Detecting Sheer Rock Outcrops Using Digital Elevation Models for Ecological Conservation

Caleb A. Aldridge, Arek Moczulski, Katy Tramel Wilson, and Evan C. Boone

***Abstract*—Rock outcrops serve as unique habitats for several specialist species, yet their identification and conservation are challenged due to logistical and computational constraints. This study aimed to improve the accuracy and efficiency of locating sheer rock outcrops (>60 cm in height) using high- and medium-resolution digital elevation models (DEMs) and conditional inference trees, a recursive binary partitioning approach to predict an outcome. In our application, the outcome is the presence of sheer rock outcrops. We used slope and topographic position index (TPI) as predictors. The conditional inference tree models estimated slope thresholds of 38.4° for high-resolution DEM and 18.14° for medium-resolution DEM as indicative of sheer rock outcrop presence. These findings align with a previous study, adding a layer of validity. In addition, a TPI >0.137 in the high-resolution DEM provided additional accuracy of sheer rock outcrop presence and the medium-resolution (∼10 m) DEM can be used when high-resolution (1 m) DEM is not available. This is particularly relevant for large-scale conservation efforts where computational resources may be limited. However, high-resolution DEM is recommended for more precise quantification of the total area of sheer rock outcrops. Our study has significant implications for efficiently locating sheer rock outcrops and identifying putative areas for protection to help conserve species that rely on these unique habitats while minimizing the logistical constraints of previous methods. The modular methodology developed here could serve as an invaluable approach for broader rock outcrop conservation. Furthermore, our findings add momentum to further innovate remote sensing of unique habitats.**

***Index Terms*—Conservation biology, habitat detection, remote sensing, spatial analysis.**

I. Introduction

ROCK outcrops are unique, localized landscape features where bedrock or consolidated sediment is exposed on Earth's surface [1], [2], [3]. Rock outcrops provide refuge from harsh climatic conditions and predators for several specialist plants and animals [2], [4]. Habitat specialization, however, often leads to range restrictions and reduced genetic diversity, increasing the risk of extirpation, further exacerbated by human disruptions such as rock climbing, timber harvest, mining, and climate change [2], [5], [6].

In Mississippi, rock outcrops are scarce as most of the state's surface geology is unconsolidated sediments [7]. Tishomingo County, in the far northeast corner of the state, is an area where rock outcrops are concentrated, especially sheer (vertical faced) rock outcrops [7], [8]. Several rock outcrop specialists occur in Tishomingo County and have been listed as species of greatest conservation need [8], including state-endangered or federally petitioned salamanders and mammals like the Green Salamander (*Aneides aeneus*), Cave Salamander (*Eurycea lucifuga*), Spring Salamander (*Gyrinophilus porphyriticus*), and Eastern Spotted Skunk (*Spilogale putorius*). It is likely the Green and Cave Salamander would be unable to persist during the hot and dry summer months (July–October) if not for cool, moist refuge in rock outcrops [9], [10], [11]. Although there is currently no legal protection for rock outcrops in Mississippi, efforts to locate them are essential for habitat and species conservation to progress, whether regulatory or voluntary [2], [8].

Traditional methods of locating rock outcrops, such as topographic surveys or visual inspection of aerial orthoimagery, can be time-consuming and ineffective [2], [12], [13], [14]. For example, dense canopy cover precludes the detection of rock outcrops in aerial orthoimagery and physical or legal restrictions prevent access to areas during topographic surveys. However, modern technologies and products, like light detection and ranging (LiDAR) and digital elevation models (DEMs), offer opportunities for efficient and cost-effective surveys of the Earth's surface. For example, Smith and Mullins [14] used freely available, high-resolution (0.75-m spatial, 1-cm vertical) LiDAR data to construct a DEM and locate putative sheer rock outcrops in the DEM by filtering for areas of high slope. As more quality-assured, high-resolution LiDAR data and DEMs become freely available, analyses and applications in sheer rock outcrop detection can be improved and streamlined.

The application by Smith and Mullins [14] is an important step towards inventorying and conserving rock outcrops [2]. Smith and Mullins [14] used a slope threshold ≥80%, or ≥38.7° = $\tan^{-1}$(80% / 100%), to detect putative sheer rock outcrops in Jefferson National Forest, Virginia. This threshold was based on measurements and expert opinion from earlier studies [11], [14], [15] and yielded a 100% accuracy during field validation. However, there is potential that a ≥38.7° threshold could omit putative sheer rock outcrops and, therefore, opportunities for habitat conservation go unrealized. Estimating a slope threshold

Manuscript received 22 November 2023; revised 4 January 2024 and 28 January 2024; accepted 22 February 2024. Date of publication 29 February 2024; date of current version 20 March 2024. This work was supported by the U.S. Fish and Wildlife Service. *(Corresponding author: Caleb A. Aldridge.)*

Caleb A. Aldridge and Evan C. Boone are with the U.S. Fish and Wildlife Service, Lower Mississippi River Fish and Wildlife Conservation Office, Tupelo, MS 38804 USA (e-mail: caleb_aldridge@fws.gov; evan_boone@fws.gov).

Arek Moczulski is a Conservation Planner with the Pennsylvania Association of Conservation Districts, Harrisburg, PA 17112 USA.

Katy Tramel Wilson is a Fisheries Specialist with the Mississippi Department of Wildlife, Fisheries, and Parks, Jackson, MS 39211 USA.

This article has supplementary downloadable material available at https://doi.org/10.1109/JSTARS.2024.3371879, provided by the authors.

Digital Object Identifier 10.1109/JSTARS.2024.3371879

through field validation at sites with a wide range of slope values would help to improve this threshold, either by supporting the threshold used by Smith and Mullins [14] or by providing an estimate less prone to type II error (omission error).

Though high-resolution remote sensing data is becoming more available and less expensive to collect, there is still a dearth of spatial coverage when compared to low- and medium-resolution data. In areas with both high- and medium-resolution remotely sensed data, DEMs can be used to identify landscape features of interest at both resolutions and comparisons made for applicability and efficiency. Furthermore, models using medium-resolution data can be better calibrated when high-resolution data is also available. From an eco-environmental perspective, this allows conservationists to locate putative landscape features (i.e., habitat) for surveillance and protection in areas lacking high-resolution data and overcome computing limitations when working with high-resolution data. However, models using medium-resolution data are presumed to be less accurate and may increase the chances of type I errors (commission error), leading to wasted time and effort when visiting sites.

In this study, we aim to advance the methods of Smith and Mullins [14] by empirically estimating a slope threshold indicative of sheer rock outcrops, utilizing a combination of remotely sensed data and field verification. While the approach of Smith and Mullins [14] effectively highlighted the contrast between slopes of sheer rock outcrops and adjacent gently sloping hillsides, it did not systematically compare sites that are completely devoid of sheer rock outcrops. Our study seeks to fill this gap by extending counterfactual analyses to include a broader range of slope values, encompassing both sites with and without sheer rock outcrops. This method not only aids in validating a slope threshold for identifying sheer rock outcrops but also enhances the robustness of the threshold estimation by integrating a diverse dataset. This inclusive and holistic strategy refines the precision of sheer rock outcrop detection and contributes to reinforcing the foundational results of Smith and Mullins [14].

We used freely available, high-resolution (1 m) DEM tiles to analyze the slope within Tishomingo State Park, Mississippi. We collected samples at sites within Tishomingo State Park across a wide range of slope values. Field sample data was used to estimate the slope threshold indicative of sheer rock outcrops and assess model accuracy for two different models. We then compared the total area of putative sheer rock outcrops in Tishomingo State Park using the 38.7° slope threshold of Smith and Mullins [14] and the estimate from the best-performing model in this study. We also modeled the slope threshold using medium-resolution (∼10 m) DEM tiles and assessed model accuracies, comparing results with those derived from high-resolution DEM tiles. We discuss the differences and improvements made through our study and outline areas for further investigation and improvement. Developing remote sensing methods that leverage freely available data to inventory rock outcrops increases the ability of conservationists to monitor species of greatest conservation need and help protect their critical habitat while reducing cost and capacity requirements.

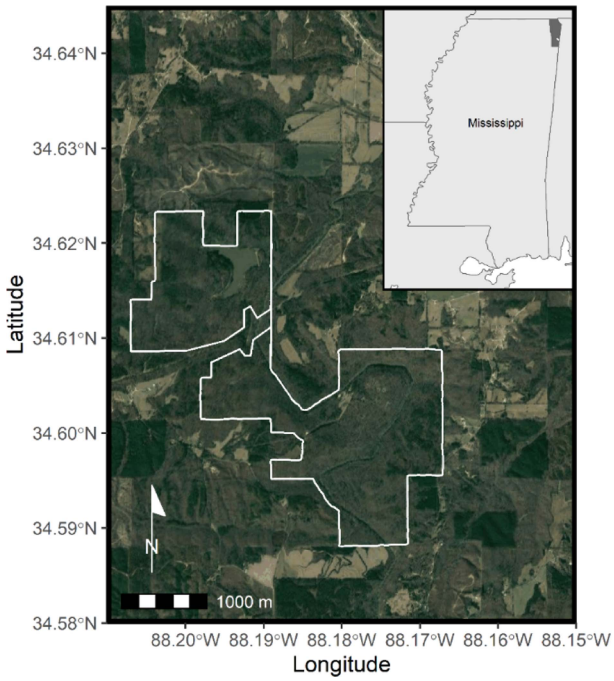

Fig. 1. Map of study area, Tishomingo State Park, in the southeast portion of Tishomingo county, Mississippi (white point within the grey polygon of inset). The white line indicates the park boundary.

II. Study Area

Tishomingo State Park is situated on the periphery of the Cumberland Plateau in northeast Mississippi (see Fig. 1). The park covers an area of approximately 600 hectares (1500 acres) and is intersected by Bear Creek, a tributary of the Tennessee River to the north, and by the National Park Service's Natchez Trace Parkway [16], [17]. Within the park bounds, rock outcrops and talus of Hartselle Sandstone and Bangor Limestone (Paleozoic Era) exhibit up to 60 m (200 feet) of relief, shaped by Bear Creek and its associated tributaries [16], [18].

Miller (1983) reported an average summer temperature of 25.6 °C (78 °F) with an average daily maximum of 32.2 °C (90 °F) and an average winter temperature of 5.6 °C (42 °F) with an average daily minimum of −0.6 °C (31 °F) for Tishomingo County. The total annual precipitation was reported as 132 cm (52 in.) with 45% (59.4 cm; 23.4 in.) falling between April and September. The majority of the vegetation in Tishomingo County and Tishomingo State Park consists of oak-hickory forest, with hydrophytic species (e.g., Bald Cypress [*Taxodium distichum*]) found in river bottoms and xerophytic species (e.g., Shortleaf Pine [*Pinus echinata*]) on drier exposures [16], [18].

III. Materials and Methods

The maps, analyses, and plots herein were produced in program R [19] using, primarily, the terra [20], caret [21], and partykit [22] packages for analyses, and the ggplot2 [23], ggmap [24], ggspatial [25], and ggparty [26] packages for maps and

plots. All of the scripts used can be accessed in the Supplemental Material.[1]

A. DEM Download and Processing

We downloaded high- (1 m) and medium-resolution (1/3 arc-sec; ∼10 m) DEM tiles (GeoTIFF) for Tishomingo County, including Tishomingo State Park, from The National Map Download Client[2] maintained by the U.S. Geological Survey. First, we selected elevation products (3DEP) under the data heading. We then selected the 1-m DEM box (or 1/3 arc-sec DEM box) and clicked Show button underneath. This allowed us to navigate to the area of interest in the interactive map pane. We then selected Selectable Polygon from the dropdown menu next to the Area of Interest heading and selected County or Equivalent from the popup dropdown. We navigated to and clicked the polygon corresponding to Tishomingo County, Mississippi on the map and then clicked the Search Products button in the side panel. This returned DEM tiles in the area of interest. The DEM tiles were downloaded by clicking the add to cart icon, clicking the Cart tab at the top left, and then following the uGet download manager instructions provided on the Cart tab page.[3] We used the terra package [20] in R to process and analyze DEM tiles. For the high-resolution DEM, we began by reading the tiles (GeoTIFF) into the R environment and then stitched the tiles together using the mosaic function. The medium-resolution DEM tile did not require stitching.

B. DEM Analysis

We used the terrain function from the terra package to calculate slope, in degrees, and topographic position index (TPI). Slope was calculated according to Horn [27], where cell slope value is determined through differentiation of elevation from the surrounding eight cells. TPI was calculated according to Wilson et al. [28], where cell TPI value is determined as the difference in a cell's elevation and the mean elevation of the surrounding eight cells. Outputs were stored as raster layers for further analysis. We included TPI in our analyses as an alternate or additive predictor but assumed slope was more critical for inclusion based on previous research [2], [3], [7], [11], [12], [14], [15]. We used Pearson's correlation coefficient (r) to measure the strength of association between slope and TPI derived from the high- and medium-resolution DEMs.

C. Field Sampling

To estimate the slope threshold indicative of sheer rock outcrops, we visited sites across a wide range of slope values and documented whether sheer rock outcrops were present or not. We focused on sites within Tishomingo State Park, cropping the DEMs to the park boundary. We then aggregated slope and TPI in the high-resolution raster layers from the 1 to 10 m scale, taking the maximum value within the 10 m × 10 m area. We did this to better match the scale of field sites and the accuracy of GPS receiver (see below) [29]. We then stratified sites by every five degrees of slope. This resulted in a total of 16 stratification groups. We then used spatially randomly sampled 10 sites in each stratum (i.e., spatial stratified sampling; spatSample function from the terra package), except for the largest slope group (>75°), which only contained two sites. This resulted in 152 potential sample sites.

Fig. 2. Examples of sample sites. Example A (top-left) is a sample site with an obvious sheer rock outcrop (>>60 cm); B (top-right) is a sample site without a sheer rock outcrop; C (bottom-left) is a sample site with a small sheer rock outcrop (>60 cm); and D (bottom-right) is a sample site with talus <60 cm in height.

Potential sample sites were uploaded to ArcGIS Field Maps[4] for use on iPhone SE Gen. 2 and iPad Gen. 7[5] (∼5 m location accuracies). We were able to visit a total of 126 sample sites between 13 January and 30 March 2023. When visiting a sample site, we recorded geolocation, a photograph of the site, and the presence or absence of sheer rock outcrop or talus >60 cm in height (see Fig. 2) within 15 m. For clarification, talus is an exposed piece of rock that has separated from the exposed bedrock. We chose a height of 60 cm based on personal observations (CAA; W. Selman, Associate Professor, Millsaps College, personal communication [virtual call], December 14, 2022) of rock outcrop specialist salamanders occurring on sheer rock outcrops of this height or higher and previous surveys of rock outcrop specialists in Tishomingo State Park [16], [17]. The radius of 15 m was based on recommendations from McCoy [29], modifying the minimum sample area to a radius r (rounded up) substituting a and b in the Pythagorean theorem with $0.5A$, $r = \sqrt{2(0.5A)^2}$, where $A = P(1+2L)$ and P is the pixel dimension (10 m as defined above) and L is the locational accuracy in fractions of pixels ($0.5 = 5\,\mathrm{m}/10\,\mathrm{m}$). Six sites were excluded because of adjacency to roads or streams, which registered as high-slope locations indicative of sheer rock outcrops, but these are false positives. The remaining 26 potential sample sites were not visited due to accessibility and resource constraints.

[1] [Online]. Available at: https://ecos.fws.gov/ServCat/Reference/Profile/155771

[2] [Online]. Available at: https://apps.nationalmap.gov/downloader/

[3] [Online]. Available at: https://apps.nationalmap.gov/uget-instructions/index.html

[4] [Online]. Available at: www.esri.com

[5] [Online]. Available at: www.apple.com

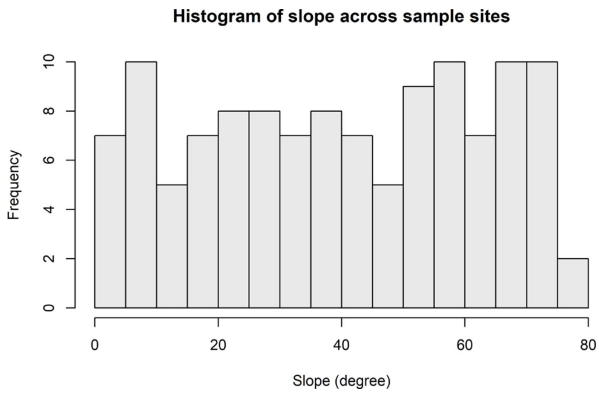

Fig. 3. Frequency distribution of slope values for the 120 sample sites visited on 13 January and 30 March, 2023, at Tishomingo State Park, Mississippi.

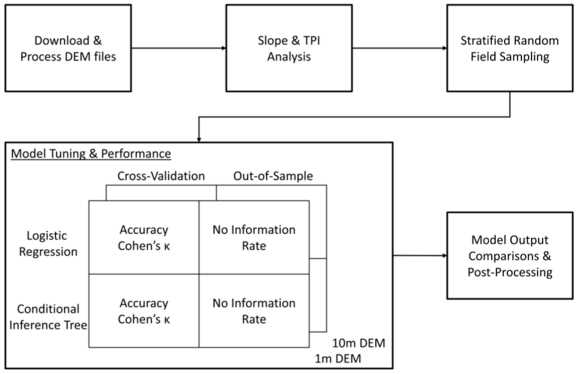

Fig. 4. Schematic of data retrieval, field sampling, and analyses.

Slope values at the 120 sites used in subsequent analysis were uniformly distributed (see Fig. 3).

D. Data Preparation and Tuning Metrics

We divided slope and TPI data derived from the high- and medium-resolutions DEMs into cross-validation and out-of-sample datasets of 100 and 20 samples—a total of four datasets (see Fig. 4). Data were partitioned to ensure uniformity and comparability between cross-validation and out-of-sample datasets. The models used in this study were tuned and compared following the same procedures for datasets derived from both the high- and medium-resolution DEMs.

The cross-validation dataset was used to determine slope and TPI parameter estimates, which maximized model accuracy. We accomplished this using repeated ten-fold cross-validation, with ten repetitions. Accuracy was defined as

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

where correct predictions include true positives and true negatives, and the total number of predictions include the former plus false positives and false negatives, with positives being the presence of a sheer rock outcrop. The most accurate parameter set was selected for further analysis. Although not used for tuning, we recorded Cohen's kappa coefficient (κ) in each iteration to characterize the relative accuracy of models. Cohen's κ reports the difference in model accuracy and the accuracy of a model with random class assignment (i.e., sheer rock outcrop present or not), proportional to class size, divided by one minus the accuracy of random class assignment [30].

E. Model Descriptions

We tuned and compared two different models. First was logistic regression, a parametric method, and the second was a nonparametric method, conditional inference trees. Logistic regression models binary outcomes (i.e., sheer rock outcrop present or not) given a linear combination of predictors using the logit link function, which transforms log odds outcomes to the probability scale [31]. Conditional inference trees use recursive binary partitioning with permutation tests as the basis for statistical inference of split given a selected α level *(α chosen as 0.05 in this study)* [22], [32], [33]. More specifically, the steps of the conditional inference tree algorithm are as follows.

1) Select the predictor with the strongest association with the outcome.
2) Split the predictor variable into two groups based on accuracy in the outcome of where the split is made. This is done iteratively and uses permutation testing for deciding when to stop.
3) Repeat steps 1 and 2 recursively until no further splits can be made.

We chose these approaches because of their ease of interpretability. However, logistic regression assumes a linear relationship and little to no interaction among predictors. Conditional inference trees, on the other hand, can handle nonlinear relationships and predictor interactions but may overfit data in certain cases [22], [32].

F. Data and Model Comparisons

To compare and contrast sites with and without sheer rock outcrops, we began with simple comparisons of slope and TPI between such sites. We calculated summary statistics (i.e., mean, standard deviation, and median) of slope and TPI for sites with and without sheer rock outcrops derived from both the high- and medium-resolution DEMs. We tested for differences in slope and TPI between sites with and without sheer rock outcrops using *t*-tests.

We used *t*-tests to test for differences in accuracy and Cohen's κ of logistic regression and conditional inference tree models. Tests were restricted to between models that used data derived from the same resolution DEM (i.e., high- or medium-resolution). We also fit models to the out-of-sample datasets and used binomial tests to compare the model accuracy to the "no information rate," or accuracy expected if samples were to be classified according to class proportions.

While our approach involved performing multiple *t*-tests on the same dataset, adjustments for multiple comparisons using the false discovery rate, which controls the proportion of false positives among all rejected tests, did not alter the overall significance of our results [34], [35]. Slope estimates from the most

TABLE I
COUNT (*n*), MEAN, STANDARD DEVIATION (IN PARENTHESES), AND MEDIAN (IN ITALICS) FOR SLOPE AND TPI OF SAMPLE SITES WITH (YES) AND WITHOUT (NO) SHEER ROCK OUTCROPS (ROCK OUTCROPS OR TALUS >60 CM IN HEIGHT) PRESENT

Outcrop present?	*n*	*High-resolution (1 m)*		*Medium-resolution (~10 m)*	
		Slope	**TPI**	**Slope**	**TPI**
Yes	64	57.4° (12.7°) *58.1°*	0.55 (0.07) *0.43*	25.5° (7.9°) *25.4°*	0.31 (1.73) *-0.12*
No	56	19.3° (11.4°) *19.5°*	0.08 (0.43) *0.07*	9.76° (7.24°) *8.87°*	0.10 (0.63) *0.05*

accurate models were interpreted as thresholds indicative of sheer rock outcrops in the high- and medium-resolution DEMs.

G. Comparisons of Outputs and Postprocessing

We compared the total area of putative sheer rock outcrops in Tishomingo State Park using the 38.7° slope threshold of Smith and Mullins [14] and that of our more accurate model that used data derived from the high-resolution DEM. We first created 25-m buffers (buffers function, terra package) around roads and streams to reduce the likelihood of biased estimation from false-positive sites. We then created polygons of locations (crop function, terra package) with the given slope thresholds and calculated the total area of the polygons (expanse function, terra package). We also calculated the total area estimated by our more accurate model using data derived from the medium-resolution DEM. We presumed that the total area estimated using the medium-resolution DEM would be higher and the percent difference would be a measure for assessing an effect of imprecision. We calculated simple statistics of absolute and relative differences between the three area estimates.

IV. RESULTS

Slope derived from the high- and medium-resolution DEMs were highly correlated ($r = 0.86$, CI = [0.80, 0.90]; $t = 18.1$, df = 118, $p = <0.001$) while TPI derived from the high- and medium-resolution DEMs were less strongly correlated ($r = 0.33$, CI = [0.16, 0.48]; $t = 3.82$, df = 118, $p = <0.001$).

A. High-Resolution (1 m) DEM Data

At the 120 sample sites visited, slope ranged from 2.4° to 78.1° with a mean and standard deviation of 39.7° and 22.6° respectively, and a median of 39.3° (see Fig. 5). TPI ranged from <0.01 to 1.76 with a mean and standard deviation of 0.33 and 0.40 respectively, and a median of 0.14 (see Fig. 6). Sheer rock outcrops were present at 64 sites (53.3%; Table I). Both slope ($\bar{x} = 38.1$, CI = [33.7, 42.4]; $t = 17.34$, df = 117.92, $p = <0.001$) and TPI ($\bar{x} = 0.47$, CI = [0.36, 0.58]; $t = 8.53$, df = 64.27, $p = <0.001$) varied between sites with and without sheer rock outcrops present.

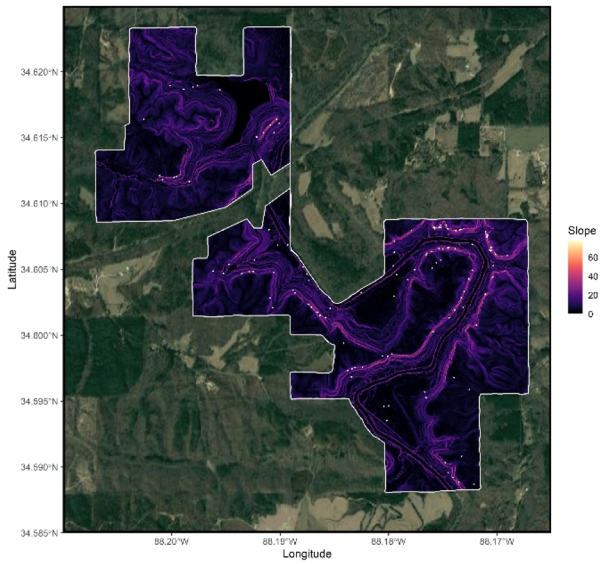

Fig. 5. Slope values within Tishomingo State Park, Mississippi. Slope values were derived from high-resolution (1 m) DEM data (downloaded from apps.nationalmap.gov). Darker colors (black and purple) represent cells with no to little slope (0°–20°) and lighter colors (yellow and orange) represent cells with moderate to high slope (60°+). White points indicate random stratified sample sites.

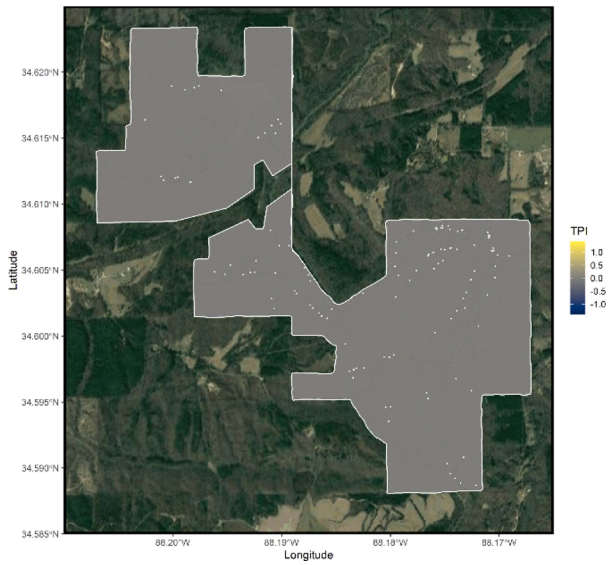

Fig. 6. TPI values across Tishomingo State Park. Values were calculated using eight neighbors at 1-m resolution.

The accuracy of the logistic regression using the cross-validation dataset ranged from 0.67 to 1.0, with a mean and standard deviation of 0.94 and 0.07, and a median of 1.0. The accuracy of the conditional inference tree using the cross-validation dataset ranged from 0.70 to 1.0, with a mean and standard deviation of 0.96 and 0.07, and a median of 1.0. Using the cross-validation dataset, Cohen's κ ranged from 0.34 to 1.0, with a mean and standard deviation of 0.88 and 0.15, and a median of 1.0 for the logistic regression; Cohen's κ ranged from 0.40 to 1.0, with a mean and standard deviation of 0.92

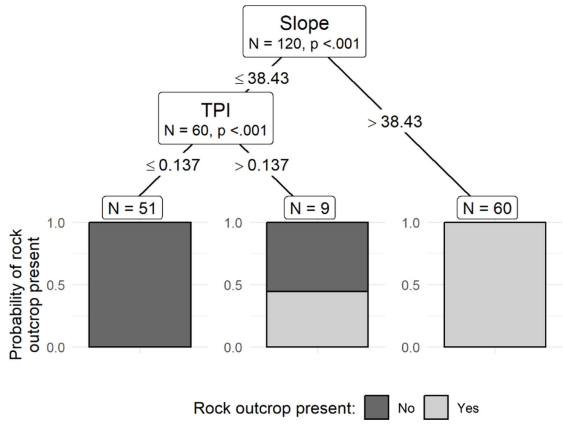

Fig. 7. Conditional inference tree model estimates and outputs for the high-resolution (1 m) DEM data (downloaded from apps.nationalmap.gov).

and 0.14, and a median of 1.0 for the conditional inference tree. Accuracy ($\bar{x} = 0.02$, CI = [−0.04, <0.01]; t = -1.75, df = 196.9, $p = 0.08$) and Cohen's κ ($\bar{x} = 0.03$, CI = [−0.07, <0.01]; $t = -1.75$, df = 196.9, $p = 0.08$) did not differ statistically between the logistic regression and conditional inference tree using the cross-validation dataset.

The conditional inference tree had a 0.95 accuracy and a 0.9 Cohen's κ for the out-of-sample dataset, while the logistic regression had a 0.9 accuracy and a 0.8 Cohen's κ. Using the out-of-sample dataset, the "no information rate" was 0.55 and binomial tests indicated that both models were statistically more accurate than this threshold—conditional inference tree accuracy (0.95, CI = [0.75, 0.99]; $x = 20$, $n = 20$, $p = <0.001$) and logistic regression (0.9, CI = [0.68, 0.98]; $x = 18$, $n = 20$, $p = <0.001$). We chose to fit the final model using the conditional inference tree because it had higher accuracy (0.95 versus 0.90) when fit to the out-of-sample dataset, had similar accuracy (0.96 versus 0.94) as logistic regression on the cross-validation dataset, is easy to interpret, perhaps more so than logistic regression, and is robust to nonlinearity and predictor interactions.

Fitting a Conditional Inference Tree the full dataset (cross-validation plus out-of-sample datasets together) had an accuracy of 0.97 and estimated a slope threshold of 38.4° where slopes >38.4° indicated that rock outcrops were present (see Fig. 7). If slope was ≤38.4° but had a TPI >0.14, then there was a 44% chance a sheer rock outcrop was present (66% chance not present). Otherwise, sheer rock outcrops were predicted to not be present (slope ≤38.4° and TPI ≤0.14).

We estimated the total area of sheer rock outcrops in the high-resolution DEM at a 38.4° threshold to be 24763.9 m^2 (266,556.4 ft^2). Comparing this to the total area estimated using the 38.7° threshold, 23989.88 m^2 (258,224.9 ft^2); our threshold estimated a 774.0 m^2 (8,331.27 ft^2) larger total area of sheer rock outcrops, a relative difference of 3.3% (see Fig. 8).

B. Medium-Resolution (∼10 m) DEM Data

At the 120 sample sites visited, slope calculated from the medium-resolution DEM ranged from 0.90° to 40.47° with a mean and standard deviation of 18.14° and 10.92° respectively, and a median of 19.43° (see Fig. 9). TPI ranged from -3.47 to 6.0 with a mean and standard deviation of 0.22 and 1.33, respectively, and a median of <0.01 (see Fig. 10). Slope varied between sites with and without sheer rock outcrops present ($\bar{x} = 15.74$, CI = [12.97, 18.45]; $t = 11.6$, df = 117.74, $p = <0.001$) while TPI did not ($\bar{x} = 0.21$, CI = [−0.25, 0.67]; $t = 0.91$, df = 81.62, $p = 0.36$).

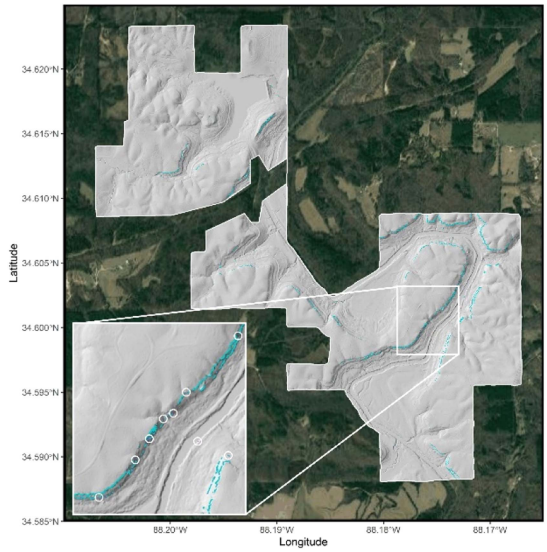

Fig. 8. Polygons of areas with slopes ≥38.4° (cyan + magenta) and ≥37.8° (cyan only) using the high-resolution DEM data (downloaded from apps.nationalmap.gov). Differnces between the two thresholds are minor; white circles indicate loactions where there are differences in area between the two thresholds.

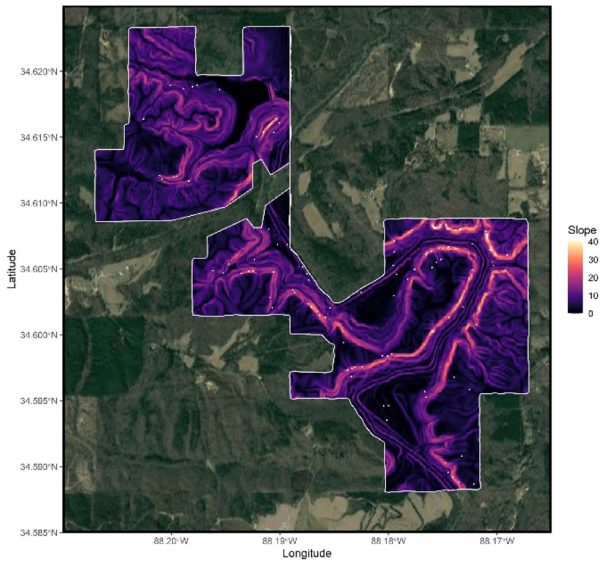

Fig. 9. Slope values within Tishomingo State Park, Mississippi. Slope values were derived from medium-resolution (∼10 m) DEM data (downloaded from apps.nationalmap.gov). Darker colors (black and purple) represent cells with no to little slope (0°–10°) and lighter colors (yellow and orange) represent cells with moderate to high slope (30°+). White points indicate random stratified sample sites.

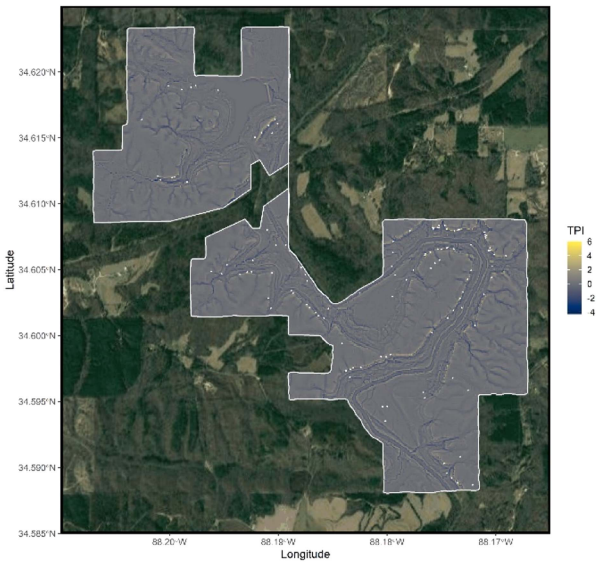

Fig. 10. TPI values across Tishomingo State Park. Values were calculated using eight neighbors at ∼10-m resolution.

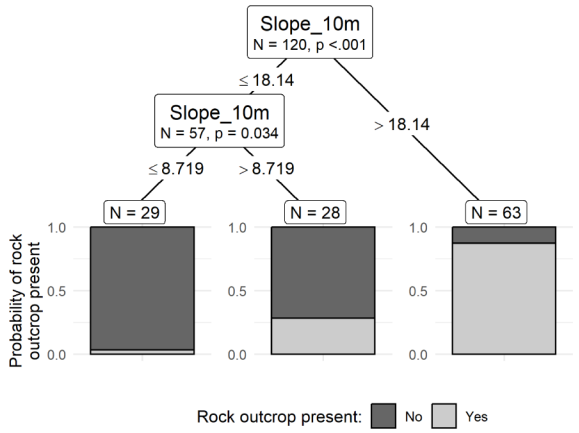

Fig. 11. Conditional inference tree model estimates and outputs for the medium-resolution (∼10 m) DEM data (downloaded from apps.nationalmap.gov).

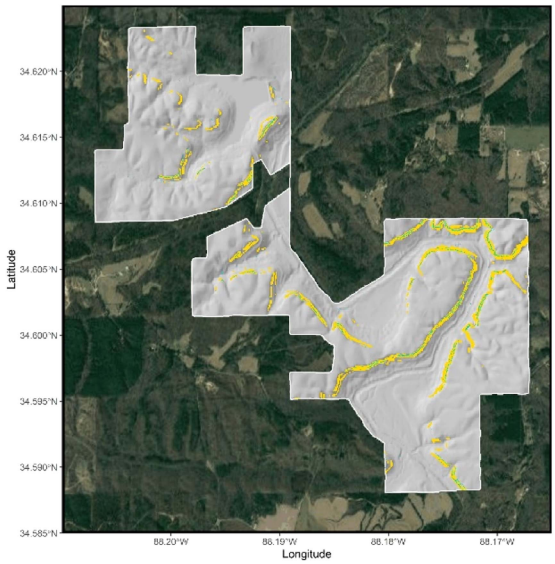

Fig. 12. Polygons of areas with slopes ≥18.14° (yellow) using the medium-resolution DEM data (downloaded from apps.nationalmap.gov). Polygons of areas with slopes ≥38.4° (cyan + magenta) and ≥37.8° (cyan only) using the high-resolution DEM data (downloaded from apps.nationalmap.gov) are plotted in magenta and cyan respectively for comparison.

Accuracy of the logistic regression using the cross-validation dataset ranged from 0.6 to 1.0, with a mean and standard deviation of 0.84 and 0.10, and a median of 0.82. The accuracy of the conditional inference tree using the cross-validation dataset ranged from 0.6 to 1.0, with a mean and standard deviation of 0.84 and 0.10, and a median of 0.82. Using the cross-validation dataset, Cohen's κ ranged from 0.2 to 1.0, with a mean and standard deviation of 0.69 and 0.19, and a median of 0.65 for the logistic regression; Cohen's κ ranged from 0.2 to 1.0, with a mean and standard deviation of 0.69 and 0.20, and a median of 0.65 for the conditional inference tree. Accuracy ($\bar{x} = <0.01$, CI $= [-0.03, 0.03]$; $t = 0.01$, df $= 197.9$, $p = 0.99$) and Cohen's κ ($\bar{x} = <0.01$, CI $= [-0.05, 0.05]$; $t = 0.01$, df $= 197.9$, $p = 0.99$) did not differ statistically between the logistic regression and conditional inference tree using the cross-validation dataset.

The logistic regression and conditional inference tree both had an accuracy of 0.8 and a Cohen's κ of 0.6 for the out-of-sample dataset. Using the out-of-sample dataset, the "no information rate" was 0.55 and binomial tests indicated that both models were statistically more accurate than this threshold—(0.8, CI $= [0.56, 0.94]$; $x = 16$, $n = 20$, $p = 0.02$). We chose to fit the final model using the conditional inference tree because it had similar cross-validation and out-of-sample dataset accuracy as logistic regression but is perhaps more easily interpreted as logistic regression and is robust to nonlinearity and predictor interactions.

The conditional inference tree fit to the original dataset (cross-validation and out-of-sample datasets together) estimated a slope threshold of 18.14° where slopes >18.14° indicated there was an 87.3% chance sheer rock outcrops were present (see Fig. 11). If slope was ≤18.14° but >8.72°, then there was a 28.6% chance a sheer rock outcrop was present. Otherwise, there was only a 3.4% chance sheer rock outcrops present (slope ≤8.72°). Model tuning nor fitting to the original dataset did not indicate that TPI was an important predictor of sheer rock outcrop presence.

We estimated the total area of sheer rock outcrops in the medium-resolution DEM at an 18.14° threshold to be 296291.9 m^2 (3,189,259.47 ft^2). Comparing this to the total area estimates using the high-resolution DEM—38.7° resulting in 23989.88 m^2 (258,224.9 ft^2) and 38.4° resulting in 24763.9 m^2 (266,556.4 ft^2)—the total area of the medium-resolution DEM was estimated 272 302 m^2 (2,931,034.33 ft^2) and 271 528 m^2 (2,922,703.07 ft^2) larger, respectively (see Fig. 12). This resulted in relative differences of 1135% and 1096.5%, respectively.

V. Discussion

In this study, we aimed to improve methods for locating sheer rock outcrops using remotely sensed data. The conditional inference tree model emerged as the most accurate and

robust model in detecting sheer rock outcrops using data derived from both high- and medium-resolution DEMs, estimating slope thresholds of 38.4° and 18.14° respectively. These results may be specific to our study area and remain to be tested in other regions. However, they do correspond to the results of Smith and Mullins [14], thereby adding a layer of external validity to our study. The importance of rock outcrops as unique habitats has been emphasized in previous studies [2], [3], making the need for accurate methods to detect rock outcrops ever more critical.

The utility of using medium-resolution (∼10 m) DEM tiles for approximating sheer rock outcrop locations when high-resolution DEM tiles are not available is a significant finding. This is particularly relevant for large-scale analyses, which may pose constraints on computational resources [29]. However, for more precise quantification of the area of sheer rock outcrops, high-resolution DEM tiles are indispensable, as highlighted by Belt and Paxton [13] and Fraser et al. [12]. The tradeoff between resolution and computational efficiency is a key consideration for future studies [36].

Though we expect the accuracy of our slope thresholds based on high- and medium-resolution DEMs to be lower outside the study area, our approach should provide the opportunity for further data collection and model tuning for other localities. Even so, we feel an error of commission in this case is less severe than that of omission. Stated another way, it is better to include more putative sites for conservation initially than completely miss opportunities to protect unique habitats and associated specialist species, especially when imperiled like in Mississippi.

While the slope is a significant predictor of sheer rock outcrops in this study, it is worth noting that other variables like geological formation, elevation, and spatial autocorrelation were not considered. These variables could potentially improve the model's predictive power and should be the focus of future research [12], [28]. In addition, combined with more field samples, future studies may be able to distinguish among different types of rocky habitats (e.g., large sheer rock outcrops versus small talus; Fig. 2). Furthermore, additional modeling techniques such as density-based clustering (DBSCAN) could offer new perspectives and information, especially for identifying high densities of high-slope cells and discretizing sheer rock outcrop units [37], [38].

The methodology in this study has significant implications for conservation efforts. Efficiently locating rock outcrops is crucial for monitoring and protecting species that rely on these unique habitats [5], [11], [15]. In fact, our approach identified sheer rock outcrops known to host specialist species [8], [16], [17]. The methods developed, once validated in other regions, could serve as invaluable tools in aiding broader conservation initiatives [2], [3]. Furthermore, when used in decision-analytic approaches (e.g., structured decision making) can prioritize conservation units based on multiple objectives and budgetary constraints [39], [40], [41]. This would facilitate a more systematic and effective approach to protect these unique habitats and reliant species [41].

In conclusion, our study successfully developed a robust and accurate model for locating sheer rock outcrops, which are critical habitats for several endangered species [5], [15], [17]. The methodology and findings of this study contribute to the existing body of literature but also provide a valuable resource for future conservation efforts [14], [29]. By setting a precedent for the use of high- and medium-resolution remotely sensed data in ecological research, we open the door for further innovations in the identification and designation of unique and critical habitats.

Acknowledgment

The authors acknowledge the collegial and technical support of A. Rodgers, W. Selman, S. Rush, A. Erves, attendees of the 2023 Mississippi Amphibian and Reptile Conservation Meeting, and Reviewers and Editors of JSTARS who all provided helpful comments. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Data available from public domain resources (https://apps.nationalmap.gov/downloader/). Steps for accessing digital elevation model data used in this study are included in the Materials and Methods.

References

[1] R. L. Bates and J. A. Jackson, *Dictionary of Geological Terms*. Garden City, NY, USA: Anchor Press, 1984.

[2] J. A. Fitzsimons and D. R. Michael, "Rocky outcrops: A hard road in the conservation of critical habitats," *Biol. Conservation*, vol. 211, pp. 36–44, Jul. 2017, doi: 10.1016/j.biocon.2016.11.019.

[3] M. L. Hunter et al., "Conserving small natural features with large ecological roles: A synthetic overview," *Biol. Conservation*, vol. 211, pp. 88–95, Jul. 2017, doi: 10.1016/j.biocon.2016.12.020.

[4] L. P. Shoo, C. Storlie, Y. M. Williams, and S. E. Williams, "Potential for mountaintop boulder fields to buffer species against extreme heat stress under climate change," *Int. J. Biometeorology*, vol. 54, pp. 475–478, 2010.

[5] A. H. Wynn, R. Highton, and J. F. Jacobs, "A new species of rock-crevice dwelling plethodon from Pigeon Mountain, Georgia," *Herpetologica*, vol. 44, no. 2, pp. 135–143, 1988.

[6] K. G. Ferris, J. P. Sexton, and J. H. Willis, "Speciation on a local geographic scale: The evolution of a rare rock outcrop specialist in Mimulus," *Philos. Trans. Roy. Soc. B, Biol. Sci.*, vol. 369, no. 1648, Aug. 2014, Art. no. 20140001, doi: 10.1098/rstb.2014.0001.

[7] D. T. Dockery and D. E. Thompson, *The Geology of Mississippi*, vol. 692. Jackson, MS, USA: Univ. Press Mississippi Jackson, 2016.

[8] "Mississippi museum of natural science," *Mississippi State Wildlife Action Plan*, Jackson, MS, USA: Mississippi Department of Wildlife, Fisheries, and Parks, Mississippi Museum of Natural Science, 2015. [Online]. Available: https://www.mdwfp.com/media/251788/mississippi_swap_revised_16_september_2016__reduced_.pdf

[9] R. E. Gordon and R. L. Smith, "Notes on the life history of the salamander aneides aeneus," *Copeia*, vol. 1949, no. 3, pp. 173–175, 1949, doi: 10.2307/1438982.

[10] C. R. Rossell Jr., J. Hicks, L. A. Williams, and S. C. Patch, "Attributes of rock crevices selected by green salamanders, aneides aeneus, on the blue ridge escarpment," *Herpetological Rev.*, vol. 40, no. 2, pp. 151–153, 2009.

[11] W. H. Smith, S. L. Slemp, C. D. Stanley, M. N. Blackburn, and J. Wayland, "Rock crevice morphology and forest contexts drive microhabitat preferences in the green salamander (aneides aeneus)," *Can. J. Zoology*, vol. 95, no. 5, pp. 353–358, 2017.

[12] O. L. Fraser, S. W. Bailey, M. J. Ducey, and K. J. McGuire, "Predictive modeling of bedrock outcrops and associated shallow soil in upland glaciated landscapes," *Geoderma*, vol. 376, Oct. 2020, Art. no. 114495, doi: 10.1016/j.geoderma.2020.114495.

[13] K. Belt and S. T. Paxton, "GIS as an aid to visualizing and mapping geology and rock properties in regions of subtle topography," *Geological Soc. Amer. Bull.*, vol. 117, no. 1/2, pp. 149–160, Jan. 2005, doi: 10.1130/B25463.1.

[14] W. H. Smith and C. Z. Mullins, "Light detection and ranging (LiDAR)-assisted detection of rock outcrops in appalachian hardwood forests," *J. Fish Wildlife Manage.*, vol. 13, no. 1, pp. 250–261, Jun. 2022, doi: 10.3996/JFWM-21-043.

[15] M. P. Hinkle, L. A. Gardner, E. White, W. H. Smith, and R. D. VanGundy, "Remnant habitat patches support green salamanders (aneides aeneus) on active and former appalachian surface mines," *Herpetological Conservation Biol.*, vol. 13, no. 3, pp. 634–641, 2018.
[16] D. E. Ferguson, "The herpetofauna of Tishomingo County, Mississippi, with comments on its zoogeographic affinities," *Copeia*, vol. 4, pp. 391–396, 1961.
[17] T. J. Rauch III, J. R. Lee, and A. Lawler, "Green salamander (aneides aeneus) populations in Mississippi: A preliminary update," *J. Mississippi Acad. Sci.*, vol. 61, no. 3, pp. 222–226, 2016.
[18] K. H. Miller, *Soil Survey of Tishomingo County, Mississippi*. Washington D.C., USA: The Service, 1983.
[19] R: A language and environment for statistical computing, "R foundation for statistical computing," 2019. [Online]. Available: https://www.r-project.org/
[20] Terra: Spatial data analysis version 1.0-10, "Comprehensive R archive network," 2021. [Online]. Available: https://CRAN.R-project.org/package=terra
[21] M. Kuhn, "Building predictive models in R using the caret package," *J. Stat. Softw.*, vol. 28, pp. 1–26, Nov. 2008, doi: 10.18637/jss.v028.i05.
[22] T. Hothorn and A. Zeileis, "partykit: A modular toolkit for recursive partytioning in R," *J. Mach. Learn. Res.*, vol. 16, no. 1, pp. 3905–3909, 2015.
[23] H. Wickham, *Ggplot2: Elegant Graphics for Data Analysis*, 2nd ed. Berlin, Germany: Springer-Verlag, 2016, doi: 10.1007/978-3-319-24277-4.
[24] D. Kahle and H. Wickham, "ggmap: Spatial visualization with ggplot2," *R J.*, vol. 5, pp. 144–161, 2013.
[25] D. Dunnington and B. Thorne, "ggspatial: Spatial data framework for ggplot2," *R package version1*, vol. 1, 2020.
[26] M. Borkovec and N. Madin, "ggparty:'ggplot' Visualizations for the 'partykit' package," *R package version*, vol. 1, 2019.
[27] B. K. P. Horn, "Hill shading and the reflectance map," *Proc. IEEE*, vol. 69, no. 1, pp. 14–47, Jan. 1981, doi: 10.1109/PROC.1981.11918.
[28] M. F. J. Wilson, B. O'Connell, C. Brown, J. C. Guinan, and A. J. Grehan, "Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope," *Mar. Geodesy*, vol. 30, no. 1/2, pp. 3–35, May 2007, doi: 10.1080/01490410701295962.
[29] R. M. McCoy, *Field Methods in Remote Sensing*. New York, NY, USA: Guilford Press, 2005.
[30] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960.
[31] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning With Applications in R*. Berlin, Germany: Springer-Verlag, 2013.
[32] T. Hothorn, K. Hornik, and A. Zeileis, "Unbiased recursive partitioning: A conditional inference framework," *J. Comput. Graphical Statist.*, vol. 15, no. 3, pp. 651–674, Sep. 2006, doi: 10.1198/106186006X133933.
[33] M. Krzywinski and N. Altman, "Classification and regression trees," *Nature Methods*, vol. 14, no. 8, Aug. 2017, pp. 758–758, doi: 10.1038/nmeth.4370.
[34] M. Krzywinski and N. Altman, "Comparing samples—Part II," *Nature Methods*, vol. 11, no. 4, pp. 355–356, 2014, doi: 10.1038/nmeth.2900.
[35] M. A. Lindquist and A. Mejia, "Zen and the art of multiple comparisons," *Psychosom. Med.*, vol. 77, no. 2, pp. 114–125, 2015, doi: 10.1097/PSY.0000000000000148.
[36] J. Vaze, J. Teng, and G. Spencer, "Impact of DEM accuracy and resolution on topographic indices," *Environ. Model. Softw.*, vol. 25, no. 10, pp. 1086–1098, Oct. 2010, doi: 10.1016/j.envsoft.2010.03.014.
[37] K. Khan, S. U. Rehman, K. Aziz, S. Fong, and S. Sarasvady, "DBSCAN: Past, present and future," presented at the Fifth Int. Conf. Appl. Digit. Inf. Web Technol. (ICADIWT 2014), IEEE, 2014, pp. 232–238.
[38] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN," *Assoc. Comput. Machinery Trans. Database Syst.*, vol. 42, no. 3, pp. 1–21, 2017.
[39] R. Gregory, L. Failing, M. Harstone, G. Long, T. McDaniels, and D. Ohlson, *Structured Decision Making: A Practical Guide to Environmental Management Choices*. Hoboken, NJ, USA: Wiley, 2012.
[40] J. E. Lyons, "Introduction to resource allocation," in *Structured Decision Making: Case Studies in Natural Resource Management*, M. C. Runge, S. J. Converse, J. E. Lyons, and D. R. Smith, Eds. Baltimore, MD, USA: Johns Hopkins Univ. Press, 2020, pp. 99–107.
[41] C. R. Margules and R. L. Pressey, "Systematic conservation planning," *Nature*, vol. 405, May 2000, Art. no. 6783, doi: 10.1038/35012251.

Caleb A. Aldridge received the B.S. degree in general biology from the University of North Alabama, Florence, AL, USA, in 2012, the M.S. degree in biological sciences from the University of Mississippi, Oxford, MS, USA, in 2017, and the Ph.D. degree in wildlife, fisheries and aquaculture from Mississippi State University, Starkville, MS, USA, in 2022.

He is a Biologist and Coordinator with the U.S. Fish and Wildlife Service, Lower Mississippi River Fish and Wildlife Conservation Office, Tupelo, MS, USA. He has previously held research, academic, and agency positions working largely in fisheries and aquatic systems and with herpetological fauna. His current research interests include incorporating decision-analytic methods and optimization routines into natural resources decision-making and management.

Arek Moczulski received the B.S. degree in environmental science from Dickson College, Carlisle, PA, USA, in 2022.

He previously held a position with the Student Conservation Association in the U.S. Fish and Wildlife Service, Lower Mississippi River Fish and Wildlife Conservation Office, Tupelo, MS, USA. He also worked as an Amphibian Surveyor with the U.S. Geological Survey, Laurel, MD, USA and is now a Conservation Planner with the Pennsylvania Association of Conservation Districts, Harrisburg, PA.

Katy Tramel Wilson received the B.S. degree in wildlife management from Delta State University, Cleveland, MS, USA, in 2022.

She previously held a position with the Student Conservation Association in the U.S. Fish and Wildlife Service, Lower Mississippi River Fish and Wildlife Conservation Office, Tupelo, MS, USA. She is currently employed as a Fisheries Specialist with the Mississippi Department of Wildlife, Fisheries, and Parks, Jackson, MS, USA.

Evan C. Boone received the B.S. degree in biological sciences from Eastern Illinois University, Charleston, IL, USA, in 2014, and the M.S. degree in biological sciences with a research focus in fisheries and parasitology from Eastern Illinois University, Charleston, IL, USA, in 2016.

He is a Biologist with the U.S. Fish and Wildlife Service, Lower Mississippi River Fish and Wildlife Conservation Office, Tupelo, MS, USA. His primary responsibilities include projects related to the spread and control of aquatic invasive species and the rehabilitation of species of greatest conservation concern.