water

Article

Precision of Headwater Stream Permanence Estimates from a Monthly Water Balance Model in the Pacific Northwest, USA

Konrad C. Hafen [1,2,*], Kyle W. Blasch [1], Paul E. Gessler [3], Roy Sando [4] and Alan Rea [5]

1 Idaho Water Science Center, U.S. Geological Survey, Boise, ID 83702, USA; kblasch@usgs.gov
2 Water Resources Program, University of Idaho, Moscow, ID 83844, USA
3 Department of Forest, Rangeland, and Fire Sciences, University of Idaho, Moscow, ID 83844, USA; paulg@uidaho.edu
4 Wyoming-Montana Water Science Center, U.S. Geological Survey, Helena, MT 59601, USA; tsando@usgs.gov
5 National Geospatial Program, U.S. Geological Survey, National Geospatial Program, Boise, ID 83702, USA; ahrea@usgs.gov
* Correspondence: khafen@usgs.gov

check for updates

Citation: Hafen, K.C.; Blasch, K.W.; Gessler, P.E.; Sando, R.; Rea, A. Precision of Headwater Stream Permanence Estimates from a Monthly Water Balance Model in the Pacific Northwest, USA. *Water* **2022**, *14*, 895. https://doi.org/10.3390/w14060895

Academic Editor: Zhenhua Zou

Received: 3 February 2022
Accepted: 10 March 2022
Published: 12 March 2022

Abstract: Stream permanence classifications (i.e., perennial, intermittent, ephemeral) are a primary consideration to determine stream regulatory status in the United States (U.S.) and are an important indicator of environmental conditions and biodiversity. However, at present, no models or products adequately describe surface water presence for regulatory determinations. We modified the Thornthwaite monthly water balance model (MWBM) with a flow threshold parameter to estimate flow permanence and evaluated the model's accuracy and precision for more than 1.3 million headwater stream reaches in the U.S. Pacific Northwest (PNW). Stream reaches were assigned to one of eight calibration groups by unsupervised classification based on sensitivity to MWBM parameters. Suitable MWBM parameter sets were identified by comparing modeled stream permanence estimates to surface water presence observations (SWPO). Parameter sets with accuracies > 65% were considered suitable. The MWBM estimated stream permanence with high precision at 40% of reaches, with poor precision at 20% of reaches, and no suitable parameter sets were identified for 40% of reaches. Results highlight the need for increased SWPO collection to improve calibration and assessment of stream permanence models. Additionally, implementation of the MWBM to estimate surface water presence indicates potential for process-based models to predict stream permanence with future development.

Keywords: surface water; stream permanence; perennial; intermittent; water balance model

1. Introduction

In the United States, stream permanence classifications (i.e., perennial, intermittent, and ephemeral) are a primary consideration to determine stream regulatory status under the Clean Water Act and are also an important indicator of environmental conditions and biodiversity [1–3]. Currently, the National Hydrography Dataset (NHD) [4] is the most comprehensive dataset describing stream permanence for the contiguous United States (CONUS) [2]. However, NHD stream permanence classifications (SPC) have been shown to exhibit up to 50% disagreement with in-situ observations [5,6] and the highest disagreement rates occur on headwater streams [7]. The US Environmental Protection Agency and Army Corps of Engineers have determined that NHD SPC, derived decades in the past, are usually not adequate for regulatory determinations, and that more reliable SPC mapping products are required that consider dynamic climate and land use in headwater and low stream order environments (https://www.epa.gov/sites/production/files/2020-01/documents/nwpr_fact_sheet_-_mapping.pdf, accessed on 27 January 2022).

A primary reason for the uncertainty of NHD SPC is that designations originated from observations made by topographic survey crews throughout the 1900s [5,7,8]. The classification methods of survey crews incorporated first-hand knowledge of stream reaches but this was limited to climatic conditions during the survey year (or recent past). Thus, climatic conditions during the survey year may not represent long-term average conditions and survey crew observations may not capture the full range of variability for surface water conditions at a stream reach [7]. Another reason for uncertainty is the SPC designations themselves. The definitions for perennial, intermittent, and ephemeral streams have changed through time, and streams were not given definitions that could be accurately assessed by survey crews over a short period of time [7–9]. For example, perennial streams were defined as "streams that flow continuously except in periods of extreme drought", but extreme drought is not defined [9]. Additionally, surveys for adjacent topographic maps were often made multiple years apart. Thus, SPC designations for part of a watershed could have been made under very different climate conditions than other portions of the same watershed [7].

Multiple efforts have sought to add dynamic context to static NHD SPC by modeling streamflow for NHD stream reaches at sub-annual time steps [10,11]. Two efforts include the NHD enhanced runoff method (EROM), which is a unit hydrograph approach that is integrated into the NHDPlus (i.e., NHD flowlines with additional descriptive attributes) itself [10], and a machine learning approach developed by the U.S. Geological Survey [11] (hereafter referred to as the "Miller data" or "Miller streamflow model"). Key to both modeling efforts were runoff estimates that are generated from the USGS Thornthwaite Monthly Water Balance Model (MWBM) [12]. The MWBM has been implemented in multiple studies to evaluate various components of the hydrological cycle, water supply, and water demand for the CONUS [13–17]. These studies demonstrate the utility of the MWBM to represent streamflow patterns over large spatial extents at a monthly time step. However, previous studies do not necessarily inform surface water presence because most of the gages used for MWBM calibration represent perennial streams draining larger basins, which gives a limited understanding of how well the model represents smaller basins with non-perennial streamflow [18].

Recently, stream permanence was assessed in Australia with a daily water balance model calibrated and validated with stream gauge data [19,20]. These results indicate that the MWBM may be useful for modeling stream permanence, especially since the MWBM can easily be applied at regional, continental, and even global extents. Because the MWBM is already integrated with the NHD and other hydrological modeling efforts [10,21], applying the MWBM to estimate stream permanence could quickly yield results over large spatial extents. However, procedures for estimating, calibrating, and assessing stream permanence with the MWBM have not been developed. With the greatest uncertainty in SPC usually occurring on headwater streams [5,7], data sources in addition to streamflow time series will be required for model calibration and assessment [22–24].

Both EROM and the Miller streamflow model are calibrated to gauged streamflow time series, which is the general standard for hydrological modeling. However, results can only be evaluated where continuous gaging stations exist. In the United States, stream gauges are strategically placed to primarily monitor water supply and flood [25]. As a result of these priorities, the majority of stream gauges are located on larger-order rivers and streams, with very few gauges in headwater catchments [18]. It is in headwater catchments where the greatest uncertainty of NHD SPC is documented and where few continuous records of streamflow exist to inform traditional modeling efforts [7,23]. As a result, performance (e.g., accuracy and uncertainty) of regional- and national-extent models in headwater streams is largely unknown. Assessing model performance in headwater streams is especially salient because more than 50% of all stream reaches are classified as headwaters [26].

To increase the spatial coverage of calibration and validation data from fixed stream gauges, some statistical and process-based models have used simple surface water presence observations (SWPO) to develop and evaluate stream permanence estimates [22,24,27–29]. SWPO are easy to collect and represent more spatial locations than streamflow gauges [30]. While SWPO represent many locations, most SWPO locations have only one observation. Thus, SWPO have much lower temporal resolution than stream gauge data [30]. SWPO are also subject to temporal and spatial bias (observations made in convenient locations at convenient times), subjectivity, and differing definitions of what constitutes a 'wet' or 'dry' observation. Nevertheless, SWPO have been effectively used to develop statistical models at local [27,28] and regional extents [17,20] and process-based models [24,29,31] at local extents where existing stream gauge networks did not provide sufficient information to model stream permanence in headwaters and other small tributary streams.

Herein, we implement the MWBM to estimate stream permanence on headwater streams of the United States' Pacific Northwest (PNW) region using SWPO for model calibration and assessment. Our primary objective is to determine if the MWBM can generate dynamic (e.g., change year-to-year) SPC estimates with accuracy better than NHD SPC (e.g., >60%) [5,7]. To accomplish this objective, we first assess the accuracy of annual stream permanence estimates generated from the MWBM to identify parameter combinations that perform at least as well as NHD SPC. Then, with the suitable parameter combinations, we calculate the precision of MWBM SPC on headwater streams throughout the PNW. This is a unique application of the MWBM and will increase knowledge about the accuracy of the MWBM in often difficult-to-model headwater streams. Additionally, this study will help establish how the MWBM can be used to assess stream permanence at large spatial extents and identify where more data and different modeling approaches may be necessary to improve stream permanence estimates.

2. Methods

2.1. Study Area

The study area consisted of the Pacific Northwest region of the United States (PNW). We define the PNW as Hydrographic Region 17 of the United States [32]. The PNW is bounded on the north by the United States–Canada border, the west by the Pacific Ocean, and the east and south by the boundaries of the Columbia River Basin, coastal watersheds, and contains several endorheic basins (Figure 1). Elevations in the region range from sea level in the coastal regions to over 4000 m above sea level in the Cascade Mountain Range. Annual precipitation ranges from 200 mm in the rain shadow of the Cascade Mountains to over 5000 mm in the coastal regions. Previous studies have examined stream permanence in this region through field observation and modeling, resulting in a rich dataset of publicly available surface water presence observations (Figure 1) that are imperative for this modeling study [7,22,33,34].

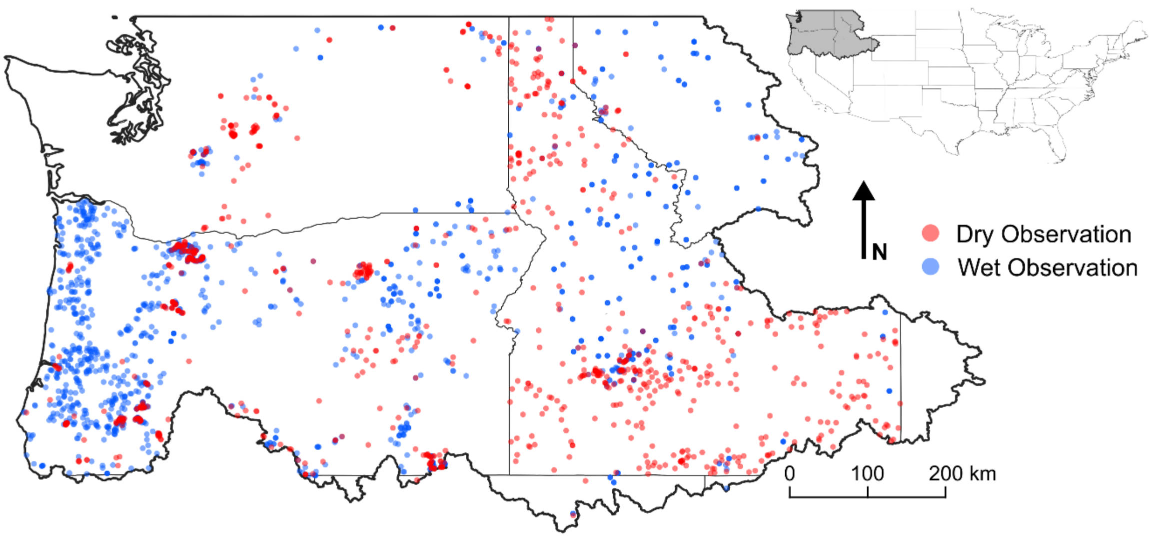

Figure 1. Location of in situ observations of surface water presence or absence on headwater streams in the Pacific Northwest from 1977 to 2019 (McShane et al., 2017).

2.2. The Monthly Water Balance Model

The MWBM has been previously implemented at global and CONUS extents to estimate monthly runoff [13,17,35,36]. The MWBM generates monthly runoff values by estimating the magnitude of hydrologic processes that supply and demand water. Equations governing hydrological processes were originally presented by [37,38] and are explained in detail by [10]. The basic model logic for the MWBM implemented herein progresses as follows and is similarly explained by [13] following the methodology of [12].

Model inputs are mean monthly temperature (*T*), total monthly precipitation (*P*), soil water holding capacity, and latitude. Monthly estimates of *P* and *T* were obtained from monthly PRISM (~4 km) data for each month from January 1975 to December 2019 [39]. PRISM climate data were used because other monthly climate datasets (e.g., Daymet, Gridmet) did not cover the entire SWPO period of record (1977–2019) with data for two prior years (1975–1976) for a model spin-up period. Soil water holding capacity estimates were obtained from STATSGO polygons [40]. STATSGO polygons were intersected with the extents of PRISM grid cells to obtain the water holding capacity corresponding to the footprint of each PRISM grid cell. A default value of 150 mm [12] was used in areas where STATSGO data were not available.

The MWBM allows precipitation to occur as rain or snow, determined by *T*. Snowfall accumulates from month to month to form a snowpack, which melts as temperatures warm. Rainfall can be converted to direct runoff, evapotranspiration (ET), soil moisture storage, and surplus water. Monthly potential evapotranspiration (PET) is determined by *T* and latitude per the Hamon equation [41]. When the sum of rainfall and snowmelt for a month is less than PET, actual evapotranspiration (AET) is the sum of rainfall, snowmelt, and the portion of water that is evaporated from the soil. When the sum of rainfall and snowmelt is greater than or equal to PET, AET is equal to PET. Water remaining after AET recharges soil water storage. Water in excess of AET and soil water storage becomes surplus. A specified proportion of surplus is converted to runoff each month, and the remaining surplus is temporarily held in storage. Thus, water is lost through AET and total runoff is the sum of direct runoff and surplus runoff. Runoff was summed for each NHD catchment and multiplied by catchment area to arrive at a runoff volume for each month. Runoff volumes were converted to mean monthly flow (volume/s).

In addition to ET, water fluxes in the MWBM are modulated by five parameters: runoff factor, direct runoff factor, snow temperature, rain temperature, and snow-melt coefficient (Table 1). Rain temperature (TR) is the temperature above which all precipitation falls as rain. Snow temperature (TS) is the temperature below which all precipitation falls as snow. When temperature is less than TR and greater than TS the proportion of rain to snow is determined by linear interpolation. Snow-melt coefficient (MC) is the maximum proportion of snow storage that can melt in a single month. Direct runoff (DR) is the proportion of precipitation and snowmelt that becomes overland runoff. Runoff factor (RF) determines the proportion of watershed storage that is converted to runoff each month.

We also added precipitation factor (PF) and temperature addition (TA) parameters to the MWBM models presented within this paper to adjust climate inputs to the MWBM, as in previous studies [13]. PF and TA were included in sensitivity and calibration analyses to identify the impact of climate inputs on MWBM performance. Descriptions, units, and value ranges of model parameters are presented in Table 1.

To convert MWBM runoff estimates to permanent or non-permanent stream classifications, we added a third parameter, flow threshold (FT), to the MWBM models presented herein (Figure 2). Streams were classified as permanent for a month (model time step) when mean monthly streamflow was greater than FT and non-permanent for a month when mean monthly streamflow was less than FT. Flow threshold has been used in previous stream permanence modeling studies to identify surface water presence and absence and flow values from these studies were used to determine the range of FT for this study [20,24,29]. We classified a stream as permanent for a calendar year (annual permanence) when the MWBM predicted a stream to be permanent for each month of the year. Calendar year was used instead of hydrological year because the irrigation season extends through October in most of the PNW. Streams the MWBM predicted to be dry at least one month of a calendar year were classified as non-permanent for the calendar year. We determined stream permanence at an annual time step to align with the observational data available for calibration (i.e., the year for many dry observations was recorded but not the month). Additionally, the current best data describing stream permanence for headwater streams (i.e., the NHD) use perennial and non-perennial classifications which describe permanence at a minimum time step of one year.

Table 1. Parameters assessed in sensitivity analysis and calibration of the monthly water balance model.

Parameter Name	ID	Description	Units	Range
Runoff Factor	RF	Proportion of catchment storage that is converted to streamflow each month	-	0.0–1.0
Direct Runoff Factor	DR	Proportion of precipitation that is converted to streamflow without infiltrating or evaporating	-	0.0–0.5
Snow Temperature	TS	Temperature below which all precipitation is snow	°C	−10.0–2.0
Rain Temperature	TR	Temperature above which all precipitation is rain	°C	0.0–10.0
Snow-Melt Coefficient	MC	The maximum proportion of snow water equivalent that can melt in a single month	-	0.0–1.0
Flow Threshold	FT	Mean monthly flow above which a stream segment is considered permanent	L/s	0.0–14.2
Precipitation Factor	PF	Multiplier for input PRISM precipitation	-	0.1–2.0
Temperature Addition	TA	Value added to increase or decrease mean monthly temperature	°C	−2.0–2.0

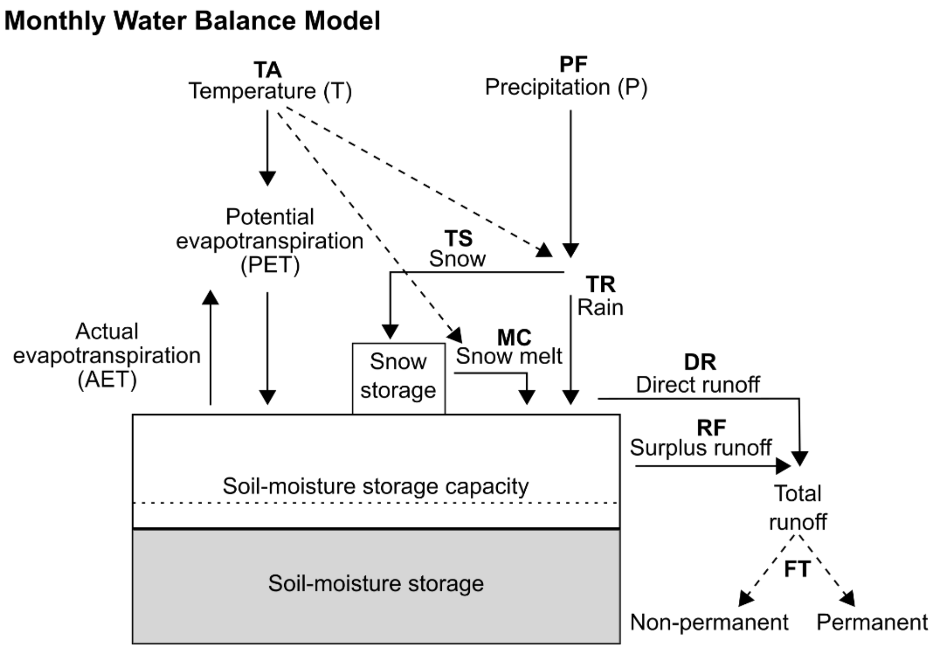

Figure 2. Diagram of the monthly water balance model. Bold parameter abbreviations correspond to the parameter descriptions and ranges in Table 1. Adapted from McCabe and Markstrom (2007) [10].

2.2.1. Model Application

Stream catchments and flowlines in the PNW were represented by the High Resolution version of the National Hydrography Dataset with added attributes (NHDPlus HR) [4]. The MWBM was applied only to headwater streams in the PNW as determined by NHDPlus HR. Modeling was limited to headwater streams because these are the portions of the stream network where NHD SPC exhibit the most uncertainty and where calibration and validation data for previous MWBM applications have been limited. Additionally, by excluding downstream segments, we eliminated the need to include flow routing in the model and could perform more model computations to evaluate model performance in the headwaters. Assessing only headwater stream segments also limited the impacts of human-altered landscapes on basin hydrology. Less than 0.1% of headwater catchments had permanent snow or ice that covered greater than 10% of the catchment area.

We defined headwater streams and catchments as those with no upstream segments or tributaries where the NHDPlus 'Start_Flag' attribute equaled 1 [42]. This included only the uppermost portions of the stream network (see example in Figure 3). With the MWBM, we calculated total monthly runoff within each headwater catchment in the PNW at the resolution of the input PRISM grid (~4 km). Total monthly runoff volume for headwater catchments was calculated by multiplying MWBM runoff depth by catchment area. When catchments intersected multiple PRISM grid cells, the volume was calculated for each intersecting area and summed to obtain the total volume. We calculated the mean monthly flow rate (L/s) at the outlet of each headwater stream segment as the catchment's total monthly runoff volume divided by the number of seconds in each month. Thus, MWBM results represent runoff estimates at the outlet (most downstream point) of each headwater stream in the PNW.

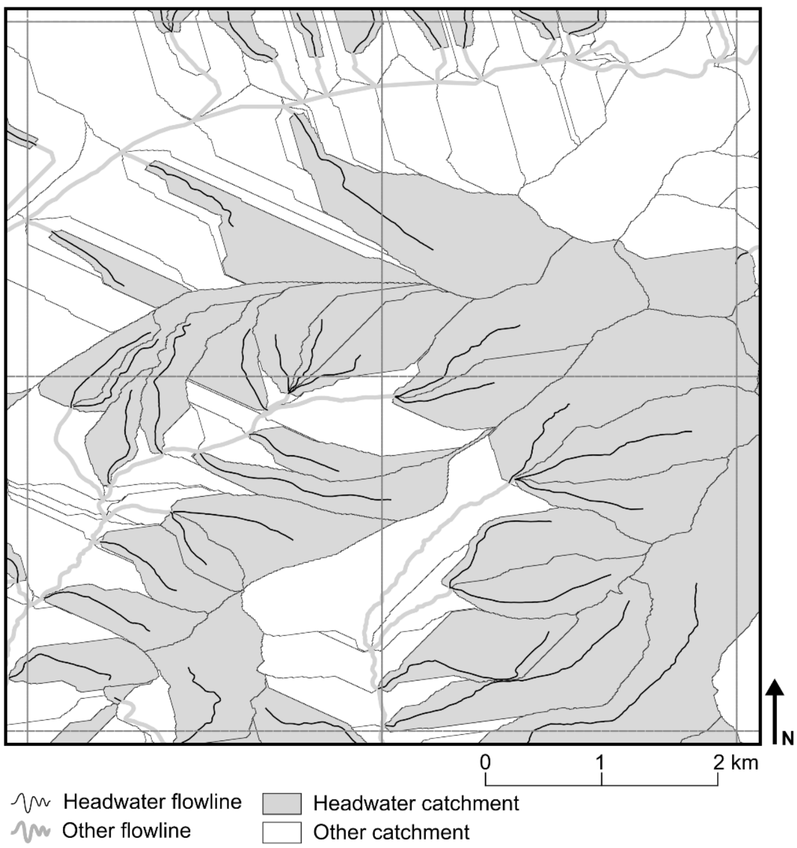

Figure 3. Example of catchments and stream segments that show the geospatial fabric to which the monthly water balance model was applied. Grid shows the extent of individual PRISM [39] grid cells (~4 km).

Monthly stream permanence was determined with the FT parameter. When mean monthly flow was less than FT the stream reach was classified as non-permanent for that month. When mean monthly flow was greater than or equal to FT the stream reach was classified as permanent for that month. Stream reaches that were classified as permanent every month of a calendar year were classified as permanent for that year. Any stream reach that was classified as non-permanent for least one month during a calendar year was classified as non-permanent for that year. This analysis grouped both intermittent and ephemeral streams into the non-permanent category.

We implemented the MWBM, as presented by [12], in the Python programming language [43] in order to adapt the MWBM to headwater streams and scale millions of model runs on a super computer. All MWBM model runs were completed on the USGS Yeti supercomputer [44].

2.2.2. Data Availability

All input data for this methodology are from publicly available sources. Final results and intermediate products from this study are available from [45].

2.3. Observation Data

Model calibration and precision analyses were conducted with observational data that described the presence or absence of surface water in headwater streams (SWPO). Only three USGS stream gauges [46] were located on headwater streams (as defined by NHDPlus HR) in the PNW. Therefore, SWPO provide the best spatial coverage and hydrological information for most headwater streams in the PNW. Observational data were primarily acquired from a dataset compiled as part of the Probability of Stream Permanence (PROSPER) [22] modeling effort in the PNW [33]. These data were supplemented with more recent SWPO collected with the FLOwPER application [34]. SWPO were made by a variety of agencies, including the Sauk-Suiattle Indian Tribe, Idaho Department of Environmental Quality, Oregon Department of Fish and Wildlife, U.S. Bureau of Land Management, U.S. Forest Service, USGS, and the U.S. Environmental Protection Agency. SWPO spanned the years 1977–2019. Each observation recorded the date (some older SWPO only specified the year) the observation was made, the geographic coordinates of the observation location, and the presence (wet) or absence (dry) of surface water in the observed stream channel (Figure 1).

All dry SWPO were used to represent streams that were annually non-permanent. Only wet SWPO made in August or September were used to represent annually permanent streams since streams in the study area that maintained flow in the driest months at the end of summer were assumed to be perennial. An observation point was assumed to represent the entire flowline (i.e., headwater stream reach). Each observation was joined to the nearest NHDPlus HR flowline. SWPO farther than 50 m from a flowline were excluded from the analysis. All SWPO greater than 20 m from a flowline were manually inspected to ensure they were associated with the correct flowline and SWPO near a confluence were inspected to ensure the observation was assigned to the correct flowline. There were no flowlines with observations from multiple years. Similar methodologies were implemented by [7] and [22] to identify SWPO for stream permanence assessment and modeling. In all, 2804 SWPO (1120 dry and 1684 wet) were used for model calibration (Figure 1). Though the MWBM produces runoff and stream permanence estimates at a monthly time step, we assessed the model on annual permanence because approximately half of the dry SWPO only specified the year of the observation and not the day and month.

2.4. Sensitivity Analysis

We assessed the relative sensitivity of annual stream permanence classification from the MWBM to each of the eight model parameters (Table 1) with the Fourier Amplitude Sensitivity Test (FAST [47,48] over the period 1977–2019. Sensitivity analysis was conducted with the SALib Python module [49]. In all, we tested 1200 parameter combinations for 1.3 million headwater catchments in the PNW. Parameter values were selected from uniform distributions that spanned the range of parameter values recommended by previous studies [12,13,17] displayed in Table 1. For many stream reaches, the sensitivity analysis produced "not-a-number" values. Upon inspection of the model runs, we observed that most of the not-a-number results occurred on stream reaches where greater than 95% of simulations predicted the same outcome (i.e., permanent, non-permanent) for all years. This indicated low sensitivity to model parameters for these stream reaches. To represent this sensitivity in further analysis, the sensitivity value for stream reaches where the initial FAST result produced not-a-number, but where greater than 99% of FAST simulations produced the same result (i.e., stream permanence classification), was set to 0.01.

2.5. Parameter Regionalization

Parameters for headwater stream reaches were regionalized by grouping reaches that responded similarly to changes in model parameters [13,50]. It is common for the optimal parameter values of hydrological models to vary spatially [51]. Parameter regionalization results in different model parameterizations for different regions to improve model outputs. We performed an unsupervised, K-means classification [52] based on MWBM parameter sensitivities to assign stream reaches to calibration groups (i.e., regions). Each headwater stream reach, as defined by NHDPlus HR, was classified based on its sensitivity to the eight MWBM parameters (as calculated in the previous section). We selected the number of calibration groups by qualitatively balancing the K-means classification error (squared distance from each point to the class center) with the number of SWPO available for each group. The number of SWPO for each group decreased as the number of calibration groups increased. We tested the K-means classification for 2–15 groups and determined that eight calibration groups best minimized K-means classification error while simultaneously maximizing the number of SWPO in each group. Physical characteristics, parameter sensitives and number of SWPO for each calibration group are presented in the Results section.

2.6. Parameter Set Selection

One million parameter sets (randomly selected using a Monte Carlo approach) from the same uniform distributions with the same ranges used for sensitivity analysis (Table 1) were evaluated for each calibration group using the spotpy python module [53]. For parameter set selection, we only ran simulations for the stream reaches where SWPO were located. Suitable parameter sets were identified by their accuracy with SWPO. The accuracy of a parameter set was determined as the sum of simulated stream permanence classifications that agreed with SWPO divided by the total number of SWPO within a given calibration group. We used previous studies that quantified the accuracy of NHD SPC on headwater streams to range from 50 to 65% [5–7] as benchmarks to identify parameter sets that produced a good model 'fit' in each calibration region. Therefore, we required suitable parameter sets to have greater than 65% overall accuracy against all SWPO and greater than 60% accuracy against wet or dry SWPO, individually. All parameter sets that produced suitable accuracy results in a calibration region were retained to assess model precision.

2.7. Precision Analysis

Modeled stream permanence was assessed on model precision (i.e., consistency). Selection of suitable parameter sets (see Section 2.6) provided an assessment of model accuracy, but the limited number (compared to implementations of the MWBM that were assessed with streamflow values where millions of observations are available) of SWPO (2357) presented opportunity for overfitting. Additionally, because all SWPO were used for model calibration we did not have an independent subset of data to assess overall model performance. Therefore, we used consistency of modeled annual stream permanence (i.e., permanent, or non-permanent) between suitable parameter sets as the model evaluation metric (referred to as precision, or model precision, hereafter).

To calculate model precision, annual stream permanence was predicted for all suitable parameter sets in each calibration region. This resulted in multiple stream permanence estimates for each headwater stream reach. Model precision was assessed by agreement of simulated stream permanence classifications between suitable parameter sets. High model precision was exhibited when all suitable parameter sets simulated the same permanence classification. Model precision decreased as different parameter sets predicted different permanence classifications for the same stream in the same year. We represented model precision as the proportion of parameter sets resulting in a permanent classification. Model precision values ranged from 0.0 to 1.0 where values of 0.0 and 1.0 indicated all parameter sets simulated non-permanent (0.0) and permanent (1.0) classifications, respectively. A value of 0.5 indicated that half of the parameter sets simulated a permanent condition, and the other half simulated a non-permanent condition, resulting in poor model precision. Precision values of 0.0–0.1 indicated high precision for non-permanent classifications and values of 0.9–1.0 indicated high precision for permanent classifications. Precision values 0.25–0.75 indicated poor precision.

For example, 20 individual stream permanence predictions would be made each year for each stream reach in a calibration group where 20 suitable model parameterizations were identified (one prediction for each parameter set). In a year where 19 of the parameter sets resulted in a permanent classification, the precision value would be 0.95. On the same stream reach, in a different year, it is possible that 10 parameter sets resulted in a permanent classification (precision value of 0.5), indicating inconsistencies in classification between parameter sets and, thus, poor model precision. In yet another year, it is possible that 0 parameter sets resulted in a permanent classification (precision value 0.0), indicating high agreement between model predictions for a non-permanent classification. Thus, the precision value is similar to estimating the probability a stream reach is permanent for a given year based on model agreement.

3. Results

3.1. Sensitivity Analysis

Parameter sensitivity values ranged from 0.0 to 1.0, where larger values indicated greater parameter sensitivity. Generally, parameter sensitivities were low for interior mountain ranges, moderate for coastal mountain ranges, and high on the interior plains and plateaus (Figure 4). Most headwater streams exhibited some degree of sensitivity to the RF and FT parameters. This is expected as RF is the main parameter controlling monthly flow rates and FT is the primary parameter controlling the model result (permanent or non-permanent).

3.2. Parameter Regionalization

The K-means classification of parameter sensitivities and the number of SWPO available for each group best supported eight calibration groups (Table 2). Headwater streams in groups 1–5 exhibited relatively high sensitivity to RF and FT and moderate sensitivity to all other parameters (Figure 5). Streams in group 6 exhibited highly variable sensitivities to FT and PF and low sensitivities to all other parameters. Headwater streams in groups 7–8 showed high sensitivity to all parameters, especially FT.

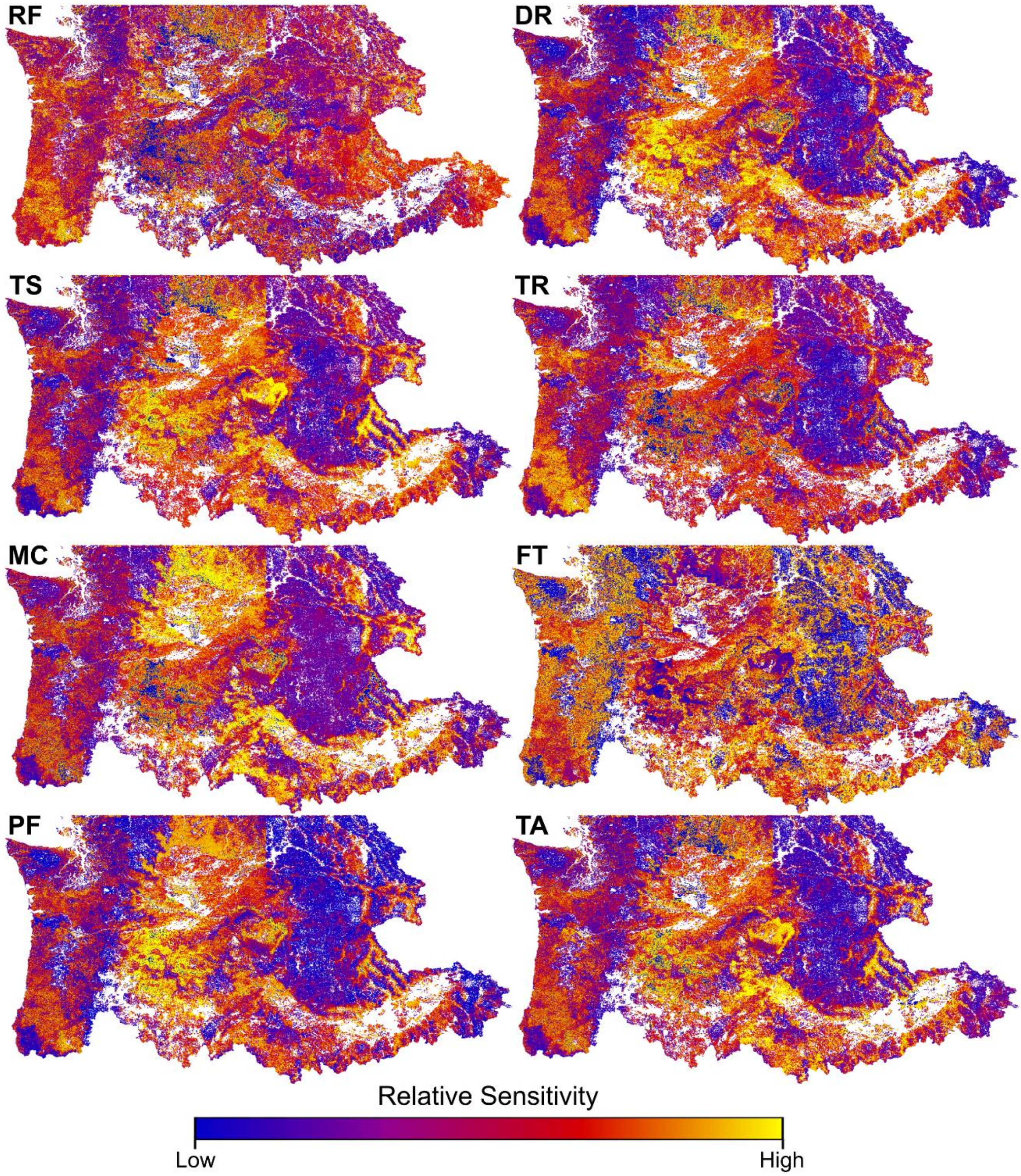

Figure 4. Spatial distribution of the relative sensitivity of stream permanence classification (permanent or non-permanent) to the monthly water balance parameters presented in Table 1.

Table 2. Number of wet and dry surface water presence observations used to calibrate the monthly water balance model for each calibration group, where %HW is the percentage of the 1.3 million headwater stream reaches within the study area assigned to each group, stream length is the total length of streams in each calibration group and drainage area is the total area drained by streams in each calibration group.

Group	%HW	Stream Length (km)	Drainage Area (km^2)	Observations	
				Dry	Wet
1	26.7	13,735	154,176	425	525
2	13.6	2502	11,236	41	119
3	13.2	2255	8087	25	145
4	2.4	69	625	3	15
5	17.1	5043	35,648	98	202
6	13.8	2096	7588	42	68
7	8.2	4332	39,436	102	100
8	5.0	4696	90,012	192	255
Total	-	34,728	346,808	928	1429

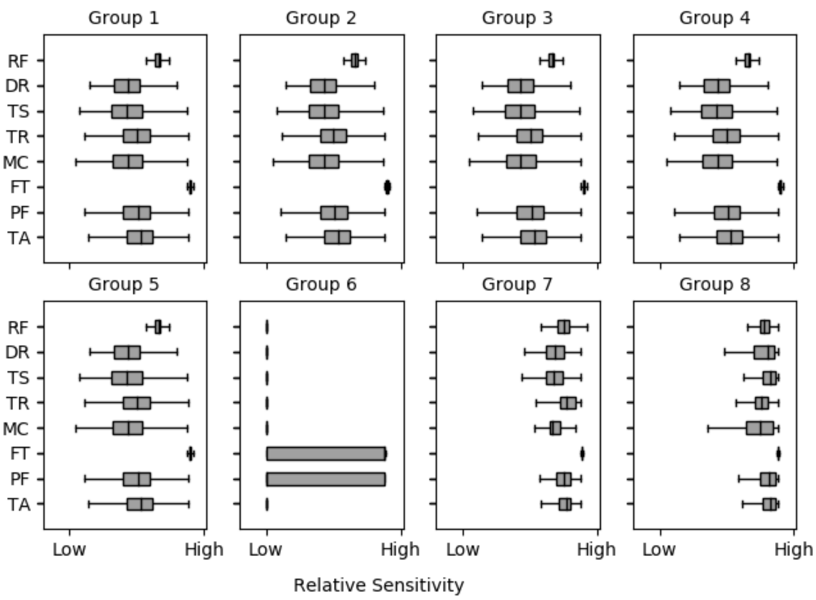

Figure 5. Distribution of relative parameter sensitivities for stream segments within each calibration group (Table 2).

SWPO were not distributed equally among calibration groups (Table 2). The majority of SWPO occurred on stream reaches in calibration groups 1 and 8. Most headwater streams in Idaho fall in these calibration groups and the majority of SWPO were located in Idaho. Only 2.4% of headwaters stream reaches were assigned to calibration group 4, and just 18 SWPO fell in this group.

Calibration groups also exhibited geographical similarities and spatial autocorrelation (Figure 6). Group 8 tended to represent mountainous regions of the interior PNW and some mountainous, coastal regions. Groups 2, 3 and 5 tended to represent coastal mountains. Groups 1, 6, and 7 tended to represent headwater streams in the plains, foothills, and plateaus of the interior PNW. While geographical similarities are apparent within calibration groups, it is also apparent that headwater streams near to each other may exhibit different sensitivities to model parameters and be classified into different calibration groups.

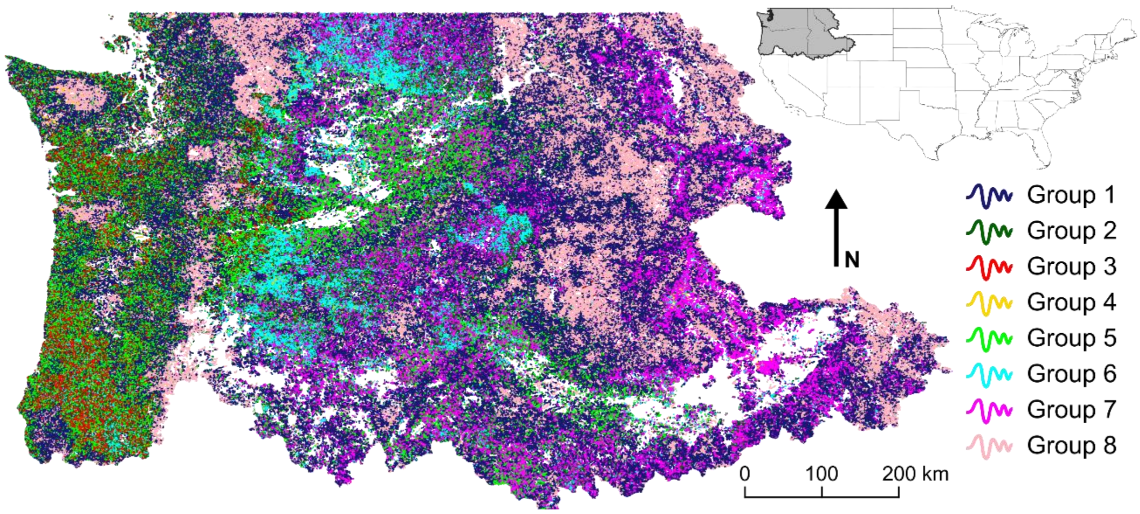

Figure 6. Spatial occurrence of stream reaches in each calibration group as determined by the parameter regionalization methods.

Calibration groups displayed some physiographic differences. Drainage areas for groups 2–6 were quite small (<0.25 km^2), but these groups were separated by differences in elevation and annual precipitation (Figure 7). While there is overlap in catchment area, maximum elevation, and annual precipitation between the eight calibration groups, it is also apparent that the groups are somewhat distinct from each other. For example, group 8 contained the largest catchments and had the greatest average maximum elevation. Group 4 primarily represented small, arid catchments. Groups 3 and 5 were similar in catchment size and elevation but had slightly different precipitation distributions (Figure 7).

3.3. Parameter Set Selection

We identified multiple (n = 13–92), suitable parameter sets for calibration groups 2–6 (Table 3). The overall accuracy of suitable parameter sets ranged from 65 to 90%. The highest accuracy was simulated for group 4, which only contained 18 SWPO. No parameter sets met accuracy constraints for calibration groups 1, 7, and 8 (Table 3). The majority of SWPO occurred in groups 1, 7, and 8 and these groups represent 39.9% of headwater streams in the PNW.

Table 3. Number (n) and accuracy range of parameter sets for each calibration group that met accuracy constraints for annual and monthly simulations. The 'wet' and 'dry' columns present the range of accuracy when MWBM stream permanence classifications are assessed on only wet SWPO or dry SWPO. The median accuracy value is displayed in parentheses.

Group	*n*	Annual Accuracy		
		Dry	Wet	Overall
1	0	-	-	-
2	32	0.61–0.78 (0.63)	0.61–0.72 (0.67)	0.65–0.70 (0.66)
3	13	0.60–0.76 (0.64)	0.64–0.74 (0.66)	0.65–0.73 (0.66)
4	92	0.67–1.00 (0.67)	0.67–0.93 (0.80)	0.67–0.89 (0.78)
5	41	0.60–0.74 (0.62)	0.61–0.75 (0.68)	0.65–0.71 (0.66)
6	19	0.67–0.79 (0.71)	0.60–0.68 (0.63)	0.65–0.69 (0.66)
7	0	-	-	-
8	0	-	-	-

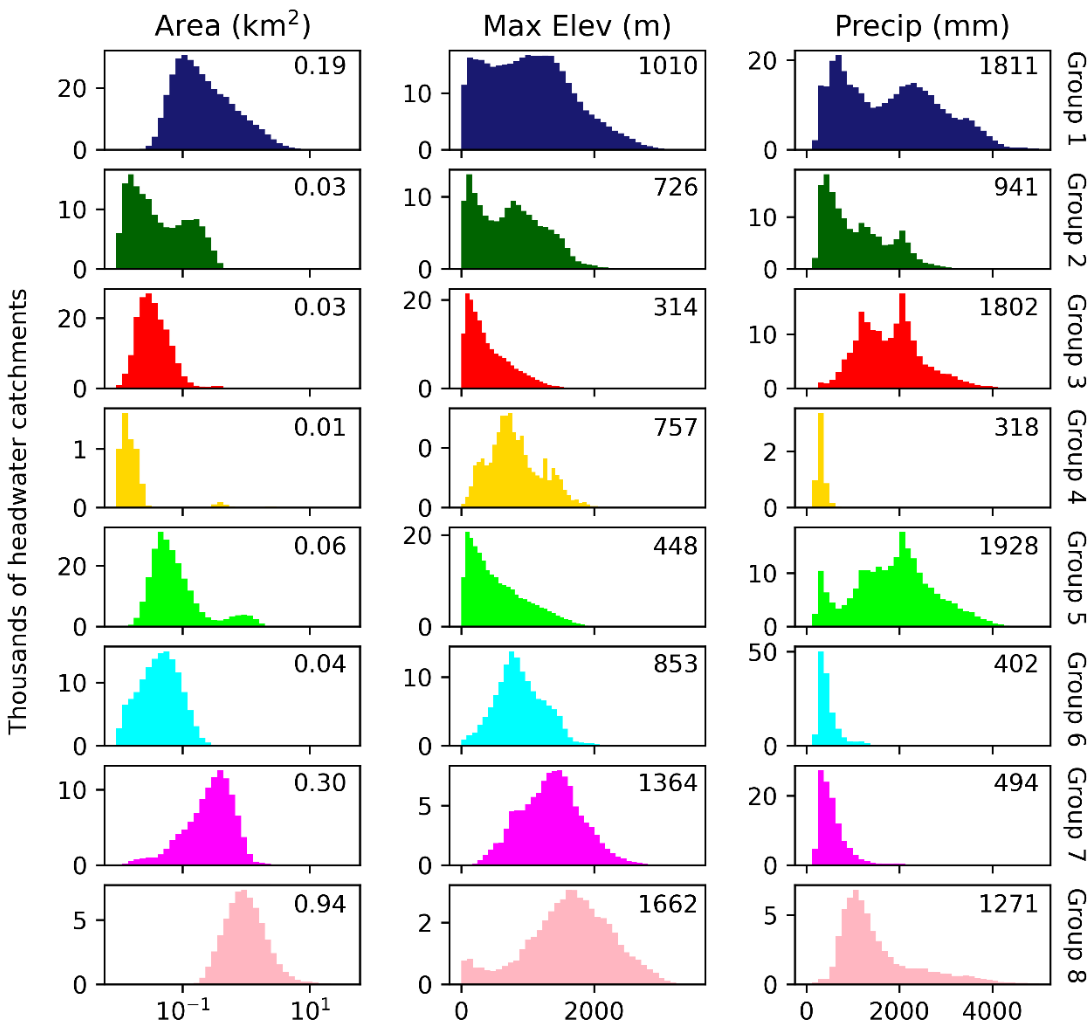

Figure 7. Distributions of drainage area (log scale), maximum elevation, and total annual precipitation for headwater catchments in each calibration group (Table 2). Median values are reported in the top-right corner of each histogram.

Suitable parameter values for groups 2–6 occurred across nearly the entire distribution for each parameter, with some distinctions among calibration groups (Figure 8). The exception was FT, which tended to values near zero. RF values tended to be less than 0.5, except for group 5 where a more uniform distribution was observed between 0.3 and 1.0. TS for groups 2 and 6 tended towards zero, while the distribution was relatively uniform for the other calibration groups. Overall, the relatively uniform distributions of suitable parameters appear to indicate a high degree of equifinality in model parameterization (Figure 8).

3.4. Model Precision

Model precision for permanent and non-permanent streams varied with climate conditions. In general, more streams showed higher model precision for permanent classification in wet years (e.g., 1997). The opposite occurred in dry years, where more streams showed higher model precision for non-permanent classifications (e.g., 1987; Figure 9). Through time, model precision was highest and most stable for group 6 (Figure 10). Group 3 exhibited a high amount of variability year to year. Precision was higher for permanent streams in groups 2–5 than for non-permanent streams. Group 6 showed slightly higher precision for non-permanent streams but was more balanced overall.

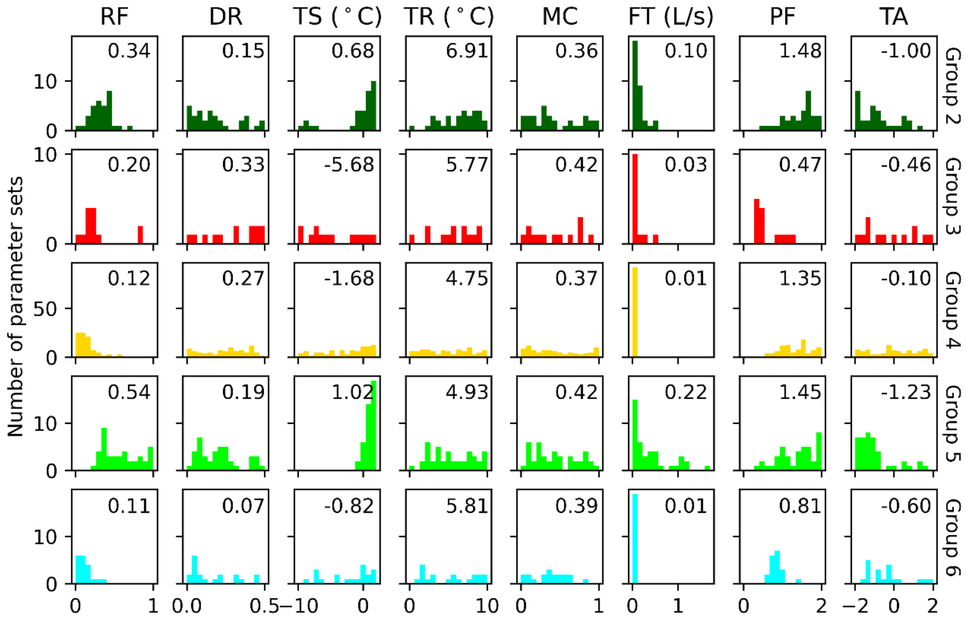

Figure 8. Distributions of suitable parameter values for parameter sets that met benchmark constraints. The x-axis represents the sampled range of parameter values (Table 1), except for FT which is subset to 0.0–2.0 L/s to represent the full range of suitable values determined during calibration. Median values are displayed in the top-right corner of each histogram.

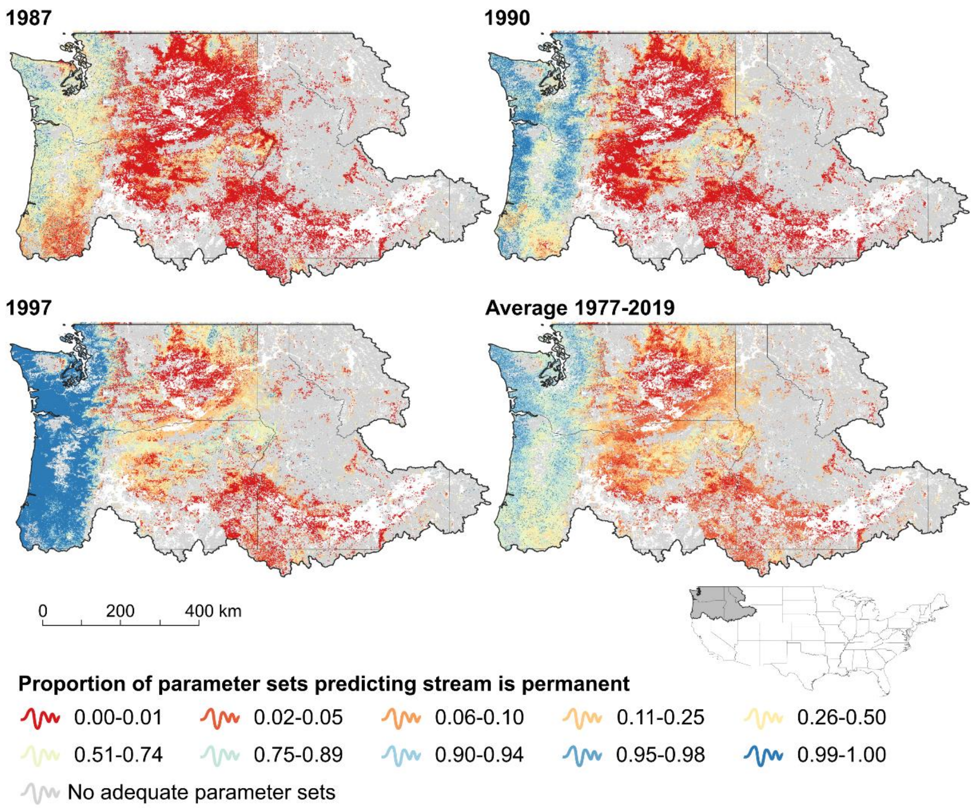

Figure 9. Model precision estimates for headwater streams in the Pacific Northwest during a drier than normal year (1987), approximately normal year (1990), wetter than normal year (1997), and the averaged precision estimates for the entire modeled period (1977–2019). Darker blue represents greater precision for a permanent classification and darker red represents greater precision for a non-permanent classification. Yellow indicates stream reaches with poor precision for the permanence classification.

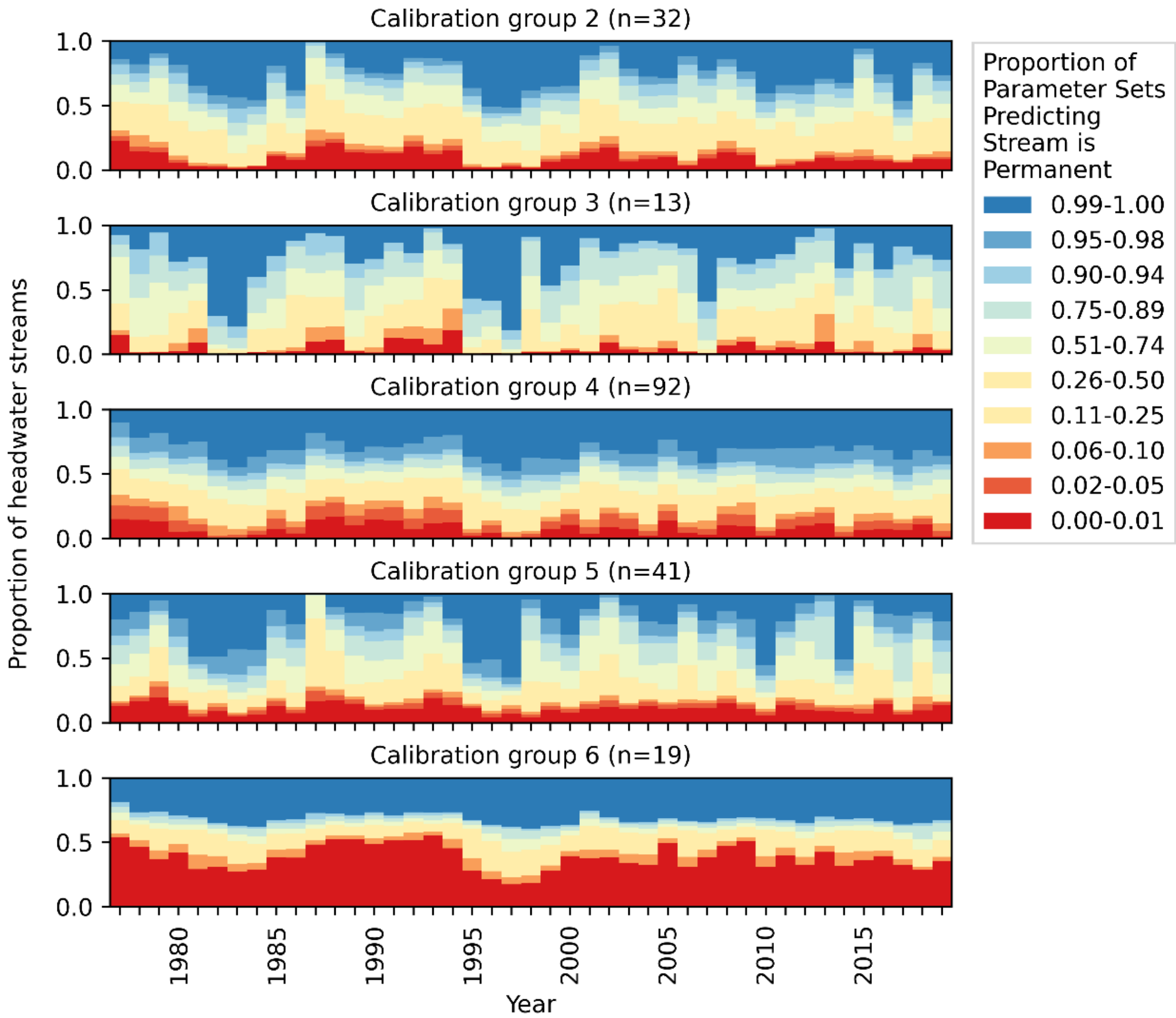

Figure 10. Estimated model precision for each calibration during each year of the study period where n is number of suitable parameter sets tested for each calibration group. Darker blue represents greater precision for a permanent classification and darker red represents greater precision for a non-permanent classification. Yellow indicates stream reaches with a poor precision for the permanence classification.

For the entire PNW region, model precision was generally greater for permanent classifications than non-permanent classifications. However, precision for non-permanent streams was less variable across years (Figures 10 and 11). Overall, the percentage of headwater streams in the PNW where model precision exceeded 90% ranged from 28 to 45%. Forty percent of headwater streams in the PNW did not have suitable parameterization so, in a given year, 15–32% of modeled stream reaches had questionable (<90% and >10%) precision. These results only describe how well different model parameterizations agreed with each other and not the accuracy of each parameterization.

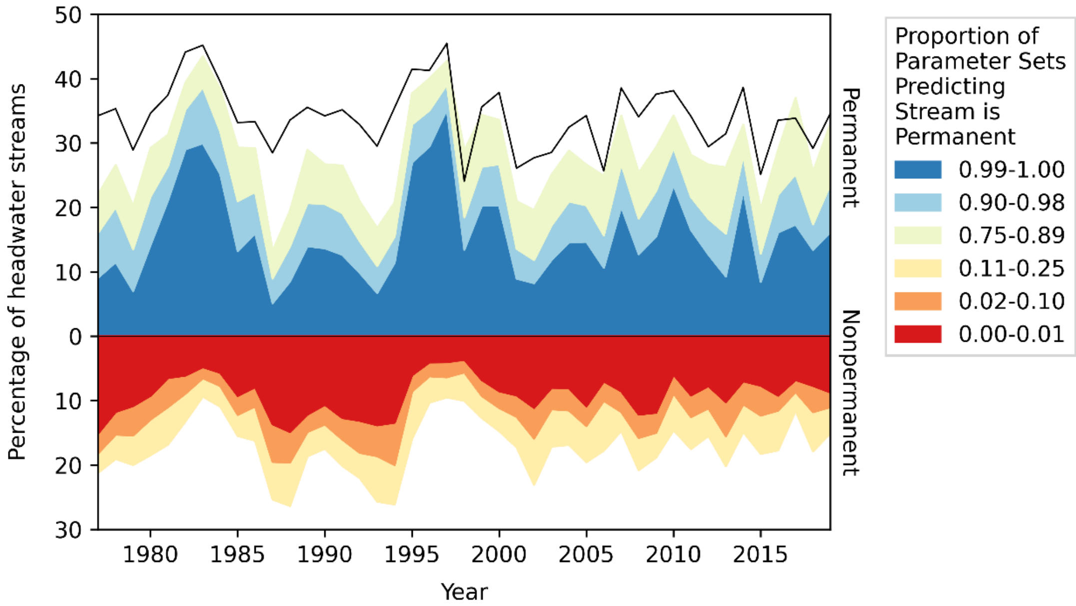

Figure 11. Estimated model precision summed for the entire study area (PNW) for each year of this study. The black line is total percentage of headwater streams with greater than 90% model precision. That is, the sum of the two darkest red and two darkest blue areas on the plot. Darker blue represents greater precision for a permanent classification and darker red represents greater precision for a non-permanent classification.

4. Discussion

While the MWBM produced precise results for only 40% of headwater streams in the PNW, it is important to consider these results in context. Ungauged headwater streams can be difficult to model even when abundant data are present to characterize the catchment [54]. The fact that the MWBM generated precise results for 40% of headwater catchments indicates the potential for development of regional stream permanence models. While the simple MWBM has its shortcomings for this application, results indicate potential for future development of simple, stream permanence models. The shortcomings identified for this particular use of the MWBM highlight important considerations and processes to include in future modeling efforts. Development of predictive stream permanence models is extremely important for assessment of how changes to land cover and climate may influence stream regulation and stream ecology in the United States and worldwide.

The primary objective of this study was to evaluate if the USGS Thornthwaite Monthly Water Balance Model (MWBM) could generate annual dynamic stream permanence classifications (SPC, e.g., perennial, non-perennial) estimates with similar accuracy to static NHD SPC. Previous studies that calibrated the MWBM to discharge data from stream gauges produced good results for large streams over regional and national extents [12,13]. These previous results indicated the MWBM has potential for modeling stream permanence over similar spatial extents for headwater streams but needed observational data to validate this possibility. By adding a flow threshold parameter to the MWBM, we were able to generate dynamic stream permanence estimates for all headwater streams in the PNW. Our methods and results highlight important considerations for future collection of SWPO, stream permanence classification with the MWBM, and stream permanence modeling in general.

Better accuracies for calibration groups that represented smaller ranges of catchment area and annual precipitation may indicate that increasing the number calibration groups would result in better parameterization. However, incorporation of additional calibration groups in this study was limited by the total number of SWPO. Accuracy and precision results were best for calibration groups 4 and 6, which, generally, represented small, arid catchments. Calibration groups 2, 3, and 5, which represented small catchments across a wide range of annual precipitation, had good accuracy results, but more variable precision estimates (Figure 10).

The spatial scale of PRISM data (~4 km) may not account for important heterogeneity in climate inputs at the scale of some headwater catchments. As shown in Figure 3, one PRISM grid cell often represented temperature and precipitation for multiple headwater catchments and small catchments were often represented by a single PRISM grid cell. Because topography is not an input to the MWBM, catchments are represented by climatic conditions and catchment area. Other studies have also indicated that PRISM data do not always capture climatic heterogeneity in mountainous regions [55,56]. Thus, differences in catchment topography and morphology could be important factors contributing to stream permanence, and incorporation of downscaled climate data or other remotely sensed data (e.g., geomorphons from lidar topography, soil moisture, and land cover) may aid in stream permanence classifications in headwater catchments.

Calibration groups were determined by how catchments responded to changes in model parameters (i.e., parameter sensitivity). Given these results, that demonstrate physiographic similarities between calibration groups, it would be useful to test a regionalization method based on physiographic characteristics of catchments—for example, grouping catchments based on a suite of factors that represent geology, temperature, precipitation regime, and land cover. We did not take this approach because, while a thorough dataset of physiographic variables has been linked to NHDPlus (medium resolution) catchments [57], a similar dataset does not yet exist for NHDPlus HR catchments, which better represent headwater streams. This could be an area of important future research that would aid in regional and continental hydrologic studies. Additionally, the number of calibration groups we considered was limited by the number of SWPO available for accuracy assessment in each calibration group. Increasing the number of SWPO would make it possible to consider additional calibration groups while still providing suitable observations for calibration and validation.

Sparse SWPO may have also prevented adequate MWBM calibration for some calibration groups. For example, only 18 SWPO occurred in calibration group 4, which could lead to model overfitting. Because SWPO use to calibrate and assess process-based models is relatively new, the effects of the sparser truth data on model performance are not well understood. This is a reason why we evaluated model performance based on precision across multiple parameter sets. While SWPO represented many more headwater locations than stream gauges, there were no repeat observations at SWPO locations. This results in a space-for-time substitution, which assumes SWPO made in different locations account for the range of conditions stream reaches experience through time and assumes that catchment physiography and climate variability are well represented. As noted above, calibration groups representing larger ranges of catchment area and total annual precipitation had poorer performance than those with narrower ranges, indicating catchment physiography and climate variability were not well represented by SWPO in each calibration group.

Furthermore, SWPO were collected by multiple agencies for multiple uses using different methods [7,33]. SWPO were often obtained as ancillary information to other objectives and, thus, studies were not designed to create a robust dataset for modeling purposes [23]. Because few SWPO were available, we did not conduct a cross-validation accuracy assessment for each calibration group, but instead used multiple suitable parameter sets to assess model precision. A larger SWPO dataset would support and encourage more robust accuracy assessment. Indeed, other stream permanence modeling efforts, conducted over smaller spatial extents, were able to achieve high accuracies with fuller underlying datasets [24,27–29]. This points to the importance of increased SWPO collection to support future stream permanence modeling efforts. For regional studies, it is also important to give attention to SWPO spatial distribution. Few SWPO for some calibration groups can present challenges with model overfitting.

Some hydrological processes and topographic features that influence surface water presence are not explicitly represented by the MWBM. Specifically, groundwater discharge is often a major contribution to summer streamflow, especially in mountainous regions [58]. However, in the MWBM, baseflow is lumped into the RF parameter, which also includes runoff contributions from lateral flow and saturation excess flow. Timing of water delivery to the stream channel is different for each of these mechanisms. With a monthly time step it may be reasonable to lump lateral flow and saturation excess flow, but the MWBM may benefit from explicit representation of baseflow for estimating stream permanence. Comparisons between the MWBM and other hydrologic models and observations in headwater streams could provide more information about the suitability of the MWBM runoff mechanisms to represent hydrology in different areas and climates. Additionally, there is no representation of valley-bottom geology in the MWBM, which controls water transport in the subsurface [59–62]. Thus, valley-bottom representation could be an important inclusion in future stream permanence models.

Local factors, such as the hydraulic conductivity of the streambed, depth to bedrock, and channel morphometry, that control surface and groundwater exchange at the reach and sub-reach scale are important for identifying the transition between surface water presence and absence [62]. These factors become increasingly important at low flows. However, reach and sub-reach-scale geologic and geomorphic data are sparse. Creation of such datasets may be important to more accurate modeling of stream permanence.

5. Conclusions

The sheer number of headwater streams in the US and the dynamic nature of streamflow precludes collection of SWPO on each individual stream. Thus, dynamic estimates of stream permanence in the headwaters will require modeling approaches to assess locations where data cannot be collected. We modified the MWBM by adding a simple flow threshold model to estimate stream permanence on over 1.3 million headwater streams in the PNW, then benchmarked MWBM stream permanence estimates against observed accuracies of NHD SPC to assess precision of the MWBM estimates.

For three of eight calibration groups—approximately 40% of headwater streams—no parameter combinations produced stream permanence results at least as accurate as NHD SPC. Precision of MWBM estimates for suitable parameter sets was poor for an additional 20% of headwater streams. Thus, the MWBM may be suitable to estimate stream permanence for approximately 40% of headwater streams in the PNW, given the data and methods used in this study. MWBM precision was best for calibration groups with more narrow distributions of catchment area and total annual precipitation.

More SWPO that are intentionally collected for modeling studies could greatly improve model development and calibration. The number of SWPO available for this study was much less than the number of modeled stream reaches. Thus, the characteristics of some stream reaches may not have been represented by the observational data used to calibrate the MWBM and assess its accuracy. Additionally, the MWBM does not explicitly represent baseflow and valley-bottom processes which can be important for stream permanence determinations. Future modeling efforts could consider the effect of these processes.

Use of a simple MWBM with primary inputs of monthly precipitation and temperature produced precise results with comparable accuracy (65%) to the NHD standard for 40% of headwater streams in the PNW. While these results indicate that the MWBM is not suitable to consistently estimate stream permanence for headwater streams, they do indicate potential for predictive models to produce reliable stream permanence estimates. Lessons learned from this modeling application will aid in development of future stream permanence models over large spatial extents.

Author Contributions: Conceptualization, K.C.H.; data curation, K.C.H. and R.S.; formal analysis, K.C.H.; funding acquisition, K.W.B. and A.R.; investigation, K.C.H.; methodology, K.C.H.; project administration, K.W.B. and P.E.G.; resources, A.R.; supervision, K.W.B. and P.E.G.; writing—original draft, K.C.H.; writing—review and editing, K.C.H., K.W.B., P.E.G., R.S. and A.R. All authors have read and agreed to the published version of the manuscript.

Funding: Research was funded by the Department of the Interior U.S. Geological Survey National Geospatial Program and the National Science Foundation Integrative Graduate Education and Research Training (IGERT) Program grant number 1249400.

Data Availability Statement: Underlying datasets used for this study are publicly available as stated in the main text. Products specific to this study can be accessed from [45].

Acknowledgments: We thank Erin Brooks and Tim Link for their advice regarding model development and application. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fritz, K.M.; Johnson, B.R.; Walters, D.M. Physical indicators of hydrologic permanence in forested headwater streams. *J. North Am. Benthol. Soc.* **2008**, *27*, 690–704. [CrossRef]
2. Nadeau, T.-L.; Rains, C.M. Hydrological connectivity between headwater streams and downstream waters: How science can inform policy. *J. Am. Water Resour. Assoc.* **2007**, *43*, 118–133. [CrossRef]
3. Vorste, R.V.; Sarremejane, R.; Datry, T. Intermittent rivers and ephemeral streams: A unique biome with important contributions to biodiversity and ecosystem services. In *Encyclopedia of the World's Biomes*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 419–429. [CrossRef]
4. U.S. Geological Survey. National Hydrography Dataset (Ver. USGS National Hydrography Dataset Best Resolution for Hydrologic Unit 17). 2019. Available online: https://www.usgs.gov/core-science-systems/ngp/national-hydrography/access-national-hydrography-products. (accessed on 1 November 2019).
5. Fritz, K.M.; Hagenbuch, E.; D'Amico, E.; Reif, M.; Wigington, P.J.; Leibowitz, S.G.; Comeleo, R.L.; Ebersole, J.L.; Nadeau, T.-L. Comparing the extent and permanence of headwater streams from two field surveys to values from hydrographic databases and maps. *J. Am. Water Resour. Assoc.* **2013**, *49*, 867–882. [CrossRef]
6. Nadeau, T.L.; Leibowitz, S.G.; Wigington, P.J.; Ebersole, J.L.; Fritz, K.M.; Coulombe, R.A.; Comeleo, R.L.; Blocksom, K.A. Validation of rapid assessment methods to determine streamflow duration classes in the Pacific Northwest, USA. *Environ. Manag.* **2015**, *56*, 34–53. [CrossRef] [PubMed]
7. Hafen, K.C.; Blasch, K.W.; Rea, A.; Sando, R.; Gessler, P.E. The influence of climate variability on the accuracy of NHD perennial and nonperennial stream classifications. *J. Am. Water Resour. Assoc.* **2020**, *56*, 903–916. [CrossRef]
8. Beaman, W. *Topographic instructions of the United States Geological Survey. Topographic Mapping*; U.S. Government Printing Office: Washington, DC, USA, 1928. [CrossRef]
9. Guptill, S.C. An enhanced digital line graph design. U.S. Geological Survey: Denver, CO, USA, 1990. [CrossRef]
10. McKay, L.; Bondelid, T.; Dewald, T.; Johnston, J.; Moore, R.; Rea, A. *NHDPlus Version 2: User Guide*; U.S. Environmental Protection Agency: Washington, DC, USA, 2012.
11. Miller, M.P.; Carlisle, D.M.; Wolock, D.M.; Wieczorek, M. A database of natural monthly streamflow estimates from 1950 to 2015 for the conterminous United States. *J. Am. Water Resour. Assoc.* **2018**, *54*, 1258–1269. [CrossRef]
12. McCabe, G.J.; Markstrom, S.L. A monthly water-balance model driven by a graphical user interface. In *Open-File Report*; U.S. Geological Survey: Reston, VA, USA, 2007.
13. Bock, A.R.; Hay, L.E.; McCabe, G.J.; Markstrom, S.L.; Atkinson, R.D. Parameter regionalization of a monthly water balance model for the conterminous United States. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 2861–2876. [CrossRef]
14. Bock, A.R.; Farmer, W.H.; Hay, L.E. Quantifying uncertainty in simulated streamflow and runoff from a continental-scale monthly water balance model. *Adv. Water Resour.* **2018**, *122*, 166–175. [CrossRef]
15. McCabe, G.J.; Wolock, D.M. Warming may create substantial water supply shortages in the Colorado River basin. *Geophys. Res. Lett.* **2007**, *34*. [CrossRef]
16. McCabe, G.J.; Wolock, D.M. Joint variability of global runoff and global sea surface temperatures. *J. Hydrometeorol.* **2008**, *9*, 816–824. [CrossRef]
17. McCabe, G.J.; Wolock, D.M. Independent effects of temperature and precipitation on modeled runoff in the conterminous United States. *Water Resour. Res.* **2011**, *47*. [CrossRef]
18. Deweber, J.T.; Tsang, Y.; Krueger, D.M.; Whittier, J.; Wagner, T.; Infante, D.M.; Whelan, G. Importance of understanding landscape biases in USGS gage locations: Implications and solutions for managers. *Fisheries* **2014**, *39*, 155–163. [CrossRef]
19. Yu, S.; Bond, N.R.; Bunn, S.; Xu, Z.; Kennard, M. Quantifying spatial and temporal patterns of flow intermittency using spatially contiguous runoff data. *J. Hydrol.* **2018**, *559*, 861–872. [CrossRef]
20. Yu, S.; Do, H.X.; van Dijk, A.I.J.M.; Bond, N.R.; Lin, P.; Kennard, M.J. Evaluating a landscape-scale daily water balance model to support spatially continuous representation of flow intermittency throughout stream networks. *Hydrol. Earth Syst. Sci.* **2020**, *24*, 5279–5295. [CrossRef]
21. Bock, A.R.; Hay, L.E.; Markstrom, S.L.; Emmerich, C.; Talbert, M. The U.S. Geological Survey Monthly Water Balance Model Futures Portal. In *Open-File Report*; U.S. Geological Survey: Reston, VA, USA, 2017. [CrossRef]

22. Jaeger, K.; Sando, R.; McShane, R.; Dunham, J.; Hockman-Wert, D.; Kaiser, K.; Hafen, K.; Risley, J.; Blasch, K. Probability of Streamflow Permanence Model (PROSPER): A spatially continuous model of annual streamflow permanence throughout the Pacific Northwest. *J. Hydrol. X* **2019**, *2*, 100005. [CrossRef]
23. Jaeger, K.; Hafen, K.; Dunham, J.; Fritz, K.; Kampf, S.; Barnhart, T.; Kaiser, K.; Sando, R.; Johnson, S.; McShane, R.; et al. Beyond streamflow: Call for a national data repository of streamflow presence for ptreams and rivers in the United States. *Water* **2021**, *13*, 1627. [CrossRef]
24. Williamson, T.N.; Agouridis, C.T.; Barton, C.D.; Villines, J.A.; Lant, J.G. Classification of Ephemeral, Intermittent, and Perennial Stream Reaches Using a TOPMODEL-Based Approach. *J. Am. Water Resour. Assoc.* **2015**, *51*, 1739–1759. [CrossRef]
25. Hester, G.; Carsell, K.; Ford, D. Benefits of USGS Streamgaging Program—Users and Uses of USGS Streamflow Data. 2006. Available online: https://water.usgs.gov/osw/pubs/nhwc_report.pdf (accessed on 10 March 2022).
26. Lowe, W.H.; Likens, G.E. Moving Headwater streams to the head of the class. *BioScience.* **2005**, *55*, 196. [CrossRef]
27. Gendaszek, A.S.; Dunham, J.B.; Torgersen, C.E.; Hockman-Wert, D.P.; Heck, M.P.; Thorson, J.; Mintz, J.; Allai, T. Land-cover and climatic controls on water temperature, flow permanence, and fragmentation of great basin stream networks. *Water* **2020**, *12*, 1962. [CrossRef]
28. Jensen, C.K.; McGuire, K.J.; Prince, P.S. Headwater stream length dynamics across four pysiographic provinces of the Appalachian highlands. *Hydrol. Process.* **2017**, *31*, 3350–3363. [CrossRef]
29. Ward, A.S.; Schmadel, N.; Wondzell, S.M. Simulation of dynamic expansion, contraction, and connectivity in a mountain stream network. *Adv. Water Resour.* **2018**, *114*, 64–82. [CrossRef]
30. Beaufort, A.; Carreau, J.; Sauquet, E. A classification approach to reconstruct local daily drying dynamics at headwater streams. *Hydrol. Process.* **2019**, *33*, 1896–1912. [CrossRef]
31. Ward, A.S.; Wondzell, S.M.; Schmadel, N.M.; Herzog, S.P. Climate change causes river network contraction and disconnection in the H.J. Andrews Experimental Forest, Oregon, USA. *Front. Water* **2020**, *2*, 7. [CrossRef]
32. Seaber, P.R.; Kapinos, F.P.; Knapp, G.L. Hydrologic Unit Maps. 1987. Available online: https://pubs.er.usgs.gov/publication/wsp2294 (accessed on 10 March 2022).
33. McShane, R.; Sando, R. Streamflow Observation Points in the Pacific Northwest, 1977–2016. Available online: https://data.usgs.gov/datacatalog/data/USGS:5a0f338de4b09af898d099b9 (accessed on 10 March 2022).
34. Jaeger, K.L.; Burnett, J.; Heaston, E.D.; Wondzell, S.M.; Chelgren, N.; Dunham, J.B.; Johnson, S.; Brown, M. FLOwPER User Guide—For Collection of Flow Permanence Field Observations. Available online: https://pubs.er.usgs.gov/publication/ofr20201075 (accessed on 10 March 2022).
35. Legates, D.R.; Mather, J.R. An evaluation of the average annual global water balance. *Geogr. Rev.* **1992**, *82*, 253. [CrossRef]
36. Legates, D.R.; McCabe, G.J., Jr. A Re-Evaluation of the average annual global water balance. *Phys. Geogr.* **2005**, *26*, 467–479. [CrossRef]
37. Thornthwaite, C.W.; Mather, J.R. The Water Balance. *Publ. Climatol.* **1955**, *8*, 1–104.
38. Thornthwaite, C.W.; Mather, J.R. Instructions and tables for computing potential evapotranspiration and the water balance. *Publ. Climatol.* **1957**, *10*, 185–311.
39. PRISM Climate Group, Oregon State University. 2020. Available online: http://prism.oregonstate.edu/ (accessed on 10 March 2022).
40. Schwarz, G.E.; Alexander, R. State Soil Geographic (STATSGO) Data Base for the Conterminous United States. 1995. Available online: https://pubs.er.usgs.gov/publication/ofr95449 (accessed on 10 March 2022).
41. Hamon, W. Estimating Potential Evaporation. *P. I. Civil Eng.-Civ. En.* **1961**, *87*, 107–120.
42. Moore, R.B.; McKay, L.D.; Rea, A.H.; Bondelid, T.R.; Price, C.V.; Dewald, T.G.; Johnston, C.M. User's Guide for the National Hydrography Dataset Plus (NHDPlus) High Resolution. Available online: https://pubs.usgs.gov/of/2019/1096/ofr20191096.pdf (accessed on 10 March 2022).
43. Python Software Foundation. Python Language Reference, Version 3.7. 2019. Available online: https://docs.python.org/3.7/reference/ (accessed on 10 March 2022).
44. USGS Advanced Research Computing. USGS Yeti Supercomputer. U.S. Geological Survey. 2021. Available online: https://www.usgs.gov/advanced-research-computing/usgs-yeti-supercomputer (accessed on 10 March 2022).
45. Hafen, K.C. *Sensitivity and Precision of Stream Permanence Estimates (1977–2019) from the USGS Thornthwaite Monthly Water Balance Model in the Pacific Northwest*; U.S. Geological Survey: Reston, VA, USA, 2021. [CrossRef]
46. Falcone, J.A. GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow. In *GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow*; U.S. Geological Survey: Reston, VA, USA, 2011. [CrossRef]
47. Cukier, R.; Fortuin, C.M.; E Shuler, K.; Petschek, A.G.; Schaibly, J.H. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *J. Chem. Phys.* **1973**, *59*, 3873–3878. [CrossRef]
48. Saltelli, A.; Tarantola, S.; Chan, K.S. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics* **1991**, *41*, 39–56. [CrossRef]
49. Herman, J.; Usher, W. SALib: An open-source Python library for sensitivity analysis. *J. Open Source Softw.* **2017**, *2*. [CrossRef]
50. Beck, H.E.; van Dijk, A.I.J.M.; de Roo, A.; Miralles, D.G.; McVicar, T.R.; Schellekens, J.; Bruijnzeel, L.A. Global-scale regionalization of hydrologic model parameters. *Water Resour. Res.* **2016**, *52*, 3599–3622. [CrossRef]

51. Patil, S.; Stieglitz, M. Comparing spatial and temporal transferability of hydrological model parameters. *J. Hydrol.* **2015**, *525*, 409–417. [CrossRef]
52. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *2*, 2825–2830.
53. Houska, T.; Kraft, P.; Chamorro-Chavez, A.; Breuer, L. SPOTting Model Parameters Using a Ready-Made Python Package. *PLoS ONE* **2015**, *10*, e0145180. [CrossRef] [PubMed]
54. Wagener, T.; Wheater, H.; Gupta, H.V. *Rainfall-Runoff Modelling in Gauged and Ungauged Catchments*; Imperial College Press: London, UK, 2004.
55. Strachan, S.; Daly, C. Testing the daily PRISM air temperature model on semiarid mountain slopes. *J. Geophys. Res. Atmos.* **2017**, *122*, 5697–5715. [CrossRef]
56. Blasch, K.W.; Hoffmann, J.P.; Graser, L.F.; Bryson, J.R.; Flint, A.L. *Hydrogeology of the Upper and Middle Verde River Watersheds, Central Arizona*; U.S. Geological Survey: Reston, VA, USA, 2006; pp. 1–114. [CrossRef]
57. Hill, R.A.; Weber, M.; Leibowitz, S.G.; Olsen, A.R.; Thornbrugh, D.J. The Stream-Catchment (StreamCat) Dataset: A database of watershed metrics for the conterminous United States. *JAWRA J. Am. Water Resour. Assoc.* **2015**, *52*, 120–128. [CrossRef]
58. Winter, T.C. The role of ground water in generating streamflow in headwater areas and in maintaining base flow. *JAWRA J. Am. Water Resour. Assoc.* **2007**, *43*, 15–25. [CrossRef]
59. Kasahara, T.; Wondzell, S.M. Geomorphic controls on hyporheic exchange flow in mountain streams. *Water Resour. Res.* **2003**, *39*, SBH 3–1. [CrossRef]
60. Packman, A.I.; Salehin, M. Relative roles of stream flow and sedimentary conditions in controlling hyporheic exchange. *Hydrobiologia* **2003**, *494*, 291–297. [CrossRef]
61. Salehin, M.; Packman, A.I.; Paradis, M. Hyporheic exchange with heterogeneous streambeds: Laboratory experiments and modeling. *Water Resour. Res.* **2004**, *40*. [CrossRef]
62. Ward, A.S.; Schmadel, N.M.; Wondzell, S.M.; Harman, C.; Gooseff, M.N.; Singha, K. Hydrogeomorphic controls on hyporheic and riparian transport in two headwater mountain streams during base flow recession. *Water Resour. Res.* **2016**, *52*, 1479–1497. [CrossRef]