AGU ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research

RESEARCH ARTICLE
10.1029/2023WR034513

Communication Distance and Bayesian Inference in Non-Perennial Streams

Ken Aho[1], Dewayne Derryberry[2], Sarah E. Godsey[3], Rob Ramos[4], Sara R. Warix[3,5], and Samuel Zipper[6]

[1]Department of Biological Sciences, Idaho State University, Pocatello, ID, USA, [2]Department of Mathematics and Statistics, Idaho State University, Pocatello, ID, USA, [3]Department of Geosciences, Idaho State University, Pocatello, ID, USA, [4]Biological Survey & Center for Ecological Research, University of Kansas, Lawrence, KS, USA, [5]Now at Hydrologic Science and Engineering, Colorado School of Mines, Golden, CO, USA, [6]Kansas Geological Survey, The University of Kansas, Lawrence, KS, USA

Key Points:

- Conventional measures of stream connectivity are often poorly adapted to dynamic non-perennial stream networks that expand and contract
- We developed Bayesian extensions to an existing stream length model, and a new Bayes-applicable metric called communication distance
- Communication distance measures increased resistance to water-borne transport in non-perennial streams, in units of stream length

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

K. Aho,
ahoken@isu.edu

Citation:

Aho, K., Derryberry, D., Godsey, S. E., Ramos, R., Warix, S. R., & Zipper, S. (2023). Communication distance and Bayesian inference in non-perennial Streams. *Water Resources Research*, *59*, e2023WR034513. https://doi.org/10.1029/2023WR034513

Received 24 JAN 2023
Accepted 21 OCT 2023

Author Contributions:

Conceptualization: Ken Aho, Dewayne Derryberry, Sarah E. Godsey
Data curation: Ken Aho
Formal analysis: Ken Aho
Investigation: Ken Aho, Dewayne Derryberry
Methodology: Ken Aho, Dewayne Derryberry, Rob Ramos
Project Administration: Sarah E. Godsey
Resources: Sara R. Warix
Software: Ken Aho, Rob Ramos
Supervision: Ken Aho

Abstract Non-perennial streams are receiving increased attention from researchers, however, suitable methods for measuring their hydrologic connectivity remain scarce. To address this deficiency, we developed Bayesian statistical approaches for measuring both average active stream length, and a new metric called average communication distance. Average communication distance is a theoretical increased *effective distance* that stream-borne materials must travel, given non-continuous streamflow. Because it is the product of the inverse probability of surface water presence and stream length, the average communication distance of a non-perennial stream segment will be greater than its actual physical length. As an application we considered Murphy Creek, a simple non-perennial stream network in southwestern Idaho, USA. We used surface water presence/absence data obtained in 2019, and priors for the probability of surface water, based on predictions from an existing regional United States Geological Survey model. Average communication distance posterior distributions revealed locations where effective stream lengths increased dramatically due to flow rarity. We also found strong seasonal (spring, summer, fall) differences in network-level posterior distributions of both average stream length and average communication distance. Our work demonstrates the unique perspectives concerning network drying provided by communication distance, and demonstrates the general usefulness of Bayesian approaches in the analysis of non-perennial streams.

Plain Language Summary We developed a new metric, communication distance, appropriate for measuring connectivity in non-perennial stream networks. Communication distance can be considered a theoretical potential distance that water borne material must travel in the absence of continuous surface flow. Communication distance will be in units of measured stream length. Nonetheless, the communication distance of a non-perennial stream segment will be greater than the actual physical length of the segment, and this distance will increase further with increased intermittency. We developed Bayesian extensions for both communication distance, and an existing stream length model of network connectivity. The use of a Bayesian approach allowed: (a) explicit consideration of the variation and uncertainty in stream segment probabilities of surface water presence, and (b) the incorporation of preexisting US Geological Survey model predictions as a framework for Bayesian priors.

1. Introduction

Non-perennial streams comprise over half of the global river network (Messager et al., 2021), are increasing in prevalence (Sauquet et al., 2021; Zipper et al., 2021), and strongly influence global water quantity and quality (Datry et al., 2014). Realization of the importance of non-perennial streams to large-scale hydrological, ecological, and biogeochemical processes has prompted increased study of these systems (Fovet et al., 2021). Nonetheless, characterization of non-perennial stream spatiotemporal dynamics remains challenging (Shanafield et al., 2021), inhibiting a clear understanding of linkages between stream drying and water quality.

1.1. Network Connectivity in Intermittent Streams

Decreased connectivity of stream segments from drying may affect water quality by preventing surface transport of materials. Numerous stream connectivity metrics exist (see reviews in Ali and Roy (2010), Bracken

Validation: Ken Aho, Dewayne Derryberry, Sarah E. Godsey
Visualization: Ken Aho
Writing – original draft: Ken Aho, Sara R. Warix
Writing – review & editing: Ken Aho, Dewayne Derryberry, Sarah E. Godsey, Rob Ramos, Sara R. Warix, Samuel Zipper

et al. (2013), and Blume and Van Meerveld (2015)), due in part to myriad perspectives concerning hydrologic connectivity (Ali & Roy, 2009). These methods, however, have limitations for describing non-perennial streams. For example, several common measures of stream connectivity are time invariant due to their reliance on Cartesian grid relationships (e.g., Larsen et al., 2012; Trigg et al., 2013), or topography and drainage area (e.g., Jensco et al., 2009; Prancevic & Kirchner, 2019). Thus, these measures may poorly describe non-perennial stream networks whose extent will vary in both time and space (Bertassello et al., 2022). Further, other stream connectivity measures, including those based on distances between "wet" locations (e.g., Ali & Roy, 2010; Western et al., 2001), or spatial autocorrelation structures (e.g., Ali & Roy, 2010; Knudby & Carrera, 2005), provide only network-scale descriptions. Thus, these methods do not consider drying patterns at the scale of individual stream segments. This latter deficiency is particularly problematic in non-perennial streams because certain locations may have inordinately large effects on stream networks (Godsey & Kirchner, 2014; Zipper et al., 2022). Of particular relevance are surface flow *bottlenecks*, that is, stream locations where surface flows often disappear, preventing the surficial flow of water from upstream to downstream locations.

1.2. Probabilistic Measures of Stream Connectivity

The variability of surface flow in non-perennial streams has driven the development and application of probabilistic models for surface water presence, often at watershed or larger spatial scales. These approaches include hidden Markov chain models based on stream temperature and conductivity (Arismendi et al., 2017), models for intermittent stream length as a function of climatic variables (Durighetto et al., 2020), logistic models for stream persistency based on intermittency sensors and geospatial and climatic data (Jensen et al., 2019; Kaplan et al., 2020), random forest classifications from remotely sensed geographic information system data (González-Ferreras & Barquín, 2017; Jaeger et al., 2019; Sando & Blasch, 2015), and probabilistic consideration of the relationship between intermittent stream length and catchment discharge (Durighetto & Botter, 2022). Importantly, Botter and Durighetto (2020) developed a probability density function (PDF) approach to define the distribution of stream network length, called the stream length duration curve (SLDC). A SLDC depicts the distribution of the "active" fraction of a stream network (i.e., the network portion with surface flow), and provides the inverse of the exceedance probability of the total length of active streams for any outlet discharge.

1.3. The Appeal of Bayesian Methods

The development of probabilistic approaches for considering surface flow in non-perennial streams is commendable. Existing work, however, generally employs a frequentist view of probability which assumes a single "true" value for the probability of water presence at a stream segment over some timespan. This view potentially ignores variation in diel and seasonal probabilities of stream segment water presence, and more importantly, may prevent assessment of uncertainty and variability in probability designations. These properties, however, can be readily considered under a Bayesian statistical approach.

Many sources of information concerning wetting and drying patterns may exist for a stream network, potentially based on multiple spatiotemporal scales and sampling schema. For example, it is possible that at a single watershed, stream surface flow has been: (a) modeled as part of subcontinent-scale research projects (e.g., USGS-PROSPER; Jaeger et al., 2019), (b) categorized into presence/absence outcomes at locations occasionally visited by local agencies or researchers, and (c) measured at a small number of locations using high-frequency intermittency sensors over days to years. Such prior information can be assimilated into Bayesian statistical analyses to inform and refine models based on current data, for example, resistivity sensor outcomes for the present water year. Weights for these *priors*, in the form of effective sample sizes (compared to current data sample sizes), can also be assigned based on prior data quality and the agreement of measurement scales of prior and current data. The *posterior distribution* represents a formal Bayesian synthesis of prior and current information. Because it will have the form of a PDF, the posterior allows straightforward assessments of uncertainty in modeled phenomena. The application of Bayesian methods seems particularly useful for depicting the probability of surface water presence in non-perennial streams, given the frequent availability of prior information, and the importance of quantifying the central tendency and variation in this parameter.

1.4. Products

In this paper we develop Bayesian statistical methods to measure stream network connectivity that allow: (a) simultaneous consideration of global (entire network) and local (stream segment) scales, (b) explicit consideration

of the variability and uncertainty in the probability of surface water presence, and (c) inclusion of prior information concerning the probability of surface water presence. We also introduce a new metric called *communication distance* that measures the theoretical *effective stream length* of segments as perceived by water-borne components requiring surface water transport. This measure may improve understanding of the balance of transport, storage, and reaction limitations within non-perennial networks and their downstream waters, whether they dry or not. Bayesian application of communication distance prompted the first reported analytical derivation of the inverse-beta distribution (the reciprocal of the conventional beta distribution) and its moments, which we also provide here.

2. Theoretical Foundations

2.1. The Stream Length Duration Curve (SLDC)

In this section we briefly review the SLDC framework of Botter and Durighetto (2020), highlighting potential extensions and refinements. For the sake of clarity and consistency, we define a stream segment as a stream section bounded by *nodes* occurring at meaningful hydrologic locations, such as sensor sites, confluences, splits, sources, and sinks (Dodds & Rothman, 2000).

Let $\boldsymbol{X}$ be a series of m Bernoulli random variables, $X_1, X_2, \ldots, X_m$ representing surface water presence or absence at segments in a stream network at the same point in time. Then, for the kth segment, $k = 1, 2, 3, \ldots, m$, we have:

$$f(x_k) = p_k^{x_k}(1 - p_k)^{1-x_k} \tag{1}$$

where p_k is the probability that the kth segment is wet, and

$$x_k = \begin{cases} 1 & \text{if the stream segment is wet} \\ 0 & \text{if the stream segment is dry} \end{cases}. \tag{2}$$

The mean and variance of X_k are

$$E(X_k) = p_k, \text{and} \tag{3}$$

$$\text{Var}(X_k) = (1 - p_k)p_k. \tag{4}$$

Jointly, $\boldsymbol{X}$ is a multivariate Bernoulli random variable, with PDF (Dai et al., 2013):

$$f(\boldsymbol{x}) = p_{0,0,\ldots,0}^{\prod_{k=1}^{m}(1-x_k)} p_{1,0,\ldots,0}^{x_1 \prod_{k=2}^{m}(1-x_k)} p_{0,1,\ldots,0}^{(1-x_1)x_2 \prod_{k=3}^{m}(1-x_k)} \cdots p_{1,1,\ldots,1}^{\prod_{k=1}^{m} x_k}. \tag{5}$$

where $p_{abc..z}$ is the joint probability of $X_1 = a$, $X_2 = b$, $X_3 = c, \ldots, X_m = z$, and $\boldsymbol{x} = (x_1, x_2, \ldots x_m)$ is a realization of $\boldsymbol{X}$.

Let Δl be a vector of individual stream lengths for the full set of stream segments: $\Delta l = \Delta l_1, \Delta l_2, \ldots, \Delta l_m$, corresponding to binary surface water presence/absence outcomes in $\boldsymbol{X}$. Then, the dot product (sum of element-wise vector products), is a random variable, L, representing wetted stream network length:

$$L = \boldsymbol{X} \bullet \Delta l. \tag{6}$$

The resulting mean wetted stream network length is

$$E(L) = \sum_{k=1}^{m} p_k \Delta l_k, \tag{7}$$

and the wetted stream network length variance is

$$\text{Var}(L) = \sum_{i=1}^{m} \sum_{j=1}^{m} \text{Cov}(L_i, L_j). \tag{8}$$

where $L_i = X_i \Delta l_i$, $L_j = X_j \Delta l_j$, and $\text{Cov}(L_i, L_j)$ denotes the covariance between stream lengths L_i and L_j. Note that for $k = i = j$, $\text{Cov}(L_k, L_k)$ is the kth segment variance, $\text{Var}(L_k) = \Delta l_k^2[p_k(1 - p_k)]$. This term will be the kth diagonal

entry in the variance covariance matrix for L, denoted Σ_L. We refer to the approach defined in Equations 1–8 as *Bernoulli stream length* due to its reliance on multivariate Bernoulli random variables. That is, Equation 7 denotes the mean Bernoulli stream network length, or the average length of the stream network that is wet, and Equation 8 represents the variance of Bernoulli stream network length.

Following Botter and Durighetto (2020), we recommend that all $p_k s$ be estimated using arithmetic means: $\hat{p}_k = n^{-1} \sum \boldsymbol{x}_k$, where $\boldsymbol{x}_k$ denotes Bernoulli surface water presence/absence data, taken over n trials, from the k th segment. Thus, we use $\boldsymbol{X}$ to represent a multivariate Bernoulli random variable describing the presence/absence of surface water across the m segments *in space*, that is, $\boldsymbol{X} = (X_1, X_2, \ldots X_m)$, $\boldsymbol{x}$ to represent a realization of $\boldsymbol{X}$ at one particular moment in time $\boldsymbol{x} = (x_1, x_2, \ldots x_m)$, and $\boldsymbol{x}_k$ to denote multiple Bernoulli (i.e., binomial) outcomes from the kth segment *over time*: $\boldsymbol{x}_k = (x_{k.1}, x_{k.2}, \ldots x_{k.n})$. As noted above, for individual time events, $x_k \in \{0, 1\}$.

Entries in Σ_L can be estimated with conventional method of moments-based variance and covariance estimators (see Aho (2014)), using observed data, although correlations (standardized covariances) for any segments i and i should have the bounds (Botter & Durighetto, 2020):

$$\rho_{i,j}^{\max} = \sqrt{\frac{p_j(1-p_i)}{p_i(1-p_j)}} \leq 1 \tag{9}$$

where $p_j \leq p_i$, and

$$\rho_{i,j}^{\min} = \left(\frac{p_j p_i}{(1-p_i)i(1-p_j)} \right)^{\beta} \geq -1 \tag{10}$$

where $\beta = 1/2$ if $p_i + p_j \leq 1$ and $\beta = -1/2$ otherwise. Generalized covariance frameworks appropriate for stream networks can also be applied (see Cressie et al. (2006) and Ver Hoef et al. (2006)).

Botter and Durighetto (2020) present the distribution of L in terms of exceedance probabilities (one minus the cumulative distribution function of L). The inverse of this sigmoidal function represents the final form of the SLDC, in reflection of the widely used flow duration curve (Castellarin et al., 2004).

2.2. Communication Distance

The concept of Bernoulli stream length is informative at a network scale. However, it may be less useful for describing internodal communication and transportation of materials. For example, for spatially adjacent nodes u and v, the drying of the connecting stream segment means that the distance from u to v with respect to surface transport of flow-borne organisms and resources has become infinite, although the Bernoulli stream length for the segment is zero. To measure resource transport constraints within stream networks we propose a new metric, *communication distance*.

Following Botter and Durighetto (2020) we represent stream segment lengths using $\Delta l_k \in \{\Delta l_1, \Delta l_2, \ldots, \Delta l_m\}$, and corresponding probabilities for surface water presence as $p_k \in \{p_1, p_2, \ldots, p_m\}$. The average communication distance of the kth segment, measured in the units provided in Δl_k, is:

$$C_k = \frac{\Delta l_k}{p_k} \tag{11}$$

Because it is the product of reciprocal probability and stream length, C_k describes the average effective stream segment length (as perceived by surface water-borne components) required for passage through the kth segment, in units of measured stream length given in Δl_k.

The average network level communication distance is:

$$C = \sum_{k=1}^{m} C_k. \tag{12}$$

Given all $p_k = 1$, the network communication distance will equal the network Bernoulli stream length, which in this case will be $\sum_{k=1}^{m} \Delta l_k$. Given any $p_k = 0$, the network communication distance becomes $\propto$. Clearly, however, to

be correctly defined as stream segments in a network, $\forall p_k > 0$ over an extended time span, making $C < \propto$. Thus, for $0 < p_k \leq 1$, the network average communication distance will be $\sum_{k=1}^{m} \Delta l_k \leq C < \propto$.

In non-perennial streams, C_k can be viewed as the average effective length of the kth segment after acknowledging intermittency. We propose that potential bottleneck locations in a network can be identified by examining differences between the segment wetted instream length (the communication distance of the segment if it were perennial) and the mean communication distance for the segment. This difference measures the increased average distance required for segment travel as a result of intermittency.

Bernoulli stream length and communication distance both provide potentially useful summaries of non-perennial stream length that acknowledge random variability in the presence of water at stream segments. However, like other current probabilistic considerations of non-perennial streams, the parameter p_k is defined (in the context of Equation 1) to be a fixed numeric constant, as is conventional under the frequentist paradigm. Below we consider formal Bayesian approaches for modeling the probability of the presence of surface water in non-perennial streams as a random variable.

2.3. Modeling the Probability of Surface Water—Bayesian Extensions to Bernoulli Stream Length and Communication Distance

2.3.1. The Probability of Surface Water as a Random Variable

Several approaches can be used to represent the probability of surface water presence at the kth segment as a random variable, θ_k, over some user-defined timescale, for example, daily, weekly, monthly, seasonally, annually. Because it is highly modifiable and bounded by [0,1], a widely used distributional model for probability is the beta PDF. If $\theta_k \sim \text{BETA}(\alpha,\beta)$, where $\alpha,\beta > 0$, and $\theta_k \in [0,1]$, it will have the PDF:

$$f(\theta_k) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta_k^{\alpha-1}(1-\theta_k)^{\beta-1}. \tag{13}$$

Other PDFs with [0,1] bounds include the triangular distribution, the two-sided power distribution, and the generalized trapezoidal distribution (see Kotz and van Dorp (2004)). In principle, any of these PDFs can replace beta distributions in analyses described here. These alternatives, however, are not members of the exponential family of distributions (Pitman, 1936), and may be non-differentiable, limiting their straightforward applicability in analyses. For example, they cannot serve as conjugate priors in Bayesian analyses (see below). A comparison of the usefulness of a large number of strictly bounded PDFs for modeling the probability of stream surface water presence is currently under development by the first author.

It is possible to define θ_k as a beta random variable with a mean defined to be some stipulated or estimated prior probability of surface water presence, $\tilde{p}_k$. Among other possibilities, the quantity $\tilde{p}_k$ can be based on pilot data sets, or surveys, or existing maps expressing outcomes from surficial water models. Specifically, let

$$\theta_k \sim \text{BETA}(\alpha, \beta = t\alpha) \tag{14}$$

for some $\alpha > 0$, where $\frac{1}{1+t}$ is equal to the stipulated prior probability, $\tilde{p}_k$, then $t = \frac{1-\tilde{p}_k}{\tilde{p}_k}$, and $E(\theta_k) = \frac{\alpha}{\alpha+\beta} = \tilde{p}_k$.

2.3.2. The Posterior Distribution of the Probability of Surface Water

Given background from Section 2.3.1, we now consider the random variable θ_k within the formal Bayesian framework:

$$f(\theta_k|\boldsymbol{x}_k) \propto f(\boldsymbol{x}_k|\theta_k)f(\theta_k) \tag{15}$$

where $f(\theta_k|\boldsymbol{x}_k)$ is the posterior density function for the probability of surface water at the kth stream segment given n observed binary presence/absence outcomes from the kth segment, where, as before, $\boldsymbol{x}k = (xk,1,xk,2,\ldots xk,n)$, $f(\boldsymbol{x}_k|\theta_k)$ is the likelihood function for the kth segment, and $f(\theta_k)$ is the kth prior density function.

A Bayesian model for θ_k, whose current data are sums of n Bernoulli trials for stream segment water presence/absence (i.e., Equation 2), would use a binomial likelihood function, resulting in, $\boldsymbol{x}_k|\theta_k \sim \text{BIN}(n,\theta_k)$. Under this framework, a beta distribution is often employed as a prior for θ_k; that is, $\theta_k \sim \text{BETA}(\alpha,\beta)$. This is because the

beta PDF is the only possible conjugate prior for a binomial likelihood function. Conjugacy results in the posterior and prior being members of the same distributional family. Thus, for the current case, both would be beta distributions. Conjugacy is useful in Bayesian applications because the prior is interpretable as additional data, and the resulting posterior distribution will have a known parametric form (Gelman et al., 2014, p. 34). The latter characteristic allows straightforward summarization of the posterior, and diminishes the need for complex numerical procedures, including Markov Chain Monte Carlo simulation. Beta priors allow a wide variety of both non-informative and informative frameworks for θ_k, as described below.

Non-informative priors, often called flat or diffuse priors, express only general information about a random variable under consideration, with the goal of "letting the data speak for themselves" (Gelman et al., 2014). Conventional non-informative beta priors include BETA(1,1), BETA(0.5, 0.5), that is, the Jeffreys prior (Jeffreys, 1946), and BETA (0, 0). All three distributions attribute equal degrees of belief to wet and dry stream outcomes. Indeed, BETA(1,1) is equivalent to a continuous uniform distribution in [0,1] and will give equal densities (of one) to all possible probabilities of surface water presence. The three distributions, however, have different prior effective sample sizes.

The prior effective sample size—a characteristic of all prior distributions—defines the effect of the prior on the posterior compared to the current data, relative to the current data sample size, n. The effective sample size for a beta prior is the sum of its hyperparameters, α and β (Morita et al., 2008). Thus, the prior distribution, BETA(0, 0), will have an effective sample size of zero, and its application will result in a beta posterior whose mean will equal the sample mean of current binomial data. Note that, to obtain a proper posterior (one with a finite integral), use of BETA(0, 0) requires that surface water presence and surface water absence outcomes are both observed at least once in current data. The prior distribution BETA(1,1) would weight current data relative to the prior by a factor of $n/2$, whereas BETA(0.5, 0.5) would weight current data relative to the prior by a factor of $n/1$. Thus, relative confidence in current data and prior data can be used to guide the parametrization of prior distributions.

Informative prior distributions should generate reasonable outcomes for a random variable under consideration, based on knowledge (and uncertainty) concerning that variable (Gelman et al., 2014). Informative beta priors can be specified, depending on the availability and quality of prior information extraneous to current data used in the likelihood. For example, as noted above, an informative beta prior can be constructed in which $E(\theta_k)$ equals a prior designation for the probability of surface water presence at the kth segment, $\tilde{p}_k$ (Equation 14). As noted in the previous paragraph, beta hyperparameters can be further modified to ensure that the prior has an appropriate effective sample size. Our approach for these implementations is fully described in Section 3.2.2.

2.3.3. The Posterior Distribution of Bernoulli Stream Length

Given beta priors and binomial likelihoods, the posterior density function of the probability of surface water presence at the kth segment will have the form $\theta_k|\boldsymbol{x}_k \sim \text{BETA}(\alpha + \sum \boldsymbol{x}_k, \beta + n - \sum \boldsymbol{x}_k)$ where α and β are the hyperparameter values defined for the beta prior distribution. Under linear transformation, the posterior distribution for the average Bernoulli length of the kth segment can be obtained by multiplying the $\theta_k|\boldsymbol{x}_k$ posterior by the constant Δl_k. That is,

$$E(L_k) = (\theta_k|\boldsymbol{x}_k) \cdot \Delta l_k \tag{16}$$

The posterior distribution of average Bernoulli stream lengths for the entire network can be obtained by taking the sum of the product $E(\theta_k|\boldsymbol{x}_k) \cdot \Delta l_k$, across all segments (cf. Eq. 5 in Botter and Durighetto (2020)). That is,

$$E(L) = \sum_{k=1}^{m} E(\theta_k|\boldsymbol{x}_k) \cdot \Delta l_k \tag{17}$$

Determining the posterior distributions of communication distance at the segment or network scale, $E(C_k)$ and $E(C)$, respectively, is less straightforward. This is because the derivation requires multiplication of the kth stream length by the reciprocal (multiplicative inverse) of the kth beta posterior.

2.3.4. The Inverse-Beta Distribution

If θ_k follows a beta distribution, then θ_k^{-1} will follow an *inverse-beta distribution*. This PDF has not been previously derived, although as a practical matter it is straightforward to obtain inverse-beta outcomes from existing computer algorithms (e.g., `1/rbeta()` in the **R** computational environment). The inverse beta PDF is distinct

from the *beta prime distribution*, also called the *inverted beta distribution*, which is used to represent the distribution of odds, that is, a probability divided by its complement (Johnson et al., 1995).

Let $\theta_k \sim \text{BETA}(\alpha, \beta)$ with $\alpha, \beta > 0$, then $\theta_k^{-1} \sim \text{BETA}^{-1}(\alpha, \beta)$ with PDF:

$$f\left(\theta_k^{-1}\right) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{1}{\theta_k^{-1}}\right)^{\alpha+1}\left(1-\frac{1}{\theta_k^{-1}}\right)^{\beta-1}, \tag{18}$$

with mean, for $\alpha > 1$,

$$E\left(\theta_k^{-1}\right) = \frac{\alpha+\beta-1}{\alpha-1}, \tag{19}$$

and variance, for $\alpha > 2$,

$$\text{Var}\left(\theta_k^{-1}\right) = \frac{\beta(\alpha+\beta-1)}{(\alpha-1)^2(\alpha-2)}. \tag{20}$$

As suggested above, the inverse beta distribution can be used to represent distributions of reciprocal probabilities which will occur in $[1, \infty)$, given probabilities in $[1, 0)$. Reciprocal probabilities are useful for measuring the *rarity* of outcomes. Specifically, the reciprocal probability, r, for an outcome A, indicates that there is a 1 in r chance that A will occur. For instance, if the probability of surface water presence a stream segment is 0.01, then one would expect that surface water will occur in 1 of 100 cases, because $r = 1/0.01 = 100$, and that surface water at the segment would be 100 times rarer than at a perennial segment. Under this framework, Equations 19 and 20 represent the mean level of surface water rarity and the variance of surface water rarity for the kth stream segment, respectively. Mathematical derivations of the inverse-beta distribution and its moments are given in Section S1 in Supporting Information S1.

2.3.5. The Posterior Distribution of Average Communication Distance

Let $(\theta_k|\boldsymbol{x}_k)^{-1}$ be an inverse beta posterior distribution representing the reciprocal probability of surface water presence at the kth stream segment, then the posterior mean communication distance for the k th segment is:

$$E(C_k) = E\left[(\theta_k|\boldsymbol{x}_k)^{-1}\right]\Delta l_k, \tag{21}$$

and posterior communication distance variance of the kth segment is:

$$\text{Var}(C_k) = \text{Var}\left[(\theta_k|\boldsymbol{x}_k)^{-1}\right]\Delta l_k^2. \tag{22}$$

In this Bayesian context, $E(C_k)$ can be viewed as the average effective stream length of the kth segment, based on both current data and prior information. Thus, $E(C_k)$ can be compared to the segment's wetted instream length (the actual physical length of the segment) to obtain an inductive measure of increased mean effective stream length due to intermittency.

The posterior average network communication distance is:

$$E(C) = \sum_{k=1}^{m} E\left[(\theta_k|\boldsymbol{x}_k)^{-1}\right]\Delta l_k. \tag{23}$$

and the posterior communication distance variance of the entire network is:

$$\text{Var}(C) = \sum_{i=1}^{m}\sum_{i=1}^{m}\text{Cov}(C_i, C_j) \tag{24}$$

where $\text{Cov}(C_i,C_j)$ denotes the covariance between communication distances C_i and C_j. For $k = i = j$, $\text{Cov}(C_k,C_k)$ is the kth marginal variance, $\text{Var}(C_k)$. We now demonstrate the use of these metrics with an example from the north central Rocky Mountains, USA.

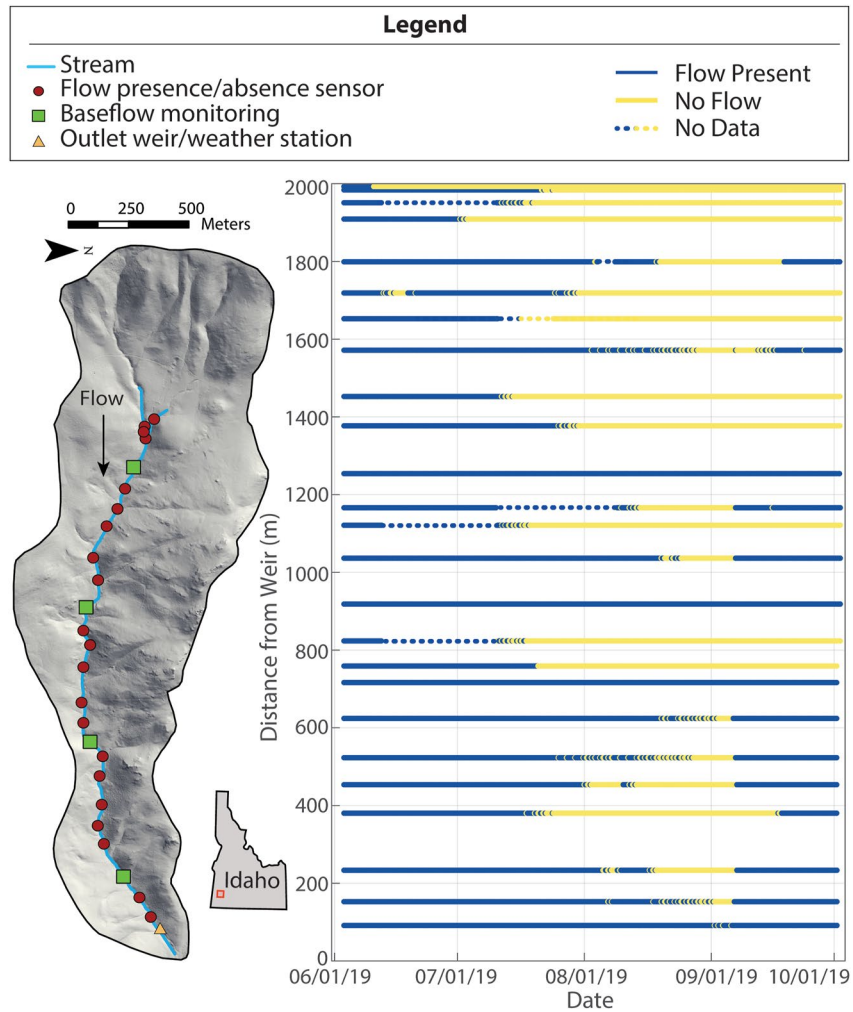

Figure 1. Instrumentation of the Murphy Creek sub-watershed in 2019 (outlet coordinates: 43.25607°N, −116.8186°W) and summary of surface water presence at each sensor. Note that missing data points occurred for some nodes over the span of sampling.

3. Materials and Methods

3.1. Field Site and Field Methods

We derived Bernoulli stream length and communication distance summaries for Murphy Creek, a simple drainage system within the larger Reynolds Creek experimental watershed in the Owyhee Mountains of southwestern Idaho, USA (see Warix et al. (2021)). Measures of surface water presence were made at 25 nodes, corresponding to 24 stream segments, every 15 min from 3 June 2019 to 2 September 2019, resulting in 11,623 repeated measures for each node (Figure 1). Missing data points, which occurred to varying levels at 16 nodes (Figure 1), constituted less than 5% of all possible time series measures. In agreement with our definition of stream segments, we designated additional (un-instrumented) nodes at the outlet and at two stream sources, resulting in a total of 27 nodes and 26 segments. At 21 nodes, surface water presence was measured with sensors (Onset HOBO Pendant/Light 64 K Datalogger sensors (UA002-64; Figure 1) that were modified to detect resistivity (Chapin et al., 2014). The resistivity sensors were placed in the deepest part of the channel and installed so that the two pole electrodes were touching the stream bed, allowing detection of the presence or absence of water at the lowest of flow conditions. At the other four instrumented sites, water levels (Onset Hobologger, U-20) and specific conductance (Onset Hobologger, U-24) were measured at 15-min intervals for baseflow monitoring, also allowing detection of the absence of stream surface water (Figure 1). Additional details can be found in Warix et al. (2021).

3.2. Statistical Methods

We created inferential models of Bernoulli stream length and network communication distance at Murphy Creek based on the entirety of the sampling period (3 June 2019–2 September 2019), and on three seasonal subsets: spring (3 June 2019–10 July 2019), summer (11 July 2019–14 August 2019), and fall (15 August 2019–2 September 2019). Seasonal cutoffs were established at dates representing approximate change points in precipitation and temperature based on long-term (1964–1996) climate data for the study area (Hanson et al., 2001), for days of the year corresponding to the 2019 sampling period (early June to early October).

3.2.1. Estimating Stream Segment Surface Water Presence/Absence

The translation of point measurements of surface flow presence/absence at sensor locations to stream segment surface flow presence/absence is a potentially difficult and contentious process. We can consider four general approaches. Under approaches one and two a sensor is defined as a node of its corresponding segment and represents singly, the stream segment surface flow presence/absence between itself and the node immediately upstream (approach 1) or downstream (approach 2). Under approach three, a pair of neighboring sensors, which are placed at the upstream and downstream ends of a stream segment, serve as the segment nodes, and surface flow presence/absence responses from both sensors are considered when defining segment surface flow presence/absence. Under approach four, a sensor is not assumed to be a segment node, but instead fully represents a segment that is bounded by unmeasured nodes halfway between each sensor (e.g., Botter & Durighetto, 2020). Approaches one, two, and four are straightforward, allow clear translation to a binary Bernoulli process, and produce data that can be used *directly* in Bayesian approaches described in Section 3.2.2.

Despite these benefits, however, we used approach three to estimate stream segment surface flow presence/absence at Murphy Creek. We based our decision on two factors. First, unlike approach four, approach three reflected our definition of stream segments (Section 2.1), and provided a clear delineation of the extent of stream segments. Second, unlike approaches one and two, approach three allowed multiple estimation points for average surface water presence, based on segment locations with the greatest potential for spatial independence (Cressie et al., 2006). Specifically, for the kth segment with bounding nodes u and v, for the ith time frame, $i = 1,2,3,\ldots,n$, we applied the following rule:

$$x_{k,i} = \begin{cases} 1.0, \text{both } u \text{ and } v \text{ wet} \\ 0.0, \text{both } u \text{ and } v \text{ dry} \\ 0.5, \text{only one of } u \text{ or } v \text{ wet} \end{cases}. \quad (25)$$

Equation 25 can be viewed as a coarse estimator of the average probability of surface water along kth segment, based on segment endpoints. Marginal (individual segment) probabilities of surface water presence and covariances among segments were both estimated using segment outcomes from Equation 25. Exceptions were the three segments associated with input and sink locations, whose extremal nodes were not instrumented (Figure 1) and time points with a missing datum at a single bounding node (8% of total segment observations). In this unusual case, surface water outcomes were based on water presence/absence outcomes at a single node (e.g., Botter & Durighetto, 2020). Situations with missing data at both bounding nodes, preventing application of Equation 25, occurred in less than 0.2% of cases.

We used the **R** package *mipfp* (Barthélemy & Suesse, 2018) to generate multivariate Bernoulli trials for water presence at segments based on estimated marginal segment probabilities of water presence and inter-segment covariances. Specifically, we generated 1,000 random multivariate Bernoulli trials, each made up of $m = 26$ potentially correlated binary outcomes, representing the simultaneous presence or absence of surface stream flow at each of the 26 designated Murphy Creek stream segments, at a particular time. We applied this approach based on estimates from data over the entire sampling period, and for separate data subsets representing spring, summer, and fall. Our simulation approach addressed the issue of potential non-binary outcomes in Equation 25, and the fact that surface water presence at segments is generally positively correlated in space. Specifically, the approach allowed the generation of large surface water presence or absence data sets for segments, made up of temporally independent (random) samples representative of *particular spans of time*, that is, the entire sampling period, spring, summer, and fall, based on the estimated marginal probabilities for stream segment

presence and the estimated spatial dependencies of segments during those periods of time. We randomly sampled with replacement with the sample size $n = 10$ from the collections of random multivariate Bernoulli outcomes 10,000 times, for the entire sampling period, and for each season, and used the numbers of successes (i.e., the number "surface water present" outcomes) from those 10 trial simulations as multivariate binomial outcomes in subsequent analyses. Our stipulation, $n = 10$ was based on general guidelines given by Gotelli and Ellison (2004, p. 150) for detection of appreciable effect sizes in ecological studies. Importantly, the reliance of a beta posterior distribution on the current sample size (compared to the prior) can be further modified with a user-defined weighting constant, w, which we define in Section 3.2.2, and consider further in Section 4.4).

3.2.2. Bayesian Methodology

Under a Bayesian framework, simulated binomial data outcomes obtained from *mipfp* algorithms were coupled with beta priors to obtain beta posteriors. Informative beta priors were defined to have a mean corresponding to the predicted probability of surface water presence from the Probability of Streamflow Permanence model (PROSPER; Jaeger et al., 2019), as reported for Murphy Creek stream segments by the United States Geological Survey (USGS) StreamStats web-based application (USGS, 2016).

The parameterization for our priors was $\theta_k \sim \text{BETA}(\alpha,\alpha t)$, for $t = \frac{1-\tilde{p}_k}{\tilde{p}_k}$, where $\tilde{p}_k$ was the PROSPER probability of surface water presence for the kth stream segment, resulting in $E(\theta_k) = \tilde{p}_k$ (Equation 14). We also required priors to have an effective sample size that was a fixed proportion, w, of n. That is, we let $\alpha + \beta = w \cdot n$. Because, under our parameterization, $\beta = \alpha \cdot t$, our weighting became $\alpha + \alpha \cdot t = w \cdot n$, resulting in the hyperparameters:

$$\alpha = w \cdot n \cdot \tilde{p}_k, \text{and}$$

$$\beta = w \cdot n(1 - \tilde{p}_k). \tag{26}$$

Thus, the posterior distribution for the probability of surface water presence at the kth segment had the form:

$$\theta_k|\boldsymbol{x}_k \sim \text{BETA}\left(w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k, w \cdot n(1 - \tilde{p}_k) + n - \sum \boldsymbol{x}_k\right), \tag{27}$$

with mean

$$E(\theta_k|\boldsymbol{x}_k) = \frac{w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k}{w \cdot n \cdot \tilde{p}_k + w \cdot n(1 - \tilde{p}_k) + n}, \tag{28}$$

and variance:

$$\text{Var}(\theta_k|\boldsymbol{x}_k) = \frac{w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k}{[w \cdot n \cdot \tilde{p}_k + w \cdot n(1 - \tilde{p}_k) + n]^2} \cdot \frac{w \cdot n \cdot (1 - \tilde{p}_k) + n - \sum \boldsymbol{x}_k}{w \cdot n \cdot \tilde{p}_k + w \cdot n(1 - \tilde{p}_k) + n + 1} \tag{29}$$

The sum of the products of the beta posterior means and respective segment lengths across all m segments, $\sum_{k=1}^{m} E(\theta_k|\boldsymbol{x}_k) \cdot \Delta l_k$, defined a posterior distribution outcome for mean Bernoulli stream length, $E(L)$ (Equation 17).

The inverse-beta posterior distribution for reciprocal probability of surface water presence at the kth segment was:

$$(\theta_k|\boldsymbol{x}_k)^{-1} \sim \text{BETA}^{-1}\left(w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k, w \cdot n(1 - \tilde{p}_k) + n - \sum \boldsymbol{x}_k\right). \tag{30}$$

Thus, the mean of the kth inverse-beta posterior was:

$$E\left[(\theta_k|\boldsymbol{x}_k)^{-1}\right] = \frac{w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k + w \cdot n(1 - \tilde{p}_k) + n - \sum \boldsymbol{x}_k - 1}{w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k - 1} \tag{31}$$

and the variance of the kth inverse-beta posterior was:

$$\text{Var}\left[(\theta_k|\boldsymbol{x}_k)^{-1}\right] = \frac{w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k + w \cdot n \cdot (1 - \tilde{p}_k) + n - \sum \boldsymbol{x}_k - 1}{w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k - 1} \cdot \frac{w \cdot n \cdot (1 - \tilde{p}_k) + n - \sum \boldsymbol{x}_k}{\left[w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k - 1\right]\left[w \cdot n \cdot \tilde{p}_k + \sum \boldsymbol{x}_k - 2\right]} \tag{32}$$

The sum of the products of the inverse-beta posterior means and respective segment lengths, $\sum_{k=1}^{m} E\left[(\theta_k|\boldsymbol{x}_k)^{-1}\right] \cdot \Delta l_k$, was used to define a posterior outcome for $E(C)$ (Equation 23), allowing expansion to a network scale.

Designation of w is an important decision. For our application, we weighted the beta priors so that they would have 50% of the weight of the sample data and, as noted above, let the current sample size be 10 random draws from a multivariate Bernoulli distribution simulated from 2019 field data (cf. Gotelli & Ellison, 2004). That is, for Equations 26–32 above, we let $w = 0.5$, $n = 10$, and, thus $\sum \boldsymbol{x}_k \in \{0,1,2,\ldots,10\}$.

Our choice of $w = 0.5$ was driven in part by parameter constraints of the inverse-beta posteriors used to calculate average seasonal communication distances. In particular, infinitely large means for inverse-beta posteriors occurred when zeroes (surface water absences) occurred for all 10 random Bernoulli observations for a segment during drier seasons when the corresponding PROSPER probability of segment surface water presence was 0.21 (the minimum PROSPER probability for the catchment) and the weighting level was $w \leq 0.47$ (Section S2 in Supporting Information S1). Larger current data sample sizes, n, and/or larger prior mean probabilities of stream presence allow greater flexibility for prior weight choices when no surface water is observed in current data (see Section S2 in Supporting Information S1). Undefined posterior means will not occur for Bayesian extensions of Bernoulli stream length for any $w > 0$. Nonetheless, the same prior weight ($w = 0.5$) was used for both communication distance and Bernoulli stream length to facilitate comparison of results under these two approaches.

3.3. Software

The **R** statistical environment (R Core Team, 2022) was used for all analyses with reliance on the package *streamDAG* (Aho, 2023; Aho et al., 2023a, 2023b), which allowed derivation of stream segment presence/absence outcomes from node data, and straightforward computation of Bernoulli stream length and communication distance posteriors, and the package *mipfp* (Barthélemy & Suesse, 2018) for simulation of multivariate Bernoulli outcomes. Several spatial and graphics packages, including *sf* (Pebesma, 2018), *ggspatial* (Dunnington, 2021), *cowplot* (Wilke, 2020), *ggplot2* (Wickham, 2016), and *gridGraphics* (Murrell & Wen, 2020) were used to visualize the results.

4. Results

4.1. Comparison of Prior, Current, and Posterior Probabilities of Surface Water Presence

Estimates for the probability of surface water presence based on data from the entirety of the 2019 field season, $\hat{p}_k$, differed considerably from designated prior distribution means, $\tilde{p}_k$, for some stream segments. Specifically, PROSPER probabilities for surface water presence were limited to the range 0.21–0.32 (Figure 2a), whereas field observations in 2019 included reaches that were (nearly) always wet or dry (Figure 2b). By definition, posterior distributions for the probability of surface water presence (Figure 2c) were a compromise between the PROSPER priors and current (2019) observations (Aho, 2014; Gelman et al., 2014).

4.2. Quantifying Uncertainty in Reach-Scale Wet/Dry Predictions: Posterior Distributions for the Probability of Surface Water for Individual Segments

Our approach allowed summaries of both intra-segment central tendency and (for the first time) the variability in the probability of surface water presence at segments (Figure 3). Posterior beta distributions of segments closer to the outlet generally had larger mean values, indicating high average probabilities of surface water presence (Figures 3a and 3b). Segments near the top and bottom of the network had smaller posterior variances for different reasons (Figure 3b). Posterior distributions variances of segments near the outlet, for example, $\overline{M91\ OUT}$, were smaller because surface water was generally present at these locations, whereas segment posteriors near inputs, for example, $\overline{INSM1993}$, had smaller variances because surface water was generally absent (Figure 3b). Critically, posterior distributions for segments near the middle of the stream were platykurtic with relatively large variances.

4.3. Rarity of Surface Water at Segments: Posterior Distributions for the Reciprocal Probability of Surface Water Presence for Individual Segments

Distributional differences between segments were even more pronounced when examining the rarity of flows. The distinctiveness of rarity responses was evident in the dramatic segment to segment variation in the symmetry

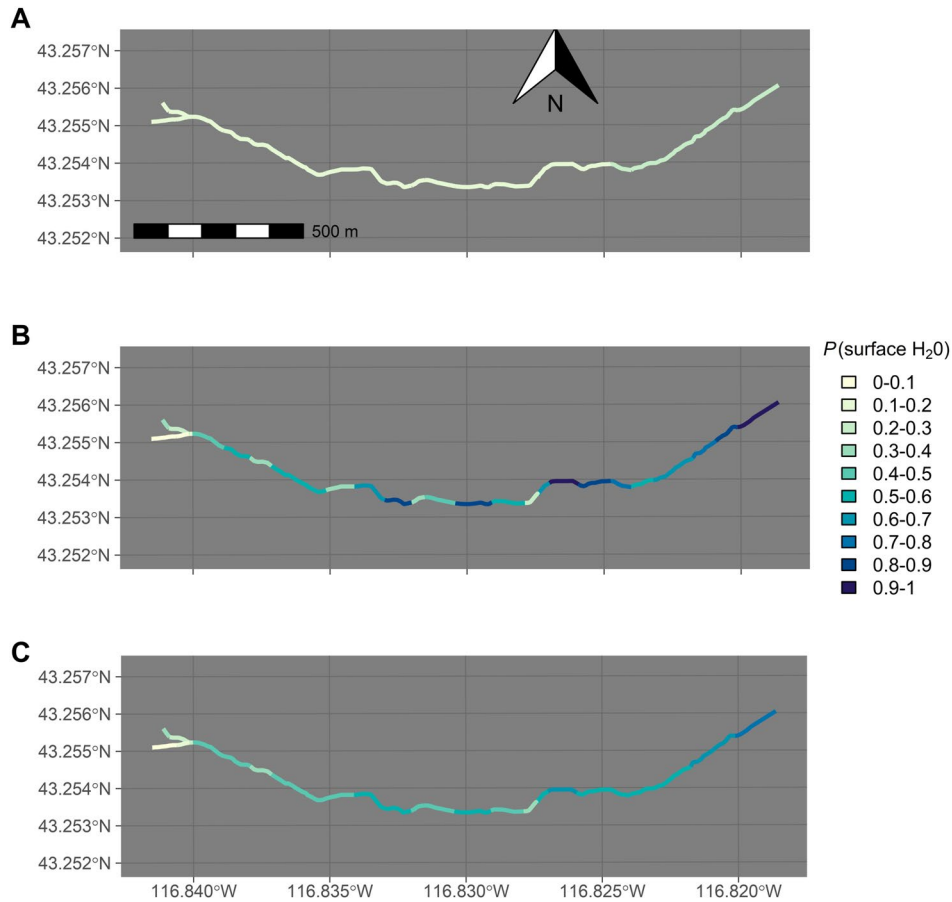

Figure 2. Spatially distributed probabilities of stream surface water presence at Murphy Creek. Flow proceeds from left to right (W to E) (see Figure 1). Panel (a) depicts stream flow presence probabilities from the USGS-PROSPER model, which were used to define means for beta distribution priors in Bayesian analyses. Panel (b) shows probabilities based on surface water data for the entire 2019 sampling period: 3 June 2019–2 September 2019. Panel (c), is based on information from both panels (a, b). Specifically, beta priors for segments whose means were predicted by the USGS-PROSPER model were coupled with binomial likelihoods based on current (simulated) data from multivariate Bernoulli distributions whose covariance parameters were estimated from data from the entire 2019 sampling period.

and kurtosis of inverse-beta posteriors (Figure 4). Segment distributions near the top of the network, for example, $\overrightarrow{INS\,M1993}$, $\overrightarrow{M1993\,M1951}$, were platykurtic with relatively large means, indicating that, on average, surface water was rare, whereas segment distributions near the outlet were leptokurtic, with much of the probability mass near one, indicating the segment resembled a perennial stream segment.

4.4. Identifying Bottleneck Locations: Posterior Distributions of Mean Communication Distance for Individual Segments

Large differences in wetted instream length and average posterior mean communication distance were evident for segments near the top of the network, indicating strong bottlenecking propensities. This was particularly true for segment $\overrightarrow{INS\,M1993}$ whose average communication distance was more than 2,000 m larger than its wetted length (Figure 5). Locations near the outlet with uncharacteristically large increases in mean effective stream length included both $\overrightarrow{M823\,M759}$ and $\overrightarrow{M380\,M233}$ (Figure 5).

4.5. Assessing Network-Scale Effects: Comparison of Mean Bernoulli Network Length, *L*, and Network Communication Distance, *C*, Over Seasons

Seasonal comparisons of network responses revealed differences in perspectives provided by network length and communication distance. The distributions of both average stream length and average communication distance changed dramatically in the spring, summer and fall (Figure 6). Greater distinctions, however, were evident for

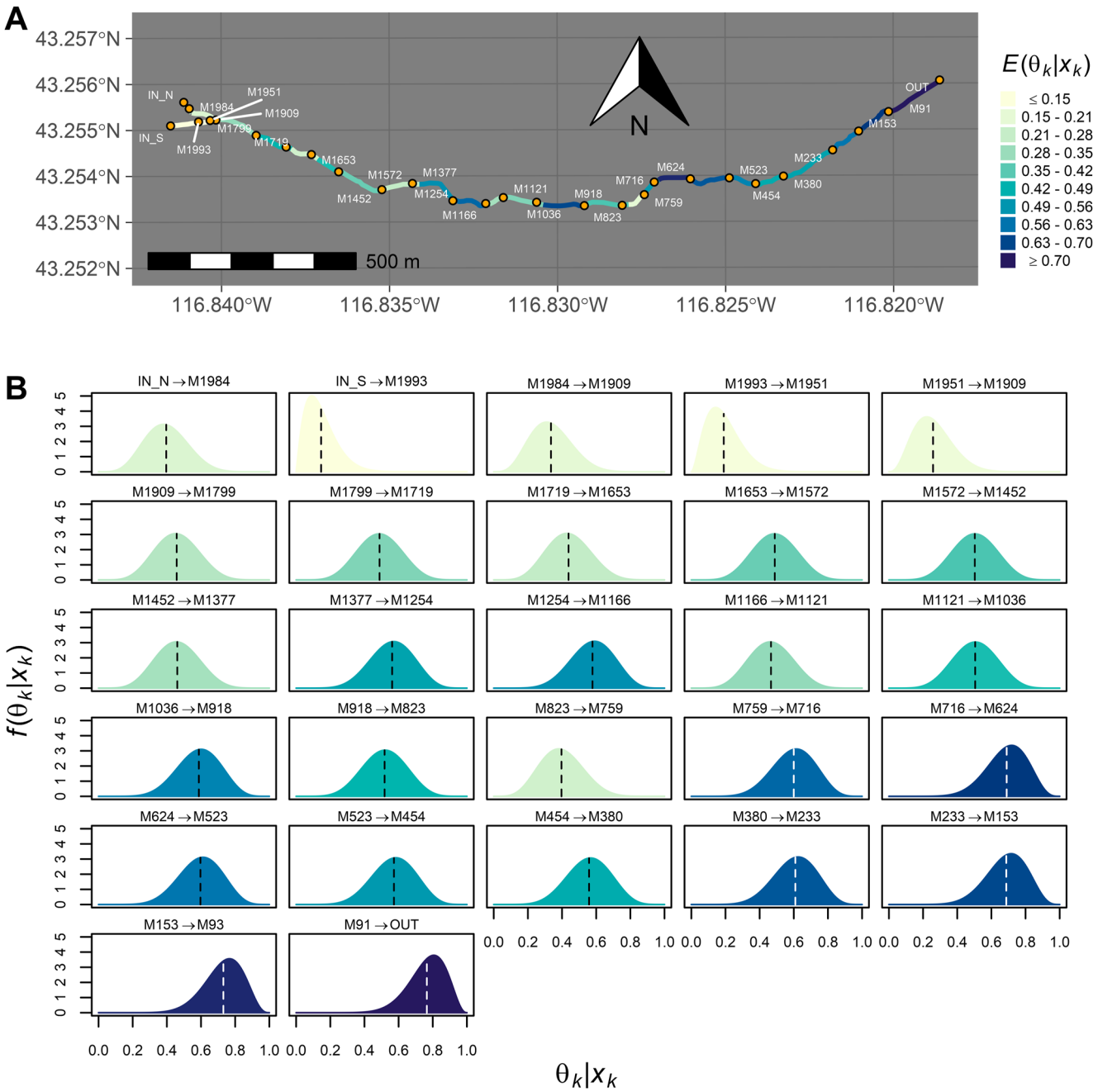

Figure 3. Posterior distributions of the *probability of surface water presence*. Summaries are based on the entire sampling season, 3 June 2019–2 September 2019. Names of segment bounding nodes correspond to meter distances from the outlet. Panel (a) locates nodes along the network. Segments are colored by their posterior distribution mean values (see key). Larger means (darker, bluer colors) indicate segments with a higher likelihood for surface water presence. Panel (b) shows beta posterior distributions for each segment. Segment posterior distributions are sorted, by row, from sources to outlet and are colored based on their mean values (see key). Posterior means of distributions are indicated with dashed lines.

the distributions of $E(C)$ compared to $E(L)$. Specifically, while the posterior distributions of $E(L)$ were symmetric across seasons (Figure 6b), posterior distributions of $E(C)$ were highly complex and asymmetric (Figure 6c). For example, the spring posterior distribution of $E(C)$ was multimodal, while the summer and fall posteriors for $E(C)$ were strongly platykurtic and negatively skewed, respectively (Figure 6c). Note that variances of seasonal distributions of mean Bernoulli network length based on current (simulated multivariate Bernoulli) data (Figure 6a) were much larger than posterior seasonal distributions of mean Bernoulli network length (Figure 6b), because of the application of reasonable informative priors for θ_k.

10.1029/2023WR034513

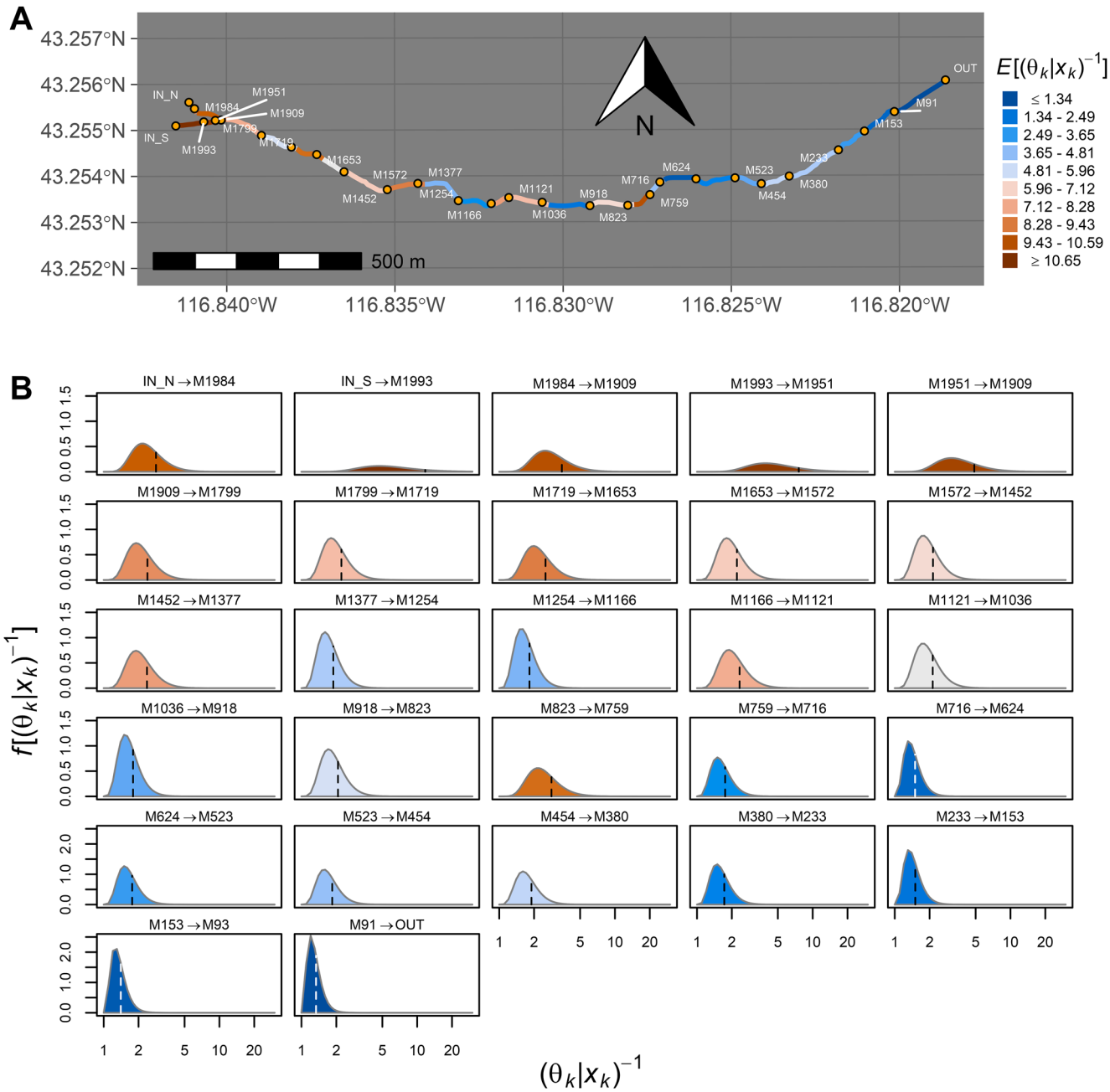

Figure 4. Posterior distributions of the *rarity of surface water presence*. Summaries are based on the entire sampling season, 3 June 2019–2 September 2019. Names of segment nodes correspond to meter distances from the outlet. Panel (a) locates nodes along the stream network. Segments are colored by their inverse beta posterior mean values (see key). Larger means (redder colors) indicate segments for which surface water is increasingly rare. Panel (b) shows inverse-beta posterior distributions for each segment. Segment posterior distributions are sorted, by row, from sources to outlet, and, again, are colored based on their mean values (see key). Posterior means of distributions are indicated with dashed lines. Note that *y*-axis limits differ for the last two rows of PDFs in panel (b) because of their leptokurtic shapes.

The highly platykurtic distribution of $E(C)$ in the summer demonstrated the possibility for good communication periods, when water was present, although the mean of the distribution of summer posterior $E(C)$ was more than 10 times as large as the wetted network length (Figure 6c). The multimodality of $E(C)$ in the spring occurred because of the strong surface flow persistence of all segments during early spring, resulting in a group of segments with smaller average communication distances, and the late spring drying of several segments, particularly $\overrightarrow{INS\,M1993}$, resulting in larger average communication distances.

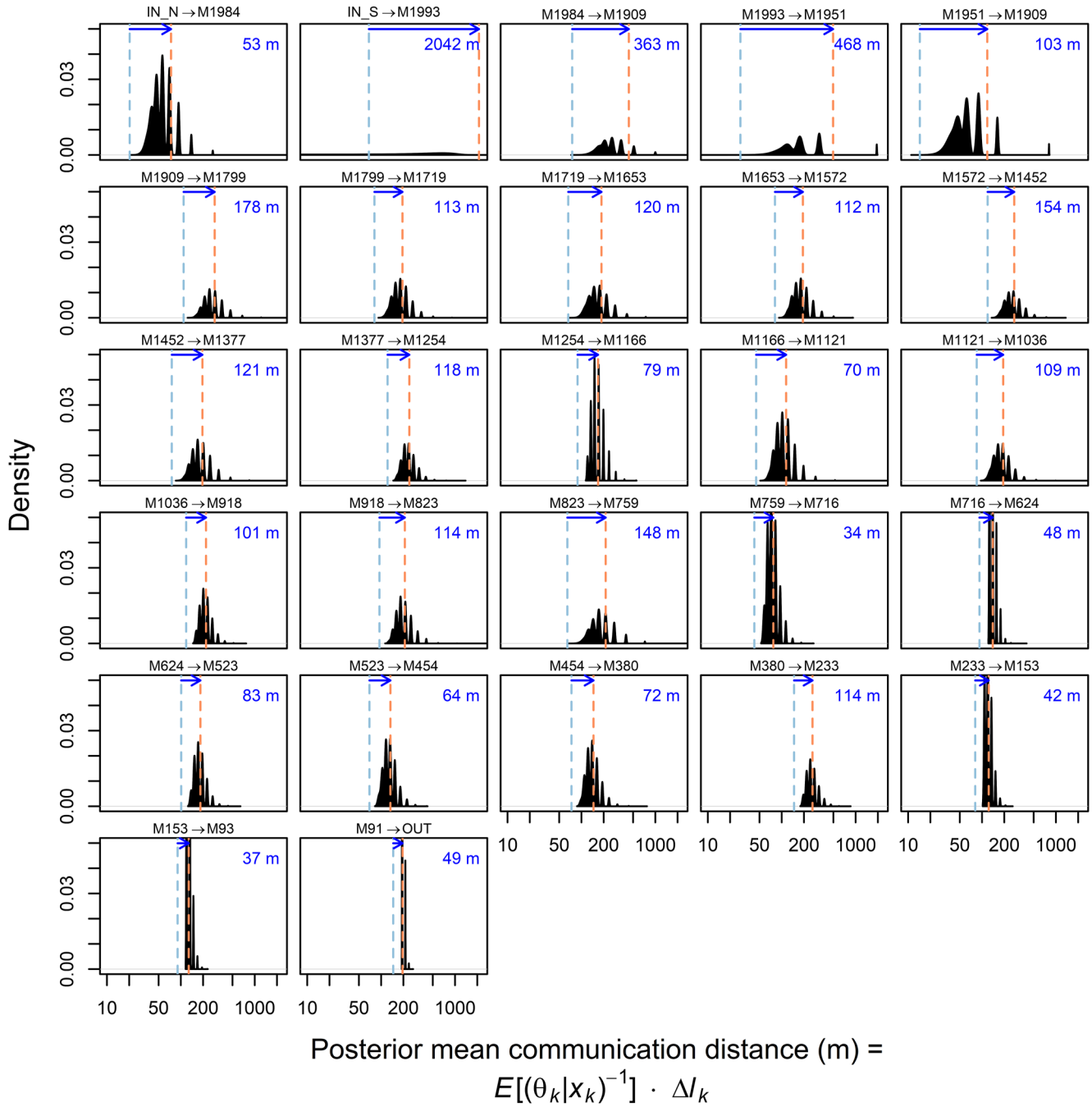

Figure 5. Distributions of posterior communication distance means for stream segments, based on the entire sampling season, 3 June 2019–2 September 2019. For each plot, the left-hand (light blue) vertical dashed line denotes the fully wetted length of the segment (communication distance if the segment were perennial), whereas the right-hand (salmon) dashed line is the average posterior mean communication distance. The difference of these values (right minus left) is given in the top-right corner of each plot. Names of segment nodes correspond to meter distances from the outlet. Segment posterior distributions are sorted, by row, from sources to outlet. Expressions of density in plots are based on a Gaussian smoothing kernel.

A fall rewet period was evident for average Bernoulli stream length, as the mean wetted extent of the fall network was longer than the mean wetted extent of the summer network (Figures 6a and 6b). This trend was not evident, however, for the average communication distance posterior distribution, as larger mean communication distances occurred in the fall compared to the summer (Figure 6c). The probability distribution of $E(C)$ in the fall appeared highly compact (Figure 6b) because of the conflation of its large range of potential outcomes (note the log-scale of the x-axis in Figure 6c), and the requirement that the area under a valid PDF be one.

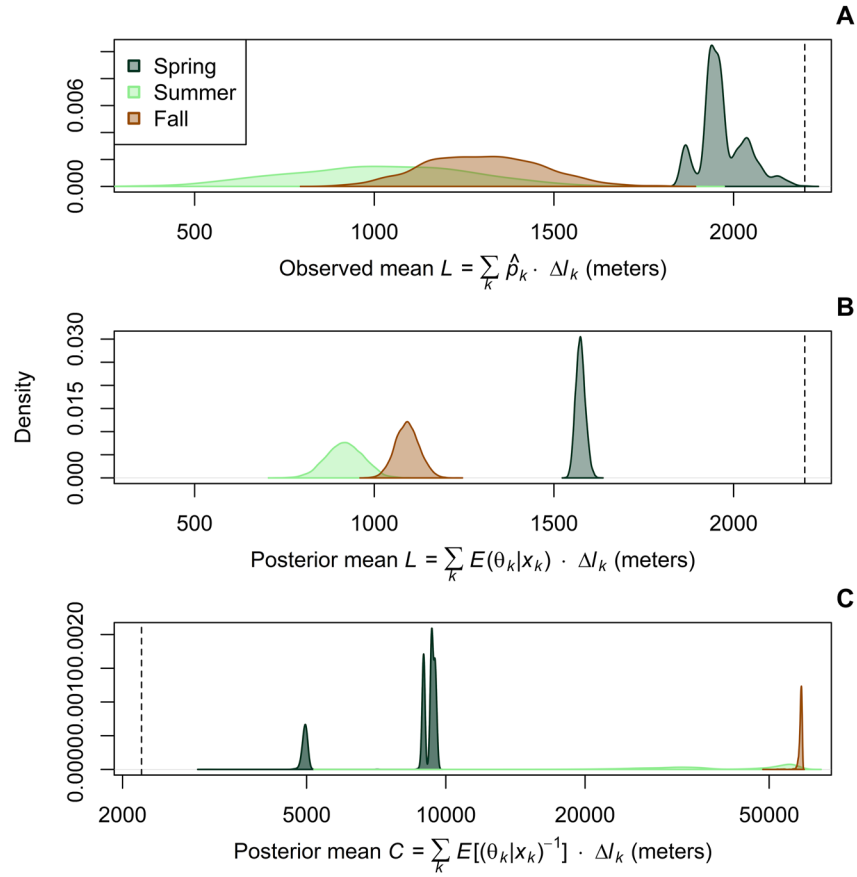

Figure 6. Seasonal distributions of (a) observed mean network length, (b) posterior mean Bernoulli stream network length, and (c) posterior mean network communication distance. All figures use random outcomes from a multivariate Bernoulli distribution with covariance parameters estimated from 2019 data (see Section 3.2). Distributions in panels (b, c) are derived under a Bayesian framework in which binomial likelihoods for segments, based on current (simulated) data, are coupled with beta priors for segments. Panel (b) distributions are the products of the means of derived posterior beta distributions for segments (Equation 17) and their corresponding stream lengths. Panel (c) distributions are the products of the means of derived inverse posterior beta distributions for segments (Equation 23) and their corresponding stream lengths. Note the log scale of the *x*-axis in panel (c). Expressions of density in plots are based on a Gaussian smoothing kernel. The fully wetted stream length of Murphy Creek is denoted with a vertical dashed line in panels (a–c).

5. Discussion

We developed measures of hydrological connectivity for non-perennial streams that: (a) allowed global (network-scale) and local (stream segment or reach-scale) perspectives concerning hydrological connectivity, (b) quantified variability in intra- and inter-segment surface flow presence probabilities, and (c) allowed the inclusion of prior information concerning probabilities of surface flow presence. Our novel contributions include Bayesian extensions of Bernoulli stream network length (Botter & Durighetto, 2020) and a new connectivity metric, communication distance.

Communication distance is a scaled (by stream segment length) measure of resistance to the passive transport of materials, for example, surface water itself, microbial organisms, and solutes, in stream networks. Because it is the product of reciprocal probability and stream length, the communication distance of a stream segment must be greater than or equal to its actual physical length. We view communication distance as a theoretical effective stream length (as perceived by surface water-borne components) required for passage through a segment. Specifically, the effective stream length of the *k*th non-perennial segment will be $1/p_k$ longer than a perennial segment of the same fully wetted length. Thus, the difference between the fully wetted length of the *k*th segment (i.e., the communication distance if the *k*th segment were perennial) and the actual average communication distance of *k*th segment can be considered a measure of increased average transport distance, as the result of intermittency (Figure 5).

As defined, communication distance considers neither flow rates nor variation in the constraints that govern the movement of particular materials in streams, as it assumes that materials of interest move freely with water. Thus, communication distance does not consider flow-driven nuances in material transport, including the fact that sediment transport is often triggered by a flow rate threshold (Pähtz et al., 2020), and microbial transport may require a delay for community establishment and growth following stream activation (Drummond et al., 2015).

Bayesian extensions of communication distance prompted the first reported derivation of the inverse-beta PDF, which can be used to represent a distribution of reciprocal probabilities of surface water presence, and directly consider the *rarity* of surface flows. This application allows inductive probabilistic consideration of important hydrological questions. For instance: "What is the probability that the average effective length of a stream network will be longer during the spring compared to summer?" Or: "What is the probability that the average effective stream length of a non-perennial segment will become more than q times as large as a comparable perennial segment?"

5.1. Inverse-Beta Distribution and Communication Distance Identify Bottleneck Locations at Segments That Are Rarely Wet

While multiplicative inverses of each other, the beta and inverse-beta distributions allow distinctive insights into non-perennial stream mechanics. For example, segments at Murphy Creek that were consistently dry, for example, $\overrightarrow{INSM1993}$, tended to have platykurtic posterior distributions of water rarity, with particularly large variances (Figure 4). On the other hand, posterior distributions of the probability of surface water presence at dry and wet segments had relatively small variances due to surface water generally being absent or present, respectively (Figure 3).

Communication distance quantifies the inhibition of surface water transport processes, due to intermittency, in units of stream segment length, potentially clarifying the effects of network drying and bottlenecking. For instance, in the Murphy Creek network, the driest segment, $\overrightarrow{INSM1993}$, had a posterior inverse-beta mean of 11.6 (Figure 4b), indicating that surface water presence at the segment would be, on average, 11.6 times rarer than a perennial segment. The average of the communication distance posterior distribution for the segment was 2,021 m longer than its wetted stream length (Figure 5). Thus, based on both current and prior information, we would conclude that the average effective stream length of the segment increased 2,021 m because of intermittency. Future hydrological research may reveal the mechanics of why particular segments in the network may act as bottlenecks.

5.2. Distinct Seasonal Variation in Active Network Length, L, and Communication Distance, C

Discrepancies in perspectives offered by network communication distance and network Bernoulli stream length were strongly evident in seasonal analyses (Figure 6). Average network Bernoulli stream lengths were longer in the fall compared to summer, indicating a fall rewet period (Figure 6b). This outcome is consistent with observations of increased fall discharge for the study region (McNamara et al., 2005). Evidence of a fall rewet period, however, was not apparent in the posterior distributions of average network communication distance. Instead, larger communication distances were more probable in the fall compared to summer. This discrepancy was due to a marked bifurcation in the behavior of stream segments in the fall. Specifically, segments near the outlet and wet spots near the stream center (e.g., $\overrightarrow{M716\ M624}$) tended to be strongly persistent ($\hat{p} \approx 1$), driving larger wetted network lengths and smaller communication distances. On the other hand, surface water at segments further from the outlet was often fully absent ($\hat{p} \approx 0$), creating potential bottlenecks. This outcome contrasted with both the summer and spring responses, but for different reasons. In the summer, communication distance posterior variances suggested that weak and strong communication outcomes had nearly equal likelihoods for most segments, because more segments varied from wet to dry during this period. In contrast, in the spring, strong persistency drove smaller effective stream lengths for all segments.

Posterior mean Bernoulli stream lengths across all seasons did not approach the full wetted length of Murphy Creek (Figure 6b), due to the moderating effect of conservative prior probabilities from the USGS-PROSPER model (Figure 2a). The PROSPER model is intended to represent the annual probability of stream segment presence, which in the seasonally dry U.S. intermountain west, is much lower in the summer than in the fall or spring (Wang et al., 2009). To address potential detrimental effects of priors on predictive accuracy, one could

decrease the weight of priors in analyses, relative to the observed (current) data, or use different priors altogether (see Section 5.3 below).

5.3. Bayesian Analytical Considerations

The USGS-PROSPER model has been validated throughout the Pacific Northwest region. However, we found the model to be suboptimal for our application (i.e., defining beta prior hyperparameters in Bayesian models). PROSPER incorporates a comprehensive suite of predictors for flow persistence including land use, land cover, soil permeability, topographic wetness index, average maximum and minimum daily temperature, and annual precipitation (Jaeger et al., 2019). PROSPER model predictions, however, are not seasonally adjusted, and are intended for regional applications rather than fine-spatial-scale predictions. We justify our use of this model to define prior means due to: (a) the lack of a better current alternative, and (b) our view of the PROSPER predictions as provisional long-term representations of the *relative* probability of surface water in stream segments at Murphy Creek. Indeed, our work now provides an updated prior framework for future Bayesian models at the study site.

We plan on further considerations of constraints to the Bayesian analysis of non-perennial streams in future papers. This work will include comparisons of the usefulness of a large number of strictly bounded PDFs (in addition to the beta distribution) for describing the probability of surface water in stream segments with varying physical properties. Other topics for consideration include the sensitivity of the posterior distributions of $\theta_k|\boldsymbol{x}_k$ and $(\theta_k|\boldsymbol{x}_k)^{-1}$ to spatiotemporal scale, sample size, and the prior weighting constant, w. A final topic of general interest is the effect of node placement and sensor number per segment for estimating the proportion of the channel that is wetted under various settings.

5.4. Further Uncertainties and Extensions

Our model predictions concern the probability of surface water presence at stream segments which may not reflect streamflow due to two factors. First, local ponding may lead to surface water without flow, though we did not observe this issue at Murphy Creek during frequent site visits, likely due to its steepness. To address this issue, researchers could combine conductivity and temperature measures (Arismendi et al., 2017). These data could then be coupled with appropriate priors to obtain posterior distributions of the probability of stream flow and the reciprocal probability of stream flow at stream segments. Second, surface networks inferred from very high-resolution topography may be relatively accurate, but those delineated with coarser topographic data that rely on a single area-based threshold (e.g., the ArcGIS Watershed Toolbox) may require ground-truthing or further consideration, particularly in headwaters, low-gradient systems, or karst regions (Yamazaki et al., 2019).

Indeed, although our demonstration considers only surficial stream networks, our field observations suggest that subsurface flow likely dominates Murphy Creek streamflow at certain times of year. In principle, one could model subsurface to surface hydrologic fluxes (e.g., vertical connectivity) and/or subsurface flow using our approaches, by considering the presence/absence of subsurface water at certain depths using an array of shallow wells. Integration of subsurface connectivity would allow more holistic considerations of human impacts on the water cycle, as groundwater pumping can cause streams to transitions from perennial to non-perennial flow regimes (Zipper et al., 2022). However, the assimilation of subsurface flow information into stream connectivity metrics such as communication distance is not a straightforward task due to the difficulty in measuring subsurface flow directions, properties, and rates (Xiao et al., 2021). Furthermore, groundwater can occupy long flowpaths, spanning years to decades (Maxwell et al., 2016), and in some settings, groundwatersheds can be substantially larger than watersheds defined by surface topography (Huggins et al., 2023).

6. Conclusions

Nuanced characterization of a non-perennial stream requires the recognition of dynamic, potentially asynchronous patterns in surface flows within the network. To help describe these properties, we considered the probability of surface water presence at non-perennial stream segments as a random variable. This approach allowed Bayesian extensions to both an existing connectivity metric, Bernoulli stream length (Botter & Durighetto, 2020), and a new metric called communication distance. Communication distance measures resistance to passive transport processes in stream networks, in units of measured stream length. For the kth non-perennial stream segment,

the difference of the mean of a posterior distribution of average communication distances and the fully wetted length of the kth segment provides an inductive measure of the increased average distance required for segment travel, as the result of intermittency. This metric may be helpful for identifying stream segments prone to network bottlenecking. Our work demonstrates the unique connectivity perspectives afforded by communication distance, and the general usefulness of Bayesian approaches in the analysis of non-perennial streams. Our work also reveals the need for the additional consideration of several topics. These efforts should include comparisons of the performance of particular PDFs as priors for the probability of surface water in stream segments, and measurement of the sensitivity of the posterior distributions of $(\theta_k|\boldsymbol{x}_k)^{-1}$ to spatiotemporal scale, sample size, and the prior weighting constant, w.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Data for the Murphy Creek network is published in Warix et al. (2021) and codified in the **R** package *streamDAG* (Aho et al., 2023a, 2023b). The *streamDAG* package is open source, and can be downloaded from the Comprehensive **R** Archive Network. Guidance and examples of *streamDAG* usage are provided in Supporting Information S1 to this manuscript, and in the vignette "Introduction to *streamDAG*" (Aho, 2023).

Acknowledgments
This work was made possible with a Grant from the National Science Foundation, Grant 2019603, RII Track-2 FEC: Aquatic Intermittency Effects on Microbiomes in Streams (AIMS) and NSF EAR1653998. Thanks to non-author AIMS personnel for their intellectual contributions. We also recognize the rigor, hard work, and expertise of three anonymous WRR reviewers whose efforts have dramatically improved this quality of the manuscript.

References

Aho, K. A. (2014). *Foundational and applied statistics for biologists using R*. CRC Press.
Aho, K. A. (2023). Introduction to the streamDAG package (ver. 1.5). https://doi.org/10.5281/zenodo.8415081
Aho, K. A., Kriloff, C., Godsey, S. E., Ramos, R., Wheeler, C., You, Y., et al. (2023). Non-perennial stream networks as directed acyclic graphs: The R-package *streamDAG*. *Environmental Modelling & Software*, *167*, 105775. https://doi.org/10.1016/j.envsoft.2023.105775
Aho, K. A., Ramos, R., Legg, A., Hale, R. L., Kraft, M., & Bond, C. T. (2023). *streamDAG*: Analytical methods for stream DAGs. R package version 1.5. Retrieved from https://CRAN.R-project.org/package=streamDAG
Ali, G. A., & Roy, A. G. (2009). Revisiting hydrologic sampling strategies for an accurate assessment of hydrologic connectivity in humid temperate systems. *Geography Compass*, *3*(1), 350–374. https://doi.org/10.1111/j.1749-8198.2008.00180.x
Ali, G. A., & Roy, A. G. (2010). Shopping for hydrologically representative connectivity metrics in a humid temperate forested catchment. *Water Resources Research*, *46*(12), W12544. https://doi.org/10.1029/2010wr009442
Arismendi, I., Dunham, J. B., Heck, M. P., Schultz, L. D., & Hockman-Wert, D. (2017). A statistical method to predict flow permanence in dryland streams from time series of stream temperature. *Water*, *9*(12), 946. https://doi.org/10.3390/w9120946
Barthélemy, J., & Suesse, T. (2018). mipfp: An R package for multidimensional array fitting and simulating multivariate Bernoulli distributions. *Journal of Statistical Software, Code Snippets*, *86*(2), 1–20. https://doi.org/10.18637/jss.v086.c02
Bertassello, L. E., Durighetto, N., & Botter, G. (2022). Eco-hydrological modelling of channel network dynamics—Part 2: Application to metapopulation dynamics. *Royal Society Open Science*, *9*(11), 220945. https://doi.org/10.1098/rsos.220945
Blume, T., & Van Meerveld, H. J. (2015). From hillslope to stream: Methods to investigate subsurface connectivity. *Wiley Interdisciplinary Reviews: Water*, *2*(3), 177–198. https://doi.org/10.1002/wat2.1071
Botter, G., & Durighetto, N. (2020). The stream length duration curve: A tool for characterizing the time variability of the flowing stream length. *Water Resources Research*, *56*(8), e2020WR027282. https://doi.org/10.1029/2020wr027282
Bracken, L., Wainwright, J., Ali, G., Tetzlaff, D., Smith, M., Reaney, S., & Roy, A. (2013). Concepts of hydrological connectivity: Research approaches, pathways and future agendas. *Earth-Science Reviews*, *119*, 17–34. https://doi.org/10.1016/j.earscirev.2013.02.001
Castellarin, A., Vogel, R. M., & Brath, A. (2004). A stochastic index flow model of flow duration curves. *Water Resources Research*, *40*(3), W03104. https://doi.org/10.1029/2003wr002524
Chapin, T. P., Todd, A. S., & Zeigler, M. P. (2014). Robust, low-cost data loggers for stream temperature, flow intermittency, and relative conductivity monitoring. *Water Resources Research*, *50*(8), 6542–6548. https://doi.org/10.1002/2013wr015158
Cressie, N., Frey, J., Harch, B., & Smith, M. (2006). Spatial prediction on a river network. *Journal of Agricultural, Biological, and Environmental Statistics*, *11*(2), 127–150. https://doi.org/10.1198/108571106x110649
Dai, B., Ding, S., & Wahba, G. (2013). Multivariate Bernoulli distribution. *Bernoulli*, *19*(4), 1465–1483. https://doi.org/10.3150/12-BEJSP10
Datry, T., Larned, S. T., & Tockner, K. (2014). Intermittent rivers: A challenge for freshwater ecology. *BioScience*, *64*(3), 229–235. https://doi.org/10.1093/biosci/bit027
Dodds, P. S., & Rothman, D. H. (2000). Geometry of river networks. I. Scaling, fluctuations, and deviations. *Physical Review E*, *63*(1), 016115. https://doi.org/10.1103/physreve.63.016115
Drummond, J. D., Davies-Colley, R. J., Stott, R., Sukias, J. P., Nagels, J. W., Sharp, A., & Packman, A. I. (2015). Microbial transport, retention, and inactivation in streams: A combined experimental and stochastic modeling approach. *Environmental Science & Technology*, *49*(13), 7825–7833. https://doi.org/10.1021/acs.est.5b01414
Dunnington, D. (2021). ggspatial: Spatial data framework for ggplot2. R package version 1.1.5. Retrieved from https://CRAN.R-project.org/package=ggspatial
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

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

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Chapman and Hall/CRC.

Godsey, S., & Kirchner, J. (2014). Dynamic, discontinuous stream networks: Hydrologically driven variations in active drainage density, flowing channels and stream order. *Hydrological Processes*, *8*(23), 5791–5803. https://doi.org/10.1002/hyp.10310

González-Ferreras, A., & Barquín, J. (2017). Mapping the temporary and perennial character of whole river networks. *Water Resources Research*, *53*(8), 6709–6724. https://doi.org/10.1002/2017wr020390

Gotelli, N. J., & Ellison, A. M. (2004). *A primer of ecological statistics* (Vol. 1). Sinauer Associates.

Hanson, C. L., Marks, D., & Van Vactor, S. S. (2001). Long-term climate database, Reynolds Creek experimental watershed, Idaho, USA. *Water Resources Research*, *37*(11), 2839–2841. https://doi.org/10.1029/2001wr000417

Huggins, X., Gleeson, T., Serrano, D., Zipper, S., Jehn, F., Rohde, M. M., et al. (2023). Overlooked risks and opportunities in groundwatersheds of the world's protected areas. *Nature Sustainability*, *6*(7), 1–10. https://doi.org/10.1038/s41893-023-01086-9

Jaeger, K., Sando, R., McShane, R. R., Dunham, J. B., Hockman-Wert, D., 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

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *186*(1007), 453–461.

Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., & Marshall, L. A. (2009). Hydrologic connectivity between landscapes and streams: Transferring reach-and plot-scale understanding to the catchment scale. *Water Resources Research*, *45*(4), W03104. https://doi.org/10.1029/2008wr007225

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

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous univariate distributions, volume 2* (Vol. 289). John Wiley & sons.

Kaplan, N. H., Blume, T., & Weiler, M. (2020). Predicting probabilities of streamflow intermittency across a temperate mesoscale catchment. *Hydrology and Earth System Sciences*, *24*(11), 5453–5472. https://doi.org/10.5194/hess-24-5453-2020

Knudby, C., & Carrera, J. (2005). On the relationship between indicators of geostatistical, flow and transport connectivity. *Advances in Water Resources*, *28*(4), 405–421. https://doi.org/10.1016/j.advwatres.2004.09.001

Kotz, S., & Van Dorp, J. R. (2004). *Beyond beta: Other continuous families of distributions with bounded support and applications*. World Scientific.

Larsen, L. G., Choi, J., Nungesser, M. K., & Harvey, J. W. (2012). Directional connectivity in hydrology and ecology. *Ecological Applications*, *22*(8), 2204–2220. https://doi.org/10.1890/11-1948.1

Maxwell, R. M., Condon, L. E., Kollet, S. J., Maher, K., Haggerty, R., & Forrester, M. M. (2016). The imprint of climate and geology on the residence times of groundwater. *Geophysical Research Letters*, *43*(2), 701–708. https://doi.org/10.1002/2015gl066916

McNamara, J. P., Chandler, D., Seyfried, M., & Achet, S. (2005). Soil moisture states, lateral flow, and streamflow generation in a semi-arid, snowmelt-driven catchment. *Hydrological Processes*, *19*(20), 4023–4038. https://doi.org/10.1002/hyp.5869

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

Morita, S., Thall, P. F., & Müller, P. (2008). Determining the effective sample size of a parametric prior. *Biometrics*, *64*(2), 595–602. https://doi.org/10.1111/j.1541-0420.2007.00888.x

Murrell, Z., & Wen, Z. (2020). gridGraphics: Redraw base graphics using 'grid' Graphics. R package version 0.5-1. Retrieved from https://CRAN.R-project.org/package=gridGraphics

Pähtz, T., Clark, A. H., Valyrakis, M., & Durán, O. (2020). The physics of sediment transport initiation, cessation, and entrainment across aeolian and fluvial environments. *Reviews of Geophysics*, *58*(1), e2019RG000679. https://doi.org/10.1029/2019rg000679

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal*, *10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Pitman, E. J. G. (1936). Sufficient statistics and intrinsic accuracy. In *Mathematical proceedings of the cambridge philosophical society* (Vol. 32, No. (4), pp. 567–579). Cambridge University Press.

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

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

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

Sauquet, E., Shanafield, M., Hammond, J. C., Sefton, C., Leigh, C., & Datry, T. (2021). Classification and trends in intermittent river flow regimes in Australia, northwestern Europe and USA: A global perspective. *Journal of Hydrology*, *597*, 126170. https://doi.org/10.1016/j.jhydrol.2021.126170

Shanafield, M., Bourke, S. A., Zimmer, M. A., & Costigan, K. H. (2021). An overview of the hydrology of non-perennial rivers and streams. *Wiley Interdisciplinary Reviews: Water*, *8*(2), e1504. https://doi.org/10.1002/wat2.1504

Trigg, M. A., Michaelides, K., Neal, J. C., & Bates, P. D. (2013). Surface water connectivity dynamics of a large-scale extreme flood. *Journal of Hydrology*, *505*, 138–149. https://doi.org/10.1016/j.jhydrol.2013.09.035

USGS. (2016). StreamStats. Retrieved from http://streamstats.usgs.gov

Ver Hoef, J. M., Peterson, E., & Theobald, D. (2006). Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, *13*(4), 449–464. https://doi.org/10.1007/s10651-006-0022-8

Wang, S. Y., Gillies, R. R., Takle, E. S., & Gutowski, W. J., Jr. (2009). Evaluation of precipitation in the Intermountain Region as simulated by the NARCCAP regional climate models. *Geophysical Research Letters*, *36*(11), L11704. https://doi.org/10.1029/2009gl037930

Warix, S. R., Godsey, S. E., Lohse, K. A., & Hale, R. L. (2021). Influence of groundwater and topography on stream drying in semi-arid headwater streams. *Hydrological Processes*, *35*(5), e14185. https://doi.org/10.1002/hyp.14185

Western, A. W., Blöschl, G., & Grayson, R. B. (2001). Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resources Research*, *37*(1), 83–97. https://doi.org/10.1029/2000wr900241

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. Retrieved from https://cran.r-project.org/package=ggplot2

Wilke, C. O. (2020). cowplot: Streamlined plot theme and plot annotations for 'ggplot2'. R package version 1.1.1. Retrieved from https://CRAN.R-project.org/package=cowplot

Xiao, D., Brantley, S. L., & Li, L. (2021). Vertical connectivity regulates water transit time and chemical weathering at the hillslope scale. *Water Resources Research*, *57*(8), e2020WR029207. https://doi.org/10.1029/2020wr029207

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., & Pavelsky, T. M. (2019). MERIT hydro: A high-resolution global hydrography map based on latest topography dataset. *Water Resources Research*, *55*(6), 5053–5073. https://doi.org/10.1029/2019wr024873

Zipper, S. C., Hammond, J. C., Shanafield, M., Zimmer, M., Datry, T., JonesAllen, C. N. D. C., et al. (2021). Pervasive changes in stream intermittency across the United States. *Environmental Research Letters*, *16*(8), 084033. https://doi.org/10.1088/1748-9326/ac14ec

Zipper, S. C., Popescu, I., Compare, K., Zhang, C., & Seybold, E. C. (2022). Alternative stable states and hydrological regime shifts in a large intermittent river. *Environmental Research Letters*, *17*(7), 074005. https://doi.org/10.1088/1748-9326/ac7539