Food Webs 33 (2022) e00249

Contents lists available at ScienceDirect

Food Webs

journal homepage: www.elsevier.com/locate/fooweb

Spatial stream modeling of Louisiana Waterthrush (*Parkesia motacilla*) foraging substrate and aquatic prey in a watershed undergoing shale gas development

Mack W. Frantz [a,*], Petra B. Wood [b], Steven C. Latta [c]

[a] West Virginia Cooperative Fish and Wildlife Research Unit, Davis College, School of Natural Resources, West Virginia University, P.O. Box 6125, Morgantown, WV, USA
[b] U.S. Geological Survey, West Virginia Cooperative Fish and Wildlife Research Unit, West Virginia University, P.O. Box 6125, Morgantown, WV, USA
[c] National Aviary, 700 Arch St., Pittsburgh, PA 15212, USA

ARTICLE INFO

Keywords:
Marcellus shale
Bioindicator
Spatial stream network model
Macroinvertebrate
Foraging

ABSTRACT

We demonstrate the use of spatial stream network models (SSNMs) to explore relationships between a semi-aquatic bioindicator songbird, Louisiana Waterthrush (*Parkesia motacilla*), and stream monitoring and benthic macroinvertebrate data in an area undergoing shale gas development. SSNMs allowed us to account for spatial autocorrelation inherent to these environmental data types and stream properties that traditional modeling approaches cannot capture to elucidate factors that affect waterthrush foraging locations. We monitored waterthrush along 58.1 km of 1st- and 2nd-order headwater stream tributaries ($n = 14$) in northwestern West Virginia over a two year period (2013–2014), sampled benthic macroinvertebrates in waterthrush territories, and collected wetted perimeter stream channel and water chemistry data along a 50 m fixed point stream grid. Spatial models outperformed traditional regression models and made a statistical difference in whether stream covariates of interest were considered relatable to waterthrush foraging. Waterthrush foraging probability index (FPI) was greater in areas where family and genus-level multi-metric indices of biotic stream integrity were higher (i.e. WVSCI and GLIMPSS). Waterthrush were found foraging both among stream flow connected and unconnected sampled sites on relatively further upstream locations where WVSCI and GLIMPSS were predicted to be highest. While there was no significant relationship found between FPI and shale gas land use on a catchment area scale, further information on waterthrush trophic dynamics and bioaccumulation of surface contaminants is needed before establishing the extent to which waterthrush foraging may be affected by shale gas development.

1. Introduction

A natural property of ecological data is autocorrelation where nearby objects are more likely to exhibit the same patterns for reasons not due to chance (Legendre, 1993). Since the advent of classical statistics that could not account for the non-independence of ecological observations (e.g. Fisher, 1935), there is an excess of spatial models for ecology that are not created equally (Dormann et al., 2007), designed mainly for terrestrial ecology (e.g. Fortin and Dale, 2005). Large, long-term datasets are being collected globally on streams as part of biomonitoring efforts to determine environmental conditions and change (Buss et al., 2015), making it increasingly important to choose appropriate statistical methods for valid assessment of stream network data (Rushworth et al., 2015). Spatial models that incorporate the unique properties of streams as dendritic networks with restricted, directed movement of resources through the landscape would be more ideal than the current trend of adopting terrestrial modeling techniques to streams (Isaak et al., 2014). In the last few years, a series of spatial stream network models (SSNMs) were created that account for stream properties (e.g. branching, flow direction and connectivity, confluences) and allow analysis of typical environmental monitoring data via stream-based spatial-weighting and autocovariance structures (Cressie et al., 2006; Ver Hoef et al., 2006; Peterson and Ver Hoef, 2010). Spatial autocorrelation is a confounding source of variability for covariates of interest on a stream network, where dismissing or ignoring it using traditional methods (i.e. generalized linear regression models that impose independence between

* Corresponding author at: West Virginia Division of Natural Resources, 738 Ward Road, Elkins, WV, USA.
E-mail address: mack.w.frantz@wv.gov (M.W. Frantz).

https://doi.org/10.1016/j.fooweb.2022.e00249
Received 31 December 2021; Received in revised form 9 June 2022; Accepted 27 July 2022
Available online 2 August 2022

observations) can lead to important information being discarded or lack of statistical inference (Legendre, 1993).

Wildlife communities in the Appalachian region, particularly in forested freshwater ecosystems (Dunscomb et al., 2014; Evans and Kiesecker, 2014), are threatened by unconventional shale gas development due to the rate at which development can outpace implementation of best management practices (Brittingham et al., 2014). Shale gas wells in the Marcellus shale region are commonly within 100–300 m of stream channels, and often even closer to headwater drainage areas (Entrekin et al., 2011). Headwater streams are the critical sources of water, sediment, organic matter, and nutrients for the rest of the system (Gomi et al., 2002), and are therefore vital for ecological integrity (Freeman et al., 2007). Furthermore, headwater streams, despite their predominance of drainage area and total stream length, are largely overlooked for protection or regulation contrary to their potential effect on downstream reaches and aquatic life (MacDonald and Coe, 2007).

The Louisiana Waterthrush (*Parkesia motacilla*), hereafter waterthrush, is an established biological indicator of aquatic stream integrity (O'Connell et al., 2000; Mulvihill et al., 2008) and species of conservation concern (USFWS, 2008). Waterthrush feed primarily on benthic macroinvertebrates (Mattsson et al., 2009) and breed along forested headwater streams, reaching some of their highest abundances in the Marcellus shale region (Sauer et al., 2014). Over a six-year period, areas disturbed by shale gas negatively affected waterthrush riparian habitat quality, nest productivity, and nest survival suggesting potential long-term population consequences (Frantz et al., 2018a). Given the propensity for shale gas in the Appalachian region to be developed on ridgetops near headwater streams (Cook et al., 2015), there is a strong need to evaluate how down-stream communities, both aquatic and terrestrial, can be affected by potential surface water pollution (Entrekin et al., 2011). In particular, the food webs along the aquatic-terrestrial interface may be indirectly influenced by surface water contamination depending on where the organisms reside or forage along the stream network. Waterthrush are known to compensate for the loss of food resources by increasing their territory sizes and foraging in nearby unimpacted areas (Mulvihill et al., 2008; Frantz et al., 2018a). As such, a more detailed study of headwater streams and foraging of a stream-dependent organism would shed some light on whether shale gas development is influencing these resources. Frantz et al. (2019b) recently found there may be a disturbance threshold at which waterthrush demography respond to aquatic prey changes using a spatial modeling approach, but did not assess locations where waterthrush were observed foraging nor used models designed specifically for stream-based spatial weighting and autocovariance structures.

In this study, we used SSNMs to evaluate their utility in quantifying characteristics of waterthrush foraging areas based on 1) water chemistry, 2) a waterthrush foraging score based on stream channel data, 3) shale gas land use based on reach contribution, and 4) multi-metric indices of biotic stream integrity at the family and genus level. We hypothesized that waterthrush foraging would more likely occur in areas with higher biotic stream integrity and with higher abundance of pollution sensitive aquatic prey such as Ephemeroptera, Plecoptera, and Trichoptera (EPT) believed to be the waterthrush's preferred prey items (Mattsson et al., 2009). We also hypothesized that waterthrush foraging activity would be negatively related to areas of higher shale gas land use and water chemistry (i.e. higher conductivity, total dissolved solids, pH, and water temperature) as surface water pollution (e.g. Latta et al., 2015) and decreased riparian habitat quality (e.g. Wood et al., 2016; Frantz et al., 2018b) from shale gas development may negatively alter aquatic prey communities (Johnson et al., 2015).

2. Material and methods

2.1. Study area

We studied waterthrush along 58.1 km of 1st- and 2nd-order headwater stream tributaries ($n = 14$) at Lewis Wetzel Wildlife Management Area (LWWMA) located in northwestern West Virginia (Fig. 1). Our waterthrush foraging study occurred in 2013 and 2014 as part of a larger waterthrush demography study over a six year period (2009–2011, 2013–2015). The study area lies within the Permian Hills subdivision of the Western Allegheny Plateau Ecoregion, an area of deeply dissected topography and relatively continuous Appalachian Oak and Mixed-Mesophytic Forest (Woods et al., 1999) with elevations of 221–480 m. The study area overlays the Marcellus-Utica shale region and occurs where waterthrush reach their highest densities within the central Appalachians (Sauer et al., 2014).

Prior to our study, LWWMA was 95% forested with the first unconventional gas well development and activity, hereafter shale gas, starting in 2007 (Farwell et al., 2016). Shale gas at our study area and within the surrounding region since then has rapidly increased (WVGES West Virginia Geological and Economic Survey, 2015). By 2015, LWWMA was 91% forested with forest loss primarily due to shale gas development (Farwell et al., 2016). Over the six year study period, gas well development activities included building of conventional and Marcellus well pads, timbering for yet unbuilt well pads, the expansion of existing road and pipeline infrastructure, and the construction of new infrastructure. Early in the study (2009–2010), the majority of Marcellus wells and their water holding ponds were located along the main stem of Buffalo Run where the majority of our headwater study streams empty. Thus, although a few Marcellus well pads were located along our study streams, they tended to primarily impact the lower portions. Between the 2010 and 2011 breeding seasons, shale gas development activities began to increase on the ridgetops (Frantz et al., 2018a, 2018b). Therefore during the waterthrush foraging study in 2013–2014 the whole downstream network of some streams became disturbed by sedimentation and surface runoff from ridgetop activity (Frantz et al., 2018a, 2018b).

2.2. Mapping of streams and disturbance

Within a Geographic Information System (GIS), we used a sequence of leaf-on and leaf-off aerial (e.g. NAIP) imagery and extensive ground-truthing to manually digitize areas of disturbance within the study area (see Frantz et al., 2018a for full description). All disturbances were classified as shale gas related (e.g. well pads and associated road and pipeline infrastructure) or as being unrelated or pre-existing (e.g. forest roads, recent even-aged timber harvests, and various types of existing clearings). We classified a few conventional impacts (i.e., stream-side vertical pump jacks) as related to shale gas development because their pads were managed in conjunction with nearby shale gas infrastructure and because their targeted formation, even though they remained shallow after development, was listed as Marcellus. Gas well records (WVDEP, 2015; WVGES West Virginia Geological and Economic Survey, 2015) were used to verify target shale formations, drilling status, and start dates for all well disturbances. Lengths of each study stream (average length 4.1 ± 0.54 km, range 0.95–7.4 km) were calculated in GIS using a 3D functional surface length tool and a 3 m resolution digital elevation model to account for topography, and study streams were defined to have a drainage basin of 9 ha (i.e. <100 ha, Swanson et al., 1998) to delineate the uppermost headwater reaches (24 k scale or higher resolution; e.g. Strager et al., 2009).

2.3. Waterthrush foraging observations

We mapped waterthrush territories along 14 streams with varying degrees of shale gas disturbance in 2013–2014. Waterthrush territories were delineated typically from early April to late June each year. Standardized territory mapping (Robbins, 1970; Bibby et al., 1992) included ≥6 (average 11.5 ± 0.6) visits along each stream reach, with visits preceding peak incubation initiation, and visits within 4 h after sunrise to ensure high rates of detection (Mattsson and Cooper, 2006).

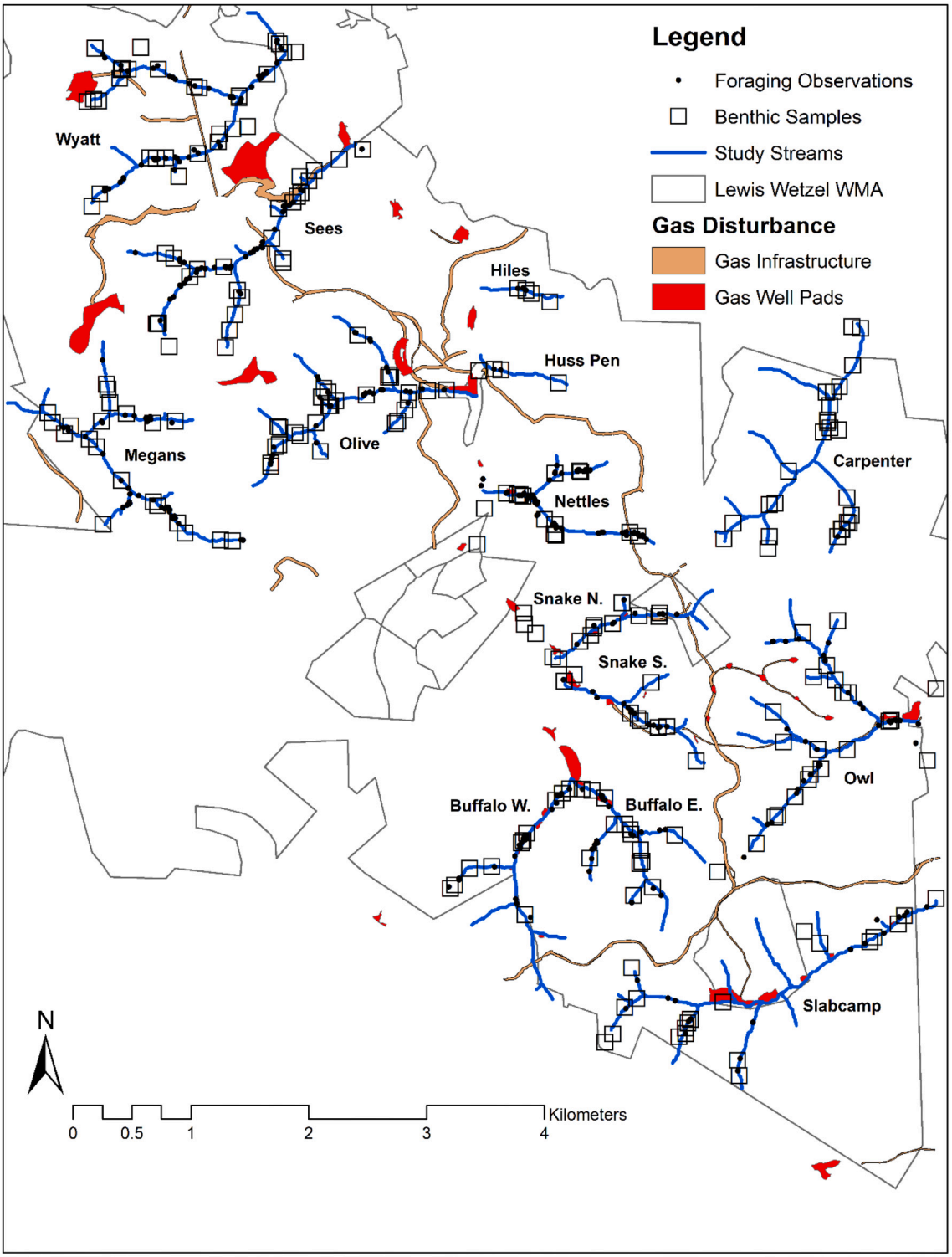

Fig. 1. Study area map. Our study area, Lewis Wetzel Wildlife Management Area (LWWMA), lies within the Marcellus-Utica shale basin. We observed Louisiana Waterthrush foraging on fourteen 1st and 2nd order headwater streams and collected benthic macroinvertebrate samples during 2013–2014.

While delineating territories, observations were made whether waterthrush were foraging or not (e.g. singing, territorial dispute, flying) and mapped with a WAAS-enabled Garmin 60CSX GPS unit with accuracy ≤5 m. We recorded observations of both male and female waterthrush since neither foraging rate nor microhabitat use differs between the sexes (Robinson, 1990). When a waterthrush was detected, we only approached close enough for observation without perceptibly influencing behavior (Vitz and Rodewald, 2010). Waterthrush are just as likely to be "loafing" as they are foraging in a given location (Robinson, 1990), therefore any observation where a waterthrush was observed to flush when first encountered was categorized as non-foraging. We varied the order and time of day we monitored study streams to prevent any time of day effects (Shields, 1977), so waterthrush observations overall should not be influenced by our presence nor time of day. We concurrently searched for and monitored waterthrush nests during these visits.

2.4. Wetted perimeter data

To evaluate in-stream riparian habitat quality for foraging waterthrush, we collected several stream channel metrics every 50 m along each stream using a wetted perimeter protocol designed to determine optimal waterthrush foraging substrate locations (Mulvihill and Latta, Unpublished results; Master et al., 2005). Sampling locations were assigned in GIS prior to entering the field using Linear Referencing, which created routes along the stream that accounted for stream bend. Points were not sampled that fell on private property ($n = 58$ points), on completely dry sections of streams ($n = 30$ points), or stream sections with water flow too high for waterthrush (n = 5 points). This resulted in sampling of 1121 points among the fourteen headwater streams. Wetted perimeter data were collected in 2013 for Olive Run and Wyatt Run, and for the remainder of the streams in 2015. The data were collected in late June–July to be representative of year-round baseflow conditions (i.e. flow between rainfall events) as verified by field technicians and the principal investigator present for multiple years of the study.

At each 50 m sampling point, a small-link metal chain was draped across the stream where water during a high flow event at the edge of the stream meets the bank. This resulted with the ends of the chain curving up each side of the stream bank and stopping at the uppermost portion at which water could flow, creating a "U-shape" with the chain. The chain conformed to all irregularities in the stream channel (e.g. rocks, logs) comprising the bottom substrate, including those sticking up above water. Using a meter tape, the lengths of chain that were above water and stretched over rocks or logs were measured (in meters) as an exposed point measurement with the number of exposed points tallied. The chain was then removed from the water and stretched to full length, with the distance between the two points that marked the edges of the stream measured. The full chain length was a wetted perimeter measurement, where a wet distance could be calculated from subtracting the exposed distance. We measured stream depth (in cm) at five regularly spaced intervals across the same start and end points of the chain: water's edge on both sides, a quarter of the way in from middle on both sides, and middle of stream. A waterthrush foraging substrate score was calculated by taking the wetted perimeter value and dividing it by the average stream depth at that point (Mulvihill and Latta, Unpublished results). A higher wetted perimeter to mean depth ratio presumably indicated relatively better foraging substrate, and smaller values poorer foraging substrate (Mulvihill and Latta, Personal communication) which corresponds to the need for sufficient water flow to maximize nest survival at the time of year collected (Frantz et al., 2018a). At every 50 m sampling point, we also collected water chemistry data in the form of pH, total dissolved solids (TDS, g/L), conductivity (μS/cm), and water temperature (°C) with a Hannah Instruments multi-parameter probe to relate results to a macroinvertebrate assemblage study conducted on the study area (Merovich Jr. et al., 2022), and because surface disturbances that can increase runoff and sedimentation increase dissolved solids and conductance (Merriam et al., 2013).

2.5. Benthic macroinvertebrate sampling

Macroinvertebrates occurring in riffle habitat the most adjacent to nest site locations were sampled using a Surber sampler in 2013 and 2014. Nest site samples ($n = 178$) were collected shortly after the nest fledged, failed, or had been abandoned (from mid-June to late July) to assess relative prey availability at the time the site was used by waterthrush. Additional macroinvertebrate samples were collected from waterthrush foraging locations ($n = 65$; average 165 ± 12.6 m from nest site locations) during two timed bouts, one each in May and June. The two bouts were later pooled into one sample after we found no taxa differences between the time periods. During sample collection, we scrubbed rock substrates and disturbed sediment 3-cm below the stream bed within the Surber frame for a total of 3 min (Mattsson and Cooper, 2006). We separated macroinvertebrates from detritus for each sample in the field and stored them in 95% or 70% ethanol.

Macroinvertebrates in benthic samples were sorted, counted, identified to genus level, and body lengths measured by an environmental scientist certified by the Society of Freshwater Scientists in macroinvertebrate identification. For each sample, we calculated several commonly used community metrics both at the family (e.g. family taxa richness) and genus (e.g. genus taxa richness) levels of resolution. These metrics were used to calculate multimetric indices of biotic integrity, one at the family level (West Virginia Stream Condition Index, WVSCI; Gerritsen et al., 2000) and one at the genus level (Genus Level Index of Most Probable Stream Status; GLIMPSS, version CF), which does not require the genus-level identification of Chironomidae (Pond et al., 2013).

2.6. GIS data preparation and stream formatting

In ArcMap GIS 10.2.2 (ESRI, 2011, Redlands, CA), foraging observations and macroinvertebrate sampling points were mapped along with the wetted perimeter point grid plotted along the headwater streams. Waterthrush typically travel no further than 60 m away from their linear territories (Mattsson and Cooper, 2009) and will forage off stream more often as the breeding season progresses (Robinson, 1990). Therefore we placed a 60 m buffer around each wetted perimeter sampling point. Using a spatial join, all foraging and non-foraging observations within those buffers were assigned to the wetted perimeter grid. Metrics from macroinvertebrate sampling points were merged into a single new output if they fell within a 60 m buffer, and then averaged if more than one point fell within a buffer. Any wetted perimeter point that did not have any foraging or non-foraging observations nor macroinvertebrate data were removed from our response variable. A foraging probability index (FPI, 0–100%), our response variable, was derived from the number of foraging observations divided by the total observations (foraging and non-foraging) for the 60 m area. Calculating FPI in this manner gave a mostly continuous index since a 60 m buffer overlapped the 50 m wetted perimeter points, meaning observations and benthic samples could be assigned to more than one nearby wetted perimeter sample point. We defined FPI as a relative index that gauged where waterthrush were most likely to be found foraging. Rather than assume any areas with no waterthrush observations (i.e. points that we did not use for FPI) were non-optimal foraging areas, we reserved these wetted perimeter points for testing model prediction.

Gas variables (% Marcellus pad, % any gas pad (unconventional and conventional), % any gas infrastructure) were created first as reach contributing area (RCA; i.e. catchment area scale) attributes using the STARS (Spatial Tools for the Analysis of River Systems) toolbox (Peterson and Ver Hoef, 2014) in ArcMap GIS as a means of calculating land use. The Accumulate Values Downstream and Watershed Attributes tools were used to create and assign the gas RCA values to sampled points on the stream. To get a percentage contribution of each gas variable, we divided the value assigned to each stream sample by the total watershed area representative of all segment watersheds encompassing the study streams (33.4 km^2 total). Percent Marcellus Pad included three retention ponds that may pose the same concerns to surface water contamination. While we include a metric that includes conventional gas well pads (% Any Gas Pad), all major landscape alterations, development, and activity seen during the duration of our study would not have occurred without shale gas at our study site (Farwell et al., 2016). Percent Any Gas Infrastructure included well pads, pipelines, retention ponds, and access roads leading to well sites.

Stream segment vectors were simplified to avoid converging streams and have minimum pseudonodes (Peterson and Ver Hoef, 2014). In the original study design, parts of Buffalo Run that the headwater streams emptied into were included as part of the boundaries of each study stream since many times a waterthrush territory would border or include part of Buffalo Run (Frantz et al., 2018a). These sections of Buffalo Run were removed from each study stream since each stream can

only have one outlet in SSNMs. Stream segments were also extensively preprocessed to ensure they were digitized in a downward flow direction and any network topology errors removed that may interfere with spatial weighting calculations.

2.7. Model spatial distance and weight preprocessing

All model analyses were done using the Spatial Stream Network (SSN) package (Ver Hoef et al., 2014) in R (R Core Team, 2014) and SSN object preprocessing for import in ArcGIS using the STARS toolbox (Peterson and Ver Hoef, 2014). In order to fit spatial models using spatial weights, we first determined stream segment proportional influence and additive function values using STARS. Stream segment proportional influence was based on RCA for each line segment watershed area (Peterson and Ver Hoef, 2010). Spatial weight was based on RCA since it serves as a surrogate for flow volume (Friedan et al., 2014). These values were contributed in R to create the spatial weights (Ver Hoef et al., 2014). Upstream distance between a stream outlet and each stream segment and sample point were calculated with STARS (Peterson and Ver Hoef, 2014) to be used in R for calculating hydrologic flow-connected and -unconnected distances (see Ver Hoef and Peterson, 2010) in R. The processed dataset was stored and displayed as a Landscape Network (LSN) that included all spatial and geographic relationships for the streams and stream dataset (Theobald et al., 2006; Peterson and Ver Hoef, 2014).

2.8. Model variables, covariance, selection, and evaluation

We initially reviewed all data graphically and through diagnostic tools to test assumptions of normality and applied data transformations if it improved approximation to normality (Zuur et al., 2010). Torgegrams (i.e. semivariograms for streams; see Zimmerman and Ver Hoef, 2017) were used to assess spatial autocorrelation which breaks up the semivariance into flow-connected and -unconnected structures (Ver Hoef et al., 2014). Based on diagnostic evaluation we added a log10 data transformation for temperature, total dissolved solids (TDS, g/L), conductivity (μS/cm), and foraging score.

Our spatial-stream network models (SSNMs) used a mixed-model autocovariance structure consisting of exponential tail-up (TU), tail-down (TD), and exponential Euclidean. TU and TD autocovariance models represent flow-connected and –unconnected relationships along the stream and is based on hydrologic (rather than "traditional" straight-line Euclidean) distance (extensive explanation can be found in Ver Hoef and Peterson, 2010). Euclidean distance was included for comparison since it is a traditional distance. Autocovariance models were not determined a priori (Friedan et al., 2014) as a partial sill, range parameter, and overall nugget effect estimated for each model helps determine relative influence of the components in individual models for either model improvement or removal (Ver Hoef and Peterson, 2010).

We used an exploratory multi-stage model selection process for model evaluation that allowed us to determine autocovariance structure and what covariates to keep for further evaluation (Friedan et al., 2014). We modeled covariates individually since we were interested in which ones had the most predictive power and were statistically significant before comparing the individual covariates in a final model set. As such, there was no need to examine correlations of covariates to avoid multicollinearity. We set $\alpha = 0.10$ to avoid missing any variables that may be of ecological relevance. We first ran a non-spatial linear regression model equivalent for each variable for comparison and evaluation of spatially-influenced properties in spatial models, and because non-spatial linear regression models are traditionally how the variables would have been modeled. Only significant variables ($P < 0.10$) from the non-spatial models were placed into spatial models. Initial covariance structures of spatial models were mixed and fixed to exponential TU, exponential TD, and exponential Euclidean since we expected variability in how spatial weights may affect each covariate. Both non-spatial and spatial models included an estimated nugget covariance effect to account for the influence of variance in geostatistical datasets (Diggle and Robeiro 2007).

Maximum likelihood (ML) estimation was used to estimate parameters of Gaussian models (response variable FPI). Akaike's Information Criterion (AIC; Akaike, 1974; Burnham and Anderson, 2002) was used to compare non-spatial and spatial models which penalized for additional spatial autocovariance structures (Ver Hoef and Peterson, 2010). We considered the model with the lowest AIC value to be the best-supported model for each candidate set, and any models with ΔAIC <2 were considered plausible (Burnham and Anderson, 2002). Including non-spatial models in initial AIC model selection allowed us to determine whether spatial models would outperform traditional regression models. R^2 and root mean square prediction error (RMSPE) based on the observed response variable and leave one out cross validation (LOOCV) predictions were also calculated. Variance decomposition was used to determine the total amount of variation associated with a response variable (Ver Hoef et al., 2014). Predictions from wetted perimeter locations with no waterthrush observations were generated using universal kriging (Cressie, 1993).

The next stage of model selection involved selecting the best autocovariance structure (Friedan et al., 2014). If exponential TU/TD models had a higher partial sill than Euclidean autocovariance, we added Mariah, Spherical, and Linear-with-sill to test before final selection of autocovariance. Final models were evaluated by AIC, ΔAIC, RMSPE, and by examining the influence of each variance component. We mapped and visually examined prediction values ± standard error (SE) as one means of determining overall model performance (Bennett et al., 2013) along with plotting of LOOCV predictions and SEs against the observed data. *Post-hoc* Spearman's Rho correlation index tests in R were used to determine what components of WVSCI, GLIMPSS, or foraging score were associated to FPI if those covariates were found important during model selection.

3. Results

During 2013–2014 we collected 948 foraging and non-foraging observations of waterthrush. Each stream had an overall average of 30.6 ± 7.2 foraging and 37.1 ± 6.2 non-foraging observations (average 67.7 ± 11.1 total observations per stream, range 4–214) collected. We had 318 60-m buffered wetted perimeter sampling points for analysis that included both waterthrush observations and benthic samples, and an additional 103 saved for prediction modeling that had benthic samples but no waterthrush observations. Each stream had an average of 22.7 ± 3.4 sampling points (range 3–42). Each sampling point contained on average 4.1 ± 0.2 waterthrush observations (range 1–30). Average foraging observations at each sampling point were 2.0 ± 0.2 (range 0–30) and average non-foraging observations were 2.0 ± 0.1 (range 0–14). Overall average foraging probability index (FPI) on each stream based on the sampling points was 47.9 ± 4.5% (range 20.3–80.4%, Fig. 2).

Stream temperature, benthic biomass, benthic density, GLIMPSS, and WVSCI were significant in explaining foraging probability index (FPI) according to the non-spatial linear regression models ($P < 0.10$, Table 1). A torgegram for FPI suggested there may be higher spatial autocorrelation between flow-connected sample points at short distances, but both flow-connected and unconnected samples have high autocorrelation (Fig. 3). The torgegram also suggested using both tail up (TU) and tail down (TD) autocovariance structures in initial spatial models to obtain the full range of autocorrelation.

The five significant variables were placed into individual spatial models where only GLIMPSS and WVSCI remained significant ($P < 0.05$, Table 2) to continue model selection and comparison. Given support that tail down (TD) models performed better than tail up (TU) models (Table 3), we added two more tail down variance components for AIC model comparison (Table 4). Traditional, non-spatial regression models

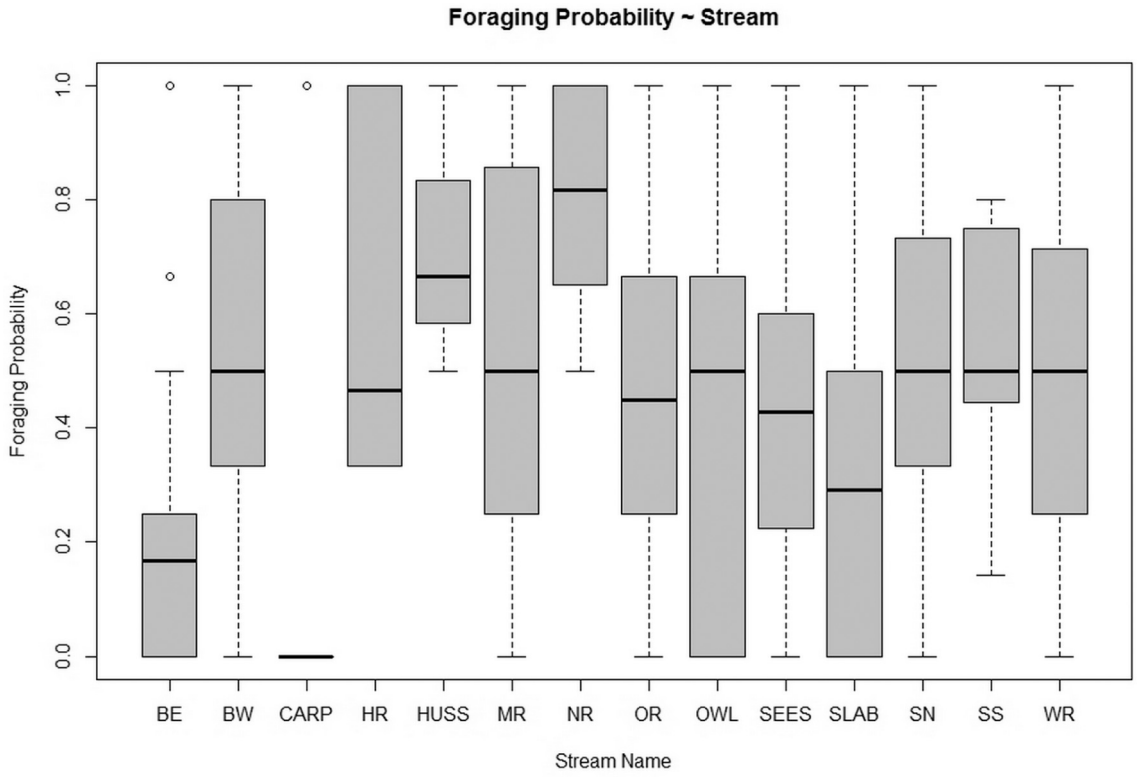

Fig. 2. Foraging probability index by stream. Overall foraging probability index (FPI, 0–100%) on each stream during 2013–2014. FPI was derived from the number of foraging observations/total observations for a 60 m area surrounding each wetted perimeter grid point on the stream. FPI was a relative index that gauged where waterthrush were most likely to be found foraging. Note Carpenter Run (CARP) only had two foraging observations, limiting the ability to make inferences about FPI on that stream.

Table 1
Initial non-spatial linear models (with nugget) to test the relationship between foraging score, water chemistry, macroinvertebrate metrics, shale gas land use and foraging probability index (FPI). Covariates with bolded *P* values were significant at $\alpha = 0.10$. R^2 is a generalized value of model fit and the partial sill (sill minus nugget) was included to assess variance of a covariate without the nugget effect.

Variable	Estimate	SE	*t* value	P value	R^2	Partial sill
Foraging Score	0.046	0.054	0.855	0.393	0.0023	0.122
Temperature (°C)	−1.073	0.619	−1.732	**0.084**	0.01	0.121
pH	0.048	0.077	0.627	0.531	0.0012	0.122
TDS, g/L	−0.052	0.106	−0.494	0.622	0.0001	0.122
µS/cm	−0.055	0.098	−0.554	0.580	0.001	0.122
GLIMPSS	0.004	0.001	3.008	**0.003**	0.03	0.119
WVSCI	0.005	0.002	2.948	**0.003**	0.03	0.119
Biomass	0.101	0.037	2.724	**0.007**	0.02	0.120
Density	0.115	0.056	2.053	**0.041**	0.01	0.121
% Marcellus Pad	−0.054	0.183	−0.294	0.769	0.0003	0.122
% All Pads	−0.062	0.175	−0.353	0.725	0.0004	0.122
% Any Gas Infrastructure	−0.018	0.086	−0.213	0.832	0.0001	0.122

had the least support compared to spatial models for GLIMPSS and WVSCI (ΔAIC values >2, Table 4).

In the AIC model comparison for WVSCI, three models had competing support (ΔAIC values >2, Table 4). For GLIMPSS, there were two models with competing support (ΔAIC values >2, Table 4). WVSCI and GLIMPSS had similar root-mean-square-prediction error (RMSPE, Table 4) and indication of both TD and Euclidean variance structure (Table 5). Final models for WVSCI and GLIMPSS had low predictive power with almost all model variance explained by the autocovariance component rather than the covariate (Table 5, Figs. 4 and 5). *Post-hoc* spearman rank correlation tests between FPI and WVSCI/GLIMPSS metrics (no. Ephemeroptera genera, no. Plecoptera genera, no. intolerant taxa tolerance value <4, and EPT richness) were all significant (Rho = 0.24, 0.20, 0.23, and 0.22 respectively, $P < 0.001$, Fig. 6).

4. Discussion

Our study is the first to apply SSNMs to relate trophic levels across the aquatic-terrestrial interface. Overall, spatial models outperformed traditional regression models, and made a statistical difference in whether stream covariates of interest were considered relatable to waterthrush foraging areas. While the spatial models had poor predictive power, SSNMs allowed us to assign variability due to spatial autocorrelation and evaluate potential trends involved in foraging on headwater streams. Stream temperature, biomass, and density were found to be significant using standard linear regression, but were no longer significant once we considered spatial autocorrelation. Therefore using standard statistical approaches could have led to making a type I error for these covariates (Dormann et al., 2007).

We did not find a relationship between foraging substrate score based on wetted perimeter data and FPI (Table 1). The protocol was designed on 1st and 2nd order waterthrush study streams in Pennsylvania (PA) that are relatively wider, less bank, and deeper water depth (Latta, 2009) than our narrow streams with steeper topography. Stucker (2000) found 1st through 3rd order waterthrush streams in Mississippi had more instream exposed rock when estimating 15 cm of additional water flow than streams without waterthrush, similar to what would create a high FPI score (i.e. higher wetted perimeter). As such, the wetted perimeter protocol may be region and stream-type specific, and water depth in relation to exposed or wetted areas that create available foraging microhabitat less important for perennial streams with many

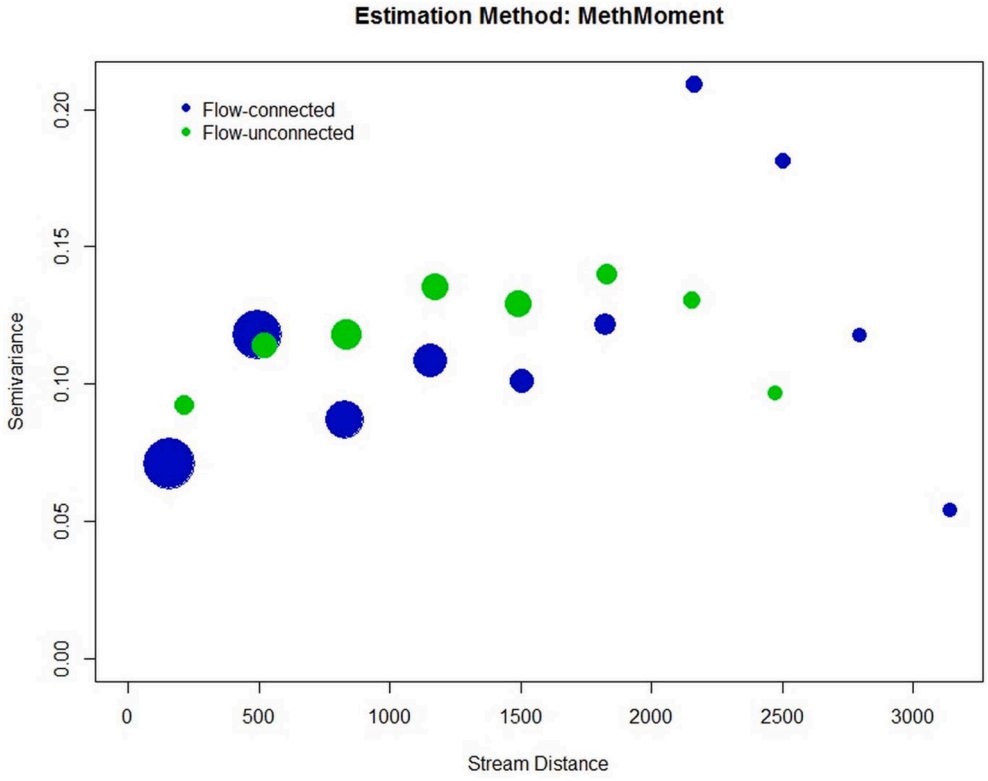

Fig. 3. Foraging probability index torgegram. An example of a torgegram for foraging probability index (FPI) which is a modified type of semivariogram. A torgegram displays semivariance (spatial autocorrelation) for samples on streams into flow-connected and -unconnected structures to assist with model fitting. Diameters of circles are proportional to the number of pairs of points in each bin.

Table 2
Initial spatial generalized linear models to test the relationship between foraging score, water chemistry, macroinvertebrate metrics, shale gas land use and foraging probability index (FPI). Covariates with bolded *P* values were significant at $\alpha = 0.10$. R^2 is a generalized value of model fit and the partial sill (sill minus nugget) was included to assess variance of a covariate without the nugget effect.

Variable	Estimate	SE	*t* value	P value	R^2	Partial sill
Temperature	−0.747	0.842	−0.887	0.376	0.002	0.0003
GLIMPSS	0.003	0.001	2.114	**0.035**	0.01	0.0003
WVSCI	0.004	0.002	2.493	**0.013**	0.02	0.0003
Biomass	0.051	0.039	1.281	0.201	0.01	0.0003
Density	0.074	0.054	1.360	0.174	0.01	0.0003

ephemeral and intermittent tributaries. For example, less availability of bank nest substrate in PA meant nests were commonly found off-stream in root balls of tree tip ups (S. Latta, Unpublished results), whereas nests at tree tip ups were rare at our study site (1 of 184 nests in 2013–2014; M. Frantz, Unpublished results). Master et al. (2005) found waterthrush densities on wintering grounds were higher on streams with higher wetted perimeter values, suggesting the need to test these protocols elsewhere in the waterthrush breeding range.

Waterthrush benthic studies were completed in 2011 and 2013–2014 during peak (2011), abated (2013), and ramped up (2014) shale gas development at our study site. We found that waterthrush territory densities were greater on streams with higher GLIMPSS scores (Wood et al., 2016), which supports our significant spatial models with GLIMPSS and WVSCI. Increasing GLIMPSS and WVSCI values increase aquatic prey metrics and indicates better riparian habitat quality for waterthrush (Frantz et al., 2018b). Our spatial model suggested at minimum a weak relationship between FPI and these multi-metric indices, at least relative to nesting locations where the majority of our benthic samples were collected. Friedan et al. (2014) used SSNMs to determine drivers of family and genus-level macroinvertebrate indices and wondered whether using coarser (mainly family-level) macroinvertebrate identification could have masked spatial patterns or reduced predictive power. In our case, using family vs. genus-level indices did not make a difference in residual variance of autocovariance (Table 5). Our results suggest family-level taxonomic resolution may be good enough to indicate most likely waterthrush foraging areas.

Table 3
Initial mixed autocovariance components (VAC) of the WVSCI and GLIMPSS spatial models. The nugget captures variability due to measurement error and/or spatial variability at less than the sampling distance. The range represents the distance at which the covariate is no longer spatially autocorrelated. The partial sill (sill minus nugget) assesses variance of a covariate without the nugget effect. Percent VAC is the percentage of residual variance accounted for by each autocovariance component. Based on higher partial sill values for Exponential TD, we added Mariah, Spherical, and Linear-with-sill TD to test before final selection of autocovariance components.

Variance component		WVSCI	GLIMPSS
Tail down (TD)	Autocovariance function	Exponential	Exponential
	Range	561.606	588.130
	Partial sill	0.073	0.067
	VAC_{TD} (%)	0.56	0.52
Tail up (TU)	Autocovariance function	Exponential	Exponential
	Range	433.213	734.33
	Partial sill	0.00000145	0.0000003
	VAC_{TU} (%)	0.00001	0.000002
Euclidean (Euc)	Autocovariance function	Exponential	Exponential
	Range	409.006	400.174
	Partial sill	0.055	0.061
	VAC_{Euc} (%)	0.42	0.47
Nugget	Nugget	0.0003	0.0003
	VAC_{Nugget} (%)	0.002	0.002

Table 4
AIC model comparison for WVSCI and GLIMPSS spatial models with Exponential Euclidean, Exponential tail down (TD), Spherical TD, and Linear plus sill TD autocovariance components in comparison to the non-spatial model with less parameters. Lowest leave-one-out cross-validation root-mean-square-prediction error (RMSPE), AIC, and ΔAIC values <2 were used to assess which models to select for final autocovariance components (VAC) model comparison.

Variance component	AIC	ΔAIC	RMSPE
WVSCI			
Exponential TD + Nugget	40.469	0.000	0.233
Exponential Euclidean + Nugget	41.250	0.781	0.233
Spherical TD + Nugget	41.971	1.502	0.236
Linear plus sill TD + Nugget	42.898	2.429	0.236
Non-spatial + Nugget	243.057	202.588	0.346
GLIMPSS			
Exponential TD + Nugget	42.981	0.000	0.234
Exponential Euclidean + Nugget	43.691	0.710	0.235
Spherical TD + Nugget	45.036	2.055	0.238
Linear plus sill TD + Nugget	45.969	2.988	0.238
Non-spatial + Nugget	243.403	200.422	0.346

Table 5
Final autocovariance components (VAC) of the WVSCI and GLIMPSS spatial models that best explain foraging probability index (FPI) based on AIC model comparison. The nugget captures variability due to measurement error and/or spatial variability at less than the sampling distance. The range represents the distance at which the covariate is no longer spatially autocorrelated. The partial sill (sill minus nugget) assesses variance of a covariate without the nugget effect. Percent VAC is the percentage of residual variance accounted for by each autocovariance component.

Variance component		WVSCI	GLIMPSS
Tail down	Autocovariance function	Exponential	Exponential
	Range	494.821	494.117
	Partial sill	0.128	0.128
	VAC_{TD} (%)	0.98	0.98
Tail down	Autocovariance function	Spherical	Spherical
	Range	282.836	NA
	Partial sill	0.129	NA
	VAC_{TD} (%)	0.97	NA
Euclidean (EUC)	Autocovariance function	Exponential	Exponential
	Range	462.148	459.309
	Partial sill	0.128	0.128
	VAC_{EUC} (%)	0.98	0.98
Nugget	Nugget	0.0003	0.0003
	VAC_{Nugget} (%)	0.002	0.002

Territory densities in 2011 were greater where Ephemeroptera, Plecoptera, and Trichoptera densities were higher, along with higher biomass (Wood et al., 2016; Merovich Jr. et al., 2022); territory density declines in 2013–2014 in part lead to the disassociation between aquatic prey biomass, density, and waterthrush demographic response at a nest and territory level (Frantz et al., 2018b). While biomass and density were no longer significant in our spatial models, there is still a positive association between richness of these sensitive taxa orders and FPI (Fig. 6). We did not assess biomass or density by size class, by which waterthrush may have shown a stronger, significant response (Wood et al., 2016) in the spatial models. Overall riparian habitat site quality may be more important to waterthrush site assessment (Frantz et al., 2018a) than benthic metrics or in-stream characteristics alone for FPI. Indeed, waterthrush appear to be able to adapt to shale gas disturbance and meet all their foraging needs until potentially a certain disturbance threshold is reached (Frantz et al., 2018b). Adaptation strategies include provisioning outside their territories and foraging on terrestrial arthropods when their preferred benthic prey is reduced (Trevelline et al., 2018). Nest survival in 2011 was best explained by Habitat Suitability Index (HSI), which describes nesting, vegetative, and foraging components important to waterthrush (Wood et al., 2016) that could also be correlated to these alternative terrestrial subsidies sought. Productivity and biomass of insects on and off stream are linked (Burdon and Harding, 2007) where further research of avian predator-prey interactions and feedback loops are needed. Since terrestrial insects are considered lower quality prey (Twining et al., 2016), waterthrush are still vulnerable to carryover effects from the breeding grounds which may result in reduced survival (Latta et al., 2016).

The three % gas land use covariates were not significant in the initial non-spatial models (Table 1). These three variables were non-normal and could not approximate normality with traditional transformations, so technically were not appropriate for the linear models. Waterthrush have a negative demographic response to the physical presence of shale gas at localized levels of the nest or territory (Frantz et al., 2018a, 2018b), as well as their aquatic prey (Frantz et al., 2018b), so it is unclear the extent to which waterthrush foraging may change due to catchment-level shale gas disturbance. However, Merovich Jr. et al. (2022) found a weak and mixed but detectable response of macroinvertebrate assemblages and water chemistry parameters at a smaller sub-catchment scale at our study site due to shale gas disturbance. The mixed benthic results and different spatial scale from Merovich Jr. et al. (2022) likely contributed to no shale gas association to FPI. GLIMPSS scores were lower downstream of shale gas disturbance in 2011 when shale gas activity was at its highest compared to sampling in 2013 and 2014 (Merovich Jr. et al., 2022). Water chemistry appeared to lag in response to disturbance in 2011 (Merovich Jr. et al., 2022), suggesting short-term chronic condition consequences to food webs from persistent land use activities in the same manner land use legacies can influence environmental variables and in turn biota (Maloney et al., 2008).

There are undisturbed stream segments and ephemeral tributaries on every stream giving opportunities to forage elsewhere (Wood et al., 2016), and waterthrush can compensate for food loss (Mulvihill et al., 2008). While this suggests to some extent adaptability in selecting foraging locations, shale gas disturbed areas have the potential to serve as sink habitats (Frantz et al., 2018a), so waterthrush may be unaware of breeding or foraging in an ecological trap (Gates and Gysel, 1978; Robertson and Hutto, 2006; Frantz et al., 2019). Additionally, waterthrush in shale gas disturbed areas have been found to bioaccumulate more heavy metals associated to the drilling process than those at undisturbed shale gas areas at our study site and elsewhere (Latta et al., 2015) with potential sex-specific influences on gene expression (Frantz et al., 2020). Foraging on macroinvertebrates is likely one way the heavy metals bioaccumulate, and we do not know how that factors into FPI. Shale gas well pad construction and drilling typical of our study site and others occur in "pulses" (Brittingham et al., 2014), making ephemeral disturbances such as sedimentation or potential runoff entering a stream system where waterthrush hold breeding territories difficult to quantify. SSNMs that can treat both spatial and temporal effects and allow non-linear data structures may be better suited to model these relationships (O'Donnell et al., 2014; Rushworth, 2014; Rushworth et al., 2015).

While our spatial models performed better than non-spatial models, they still had poor predictive power (Table 5, Fig. 4). Aquatic prey community responses were weaker in 2013–2014 at a nest and territory level than in 2011 in relation to shale gas activity levels (Frantz et al., 2018b) where this may also be true of predictive power at the catchment-level. Additionally, Friedan et al. (2014) found that spatial-weighting schemes made a substantial difference in model performance and affected variables differently. Our only spatial weighting scheme consisted of reach contributing area (RCA) (i.e. catchment area, Horizon Systems Corporation, 2007) and represented the aerial extent that contributes overland flow to a stream line segment. Some other spatial-weighting options to consider are Shreve (1967) or consideration of slope at stream segments. While catchment area may have been an appropriate scale for gas land use variables, a spatial weighting scheme such as slope that reflects local scale variability may have been more appropriate for headwater streams (Friedan et al., 2014). Our headwater streams have steep topography and many ephemeral tributaries, and

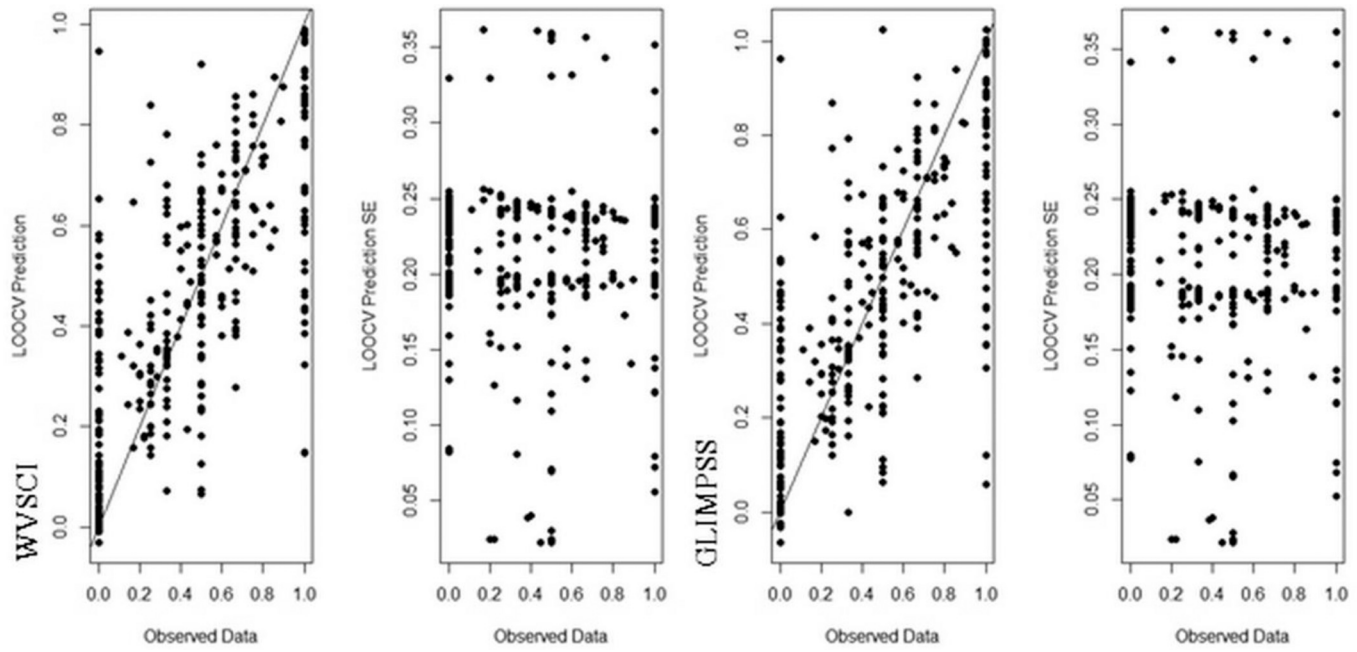

Fig. 4. Model performance. Leave one out cross validation predictions (LOOCV) and standard error (SE) against the observed data for the top WVSCI and GLIMPSS spatial models as one means of assessing model performance.

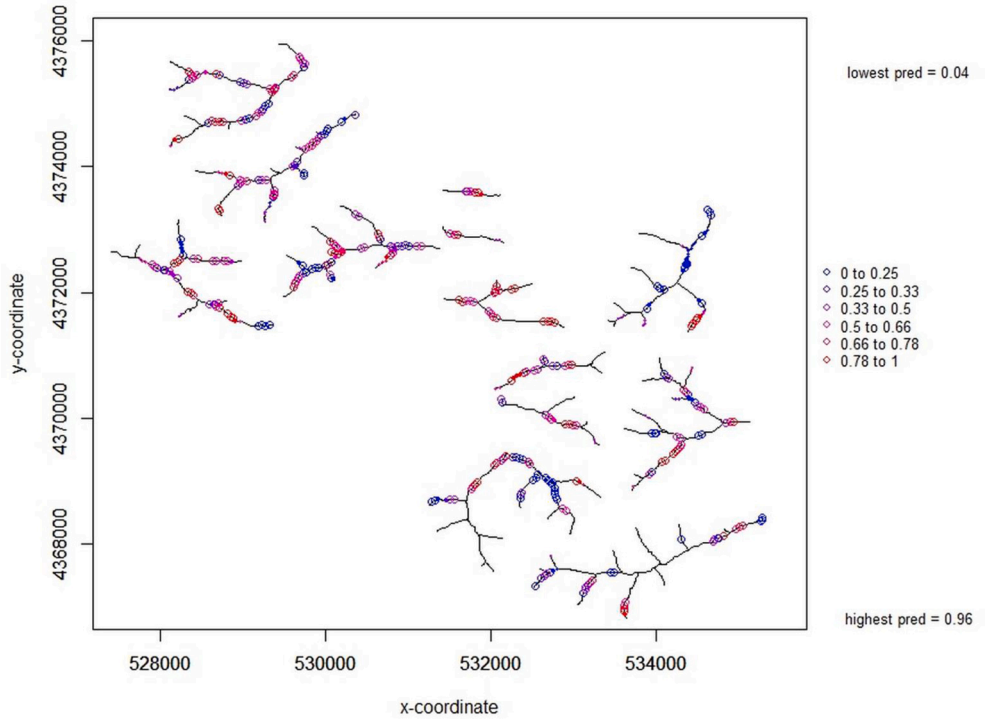

Fig. 5. Prediction map. An example of prediction values mapped for WVSCI (solid circles) in relation to collected WVSCI data (open circles). The larger the solid circle, the more confidence in the prediction value (note most circles are small). Red values have a higher foraging probability index (FPI) than blue values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

headwaters are known to have high between stream variability of habitat and high macroinvertebrate beta diversity within and among catchments (Clarke et al., 2008).

Waterthrush have linear territories on the stream (Mulvihill et al., 2008) and typically fly up and down the stream corridor rather than around it (M. Frantz, *Personal observation*). As such it makes sense that tail-down (TD) autocovariance structure explained the most model variability as it allows correlation of samples between flow-connected and -unconnected stream segments. Euclidean distance having almost equal explanation of model variability likely reflects some combination of terrestrial components to waterthrush ecology and territory-scale or higher attributes. Trevelline et al. (2016) found that terrestrial Lepidoptera was in 92% of waterthrush nestling diets where phenological shifts in the availability of terrestrial insects may play an additional

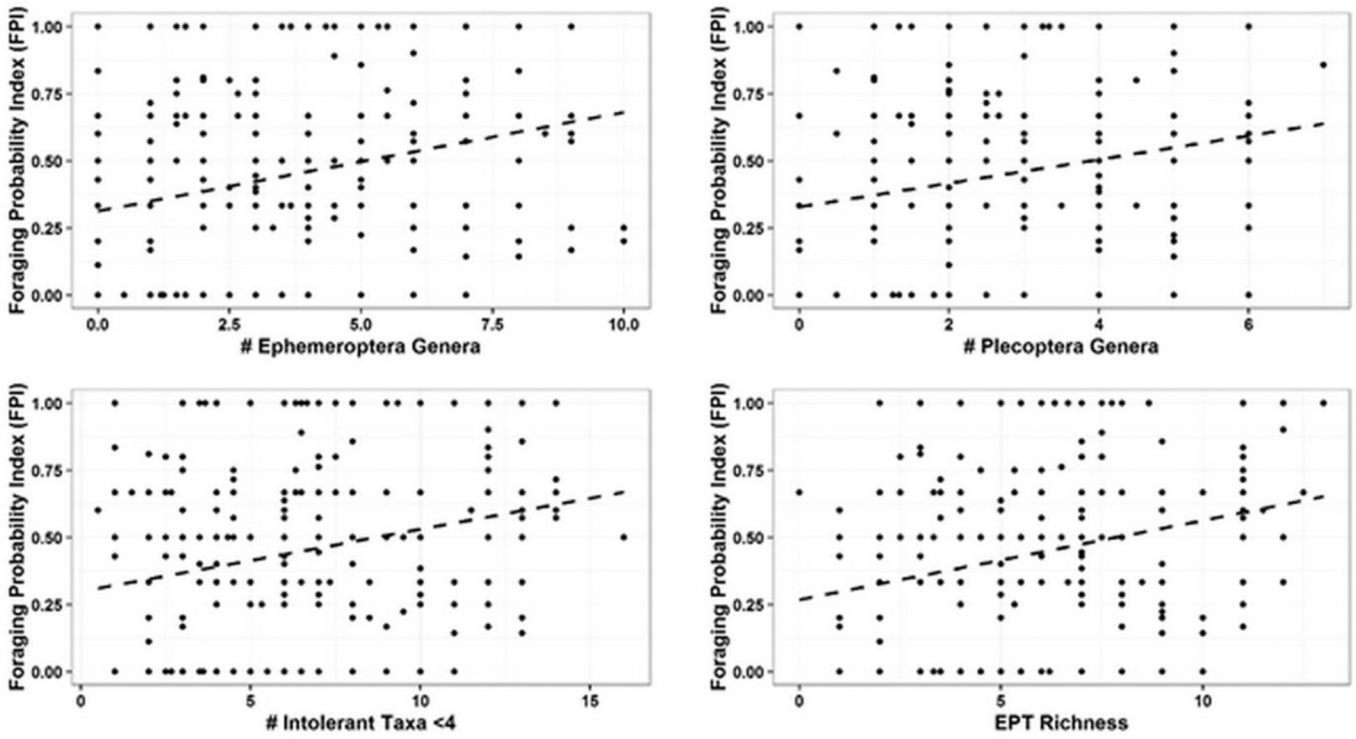

Fig. 6. Foraging probability index correlations. *Post-hoc* spearman rank correlation tests between foraging probability index (FPI) and WVSCI/GLIMPSS metrics (no. Ephemeroptera genera, no. Plecoptera genera, no. intolerant taxa tolerance value <4, and EPT richness). All tests were significant (Rho = 0.24, 0.20, 0.23, & 0.22 respectively, $P < 0.001$).

unfactored role here. However predictive power depends not only on autocovariance structure but the covariates themselves (Friedan et al., 2014). For instance, FPI was likely not representative on Carpenter Run which could have introduced some unnecessary variability in the data (Fig. 2). Priority during our six year (2009–2011, 2013–2015) waterthrush demography project was to accurately delineate waterthrush territories, with emphasis on collecting new location points that reflected all boundaries of a waterthrush territory. While we monitored Carpenter Run for nests in 2013, we did not start collecting territory and foraging observations until 2014. Introducing stream ($n = 14$) as a random effect *post-hoc* in the WVSCI and GLIMPSS models may have explained between 4.6 and 6.5% of model variance (e.g. AIC = 41.81, RMSPE = 0.232 for WVSCI with Exponential TD autovariance + Stream). This emphasizes the importance of not only thorough observation collection but accounting for headwater stream heterogeneity.

Physical features of the landscape affect aquatic-terrestrial food web interactions (Witman et al., 2004) like forest streamside vegetation (Sweeney, 1993). While we did not include habitat covariates in our models to relate to the stream channel components that compose FPI, our study site remained relatively forested (>95%) and intact despite localized landscape disturbance from shale gas (Merovich Jr. et al., 2022). Waterthrush were previously found to have a weak positive relationship to forest canopy disturbance due to shale gas development (Frantz et al., 2018) which could be related to increased net primary production (Johnson et al., 2015) or increased abundance of specific prey items (Barton, 2016). Merovich Jr. et al. (2022) found the strongest macroinvertebrate assemblage dissimilarities upstream and downstream of shale gas disturbance in 2011 using GLIMPSS when shale gas intensity was highest, but not during our foraging study except for specific indicator genera. Where waterthrush were found foraging in our study demonstrate how macroinvertebrate prey is controlled by local environmental conditions and how placement or emergence of these food resources in part can explain insectivore densities and distributions (Gray, 1993).

Collecting large data sets due to stream monitoring programs is becoming commonplace (Rushworth et al., 2015), stressing the need to use the proper statistical tools that will provide optimal performance and prediction power. While our spatial models had poor performance power, we can still produce predictive maps that can direct us to potentially important waterthrush foraging areas to evaluate further such as upper reaches of headwater tributaries (Fig. 5). The utility of SSNMs have been used previously to predict fish densities (Isaak et al., 2016), and now for an apex avian predator that habits the aquatic-terrestrial interface, and thus has the potential for land managers with waterthrush occurrence data to prioritize management or conservation areas given the waterthrush's role as a bioindicator of aquatic stream integrity (O'Connell et al., 2000; Mulvihill et al., 2008). Our exploratory SSNM analyses are a starting point to inquire further into food-web interactions between waterthrush, macroinvertebrates, and potential surface water contamination, and serves as an example of how spatial autocorrelation coming from multiple sources and scales may influence study implications.

Funding

Our research was funded by the West Virginia Division of Natural Resources (WVDNR), West Virginia University (WVU), National Aviary's Avian Conservation Endowment, and Appalachian Stewardship Foundation (ASF).

Declaration of Competing Interest

There are no conflicts of interest concerning our article. The corresponding author's current affiliation is with the West Virginia Division of Natural Resources (WVDNR) as of 2018. One funding source for the project was the WVDNR prior to 2018 in years 2013 and 2014 when the corresponding author was a graduate student. The corresponding author receives no form of compensation for publication. Time spent on the manuscript is volunteered and not on work time.

Acknowledgements

West Virginia Division of Natural Resources (WVDNR) provided access to the study area and Wheeling Jesuit University provided access to field housing. Our research was funded by WVDNR, West Virginia University, National Aviary, and Appalachian Stewardship Foundation. We thank Jack Toriello, Blake Hepner, Darin Blood, Jerry Kreiser, Steve Daniels, Tasso Cocoves, Nick Glover, and Ashley Akers for collecting data for this project. We additionally thank Jim Sheehan for project assistance and Ross Andrew for manuscript review. This study was completed under the auspices of West Virginia University IACUC protocol #04-0302, 07-0303. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. 19, 716–722.

Barton, J.A., 2016. MS thesis In-stream leaf decomposition as an indicator of Marcellus shale impairment across a land use gradient. Bucknell University, Lewisburg, PA, USA.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S. A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Softw. 40, 1–20.

Bibby, C.J., Burgess, N.D., Hill, D.A., 1992. Bird Census Techniques. Academic Press, London, p. 257.

Brittingham, M.C., Maloney, K.O., Farag, A.M., Harper, D.D., Bowen, Z.H., 2014. Ecological risks of shale oil and gas development to wildlife, aquatic resources and their habitats. Environ. Sci. Technol. 48, 11034–11047.

Burdon, F.J., Harding, J.S., 2007. The linkage between riparian predators and aquatic insects across a stream-resource spectrum. Freshw. Biol. 53, 330–346.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: An Information-Theoretic Approach. Springer-Verlag, New York, New York, USA.

Buss, D.F., Carlisle, D.M., Chon, T.-S., Culp, J., Harding, J.S., Keizer-Vlek, H.E., Robinson, W.A., Strachan, S., Thirion, C., Hughes, R.M., 2015. Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. Environ. Monit. Assess. 187, 4132.

Clarke, A., Mac Nally, R., Bond, N., Lake, P.S., 2008. Macroinvertebrate diversity in headwater streams: a review. Freshw. Biol. 53, 1707–1721.

Cook, N.A., Sarver, E.A., Krometis, L.H., Huang, J., 2015. Habitat and water quality as drivers of ecological system health in central Appalachia. Ecol. Eng. 84, 180–189.

Cressie, N., 1993. Statistics for Spatial Data, revised edition. John Wiley and Sons, New York.

Cressie, N., Frey, J., Harch, B., Smith, M., 2006. Spatial prediction on a river network. J. Agric. Biol. Environ. Stat. 11, 127–150.

Diggle, P.J., Ribeiro Jr., P.J., 2007. Model-based geostatistics. 2007. Springer, New York. ISBN 0387329072.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M., Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30, 609–628.

Dunscomb, J.K., Evans, J.S., Strager, J.M., Strager, M.P., Kiesecker, J.M., 2014. Assessing future energy development across the Appalachian Landscape Conservation Cooperative. In: The Nature Conservancy, Charlottesville, VA. 48 pp with appendices. Appalachian Landscape Conservation Cooperative Grant #2012-02.

Entrekin, S., Evans-White, M., Johnson, B., Hagenbuch, E., 2011. Rapid expansion of natural gas development poses a threat to surface waters. Front. Ecol. Environ. 9, 503–511.

ESRI, 2011. ArcMap GIS Desktop. Release 10.2.2. Environmental Systems Research Institute, Redlands, CA.

Evans, J.S., Kiesecker, J.M., 2014. Shale gas, wind and water: assessing the potential cumulative impacts of energy development on ecosystem services within the Marcellus play. PLoS One 9, e89210.

Farwell, L.S., Wood, P.B., Sheehan, J., George, G.A., 2016. Shale gas development effects on the songbird community in a central Appalachian forest. Biol. Conserv. 201, 78–91.

Fisher, R.A., 1935. The Design of Experiments. Oliver and Boyd, Edinburgh.

Fortin, M.J., Dale, M.R., 2005. Spatial Analysis: A Guide for Ecologists. Cambridge University Press, Cambridge.

Frantz, M.W., Wood, P.B., Sheehan, J., George, G., 2018a. Demographic response of a migratory songbird of conservation concern to shale gas development. Condor: Ornithol. Appl. 120, 265–282.

Frantz, M.W., Wood, P.B., Merovich Jr., G.T., 2018b. Demographic characteristics of an avian predator, Louisiana Waterthrush (*Parkesia motacilla*), in response to its aquatic prey in a central Appalachian USA watershed impacted by shale gas development. PLoS One 13, e0206077.

Frantz, M.W., Wood, P.B., Sheehan, J., George, G., 2019. Louisiana Waterthrush (*Parkesia motacilla*) survival and site fidelity in an area undergoing shale gas development. Wilson J. Ornithol. 131, 84–95.

Frantz, M.W., Wood, P.B., Latta, S.C., Welsh, A.B., 2020. Epigenetic response of Louisiana Waterthrush Parkesia motacilla to shale gas development. Ibis 162, 1211–1224. https://doi.org/10.1111/ibi.12833.

Freeman, M.C., Pringle, C.M., Jackson, C.R., 2007. Hydrologic connectivity and the contribution of stream headwaters to ecological integrity at regional scales. J. Am. Water Resour. Assoc. 43, 5–14.

Friedan, J.C., Peterson, E.E., Webb, J.A., Negus, P.M., 2014. Improving the predictive power of spatial statistical models of stream macroinvertebrates using weighted autocovariance functions. Environ. Model. Softw. 60, 320–330.

Gates, J.E., Gysel, L.W., 1978. Avian nest dispersion and fledging success in field-forest ecotones. Ecology 59, 871–883.

Gerritsen, J., Burton, J., Barbour, M.T., 2000. A Stream Condition Index for West Virginia Wadeable Streams. Tetra Tech Inc., Owing Mills.

Gomi, T., Sidle, R.C., Richardson, J.S., 2002. Understanding processes and downstream linkages of headwater systems. Bioscience 52, 905–916.

Gray, L.J., 1993. Response of insectivorous birds to emerging aquatic insects in riparian habitats of a tallgrass prairie system. Am. Midl. Nat. 129, 288–300.

Isaak, D.J., Peterson, E.E., Ver Hoef, J.M., Wenger, S.J., Falke, J.A., Torgersen, C.E., Sowder, C., Steel, E.A., Marie-Josee Fortin, M.-J., Jordan, C.E., Ruesch, A.S., Som, N., Monestiez, P., 2014. Applications of spatial statistical network models to stream data. WIREs Water 1, 277–294.

Isaak, D.J., Ver Hoef, J.M., Peterson, E.E., Horan, D.L., Nagel, D.E., 2016. Scalable population estimates using spatial-stream-network SSN models, fish density surveys, and national geospatial database frameworks for streams. Can. J. Fish. Aquat. Sci. 74, 147–156.

Johnson, E., Austin, B.J., Inlander, E., Gallipeau, C., Evans-White, M.A., Entrekin, S., 2015. Stream macroinvertebrate communities across a gradient of natural gas development in the Fayetteville shale. Sci. Total Environ. 530–531, 323–332.

Latta, K., 2009. What Determines Success? Breeding Habitat Characteristics of the Louisiana Waterthrush (*Seirus motacilla*). Senior Thesis. University of Michigan, p. 16.

Latta, S.C., Marshall, L.C., Frantz, M.W., Toms, J., 2015. Evidence from two shale regions that a riparian songbird accumulates metals associated with hydraulic fracturing. Ecosphere 6, 144.

Latta, S.C., Cabezas, S., Mejia, D.A., Paulino, M.M., Almonte, H., Miller-Butterworth, C. M., Bortolotti, G.R., 2016. Carry-over effects provide linkages across the annual cycle of a Neotropical migratory bird, the Louisiana Waterthrush Parkesia motacilla. Ibis 158, 395–406.

Legendre, P., 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74, 1659–1673.

MacDonald, L.H., Coe, D., 2007. Influence of headwater streams on downstream reaches in forested areas. For. Sci. 53, 148–168.

Maloney, K.O., Feminella, J.W., Mitchell, R.M., Miller, S.A., Mulholland, P.J., Houser, J. N., 2008. Landuse legacies and small streams: identifying relationships between historical land use and contemporary stream conditions. J. North Am. Benthol. Soc. 27, 280–294.

Master, T.L., Mulvihill, R.S., Leberman, R.C., Sanchez, J., Carman, E., 2005. A Preliminary Study of Riparian Songbirds in Costa Rica, with Emphasis on Wintering Louisiana Waterthrushes. USDA Forest Service Gen. Tech. Rep. PSW-GTR-191, pp. 528–532.

Mattsson, B.J., Cooper, R.J., 2006. Louisiana waterthrushes (*Seiurus motacilla*) and habitat assessment as cost-effective indicators of instream biotic integrity. Freshw. Biol. 51, 1941–1958.

Mattsson, B.J., Cooper, R.J., 2009. Multiscale analysis of the effects of rainfall extremes on reproduction by an obligate riparian bird in urban and rural landscapes. Auk 126, 64–76.

Mattsson, B.J., Master, T.L., Mulvihill, R.S., Robinson, W.D., 2009. Louisiana Waterthrush (*Parkesia motacilla*). In: Poole, A.F. (Ed.), Birds of North America. Cornell Lab of Ornithology, Ithaca, NY. https://doi.org/10.2173/bna.151.

Merovich Jr., G.T., Frantz, M.W., Wood, P.B., 2022. Patterns in benthic macroinvertebrate assemblages in ana active region of unconventional shale-gas development in the western Appalachian Plateau of West Virginia, USA. Environ. Monit. Assess. 194, 368.

Merriam, E.R., Petty, J.T., Strager, M.P., Maxwell, A.E., Ziemkiewicz, P.F., 2013. Scenario analysis predicts context-dependent stream response to landuse change in a heavily mined central Appalachian watershed. Freshw. Sci. 32, 1246–1259.

Mulvihill, R.S., Newell, F.L., Latta, S.C., 2008. Effects of acidification on the breeding ecology of a stream-dependent songbird, the Louisiana Waterthrush (*Seiurus motacilla*). Freshw. Biol. 53, 2158–2169.

O'Connell, T.J., Jackson, L.E., Brooks, R.P., 2000. Bird guilds as indicators of ecological condition in the Central Appalachians. Ecol. Appl. 10, 1706–1721.

O'Donnell, D., Rushworth, A., Bowman, A.W., Scott, E.M., 2014. Flexible regression models over river networks. J. R. Stat. Soc.: Ser. C: Appl. Stat. 63 (Part I), 47–63.

Peterson, E.E., Ver Hoef, J.M., 2010. A mixed-model moving-average approach to geostatistical modeling in stream networks. Ecology 91, 644–651.

Peterson, E.E., Ver Hoef, J.M., 2014. STARS: an ArcGIS toolset used to calculate the spatial information needed to fit spatial statistical models to stream network data. J. Stat. Softw. 56, 1–17.

Pond, G.J., Bailey, J.E., Lowman, B.M., Whitman, M.J., 2013. Calibration and validation of a regionally and seasonally stratified macroinvertebrate index for West Virginia wadeable streams. Environ. Monit. Assess. 185, 1515–1540.

R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.Rproject.org/.

Robbins, C.S., 1970. Recommendations for an international standard for a mapping method in bird census work. Audubon Field Notes 24, 723–726.
Robertson, B.A., Hutto, R.L., 2006. A framework for understanding ecological traps and an evaluation of existing evidence. Ecology 87, 1075–1085.
Robinson, W.D., 1990. Louisiana Waterthrush Foraging Behavior and Microhabitat Selection in Southern Illinois. Masters Thesis. Southern Illinois University, Carbondale, IL.
Rushworth, A.M., 2014. smnet: Smoothing for stream network data. R package version 1.0.
Rushworth, A.M., Peterson, E.E., Ver Hoef, J.M., Bowman, A.W., 2015. Validation and comparison of geostatistical and spline models for spatial stream networks. Environmetrics 26, 327–338.
Sauer, J.R., Hines, J.E., Fallon, J.E., Pardieck, K.L., Ziolkowski, D.J., Link, W.A., 2014. The North American Breeding Bird Survey, Results and Analysis 1966–2013 Version 01.30.2015. USGS Patuxent Wildlife Research Center, Laurel, Maryland. http://www.mbr-pwrc.usgs.gov/bbs/.
Shields, W.M., 1977. The effect of time of day on avian census results. Auk 2, 380–383.
Shreve, R.L., 1967. Infinite topologically random channel networks. J. Geol. 75, 178–186.
Strager, M.P., Petty, J.T., Strager, J.M., Barker-Fulton, J., 2009. A spatially explicit framework for quantifying downstream hydrologic conditions. J. Environ. Manag. 90, 1854–1861.
Stucker, J.H., 2000. Biodiversity of Southeastern Minnesota Forested Streams: Relationships between Trout Habitat Improvement Practices, Riparian Communities and the Louisiana Waterthrush. Natural Heritage and Nongame Wildlife Program, Minnesota Department of Natural Resources, p. 158.
Swanson, F.J., Johnson, S.L., Gregory, S.V., Acker, S.A., 1998. Flood disturbance in a forested mountain landscape. BioScience 48, 681–689.
Sweeney, B.W., 1993. Effects of streamside vegetation on macroinvertebrate communities of White Clay Creek in eastern North America. Proceed. Acad. Sci. Philadelphia 144, 291–340.
Theobald, D.M., Norman, J.B., Peterson, E., Ferraz, S., Wade, A., Sherburne, M.R., 2006. Functional Linkage of Water Basins and Streams FLoWS v1 Users Guide: ArcGIS Tools for Network-based Analysis of Freshwater Ecosystems. Natural Resource Ecology Lab, Colorado State University, Fort Collins, CO, p. 43.
Trevelline, B.K., Latta, S.C., Marshall, L.C., Nuttle, T., Porter, B.A., 2016. Molecular analysis of nestling diet in a long-distance Neotropical migrant, the Louisiana Waterthrush (*Parkesia motacilla*). Auk Ornithol. Adv. 133, 415–428.
Trevelline, B.K., Nuttle, T., Porter, B.A., Brouwer, N.L., Hoenig, B.D., Steffensmeier, Latta, S.C., 2018. Stream acidification and reduced aquatic prey availability are associated with dietary shifts in an obligate riparian Neotropical migratory songbird. PeerJ 6, e5141.
Twining, C.W., Brenna, J.T., Lawrence, P., Shipley, J.R., Tollefson, T.N., Winkler, D.W., 2016. Omega-3 long-chain polyunsaturated fatty acids support aerial insectivore performance more than food quantity. PNAS 113, 10920–10925.
USFWS (United States Fish and Wildlife Service), 2008. Birds of Conservation Concern 2008. United States Department of Interior, Fish and Wildlife Service, Division of Migratory Bird Management, Arlington, Virginia, p. 85 [Online]. http://www.fws.gov/migratorybirds.
Ver Hoef, J.M., Peterson, E.E., 2010. A moving average approach for spatial statistical models of stream networks. J. Am. Stat. Assoc. 105, 6–18.
Ver Hoef, J.M., Peterson, E.E., Clifford, D., Shah, R., 2014. The SSN package: an R package used to fit spatial statistical models to stream network data. J. Stat. Softw. 56, 1–45.
Ver Hoef, J.M., Peterson, E., Theobald, D., 2006. Spatial statistical models that use flow and stream distance. Environ. Ecol. Stat. 13, 449–464.
Vitz, A.C., Rodewald, A.D., 2010. Movements of fledgling ovenbirds (*Seirus aurocapilla*) and worm-eating warblers (*Helmitheros vermivorum*) within and beyond the natal home range. Auk 127, 364–371.
Witman, J.D., Ellis, J.C., Anderson, W.B., 2004. In: Polis, G.A., Power, M.E., Huxel, G.R. (Eds.), The Influence of Physical Processes, Organisms, and Permeability on Cross-Ecosystem Fluxes. Food Webs at the Landscape Level. The University of Chicago Press, Chicago, IL, pp. 335–358.
Wood, P., Frantz, M., Becker, D., 2016. Louisiana Waterthrush and benthic macroinvertebrate response to shale gas development. J. Fish Wildl. Manag. 7, 423–433.
Woods, A.J., Omernik, J.M., Brown, D.D., 1999. Level III and IV Ecoregions of Delaware, Maryland, Pennsylvania, Virginia, and West Virginia, U.S. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Corvallis, OR [Online]. ftp://ftp.epa.gov/wed/ecoregions/reg3/reg3_eco_desc.doc.
WVDEP (West Virginia Department of Environmental Protection), 2015. Oil and Gas Production Data. State of West Virginia, [Online]. http://www.dep.wv.gov/oil-and-gas/databaseinfo/Pages/default.aspx.
WVGES (West Virginia Geological and Economic Survey), 2015. WV Marcellus Shale Wells: Completed, Permitted, and Cancelled Episodes. [Online] http://www.wvgs.wvnet.edu/www/datastat/devshales.htm. WVGES.
Zimmerman, D.L., Ver Hoef, J.M., 2017. The torgegram for fluvial variography: characterizing spatial dependence on stream networks. J. Comput. Graph. Stat. 26 (2), 253–264.
Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3–14.