

REGULAR ARTICLE

# RIVERSCAPE-SCALE MODELING OF FUNDAMENTALLY SUITABLE HABITAT FOR MUSSEL ASSEMBLAGES IN AN OZARK RIVER SYSTEM, MISSOURI

Kayla N. Key<sup>\*1,2</sup>, Amanda E. Rosenberger<sup>3</sup>, Garth A. Lindner<sup>†4</sup>, Kristen Bouska<sup>5</sup>, and Stephen E. McMurray<sup>6</sup>

<sup>1</sup> West Tennessee River Basin Authority, Tennessee Department of Environment and Conservation, Humboldt, TN 38343 USA

<sup>2</sup> Tennessee Cooperative Fishery Research Unit, Department of Biology, Tennessee Tech University, Cookeville, TN 38505 USA

<sup>3</sup> U.S. Geological Survey, Tennessee Cooperative Fishery Research Unit, Department of Biology, Tennessee Tech University, Cookeville, TN 38505 USA

<sup>4</sup> University of Missouri, School of Natural Resources, Columbia, MO 65211 USA

<sup>5</sup> U.S. Geological Survey, Upper Midwest Environmental Sciences Center, La Crosse, WI 54603 USA

<sup>6</sup> Missouri Department of Conservation, Science Branch, Columbia, MO 65201 USA

---

## ABSTRACT

Identifying the physical habitat characteristics associated with riverine freshwater mussel assemblages is challenging but crucial for understanding the causes of mussel declines. The occurrence of mussels in multispecies beds suggests that common physical factors influence or limit their occurrence. Fine-scale geomorphic and hydraulic factors (e.g., scour, bed stability) are predictive of mussel-bed occurrence, but they are computationally challenging to represent at intermediate or riverscape scales. We used maximum entropy (MaxEnt) modeling to evaluate associations between riverscape-scale hydrogeomorphic variables and mussel-bed presence along 530 river km of the Meramec River basin, USA, to identify river reaches that are fundamentally suitable for mussels as well as those that are not. We obtained the locations of mussel beds from an existing, multiyear dataset, and we derived river variables from high-resolution, open-source datasets of aerial imagery and topography. Mussel beds occurred almost exclusively in reaches identified by our model as suitable; these were characterized by laterally stable channels, absence of adjacent bluffs, proximity to gravel bars, higher stream power, and larger areas of contiguous water (a proxy for drought vulnerability). We validated our model findings based on model sensitivity using a set of mussel-bed locations not used in model development. These findings can inform how resource managers allocate survey, monitoring, and conservation efforts.

**KEY WORDS:** freshwater mussels, conservation planning, unionids, hydrogeomorphology, MaxEnt, riverscape

---

## INTRODUCTION

Our knowledge of the physical habitat characteristics associated with mussel beds remains incomplete (Haag 2012).

Understanding the habitat associations of mussels is foundational for further understanding the threats to mussel populations and enacting regional and strategic conservation efforts to improve the status of mussels nationwide (FMCS 2016). Mussel presence has been predicted most successfully at large spatial scales, using variables such as watershed geology, soils, land use, and topography (Strayer 1983;

---

<sup>†</sup>Current affiliation: Indiana Department of Natural Resources, Indianapolis, IN 46204 USA

\*Corresponding Author: kaylakey@me.com

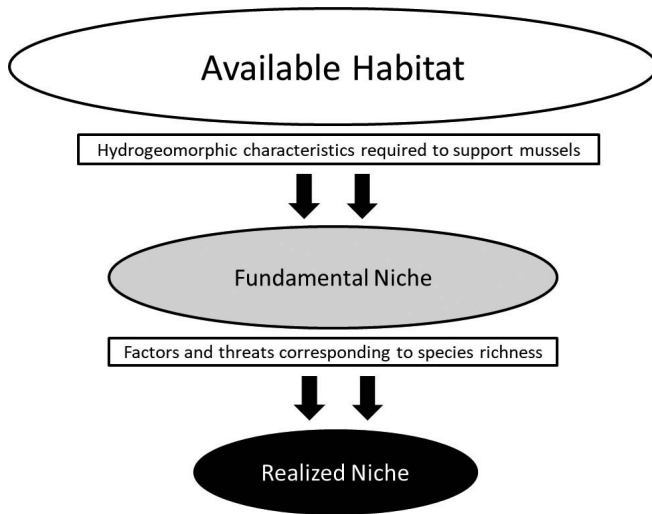


Figure 1. Conceptual model of the approach used to depict habitat requirements of freshwater mussels, starting with the identification of the fundamental habitat requirements through a description of hydrogeomorphic features of the river system. Based on this, one can then investigate other factors limiting mussel species, including anthropogenic threats.

Arbuckle and Downing 2002; McRae et al. 2004; Daniel and Brown 2014; Walters et al. 2017). However, large-scale factors such as geology and topography are not tractable for management actions.

Reach-level factors are influenced more easily by management, but efforts to predict mussel occurrence at this scale have been less successful (Strayer 2008). Hydrogeomorphic variables related to substrate stability (e.g., shear stress, channel stability) show promise for predicting where mussels occur (Hardison and Layzer 2001; Allen and Vaughn 2010; Drew et al. 2018). Unfortunately, generalizing these results to other streams can be difficult due to varying methods and differences in size and geomorphological processes among lotic systems (Layzer and Madison 1995; Steuer et al. 2008; Pandolfo et al. 2016). Furthermore, measurements of hydraulic variables at small scales are time-consuming, making it impractical to generate data needed for watershed-scale inference, which is most useful for resource managers (Fausch et al. 2002).

An investigation of mussel habitat associations at a riverscape scale (continuous spatial data across a longitudinal river gradient) may be most useful to prioritize management efforts within a watershed (Bouska et al. 2018). For such a framework, the first step is inferring what combinations of reach-scale habitat factors provide fundamentally suitable conditions for mussels. This is analogous to the single-species, fundamental niche concept, which describes “a state of environment which would permit the species to exist” (Hutchinson 1957; Fig. 1). The advantage of the riverscape scale is that we can use those data to spatially identify where these combinations of habitat factors do and do not occur in the river system of interest. However,

at a riverscape scale, it is time-consuming and expensive to generate the hydrogeomorphic data needed to predict mussel occurrence (e.g., shear stress). Easily obtainable riverscape-scale data are needed that represent or are responsive to hydrogeomorphic processes and can predict mussel occurrence in similar ways. If the fundamental niche can be described or quantified, it can be contrasted with the realized niche, which describes where animals actually occur according to the influence of other factors, such as competition or anthropogenic impacts. Such riverscape-scale habitat models can allow inference about whether mussel declines or absence in a particular stream reach are due to habitat factors or other factors, such as human impacts on water quality (Bouska et al. 2018; Fig. 1).

Mussels often occur in multispecies beds, suggesting that common physical factors influence or limit the occurrence of multiple species (Vaughn 1997), and the probability of occurrence of species of management concern increases with increased assemblage richness (Zipkin et al. 2009; Lueckenhoff 2015). For these reasons, it is reasonable—and most informative to managers—to assess habitat relationships associated with entire, multispecies mussel beds, rather than using traditional species-specific habitat models. It also may be most useful to depict such associations along a continuous riverscape scale instead of imposing preconceived spatial scales (e.g., hydrologic units) of unknown relevance to the organisms; such an approach allows scale to arise from analytical results (Parsons et al. 2004) and better represents multiple levels of ecological organization.

The goal of this study was to delineate fundamentally suitable and unsuitable river reaches for mussels based on habitat variables (i.e., in the absence of other limiting factors, such as water quality or fish hosts) in the Meramec River basin of Missouri on a riverscape scale. We confined our study to midsized river reaches suitable for snorkel-based sampling, such as timed visual searches. We evaluated habitat features associated with mussel beds, with a focus on riverscape-scale hydrogeomorphic variables associated with water availability, channel stability, and stable gravel substrate. Our objectives were to (1) generate a dataset of spatial layers representing hydrogeomorphic stream characteristics relevant to mussel ecology derived from open-source, remotely sensed data, (2) develop a fundamental niche habitat model to evaluate associations between the presence of mussel beds and hydrogeomorphic stream characteristics, and (3) validate the model by testing its overall sensitivity and evaluate the occurrence of mussel beds outside of fundamentally suitable habitat as delineated by the model. Because the model is intended only to delineate, in general, fundamentally suitable and unsuitable habitat, rather than to predict actual bed location, we did not evaluate model specificity. We hypothesized that specific combinations of physical features would be associated with mussel beds and the absence of these features would characterize sizeable portions of the river as fundamentally unsuitable for mussel beds.

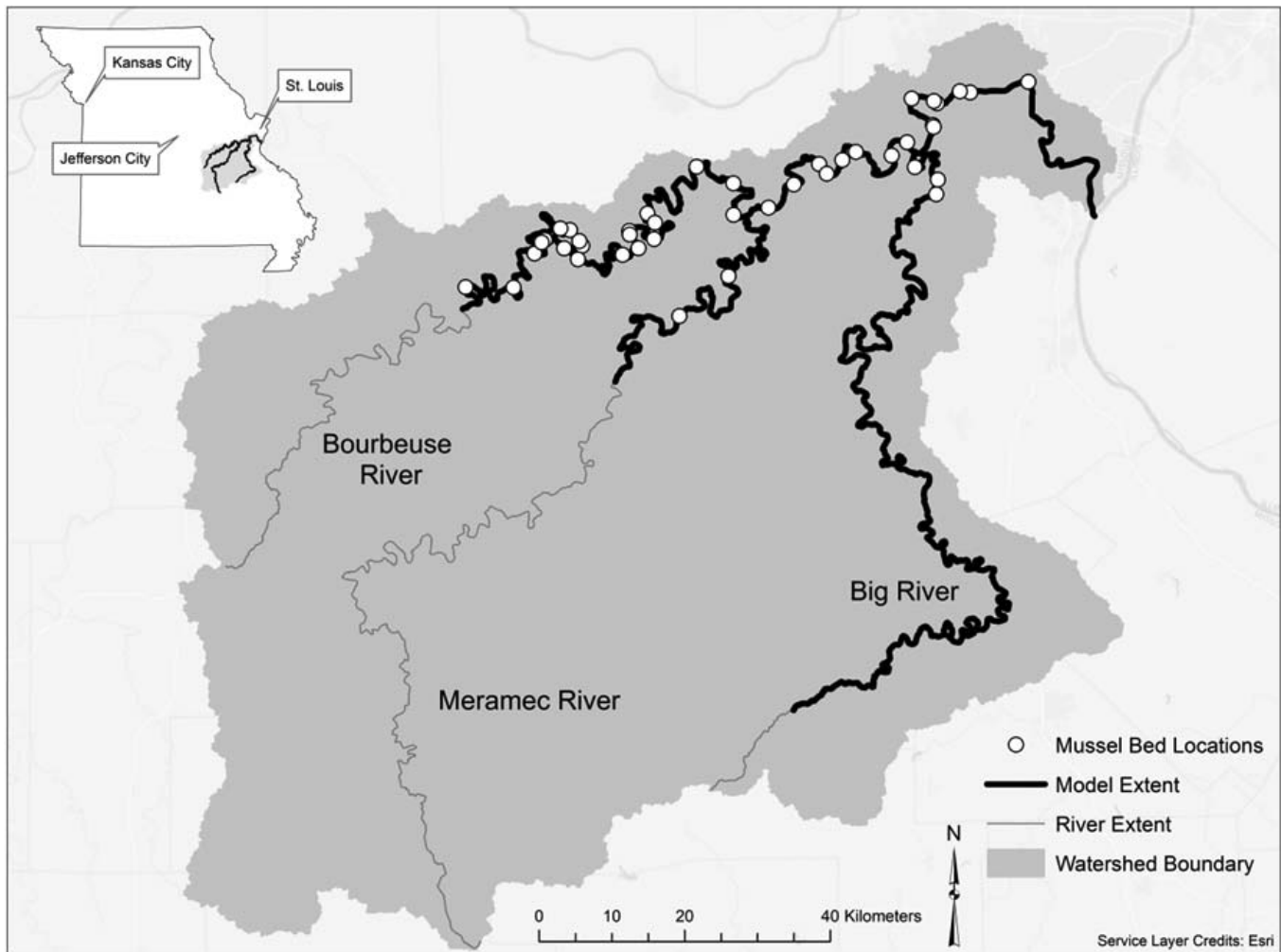


Figure 2. Map of the Meramec River basin showing the Meramec River and its two largest tributaries, the Big and Bourbeuse rivers. The inset map shows the location of the Meramec River basin in Missouri, USA. The study reach (model extent) is indicated by bolded streamlines and mussel-bed locations used in modeling and validation are shown in white circles.

## METHODS

### Study Area

The study was focused on the Meramec River basin (the Meramec, Big, and Bourbeuse rivers, “MBB” hereafter as an acronym for this basin; Fig. 2) in the northeastern Ozark Plateau in Missouri. The MBB is located within the Upper Mississippi mussel faunal province (Haag 2010), and it is a hotspot of mussel diversity in the midwestern United States (Roberts and Bruenderman 2000; Hinck et al. 2012). The Meramec River has two major tributaries, the Big River (2,473 km<sup>2</sup>) and the Bourbeuse River (2,183 km<sup>2</sup>). Stream substrates in the basin are dominated by gravel (Jacobson and Primm 1997). Mussel diversity has declined in the MBB, potentially due to altered floodplains, channel incision, invasive species, and water pollution (TNC 2014). The basin has a drainage area of 10,308 km<sup>2</sup> and is approximately 70% forested (MSDIS 2011). The remainder of the watershed is composed of row-crop agriculture, pasture, and low-density industry and

urbanization, with the exception of the heavily urbanized St. Louis metropolitan region, near the point at which the Meramec River flows into the Mississippi River (Homer et al. 2012). Our study was located along 530 km of the main channels of the Meramec, Big, and Bourbeuse rivers (the MBB) upstream of the urbanized part of the watershed (Fig. 2).

### Mussel Survey Dataset

We used the Missouri Department of Conservation (MDC) mussel database (data available upon request to and subject to the approval of the Missouri Department of Conservation, 3500 East Gans Road, Columbia, MO, 65201) to extract geographic locations from mussel surveys completed by the department. This dataset is a large, statewide, and long-term database managed by MDC biologists and includes survey information for specific mussel-bed locations across Missouri. Information on these mussel beds includes GPS points, survey method used, list of species found, and number of individuals found, but it

includes little or no information on survey design. To be consistent with the dates of hydrogeomorphic data used in this study (2012 and 2014), we extracted survey information for sites in the MBB sampled after 1993 by wading, snorkeling, or diving. We removed from the dataset sites where mussels were not reported, sites that were represented by haphazard or incidental collections, or sites that were sampled only by tactile search methods. Sites located within 180 m of another site were considered one collective bed. We chose this distance based on previous studies of average mussel-bed length in the MBB (Lueckenhoff 2015; Schrum 2017). The remaining dataset included 106 unique mussel-bed locations. To build our model, we then selected a subset of 42 mussel-bed locations representing the highest-richness sites; these beds individually contained 14–31 species and, collectively, the entire Meramec River mussel fauna (Key 2019). The remaining 64 mussel-bed locations individually contained 2–26 species and a total of 39 species. We used these beds to determine the model's sensitivity by assessing how many of these beds fell within areas that our model defined as suitable habitat. This validation is an additional step beyond the standard validation methods used regularly in model development, such as AUC values described subsequently (see “Model Building”).

### Generation of Hydrogeographic Variables

A critical step for setting up the workflow of generating the spatial hydrogeomorphic variables is to define the stream dimensions and the location of the stream channel. To do so, we used the definition of bankfull flow as the minimum width-to-depth ratio for any location along the stream channel (Wolman 1955). We used two readily available spatial data sources: light detection and ranging (LiDAR) and National Hydrography Dataset (NHD). LiDAR coverage was accessed through the Missouri Spatial Data Information Service (MSDIS 2011) and flown between 2012 and 2014. LiDAR tiles (1-m horizontal resolution) were mosaicked into a single, seamless digital elevation model (DEM) and resampled to 10-m resolution. LiDAR does not penetrate water, and bathymetric data for the entire watershed do not exist and would be extremely costly to generate. Therefore, we used available elevation data from DEMs and minimum stream-bottom elevation from the National Hydrography Dataset (USGS 2004) as a measure of channel depth.

Throughout the length of the MBB study reaches, we generated 200-m cross sections at 10-m intervals at right angles to the stream centerline. Using these cross sections, the elevations of the underlying DEM were assigned to points along each cross section at 3-m intervals. The width-to-depth ratio was determined by first identifying the minimum elevation along a cross section taken from the underlying LiDAR. We determined that the minimum elevations from LiDAR represented either an in-channel gravel bar or a value generated from the hydro-flattening; either way, the values provide the best available method for determining channel-bank heights over large spatial domains. Using MATLAB

(MathWorks 2016), the horizontal and vertical distances of this minimum elevation point were then calculated relative to all other points in that same cross section. The minimum value of the ratios of the horizontal distances relative to the vertical distances identified the depth of water at which bankfull flows occur (Fig. S1). The resulting water depth was added to the minimum elevation taken from the LiDAR to define the bankfull elevation for each cross section. The bankfull elevations for the cross sections were interpolated into a continuous grid using natural neighbor interpolation. The DEM was then subtracted from the bankfull elevation grid to approximate the channel polygon, where positive values are the channel and negative values are outside the channel. Areas deemed outside the channel were removed through visual inspection (e.g., near-channel gravel mines). The channel polygon was then smoothed, and a stream centerline was derived. Cross sections for each point at 10-m intervals along the derived centerline were then created and clipped to the channel polygon to determine channel width at 10-m intervals (Fig. S1).

After we defined the stream dimensions and location of the channel, we developed four hydrogeomorphic variables generated solely from LiDAR coverage, including two bluff adjacency variables (bluff adjacency, binary, “ba”; bluff adjacency area, continuous, “baa”) and two stream power index variables (stream power class, binary, “spc”; stream power index, continuous, “spi”) (Table 1 and Fig. S2). We identified bluffs as steep cliffs, usually along stream meanders; bluffs are found throughout the MBB. The bluff adjacency variables were created by classifying each pixel in the DEM using a neighborhood search criterion for change in elevation from each pixel centroid. Here pixels were stratified by range criteria to classify those areas of the landscape with a slope equal to or greater than 100% or 45°. Pixels with a range, or change in elevation, from the focal pixel to one of the surrounding eight pixels greater than or equal to a 10-m change in elevation were classified as bluffs. Having classified areas of the landscape as high-slope or bluffs, we then used the stream centerline to search at 10-m intervals for adjacent bluffs within a buffer of 1.5 channel widths (one channel width from each bank). Each point along the centerline was classified as either adjacent or not adjacent to a bluff as the ba variable. The variable baa represented the total bluff area within 1.5 channel widths for each stream centerline point at 10-m intervals. The value or class for bluff adjacency was attributed to each cross section at 10-m intervals and then interpolated using natural neighbors into a continuous gridded variable (Fig. S2).

The stream power variables were created by defining stream power as an index,  $\text{spi} = \ln(A_d) \times S_{500}$ , where spi is the stream power index,  $A_d$  is the total drainage area upstream of the site, and  $S_{500}$  is the slope over 500 m (Moore et al. 1991). Drainage area was solved by burning the stream centerline into the DEM, which enforces correct drainage of the flow direction and accumulation grids, then attributing the drainage area for the centerline points. For each centerline point at 10-m intervals, slope was calculated over a 500-m interval, spanning



Table 1. List of hydrogeomorphic variables generated, including the abbreviated names, the type of layer (continuous/binary), ecological justification, methodological description, and hypotheses of where mussels are expected. \* denotes layers used in our model.

Habitat Characteristic ("layer name"): type	Justification	Description	Hypothesis
*Bluff adjacency area ("baa"): continuous	Conversations with malacologists indicate that mussel beds are usually found in the vicinity of bluffs	Total bluff area (m <sup>2</sup> ) within one channel width of each bank	The probability of mussel presence increases with increasing bluff area adjacent to the channel
Bluff adjacency ("ba"): binary	adjacent to the stream channel	Whether there is a bluff within one channel width of each bank	The probability of mussel presence increases in channels adjacent to bluffs
*Stream power index ("spi"): continuous	Stream power is a major control of slope toe erosion (Nefeslioglu et al. 2008), which can have negative effects on mussels (Hartfield 1993)	Index of potential energy of water in the channel, using $spi = \ln(A_d) \times S_{500}$	The probability of mussel presence increases in areas with moderate stream power
Stream power class ("spc"): binary		Potential energy of water in the stream channel, classed as either high or low, based on spi	The probability of mussel presence increases in areas with low stream power
*Lateral channel stability ("lcs"): binary	Lateral channel movement and bank erosion could disrupt substrate stability and mussel occurrence (Strayer 1999; Strayer et al. 2004)	Lateral channel movement of > 10 m in 17 years classed as unstable, all else classed as stable	The probability of mussel presence increases in stable channels
*Gravel/pool class ("gpc"): binary	(1) Conversations with malacologists indicate that mussels are frequently found near gravel bars, and (2) areas with persistent gravel bars indicate areas that have stable beds, a necessary condition for mussel persistence (Bates 1962; Peck 2005; Zigler et al. 2008)	Reaches dominated by gravel are classed as gravel, all else classed as pool reaches	The probability of mussel presence increases within gravel class reaches
Gravel bar proximity ("gbp"): binary		All areas within 100 m of a gravel bar are classed as adjacent to a gravel bar, all else classed as not adjacent to a gravel bar	The probability of mussel presence increases within 100 m of gravel reaches
*Distance to gravel bar ("dgb"): continuous		Euclidean distance (m) to nearest gravel bar	The probability of mussel presence increases in areas with close proximity to gravel reaches
*Low-flow surface availability index ("lwai"): continuous	Refuge during drought periods is necessary for mussel survival (Golladay et al. 2004)	Cross-sectional average of the area of water pixels surrounding each cell, normalized by stream width	The probability of mussel presence increases in areas with higher low-water availability index values
Low-flow surface water availability class ("lwac"): binary		Cross-sectional average of the area of water pixels surrounding each cell, normalized by stream width, classed as high or low	The probability of mussel presence increases in areas with high low-water availability classification

$A_d$  is the total drainage area upstream of the site, and  $S_{500}$  is the slope over 500 m.

250 m upstream and downstream of every point. A moving average of 50 m was used to smooth the estimates of slope. The value of spi for each point was attributed to the corresponding cross section at 10-m intervals and then interpolated using natural neighbors into a continuous gridded variable. The gridded variable was classified into a binary variable of high and low stream power classes using the mean value as the break between the two classes (spc; Fig. S2).

Aerial imagery was used as the primary source for deriving six additional hydrogeomorphic variables (Table 1). Three variables characterized the channel based on the properties of gravel bars in the imagery (gravel/pool class, binary, "gpc"; gravel bar proximity, binary, "gbp"; distance to gravel bar, continuous, "dgb") (Fig. S3). Because bathymetry data were not available, we created two variables as proxies for water availability during low flows (low-flow surface water

availability index, continuous, “lwai”; low-flow surface water availability class, binary, “lwac”) (Fig. S4). The final variable represented lateral channel stability (“lcs”) (Fig. S5). To derive the gravel variables, unsupervised classifications were performed on images from the National Agriculture Imagery Program (NAIP; 1-m horizontal resolution) (MSDIS 2011). This is leaf-on imagery gathered over six days in June 2012 and seven days in July 2014 at low water conditions, which were conducive for identifying exposed gravel bars. The raster grid was converted to a polygon vector layer and polygons were classified as either gravel, water, or other by using the ISO Cluster Unsupervised Classification in the Spatial Analyst Toolbox in ArcMap 10.7. The classification results were processed by filtering, smoothing, and cleaning boundaries using respective tools in the Spatial Analyst Toolbox. The postprocessed classification was visually inspected for misclassified regions by comparing to the original aerial imagery. We identified stable gravel bars by using only pixels that were identified as gravel in both 2012 and 2014 imagery; all others were identified as pool. These data worked well for our study because a near-record flood occurred in 2013; therefore, those gravel bars that remained after that flood were considered stable. The cross sections derived from LiDAR were then used to classify the reaches as either a gravel bar or pool. If a cross section was intercepted by a gravel polygon, then that entire cross section was classified as gravel. Due to the nature of the error in data collection for in-channel habitat characteristics from left bank to right bank and the scale of our environmental data, we did not distinguish between suitable and unsuitable habitat laterally within the channel. Instead, we focused on delineating whole reaches longitudinally along the river as suitable or unsuitable habitat.

The variable gpc was derived by interpolating using natural neighbors to create a continuous, longitudinal representation of gravel and pool classes (Fig. S3). The variable gbp was derived by classifying each centerline point as a distance either greater than or less than 100 m to a gravel reach (Fig. S3). The dgb variable was the continuous representation of the Euclidean distance to gravel bars for every centerline point (Fig. S3).

We derived low-water availability variables (lwai and lwac) by performing an eight-cell focal search window on each cell within the channel to assign the total area of water pixels surrounding each cell, including that cell’s area. The cross sections were assigned the average value of the pixels intercepting each cross section, which was then divided by channel width to account for changes in channel size due to drainage-area scaling. The variable lwai was derived by interpolating cross sections to generate a continuous grid of low-flow surface water availability index. The continuous lwai grid was classified into a binary variable of high- and low-flow water surface availability (lwac) using the median value as the threshold between the classes (Fig. S4). We generated 100 random points within our channel to validate the unsupervised classifications. Each point was assigned the class code corresponding with the location of the classified layer, with

92% of the points correctly classified through visual inspection of aerial imagery. We acknowledge that this variable does not perfectly represent the vulnerability of any given point to dewatering during low flow because we do not have depth data. However, at the riverscape scale in which we are modeling, our low-flow water availability variables can help identify areas likely to contain water during low flow periods, making them potentially important target locations for managers.

We used aerial imagery to classify the channel into one of two lateral channel stability classes (lcs: laterally stable and laterally unstable), using Leaf-off Digital Orthophoto Quarter Quads (DOQQs) from 1990 (1-m horizontal resolution) and 2007 (0.6-m horizontal resolution) (MSDIS 2011). We used this set of imagery because the leaf-off aspect makes delineating the stream channel less challenging. The approximate bankfull lines for each year were digitized using visual clues, including shadows, breaks in vegetation and substrate, scour lines from high flows, and an overlain semitransparent hillshade from the DEM. Areas where the lines diverged by more than 10 m between the two sets of images were classified as laterally unstable, while areas where the lines did not diverge by more than 10 m were classified as laterally stable (Fig. S5).

### Model Building

The maximum entropy modeling method known as MaxEnt (Phillips 2017) was used to generate habitat models for mussel beds in the MBB. This method uses presence-only data to find the probability distribution of maximum entropy (i.e., closest to uniform) given constraints of known locations and hydrogeomorphic variables relative to the spatial extent of the analysis (Raxworthy et al. 2007). Because absence points in our case could not be generated with certainty, MaxEnt generated 10,000 pseudo-absence points automatically. MaxEnt generated a map showing predicted habitat suitability for each area of the landscape (given the spatial grain size) with values ranging from 0 to 1, where 0 is the most unsuitable and 1 is the most suitable habitat. A portion of mussel-bed locations ( $n = 42$ ) were used in the MaxEnt model. All continuous hydrogeomorphic variables used in the final model were not highly correlated. Most variables had a correlation coefficient of  $< 0.29$ ; however, the variables dgb and gpc had a slightly higher correlation coefficient of 0.58. The relative contribution of hydrogeomorphic variables in the habitat suitability model was assessed in MaxEnt via Jackknife analysis. Run type was set to bootstrap to generate test data with 20% of the presence data, and random seed was chosen to randomize test data. Replicates were set to 50 and iterations set to 5,000. All other settings in MaxEnt were set to default. The models were built using 80% of the total amount of presence data used in the model ( $n = 42$ ), referred to as “training data,” to generate the algorithms relating the hydrogeomorphic variables to the habitat suitability of every parcel on the landscape. The remaining 20% of mussel-bed locations that were withheld from MaxEnt were used to test the model. The

Table 2. Summary of values for each of the two classes of the six binary hydrogeomorphic variables and the values of the sample mussel beds for the same six layers. Included for both the six layers and the sample points are the percentages for each class, the minimum and maximum lengths (m) of each reach class, and the mean and standard deviation of each reach class.

Habitat characteristic	Class	Percent of layers	Percent of samples	Reach length minimum	Reach length maximum	Mean reach length	Reach length standard deviation
Bluff adjacency (“ba”)	Adjacent	40.6	40.6	30	3,885	615	523
	Not adjacent	59.4	59.4	10	9,977	886	1,159
Gravel/pool class (“gpc”)	Gravel	51.3	59.4	37	9,943	603	874
	Pool	48.7	40.6	50	4,392	571	498
Gravel bar proximity (“gbp”)	<100 m	67.3	71.0	61	1,1761	845	1,028
	>100 m	32.7	29.0	10	4,187	412	502
Lateral channel stability (“lcs”)	Stable	85.4	88.4	148	47,643	4,464	6,395
	Unstable	14.6	11.6	73	4,545	784	869
Low-flow surface water availability class (“lwac”)	High	58.1	55.1	95	7,216	1,392	1,304
	Low	41.9	44.9	113	11,819	989	1,352
Stream power class (“spc”)	High	43.1	50.7	47	6,661	1,014	900
	Low	56.9	49.3	142	7,427	1,365	1,334

area under the receiver operating curve (AUC) measures the probability that presence locations have a higher habitat suitability score than randomly chosen pseudo-absence points (Phillips and Dudik 2008). Models with average AUC values greater than 0.5 are considered sufficiently better than a random model at distinguishing among habitat suitability of presence locations from pseudo-absence points (Elith 2002). We used the test gain value to assess which environmental variables were the most important for model fit (Phillips 2017). Gain is a likelihood (deviance) statistic that maximizes the probability of the presence in relation to the background (pseudo-absence) data. Taking the exponent of the final gain gives the (mean) probability of the presence sample(s) compared to the pseudo-absences. We developed response curves to investigate the relationships between specific values within hydrogeomorphic variables and suitable/unsuitable reaches. In the absence of information-theoretic approaches available for MaxEnt model selection, we selected the best-fit model as our final model using our 10 variables (Table 1) and a stepwise model selection approach, preferentially selecting variables in a stepwise manner leading to high total model AUC values and containing only those variables with sizeable individual effects on model results when other variables are removed (following Elith 2002).

The raw model results were converted to a binary map of suitable and unsuitable reaches using the equal test sensitivity and specificity logistic threshold of 0.45. This commonly used threshold sets sensitivity equal to specificity (Phillips 2017; Cao et al. 2013). Once reaches were delineated, a buffer of 40 m was used to separate suitable and unsuitable habitats to account for areas of transition. The buffer size selected represents the average length of transitional values of habitat suitability seen across the watershed between long, continuous reaches of high habitat suitability values versus low suitability values. The portions of river highlighted as suitable habitat by this approach are measurable reaches of suitable habitat. These

reaches vary in length, with the length depending on the continuity of stream characteristics and the relationship of mussel presence to those characteristics.

All spatial analyses were performed in ArcGIS and projected to NAD 1983 UTM Zone 15N (ESRI 2011). Due to computational limitations, we were not able to perform the image classification analysis on aerial imagery at 1 m. The finest resolution we were able to classify and process was 10 m. Therefore, all other layers were resampled to a 10-m resolution using majority setting on the resample tool in ArcMap, and the final product has a 10-m resolution.

### Model Validation

The remaining mussel-bed locations that were not used in model development ( $n = 64$ ) were used to validate the model. The location within the mussel bed at which these GPS points were taken is unknown in this dataset. To reduce the potential GPS errors on our results, we considered points that were located within 180 m (average mussel-bed length in the MBB; see above) of a suitable reach to be within suitable habitat. We reported the percentage of validation mussel-bed locations that were within a suitable reach (predicted by the MaxEnt model). We used a Pearson’s chi-squared test to determine if validation mussel-bed locations were found disproportionately in suitable versus unsuitable reaches.

## RESULTS

### Hydrogeomorphic Variables

*LiDAR derived variables.*—Bluff adjacency area (baa) ranged from 0 to 28,173 m<sup>2</sup>, with a mean of 1,393 m<sup>2</sup>. For bluff adjacency (ba), 41% of total channel length was adjacent to a bluff, and 59% was not (Table 2). The baa for the entire

Table 3. Minimum, maximum, mean, and standard deviation for the four continuous habitat layers that we generated and the values of the layers at mussel-bed locations.

Habitat characteristic	Layer/Location	Minimum	Maximum	Mean	Standard deviation
Bluff adjacency area ("baa")	Layer	0	28,173	1,393	2,769
	Mussel location	0	17,678	1,129	2,672
Distance to gravel bar ("dgb")	Layer	0	1,820	121	222
	Mussel location	0	371	67	105
Low-flow surface water availability index ("lwai")	Layer	0	13.30	4.53	1.98
	Mussel location	0.78	8.55	4.34	1.73
Stream power index ("spi")	Layer	-0.0146	0.0243	0.0035	0.0031
	Mussel location	-0.0073	0.0120	0.0037	0.0029

model extent and mussel-bed locations had a 21% difference between the means (Table 3). For ba, the percentages representing the length of each class comprising the entire model extent and mussel-bed locations were identical. Stream power index (spi) values ranged from -0.0146 to 0.0243, with a mean of 0.0035. The spi for the entire model extent and mussel-bed locations had a 6% difference between the means (Table 3). For stream power class, (spc), 43% of total channel length had high stream power, and 57% had low stream power (Table 2). The spc had a 7.6% difference for the percentage of each class comprising the entire model extent and mussel-bed locations.

*Aerial-imagery-derived variables.*—The gravel class constituted 51% of the 530 river km, while pool class was 49% of the total channel length. The gpc had the largest discrepancy between entire model extent and mussel-bed locations with an 8.1% difference between the two. The gbp variable had a 3.7% difference for the percentage of each class comprising the entire model extent and mussel-bed locations. The percent difference between the means of the entire model extent and mussel-bed locations was 58% for the dgb, which was the greatest difference between the variables and bed locations among the continuous variables (Table 3).

The continuous representation of the low-flow surface water availability index values ranged from 0 to 13.3 and had an average index value of 4.5 in the lwai variable. The lwai and mussel-bed locations had a 4% difference between the means, which was the lowest among the continuous variables (Table 3). For the binary lwac variable, average reach length was 1,392 m. The high class in the lwac was 58% of the total reach length, while the low class was 42% (Table 2). The lcs variable was classified as either laterally stable or unstable, where 85% of the 530 river km were classified as stable and 15% as unstable. The average reach length for the stable class was 4,464 m and for the unstable class was 784 m (Table 2). The lwai and lcs both had a 3% difference for the percentage of each class comprising the entire model extent and mussel-bed locations.

## Model Results

*Model generation.*—The training and test AUC values of the top model were 0.75 and 0.62, respectively. Six hydrogeomor-

phic variables were used in the best model: lcs, dgb, gpc, spi, baa, and lwai (Table 1). The model results show separation of habitat into suitable and unsuitable habitat (Fig. 3). Jackknife analysis indicated that lcs, dgb, lwai, and spi contributed significantly to the final model, while baa and gpc did not (Fig. 4). Response curves indicated that suitable reaches differed from unsuitable reaches in values of baa, lcs, dgb, spi, and lwai (Fig. 5). However, all available spi values in the response curve were above the 0.5 probability of presence; therefore, we cannot report definitively relationship of mussel presence with this variable. More specifically, areas predicted as suitable were in reaches with laterally stable channels, with zero or very low bluff area, near gravel bars, with slightly higher stream power, and with greater areas of contiguous water (Fig. 5).

*Model validation.*—Eighty-three percent (53 of 64) of the validation mussel beds fell within a suitable reach in the binary suitability model, and mussel beds were found disproportionately within suitable reaches ( $\chi^2 = 18.77$ , 1 df,  $P < 0.05$ ). Eleven beds fell within reaches that the model predicted as unsuitable. Seven of these beds were within a river reach with a split channel (see Discussion) and one was in a relatively short, unsuitable reach between two long, suitable reaches. The remaining three beds were centrally located within an unsuitable habitat reach and had no evident characteristics that could help inform future modeling efforts.

## DISCUSSION

Our study successfully delineated suitable habitat for mussel presence in the MBB. Our model results suggest that mussel beds are dependent on multiple channel features, particularly the presence of laterally stable channels, proximity to gravel bars, and greater contiguous expanses of water. One potential source of bias in our dataset is that mussel biologists typically focus their sampling effort on gravel bars and avoid less-accessible, deeper areas of rivers, and our dataset suffers from this source of bias. Sampling a randomly selected set of stream reaches identified by our models as unsuitable is necessary to ground truth our model results. Nevertheless, our approach shows that readily available aerial imagery and topography data can provide useful description of stream-



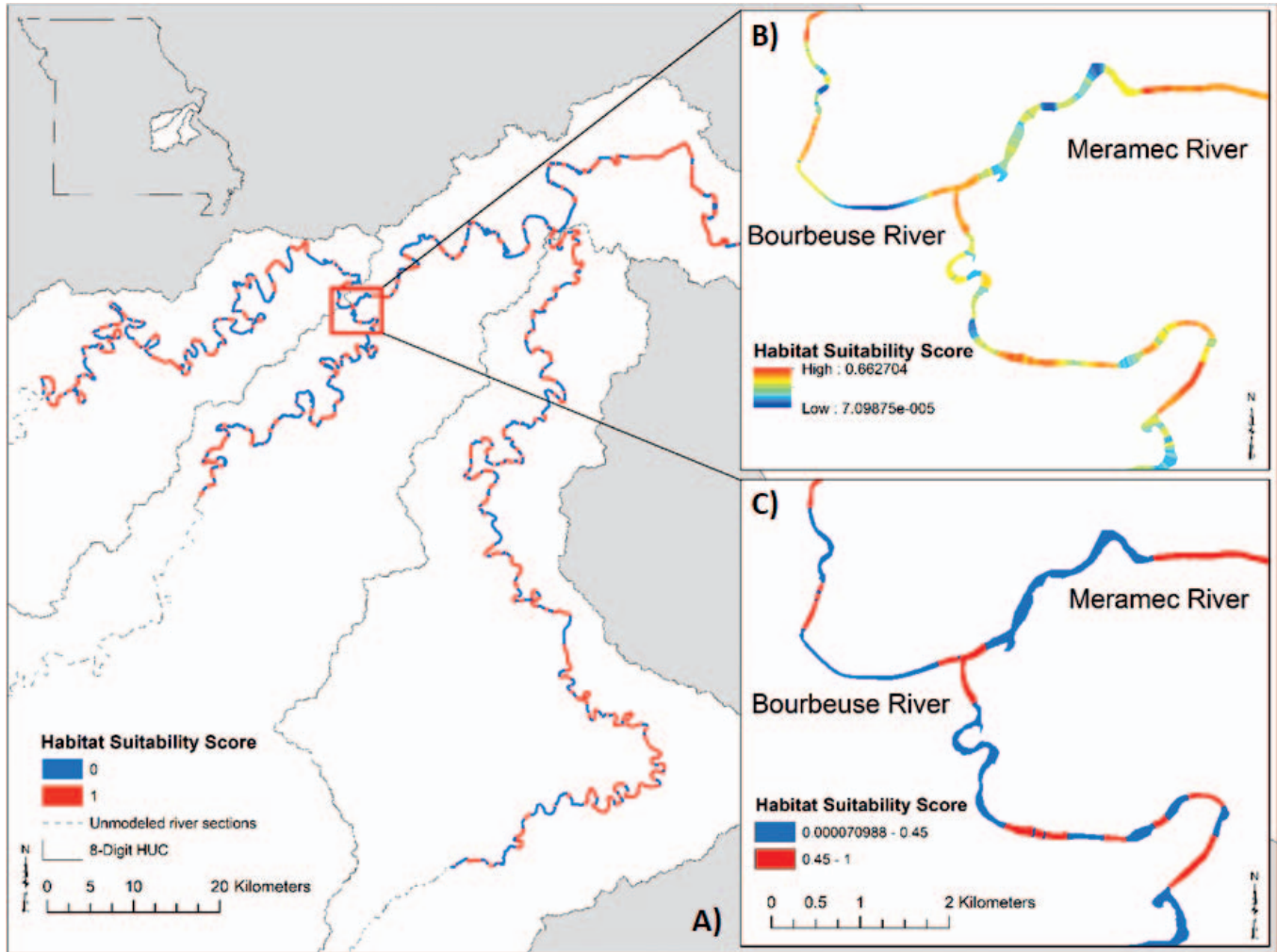


Figure 3. Model results for classifying stream reaches in the Meramec River basin with regard to their suitability for mussel beds. (A) Map of the entire study area showing a binary classification (suitable or unsuitable). (B) Detailed map of an example section of the watershed showing continuous suitability scores. (C) Detailed map of an example section of the watershed showing binary scores.

reach characteristics associated with the occurrence of mussel beds at a riverscape scale.

Other aspects of our model findings require ground-truthing due to the nature of available large-scale hydrogeomorphic data. For example, we used aerial imagery to

define low-flow surface water availability as a proxy for identifying potential drought refugia. Ground-truthing based on bathymetric data is necessary to evaluate the extent to which low-flow surface water availability represents vulnerability to emersion during drought. Similarly, lateral bank stability may or may not reflect the type of substrate stability that other papers have associated with mussel occurrence on a smaller spatial scale. Evaluating these factors can help refine approaches for utilizing large-scale watershed data. Other challenges associated with using riverscape-scale hydrogeomorphic data include lack of availability of such data in some areas, extensive canopy shading of smaller streams, and the computational power required to run models with detailed spatial information.

Many factors that influence mussel presence were not included in our model, such as fish-host relationships, species-specific differences in habitat requirements, and anthropogenic influences. However, identifying minimum characteristics of habitats necessary to support mussel presence provides a

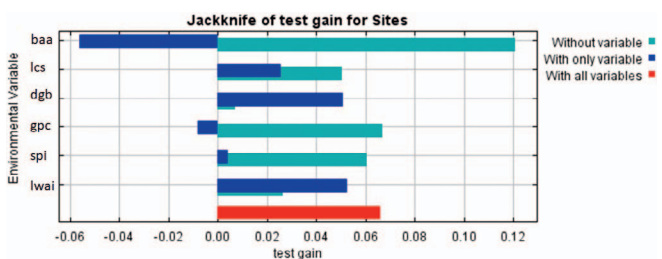


Figure 4. Results of jackknife analysis of the contribution of individual habitat variables to model performance. Final model variables include bluff area adjacency (baa), lateral channel stability (lcs), distance to gravel bar (dgb), gravel/pool class (gpc), stream power index (spi), and low water availability (lwai).

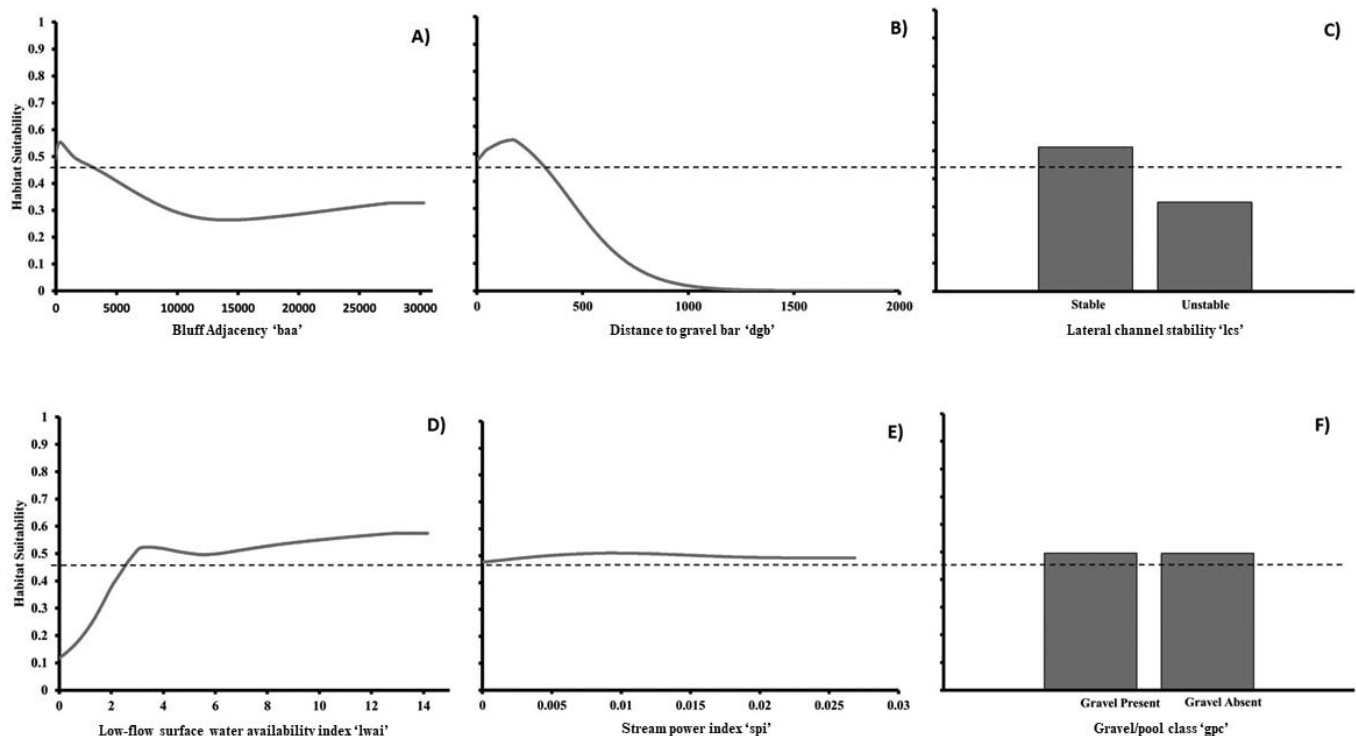


Figure 5. Response curves for hydrogeomorphic variables used in the habitat model. (A) bluff area adjacency (baa), (B) distance to gravel bar (dgb), (C) lateral channel stability (lcs), (D) low water availability (lwai), (E) stream power index (spi), and (F) gravel/pool class (gpc). The dashed line represents the equal sensitivity and specificity logistic threshold used to delineate habitats predicted as suitable ( $\geq 0.45$ ) and unsuitable ( $< 0.45$ ).

baseline that can allow the effects of other factors to be evaluated (Bouska et al. 2018; Fig. 1). Although the datasets we used are specific to the Meramec River basin, our approach is amenable to other river systems where researchers have access to comprehensive mussel-survey data. LiDAR is quickly becoming ubiquitous in the contiguous United States, and high-resolution aerial imagery for some areas often can be found dating back decades, making this approach a reasonable option in some areas where data are available. Our methods also may be applicable for other benthic organisms, such as snails, insects, or crayfish.

We developed an approach that uses readily available, continuous, longitudinal data to describe associations of hydrogeomorphic features and the presence of mussel beds to identify suitable mussel habitat at a large scale. We specifically address a recognized knowledge gap for understanding mussel-habitat distributions (FMCS 2016) by using hydrogeomorphic variables and scaling them to the channel in a continuous, longitudinal manner. By using riverscape-scale, fundamental niche habitat modeling, managers can target specific river reaches for mussel surveys, reintroduction efforts, and other management activities.

## ACKNOWLEDGMENTS

We thank the following individuals for contributions throughout various stages of the project: Scott Faiman (MDC), Ivan Vining (MDC), Rob Pulliam (MDC), Andy

Roberts and Joshua Hundley (US Fish and Wildlife Service), Craig Scroggins (MDC), and Jodi Whittier (MU). We also would like to thank the reviewers of this manuscript for their comments and suggestions. This work was funded by the Missouri Department of Conservation. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

## LITERATURE CITED

- Allen, D. C., and C. C. Vaughn. 2010. Complex hydraulic and substrate variables limit freshwater mussel species richness and abundance. *Journal of North American Benthological Society* 29:383–394.
- Arbuckle, K. E., and J. Downing. 2002. Freshwater mussel abundance and species richness: GIS relationships with watershed land use and geology. *Canadian Journal of Fisheries and Aquatic Sciences* 59:310–316.
- Bates, J. M. 1962. The impact of impoundment on the mussel fauna of Kentucky Reservoir, Tennessee River. *American Midland Naturalist* 68:232–236.
- Bouska, K., A. Rosenberger, S. McMurray, G. Linder, and K. N. Key. 2018. State-level mussel management and conservation: Current status and a strategic path forward. *Fisheries* 43:345–360.
- Cao, Y., E. R. DeWalt, J. L. Robinson, T. Tweddle, L. Hinz, and M. Pessino. 2013. Using Maxent to model the historic distributions of stonefly species in Illinois streams: The effects of regularization and threshold selections. *Ecological Modelling* 259:30–39.
- Daniel, W., and K. M. Brown. 2014. The role of life history and behavior in explaining unionid mussel distributions. *Hydrobiologia* 734:57–68.
- Drew, C. A., M. Eddy, T. J. Kwak, W. G. Cope, and T. Augspurger. 2018. Hydrologic characteristics of freshwater mussel habitat: Novel insights from modeled flows. *Freshwater Science* 37:343–356.

- Elith, J. 2002. Quantitative methods for modeling species habitat: Comparative performance and an application to Australian plants. Pages 39–59 in S. Ferson and M. Burgman, editors. *Quantitative Methods for Conservation Biology*. Springer, New York.
- ESRI (Environmental Systems Research Institute). 2011. ArcGIS Desktop: Release 10. ESRI, Redlands, CA.
- Fausch, K. D., C. E. Torgersen, C. V. Baxter, and H. W. Li. 2002. Landscapes to riverscapes: Bridging the gap between research and conservation of stream fishes. *BioScience* 52:483–498.
- FMCS (Freshwater Mollusk Conservation Society). 2016. A national strategy for the conservation of native freshwater mollusks. *Freshwater Mollusk Biology and Conservation* 19:1–21.
- Golladay, S. W., P. Gagnon, M. Kerans, J. M. Battle, and D. W. Hicks. 2004. Response of freshwater assemblages (*Bivalvia*: Unionidae) to a record drought in the Gulf Coastal Plain of southwestern Georgia. *Journal of the North American Benthological Society* 23:494–506.
- Haag, W. R. 2010. A hierarchical classification of freshwater mussel diversity in North America. *Journal of Biogeography* 37:12–26.
- Haag, W. R. 2012. *North American Freshwater Mussels: Natural History, Ecology, and Conservation*. Cambridge University Press, Cambridge, UK.
- Hardison, B. S., and J. B. Layzer. 2001. Relations between complex hydraulics and the localized distribution of mussels in three regulated rivers. *Regulated Rivers: Research & Management* 17:77–84.
- Hartfield, P. 1993. Headcuts and their effect on freshwater mussels. Pages 131–141 in K. S. Cummings, A. C. Buchanan, and L. M. Koch, editors. *Conservation and Management of Freshwater Mussels*. Upper Mississippi River Conservation Committee, Rock Island, IL.
- Hinck, J. E., S. E. McMurray, A. D. Roberts, M. C. Barnhart, C. G. Ingersoll, N. Wang, and T. Augspurger. 2012. Spatial and temporal trends of freshwater mussel assemblages in the Meramec River Basin, Missouri, USA. *Journal of Fish and Wildlife Management* 3:319–331.
- Homer, C. H., J. A. Fry, and C. A. Barnes. 2012. The National Land Cover Database, U.S. Geological Survey Fact Sheet 2012–302. Available at <https://www.mrlc.gov/data?f%5B0%5D=category%3Aland%20cover> (accessed November 2015).
- Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harbour Symposium on Quantitative Biology* 22:415–427.
- Jacobson, R. B., and A. T. Primm. 1997. Historical land-use changes and potential effects on stream disturbance in the Ozark Plateaus, Missouri. US Geological Survey Water-Supply Paper 2484, Denver, CO. Available at <https://pubs.usgs.gov/wsp/2484/report.pdf>, accessed August 11, 2021.
- Key, K. N. 2019. A spatial assessment of the status and risks to mussel concentrations in the Meramec Drainage of Missouri. PhD dissertation, Tennessee Technological University, Cookeville.
- Layzer, J. B., and L. M. Madison. 1995. Microhabitat use by freshwater mussels and recommendations for determining their instream flow needs. *Regulated Rivers: Research and Management* 10:329–345.
- Luekenhoff, L. 2015. Development of standardized visual sampling methods for assessing community metrics of unionoid mussel species and tribal groups in Missouri. MS thesis, University of Missouri, Columbia.
- MathWorks. 2016. MATLAB version 9.0. MathWorks, Natick, MA.
- McRae, S. E., J. D. Allan, and J. B. Burch. 2004. Reach- and catchment-scale determinants of the distribution of freshwater mussels in south-eastern Michigan, U.S.A. *Freshwater Biology* 49:127–142.
- Moore, I. D., R. B. Grayson, and A. R. Ladson. 1991. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrological Processes* 5:3–30.
- MSDIS (Missouri Spatial Data Information Service). 2011. Curators of the University of Missouri-Columbia. Available at <http://msdis.missouri.edu/> (accessed March 2015).
- Nefeslioglu, H. A., T. Y. Duman, and S. Durmaz. 2008. Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea Region of Turkey). *Geomorphology* 94:401–418.
- Pandolfo, T. J., T. J. Kwak, W. G. Cope, R. J. Heise, R. B. Nichols, and K. Pacifici. 2016. Species traits and catchment-scale habitat factors influence the occurrence of freshwater mussel populations and assemblages. *Freshwater Biology* 61:1671–1684.
- Parsons, M., M. C. Thoms, and R. H. Norris. 2004. Using hierarchy to select scales of measurement in multiscale studies of stream macroinvertebrate assemblages. *Freshwater Science* 23:157–170.
- Peck, A. J. 2005. A reach scale comparison of fluvial geomorphological conditions between current and historic freshwater mussel beds in the White River, Arkansas. Master's thesis, Arkansas State University, Jonesboro.
- Phillips, S. J. 2017. A brief tutorial on MaxEnt. Available from url: [http://biodiversityinformatics.amnh.org/open\\_source/maxent/](http://biodiversityinformatics.amnh.org/open_source/maxent/). Accessed August 11, 2021.
- Phillips, S. J., and M. Dudik. 2008. Modeling of species distributions with MaxEnt: New extensions and a comprehensive evaluation. *Ecography* 31:161–175.
- Raxworthy, C. J., C. M. Ingram, N. Rabibisoa, and R. G. Pearson. 2007. Applications of ecological niche modeling for species delimitation: A review and empirical evaluation using day geckos (*Phelsuma*) from Madagascar. *Systematic Biology* 56:907–923.
- Roberts, A. D., and S. Bruenderman. 2000. A reassessment of the status of freshwater mussels in the Meramec River basin, Missouri. Report prepared for the US Fish and Wildlife Service, Fort Snelling, MN.
- Schrum, M. C. 2017. Development and validation of standardized sampling protocols for assessing freshwater mussel populations in Missouri. Master's thesis, University of Missouri, Columbia.
- Steuer, J. J., T. J. Newton, and S. J. Zigler. 2008. Use of complex hydraulic variables to predict the distribution and density of unionids in a side channel of the Upper Mississippi River. *Hydrobiologia* 610:67–82.
- Strayer, D. L. 1983. The effects of surface geology and stream size on freshwater mussel (*Bivalvia*: Unionidae) distribution in southeastern Michigan, USA. *Freshwater Biology* 13:253–264.
- Strayer, D. L. 1999. Use of flow refuges by unionid mussels in rivers. *Journal of North American Benthological Society*. 18:468–476.
- Strayer, D. L. 2008. *Freshwater Mussel Ecology: A Multifactor Approach to Distribution and Abundance*. University of California Press, Berkeley, CA.
- Strayer, D. L., J. A. Downing, W. R. Haag, T. L. King, J. B. Layzer, T. J. Newton, and S. Jerrine Nichols. 2004. Changing perspectives on pearly mussels, North America's most imperiled animals. *Bioscience* 54:429–439. Accessed August 11, 2021.
- TNC (The Nature Conservancy). 2014. Meramec River Conservation Plan. The Nature Conservancy, Missouri Chapter, St. Louis, MO. Available at <https://www.nature.org/media/missouri/meramec-river-conservation-action-plan.pdf>.
- USGS (US Geological Survey). 2004. National Hydrography Dataset. USGS, Reston, VA. Available at [https://www.usgs.gov/core-science-systems/ngp/national-hydrography/national-hydrography-dataset?qt-science\\_support\\_page\\_related\\_con=0#qt-science\\_support\\_page\\_related\\_con](https://www.usgs.gov/core-science-systems/ngp/national-hydrography/national-hydrography-dataset?qt-science_support_page_related_con=0#qt-science_support_page_related_con), accessed July 28, 2021.
- Vaughn, C. C. 1997. Regional patterns of mussel species distributions in North American rivers. *Ecography* 20:107–115.
- Walters, A. D., D. Ford, E. R. Chong, M. G. Williams, N. B. Ford, L. R. Williams, and J. A. Banta. 2017. High-resolution ecological niche modeling of threatened freshwater mussels in east Texas, USA. *Aquatic Conservation: Marine and Freshwater Ecosystems* 27:1251–1260. doi: 10.1002/aqc.2784
- Wolman, M.G. 1955. The natural channel of Brandywine Creek, Pennsylvania, U.S. Geological Survey Professional Paper 271. Available at <https://doi.org/10.3133/pp271>, accessed July 28, 2021.
- Zigler, S. J., T. J. Newton, J. J. Steuer, M. R. Bartsch, and J. S. Sauer. 2008. Importance of physical and hydraulic characteristics to unionid mussels: A retrospective analysis in a reach of large river. *Hydrobiologia* 598:343–360.
- Zipkin, E. F., A. DeWan, and J. A. Royle. 2009. Impacts of forest fragmentation on species richness: A hierarchical approach to community modelling. *Journal of Applied Ecology* 46:815–882.



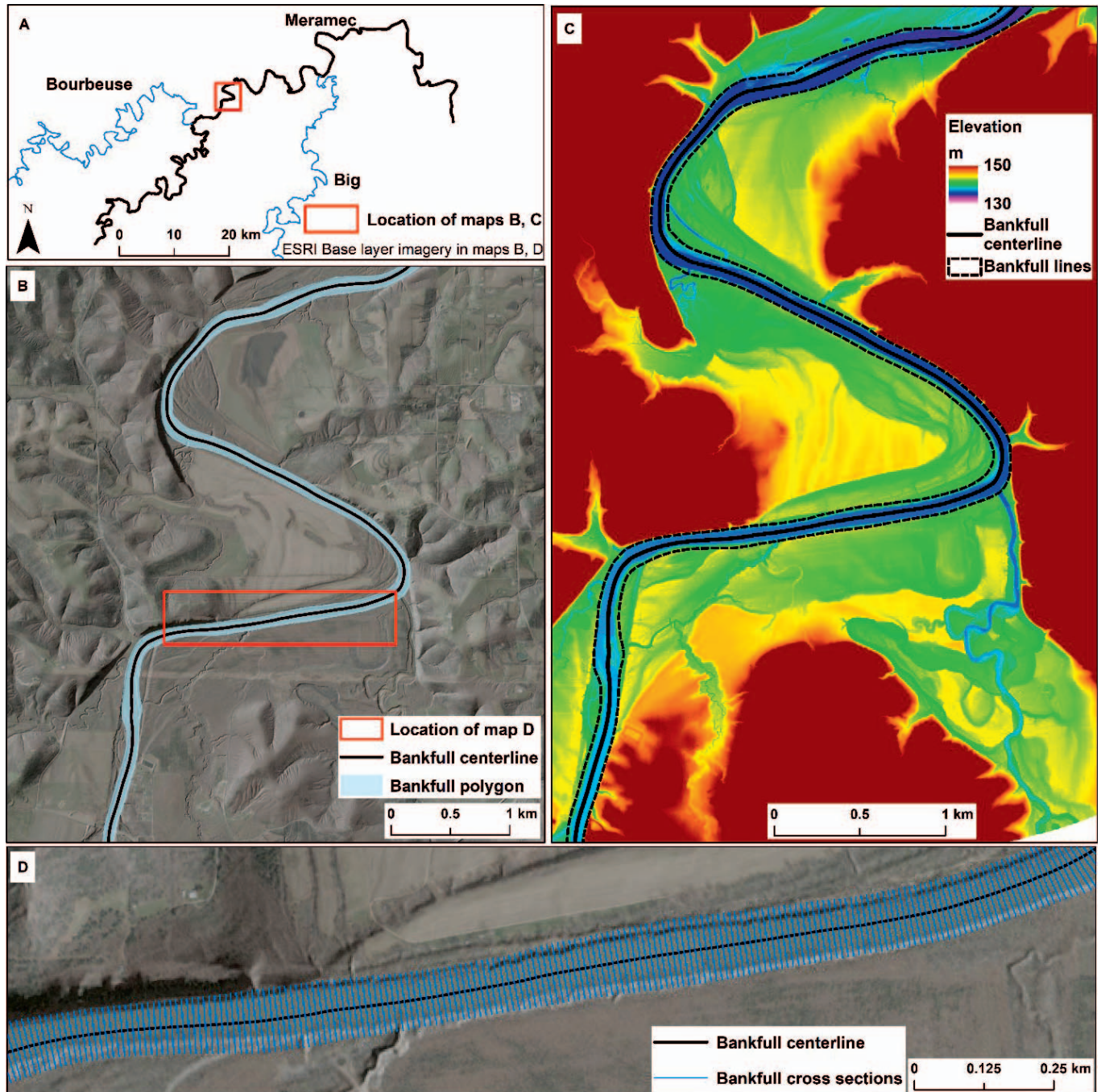


Figure S1. Examples of the stream channel delineation from LiDAR: A) main channel network of the Meramec River watershed, B) the stream centerline and the bankfull polygon superimposed over aerial imagery and a semi-transparent hillshade of the LiDAR digital elevation model, C) the stream centerline and the bank lines superimposed over the DEM used for the channel delineation, and D) the stream centerline and bankfull cross sections superimposed over aerial imagery and a semi-transparent hillshade of the LiDAR digital elevation model.



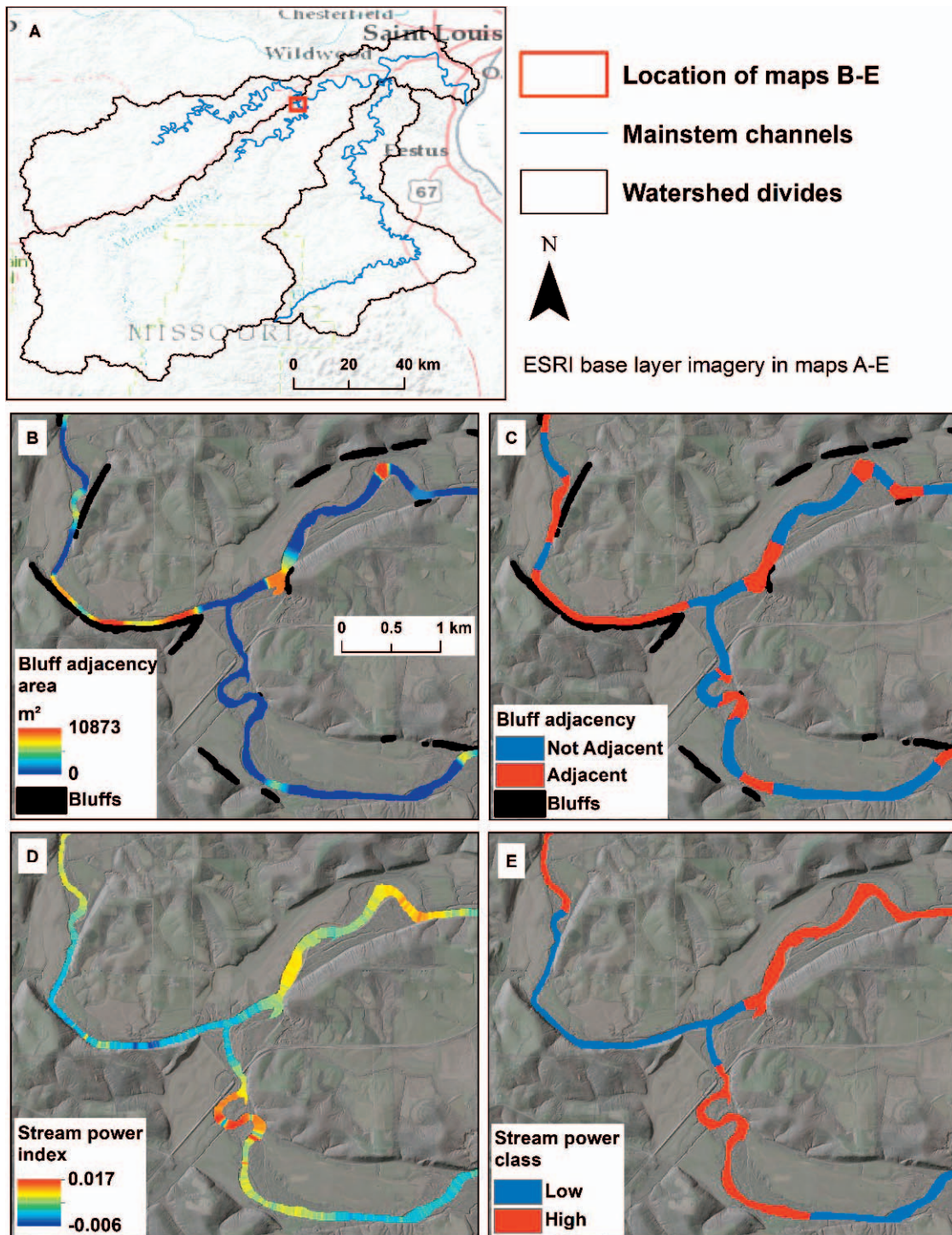


Figure S2. Example maps of bluff adjacency and stream power layers: A) map of the Meramec River watershed with locations of the subsequent detailed maps, B) bluff adjacency area (baa) continuous layer with classified bluffs, C) bluff adjacency (ba) binary layer with classified bluffs, D) stream power index (spi) continuous layer, and E) stream power class (spc) binary layer. The base layers for maps B-E are composed of aerial imagery overlaid with a semi-transparent hillshade derived from the LiDAR digital elevation model.



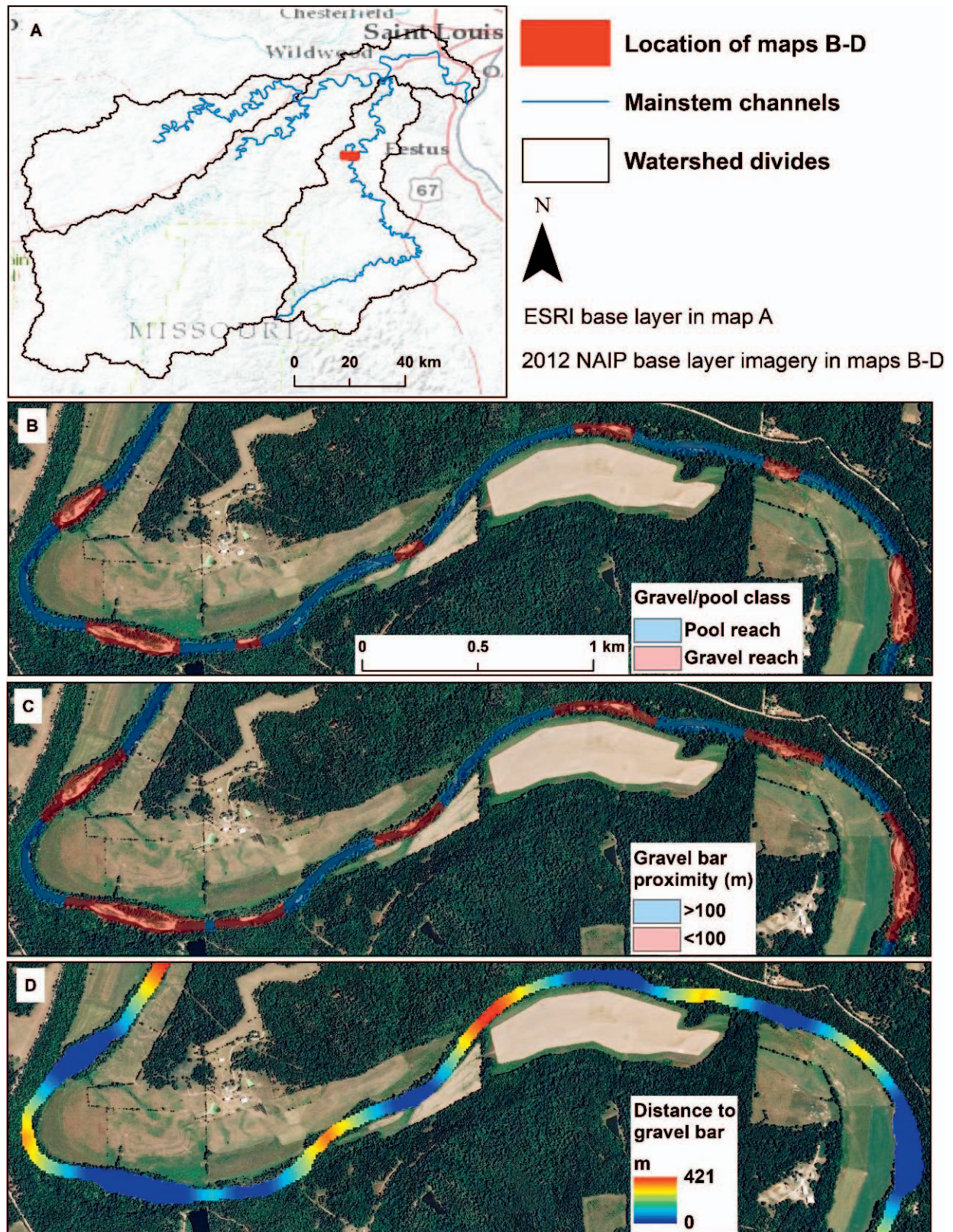


Figure S3. Example maps of gravel layers: A) map of the Meramec River watershed with locations of the subsequent detailed maps, B) semi-transparent gravel/pool class (gpc) binary layer superimposed over aerial imagery to show underlying stream channel, C) semi-transparent gravel bar proximity (gbp) binary layer, and D) distance to gravel bar (dgb) continuous layer.



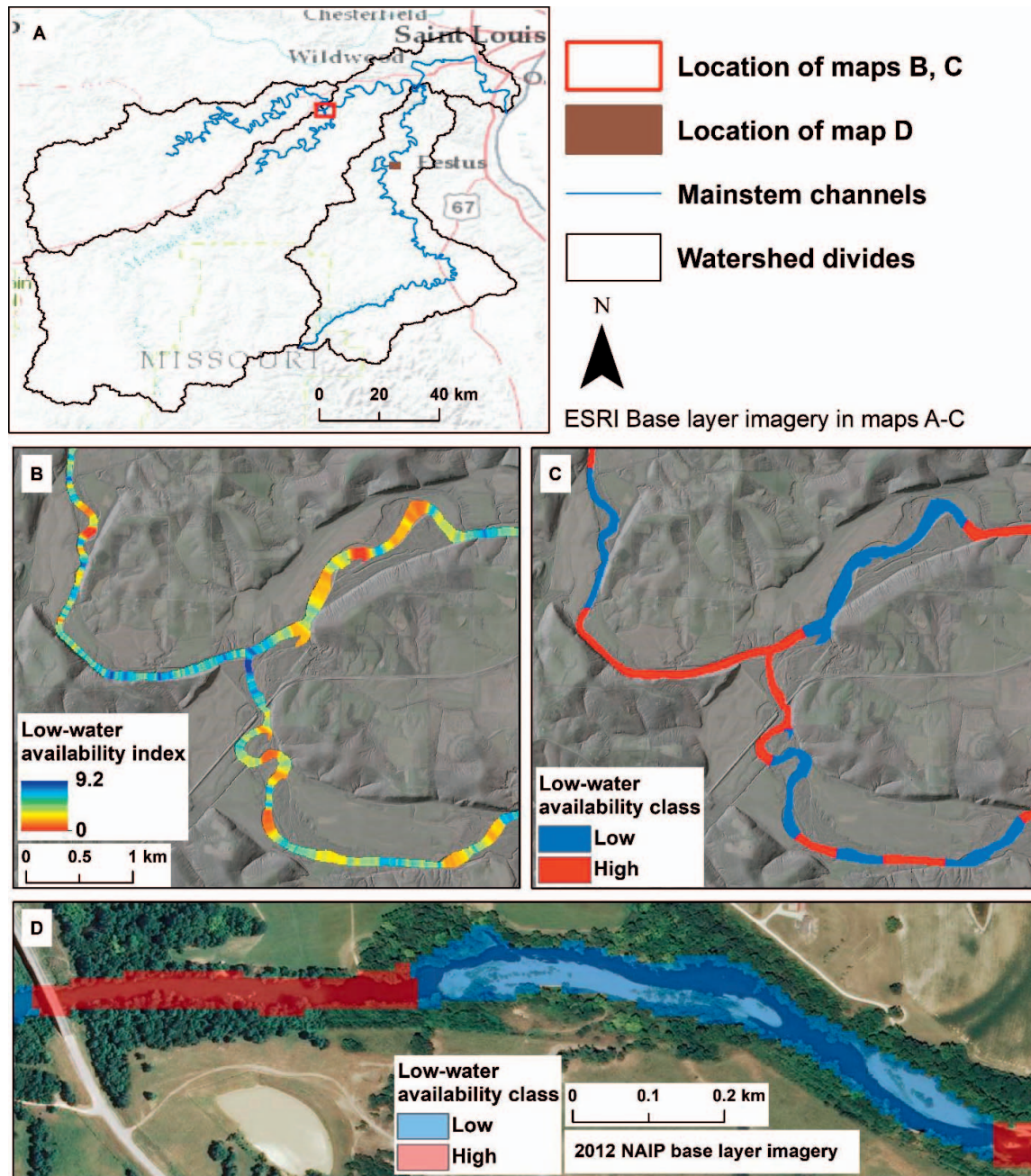


Figure S4. Example maps of low-flow surface water availability layers: A) map of the Meramec River watershed with locations of the subsequent detailed maps, B) low-flow surface water availability index (lwai) continuous layer, C) low-flow water availability class (lwac) binary layer, and D) large-scale example of semi-transparent low-flow surface water availability class binary layer showing the underlying stream channel classified as either low low-water surface availability or high low-flow surface water availability. The base layers for maps B and C are composed of aerial imagery overlaid with a semi-transparent hillshade derived from the LiDAR digital elevation model.

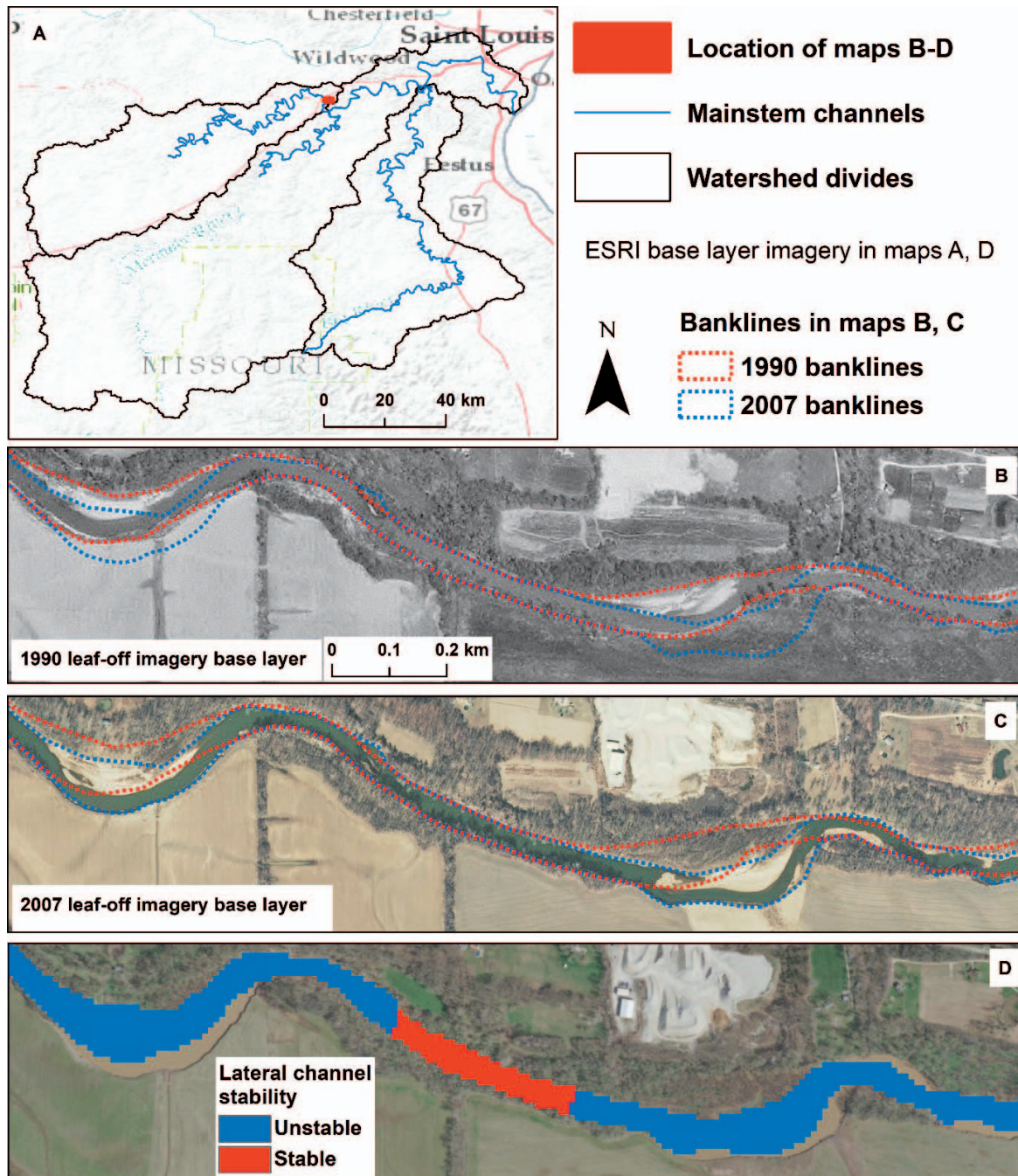


Figure S5. Example maps of lateral channel stability layer: A) map of the Meramec River watershed with locations of the subsequent detailed maps, B) digitized bank lines from 1990 and 2007 superimposed over 1990 leaf-off aerial imagery, C) digitized bank lines from 1990 and 2007 superimposed over 2007 leaf-off imagery, and D) lateral channel stability binary layer from the digitized bank lines.