18 m, but soil depths vary from 2 m to 46 m in this dataset (Hengl et al. 2017) in the upper CLP basin. The dominant hydrologic soil group is D, characterized as shallow, with very slow infiltration and transmission rates and high runoff potential when saturated (Soil Survey Staff, NRCS, 2016). The lithology of the upper basin is dominated by Precambrian granitic and metamorphic rocks, and the lower basin consists of younger Permian and Cretaceous sedimentary shales and sandstones (State Geologic Map Compilation, Horton et al., 2017). There is considerable bedrock jointing in the Front Range that can impact subsurface flow and channel flow initiation (Wohl, 2008; Whiting and Godsey, 2016). Land cover changes with elevation, with tundra at the highest elevations, subalpine and montane coniferous forest in intermediate elevations, down to grassland in lower elevations.

2.2 Streamflow Data

The Stream Tracker project, developed at Colorado State University, combines a network of sensors and citizen science data collection to assess streamflow intermittence. Up to five years of streamflow data from water years 2016-2020 are available for over 200 streams in the CLP watershed. Observers report if streams have flow, no flow, or standing water by visual assessment, where flow is considered a connected surface flow path present from above to below the monitoring point. Observers can establish new sites or record observations at existing sites, which are located at road or trail stream crossings. Data are collected via a mobile phone app or uploaded onto a website (anecdata.org or citsci.org).

Stream Tracker observations are most frequent from April to September, when streams are not snow- or ice-covered. This period encompasses typical peak flows in the spring and low flows in the fall. Because these are visual observations, they are not collected at regular time intervals. To create a database with consistent month-to-month average flow fraction for all sites, we filled gaps between Stream Tracker observations by relating each stream’s flow record to streamflow sensor data. Stream stage sensors were installed in headwater catchments of the CLP and neighboring basins to record water levels across a gradient of elevation and snow persistence. At these sites, stream stage is measured with either a capacitance rod or pressure transducer. We also included United States Geological Survey (USGS) and Colorado Division of Water Resources (CDWR) sites in the study area.

We applied a gap-filling procedure to fill in flow status (flow, standing water, no flow) for days without observations and then computed the April-September fraction of days with flow for each year from 2016-2020. Stream Tracker observations were split into training (2016-2018) and testing (2019-2020) datasets. We used a rule-based procedure to identify no flow and perennial streams for the training dataset. If flow was recorded for all observations, and at least one of these observations was collected during the driest month, September, the site was considered to have perennial flow (flow fraction = 1). If no flow was ever recorded at the site, and observations of no flow were collected in the wettest months, April or May, then we assumed all days over the analysis period had no flow (flow fraction = 0). For all other sites, we used flexible discriminant analysis (Hastie et al. 1994) to develop models for predicting flow, standing water, or no flow for each day with the mda package in R (Hastie and Tibshirani 2020). Predictor variables were normalized streamflow from stream stage sensors and day of year. Models predicted the flow, standing water, or no flow condition at each Stream Tracker site at a daily time step. This approach produced models with a Nash-Sutcliffe efficiency coefficient of 0.88 for the training time period and 0.85 for the testing time period, when comparing the observed and simulated April-September flow fractions across all sites. This process resulted in 82 gap-filled Stream Tracker sites (Figure 1); sites that did not have observations during multiple years and multiple times of year could not be gap-filled.

The gap-filled data were summarized into 2016-2020 April-September average flow fraction (referred to as mean annual) and yearly average (annual) flow fraction (ff), defined as the number of days with flow divided by the total number of days in a given time period. Flow fraction is, therefore, directly related to flow duration. Standing water predictions were grouped with no flow rather than flow. We designated flow fraction classifications as nonperennial (ff<0.92), and perennial (ff>=0.92) for later comparison of predicted flow fractions to NHD stream type classifications. The cutoff for perennial streams was chosen because known perennial streams, such as the Cache La Poudre river, had predicted flow fractions >=0.92.

2.3 Basin Attribute Data

To evaluate drivers of streamflow duration, we compiled climatic, topographic, subsurface, and land cover attributes for the CLP and neighboring watersheds in north-central Colorado. A total of 268 basin attributes were compiled as potential predictors of flow fraction, drawing from freely available data sources (Table 1). For the mean annual model, climate attributes included mean annual, mean seasonal, and mean monthly values of precipitation, temperature, and snow persistence. The annual model included annual and seasonal values for each year with Stream Tracker data (2016-2020). Physiographic and land cover attributes were the same for both models. We chose to evaluate attributes at a 30 m resolution, selected over higher resolutions because it results in fewer problems with disconnected flow lines (Hastings and Kampf, 2014) and is consistent with the medium-resolution NHDPlus V2 (from here on referred to as NHD MR). The watershed boundaries used to delineate the study area were 30 m resolution HUC8 subbasins from the Watershed Boundaries Dataset (WBD) included in the NHD MR. The 10 m flow accumulation and flow direction data from the higher-resolution NHDPlus HR (or NHD HR), in beta at the time of this study, were not readily available for our use.

2.3.1 Climate

In the mountainous watersheds of the northern Colorado, snow is a dominant driver of streamflow (Kampf and Lefsky, 2016; Doesken and Judson, 1996; Stewart et al. 2004). Consequently, the snow cover-derived attribute snow persistence (SP), defined as the fraction of time that snow is present on the ground, can be a strong predictor of streamflow (Eurich et al. 2021; Hammond et al. 2018). We used the USGS mean annual 500 m-resolution annual average SP product (Hammond, 2020), which is averaged over the 2001-2020 period of record, corresponding to MODIS snow cover data availability (Hall & Riggs, 2016). We also used mean monthly SP data for each month from January-June; these were averaged only over the five-year period of 2016-2020.

Precipitation (P) and temperature (T) data from the PRISM Climate Group, and potential evapotranspiration (PET) data from gridMET, all at 4 km spatial resolution, were also computed for the 2001-2020 period of record. Annual values were assessed by water year (Oct 1 of the previous year to Sept 30) instead of calendar year because water years are better suited to represent the timing of snow accumulation and melt. Seasonal values were computed for spring (March-May), summer (June-September), fall (October-November) and winter (December-February). September was included in the summer season to avoid bridging multiple water years for the fall calculations. We also derived surplus/deficit (SD = P-PET) and aridity index (AI = P/PET) from P and PET data. To reduce the total number of inputs in our annual model, we summarized each attribute seasonally rather than monthly (Table 1).

2.3.2 Topography

Topographic basin characteristics such as drainage area, average catchment elevation, and dominant aspect, slope, and curvature can influence flow generation (Montgomery and Dietrich, 1988; Tarboton et al., 1991; Heine et al., 2004; Clubb et al., 2014). For this study these elements were derived from a USGS National Elevation Dataset (NED) 1 arc-second (30 m ground distance) DEM. To represent channel topographic characteristics at the point of measurement, we included topographic wetness index (TWI; Williamson et al., 2015), slope, and geomorphon shape for the pixel representing each stream tracker point. TWI was calculated using the NHD MR flow accumulation layer and the slope layer derived from the 30 m DEM (Beven and Kirkby 1979; Sørensen et al., 2006). Geomorphons describe the morphology of bare-earth terrain as one of 498 unique shapes, the 10 most common being: flat, peak, ridge, shoulder, spur, slope, pit, valley, footslope, and hollow (Jaziewicz and Stepinski 2013). We processed the 30 m DEM using the 'r.geomorphon' extension in GRASS GIS open-source software (GRASS Development Team 2017), following the method described by Baker et al. (2018) in their analysis of the Chesapeake Bay watershed. Instead of compiling the dominant geomorphon shape of the catchment, we used a 180 m (6 pixel) and 300 m (10 pixel) buffer/search of geomorphon type to describe the potential for local flow.

2.3.3 Subsurface Properties

We characterized underlying geology by bedrock type (igneous, metamorphic, sedimentary, and unconsolidated). Other studies have attempted to include a factor of bedrock permeability (Jaeger et al., 2019; Kaplan et al., 2021), but we chose to use generalized bedrock type as a predictor because a recent study in the CLP watershed found that flow duration changed at the interface of metamorphic and sedimentary bedrock (Martin et al., 2021).

Depth to bedrock (ISRIC SoilGrids250m, Poggio et al., 2021) and available water storage (Soil Survey Staff, gSSURGO) account for groundwater storage potential. Other soil characteristics that we included from the gSSURGO database were hydraulic conductivity, soil texture indicators (percent sand, silt, and clay), and the NRCS-defined hydrologic soil group (HSG), which characterizes the infiltration capacity of soils.

2.3.4 Land Cover

To categorize land cover types, the Landscape Fire and Resource Management Planning Tools' (LANDFIRE) existing vegetation type categories were condensed into eight lifeform groups: tree, shrub, herb, sparse, barren, developed, agriculture, and open water.

Monitoring Trends in Burn Severity (MTBS) wildfire and prescribed burn boundaries were included in this study and split into five-year age groups. Of the nine fires included in our study area, The High Park fire of 2012 is the largest, encompassing 90,769 acres. The oldest fire boundary is a small 1547 acre burn from 1989. Six predictor variables were compiled: all boundaries and age groups 0-5, 6-10, 11-15, 16-20, and 25+ years old.

Table 1. Basin attributes as possible predictor variables and data sources.

Characteristic	***Attribute***	***Source***
Climate	Precipitation (mm), Temperature (°C) (mean annual, mean seasonal, mean monthly, annual, & seasonal)	PRISM (Daly, 2013)
	Potential evapotranspiration (mm) (mean annual, mean seasonal, mean monthly, annual, & seasonal)	gridMET (Abatzoglou, 2013)
	Aridity index (unitless), surplus/deficit (mm) (mean annual, mean seasonal, annual, & seasonal)	Calculated from PRISM and gridMet data
	Snow persistence (%) (mean annual, mean monthly, annual, & monthly)	Hammond, 2020
Topography	Mean curvature, dominant aspect (0-360°), mean elevation (m), mean slope (m/m), local slope (m/m), drainage area (km^2)	USGS National Elevation Dataset (NED) 30 m (Gesch et al., 2002)
	Local topographic wetness index	Calculated from USGS NHD Plus V2 and NED
	Local geomorphons	NED, processed in GRASS
Geology	Igneous, metamorphic, sedimentary, unconsolidated (%)	State Geologic Map Compilation (Horton et al., 2017)
	Depth to bedrock (cm)	ISRIC SoilGrids250m (Poggio et al., 2021)

	Available Water Storage (mm), hydraulic conductivity (μm/s), percent sand, silt, and clay, hydrologic soil group (%)	gSSURGO
Land Cover	Tree, shrub, herb, sparse, barren, developed, agriculture, open water (%)	LANDFIRE
	Fire Boundaries (%) (all boundaries, 0-5, 6-10, 11-15, 16-20, 25+ age bins)	MTBS

2.4 Compiling Watershed Attributes

Average values for each attribute were compiled for the drainage areas above each flow observation point. The Flow Accumulation tool in the ESRI ArcMap Spatial Analyst toolbox was used to conduct weighted flow accumulations on all watershed attribute rasters. Attributes such as temperature and surplus/deficit that contained negative values were processed through T.E. Dilts' (2015) Flow Accumulation for Both Positive and Negative Values Toolbox, developed at the University of Nevada Reno. Output values were divided over the number of cells that flow into each cell (the NHD MR standard flow accumulation) to compute the average value of attributes for each pixel's drainage area. These averaged flow accumulations have been coined Continuous Parameter Grids (CPGs) in similar studies (Sando et al., 2018; Jaeger and Sando, 2019). Local slope, TWI, and geomorphons were the only predictors not processed through weighted flow accumulations; these values were taken from the pixel containing the stream monitoring locations.

All CPGs were snapped to the same extent, pixel size, and geographic coordinate system as the NHD MR flow accumulation to align the pixels for accurate attribute extraction and geospatial analysis. We extracted all attribute values underlying each observation point and used these values for model training and testing.

2.5 Model Development

We used random forest regression models (Breimen, 2001) to predict mean annual and annual flow fractions, following the approach of similar studies (Jaeger et al., 2019; Price et al., 2021; Moidu et al., 2021), and specifically the model structure used by Zipper et al. (2021b). This approach has been useful for hydrological modeling because random forests handle large numbers of predictors well without overfitting and allow easy interpretation of variable importance (Eng et al., 2017; Addor et al., 2018; Miller et al., 2018).

Stream Tracker observation sites were randomly split into 80% training and 20% testing datasets. Separately, the streamflow sensor locations were also split in an 80/20 ratio and added to the training and testing datasets. These random assignments were edited slightly to ensure that both training and testing datasets had even representation across the elevation gradient because the Stream Tracker dataset is imbalanced in the number of sites for lower and higher elevations.

Random forest models are biased toward categorical variables with large numbers of levels (Strobl et al., 2007). Therefore, to reduce bias and complexity of the model, we converted categorical data such as bedrock, land cover, and hydrologic soil type to continuous rasters with binary presence/absence values and calculated their percent coverage with weighted flow accumulations. For instance, instead of having one "Geology" variable with multiple categories, we isolated each bedrock type before processing it into an individual predictor variable representing the percent of the contributing area contacting that rock type. This also allowed for easier isolation of streamflow drivers when completing analysis.

Predictor variables were checked for collinearity before being included in the models. We ranked variables by their correlation to the flow fraction of interest; for each group of highly correlated variables (Pearson correlation values >0.9), we retained the one with the highest correlation to flow fraction and dropped the others. Climate variables showed the highest instances of collinearity, while nearly all topographic, subsurface, and land cover variables had cross-correlations <0.9 and were retained. Elevation had the greatest correlation to flow fraction in the annual model, but was excluded to better evaluate the influence of climate variables that influence interannual variability in flow.

After adjusting for collinearity, 29 of the 105 predictor variables were initially included in the mean annual model. The number of variables was further reduced by ranking the variables in order of importance and determining how many were necessary to minimize model error, following Zipper et al. (2021; Figure 2). Of these, 11 predictors were selected to optimize model performance (Figure 2a). Similarly, 49 of the 69 predictor variables were included in the annual model after collinearity analysis. Error generally reduced with more predictors, but that reduction was minimal after 23 predictors (Figure 2b).

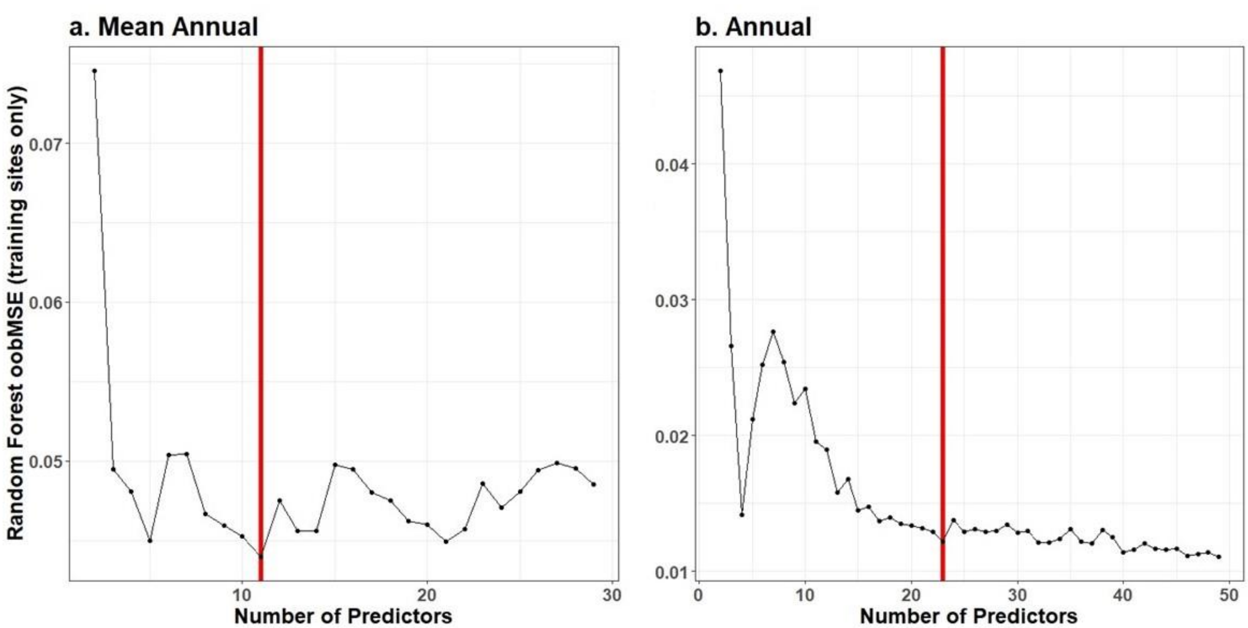

Figure 2. Random Forest out of bag mean squared error (oobMSE) as a function of the number of predictors included. The red line indicates the number of predictors selected for (a) the mean annual model, 11 predictors, and (b) the annual model, 23 predictors.

The random forest models were tuned and executed using the ‘tidymodels’ package in R (Kuhn & Wickham, 2020). We then applied the models using the ‘raster’ package (Hijmans,

2021) to output predictions of annual and mean annual flow fractions across the study area. The prediction ouptut files were geotiffs with an April-September flow fraction prediction for each 30 m cell in the study area. We then trimmed the output down to a 60 m (2 pixel) buffer around NHD HR flowlines to display the results and compare with flow classifications in the NHD.

3. RESULTS

3.1 Model Accuracy

The random forest regression models we developed to predict mean annual and annual flow fraction showed strong performance. The mean annual model for the training sites had a Nash-Sutcliffe efficiency (NSE) = 0.96 and percent bias (PBIAS) = 0. For the testing sites, this model had an NSE = 0.88 and PBIAS = -3.9. The annual model for the training sites had an NSE = 0.99 and PBIAS = -0.2, and the test data had an NSE of 0.81 and PBIAS = -6.8 (Figure 3). Negative PBIAS values indicate slight overestimation biases in both models (Moriasi et al., 2007). The models successfully predicted flow fraction along the full range of zero to one, performing well on mid-level flow fractions that represent intermittent flow regimes.

Both models were applied for the same set of 90 flow observation sites, but the annual model's inputs included flow fractions and annual climate variables for each of the five years in the study (five sites had only two years of data), resulting in 435 unique inputs (Figure 3b).

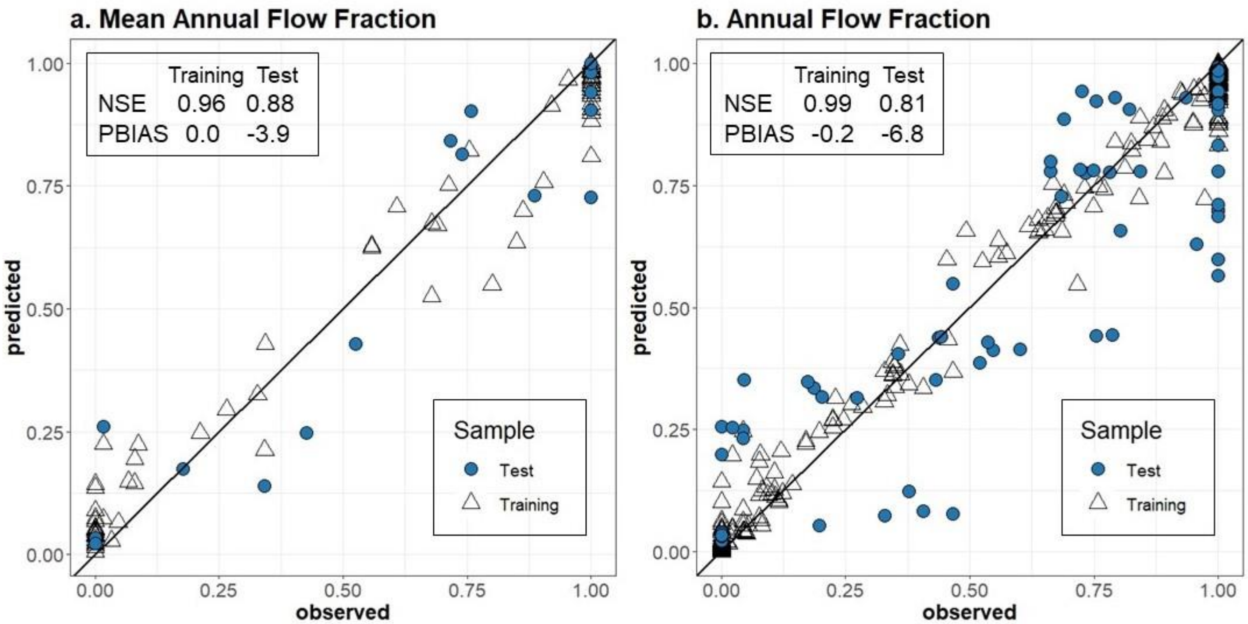

Figure 3. Observed vs predicted flow fraction comparison for mean annual and annual models. Hollow triangles represent training data, and blue circles represent test data. The black line is a 1:1 line to show model fit.

3.2 Identifying Significant Drivers of Flow Duration

Both mean annual and annual flow fraction models identified climate attributes as significant drivers of streamflow duration in north-central Colorado. Most notably, a factor of SP ranks as the most important predictor in both models. The mean annual results show that 6 of the 11 predictor variables (55%), including the top five, are climatic variables (Figure 4a). SP and P were positively correlated with flow fraction, whereas PET was negatively correlated with flow fraction. Forested and herbaceous land cover also influenced the mean annual model, as the 6^{th} and 7^{th} most significant drivers, respectively. Forested land cover was positively correlated, while herbaceous was negatively correlated to flow fraction. Ten of the 23 predictors (44%) in the annual model are climatic, including the top two most significant predictors (Figure 4b). Again, herbaceous and forested land cover influence this model as well.

Subsurface characteristics were also identified as important drivers of flow duration. These include hydraulic conductivity, available water storage, drainage area, hydrologic soil group, and depth to bedrock. Higher flow fractions were found with greater hydraulic conductivity and HSG A, whereas depth to bedrock, available water storage, HSG D and C were negatively correlated to flow fraction. Topographic characteristics selected include curvature and local TWI in the annual model, but none ranked in the top 11 predictors for the mean annual model. Bedrock type was not selected for the models, with the exception of percent sedimentary bedrock in the mean annual model and percent metamorphic in the annual model, but these variables had minimal effect on model MSE.

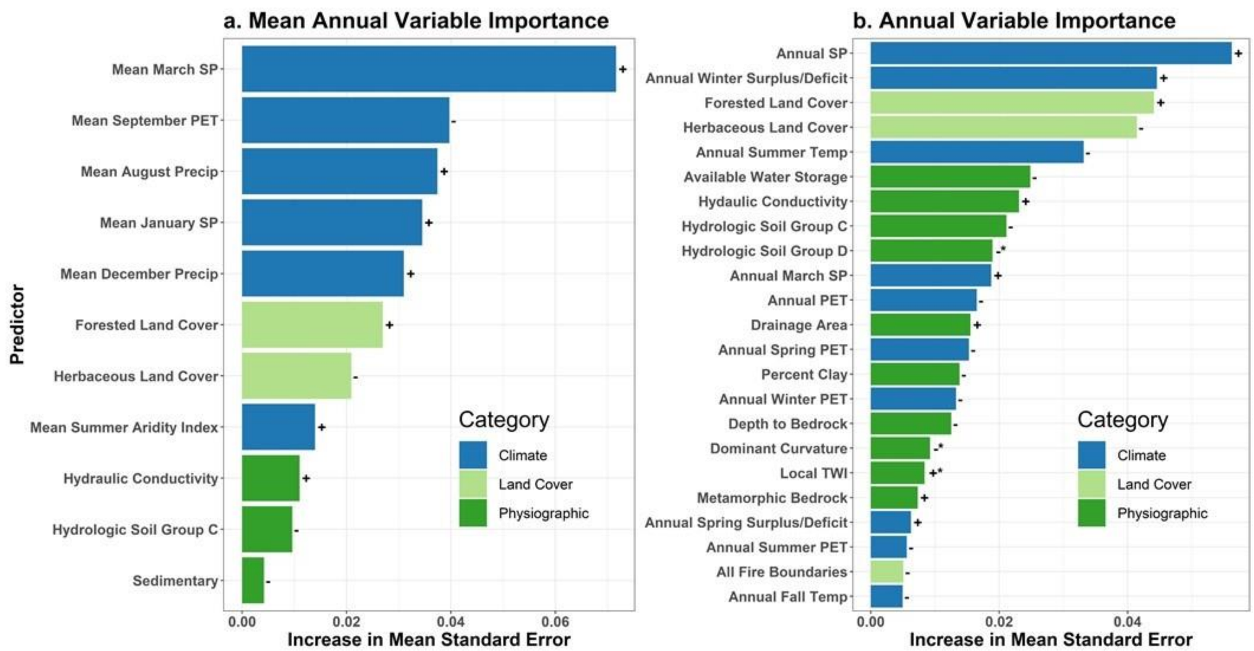

Figure 4. Mean annual and annual variable importance for predicting flow fraction, displayed as the increase in mean squared error when removed from the model. "+" indicates positive correlation to flow fraction, "-" indicates a negative correlation, and correlations marked with "*" are insignificant. For plotted predictor correlation to flow fraction, see Appendix A.

The influence of climate variables on simulated flow fractions is evident in the semi-gridded pattern of flow predictions in some areas; this pattern matches the 4 km pixel resolution of PRISM precipitation data (Figure 5). Finer, 800 m-resolution PRISM climate datasets are available, but because they are much larger files and require an acquisition fee, we elected to use the coarser-resolution product. The resolution effect along stream lines is minor, and we did not see abrupt or extreme changes in predicted flow fraction along the majority of PRISM pixel boundaries (Figure 5c).

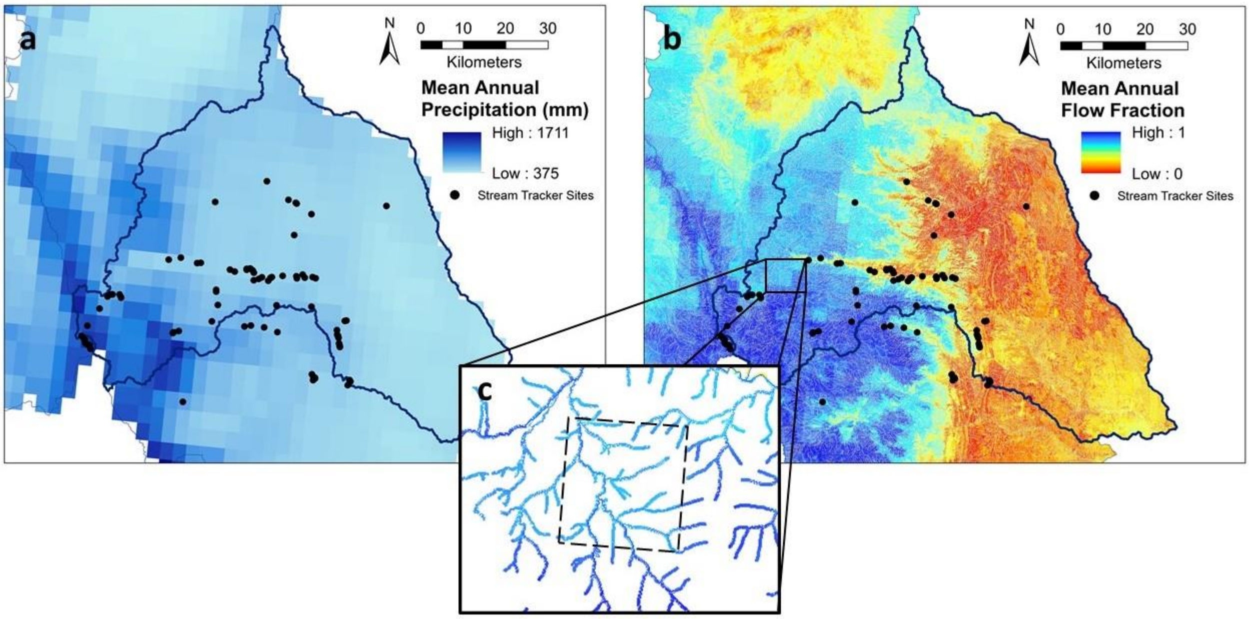

Figure 5. The 4 km-resolution PRISM mean annual precipitation (A) compared to mean annual flow fraction predictions (B), shown here prior to being clipped to flowlines to emphasize the pattern and distribution of flow fraction predictions. Panel C shows predictions along flowlines at a PRISM pixel boundary. Flow fraction predictions along larger streams are not disrupted at pixel boundaries (blue line of perennial flow), but predictions along tributaries show some abrupt changes at these boundaries.

3.3 NHD Assessment

To isolate predictions along flowlines, we clipped to NHD HR flowlines with a 60 m (2 pixel) buffer. Because our model used the NHD MR flow accumulation and flow direction datasets, our predicted flowlines do not precisely line up with the NHD HR data. However, the NHD MR data lacks flowlines that correspond to our sensors and ST sites on headwater streams. After various trials, we deemed this NHD HR 60 m buffer method to be the most effective at displaying our overall results (Figure 6). Potential predicted flowlines that are not included in the NHD HR dataset are unfortunately not displayed using this method. Further, the misalignment of our model's and the NHD HR's flowlines, along with the complexity of the 60 m buffer, makes comparing the length of perennial, intermittent, and ephemeral channels between NHD HR and our model extremely difficult. For this reason, we did not compare the NHD HR dataset's flow classifications to our predictions in this study. However, we conducted a basic analysis of the NHD MR.

Table 2. Matrix of NHD MR classifications vs the mean annual model predictions in the CLP basin. Perennial FCode = 46006, intermittent FCode = 46003, and ephemeral FCode = 46007 (Moore et al., 2019).

		NHD Classifications			
		per	**int + eph**	Model total	**Model percent**
Mean Annual Predctions	**per**	10952	1326	12278	10
	int + eph	35138	71332	106470	90
	NHD total	46090	72658	118748	
	NHD percent	39	61		
	NHD Accurate	89	67		
	NHD False Positive (%)	286	1		
	NHD False Negative (%)	7	34		

Overall accuracy = 69%

We converted the NHD MR flowlines to 118,748 points spaced 30 m apart and extracted the mean annual model predictions to those points for comparison (Table 2). We defined ephemeral streams as having flow fractions < 0.05, intermittent flow fractions between 0.05 and 0.92, and perennial flow fractions >= 0.92. The upper cutoff for perennial streams was chosen because known perennial streams, such as the Cache La Poudre river, had predicted flow fractions >=0.92. Because the NHD MR was designed to exclude ephemeral streams, we grouped intermittent and ephemeral classifications for our analysis. In the CLP basin, NHD MR flow classifications matched our mean annual model's classifications 89% of the time for perennial streams and 67% for nonperennial streams. The NHD predicted nearly four times more perennial streams than our model, while only 11% of the perennial streams identified by our model were classified as nonperennial in the NHD. The NHD and our model matched class for the majority of nonperennial streams, but 33% were misclassified as perennial in the NHD.

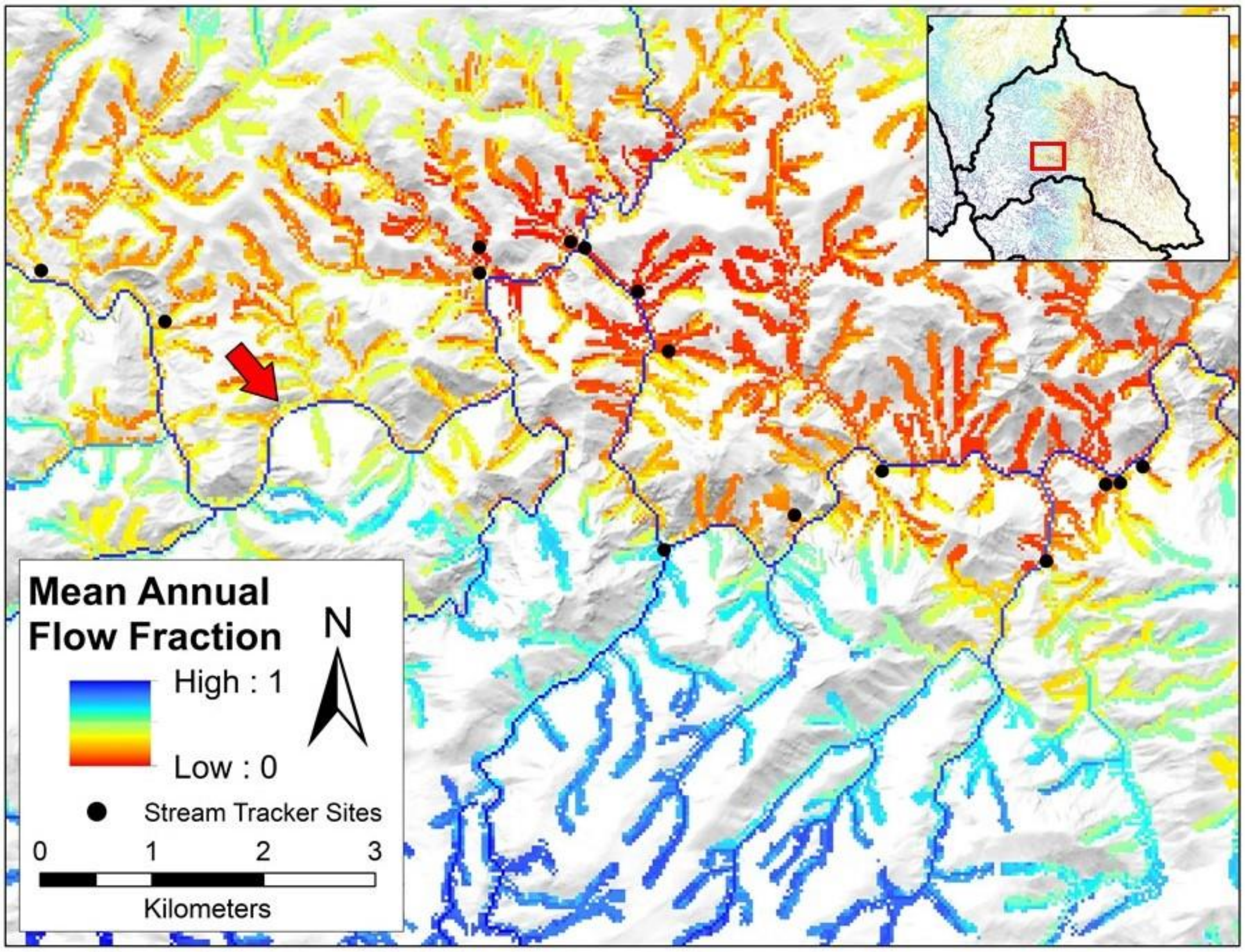

Figure 6. Mean annual flow fraction predictions clipped to a 60 m buffer around NHDPlus HR flowlines. This buffer captures most of the important flowlines, but can miss them in spots (red arrow). Missing pixels within the buffer area represent "no data" values.

3.4 Wet vs. Dry Water Years

The goal for our annual model was to capture interannual variability in flow fraction, which would allow it to be used as a predictive tool in the coming years. To assess the applicability of our model for representing this variability between years, we compared the output from wet and dry years. The average precipitation for the CLP basin in 2017 was 515 mm, and in 2020 it was 396 mm, respectively the highest and lowest totals during our study period (Table 3; PRISM Climate Group 2021). Although 2017 had the highest P total, it also had the lowest SP of the five years. 2016 had both high P and SP values, therefore, we chose to compare 2016 and 2017 (wet years) to 2020 (dry year). The wettest year based on AI (2017) had the highest flow at a low elevation stream gauge (Bighorn), while the wettest year in terms of SD (2019) had the highest discharge at a high elevation site (Michigan). The driest year for AI and SD (2020) was also the driest year for Michigan, but not for Bighorn (Table 4).

Table 3. Interannual variability of climate characteristics in the Cache La Poudre basin during the 2016-2020 study period. Aridity Index (AI) is a unitless metric where values closer to 0 indicate greater aridity. Surplus/deficit is a moisture balance calculation of precipitation minus potential evapotranspiration, and more negative values suggest greater water losses.
* indicates wettest year based on the column variable
† indicates driest year based on the column variable

Year	*P (mm)*	*PET (mm)*	*T (°C)*	*SP (%)*	*AI*	*SD (mm)*
2001-2020	550	1234	6.5	35	0.46	-684
2016	487	1235	6.9	42*	0.39	-748
2017	515	1241	7.3	29†	0.42*	-726
2018	450	1307	7.3	31	0.35	-857
2019	468	1187	5.9	38	0.39	-719*
2020	396	1286	6.4	37	0.31†	-890†

Table 4. Annual discharge at Bighorn and Michigan flow sensors during the 2016-2020 study period. The other six sensors did not have complete records for all five years. The Bighorn sensor is at 2,922 m elevation and drains a catchment area of 3.4 km^2. The Michigan sensor is at 3421 m and drains 4.0 km^2.

Year	*Bighorn Q (mm)*	*Michigan Q (mm)*
2016-2020	26	606
2016	35	552
2017	37*	616
2018	20	559
2019	16†	767*
2020	22	537†

Because the wet and dry years were not consistent across the study area, the comparisons between years show both positive and negative differences in flow fraction. Subtracting the 2020 flow fraction from 2016 resulted in a difference range from -38% to 49%, where negative values indicate an increase in flow fraction for 2020. Comparing 2017 and 2020 results in a larger range: -44% to 64%. The majority of pixels show close to zero difference between wet and dry years, but our model was able to capture isolated areas at both extremes (Figure 7c and 7d).

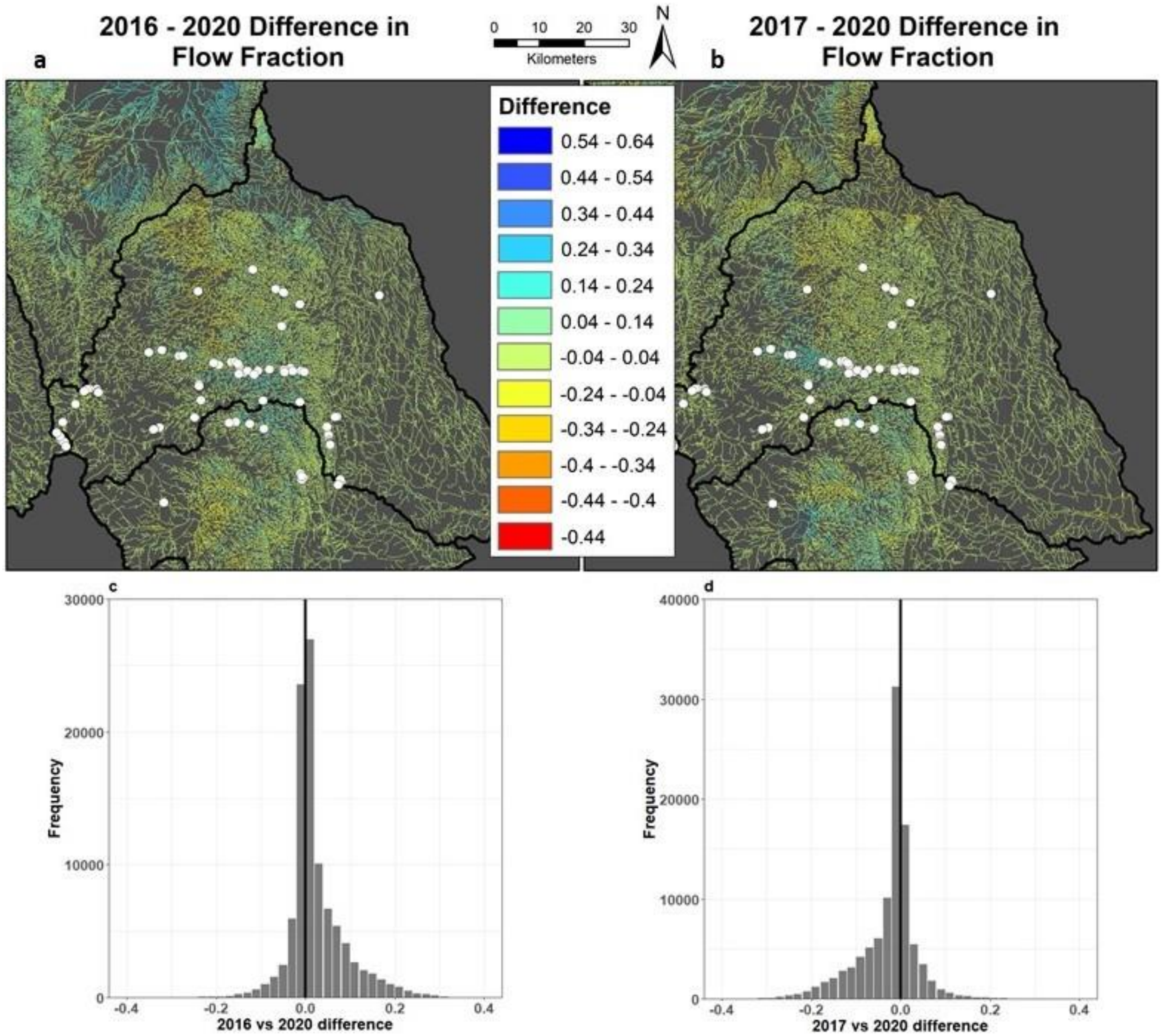

Figure 7. Wet and dry year comparisons. Panels a and c show the difference between 2016 and 2020 flow fraction values, while b and d compare 2017 and 2020. Positive values indicate higher predicted flow fraction in the wet year, and negative values indicate higher flow fraction in 2020, the dry year. Panels c and d show a sample subset of pixels in each raster.

4. DISCUSSION

4.1 Stream Tracker Data

Our first objective for this study was to quantify the duration of flow in headwater tributaries across the elevation gradient of the Cache la Poudre watershed. We accomplished this by using both streamflow observations from the Stream Tracker program and eight flow sensors. Our findings indicate that crowdsourced flow presence/absence data, when bolstered by a small network of sensors, can be effectively used to improve our knowledge of headwater stream intermittency. Crowdsourcing is a cost-effective method for gathering large amounts of data, while engaging the public as citizen scientists and stream stewards. The Stream Tracker program captures flow data from small, ungauged streams that can be difficult to monitor regularly. By implementing a gap-filling approach using accompanying sensors, we created a larger dataset of streams for which we have reasonable estimates of mean annual and annual flow duration.

We also evaluated whether this approach could be used to reconstruct mean monthly flow fractions from April to September, but this yielded poor results. We attribute this to a smaller sample size of monthly Stream Tracker observations to assess the model as well as to the seasonal patterns of flow in this region. Wet months tended to have only flow observations whereas dry months had only no flow observations; this gave less information to train models for flow fractions between 0 (no flow) and 1 (all flow).

Another limitation of crowdsourced data is that observation sites tend to be located primarily in easy-to-access locations. There are relatively few roads in the Colorado Front Range, and Figure 1 shows most of our ST points are located along the CLP canyon, where observations can be made close to a road; this data gathering strategy misses remote areas. The location of flow observations influences our model, as we had a much smaller model training dataset at higher elevations (Figure 5). Deploying a larger sensor array in these remote areas could help improve our output. ER sensors may be useful in this capacity because their small size and portability lends them to widespread deployment (Chapin et al., 2014; Costigan et al., 2015).

4.2 Drivers of Flow Duration

Our next objective was to identify and evaluate drivers of streamflow duration / flow fraction in north-central Colorado. Random forest modeling results indicate that climate attributes were the most significant drivers of flow duration in our models, which is in agreement with most other studies of nonperennial streamflow (Sando and Blasch, 2015; Jaeger and Sando, 2019; Hammond et al., 2021; Zipper et al., 2021; Moidu et al., 2021). This is in contrast to Price et al. (2021), who found land use and physiography as more significant drivers of flow on nonperennial gaged streams across the U.S.

The highly significant influence of snow persistence (SP) in both models in this study corroborate the findings of Eurich et al. (2021) and Hammond et al. (2018), who found that SP can be effectively used as the sole snow-related variable in streamflow prediction. This has implications for addressing climate change; monitoring SP can tell us where streams may be vulnerable to increased drying due to declining snow cover (Hammond et al., 2018). SP also relates to streamflow timing, where streams with earlier snow loss tend to have earlier streamflow rise in spring (Harrison et al., 2021).

In our mean annual model, we found a strong influence of precipitation and potential evapotranspiration in the summer months (June-September); we identified mean September PET, mean August P, and the P- and PET-derived aridity index for summer as significant drivers.

Streams in this region tend to dry during the summer months, and the extent of rainfall inputs and PET during this season can best help us predict mean flow fraction. In contrast, the annual model was influenced by primarily winter and spring climate attributes, suggesting that variability in winter precipitation is more important when predicting flow fraction from year to year.

Subsurface characteristics displayed a variety of influences on flow duration. Both depth to bedrock and available water storage are negatively correlated to annual flow fraction (Figures A2p and A2f). Greater water storage can allow less water to be exported to streams, which decreases flow duration. The extent of this effect may be dependent on soil characteristics. Both models were influenced by hydraulic conductivity, which is positively correlated to flow fraction. In addition, the hydrologic soil group (HSG) C, which is characterized by fine-textured soils and very slow infiltration and transmission rates (Soil Survey Staff, NRCS, 2016) was also utilized by our models and has a negative correlation to flow fraction. Ultimately, this suggests that the ability to transmit water through the subsurface helps sustain streamflow.

Bedrock permeability and jointing can impact flow paths throughout a watershed (Whiting and Godsey, 2016), but we did not include proximity to joints or faults in this study because of the difficulty of compiling that information into CPGs required by our models. We did evaluate bedrock type, and sedimentary and metamorphic bedrock were low on the list of variable importance. Percent sedimentary bedrock was included in the mean annual model, showing negative correlation to flow fraction, and metamorphic bedrock was included in the annual model and positively correlated with flow fraction. This result is supported by Martin et al. (2021), who found a stark change in flow at the boundary of metamorphic and sedimentary bedrock, where flow initiated at a fault line in the metamorphic bedrock, and upon reaching the more permeable sedimentary conglomerate, flow ceased.

Local slope and geomorphons were not significant predictors of flow duration in our models, and although local TWI was included in the annual model, it had an insignificant correlation to flow fraction. Our evaluation of these variables may have been too localized, and they may need to be assessed at the stream reach scale to be captured in our models. Alternatively, because faults and other subsurface qualities are an important source of streamflow in our study area, terrain indices may not work as well for predicting flow duration as they do in wetter climates with more regionally continuous water tables. Also, our sampling may have been biased relative to the distribution of geomorphons because flow observations were often concentrated along roads, which are typically at valley bottoms.

Forest and herbaceous land cover categories were strong predictors in both of our models, where forested land cover had a positive influence on flow fraction, and herbaceous had a negative influence. The division between forest and herbaceous vegetation in this region follows an elevation gradient, so the increased flow fractions in forested areas also correspond to higher elevations and greater snowfall influences. The other land cover factor included in this study, fire boundaries, had a negligible influence on our annual model. This study’s data collection started in 2016, four years after the High Park Fire; since fire effects on streamflow in this region tend to persist for about three years post-fire (Wilson et al. 2018), the data collection period may have been too late to detect a fire effect. Burn severity was not included in this study, but based on our results, we do not expect it would be a significant predictor of streamflow duration. Even so, the flow prediction output of this model will likely be impacted by the 2020 Cameron Peak Fire that burned over 200,000 acres in the CLP and BT watersheds. Including this fire boundary in future studies could be useful for determining whether fire does affect flow duration in the years immediately following the burn.

4.3 Applicability of Models

4.3.1 Flow Duration Predictions and Spatial Output

The final objective of this study was to predict mean annual and annual streamflow duration across the study region. Both models accurately predicted flow fraction for both training and testing sites with strong performance (Figure 3; Moriasi et al., 2007). Based on visual assessment of Figure 3, the models tended to overpredict low flow fractions, and underpredict higher flow fractions, suggesting that intermediate flow fractions corresponding to intermittent flow regimes are best captured by these models.

Interannual variability is captured in our annual model, although when comparing wet and dry years, many parts of the study area had very little variability in flow fraction between years. We could, however, focus on the small number of pixels with greater variability as areas most vulnerable to climate change. These areas are primarily in mid-to-low elevations (Figures B1 and B2), where it is expected that stream flow is more sensitive to timing of input and to individual storms. Our study period overall had below-average precipitation in all years (Table 3), so this likely means our observation data were biased toward low flow fractions. As data are collected over more years in the future, this analysis could be updated to determine how the model is affected by a broader set of climate conditions in the training dataset.

Our comparisons of wet vs dry years highlighted the spatial variability of winter precipitation and its influence on flow fraction. Areas with greater SP and winter surplus-deficit in 2020, the dry year, had greater flow fraction when compared to 2016 and 2017, the wet years. Understanding this variability in snowfall and snow persistence is key in addressing the effects of climate change on streamflow (Hammond et al., 2018).

Intermittent streams are often characterized by disconnected flow paths, and dry patches can vary widely over space (Godsey and Kirchner, 2014). Our spatial output is at the NHD MR 30 m pixel scale and lacks the resolution needed to show variability in flow duration along most stream lengths. Our model does not represent short drying patches (e.g. a 4 pixel, or 120 m segment), as our predictions are fairly uniform for entire headwater stream reaches. Future work could merge this coarse modeling with finer-resolution topographic information to predict intermittency within stream reaches.

4.3.2 NHD Accuracy Assessment

The models in this study were developed to improve flow classification along existing NHD flowlines. The NHD MR's flowlines were primarily derived from hand-delineated stream lines on 1:24,000- and 1:100,000-scale topographic maps (Dewald 2017; Fritz et al. 2013). This method excluded most ephemeral channels and some short stream segments deemed too small or too close to watershed boundaries for mapping (Chorley and Dale 1972; Drummond 1974; Mark 1983, USGS 1999, 2009). The NHD MR has been shown to greatly underestimate the extent of headwater streams (Fritz et al., 2013), while the NHD HR has been shown to overpredict these streams in our study area (Martin et al., 2021). Even where channels are mapped correctly, they can have inaccurate flow classifications in the NHD (Fritz et al., 2013). The NHD HR dataset improves upon the NHD MR by using only the NDH MR's high-resolution data (1:24,000-scale or better) and incorporating the national Watershed Boundary Dataset (WBD), and a ⅓ arc-second (10 m) 3D Elevation Program (3DEP) DEM (Moore et al 2019). However, DEM-defined channel initiation points do not always align well with channel heads identified in the field

(Hastings and Kampf, 2014), and legacy NHD inaccuracies of headwater stream extents and classifications persist in this new dataset. The models we developed do not have the capacity to define flow paths that could improve the mapping of headwater channels, but, as stated above, finer-resolution topographic information could be utilized to solve this issue.

We predicted nonperennial flow for 76% of NHD MR-designated perennial streams (Table 2). This result supports previous studies that point out the inaccuracy of NHD flow classifications in headwater streams (Fritz et al., 2013; Hafen et al., 2020). Based on our predictions, the NHD MR under-represents intermittent and ephemeral streams by 32% and over-represents perennial streams by three times in the CLP basin. This type of comparison could also be conducted with NHD HR if the issues with inconsistencies in flow paths between NHD HR and our predicted flow paths can be resolved. Ultimately, our findings can be used to update the NHD with improved stream classifications.

5. CONCLUSIONS

In this study, we predicted mean annual and annual flow duration across the stream networks of eight subbasins in north-central Colorado by utilizing a crowdsourced dataset of flow observations. We used a random forest regression modeling approach to create a map of flow fractions from April-September, and we ranked the importance of predictor variables to evaluate how climatic, physiographic, and land cover drivers affect flow duration.

Climate variables were the most significant predictors of flow fraction in each model. Flow fraction increased with snow persistence, precipitation, moisture surplus/deficit, and forest cover, and decreased with potential evapotranspiration, temperature, and herbaceous land cover. The subsurface characteristics of hydraulic conductivity, hydrologic soil group, depth to bedrock, and available water storage demonstrated that higher transmissivity and lower water storage resulted in higher flow fractions. Underlying bedrock had minimal influence on our models, but sedimentary bedrock, which is typically permeable, was negatively correlated with flow fraction. Our mean annual model predictions suggest that the medium-resolution NHDPlus V2 underrepresents nonperennial streams in the Cache la Poudre watershed, and that 76% of streams classified as perennial in the NHD are likely intermittent or ephemeral (Table 2). The annual model we developed showed little variability in flow fraction between years, with the exception of isolated areas at mid-to-low elevations, where annual variability in flow fraction was driven by spatial variability of winter precipitation. We found that even in a single region, the years with high and low flow fractions are not consistent between streams.

Future studies would benefit from bedrock permeability predictors and an assessment of jointing in the study area. To further improve upon the NHD, future studies could incorporate higher-resolution topographical information to predict channel locations as well as flow duration. Our mean annual model could be applied to the high-resolution NHDPlus dataset for further analysis of headwater stream classifications in the region. Crowdsourced flow data gathering could be implemented anywhere that lacks sufficient gaging data to assess nonperennial streamflow. The data need to be supplemented by a strategic deployment of flow sensors to fill data gaps and capture harder-to-access areas. The annual model we developed would benefit from more years of data that could capture a variety of climate conditions. Increased observations could also allow a mean monthly model to be developed for April-September to better capture the timing of stream drying.

As climate regimes become less predictable, it is crucial for resource managers to understand streamflow dynamics when identifying streams most vulnerable to drying. Intermittent flow regimes vary widely both spatially and temporally, but for scientific, policy, and management purposes, it is important to consistently and correctly classify streams that exhibit these flow regimes. Predicting flow duration across networks helps accomplish this and allows us to evaluate the degree of intermittency in headwater streams and monitor trends over time. Improving upon the approach in this study, we can better understand the mechanisms behind nonperennial headwater flow duration and effectively revise the NHD to create a more reliable dataset for headwater streams.

BIBLIOGRAPHY

Acuña, Vicenç & Datry, Thibault & Marshall, Jonathan & Barcelo, Daniel & Dahm, Clifford & Ginebreda, Antonio & McGregor, Glenn & Sabater, Sergi & Tockner, Klement & Palmer, Margaret. (2014). Why should we care about temporary waterways?. *Science*, *343*, 1080-1. https://doi.org/10.1126/science.1246666

Addor, N., Nearing, G., Prieto, C., Newman, A.J., Le Vine, N., & Clark, M.P. (2018). A ranking of hydrological signatures based on their predictability in space. *Water Resources Research*, *54*, 8792–8812. https://doi.org/10.1029/2018WR022606

Bishop, K., Buffam, I., Erlandsson, M., Fölster, J., Laudon, H., Seibert, J., & Temnerud, J. (2008). Aqua incognita: The unknown headwaters. *Hydrological Processes*, *22*(8), 1239-1242.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324

Capesius, J.P., & Stephens, V.C. (2009). Regional regression equations for estimation of natural streamflow statistics in colorado. *Scientific Investigations Report,* 2009-5136. http://pubs.usgs.gov/sir/2009/5136/

Chapin, T.P., Todd, A.S., & Zeigler, M.P. (2014). Robust, low-cost data loggers for stream temperature, flow intermittency, and relative conductivity monitoring, *Water Resources Research*, *50*, 6542–6548. doi:10.1002/ 2013WR015158.

Costigan, K.H., Jaeger, K.L., Goss, C.W., Fritz, K.M., & Goebel, P.C. (2016). Understanding controls on flow permanence in intermittent rivers to aid ecological research: Integrating meteorology, geology, and landcover. *Ecohydrology, 9*, 1141-1153. https://doi.org/10.1002/eco.1712

Dewald, T. (2017). Making the Digital Water Flow: The evolution of geospatial surface water frameworks; *USEPA Office of Water*, Washington, DC, USA, 8.

Datry, T., Bonada, N., & Boulton, A.J. (2017). Intermittent rivers and ephemeral streams: Ecology and management. *Academic Press*, London.

Datry, T., Boulton, A.J., Bonada, N., Fritz, K., Leigh, C., Sauquet, E., Tockner, K., Hugueny, B., & Dahm, C.N. (2018). Flow intermittence and ecosystem services in rivers of the Anthropocene. *Journal of Applied Ecology*, *55*, 353–364.

Downing, J.A., Cole, J.J., Duarte, C.M., Middelburg, J.J., Melack, J.M, Prairie, Y.T., Kortelainen, P., Striegl, R.G., McDowell, W.H., & Tranvik, L.J. (2012) Global abundance and size distribution of streams and rivers. *Inland Waters, 2*(4), 229-236. https://doi.org/10.5268/IW-2.4.502

Fritz, K.M., & Dodds, W.K. (2005). Harshness: characterisation of intermittent stream habitat over space and time. *Marine and Freshwater Research, 56,* 13-23. https://doi.org/10.1071/MF04244

Fritz, K.M., Hagenbuch, E., D'Amico, E., Reif, M., Wigington Jr., P.J., Leibowitz, S.G., Comeleo, R.L., Ebersole, J.L., & Nadeau, T.L. (2013). Comparing the extent and permanence of headwater streams from two field surveys to values from hydrographic databases and maps. *Journal of the American Water Resources Association (JAWRA), 49*(4), 867-882. https://doi.org/10.1111/jawr.12040

Godsey S.E., & Kirchner J.W. (2014). Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density, flowing channels and stream order. *Hydrological Processes, 28*(23): 5791–580. https://doi.org/10.1002/hyp.10310

Hafen, K.C., Blasch, K.W., Rea, A., Sando, R., & Gessler, P.E. (2020). The influence of climate variability on the accuracy of NHD perennial and nonperennial stream classifications." *Journal of the American Water Resources Association (JAWRA), 56*(5), 903–916. https://doi.org/10.1111/1752-1688.12871

Hammond, J. C., Saavedra, F. A., & Kampf, S. K. (2018). How does snow persistence relate to annual streamflow in mountain watersheds of the Western U.S. with wet maritime and dry continental climates? *Water Resources Research*, *54*, 2605–2623. https://doi.org/10.1002/2017WR021899

Hammond, J.C. (2020). Contiguous U.S. annual snow persistence and trends from 2001-2020: U.S. Geological Survey data release. Accessed 13 February, 2021. https://doi.org/10.5066/P9U7U5FP

Hammond, J.C., Zimmer, M., Shanafield, M., Kaiser, K., Godsey, S.E., Mims, M.C., et al. (2021). Spatial patterns and drivers of non-perennial flow regimes in the contiguous United States. *Geophysical Research Letters*, *48*, e2020GL090794. https://doi.org/10.1029/2020GL090794

Hastie, T., Tibshirani, R., & Buja, A. (1994). Flexible discriminant analysis by optimal scoring. *Journal of the American Statistical Association, 89*(428), 1255-1270. https://doi.org/10.1080/01621459.1994.10476866

S original by Trevor Hastie & Robert Tibshirani. Original R port by Friedrich Leisch, Kurt Hornik and Brian D. Ripley. Balasubramanian Narasimhan has contributed to the upgrading of the code. (2020). mda: Mixture and Flexible Discriminant Analysis. R package version 0.5-2. https://CRAN.R-project.org/package=mda

Hastings, B.E., & Kampf, S.K. (2014). Evaluation of digital channel network derivation methods in a glaciated subalpine catchment. *Earth Surface Processes and Landforms, 39*(13), 1790-1802.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., et al. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE, 12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Hijmans, R.J. (2021). raster: Geographic data analysis and modeling. R package version 3.4-10. https://CRAN.R-project.org/package=raster

Hothorn, T., & Zeileis, A. (2015). partykit: A modular toolkit for recursive partytioning in R. *Journal of Machine Learning Research*, *16*, 3905-3909. https://jmlr.org/papers/v16/hothorn15a.html

Jaeger, K.L., Olden, J.D., & Pelland, N.A. (2014). Climate change poised to threaten hydrologic connectivity and endemic fishes in dryland streams. *Proceedings of the National Academy of Sciences of the USA, 111*, 13894–13899.

Jaeger, K.L., Sando, R., McShane, R.R., Dunham, J.B., Hockman-Wert, D.P., Kaiser, K.E., Hafen, K., Risley, J.C., & Blasch, K.W. (2019). Probability of streamflow permanence model (PROSPER): A spatially continuous model of annual streamflow permanence throughout the Pacific Northwest. *Journal of Hydrology, X*(2), 100005-100023. https://doi.org/10.1016/j.hydroa.2018.100005

Kampf, S., Strobl, B., Hammon, J., Anenberg, A., Etter, S., Martin, C., Puntenney-Desmond K., Seibert, J., & van Meerveld, I. (2018). Testing the waters: Mobile apps for crowdsourced streamflow data. *Eos*, 99. https://doi.org/10.1029/2018EO096355

Kaplan, N.H., Blume, T., & Weiler, M. (2020). Predicting probabilities of streamflow intermittency across a temperate mesoscale catchment. *Hydrology and Earth System Sciences, 24*, 5453–5472. https://doi.org/10.5194/hess-24-5453-2020

Kingsford, R., Georges, A., & Unmack, P. (2006). Vertebrates of desert rivers: Meeting the challenges of temporal and spatial unpredictability. In: Kingsford, R. (Ed.), Ecology of desert rivers. Cambridge University Press, 154–200.

Koundouri, P., Boulton, A.J., Datry, T., & Souliotis, I. (2017). Chapter 5.2 - Ecosystem services, values, and societal perceptions of intermittent rivers and ephemeral streams, Editor(s): Datry, T., Bonada, N., & Boulton, A. Intermittent rivers and ephemeral streams. Academic Press, 455-476, ISBN 9780128038352. https://doi.org/10.1016/B978-0-12-803835-2.00018-8

Kuhn, et al., (2020). Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. https://www.tidymodels.org

Leibowitz, S.G., Wigington, P.J., Jr, Rains, M.C. & Downing, D.M. (2008). Non-navigable streams and adjacent wetlands: Addressing science needs following the Supreme Court's *Rapanos* decision. *Frontiers in Ecology and the Environment, 6*(7): 364–371. https://doi.org/10.1890/070068

Marshall, J.C., Acuña, V., Allen, D.C., Bonada, N., Boulton, A.J., Carlson, S.M., Dahm, C.N., Datry, T., Leigh, C., Negus, P., Richardson, J.S., Sabater, S., Stevenson, R.J., Steward, A.L., Stubbington, R., Tockner, K., & Vander Vorste, R. (2018). Protecting U.S. temporary waterways. *Science, 361*, 856–857.

Martin, C., Kampf, S.K., Hammond, J.C., Wilson, C., Anderson, S.P. (2021). Controls on streamflow densities in semiarid Rocky Mountain catchments. *Water*, *13*, 521. https://doi.org/10.3390/w13040521

Miller, M.P., Carlisle, D.M., Wolock, D.M., & Wieczorek, M. (2018). A database of natural monthly streamflow estimates from 1950 to 2015 for the conterminous United States. *Journal of the American Water Resources Association, 54*(6): 1258–1269. https://doi.org/10. 1111/1752-1688.12685

Moore, R.B., McKay, L.D., Rea, A.H., Bondelid, T.R., Price, C.V., Dewald, T.G., & Johnston, C.M. (2019). User's guide for the national hydrography dataset plus (NHDPlus) high resolution: U.S. Geological Survey Open-File Report 2019–1096, 66. https://doi.org/10.3133/ofr20191096.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., & Veith, T.L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE, 50*(3), 885-900.

Nadeau, T.L., & Rains, M.C. (2007). Hydrological connectivity between headwater streams and downstream waters: How science can inform policy. *Journal of the American Water Resources Association (JAWRA), 43*(1): 118-133. https://doi.org/10.1111/j.1752-1688.2007.00010.x

Poggio, L., de Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty, *SOIL, 7*, 217–240, https://doi.org/10.5194/soil-7-217-2021, 2021.

PRISM Climate Group. (2021). PRISM. PRISM Climate Group, Oregon State University. https://prism.oregonstate.edu

Sando, R., & Blasch, K. (2015). Predicting alpine headwater stream intermittency: A case study in the northern Rocky Mountains. *Ecohydrology & Hydrobiology, 15*, 68-80. https://doi.org/10.1016/j.ecohyd.2015.04.002

Sando, R., Olsen, T.D., Kaiser, K.E., Haluska, T.L., & Hockman-Wert, D. P. (2018). Probability of streamflow permanence (PROSPER) continuous parameter grids (CPGs): U.S. Geological Survey data release. https://doi.org/10.5066/F7CC0ZXH

Schilling, O.S., Cook, P.G., Grierson, P.F., Dogramaci, S., & Simmons, C.T. (2021). Controls on interactions between surface water, groundwater and riverine vegetation along intermittent rivers and ephemeral streams in arid regions. *Water Resources Research, 57*, e2020WR028429. https://doi.org/10.1029/2020WR028429

Shumilova, O., Zak, D., Datry, T., von Schiller, D., Corti, R., Foulquier, A., Obrador, B., Tockner, K., Allan, D.C., Altermatt, F., Arce, M.I., et al. (2019). Simulating rewetting events in intermittent rivers and ephemeral streams: A global analysis of leached nutrients and organic matter. *Global Change Biology 25*, 1591–1611.

Soil Survey Staff, Natural Resources Conservation Service (NRCS), United States Department of Agriculture (USDA). (2016). Web soil survey: Updated hydrologic soils group (HSG) questions & answers. https://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=stelprdb1262857&ext=pdf

Soil Survey Staff, United States Department of Agriculture (USDA), Natural Resources Conservation Service (NRCS). (2020). Gridded Soil Survey Geographic (gSSURGO) Database for the Conterminous United States. Accessed 31 August, 2021. https://gdg.sc.egov.usda.gov/

Strobl, C., Boulesteix, A.L., Zeileis, A. et al. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics, 8,* 25. https://doi.org/10.1186/1471-2105-8-25

Tennant, C.J., Crosby, B.T., & Gosey, S.E. (2015). Elevation-dependent responses of streamflow to climate warming. *Hydrological Processes, 29*, 991-1001. https://doi.org/10.1002/hyp.10203

U.S. Department of Agriculture, Geological Survey Mineral Resources. (2005). State Geologic Maps. USDA-USGS Mineral Resources. http://pubs.usgs.gov/of/2005/1351/

U.S. Department of the Interior, Geological Survey, and U.S. Department of Agriculture. (2016). Existing Vegetation Type Layer, LANDFIRE 2.0.0. Accessed 02 September, 2021. http://landfire.cr.usgs.gov/viewer/

U.S. Environmental Protection Agency (USEPA), & U.S. Geological Survey (USGS) (2012). National Hydrography Dataset Plus (NHDPlus). Version 2.10. Accessed 03 November, 2020. https://www.epa.gov/waterdata/get-nhdplus-national-hydrography-dataset-plus-data

U.S. Forest Service, & U.S. Geological Surevey. (2021). MTBS data access: Fire level geospatial data. MTBS Project. Accessed 09 September, 2021. https://www.mtbs.gov/

U.S. Geological Survey. (2019). National Hydrography Dataset Plus (NHDPlus). Version Best Resolution (NHD) for Colorado. Accessed 03 November, 2020. https://www.usgs.gov/core-science-systems/ngp/national-hydrography/access-national-hydrography-products

von Schiller, D., Bernal, S., Dahm, C.N., & Martí, E. (2017). Chapter 3.2 - Nutrient and organic matter dynamics in intermittent rivers and ephemeral streams. Editor(s): Datry, T., Bonada, N., & Boulton, A., Intermittent rivers and ephemeral streams. Academic Press, 135-160, ISBN 9780128038352. https://doi.org/10.1016/B978-0-12-803835-2.00006-1

Ward, A.S., Wondzell, S.M., Schmadel, N.M., & Herzog, S.P. (2020). Climate change causes river network contraction and disconnection in the H.J. Andrews Experimental Forest, Oregon, USA. *Frontiers in Water, 2*, 7. https://doi.org/10.3389/frwa.2020.00007

Wilson, C., Kampf, S.K., Wagenbrenner, J.W., & MacDonald, L.H. (2018). Rainfall thresholds for post-fire runoff and sediment delivery from plot to watershed scales. *Forest Ecology and Management, 430,* 346-356.

Wohl, Ellen, (2008). The effect of bedrock jointing on the formation of straths in the Cache la Poudre River drainage, Colorado Front Range. *Journal of Geophysical Research, 113*, F01007. https://doi.org/10.1029/2007JF000817

Zipper, S.C., Hammond, J., Shanafield, M., Zimmer, M., Datry, T., Jones, N., Kaiser, K., Godsey, S., Burrows, R.M., Blaszczak, J.R., Busch, M.H., Price, A.N., Boersma, K., Ward, A. S., Costigan, K., Allen, G., Krabbenhoft, C., Dodds, W.K., Mims, M.C., Olden, J.D., Kampf, S., Burgin, A.J., & Allen, D.C. (2021a). Pervasive changes in stream intermittency across the United States. Environmental Research Letters, 16, 084033. https:doi.org/10.1088/1748-9326/ac14ec

Zipper, S.C., Hammond, J., Shanafield, M., Zimmer, M., Datry, T., Jones, N., Kaiser, K., Godsey, S., Burrows, R.M., Blaszczak, J.R., Busch, M.H., Price, A.N., Boersma, K., Ward, A. S., Costigan, K., Allen, G., Krabbenhoft, C., Dodds, W.K., Mims, M.C., Olden, J.D., Kampf, S., Burgin, A.J., & Allen, D.C. (2021b). Data and code release: Pervasive changes in stream intermittency across the United States (published in Environmental Research Letters), HydroShare. https://doi.org/10.4211/hs.fe9d240438914634abbfdcfa03bed863

APPENDIX A
PREDICTOR CORRELATION TO FLOW FRACTION

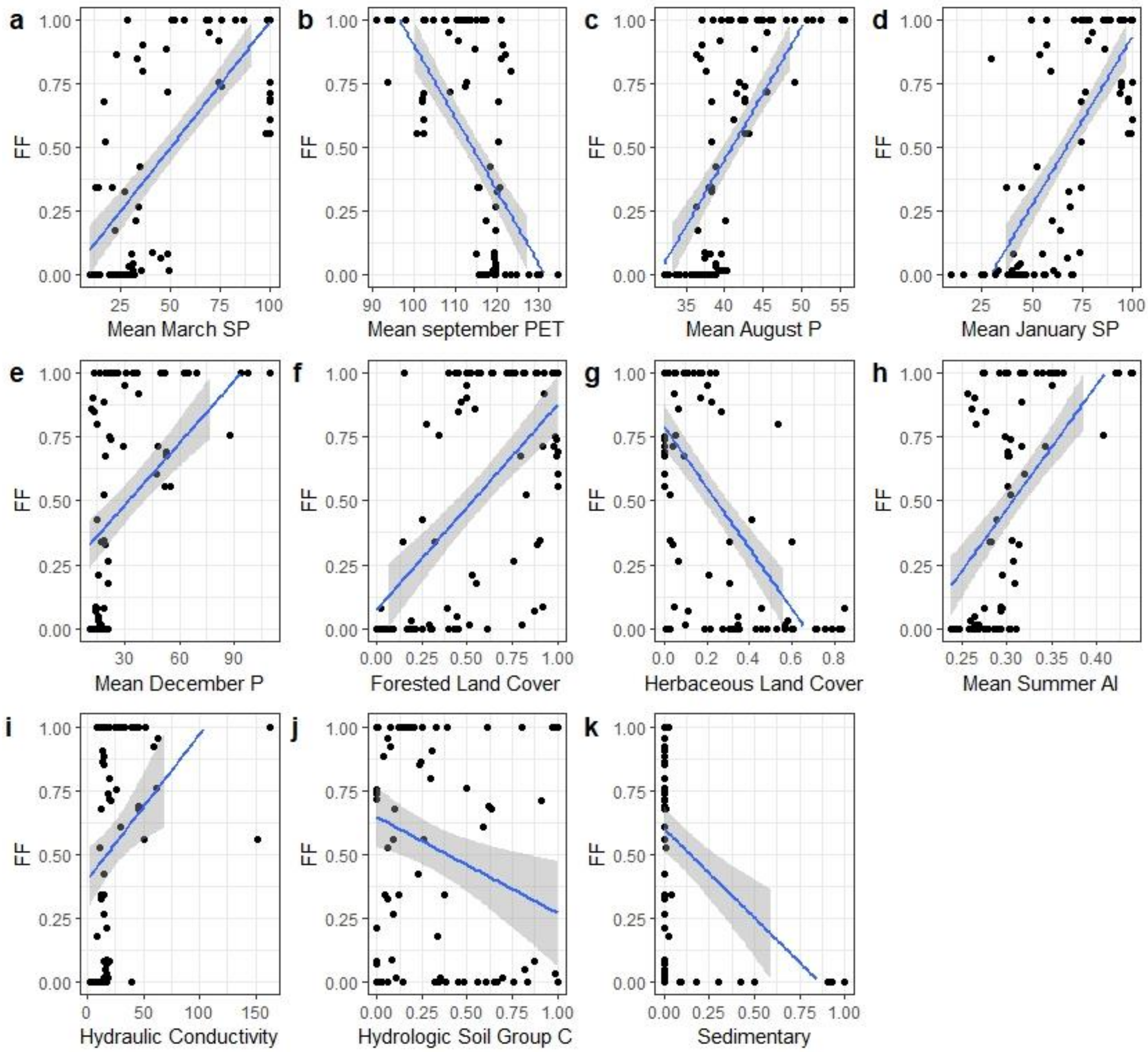

Figure A1. Observed mean annual flow fraction (FF) vs predictor variables.

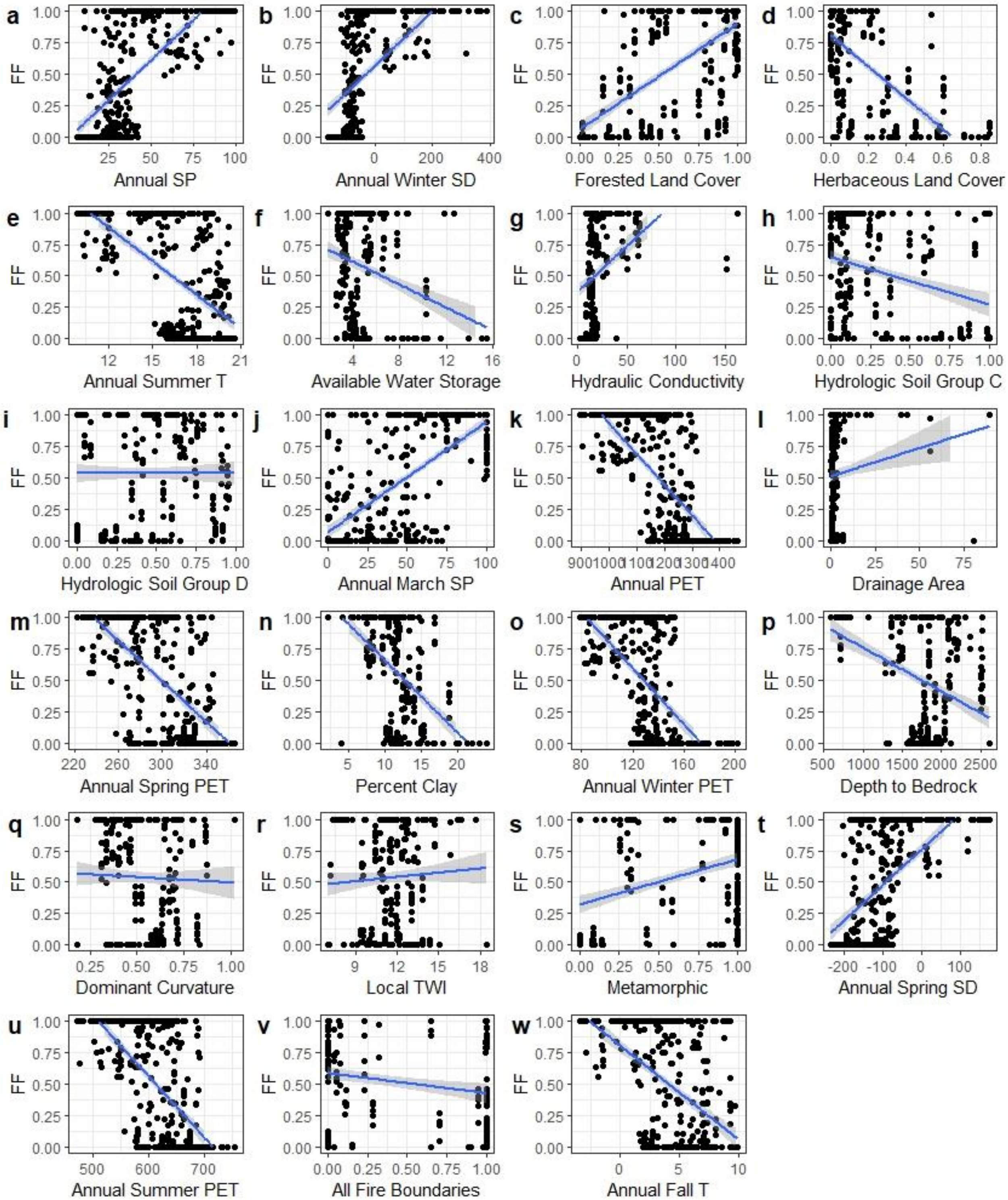

Figure A2. Observed annual flow fraction (FF) vs predictor variables.

APPENDIX B

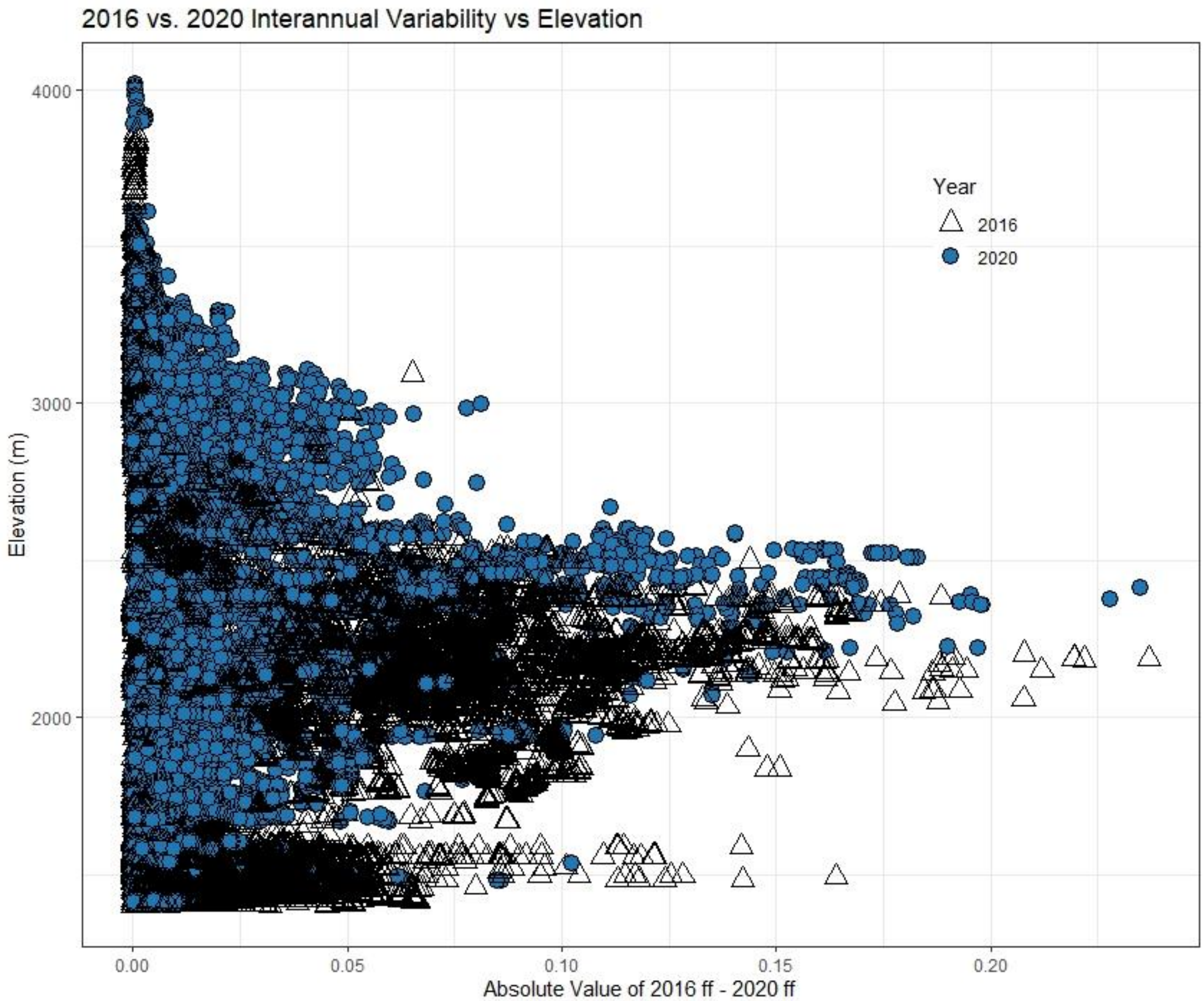

Figure B1. Elevation plotted against the absolute value of the difference between 2016 flow fraction and 2020 flow fraction. Blue circles indicate points that are wetter in the dry year (2020). Points that are wetter in the wet year (2016) are indicated by hollow triangles. Areas of most extreme interannual variability are in mid and low elevations.

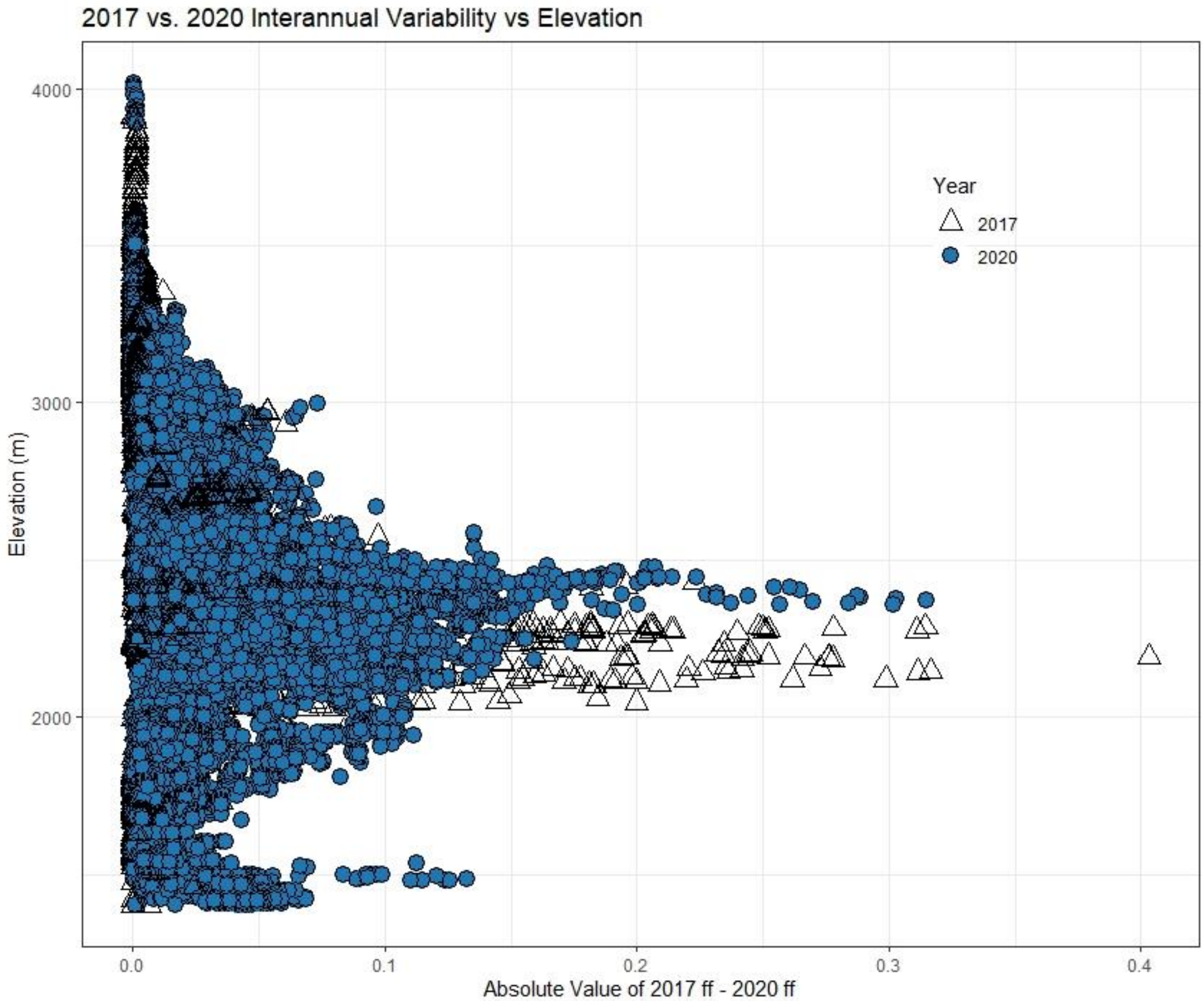

Figure B2. Elevation plotted against the absolute value of the difference between 2017 flow fraction and 2020 flow fraction. Blue circles indicate points that are wetter in the dry year (2020). Points that are wetter in the wet year (2017) are indicated by hollow triangles. Areas of most extreme interannual variability are in mid elevations.