ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research

RESEARCH ARTICLE
10.1029/2022WR034169

Mapping Surface Water Presence and Hyporheic Flow Properties of Headwater Stream Networks With Multispectral Satellite Imagery

D. N. Dralle[1], D. A. Lapides[1,2], D. M. Rempe[3], and W. J. Hahm[2]

[1]Pacific Southwest Research Station, United States Forest Service, Davis, CA, USA, [2]Simon Fraser University, Burnaby, BC, Canada, [3]University of Texas at Austin, Austin, TX, USA

Key Points:

- Wetted channels can be mapped with high accuracy using machine learning and high spatio-temporal resolution multispectral imagery
- Coincident observations of runoff and wetness state enable estimation of hydraulic properties of the hyporheic zone
- The scaling of hyporheic properties with contributing area exhibits punctuated break points in relation to stream network topology

Correspondence to:
D. N. Dralle,
david.dralle@usda.gov

Citation:
Dralle, D. N., Lapides, D. A., Rempe, D. M., & Hahm, W. J. (2023). Mapping surface water presence and hyporheic flow properties of headwater stream networks with multispectral satellite imagery. *Water Resources Research*, *59*, e2022WR034169. https://doi.org/10.1029/2022WR034169

Received 22 NOV 2022
Accepted 7 AUG 2023

Abstract Growth and contraction of headwater stream networks determine habitat extent, and open a window to the hyporheic zone. A fundamental challenge is observation of this process: wetted channel extent is dynamic in space and time, with wetted channel length varying by orders of magnitude over the course of a single storm event in headwater catchments. To date, observational data sets are produced from boots-on-the-ground campaigns, drone imaging, or flow presence sensors, which are often laborious and limited in their spatial and temporal extents. Here, we evaluate satellite imagery as a means to detect wetted channel extent via machine learning methods trained on local surveys of wetted channel extent. Even where channel features are smaller than the imagery's spatial resolution, the presence of surface water may be imprinted upon the spectral signature of an individual pixel. For two catchments in northern California with minimal riparian canopy cover and highly dynamic wetted channel extent, we train a random forest model on RapidEye imagery captured contemporaneously with the existing surveys to predict wetted channel extent (accuracy >91%). The model is used to produce length-discharge (L-Q) relations and to calculate spatially distributed estimates of channel hyporheic flow capacity and exchange. A sharp break in hyporheic flow capacity occurs from main stem channels to lower order tributaries, resulting in a stepped L-Q relationship that cannot be captured by traditionally used power law models. Remotely sensed imagery is a powerful tool for mapping wetted channels at high spatial resolution.

Plain Language Summary Streams contract, expand

Flow exceeds capacity

Detected, from space

1. Introduction

Stream networks expand and contract through time, yielding insight into how hillslope runoff generation interacts with channel hyporheic flow properties to create aquatic ecosystem habitat. Ephemeral and intermittent streams constitute much of Earth's fluvial channel network (Datry et al., 2007; Kampf et al., 2021; Messager et al., 2021), and are increasingly being studied due to their role in a wide range of earth system processes (Fovet et al., 2021), such as carbon transport (e.g., Wondzell & Ward, 2022), water transit times (e.g., Lapides et al., 2022), metapopulation dynamics (Durighetto et al., 2022; Giezendanner et al., 2021), and water-borne disease transmission (e.g., Perez-Saez et al., 2017). Historically, time-consuming walking surveys have provided the observational basis for our understanding of the dynamic extent of wetted stream channels in headwater catchments (e.g., Godsey & Kirchner, 2014; Lovill et al., 2018; Whiting & Godsey, 2016). However, these surveys are limited in their spatiotemporal coverage (Lapides et al., 2021), underscoring the status of headwater stream networks as *aqua incognita* (Bishop et al., 2008). Sparse observations have limited exploration of the physical controls on channel growth and contraction (Moidu et al., 2021). Nevertheless, availability of wetted channel data is increasing, and advances in technology, especially within the realms of unmanned aerial systems (e.g., Micieli et al., 2022) and high resolution satellite imaging (e.g., Vanderhoof & Burt, 2018), will continue to improve monitoring efforts.

What determines whether the surface is wetted along a particular reach? Carlston (1963) articulated a general principle:

> "…as transmissibility decreases, the amount or rate of movement of ground water passing through the system decreases and a proportionately greater percentage of the precipitation is forced to flow directly into the streams in flow over the land surface."

This general idea also forms the basis for the variable source area concept described by Hewlett and Hibbert (1967), who noted that:

> "…when the subsurface flow of water from upslope exceeds the capacity of the soil profile to transmit it, the water will come to the surface and channel length will grow."

This flow-emergence principle is applicable on both hillslopes—where saturation overland flow may be generated (Beven & Kirkby, 1979)—as well as in channels—where the presence of water at the surface depends on whether the up-network delivery of water to a point exceeds the subsurface flow capacity of the hyporheic zone (equal to the product of the local slope and cross-sectional area-average conductivity of the material in the hyporheic zone). The idea appeared again in the context of network-scale wetted channel extent mapping (Godsey & Kirchner, 2014), re-invigorating the study of process controls on stream network dynamics. We use the term hyporheic zone in a broad sense to refer to the saturated subsurface of the river corridor that is capable of conveying (and exchanging) flow down the river network.

An important implication of the flow-emergence principle is that when a reach transitions in time from wet to dry, the flow being conveyed by the channel at that point equals the hyporheic flow capacity (Durighetto et al., 2020; Godsey & Kirchner, 2014; Prancevic & Kirchner, 2019). If runoff generation is uniform in space, then area-normalized discharge (unit runoff, Q (L T^{-1}); see Table 2 for a complete description of variables) at the catchment outlet can be used as an estimate for runoff at any point in the watershed. Thus, instantaneous unit runoff measured at the outlet can be used to approximate hyporheic flow capacity (ρ (L T^{-1})) at points in the network that are transitioning from wet to dry (Durighetto & Botter, 2022). Because flow capacity varies throughout the network (due to local topographic and hyporheic properties), a range of wetted channel extent maps are required to identify the flow thresholds that delineate wet and dry states throughout the watershed. Paired with the flow-emergence principle and an assumption of spatially uniform unit runoff, surface water presence-absence dynamics provide a unique window into the hyporheic zone, which has historically been difficult to characterize (Ward et al., 2018; Wondzell, 2011).

Establishing a record of wetted channel extent across the full range of observed flows remains a challenging task (Jaeger et al., 2021). Recent methodological developments, such as the deployment of flow presence-absence sensors and drone surveys (e.g., Carbonneau et al., 2020; Dugdale et al., 2022; Jensen et al., 2019; Micieli et al., 2022; Zanetti et al., 2022), provide important constraints, but still tend to be limited in space or time (at least relative to other hydrological data, such as flow observations). There are some data-driven approaches that classify the intermittency of a particular reach (e.g., Jaeger et al., 2019; Sando & Blasch, 2015) using climatic and geological correlates, but attempts to remotely observe water presence/absence are limited to large open-water (Wang et al., 2022) bodies or main stem river reaches with widths greater than ∼10–30 m (Li et al., 2020; Qin et al., 2021; Verma et al., 2021; Wang et al., 2022). Headwater stream widths are typically much smaller (Allen et al., 2018), below the resolution of most satellite data products. However, there is evidence that, even if a channel is smaller than the satellite image pixel scale or is partially obscured by vegetation, individual pixels themselves may contain spectral information on sub-canopy conditions or transitions in cover type (e.g., wet to dry, or in an analogous problem of classifying vegetation cover, forested to not forested Carbonneau et al., 2020; Chambers et al., 2009; Eriksson et al., 2006; Linderman et al., 2004; Ling et al., 2020; Xue et al., 2022). These are also referred to as "mixed pixels."

Here, we explore the ability of a random forest machine learning model to identify the presence of wetted channels at the sub-pixel scale with relatively high resolution (5 m pixel) satellite imagery trained on existing wetted channel surveys in small headwater catchments with a highly dynamic stream extent. We use the resulting predictive model to generate ∼ weekly maps of wetted channel extent at 10 m spaced points along the stream. We then identify flow thresholds from the outlet hydrograph that delineate wet and dry states across the geomorphic channel network, thus producing spatially distributed estimates of hyporheic zone flow properties at the sub-reach scale.

10.1029/2022WR034169

2. Methods

2.1. Site Description

The study catchments (Dry Creek, 3.54 km^2, and Hank Creek, 5.59 km^2, the southern and northern catchments shown in Figure 1, respectively) are located within the Eel River watershed in Mendocino County, California. Average annual rainfall is approximately 1,800 mm, mostly delivered during a winter wet season (typically November through April), followed by a warm dry season (May through October) (Dralle et al., 2018). Snow is rare at the sites. The sites are situated within relict deep-seated earthflow terrain of the central belt melange of the Franciscan complex (Blake et al., 1985; Langenheim et al., 2013), a geological assemblage made up of three roughly north-south trending belts (coastal, central, eastern). The melange contains rocks of mixed lithology and size suspended within a clay-like, shale-derived matrix. Weathering profiles in the melange are thin, and a perennial water table can be found at depths typically less than 3 m, even at the end of the dry season (Hahm et al., 2019). In the early winter months, infiltrating rainfall rapidly replenishes root-zone water storage deficits, leading to recharge of groundwater tables that rise to the ground surface, typically within 200 mm of seasonal total rainfall (Dralle et al., 2018). Once water tables intersect the ground surface, saturation overland flow is widespread and channel networks rapidly expand, with flows that can exceed 50 mm per day (Lapides et al., 2022). Rapid flow increases are followed by quick recessions with attendant contraction of wetted channel extent. Runoff responds rapidly to precipitation (typical lag-to-peaks in Dry and Hank Creeks of only 2–3 hr [Lapides et al., 2022]). At runoff rates >10 mm/day, most of the channel network has flowing water and saturation overland flow extends up adjacent hillslopes (Lapides et al., 2022). Thin weathering profiles with small water storage capacity also impact the site's plant community—a relatively sparse oak savanna, comprised mainly of non-native annual grasses and Oregon white oak (*Quercus garryanna*) (Hahm et al., 2017, 2018, 2019).

The geomorphic channel drainage density is 16.9 km/km^2, with an average upslope contributing area at the channel heads of 1,085 m^2 (Lovill et al., 2018). Wetted channel widths at Dry and Hank Creeks vary from zero (at the geomorphic channel heads) to typical winter storm values of 5 m (for Dry) and 8 m (for Hank) at the staff gauge locations near their outlets. The upper portions of the channel network (between contributing areas of 1,085 and 10,000 m^2) have very narrow channels (generally less than 1 m width), and their wetted dynamics are not considered in this study. For drainage areas above 10,000 m^2, a single power law relationship describes well the geomorphic channel slope as a function of drainage area in both the Dry and Hank Creek networks (Lovill et al., 2018). To increase the likelihood that the uniform unit runoff assumption is valid, we analyzed channels that integrate over relatively large contributing areas; in this case, at least an order of magnitude (10×) greater than the characteristic channel initiation threshold area. We performed our analyses at smaller and larger scales and results were not particularly sensitive to this choice. Average hillslope gradients (calculated from 1 m pixels using data from Lovill et al. (2018)) are 28%, and landscape-wide cosmogenic nuclide-inferred erosion rates are between 0.12 and 0.16 mm/year (Hahm et al., 2019).

2.2. Data Sources

2.2.1. Streamflow

Water levels recorded at 15 min intervals as part of the Eel River Critical Zone Observatory monitoring program are converted to streamflow data near the outlet of Dry Creek (near gaging station; Figure 1). Rating curve development is detailed in Hahm et al. (2019), and accuracy is estimated to be between 5% and 10% of the true value of Q. Because lag to peak response timescales in Dry Creek are much less than a day (Lapides et al., 2022), the flow measured at the gage is considered to be representative of flow conditions throughout the watershed at the daily timescale. Field measurements of discharge were also made in Hank Creek between 2015 and 2019 ($n = 24$, data not shown) which established that Dry and Hank Creek have nearly identical instantaneous unit runoff (discharge normalized by catchment area), leading to the use of Dry Creek's runoff (calculated with a more frequently updated rating curve) as a proxy for runoff at Hank Creek. Unless otherwise stated, reported runoff data comes from the Dry Creek gauging station.

2.2.2. Topography

A bare earth digital elevation model (DEM) at 1 m resolution was generated from LIDAR data collected by the National Center for Airborne Laser Mapping in 2015 (https://doi.org/10.5069/G9WH2N2P). Geomorphic

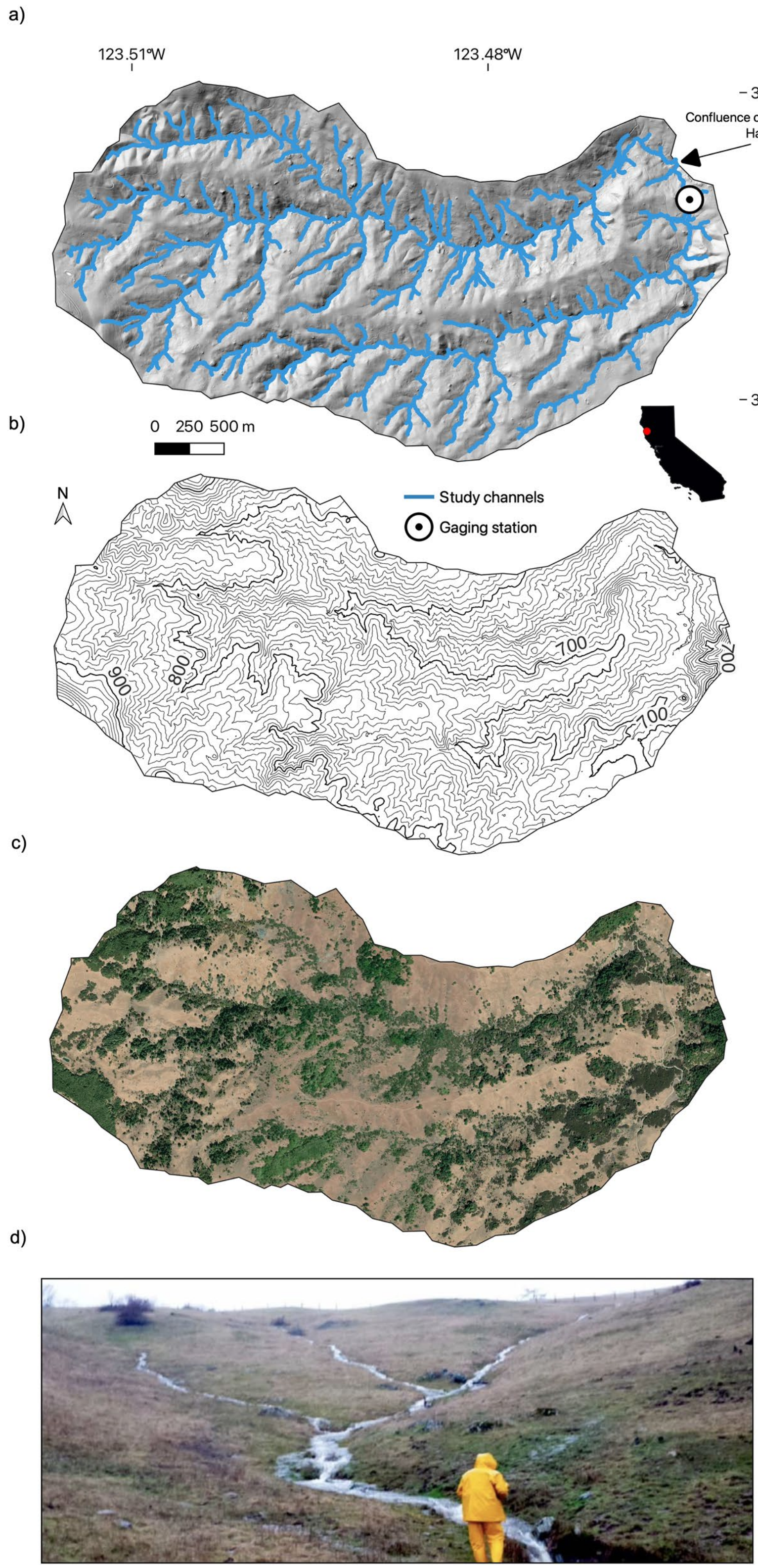

Figure 1. Hillshade of adjacent study catchments with study channels and gauging station location (a) and topographic map with 10 m spaced elevation contours (b, elevations labeled in meters) both derived from a Lidar digital elevation model. Panel (c) shows aerial imagery (bottom, Esri "World Imagery," accessed 27 September 2022). Panel (d) is a photograph taken of a Dry Creek sub-catchment during a high flow event, from Hahm et al. (2019).

channel networks and upstream contributing areas were mapped from this DEM, which also served as the basemap for wetted channel mapping.

2.2.3. Satellite Imagery

We acquired 217 scenes of cloud- and snow-free RapidEye satellite imagery (5 m pixel scale, 8 m orthorectification accuracy) from Planet Labs (Planet Team, 2017). The imagery contains five spectral bands: 440–510 nm (blue), 520–590 nm (green), 630–685 nm (red), 690–730 nm (red edge) and 760–850 nm (near infrared). All images were visually inspected for artifacts and other visible irregularities. Imagery were downsampled to 10 m resolution to account for lower orthorectification accuracy.

2.2.4. Wetted Channel Extents

Data on wetted channel extents is derived from four primary maps (which include two distinct dry season surveys (negligible runoff) and two distinct wet season surveys (high and low runoff) that are mapped in the left column of Figure 4) and multiple inferential maps. These training data are summarized in Table 1 and described in detail as follows:

Walking surveys Two boots-on-the ground walking surveys of the entirety of the Hank and Dry channel networks were performed by Lovill et al. (2018) in 2015. RapidEye imagery paired with the first walking survey was collected approximately 4 days (4 June 2015) after the end of the 5 day long survey (26 May through 31 May 2015). It is possible that some pixels classified as wet during the walking survey may have dried between the end of the walking survey and the imagery date. The second survey occurred 20–24 August 2015, and was ideally paired with imagery during the middle of the survey on 22 August 2015.

Drone survey During a wet-season dry spell in February of 2018, an unmanned aerial vehicle survey on 4 February 2018 (8 days after the most recent rainfall) revealed that channels with contributing area less than 20,000 m^2 were entirely dry. RGB drone imagery (collected using a DJI Mavic Pro) was flown over approximately 1/2 of the total catchment area, and channel condition was assessed manually by video review. Channels are clearly aerially visible throughout the watershed (see Figure 1d), even in locations with riparian vegetation, which does not significantly overhang channels. RapidEye imagery from the nearest following image date (12 February 2018 with intervening rainfall of only about 1 mm) therefore provides "dry" observations of smaller channels (between contributing areas of 10,000 and 20,000 m^2) during the typical wet season months. This is important for training the random forest model, as it helps disentangle spectral signatures that might correlate with large extent of wetted channel (e.g., high greenness from grasses during wet periods) from true spectral indicators of channel wetness.

High flow survey We identified one cloud-free image that coincides with a very high-flow rate of 34 mm/day. Field visits indicate that the channel network considered here can be assumed to be fully wetted at flow rates exceeding 10 mm/day (Lapides et al., 2022). The channel network on the high flow day is therefore assumed to be fully wetted.

Extrapolative/interpolative surveys Machine learning approaches can require significant amounts of data for training and validation. To increase the amount of imagery data available for training, we rely on the "hierarchical principle" described by Botter et al. (2021), which states that, "…temporary stream activation follows a fixed and repeatable sequence, in which the least persistent sections activate only when the most persistent ones are already flowing." The hierarchical principle is implicit in (though not exclusive to) previous frameworks (e.g., Prancevic & Kirchner, 2019), and is a consequence of the two process-based assumptions stated in the Introduction: (a) Surface flow emerges when up-network discharge equals the hyporheic flow capacity at a point in the stream network, and (b) unit runoff is spatially uniform. If these are true, the hierarchical principle follows. These ideas suggest two approaches for extrapolating or interpolating wetted channel surveys to obtain additional training data. The first approach, which we refer to as "interpolative," involves interpolating between the two walking surveys. These surveys capture the summer recession, and we reason that any reaches dry at the beginning of the summer remain dry throughout the whole (rainless) summer. Conversely, any reaches that are still wet at the end of the summer are wet for the duration of time between the surveys. We further extend these data using the following "extrapolative" method. First, we know that there was essentially no rainfall between 15 March 2015 and the first survey date (26–31 May 2015). Furthermore, a nearby (20 km) stream—Elder Creek, which has flow that is strongly correlated to Dry Creek (Hahm et al., 2019)—was monotonically decreasing during this period. We use these two pieces of information to infer that all wet reaches during the first survey were wet for the entire period from 15 March to 26 May. We further noted that there was essentially no rainfall during the months of July–October

10.1029/2022WR034169

Table 1
Dates of Walking Surveys and RapidEye Imagery Collection

Walking survey date	RapidEye imagery date	Scenes	Wetted channel drainage density (km/km^2)	Q (mm/day)	Mapping method	Notes	LOO accuracy (%)
26–31 May 2015	4 June 2015	1	3.4	N/A	Walking	Early dry season (Survey 1)	77
20–24 August 2015	22 August 2015	1	1.0	N/A	Walking	Late dry season (Survey 2)	84
4 February 2018	12 February 2018	1	N/A	5	Drone	Imagery used only for points with accumulated area <20,000 m^2, which are dry. Full drainage density unknown	54
N/A	7 March 2016	1	16.9	44	High-flow	Inferred fully wetted network based on prior wetted surveys and stream discharge	60
N/A	4 June to 22 August 2015	4	N/A	N/A	Survey 2 wet interpolation	Interpolation of wet reaches from Survey 2	8
N/A	4 June to 22 August 2015	4	N/A	N/A	Survey 1 dry interpolation	Interpolation of dry reaches from Survey 1	93
N/A	15 March to 26 May 2015	1	N/A	N/A	Survey 1 wet extrapolation	Extrapolation of wet reaches from Survey 1	66
N/A	All July–October imagery	22	N/A	Multiple	Survey 2 dry extrapolation	Extrapolation of dry reaches from Survey 2	78
N/A	Dates when Elder Creek streamflow ≥0.9 mm/day	29	N/A	Multiple	High-flow extrapolation	Extrapolation of wet reaches from Survey 1	19

Note. LOO Accuracy column reports leave-one-out random forest model accuracy on training data in each row when model is trained on all training data except that described in the row or directly inferred from data in that row.

Table 2
Description of Variables

Variable	Dimensions	Description
L	(–)	Wetted channel drainage density (sum of length of wetted reaches normalized by total channel length)
A	L^2	Drainage area
A_0	L^2	Drainage area at outlet
D	L^3T^{-1}	Discharge at the outlet
D_{exf}	L^3T^{-1}	Exfiltration exchange flux from hyporheic zone to stream integrated across wetted channel network
D_{hef}	L^3T^{-1}	Exchange flux between stream and hyporheic zone integrated across wetted channel network; calculated as $\left(D_{exf} + D_{inf}\right)/2$
D_{inf}	L^3T^{-1}	Infiltration exchange flux from stream to hyporheic zone integrated across wetted channel network
H	L	Average local reach hyporheic zone thickness
$\mathcal{P}$	L^3T^{-1}	Reach hyporheic flow capacity, expressed as volume per time
K	$L\,T^{-1}$	Average local reach flow-parallel hydraulic conductivity
q_{gw}	L^2T^{-1}	Along-reach specific groundwater inflow
q_{hef}	L^2T^{-1}	Along-reach specific exchange flux between stream and hyporheic zone
Q	$L\,T^{-1}$	Upstream-area normalized discharge (i.e., runoff) at the outlet (D/A_0)
S	(–)	Local reach slope
W	L	Average local reach hyporheic zone width
x	L	Along-reach (longitudinal) channel coordinate
α	$\left(T L^{-1}\right)^{\beta}$	Scaling intercept for L-Q relationship
β	(–)	Scaling exponent for L-Q relationship. Fraction by which L changes for a change in Q
ρ	$L\,T^{-1}$	Reach hyporheic flow capacity, expressed as volume per time normalized by upstream area

during the study period, so we inferred that dry reaches during the first walking survey remained dry during all summer months in the study period. Finally, given the correspondence between Elder Creek and Dry Creek, we inferred that any wet reaches during the first walking survey (26–31 May 2015, when Elder Creek discharge was approximately 0.3 mm/day) would also be wet on any date on which Elder Creek runoff was 0.9 mm/day (i.e., a threefold safety factor). At such times, we infer that Dry Creek discharge (which was not measured in early 2015) was larger than during the May walking survey.

2.3. Random Forest Model and Training Data Pre-Processing

We trained a random forest machine learning classification model (Belgiu & Drăguţ, 2016) to identify wetted channel reaches from the satellite imagery, implemented in Python via the Scikit-Learn package (Pedregosa et al., 2011). Random forests are ensembles of decision trees, each of which is classified on a subset of the training data in order to reduce overfitting.

To create the data sets for the random forest modeling, we extracted equally spaced 10 m nodes along the channel network. Survey data from walking surveys Lovill et al. (2018) were extracted from polylines to nodes using a 1.5 m buffer. Labels for the drone survey and high-flow survey were applied directly to these nodes based on drainage area threshold criteria described above. Extrapolated and interpolated data used extracted points from other survey dates. To improve the strength of spectral signal originating from within the channel itself (as opposed to on side slopes), the channel network nodes were clipped to drainage areas greater than 10,000 m^2, approximately 10 times larger than the average drainage area required for channel initiation. For this reason, predicted wetted channel extents are underestimates of the true extent of wetted channel at high flow values. Each node for the relevant portions of the wetted channel maps described above was assigned a 1 (wetted) or 0 (dry) target prediction label. All RapidEye pixel band values were extracted at the location of each node for each

10.1029/2022WR034169

RapidEye scene for use as input features. Predictors used for the random forest model include: blue, green, red, rededge, and near infrared bands from RapidEye pixels and the normalized difference water index ([NDWI], Gao, 1996) calculated from RapidEye pixel values.

We first split all of the available data into randomly selected training (75% of the data) and testing (25% of the data) groups. The random forest was trained on the training data group initially, and predictions were made for the test data group to obtain accuracy metrics. We also tested the importance of each type of training data by performing leave-one-out (LOO) accuracy tests for each row of Table 1 (including any other extrapolated or interpolated data derived from that row using the methods described in Section 2.2.4). The entirety of the wetted channel data set was then used to train a final classifier that was used to predict wetness states for each RapidEye image in the collection. Default Scikit-learn v1.0.1 Random Forest Classifier parameters were used: 100 trees in the forest, Gini impurity to measure the quality of splits, no maximum tree depth, two samples required per split, one minimum sample in each leaf, and the square root of the number of features considered for each split. We performed hyperparameter tuning on tree-depth and found no significant difference in model performance on the test set across a range of values and therefore chose the default (and most straightforward to present) parameterization.

2.4. Power Law Model to Relate Runoff and Wetted Channel Length

With 217 RapidEye scenes, our wetted channel maps cover a large range of conditions, but the distribution of flows on dates for which we have RapidEye scenes is not the same as the full natural distribution (see Figure 9d), due largely to the fact that clouds are more common in the wet season at higher flows. To interpolate across all possible network states, Godsey and Kirchner (2014) demonstrated that a power-law relationship may be appropriate to relate runoff at the outlet to wetted channel length. We fitted a power law curve of the form:

$$L = \alpha Q^{\beta}, \tag{1}$$

where L [·] is the wetted channel length as a fraction of the maximum observed wetted channel length (also plotted in length units in Figure 5b), α is a positive constant, Q (mm/day) is runoff at the outlet, and $0 \leq \beta \leq 1$. We used this relationship with the full distribution of daily streamflow during the study period to infer the full distribution of wetted channel extents. We nevertheless recognize this may be an underestimate of the full extent of wetted channel at high flow values, where surface flow extends below the 10,000 m^2 contributing area threshold used to identify study reaches (Hahm et al., 2019; Lapides et al., 2022). Additionally, Durighetto and Botter (2022) clarify that the power law model may be problematic for other reasons; for example, it is not bounded with respect to Q. The power law model also does not allow for “stepped” behavior in the L-Q relationship. The power law fit should be taken as a rough approximation of the true L-Q relationship, and we include it primarily for the purpose of comparison against extant studies which have commonly reported exponent values.

2.5. Logistic Regression Model to Estimate Hyporheic Flow Capacity

The occurrence of water at the surface is observed when the flow rate at a specific location surpasses the subsurface's ability to transport that flow (Godsey & Kirchner, 2014; Hewlett & Hibbert, 1967). As such, the runoff at the transition in time between a dry and wet state at a particular node is equal to the hyporheic zone's subsurface flow capacity at that location (Durighetto & Botter, 2022; Godsey & Kirchner, 2014; Prancevic & Kirchner, 2019). To model this relationship, logistic regression models were fitted to the random forest predictions of wet versus dry conditions as a function of instantaneous runoff, with each node analyzed separately.

Subsequently, the flow capacity (ρ, expressed in mm/day) at each node was estimated as the value where the logistic function first predicts the reach to be wetted. The hyporheic capacity, expressed as discharge in m^3/day ($\mathcal{P}$), was then calculated by multiplying the ρ value by the drainage area at each node. Due to uneven distribution of predictions between wet and dry conditions at nodes, logistic regression predictors were weighted by the inverse sizes of the dry and wet sample sets, ensuring that each sample had an equal influence on the fit.

Additionally, the logistic regression model can be employed to extend random forest predictions to all dates with runoff observations. By utilizing the estimated flow capacity at each location, it is possible to determine the number of wetted nodes for each daily streamflow value during the study period. Specifically, if the ρ value derived from the logistic regression is lower than the daily runoff, the channel is considered to be wet on that day.

10.1029/2022WR034169

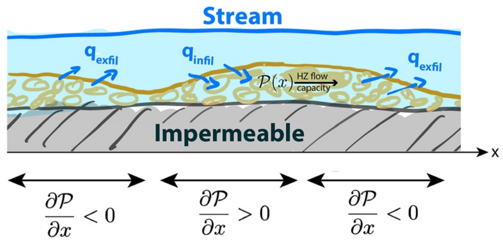

Figure 2. Simplified stream profile illustrating drivers of hyporheic exchange flows. Where volumetric flow capacity ($\mathcal{P}$) in the HZ decreases in the downstream (x) direction $\left(\frac{\partial \mathcal{P}}{\partial x} < 0\right)$, there will be hyporheic exfiltration into the stream. Where flow capacity increases with x, there will be infiltration from the stream into the HZ.

The total wetted channel extent can be obtained by multiplying the sum of wetted nodes by the inter-node length (10 m).

2.6. Hyporheic Exchange Flows

Longitudinal (along-stream) gradients in flow capacity estimated from the logistic regression approach may also be used to constrain hyporheic exchange flows throughout the network. We follow the model for wetted channel expansion/contraction developed by Ward et al. (2018) (see Figure 2), who posit two laterally homogeneous, parallel domains representing the surface stream environment and subsurface hyporheic zone. Surface flow only occurs where runoff exceeds ρ; at these places in the subsurface hyporheic zone, the continuity equation requires:

$$q_{hef} = q_{gw} - \frac{\partial \mathcal{P}}{\partial x}, \tag{2}$$

where x is defined as positive in the down-valley direction, q_{gw} is the per-channel-length contribution of groundwater from channel-adjacent hillslopes (units of m^2/day), $\mathcal{P}$ is the hyporheic flow capacity expressed in volumetric flow units (m^3/day) obtained by multiplying ρ at a point by upslope contributing area at that point ($\mathcal{P} = \rho \cdot A$), and q_{hef} (converted to m^2/day) is the channel-specific hyporheic exchange flow. If there is no gradient in hyporheic flow capacity (i.e., the right-most term in Equation 2 is zero), then continuity implies that hyporheic exchange must be equal to the incoming local groundwater flux. If instead q_{gw} is negligible, Equation 2 states that if volumetric flow capacity decreases in the downstream direction (i.e., $\frac{\partial \mathcal{P}}{\partial x} < 0$), there must be exfiltration ($q_{hef} > 0$) of water from the hyporheic zone into the surface environment (see Figure 2). Conversely, where flow capacity increases in the downstream direction $\left(\frac{\partial \mathcal{P}}{\partial x} > 0\right)$, water must infiltrate into the hyporheic zone ($q_{hef} < 0$). Thus at any point in the network where flow exceeds capacity, spatial gradients in flow capacity can dictate whether surface flows are infiltrating or exfiltrating from the hyporheic zone.

Numerous metrics have been proposed to quantify hyporheic exchange flows (Kasahara & Wondzell, 2003; Wondzell, 2011). Here, we calculate the average (D_{hef} m^3/day) of the magnitudes of total network-integrated exfiltrating (D_{exf}, m^3/day) and infiltrating ($D_{\inf}$) exchange flows, and report the ratio of D_{hef} to total volumetric discharge in the stream at the outlet (D, m^3/day) across a range of flow values (D_{hef} will change with catchment discharge because wetted extent, and thus the integration domain for exchange flows, changes with catchment discharge). To place a lower bound on these exchange fluxes, we note that where flow capacity increases in the downstream direction (i.e., hyporheic infiltration is possible), the additional flow capacity may entirely be occupied by incoming groundwater fluxes, q_{gw}. Therefore, a lower bound on hyporheic infiltration ($q_{\inf}$) may be calculated as:

$$q_{\inf} = \max\left[0, \frac{\partial \mathcal{P}}{\partial x} - q_{gw}\right] \quad \text{where}: \quad \frac{\partial \mathcal{P}}{\partial x} > 0. \tag{3}$$

To calculate q_{gw} along each 10 m reach between prediction points, we follow Schmadel et al. (2017) and multiply unit runoff (Q) by the contributing area difference between points, then divide by 10 m (the inter-node length), thus obtaining a channel-length specific groundwater efflux in units of m^2/day. Where $\frac{\partial \mathcal{P}}{\partial x} < 0$, decreasing flow capacity requires that a minimum of $\frac{\partial \mathcal{P}}{\partial x}$ must exfiltrate from water stored in the hyporheic zone, in addition to exfiltration driven by q_{gw} (which, we note may be the groundwater itself, or exfiltrating hyporheic storage displaced by incoming groundwater). Thus, a lower bound on exfiltration of hyporheic storage is:

$$q_{exf} = -\frac{\partial \mathcal{P}}{\partial x} \quad \text{where}: \quad \frac{\partial \mathcal{P}}{\partial x} < 0. \tag{4}$$

We integrate these length-specific rates of exchange along all wetted channel paths (i.e., where $D > \rho$), obtaining volumetric rates of infiltration ($D_{\inf}$) and exfiltration (D_{exf}) from the hyporheic zone. Finally, following Wondzell (2011), we calculate D_{hef} as the average of these two rates:

10.1029/2022WR034169

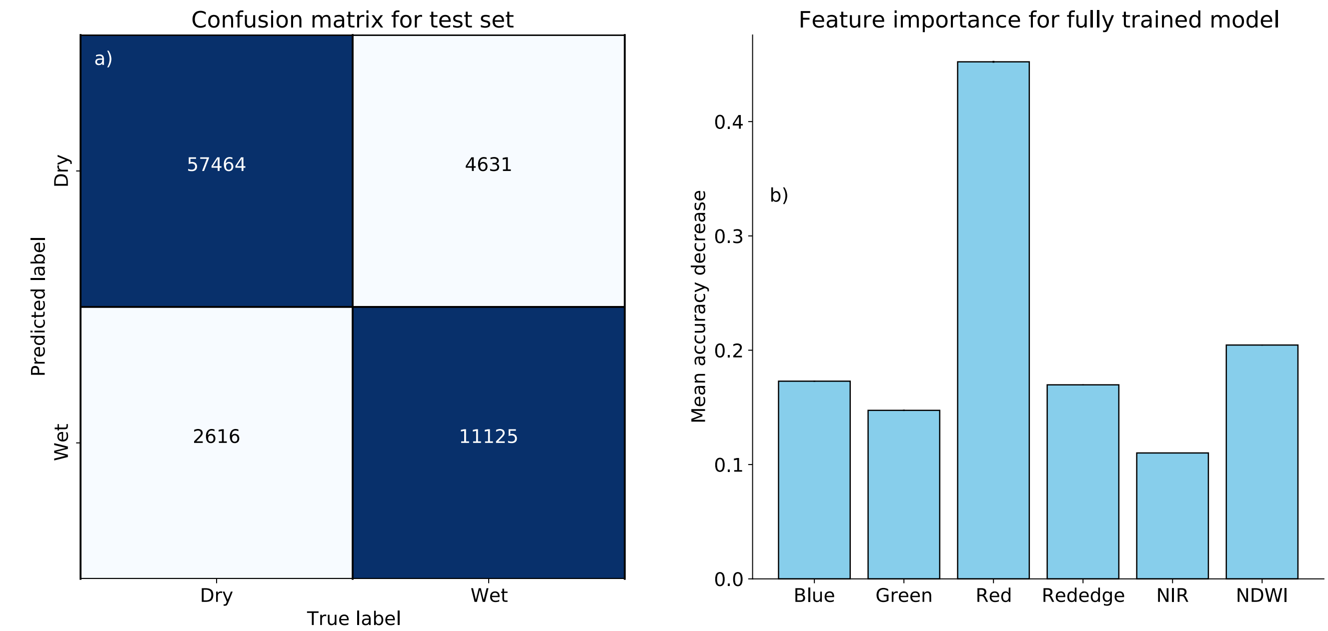

Figure 3. (a) Confusion matrix illustrating prediction accuracy of random forest classifier model (trained on 75% of original data) on test data (25% with-held data points). (b) Permutation feature importance for features in model trained on full training data set.

$$D_{hef} = (D_{\text{inf}} + D_{exf})/2 \tag{5}$$

The ratio we report ($D_{hef}: D$) is somewhat different from Wondzell (2011), who compute D_{hef} as the channel-length specific flux (and thus their ratio has units of m^{-1}). Here, ($D_{hef}: D$) is dimensionless, and can be interpreted as the ratio of the average gross volumetric flux between the hyporheic zone and the stream environment (D_{hef}) to the total volumetric flux exiting the watershed (D).

3. Results

In the Results, we first describe (a) the random forest model's ability to infer wetness state from satellite imagery, which makes possible a description of (b) temporal dynamics of the wetted channel network, and finally (c) the inferred catchment-scale hyporheic flow capacity, wetted channel network length-discharge relationship, and hyporheic exchange flows.

3.1. Random Forest Performance

Overall accuracy of the random forest model in validation is 91%. The confusion matrix in Figure 3a illustrates how this error is partitioned among false positives in the lower left corner (channel is classified as wet when it is actually dry) and false negatives in the upper right corner (channel is classified as dry when it is actually wet). The dark-colored diagonal of the confusion matrix contains the number of correct classifications. In general, false negatives (predicted dry when actually wet) are much more common than false positives. The relative prevalence of false negatives is also apparent in Figure 4 where many wet channels are classified as dry (dark blue in subplots i, f, and j). Even in this 100% wet training sample, however, prediction accuracy is quite good; 94% are predicted to be wet. The remaining rows in Figure 4 depict predictions and prediction error across other illustrative training data dates. The first two rows (a–c and d–f) illustrate predictions on the two walking survey dates from Lovill et al. (2018). The bottom row (j–l) illustrates prediction accuracy during the single drone survey date from a wet time of year (February is a peak wet season month) with channels between contributing areas of 10,000 and 20,000 m^2 that are nevertheless dry.

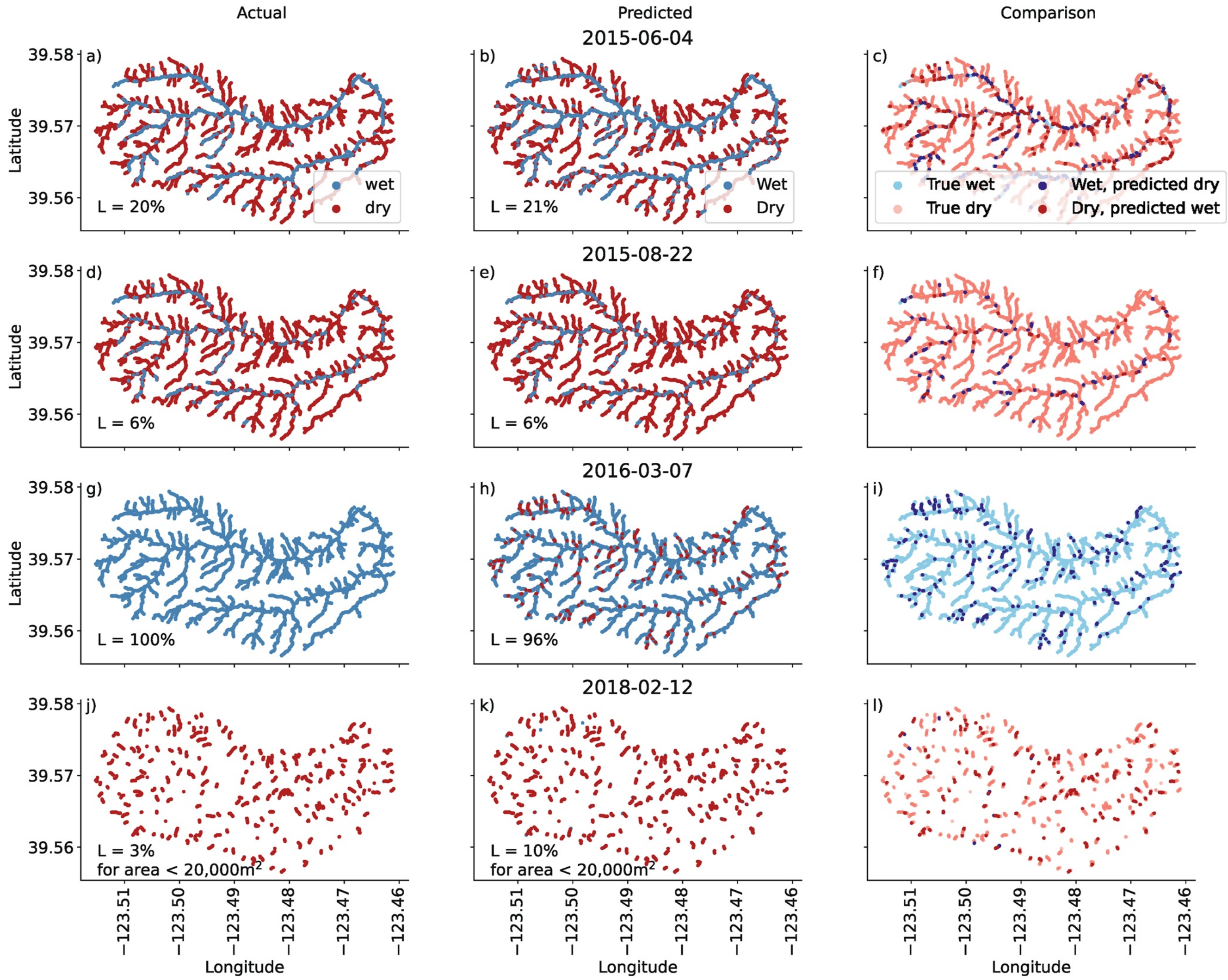

Figure 4. Visual representation of model performance on the four primary training dates. In each subplot, L expressed the extent of active channel as a percent of total channel length (wet plus dry). The top two rows are the surveyed dates in summer 2015. The third row represents a fully wetted network during a wet season peak flow event, and the bottom (drone survey data) represents an image during the wet season when many small channels are nevertheless dry. The fourth row does not illustrate continuous networks because we only have suitable training data for contributing areas less than 20,000 m^2 on this particular date.

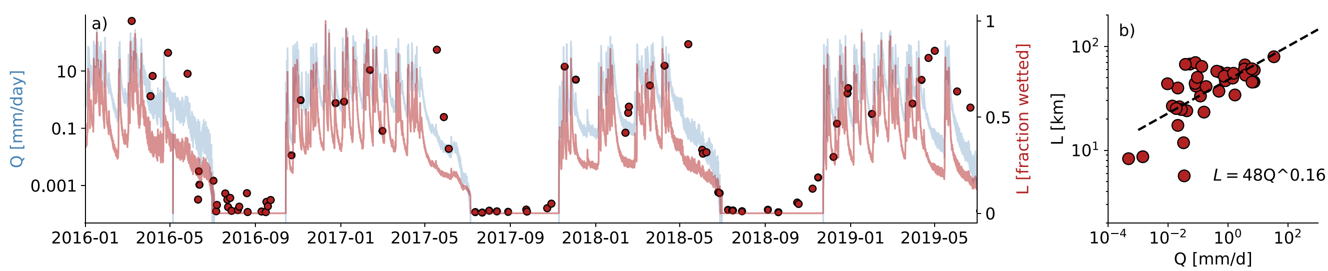

Figure 5. (a) Timeseries of Dry Creek runoff (blue) shown with timeseries of wetted channel length normalized by the maximum predicted length as predicted by a power law fit (red). Red points are discrete values of L calculated from random forest model predictions on RapidEye satellite imagery. (b) The predicted wetted channel extent length as a function of outlet runoff (the power law fit excludes zero flow predictions; $R^2 = 0.44$). The continuous wetted channel prediction (red line in (a)), is estimated from the power law fit in panel (b) using the continuous streamflow timeseries.

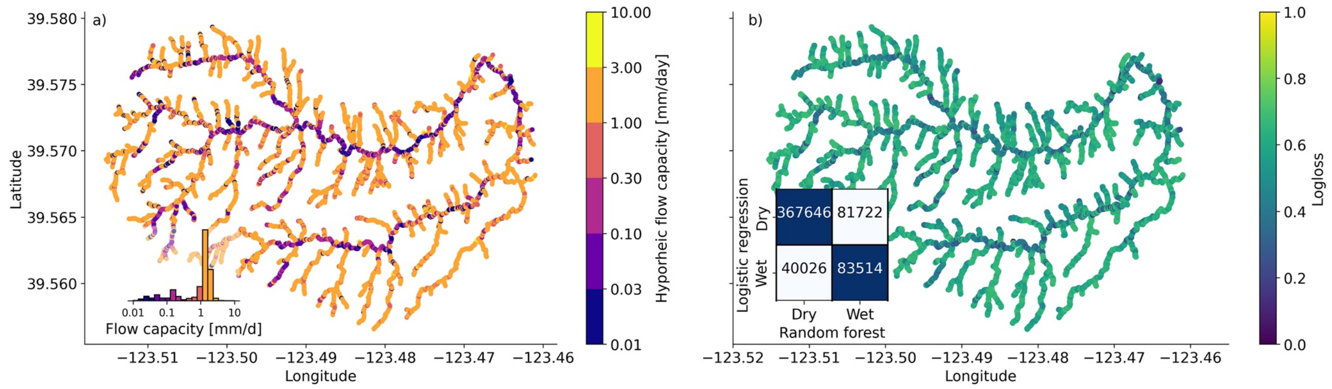

Figure 6. (a) Map of inferred hyporheic flow capacity shows decreased area-normalized (L T^{-1} units) flow capacity at higher drainage areas (main channel stems). Inset shows histogram of flow capacities. (b) Fit quality information for flow capacity logistic regressions (see Methods), including a map of log loss and (inset) a confusion matrix of logistic regression predictions compared to the random forest predictions to which they are fit. Fit quality is lower at higher logloss.

The LOO error analysis (last column of Table 1) indicates that wet training data from different times of year are extremely important for training an accurate random forest predictor. Including the single high-flow date was not adequate to train the model to recognize wetted reaches in general. Wet training data during the dry season (fifth row of Table 1) and high-flow dates (last row of Table 1) have by far the lowest LOO accuracies at 8% and 19%.

In the data repository (Dralle et al., 2022), we include additional analysis of the random forest model output and accuracy/uncertainty metrics. For example, we report on the agreement of trees within the random forest at different times of year under different conditions, demonstrating that inter-tree agreement is generally highest during the dry season (i.e., prediction confidence is highest), and lowest from the end of the dry season through the wet season and for wet predictions in general.

3.2. Wetted Channel Dynamics and Scaling

Wetted network extents predicted by the random forest (red points in Figure 5a) have seasonal patterns showing a mostly dry network in the summer when streamflow (light blue curve) is low, and variable extent during the wet season when flow varies over a few orders of magnitude (from 0.01 mm/day to nearly 50 mm/day). Consistent with theoretical expectations (Godsey & Kirchner, 2014; Prancevic & Kirchner, 2019), wetted extent generally exhibits power law scaling with runoff (Figure 5b), with a power law exponent $\beta = 0.16$. This exponent is likely an underestimate, as we do not predict wetted extent below contributing areas of 10,000 m^2, despite the fact that the network occasionally expands beyond this threshold during high flow periods (Lapides et al., 2022). The power law fit is also used to extrapolate wetted extent in Figure 5a (light red curve).

3.3. Estimates of Hyporheic Flow Capacity

For each point throughout the network, logistic regression of the random forest predictions (wet or dry) onto runoff provides an estimate of hyporheic flow capacity (in runoff units), which we map in Figure 6a (inset illustrates the probability distribution function (PDF) of flow capacities throughout the network, expressed in runoff units). Figure 6b illustrates goodness of fit of the logistic regression, expressed as logloss in the map, and via a confusion matrix in the inset. The confusion matrix illustrates whether the regression properly classifies channel wetness state under different flow conditions. As represented by logloss, fits are fair with better performance in larger channels, likely because larger channels are easier to observe at the pixel scale of RapidEye satellites. This effect would seem to overcome more limited training data in larger channels compared to smaller channels, which are more numerous.

Generally, flow capacities expressed as runoff units are lowest in mainstem channels (larger areas in Figure 7b, and lower values in Figure 6a inset PDF, typically between 0.01 mm/day and 1 mm/day) and highest in smaller tributaries (peak in the inset PDF between 1 and 10 mm/day), consistent with the expectation that the wetted network expands toward channel heads with increasing runoff at the outlet. Clear multi-modality of the runoff flow capacity PDF (inset of Figure 6a) suggests that activation of channels is punctuated at different flow levels,

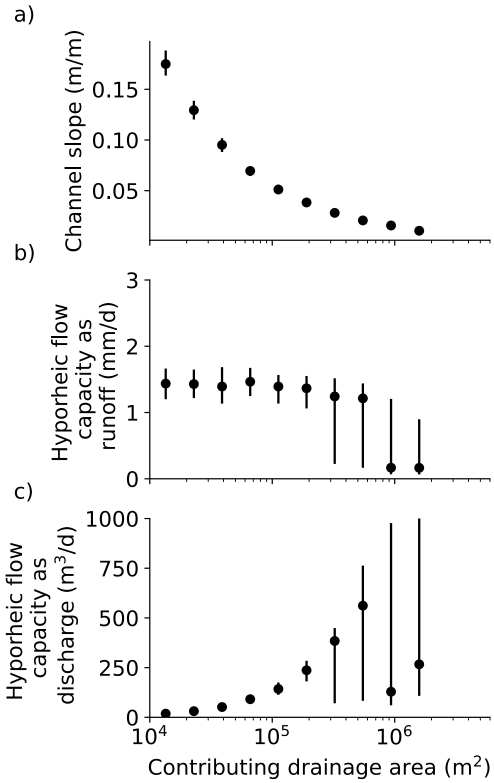

Figure 7. Network hydraulic scaling relationships: With increasing upstream contributing area, (a) channel slope decreases, (b) hyporheic flow capacity expressed as runoff (ρ) decreases, (c) hyporheic flow capacity expressed as discharge ($\mathcal{P}$) increases but then drops at high contributing areas, and (d) cross-sectional area transmissivity follows a similar increasing then decreasing pattern. Points and error bars show bin medians and the interquartile range.

with a large increase in channel length occurring near 1 mm/day when side channels activate. However, when expressed as absolute hyporheic flow capacity (discharge units), flow capacity increases with increasing drainage area, before becoming highly variable (with a smaller median value) in the main stem (Figure 7c).

3.4. Hyporheic Exchange Flows

Figure 8 plots the hyporheic exchange flux (D_{hef}) relative to volumetric discharge at the outlet (D) across a range of flow exceedance probabilities. The colorbar plots the correspondence between exceedance probability and runoff. For small exceedances (large flows), hyporheic exchange fluxes are comparable in magnitude to outlet discharge ($D_{hef} : D \approx 1$), whereas at low flows, the exchange flux magnitude increases to >100 times discharge at the outlet.

3.5. L-Q Relations and the Persistence of Wetted Channel Extent

Random forest prediction on imagery dates, the power law L-Q fit in Figure 5b, and the logistic regression provide three ways to explore the relationship between L and Q. We illustrate these relationships in Figure 9b, which shows that the logit and power law inferred L-Q relations fall in the point cloud of extents predicted with the random forest, although the functional forms are quite different. The logit predicts a sudden increase in wetted channel length between 1 and 2 mm/day, primarily due to expansion of the network out of the mainstem into side channels. These different methods of exploring L-Q relationships result in different probability distributions for L, plotted as cumulative distribution functions (CDFs) in Figure 9a. Differences arise because cloud-free imagery may be more readily available during dry periods, which would bias distributions inferred from imagery dates alone toward smaller wetted channel extents. This is made clear in Figure 9a, in which the CDF inferred from random forest imagery dates falls below the extents predicted from the logit and power law (which produce extent predictions on all days of the year solely as a function of discharge). Observational bias in the imagery toward dry dates is also apparent in Figure 9d, where the flow CDF predicts a higher likelihood of low flows when computed only using days on which imagery is available.

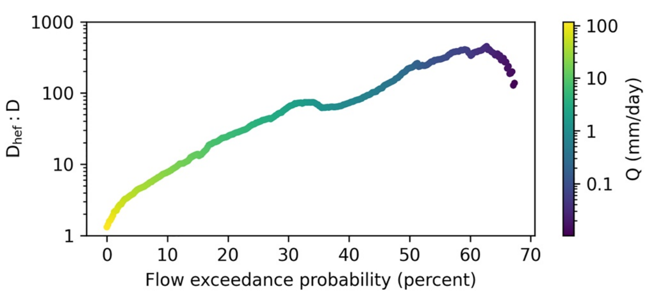

Figure 8. Magnitude of inferred, network-integrated hyporheic exchange flows relative to volumetric discharge at the catchment outlet across a range of flow values (expressed as runoff (mm/day) and as a flow exceedance probability).

4. Discussion

4.1. Surface Water Presence Detection in Small Headwater Stream Networks

The approach outlined here can be used to aid ecological monitoring of channel networks. As hydrological regimes shift in response to increasing anthropogenic pressures and a changing climate, so too will the wetted extent of channel networks (Lapides et al., 2021), with consequences for food webs, sedimentation, riverine nutrient cycling, habitat extent/quality, and other ecological processes (Arthington et al., 2005; Bernal et al., 2004; Hwan & Carlson, 2016; Larned et al., 2010; Sabo et al., 2010). A remote-sensing based framework for detecting the absence or presence of water in (often difficult-to-access) headwater stream networks would contribute to ongoing efforts to address this significant challenge in watershed management (Moidu et al., 2021).

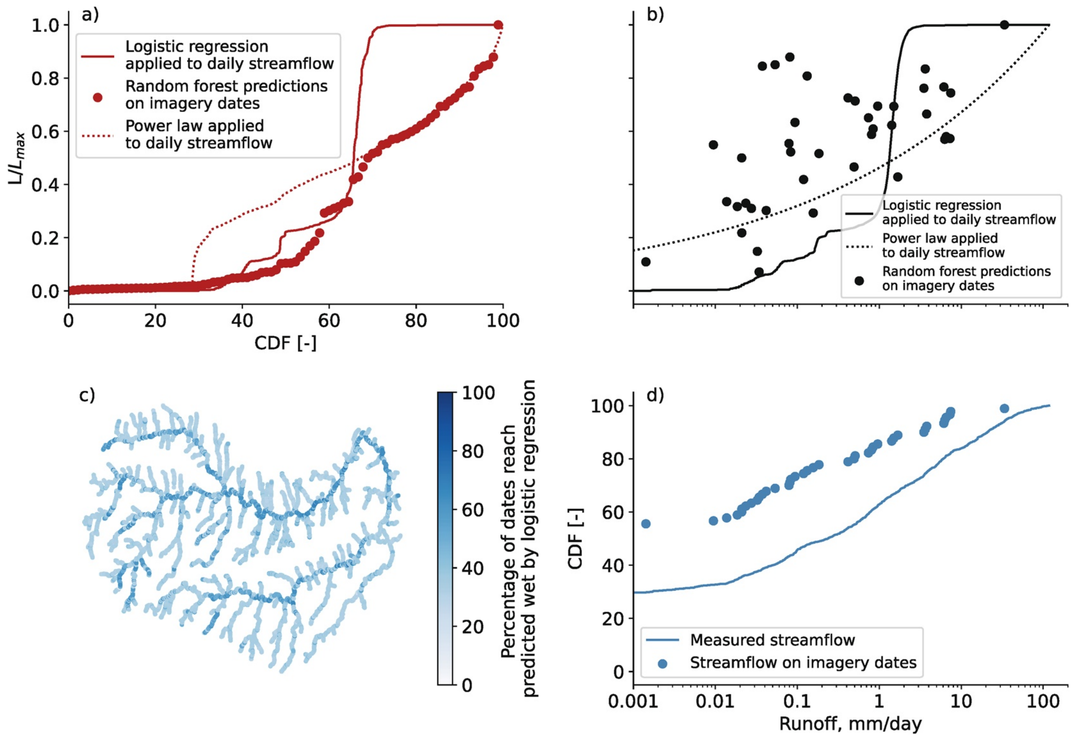

Figure 9. (a) Empirical cumulative distribution function (CDF) of wetted channel length from three methods: random forest-estimated wetted network length on imagery dates (scatter points), power law model trained on random forest results aggregated to total network length applied to all daily streamflow (dotted line), and logistic regression trained at each reach on random forest results applied to all daily streamflow (solid line). (b) Relationship between runoff and wetted channel length as estimated by each of the three methods. (c) Map of reach persistence based on logistic regression applied to all daily streamflow. (d) CDF of runoff using (solid line) all streamflow data during the study period and (scatter points) only streamflow on imagery dates. The CDF does not go to zero at the lowest non-zero flows because flow is zero (and thus not plot-able on a log scale) for a large fraction (roughly 30%) of the year.

4.2. Hyporheic Flow Properties Across River Networks

Drivers of surface water presence throughout headwater stream networks include: upstream runoff production, hyporheic zone transmissivity (hydraulic conductivity times average conductive depth), channel width, and slope (Ward et al., 2018). Channel slope can be approximated across landscapes using DEMs. However, both the spatial pattern of width-integrated transmissivity and the variation in runoff production are poorly constrained by our current data sets and understanding (Prancevic & Kirchner, 2019; Thompson et al., 2011). Timeseries of wetted channel extent reflect the spatial patterns in both of these fundamental but difficult-to-measure hydrological variables. Given an assumption about the pattern of transmissivity (such as the scaling relationship proposed in Prancevic & Kirchner, 2019), runoff production can be inferred from wetted channel maps. Conversely, given the assumption of spatial uniformity of runoff (Durighetto & Botter, 2022), the pattern of transmissivity (and thus hyporheic zone flow capacity, ρ) can be inferred with observations of water absence/presence in the channel that span a range of flows. The latter assumption is applied in this work to map ρ throughout the Dry and Hank Creek channel networks.

Estimating ρ using the presented method requires relatively high spatio-temporal resolution observations of wetted channel extent, but avoids unnecessary assumptions about the contributing-area scaling of hyporheic zone properties, thus generalizing the functional relationship (Figure 9b) between wetted channel length and discharge (Durighetto & Botter, 2022). The approach can account for the often discrete and discontinuous properties inherent in the geomorphology and geometry of channel networks. For example, an abrupt transition from a pool to a riffle, or from a tributary to a mainstem channel, might be accompanied by a large change in ρ (Käser et al., 2009; Schmadel et al., 2017). Here, the method revealed punctuated activation of different channels (mainstem vs. side channels) that resulted in a stepped L versus Q relationship (Figure 9b), which cannot be captured by the power-law L-Q model that emerges from presumed scaling relationships between contributing area and hyporheic zone transmissivity (Prancevic & Kirchner, 2019).

Estimates of ρ may also be useful in surface-groundwater exchange models, where parameters representing subsurface properties can influence understanding of hyporheic zone processes (Schmadel et al., 2017). In Figure 8, we used distributed ρ estimates to calculate that the magnitude of exchange fluxes relative to discharge ($D_{hef} : D$) increases significantly as flows decline, supporting the expectation that the influence of hyporheic processes on water quality (e.g., temperature, chemistry) is greater at low flows (Wondzell, 2011). Estimates of ρ could straightforwardly be used to parameterize subsurface elements of spatially distributed hyporheic zone models (e.g., Ward et al., 2018).

Summary of approach for predicting spatially resolved channel network wetness state across flow regime and spatially resolved hyporheic flow capacity: The workflow and heuristics developed here can be generally applied to any catchment with comparable data availability (and are not limited to satellite imagery; drones could also be used). Here, we provide a brief step-by-step overview of the methodology that can serve as a "recipe" for future applications in new catchments:

1. At the outlet of a focus catchment, acquire a streamflow discharge record of sufficient duration to include a large range of flows (and thus a large range of potential wetted channel extents).
2. Obtain local point observations of the extent of wetted channels across a range of streamflows when remotely sensed channel imagery is available (see following). It is not necessary to observe all points in the channel network, but the more data across a range of contributing areas, channel types, and climate and phenological conditions, the better. Section 2.2.4 describes heuristics for extrapolating and interpolating sparse observations to obtain a larger observational data set of wetted channel dynamics.
3. Obtain a vector of pixel spectral band values at each point in the channel network from multispectral remotely sensed imagery, including at locations and times where channel wetness states have been observed.
4. Train a statistical or machine learning classification model to identify channel wetness state (the target, wet or dry) as a function of spectral inputs (the predictors). We used a random forest model, described in Section 2.3.
5. Use the trained model and all available remotely sensed imagery to predict wetted state throughout the channel network at various points in time (and therefore across a range streamflow values).
6. Area normalize the outlet discharge (i.e., calculate instantaneous unit runoff) and match in time with the distributed predictions of channel wetness state to obtain pairs of runoff and wetness state at each point along the channel network for all dates with available imagery for generating wetness state predictions.
7. For each point in the channel network, perform logistic regression on the paired runoff and wetness state values to identify the streamflow threshold at which flow appears on the surface (Section 2.5). Under the assumptions outlined in Section 2.6, this flow threshold provides an estimate of the hyporheic flow capacity at that point in the network.

After completion of the workflow described above, additional analyses may be performed. For example, predictions of wetted channel length (the sum of lengths of all segments in the network predicted as wet by the random forest model) and flow can be used to estimate an L-Q relationship and then generate a continuous timeseries of wetted channel length (Section 2.4) and hyporheic flow properties (Section 2.5). Predictions of hyporheic flow capacity can be used to both generate continuous timeseries of channel wetness across the whole network and to constrain hyporheic exchange flows (e.g., Wondzell, 2011) as a function of total discharge at the catchment outlet (Section 2.6).

4.3. Challenges and Opportunities

We introduced a generic workflow that nevertheless remains untested in different environments. Perhaps most obvious is the need to try the approach in more heavily forested watersheds where the channel may not be so easily observed with satellite imagery. Even in such forested watersheds, higher resolution remote sensing data products with different sensing capabilities (e.g., thermal) and rapidly advancing unmanned aerial systems may make it possible to capture glimpses of channels through thick canopy.

Another limitation of the method is the availability of training data. Machine learning approaches are data hungry, and the LOO exercise in Table 1 reveals the relative importance of different training data in our seasonal watershed. However, we developed reasonable heuristics to increase the size of a training data set in data sparse environments, or where channel surveys are infrequent. For example, if a reach is mapped as dry for $q = q_0$, it stands to reason that reach will remain dry for $q = q_1 \ll q_0$, making it possible to utilize imagery on various dates for training a machine learning model. Analogously, if a reach is mapped as wet for $q = q_0$, it likely remains wet for $q = q_2 \gg q_0$. These heuristics (which follow from the flow emergence principle and the uniform runoff assumption) make it possible to expand sparse training data sets to include a wider range of environmental and flow conditions.

Higher data availability and quality may never answer whether machine learning models, which are difficult to interpret mechanistically, are getting the right answers for the right reasons. Is our random forest model truly "seeing" the water in the channels? There are promising indicators. The model performs very well in validation, and coherent scaling relationships between contributing area and hyporheic flow properties emerge (e.g., mainstems have demonstrably lower ρ; Figure 7). The latter is promising considering we did not use contributing area as a predictor; when contributing area is included, results are generally similar. A random forest model is also among the less sophisticated machine learning models; more complex methods (e.g., convolutional neural networks) may provide additional support for the validity of the general approach, and may in fact be necessary in more challenging settings where, for example, canopy cover obscures channels.

For this technique to work a series of assumptions is required. Our interpolation and extrapolation approaches were designed to be as conservative as possible, and were chosen to be consistent with the hierarchical principle described by Botter et al. (2021) and current understanding of the physical drivers of wetted channel dynamics (e.g., Durighetto & Botter, 2022; Prancevic & Kirchner, 2019). The assumption of spatially uniform unit runoff may be problematic when attempting to use outlet discharge across larger catchments, where climate gradients and geological or vegetation heterogeneity might impact water balance within subwatersheds. Even in our small catchment, distinct lithologic blocks within the melange likely exhibit different flow generation properties Lovill et al. (2018). Measuring heterogeneity of runoff generation and flow recessions at sub-catchment scales is an outstanding problem in hydrology (e.g., Harman et al., 2009), and the assumption of uniform specific runoff is nearly ubiquitous across present mechanistic frameworks for predicting wetted channel dynamics (Durighetto & Botter, 2022; Prancevic & Kirchner, 2019; Ward et al., 2018). Additionally, we restricted our analysis to channels with larger contributing areas >10,000 m^2, hoping to integrate enough area to smooth out/overwhelm any issues related to small-scale runoff generation variability that might stem from specific hillslope, vegetation, or pedologic idiosyncrasies.

5. Conclusion

We demonstrate a proof-of-concept approach for using multi-spectral imagery and machine learning trained on observational data to monitor the growth and contraction of a headwater stream at high spatial and temporal resolution. The method predicts water presence with 91% accuracy. Assuming unit runoff is spatially uniform, and that water emerges in the channel when up-network runoff production exceeds the flow capacity of the hyporheic zone, we use predicted maps of channel extent to estimate hyporheic hydrogeologic properties and hyporheic exchange. The approach has promising applications in environmental monitoring, and details a prototypical workflow for potential applications in other environments.

Data Availability Statement

Data and code for this publication are available in an online data repository (Dralle et al., 2022).

Acknowledgments
We would like to thank Planet Labs for providing high-resolution imagery through their Education and Research Program (Planet Team, 2017) and Marilyn Russell for site access. Funding was provided by Simon Fraser University, a Natural Sciences and Engineering Research Council of Canada Discovery Grant, the Canadian Foundation for Innovation, and the USDA Forest Service Pacific Southwest Research Station with funds administered by Oak Ridge Institute for Science and Education (ORISE).

References

Allen, G. H., Pavelsky, T. M., Barefoot, E. A., Lamb, M. P., Butman, D., Tashie, A., & Gleason, C. J. (2018). Similarity of stream width distributions across headwater systems. *Nature Communications*, *9*(1), 1–7. https://doi.org/10.1038/s41467-018-02991-w

Arthington, A. H., Balcombe, S. R., Wilson, G. A., Thoms, M. C., & Marshall, J. (2005). Spatial and temporal variation in fish-assemblage structure in isolated waterholes during the 2001 dry season of an arid-zone floodplain river, Cooper Creek, Australia. *Marine and Freshwater Research*, *56*(1), 25–35. https://doi.org/10.1071/mf04111

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Bernal, S., Butturini, A., Riera, J., Vázquez, E., & Sabater, F. (2004). Calibration of the INCA model in a Mediterranean forested catchment: The effect of hydrological inter-annual variability in an intermittent stream. *Hydrology and Earth System Sciences*, *8*(4), 729–741. https://doi.org/10.5194/hess-8-729-2004

Beven, K. J., & 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. *Hydrological Sciences Journal*, *24*(1), 43–69. https://doi.org/10.1080/02626667909491834

Bishop, K., Buffam, I., Erlandsson, M., Fölster, J., Laudon, H., Seibert, J., & Temnerud, J. (2008). *Aqua Incognita*: The unknown headwaters. *Hydrological Processes*, *22*(8), 1239–1242. https://doi.org/10.1002/hyp.7049

Blake, M. C., Jr., Jayko, A., & McLaughlin, R. (1985). Tectonostratigraphic terranes of the northern coast ranges, California. US Geological Survey. Circum Pacific Council Publications. Retrieved from https://archives.datapages.com/data/circ_pac/1/159_b.htm

Botter, G., Vingiani, F., Senatore, A., Jensen, C., Weiler, M., McGuire, K., et al. (2021). Hierarchical climate-driven dynamics of the active channel length in temporary streams. *Scientific Reports*, *11*(1), 21503. https://doi.org/10.1038/s41598-021-00922-2

Carbonneau, P. E., Belletti, B., Micotti, M., Lastoria, B., Casaioli, M., Mariani, S., et al. (2020). UAV-based training for fully fuzzy classification of sentinel-2 fluvial scenes. *Earth Surface Processes and Landforms*, *45*(13), 3120–3140. https://doi.org/10.1002/esp.4955

Carlston, C. W. (1963). *Drainage density and streamflow*. US Government Printing Office.

Chambers, J. Q., Robertson, A. L., Carneiro, V., Lima, A. J., Smith, M.-L., Plourde, L. C., & Higuchi, N. (2009). Hyperspectral remote detection of niche partitioning among canopy trees driven by blowdown gap disturbances in the central amazon. *Oecologia*, *160*(1), 107–117. https://doi.org/10.1007/s00442-008-1274-9

Datry, T., Larned, S. T., & Scarsbrook, M. R. (2007). Responses of hyporheic invertebrate assemblages to large-scale variation in flow permanence and surface–subsurface exchange. *Freshwater Biology*, *52*(8), 1452–1462. https://doi.org/10.1111/j.1365-2427.2007.01775.x

Dralle, D. N., Hahm, W. J., Rempe, D. M., Karst, N. J., Thompson, S. E., & Dietrich, W. E. (2018). Quantification of the seasonal hillslope water storage that does not drive streamflow. *Hydrological Processes*, *32*(13), 1978–1992. https://doi.org/10.1002/hyp.11627

Dralle, D. N., Lapides, D. A., Rempe, D. M., & Hahm, W. J. (2022). Harnessing hyperspectral imagery to map surface water presence and hyporheic flow properties of headwater stream networks [Dataset and Code. https://doi.org/10.5281/zenodo.7199463

Dugdale, S. J., Klaus, J., & Hannah, D. M. (2022). Looking to the skies: Realising the combined potential of drones and thermal infrared imagery to advance hydrological process understanding in headwaters. *Water Resources Research*, *58*(2), e2021WR031168. https://doi.org/10.1029/2021wr031168

Durighetto, N., Bertassello, L. E., & Botter, G. (2022). Eco-hydrological modelling of channel network dynamics—Part 1: Stochastic simulation of active stream expansion and retraction. *Royal Society Open Science*, *9*(11), 220944. https://doi.org/10.1098/rsos.220944

Durighetto, N., & Botter, G. (2022). On the relation between active network length and catchment discharge. *Geophysical Research Letters*, *49*(14), e2022GL099500. https://doi.org/10.1029/2022gl099500

Durighetto, N., Vingiani, F., Bertassello, L. E., Camporese, M., & Botter, G. (2020). Intraseasonal drainage network dynamics in a headwater catchment of the Italian Alps. *Water Resources Research*, *56*(4), e2019WR025563. https://doi.org/10.1029/2019WR025563

Eriksson, H. M., Eklundh, L., Kuusk, A., & Nilson, T. (2006). Impact of understory vegetation on forest canopy reflectance and remotely sensed LAI estimates. *Remote Sensing of Environment*, *103*(4), 408–418. https://doi.org/10.1016/j.rse.2006.04.005

Fovet, O., Belemtougri, A., Boithias, L., Braud, I., Charlier, J.-B., Cottet, M., et al. (2021). Intermittent rivers and ephemeral streams: Perspectives for critical zone science and research on socio-ecosystems. *Wiley Interdisciplinary Reviews: Water*, *8*(4), e1523. https://doi.org/10.1002/wat2.1523

Gao, B.-C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266. https://doi.org/10.1016/s0034-4257(96)00067-3

Giezendanner, J., Benettin, P., Durighetto, N., Botter, G., & Rinaldo, A. (2021). A note on the role of seasonal expansions and contractions of the flowing fluvial network on metapopulation persistence. *Water Resources Research*, *57*(11), e2021WR029813. https://doi.org/10.1029/2021wr029813

Godsey, S. E., & Kirchner, J. W. (2014). Dynamic, discontinuous stream networks: Hydrologically driven variations in active drainage density, flowing channels and stream order. *Hydrological Processes*, *28*(23), 5791–5803. (00010). https://doi.org/10.1002/hyp.10310

Hahm, W. J., Dietrich, W. E., & Dawson, T. E. (2018). Controls on the distribution and resilience of *Quercus garryana*: Ecophysiological evidence of oak's water-limitation tolerance. *Ecosphere*, *9*(5), e02218. https://doi.org/10.1002/ecs2.2218

Hahm, W. J., Dralle, D. N., Lovill, S. M., Rose, J., Dawson, T. E., & Dietrich, W. E. (2017). Exploratory tree survey (2016 - Eel river critical zone observatory - Sagehorn - Central belt melange, Franciscan complex, northern California coast ranges, USA). HydroShare, https://doi.org/10.4211/hs.7881821a5c0e4ae3822b96a59f4bf8b6

Hahm, W. J., Rempe, D. M., Dralle, D. N., Dawson, T. E., Lovill, S. M., Bryk, A. B., et al. (2019). Lithologically controlled subsurface critical zone thickness and water storage capacity determine regional plant community composition. *Water Resources Research*, *55*(4), 3028–3055. https://doi.org/10.1029/2018wr023760

Harman, C., Sivapalan, M., & Kumar, P. (2009). Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics. *Water Resources Research*, *45*(9), W09404. https://doi.org/10.1029/2008wr007392

Hewlett, J. D., & Hibbert, A. R. (1967). Factors affecting the response of small watersheds to precipitation in humid areas. *Forest Hydrology*, *1*, 275–290.

Hwan, J., & Carlson, S. (2016). Fragmentation of an intermittent stream during seasonal drought: Intra-annual and interannual patterns and biological consequences. *River Research and Applications*, *32*(5), 856–870. https://doi.org/10.1002/rra.2907

Jaeger, K. L., Hafen, K. C., Dunham, J. B., Fritz, K. M., Kampf, S. K., Barnhart, T. B., et al. (2021). Beyond streamflow: Call for a national data repository of streamflow presence for streams and rivers in the United States. *Water*, *13*(12), 1627. https://doi.org/10.3390/w13121627

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

Jensen, C. K., McGuire, K. J., McLaughlin, D. L., & Scott, D. T. (2019). Quantifying spatiotemporal variation in headwater stream length using flow intermittency sensors. *Environmental Monitoring and Assessment*, *191*(4), 226. https://doi.org/10.1007/s10661-019-7373-8

Kampf, S. K., Dwire, K. A., Fairchild, M. P., Dunham, J., Snyder, C. D., Jaeger, K. L., et al. (2021). Managing nonperennial headwater streams in temperate forests of the United States. *Forest Ecology and Management*, *497*, 119523. https://doi.org/10.1016/j.foreco.2021.119523

Kasahara, T., & Wondzell, S. M. (2003). Geomorphic controls on hyporheic exchange flow in mountain streams. *Water Resources Research*, *39*(1), SBH3-1–SBH3-14. https://doi.org/10.1029/2002wr001386

Käser, D. H., Binley, A., Heathwaite, A. L., & Krause, S. (2009). Spatio-temporal variations of hyporheic flow in a riffle-step-pool sequence. *Hydrological Processes: An International Journal*, *23*(15), 2138–2149. https://doi.org/10.1002/hyp.7317

Langenheim, V., Jachens, R. C., Wentworth, C. M., & McLaughlin, R. J. (2013). Previously unrecognized regional structure of the coastal belt of the Franciscan complex, northern California, revealed by magnetic data. *Geosphere*, *9*(6), 1514–1529. https://doi.org/10.1130/ges00942.1

Lapides, D. A., Hahm, W. J., Rempe, D. M., Dietrich, W. E., & Dralle, D. N. (2022). Controls on stream water age in a saturation overland flow-dominated catchment. *Water Resources Research*, *58*(4), e2021WR031665. https://doi.org/10.1029/2021wr031665

Lapides, D. A., Leclerc, C. D., Moidu, H., Dralle, D. N., & Hahm, W. J. (2021). Variability of stream extents controlled by flow regime and network hydraulic scaling. *Hydrological Processes*, *35*(3), e14079. https://doi.org/10.1002/hyp.14079

Larned, S. T., Datry, T., Arscott, D. B., & Tockner, K. (2010). Emerging concepts in temporary-river ecology. *Freshwater Biology*, *55*(4), 717–738. https://doi.org/10.1111/j.1365-2427.2009.02322.x

Li, D., Wu, B., Chen, B., Qin, C., Wang, Y., Zhang, Y., & Xue, Y. (2020). Open-surface river extraction based on sentinel-2 MSI imagery and DEM data: Case study of the upper yellow river. *Remote Sensing*, *12*(17), 2737. https://doi.org/10.3390/rs12172737

Linderman, M., Liu, J., Qi, J., An, L., Ouyang, Z., Yang, J., & Tan, Y. (2004). Using artificial neural networks to map the spatial distribution of understorey bamboo from remote sensing data. *International Journal of Remote Sensing*, *25*(9), 1685–1700. https://doi.org/10.1080/01431160310001598971

Ling, F., Li, X., Foody, G. M., Boyd, D., Ge, Y., Li, X., & Du, Y. (2020). Monitoring surface water area variations of reservoirs using daily MODIS images by exploring sub-pixel information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *168*, 141–152. https://doi.org/10.1016/j.isprsjprs.2020.08.008

Lovill, S. M., Hahm, W. J., & Dietrich, W. E. (2018). Drainage from the critical zone: Lithologic controls on the persistence and spatial extent of wetted channels during the summer dry season. *Water Resources Research*, *54*(8), 5702–5726. https://doi.org/10.1029/2017WR021903

Messager, M. L., Lehner, B., Cockburn, C., Lamouroux, N., Pella, H., Snelder, T., et al. (2021). Global prevalence of non-perennial rivers and streams. *Nature*, *594*(7863), 391–397. https://doi.org/10.1038/s41586-021-03565-5

Micieli, M., Botter, G., Mendicino, G., & Senatore, A. (2022). UAV thermal images for water presence detection in a Mediterranean headwater catchment. *Remote Sensing*, *14*(1), 108. https://doi.org/10.3390/rs14010108

Moidu, H., Obedzinski, M., Carlson, S., & Grantham, T. (2021). Spatial patterns and sensitivity of intermittent stream drying to climate variability. *Water Resources Research*, *57*(11), e2021WR030314. https://doi.org/10.1029/2021wr030314

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Perez-Saez, J., Mande, T., Larsen, J., Ceperley, N., & Rinaldo, A. (2017). Classification and prediction of river network ephemerality and its relevance for waterborne disease epidemiology. *Advances in Water Resources*, *110*, 263–278. https://doi.org/10.1016/j.advwatres.2017.10.003

Planet Team. (2017). Planet application program interface: In space for life on earth. Retrieved from https://api.planet.com

Prancevic, J. P., & Kirchner, J. W. (2019). Topographic controls on the extension and retraction of flowing streams. *Geophysical Research Letters*, *46*(4), 2084–2092. https://doi.org/10.1029/2018GL081799

Qin, P., Cai, Y., & Wang, X. (2021). Small waterbody extraction with improved U-Net using Zhuhai-1 hyperspectral remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5. https://doi.org/10.1109/lgrs.2020.3047918

Sabo, J. L., Finlay, J. C., Kennedy, T., & Post, D. M. (2010). The role of discharge variation in scaling of drainage area and food chain length in rivers. *Science*, *330*(6006), 965–967. https://doi.org/10.1126/science.1196005

Sando, R., & Blasch, K. W. (2015). Predicting alpine headwater stream intermittency: A case study in the northern Rocky Mountains. *Ecohydrology and Hydrobiology*, *15*(2), 68–80. https://doi.org/10.1016/j.ecohyd.2015.04.002

Schmadel, N. M., Ward, A. S., & Wondzell, S. M. (2017). Hydrologic controls on hyporheic exchange in a headwater mountain stream. *Water Resources Research*, *53*(7), 6260–6278. https://doi.org/10.1002/2017wr020576

Thompson, S. E., Harman, C. J., Troch, P. A., Brooks, P. D., & Sivapalan, M. (2011). Spatial scale dependence of ecohydrologically mediated water balance partitioning: A synthesis framework for catchment ecohydrology. *Water Resources Research*, *47*(10), W00J03. https://doi.org/10.1029/2010wr009998

Vanderhoof, M. K., & Burt, C. (2018). Applying high-resolution imagery to evaluate restoration-induced changes in stream condition, Missouri river headwaters basin, Montana. *Remote Sensing*, *10*(6), 913. https://doi.org/10.3390/rs10060913

Verma, U., Chauhan, A., MM, M. P., & Pai, R. (2021). DeepRivWidth: Deep learning based semantic segmentation approach for river identification and width measurement in SAR images of coastal Karnataka. *Computers & Geosciences*, *154*, 104805. https://doi.org/10.1016/j.cageo.2021.104805

Wang, X., Gong, J., Zhang, Y., & Atkinson, P. M. (2022). Near real-time surface water extraction from GOES-16 geostationary satellite ABI images by constructing and sharpening the green-like band. *Science of Remote Sensing*, *5*, 100055. https://doi.org/10.1016/j.srs.2022.100055

Ward, A. S., Schmadel, N. M., & Wondzell, S. M. (2018). Simulation of dynamic expansion, contraction, and connectivity in a mountain stream network. *Advances in Water Resources*, *114*, 64–82. https://doi.org/10.1016/j.advwatres.2018.01.018

Whiting, J. A., & Godsey, S. E. (2016). Discontinuous headwater stream networks with stable flowheads, Salmon River basin, Idaho. *Hydrological Processes*, *30*(13), 2305–2316. https://doi.org/10.1002/hyp.10790

Wondzell, S. M. (2011). The role of the hyporheic zone across stream networks. *Hydrological Processes*, *25*(22), 3525–3532. https://doi.org/10.1002/hyp.8119

Wondzell, S. M., & Ward, A. S. (2022). The channel-source hypothesis: Empirical evidence for in-channel sourcing of dissolved organic carbon to explain hysteresis in a headwater mountain stream. *Hydrological Processes*, *36*(5), e14570. https://doi.org/10.1002/hyp.14570

Xue, Y., Qin, C., Wu, B., Li, D., & Fu, X. (2022). Automatic extraction of mountain river surface and width based on multisource high-resolution satellite images. *Remote Sensing*, *14*(10), 2370. https://doi.org/10.3390/rs14102370

Zanetti, F., Durighetto, N., Vingiani, F., & Botter, G. (2022). Analyzing river network dynamics and the active length–discharge relationship using water presence sensors. *Hydrology and Earth System Sciences*, *26*(13), 3497–3516. https://doi.org/10.5194/hess-26-3497-2022