Received: 22 May 2023 | Accepted: 25 March 2024

DOI: 10.1111/1365-2664.14665

Journal of Applied Ecology | BRITISH ECOLOGICAL SOCIETY

RESEARCH ARTICLE

Using multi-scale spatial models of dendritic ecosystems to infer abundance of a stream salmonid

Xinyi Lu[1,2] | Yoichiro Kanno[2,3] | George P. Valentine[2,3] | Jacob M. Rash[4] | Mevin B. Hooten[2,5]

[1]Mathematics and Statistics Department, Utah State University, Logan, Utah, USA

[2]Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, Colorado, USA

[3]Graduate Degree Program in Ecology, Colorado State University, Fort Collins, Colorado, USA

[4]North Carolina Wildlife Resources Commission, Marion, North Carolina, USA

[5]Department of Statistics and Data Sciences, The University of Texas at Austin, Austin, Texas, USA

Correspondence
Xinyi Lu
Email: lucy.lu@usu.edu

Funding information
U.S. Fish and Wildlife Service, Grant/Award Number: F20AC11372; U.S. Geological Survey Southeast Climate Adaptation Science Center, Grant/Award Number: G21AC1005

Handling Editor: Nubia Marques

Abstract

1. Understanding patterns of species abundance is essential for planning landscape-level conservation. The complex hierarchies of dendritic ecosystems result in different levels of heterogeneity at distinct geographic scales. Species responses to dynamic environmental drivers may also vary spatially depending on their interactions with landscape features. Monitoring abundance by explicitly quantifying their spatial and temporal variation is important for strategic management.
2. We analysed brook trout (*Salvelinus fontinalis*) count data collected from 173 sites in western North Carolina between 1989 and 2015. We developed a Bayesian hierarchical model that used single- and multi-pass electro-fishing data and characterized their respective capture probabilities. We quantified spatial variation using a multi-scale process model representative of the nested stream habitats, and we investigated differences in population temporal trends and responses to seasonal weather patterns by space and life stage.
3. Trout abundance was lower on the Atlantic slope of the Eastern Continental Divide than in the interior, on average, and the Atlantic slope juveniles were more adversely affected by high winter flows. However, Atlantic slope populations of both lifestages demonstrated positive temporal trends, whereas Interior juveniles demonstrated a negative trend. We found higher spatial variation than temporal variation in abundance when conditioned on the covariates, where the primary source of spatial heterogeneity was revealed at the segment level, compared to watershed or network levels.
4. Our multi-scale spatial model outperformed simpler models in abundance estimation and out-of-sample prediction. The inferred per-pass capture probabilities indicated that single-pass surveys were as efficient as multi-pass surveys.
5. *Synthesis and applications*. Our study suggested conservation priority should involve multiple criteria, including present-day abundance, temporal trend and sensitivity to environmental drivers. Based on the inferred scale-specific variations in trout abundance, we recommend that future surveys strategically combine single-pass surveys with multi-pass surveys to optimize abundance estimation.

Our approach is widely applicable to other species and ecosystems occupying dendritic habitats.

KEYWORDS
Bayesian, brook trout, climate change, data fusion, depletion sampling, network, N-mixture

1 | INTRODUCTION

Understanding patterns of species abundance in space provides an essential guide to management and conservation. A suite of population models exist for major types of survey data, including presence/absence, count, and capture-recapture data. For lotic species, occupancy models have been widely used to characterize population distributions because of generally available presence/absence data and a flexible paradigm for parameter estimation (DeWeber & Wagner, 2015; MacKenzie et al., 2002). Capture-recapture studies enable in-depth inference about population vital rates, such as survival and movement (Letcher et al., 2015; Terui et al., 2021). However, conclusions from these studies may be bound by the local ecological systems and difficult to replicate due to intensive sampling demands. In comparison, count data balance between information entropy and sampling effort, thereby facilitating our learning of a species' abundance across its range. Traditional population models for stream fish typically describe density (i.e. per-unit abundance, where the unit may be defined by length, area, or volume) as a linear function of habitat covariates (Rosenfeld, 2003). Recent studies have utilized random effects in generalized linear mixed models to quantify variation due to unobserved covariates (Helser et al., 2004; Wenger et al., 2022). Although stream fish abundance has been studied at many geographic scales from stream reach to watershed (Deschênes & Rodrguez, 2007; Ebersole et al., 2009; Jackson et al., 2001), common approaches have rarely quantified habitat hierarchy in an integrative manner.

Patterns of abundance are influenced by a combination of ecological processes operating at various spatial scales (Brown, 2014; Ricklefs, 1987). In stream ecosystems, watershed characteristics are shaped by geology and soil, and they provide a broad-scale template on which stream networks develop (Frissell et al., 1986). In stream networks, local stream habitats change due to geomorphic heterogeneity and influence of tributaries (Benda et al., 2004; Fausch et al., 2002). We accounted for this nested hierarchy using a combination of spatial models, including a class of recently developed spatial stream network (SSN) models (Peterson et al., 2013; Ver Hoef & Peterson, 2010). SSN models explicitly characterize the configurations, connectivities, and flow directions of streams, and are increasingly applied to studies of lotic species (e.g. Isaak et al., 2017). In addition to dependence in space, we also accounted for temporal trend and autocorrelation in abundance. Delineating spatiotemporal autocorrelation in population models can reduce estimation bias and improve uncertainty quantification (Hocking et al., 2018), and space–time models concerning stream networks are an area of active research (Santos-Fernandez et al., 2022).

Capture probabilities in abundance models are different from detection probabilities in occupancy models in that the former are a direct measure of sampling efficiency while the latter are conditioned on species presence (Royle, 2006). Observed abundance has been used as an index for species distribution when capture probability cannot be estimated (Matson et al., 2018). In the context of stream fish population studies, electro-fishing is a common survey method used by fisheries ecologists and managers to collect count data (Bohlin et al., 1989). Kruse et al. (1998) used single-pass electro-fishing data to demonstrate relative abundance in mountain streams. Similarly, Bertrand et al. (2006) showed that single-pass samples were representative of species richness and trends in intermittent prairie streams. However, capture probabilities can change over space and time (Kanno et al., 2015; Rosenberger & Dunham, 2005), and the utility of single-pass data as abundance indices may be limited without standardized effort (Bertrand et al., 2006). On the other hand, capture probability can be estimated from multi-pass removal data under assumptions of population closure and constant capture probability (Peterson et al., 2004; Zippin, 1958). To leverage the spatial coverage of single-pass data and the reliability of multi-pass data, we developed a modelling framework that accommodates both sampling protocols while accounting for their respective capture probabilities.

Our objective was to develop a multi-scale spatial modelling framework for abundance estimation within dendritic ecosystems. Specifically, we sought to validate our proposed method via simulation, and to demonstrate improved estimation and predictive performances compared to a simpler alternative. Although our framework has potential utility for many lotic species, we illustrated an application with native brook trout (*Salvelinus fontinalis*) populations in western North Carolina, USA.

2 | MATERIALS AND METHODS

2.1 | Study species

Brook trout are an iconic cold-water fish native to eastern North America. Native brook trout have experienced substantial declines due to anthropogenic activities (Hudy et al., 2008), and remnant populations inhabit isolated headwaters with limited dispersal potential (Kazyak et al., 2022) at their southern range including

our study area. Brook trout are typically short lived (up to 4 years; Larson & Moore, 1985). They spawn during fall, and fertilized eggs develop in redds (depressions in gravel substrates) and hatch in late winter to early spring (Hazzard, 1932). Brook trout are sensitive to stream temperature and flow. Prolonged high summer temperatures can cause physiological stress and adversely affect spawning (Warren et al., 2012; Xu et al., 2010), and high flows during winter and spring can result in bed scouring events that cause high mortality in young-of-the-year (YOY) (Kanno et al., 2015; Roghair et al., 2002). However, there is limited knowledge on how differently trout populations respond to seasonal weather patterns along the environmental gradient across their native range (Valentine et al., 2024).

2.2 | Data

We analysed brook trout count data collected from 173 sites in western North Carolina between 1989 and 2015 by the North Carolina Wildlife Resources Commission (NCWRC). Our study did not require ethical approval. The study sites were distributed geographically across the major watersheds delineated by the hydrologic unit code (HUC) 8 in the southern Appalachian Mountains region (Figure 1). Our multi-scale study system consisted of 11 watersheds, within which 14 distinct stream networks and 109 unique stream segments were nested (see Figures 2 and 3 for illustration). A segment is a portion of a stream connecting two confluences (or its beginning and a confluence), and a network is a collection of hydrologically connected segments. Different networks were separated by physical barriers in our study region. The Eastern Continental Divide (ECD) functions as a natural barrier rendering diverse trout habitat characteristics on either side. The study area is underlain by granitic and metamorphic rocks on both sides of the Divide; however, topographic relief and erosion rates vary by side, resulting in broad-scale differences in geomorphology, soil layers, and water chemistry (Johnson, 2020; Rissler et al., 2004). The Atlantic slope sites were lower in elevation and warmer than the Interior sites (Table 1). Maximum summer air temperature differed by approximately 2 degrees Celsius and maximum winter air temperature differed by approximately 1.5 degrees Celsius, on average, between the two groups of sites. The Interior sites demonstrated higher variation in watershed area and channel gradient than the Atlantic slope sites, albeit the mean values were similar between slopes. Anecdotally, local fisheries biologists have

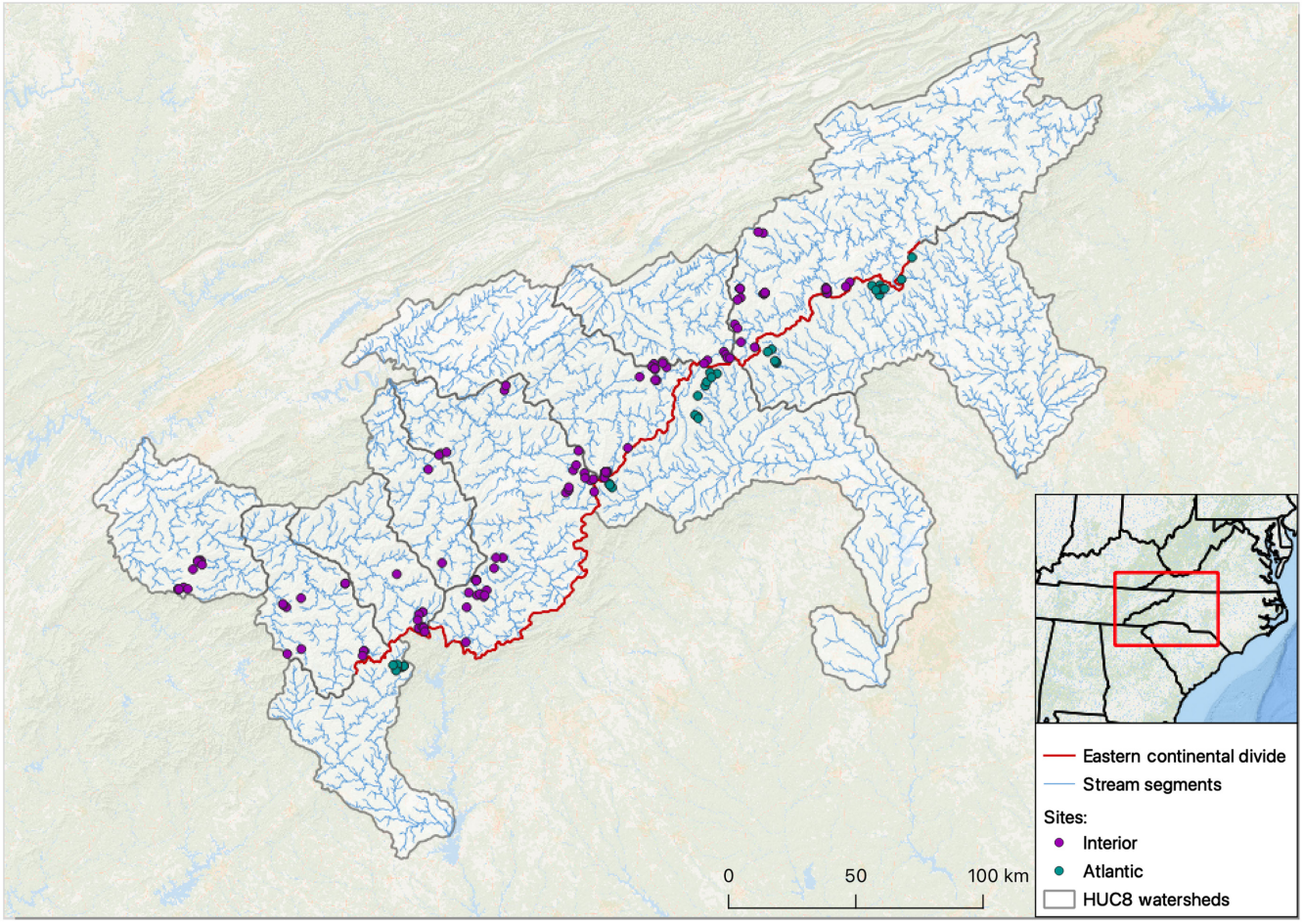

FIGURE 1 Map of the 173 study sites across 11 watersheds in western North Carolina, USA.

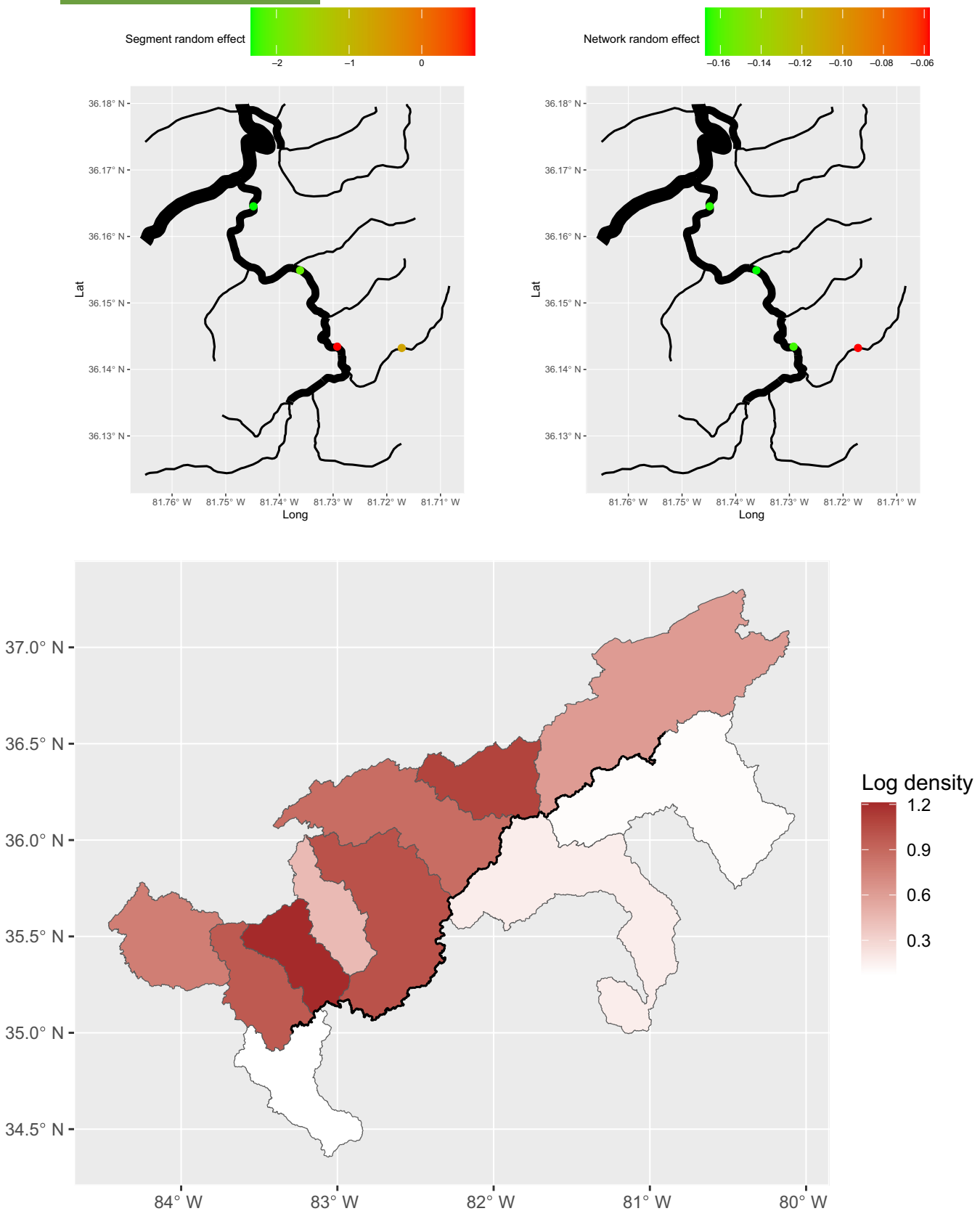

FIGURE 2 Map of estimated posterior mean random effects for YOY at the watershed level (bottom, combined with estimated β_0), the segment level (top left) and the network level (top right). The Eastern Continental Divide is represented by the bold line (bottom). A set of segments are illustrated for a portion of a single network (top), where line width represents stream order and points represents study sites. The bottom panel suggests that the Atlantic slope supported fewer trout than the Interior, on average; however, there was notable variation within watersheds, at the segment level (top left), as is demonstrated by the wider range of values in the legend.

Journal of Applied Ecology BRITISH ECOLOGICAL SOCIETY

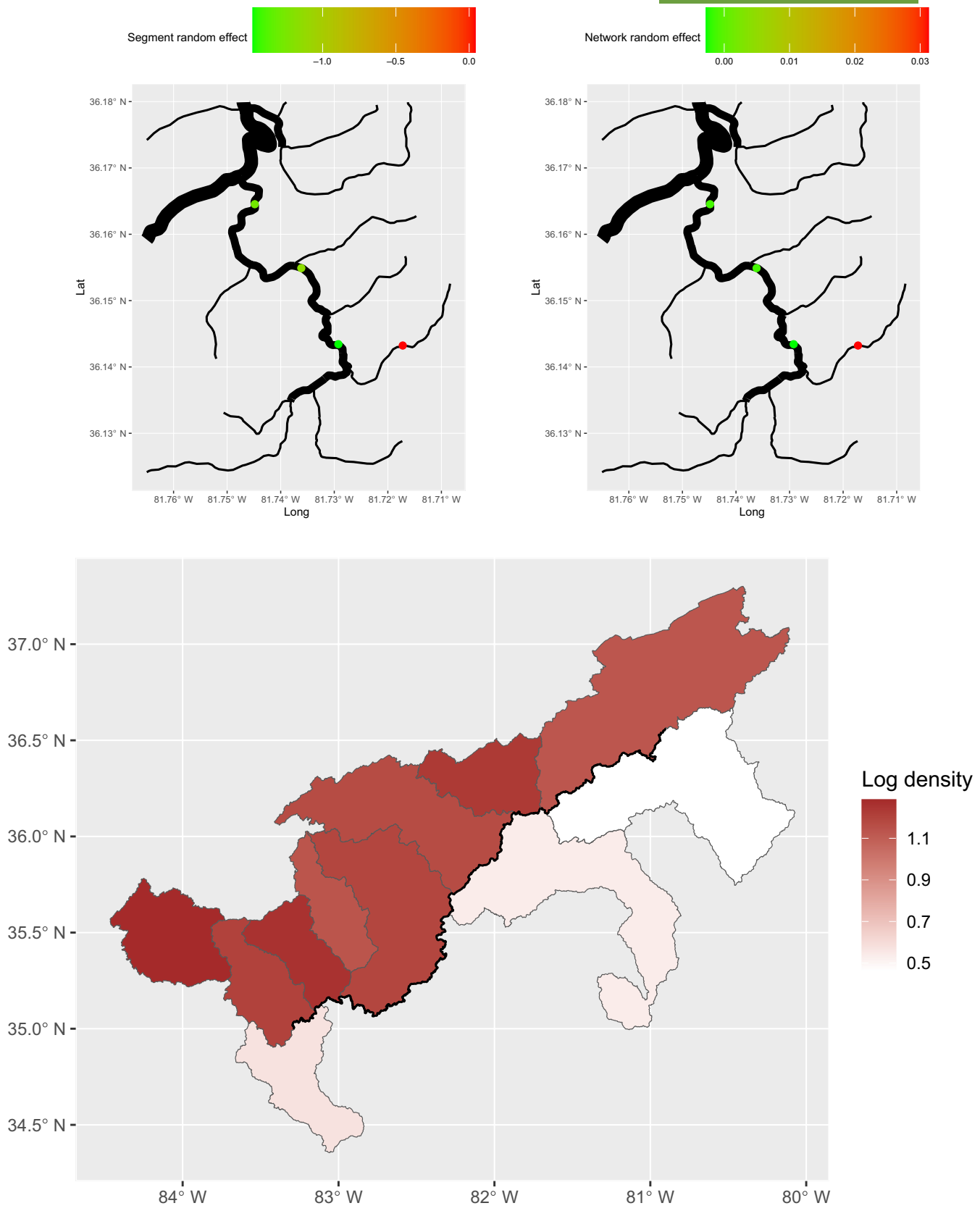

FIGURE 3 Map of estimated posterior mean random effects for adults at the watershed level (bottom, combined with estimated β_0), the segment level (top left), and the network level (top right). See Figure 2 caption for additional details.

TABLE 1 Summary of study site features on either side of the Divide (mean [range]). Elevation was obtained from NCWRC trout surveys. Watershed area and channel gradient were obtained from NHDPlus (U.S. Geological Survey, 2016), and maximum winter and summer air temperatures were obtained from Daymet (Thornton et al., 2022).

	Interior (136 sites)	Atlantic (37 sites)
Elevation (m)	991 (666, 1376)	670 (390, 1074)
Watershed area (km^2)	12.52 (1.87, 79.86)	13.27 (2.56, 65.15)
Channel gradient (%)	6.45 (0.30, 23.05)	6.49 (0.62, 16.12)
Max winter air temperature (°C)	7.44 (4.85, 10.13)	8.89 (5.90, 10.03)
Max summer air temperature (°C)	24.48 (21.71, 26.86)	26.62 (23.06, 28.74)

observed that the Interior provided more productive cold-water habitats than the Atlantic slope, although quantitative assessments have not been conducted to date.

Electro-fishing surveys of brook trout populations occurred between June and October using two sampling protocols. Depletion surveys were conducted at two-thirds of the study sites (Group A), where block nets or natural barriers were used to ensure population closure during sampling, and three successive electro-fishing passes were conducted with one backpack electrofisher and one netter for every three meters of stream width. The remaining sites (Group B) were surveyed with single-pass electrofishing. Single-pass surveys complemented the depletion surveys by broadening the spatial coverage with less rigorous sampling effort per site, and were intended as a rapid population assessment. During the study period, multi-pass sites had between one to 10 temporal replicates, whereas single-pass sites had no more than two temporal replicates. Brook trout collected during the survey were weighed (g) and measured (mm in total length [TL]), and segregated into YOY (<90 mm TL) and adults (≥90 mm TL) for subsequent analysis.

For weather-related covariates at the survey sites, we incorporated data from Daymet (Thornton et al., 2022) and the National Hydrography Dataset Plus (NHDPlus, U.S. Geological Survey, 2016). Daymet provided daily estimates of temperature on a 1 km × 1 km gridded surface over continental North America, and we used the average of daily maximum temperatures between June and September of each year to represent high summer temperatures at each site. NHDPlus provides monthly estimates of flow at stream segments (uniquely identified by COMID) across the United States, and we used the highest available flow percentiles (90th percentile) between December and February to represent high winter flows at each segment. Similarly, we used the highest available flow percentiles between March and May to represent high spring flows. Further, we used topological data from NHDPlus on watersheds and stream networks in the study region to specify the spatial structure of our model.

2.3 | Model

Our hierarchical framework consisted of the data, process, and parameter models following the convention of Berliner (1996) and Wikle et al. (1998). We denote $y_{i,t,j}$ as the observed count at site i, $i = 1, \ldots, n$, year t, $t = 1, \ldots, T$, and pass j, $j = 1,2,3$. The data model characterizes the depletion sampling protocol as follows,

$$y_{i,t,j} \sim \begin{cases} \text{Binomial}\left(N_{i,t},\ p_{i,t}\right),\ j=1, \\ \text{Binomial}\left(N_{i,t} - \sum_{l=1}^{j-1} y_{i,t,l},\ p_{i,t}\right),\ j>1. \end{cases} \quad (1)$$

True abundance at site i in year t is represented by $N_{i,t}$, and capture probability $p_{i,t}$ is p_A if site i was in Group A, and p_B if in Group B. Single-pass capture probability (p_B) is estimated from its prior specification and the modelled abundance under the hierarchical framework. Further, we specified

$$N_{i,t} \sim \text{Poisson}\left(\frac{A_i \lambda_{i,t}}{1000}\right), \quad (2)$$

where A_i denotes the survey area and $\lambda_{i,t}$ denotes areal density (fish/m^2). Survey areas were adjusted by a fraction of 1000 to facilitate numerical stability.

Our process model characterized density using a log-linear function of covariates and random effects as follows,

$$\log\left(\lambda_{i,t}\right) = \boldsymbol{x}_i'\boldsymbol{\beta} + \boldsymbol{h}_{i,t}'\boldsymbol{\theta}_i + \varepsilon_i + \nu_i + \psi_i + \gamma_i \log t + \eta_t. \quad (3)$$

Our landscape covariates, $\boldsymbol{x}_i$, at site i included an intercept, latitude, elevation, and an indicator for either slope of the ECD (0 = Interior, 1 = Atlantic). Latitude and elevation were standardized over the study region. Our weather-related covariates, $\boldsymbol{h}_{i,t}$, at site i in year t included high summer temperature (in year $t-1$), high winter flow, and high spring flow. Because fish sampling occurred between June and October, which overlapped our definition of summer (June–September), we used previous summer temperatures to ensure they were past conditions. The weather-related covariates were standardized at each site, and their effects were modelled by $\boldsymbol{\theta}_i = \boldsymbol{\theta}_{\text{Interior}}$ if site i is Interior, and $\boldsymbol{\theta}_i = \boldsymbol{\theta}_{\text{Atlantic}}$ if site i is on the Atlantic slope.

We specified multi-scale spatial random effects to represent the nested architecture of stream habitats. This specification provides scale-specific inference on spatial variation and dependence in abundance. It also mitigates computational barriers by enabling block updates of large covariance matrices (Katzfuss, 2017; Lu et al., 2023). At the watershed level, we used a conditional autoregressive model (CAR),

$$\boldsymbol{\varepsilon} \sim \text{N}\left(\boldsymbol{0}, \sigma_\varepsilon^2 \left(\boldsymbol{R} - \rho_\varepsilon \boldsymbol{W}\right)^{-1}\right), \quad (4)$$

where $\boldsymbol{W}$ denotes a weight matrix defined by inverse Euclidean distances between pairs of watershed centers, and $\boldsymbol{R}$ is the diagonal

matrix of row sums of $\boldsymbol{W}$ (Ver Hoef et al., 2018). We set $\rho_\epsilon = 0.99$ to allow moderate spatial correlation among watersheds (Besag & Kooperberg, 1995). At the segment level, we used a geostatistical model with exponential covariance function,

$$\mathrm{Cov}(v_i, v_j) = \begin{cases} \sigma_v^2 \exp(-d_{ij}/\phi_v), & \text{if sites } i \text{ and } j \text{ were located on the same segment,} \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

where d_{ij} denotes the Euclidean distance between sites i and j, and ϕ_v denotes the range of spatial dependence along the segment.

Euclidean-distance-based covariance models may be inadequate for ecological processes within dendritic networks (Peterson et al., 2013; Ver Hoef & Peterson, 2010). Therefore, we specified a network level random effect, $\boldsymbol{\psi}$, using an SSN model based on hydrologic distances and stream topology. Ver Hoef and Peterson (2010) proposed three classes of SSN models distinguished by their relative direction of stream flow to that of correlation decay. For our application, we adopted a tail-down model because it correlates merging segments and can be used to account for dependence due to fish movement both upstream and downstream (e.g. movement between main stem and tributaries during different seasons Letcher et al., 2015; Petty et al., 2012; Thomaz et al., 2016). We defined

$$\mathrm{Cov}(\psi_i, \psi_j) = \begin{cases} \sigma_\psi^2 \exp(-s_{ij}/\phi_\psi), & \text{if sites } i \text{ and } j \text{ were located in the same network,} \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

The network distance between sites i and j, s_{ij}, was measured by the downstream distance of site j from site i, if the sites were flow-connected; if the sites were flow-unconnected, s_{ij} was measured by the sum of the downstream distances from sites i and j to their common junction (see Figure B1 in Appendix B for illustration). Although the network covariance (Equation 6) shares the same functional form as the segment covariance (Equation 5), the unique definition of network distance and non-zero correlation between different segments within a network (Equation 6) allowed us to distinguish network random effects from segment random effects. We delineated stream networks in our study region using the STAR package (Peterson & Ver Hoef, 2014) in ArcGIS (ESRI, 2011; see Figures 2 and 3 for illustration), and extracted pairwise network distances between study sites using the SSN package in R (Ver Hoef et al., 2014).

We modelled linear trend on log year by letting $\gamma_i = \gamma_{\text{Interior}}$ if site i is Interior, and $\gamma_i = \gamma_{\text{Atlantic}}$ if site i is on the Atlantic slope. A first-order autoregressive model was used for additional temporal variation, so that $\eta_t \sim \mathrm{N}\left(\rho_\eta \eta_{t-1},\ \sigma_\eta^2\right)$, for $t = 2, \ldots, T$, where ρ_η is the autocorrelation parameter bounded between −1 and 1 to ensure stationarity, and η_1 is the Guassian distributed initial condition with mean zero and variance σ_η^2.

We completed the hierarchy by specifying prior distributions in our parameter model. Capture efficiencies (p_A and p_B) were assigned informative priors based on previous studies (Kanno et al., 2015) and the NCWRC 2022 calibration surveys (J.M. Rash, unpublished data). The other parameters were given diffuse priors. We provided a full list of prior distributions for our case study in Appendix A. Our model was implemented using a Markov Chain Monte Carlo sampling algorithm in R (R Core Team, 2021), and fit to YOY and adult data separately. We ran the algorithm for 20,000 iterations and discarded 10,000 burn-in samples. In a simulation study, we demonstrated that our model was able to recover all parameters used to generate data with their respective estimated 95% credible intervals. We also compared our full process model to models with partial or no spatial structure using a five-fold cross-validation. Our multi-scale spatial model produced the highest posterior predictive log likelihood and the lowest prediction root mean squared error (RMSE), thereby exhibiting the best predictive performance among contending models (see Appendix B for details).

3 | RESULTS

We evaluated our model on the brook trout data in comparison to a simplistic model with no spatial structure. Our model produced an estimation RMSE of 1.81 (in count unit) for YOY and 1.31 for adults, and an out-of-sample prediction RMSE of 21.67 for YOY and 14.06 for adults, both lower than the simplistic alternative (see Appendix C for details). Our results indicated that trout abundance was positively related to elevation, and the Interior supported more trout than the Atlantic slope. The estimated posterior mean β_0 was higher for Interior populations than their Atlantic counterparts. When all other covariates are held at their standardized means and the temporal trend and random effects at zero, the posterior mean estimated density was 0.96 per m^2 for Interior populations (YOY and adult combined), and 0.05 per m^2 for Atlantic slope populations. Nonetheless, both Atlantic slope YOY and adult demonstrated positive temporal trends, whereas Interior YOY demonstrated a negative trend (Table 2).

The effects of seasonal weather patterns were stronger on YOY than on adults. Posterior estimates of $\boldsymbol{\theta}$ indicated the impacts of high winter and spring flows on YOY differed by slope. Atlantic YOY were most adversely affected by high winter flow (θ_2). Interior YOY, on the other hand, were most adversely affected by high spring flow (θ_3).

TABLE 2 Estimated posterior means (95% credible intervals) of covariate effects on log abundance. Statistical significance was declared for parameters whose credible intervals did not include zero.

	YOY		Adult	
Parameter	**Atlantic**	**Interior**	**Atlantic**	**Interior**
β_1 (Lat.)	−0.06 (−0.44, 0.15)		0.04 (−0.24, 0.35)	
β_2 (Elev.)	1.27 (1.05, 1.45)		1.32 (1.14, 1.46)	
β_0 (Mean)	−0.11 (−0.64, 0.44)	1.01 (0.81, 1.19)	0.57 (−0.00, 1.06)	1.24 (0.89, 1.51)
θ_1 (Temp)	0.04 (−0.10, 0.15)	−0.07 (−0.23, 0.02)	0.07 (−0.03, 0.19)	0.04 (−0.06, 0.18)
θ_2 (W. Flow)	−0.87 (−1.03, −0.69)	−0.25 (−0.38, −0.16)	−0.08 (−0.23, 0.05)	0.03 (−0.05, 0.12)
θ_3 (S. Flow)	−0.43 (−0.59, −0.29)	−0.44 (−0.54, −0.35)	0.12 (0.01, 0.23)	0.07 (−0.01, 0.15)
γ (Trend)	0.59 (0.25, 0.94)	−0.35 (−0.50, −0.18)	0.52 (0.37, 0.66)	−0.09 (−0.20, 0.06)

TABLE 3 Estimated posterior means (95% credible intervals) of capture probability and random effect parameters.

Parameter	**YOY**	**Adult**
p_A (Group A capture)	0.57 (0.56, 0.59)	0.70 (0.69, 0.71)
p_B (Group B capture)	0.56 (0.46, 0.66)	0.70 (0.60, 0.79)
σ^2_ε (watershed variance)	0.04 (0.01, 0.09)	0.005 (0.001, 0.017)
σ^2_ν (segment variance)	5.55 (4.21, 6.85)	3.13 (2.42, 3.88)
ϕ_ν (segment range)	2.97 (1.50, 4.73)	3.21 (1.28, 5.81)
σ^2_ψ (network variance)	0.008 (0.002, 0.021)	0.002 (0.000, 0.005)
ϕ_ψ (network range)	7.49 (3.93, 9.90)	6.66 (2.41, 9.92)
σ^2_η (temporal variance)	0.42 (0.23, 0.74)	0.04 (0.01, 0.09)
ρ_η (autocorrelation)	0.31 (0.03, 0.71)	0.32 (0.02, 0.88)

High summer temperature did not have significant effects on YOY (θ_1), and all weather-related covariates were less influential for adults than for YOY (Table 2).

Despite reduced overall sampling effort in Group B, our model did not show significant differences between Group A and Group B capture efficiencies per-pass, for YOY or adult (Table 3). The estimated 95% credible intervals of capture probabilities were wider for Group B than for Group A because Group B consisted of fewer sites. Among the three levels of spatial random effects, most variation was detected between segments (Equation 5), and less variation was detected between watersheds or networks (Equation 6). The network-scale random effects had higher estimated ranges (ϕ_ψ) than the corresponding segment-scale ranges (ϕ_ν), indicating that correlation decay was slower along stream networks. We provided conceptual illustrations of estimated posterior mean densities at the three spatial scales (Figures 2 and 3).

The estimated temporal variance (σ^2_η) was smaller than the segment-scale spatial variance (σ^2_ν), indicating the brook trout populations demonstrated higher spatial variability across the study region than temporal dynamics over the study period, conditioned on the covariates. We estimated greater temporal variance for YOY than for adults. Temporal autocorrelation was comparable between YOY and adults (Table 3).

4 | DISCUSSION

We developed a multi-scale spatial modelling framework for abundance estimation of animals occupying dendritic ecosystems. Our proposed method was validated via simulation, and we demonstrated its advantages in model fitness and predictive ability over a simpler alternative using brook trout count data. Compared to population models with extensive variable selection (Cao et al., 2016; Dunham & Rieman, 1999), our framework prioritizes decomposing spatial variation at different geographic scales using random effects. This strategy operationalizes the concept that stream habitats are spatially nested (Fausch et al., 2002; Frissell et al., 1986) and helps reveal the scale where the largest variation in abundance exists. Subsequent analyses based on our inference can focus on habitat features at the corresponding scale to investigate drivers of brook trout distribution systematically. Beyond inferring abundance, our model incorporated both single- and multi-pass electro-fishing data and distinguished their respective capture efficiencies. The framework expanded the spatial range of inference by integrating data collected using different methods, which is useful for multi-state, multi-agency conservation planning. Extensions to this framework can incorporate other sources of fish count data besides electrofishing data, such as snorkelling (Thurow & Schill, 1996) or eDNA

data (Lacoursière-Roussel et al., 2016), by developing survey-specific data models that share the same process model for density, similar to examples of integrated population models (IPMs; Abadi et al., 2010; Schaub & Abadi, 2011; Scheuerell et al., 2021). These extensions will benefit from data models calibrated to the empirical distributions of observed abundance (e.g. zero-inflated Poisson models for excessive zero counts, or negative-Binomial models for over-dispersed counts) and informative prior specifications when capture probabilities are expected to be low.

Because the estimated per-pass capture efficiencies were comparable between single- and multi-pass surveys for brook trout in western North Carolina, our analysis informs that less intensive efforts during single-pass surveys may be leveraged to increase the number of monitored streams. Increasing spatial coverage of population surveys along with incorporating segment-specific covariates, observed or modelled (Kanno et al., 2014), will assist in spatial prediction and uncertainty quantification (Leach et al., 2022). Much spatial variation was revealed among segments, highlighting the need for characterizing populations at the fine resolution for brook trout management. In fact, the Eastern Brook Trout Joint Venture, a multi-agency partnership for conservation, currently aims to assess and update population status at comparable resolutions (catchments or groups of spatially contiguous catchments defined as "patches") in the eastern range based on data and expert knowledge (EBTJV (Eastern Brook Trout Joint Venture), 2018). Our findings justify this level of detailed effort, although geographically larger units have been used in some cases due to data insufficiency (Hudy et al., 2008). As on-going brook trout surveys collect more data in space and time, our analytical framework can incorporate more flexible temporal dynamics such as density-regulated growth and/or Allee effects that characterize inter-specific competition (Morita, 2018) and angling. Further, we can extend the multi-scale specification by letting the weather effects and temporal trends vary by watershed and segment, and explore the interplay between spatial and temporal variability to fully leverage the information in abundance surveys.

Our analysis showed that seasonal weather affected brook trout populations prominently at the early life stage (Kanno et al., 2015; Kovach et al., 2016). The negative impact of high winter flows is concerning for population persistence because the study area is projected to experience higher precipitation, particularly during winters (Ingram et al., 2013). In our case, Interior habitats supported larger populations and Interior YOY were less sensitive to winter flow. However, the temporal trends of both life stages were significantly positive in the Atlantic slope, whereas Interior YOY trended negatively. These results suggest that conservation priority should be determined with caution, and considerations should involve multiple criteria including temporal trend, variation, and sensitivity to environmental stressors, in addition to present-day abundance. Beyond the broad-scale contrast in abundance between Interior and Atlantic slope populations, a major source of spatial variation was found among segments, whose range of autocorrelation most closely matched that of brook trout movement in the southern Appalachian Mountains (typically limited to a few hundred meters, Hudy et al., 2008; Kazyak et al., 2022). Our network specification could be more informative for mobile organisms, such as brook trout occupying larger and more connected stream habitats (Huntsman et al., 2016; White et al., 2020). The framework could also be useful for aquatic organisms that travel both along the waterway (characterized via network models) and over land (characterized via geostatistical models), such as amphibians (Campbell Grant et al., 2010; Miller et al., 2015) and aquatic insects with a terrestrial life stage (Chaput-Bardy et al., 2008; Uno & Power, 2015).

AUTHOR CONTRIBUTIONS

All authors contributed to the study conception. Material preparation and data collection were performed by Jacob M. Rash and North Carolina Wildlife Resources Commission staff; data curation was performed by George P. Valentine; model development and data analysis were performed by Xinyi Lu, Yoichiro Kanno and Mevin B. Hooten. The first draft of the manuscript was written by Xinyi Lu and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

FUNDING INFORMATION

This research was funded by Cooperative Agreement No. G21AC1005 from the U.S. Geological Survey Southeast Climate Adaptation Science Center, and Cooperative Agreement no. F20AC11372 from the U.S. Fish and Wildlife Service.

CONFLICT OF INTEREST STATEMENT

The authors have no relevant financial or non-financial interests to disclose.

DATA AVAILABILITY STATEMENT

Data available via ScienceBase https://www.sciencebase.gov/catalog/item/6439af5dd34ee8d4ade231f8 (https://doi.org/10.5066/P9DQID6G; Lu et al., 2024).

ORCID

Xinyi Lu https://orcid.org/0000-0003-0271-4878
Yoichiro Kanno https://orcid.org/0000-0001-8452-5100

REFERENCES

Abadi, F., Gimenez, O., Ullrich, B., Arlettaz, R., & Schaub, M. (2010). Estimation of immigration rate using integrated population models. *Journal of Applied Ecology*, *47*, 393–400.

Benda, L., Poff, N. L., Miller, D., Dunne, T., Reeves, G., Pess, G., & Pollock, M. (2004). The network dynamics hypothesis: How channel networks structure riverine habitats. *BioScience*, *54*, 413–427.

Berliner, L. M. (1996). Hierarchical Bayesian time series models. In K. M. Hanson & R. N. Silver (Eds.), *Maximum entropy and Bayesian methods* (Vol. *79*, pp. 15–22). Springer.

Bertrand, K. N., Gido, K. B., & Guy, C. S. (2006). An evaluation of single-pass versus multiple-pass backpack electrofishing to estimate trends in species abundance and richness in prairie streams. *Transactions of the Kansas Academy of Science*, *1903*, 131–138.

Besag, J., & Kooperberg, C. (1995). On conditional and intrinsic autoregressions. *Biometrika*, *82*, 733–746.

Bohlin, T., Hamrin, S., Heggberget, T. G., Rasmussen, G., & Saltveit, S. J. (1989). Electrofishing—Theory and practice with special emphasis on salmonids. *Hydrobiologia*, *173*, 9–43.

Brown, J. H. (2014). Why are there so many species in the tropics? *Journal of Biogeography*, *41*, 8–22.

Campbell Grant, E. H., Nichols, J. D., Lowe, W. H., & Fagan, W. F. (2010). Use of multiple dispersal pathways facilitates amphibian persistence in stream networks. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 6936–6940.

Cao, Y., Hinz, L., Metzke, B., Stein, J., & Holtrop, A. (2016). Modeling and mapping fish abundance across wadeable streams of Illinois, USA, based on landscape-level environmental variables. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*, 1031–1046.

Chaput-Bardy, A., Lemaire, C., Picard, D., & Secondi, J. (2008). In-stream and overland dispersal across a river network influences gene flow in a freshwater insect, *Calopteryx splendens*. *Molecular Ecology*, *17*, 3496–3505.

Deschênes, J., & Rodrguez, M. A. (2007). Hierarchical analysis of relationships between brook trout (*Salvelinus fontinalis*) density and stream habitat features. *Canadian Journal of Fisheries and Aquatic Sciences*, *64*, 777–785.

DeWeber, J. T., & Wagner, T. (2015). Predicting brook trout occurrence in stream reaches throughout their native range in the eastern United States. *Transactions of the American Fisheries Society*, *144*, 11–24.

Dunham, J., & Rieman, B. (1999). Metapopulation structure of bull trout: Influences of physical, biotic, and geometrical landscape characteristics. *Ecological Applications*, *9*, 642–655.

Ebersole, J. L., Colvin, M. E., Wigington, P. J., Jr., Leibowitz, S. G., Baker, J. P., Church, M. R., Compton, J. E., & Cairns, M. A. (2009). Hierarchical modeling of late-summer weight and summer abundance of juvenile coho salmon across a stream network. *Transactions of the American Fisheries Society*, *138*, 1138–1156.

EBTJV (Eastern Brook Trout Joint Venture). (2018). Eastern brook trout roadmap to conservation. https://easternbrooktrout.org/science-data/reports/eastern-brook-trout-roadmap-to-conservation-2018/view

ESRI. (2011). *ArcGIS Desktop: Release 10*. Environmental Systems Research Institute.

Fausch, K. D., Torgersen, C. E., Baxter, C. V., & Li, H. W. (2002). Landscapes to riverscapes: Bridging the gap between research and conservation of stream fishes: A continuous view of the river is needed to understand how processes interacting among scales set the context for stream fishes and their habitat. *BioScience*, *52*, 483–498.

Frissell, C. A., Liss, W. J., Warren, C. E., & Hurley, M. D. (1986). A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. *Environmental Management*, *10*, 199–214.

Hazzard, A. (1932). Some phases of the life history of the eastern brook trout, Salvelinus fontinalis Mitchell. *Transactions of the American Fisheries Society*, *62*, 344–350.

Helser, T. E., Punt, A. E., & Methot, R. D. (2004). A generalized linear mixed model analysis of a multi-vessel fishery resource survey. *Fisheries Research*, *70*, 251–264.

Hocking, D. J., Thorson, J. T., O'Neil, K., & Letcher, B. H. (2018). A geostatistical state-space model of animal densities for stream networks. *Ecological Applications*, *28*, 1782–1796.

Hudy, M., Thieling, T. M., Gillespie, N., & Smith, E. P. (2008). Distribution, status, and land use characteristics of subwatersheds within the native range of brook trout in the eastern United States. *North American Journal of Fisheries Management*, *28*, 1069–1085.

Huntsman, B. M., Petty, J. T., Sharma, S., & Merriam, E. R. (2016). More than a corridor: Use of a main stem stream as supplemental foraging habitat by a brook trout metapopulation. *Oecologia*, *182*, 463–473.

Ingram, K. T., Dow, K., Carter, L., Anderson, J., & Sommer, E. K. (2013). *Climate of the Southeast United States: Variability, change, impacts, and vulnerability*. Springer.

Isaak, D. J., Ver Hoef, J. M., Peterson, E. E., Horan, D. L., & Nagel, D. E. (2017). Scalable population estimates using spatial-stream-network (SSN) models, fish density surveys, and national geospatial database frameworks for streams. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*, 147–156.

Jackson, D. A., Peres-Neto, P. R., & Olden, J. D. (2001). What controls who is where in freshwater fish communities–the roles of biotic, abiotic, and spatial factors. *Canadian Journal of Fisheries and Aquatic Sciences*, *58*, 157–170.

Johnson, B. (2020). Stream capture and the geomorphic evolution of the Linville Gorge in the southern Appalachians, USA. *Geomorphology*, *368*, 107360.

Kanno, Y., Letcher, B. H., Hitt, N. P., Boughton, D. A., Wofford, J. E., & Zipkin, E. F. (2015). Seasonal weather patterns drive population vital rates and persistence in a stream fish. *Global Change Biology*, *21*, 1856–1870.

Kanno, Y., Vokoun, J., & Letcher, B. (2014). Paired stream–air temperature measurements reveal fine-scale thermal heterogeneity within headwater brook trout stream networks. *River Research and Applications*, *30*, 745–755.

Katzfuss, M. (2017). A multi-resolution approximation for massive spatial datasets. *Journal of the American Statistical Association*, *112*, 201–214.

Kazyak, D. C., Lubinski, B. A., Kulp, M. A., Pregler, K. C., Whiteley, A. R., Hallerman, E., Coombs, J. A., Kanno, Y., Rash, J. M., & Morgan, R. P. (2022). Population genetics of brook trout in the southern Appalachian Mountains. *Transactions of the American Fisheries Society*, *151*, 127–149.

Kovach, R. P., Muhlfeld, C. C., Al-Chokhachy, R., Dunham, J. B., Letcher, B. H., & Kershner, J. L. (2016). Impacts of climatic variation on trout: A global synthesis and path forward. *Reviews in Fish Biology and Fisheries*, *26*, 135–151.

Kruse, C. G., Hubert, W. A., & Rahel, F. J. (1998). Single-pass electrofishing predicts trout abundance in mountain streams with sparse habitat. *North American Journal of Fisheries Management*, *18*, 940–946.

Lacoursière-Roussel, A., Côté, G., Leclerc, V., & Bernatchez, L. (2016). Quantifying relative fish abundance with eDNA: A promising tool for fisheries management. *Journal of Applied Ecology*, *53*, 1148–1157.

Larson, G. L., & Moore, S. E. (1985). Encroachment of exotic rainbow trout into stream populations of native brook trout in the southern Appalachian Mountains. *Transactions of the American Fisheries Society*, *114*, 195–203.

Leach, C. B., Williams, P. J., Eisaguirre, J. M., Womble, J. N., Bower, M. R., & Hooten, M. B. (2022). Recursive Bayesian computation facilitates adaptive optimal design in ecological studies. *Ecology*, *103*, e03573.

Letcher, B. H., Schueller, P., Bassar, R. D., Nislow, K. H., Coombs, J. A., Sakrejda, K., Morrissey, M., Sigourney, D. B., Whiteley, A. R., O'Donnell, M. J., & Dubreuil, T. L. (2015). Robust estimates of environmental effects on population vital rates: An integrated capture–recapture model of seasonal brook trout growth, survival and movement in a stream network. *Journal of Animal Ecology*, *84*, 337–352.

Lu, X., Hooten, M. B., Raiho, A. M., Swanson, D. K., Roland, C. A., & Stehn, S. E. (2023). Latent trajectory models for spatio-temporal dynamics in Alaskan ecosystems. *Biometrics*, *79*, 3664–3675.

Lu, X., Kanno, Y., Valentine, G. P., Rash, J. M., & Hooten, M. B. (2024). Data from: Using multi-scale spatial models of dendritic ecosystems to infer abundance of a stream salmonid. https://doi.org/10.5066/P9DQID6G

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, *83*, 2248–2255.

Matson, R., Delanty, K., Shephard, S., Coghlan, B., & Kelly, F. (2018). Moving from multiple pass depletion to single pass timed electrofishing for fish community assessment in wadeable streams. *Fisheries Research*, *198*, 99–108.

Miller, W. L., Snodgrass, J. W., & Gasparich, G. E. (2015). The importance of terrestrial dispersal for connectivity among headwater salamander populations. *Ecosphere*, *6*, 1–9.
Morita, K. (2018). Assessing the long-term causal effect of trout invasion on a native charr. *Ecological Indicators*, *87*, 189–192.
Peterson, E., & Ver Hoef, J. (2014). Stars: An ArcGIS toolset used to calculate the spatial information needed to fit spatial statistical models to stream network data. *Journal of Statistical Software*, *56*, 1–17.
Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M.-J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Ashley Steel, E., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: An integrated network perspective. *Ecology Letters*, *16*, 707–719.
Peterson, J. T., Thurow, R. F., & Guzevich, J. W. (2004). An evaluation of multipass electrofishing for estimating the abundance of stream-dwelling salmonids. *Transactions of the American Fisheries Society*, *133*, 462–475.
Petty, J. T., Hansbarger, J. L., Huntsman, B. M., & Mazik, P. M. (2012). Brook trout movement in response to temperature, flow, and thermal refugia within a complex Appalachian riverscape. *Transactions of the American Fisheries Society*, *141*, 1060–1073.
R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
Ricklefs, R. E. (1987). Community diversity: Relative roles of local and regional processes. *Science*, *235*, 167–171.
Rissler, L. J., Wilbur, H. M., & Taylor, D. R. (2004). The influence of ecology and genetics on behavioral variation in salamander populations across the Eastern Continental Divide. *The American Naturalist*, *164*, 201–213.
Roghair, C. N., Dolloff, C. A., & Underwood, M. K. (2002). Response of a brook trout population and instream habitat to a catastrophic flood and debris flow. *Transactions of the American Fisheries Society*, *131*, 718–730.
Rosenberger, A. E., & Dunham, J. B. (2005). Validation of abundance estimates from mark–recapture and removal techniques for rainbow trout captured by electrofishing in small streams. *North American Journal of Fisheries Management*, *25*, 1395–1410.
Rosenfeld, J. (2003). Assessing the habitat requirements of stream fishes: An overview and evaluation of different approaches. *Transactions of the American Fisheries Society*, *132*, 953–968.
Royle, J. A. (2006). Site occupancy models with heterogeneous detection probabilities. *Biometrics*, *62*, 97–102.
Santos-Fernandez, E., Ver Hoef, J. M., Peterson, E. E., McGree, J., Isaak, D. J., & Mengersen, K. (2022). Bayesian spatio-temporal models for stream networks. *Computational Statistics & Data Analysis*, *170*, 107446.
Schaub, M., & Abadi, F. (2011). Integrated population models: A novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, *152*, 227–237.
Scheuerell, M. D., Ruff, C. P., Anderson, J. H., & Beamer, E. M. (2021). An integrated population model for estimating the relative effects of natural and anthropogenic factors on a threatened population of steelhead trout. *Journal of Applied Ecology*, *58*, 114–124.
Terui, A., Kim, S., Pregler, K. C., & Kanno, Y. (2021). Non-random dispersal in sympatric stream fishes: Influences of natural disturbance and body size. *Freshwater Biology*, *66*, 1865–1875.
Thomaz, A. T., Christie, M. R., & Knowles, L. L. (2016). The architecture of river networks can drive the evolutionary dynamics of aquatic populations. *Evolution*, *70*, 731–739.
Thornton, M., Shrestha, R., Wei, Y., Thornton, P., Kao, S., & Wilson, B. (2022). *Daymet: Daily surface weather data on a 1-km grid for North America, version 4 R1*. ORNL DAAC.
Thurow, R. E., & Schill, D. J. (1996). Comparison of day snorkeling, night snorkeling, and electrofishing to estimate bull trout abundance and size structure in a second-order Idaho stream. *North American Journal of Fisheries Management*, *16*, 314–323.
U.S. Geological Survey. (2016). NHDPlus. Version 2.
Uno, H., & Power, M. E. (2015). Mainstem-tributary linkages by mayfly migration help sustain salmonids in a warming river network. *Ecology Letters*, *18*, 1012–1020.
Valentine, G. P., Lu, X., Childress, E. S., Dollof, A. C., Hitt, N. P., Kulp, M. A., Letcher, B. H., Pregler, K. C., Rash, J. M., Hooten, M. B., & Kanno, Y. (2024). Spatial asynchrony and cross-scale climate interactions in populations of a cold-water stream fish. *Global Change Biology*, *30*, e17029.
Ver Hoef, J. M., & Peterson, E. E. (2010). A moving average approach for spatial statistical models of stream networks. *Journal of the American Statistical Association*, *105*, 6–18.
Ver Hoef, J. M., Peterson, E. E., Clifford, D., & Shah, R. (2014). SSN: An R package for spatial statistical modeling on stream networks. *Journal of Statistical Software*, *56*, 1–45.
Ver Hoef, J. M., Peterson, E. E., Hooten, M. B., Hanks, E. M., & Fortin, M.-J. (2018). Spatial autoregressive models for statistical inference from ecological data. *Ecological Monographs*, *88*, 36–59.
Warren, D. R., Robinson, J. M., Josephson, D. C., Sheldon, D. R., & Kraft, C. E. (2012). Elevated summer temperatures delay spawning and reduce redd construction for resident brook trout (*Salvelinus fontinalis*). *Global Change Biology*, *18*, 1804–1811.
Wenger, S. J., Stowe, E. S., Gido, K. B., Freeman, M. C., Kanno, Y., Franssen, N. R., Olden, J. D., Poff, N. L., Walters, A. W., & Bumpers, P. M. (2022). Simple statistical models can be sufficient for testing hypotheses with population time-series data. *Ecology and Evolution*, *12*, e9339.
White, S. L., Hanks, E. M., & Wagner, T. (2020). A novel quantitative framework for riverscape genetics. *Ecological Applications*, *30*, e02147.
Wikle, C. K., Berliner, L. M., & Cressie, N. (1998). Hierarchical Bayesian space-time models. *Environmental and Ecological Statistics*, *5*, 117–154.
Xu, C., Letcher, B. H., & Nislow, K. H. (2010). Context-specific influence of water temperature on brook trout growth rates in the field. *Freshwater Biology*, *55*, 2253–2264.
Zippin, C. (1958). The removal method of population estimation. *The Journal of Wildlife Management*, *22*, 82–90.

How to cite this article: Lu, X., Kanno, Y., Valentine, G. P., Rash, J. M., & Hooten, M. B. (2024). Using multi-scale spatial models of dendritic ecosystems to infer abundance of a stream salmonid. *Journal of Applied Ecology*, *61*, 1703–1715. https://doi.org/10.1111/1365-2664.14665

APPENDIX A

Prior distributions

$$p_A,\ p_B \sim \text{Beta}(60, 40),\ \text{for YOY},$$

$$p_A,\ p_B \sim \text{Beta}(70, 30),\ \text{for adults},$$

$$\boldsymbol{\beta},\ \boldsymbol{\theta},\ \boldsymbol{\mu} \sim \text{N}(\mathbf{0}, I),$$

$$\log(\sigma_\epsilon^2),\ \log(\sigma_\nu^2),\ \log(\sigma_\psi^2),\ \log(\sigma_\gamma^2),\ \log(\sigma_\eta^2) \sim \text{N}(0, 1),$$

$$\phi_\nu,\ \phi_\psi \sim \text{Unif}(0, 10).$$

APPENDIX B

Simulation

We demonstrated using a simulation study that our hierarchical model was able to recover data-generating parameters and showed desirable predictive performance. Because the focus of this simulation was on identifying spatial parameters and prediction, we defined a uniform detection probability and omitted temporal trends and random effects. We generated $n = 63$ sampling sites and 4 hydrologically separated spatial networks using the createSSN function from the SSN package in R (Ver Hoef et al., 2014; Figure B1). We simulated spatial covariates $\mathbf{x}$ and spatio-temporal covariates $\boldsymbol{h}$ for $T = 10$ years from standard normal distributions. For the multi-scale spatial structure, we designated each network to a distinct watershed and simulated watershed level random effects using Equation 4, where $\boldsymbol{W}$ is defined using inverse Euclidean distances between mean site locations within each network. We simulated segment level random effects using Equation 5 and network level random effects using Equation 6. We obtained population densities, λ, using Equation 3. We let $A_i = 1$ for all sites and sampled true abundance, N, using Equation 2. We then sampled observed abundance, y, using Equation 1. We fit our model to the simulated data using an MCMC algorithm and summarized the marginal posterior distributions in Table B1.

We compared the predictive performance of our multi-scale process model (Equation 3) to models with partial or no random effects using a five-fold cross validation. At each fold, we separated the data into a training set and a test set. We fit the contending models to the training set and obtained predicted log densities on the test set using universal kriging (kriging was not necessary for the model without random effects). Model performance was then measured by posterior predictive log likelihood (PPLL) and root mean squared error (RMSE), and averaged across the folds. The full model (watershed + stream + network) demonstrated the best predictive performance (Table B2), thereby substantiating our multi-scale spatial specification.

TABLE B1 True parameters and their marginal posterior distributions for the simulation.

Parameter	True	Posterior mean (95% CI)
p	0.7	0.69 (0.65, 0.72)
β_1	1	1.01 (0.94, 1.08)
β_2	−0.5	−0.49 (−0.58, −0.40)
θ_1	0.4	0.43 (0.36, 0.50)
θ_2	0.8	0.76 (0.69, 0.83)
σ_ϵ^2	0.01	0.010 (0.007, 0.014)
σ_ν^2	0.01	0.009 (0.006, 0.013)
σ_ψ^2	0.01	0.010 (0.007, 0.014)
ϕ_ν	2	3.55 (1.37, 4.94)
ϕ_ψ	10	11.84 (6.88, 14.83)

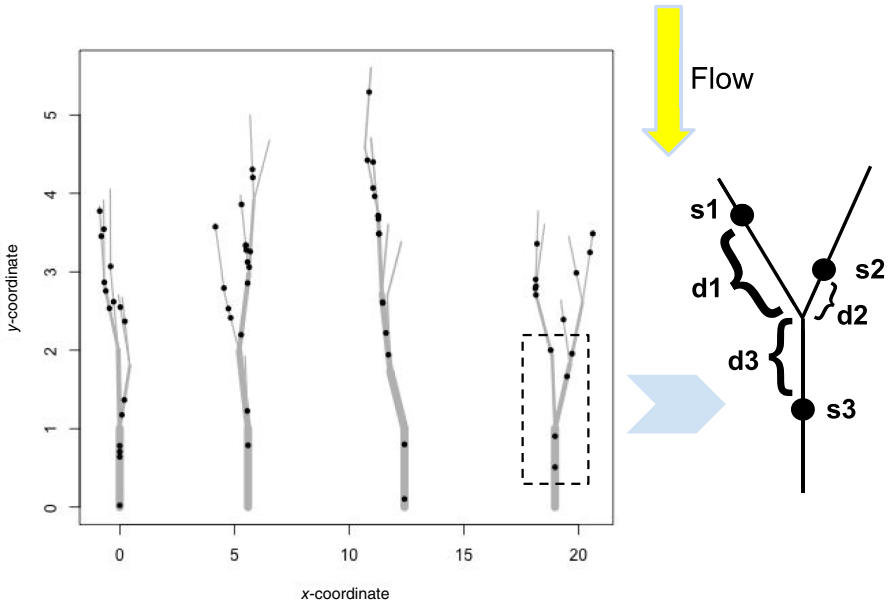

FIGURE B1 Simulated network structure (left) where sampling sites are represented by points, and a conceptual representation of the tail-down SSN model (right). Sites s1 and s3 are flow-connected and their network distance is measured by d1 + d3; sites s1 and s2 are flow-unconnected and their network distance is measured by d1 + d2.

Journal of Applied Ecology BRITISH ECOLOGICAL SOCIETY

TABLE B2 Predictive scores of models with different spatial structures. Higher scores in PPLL and lower scores in RMSE indicate better predictive performance.

Model	PPLL	RMSE
Watershed + Segment + Network	−1.82	1.98
Watershed + Segment	−1.82	1.99
Network	−2.02	2.17
No random effect	−2.10	2.18

TABLE C1 Estimation and predictive scores of the proposed and the naive models using brook trout count data by life-stage. Lower scores in RMSE indicate better performance.

	$RMSE_{est}$		$RMSE_{pred}$	
Model	YOY	Adult	YOY	Adult
Proposed	1.81	1.31	21.67	14.06
Simplistic	6.95	4.36	21.98	17.21

APPENDIX C

Model comparison

We compared our proposed method in Section 2.3 to a hierarchical model with the same data model and prior specifications, but a simplistic process model as follows,

$$\log(\lambda_{i,t}) = \boldsymbol{x}_i'\boldsymbol{\beta} + \boldsymbol{h}_{i,t}'\boldsymbol{\theta}_i + \gamma_i \log t + \eta_t.$$

We first compared estimation errors between the two models using all brook trout data, for YOY and adult, respectively. Although true abundance was unknown, we estimated observed abundance by sampling from its posterior predictive distribution,

$$\tilde{y}_{i,t}^{(q)} \sim \text{Binomial}\left(N_{i,t}^{(q)}, 1 - \left(1 - p_{i,t}^{(q)}\right)^{J_{i,t}}\right), \tag{7}$$

where $J_{i,t}$ denotes the number of survey passes at site i in year t, $\tilde{y}_{i,t} \equiv \sum_{j=1}^{J_{i,t}} y_{i,t,j}$ denotes the total number of fish captured across passes, and $N_{i,t}^{(q)}$ and $p_{i,t}^{(q)}$ are the qth posterior samples for $q = 1, \ldots, Q$. The estimation RMSE was calculated as follows,

$$\text{RMSE}_{\text{est}} = \frac{1}{Q}\sum_{q=1}^{Q}\left(\frac{1}{|\mathcal{O}|}\sum_{(i,t)\in\mathcal{O}}\left(\%\tilde{y}_{i,t} - \%\tilde{y}_{i,t}^{(q)}\right)^2\right)^{1/2},$$

where $\mathcal{O}$ denotes the set of site-year combinations that underwent survey. We then compared predictive performance via a ten-fold cross-validation. At each non-overlapping fold, we designated 90% of the observed data as the training set, and the remaining 10% as the test set. We fit both models to the training set and used posterior samples to predict abundance in the test set using Equation 7. The predictive RMSE was calculated as follows,

$$\text{RMSE}_{\text{pred}} = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{Q}\sum_{q=1}^{Q}\left(\frac{1}{|\mathcal{T}_m|}\sum_{(i,t)\in\mathcal{T}_m}\left(\%\tilde{y}_{i,t} - \%\tilde{y}_{i,t}^{(q)}\right)^2\right)^{1/2},$$

where $m = 1, \ldots, M$ denotes the fold, and $\mathcal{T}_m$ denotes the site-year combinations in the mth test set. Our proposed model outperformed the simplistic model in both abundance estimation and out-of-sample prediction (Table C1).