1-1-2019

Probability of Streamflow Permanence Model (Prosper): A Spatially Continuous Model of Annual Streamflow Permanence Throughout the Pacific Northwest

K. E. Kaiser
Boise State University

Research papers

Probability of Streamflow Permanence Model (PROSPER): A spatially continuous model of annual streamflow permanence throughout the Pacific Nortwest

K.L. Jaeger [a,*], R. Sando [b], R.R. McShane [b], J.B. Dunham [c], D.P. Hockman-Wert [c], K.E. Kaiser [d,e], K. Hafen [d,f], J.C. Risley [g], K.W. Blasch [d]

[a] U.S. Geological Survey Washington Water Science Center, Tacoma, WA, United States
[b] U.S. Geological Survey Wyoming-Montana Water Science Center, Helena, MT, United States
[c] U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, Corvallis, OR, United States
[d] U.S. Geological Survey Idaho Water Science Center, Boise, ID, United States
[e] Boise State University, Department of Geosciences, Boise, ID, United States
[f] University of Idaho, Water Resources Program, Moscow, ID, United States
[g] U.S. Geological Survey Oregon Water Science Center, Portland, OR, United States

ARTICLE INFO

ABSTRACT

The U.S. Geological Survey (USGS) has developed the PRObability of Streamflow PERmanence (PROSPER) model, a GIS raster-based empirical model that provides streamflow permanence probabilities (probabilistic predictions) of a stream channel having year-round flow for any unregulated and minimally-impaired stream channel in the Pacific Northwest region, U.S. The model provides annual predictions for 2004–2016 at a 30-m spatial resolution based on monthly or annually updated values of climatic conditions and static physiographic variables associated with the upstream basin. Predictions correspond to any pixel on the channel network consistent with the medium resolution National Hydrography Dataset channel network stream grid. Total annual precipitation and percent forest cover were consistently the most important predictor variables among global and most subregional models, which had error rates between 17 and 22%. Probabilities were converted to wet and dry streamflow permanence classes with an associated confidence. Wet and dry classifications were used to derive descriptors that characterize the statistical and spatial distribution of streamflow permanence in three focal basins. Predicted dry channel segments account for 52–92% of the stream network across the three focal basins; streamflow permanence decreased during climatically drier years. Predictions are publicly available through the USGS StreamStats platform. Results demonstrate the utility of the PROSPER model as a tool for identifying areas that may be resilient or sensitive to drought conditions, allowing for management efforts that target protecting critical reaches. Importantly, PROSPER's successful predictive performance can be improved with new datasets of streamflow permanence underscoring the importance of field observations.

© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Streamflow permanence, defined as the degree to which rivers and streams maintain surface flow conditions (Costigan et al., 2016; Datry et al., 2017), exerts primary control on the transfer of energy and materials (e.g., nutrients and organisms) by surface water through the river network (*sensu* Pringle, 2003) and is a fundamental driver of riverine ecosystems (Poff et al., 1997; Stanley et al., 1997; Datry et al., 2017). Streamflow permanence classification (perennial or non-perennial) is a major component in aquatic and terrestrial species vulnerability assessments (Poff et al., 2010; Donnelly et al., 2016; Perkin et al., 2017), land management activities (e.g., Jaeger et al., 2007; Michael, 2004), and water quality regulations (e.g., Fritz et al., 2013; Acuña et al., 2014; Caruso, 2014). Inaccurate streamflow permanence classifications can therefore have important and wide-ranging consequences for management of water resources.

* Corresponding author.
 E-mail address: kjaeger@usgs.gov (K.L. Jaeger).

Despite the importance of accurate streamflow permanence classifications, our understanding and available observations of streamflow permanence at a regional extent is surprisingly incomplete. Currently (2018), the most comprehensive dataset containing nationally consistent streamflow permanence classifications is the National Hydrography Dataset Plus (NHDPlus) (https://www.epa.gov/waterdata/nhdplus-national-hydrography-dataset-plus). Streamflow permanence classifications contained in the NHDPlus are based on one-time field surveys, typically conducted in the mid- to late-1900s, although contemporary stewardship efforts update NHDPlus classifications (McKay et al., 2012) using various methods. In areas where the accuracy of the NHDPlus streamflow classifications has been analyzed, results showed that misclassifications can be as high as 50% (Fritz et al., 2013; Ebersole et al., 2014) and that flow permanence may change through time (Eng et al., 2016). Additionally, there is increasing recognition of the complex spatiotemporal dynamics of streamflow permanence. Streamflow permanence patterns may be conceptualized as a spatial and temporal gradient (Boulton et al., 2017; Costigan et al., 2017) that includes substantial inter-annual variability of stream drying patterns (Jaeger and Olden, 2012; Godsey and Kirchner, 2014; Jensen et al., 2017).

As the need for accurate and up-to-date streamflow permanence classifications has been recognized, locally scoped efforts have gained popularity. These efforts include conducting field mapping of streamflow presence through direct observation (Turner and Richter, 2011; Jensen et al., 2017), using temperature or other sensors as proxies for streamflow or surface water presence (Blasch et al., 2002; Chapin et al., 2014; Gungle, 2006; Bhamjee et al., 2015; Arismendi et al., 2017), and identifying indirect physical or biological indicators of streamflow permanence (Fritz et al., 2006; 2008; Nadeau et al., 2015). Streamflow presence has also been interpreted from the intensity of ground returns of topographic airborne Light Detection And Ranging (LiDAR) systems (Hooshyar et al., 2015), and unmanned aerial vehicles may provide a cost-effective method to field mapping, although this method remains experimental (Spence and Mengistu, 2016). Whereas these approaches can provide valuable insight for specific locations or basins, it is also important to develop approaches for analyzing streamflow permanence at a regional extent.

In the current (2018) absence of exhaustive streamflow permanence observations at regional scales, pairing streamflow permanence observations, where they do exist, with statistical or physical modeling approaches can yield streamflow permanence classifications for areas that lack field observations (Sando and Blasch, 2015; González-Ferreras and Barquín, 2017). With the widespread availability of moderate-resolution, remotely sensed and geographic information system (GIS)-derived datasets at national scales, coupled with assemblages of streamflow observation datasets, opportunities have emerged to develop relatively high resolution, spatially explicit classifications of streamflow permanence at regional extents that also account for prevailing hydrologic conditions. The prospect of temporally dynamic streamflow permanence classifications at a regional scale consequently not only allows for improved understanding of the spatiotemporal dynamics of streamflow permanence and the physiographic and hydroclimatic variables that control streamflow permanence (Costigan et al., 2016; Datry et al., 2018; Eng et al., 2016), but also serves as an immediate tool to water resource managers who are increasingly challenged by limited knowledge of where and when streams and rivers maintain streamflow (Sando and Blasch, 2015).

1.1. The PROSPER model

The U.S. Geological Survey (USGS) created a 30-m resolution, temporally resolved model of streamflow permanence for the Pacific Northwest Region, U.S. The PRObability of Streamflow PERmanence (PROSPER) model incorporates empirical data, static physiographic variables, and monthly to annual climatic data to predict the annual probability of year-round streamflow for any unregulated and minimally-impaired stream channel. Predictions extend to channels draining 0.09 km^2, or greater, and are concurrent with medium resolution NHDPlus version 2 grids. Predictions are made for greater stream length than is represented by the NHDPlus stream channel network, but it is widely acknowledged that small headwater streams are underrepresented by NHD (Fritz et al., 2013; Benstead and Leigh, 2012). Predicted annual probabilities of streamflow permanence for 2004–2016 are publicly available through the interactive USGS StreamStats platform (https://water.usgs.gov/osw/streamstats/) and as a USGS Science-Base Data Release (Sando and Hockman-Wert, 2018).

The objectives of this paper are to 1) introduce the PROSPER model and describe model output of streamflow permanence probabilities and streamflow permanence classes and 2) demonstrate applications of the PROSPER model to quantitatively describe streamflow permanence using three focal basins. This work provides end users with streamflow permanence conditions for the Pacific Northwest Region at unprecedented spatial extent and temporal resolution, representing the largest dynamic stream mapping effort the authors are aware of, and which is readily available for applications in water resources management or species conservation.

2. Materials and methods

2.1. Study area

A global PROSPER model was developed for 2-digit Hydrologic Unit Code (HUC2) 17 Pacific Northwest Region, U.S., which encompassed all of Washington, most of Oregon and Idaho, western Montana, and smaller portions of Wyoming, Utah, and Nevada along state boundaries (Fig. 1). Post hoc analysis of the variability in the PROSPER model output was evaluated by partitioning the study area into four subbasin regions (17a, b, c, and d) based on approximate partitioning between HUC4 basins (Fig. 1). Increases in digit numbers associated with the HUC correspond to smaller basin size. The HUC4 watershed boundaries presumably reflect similarity in climate and physiography and therefore are a reasonable approach to evaluate subregional variability as well as a convenient, straightforward approach to break up the relatively large study area for a post hoc subregional analysis.

Topographically, the study area includes mountain chains of varying spatial extent, broad valley and lowland areas, and extensive plateau regions (Fig. 1). The study area occurs in four geologic provinces that include the Cascade Volcanoes, the Columbia Plateau, the North Cascades, and the Coast Mountains. The region is mostly underlain by volcanic rock, but also includes areas of granite, metamorphic, and mixed sedimentary and volcanic rock (Schruben et al., 1994). Portions of the study area are strongly influenced by volcanism and the last glacial maximum.

The study area represents a broad range of climates from inland arid and semi-arid to coastal humid regions that support temperate rainforests (Wolock, 2003a; Leibowitz et al., 2016). Normal mean annual precipitation totals range from approximately 310 mm in inland low elevation areas to 2400 mm along the coast (PRISM Climate Group, 2004). Normal mean annual temperatures (1981–2010) range from approximately 4 °C in inland areas to 11 °C in coastal areas (PRISM Climate Group, 2004). Summers are dry and warm; winters range from mild temperatures and rain-dominated precipitation along the coast and at low elevations to cold temperatures and snow-dominated precipitation at higher elevations (Griffith, 2010).
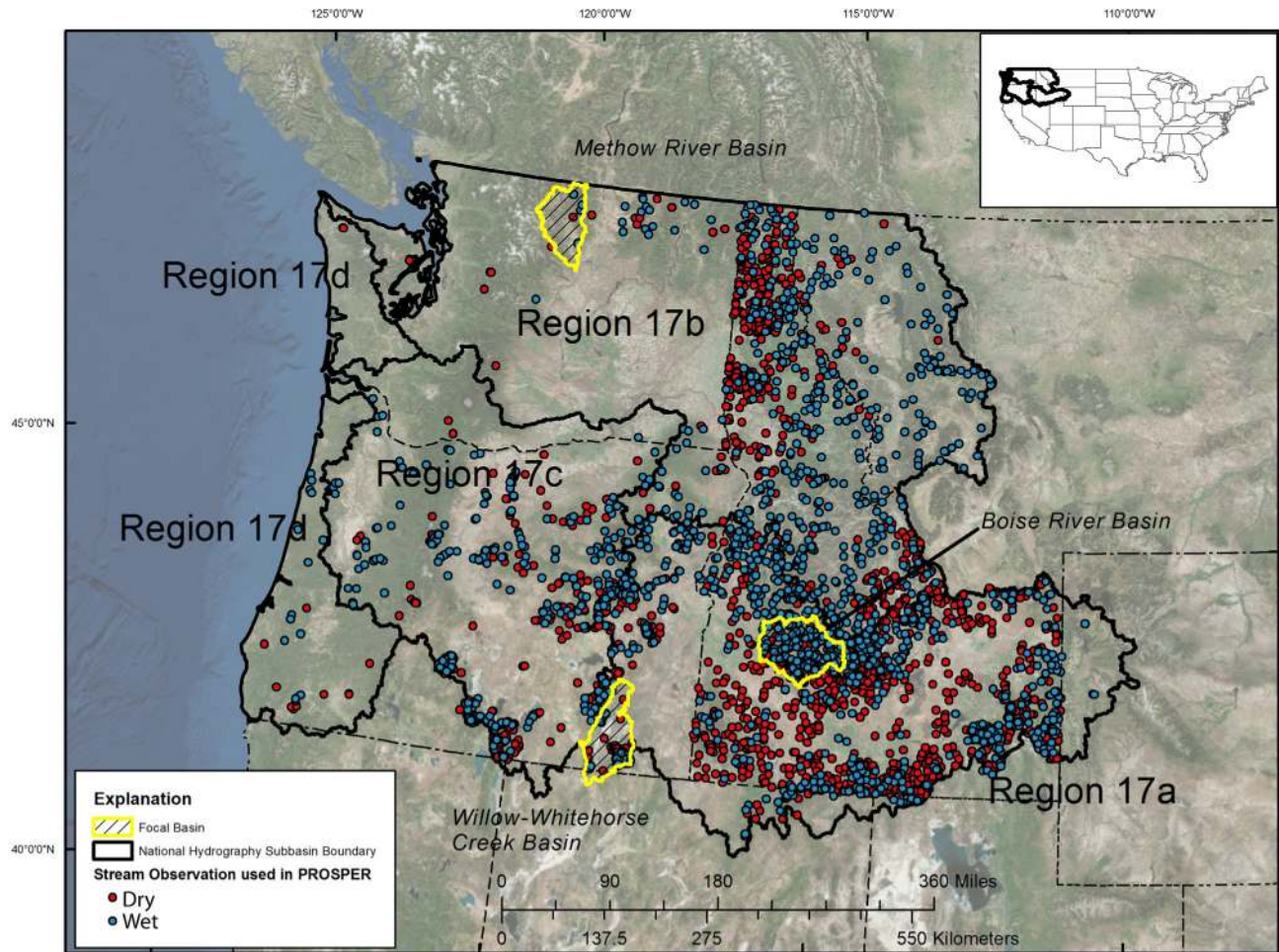
Fig. 1. Map of the study area, which corresponds to the Hydrologic Unit Code (HUC) 17 Pacific Northwest Region including the four regional subbasins 17a-d, watershed boundaries for the three focal basins, and filtered streamflow observations used in PROSPER model development.

The Methow River basin in Washington, the adjacent basins of Willow and Whitehorse Creeks in Oregon, and the Boise River basin in Idaho were selected to serve as examples for a more detailed evaluation of the spatiotemporal variability of streamflow permanence predictions within the context of basin-scale hydroclimatic conditions (Fig. 1, Table SI.1). These focal basins represent the broad range of hydroclimatic and physical conditions in the study area and illustrate the range in predictive ability of PROSPER by including a gradient of streamflow permanence conditions (Konrad et al., 2003; Wood et al., 2009; Schultz et al., 2017). In particular, the Methow River basin is known to have strong surface-groundwater interactions (Konrad et al., 2003), which are processes not directly captured by PROSPER. The Boise River and Willow and Whitehorse Creek basins have similar geologic and climatic conditions, but Willow and Whitehorse Creeks tend towards reduced streamflow permanence compared to the Boise River.

2.2. Data

To build the PROSPER model, we used streamflow permanence observations aggregated from various datasets that were part of previous or ongoing field data collection efforts (McShane et al., 2017). A broad suite of GIS-derived climate and physiographic characteristics were used as predictor variables for streamflow permanence probabilities.

2.2.1. Streamflow observations

A total of 3878 observations (1941 wet, 1937 dry) were used in the PROSPER model (Fig. 1). Final observations used in the PROSPER model were filtered (methods detailed in Sando and Hockman-Wert, 2018) from a larger dataset of 24,316 streamflow observations that occurred from 1977 to 2016. The larger observation dataset was compiled or derived from 11 datasets that were part of independent projects that included aquatic species habitat surveys, wet/dry stream channel mapping, and beneficial use reconnaissance surveys, or were collected specifically for the PROSPER project (McShane et al., 2017). Observations were distributed across a range of drainage areas (0.092–24,300 km^2) with a focus on small streams. More than one third of the observations occurred in streams with a drainage area less than 10 km^2. Streamflow observations included one-time surveys and repeat surveys extending over several years, as well as discrete locations or continuous sections of a stream channel reach. The streamflow observations were processed into a single consistent binary dataset with classifications of "wet" or "dry". A site classified as wet required the presence of surface water that could either be flowing or standing water in pools; a site classified as dry had no surface water present. Wet classifications required that the observation occurred after July 1 to coincide with the hydroclimatically driest time of the year; dry classifications could occur any time within the year. Repeat streamflow observations within a single year at the same location were considered "dry" if any of the observations were

dry or "wet" if all observations were wet and an observation occurred after July 1.

2.2.2. Climatic and physiographic variables

To capture the mechanisms that influence streamflow permanence in the study area, a total of 257 climatic and 35 physical predictor variables (292 total) were considered for inclusion in the statistical model development (Table 1). NHDPlus flow direction grids were used to summarize each predictor variable upstream of every channel grid cell on the river network. We refer to these predictor variable grids (e.g. monthly precipitation, percent forest cover) as Continuous Parameter Grids (CPGs; Sando et al., 2018). The inter-annual variability in streamflow permanence conditions is important in studying annual patterns for a variety of ecological phenomena (Datry et al., 2017). Thus, values for many of the climatic variables represented monthly or annual conditions, rather than 30-year normal conditions.

Climatic predictor variables included in the PROSPER model represented annual or monthly values for each year from 2004 through 2016. This time period represents the longest continuous period of existing data for all of the climatic variables. Physiographic characteristics considered as potential predictor variables in PROSPER were selected to capture mechanisms that directly affect streamflow permanence while limiting redundancy. In particular, geologic characteristics that are presumed to influence flow permanence are captured in predictor variables of permeability, topography, and soil characteristics. Surface geology was not explicitly included as a predictor variable. Permeability was derived based on the surficial geology presented as percent contributing area permeable versus non-permeable surficial geology. While basin elevation and drainage area are highly correlated with streamflow permanence (Sun et al., 2011; González-Ferreras and Barquín, 2017), they were excluded from the model to avoid reducing the statistical influence of temporally variable predictor variables (e.g. precipitation and temperature), which may better describe year-to-year changes in streamflow permanence. The topographic wetness index incorporates local slope and drainage area and was included as a predictor variable to capture the topographic control on hydrologic condition (Beven and Kirkby, 1979; Sörensen et al., 2006; Williamson et al., 2015).

2.3. Analysis

A series of steps were employed to provide annual predictions of streamflow permanence probabilities (Fig. 2). The PROSPER model was developed using random forest classification to produce
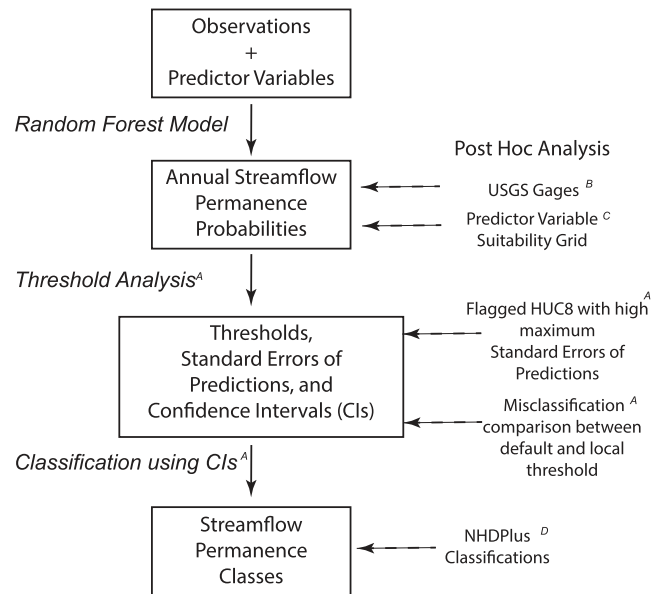

Fig. 2. Schematic of workflow including validation measures represented by dashed arrows for PROSPER model used in this study. Superscript letters correspond to the Appendix in which methods and results are detailed.

Table 1
List of predictor variables considered for use in the random forest model.

Characteristic		Number of Rasters	Source
Physiographic			
Land Use – Land Cover	Percent forest, agriculture, developed, wetland, open water (2001, 2006, 2011)	15	National Landcover Dataset (NLCD; Fry et al., 2011; Homer et al., 2007; Homer et al., 2015)
	Percent Irrigated Land (2002, 2007, 2012)	3	Moderate Resolution Imaging Spectroradiometer (MODIS) Irrigated Agriculture Dataset for the United States (MIrAD-US; Pervez and Brown, 2010)
Topography	Mean stream slope, mean basin slope	2	National Hydrography Dataset Plus version 2 (NHDPlus v2; USEPA and USGS, 2012)
	Elevation, contributing area	2	
	Topographic wetness index	1	
Soils	Mean basin percent hydrologic soil group A, B, C, D	4	STATSGO2 (Soil Survey Staff, 2016)
	Depth to water table	1	
	Available water capacity	1	
	Mean basin percent clay, sand, silt	3	
Permeability	Percent permeable, percent impermeable	2	Modified from state geological maps (Ludington et al., 2007; Stoeser et al., 2005)
Baseflow index	Percent of streamflow defined as baseflow for a given location	1	Wolock (2003b)
Climatic			
Temperature	Annual mean daily maximum (2000–2016)	17	PRISM Climate Group (2004)
	Annual mean daily minimum (2000–2016)	17	
	1981–2010 normal annual daily maximum	1	
	1981–2010 normal annual daily minimum	1	
Precipitation	Total annual (2000–2016)	17	PRISM Climate Group (2004)
	1981–2010 normal annual	1	
Snow Water Equivalent	First-of-month for Mar-Aug (2004–2016)	78	Snow Data Assimilation System (SNODAS) Version 1 (Barrett, 2003)
	2004–2016 average first-of-month (Mar-Aug)	6	
Evapotranspiration	Monthly mean for Mar-Sep (2000–2015)	112	Operational Simplified Surface Energy Balance (SSEBop; Senay et al., 2013)
	2000–2015 monthly mean for Mar-Sep	7	

annual streamflow permanence probabilities for every stream channel pixel within the study area. A localized threshold analysis was then conducted to identify threshold values that classify the streamflow permanence probabilities as wet or dry at each stream channel pixel. Confidence intervals were constructed for the threshold values and their associated mean standard errors of predictions. Finally, the confidence intervals were used to categorize the streamflow permanence probabilities at each pixel into a streamflow permanence class, which consisted of a wet or dry classification with an associated confidence (e.g., dry with 95% confidence). Additional validation steps were included to evaluate the reliability of streamflow permanence probabilities and classes. Methods to develop the random forest classification model are detailed below. Methods to translate streamflow permanence probabilities to classes and post hoc validation analyses are detailed in appendices.

2.3.1. PROSPER model development and validation

Random forest classification (Breiman, 2001) was used as the statistical modeling framework. Random forest classification constructs several decision trees (e.g., a forest) on a training dataset and outputs the class, in this case, wet or dry, that is the mode of the individual trees as well as the proportion of total trees represented by each class. The 'randomForest' package (Liaw and Wiener, 2002) was used in R (R Core Team, 2015). The number of trees built was set to 500 and the number of predictor variables randomly selected to split the data at each node was set to five.

For this study, bootstrapped methods (Breiman, 2001) that included a random subsample of about two-thirds of the data were employed. Each individual classification tree in the random forest model was developed using one of the subsamples obtained from the bootstrapping process. The submodel was then applied to each observation that was not included in the subsample and a class ("wet" or "dry") was predicted. For each individual classification tree, the number of misclassified observations was then divided by the total number of observations to obtain an out-of-bag error estimate, which is the standard metric to evaluate the strength of the RF model as a classifier (Breiman, 2001).

To obtain a final classification for each observation, first, a probability was produced by running the predictor values through each classification tree and dividing the frequency of predicted wet classes by the total number of trees (500). We call this probability for each observation, the streamflow permanence probability. While we use the term "probability" to describe the outcome, the prediction can be considered a relative degree to which a particular observation or pixel is more statistically similar, in terms of predictor variables, to either the population of wet or dry observations. Higher streamflow permanence probabilities represent statistical similarities to the hydrologic and physiographic conditions where wet observations occurred; lower streamflow permanence probabilities represent statistical similarities to the hydrologic and physiographic conditions where dry observations occurred in that given year. The streamflow permanence probability for each observation was then converted to a final predicted class based on the default approach of the predicted class with the highest proportion of total predictions. Specifically, a final prediction of 'wet' was assigned if the probability for an observation was greater than 0.5; conversely a final prediction of "dry" was assigned for probability values less than 0.5. The default out-of-bag error rate was determined by averaging the out-of-bag error rates from all the random forest classification trees. The importance of each explanatory variable was quantified by calculating the percent decrease of default classification error associated with the inclusion of each predictor variable (Liaw and Wiener, 2002).

An additional analysis was conducted to determine if classification accuracy could be improved in some or all of the study area by empirically adjusting the threshold from a default of 0.5 to a value that more accurately reflects local conditions. This analysis, termed the localized threshold analysis, is described in Appendix A. The local threshold error rate was determined by analyzing the models ability to correctly classify both true positives (sensitivity) and negatives (specificity). Performance of the local threshold was evaluated by analyzing the resulting changes in probability of misclassification of wet and dry observations for each HUC-8 region in the study area.

A global random forest PROSPER model was developed for the entire study area with subsequent analysis of the spatial characteristics of the residuals at a subregional scale. A global model was chosen because the density of streamflow observations was highly variable across the study area and the influence of statistical and spatial distributions of the predictor variable data on streamflow permanence was not well known. Three subregional models were also developed, which included subregion 17a, 17b, and a combined subregion of 17c and 17d, to assess in a post hoc analysis how predictor variable importance varies among these different regions (Fig. 1). Subregion 17d was combined with 17c because it contained only 28 streamflow observations, which was insufficient for a stand-alone submodel.

The final global random forest model, developed and calibrated using the available observation data, was used to provide streamflow permanence probabilities for each year (2004–2016) at each 30-m grid cell in the study area (570,400,479 total predictions, representing a total of approximately 1.84 million km in stream channel length). This was done by calculating predictor variable values for each pixel location in the study area and running those data through the RF classification model.

The streamflow permanence probabilities were then converted into streamflow permanence classes of either wet or dry binary classification with an associated confidence (Appendix A). A total of 10 streamflow permanence classes range from −5 (dry classification with 95% confidence or greater) to 5 (wet classification with 95% confidence or greater).

Three post hoc, supplemental validation exercises were conducted in addition to validation internal to the random forest modeling approach. First, streamflow permanence probability values were compared to streamflow statistics at USGS gages stratified by six subregional climate classes within the study area (Appendix B). The six climate classes ranged from "arid" to "very wet" as defined by Leibowitz et al. (2016) and based on the Feddema (2005) Moisture Index, which is a ratio between precipitation and evapotranspiration. Second, a Predictor Variable Suitability Grid (Appendix C) was created to evaluate how well the study area is represented by the streamflow observation dataset. Third, streamflow permanence classes were compared to NHDplus classifications to evaluate the reliability of streamflow permanence classifications (Appendix D).

2.3.2. PROSPER-derived descriptors of streamflow permanence

Streamflow permanence classes were extracted for each year (2004–2016) for the Methow River, Willow and Whitehorse Creeks, and Boise River focal basins (Fig. 1) for further evaluation of spatiotemporal variability in streamflow permanence. We used standardized total annual precipitation as an initial approach to investigate potential relationships between streamflow permanence and basin-scale drought conditions. Specifically, we assessed 1) the year-to-year change in the statistical distribution of contiguous wet and dry stream segments, 2) how continuity of wet or dry stream segments, as a proxy for streamflow fragmentation, varies with differences in total annual precipitation, and 3) how the position of wet and dry segments in a stream network varies with differences in total annual precipitation. This demonstrates one of the potential ways to analyze model output. A comprehensive analysis

of each predictor variable in relation to streamflow permanence predictions was beyond the scope of this project, but it is important to note that an analysis is only meaningful for those predictor variables that were most influential in the global random forest model (e.g. total annual precipitation).

We assessed the relation between basin dryness and streamflow fragmentation to evaluate how the two vary in relation to inter-annual precipitation variability. We use the proportion of the channel network that is predicted to be dry as a descriptor of overall basin dryness. Streamflow fragmentation is represented by the ratio of contiguously wet stream segments to contiguously dry stream segments. Low ratios correspond to higher streamflow fragmentation as a result of higher frequencies of dry channel segments. An increase in the proportion of the network with dry stream segments and potentially increased frequency of dry stream segments would be expected in years of lower precipitation. Consequently, basin dryness and streamflow fragmentation would be expected to increase.

Changes in the position of wet and dry channel segments within the river network were evaluated in terms of the elevation and drainage area of wet and dry stream segments. When considering the topology of a river network, stream segments with smaller cumulative drainage areas represent headwater streams, which can occur at either high or low elevations. Moving in the downstream direction, stream segments of increasing cumulative drainage area correspond to decreasing elevations. For example, in a catchment that contains dry headwater channels and progressive downstream wetting of channels, dry headwater channels occur at higher elevations and small drainage areas; wet channels occur downstream at lower elevations and progressively larger drainage areas. Because the wet stream segments occur at low elevations, the ratio of elevations of wet stream segments to dry stream segments would also be low. Conversely, the ratio of cumulative drainage area of wet stream segments to dry stream segments would be high. In climatically drier years, an increase in the elevation ratio could reflect lower elevation wet segments that remain stable while dry channel segments expand downstream. A decrease in the drainage area ratio could reflect a downstream expansion of dry channel segments.

3. Results

3.1. Model performance and PROSPER streamflow permanence probabilities

Streamflow permanence probabilities were calculated for the 3878 observations that met the data filtering criteria. The mean out-of-bag (OOB) error rate for the global model was approximately 20%, and ranged from 17 to 22% for the subregional models (Table 2). The error rate was lowest for the 17a subregional model, located in the southeastern portion of the study area which is largely located within the semiarid climate class (Leibowitz et al., 2016). The highest error rate was for the 17b subregional model, which corresponds to the northern third of the study area and includes a strong climatic gradient of arid to very wet.

A total of 29 predictor variables were included in the global model and the three subregional models used for post hoc analysis. Total annual precipitation and percent forest cover were two of the top three most important predictor variables for the global model and two of the three submodels (17a and 17cd, Fig. 3, Table 2). If precipitation or forest cover was not included in the models, standardized mean accuracy decreased by 30–50% and 25–40%, respectively (Fig. 3). Annual mean monthly minimum temperature and evapotranspiration (ET) in May and August were also included as important predictor variables for these three models. Standardized mean accuracy decreased by 22–40% if these predictor variables were not included in the model (Fig. 3). A decrease in accuracy for individual predictor variables does not directly correspond to the overall OOB error rate (Strobl et al., 2008). Snow water equivalent (SWE) on April 1 and May 1 and annual mean monthly minimum temperature were the most important predictor variables for submodel 17b (Table 2). Higher values of total annual precipitation, percent forest cover, ET and SWE corresponded to more wet classifications.

Streamflow permanence probabilities had high variability among subregional climate classes within the study area in addition to inter-annual variability within each climate class (Fig. 4). Notably, lower streamflow permanence probabilities corresponded to the climatically drier regions of the study area, which also generally had lower standard deviations in annual predictions (Fig. 4B). Higher streamflow permanence probabilities were concentrated in coastal and higher-elevation mountain regions, which had moderate inter-annual standard deviations. Relatively high standard deviations (e.g., 17 to 18%) were limited to a few isolated areas in the interior regions of the study area associated with transitional climatic zones (i.e., wet to dry or dry to wet transitions) (Fig. 4B).

The hydroclimatic conditions of the modeling period represent the wide range of variability that characterizes the study area (SI.1). Hydroclimatic conditions were broadly evaluated using the Self Calibrating Palmer Drought Severity Index (scPDSI, Wells et al., 2004), a common index of meteorological drought used in the U.S. The scPDSI serves as a convenient metric to evaluate hydroclimatic conditions because it integrates a suite of hydrologic conditions that presumably influence streamflow permanence including precipitation, temperature, and evapotranspiration. The scPDSI values are calculated using data over the entire period of record, in this case 1895–2016. Values range from about −10 to 10 where negative values reflect drought conditions. Values of −2 indicate moderate drought; values less than −4 indicate extreme drought. The scPDSI was used to identify the hydroclimatically driest and wettest years during the modeling period for further analysis of streamflow permanence in the three focal basins.

Table 2
Consolidated classification tables summarizing performance of final random forest global model and three subregional models.

Model		Predicted Dry	Predicted Wet	OOB estimate of error rate	Top 3 variables
Global	Observed Dry	1538	399	19.55%	Precip, Min Temp, Forest
	Observed Wet	359	1582		
17a	Observed Dry	843	189	17.68%	Precip, Forest, ET in May
	Observed Wet	176	856		
17b	Observed Dry	416	115	21.90%	SWE on May 1, Min Temp, SWE on April 1
	Observed Wet	118	415		
17cd	Observed Dry	283	91	20.40%	Precip, Forest, ET in August
	Observed Wet	62	314		
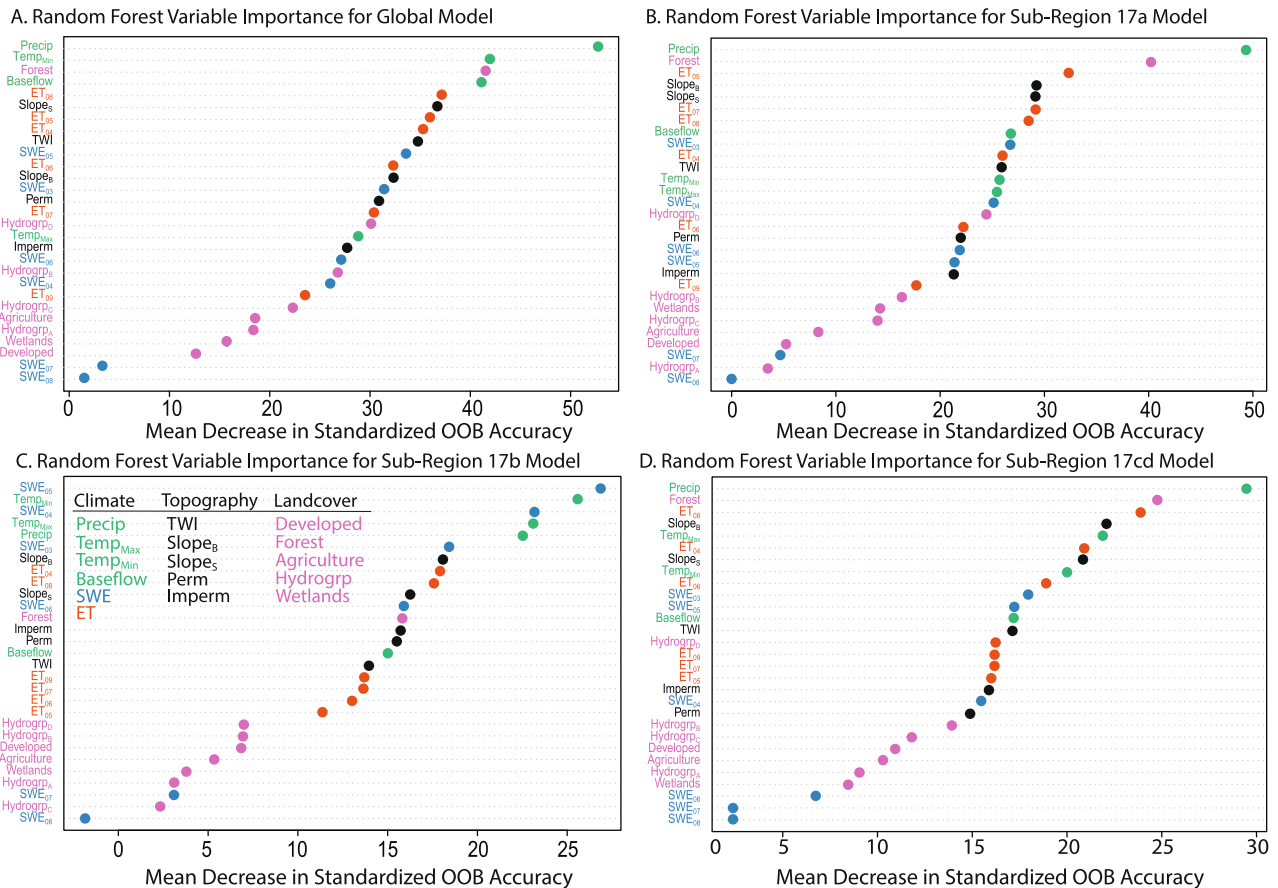
Fig. 3. Predictor variable importance plot for the global model of the entire study area (A) and 3 subregional models 17a (B), 17b (C), and 17cd (D) as a function of standardized mean decrease in accuracy. Mean percent of a given hydrologic soil group in the basin is denoted by Hydrogrp A/B/C/D, SlopeB is the average basin slope, SlopeS is the mean channel slope, and TWI is the basin average TWI. Subscripts on ET and SWE denote the year of measurement.

Additional analysis beyond the scope of this project is needed to evaluate potential patterns in streamflow permanence probabilities and year-to-year climate conditions.

The reliability of threshold values was assessed through evaluation of the predictions' maximum standard error of prediction (Appendix A) and comparison of streamflow permanence classifications with NHDPlus classifications (Appendix D). This resulted in 29 out of a total of 220 HUC8 basins being flagged as potentially unreliable (22 basins flagged with high maximum standard error of prediction values; seven basins flagged because of disagreement between NHDPlus classifications).

3.2. Using PROSPER to quantify flow permanence in focal basins

Distinct differences in predicted wet and dry conditions were evident between climatically drier (2004) and wetter (2011) years highlighting the sensitivity of flow permanence to hydroclimatic conditions (Fig. 5). Inter-annual variability in flow permanence had subsequent implications on the character and basin-scale configuration of wet and dry stream segments that varied across the three focal basins (Fig. 6). During years with less precipitation, the frequency and length of dry channel segments generally increased across the three focal basins, resulting in an overall increase in basin dryness and increased streamflow fragmentation (Fig. 6A and B). However, the inter-annual change in the position of dry channel segments within the river network was inconsistent across basins and years of different precipitation (Fig. 6C).

Annual tallies of the contiguous wet and dry stream segments for each of the three focal basins over the 13-year model period indicated that, on average, dry stream segments accounted for 55–92% of the stream network length within each basin (Fig. 6A). The Willow and Whitehorse Creeks were exceptionally dry; predicted dry stream segments accounted for more than 82% of the stream network in any year, while in the Methow and Boise Rivers basins dry stream segments were at least 50 and 42% of the stream network in any year, respectively. Predicted length of contiguous wet and dry stream segments ranged from 30 m (individual pixel) to 7230 km (Fig. 6A). The statistical distribution of the contiguous wet and dry stream segments is a reflection that the Methow and Boise Rivers are composed of long wet stream segments that correspond to the mainstem that extend into the lower reaches of tributaries, and short wet streams segments that are interspersed with longer dry stream segments in the upper reaches of tributaries. Willow and Whitehorse Creeks are drier versions of the Methow and Boise Rivers, in which the longest contiguous wet stream segments also corresponded to the mainstems, but the majority of the river network was dominated by longer (maximum of more than 500 km) contiguous dry stream segments relative to the Methow and Boise Rivers (Fig. 6A). The inter-annual variability in the statistical distribution of wet and dry stream segments was greatest for the Boise River basin; the range in annual proportion of predicted dry segments was 27% in the Boise River basin compared to 19% in the Methow basin and 16% for the Willow and Whitehorse basins. For the Methow and Boise River basins, the predicted lengths of dry stream segments appeared more variable relative to wet
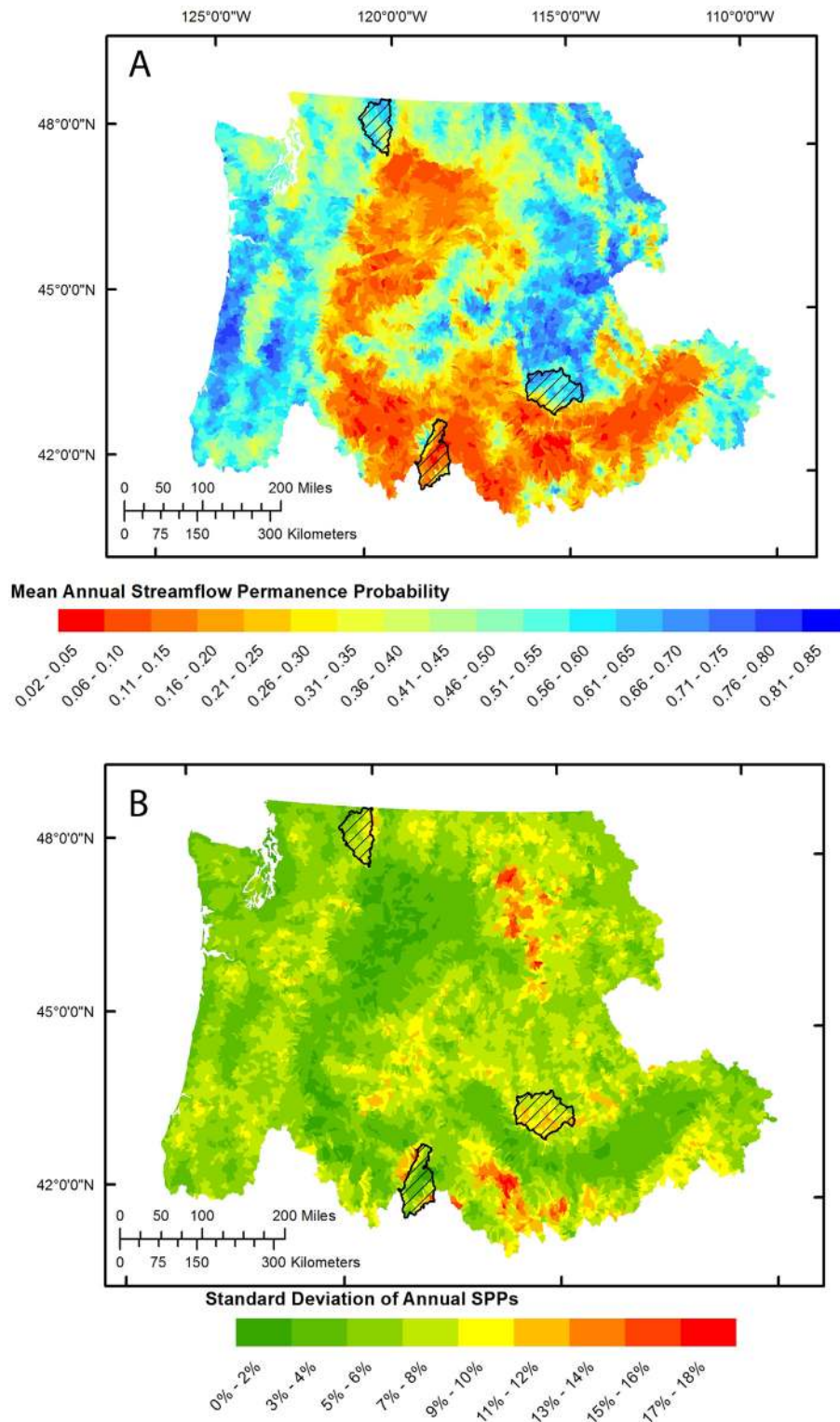
Fig. 4. Mean PROSPER model predictions of streamflow permanence probabilities (A) and associated standard deviation (B) for the entire study area for the 2004–2016 modeling period. Watershed boundaries for three focal basins are in black.

stream segments; however, wet stream segments are orders of magnitudes longer in these two rivers, with the mainstem and connected tributaries comprising a large and relatively consistent portion of the wetted networks.

Years with less precipitation generally resulted in greater overall basin dryness and generally followed the expected trend of

decreased wet-to-dry segment ratio, and consequently increased streamflow fragmentation, in the Methow River and Willow and Whitehorse Creek basins (Fig. 6B). This result is attributed to increased frequency of dry stream segments in tributaries in drier years. The Boise River basin exhibited a greater range in overall dryness of the basin relative to the other two focal basins that only
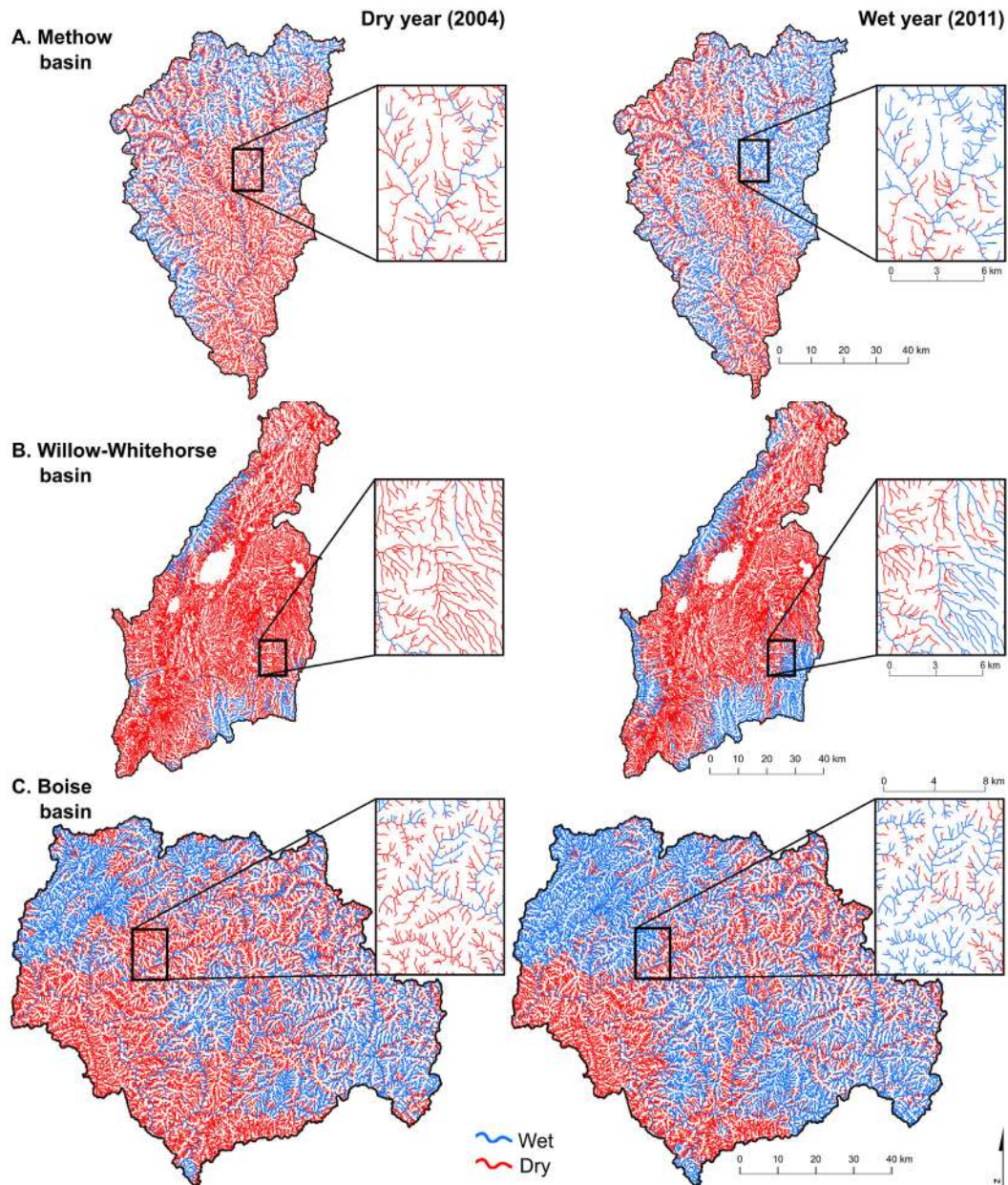
Fig. 5. Predicted wet and dry channel segments for the Methow River (A), Willow-Whitehorse Creek (B), and Boise River focal basins for 2004 and 2011, which represent the hydroclimatically driest and wettest years, respectively, over the 13-year modeling period.

moderately aligned with precipitation of individual years. However, year-to-year changes in streamflow fragmentation were limited in the Boise River basin, which is attributed to consistent contiguous wet segments that extend into the tributaries.

Finally, climatically drier years resulted in lower elevations having more dry segments, and consequently the remaining wet segments being higher elevation, as reflected in the increased ratio of mean elevation of wet and dry segments, in the Methow River and Willow and Whitehorse Creek basins, with little change in the Boise River basin (Fig. 6C). The Willow and Whitehorse Creek basins showed an increase in drainage area of dry segments relative to wet segments in years with less precipitation. However, the Methow River basin exhibited virtually no change in the drainage area ratio, along with the Boise River basin. Although the Boise River basin exhibited minimal inter-annual variability in both ratio descriptors, the slight increase of the elevation ratio in years of low precipitation adheres to expected patterns (Fig. 6C).

4. Discussion

The PROSPER model was built to provide predictions of annual streamflow permanence for 2004–2016 for all minimally impaired and unregulated streams and rivers in the Pacific Northwest concurrent with NHDPlus grids. We also used a localized threshold analysis to translate streamflow permanence probabilities produced with the PROSPER model back into discrete classes of wet or dry conditions with an associated confidence. We use these interpreted wet-dry conditions to derive additional descriptors of annual streamflow permanence conditions in three focal basins to demonstrate model utility for applied and basic science questions.

4.1. Model characteristics

Streamflow generation is the result of interactions between precipitation and the physiographic conditions of geology, topography,
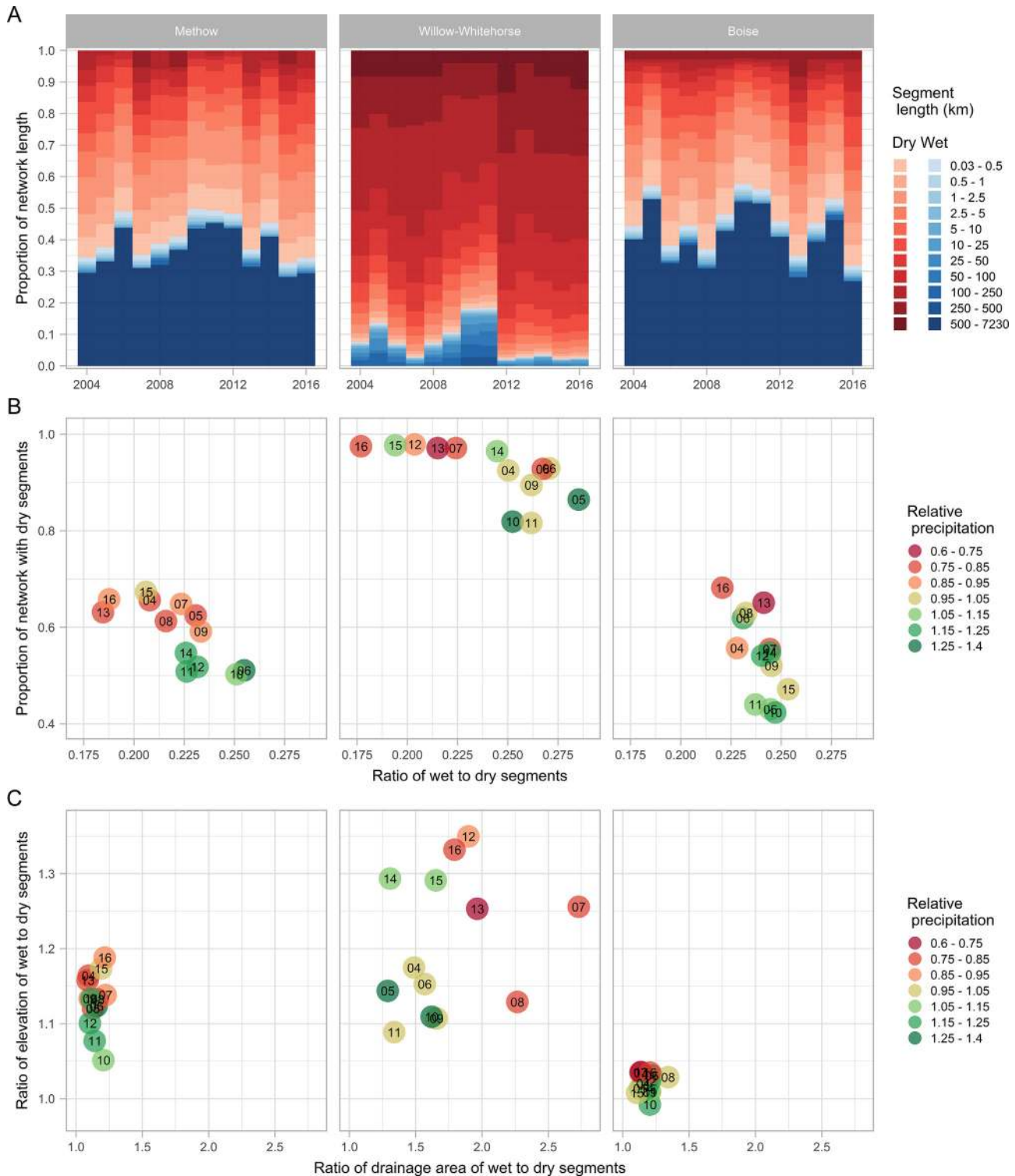
Fig. 6. Spatiotemporal variability of predicted streamflow permanence for the Methow River, Willow-Whitehorse Creek, and Boise River focal basins for the model period (2004–2016). A) The distribution of predicted contiguously wet or dry stream segment lengths and the cumulative proportion of the channel network for each year (A). The relation between the proportion of the channel network predicted to be dry and the ratio of the frequency of contiguously wet segments to the frequency of contiguously dry segments for each year (B). The relation between the ratio of contiguously wet segments to continuously dry segments in terms of both mean elevation and drainage area at the downstream extent for each year (C). Individual years are colored by total annual precipitation normalized to the 13-year modeling period for each focal basin.

soil type, and land cover (Winter, 2007; Buttle et al., 2012; Sayama et al., 2011). These interactions exhibit high spatial heterogeneity and vary considerably across the spatial extent of the study area (Tromp-van Meerveld and McDonnell, 2006). However, the inclusion of total annual precipitation as the most important predictor variable in the global model and two of the three subregional models underscores how precipitation exerts primary control on surface flow conditions at a regional scale (Buttle et al., 2012; Wenger et al., 2010) and can be an important determinant on baseflow conditions at smaller catchment scales (Belmar et al., 2016).

The importance of May 1 SWE for submodel 17b is more consistent with Sando and Blasch (2015), who identified average snow extent in late spring (March through July) as the most important variable in a Rocky Mountain watershed. This result may be driven by the high density of streamflow observations in the mountainous, snow-dominated geography of northern Idaho, which is similar to the catchment conditions of study sites used by Sando and Blasch (2015). The importance of May 1 SWE might also be augmented by the relative sparseness of streamflow observations in the western region of 17b, which includes rain-dominated and rain-snow mix climatic regimes in the Cascade Mountains and Columbia Plateau region (Leibowitz et al., 2016). Although the other subregions of the study area include some snow-dominated areas (e.g., eastern Oregon mountains), most precipitation falls as rain (temperate Coast Range, and more arid, low elevation inland areas). As a result, annual precipitation, which includes snow, more accurately describes the study area as a whole and the subregions 17a, c, and d.

Percent forest cover was a consistently important predictor variable, included in the global and two of the three subregional models. Although percent forest cover is correlated with precipitation, these results are consistent with findings from González-Ferreras and Barquín (2017), who employed random forest classification to map streamflow permanence probabilities in a Mediterranean catchment in Spain. Similar to the PROSPER model, increase in broadleaf forest cover increased classification of stream reaches as perennial (González -Ferreras and Barquín, 2017). The study area of Sando and Blasch (2015) was located on heavily-forested U.S. Forest Service lands and may have lacked the spatial heterogeneity in land cover for percent forest cover to be included as a distinguishing factor of flow permanence. Alternatively, differences in important variables on flow permanence may also be attributed to the differences in spatial scale between the two studies. Nonetheless, inclusion of percent forest cover as one of the most important predictor variables may reflect the ability of forest soils to regulate hydrologic conditions, including sustaining baseflows (Belmar et al., 2016). In particular, characteristically high infiltration rates of forest soils can facilitate both high soil water that contributes to baseflows along shallower pathways (Sayama et al., 2011) and deeper percolation to groundwater that also contributes to baseflow (Hewlett, 1961; Winter, 2007).

The absence of topographic wetness index as an important predictor variable was unexpected (Table 3). However, this variable was a basin average and it is likely to have greater influence on local conditions. Therefore, a basin average or local topographic wetness index may be more important for streamflow permanence predictions for basin model domains smaller than a regional or sub-regional scale.

4.2. Variability and uncertainty in model predictions

Comparison of streamflow permanence probabilities at identified flowing – and no flow-classified USGS gage locations across the range of climate classes of the study area provided a convenient method to validate PROSPER output in parts of the study area that lacked streamflow observations (Fig B.1 and B.2). We recognize that the density and spatial extent of USGS gage locations imposes limitations on this analysis as a comprehensive model validation tool. Nonetheless, consistent, statistically significant differences in streamflow permanence probabilities between flowing- and no flow-classified USGS stream gages across five of the six climate classes provided support for the validity of the relative streamflow permanence probabilities within a given climate class. The lack of statistically significant differences in streamflow permanence probabilities between flowing- and no flow-classified USGS stream gage locations for the wettest climate class is attrib-

uted to a limited number of dry observations relative to wet observations (27 vs 277) in this climate class and thus an inability to discern potential differences in streamflow permanence probabilities.

The variability of streamflow permanence probabilities for USGS gages across climate classes, however, illuminated the challenge of a global model to represent the geographic heterogeneity of the study area. In particular, because precipitation, forest cover, and temperature were the most important variables in the global model, stream channel pixels in arid climate classes, including pixels with consistent year-round flow, will have lower streamflow permanence probabilities compared to wetter climate classes. This is a function of the inherent arid-region characteristics of lower annual precipitation, lower forest coverage, and higher minimum temperatures. Smaller subregional or local watershed models may result in a consistent range of streamflow permanence probabilities that correspond to wet or dry streamflow conditions (Sando and Blasch, 2015; González-Ferreras and Barquín, 2017), but are likely constrained to a homogeneous climate.

The geographic variability that occurs in the threshold values, the reliability of threshold values at a location, and the resulting streamflow permanence classification of wet or dry with associated confidence are a result of two factors. First, the geographic variability reflects subregional- or local-scale processes not captured in the global PROSPER model, but which influence streamflow permanence. Second, adequate representation of both wet and dry conditions in the streamflow observation data result, at least in part, in threshold values that are considered reliable.

Local-scale controls can often be the dominating factor on flow permanence (Whiting and Godsey, 2016). Processes that include local surficial or hydrogeologic controls that can either contribute to sustained year-round flow in arid climates (Winter, 2007) or result in streams going dry in wetter climates (Jaeger et al., 2007) are not represented in the global PROSPER model at this time. As a result, it is likely that locations with stronger hydrogeologic controls on flow permanence may require smaller-scale models that are specific to those locations for more accurate prediction of flow permanence. In particular, the flagged HUC8 basins with high maximum standard error of prediction in the southwestern region of 17b (Fig. 5B) are candidate areas for smaller-scale models that may better capture local controls on flow permanence.

Sparsity of field observations particularly contributes to the variability in the uncertainty of threshold values and the consequent streamflow permanence classes that have lower confidence levels. For example, the HUC8 basins flagged with high maximum standard error of prediction that are concentrated along the Oregon Coast Range (Region 17d and western portion of Region 17c, Fig. 5B) lacked dry streamflow observations. Although this area is located in a very wet climate class, dry streamflow conditions occur, but in limited frequency relative to year-round flowing conditions (May and Gresswell, 2004). Therefore, concerted efforts to identify river networks in these HUC8 basins in which dry streamflow conditions occur have the potential to improve PROSPER's predictive ability in these basins.

In the absence of local-scale flow permanence prediction models, PROSPER streamflow permanence classifications in flagged HUC8 basins with high maximum standard error of prediction should be considered with caution by the end user. Wet and dry classification that have lower associated confidence within the study area are useful predictions depending on the objective of the end user. However, reliance on these predictions would benefit from ground-truthing field efforts. In these areas of less reliable threshold values, relative comparison of streamflow permanence probabilities between streams of interest and known perennial systems in the area may be a useful approach to evaluate streamflow permanence.

4.3. Variability in flow permanence characteristics in three focal basins

Results from comprehensive mapping of the three focal basins highlight the predominance of dry stream segments in these basins, but also yield information on how streamflow permanence expands and contracts under different hydroclimatic conditions. Initial descriptors of basin dryness, streamflow fragmentation, and position in the river network employed here serve as potential examples of how PROSPER output can be used to generate a myriad of metrics that quantitatively describe the temporal changes and spatial organization of wet and dry stream segments at local (e.g., 30-m^2) to basin (10^1 km^2) to regional landscape (10^{2-3} km^2) scales and which are germane to a wide range of applications. In addition to PROSPER-derived descriptors, the total variance in predictions, and degree in fluctuation in annual streamflow permanence probabilities can be used to develop and test hypotheses regarding flow permanence. In particular, PROSPER output can be used to identify sections of the river network that may be resilient or sensitive to drought conditions, allowing for management efforts that target protection of critical reaches (Sando and Blasch, 2015; Isaak et al., 2016).

4.4. Limitations of PROSPER

It is important to recognize the known limitations of the PROSPER model to help users avoid misinterpreting streamflow permanence conditions in some areas across the Pacific Northwest. A main limitation is that PROSPER does not account for the effects of streamflow regulation (dams) or diversions. Thus, for stream segments downstream of reservoirs, it is likely that streamflow permanence conditions are more stable and potentially more likely to flow year-round than what is predicted by PROSPER. Conversely, diversions and withdrawals from a stream network are likely to reduce the probability of streamflow permanence for downstream locations and should be considered in addition to the probability of streamflow permanence predicted by PROSPER.

Additionally, preliminary analysis of PROSPER results shows that spring-fed streams in arid climate regions tended to be biased toward low streamflow permanence probabilities, resulting in erroneous dry classifications. Future work for improving the PROSPER model includes adding a springs dataset obtained from the High Resolution NHD as a predictor variable in the model. Other predictor variables may also be included in future iterations of PROSPER that better capture local processes and conditions that influence streamflow permanence. Updated PROSPER models and associated predictions is publicly available through the USGS StreamStats platform (https://water.usgs.gov/osw/streamstats/) and in updated USGS ScienceBase Data Releases (Sando et al., 2018; Sando and Hockman-Wert, 2018). All data and processing scripts necessary for reproducing the results presented in this manuscript are permanently archived by the USGS, and are available upon request to the Wyoming-Montana Water Science Center.

Finally, PROSPER model output does not provide information on the hydroperiod of flowing conditions in terms of the timing and duration of flowing or no flow conditions for a given pixel. While it is assumed that dry classifications correspond to the late summer baseflow period that is typical of the Pacific Northwest region, further work is necessary to characterize the flow permanence hydroperiod (Arismendi et al., 2017), which has been shown to vary both regionally and through time (Eng et al., 2016).

While PROSPER is an unprecedented and valuable effort to model annual streamflow permanence, we recognize that it is not a substitute for on-the-ground local knowledge of hydrologic systems in the Pacific Northwest. Rather, our intent is for the model to serve as a supplemental dataset to help understand the dynamic nature of regional hydrology, particularly as it is affected by changing climatic conditions.

5. Conclusions

The PROSPER model is a moderate resolution, temporally resolved predictive model of streamflow permanence for the Pacific Northwest Region, U.S. The model provides annual predictions of streamflow permanence probabilities and wet or dry classifications at a 30-m spatial resolution for streams that correspond to the NHDPlus stream grid. Data in this form are publicly available (Appendix E) and have potential use in a wide array of applications across a range of spatial scales that can provide a better understanding of controls on streamflow permanence, and also allow for quantitative characterization of the spatiotemporal dynamism of streamflow permanence and its sensitivity or resilience to physiographic and hydroclimatic conditions.

Total annual precipitation was identified as one of the most important predictor variables in the global model and two of the three subregional models, which underscores the importance of precipitation as a primary control on surface flow conditions at a regional scale and baseflow conditions at smaller catchment scales. However, local-scale controls on flow permanence, including local surficial or hydrogeologic controls, are not yet adequately represented in the PROSPER model. Despite this, the PROSPER model delivered approximately 80% accuracy in correct classification of wet or dry streamflow observations for the Pacific Northwest study area. Some HUC8 basins with local controls on flow permanence may be candidate areas for smaller-scale models. However, increased observation data of no flow conditions that occur but may be infrequent in very wet climates such as the Oregon Coast Range could improve PROSPER prediction accuracy in basins.

An initial analysis of PROSPER predictions for three focal basins in the study area illustrate the ubiquity of dry stream segments, but also the year-to-year changes in the spatial composition and configuration of surface flow in these networks. Under drier climatic conditions, the proportion of wet segments decreased substantially in all three focal basins, especially in the most arid basin. The average elevation of both wet and dry segments in all three basins increased in years with less precipitation, but more so for wet segments, particularly in the wettest focal basin. Also, in years with low precipitation, only in the driest focal basin did the average drainage area of wet segments increase considerably. The mainstems and lower extremities of larger tributaries in wetter river basins may be able to persistently flow through still considerable portions of the river network during drought conditions, but in more arid river basins, only stream segments with increasingly larger catchments may be able to sustain year-round streamflow if drought conditions were to worsen, and even then, for a relatively minor portion of the river network.

Appendices: Introduction

The material presented in the following appendices support the PRObability of Streamflow PERmanence (PROSPER) model. Specifically, the appendices include methods that describe the following analyses and processes: a threshold analysis to support conversion of streamflow permanence probabilities into "wet" and "dry" streamflow permanence classifications (Appendix A), a comparison between PROSPER streamflow permanence probabilities and USGS streamgages (Appendix B), a supplementary product to evaluate PROSPER predictor variables (Appendix C), a comparison of PROSPER classifications with NHD classifications (Appendix D), and detailed directions on how to obtain PROSPER output for end-users (Appendix E).

Appendix A. A local threshold analysis to translate streamflow permanence probabilities to wet and dry classes

Methods

Streamflow permanence probabilities are determined by the relation between predictor variables and the streamflow permanence conditions (i.e., whether a location is in a wet or dry state) during assumed annual low-flow periods. These probabilities are defined in the random forest classification model as the ratio of classification trees that predict a site to be wet divided by the total number of classification trees (500). All streamflow observation locations used to build PROSPER therefore have a streamflow permanence probability between 0 and 1. Theoretically, dry-state observations have corresponding lower streamflow permanence probabilities while wet-state observations have higher streamflow permanence probabilities. However, in some applications, random forest has been shown to produce bias in the predictions. This bias, which can be generally described as over-predicting extreme low values and under-predicting extreme high values, has been demonstrated and described extensively in applications of random forest regression analysis (Xu, 2013; Zhang and Lu, 2012). This bias can result in over-predicting streamflow permanence probabilities at dry locations in generally wet environments, and under-predicting streamflow permanence probabilities at wet locations in generally dry environments. While this issue has received much attention in the context of random forest regression analysis (e.g., Xu, 2013; Zhang and Lu, 2012), there has been little documentation of similar bias in random forest classification.

To provide an alternative method for assigning a final class to predicted probabilities, a locally optimized (Singh et al., 2012; Kang et al., 2014), but regionally consistent, probability threshold analysis was developed. Rather than assume the probabilities are equally distributed around 0.5 everywhere in the study area, this analysis identifies threshold streamflow permanence probability values that effectively determine if the streamflow permanence probability for any given pixel is either statistically more similar to a wet site (above the threshold) or a dry site (below the threshold) in that local area. The resulting wet or dry classification is distinct from perennial and non-perennial classifications used by NHDPlus, in that it is variable not just in space, but also in time, capturing the annual conditions as a consequence of inclusion of climatic predictor variables. This analysis allows us to statistically categorize annual streamflow permanence probabilities into wet or dry classes for each year that are representative for different hydroclimatic regions, which can then be used to compare predictions across the study area. Translating probabilities to a meaningful descriptor of streamflow permanence condition (e.g., a wet or dry classification) facilitates quantitative description of the year-to-year variability of streamflow presence or absence at a variety

of spatial scales that extend from the local (30-m grid cell) to catchment or regional scale. The localized threshold analysis represents a novel approach to provide a spatially adaptive classification threshold in the random forest method.

The probabilistic threshold analysis, hereafter referred to as the threshold analysis, was conducted using the following steps. The predicted mean (2004–2016) streamflow permanence probabilities for separate populations of wet observations and dry observations were spatially interpolated across the study area using empirical Bayesian kriging (Pilz and Spöck, 2008), resulting in a mean wet and mean dry probability grid, respectively. The parameters of the semivariogram (Eq. (1)) used to interpolate the data were estimated for this study empirically by taking 100 bootstrap samples consisting of 30 observations and calculating semivariograms for each sample. For each prediction location, the prediction is calculated using a semivariogram distribution generated by a likelihood-based sampling of semivariograms in the neighborhood (345 km) of the location (Krivoruchko and Gribov, 2014).

The semivariogram model used in the empirical Bayesian kriging is shown as

$$\gamma(h) = Nugget + b|h|^\alpha \tag{1}$$

where

γ is the variance measured as a function of h

h is a given distance (maximum of 345 km for this analysis),

Nugget is the error associated with the modeled data

b is the slope, and

α is a value ranging between 0.25 and 1.75.

A threshold grid was created by averaging the mean wet probability grid and the mean dry probability grid, which represents the threshold value that will classify the streamflow permanence probability as wet or dry for that location. To evaluate the uncertainty around the threshold values, standard error of prediction grids for both the mean wet probability grid and the mean dry probability grid were created using the square root of the mean variance calculated from subsets used in the streamflow permanence probability grid generated by empirical Bayesian kriging. A mean standard error of prediction grid was then created by averaging the standard error of prediction grids for the mean wet and the mean dry probability grids.

Because the threshold and associated mean standard error of prediction grids are expected to represent catchment-scale, rather than site-scale, conditions, the grids were smoothed to avoid introducing abrupt localized fluctuations in the data at observations that might not be representative of the catchment-scale conditions. To smooth the grids, the average values for individual HUC8 watersheds within the study area were calculated. The HUC8 polygons were then converted to points at their centroids. The threshold and associated mean standard error of prediction values were then re-interpolated with a spatial resolution of 4 km. A series of confidence intervals (70%, 80%, 90%, and 95%) were constructed using the equation

$$CI_i = T_i \pm \frac{z * SEP_i}{\sqrt{n}} \tag{2}$$

where,

CI_i is the confidence interval value at pixel i

T_i is the threshold grid value at pixel i

z is the critical value for a 1-tailed Student's t test,

SEP_i is the mean standard error of prediction value at pixel i, and

n is the sample size.

The confidence intervals were used to categorize the streamflow permanence probability at each pixel into a streamflow permanence class, which consisted of a wet or dry classification

with an associated confidence represented as one of the 5 intervals. A total of 10 streamflow permanence classes ranged from −5 (dry classification with 95% confidence or greater) to 5 (wet classification with 95% confidence or greater). The streamflow permanence classes represent spatially explicit, binary categorical values.

High mean standard error of prediction values could indicate areas with potentially unreliable threshold values that are used to translate probabilities into classifications. To provide some qualifying measure of threshold value reliability, HUC8 regions with relatively high mean standard error of prediction values were identified and flagged to assess threshold value reliability. Specifically, standard error of prediction grids for the mean dry and the mean wet probability grids were calculated for each HUC8 region, and the maximum of the two means was retained to generate a grid of maximum mean standard error of prediction. Any HUC8 region with a maximum mean standard error of prediction value in the 90th percentile (0.19) or higher was flagged as having a potentially unreliable threshold value as a consequence of relatively large uncertainty around the threshold values in that HUC8 region.

The default method of using the majority class for assigning a final prediction class in random forest typically results in the best overall model performance. However, gaining accuracy in overall model performance can also mean sacrificing accuracy for subsets of data, if the response variable being predicted is not randomly distributed. To determine the effect on model performance of using the local threshold compared to the default 0.5 threshold, the out-of-bag classification accuracy associated with each threshold method was compared for the overall model, as well as for each HUC8. The differences in classification accuracy for the two classes, wet and dry, were visualized as a function of the proportion of streamflow observations in each HUC8 basin represented by that class and fit with a linear function using a generalized additive model (GAM) technique (Hastie, 2017) as part of the 'gam' package developed for R. Both GAM models were statistically significant (*p*-values less than 0.01).

The relationship between the standard error of prediction, local threshold value and associated confidence intervals

For a given stream grid pixel, a predicted streamflow permanence probability that is much greater or less than the threshold value at that location will tend to be assigned a streamflow permanence classification with a high associated confidence (e.g. −5 for dry with 95% confidence; 5 for wet with 95% confidence). However, the mean standard error of prediction directly affects the width of confidence intervals around threshold values. Lower mean standard error of prediction results in a narrower confidence interval; higher mean standard error of prediction results in a wider confidence interval. Consequently, for a given location with a low mean standard error of prediction associated with the threshold value, a predicted streamflow permanence probability for a stream pixel that is not necessarily far from the threshold value in that location can be assigned a streamflow permanence classification with high confidence. In contrast, threshold values with high mean standard error of prediction require streamflow permanence probability values to be farther from the threshold value in order to be assigned a streamflow permanence classification with high confidence. Therefore, locations in the study area with low mean standard error of prediction consequently have a higher frequency of a streamflow permanence classification at 95% confidence. Locations in the study area with high mean standard error of prediction have a higher frequency of lower confidence streamflow permanence classification (e.g. −3 to 3).

Results

The threshold analysis was conducted for the entire study area at a 4 km resolution, which classified streamflow permanence probabilities into either a wet or dry streamflow permanence class with associated confidence (Fig. A1). Localized threshold values that determined wet or dry classes ranged from 0.32 to 0.77 with a mean of 0.53. Variability in threshold values appeared to be regionally coherent. Lower threshold values occurred in the arid, lower elevation regions of the study area including the Columbia Plateau and Snake River plain; the highest threshold values occurred in the interior, high elevation mountain region that is generally associated with the three wetter climate classes. Uncertainty in the threshold values, as represented by the standard error of prediction, was a function of streamflow observation density. The lowest mean standard error of prediction occurred in the eastern half of the study area, where observational data density was the highest. Correspondingly, the highest standard error of prediction values occurred in the western third of the study area, where data density was lower, and which includes subre-
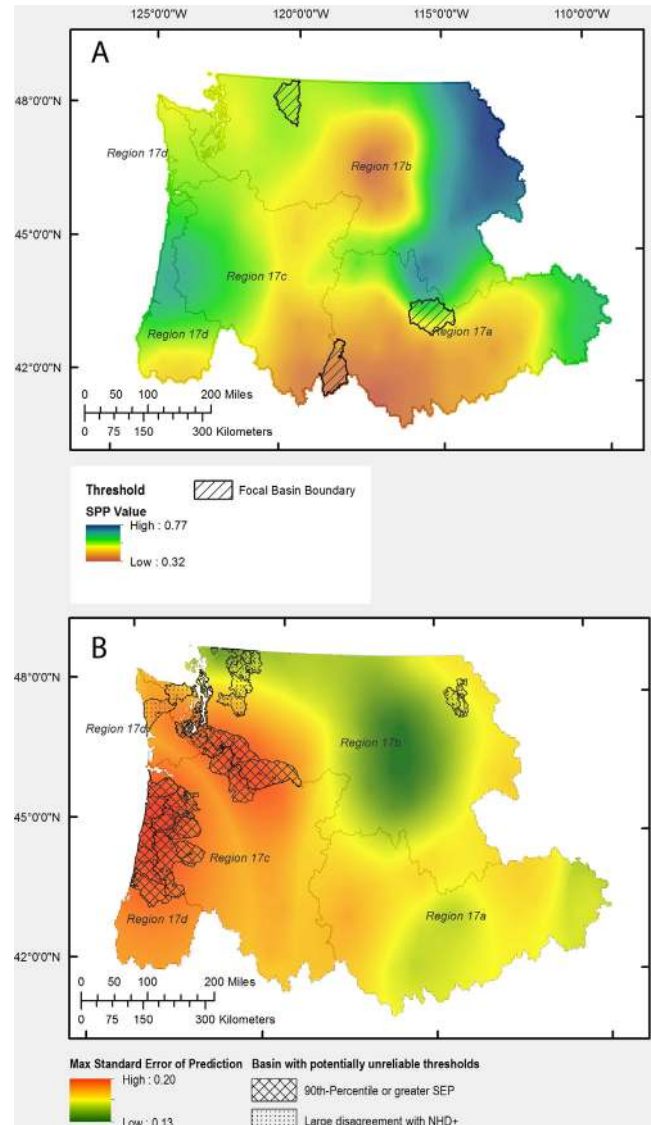
Fig. A1. Mean threshold streamflow probability prediction (SPP) that determines if a given 30-m grid cell is wet or dry for the study area (A). Mean standard error of prediction (SEP) for each threshold value for the study area (B).

gion 17d and the western portion of subregions 17b and 17c (Figs. 1 and A1). On average, 82% of the pixels classified as dry were predicted with 95% confidence over the study period with little year-to-year variability (16%) (Table A1). The percentage of wet pixels predicted with 95% confidence accounted for 46% of all predicted wet pixels with substantially more (37%) year-to-year variability.

The out-of-bag classification accuracy was essentially the same using the two threshold methods, with slightly better performance using the default threshold (0.5). When the default threshold was used to assign a class, the accuracy was 81% for wet observations and 79% for dry observations. When the locally optimized threshold was used, the accuracy was 80% for wet observations and 78% for dry observations. However, when the classification accuracy for each method was first stratified by HUC8 regions, results

indicate that the mean probability of accurately classifying dry observations increased by 7%, while the mean probability of accurately classifying wet observations decreased by 6%. Furthermore, when the change in classification accuracy caused by using the local threshold is modeled as a function of normal (1981–2010) annual precipitation (PRISM Climate Group) averaged for each HUC8 basin, results show a marked increase in classification accuracy of dry observations in HUC8 regions that typically receive more than about 500 mm/year of precipitation (wet basins) (Fig. A2). There is also a slight increase in classification accuracy of wet observations in HUC8 regions that typically receive less than 500 mm/year of precipitation (dry basins). Conversely, there is a slight decrease in classification accuracy of dry observations in dry basins, and a larger decrease in classification accuracy of wet observations in wet basins. Twenty-two HUC8 regions were

Table A1
Annual proportions of all wet (streamflow permanence classes 1–5) and dry (streamflow permanence classes -1 - -5) pixels predicted at each confidence level.

SPC	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Mean
−5	0.82	0.79	0.84	0.85	0.83	0.82	0.80	0.79	0.84	0.84	0.83	0.81	0.85	0.82
−4	0.05	0.05	0.04	0.04	0.04	0.05	0.05	0.05	0.04	0.04	0.04	0.05	0.04	0.05
−3	0.05	0.06	0.05	0.05	0.05	0.06	0.06	0.06	0.05	0.05	0.05	0.06	0.04	0.06
−2	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.03	0.02	0.03	0.03	0.02	0.03
−1	0.05	0.06	0.05	0.04	0.05	0.05	0.06	0.06	0.05	0.04	0.05	0.05	0.04	0.05
1	0.16	0.16	0.16	0.17	0.16	0.15	0.14	0.15	0.15	0.17	0.16	0.17	0.16	0.18
2	0.08	0.08	0.08	0.08	0.08	0.08	0.07	0.07	0.08	0.08	0.08	0.09	0.08	0.09
3	0.14	0.14	0.14	0.15	0.14	0.13	0.13	0.13	0.14	0.15	0.14	0.15	0.14	0.16
4	0.10	0.10	0.11	0.11	0.10	0.10	0.10	0.10	0.10	0.11	0.11	0.11	0.11	0.11
5	0.52	0.51	0.51	0.49	0.53	0.54	0.57	0.55	0.54	0.49	0.51	0.49	0.51	0.46
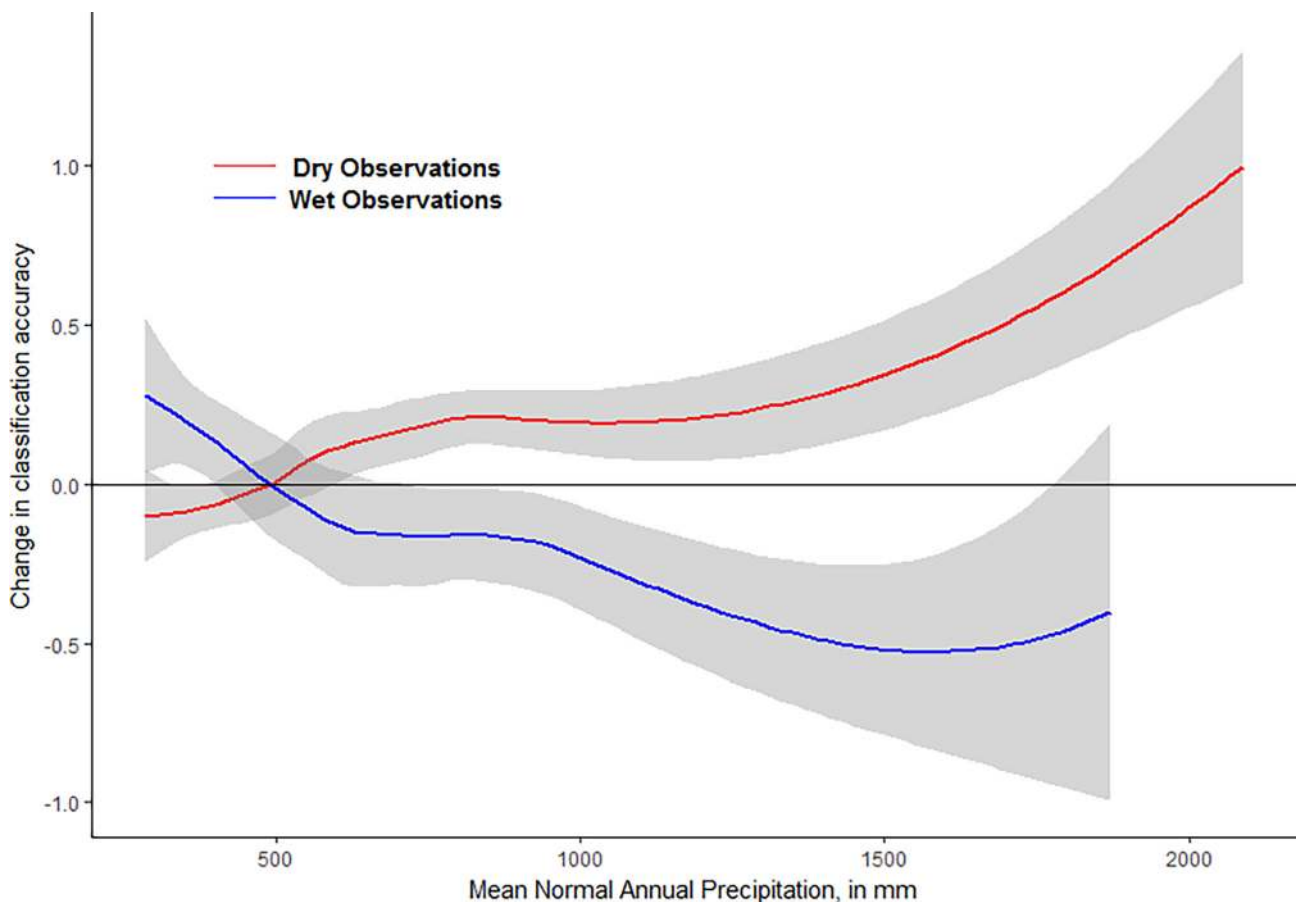
Fig. A2. Differences in classification accuracy using the local threshold versus the 0.5 default for wet observations and dry observations against mean normal annual precipitation for the HUC8 basins in the study area in which classification accuracy differences exist.

Fig. B2. Boxplots of mean PROSPER streamflow permanence probabilities (SPP) for the 2004–2016 modeling period at U.S. Geological Survey stream gages classified as either with flow (1st percentile of daily streamflow less than 0.0283 m^3s^{-1}) or with no flow (1st percentile of daily streamflow greater than or equal to 0.0283 m^3s^{-1}). USGS stream gages were distributed across six different climate classes defined for Washington, Oregon and Idaho (Leibowitz et al., 2016); gages located outside this three-state area were not assigned a climate class. Results of Welch's unequal variance t-test for significant differences are reported for each climate class in addition to the number of USGS gages in each gage class within each climate class.

Results

Streamflow permanence probabilities were significantly lower for no flow-classified USGS gages across the climate classes (Fig. B2). The Very Wet climate class was the only exception in which statistical differences were not detected between streamflow permanence probabilities of flow- and no flow-classified gages. Differences in the range of streamflow permanence probabilities across climate classes highlighted the variability in how streamflow permanence probabilities represent streamflow permanence. For example, streamflow permanence probabilities that correspond to flow-classified USGS gages in the arid climate class were low enough to be well within the range of streamflow permanence probabilities that correspond to no flow-classified USGS gages in other climate classes, thus presenting a challenge for a global interpretation of streamflow permanence probabilities. Recognition of this variability consequently necessitated the subsequent threshold analysis that allows translation of streamflow permanence probabilities back into wet or dry conditions (Appendix A).

Appendix C. A predictor variable suitability grid to evaluate PROSPER streamflow permanence predictions

A predictor variable suitability grid was created to identify pixels for which predictor variable values extend beyond those associated with streamflow observation point locations that were used in the model development. The number of predictor variables with values that fell outside of the range of values used for model development was compiled for each pixel to create the Predictor Variable Suitability Grid. Larger numbers reflect more predictor variables at that 30-m location that were outside of predictor variable values used in model development and thus might indicate potentially less reliable predictions. A complementary attribute includes the proportion of predictor variables with values within the range of predictor variable values that were used in

model development. Larger proportions reflect more predictor variables at that 30-m location that were within the range of predictor variable values used in model development and thus can indicate more reliable predictions. While portions of the study area might have a lower density of streamflow observations, for example central and eastern Washington, these areas are potentially still well represented if they are statistically similar (defined in predictor variable space) to other locations that are represented in the streamflow observation dataset used to calibrate the random forest model.

Approximately 38% of pixels within the study area have at least one predictor variable that was outside the range of values used in model development. However, more than 80% of those pixels have less than 5 variables outside of the model range. The maximum number of predictor variables for an individual grid cell pixel at which values extended beyond the range of model development is 101.

Currently, the predictor variable suitability grid does not identify the specific predictor variables whose values are beyond those used in the model; future work should identify specific predictor variables in the grid and their spatial distribution in the study area. However, in the absence of this analysis, a reasonable approach for end users may be to simply consider these areas that have a relatively high number value in the grid (e.g., more than 5) as potentially not well represented by the model. High number values in the grid results in higher uncertainty in streamflow permanence probabilities at those locations and may account for streamflow permanence probabilities and classifications that are not aligned with on-the-ground conditions.

Appendix D. Comparison of PROSPER streamflow permanence classifications with NHDPlus streamflow classifications

Streamflow permanence classifications were compared to NHDPlus streamflow classifications (perennial or intermittent) for HUC8 watersheds in the study area as another method to quan-

tify the reliability of identified threshold values (Appendix A). Seven HUC8 watersheds, mostly clustered in the North Cascades Mountain Region along the northern border of the study area (Fig. A1), were flagged because of disagreement with NHDPlus classifications. Although there is recognition that NHDPlus classifications are an imperfect baseline for comparison based on the known errors in the NHDPlus classification system (Fritz et al., 2013), it is nevertheless the most spatially comprehensive streamflow permanence dataset for the study area, particularly for data-sparse regions. There was an approximately 80% agreement between streamflow field observations used in the model development and the NHD classifications, using both the medium and high resolution NHD. The streamflow permanence classifications were extracted from the PROSPER stream grid cells using a 15-meter buffer around the NHDPlus flowlines. For each HUC8 region, the proportion of 30-m grid cells classified as wet, based on the streamflow permanence classifications averaged over the years 2004–2016, was compared to the proportion of stream length classified as perennial in the NHDPlus dataset. The HUC8 regions were ranked and sorted according to both the proportion of pixels classified as wet by the streamflow permanence classification and the proportion of streams classified as perennial by NHDPlus. The HUC8 streamflow permanence classification ranks were plotted against NHDPlus ranks and a 90% confidence interval ellipse was constructed. The HUC8 regions outside the 90% confidence interval with NHDPlus were flagged as having potentially unreliable threshold values.

Appendix E. Obtaining PROSPER predictions in StreamStats

StreamStats is a Web-based GIS application that was created by the USGS to provide users with access to an assortment of analytical tools and datasets that are useful for water-resource planning and management (U.S. Geological Survey, 2016). StreamStats, as well as a brief description of the application and links to user instructions, definitions, fact sheets, and other information, can be accessed at http://water.usgs.gov/osw/streamstats/. It is recommended that, in addition to the application description and user instructions, users read the limitations for the StreamStats application before attempting to use StreamStats.

The PROSPER data are publicly available through the StreamStats Web App, as well as in ScienceBase as downloadable GeoTiffs. It is recommended that StreamStats is used for obtaining PROSPER predictions at individual locations or a subset of locations, while ScienceBase is used for obtaining PROSPER data for large regions. In StreamStats, the user can view each of the annual PROSPER streamflow permanence probability grids for the Pacific Northwest. When a particular pixel (stream location) is selected, a pop up window will contain the streamflow permanence probability for each annual probability grid selected by the user, as well as the respective wet or dry classifications made using the local threshold analysis (Appendix A).

For a single location:

1. Go to the StreamStats Web App (https://streamstats.usgs.gov/ss/) and click the "Exploration Tools" link in the upper left corner application window. Select "PROSPER Tool".
2. Select the years of data you want to analyze in the "Include in query" boxes and change the "Displayed layer" to the dataset you want to visualize in the application window.
3. Zoom to your area of interest, or type a location in the box provided in the table of contents on the left side of the application window.
4. Click on a pixel to show the PROSPER predictions for the years included in the query.

5. If you wish to adjust the configuration, click "Configure" in the "Exploration Tools" window.
6. When you are satisfied with the location and data you have selected, Click "Continue" under the "Build Report" tab (it should be shown automatically).

Upon completing steps 1–6, a report should be generated that provides the data specified, as well as any warnings or flags associated with the location. In the report are options to download the data as a comma-delimited text file (.csv).

Appendix F. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.hydroa.2018.100005.

References

Acuña, V. et al., 2014. Why should we care about temporary waterways? Science 343 (6175), 1080–1081.

Arismendi, I., Dunham, J.B., Heck, M.P., Schultz, L.D., Hockman-Wert, D., 2017. A statistical method to predict flow permanence in dryland streams from time series of stream temperature. Water 9 (12), 946.

Barrett, A.P., 2003. National operational hydrologic remote sensing center snow data assimilation system (SNODAS) products at NSIDC. National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences Boulder, CO.

Belmar, O., Barquín, J., Álvarez-Martínez, J.M., Peñas, F.J., Del, M., The role of forest maturity on catchment hydrologic stability. 2016.

Benstead, J.P., Leigh, D.S., 2012. An expanded role for river networks. Nature Geosci. 5 (10), 678.

Beven, K., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrol. Sci. J. 24 (1), 43–69.

Bhamjee, R., Lindsay, J.B., Cockburn, J., 2015. Monitoring ephemeral headwater streams: a paired-sensor approach. Hydrol. Process.

Blasch, K.W., Ferré, T., Christensen, A.H., Hoffmann, J.P., 2002. New field method to determine streamflow timing using electrical resistance sensors. Vadose Zone J. 1 (2), 289–299.

Boulton, A.J., Rolls, R.J., Jaeger, K.L., Datry, T., 2017. Hydrological connectivity in intermittent rivers and ephemeral streams. In: Intermittent Rivers Ephemeral Streams. Elsevier, pp. 79–108.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Buttle, J. et al., 2012. An overview of temporary stream hydrology in Canada. Can. Water Res. J. 37 (4), 279–310.

Caruso, B.S., 2014. GIS-based stream classification in a mountain watershed for jurisdictional evaluation. JAWRA J. Am. Water Res. Assoc. 50 (5), 1304–1324.

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 Resour. Res. 50 (8), 6542–6548.

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 land cover. Ecohydrology 9 (7), 1141–1153.

Costigan, K.H. et al., 2017. Flow regimes in intermittent rivers and ephemeral streams. In: Intermittent Rivers Ephemeral Streams. Elsevier, pp. 51–78.

Datry, T., Bonada, N., Boulton, A.J., 2017. Chapter 1 – General Introduction, Intermittent Rivers and Ephemeral Streams. Academic Press, pp. 1–20. https://doi.org/10.1016/B978-0-12-803835-2.00001-2.

Datry, T. et al., 2018. Flow intermittence and ecosystem services in rivers of the Anthropocene. J. Appl. Ecol.

Donnelly, J., Naugle, D.E., Hagen, C.A., Maestas, J.D., 2016. Public lands and private waters: scarce mesic resources structure land tenure and sage-grouse distributions. Ecosphere 7 (1).

Ebersole, J.L., Wigington Jr, P.J., Leibowitz, S.G., Comeleo, R.L., Sickle, J.V., 2014. Predicting the occurrence of cold-water patches at intermittent and ephemeral tributary confluences with warm rivers. Freshwater Sci. 34 (1), 111–124.

Eng, K., Wolock, D.M., Dettinger, M., 2016. Sensitivity of intermittent streams to climate variations in the USA. River Res. Appl. 32 (5), 885–895.

Feddema, J.J., 2005. A revised Thornthwaite-type global climate classification. Phys. Geogr. 26 (6), 442–466.

Fritz, K.M. et al., 2013. Comparing the extent and permanence of headwater streams from two field surveys to values from hydrographic databases and maps. JAWRA J. Am. Water Resour. Assoc. 49 (4), 867–882.

Fritz, K.M., Johnson, B.R., Walters, D.M., 2006. Field operations manual for assessing the hydrologic permanence and ecological condition of headwater streams. United States Environmental Protection Agency Office of Research and Development.

Fritz, K.M., Johnson, B.R., Walters, D.M., 2008. Physical indicators of hydrologic permanence in forested headwater streams. J. North Am. Benthol. Soc. 27 (3), 690–704.

Fry, J. et al., 2011. Completion of the 2006 national land cover database for the conterminous United States. Photogramm. Eng. Remote Sens. 77 (9), 858–864.

Godsey, S., Kirchner, J., 2014. Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density, flowing channels and stream order. Hydrol. Process. 28 (23), 5791–5803.

González-Ferreras, A., Barquín, J., Mapping the temporary and perennial character of whole river networks. Water Resources Research. 2017.

Griffith, G., 2010. Level III North American Terrestrial Ecoregions: United States Descriptions. Prepared for the North American Commission for Environmental Cooperation (www. cec. org), version May, 11: 2010.

Gungle, B., 2006. Timing and Duration of Flow in Ephemeral Streams of the Sierra Vista Subwatershed of the Upper San Pedro Basin, Cochise County, Southeastern Arizona. 2328-0328.

Hastie, T.J., 2017. Generalized additive models. In: Statistical Models in S. Routledge, pp. 249–307.

Hewlett, J.D., 1961. Soil Moisture as a Source of Base Flow from Steep Mountain Watersheds. US Department of Agriculture Forest Service, Southeastern Forest Experiment Station.

Homer, C. et al., 2007. Completion of the 2001 national land cover database for the counterminous United States. Photogramm. Eng. Remote Sens. 73 (4), 337.

Homer, C. et al., 2015. Completion of the 2011 National Land Cover Database for the conterminous United States–representing a decade of land cover change information. Photogramm. Eng. Remote Sens. 81 (5), 345–354.

Hooshyar, M., Kim, S., Wang, D., Medeiros, S.C., 2015. Wet channel network extraction by integrating LiDAR intensity and elevation data. Water Resour. Res. 51 (12), 10029–10046.

Isaak, D.J. et al., 2016. Slow climate velocities of mountain streams portend their role as refugia for cold-water biodiversity. Proc. Natl. Acad. Sci. 113 (16), 4374–4379.

Jaeger, K.L., Montgomery, D.R., Bolton, S.M., 2007. Channel and perennial flow initiation in headwater streams: management implications of variability in source-area size. Environ. Manage. 40 (5), 775–786.

Jaeger, K.L., Olden, J., 2012. Electrical resistance sensor arrays as a means to quantify longitudinal connectivity of rivers. River Res. Appl. 28 (10), 1843–1852.

Jensen, C.K., McGuire, K.J., Prince, P.S., 2017. Headwater stream length dynamics across four physiographic provinces of the Appalachian Highlands. Hydrol. Process.

Kang, J., Zhang, N., Shi, R., 2014. A Bayesian nonparametric model for spatially distributed multivariate binary data with application to a multidrug-resistant tuberculosis (MDR-TB) study. Biometrics 70 (4), 981–992.

Konrad, C.P., Drost, B.W., Wagner, R.J., 2003. Hydrogeology of the Unconsolidated Sediments, Water Quality, and Ground-water/surface-Water Exchanges in the Methow River Basin, Okanogan County. US Department of the Interior, US Geological Survey, Washington.

Krivoruchko, K., Gribov, A., 2014. Pragmatic Bayesian kriging for non-stationary and moderately non-Gaussian data. In: Mathematics of Planet Earth. Springer, pp. 61–64.

Leibowitz, S.G. et al., 2016. Hydrologic Landscape Characterization for the Pacific Northwest, USA. JAWRA J. Am. Water Resour. Assoc. 52 (2), 473–493.

Liaw, A., Wiener, M., 2002. Classification and regression by Random Forest. R News 2 (3), 18–22.

Ludington, S., et al., 2007. Preliminary integrated geologic map databases for the United States. California, Nevada, Arizona, Washington, Oregon, Idaho, and Utah. Version, Western States, 1.

May, C.L., Gresswell, R.E., 2004. Spatial and temporal patterns of debris-flow deposition in the Oregon Coast Range USA. Geomorphology 57 (3), 135–149.

McKay, L. et al., 2012. NHDPlus Version 2: User Guide. National Operational Hydrologic Remote Sensing Center, Washington, DC.

McShane, R.R., Sando, R., Hockman-Wert D.P., 2017. Streamflow Observation Points in the Pacific Northwest, 1977-2016: U.S. Geological Survey data release, https://doi.org/10.5066/F7BV7FSP.

Michael, J.L., 2004. Best management practices for silvicultural chemicals and the science behind them. Water Air Soil Pollut. Focus 4 (1), 95–117.

Nadeau, T.-L. et al., 2015. Validation of rapid assessment methods to determine streamflow duration classes in the Pacific Northwest USA. Environ. Manage. 56 (1), 34–53.

Perkin, J.S. et al., 2017. Groundwater declines are linked to changes in Great Plains stream fish assemblages. Proc. Natl. Acad. Sci. 114 (28), 7373–7378.

Pervez, M.S., Brown, J.F., 2010. Mapping irrigated lands at 250-m scale by merging MODIS data and national agricultural statistics. Remote Sens. 2 (10), 2388–2412.

Pilz, J., Spöck, G., 2008. Bayesian spatial sampling design. na.

Poff, N.L. et al., 1997. The natural flow regime. Bioscience 47 (11), 769–784.

Poff, N.L., Pyne, M.I., Bledsoe, B.P., Cuhaciyan, C.C., Carlisle, D.M., 2010. Developing linkages between species traits and multiscaled environmental variation to explore vulnerability of stream benthic communities to climate change. J. North Am. Benthol. Soc. 29 (4), 1441–1458.

Pringle, C., 2003. What is hydrologic connectivity and why is it ecologically important? Hydrol. Process. 17 (13), 2685–2689.

PRISM Climate Group, 2004. Oregon State University accessed February 2017. http://prism.oregonstate.edu, .

Sando, R., Blasch, K.W., 2015. Predicting alpine headwater stream intermittency: a case study in the northern Rocky Mountains. Ecohydrol. Hydrobiol. 15 (2), 68–80.

Sando, R., Hockman-Wert, D.P., 2018. Probability of Streamflow Permanence (PROSPER) Model Output Layers: U.S. Geological Survey data release, https://doi.org/10.5066/F77M0754.

Sando, R., Olsen, T.D., Kaiser, K.E., Haluska, T., Hockman-Wert, D.P., 2018. In: Probability of Streamflow Permanence (PROSPER) Continuous Parameter Grids (CPGs). U.S. Geological Survey data release. https://doi.org/10.5066/F7CC0ZXH.

Sayama, T., McDonnell, J.J., Dhakal, A., Sullivan, K., 2011. How much water can a watershed store? Hydrol. Process. 25 (25), 3899–3908.

Schruben, P.G.A., Bawiec, R.E., King, W.J., Beikman, P.B., Helen, M., 1994. Geology of the Conterminous United States at 1: 2,500,000 Scale–A Digital Representation of the 1974 PB King and HM Beikman Map.

Schultz, L. et al., 2017. Spatial and temporal variability in the effects of wildfire and drought on thermal habitat for a desert trout. J. Arid Environ. 145, 60–68.

Senay, G.B. et al., 2013. Operational evapotranspiration mapping using remote sensing and weather datasets: a new parameterization for the SSEB approach. JAWRA J. Am. Water Resour. Assoc. 49 (3), 577–591.

Singh, T.R., Roy, S., Singh, O.I., Sinam, T., Singh, K., 2012. A new local adaptive thresholding technique in binarization. arXiv preprint arXiv:1201.5227.

Sörensen, R., Zinko, U., Seibert, J., 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrol. Earth Syst. Sci. Discuss. 10 (1), 101–112.

Spence, C., Mengistu, S., 2016. Deployment of an unmanned aerial system to assist in mapping an intermittent stream. Hydrol. Process. 30 (3), 493–500.

Soil Survey Staff, 2016. Natural Resources Conservation Service, United States Department of Agriculture. US General Soil Map (STATSGO2).

Stanley, E.H., Fisher, S.G., Grimm, N.B., 1997. Ecosystem expansion and contraction in streams. Bioscience 47 (7), 427–435.

Stoeser, D.B. et al., 2005. Preliminary Integrated Geologic Map Databases for the United States. US Geological Survey Open-File Report, p. 1351.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinf. 9 (1), 307.

Sun, X., Thompson, C., Croke, B., 2011. Using a logistic regression model to delineate channel network in southeast Australia. Proc. of 19th International Congress on Modelling and Simulation, Perth (Australia).

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2014.

Tromp-van Meerveld, H., McDonnell, J., 2006. Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. Water Resour. Res. 42 (2).

Turner, D.S., Richter, H.E., 2011. Wet/dry mapping: using citizen scientists to monitor the extent of perennial surface flow in dryland regions. Environ. Manage. 47 (3), 497–505.

U.S. Environmental Protection Agency (USEPA) and the U.S. Geological Survey (USGS), 2012. National Hydrography Dataset Plus NHDPlus. Version 2.10. Accessed 27 June 2016. https://www.epa.gov/waterdata/nhdplus-national-hydrography-dataset-plus.

Welch, B.L., 1947. The generalization ofstudent's' problem when several different population variances are involved. Biometrika 34 (1/2), 28–35.

Wells, N., Goddard, S., Hayes, M.J., 2004. A self-calibrating Palmer drought severity index. J. Clim. 17 (12), 2335–2351.

Wenger, S.J., Luce, C.H., Hamlet, A.F., Isaak, D.J., Neville, H.M., 2010. Macroscale hydrologic modeling of ecologically relevant flow metrics. Water Resour. Res. 46 (9).

Whiting, J.A., Godsey, S.E., 2016. Discontinuous headwater stream networks with stable flowheads, Salmon River basin Idaho. Hydrol. Process. 30 (13), 2305–2316.

Williamson, T.N., Agouridis, C.T., Barton, C.D., Villines, J.A., Lant, J.G., 2015. Classification of ephemeral, intermittent, and perennial stream reaches using a TOPMODEL-based approach. JAWRA J. Am. Water Resour. Assoc. 51 (6), 1739–1759.

Winter, T.C., 2007. The role of ground water in generating streamflow in headwater areas and in maintaining base flow. JAWRA J. Am. Water Resour. Assoc. 43 (1), 15–25.

Wolock, D.M., 2003a. Hydrologic Landscape Regions of the United States. 2331–1258. US Geological Service.

Wolock, D.M., 2003b. Base-flow index grid for the conterminous United States: US Geological Survey open-file report 03-263, digital data set.

Wolock, D.M., 2003c. Flow characteristics at US Geological Survey streamgages in the conterminous United States: US Geological Survey open-file report 03-146, digital data set.

Wood, M.S., Rea, A., Skinner, K.D., Hortness, J.E., 2009. Estimating Locations of Perennial Streams in Idaho Using a Generalized Least-Squares Regression Model of 7-Day, 2-Year Low Flows. US Geological Survey. 2328-0328.

Xu, R., 2013. Improvements to Random Forest Methodology.

Zhang, G., Lu, Y., 2012. Bias-corrected random forests in regression. J. Appl. Stat. 39 (1), 151–160.