Midwest Fish Habitat Partnership Fish Habitat Modeling Results

Ohio River Basin and Southeast Aquatic Resources Partnership

Model Summaries
1/24/2012

Small Streams Signature Fish Index: Score
Modified Index of Centers of Diversity: Score
Smallmouth Bass (*Micropterus dolomieu)*: Probability of Presence
Redhorse (*Moxostoma spp.*): Probability of Presence
Percent Intolerant Fish: Percent of Individuals
Great Rivers Species: Probability of Presence
Intolerant Mussels: Probability of Presence

Authors
Roy Martin, Aquatic Ecologist/Modeler – Downstream Strategies
Todd Petty, Aquatic Ecologist – West Virginia University
Jason Clingerman, Aquatic Ecologist/ Modeler – Downstream Strategies
Fritz Boettner, Project Manager/GIS – Downstream Strategies
Sally Letsinger, GIS – Geodata Basics
Jackie Strager, GIS – West Virginia University
Anne Hereford, Environmental Scientist – Downstream Strategies
Evan Hansen, Chief Scientist – Downstream Strategies

Downstream Strategies
295 High Street, Suite 3
Morgantown, WV 26505
www.downstreamstrategies.com

Report date:
January 24, 2012

ACKNOWLEDGEMENTS

This project would not have been possible without the vision and guidance provided by Maureen Gallagher. We would like to thank Ball State University (BSU), California University of Pennsylvania (CUPA), Illinois Environmental Protection Agency (ILEPA), Indiana Department of Environmental Management (IDEM), Kentucky Division of Water (KDOW), Marshall University (MU), Maryland Department of Natural Resources (MDDNR), Midwest Biodiversity Institute (MBI), New York state Department of Environmental Conservation (NYSDEC), North Carolina Division of Water Quality (NCDWQ), North Carolina Wildlife Resources Commission (NCWRC), Northern Kentucky Sanitation District #1 (NKSD1), Ohio Environmental Protection Agency (OEPA), Ohio River Valley Water Sanitation Commission (ORSANCO), Pennsylvania Department of Environmental Protection (PADEP), Pennsylvania Fish and Boat Commission (PAFBC), South Carolina Department of Natural Resources (SC DNR), Tennessee Department of Environmental Conservation (TDEC), Tennessee Valley Authority (TVA), Tennessee Wildlife Resources Agency (TWRA), The Nature Conservancy (TNC), United States Army Corps of Engineers (USACE), United States Fish and Wildlife Service (USFWS), Virginia Department of Environmental Quality (VADEQ), West Virginia Department of Environmental Protection (WVDEP), and Western Pennsylvania Conservancy (WPC) for their resources, time, and expertise.

We would like to specifically recognize and thank all those who provided data, feedback, and local expertise. These invaluable contributors include Sue Thompson (CUPA), Jim Stahl (IDEM), Mark Joseph and Roy Smogor (ILEPA), John Brumley and Eric Eisiminger (KDOW), Michael Kashiwagi (MDDNR), David Yow (NCWRC), Matt Wooten and Emily Heppner (NKSD1), Casey Festa (NYSDEC), Dennis Mishne (OEPA), Ryan Argo and Rob Tewes (ORSANCO), Rick Spear (PADEP), Russell Burman and Chris Urban (PAFBC), Mark Scott (SCDNR), Scott Robinson (SARP), Linda Cartwright *(TNDEC)* (TDEC), Charlie Saylor (TVA), Jason Henegar and Frank Fiss (TWRA), Sally Palmer (TNC), Chip Hall and Matthew Granstaff (USACE), Jeromy Applegate, Bob Butler, Will Duncan, John Faustini, Adam Kaeser, Peggy Shute, and Jerry Ziewitz (USFWS), Aimee Budd, Gabriel Darkwah, and Jason Hill, (VADEQ), Eric Chapman (WPC), and Robert Row (WVDEP).

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

ABBREVIATIONS

BRT	boosted regression tree
CASI	cumulative anthropogenic stress index
CNQI	cumulative natural quality index
CV	cross validation
DS	Downstream Strategies
FHP	Fish Habitat Partnership
GIS	geographic information systems
NHD	National Hydrography Dataset
ORB	Ohio River Basin Fish Habitat Partnership
ROC	receiver operating characteristic
SARP	Southeast Aquatic Resources Partnership Restoration Effort
USFWS	United States Fish and Wildlife Service

1. INTRODUCTION

1.1 Project background

Fishery and aquatic scientists often assess habitats to understand the distribution, status, stressors, and relative abundance of aquatic resources. Due to the spatial nature of aquatic habitats and the increasing scope of management concerns, using traditional analytical methods for assessment is often difficult. However, advancements in the geographic information systems (GIS) field and related technologies have enabled scientists and managers to more effectively collate, archive, display, analyze, and model spatial and temporal data. For example, spatially explicit habitat assessment models allow for a more robust interpretation of many terrestrial and aquatic datasets, including physical and biological monitoring data, habitat diversity, watershed characteristics, and socioeconomic parameters.

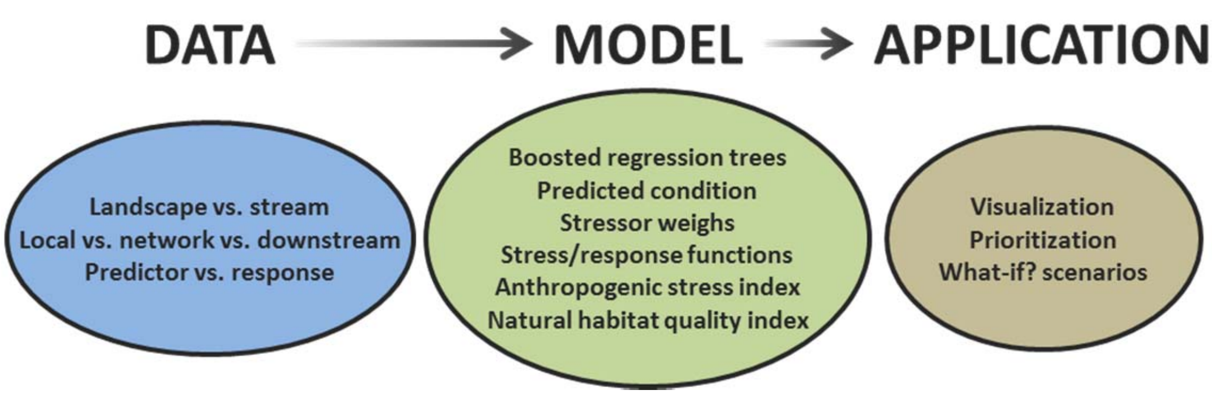
For this project, Downstream Strategies (DS) was contracted by the United States Fish and Wildlife Service (USFWS) to create a spatially explicit data analysis and modeling system for assessing fish habitat condition across the Midwest based on a range of metrics. The data and tools developed as part of the project will be applicable to watersheds, streams, rivers, and lakes within the boundaries of the USFWS's Midwest Fish Habitat Partnership (FHP) and scalable to the national level.

Generally, the models, analyses, and data produced as a result of this project are intended to enable a unique, broad, and spatially explicit understanding of the links between natural habitat conditions, human influences on aquatic habitats, and aquatic health. Specifically, the outcomes will be utilized to conduct fish habitat condition assessments based on a range of stakeholder-specified metrics and modeling endpoints that will help determine the natural drivers of aquatic conditions as well as the major stressors at various spatial scales in specific FHP regions. Additionally, a geospatial decision support tool will be developed to give users the ability to understand habitat conditions and stressors based on the status and severity of threats at specified locations. The ultimate goal is to improve understanding of how local and regional processes influence stream conditions in the region and to provide FHPs with additional knowledge, data, and tools to help prioritize and drive conservation action in the Midwest.

1.2 Overview of the assessment process

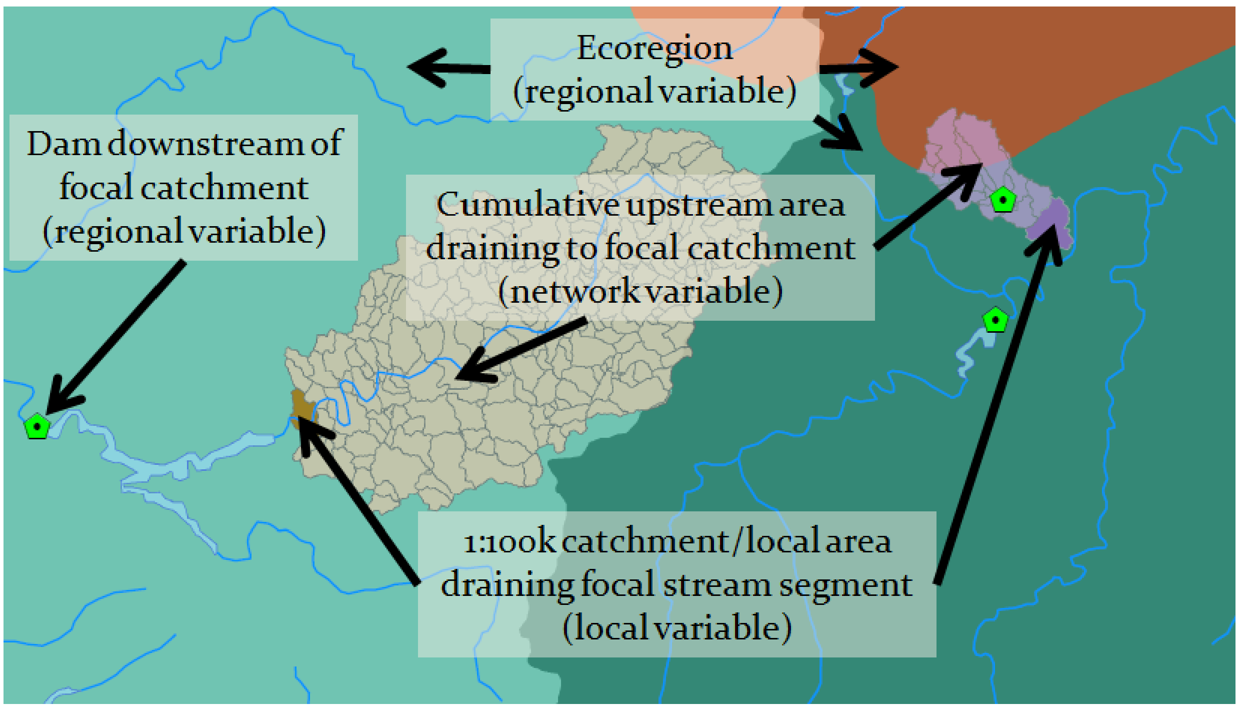
A diagram of the general assessment process is outlined in Figure 1. DS received landscape and aquatic data specified and provided by the individual FHPs to develop models and tools for use in visualizing expected current and potential future conditions and prioritizing management actions.

Figure 1: Diagram of the habitat assessment process



The data provided by FHPs for use in the modeling process can be broken down into two categories: response variables and predictor variables. The response variables are typically instream measures of condition, including biological measures such as species abundance, presence, or richness; or instream physicochemical measures, such as pH, conductivity, or physical habitat scores. Each FHP provided five to seven individual response variables for use in the assessment. Each response variable represents a separate model and assessment and each of those models has a summary report. The predictor variables are typically measures of land use or land cover derived from GIS, such as percent impervious surface area or road crossing density. Although the response variable is always measured at the same local scale (i.e., individual sample site on a stream), the predictor variables are compiled at multiple scales (Figure 2), including the local scale (i.e., single 1:100k National Hydrography Dataset (NHD) stream catchment), the network scale (i.e., all upstream catchments and the local catchment), or the regional scale (e.g., ecoregion).

Figure 2: Diagram and examples of different scales of data used for predictor variables



The process then employs a statistical modeling approach, called boosted regression trees (BRTs), to relate the instream response variable to the landscape-based predictor variables. This process results in a series of quantitative outcomes, including predictions of expected current conditions to all catchments in the FHP (on the scale of the response), measures of the accuracy of those predictions, a quantification of each predictor variable's relative influence on the predictions (i.e., variable importance), and a series of plots illustrating the modeled functional relationship between each predictor and the response (e.g., plot of impervious area vs. presence-absence).

Predictive accuracy is quantified using an internal cross-validation (CV) method (Elith et al., 2008). The method consists of randomly splitting the input dataset into ten equally-sized subsets, developing a BRT model on a single subset and testing its performance on the remaining nine, and then repeating that process for the remaining nine subsets. Thus, the accuracy measures, such as the CV receiver operating characteristic (ROC) score (for presence-absence responses) or the CV correlation coefficient, are actually averages of ten separate ROC or correlation measurements. A standard error for the ten estimates is also given. CV measures are designed to estimate how well the model will perform using independent data.

The outcomes of the statistical modeling component are then used to generate the application portion of the assessment. In this portion, predictions of current conditions and model accuracy can be visualized in a spatially explicit manner using GIS. Likewise, the statistical outcomes are used to generate the post-modeling indices of anthropogenic stress and natural habitat quality. These indices are derived directly from the measures of variable influence and their functional relationships with the response. Specifically, each predictor variable in the statistical model is extracted, along with its importance value and functional plot, to generate an individual metric for use in calculating a cumulative index of stress or natural quality. The individual predictors that are anthropogenic in nature (e.g., impervious surface cover) are used to generate anthropogenic stress metrics and the cumulative anthropogenic stress index (CASI), whereas predictors that are of natural origin (e.g., bedrock geology) are used to generate natural quality metrics and the cumulative natural quality index (CNQI). These metrics and indices are generated at the 1:100k NHD catchment scale and can be mapped in GIS. Once developed, CNQI and CASI are used to generate and visualize restoration and protection priorities. For example, areas of high natural quality (i.e., high CNQI score) and low stress (i.e., low CASI score) could represent protection priorities, whereas areas of high natural quality and high stress may represent restoration priorities.

All of the aforementioned outputs will be integrated into a GIS-based decision support system for use in visualizing current conditions and for use in forecasting and visualizing scenarios based on expected changes in anthropogenic stressors or socioeconomic factors.

1.3 Document outline

This report provides a summary of the key outcomes resulting from models developed by DS for use in assessing aquatic habitats in the Ohio River Basin FHP (ORB) and Southeast Aquatic Resources Partnership Restoration Effort (SARP). The appendices provide additional maps, charts, and metadata useful for evaluating the results of the models.

This document is divided into nine major sections. This section, Section 1, summarizes the project goals, structure, and methodology. Section 2-Section 8 summarize the model input and results for each response variable. Section 9 summarizes some of the limitations to this modeling effort, and outlines suggestions for future similar works.

The following are included for each model's results summary. Subsection one, Modeling inputs, discusses details of the predictor and response variables used in the analyses. Subsection two, Modeling process, covers the basic details and outcomes of the statistical modeling process using boosted regression trees, including information on model certainty. Variable influence and functional relationships between predictor and response variables are included under corresponding headings as well. Subsection three, Post-modeling, contains information resulting from the post-modeling process, including information on the top stressors and natural habitat variables and their relative weights in the calculation of the final indices. Section four, Mapped Results, contains maps for visualizing conditions at the 1:100k catchment scale and includes maps of expected current diversity index condition, stress, and natural quality; it also provides an example of how the two post-modeling indices (i.e., CNQI and CASI) can be combined to arrive at restoration priorities and how those priorities can be visualized in a spatially explicit manner.

2. SMALL STREAMS SIGNATURE FISH INDEX

2.1 Modeling inputs

DS used a list of predictor variables selected by ORB and SARP to develop a ten-fold CV BRT model for the small streams signature fish index ("stream index") at the 1:100k catchment scale. The model was used to produce maps of expected current stream index scores and of expected current natural habitat quality and anthropogenic stress at the 1:100k scale throughout the extents of both FHPs.

DS cooperated with ORB and SARP to arrive at a list of landscape-based habitat variables used to predict stream index throughout the region; ultimately, those variables were also used for characterizing habitat quality and anthropogenic stress. From an initial suite of 372 catchment attributes, DS and the FHPs compiled a list of 92 predictors for evaluation. From that list, 48 variables were removed due to statistical redundancy ($r > 0.6$) or logical redundancy, resulting in a final list of 44 predictor variables for the BRT model and assessment. See Appendix A for a full data dictionary and the metadata document for variable processing notes.

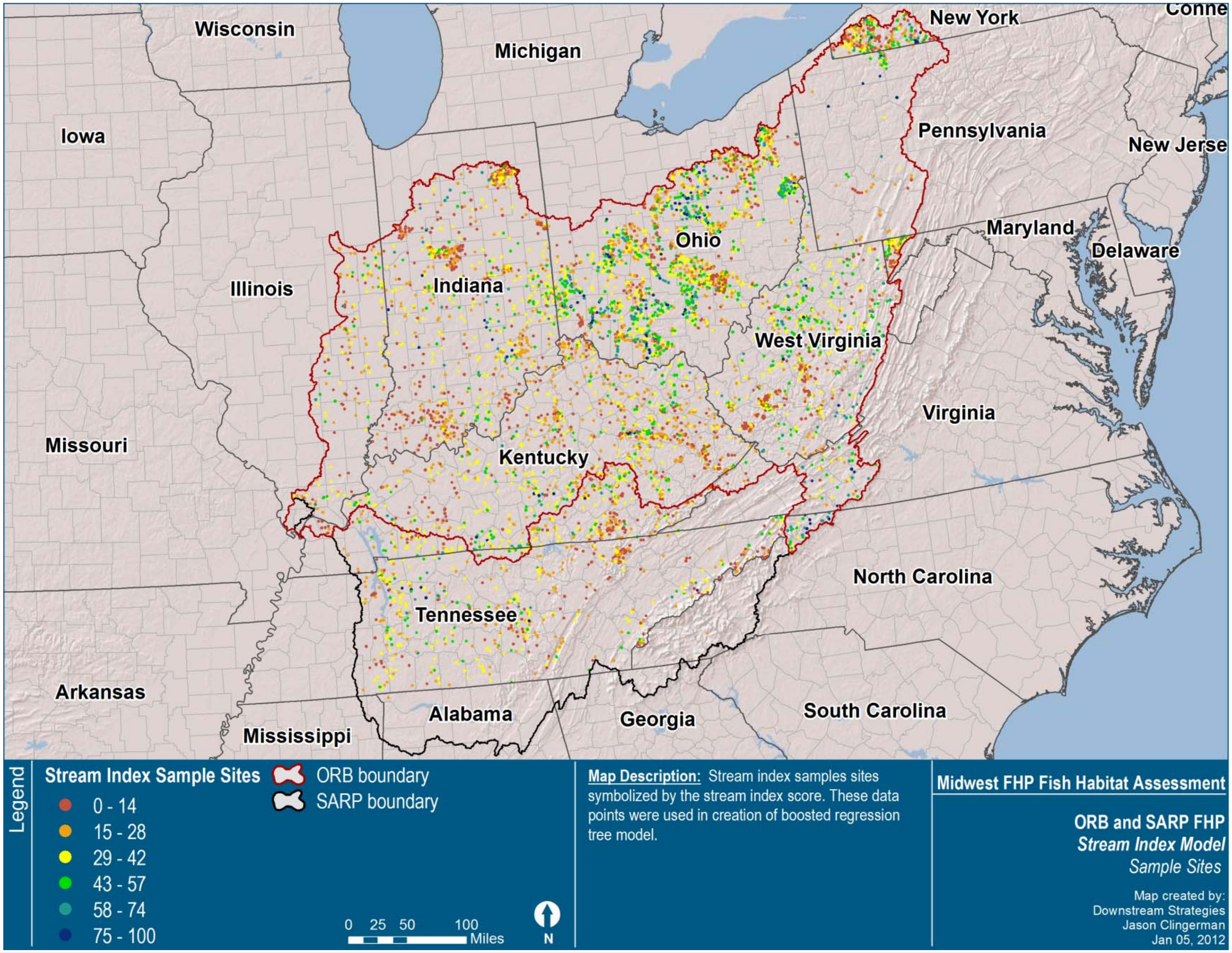
ORB and SARP provided DS with a fish collection dataset comprised of 5,346 observations from 1996 to 2010 from catchments with a stream order of less than six. These 5,346 observations were used to construct the stream index BRT model. Figure 3 maps all of the sampling sites that were used to construct the model and outlines all of the 1:100k catchments to which the modeling outputs were applied.

The stream index variable was fully processed by ORB and SARP prior to being transferred to DS. Stream index values were calculated for each sample site in the dataset using abundance data for 12 selected small stream species and three selected small stream genera (Table 1). Each species or taxa scored 0-5 for each site based on 5th, 25th, 50th, and 75th percentile values of the species' abundance from all events (based on box plots). Then, species scores were added to produce a final stream index score for each site. Scores were stratified by stream size and Omernik's Level II Ecoregions. Sites were divided into small rivers (fourth and fifth order) and headwaters (first through third order). Scores were standardized from each stream size and ecoregion by dividing individual scores by the maximum observed for each stratum.

Table 1: Small streams signature fish index taxa list with scoring thresholds based on abundance

		Stream order				
Scientific name	**Common name**	**1**	**2**	**3**	**4**	**5**
Ammocrypta spp	Sand darters	1	1	2	5	11
Aphredoderus sayanus	Pirate perch	1	1	3	8	18
Clinostomus spp	Dace	1	2	6	27	63
Cottus spp	Sculpins	1	4	17	61	146
Elassoma zonatum	Banded pygmy sunfish	1	1	1	1	1
Esox niger	Chain pickerel	1	1	1	2	2
Etheostoma spp	Darters	1	3	8	24	55
Forbesichtys agassizii	Spring cavefish	1	1	1	1	1
Lepomis punctatus	Spotted sunfish	1	1	1	1	1
Lota lota	Burbot	1	1	1	2	2
Micropterus dolomieu	Smallmouth bass	1	2	5	13	29
Noturus phaeus	Brown madtom	1	1	5	11	22
Percina spp	Darters	1	3	6	13	21
Phoxinus spp	Dace	1	2	9	35	83
Salvenius fontinalis	Brook trout	1	3	9	21	48
Umbra limi	Central mudminnow	1	1	2	7	16

Figure 3: Small streams signature fish index modeling area and sampling sites



2.2 Modeling process

2.2.1 *Predictive performance*

The final selected model was comprised of 5,700 trees. The model had a CV deviance statistic of 240.138±4.583 and a CV correlation statistic of 0.619±0.007.

2.2.2 *Variable influence*

The BRT output includes a list of the predictor variables used in the model, ordered and scored by their relative importance. The relative importance values are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman and Meulman, 2003). The relative influence score is scaled so that the sum of the scores for all variables is 100, where higher numbers indicate greater influence. Of the 44 predictor variables used to develop the stream index model, 43 had a relative influence value greater than zero (

Table 2). The five most influential predictors, which accounted for over 47% of the total influence in the model, were:

- network drainage area,
- Level III Ecoregion,
- network wetland land cover,
- mean annual air temperature, and
- network mean baseflow index.

The five most influential anthropogenic stressors, which accounted for almost 23% of the total influence, were:

- network wetland land cover,
- network density of cattle,
- network riparian disturbance,
- network impervious surface cover, and
- network pasture land cover.

The five most influential natural habitat variables, which contributed almost 43% of the total influence, were:

- network drainage area,
- Level III Ecoregion,
- mean annual air temperature,
- network mean baseflow index, and
- network C, C/D soil cover.

Network drainage area, the single most important variable in terms of relative influence, contributed nearly 16% of the total influence.

Table 2: Relative influence of all variables in the final stream index model

Variable code	Variable description	Relative influence
cumdrainag	Network drainage area	15.94
eco_code3	Level III Ecoregion	12.94
wetlandpc	Network wetland land cover	6.64
temp	Mean annual air temperature	6.18
bfi_meanc	Network mean baseflow index	5.85
cattlec	Network density of cattle	4.36
ripdispc	Network riparian disturbance index	4.31
imp06c	Network impervious surface cover	3.87
pastpc	Network pasture land cover	3.80
roadcrc_den	Network density of road crossings	3.10
cropspc	Network rowcrop land cover	2.92
water_swc	Network surface water consumption	2.70
water_gwc	Network groundwater consumption	2.22
grasspc	Network grassland land cover	2.19
devp	Local developed land cover	2.10
popdens	Local population density	2.09
soil3pc	Network soil group C,C/D cover	2.06
roadcr_den	Local density of road crossings	1.78
tric_den	Network density of Toxic Release Inventory sites	1.69
damsc_den	Local wetland land cover	1.67
ripdisp	Local riparian disturbance	1.45
brock7pc	Network shale bedrock geology cover	1.45
brock1pc	Network carbonate bedrock geology cover	1.03
minesc_den	Network density of mines	0.94
brock6pc	Network sandstone bedrock geology land cover	0.80
soil4pc	Network soil group D cover	0.78
brock4pc	Network metamorphic bedrock geology cover	0.75
npdesc_den	Network density of National Pollutant Discharge Elimination System permits	0.55
surf4pc	Network lacustrine surficial geology cover	0.51
soil1pc	Network soil group A, A/D cover	0.35
surf2pc	Network outwash surficial geology cover	0.35
brock3pc	Network mafic/igneous bedrock geology cover	0.33
surf7pc	Network clay surficial geology cover	0.32
tri_den	Local density of Toxic Release Inventory sites	0.32
surf6pc	Network residuum surficial geology cover	0.25
surf3pc	Network alluvium surficial geology cover	0.23
brock5pc	Network sand/gravel bedrock geology cover	0.22
brock2pc	Network felsic/igneous bedrock geology cover	0.21
mines_den	Local density of mines	0.21
npdes_den	Local density of National Pollutant Discharge Elimination System permits	0.20
cercc_den	Network density of Superfund sites	0.14
dams_den	Local density of dams	0.13
surf5pc	Network loess surficial geology cover	0.10
cerc_den	Local density of Superfund sites	0.00

Note: Individual variables are highlighted according to whether they were determined to be anthropogenic in nature (red highlight) or natural (green highlight).

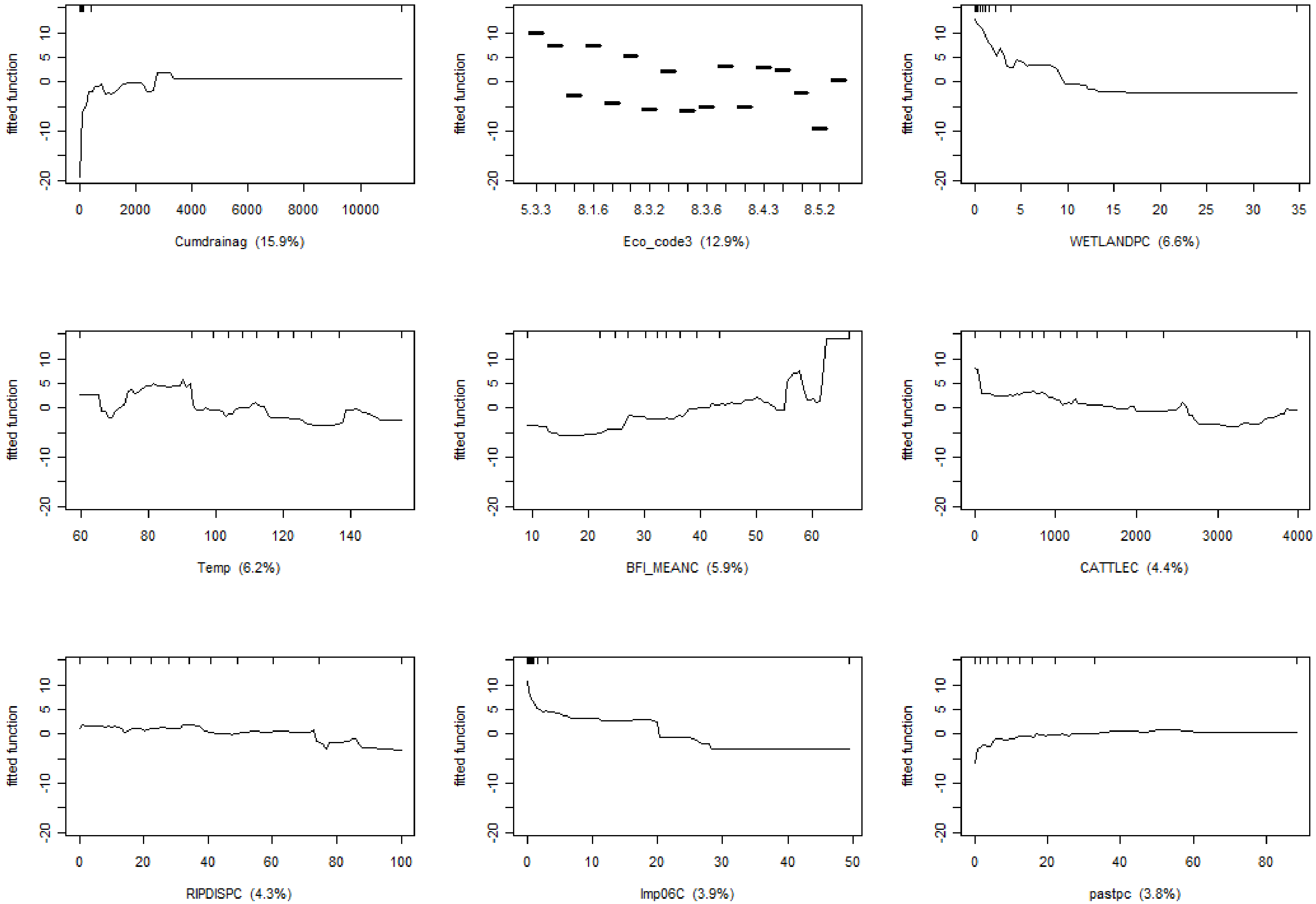
2.2.3 *Variable functions*

The BRT output also contains quantitative information on partial dependence functions that can be plotted to visualize the effect of each individual predictor variable on the response after accounting for all other variables in the model. Similar to the interpretation of traditional regression coefficients, the function plots are not always a perfect representation of the relationship for each variable, particularly if interactions are strong or predictors are strongly correlated. However, they do provide a useful and objective basis for interpretation (Friedman, 2001; Friedman and Meulman, 2003).

These plots show the trend of the response variable (y-axis) as the predictor variable (x-axis) changes. The response variable is transformed (usually to the logit scale) so that the magnitude of trends for each predictor variable's function plot can be accurately compared. The dash marks at the top of each function represent the deciles of the data used to build the model. The function plots for the nine most influential variables in the stream index model (see

Table 2 for reference) are illustrated in Figure 4 below. The plots for all 44 variables are shown in Appendix B.

Figure 4: Functional responses of the dependent variable to individual predictors of stream index



Note: Only the top nine predictors, based on relative influence (shown in parentheses; see Appendix A for descriptions of variable codes), are shown here. See Appendix B for plots of remaining predictor variables.

2.3 Post-modeling

The variable importance table and partial dependence functions of the final BRT model were used to create the post-modeling indices of natural habitat quality and anthropogenic stress for stream index. The CNQI was comprised of 20 variables with relative influence greater than zero that were classified as natural habitat features (Table 3). The CASI was comprised of ten variables with relative influence greater than zero that were classified as anthropogenic habitat features (Table 4). To calculate the cumulative indices (i.e., CNQI and CASI), each of the individual natural or anthropogenic variables used in the two indices was converted to a metric by first applying the appropriate transformations, based on their function plots, and then rescaling the transformed measures to a 0 to 100 scale. To calculate the cumulative index from the individual metrics, the metrics were first multiplied by their appropriate weighting factors (Table 3 and Table 4) and then summed. The CNQI and CASI scores were a result of a rescaling of those weighted and summed metrics, again from 0 to 100.

2.3.1 *Variable weights*

Table 3 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CNQI. The five most influential factors in the CNQI were:

- network drainage area,
- Level III Ecoregion,
- mean annual air temperature,
- network baseflow index, and
- network C, C/D soil cover.

Table 4 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CASI. The five most influential factors in the CASI were:

- network density of cattle,
- network riparian disturbance,
- network impervious surface cover,
- network rowcrop land cover, and
- network grassland land cover.

Table 3: Relative influence and weights for natural variables on stream index

Variable	Variable description	Relative influence	Weighting factor
cumdrainag	Network drainage area	15.94	1
eco_code3	Level III Ecoregion	12.94	0.81
temp	Mean annual air temperature	6.18	0.39
BFI_meanc	Network mean baseflow index	5.85	0.37
soil3pc	Network soil group C, C/D cover	2.06	0.13
brock7pc	Network shale bedrock geology cover	1.45	0.09
brock1pc	Network carbonate bedrock geology cover	1.03	0.07
brock6pc	Network sandstone bedrock geology land cover	0.8	0.05
soil4pc	Network soil group D cover	0.78	0.05
brock4pc	Network metamorphic bedrock geology cover	0.75	0.05
surf4pc	Network lacustrine surficial geology cover	0.51	0.03
soil1pc	Network soil group A, A/D cover	0.35	0.02
surf2pc	Network outwash surficial geology cover	0.35	0.02
brock3pc	Network mafic/igneous bedrock geology cover	0.33	0.02
surf7pc	Network clay surficial geology cover	0.32	0.02
surf6pc	Network residuum surficial geology cover	0.25	0.02
surf3pc	Network alluvium surficial geology cover	0.23	0.01
brock5pc	Network sand/gravel bedrock geology cover	0.22	0.01
brock2pc	Network felsic/igneous bedrock geology cover	0.21	0.01
surf5pc	Network loess surficial geology cover	0.1	0.01

Table 4: Relative influence and weights for anthropogenic variables on stream index

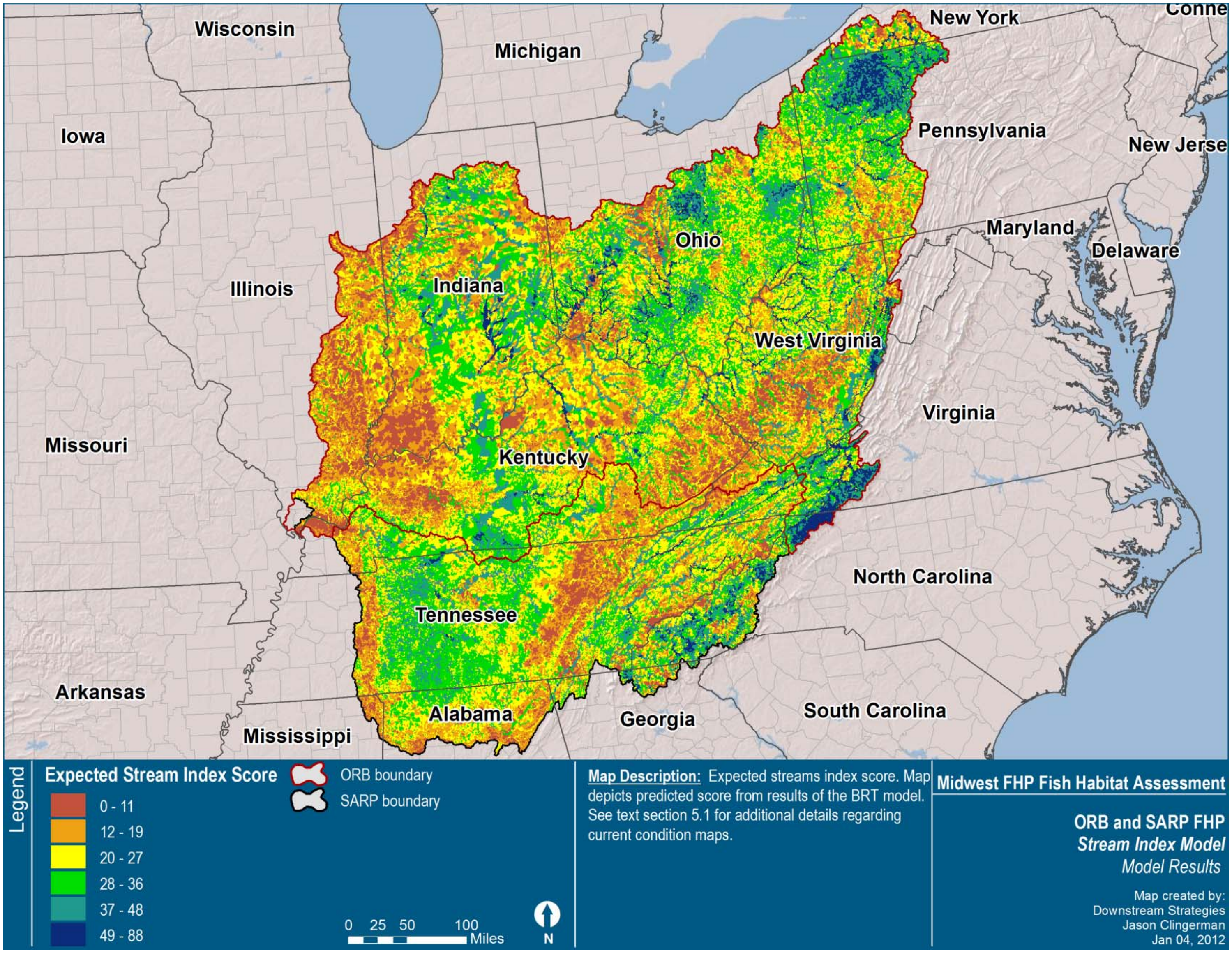
Variable	Variable description	Relative influence	Weighting factor
cattlec	Network density of cattle	4.36	1
ripdispc	Network riparian disturbance index	4.31	0.99
imp06c	Network impervious surface cover	3.87	0.89
cropspc	Network rowcrop land cover	2.92	0.67
grasspc	Network grassland land cover	2.19	0.5
TRIc_den	Network density of Toxic Release Inventory sites	1.69	0.39
damsc_den	Local wetland land cover	1.67	0.38
ripdisp	Local riparian disturbance	1.45	0.33
NPDESc_den	Network density of National Pollutant Discharge Elimination System permits	0.55	0.13
dams_den	Local density of dams	0.13	0.03

2.4 Mapped Results

2.4.1 *Expected current conditions*

Stream index was calculated for all 1:100k stream catchments in the study area using the BRT model. The predicted probability values ranged from 0 to 87.9. The mean predicted stream index score for the total 226,919 catchments was 24.5. These results are mapped in Figure 5.

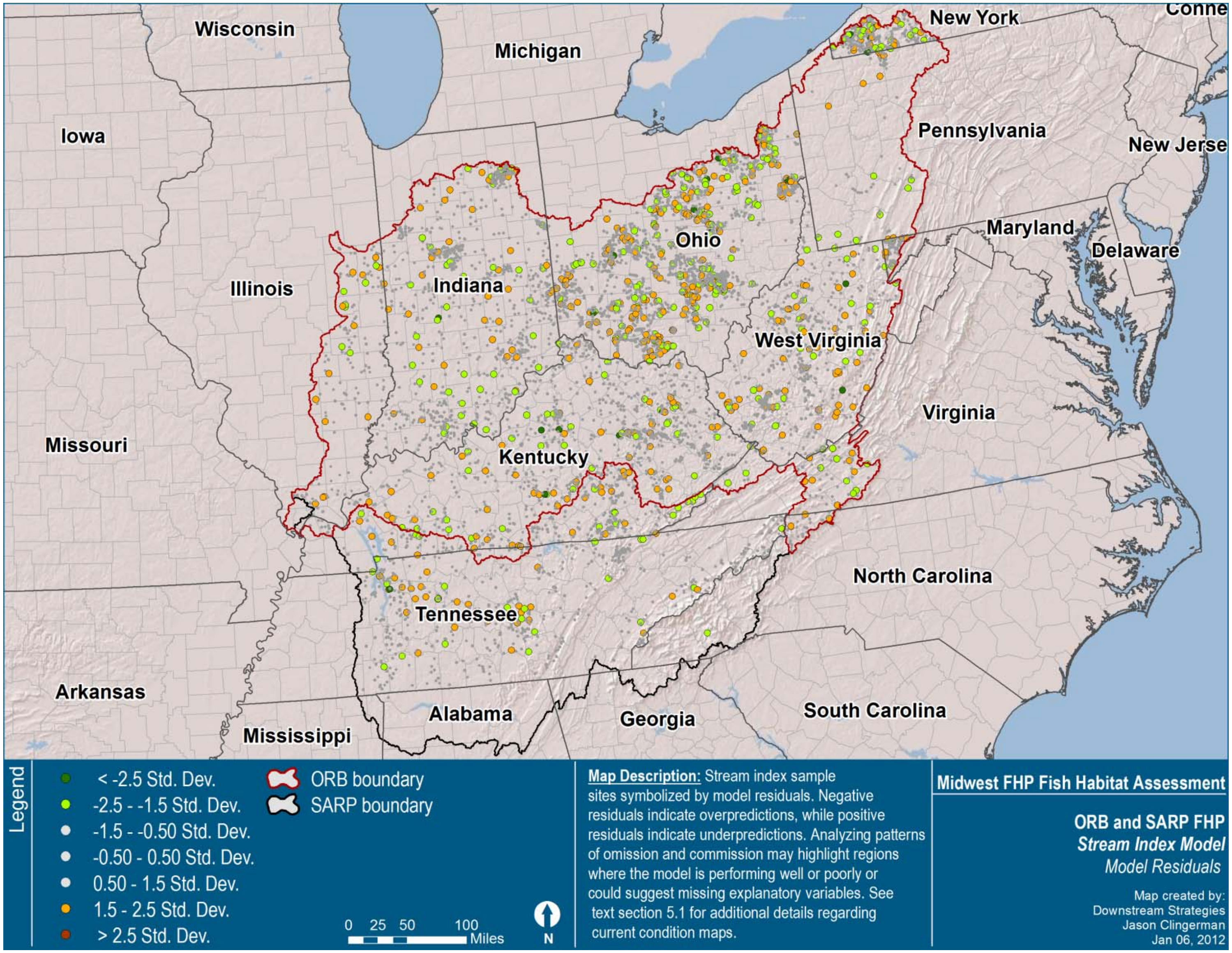
Figure 5: Expected stream index score



2.4.2 *Spatial variability in predictive performance*

Analyzing patterns of omission and commission may highlight regions where the model is performing well or poorly or could suggest missing explanatory variables (Figure 6). To assess omission and commission, residuals are also calculated by the BRT model. The residuals are a measure of the difference in the measured and modeled values (measured value *minus* modeled value). Negative residuals indicate overpredictions (predicting higher values than are true), while positive residuals indicate underpredictions (predicting lower values than are true).

Figure 6: Distribution of stream index model residuals by sampling site



2.4.3 *Indices of stress and natural quality*

Maps of CNQI and CASI illustrate the spatial distribution of natural habitat potential (i.e., CNQI score) and anthropogenic stress (i.e., CASI score) in the ORB and SARP. CNQI and CASI scores are mapped in Figure 7 and Figure 8, respectively. The top five most influential variables toward the calculation of CNQI are shown in Figure 9-Figure 13. The top five variables in the calculation of CASI are mapped in Figure 14-Figure 18. CNQI, CASI, and their metrics are all scaled on a 0-100 scale (see Section 2.3 for more details on CNQI and CASI calculation). For CNQI, higher values indicate higher natural quality, while higher values for CASI indicate higher levels of anthropogenic stress.

Figure 7: Cumulative natural quality index for small streams signature fish index

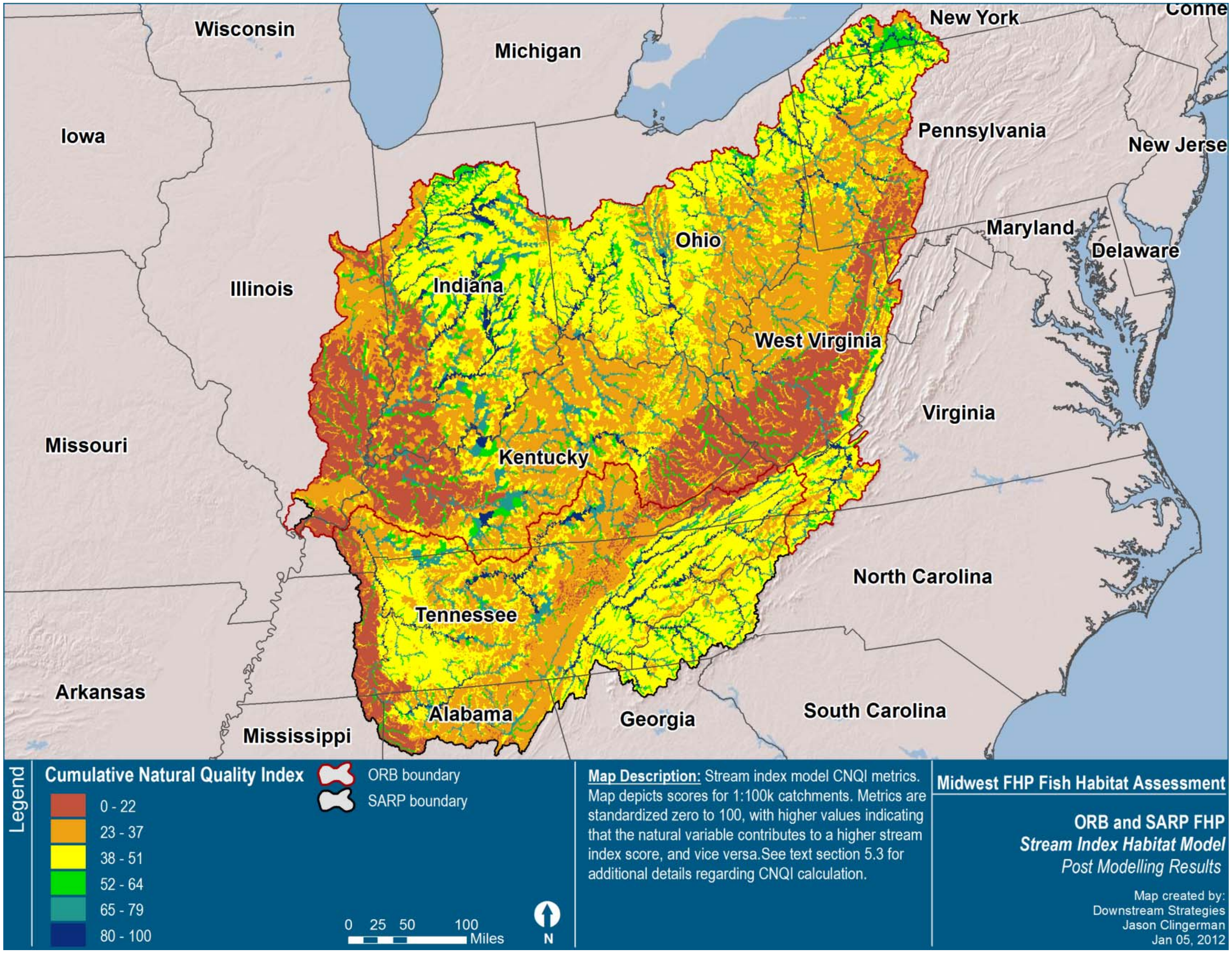


Figure 8: Cumulative anthropogenic stress index for small streams signature fish index

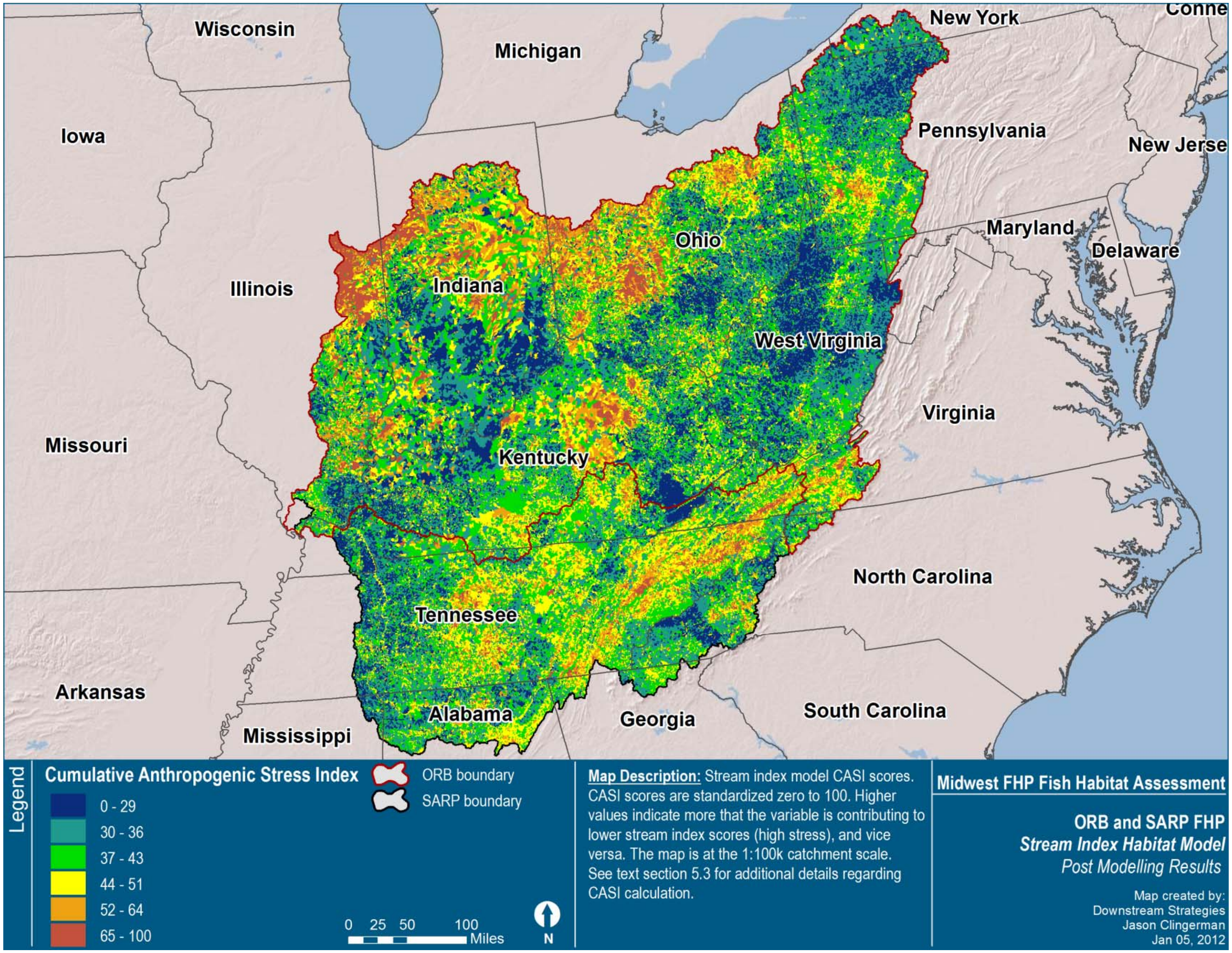


Figure 9: Most influential natural index metric for small streams signature fish index

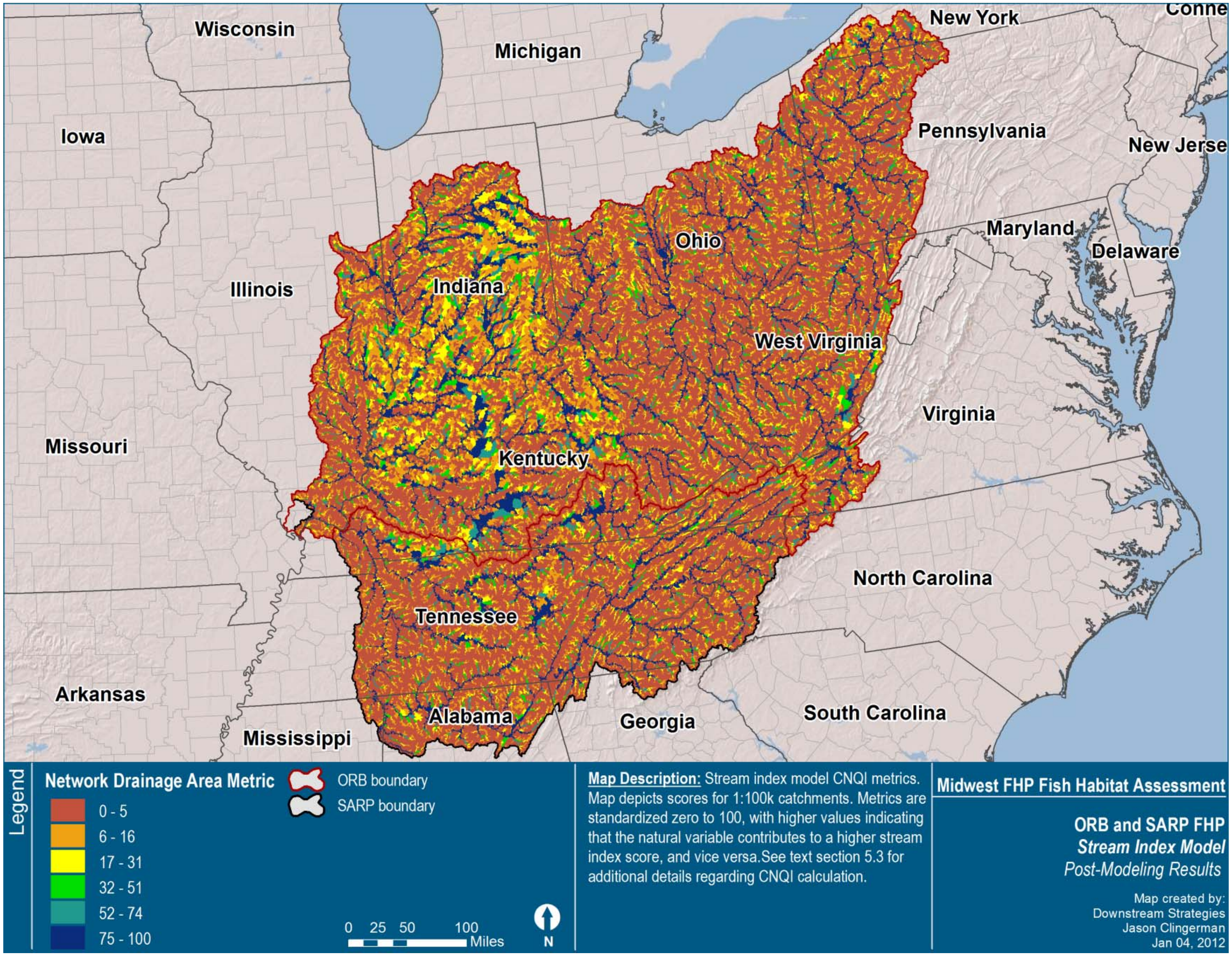


Figure 10: Second most influential natural index metric for small streams signature fish index

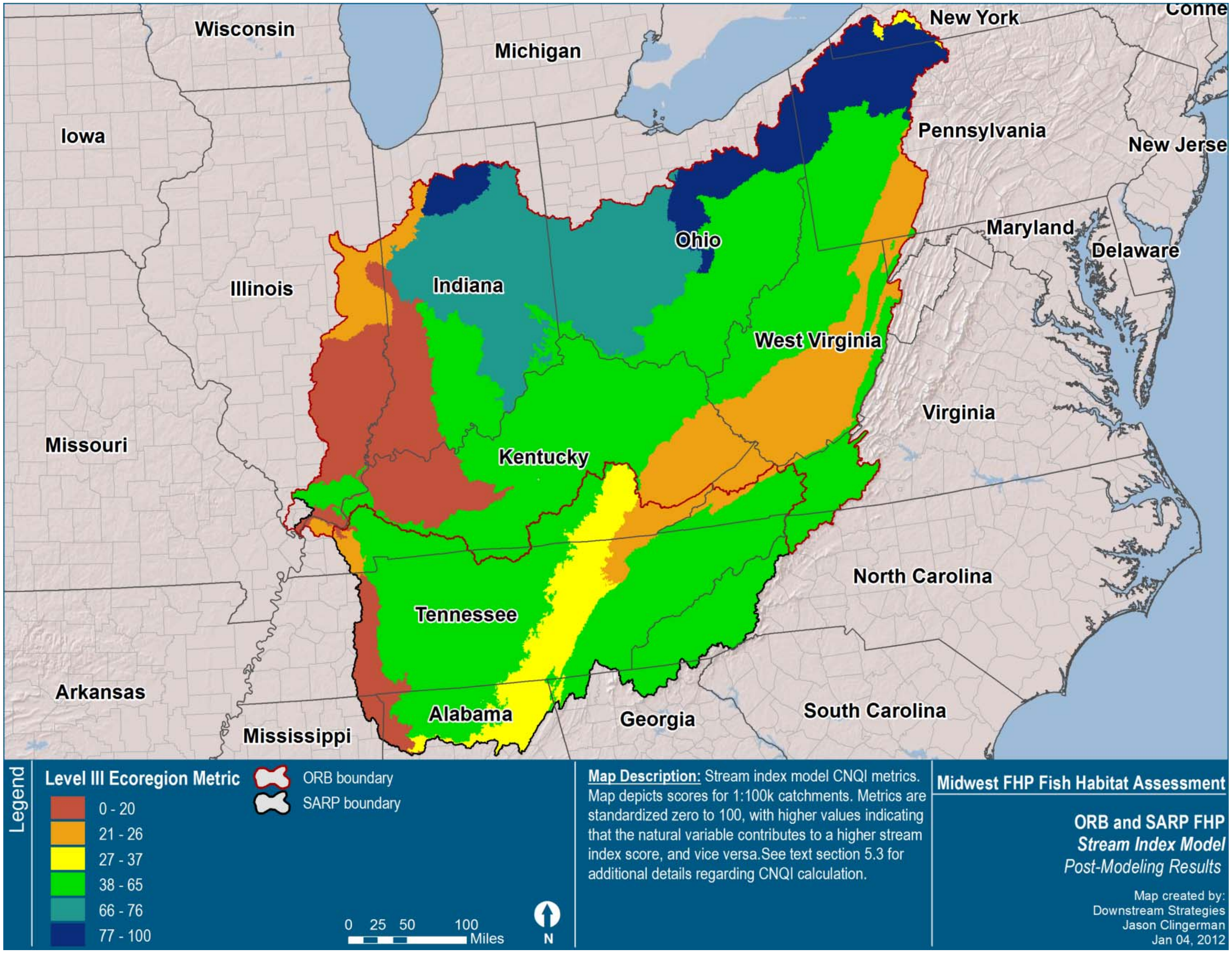


Figure 11: Third most influential natural index metric for small streams signature fish index

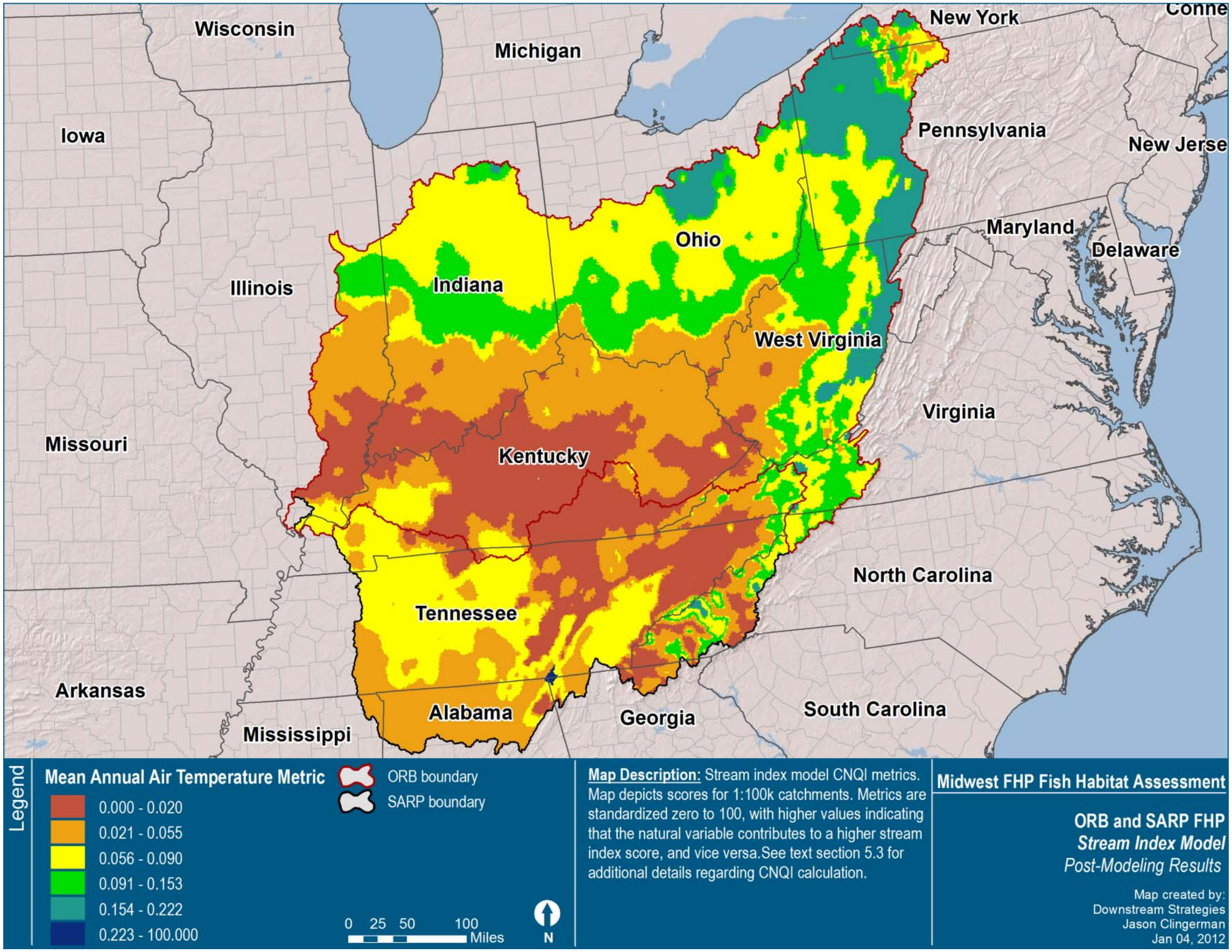


Figure 12: Fourth most influential natural index metric for small streams signature fish index

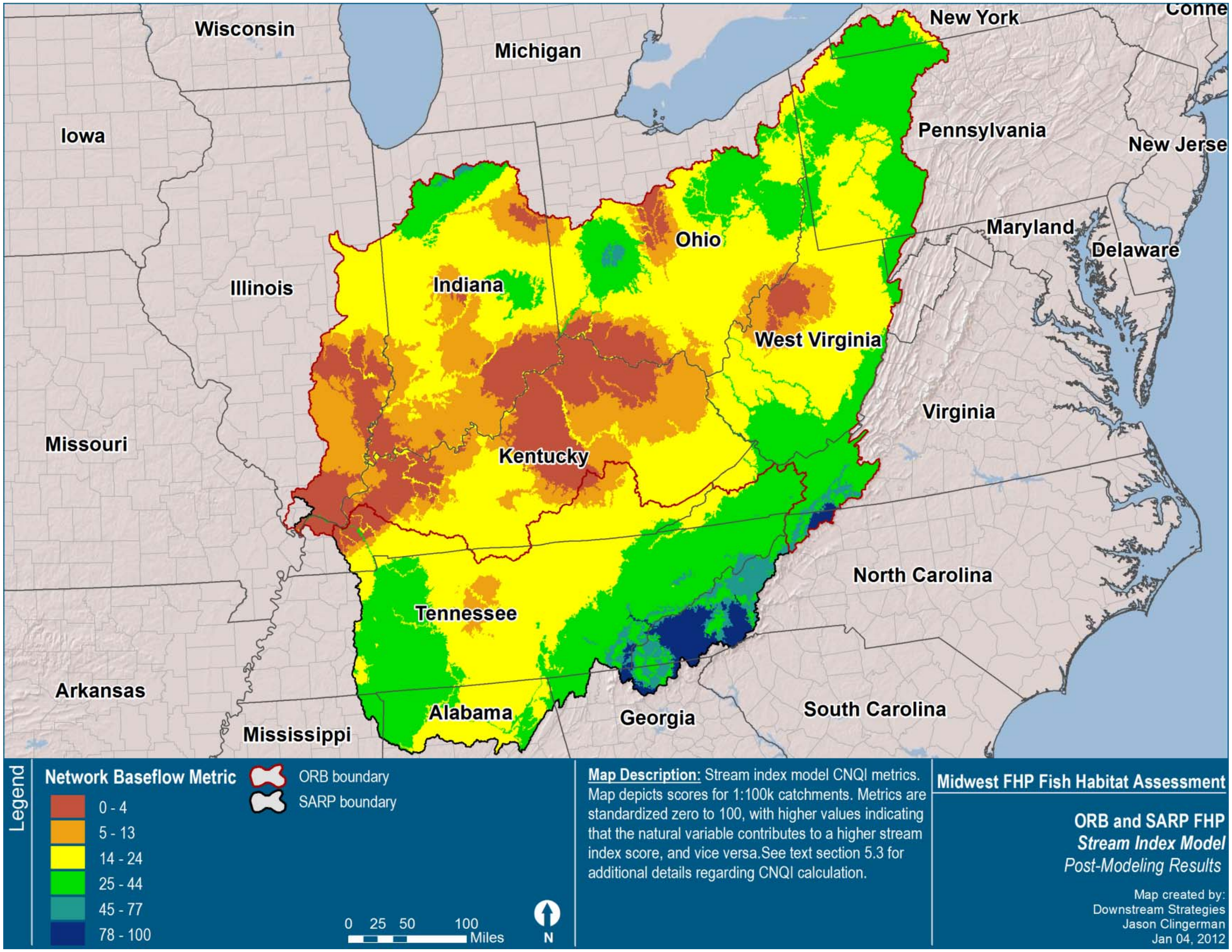


Figure 13: Fifth most influential natural index metric for small streams signature fish index

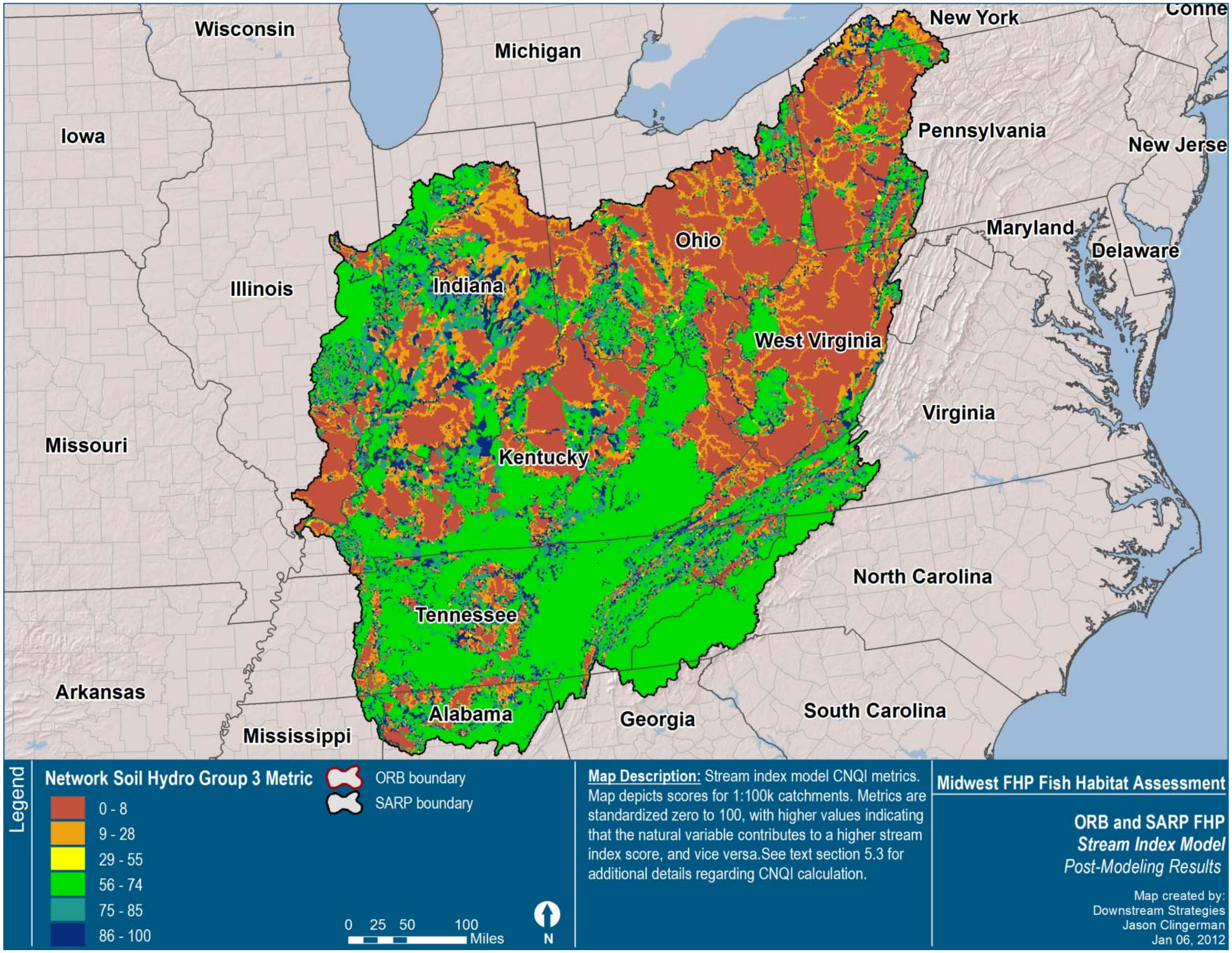


Figure 14: Most influential anthropogenic index metric for small streams signature fish index

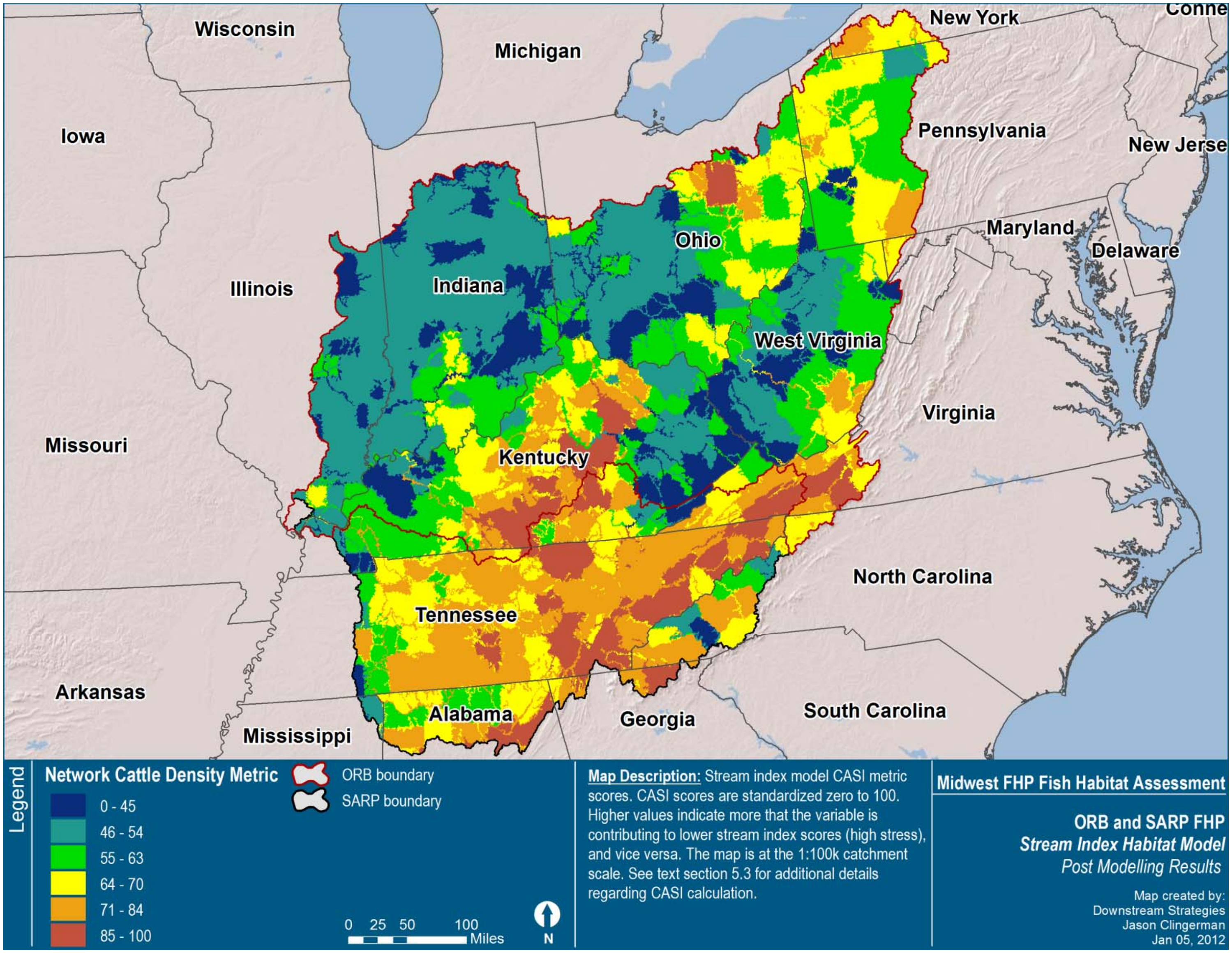


Figure 15: Second most influential anthropogenic index metric for small streams signature fish index

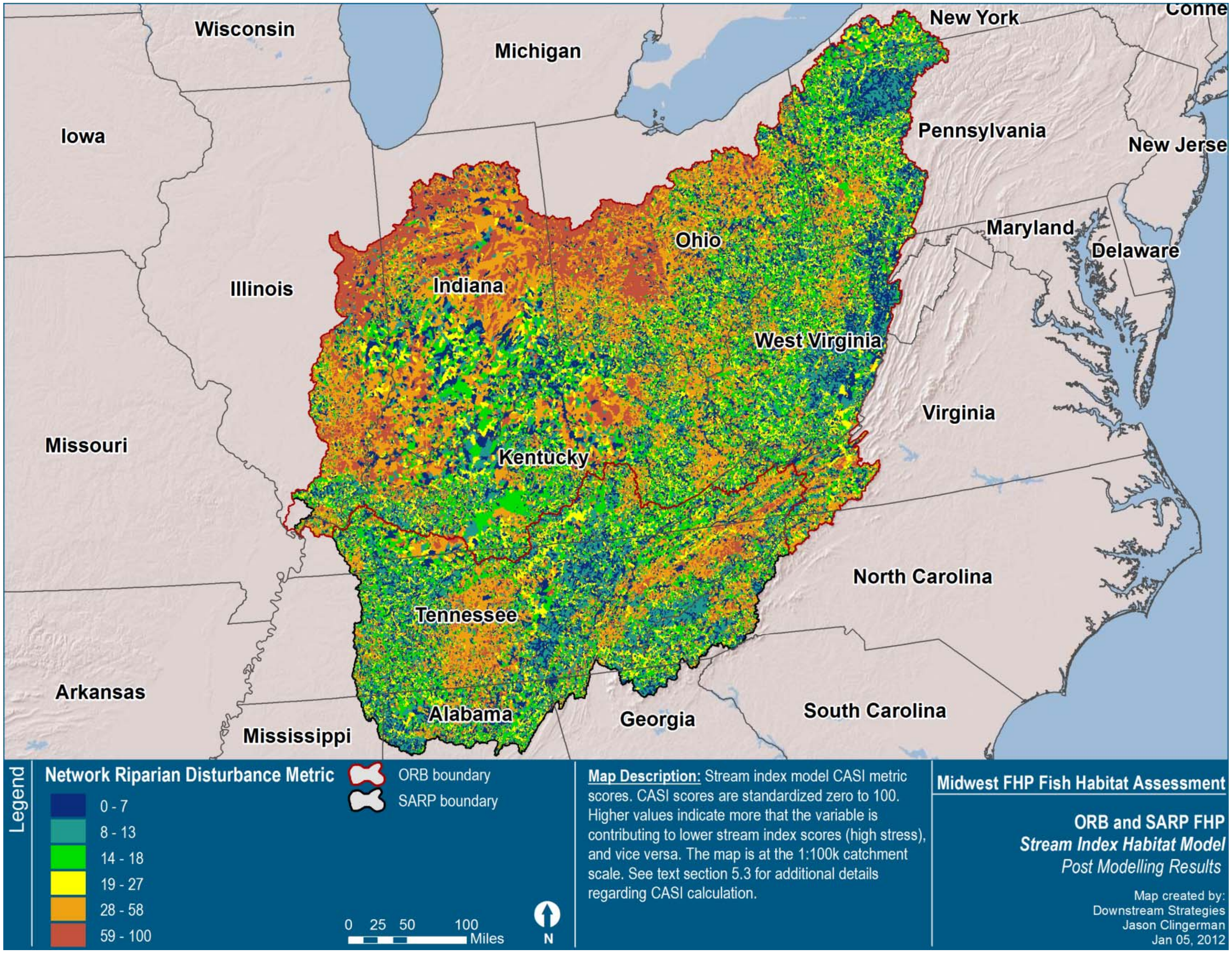


Figure 16: Third most influential anthropogenic index metric for small streams signature fish index

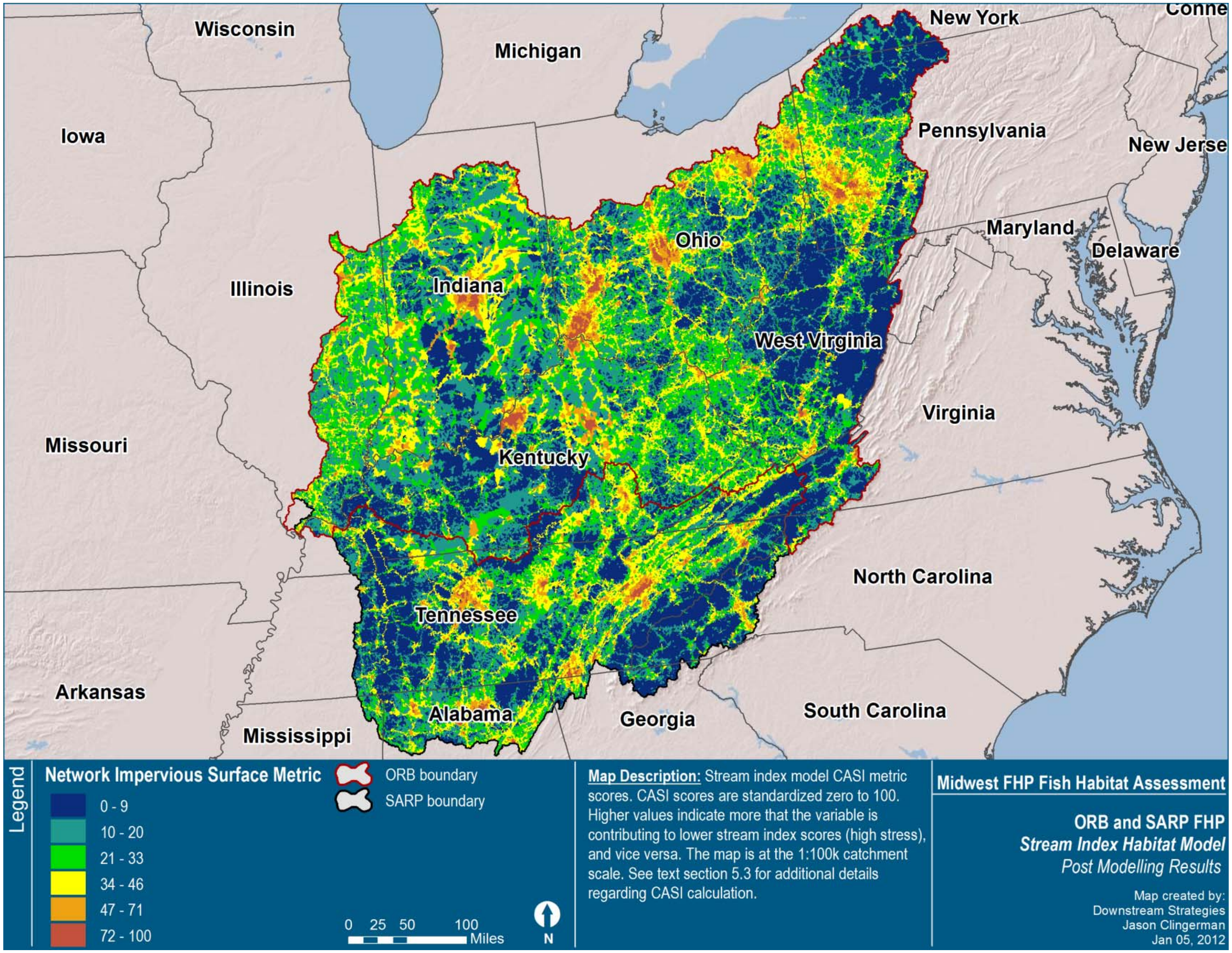


Figure 17: Fourth most influential anthropogenic index metric for small streams signature fish index

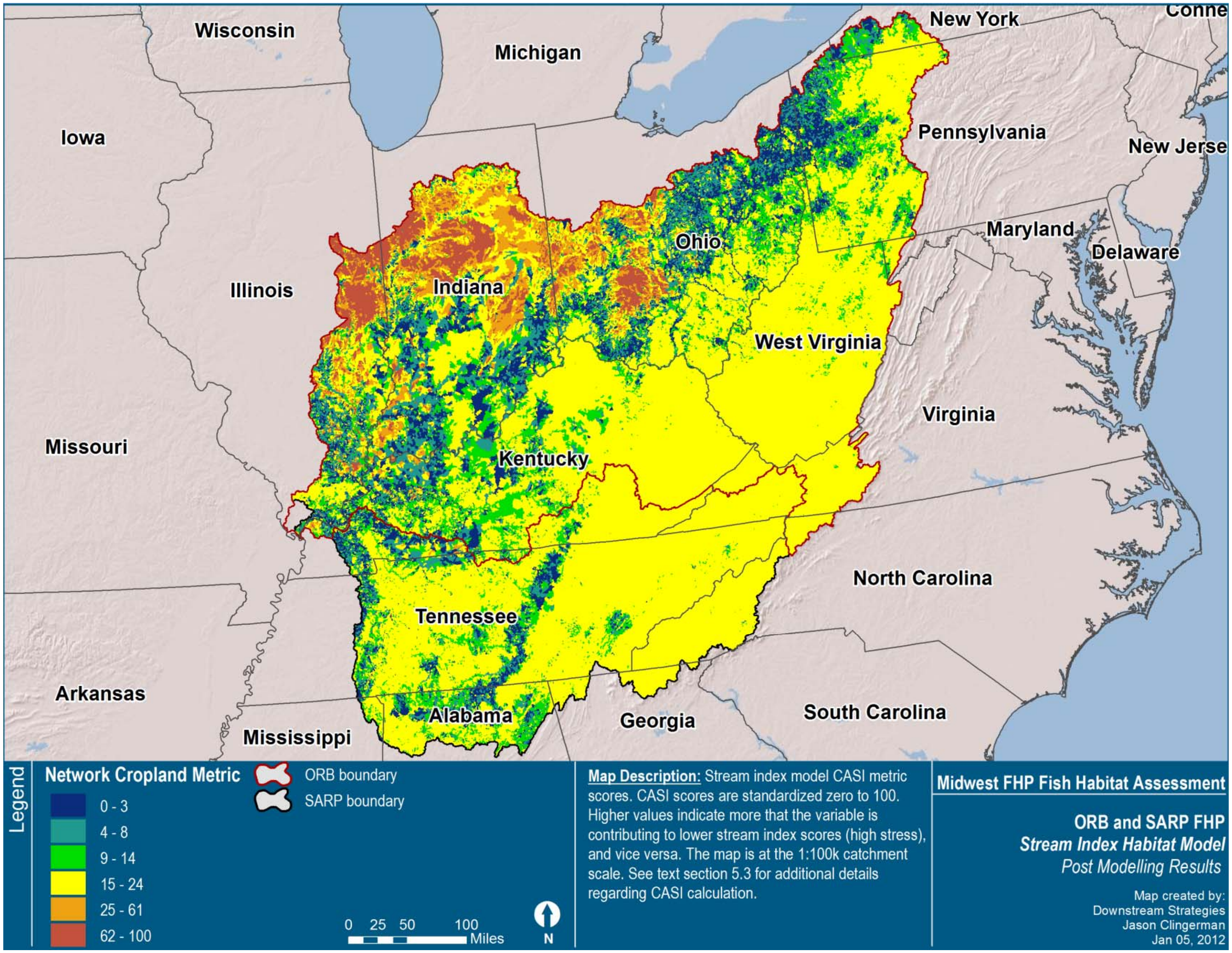
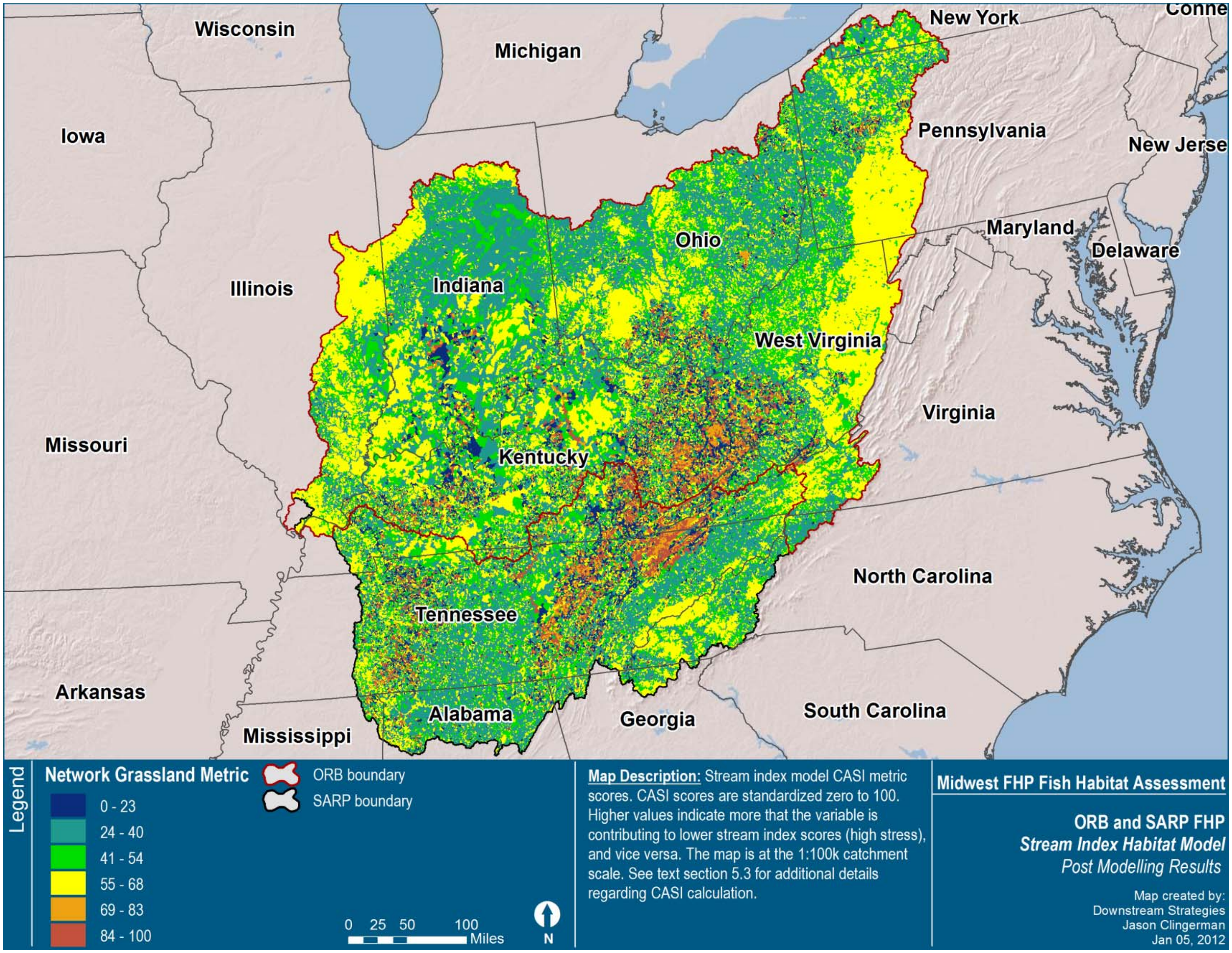


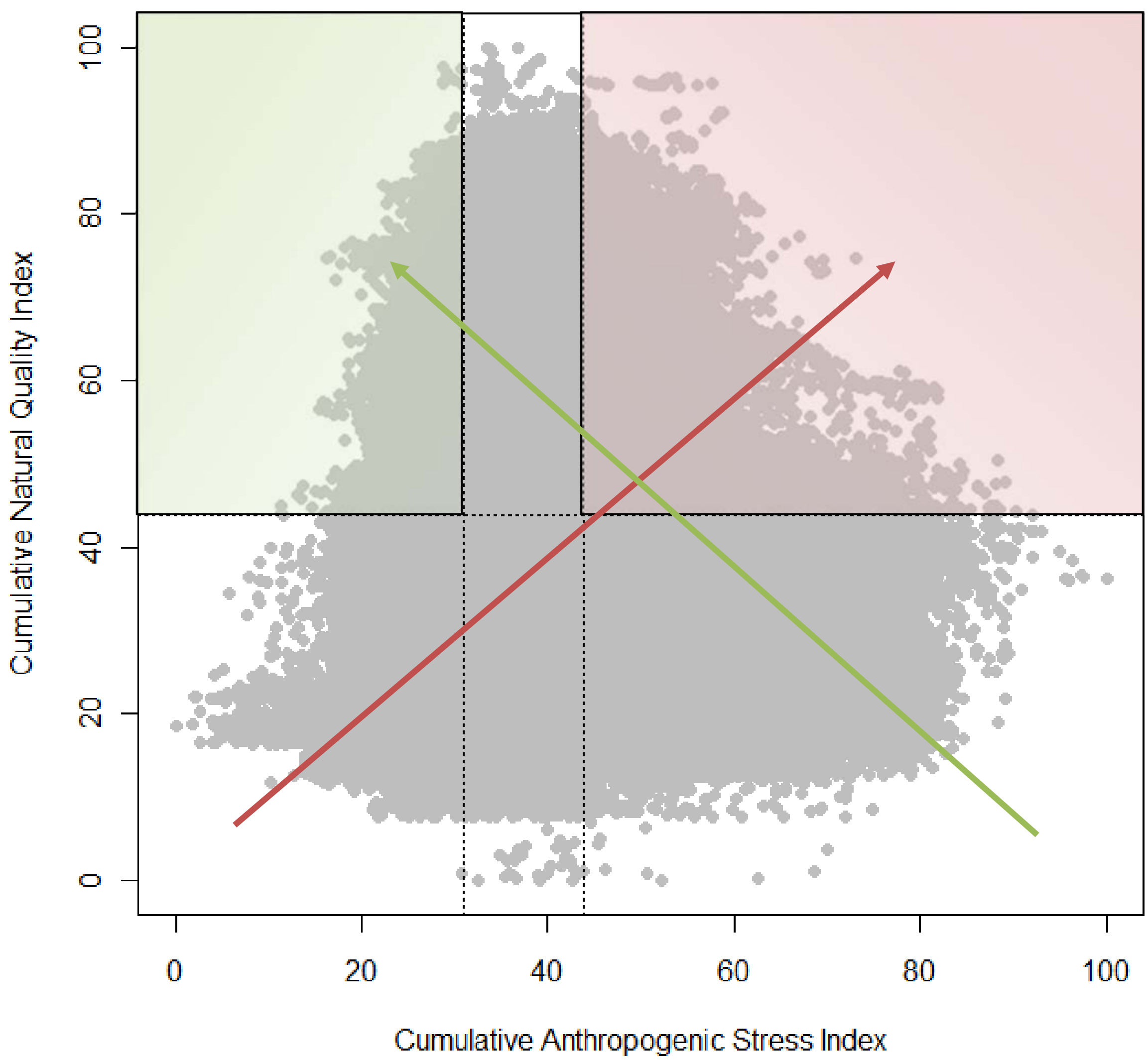
Figure 18: Fifth most influential anthropogenic index metric for small streams signature fish index



2.4.4 *Restoration and protection priorities*

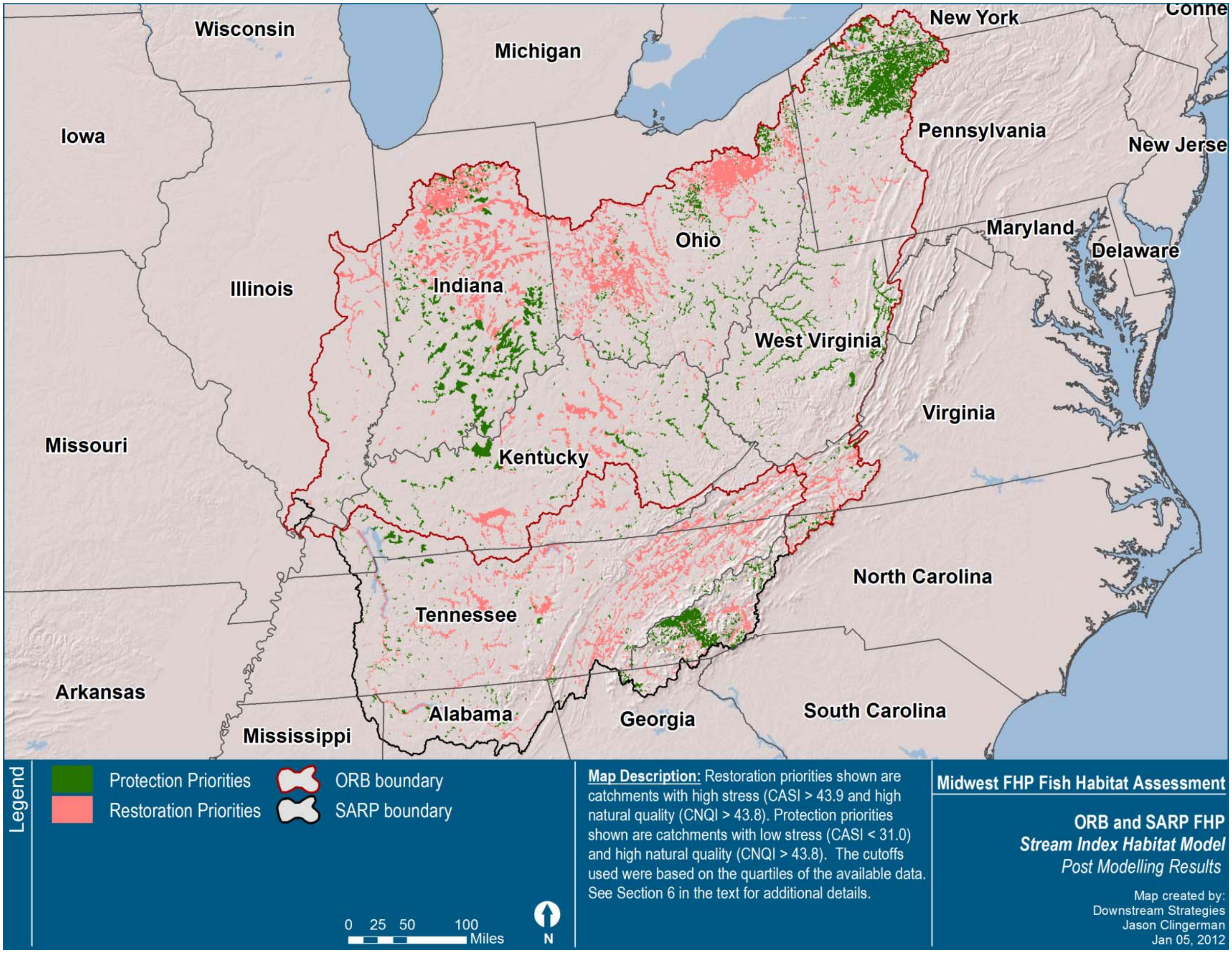
A plot of CNQI versus CASI values for all catchments in the study area (Figure 19) can be used as a reference when defining thresholds for categories of CNQI and CASI scores for use in the development of restoration and protection priorities. In the example shown (Figure 20), thresholds for restoration (high natural potential coupled with high anthropogenic stress) were set to CNQI greater than 43.8 and CASI greater than 43.9 (third quartiles). The thresholds used for protection priorities (high natural potential and low anthropogenic stress) were CNQI greater than 43.8 and CASI less than 31.0 (first quartile).

Figure 19: CNQI versus CASI values for all catchments for small streams signature fish index



Note: Breakpoints for CNQI and CASI classes in this example are denoted by dashed lines. The arrows indicate the directions of increasing potential protection (green arrow) or restoration (red arrow) priority. The red box indicates catchments defined as restoration priorities under the example scenario. The green box indicates catchments defined as protection priorities under the same scenario.

Figure 20: Restoration and protection priorities for small streams signature fish index



3. MODIFIED INDEX OF CENTERS OF DIVERSITY

3.1 Modeling inputs

DS used a list of predictor variables selected by ORB and SARP to develop a ten-fold CV BRT model for the modified index of centers of diversity ("diversity index") at the 1:100k catchment scale. The model was used to produce maps of expected current diversity index scores and maps of expected current natural habitat quality and anthropogenic stress at the 1:100k scale throughout the extents of both FHPs.

DS cooperated with ORB and SARP to arrive at a list of landscape-based habitat variables used to predict the diversity index throughout the region; ultimately, those variables were also used for characterizing habitat quality and anthropogenic stress. From an initial suite of 372 catchment attributes, DS and the FHPs compiled a list of 103 predictors for evaluation. From that list, 58 variables were removed due to statistical redundancy ($r > 0.6$) or logical redundancy, resulting in a final list of 45 predictor variables for the BRT model and assessment. See Appendix A for a full data dictionary and the metadata document for variable processing notes.

ORB and SARP provided DS with a fish collection dataset comprised of 6,193 observations from 1996 to 2010. From this pool of data, DS removed three observation points with duplicate observations per catchment. After applying this initial filter, 6,190 observations were used to construct the diversity index BRT model. Figure 21 maps all of the sampling sites that were used to construct the model and outlines all of the 1:100k catchments to which the modeling outputs were applied.

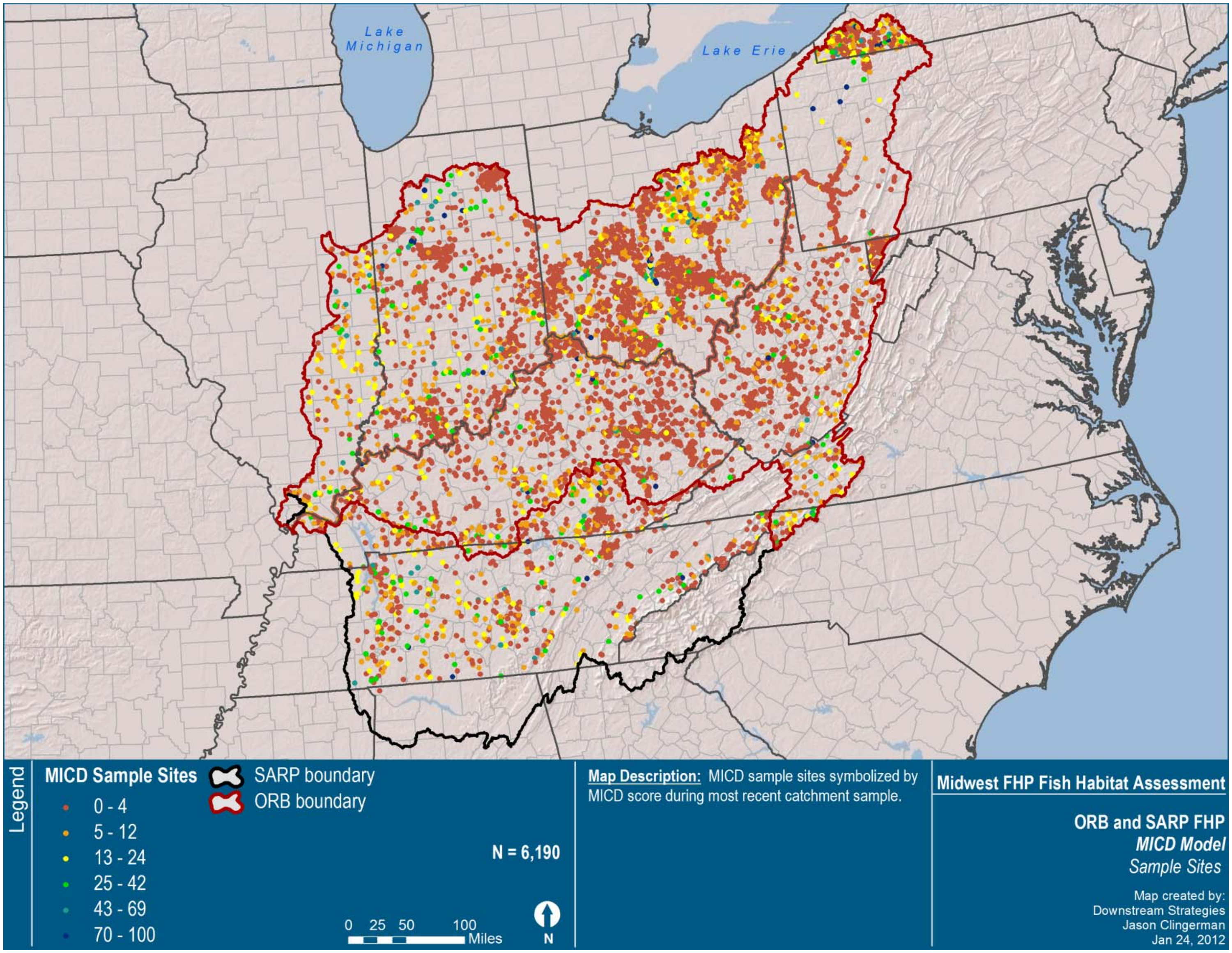
The diversity index variable was fully processed by ORB and SARP prior to being transferred to DS. To calculate the diversity index for each sample site in the data, a value was assigned based on the densities of species at the site relative to the density of species across all sites used in the analysis (Bear, 2006; Patton, 2001). The diversity index was calculated based on Equation 1. The diversity index formula was calculated separately for each stream size class with each Level II Ecoregion. The index included only native species.

Equation 1: Modified index of centers of diversity

$$Relative\ density = \frac{density\ of\ a\ species\ at\ a\ site}{sum\ density\ of\ the\ species\ at\ all\ sites\ within\ an\ ecoregion\ or\ stream\ order}$$

$$Modified\ Index\ of\ Centers\ of\ Diversity = sum\ of\ the\ relative\ densities\ for\ all\ species\ at\ a\ site$$

Figure 21: Modified index of centers of diversity modeling area and sampling sites



3.2 Modeling process

3.2.1 *Predictive performance*

The final selected model was comprised of 4,150 trees. The model had a CV deviance statistic of 0.577±0.017 and a CV correlation statistic of 0.636±0.010.

3.2.2 *Variable influence*

The BRT output includes a list of the predictor variables used in the model ordered and scored by their relative importance. The relative importance values are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman and Meulman, 2003). The relative influence score is scaled so that the sum of the scores for all variables is 100, where higher numbers indicate greater influence. Of the 45 predictor variables used to develop the diversity index model, 44 had a relative influence value greater than zero (Table 5). The five most influential predictors were all natural habitat variables and accounted for almost 54% of the total influence in the model:

- network drainage area,
- Level III Ecoregion,
- mean annual air temperature,
- catchment slope, and
- network mean baseflow index.

The five most influential anthropogenic stressors, which accounted for almost 13% of the total influence, were:

- network rowcrop land cover,
- network density of cattle,
- network impervious surface cover,
- network wetland land cover, and
- network pasture land cover.

Network drainage area, the single most important variable in terms of relative influence, contributed more than 25% of the total influence.

Table 5: Relative influence of all variables in the final diversity index model

Variable code	Variable description	Relative influence
cumdrainag	Network drainage area	25.04
eco_code3	Level III Ecoregion	17.66
temp	Mean annual air temperature	5.06
slope	Slope of catchment flowline	3.15
BFI_MEANC	Network mean baseflow index	2.89
cropspc	Network rowcrop land cover	2.83
brock6pc	Network sandstone bedrock geology land cover	2.80
cattlec	Network density of cattle	2.80
precip	Mean annual precipitation	2.66
imp06c	Network impervious surface cover	2.56
wetlandpc	Network wetland land cover	2.50
minelevraw	Minimum catchment elevation	2.49
pastpc	Network pasture land cover	2.20
water_swc	Network surface water consumption	2.08
grasspc	Network grassland land cover	1.94
ripdispc	Network riparian disturbance index	1.75
water_gwc	Network groundwater consumption	1.45
forpc	Network forested land cover	1.43
roadcr_den	Local density of road crossings	1.42
imp06	Local impervious surface cover	1.40
brock7pc	Network shale bedrock geology cover	1.30
soil3pc	Local soil group C,C/D cover	1.13
ripdisp	Local riparian disturbance index	1.10
brock1pc	Network carbonate bedrock geology cover	1.02
surf7pc	Network clay surficial geology cover	1.01
surf3pc	Network alluvium surficial geology cover	0.98
surf2pc	Network outwash surficial geology cover	0.93
soil2pc	Network soil group B, B/D cover	0.92
surf5pc	Network loess surficial geology cover	0.91
soil1pc	Network soil group A, A/D cover	0.83
surf4pc	Network lacustrine surficial geology cover	0.64
soil4pc	Network soil group D cover	0.54
brock5pc	Network sand/gravel bedrock geology cover	0.52
brock4pc	Network metamorphic bedrock geology cover	0.34
brock5p	Local sand/gravel bedrock geology cover	0.31
TRI_den	Local density of Toxic Release Inventory sites	0.29
surf6pc	Network residuum surficial geology cover	0.27
surf3p	Local alluvium surficial geology cover	0.22
surf8p	Local colluvium surficial geology cover	0.19
dams_den	Local density of dams	0.15
mines_den	Local density of mines	0.10
brock2pc	Network felsic/igneous bedrock geology cover	0.10
brock3pc	Network mafic/igneous bedrock geology cover	0.06
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.06
CERC_den	Local density of Superfund sites	0.00

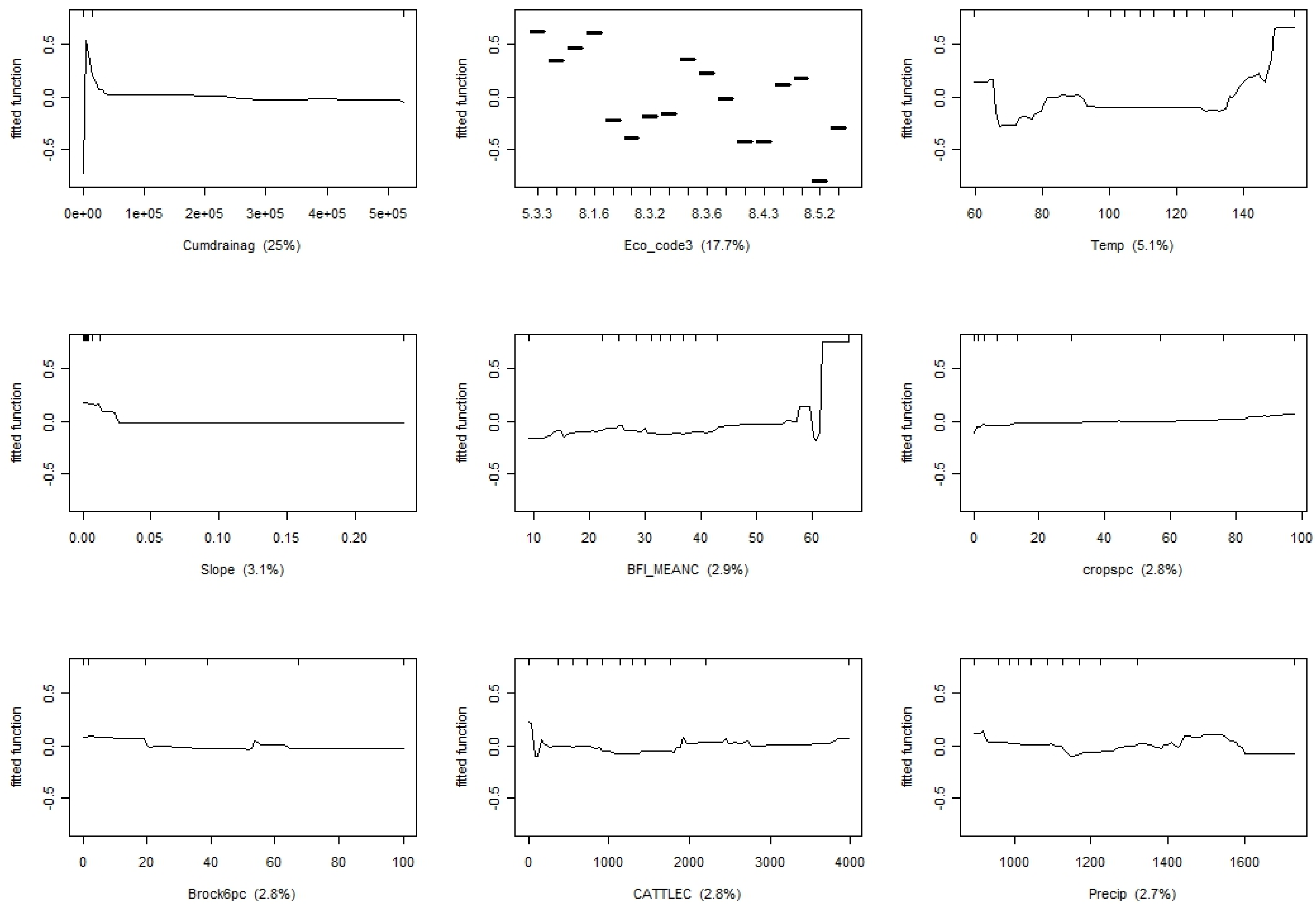
Note: Individual variables are highlighted according to whether they were determined to be anthropogenic in nature (red highlight) or natural (green highlight).

3.2.3 *Variable functions*

The BRT output also contains quantitative information on partial dependence functions that can be plotted to visualize the effect of each individual predictor variable on the response after accounting for all other variables in the model. Similar to the interpretation of traditional regression coefficients, the function plots are not always a perfect representation of the relationship for each variable, particularly if interactions are strong or predictors are strongly correlated. However, they do provide a useful and objective basis for interpretation (Friedman, 2001; Friedman and Meulman, 2003).

These plots show the trend of the response variable (y-axis) as the predictor variable (x-axis) changes. The response variable is transformed (usually to the logit scale) so that the magnitude of trends for each predictor variable's function plot can be accurately compared. The dash marks at the top of each function represent the deciles of the data used to build the model. The function plots for the nine most influential variables in the diversity index model (see Table 5 for reference) are illustrated in Figure 22 below. The plots for all 45 variables are shown in Appendix B.

Figure 22: Functional responses of the dependent variable to individual predictors of diversity index



Note: Only the top nine predictors, based on relative influence (shown in parentheses; see Appendix A for descriptions of variable codes), are shown here. See Appendix B for plots of remaining predictor variables.

3.3 Post-modeling

The variable importance table and partial dependence functions of the final BRT model were used to create the post-modeling indices of natural habitat quality and anthropogenic stress for diversity index. The CNQI was comprised of 27 variables with relative influence greater than zero that were classified as natural habitat features (Table 6). The CASI was comprised of ten variables with relative influence greater than zero that were classified as anthropogenic habitat features (Table 7). To calculate the cumulative indices (i.e., CNQI and CASI), each of the individual natural or anthropogenic variables used in the two indices was converted to a metric by first applying the appropriate transformations, based on their function plots, and then rescaling the transformed measures to a 0 to 100 scale. To calculate the cumulative index from the individual metrics, the metrics were first multiplied by their appropriate weighting factors and then summed. The CNQI and CASI scores were a result of a rescaling of those weighted and summed metrics, again from 0 to 100.

3.3.1 *Variable weights*

Table 6 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CNQI. The five most influential factors in the CNQI were:

- network drainage area,
- Level III Ecoregion,
- mean annual air temperature,
- slope of catchment flowline, and
- network baseflow index.

Table 7 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CASI. The five most influential factors in the CASI were:

- network impervious surface cover,
- network wetland land cover,
- network surface water consumption,
- network grassland land cover, and
- network groundwater consumption.

Table 6: Relative influence and weights for natural variables on diversity index

Variable	Variable description	Relative influence	Weighting factor
cumdrainag	Network drainage area	25.04	1.00
eco_code3	Level III Ecoregion	17.66	0.71
temp	Mean annual air temperature	5.06	0.20
slope	Slope of catchment flowline	3.15	0.13
BFI_meanc	Network mean baseflow index	2.89	0.12
brock6pc	Network sandstone bedrock geology land cover	2.80	0.11
precip	Mean annual precipitation	2.66	0.11
minelevraw	Minimum catchment elevation	2.49	0.10
brock7pc	Network shale bedrock geology cover	1.30	0.05
soil3pc	Local soil group C,C/D cover	1.13	0.04
brock1pc	Network carbonate bedrock geology cover	1.02	0.04
surf7pc	Network clay surficial geology cover	1.01	0.04
surf3pc	Network alluvium surficial geology cover	0.98	0.04
surf2pc	Network outwash surficial geology cover	0.93	0.04
soil2pc	Network soil group B, B/D cover	0.92	0.04
surf5pc	Network loess surficial geology cover	0.91	0.04
soil1pc	Network soil group A, A/D cover	0.83	0.03
surf4pc	Network lacustrine surficial geology cover	0.64	0.03
soil4pc	Network soil group D cover	0.54	0.02
brock5pc	Network sand/gravel bedrock geology cover	0.52	0.02
brock4pc	Network metamorphic bedrock geology cover	0.34	0.01
brock5p	Local sand/gravel bedrock geology cover	0.31	0.01
surf6pc	Network residuum surficial geology cover	0.27	0.01
surf3p	Local alluvium surficial geology cover	0.22	0.01
surf8p	Local colluvium surficial geology cover	0.19	0.01
brock2pc	Network felsic/igneous bedrock geology cover	0.10	0.00
brock3pc	Network mafic/igneous bedrock geology cover	0.06	0.00

Table 7: Relative influence and weights for anthropogenic variables on diversity index

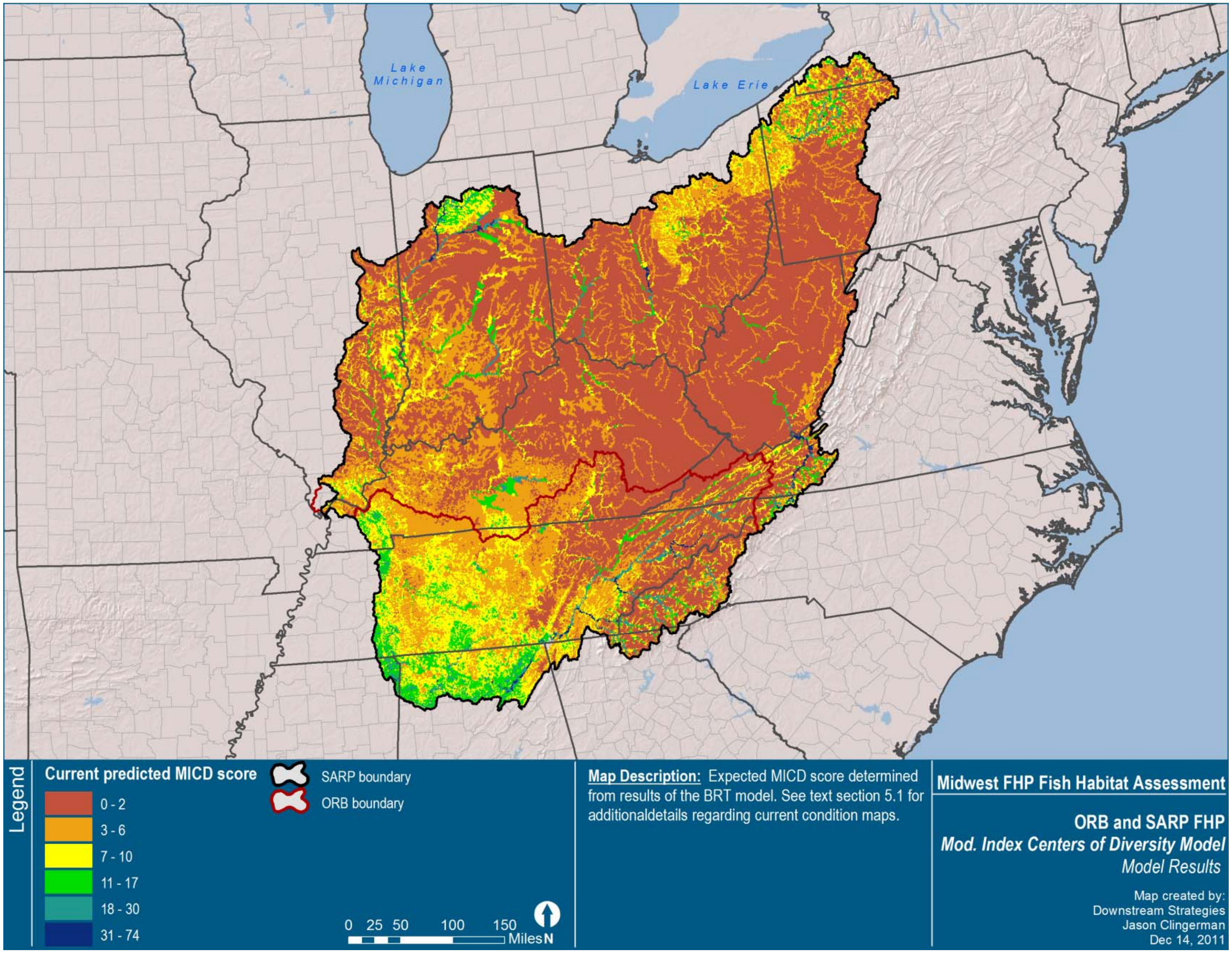
Variable	Variable description	Relative influence	Weighting factor
imp06c	Network impervious surface cover	2.56	1.00
wetlandpc	Network wetland land cover	2.50	0.98
water_swc	Network surface water consumption	2.08	0.81
grasspc	Network grassland land cover	1.94	0.76
water_gwc	Network groundwater consumption	1.45	0.57
forpc	Network forested land cover	1.43	0.56
imp06	Local impervious surface cover	1.40	0.55
ripdisp	Local riparian disturbance index	1.10	0.43
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.06	0.02
CERC_den	Local density of Superfund sites	0.00	0.00

3.4 Mapped Results

3.4.1 *Expected current conditions*

Diversity index scores were predicted for all 1:100k stream catchments in the study area using the BRT model. The predicted probability values ranged from 0 to 100. The mean predicted value of the 225,541 total catchments was 3.75. These results are mapped in Figure 23.

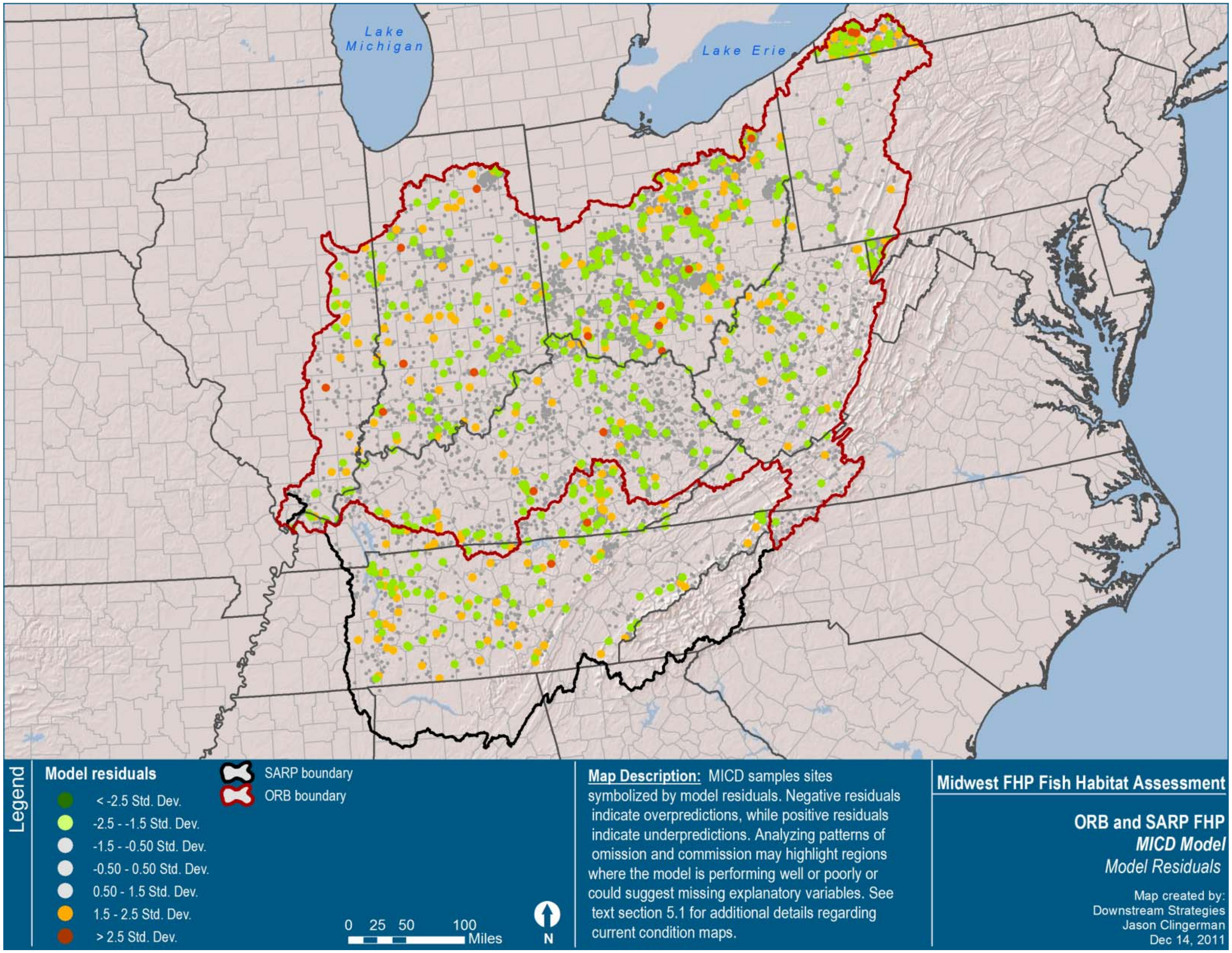
Figure 23: Expected diversity index score



3.4.2 ***Spatial variability in predictive performance***

Analyzing patterns of omission and commission may highlight regions where the model is performing well or poorly or could suggest missing explanatory variables (Figure 24). To assess omission and commission, residuals are also calculated by the BRT model. The residuals are a measure of the difference in the measured and modeled values (measured value *minus* modeled value). Negative residuals indicate overpredictions (predicting higher values than are true), while positive residuals indicate underpredictions (predicting lower values than are true).

Figure 24: Distribution of diversity index model residuals by sampling site



3.4.3 *Indices of stress and natural quality*

Maps of CNQI and CASI illustrate the spatial distribution of natural habitat potential (i.e., CNQI score) and anthropogenic stress (i.e., CASI score) in the ORB and SARP. CNQI and CASI scores are mapped in Figure 25 and Figure 26, respectively. The top five most influential variables toward the calculation of CNQI are shown in Figure 27-Figure 31. The top five variables contributing toward the calculation of CASI are mapped in Figure 32-Figure 36. CNQI, CASI, and their metrics are all scaled on a 0-100 scale (see Section 3.3 for more details on CNQI and CASI calculation). For CNQI, higher values indicate higher natural quality, while higher values for CASI indicate higher levels of anthropogenic stress.

Figure 25: Cumulative natural quality index for diversity index

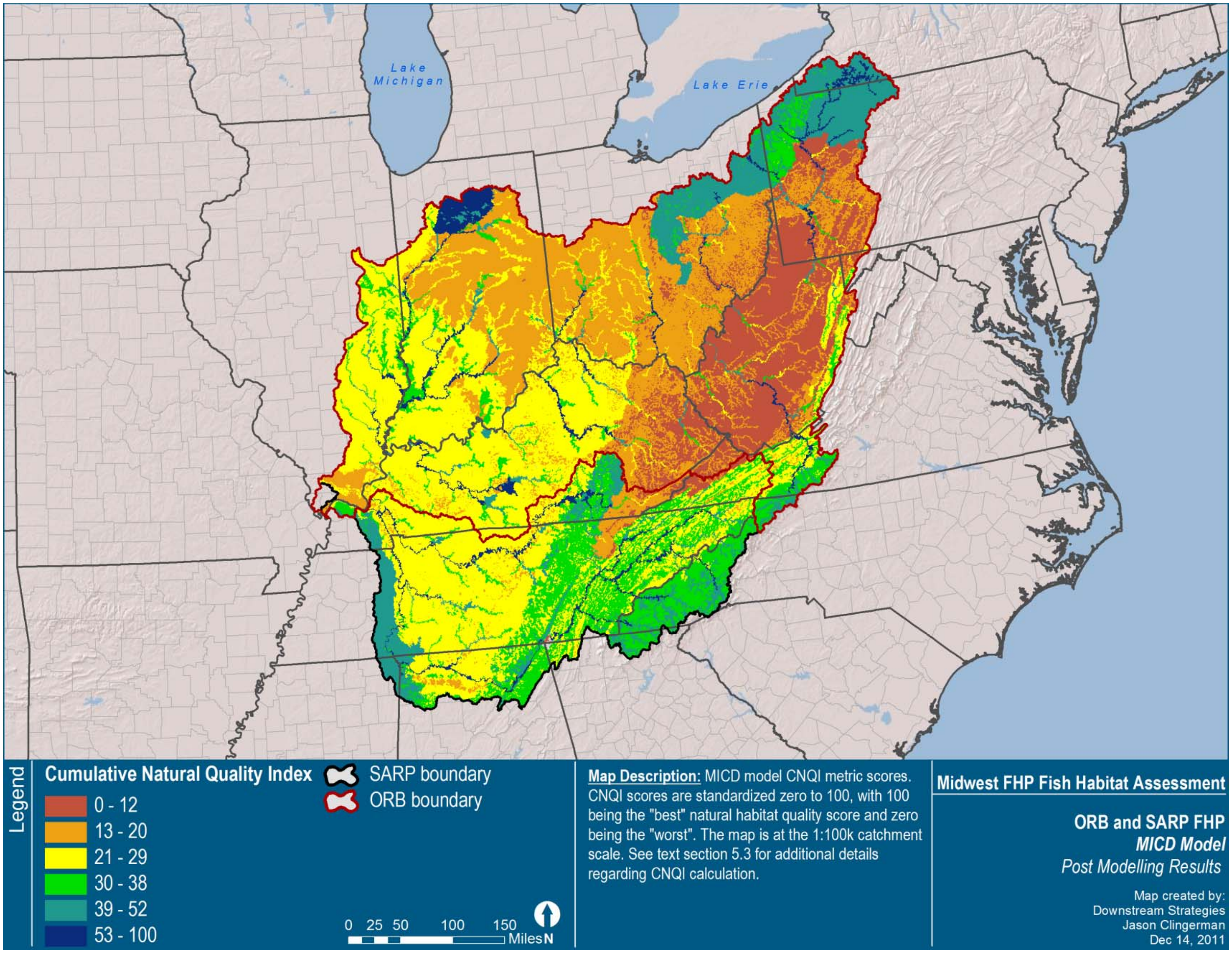


Figure 26: Cumulative anthropogenic stress index for diversity index

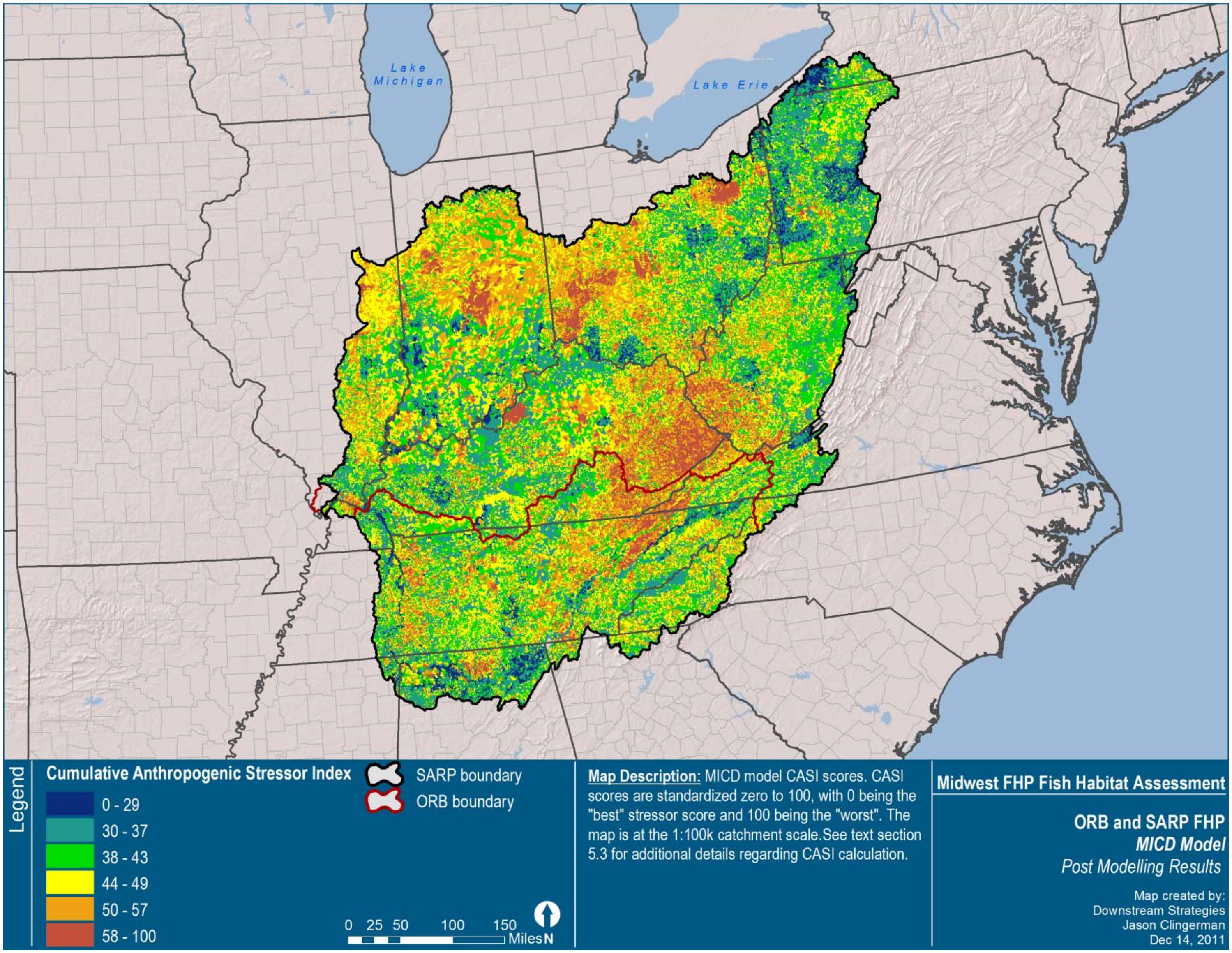


Figure 27: Most influential natural index metric for diversity index

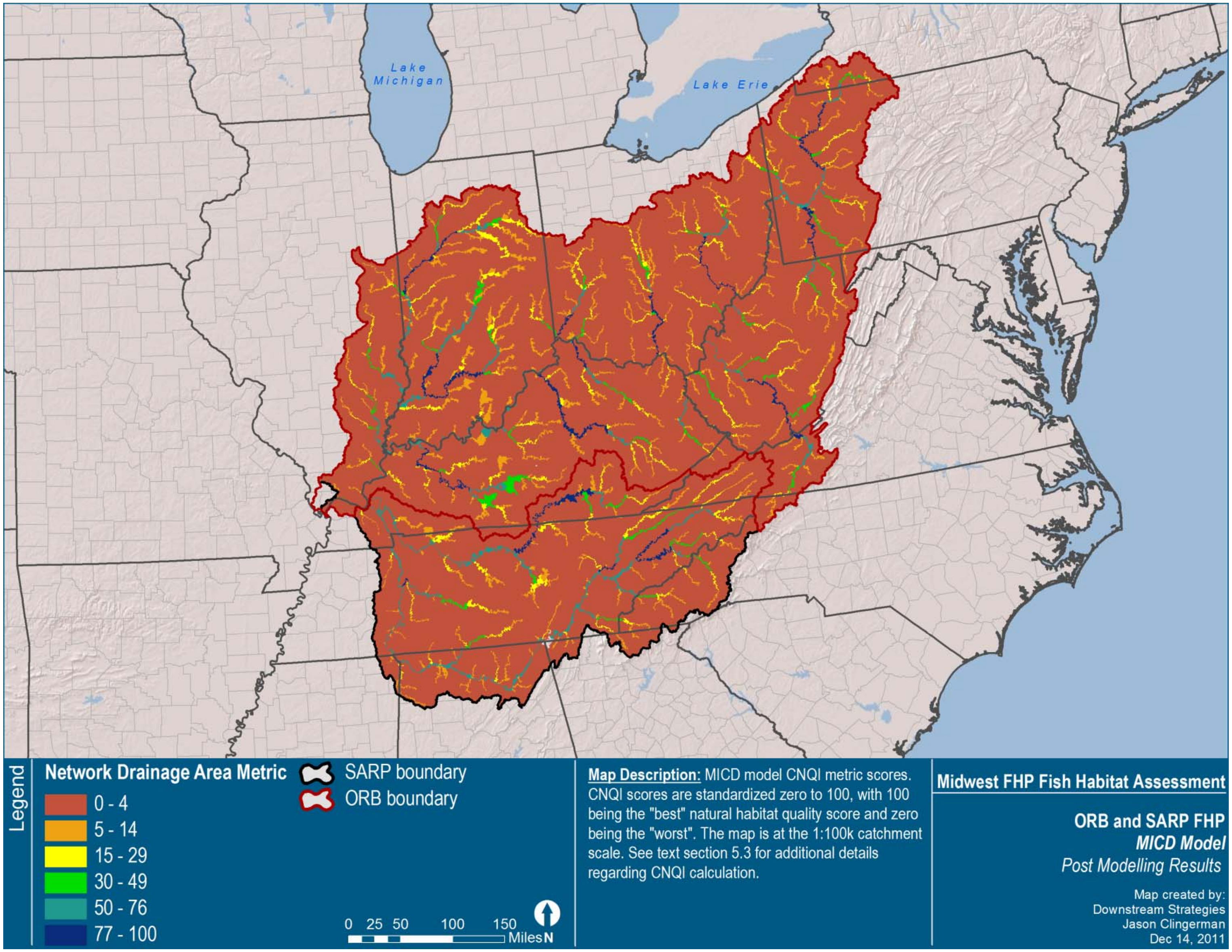


Figure 28: Second most influential natural index metric for diversity index

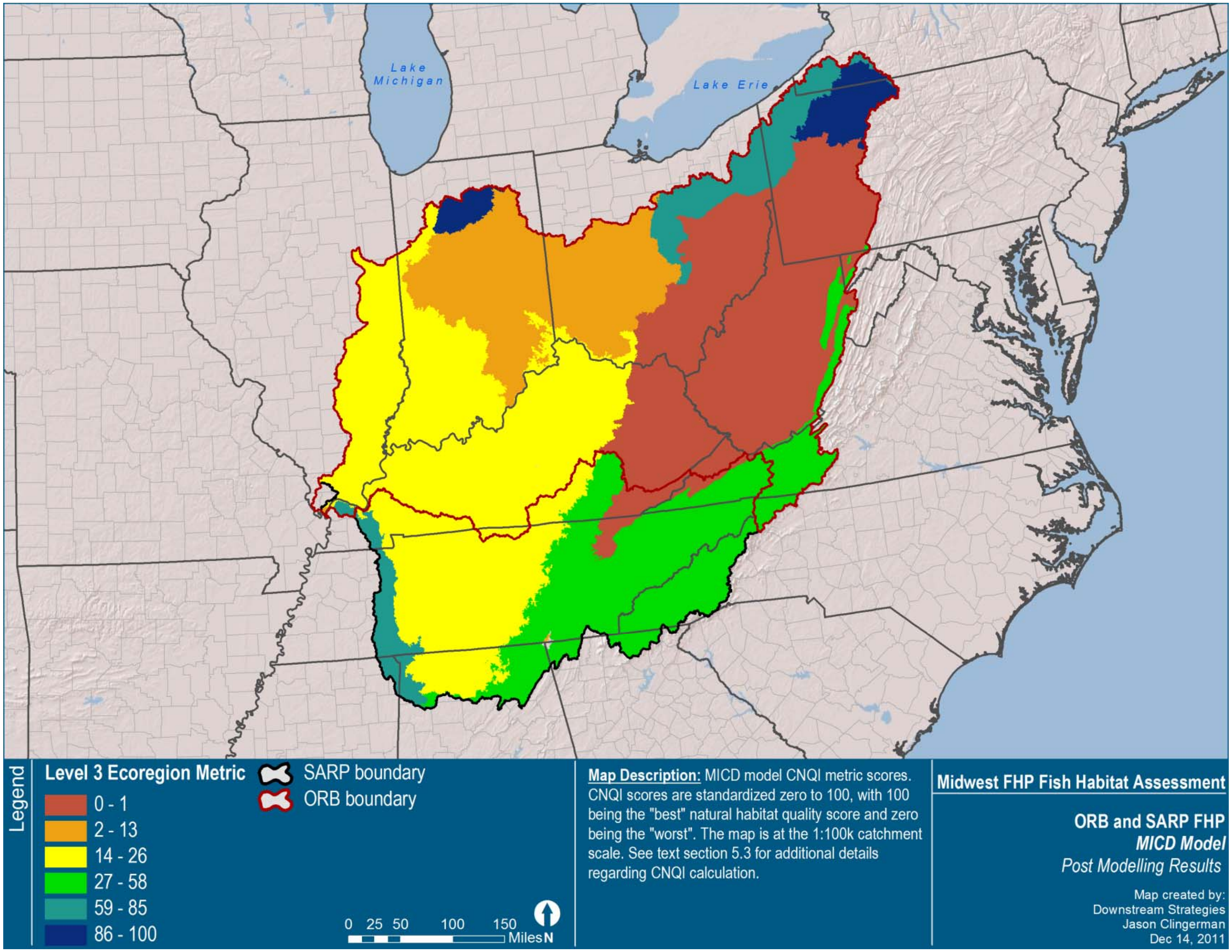


Figure 29: Third most influential natural index metric for diversity index

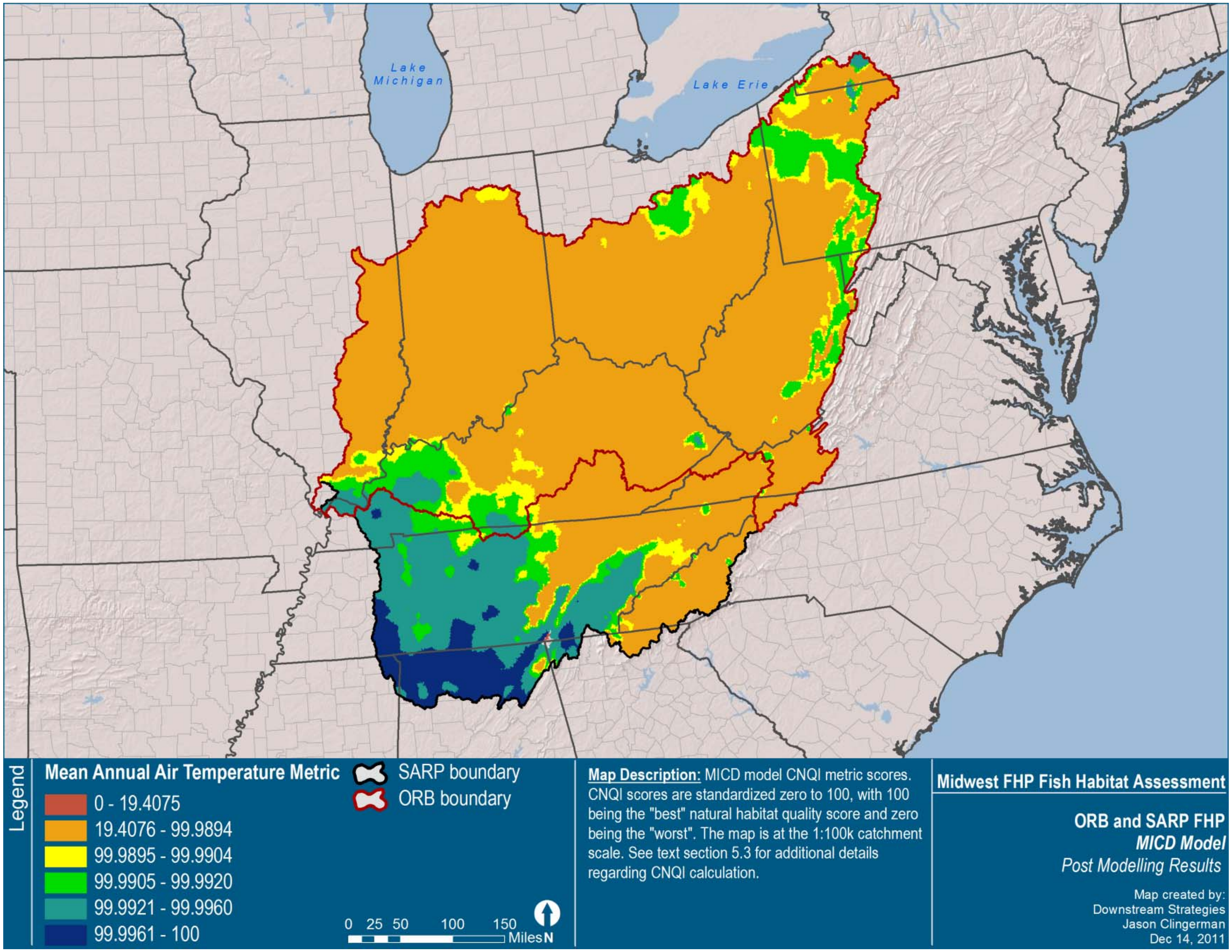


Figure 30: Fourth most influential natural index metric for diversity index

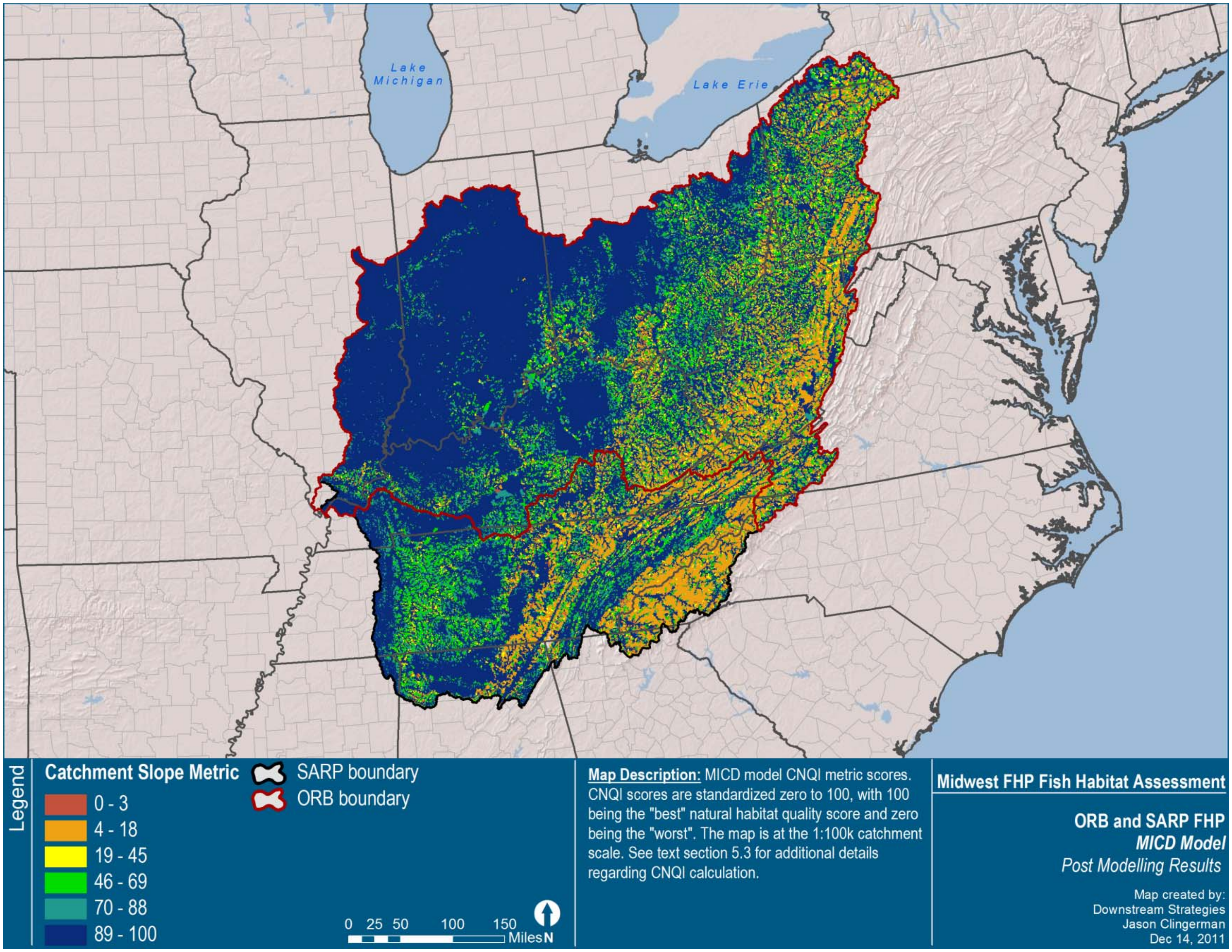


Figure 31: Fifth most influential natural index metric for diversity index

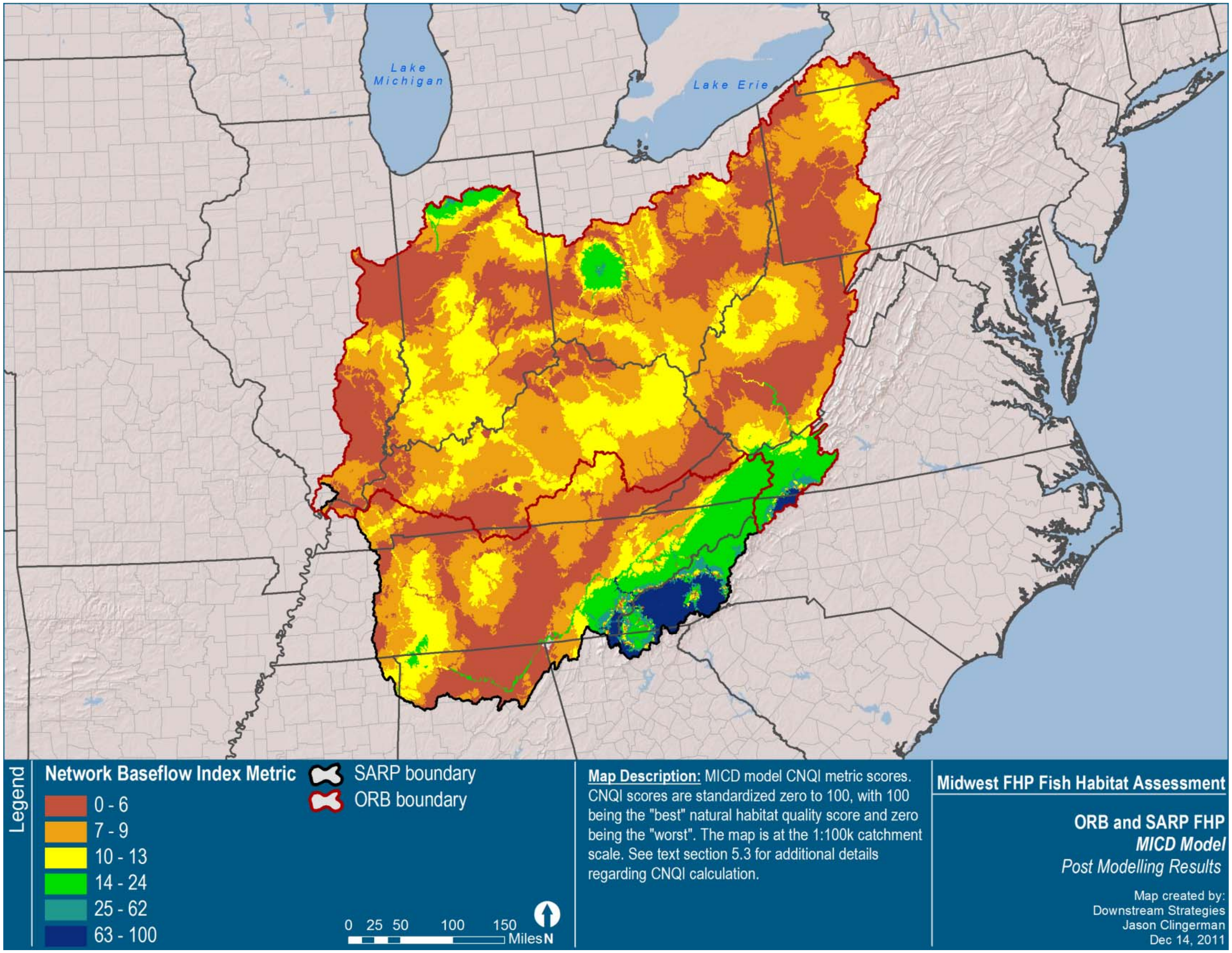


Figure 32: Most influential anthropogenic index metric for diversity index

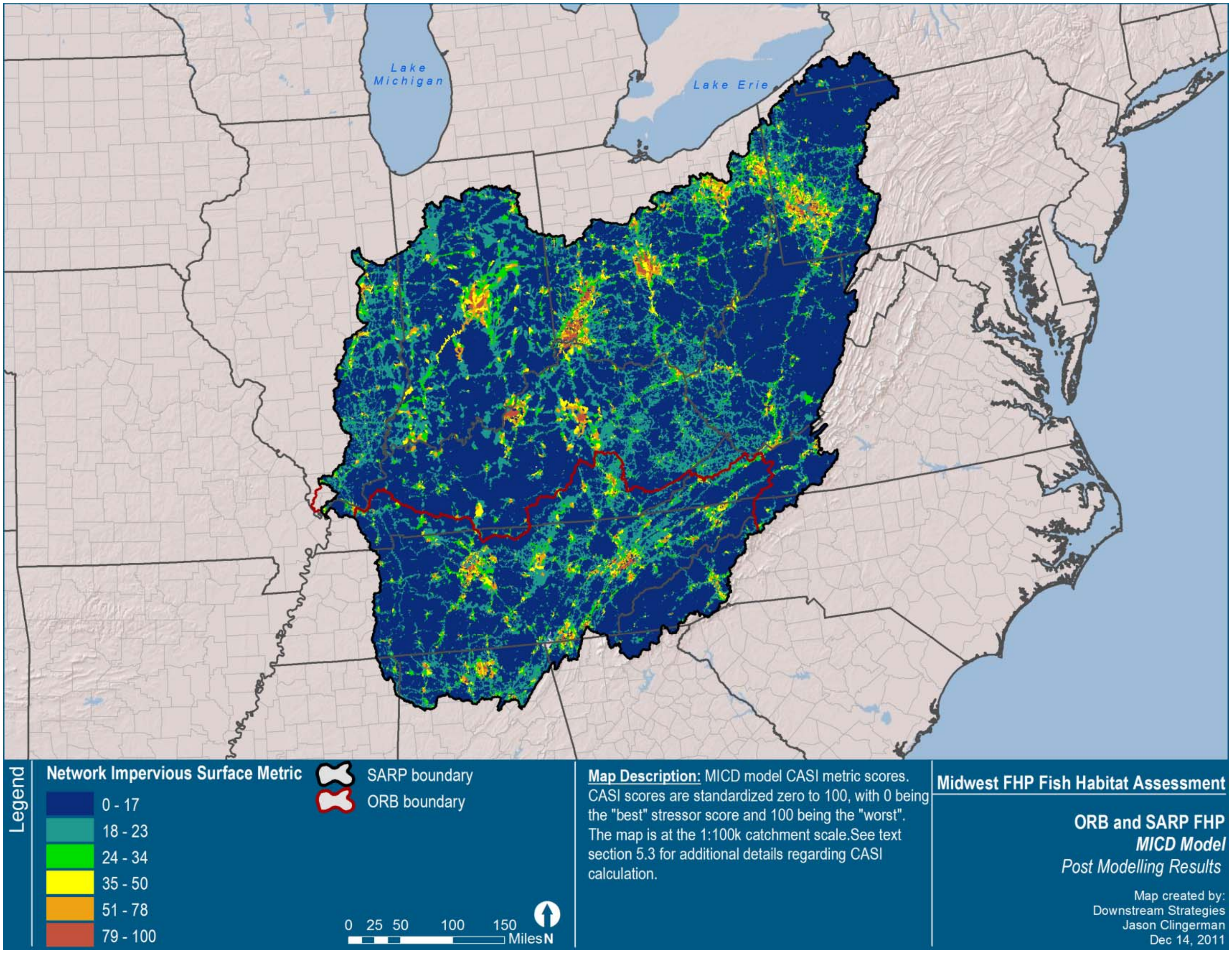


Figure 33: Second most influential anthropogenic index metric for diversity index

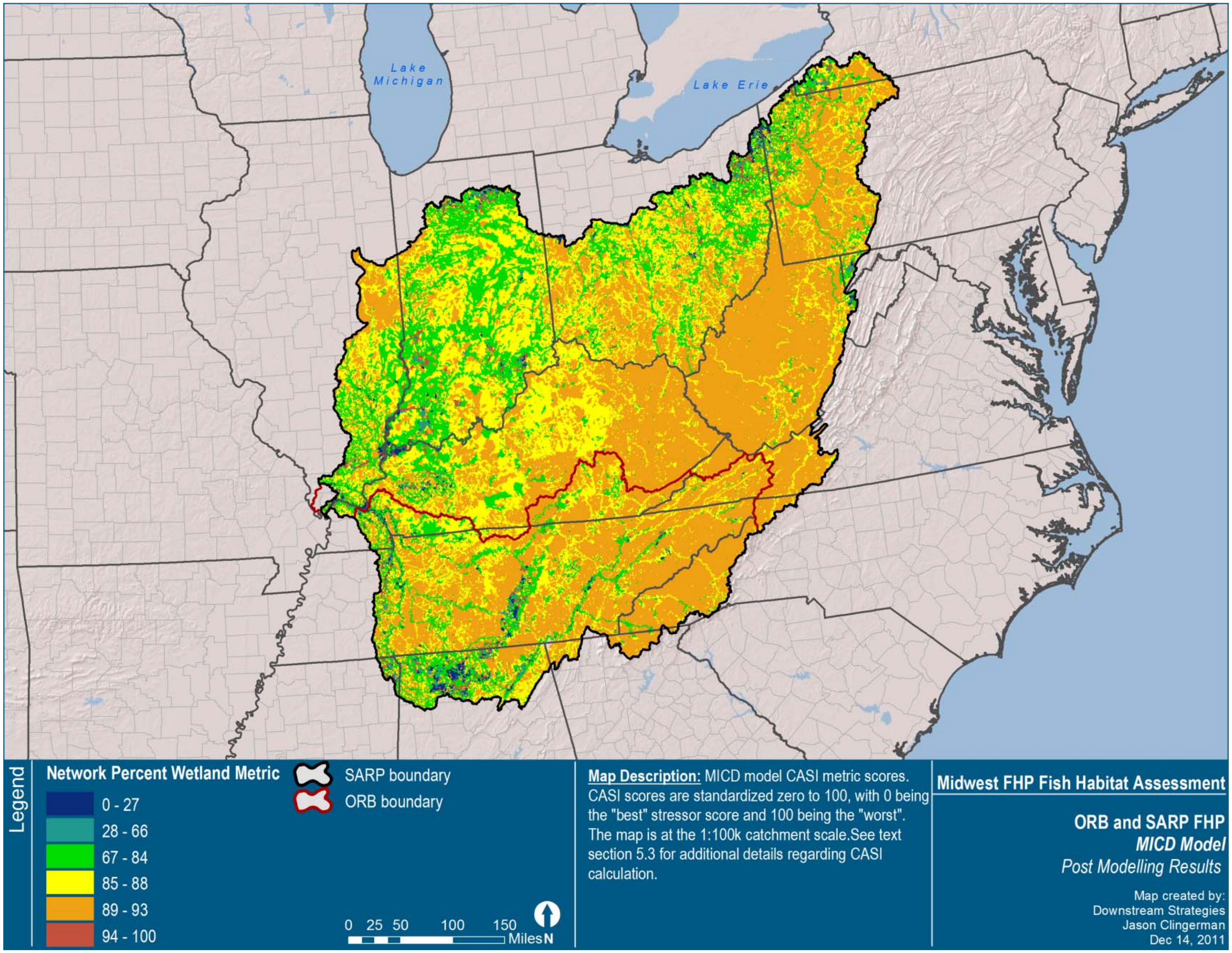


Figure 34: Third most influential anthropogenic index metric for diversity index

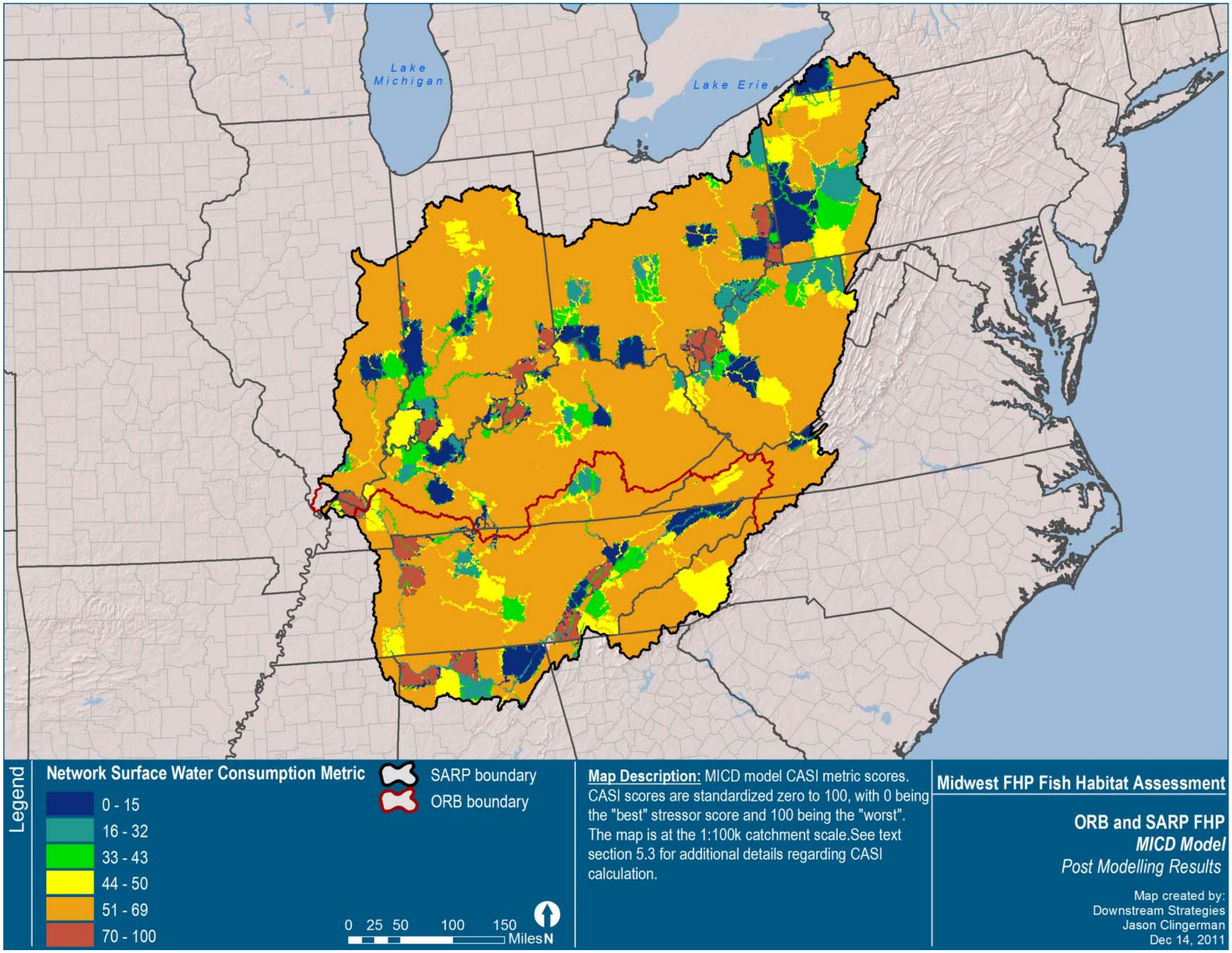


Figure 35: Fourth most influential anthropogenic index metric for diversity index

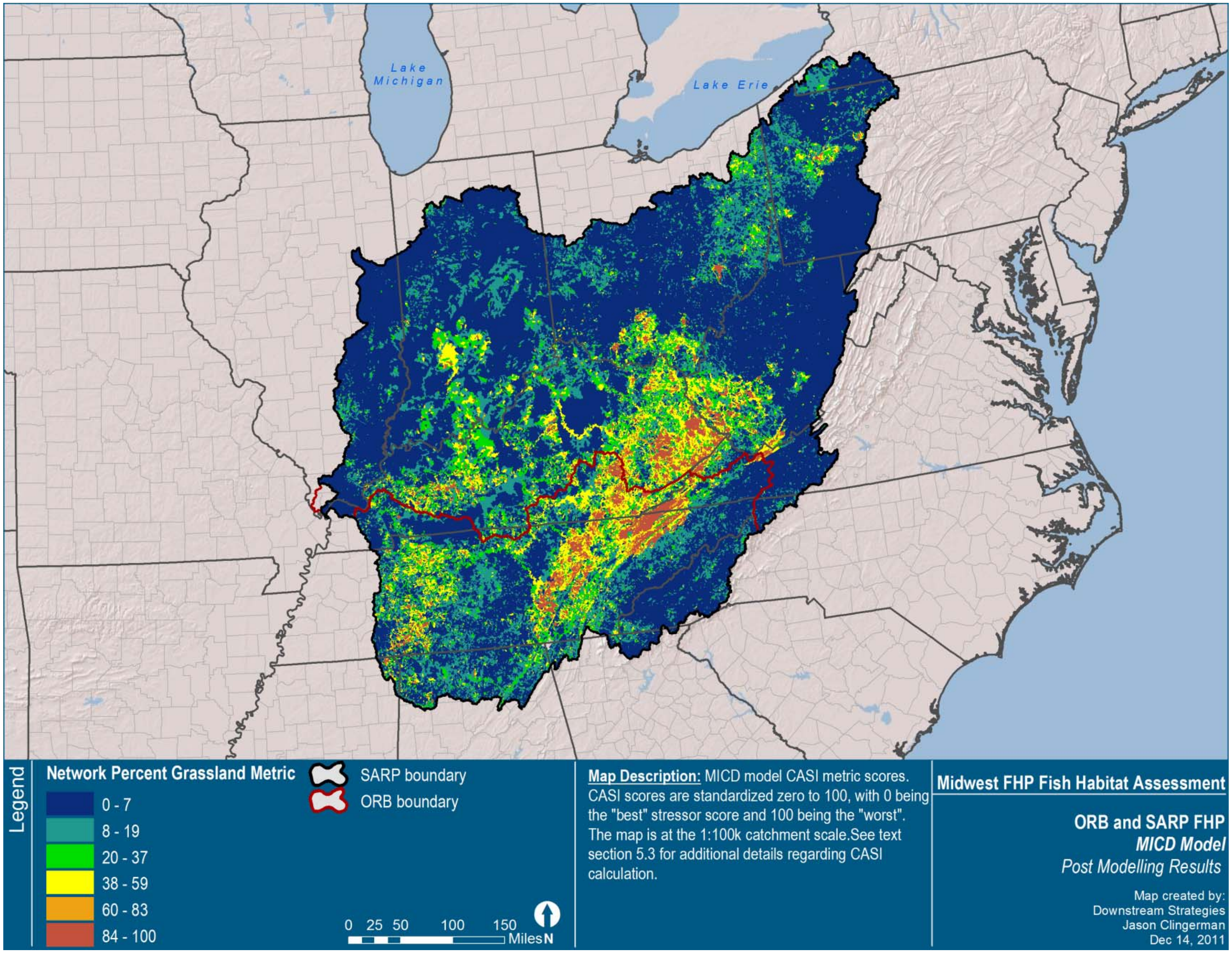
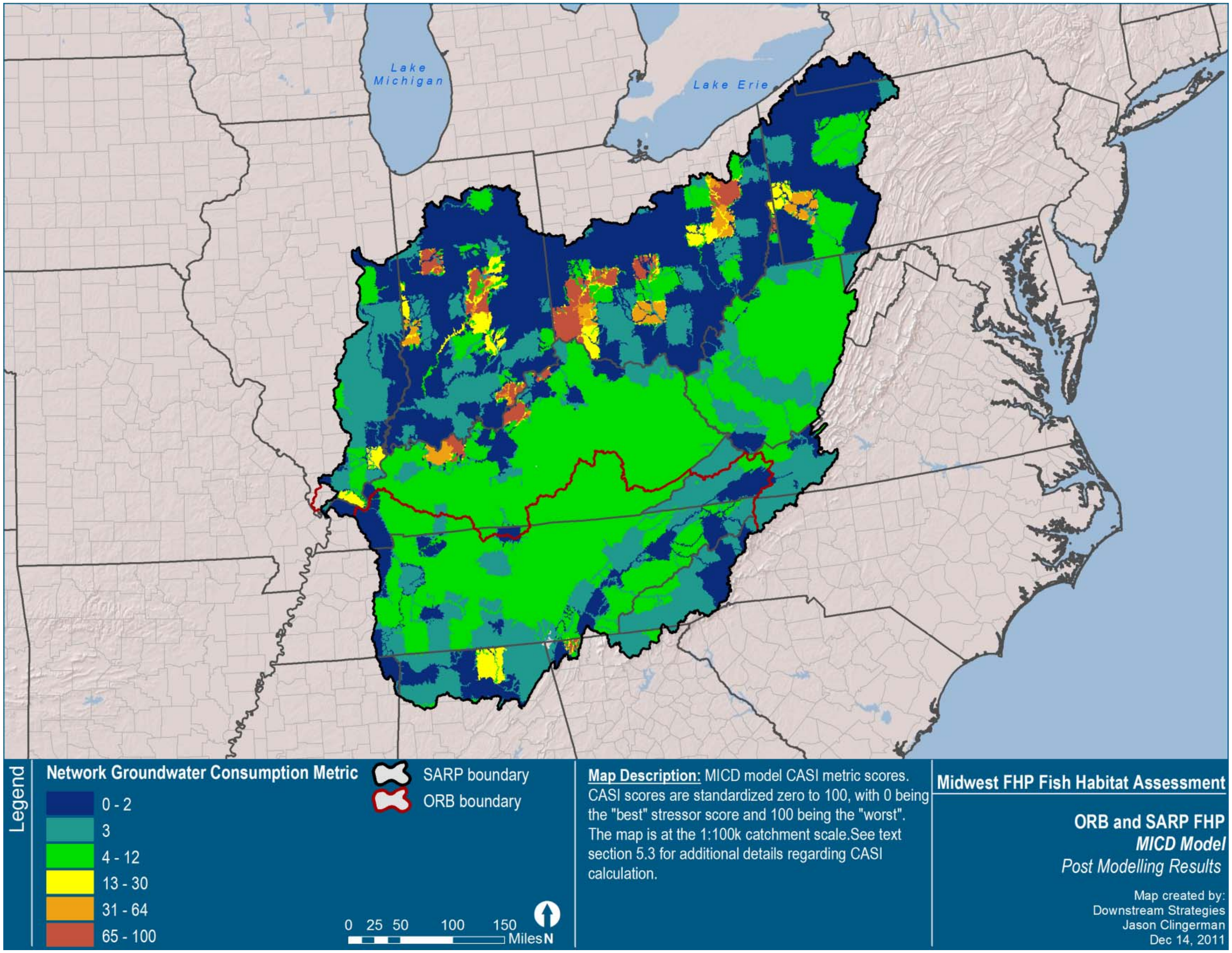


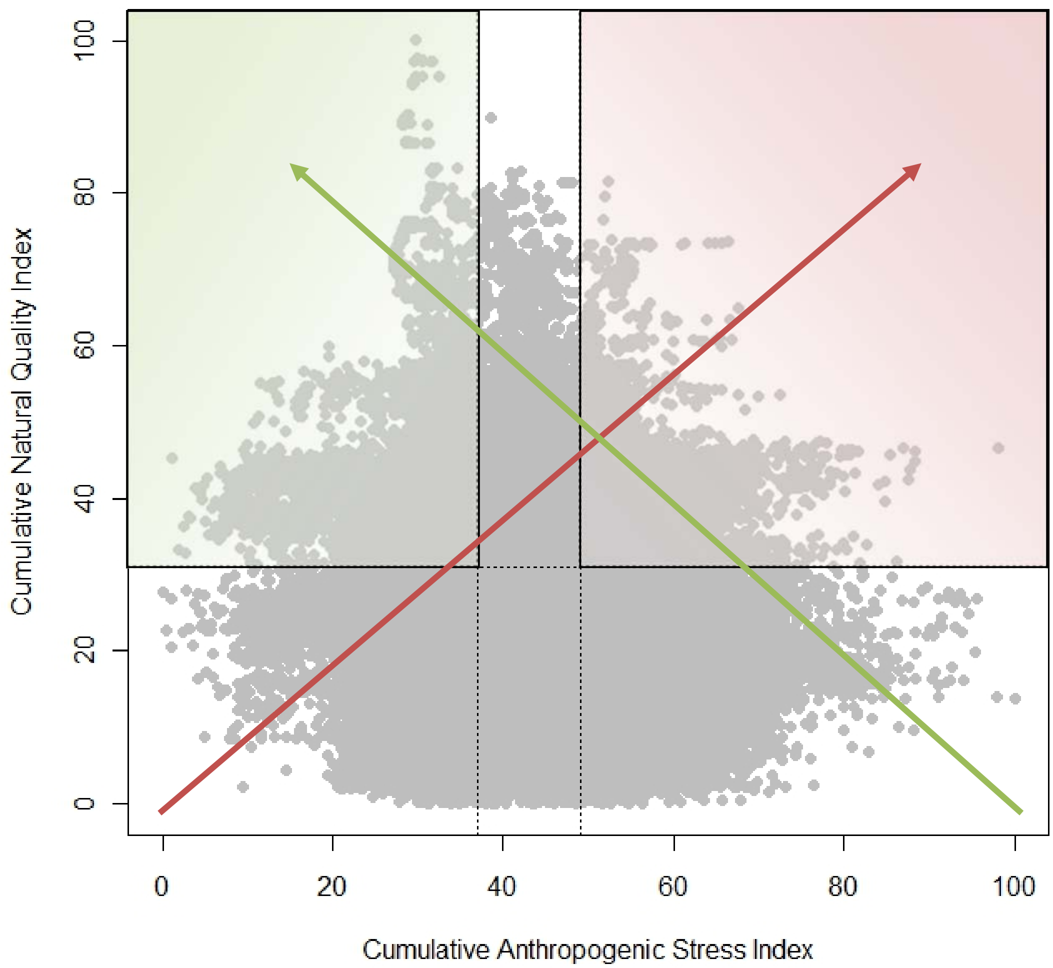
Figure 36: Fifth most influential anthropogenic index metric for diversity index



3.4.4 *Restoration and protection priorities*

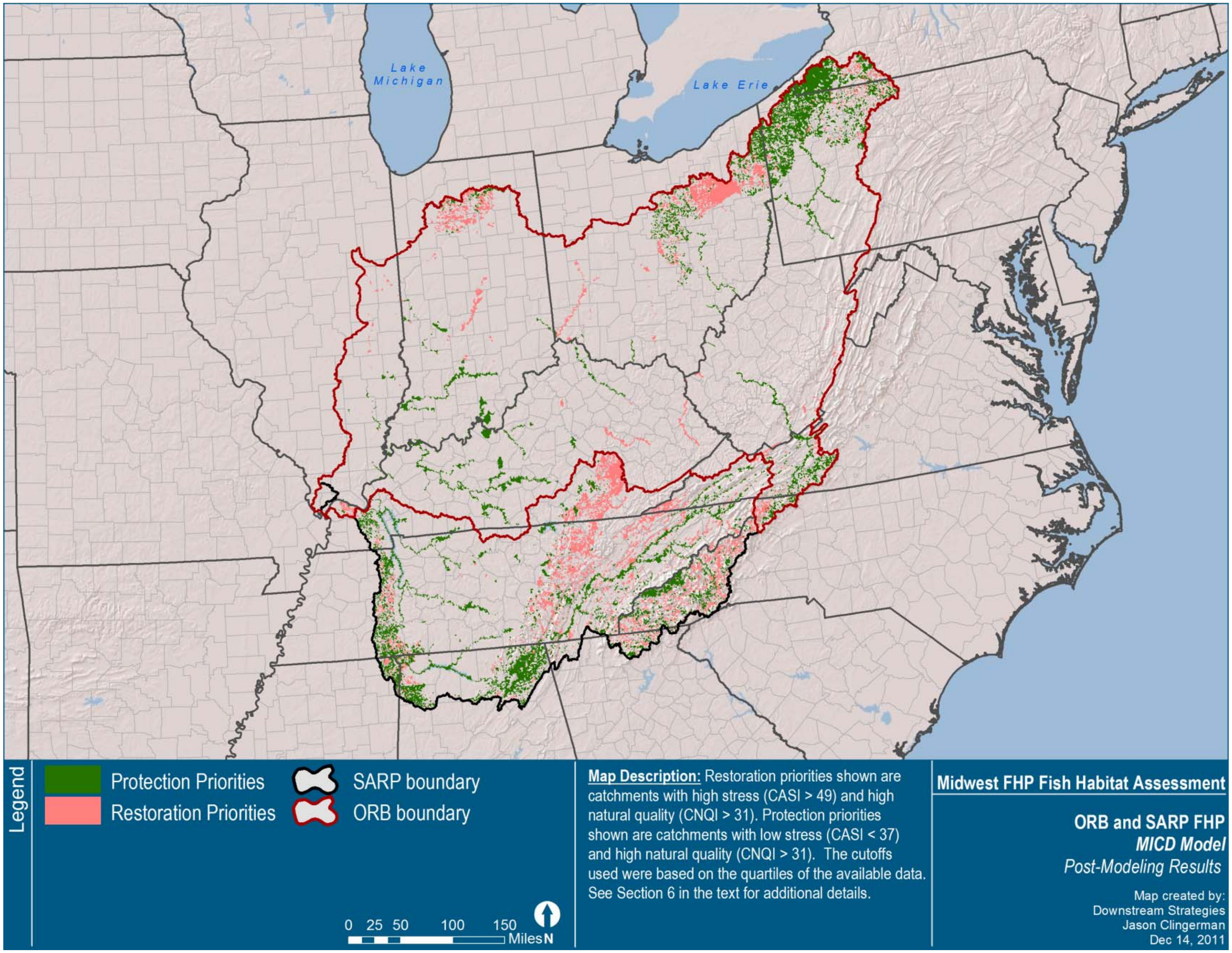
A plot of CNQI versus CASI values for all catchments in the study area (Figure 37) can be used as a reference when defining thresholds for categories of CNQI and CASI scores for use in the development of restoration and protection priorities. In the example shown (Figure 38), thresholds for restoration (high natural potential coupled with high anthropogenic stress) were set to CNQI greater than 31 (third quartile) and CASI greater than 49 (third quartile). The thresholds used for protection priorities (high natural potential and low anthropogenic stress) were CNQI greater than 31 and CASI less than 37 (first quartile).

Figure 37: CNQI versus CASI values for all catchments for diversity index



Note: Breakpoints for CNQI and CASI classes in this example are denoted by dashed lines. The arrows indicate the directions of increasing potential protection (green arrow) or restoration (red arrow) priority. The red box indicates catchments defined as restoration priorities under the example scenario. The green box indicates catchments defined as protection priorities under the same scenario.

Figure 38: Restoration and protection priorities for modified index of centers of diversity



4. SMALLMOUTH BASS

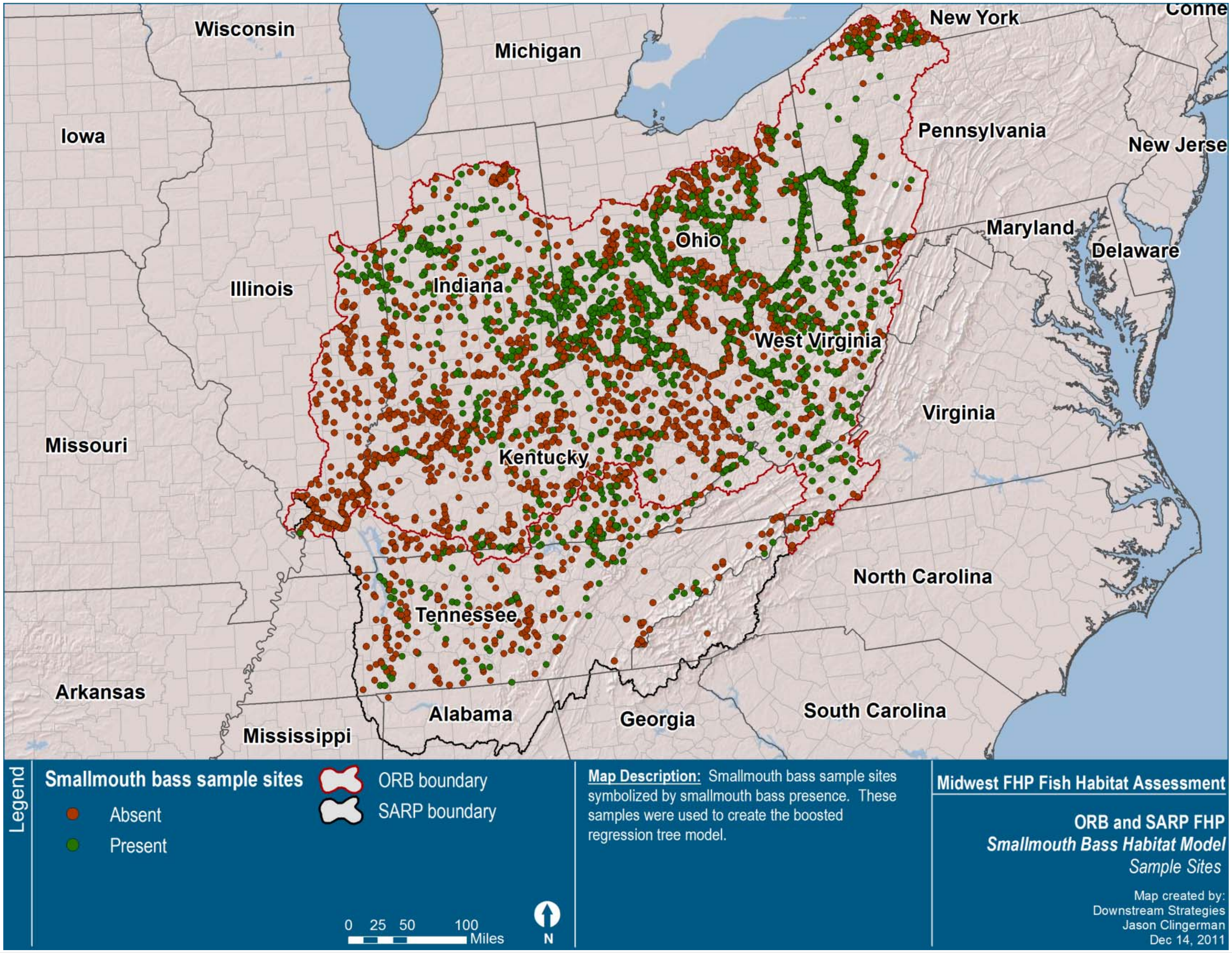
4.1 Modeling inputs

DS used a list of predictor variables selected by ORB and SARP to develop a ten-fold CV BRT model for smallmouth bass at the 1:100k catchment scale. The model was used to produce maps of expected current smallmouth bass distribution and maps of expected current natural habitat quality and anthropogenic stress at the 1:100k scale throughout the extents of both FHPs.

DS cooperated with ORB/SARP to arrive at a list of landscape-based habitat variables used to predict the presence of smallmouth bass throughout the region; ultimately, those variables were also used for characterizing habitat quality and anthropogenic stress. From an initial suite of 372 catchment attributes, DS and the FHPs compiled a list of 92 predictors for evaluation. From that list, 46 variables were removed due to statistical redundancy ($r > 0.6$) or logical redundancy, resulting in a final list of 46 predictor variables for the BRT model and assessment. See Appendix A for a full data dictionary and the metadata document for variable processing notes.

ORB and SARP provided DS with a presence-absence dataset for smallmouth bass comprised of 3,878 observations collected in catchments larger than 30 square kilometers in drainage area, but in segments with stream order less than nine. Catchments less than 30 square kilometers in drainage area and catchments with stream order greater than nine are known to be poor habitat for smallmouth bass. Samples were taken over a time frame spanning 1996 to 2010. Figure 39 maps all of the sampling sites that were used to construct the model and indicates the ORB and SARP boundaries to which the modeling outputs were applied.

Figure 39: Smallmouth bass modeling area and sampling sites



4.2 Modeling process

4.2.1 *Predictive performance*

The final selected model was comprised of 2,850 trees. The model had a CV correlation statistic of 0.598±0.016, and a CV ROC score of 0.846±0.008. The CV correlation statistic indicates the mean correlation resulting from each fold (ten in this case) of the cross-validation process.

4.2.2 *Variable influence*

The BRT output includes a list of the predictor variables used in the model ordered and scored by their relative importance. The relative importance values are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman and Meulman, 2003). The relative influence score is scaled so that the sum of the scores for all variables is 100, where higher numbers indicate greater influence. Of the 46 predictor variables used to develop the smallmouth bass model, 43 had a relative influence value greater than zero (Table 8). The five most influential predictors were all natural habitat variables and accounted for almost 53% of the total influence in the model:

- network drainage area,
- Level III Ecoregion,
- mean annual air temperature,
- network wetland land cover, and
- minimum catchment elevation.

The five most influential anthropogenic stressors, which accounted for over 11% of the total influence, were:

- network density of cattle,
- Local groundwater consumption,
- network surface water consumption,
- network forested land cover, and
- network density of dams.

Network drainage area, the single most important variable in terms of relative influence, contributed almost 18% of the total influence.

Table 8: Relative influence of all variables in the final smallmouth bass model

Variable code	Variable description	Relative influence
cumdrainag	Network drainage area	17.70
eco_code3	Level III Ecoregion	12.63
temp	Mean annual air temperature	11.93
wetlandpc	Network wetland land cover	5.33
minelevraw	Minimum catchment elevation	5.15
brock7pc	Network shale bedrock geology cover	3.38
slope	Slope of catchment flowline	3.18
cattlec	Network density of cattle	2.97
BFI_meanc	Network mean baseflow index	2.74
water_gw	Local groundwater consumption	2.63
water_swc	Network surface water consumption	2.51
forpc	Network forested land cover	2.34
damsc_den	Network density of dams	2.13
cropspc	Network rowcrop land cover	1.95
soil2pc	Network soil type B, B/D cover	1.95
grasspc	Network grassland land cover	1.86
brock1pc	Network carbonate bedrock geology cover	1.85
soil4pc	Network soil type D cover	1.84
pastpc	Network pasture land cover	1.78
roadcrc_den	Network density of road crossings	1.70
imp06c	Network impervious surface cover	1.70
grassp	Network grassland land cover	1.51
pastp	Local pasture land cover	1.44
roadcr_den	Local density of road crossings	1.36
imp06	Local Impervious surface cover	1.18
soil3pc	Network soil type C,C/D cover	1.03
surf7pc	Network clay surficial geology cover	0.81
soil4p	Local soil type D cover	0.64
surf3pc	Network alluvium surficial geology cover	0.58
soil1p	Local soil type A, A/D cover	0.49
brock3pc	Network mafic/igneous bedrock geology cover	0.34
soil1pc	Network soil type A, A/D cover	0.23
surf6pc	Network residuum surficial geology cover	0.21
dams_den	Local density of dams	0.20
surf2pc	Network Outwash surficial geology cover	0.18
brock5pc	Network Sand/Gravel bedrock geology cover	0.17
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.10
mines_den	Local density of mines	0.09
surf5pc	Network loess surficial geology cover	0.09
surf4p	Local lacustrine surficial geology cover	0.04
TRI_den	Local density of Toxic Release Inventory sites	0.04
brock4p	Local metamorphic bedrock geology cover	0.01
CERC_den	Local density of Superfund sites	0.01
surf2p	Local outwash surficial geology cover	0.00
brock2p	Local felsic/igneous bedrock geology cover	0.00
brock3p	Local mafic/igneous bedrock geology cover	0.00

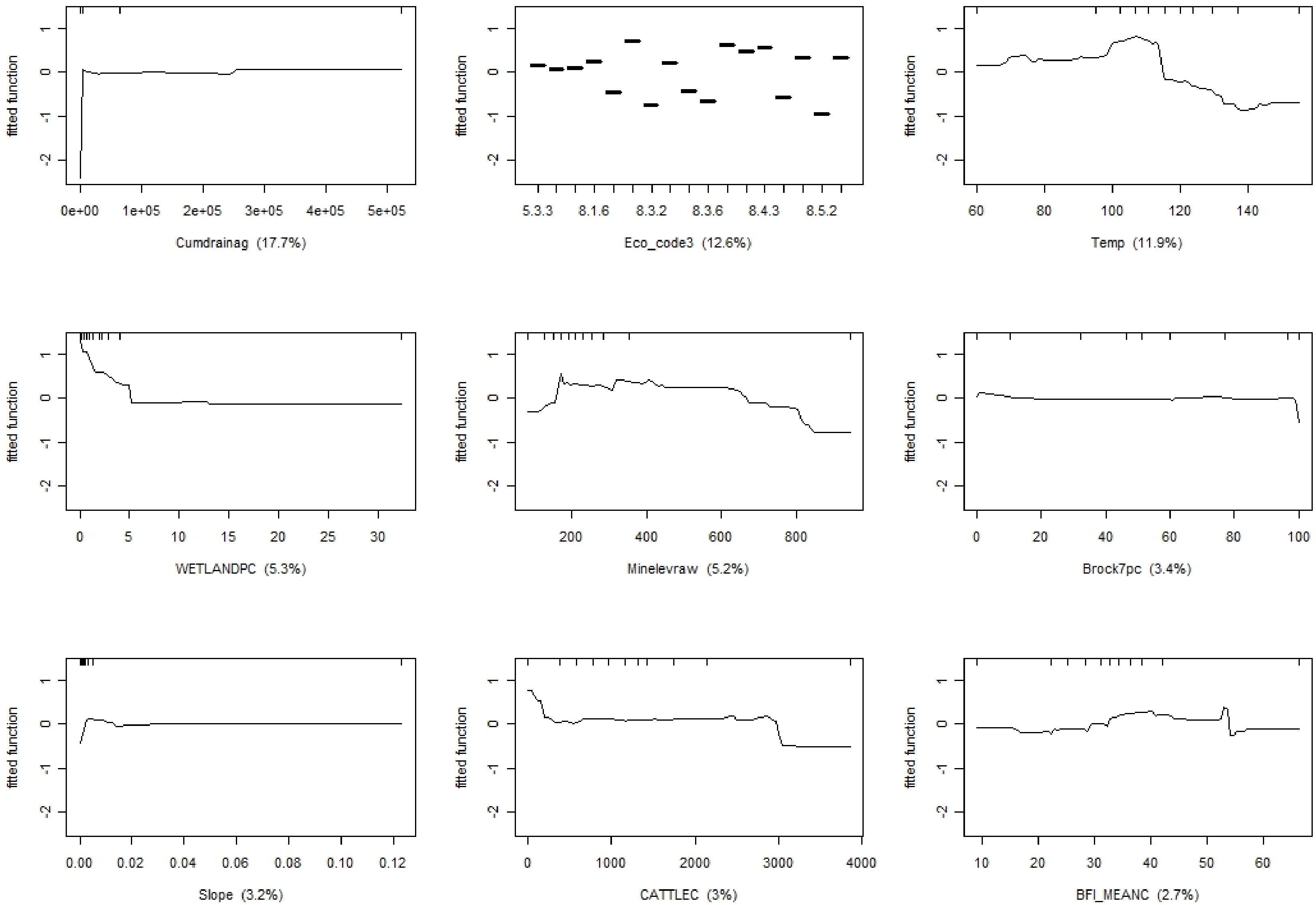
Note: Individual variables are highlighted according to whether they were determined to be anthropogenic in nature (red highlight) or natural (green highlight).

4.2.3 ***Variable functions***

The BRT output also contains quantitative information on partial dependence functions that can be plotted to visualize the effect of each individual predictor variable on the response after accounting for all other variables in the model. Similar to the interpretation of traditional regression coefficients, the function plots are not always a perfect representation of the relationship for each variable, particularly if interactions are strong or predictors are strongly correlated. However, they do provide a useful and objective basis for interpretation (Friedman, 2001; Friedman and Meulman, 2003).

These plots show the trend of the response variable (y-axis) as the predictor variable (x-axis) changes. The response variable is transformed (usually to the logit scale) so that the magnitude of trends for each predictor variable's function plot can be accurately compared. The dash marks at the top of each function represent the deciles of the data used to build the model. The function plots for the nine most influential variables in the smallmouth bass model (see Table 8 for reference) are illustrated in Figure 40 below. The plots for all 46 variables are shown in Appendix B.

Figure 40: Functional responses of the dependent variable to individual predictors of smallmouth bass



Note: Only the top nine predictors, based on relative influence (shown in parentheses; see Appendix A for descriptions of variable codes), are shown here. See Appendix B for plots of remaining predictor variables.

4.3 Post-modeling

The variable importance table and partial dependence functions of the final BRT model were used to create the post-modeling indices of natural habitat quality and anthropogenic stress for smallmouth bass. The CNQI was comprised of 24 variables with relative influence greater than zero that were classified as natural habitat features (Table 9). The CASI was comprised of 13 variables with relative influence greater than zero that were classified as anthropogenic habitat features (Table 10). To calculate the cumulative indices (i.e., CNQI and CASI), each of the individual natural or anthropogenic variables used in the two indices was converted to a metric by first applying the appropriate transformations, based on their function plots, and then rescaling the transformed measures to a 0 to 100 scale. To calculate the cumulative index from the individual metrics, the metrics were first multiplied by their appropriate weighting factors and then summed. The CNQI and CASI scores were a result of a rescaling of those weighted and summed metrics, again from 0 to 100.

4.3.1 *Variable weights*

Table 9 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CNQI. The five most influential factors in the CNQI were:

- network drainage area,
- Level III Ecoregion,
- mean annual air temperature,
- network wetland land cover, and
- minimum catchment elevation.

Table 10 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CASI. The five most influential factors in the CASI were:

- network density of cattle,
- network forested land cover,
- network density of dams,
- network rowcrop land cover, and
- network impervious surface cover.

Table 9: Relative influence and weights for natural variables on smallmouth bass

Variable code	Variable description	Relative influence	Weighting factor
cumdrainag	Network drainage area	17.70	1.00
eco_code3	Level III Ecoregion	12.63	0.71
temp	Mean annual air temperature	11.93	0.67
wetlandpc	Network wetland land cover	5.33	0.30
minelevraw	Minimum catchment elevation	5.15	0.29
brock7pc	Network shale bedrock geology cover	3.38	0.19
slope	Slope of catchment flowline	3.18	0.18
BFI_meanc	Network mean baseflow index	2.74	0.15
soil2pc	Network soil type B, B/D cover	1.95	0.11
brock1pc	Network carbonate bedrock geology cover	1.85	0.10
soil4pc	Network soil type D cover	1.84	0.10
soil3pc	Network soil type C, C/D cover	1.03	0.06
surf7pc	Network clay surficial geology cover	0.81	0.05
soil4p	Local soil type D cover	0.64	0.04
surf3pc	Network alluvium surficial geology cover	0.58	0.03
soil1p	Local soil type A, A/D cover	0.49	0.03
brock3pc	Network mafic/igneous bedrock geology cover	0.34	0.02
soil1pc	Network soil type A, A/D cover	0.23	0.01
surf6pc	Network residuum surficial geology cover	0.21	0.01
surf2pc	Network outwash surficial geology cover	0.18	0.01
brock5pc	Network sand/gravel bedrock geology cover	0.17	0.01
surf5pc	Network loess surficial geology cover	0.09	0.01
surf4p	Local lacustrine surficial geology cover	0.04	0.00
brock4p	Local metamorphic bedrock geology cover	0.01	0.00

Table 10: Relative influence and weights for anthropogenic variables on smallmouth bass

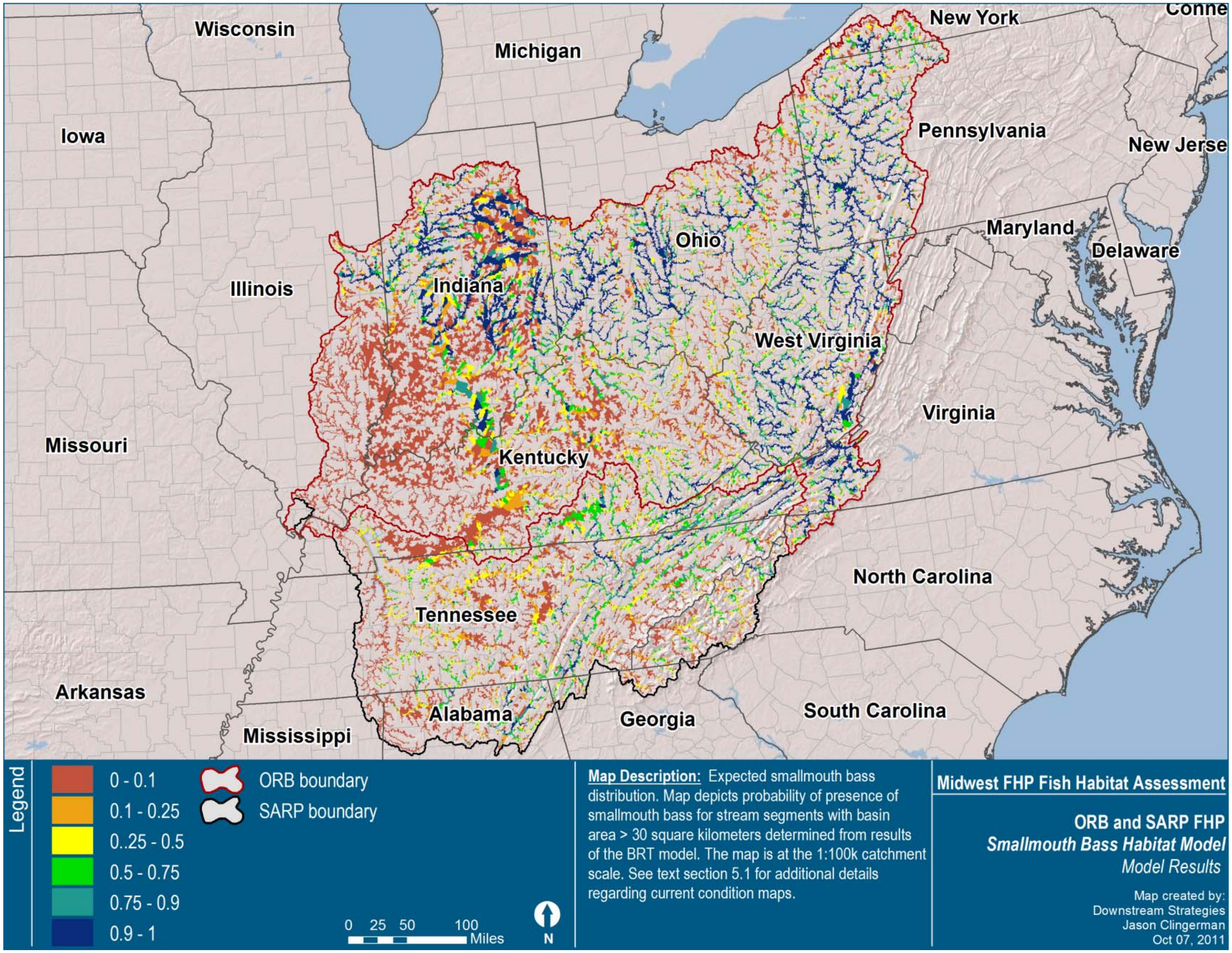
Variable code	Variable description	Relative influence	Weighting factor
cattlec	Network density of cattle	2.97	1.00
forpc	Network forested land cover	2.34	0.79
damsc_den	Network density of dams	2.13	0.72
cropspc	Network rowcrop land cover	1.95	0.66
imp06c	Network impervious surface cover	1.70	0.57
grassp	Network grassland land cover	1.51	0.51
pastp	Local pasture land cover	1.44	0.49
roadcr_den	Local density of road crossings	1.36	0.46
imp06	Local Impervious surface cover	1.18	0.40
dams_den	Local density of dams	0.20	0.07
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.10	0.03
mines_den	Local density of mines	0.09	0.03
CERC_den	Local density of Superfund sites	0.01	0.00

4.4 Mapped Results

4.4.1 *Expected current conditions*

Smallmouth bass probability of presence was calculated for all 1:100k stream catchments in the study area using the BRT model. The predicted probability values ranged from 0.002 to 1. The mean predicted probability value for the 56,370 total catchments larger than 30 square kilometers in drainage area was 0.402. There were 16,170 catchments larger than 30 square kilometers in drainage area with a predicted probability of presence greater than 0.75, and 6,527 catchments larger than 30 square kilometers in drainage area where the probability of presence was between 0.5 and 0.75. These results are mapped in Figure 41.

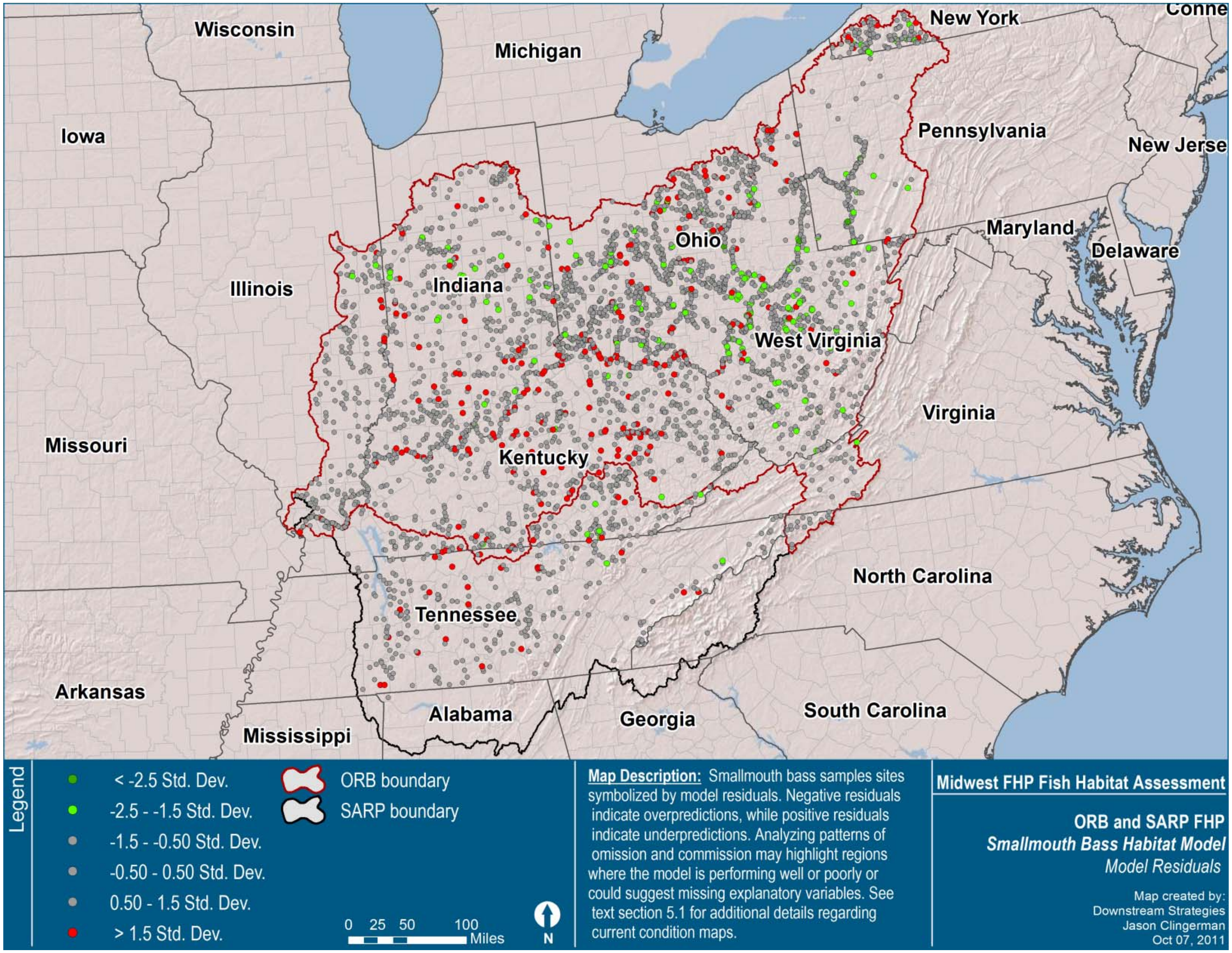
Figure 41: Expected smallmouth bass probability of presence



4.4.2 ***Spatial variability in predictive performance***

Analyzing patterns of omission and commission may highlight regions where the model is performing well or poorly or could suggest missing explanatory variables (Figure 42). To assess omission and commission, residuals are also calculated by the BRT model. The residuals are a measure of the difference in the measured and modeled values (measured value *minus* modeled value). Negative residuals indicate overpredictions (predicting higher values than are true), while positive residuals indicate underpredictions (predicting lower values than are true).

Figure 42: Distribution of smallmouth bass model residuals by sampling site



4.4.3 *Indices of stress and natural quality*

Maps of CNQI and CASI illustrate the spatial distribution of natural habitat potential (i.e., CNQI score) and anthropogenic stress (i.e., CASI score) in ORB and SARP. CNQI and CASI scores are mapped in Figure 43 and Figure 44, respectively. The top five most influential variables toward the calculation of CNQI are shown in Figure 45-Figure 49. The top four variables contributing toward the calculation of CASI are mapped in Figure 50-Figure 54. CNQI, CASI, and their metrics are all scaled on a 0-100 scale (see Section 4.3 for more details on CNQI and CASI calculation). For CNQI, higher values indicate higher natural quality, while higher values for CASI indicate higher levels of anthropogenic stress.

Figure 43: Cumulative natural quality index for smallmouth bass

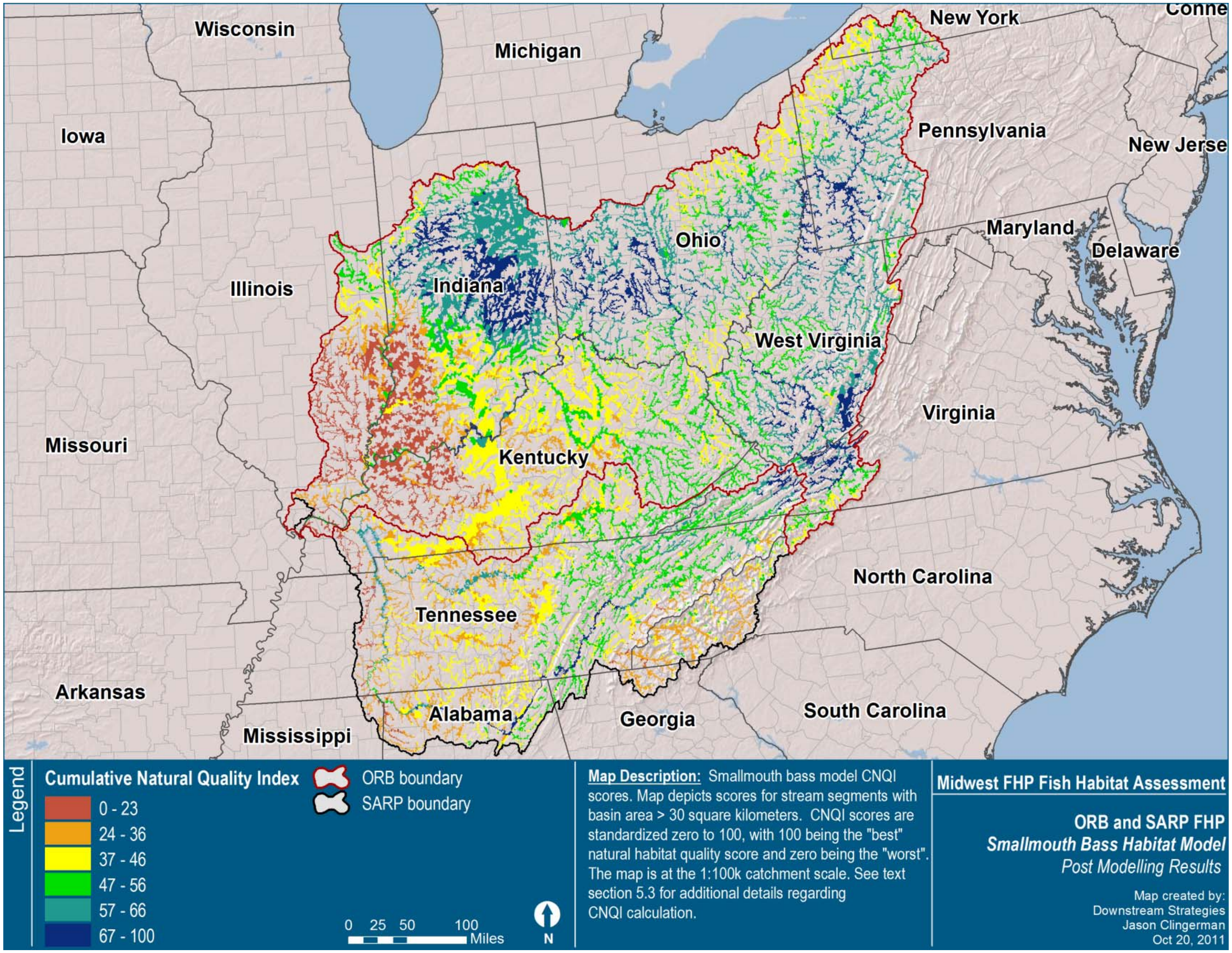


Figure 44: Cumulative anthropogenic stress index for smallmouth bass

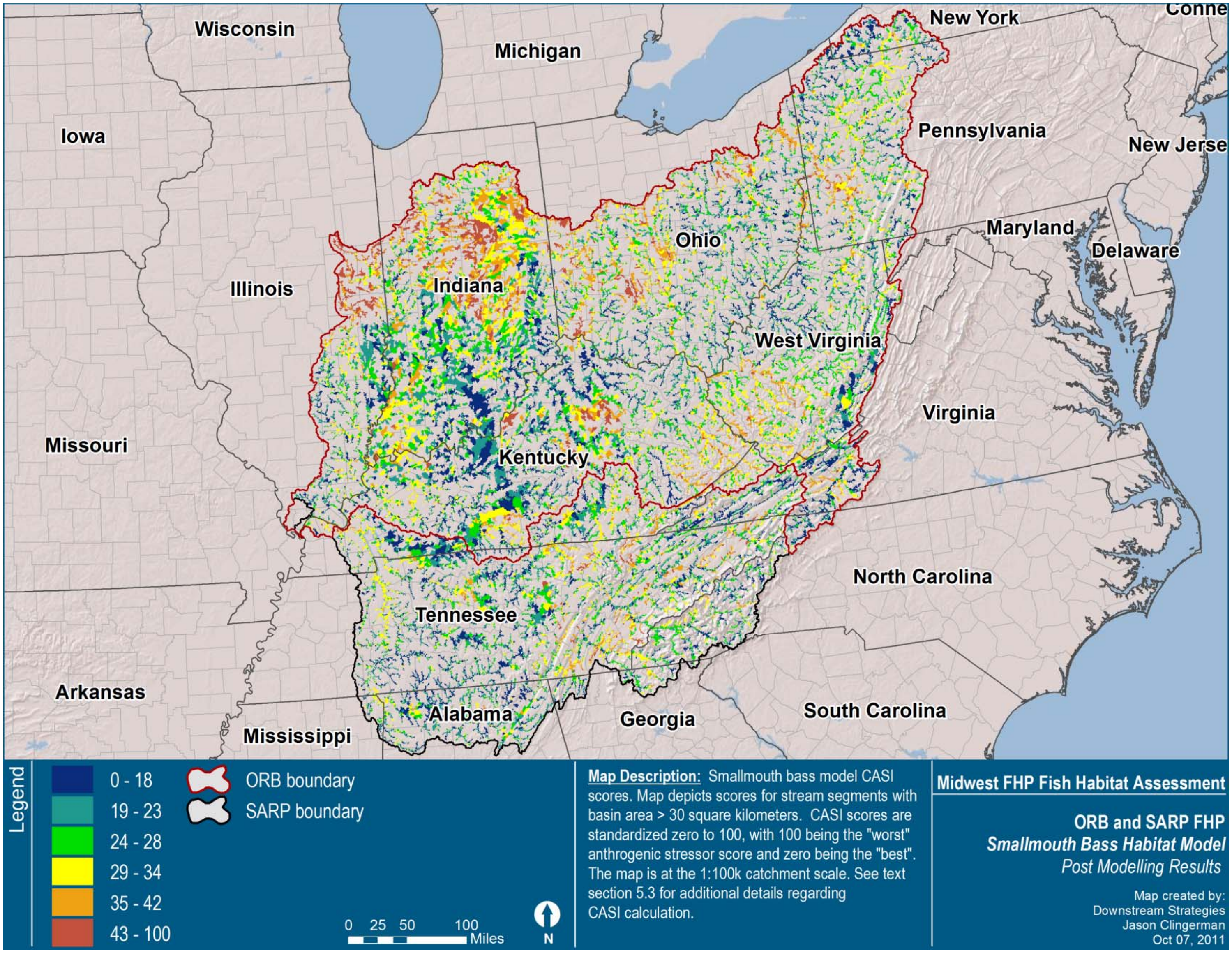


Figure 45: Most influential natural index metric for smallmouth bass

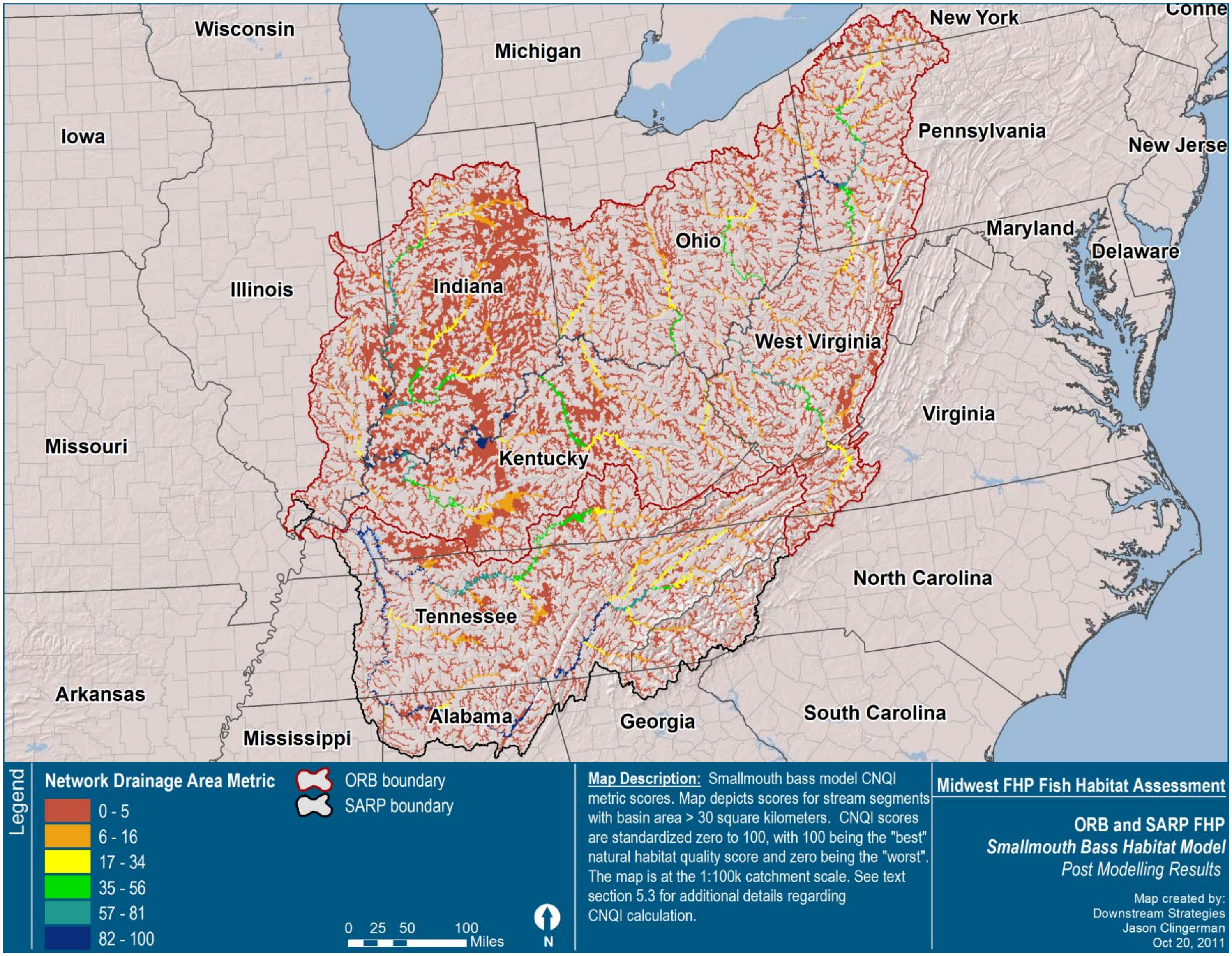


Figure 46: Second most influential natural index metric for smallmouth bass

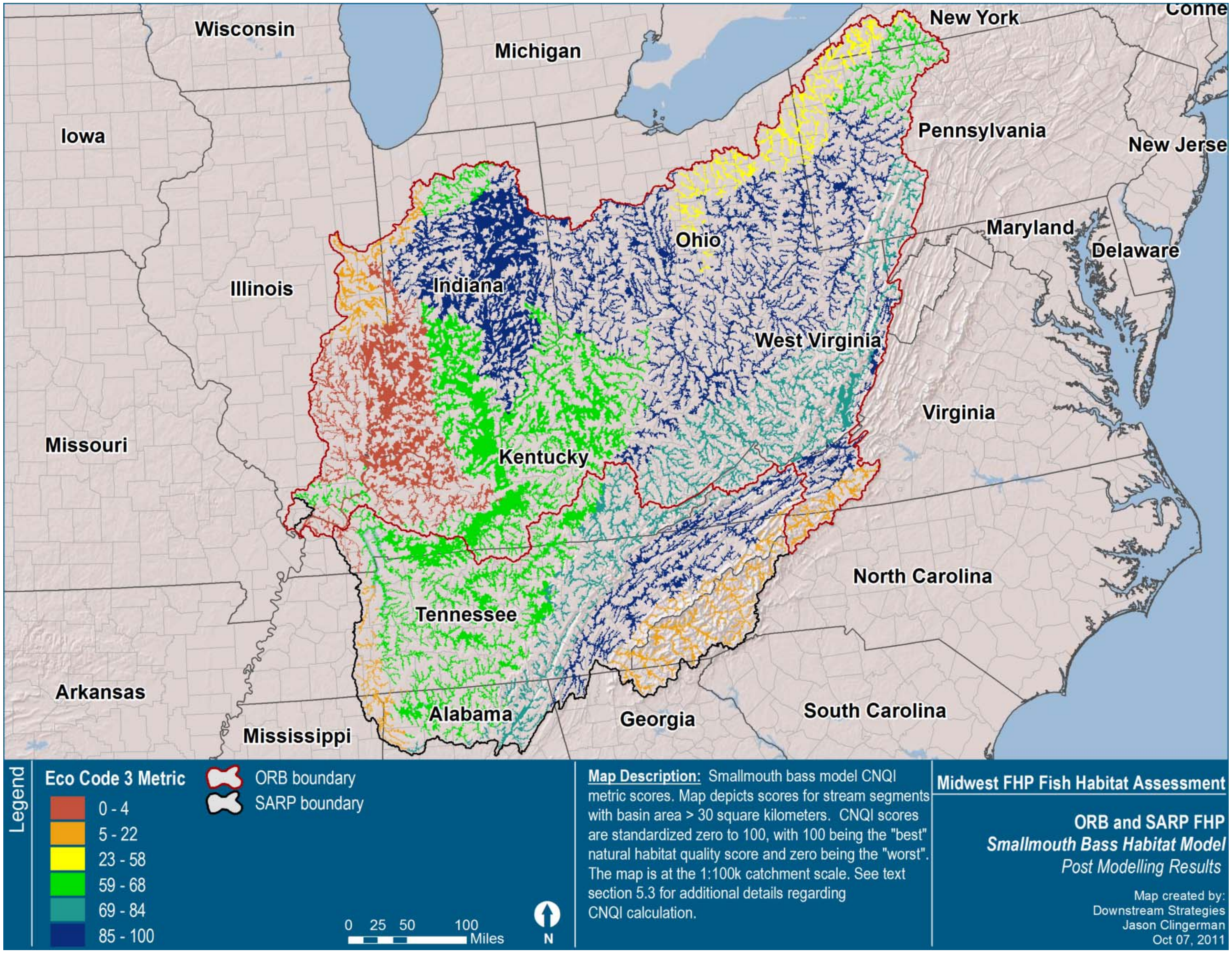


Figure 47: Third most influential natural index metric for smallmouth bass

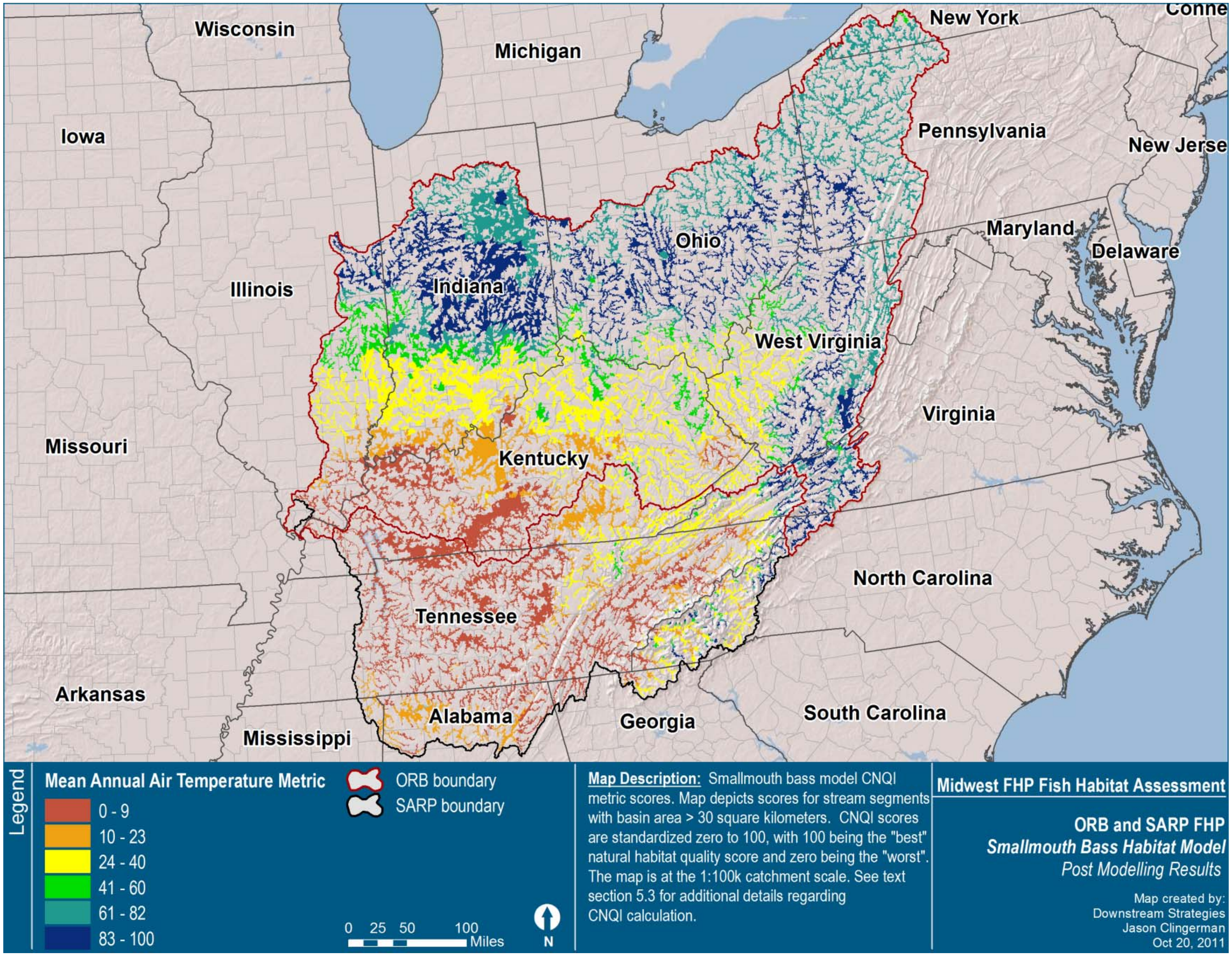


Figure 48: Fourth most influential natural index metric for smallmouth bass

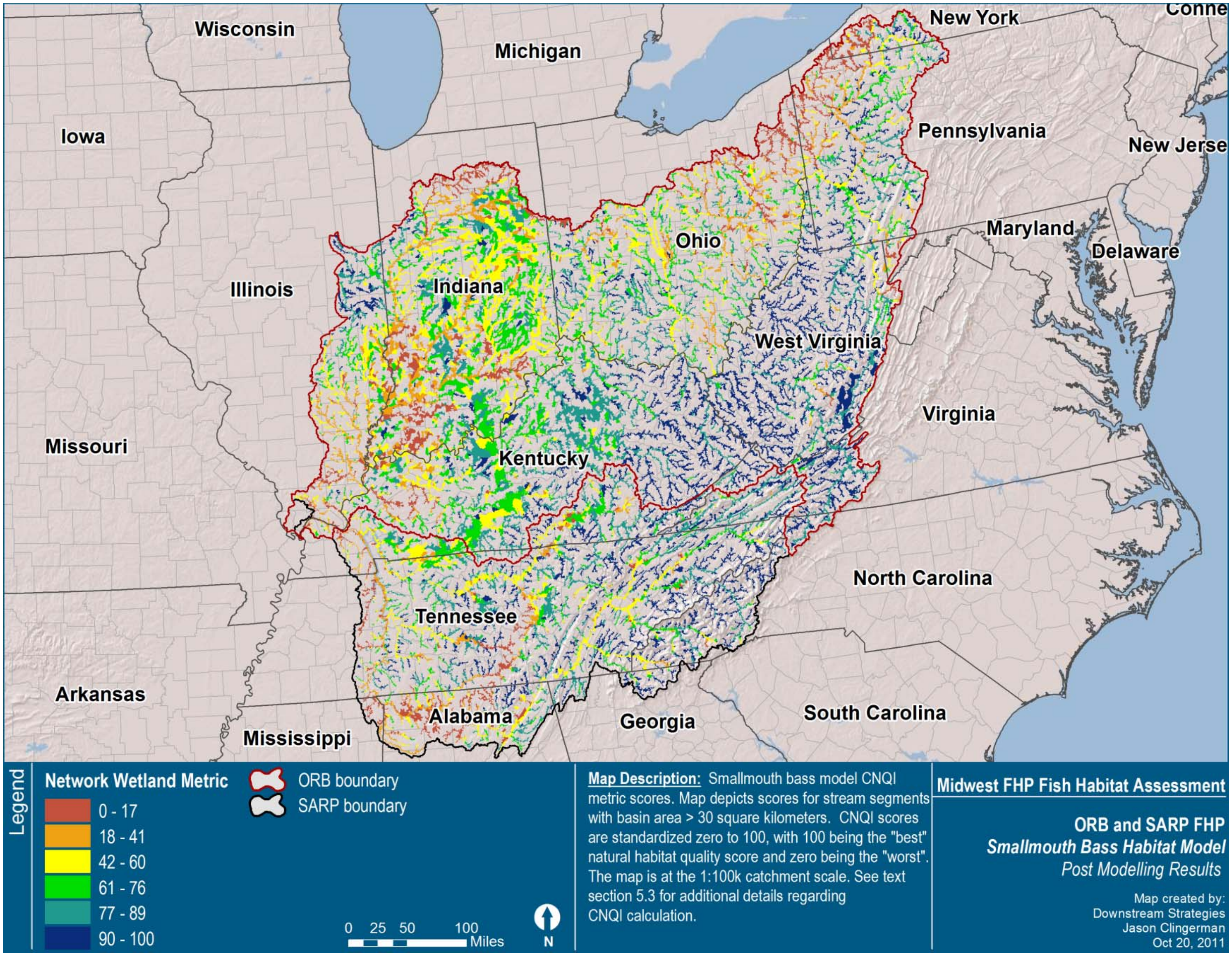


Figure 49: Fifth most influential natural index metric for smallmouth bass

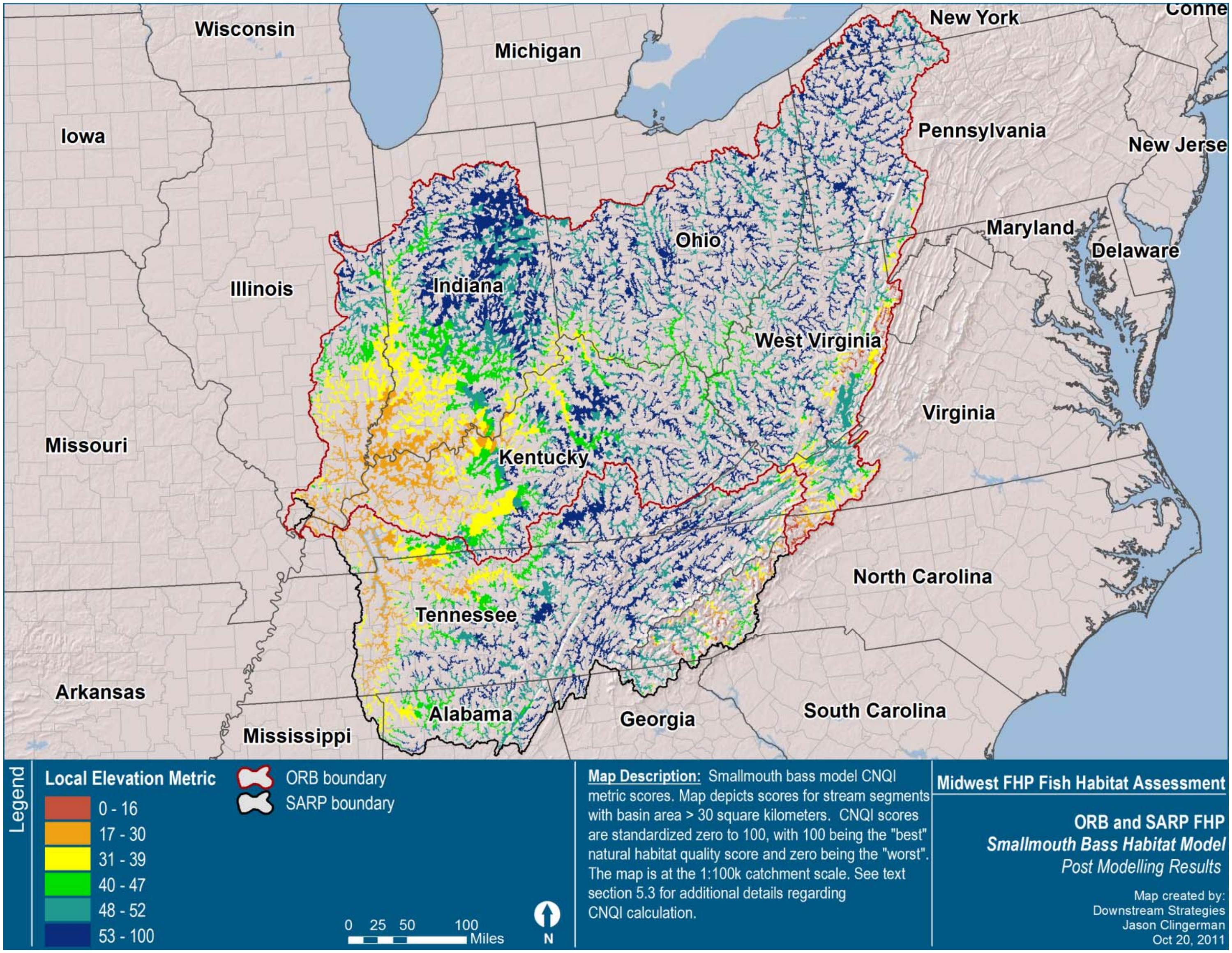


Figure 50: Most influential anthropogenic index metric for smallmouth bass

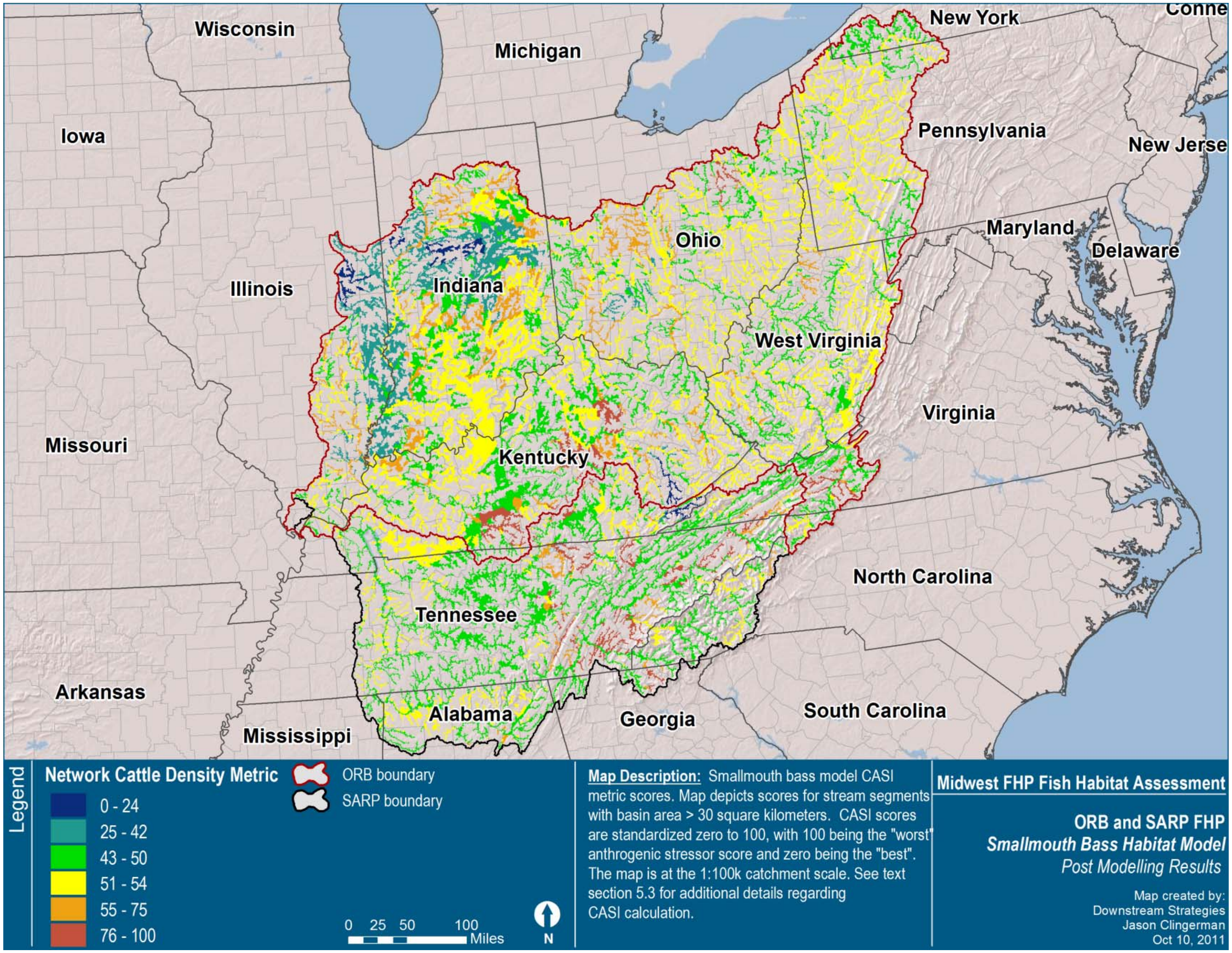


Figure 51: Second most influential anthropogenic index metric for smallmouth bass

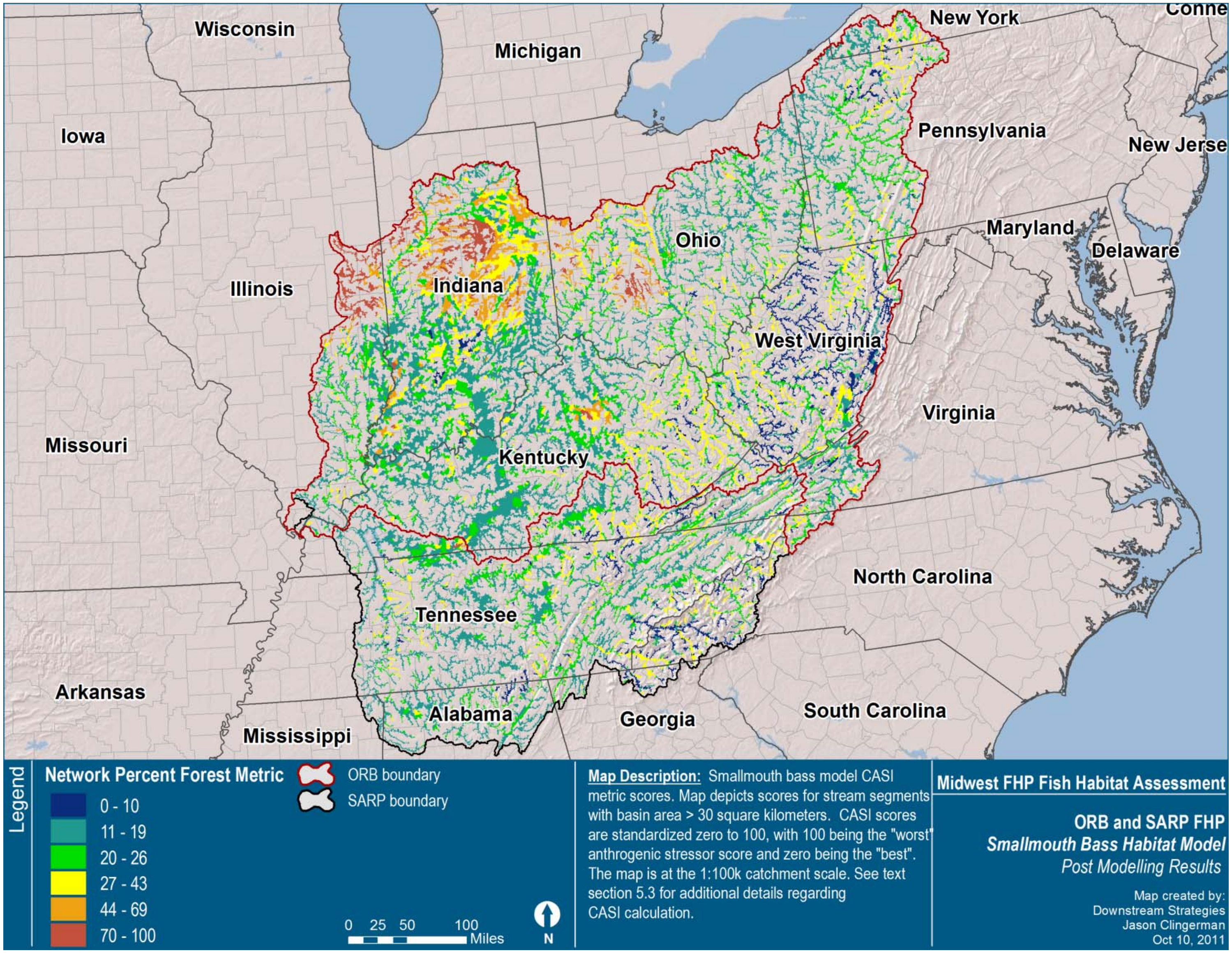


Figure 52: Third most influential anthropogenic index metric for smallmouth bass

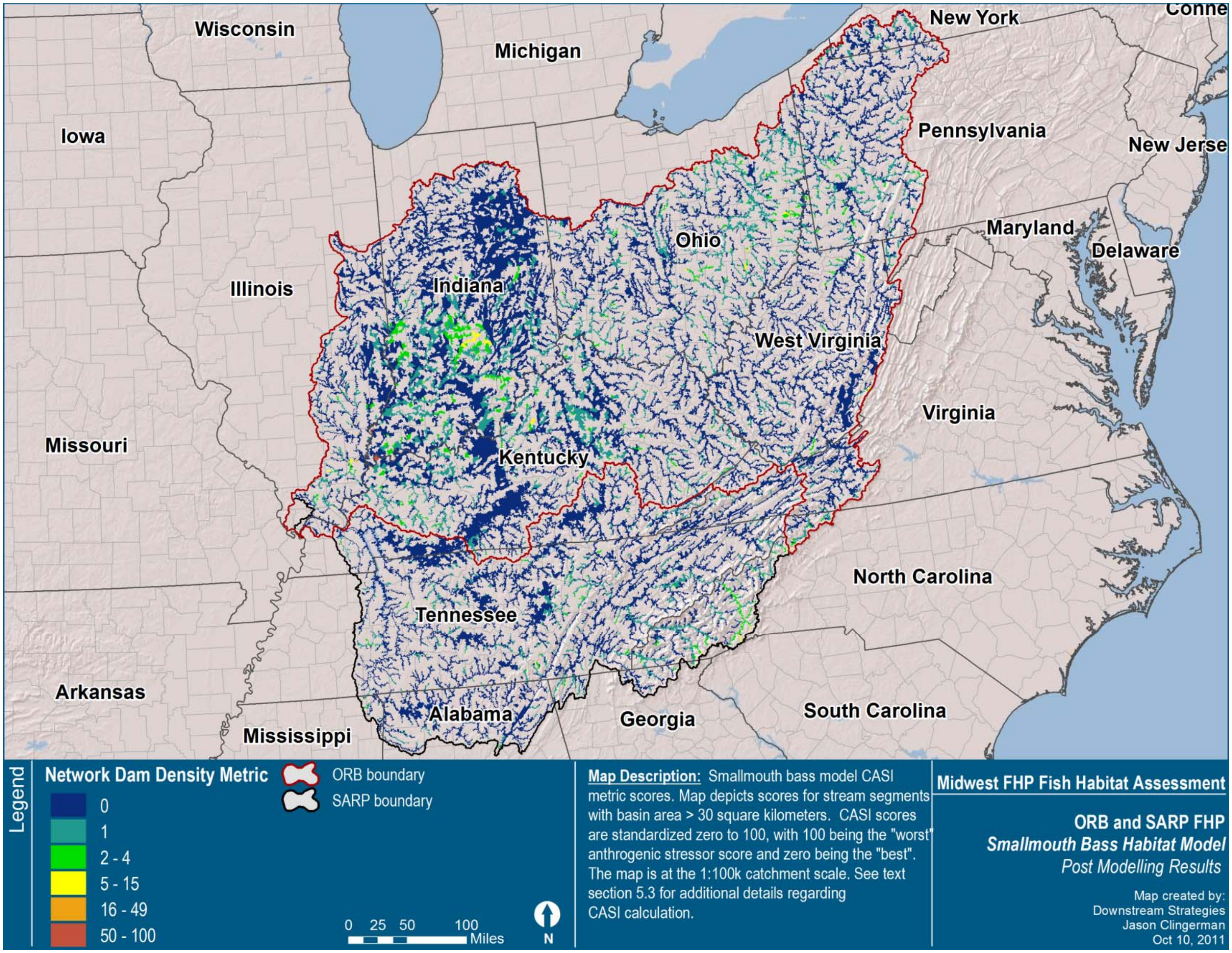


Figure 53: Fourth most influential anthropogenic index metric for smallmouth bass

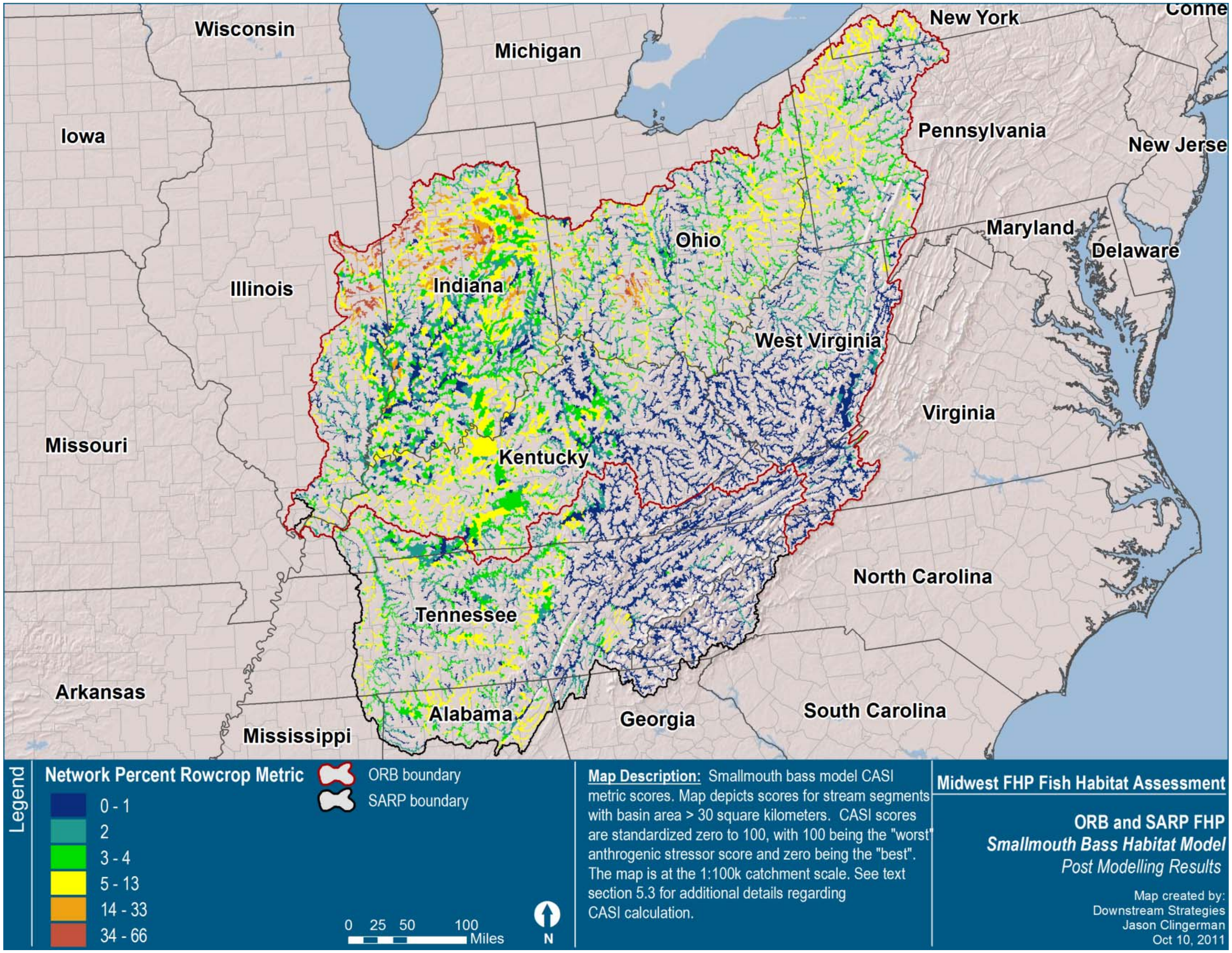
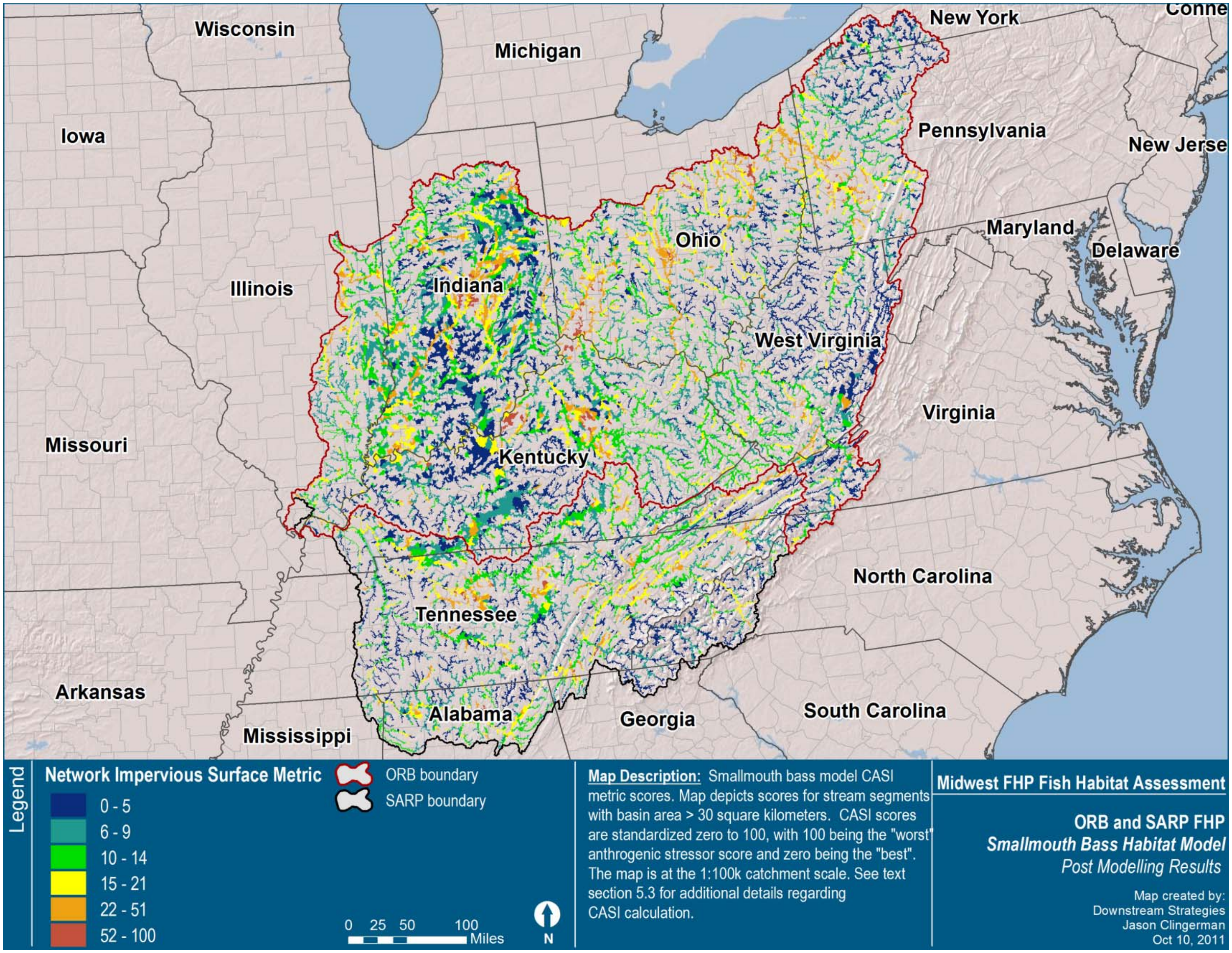


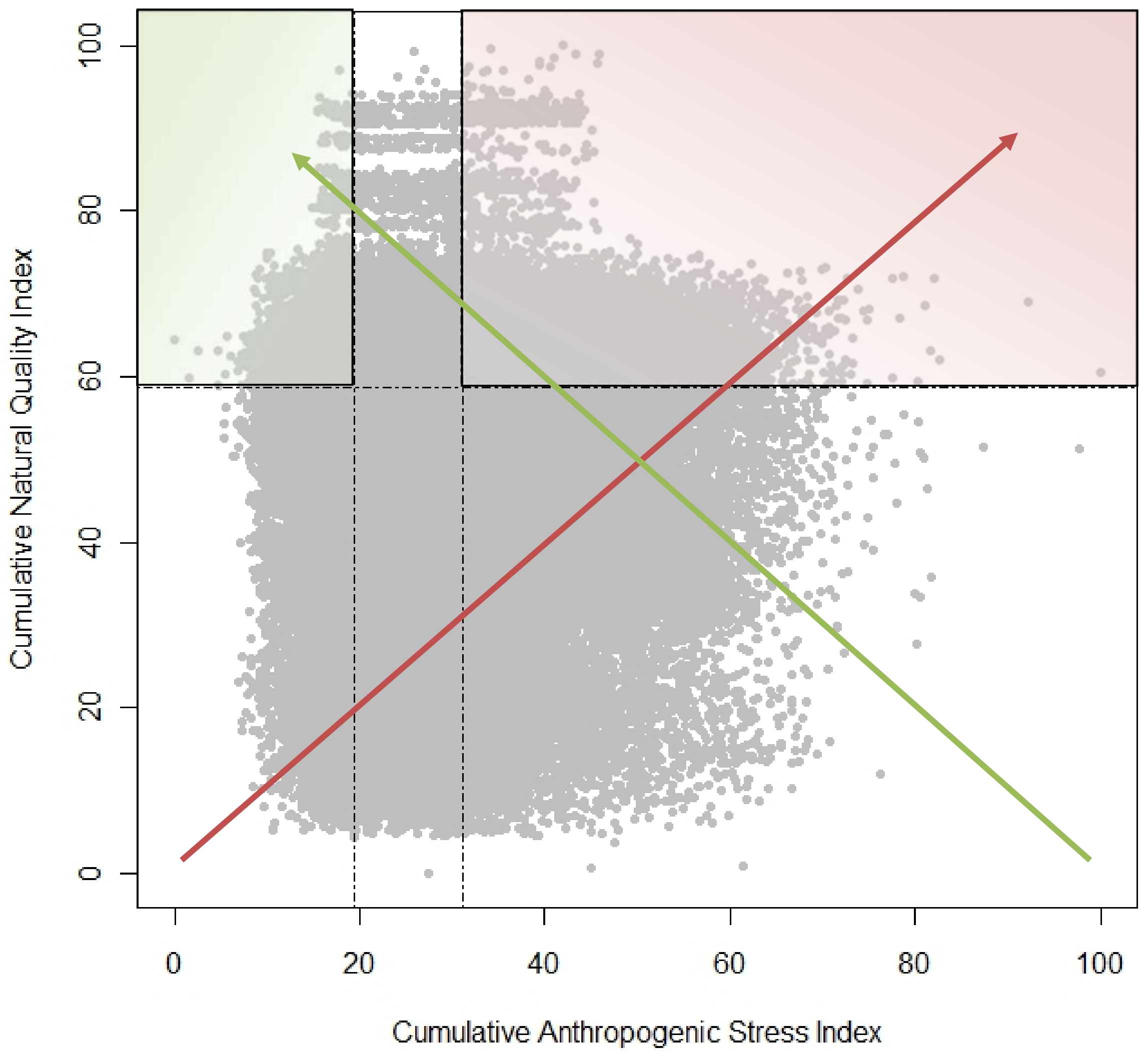
Figure 54: Fifth most influential anthropogenic index metric for smallmouth bass



4.4.4 *Restoration and protection priorities*

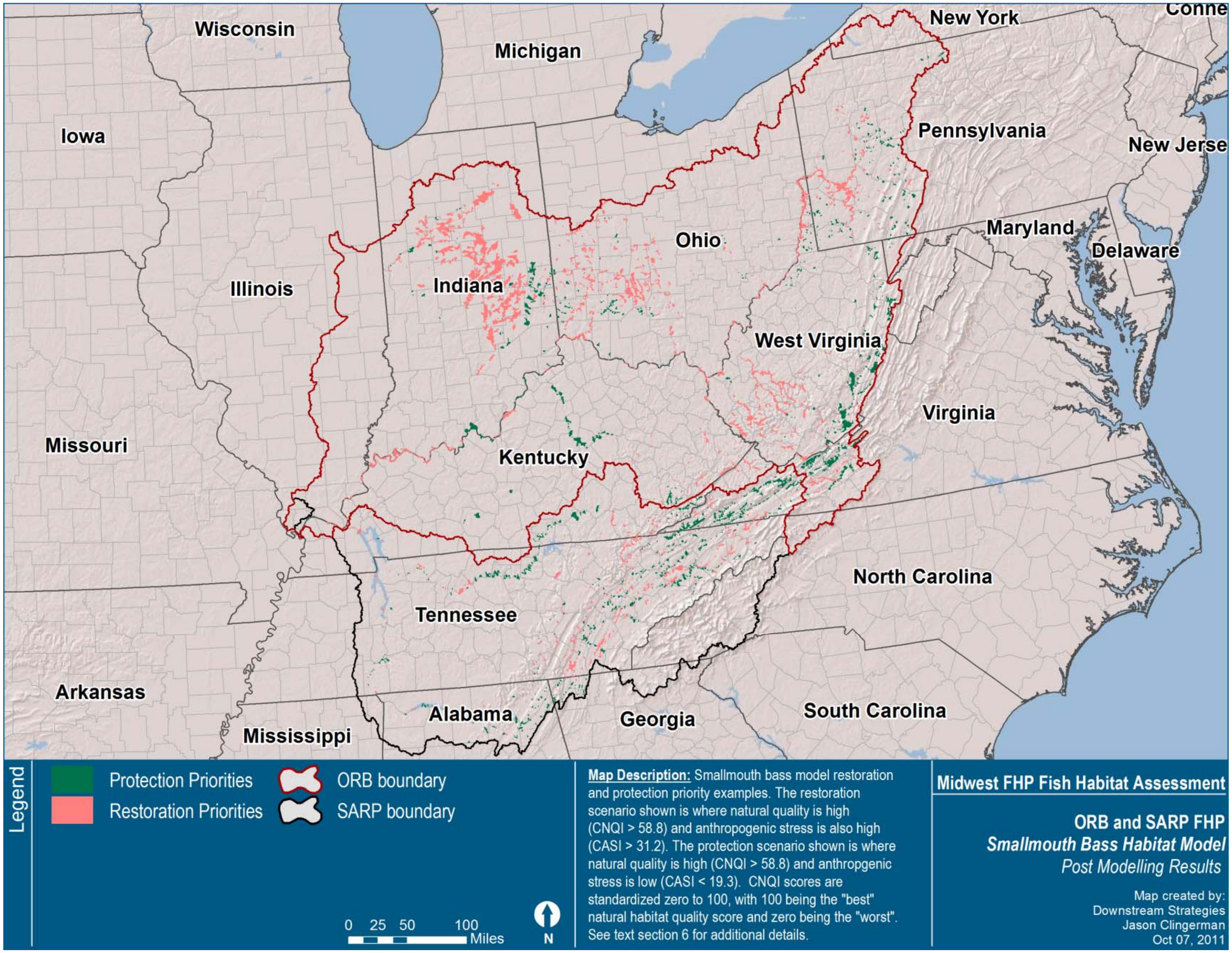
A plot of CNQI versus CASI values for all catchments in the study area (Figure 55) can be used as a reference when defining thresholds for categories of CNQI and CASI scores for use in the development of restoration and protection priorities. In the example shown (Figure 56), thresholds for restoration (high natural potential coupled with high anthropogenic stress) were set to CNQI greater than 58.8 and CASI greater than 31.2 (third quartiles). The thresholds used for protection priorities (high natural potential and low anthropogenic stress) were CNQI greater than 58.8 and CASI less than 19.3 (first quartile).

Figure 55: CNQI versus CASI values for all catchments for smallmouth bass



Note: Breakpoints for CNQI and CASI classes in this example are denoted by dashed lines. The arrows indicate the directions of increasing potential protection (green arrow) or restoration (red arrow) priority. The red box indicates catchments defined as restoration priorities under the example scenario. The green box indicates catchments defined as protection priorities under the same scenario.

Figure 56: Restoration and protection priorities for smallmouth bass



5. REDHORSE

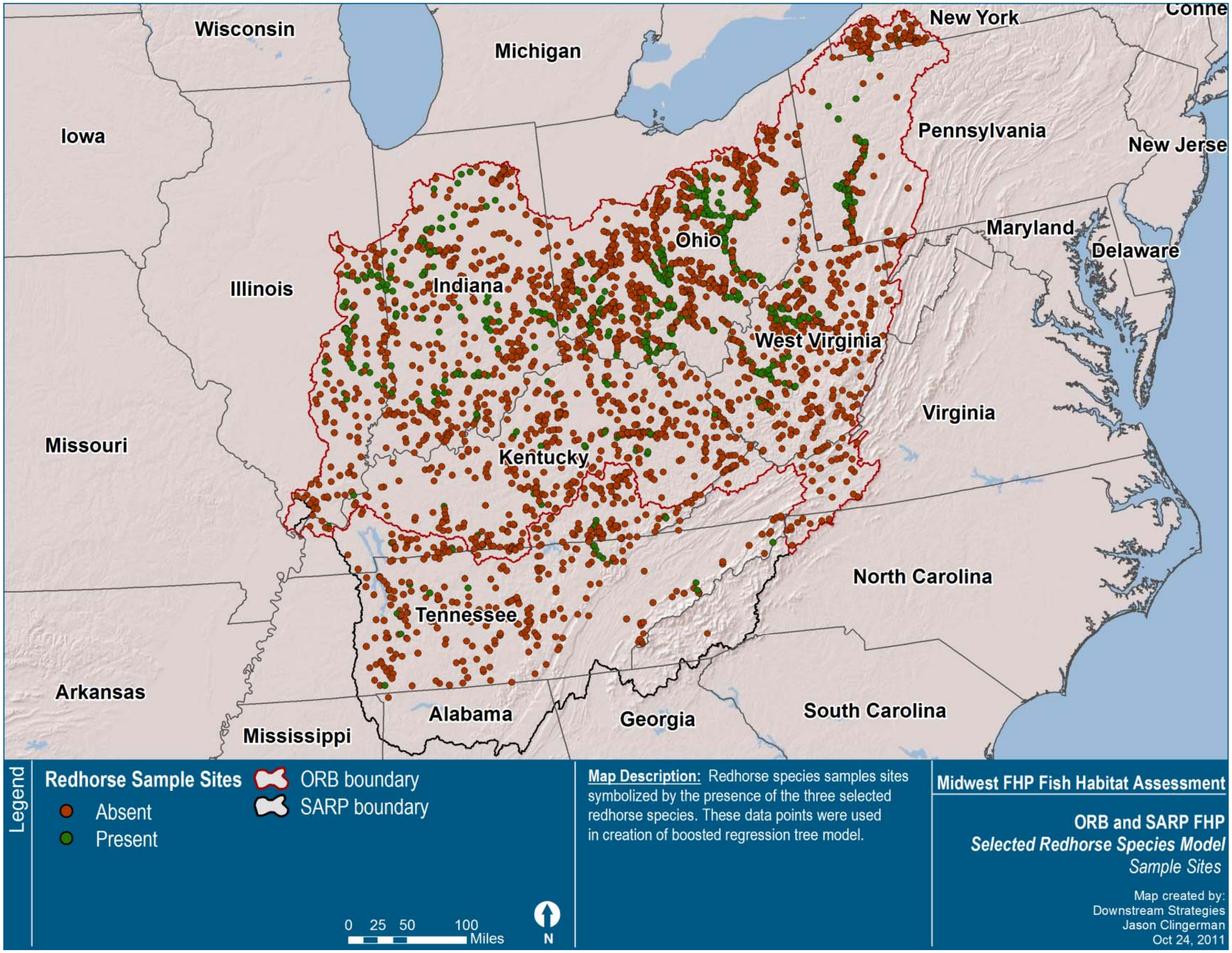
5.1 Modeling inputs

DS used a list of predictor variables selected by ORB and SARP to develop a 10-fold CV BRT model for redhorse at the 1:100k catchment scale. The model was used to produce maps of expected current redhorse distribution and maps of expected current natural habitat quality and anthropogenic stress at the 1:100k scale throughout the extents of both FHPs.

DS cooperated with ORB and SARP to arrive at a list of landscape-based habitat variables used to predict the presence of smallmouth (*Moxostoma breviceps*), shorthead (*Moxostoma macrolepidotum*), and river redhorse (*Moxostoma carinatum*) throughout the region; ultimately, those variables were also used for characterizing habitat quality and anthropogenic stress. From an initial suite of 372 catchment attributes, DS and the FHPs compiled a list of 92 predictors for evaluation. From that list, 42 variables were removed due to statistical redundancy ($r > 0.6$) or logical redundancy, resulting in a final list of 50 predictor variables for the BRT model and assessment. See Appendix A for a full data dictionary and the metadata document for variable processing notes.

ORB and SARP provided DS with a presence-absence dataset for redhorse comprised of 2,857 observations collected in streams between 47 and 45,000 square kilometers in drainage area and over a time frame spanning 1996 to 2010. The goal of this model was to identify smaller streams that are important to redhorse species. Taking that into consideration, streams with a drainage area less than 47 square kilometers were excluded since they would not be expected to have redhorse species, and streams with a drainage area greater than 45,000 square kilometers were excluded since redhorse species would naturally be expected there. Figure 57 maps all of the sampling sites that were used to construct the model and outlines all of the 1:100k catchments to which the modeling outputs were applied.

Figure 57: Redhorse modeling area and sampling sites



5.2 Modeling process

5.2.1 *Predictive performance*

The final selected model was comprised of 1,650 trees. The model had a CV correlation statistic of 0.676±0.018, and a CV ROC score of 0.921±0.007.

5.2.2 *Variable influence*

The BRT output includes a list of the predictor variables used in the model ordered and scored by their relative importance. The relative importance values are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman and Meulman, 2003). The relative influence score is scaled so that the sum of the scores for all variables is 100, where higher numbers indicate greater influence. Of the 50 predictor variables used to develop the redhorse model, 44 had a relative influence value greater than zero (Table 11). The five most influential predictors, which accounted for more than 64% of the total influence in the model, were:

- network drainage area,
- minimum catchment elevation,
- mean annual air temperature,
- Level III Ecoregion, and
- network density of Superfund sites.

The five most influential anthropogenic stressors, which accounted for nearly 11% of the total influence, were:

- network density of Superfund sites,
- network density of cattle,
- network surface water consumption,
- network density of dams, and
- network density of road crossings.

The five most influential natural habitat variables, which contributed nearly 64% of the total influence, were:

- network drainage area,
- minimum catchment elevation,
- mean annual air temperature,
- Level III Ecoregion, and
- network wetland land cover.

Network drainage area, the single most important variable in terms of relative influence, contributed almost 43% of the total influence.

Table 11: Relative influence of all variables in the final redhorse model

Variable code	Variable description	Relative influence
cumdrainag	Network drainage area	42.79
minelevraw	Minimum catchment elevation	7.37
temp	Mean annual air temperature	5.83
eco_code3	Level III Ecoregion	5.69
cercc_den	Network density of Superfund sites	2.49
cattlec	Network density of cattle	2.31
water_swc	Network surface water consumption	2.29
damsc_den	Network density of dams	2.09
wetlandpc	Network wetland land cover	1.86
brock5pc	Network sand/gravel bedrock geology cover	1.75
roadcrc_den	Network density of road crossings	1.73
brock7pc	Network shale bedrock geology cover	1.69
soil2pc	Network soil type B, B/D cover	1.66
cropspc	Network rowcrop land cover	1.64
forpc	Network forested land cover	1.62
imp06	Local impervious surface cover	1.36
minesc_den	Network density of mines	1.15
ripdisp	Local riparian disturbance score	1.12
BFI_meanc	Network mean baseflow index	1.11
imp06c	Network impervious surface cover	1.09
slope	Slope of catchment flowline	0.95
npdesc_den	Network density of National Pollutant Discharge Elimination System permits	0.91
grasspc	Network grassland land cover	0.91
grassp	Network grassland land cover	0.88
pastp	Local pasture land cover	0.83
roadcr_den	Local density of road crossings	0.70
surf5pc	Network loess surficial geology cover	0.65
soil3pc	Network soil type C, C/D cover	0.63
TRI_den	Local density of Toxic Release Inventory sites	0.63
brock3pc	Network mafic/igneous bedrock geology cover	0.54
soil4pc	Network soil type D cover	0.47
brock6p	Network sandstone bedrock geology cover	0.46
pastpc	Network pasture land cover	0.45
brock1pc	Network carbonate bedrock geology cover	0.39
surf3pc	Network alluvium surficial geology cover	0.36
surf2pc	Network outwash surficial geology cover	0.29
soil1pc	Network soil type A, A/D cover	0.26
mines_den	Local density of mines	0.25
surf4pc	Network lacustrine surficial geology cover	0.22
surf7pc	Network clay surficial geology cover	0.17
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.15
dams_den	Local density of dams	0.12
surf6pc	Network residuum surficial geology cover	0.07
soil4p	Local soil type D cover	0.07
CERC_den	Local density of Superfund sites	0.00
soil1p	Local soil type A, A/D cover	0.00

brock2p	Local felsic/igneous bedrock geology cover	0.00
brock3p	Local mafic/igneous bedrock geology cover	0.00
brock4p	Local metamorphic bedrock geology cover	0.00
surf4p	Local lacustrine surficial geology cover	0.00

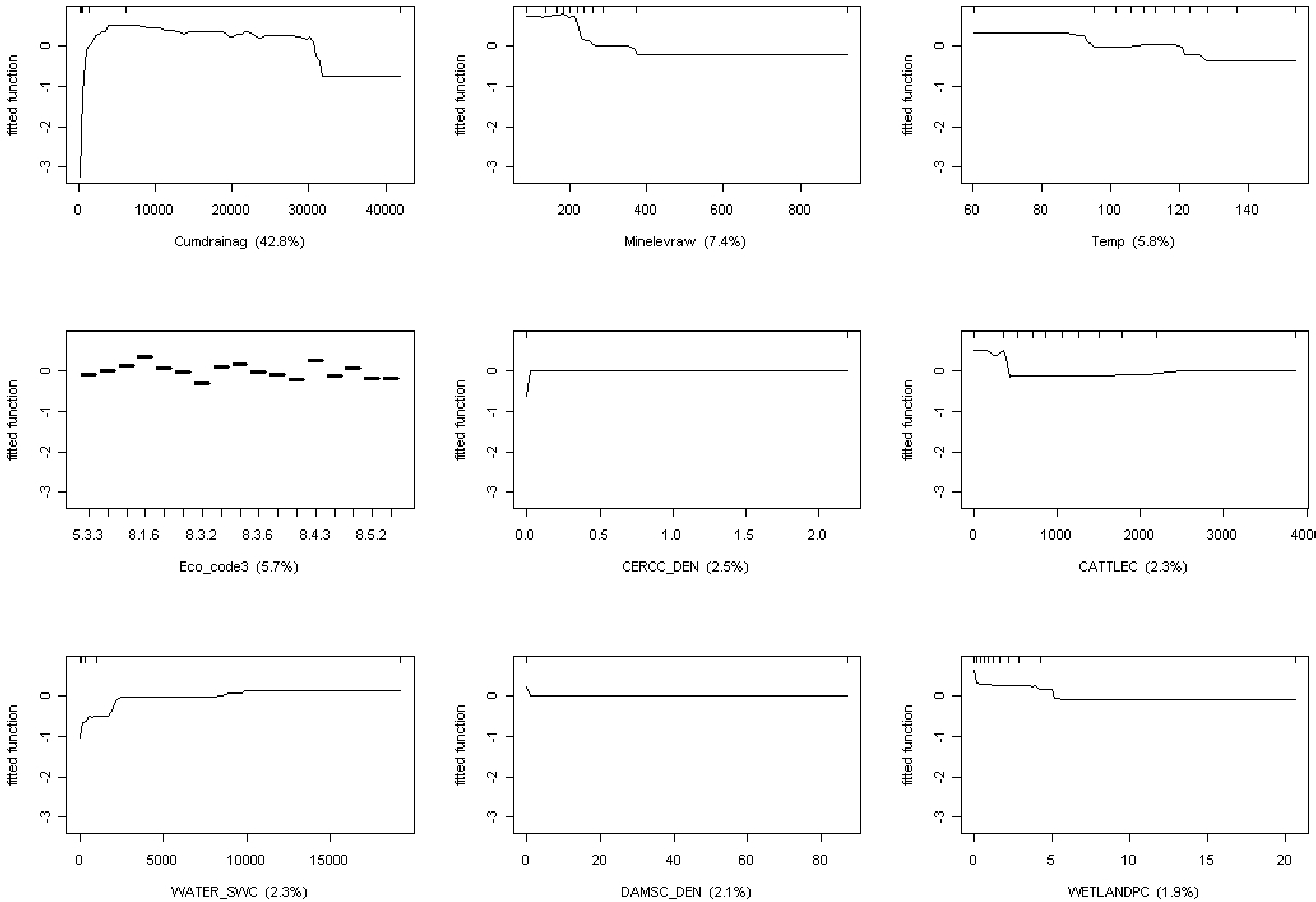
Note: Individual variables are highlighted according to whether they were determined to be anthropogenic in nature (red highlight) or natural (green highlight).

5.2.3 *Variable functions*

The BRT output also contains quantitative information on partial dependence functions that can be plotted to visualize the effect of each individual predictor variable on the response after accounting for all other variables in the model. Similar to the interpretation of traditional regression coefficients, the function plots are not always a perfect representation of the relationship for each variable, particularly if interactions are strong or predictors are strongly correlated. However, they do provide a useful and objective basis for interpretation (Friedman, 2001; Friedman and Meulman, 2003).

These plots show the trend of the response variable (y-axis) as the predictor variable (x-axis) changes. The response variable is transformed (usually to the logit scale) so that the magnitude of trends for each predictor variable's function plot can be accurately compared. The dash marks at the top of each function represent the deciles of the data used to build the model. The function plots for the nine most influential variables in the redhorse model (see Table 11 for reference) are illustrated in Figure 58 below. The plots for all 50 variables are shown in Appendix B.

Figure 58: Functional responses of the dependent variable to individual predictors of redhorse



Note: Only the top nine predictors, based on relative influence (shown in parentheses; see Appendix A for descriptions of variable codes), are shown here. See Appendix B for plots of remaining predictor variables.

5.3 Post-modeling

The variable importance table and partial dependence functions of the final BRT model were used to create the post-modeling indices of natural habitat quality and anthropogenic stress for redhorse. The CNQI was comprised of 23 variables with relative influence greater than zero that were classified as natural habitat features (Table 12). The CASI was comprised of 10 variables with relative influence greater than zero that were classified as anthropogenic habitat features (Table 13). To calculate the cumulative indices (i.e., CNQI and CASI), each of the individual natural or anthropogenic variables used in the two indices was converted to a metric by first applying the appropriate transformations, based on their function plots, and then rescaling the transformed measures to a 0 to 100 scale. To calculate the cumulative index from the individual metrics, the metrics were first multiplied by their appropriate weighting factors and then summed. The CNQI and CASI scores were a result of a rescaling of those weighted and summed metrics, again from 0 to 100.

5.3.1 *Variable weights*

Table 12 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CNQI. The five most influential factors in the CNQI were:

- network drainage area,
- minimum catchment elevation,
- mean annual air temperature,
- Level III Ecoregion, and
- network wetland land cover.

Table 13 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CASI. The five most influential factors in the CASI were:

- network density of cattle,
- network density of dams,
- local impervious surface cover,
- local riparian disturbance, and
- network impervious surface cover.

Table 12: Relative influence and weights for natural variables on redhorse

Variable code	Variable description	Relative influence	Weighting factor
cumdrainag	Network drainage area	42.79	1.00
minelevraw	Minimum catchment elevation	7.37	0.17
temp	Mean annual air temperature	5.83	0.14
eco_code3	Level III Ecoregion	5.69	0.13
wetlandpc	Network wetland land cover	1.86	0.04
brock5pc	Network sand/gravel bedrock geology cover	1.75	0.04
brock7pc	Network shale bedrock geology cover	1.69	0.04
soil2pc	Network soil type B, B/D cover	1.66	0.04
BFI_meanc	Network mean baseflow index	1.11	0.03
slope	Slope of catchment flowline	0.95	0.02
surf5pc	Network loess surficial geology cover	0.65	0.02
soil3pc	Network soil type C, C/D cover	0.63	0.01
brock3pc	Network mafic/igneous bedrock geology cover	0.54	0.01
soil4pc	Network soil type D cover	0.47	0.01
brock6p	Network sandstone bedrock geology cover	0.46	0.01
brock1pc	Network carbonate bedrock geology cover	0.39	0.01
surf3pc	Network alluvium surficial geology cover	0.36	0.01
surf2pc	Network outwash surficial geology cover	0.29	0.01
soil1pc	Network soil type A, A/D cover	0.26	0.01
surf4pc	Network lacustrine surficial geology cover	0.22	0.01
surf7pc	Network clay surficial geology cover	0.17	0.00
surf6pc	Network residuum surficial geology cover	0.07	0.00
soil4p	Local soil type D cover	0.07	0.00

Table 13: Relative influence and weights for anthropogenic variables on redhorse

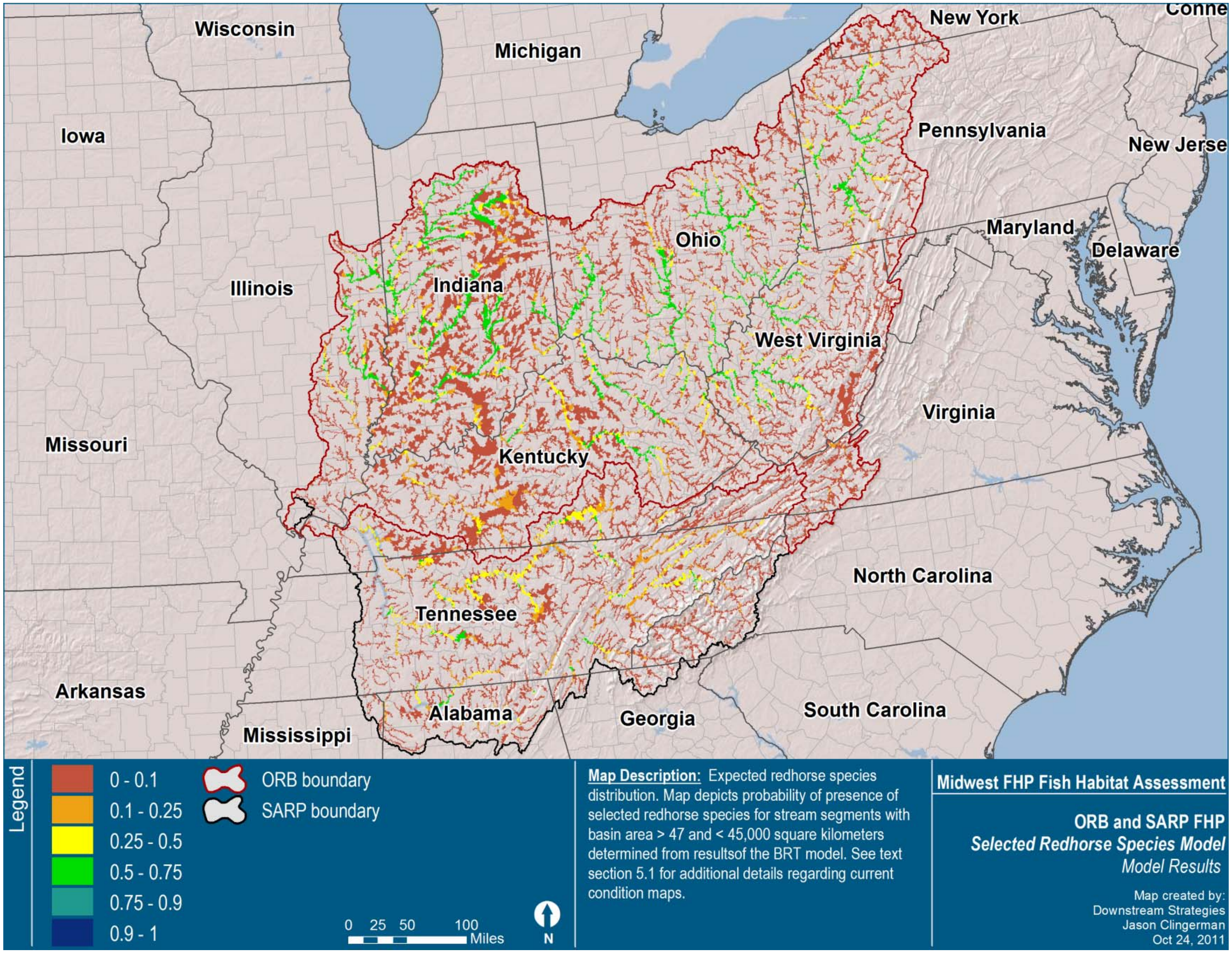
Variable code	Variable description	Relative influence	Weighting factor
cattlec	Network density of cattle	1.00	1.00
damsc_den	Network density of dams	0.91	0.91
imp06	Local Impervious surface cover	0.59	0.59
ripdisp	Local riparian disturbance score	0.49	0.49
imp06c	Network impervious surface cover	0.47	0.47
grasspc	Network grassland land cover	0.40	0.40
TRI_den	Local density of Toxic Release Inventory sites	0.27	0.27
pastpc	Network pasture land cover	0.19	0.19
mines_den	Local density of mines	0.11	0.11
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.06	0.06

5.4 Mapped Results

5.4.1 *Expected current conditions*

Redhorse probability of presence was calculated for all 1:100k stream catchments in the study area using the BRT model. The predicted probability values ranged from 0.006 to 1. The mean predicted probability value for the 45,450 total catchments larger than 47 square kilometers drainage area and smaller than 45,000 square kilometers drainage area was 0.114. There were 3,198 catchments larger than 47 square kilometers in drainage area and smaller than 45,000 square kilometers with a predicted probability of presence greater than 0.75; and 4,390 catchments larger than 47 square kilometers and smaller than 45,000 square kilometers where the probability of presence was between 0.5 and 0.75. These results are mapped in Figure 59.

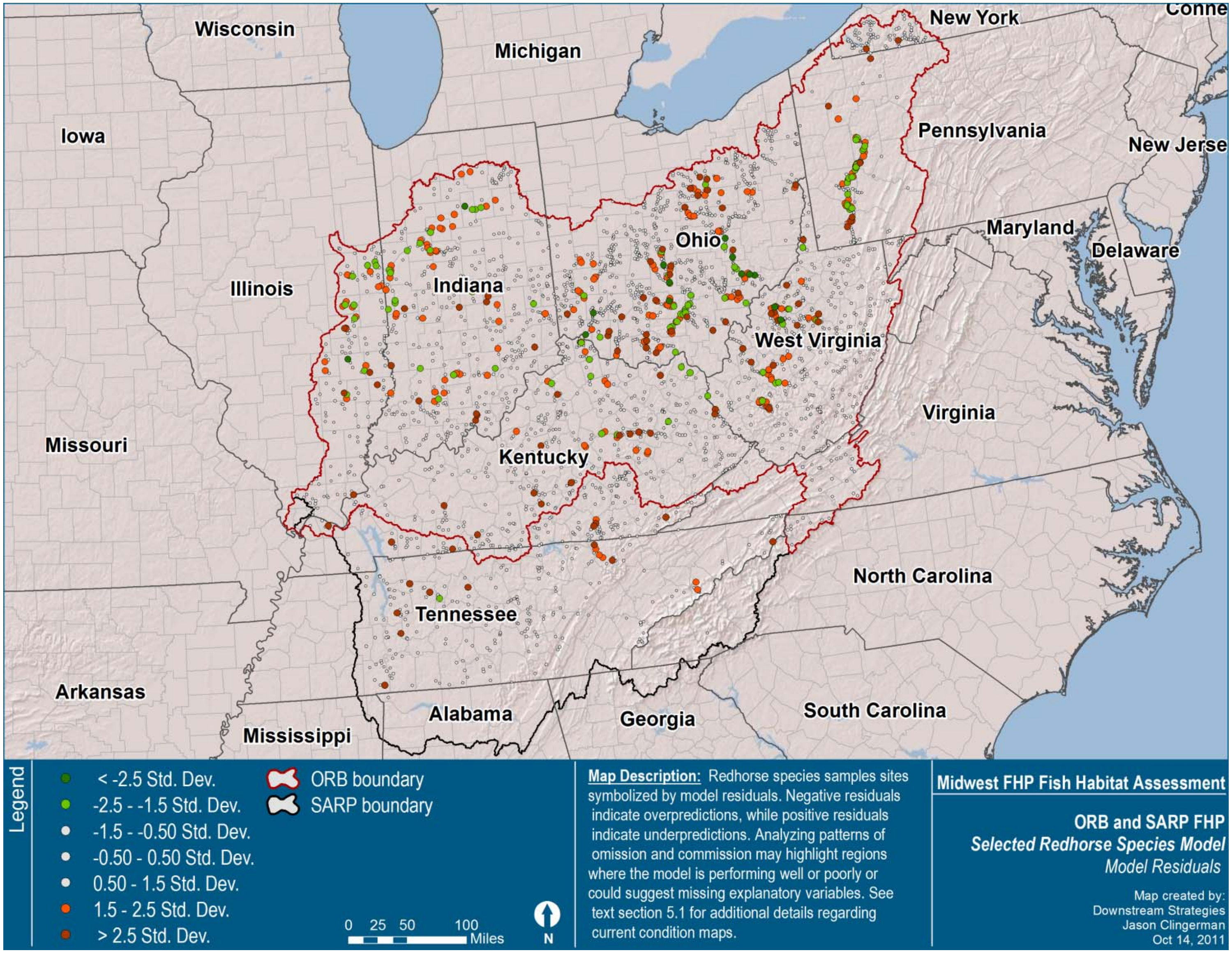
Figure 59: Expected redhorse score



5.4.2 *Spatial variability in predictive performance*

Analyzing patterns of omission and commission may highlight regions where the model is performing well or poorly or could suggest missing explanatory variables (Figure 60). To assess omission and commission, residuals are also calculated by the BRT model. The residuals are a measure of the difference in the measured and modeled values (measured value *minus* modeled value). Negative residuals indicate overpredictions (predicting higher values than are true), while positive residuals indicate underpredictions (predicting lower values than are true).

Figure 60: Distribution of redhorse model residuals by sampling site



5.4.3 *Indices of stress and natural quality*

Maps of CNQI and CASI illustrate the spatial distribution of natural habitat potential (i.e., CNQI score) and anthropogenic stress (i.e., CASI score) in the ORB and SARP. CNQI and CASI scores are mapped in Figure 61 and Figure 62, respectively. The top five most influential variables toward the calculation of CNQI are shown in Figure 63-Figure 67. The top five variables contributing toward the calculation of CASI are mapped in Figure 68-Figure 72. CNQI, CASI, and their metrics are all scaled on a 0-100 scale (see Section 5.3 for more details on CNQI and CASI calculation). For CNQI, higher values indicate higher natural quality, while higher values for CASI indicate higher levels of anthropogenic stress.

Figure 61: Cumulative natural quality index for redhorse

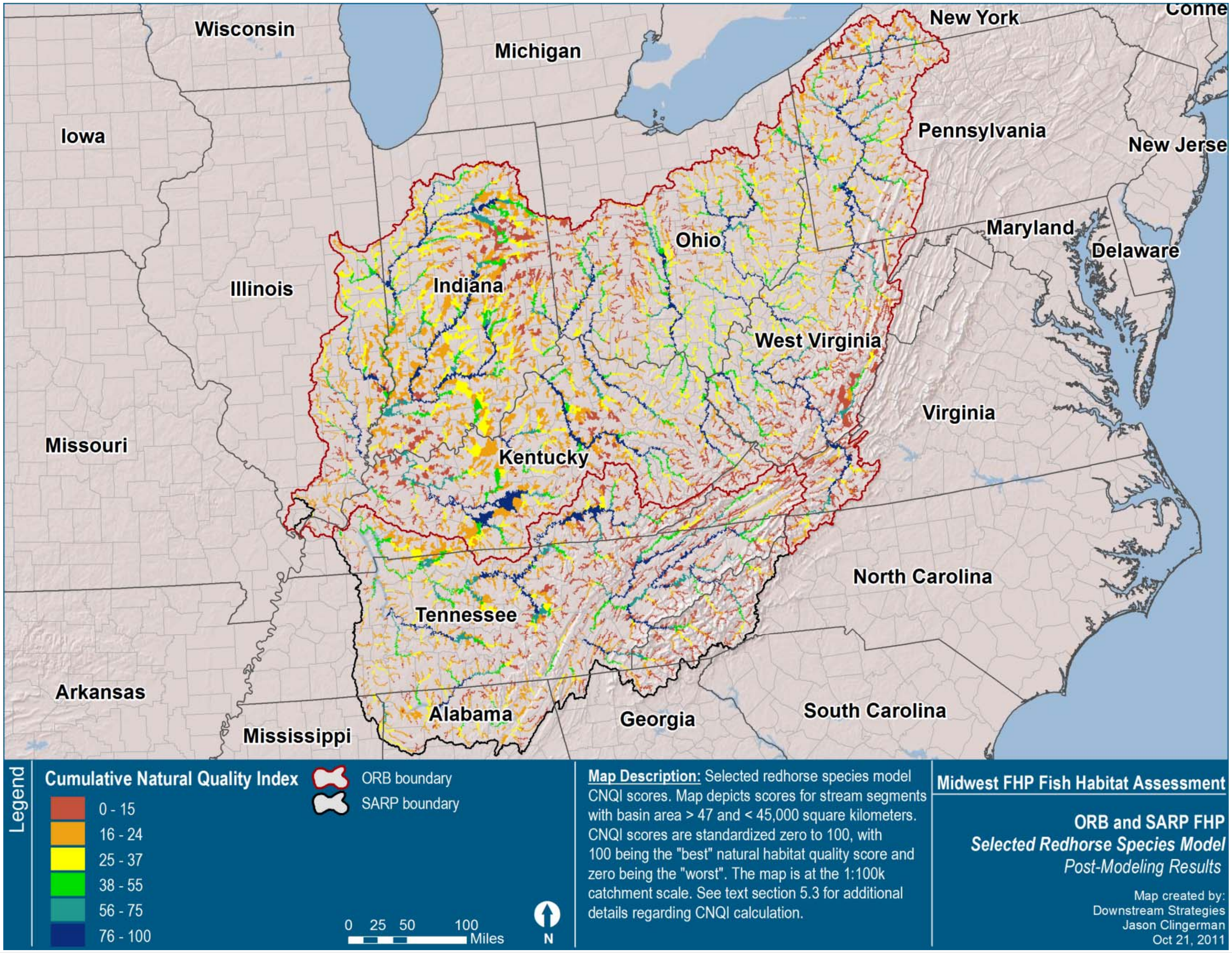


Figure 62: Cumulative anthropogenic stress index for redhorse

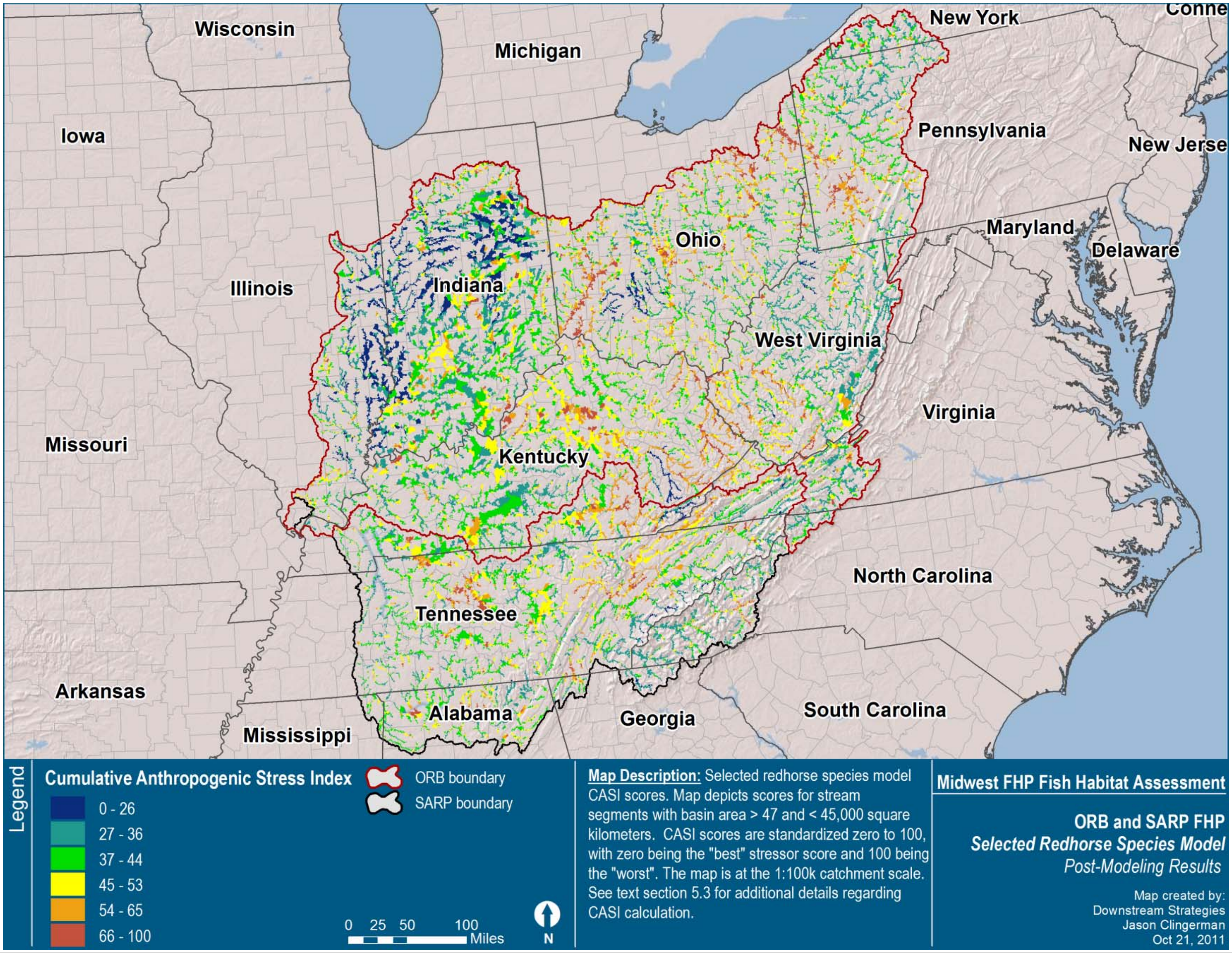


Figure 63: Most influential natural index metric for redhorse

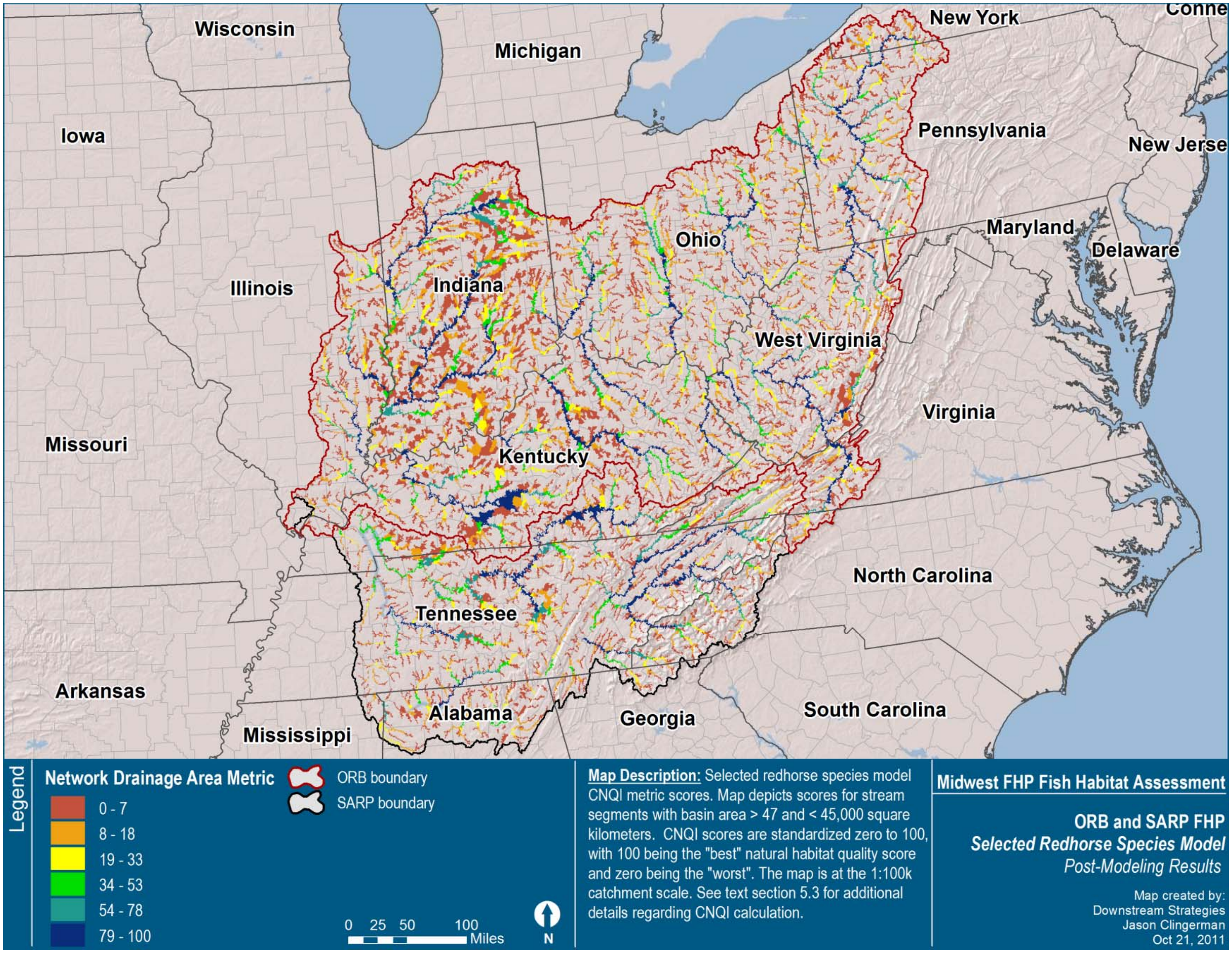


Figure 64: Second most influential natural index metric for redhorse

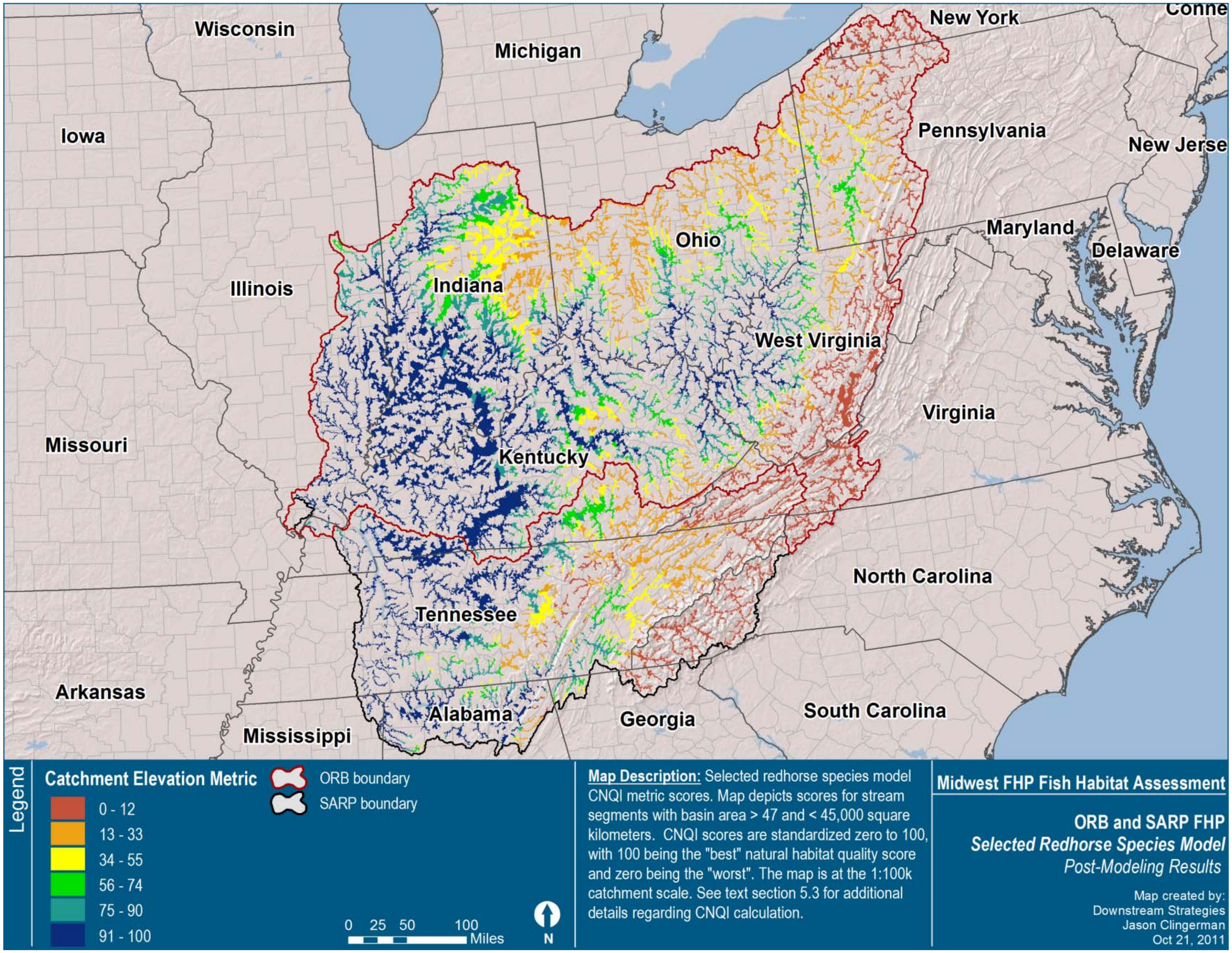


Figure 65: Third most influential natural index metric for redhorse

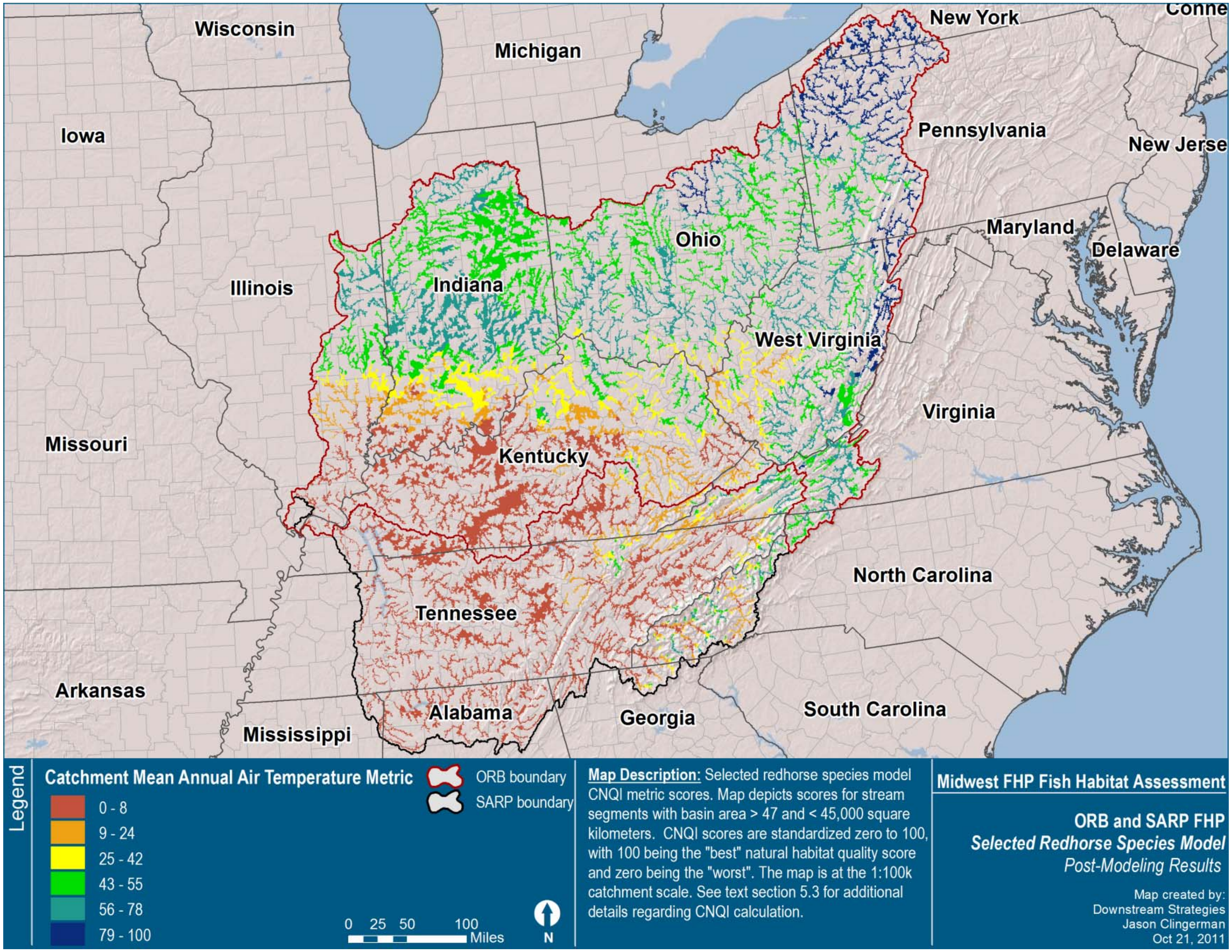


Figure 66: Fourth most influential natural index metric for redhorse

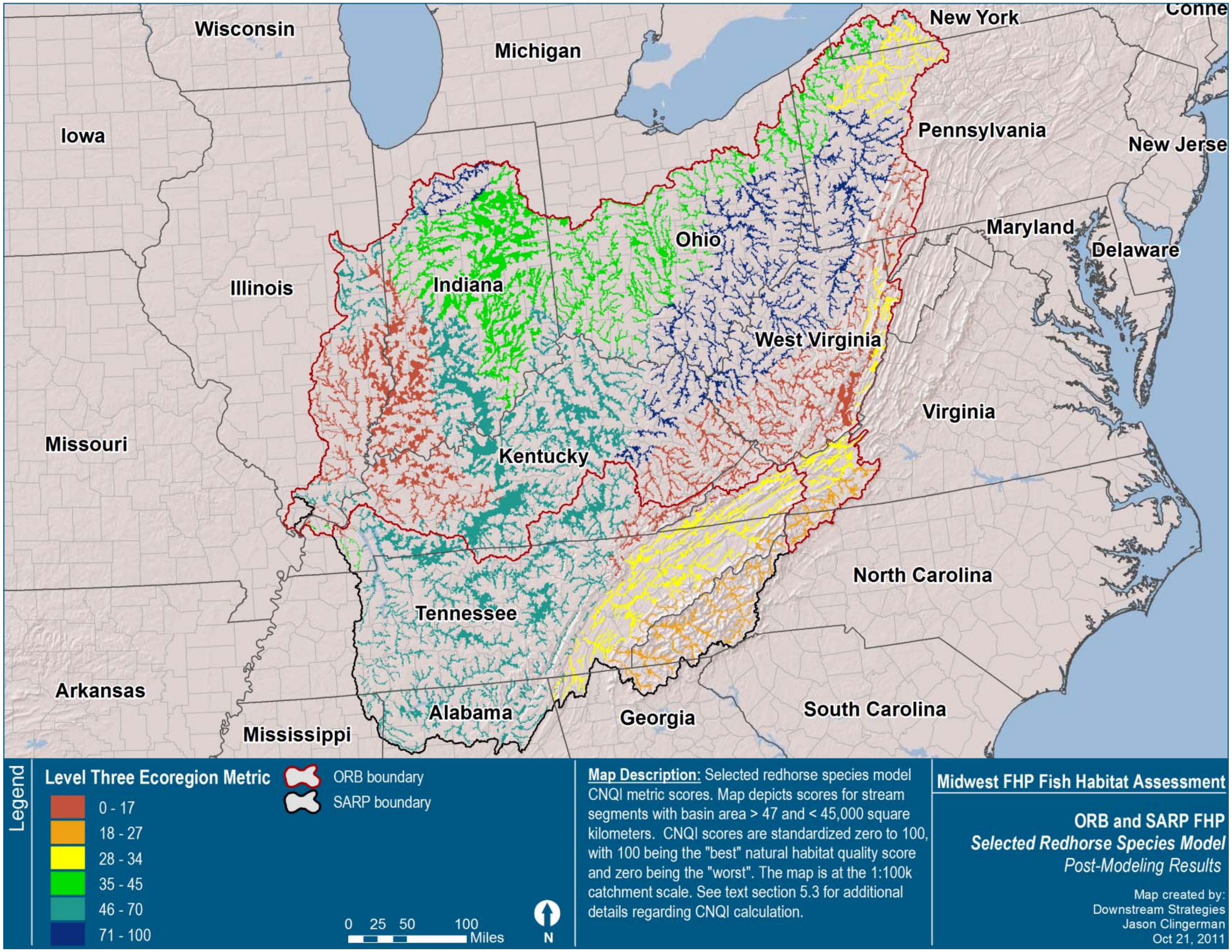


Figure 67: Fifth most influential natural index metric for redhorse

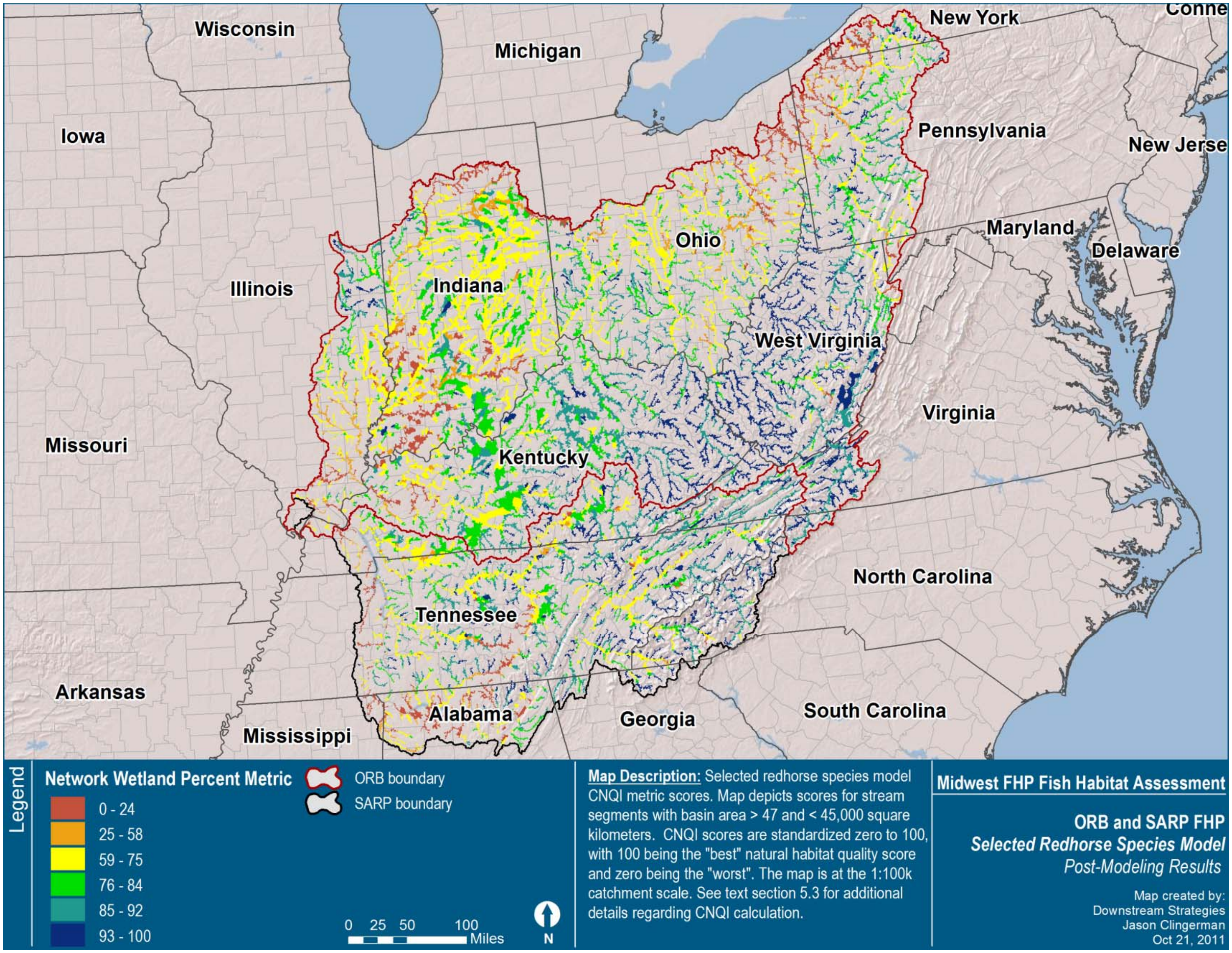


Figure 68: Most influential anthropogenic index metric for redhorse

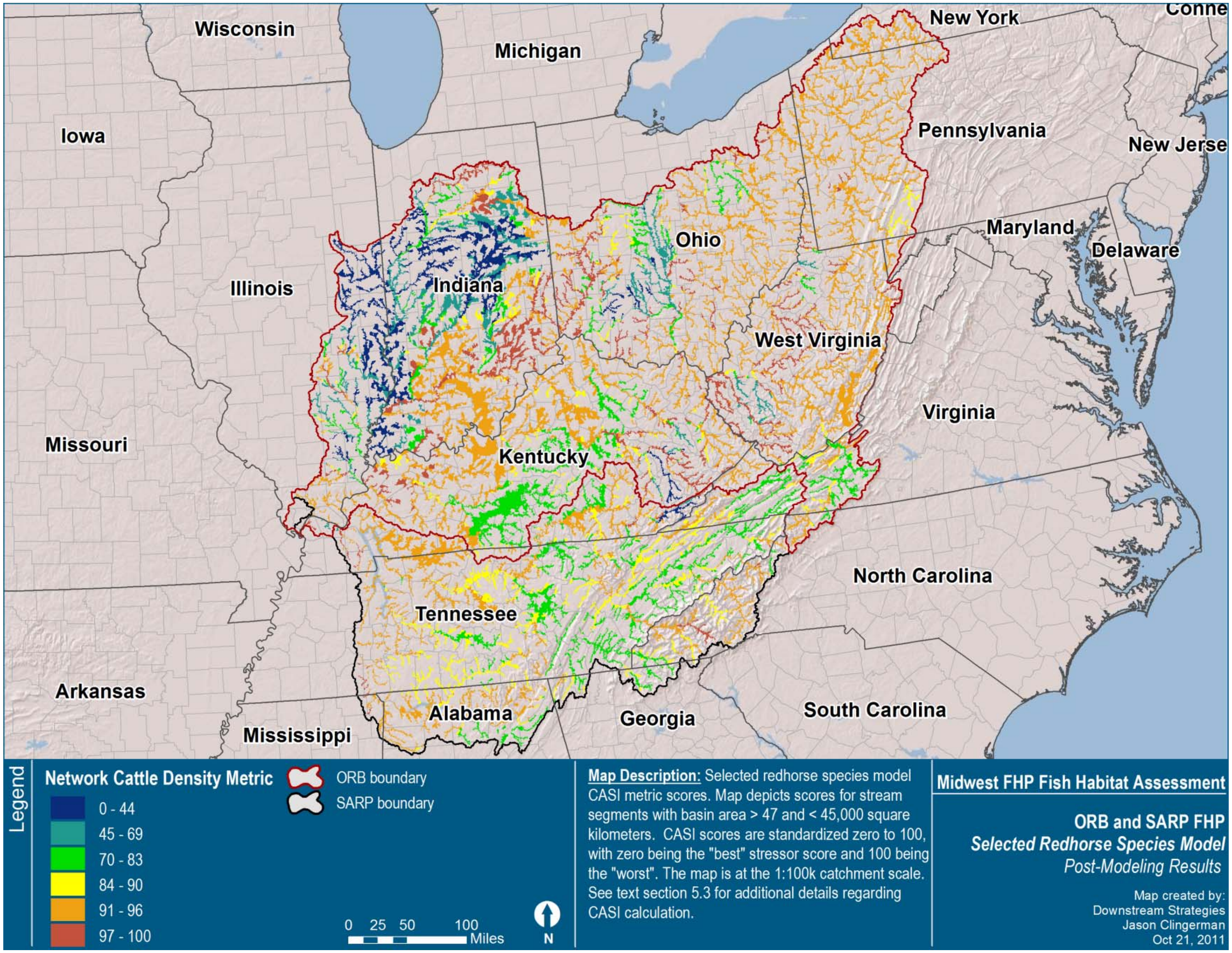


Figure 69: Second most influential anthropogenic index metric for redhorse

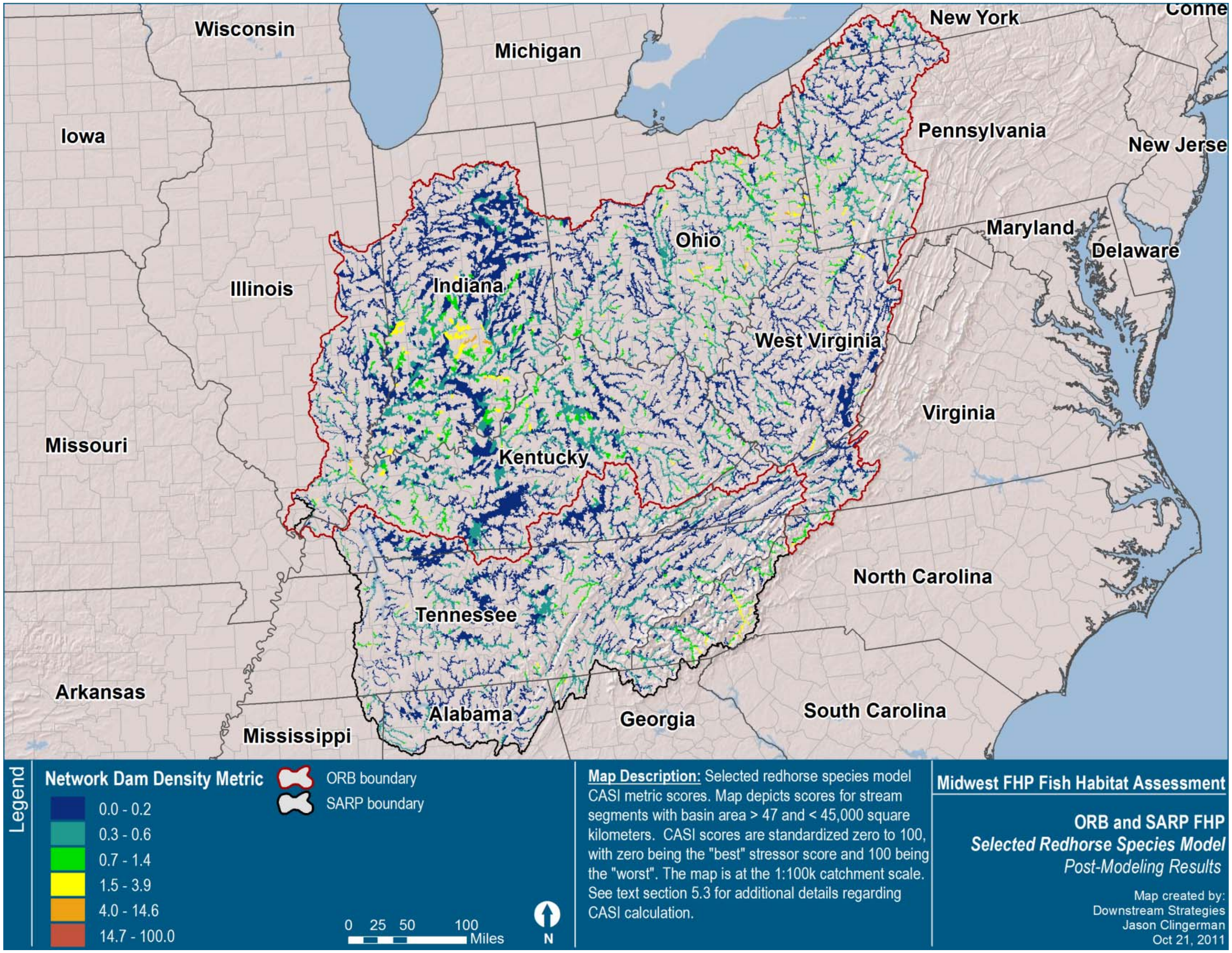


Figure 70: Third most influential anthropogenic index metric for redhorse

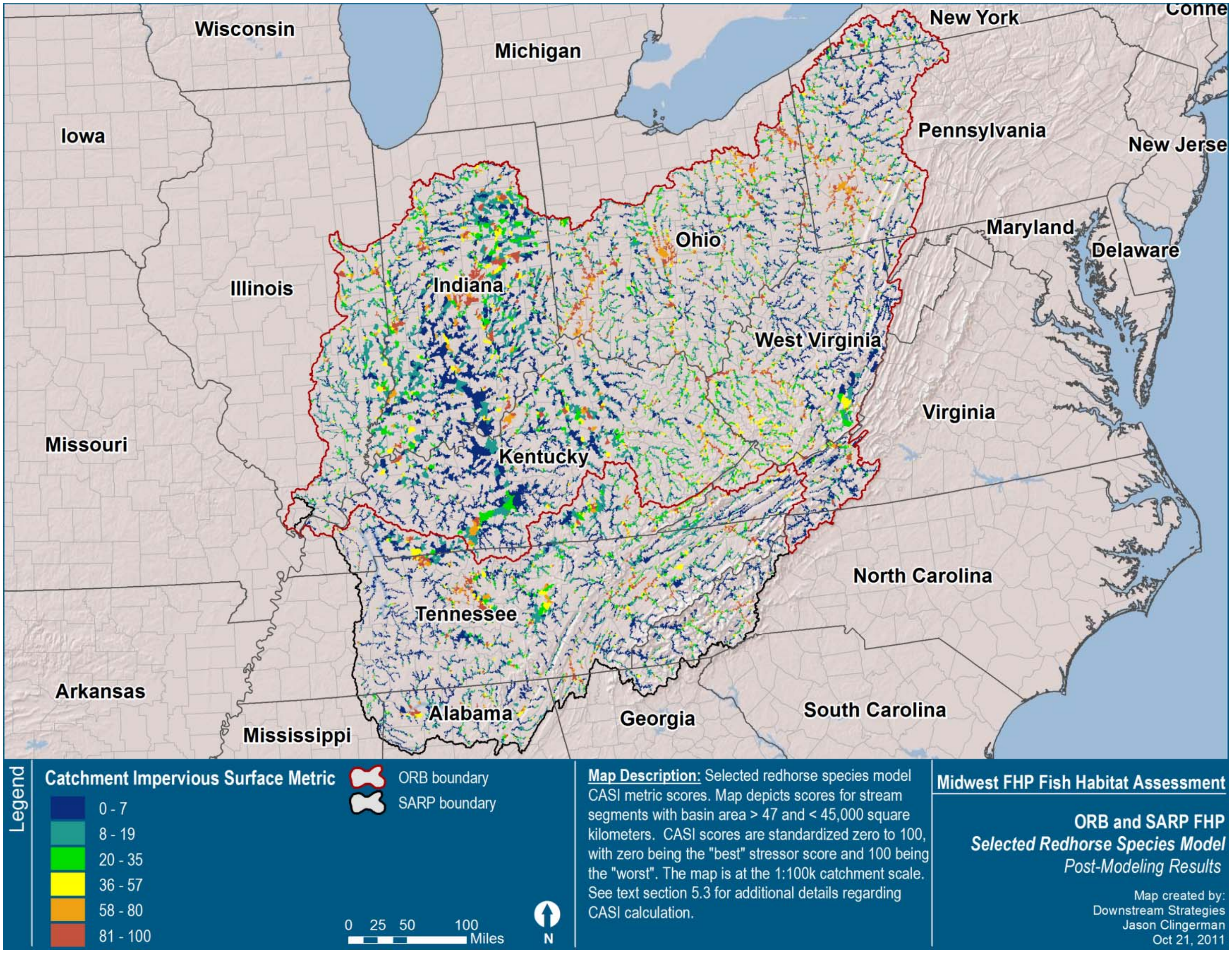


Figure 71: Fourth most influential anthropogenic index metric for redhorse

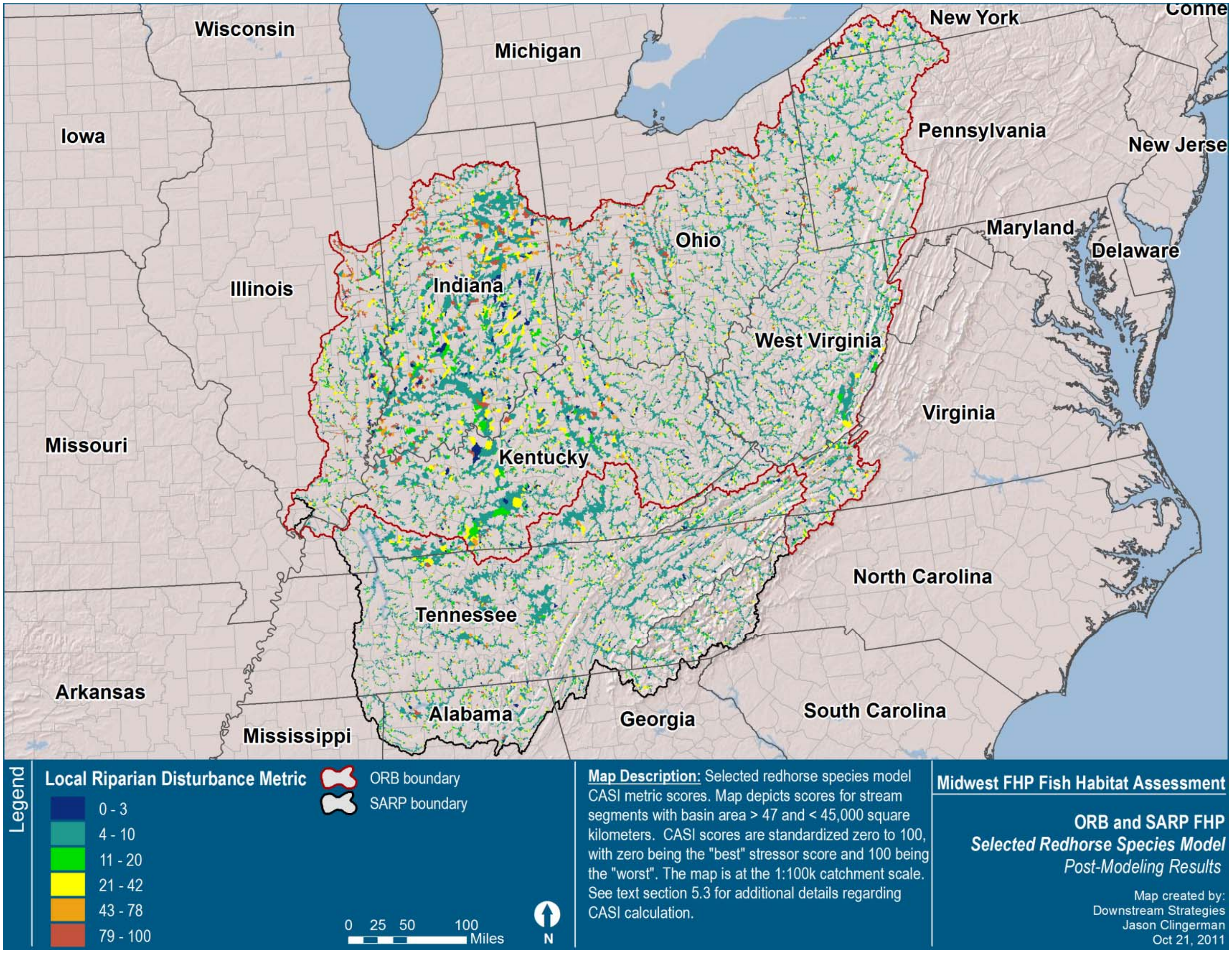
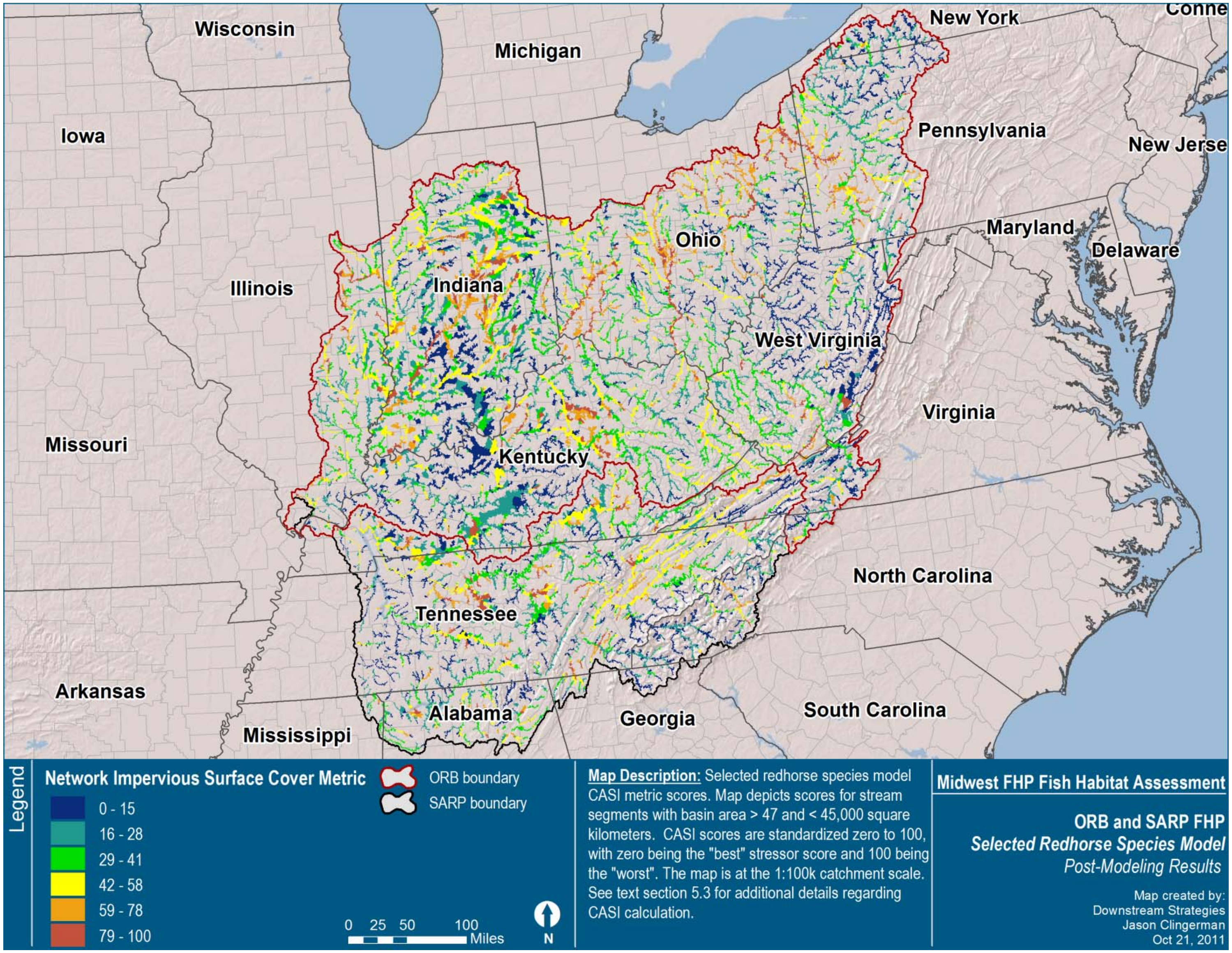


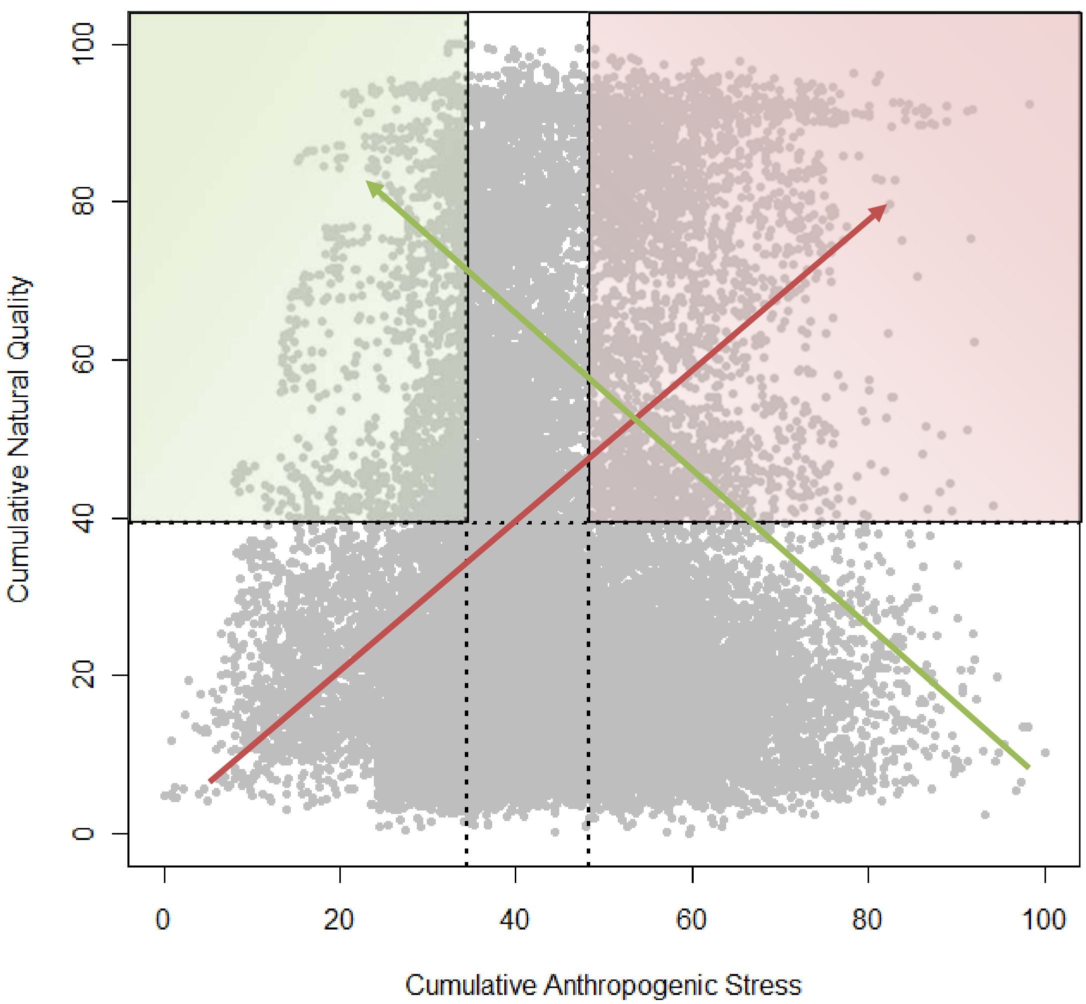
Figure 72: Fifth most influential anthropogenic index metric for redhorse



5.4.4 *Restoration and protection priorities*

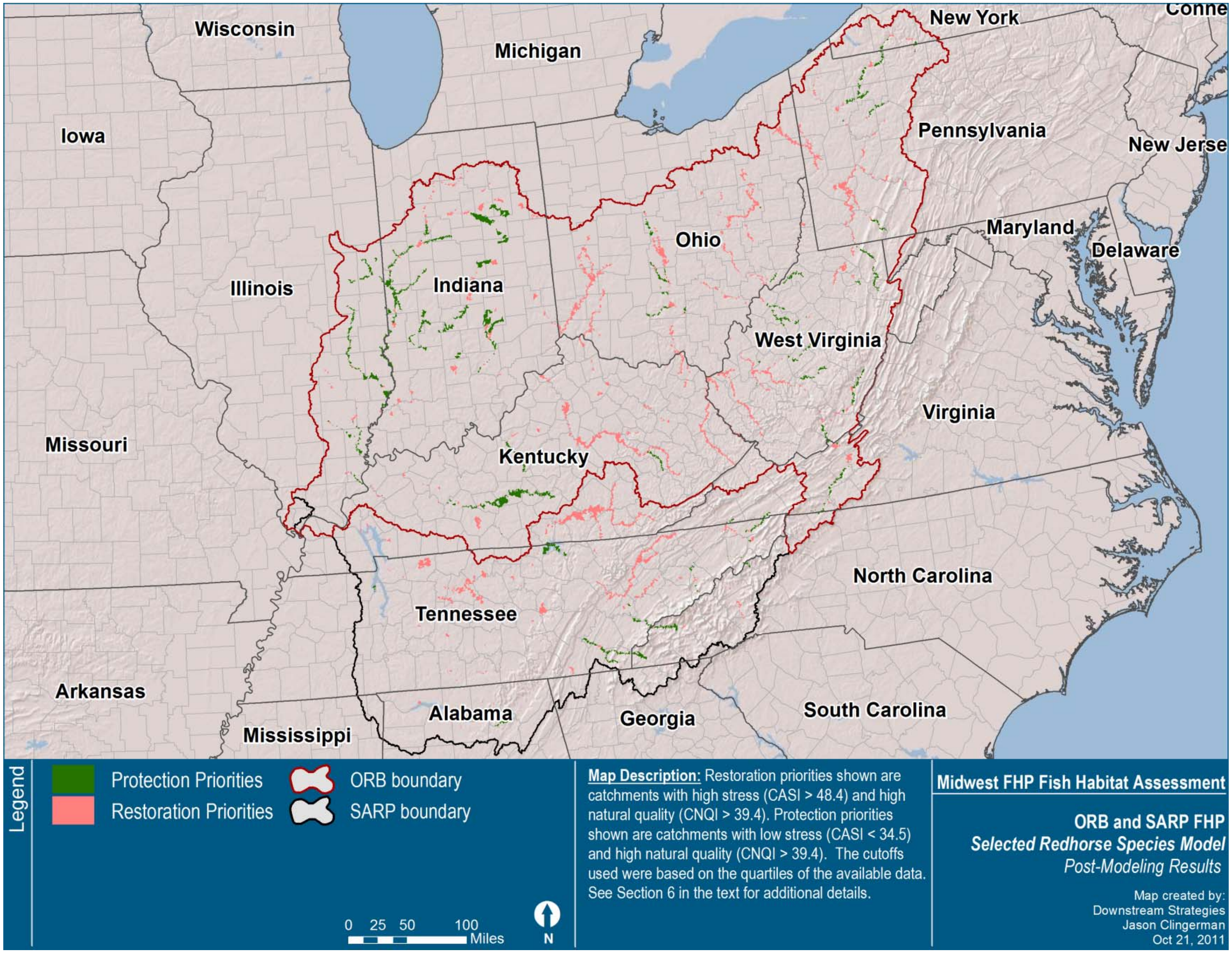
A plot of CNQI versus CASI values for all catchments in the study area (Figure 73) can be used as a reference when defining thresholds for categories of CNQI and CASI scores for use in the development of restoration and protection priorities. In the example shown (Figure 74), thresholds for restoration (high natural potential coupled with high anthropogenic stress) were set to CNQI greater than 39.4 and CASI greater than 48.4 (third quartiles). The thresholds used for protection priorities (high natural potential and low anthropogenic stress) were CNQI greater than 39.4 and CASI less than 34.5 (first quartile).

Figure 73: CNQI versus CASI values for all catchments for redhorse



Note: Breakpoints for CNQI and CASI classes in this example are denoted by dashed lines. The arrows indicate the directions of increasing potential protection (green arrow) or restoration (red arrow) priority. The red box indicates catchments defined as restoration priorities under the example scenario. The green box indicates catchments defined as protection priorities under the same scenario.

Figure 74: Restoration and protection priorities for redhorse



6. PERCENT INTOLERANT FISH

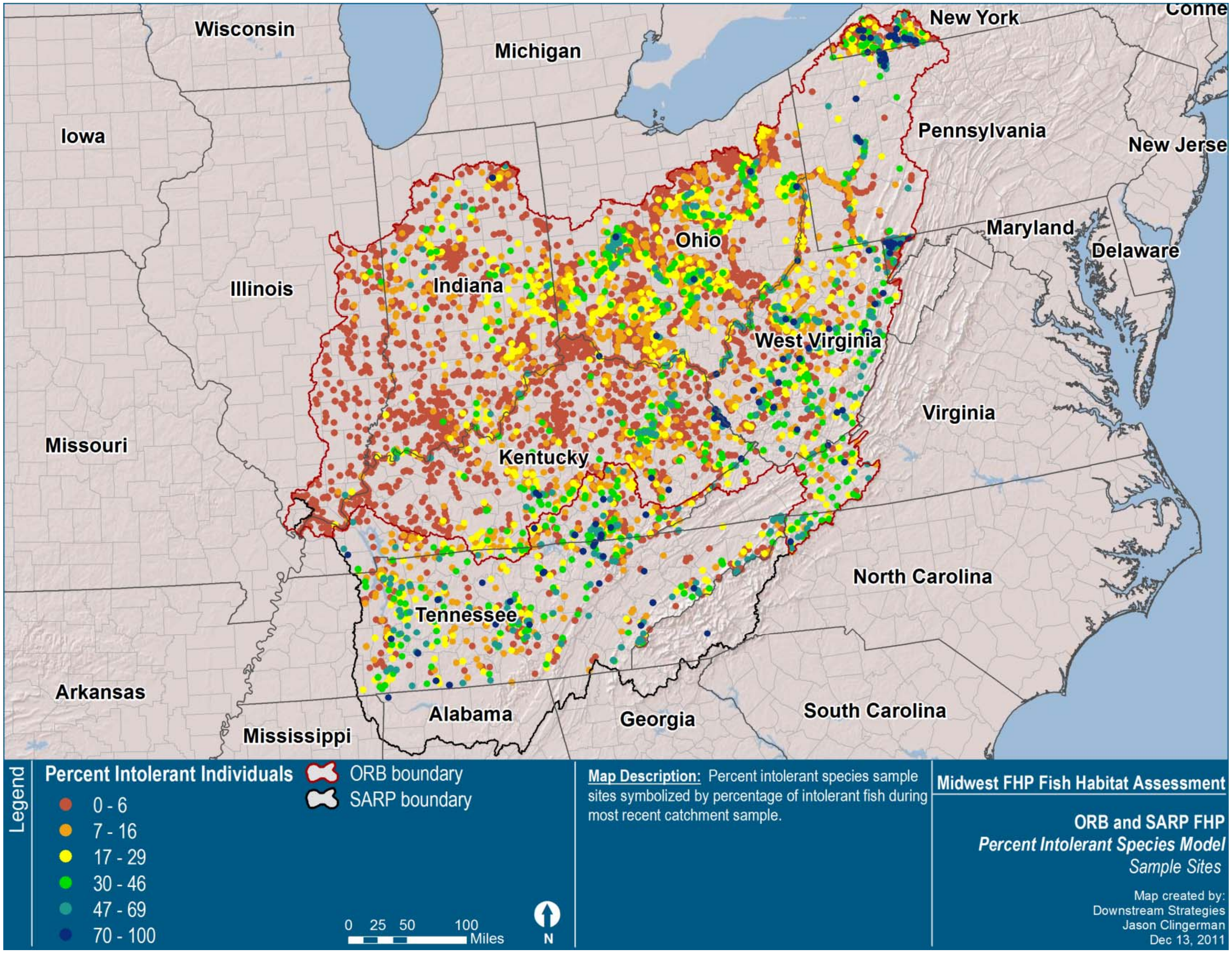
6.1 Modeling inputs

DS used a list of predictor variables selected by ORB and SARP to develop a ten-fold CV BRT model for percent intolerant fish at the 1:100k catchment scale. The model was used to produce maps of expected current percent intolerant fish distribution and maps of expected current natural habitat quality and anthropogenic stress at the 1:100k scale throughout the extents of both FHPs.

DS cooperated with ORB and SARP to arrive at a list of landscape-based habitat variables used to predict percent intolerant fish throughout the region; ultimately, those variables were also used for characterizing habitat quality and anthropogenic stress. From an initial suite of 372 catchment attributes, DS and the FHPs compiled a list of 92 predictors for evaluation. From that list, 46 variables were removed due to statistical redundancy ($r > 0.6$) or logical redundancy, resulting in a final list of 46 predictor variables for the BRT model and assessment. See Appendix A for a full data dictionary and the metadata document for variable processing notes.

ORB and SARP provided DS with a dataset for percent intolerant fish comprised of 6,186 observations collected over a time frame spanning 1996 to 2010. Intolerant fish in this analysis includes species that are intolerant to human disturbance; a list of species considered intolerant can be found in Table 24 (located in Appendix A). Figure 75 maps all of the sampling sites that were used to construct the model and outlines all of the 1:100k catchments to which the modeling outputs were applied.

Figure 75: Percent intolerant fish modeling area and sampling sites



6.2 Modeling process

6.2.1 *Predictive performance*

The final selected model was comprised of 4,750 trees. The model had a CV correlation statistic of 0.683±0.005.

6.2.2 *Variable influence*

The BRT output includes a list of the predictor variables used in the model ordered and scored by their relative importance. The relative importance values are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman and Meulman, 2003). The relative influence score is scaled so that the sum of the scores for all variables is 100, where higher numbers indicate greater influence. Of the 46 predictor variables used to develop the percent intolerant fish model, 45 had a relative influence value greater than zero (Table 14). The five most influential predictors were all natural habitat variables and accounted for over 51% of the total influence in the model:

- network drainage area,
- Level III Ecoregion,
- mean annual precipitation,
- network baseflow index, and
- slope of catchment flowline.

The five most influential anthropogenic stressors, which accounted for over 15% of the total influence, were:

- network forested land cover,
- network pasture land cover,
- network impervious surface cover,
- network surface water consumption, and
- network density of cattle.

Network drainage area, the single most important variable in terms of relative influence, contributed almost 15% of the total influence.

Table 14: Relative influence of all variables in the final percent intolerant fish model

Variable code	Variable description	Relative influence
cumdrainag	Network drainage area	14.77
eco_code3	Level III Ecoregion	12.85
precip	Mean annual precipitation	8.95
BFI_meanc	Network mean baseflow index	8.58
slope	Slope of catchment flowline	6.08
forpc	Network forested land cover	5.38
minelevraw	Minimum catchment elevation	5.29
wetlandpc	Network wetland land cover	3.55
pastpc	Network pasture land cover	2.74
temp	Mean annual air temperature	2.61
imp06c	Network impervious surface cover	2.52
water_swc	Network surface water consumption	2.44
soil2pc	Network soil group B, B/D cover	2.43
cattlec	Network density of cattle	2.41
ripdisp	Local riparian disturbance index score	2.31
water_gwc	Network groundwater consumption	2.19
roadcrc_den	Network density of road crossings	1.83
brock6pc	Network sandstone bedrock geology land cover	1.78
grasspc	Network grassland land cover	1.37
devp	Local developed land cover	1.19
roadcr_den	Local density of road crossings	1.07
cropspc	Network rowcrop land cover	1.04
brock7pc	Network shale bedrock geology cover	0.95
soil3pc	Network soil group C,C/D cover	0.68
surf3pc	Network alluvium surficial geology cover	0.67
damsc_den	Network density of dams	0.55
brock1pc	Network carbonate bedrock geology cover	0.48
surf4pc	Network lacustrine surficial geology cover	0.43
brock4pc	Network metamorphic bedrock geology cover	0.33
brock5pc	Network sand/gravel bedrock geology cover	0.31
soil4pc	Network soil group D cover	0.31
surf4p	Local lacustrine surficial geology cover	0.29
surf6pc	Network residuum surficial geology cover	0.27
surf7pc	Network clay surficial geology cover	0.26
dams_den	Local density of dams	0.19
surf8p	Local colluvium surficial geology cover	0.18
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.16
soil4p	Local soil group D cover	0.14
brock2p	Local felsic/igneous bedrock geology cover	0.10
mines_den	Local density of mines	0.09
surf5pc	Network loess surficial geology cover	0.08
TRI_den	Local density of Toxic Release Inventory sites	0.08
surf2p	Local outwash surficial geology cover	0.04
brock3p	Local mafic/igneous bedrock geology cover	0.01
brock4p	Local metamorphic bedrock geology cover	0.01
CERC_den	Local density of Superfund sites	0.00

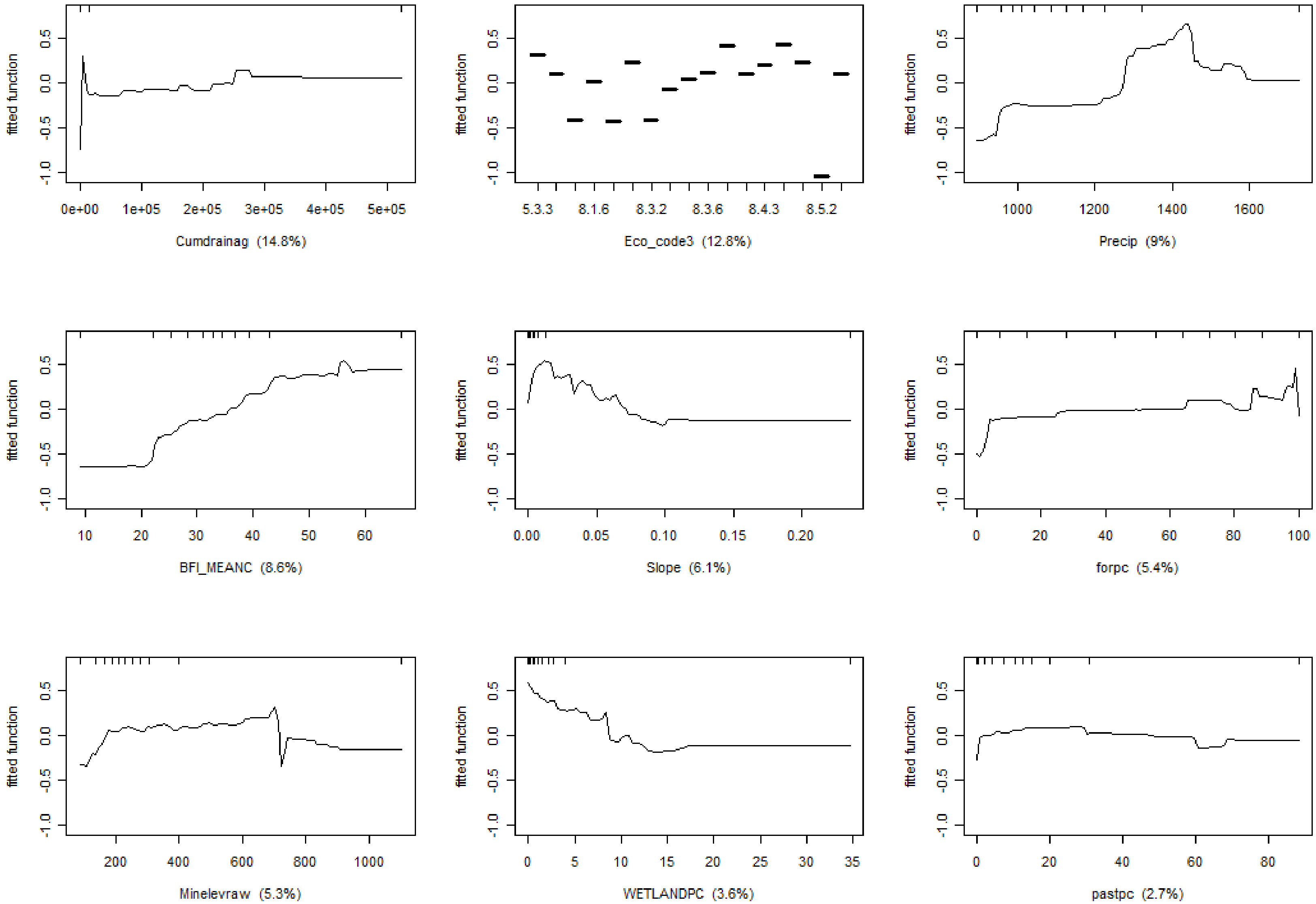
Note: Individual variables are highlighted according to whether they were determined to be anthropogenic in nature (red highlight) or natural (green highlight).

6.2.3 ***Variable functions***

The BRT output also contains quantitative information on partial dependence functions that can be plotted to visualize the effect of each individual predictor variable on the response after accounting for all other variables in the model. Similar to the interpretation of traditional regression coefficients, the function plots are not always a perfect representation of the relationship for each variable, particularly if interactions are strong or predictors are strongly correlated. However, they do provide a useful and objective basis for interpretation (Friedman, 2001; Friedman and Meulman, 2003).

These plots show the trend of the response variable (y-axis) as the predictor variable (x-axis) changes. The response variable is transformed (usually to the logit scale) so that the magnitude of trends for each predictor variable's function plot can be accurately compared. The dash marks at the top of each function represent the deciles of the data used to build the model. The function plots for the nine most influential variables in the percent intolerant fish model (see Table 14 for reference) are illustrated in Figure 76 below. The plots for all 46 variables are shown in Appendix B.

Figure 76: Functional responses of the dependent variable to individual predictors of intolerant fish



Note: Only the top nine predictors, based on relative influence (shown in parentheses; see Appendix A for descriptions of variable codes), are shown here. See Appendix B for plots of remaining predictor variables.

6.3 Post-modeling

The variable importance table and partial dependence functions of the final BRT model were used to create the post-modeling indices of natural habitat quality and anthropogenic stress for percent intolerant fish. The CNQI was comprised of 28 variables with relative influence greater than zero that were classified as natural habitat features (Table 15). The CASI was comprised of 8 variables with relative influence greater than zero that were classified as anthropogenic habitat features (Table 16). To calculate the cumulative indices (i.e., CNQI and CASI), each of the individual natural or anthropogenic variables used in the two indices was converted to a metric by first applying the appropriate transformations, based on their function plots, and then rescaling the transformed measures to a 0 to 100 scale. To calculate the cumulative index from the individual metrics, the metrics were first multiplied by their appropriate weighting factors and then summed. The CNQI and CASI scores were a result of a rescaling of those weighted and summed metrics, again from 0 to 100.

6.3.1 *Variable weights*

Table 15 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CNQI. The five most influential factors in the CNQI were:

- network drainage area,
- Level III Ecoregion,
- mean annual precipitation,
- network baseflow index, and
- slope of catchment flowline.

Table 16 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CASI. The five most influential factors in the CASI were:

- network forested land cover,
- network impervious surface cover,
- network density of cattle,
- network groundwater consumption, and
- local developed land cover.

Table 15: Relative influence and weights for natural variables on intolerant fish

Variable code	Variable description	Relative influence	Weighting factor
cumdrainag	Network drainage area	14.77	1.00
eco_code3	Level III Ecoregion	12.85	0.87
precip	Mean annual precipitation	8.95	0.61
BFI_meanc	Network mean baseflow index	8.58	0.58
slope	Slope of catchment flowline	6.08	0.41
minelevraw	Minimum catchment elevation	5.29	0.36
wetlandpc	Network wetland land cover	3.55	0.24
temp	Mean annual air temperature	2.61	0.18
soil2pc	Network soil group B, B/D cover	2.43	0.16
brock6pc	Network sandstone bedrock geology land cover	1.78	0.12
brock7pc	Network shale bedrock geology cover	0.95	0.06
soil3pc	Network soil group C,C/D cover	0.68	0.05
surf3pc	Network alluvium surficial geology cover	0.67	0.05
brock1pc	Network carbonate bedrock geology cover	0.48	0.03
surf4pc	Network lacustrine surficial geology cover	0.43	0.03
brock4pc	Network metamorphic bedrock geology cover	0.33	0.02
brock5pc	Network sand/gravel bedrock geology cover	0.31	0.02
soil4pc	Network soil group D cover	0.31	0.02
surf4p	Local lacustrine surficial geology cover	0.29	0.02
surf6pc	Network residuum surficial geology cover	0.27	0.02
surf7pc	Network clay surficial geology cover	0.26	0.02
surf8p	Local colluvium surficial geology cover	0.18	0.01
soil4p	Local soil group D cover	0.14	0.01
brock2p	Local felsic/igneous bedrock geology cover	0.10	0.01
surf5pc	Network loess surficial geology cover	0.08	0.01
surf2p	Local outwash surficial geology cover	0.04	0.00
brock3p	Local mafic/igneous bedrock geology cover	0.01	0.00
brock4p	Local metamorphic bedrock geology cover	0.01	0.00

Table 16: Relative influence and weights for anthropogenic variables on intolerant fish

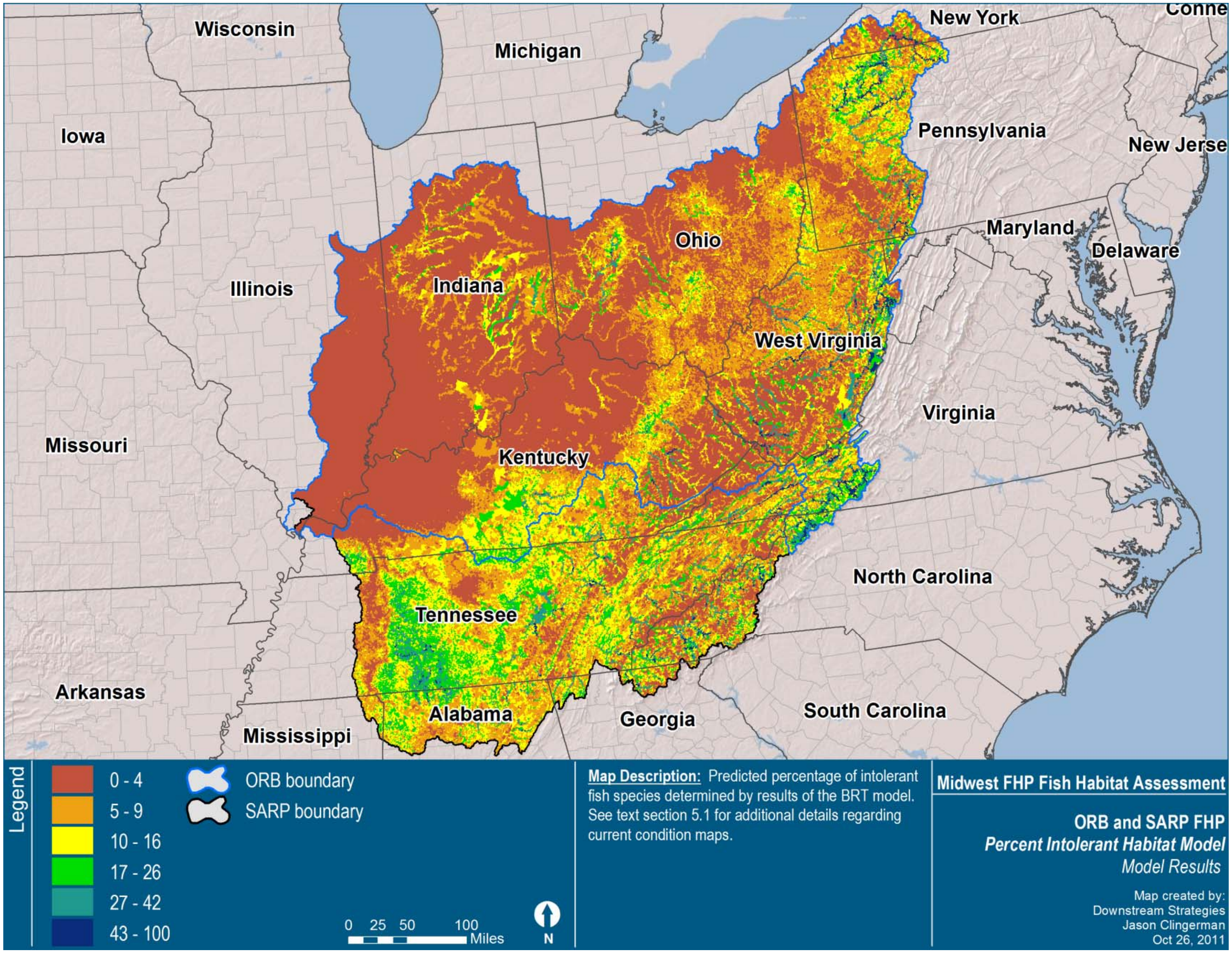
Variable code	Variable description	Relative influence	Weighting factor
forpc	Network forested land cover	5.38	1.00
imp06c	Network impervious surface cover	2.52	0.47
cattlec	Network density of cattle	2.41	0.45
water_gwc	Network groundwater consumption	2.19	0.41
devp	Local developed land cover	1.19	0.22
cropspc	Network rowcrop land cover	1.04	0.19
mines_den	Local density of mines	0.09	0.02
TRI_den	Local density of Toxic Release Inventory sites	0.08	0.01

6.4 Mapped Results

6.4.1 *Expected current conditions*

Percent intolerant fish was calculated for all 1:100k stream catchments in the study area using the BRT model. The predicted values ranged from 0 to 100. The mean predicted value was for the 226,919 total catchments was 7.5%. These results are mapped in Figure 77.

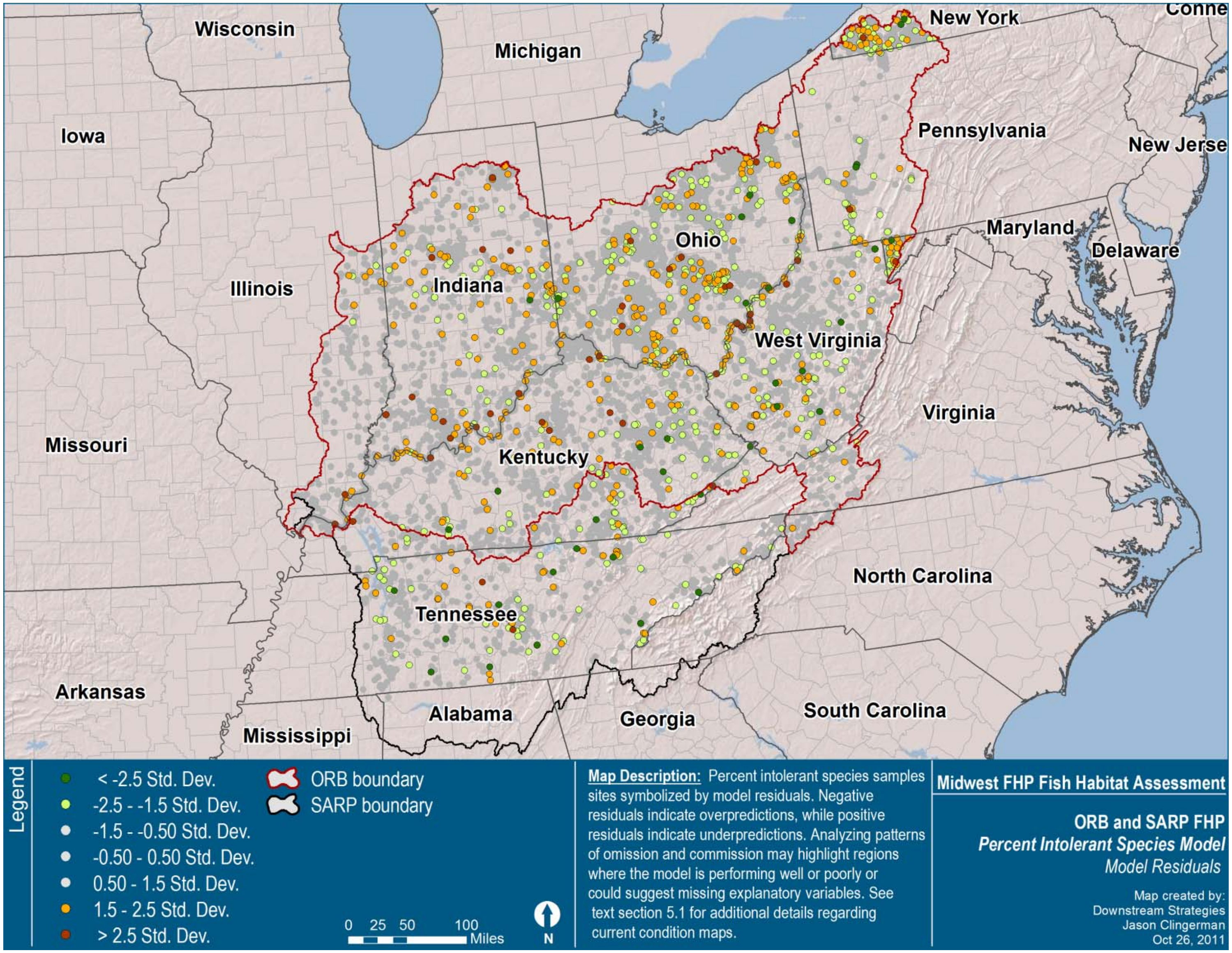
Figure 77: Expected percent intolerant fish distribution



6.4.2 ***Spatial variability in predictive performance***

Analyzing patterns of omission and commission may highlight regions where the model is performing well or poorly or could suggest missing explanatory variables (Figure 78). To assess omission and commission, residuals are also calculated by the BRT model. The residuals are a measure of the difference in the measured and modeled values (measured value *minus* modeled value). Negative residuals indicate overpredictions (predicting higher values than are true), while positive residuals indicate underpredictions (predicting lower values than are true).

Figure 78: Distribution of percent intolerant fish model residuals by sampling site



6.4.3 ***Indices of stress and natural quality***

Maps of CNQI and CASI illustrate the spatial distribution of natural habitat potential (i.e., CNQI score) and anthropogenic stress (i.e., CASI score) in the ORB and SARP. CNQI and CASI scores are mapped in Figure 79 and Figure 80, respectively. The top five most influential variables toward the calculation of CNQI are shown in Figure 81-Figure 85. The top five variables contributing toward the calculation of CASI are mapped in Figure 86-Figure 90. CNQI, CASI, and their metrics are all scaled on a 0-100 scale (see Section 6.3 for more details on CNQI and CASI calculation). For CNQI, higher values indicate higher natural quality, while higher values for CASI indicate higher levels of anthropogenic stress.

Figure 79: Cumulative natural quality index for intolerant fish

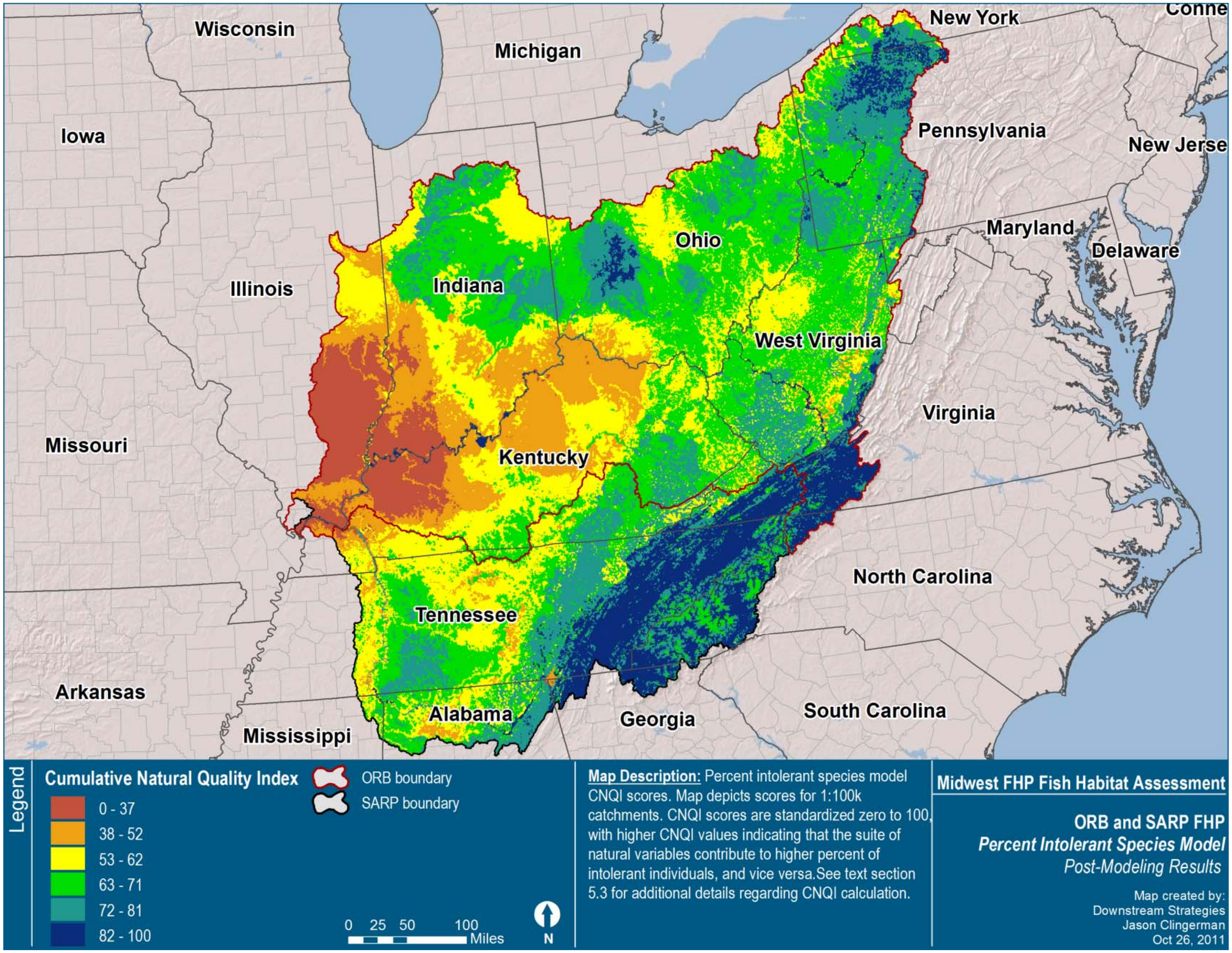


Figure 80: Cumulative anthropogenic stress index for intolerant fish

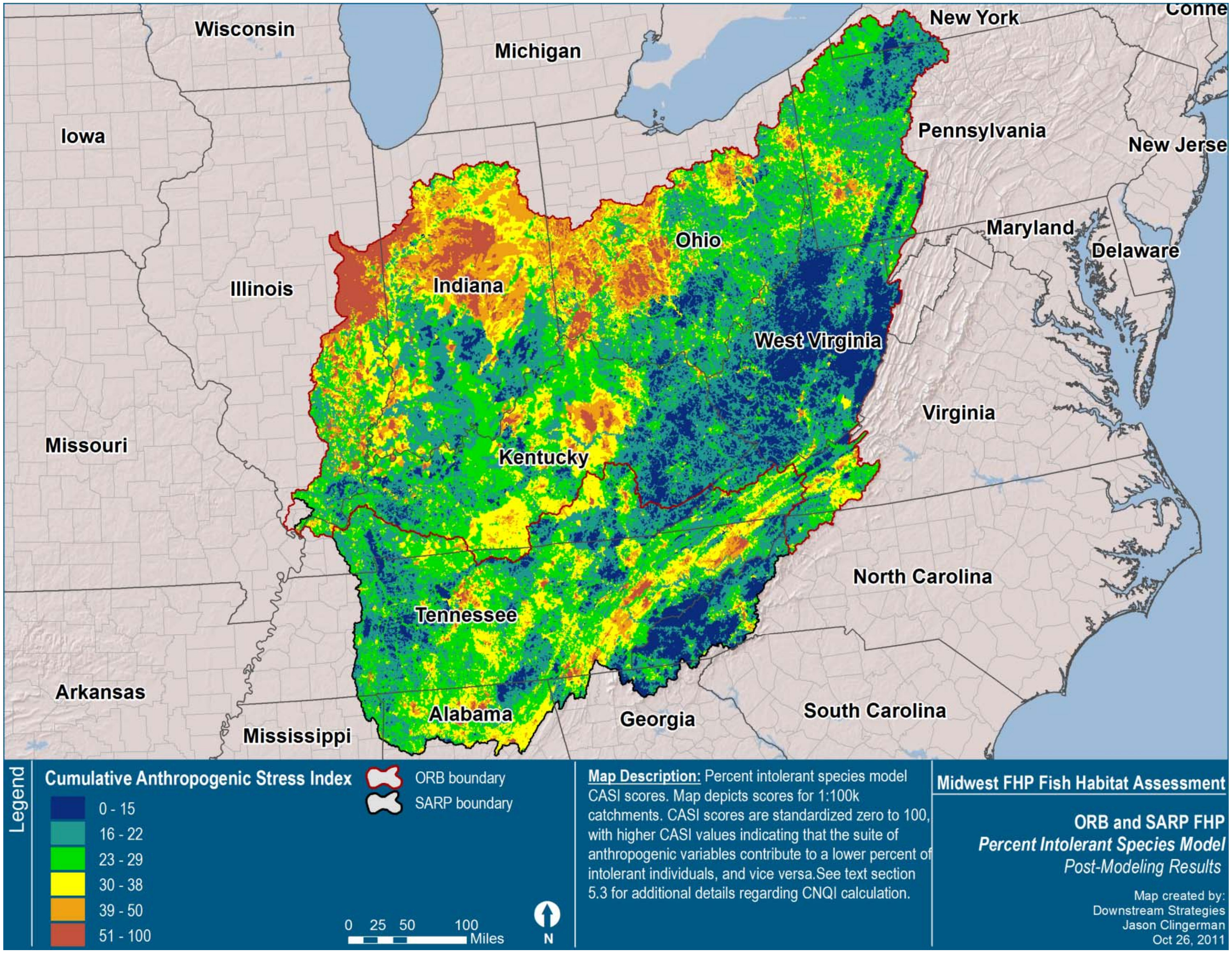


Figure 81: Most influential natural index metric for intolerant fish

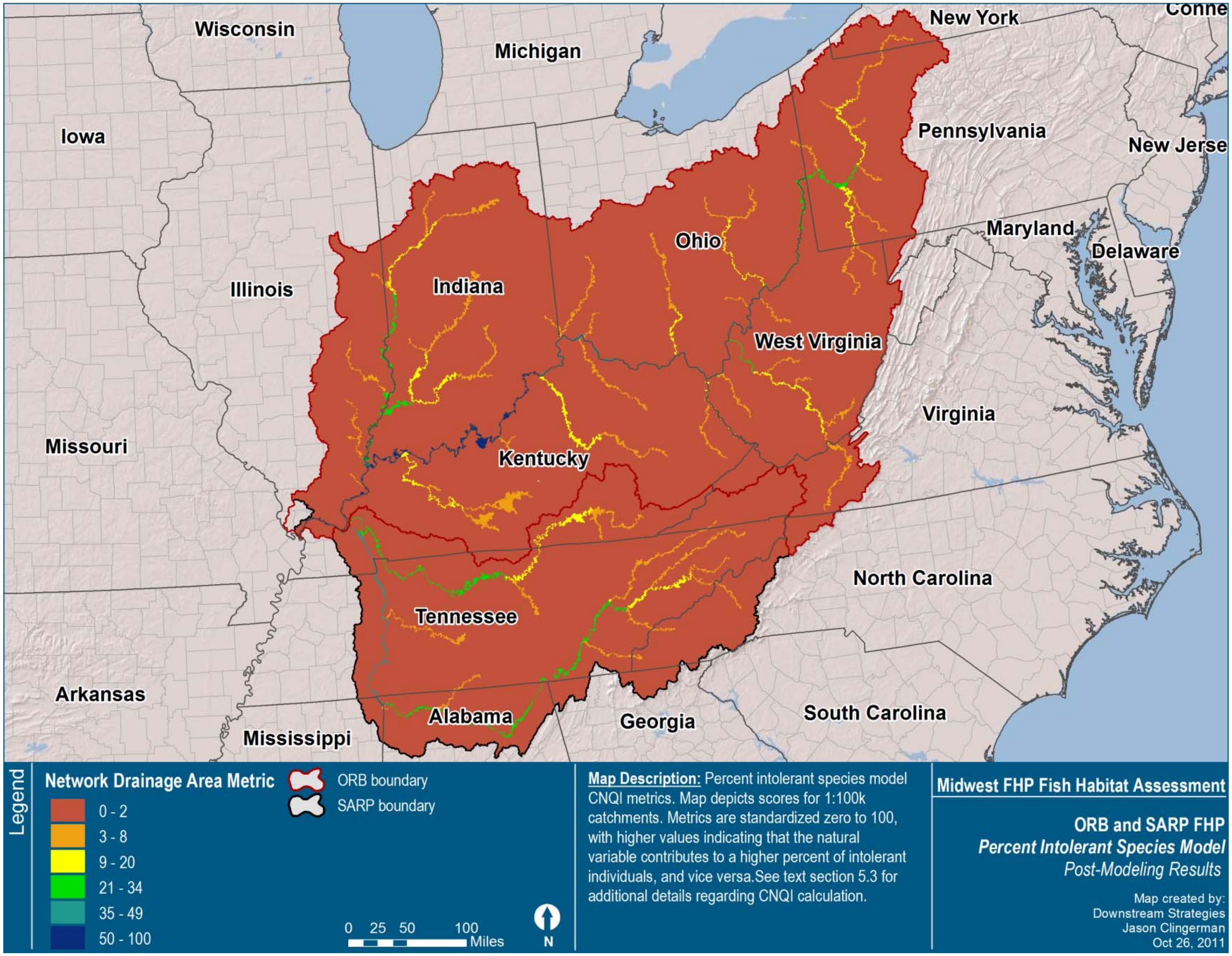


Figure 82: Second most influential natural index metric for intolerant fish

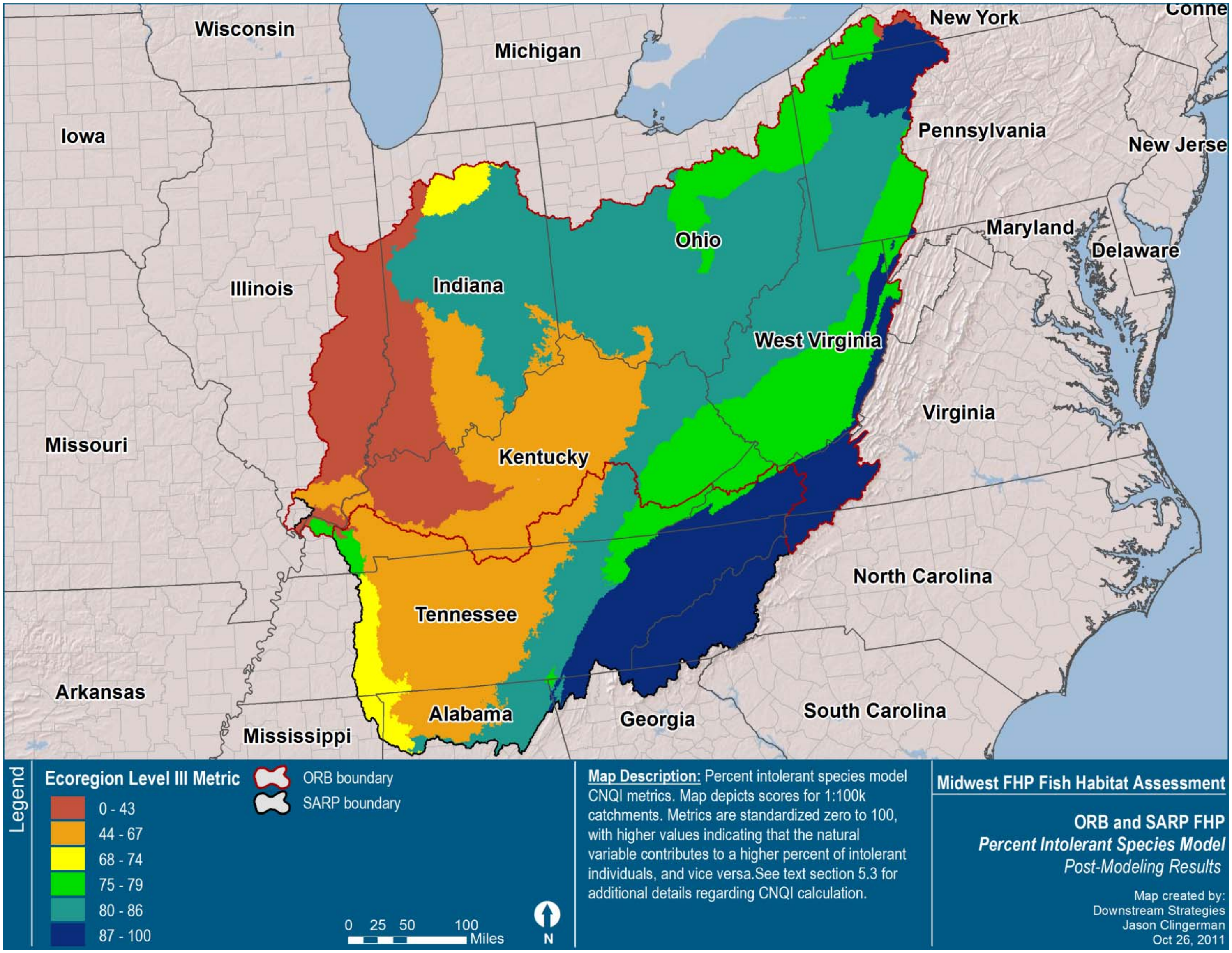


Figure 83: Third most influential natural index metric for intolerant fish

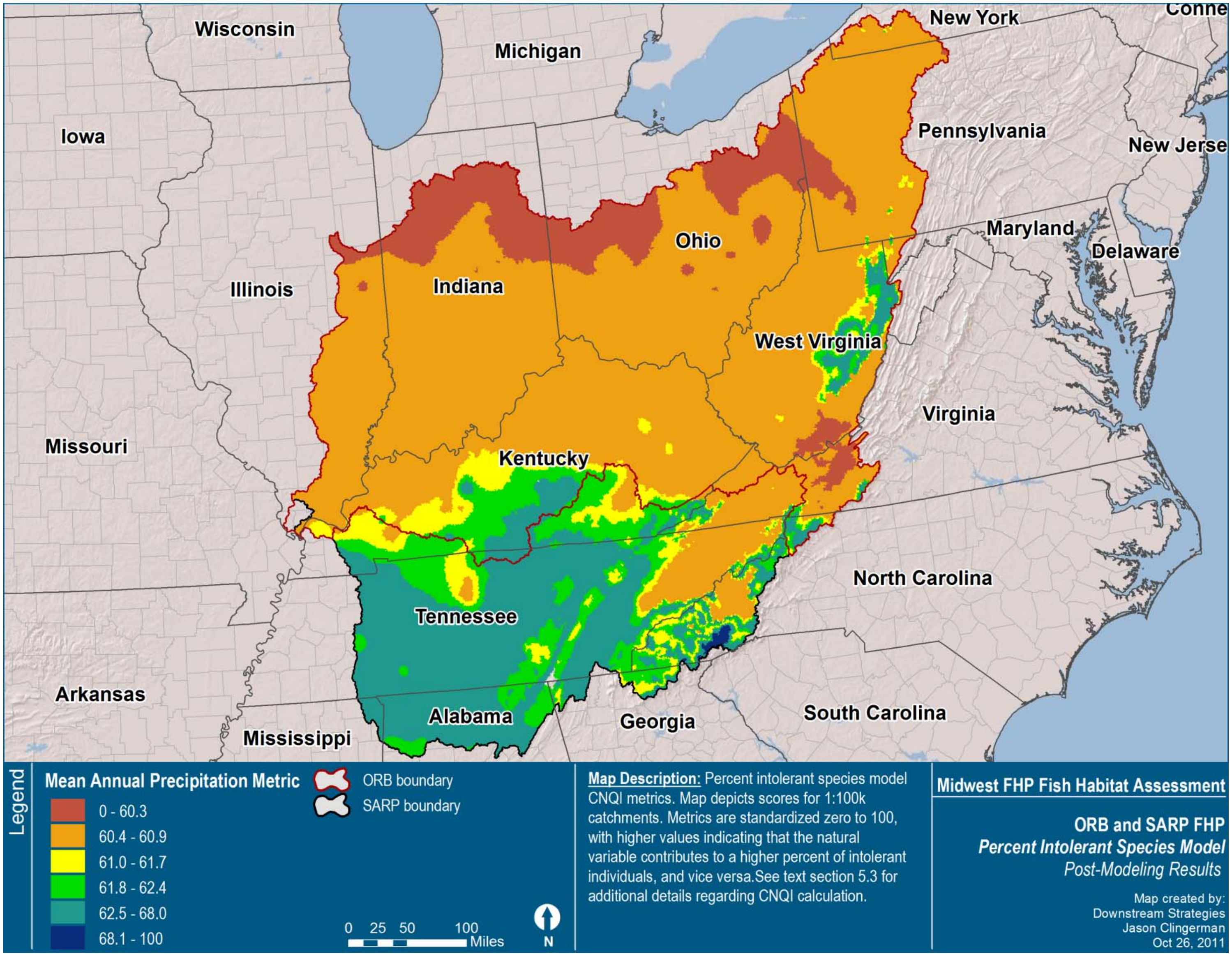


Figure 84: Fourth most influential natural index metric for intolerant fish

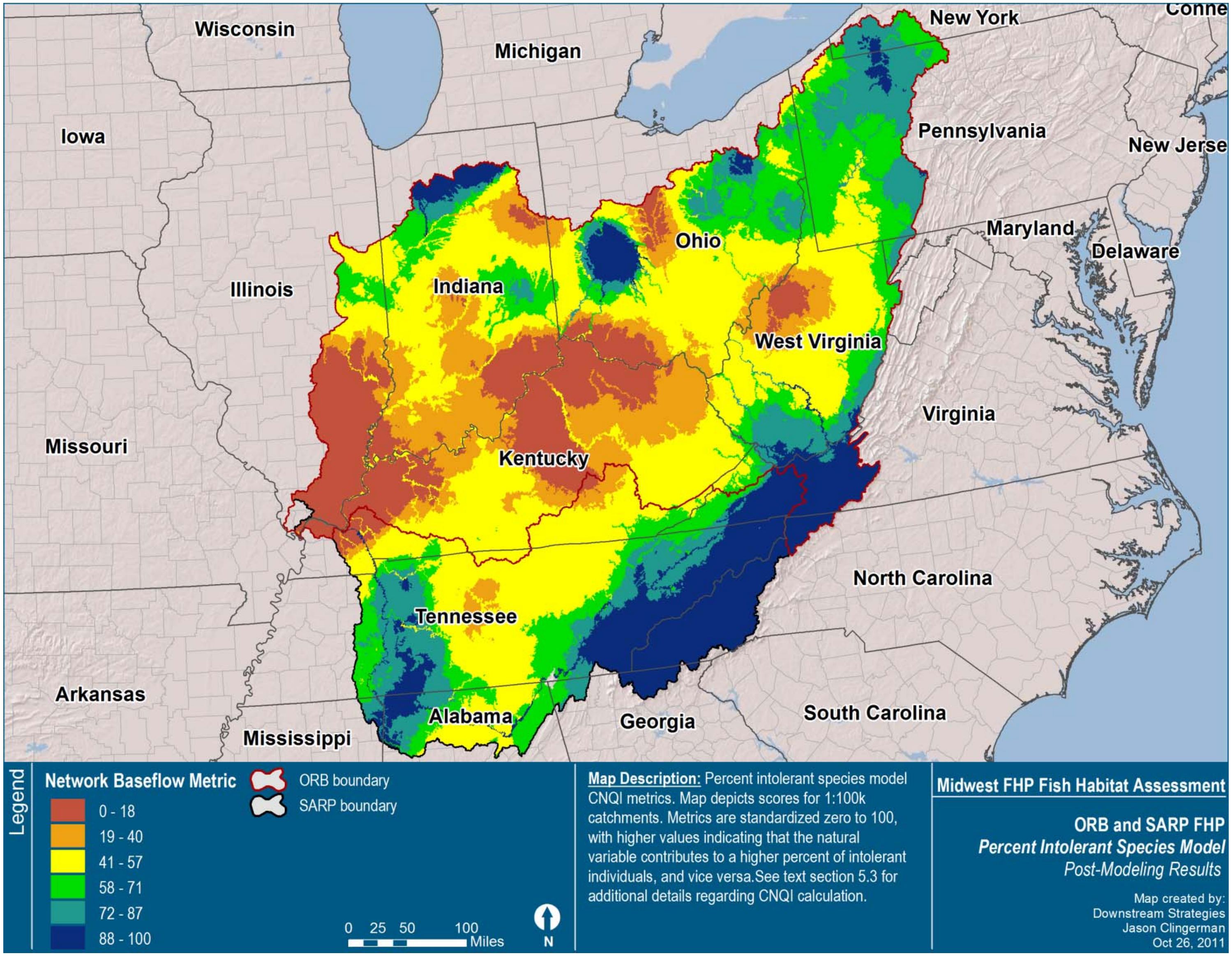


Figure 85: Fifth most influential natural index metric for intolerant fish

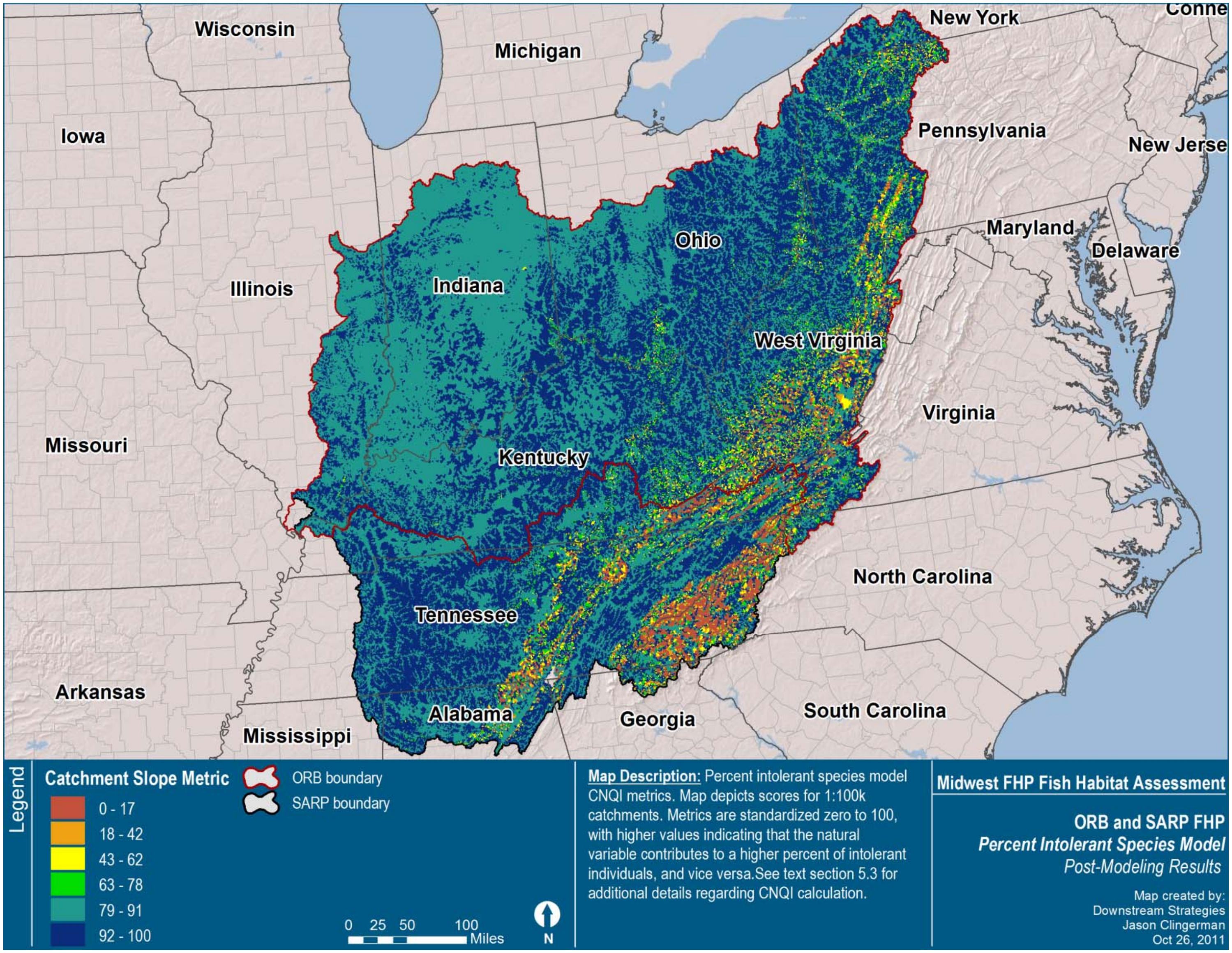


Figure 86: Most influential anthropogenic index metric for intolerant fish

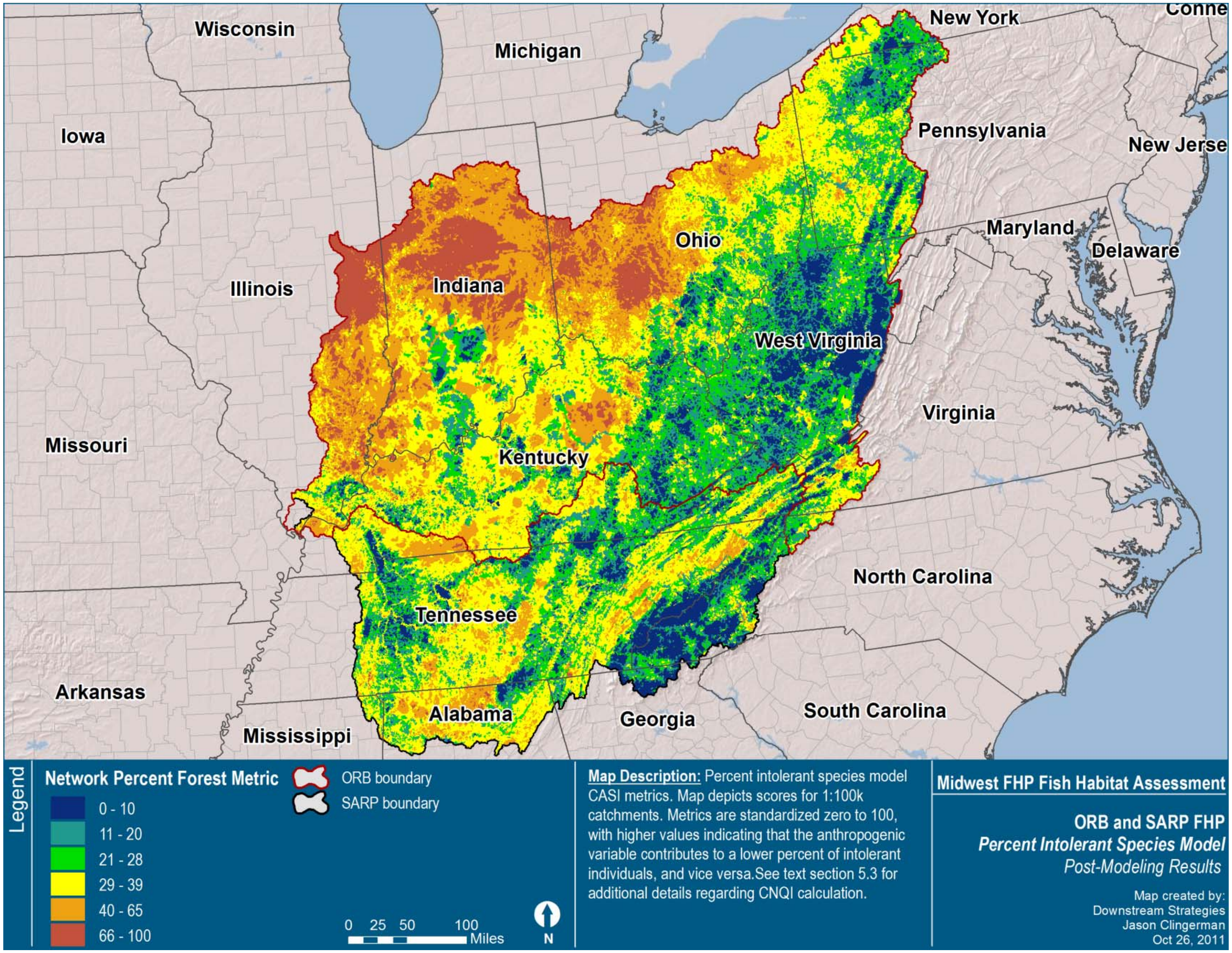


Figure 87: Second most influential anthropogenic index metric for intolerant fish

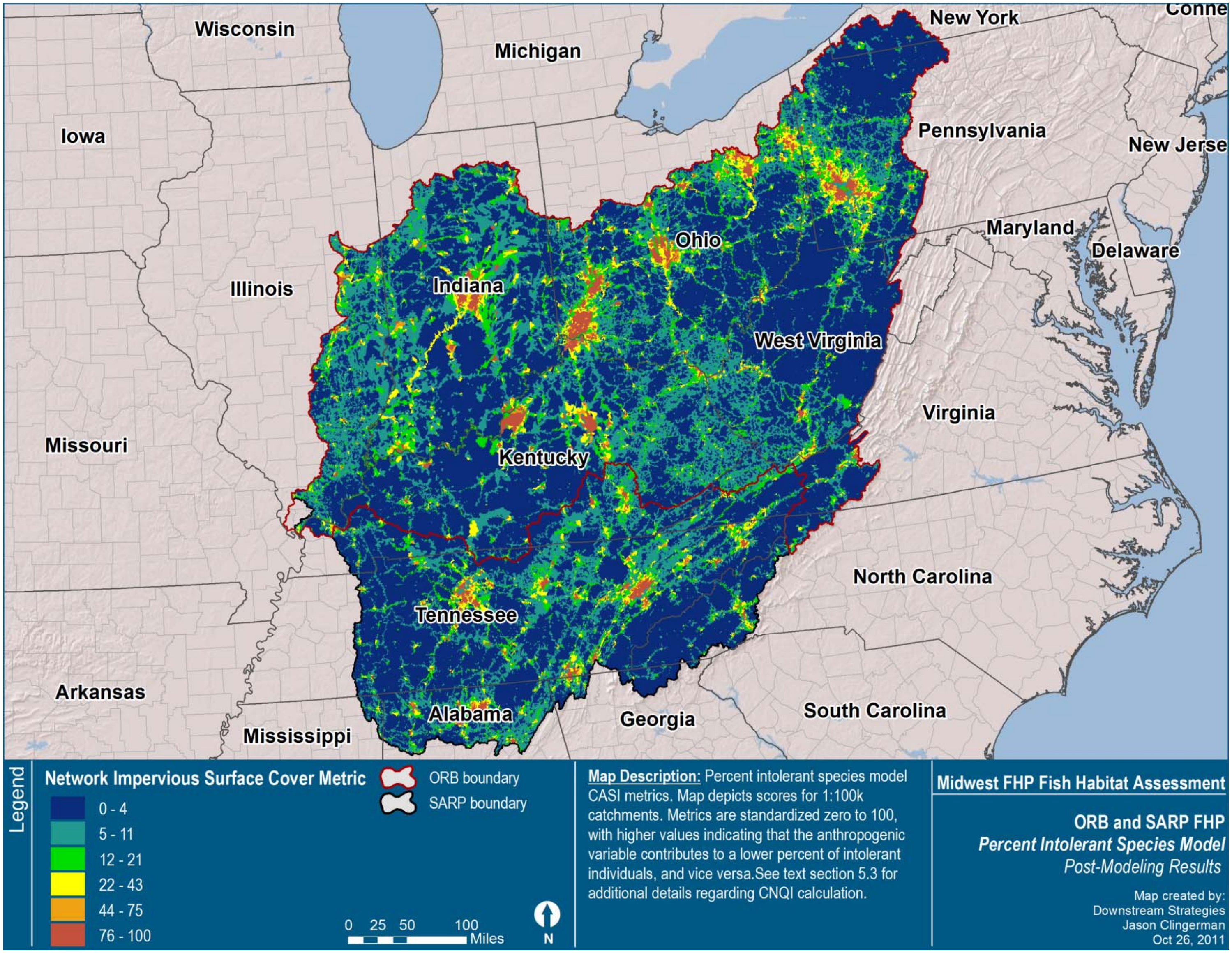


Figure 88: Third most influential anthropogenic index metric for intolerant fish

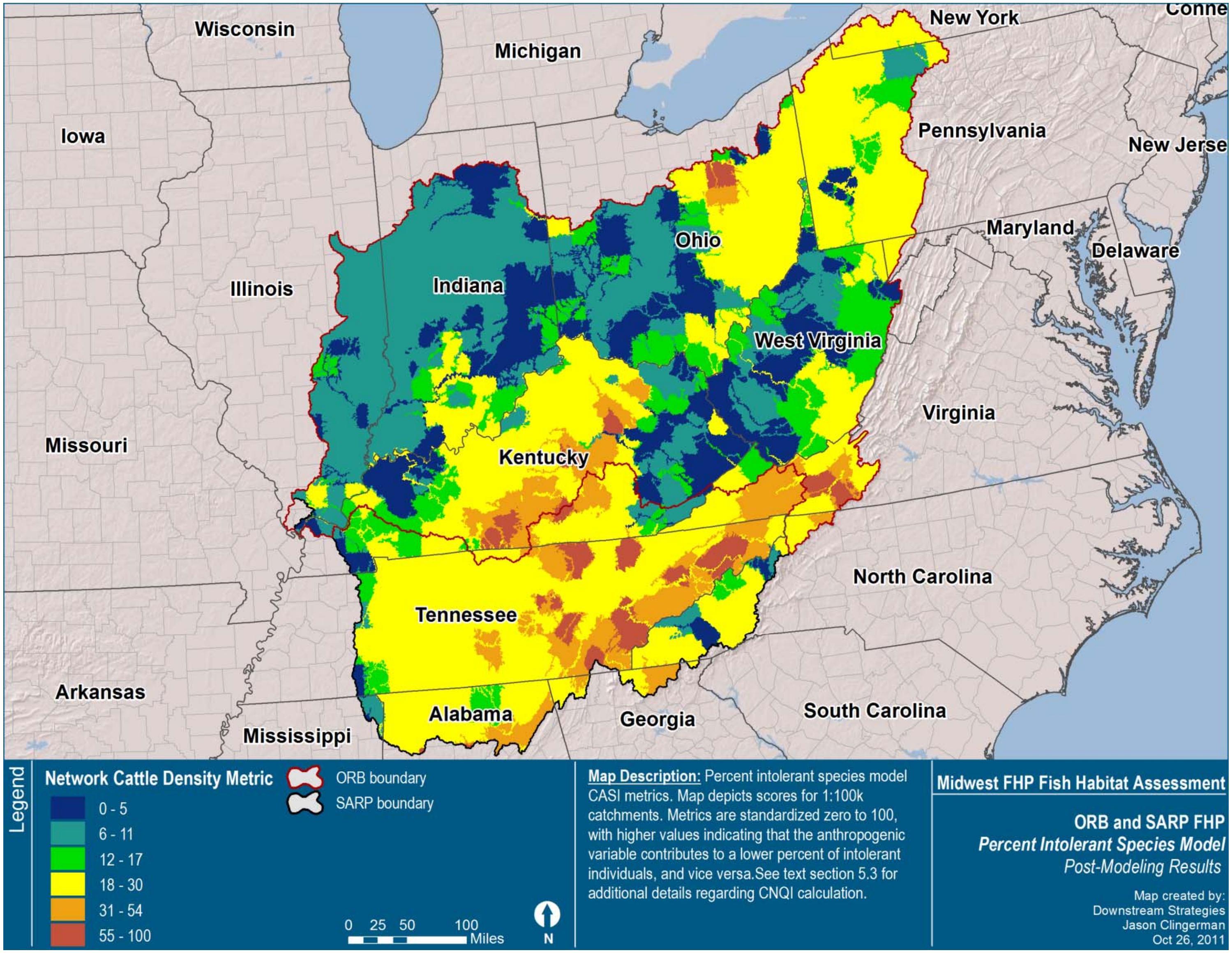


Figure 89: Fourth most influential anthropogenic index metric for intolerant fish

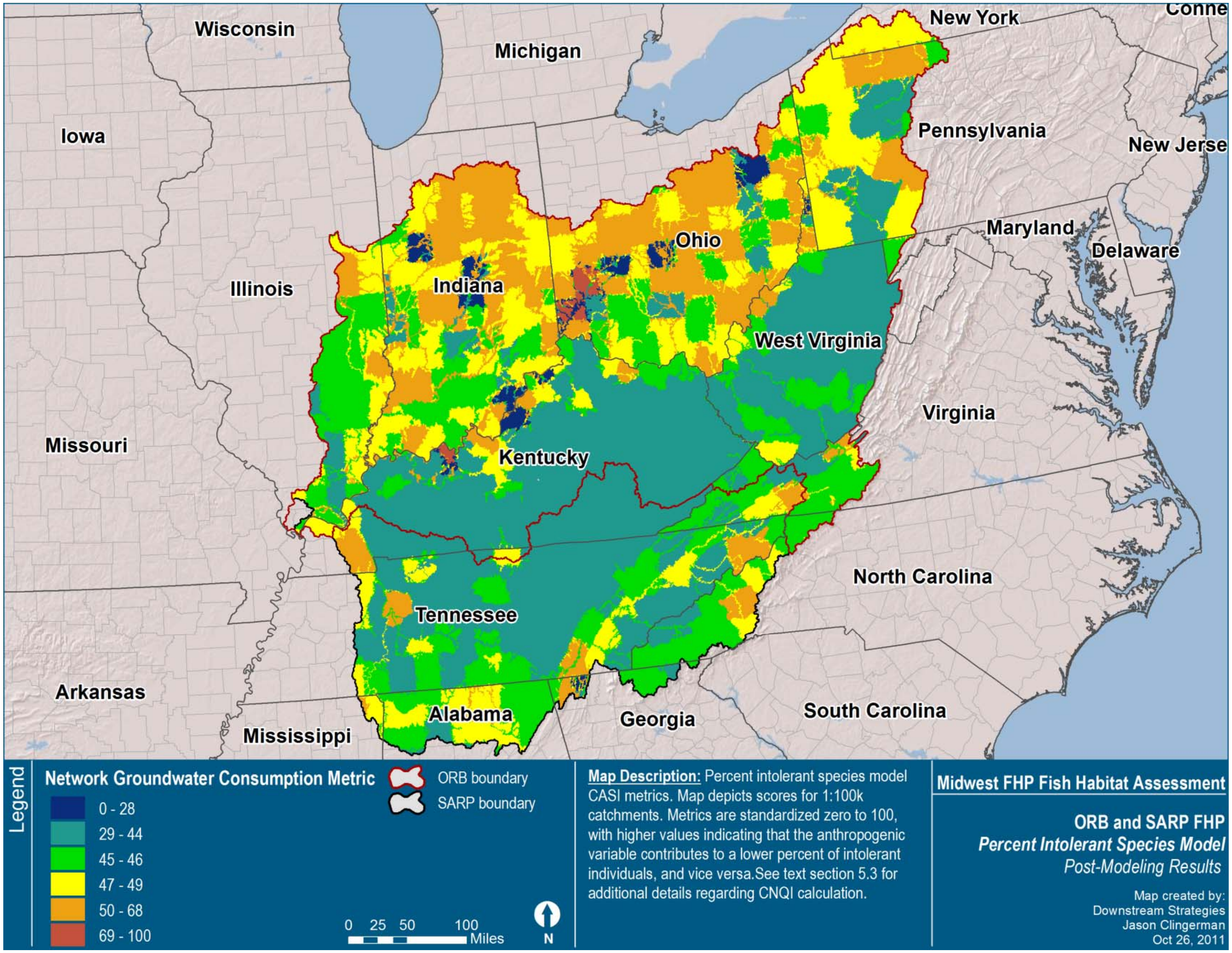
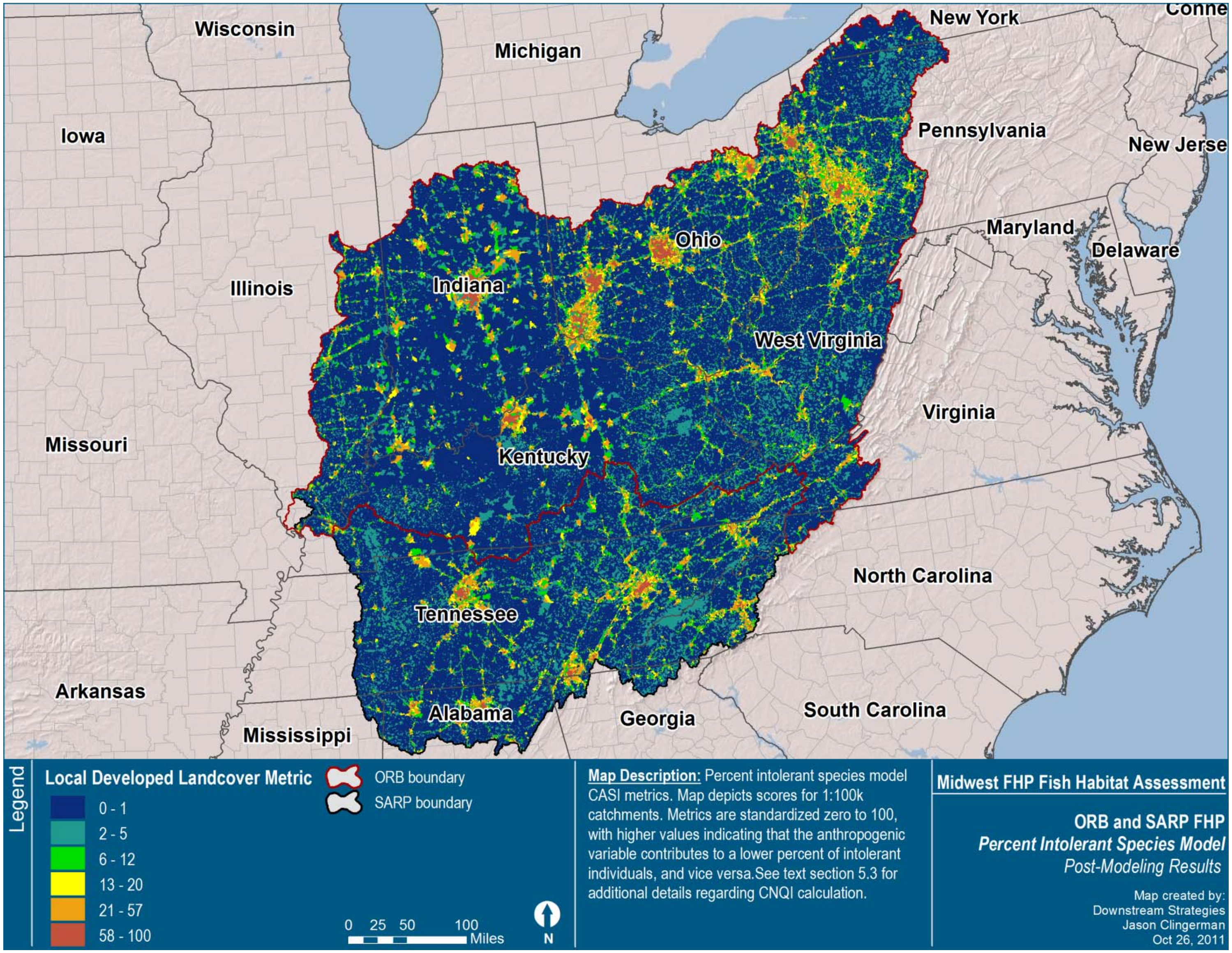


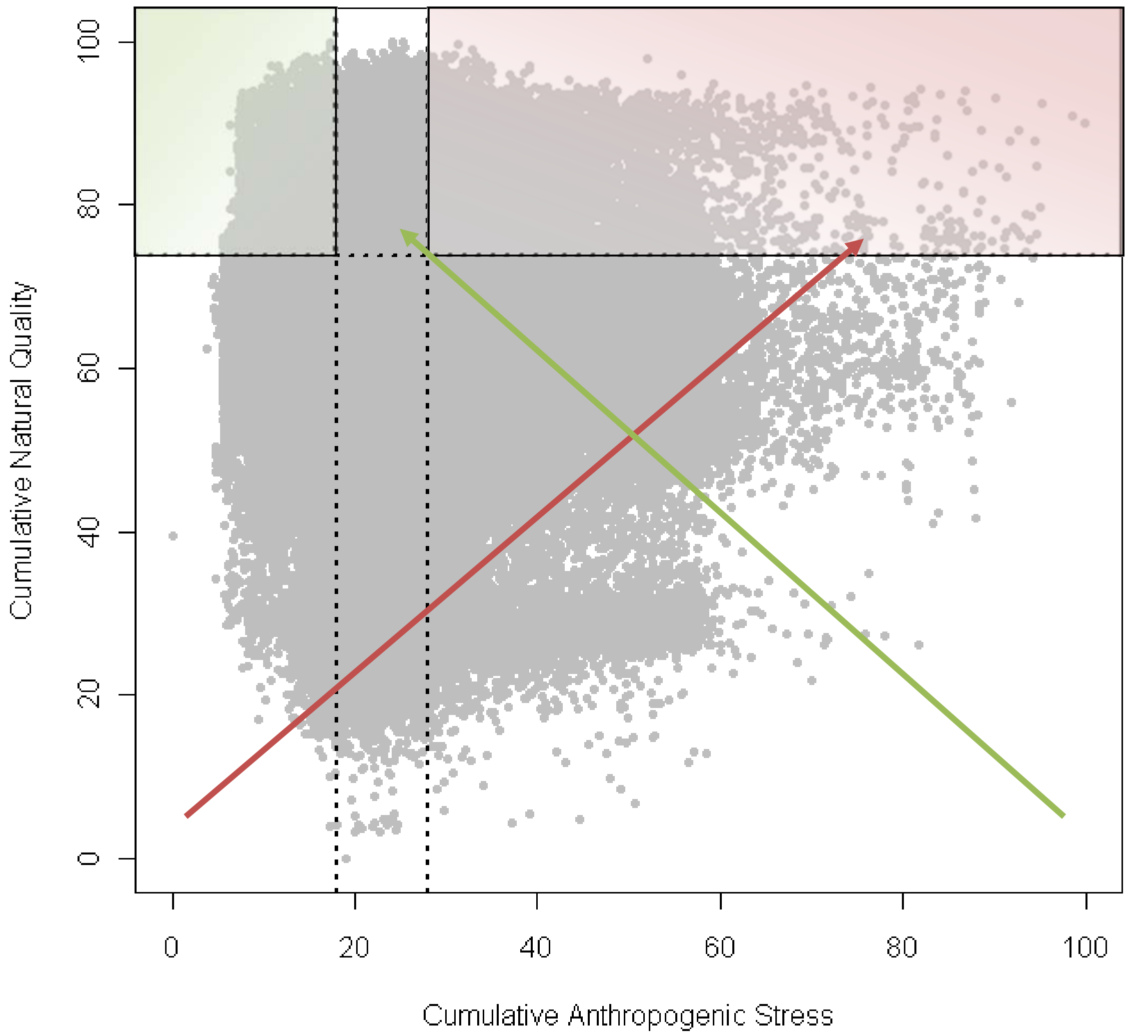
Figure 90: Fifth most influential anthropogenic index metric for intolerant fish



6.4.4 *Restoration and protection priorities*

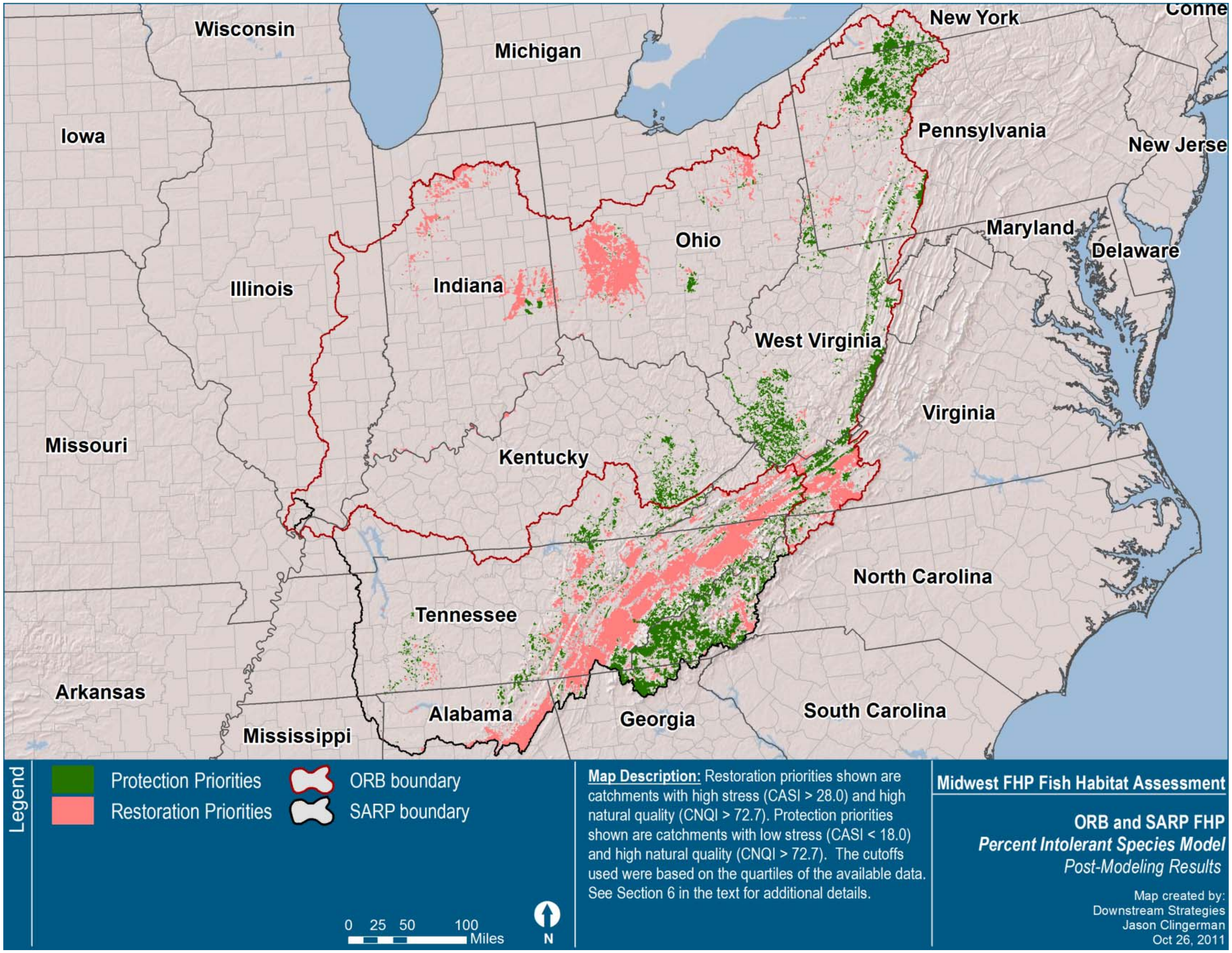
A plot of CNQI versus CASI values for all catchments in the study area (Figure 91) can be used as a reference when defining thresholds for categories of CNQI and CASI scores for use in the development of restoration and protection priorities. In the example shown (Figure 92), thresholds for restoration (high natural potential coupled with high anthropogenic stress) were set to CNQI greater than 72.7 and CASI greater than 28.0 (third quartiles). The thresholds used for protection (high natural potential and low anthropogenic stress) priorities were CNQI greater than 72.7 and CASI less than 18.0 (first quartile).

Figure 91: CNQI versus CASI values for all catchments for intolerant fish



Note: Breakpoints for CNQI and CASI classes in this example are denoted by dashed lines. The arrows indicate the directions of increasing potential protection (green arrow) or restoration (red arrow) priority. The red box indicates catchments defined as restoration priorities under the example scenario. The green box indicates catchments defined as protection priorities under the same scenario.

Figure 92: Restoration and protection priorities for intolerant fish



7. GREAT RIVERS SPECIES

7.1 Modeling inputs

DS used a list of predictor variables selected by ORB and SARP to develop a ten-fold CV BRT model for great rivers species at the 1:100k catchment scale. The model was used to produce maps of expected current great rivers species distribution and maps of expected current natural habitat quality and anthropogenic stress at the 1:100k scale throughout the extents of both FHPs.

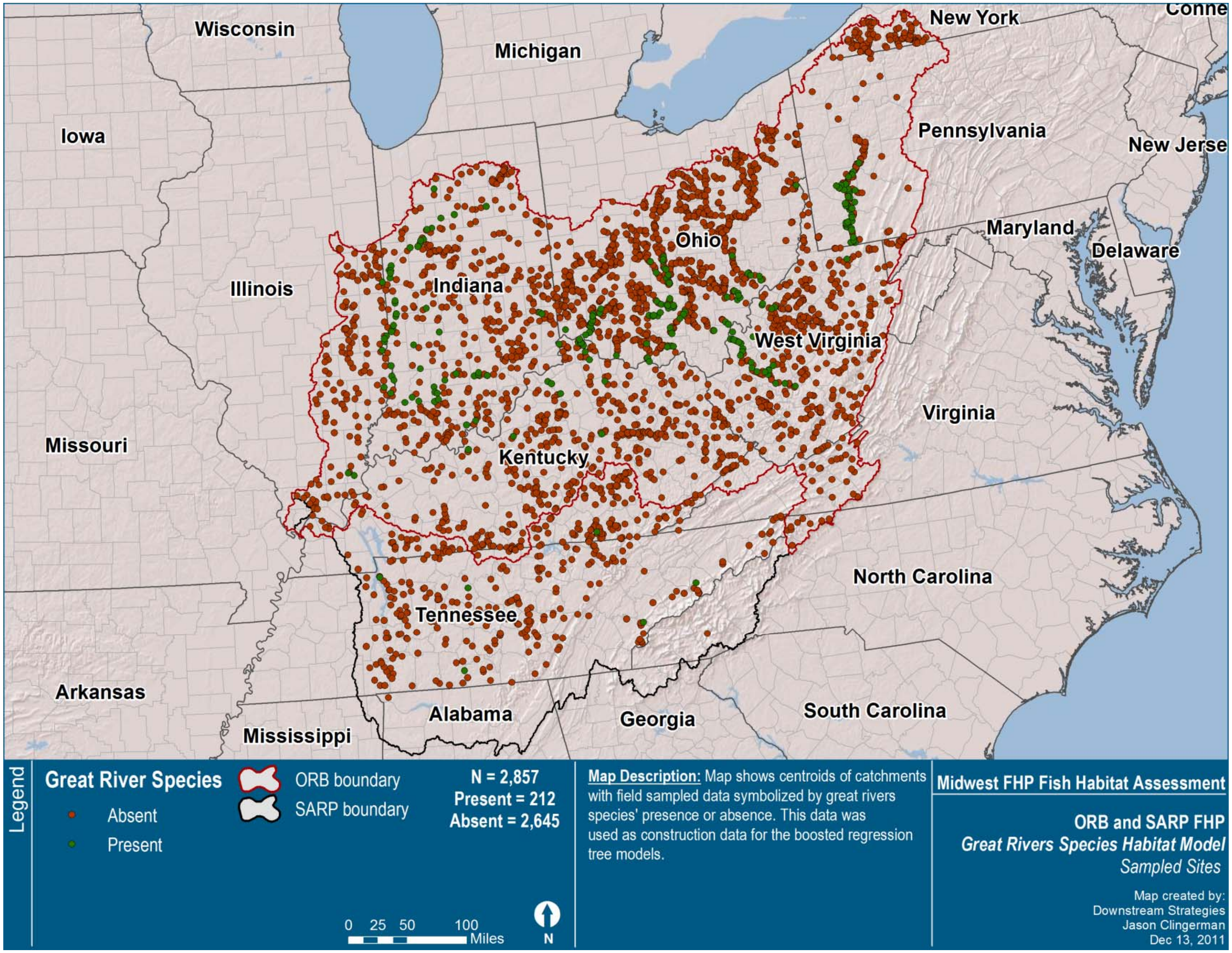
DS cooperated with ORB to arrive at a list of landscape-based habitat variables used to predict great rivers species throughout the region; ultimately, those variables were also used for characterizing habitat quality and anthropogenic stress. From an initial suite of 372 catchment attributes, DS and the FHPs compiled a list of 91 predictors for evaluation. From that list, 41 variables were removed due to statistical redundancy ($r > 0.6$) or logical redundancy, resulting in a final list of 50 predictor variables for the BRT model and assessment. See Appendix A for a full data dictionary and the metadata document for variable processing notes.

ORB provided DS with a presence-absence dataset for great rivers species comprised of 2,857 observations collected in streams between than 47 and 45,000 square kilometers in drainage area and over a time frame spanning 1996 to 2010. The goal of this model was to identify smaller streams that are important to great rivers species. Taking that into consideration, streams with drainage area less than 47 square kilometers were excluded since they would not be expected to have great rivers species, and streams with drainage area greater than 45,000 square kilometers were excluded since great rivers species would naturally be expected there. Table 17 lists the common and scientific names of the species considered as great rivers species. If any one of the species on this list was encountered in a sample, the stream reach was considered to be a presence. Figure 93 maps all of the sampling sites that were used to construct the model and outlines all of the 1:100k catchments to which the modeling outputs were applied.

Table 17: Great rivers species list

Scientific name	Common name
Acipenser fulvescens	Lake sturgeon
Anguilla rostrata	American eel
Atractosteus spatula	Alligator gar
Cycleptus elongatus	Blue sucker
Polyodon spathula	Paddlefish
Sander canadensis	Sauger
Scaphirhynchus albus	Pallid sturgeon
Scaphirhynchus platorhynchus	Shovelnose sturgeon

Figure 93: Great rivers species modeling area and sampling sites



7.2 Modeling process

7.2.1 *Predictive performance*

The final selected model was comprised of 3,050 trees. The model had a CV deviance statistic of 0.300±0.024, a CV correlation statistic of 0.621±0.016, and a CV ROC score of 0.934±0.005.

7.2.2 *Variable influence*

The BRT output includes a list of the predictor variables used in the model ordered and scored by their relative importance. The relative importance values are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman and Meulman, 2003). The relative influence score is scaled so that the sum of the scores for all variables is 100, where higher numbers indicate greater influence. Of the 50 predictor variables used to develop the great rivers model, 43 had a relative influence value greater than zero (Table 18). The five most influential predictors, which accounted for nearly 72% of the total influence in the model, were:

- network drainage area,
- local riparian disturbance,
- network carbonate bedrock,
- minimum catchment elevation, and
- mean annual air temperature.

The five most influential anthropogenic stressors, which accounted for over 22% of the total influence, were:

- local riparian disturbance,
- network surface water consumption,
- network density of Superfund sites,
- local impervious surface cover, and
- network pasture land cover.

The five most influential natural habitat variables, which contributed over 59% of the total influence, were:

- network drainage area,
- network carbonate bedrock,
- minimum catchment elevation,
- mean annual air temperature, and
- network B, B/D soil cover.

Network drainage area, the single most important variable in terms of relative influence, contributed almost 35% of the total influence.

Table 18: Relative influence of all variables in the final great rivers species model

Variable code	Variable description	Relative influence
cumdrainag	Network drainage area	34.35
ripdisp	Local riparian disturbance index score	14.48
brock1pc	Network carbonate bedrock geology cover	13.79
minelevraw	Minimum catchment elevation	5.98
temp	Mean annual air temperature	3.35
water_swc	Network surface water consumption	2.35
cercc_den	Network density of Superfund sites	2.22
imp06	Local impervious surface cover	2.04
soil2pc	Network soil type B, B/D cover	1.84
eco_code3	Level III Ecoregion	1.84
wetlandpc	Network wetland land cover	1.60
pastpc	Network pasture land cover	1.49
surf3pc	Network alluvium surficial geology cover	1.27
soil3pc	Network soil type C, C/D cover	0.92
imp06c	Network impervious surface cover	0.85
BFI_meanc	Network mean baseflow index	0.83
npdesc_den	Network density of National Pollutant Discharge Elimination System permits	0.81
cropspc	Network rowcrop land cover	0.79
dams_den	Local density of dams	0.77
roadcrc_den	Network density of road crossings	0.70
grasspc	Network grassland land cover	0.64
TRI_den	Local density of Toxic Release Inventory sites	0.62
slope	Slope of catchment flowline	0.60
cattlec	Network density of cattle	0.58
damsc_den	Network density of dams	0.49
brock3pc	Network mafic/igneous bedrock geology cover	0.49
pastp	Local pasture land cover	0.49
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.45
soil4pc	Network soil type D cover	0.44
forpc	Network forested land cover	0.43
brock7pc	Network shale bedrock geology cover	0.41
grassp	Local grassland land cover	0.40
soil1pc	Network soil type A, A/D cover	0.35
brock5pc	Network Sand/Gravel bedrock geology cover	0.35
roadcr_den	Local density of road crossings	0.33
minesc_den	Network density of mines	0.21
soil4p	Local soil type D cover	0.15
mines_den	Local density of mines	0.06
surf4pc	Network lacustrine surficial geology cover	0.06
surf2pc	Network outwash surficial geology cover	0.05
surf7pc	Network clay surficial geology cover	0.05
brock6p	Local sandstone bedrock geology cover	0.04
surf6pc	Network residuum surficial geology cover	0.02
CERC_den	Local density of Superfund sites	0.00
soil1p	Local soil type A, A/D cover	0.00
brock2p	Local felsic/igneous bedrock geology cover	0.00

brock3p	Local mafic/igneous bedrock geology cover	0.00
brock4p	Local metamorphic bedrock geology cover	0.00
surf4p	Local lacustrine surficial geology cover	0.00
surf5pc	Network loess surficial geology cover	0.00

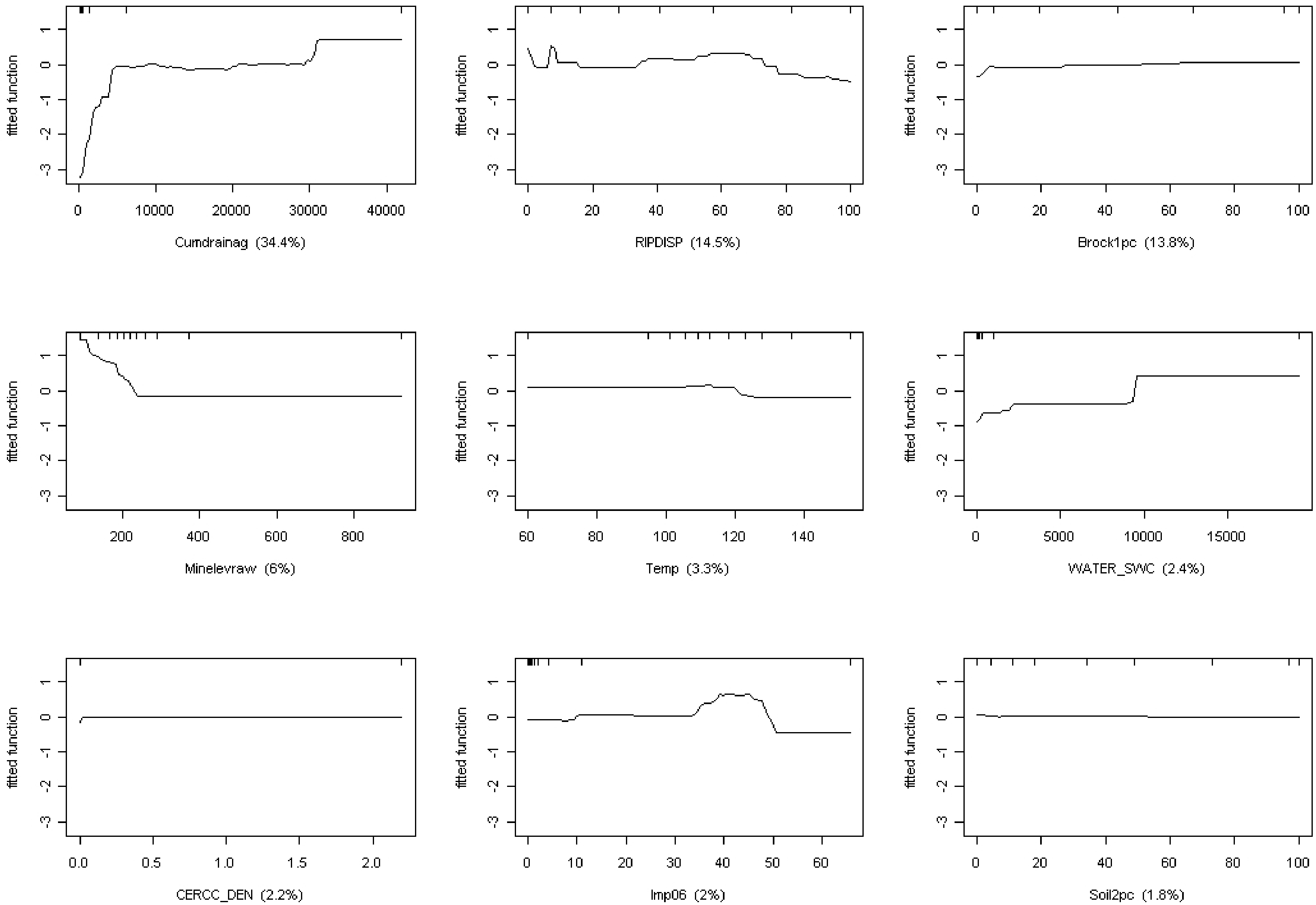
Note: Individual variables are highlighted according to whether they were determined to be anthropogenic in nature (red highlight) or natural (green highlight).

7.2.3 *Variable functions*

The BRT output also contains quantitative information on partial dependence functions that can be plotted to visualize the effect of each individual predictor variable on the response after accounting for all other variables in the model. Similar to the interpretation of traditional regression coefficients, the function plots are not always a perfect representation of the relationship for each variable, particularly if interactions are strong or predictors are strongly correlated. However, they do provide a useful and objective basis for interpretation (Friedman, 2001; Friedman and Meulman, 2003).

These plots show the trend of the response variable (y-axis) as the predictor variable (x-axis) changes. The response variable is transformed (usually to the logit scale) so that the magnitude of trends for each predictor variable's function plot can be accurately compared. The dash marks at the top of each function represent the deciles of the data used to build the model. The function plots for the nine most influential variables in the great rivers model (see Table 18 for reference) are illustrated in Figure 94 below. The plots for all 50 variables are shown in Appendix B.

Figure 94: Functional responses of the dependent variable to individual predictors of great rivers species



Note: Only the top nine predictors, based on relative influence (shown in parentheses; see Appendix A for descriptions of variable codes), are shown here. See Appendix B for plots of remaining predictor variables.

7.3 Post-modeling

The variable importance table and partial dependence functions of the final BRT model were used to create the post-modeling indices of natural habitat quality and anthropogenic stress for great rivers. The CNQI was comprised of 22 variables with relative influence greater than zero that were classified as natural habitat features (Table 19). The CASI was comprised of 9 variables with relative influence greater than zero that were classified as anthropogenic habitat features (Table 20). To calculate the cumulative indices (i.e., CNQI and CASI), each of the individual natural or anthropogenic variables used in the two indices was converted to a metric by first applying the appropriate transformations, based on their function plots, and then rescaling the transformed measures to a 0 to 100 scale. To calculate the cumulative index from the individual metrics, the metrics were first multiplied by their appropriate weighting factors and then summed. The CNQI and CASI scores were a result of a rescaling of those weighted and summed metrics, again from 0 to 100.

7.3.1 *Variable weights*

Table 19 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CNQI. The five most influential factors in the CNQI were:

- network drainage area,
- network carbonate bedrock,
- minimum catchment elevation,
- mean annual air temperature, and
- network B, B/D soil cover.

Table 20 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CASI. The five most influential factors in the CASI were:

- local riparian disturbance,
- local impervious surface cover,
- network pasture land cover,
- network impervious surface cover, and
- network grassland land cover.

Table 19: Relative influence and weights for natural variables on great rivers species

Variable code	Variable description	Relative influence	Weighting factor
Cumdrainag	Network drainage area	34.35	1.00
Brock1pc	Network carbonate bedrock geology cover	13.79	0.40
Minelevraw	Minimum catchment elevation	5.98	0.17
Temp	Mean annual air temperature	3.35	0.10
Soil2pc	Network soil type B, B/D cover	1.84	0.05
Eco_code3	Level III Ecoregion	1.84	0.05
WETLANDPC	Network wetland land cover	1.60	0.05
Surf3pc	Network alluvium surficial geology cover	1.27	0.04
Soil3pc	Network soil type C, C/D cover	0.92	0.03
BFI_MEANC	Network mean baseflow index	0.83	0.02
Slope	Slope of catchment flowline	0.60	0.02
Brock3pc	Network mafic/igneous bedrock geology cover	0.49	0.01
Soil4pc	Network soil type D cover	0.44	0.01
Brock7pc	Network shale bedrock geology cover	0.41	0.01
Soil1pc	Network soil type A, A/D cover	0.35	0.01
Brock5pc	Network sand/gravel bedrock geology cover	0.35	0.01
Soil4p	Local soil type D cover	0.15	0.00
Surf4pc	Network lacustrine surficial geology cover	0.06	0.00
Surf2pc	Network outwash surficial geology cover	0.05	0.00
Surf7pc	Network clay surficial geology cover	0.05	0.00
Brock6p	Local sandstone bedrock geology cover	0.04	0.00
Surf6pc	Network residuum surficial geology cover	0.02	0.00

Table 20: Relative influence and weights for anthropogenic variables on great rivers species

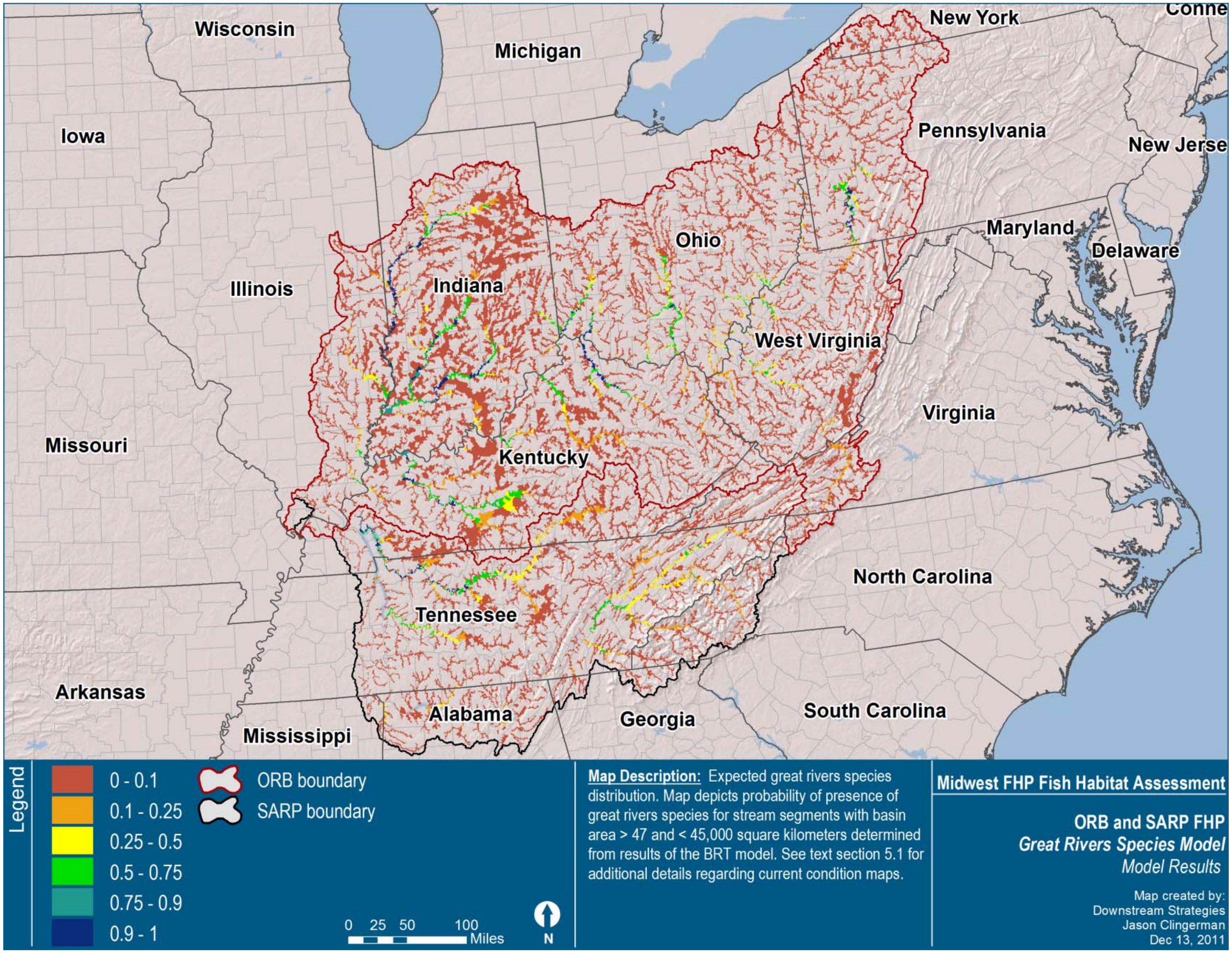
Variable code	Variable description	Relative influence	Weighting factor
Ripdisp	Local riparian disturbance index score	14.48	1.00
imp06	Local impervious surface cover	2.04	0.14
Pastpc	Network pasture land cover	1.49	0.10
imp06c	Network impervious surface cover	0.85	0.06
Grasspc	Network grassland land cover	0.64	0.04
TRI_den	Local density of Toxic Release Inventory sites	0.62	0.04
Cattlec	Network density of cattle	0.58	0.04
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.45	0.03
Forpc	Network forested land cover	0.43	0.03

7.5 Mapped Results

7.5.1 *Expected current conditions*

Great rivers species probability of presence was calculated for all 1:100k stream catchments in the study area using the BRT model. The predicted probability values ranged from 0 to 1. The mean predicted probability value for the 45,450 total catchments larger than 47 square kilometers drainage area and less than 45,000 square kilometers was 0.017. There were 1,314 catchments larger than 47 square kilometers and less than 45,000 square kilometers with a predicted probability of presence greater than 0.75, and 1,595 catchments larger than 47 square kilometers and less than 45,000 square kilometers where the probability of presence was between 0.5 and 0.75. These results are mapped in Figure 95.

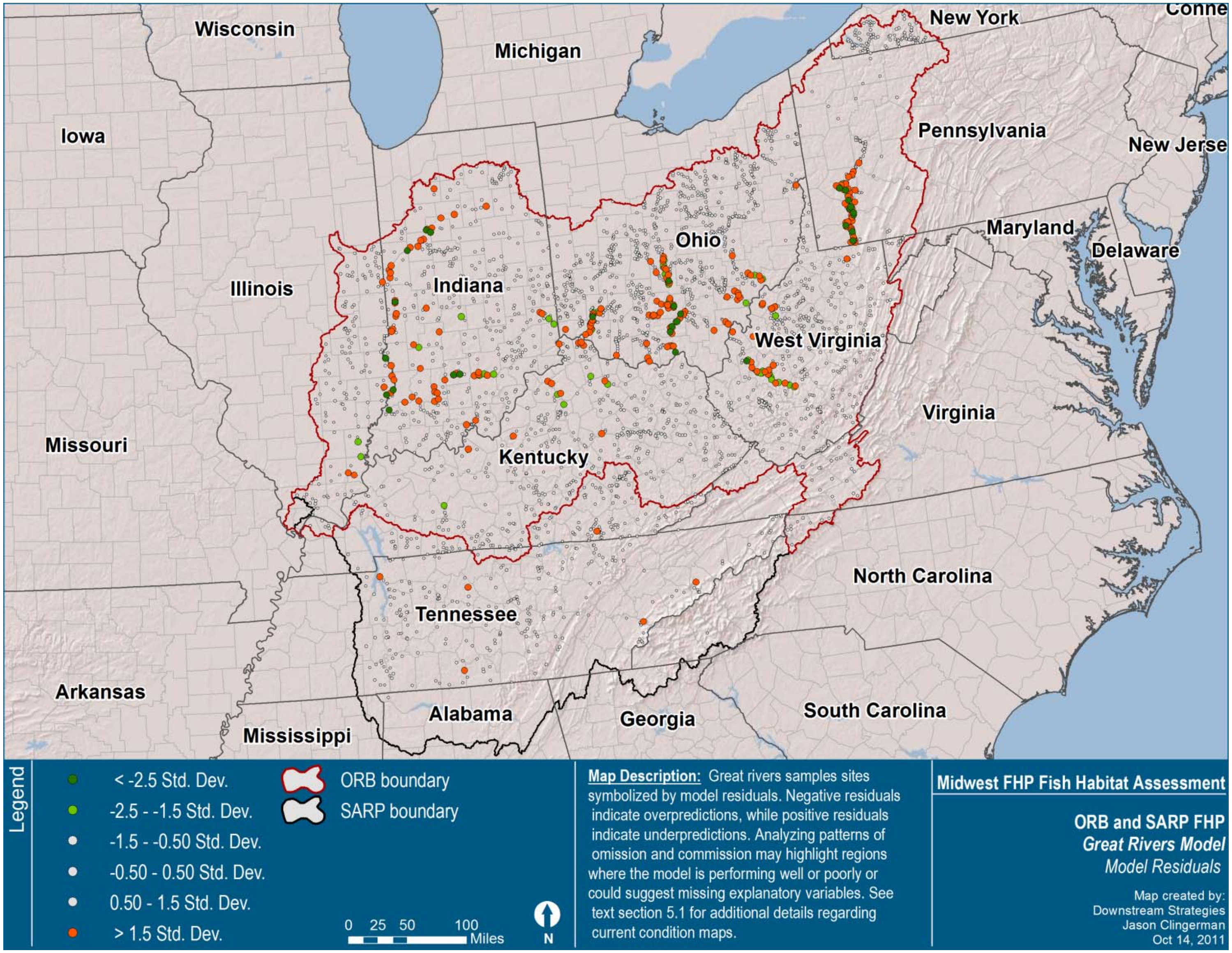
Figure 95: Expected great rivers species distribution



7.5.2 *Spatial variability in predictive performance*

Analyzing patterns of omission and commission may highlight regions where the model is performing well or poorly or could suggest missing explanatory variables (Figure 96). To assess omission and commission, residuals are also calculated by the BRT model. The residuals are a measure of the difference in the measured and modeled values (measured value *minus* modeled value). Negative residuals indicate overpredictions (predicting higher values than are true), while positive residuals indicate underpredictions (predicting lower values than are true).

Figure 96: Distribution of great rivers species model residuals by sampling site



7.5.3 *Indices of stress and natural quality*

Maps of CNQI and CASI illustrate the spatial distribution of natural habitat potential (i.e., CNQI score) and anthropogenic stress (i.e., CASI score) in the ORB and SARP. CNQI and CASI scores are mapped in Figure 97 and Figure 98, respectively. The top five most influential variables toward the calculation of CNQI are shown in Figure 99-Figure 103. The top five variables contributing toward the calculation of CASI are mapped in Figure 104-Figure 108. CNQI, CASI, and their metrics are all scaled on a 0-100 scale (see Section 7.3 for more details on CNQI and CASI calculation). For CNQI, higher values indicate higher natural quality, while higher values for CASI indicate higher levels of anthropogenic stress.

Figure 97: Cumulative natural quality index for great rivers species

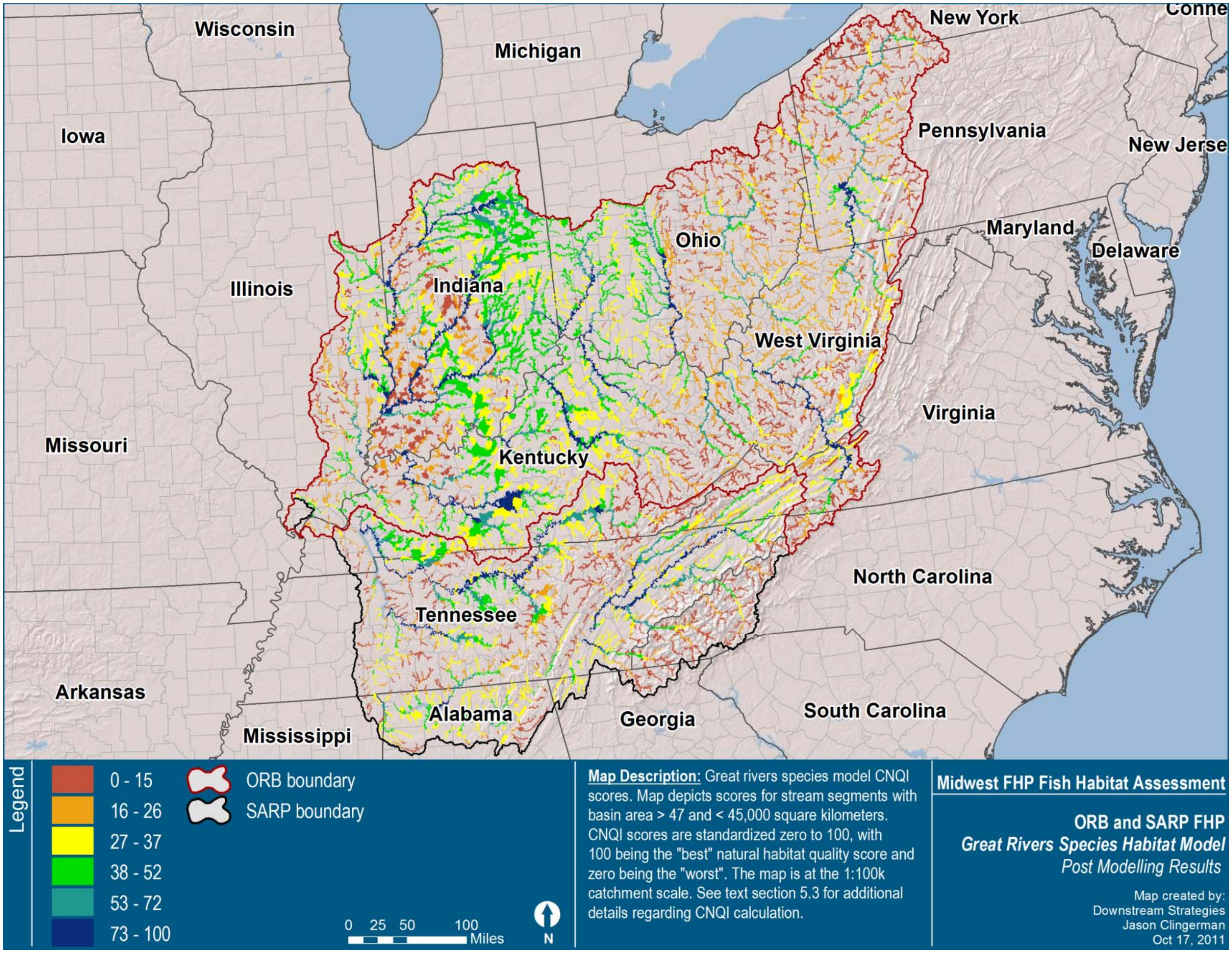


Figure 98: Cumulative anthropogenic stress index for great rivers species

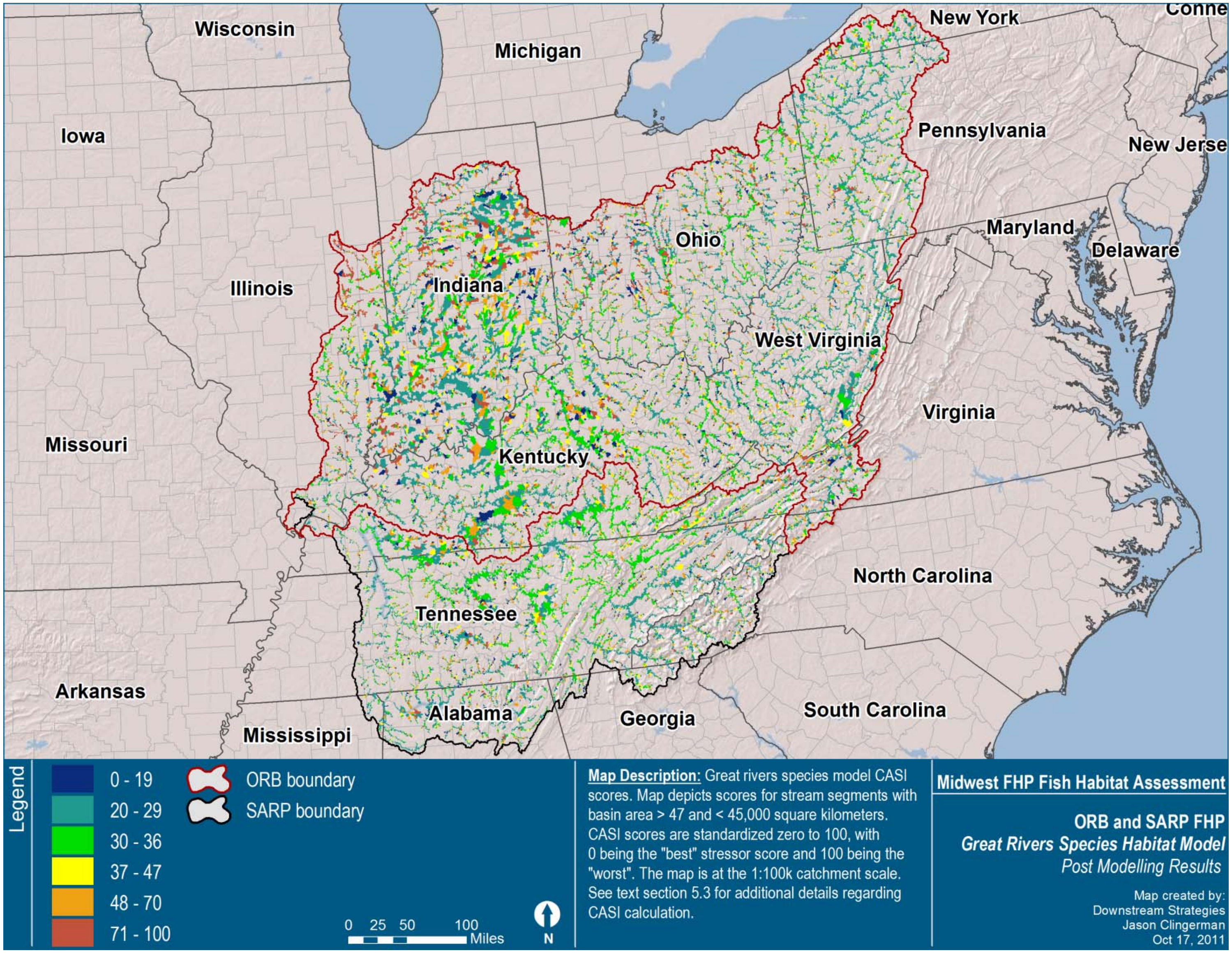


Figure 99: Most influential natural index metric for great rivers species

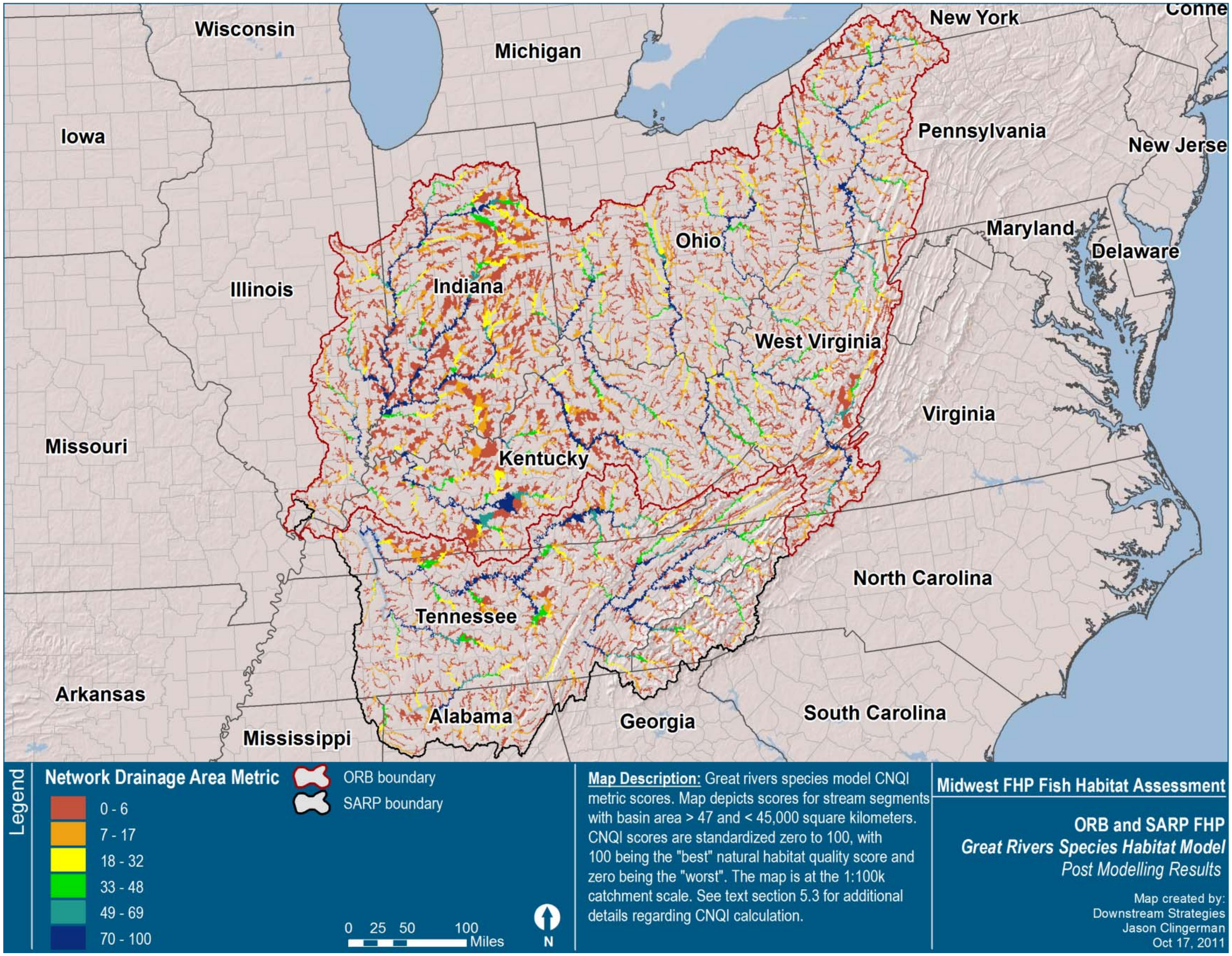


Figure 100: Second most influential natural index metric for great rivers species

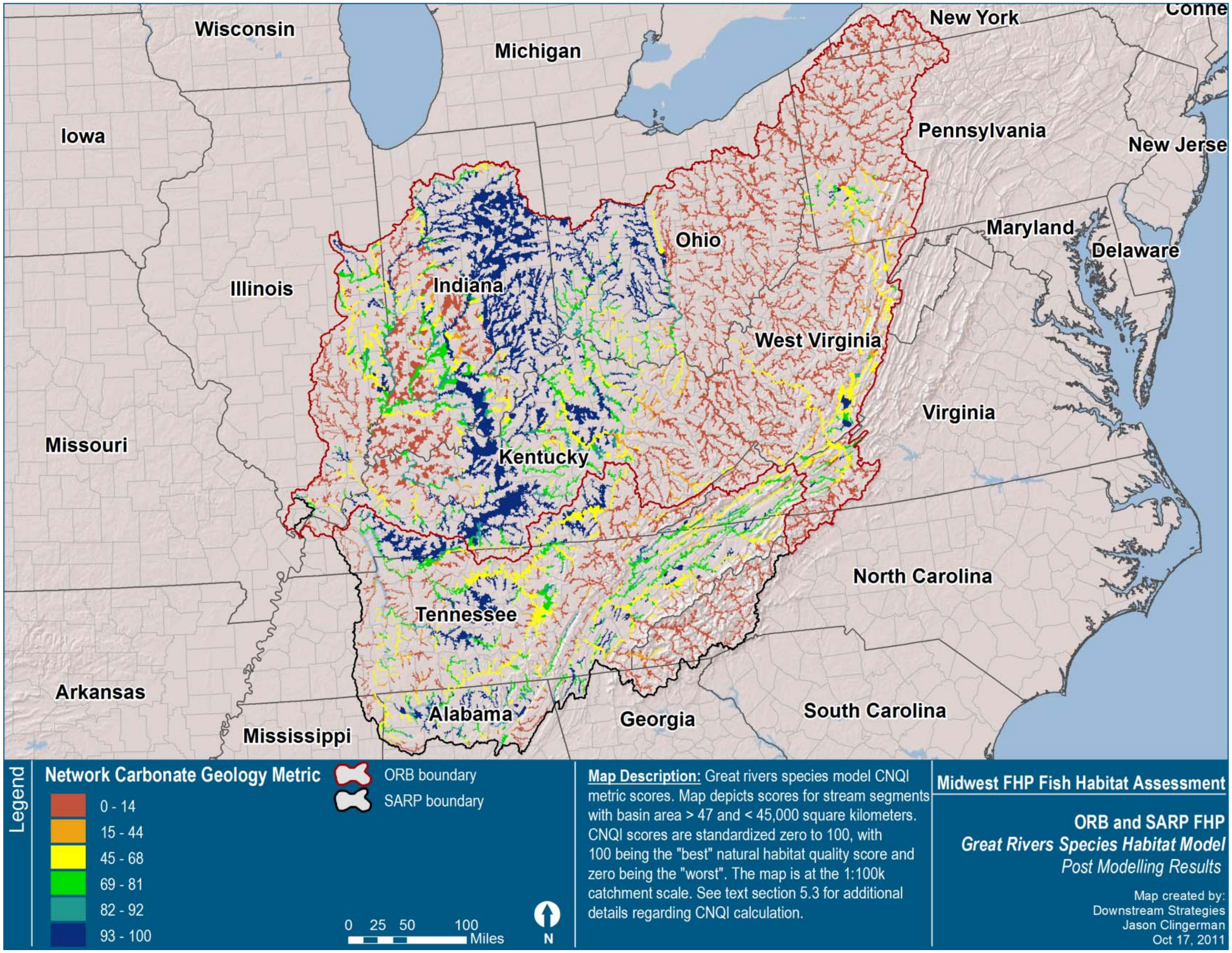


Figure 101: Third most influential natural index metric for great rivers species

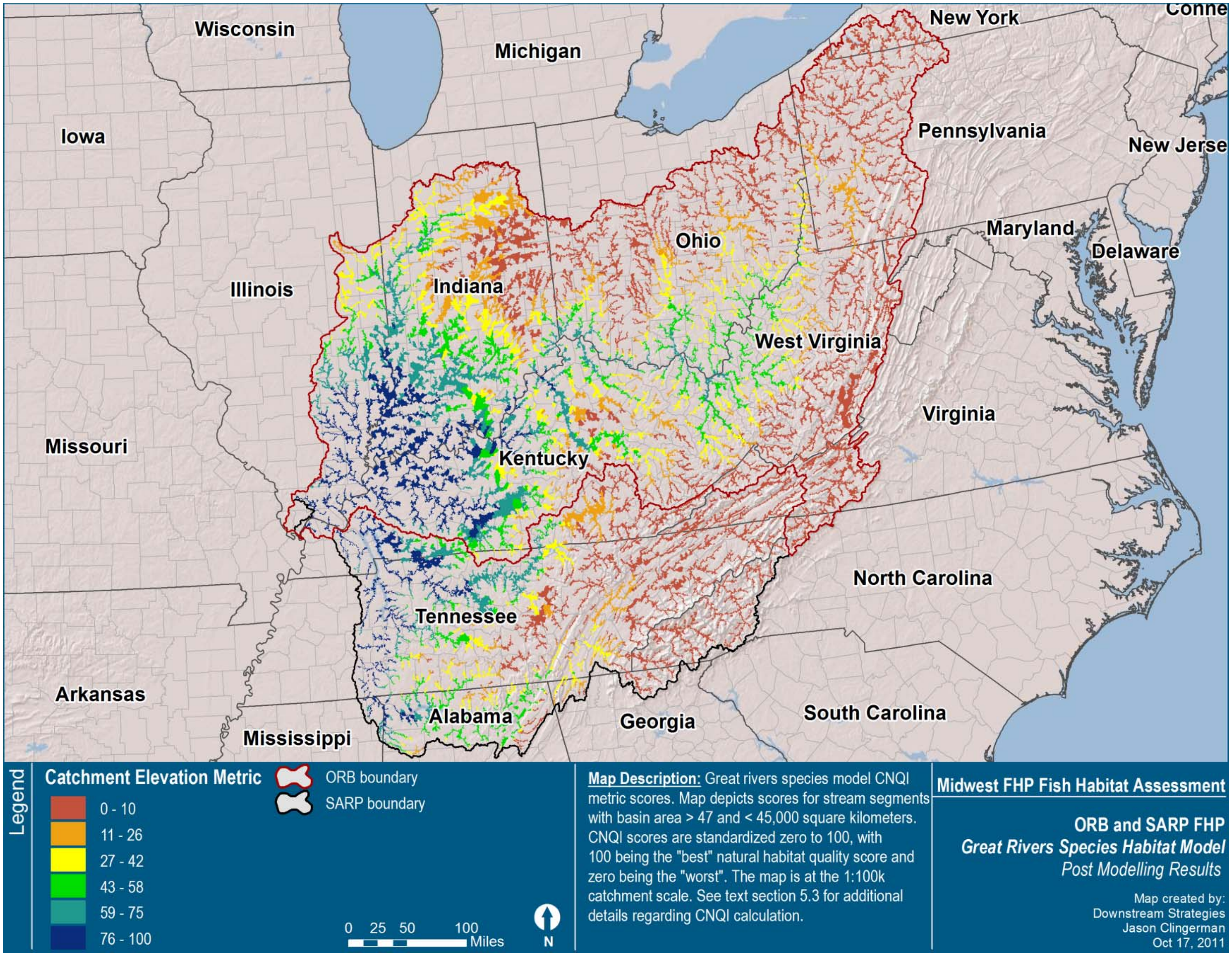


Figure 102: Fourth most influential natural index metric for great rivers species

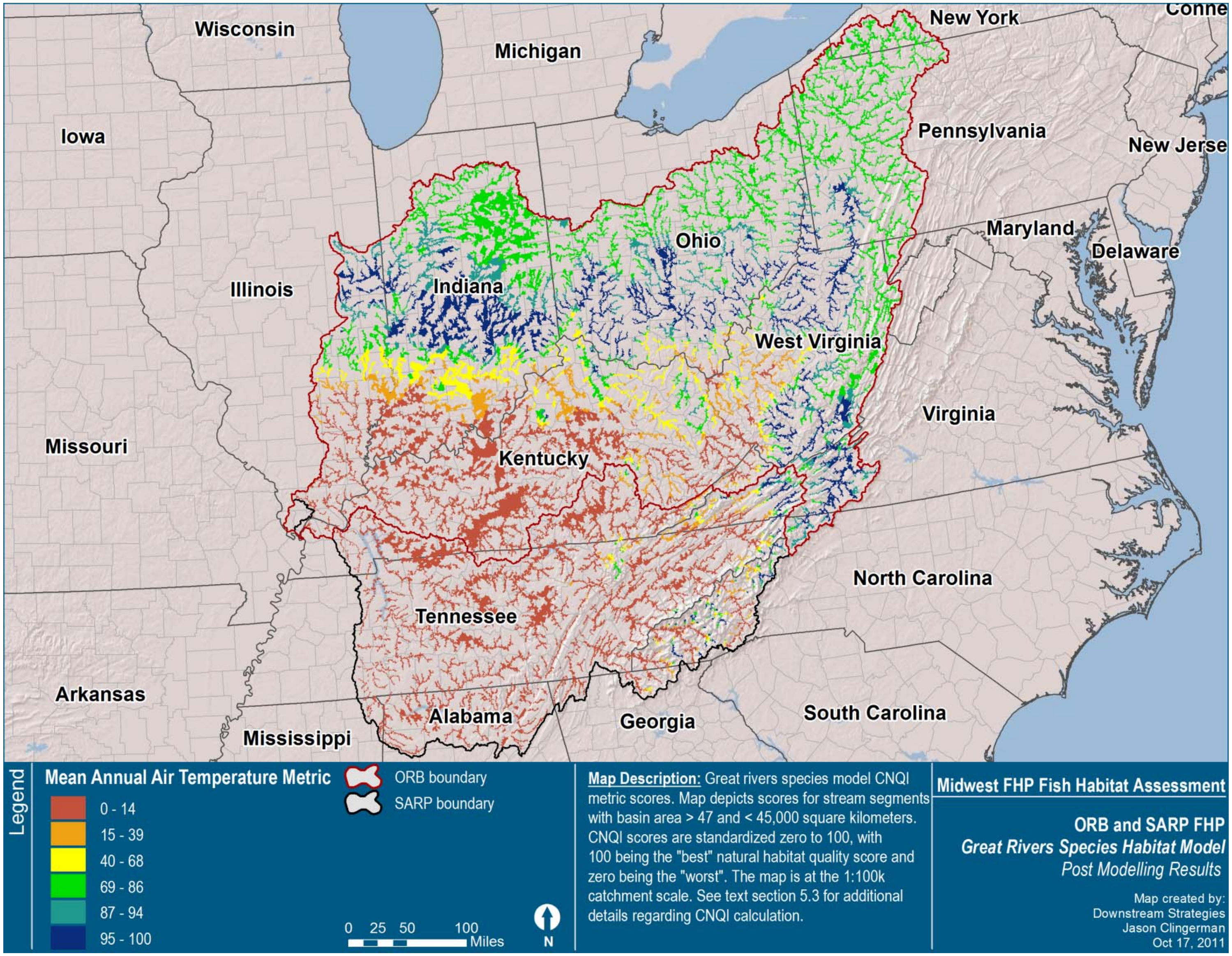


Figure 103: Fifth most influential natural index metric for great rivers species

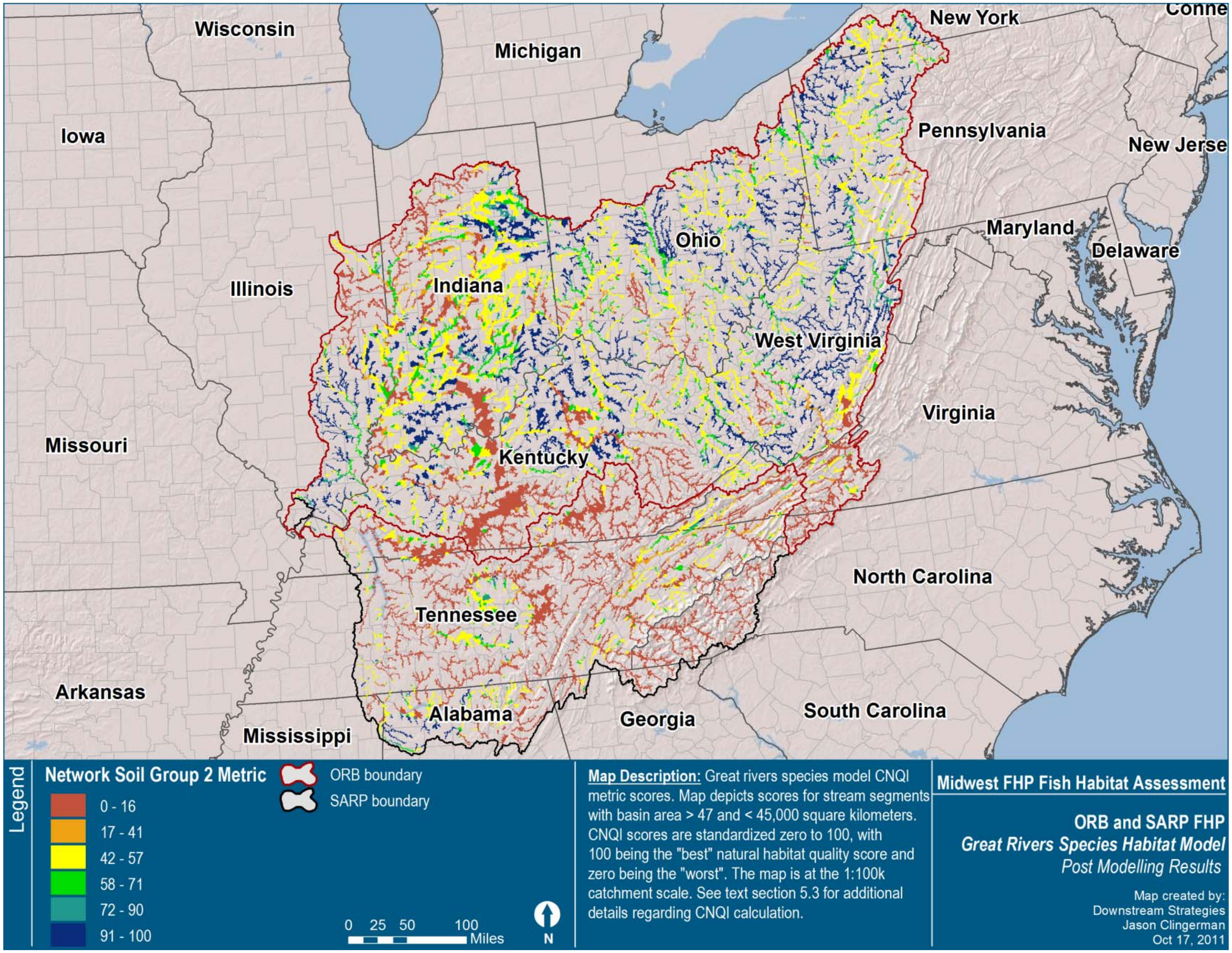


Figure 104: Most influential anthropogenic index metric for great rivers species

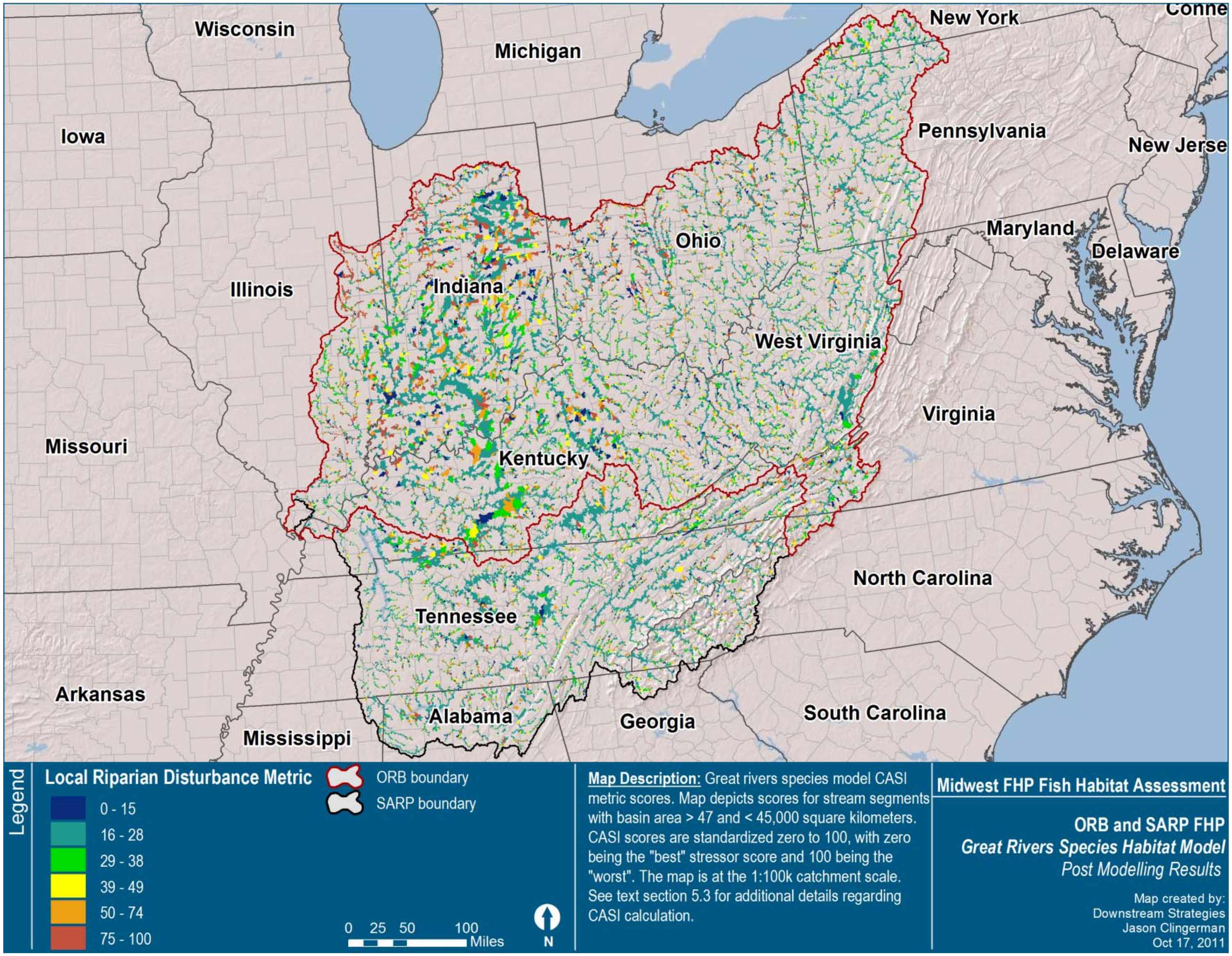


Figure 105: Second most influential anthropogenic index metric for great rivers species

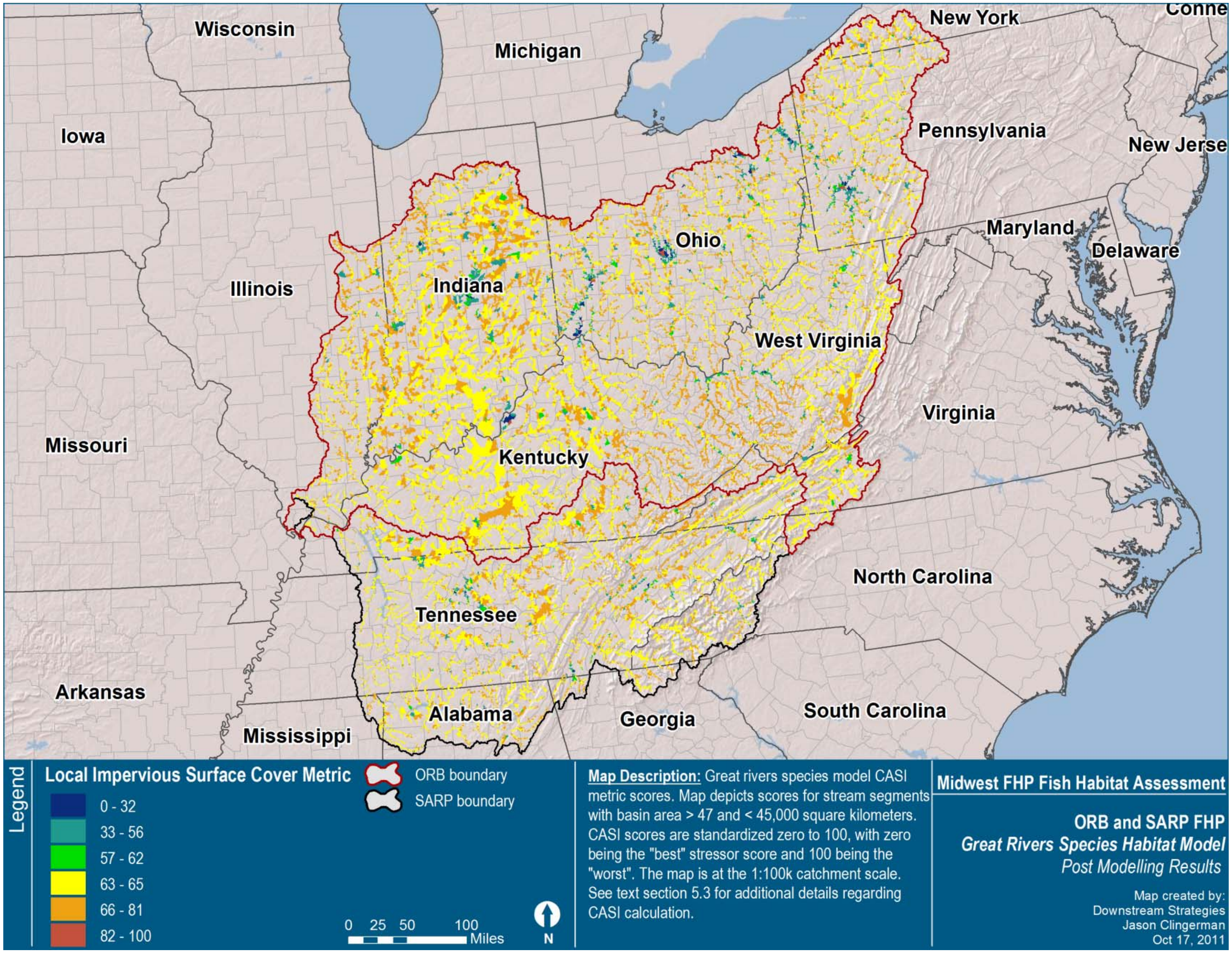


Figure 106: Third most influential anthropogenic index metric for great rivers species

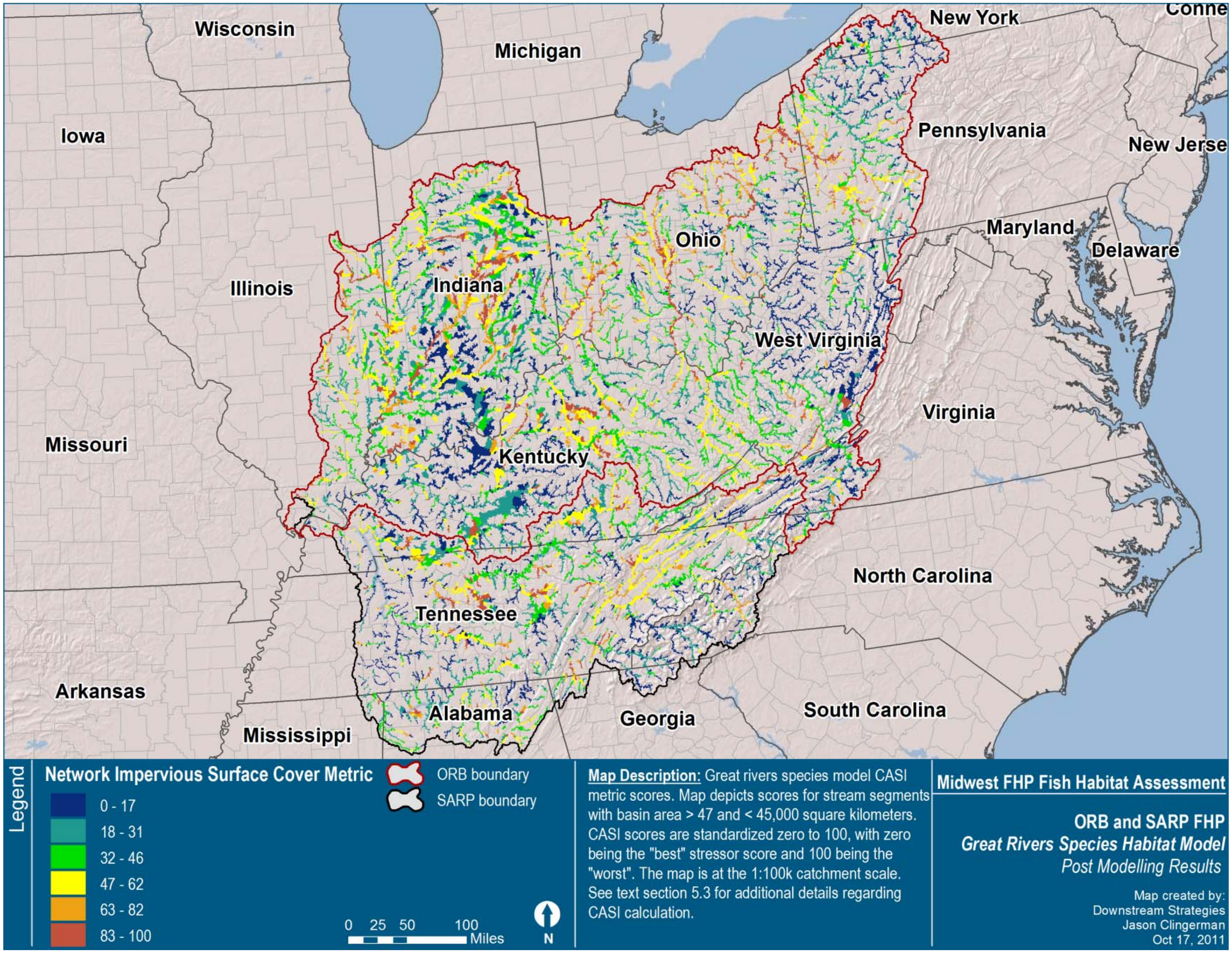


Figure 107: Fourth most influential anthropogenic index metric for great rivers species

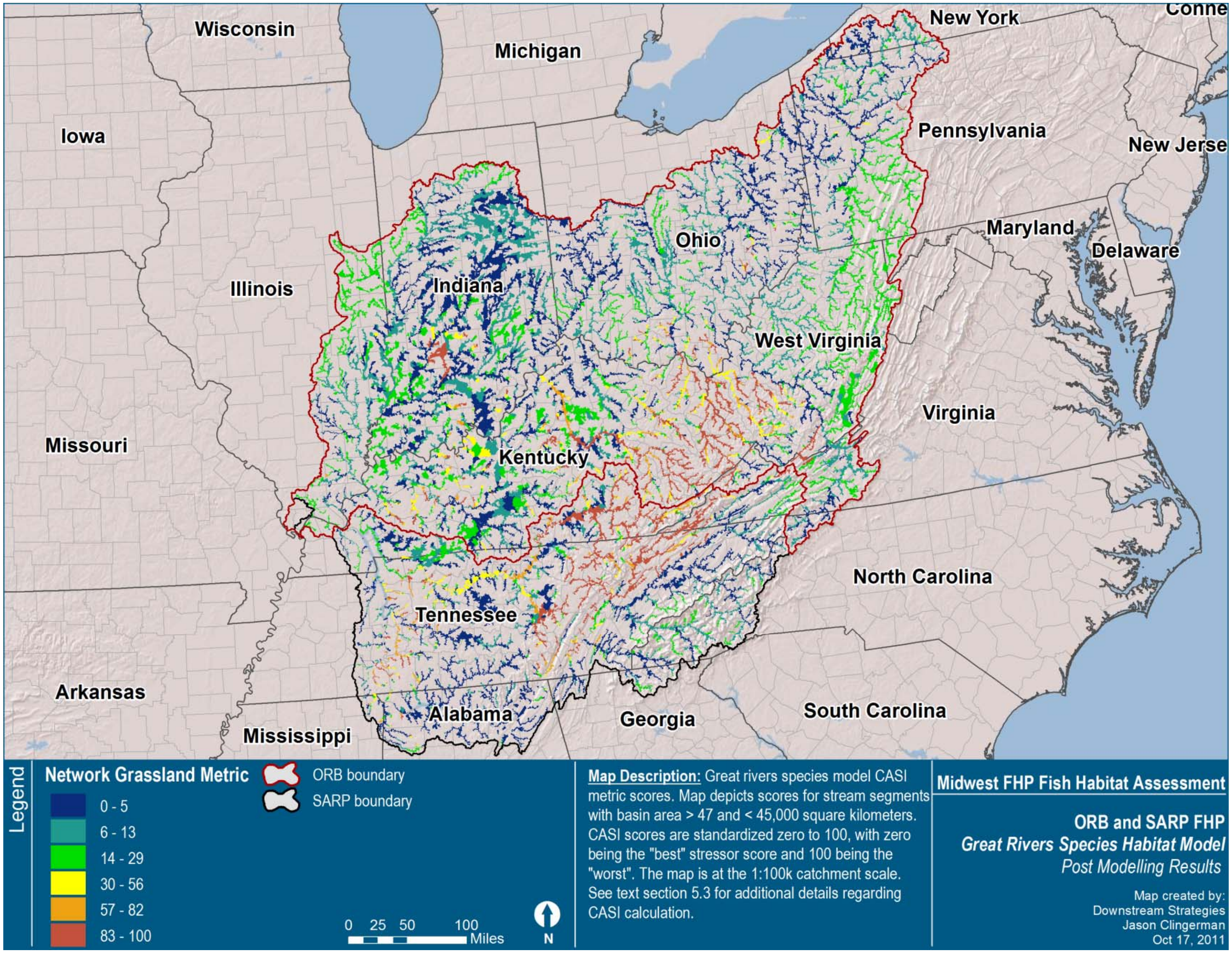
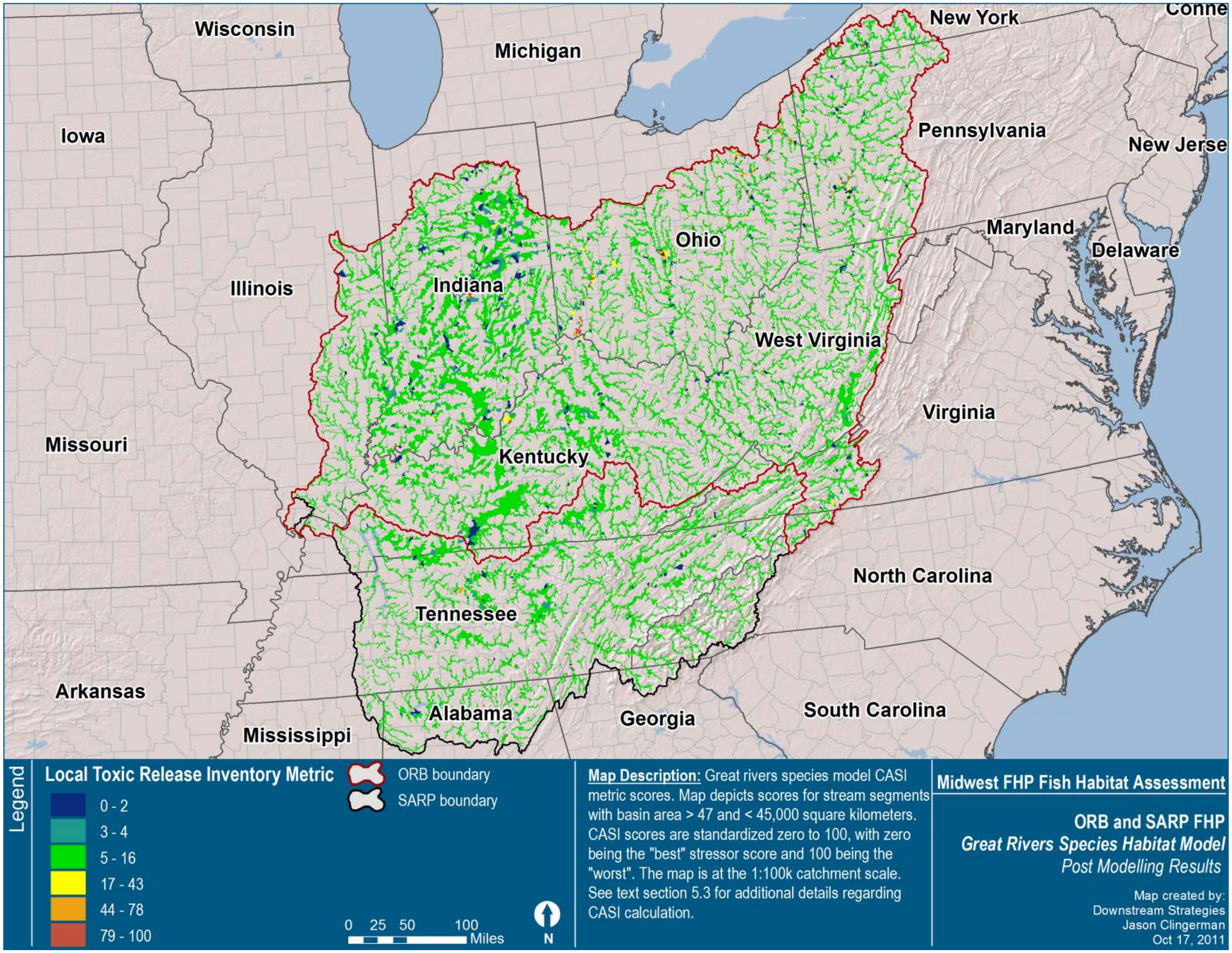


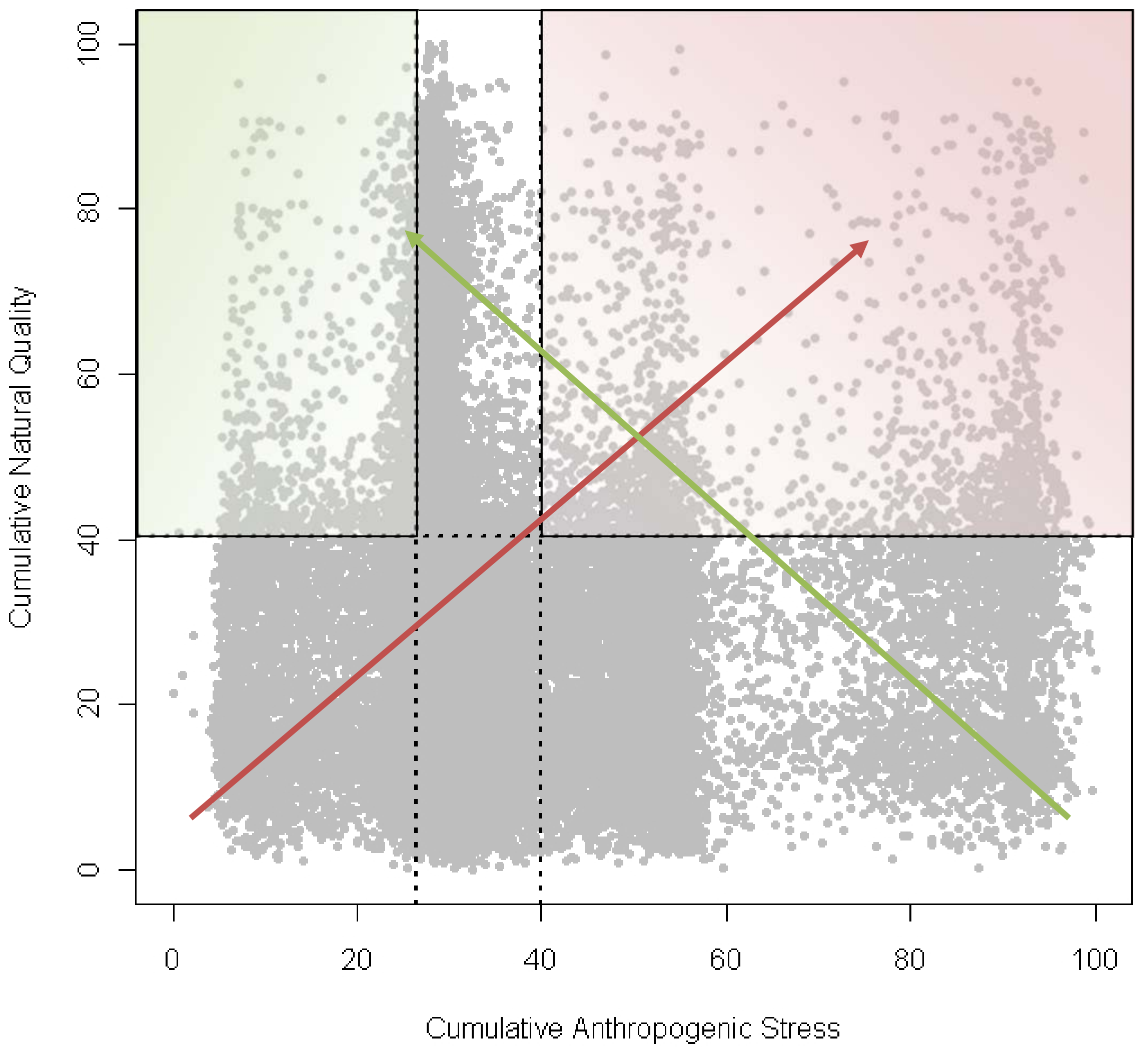
Figure 108: Fifth most influential anthropogenic index metric for great rivers species



7.5.4 *Restoration and protection priorities*

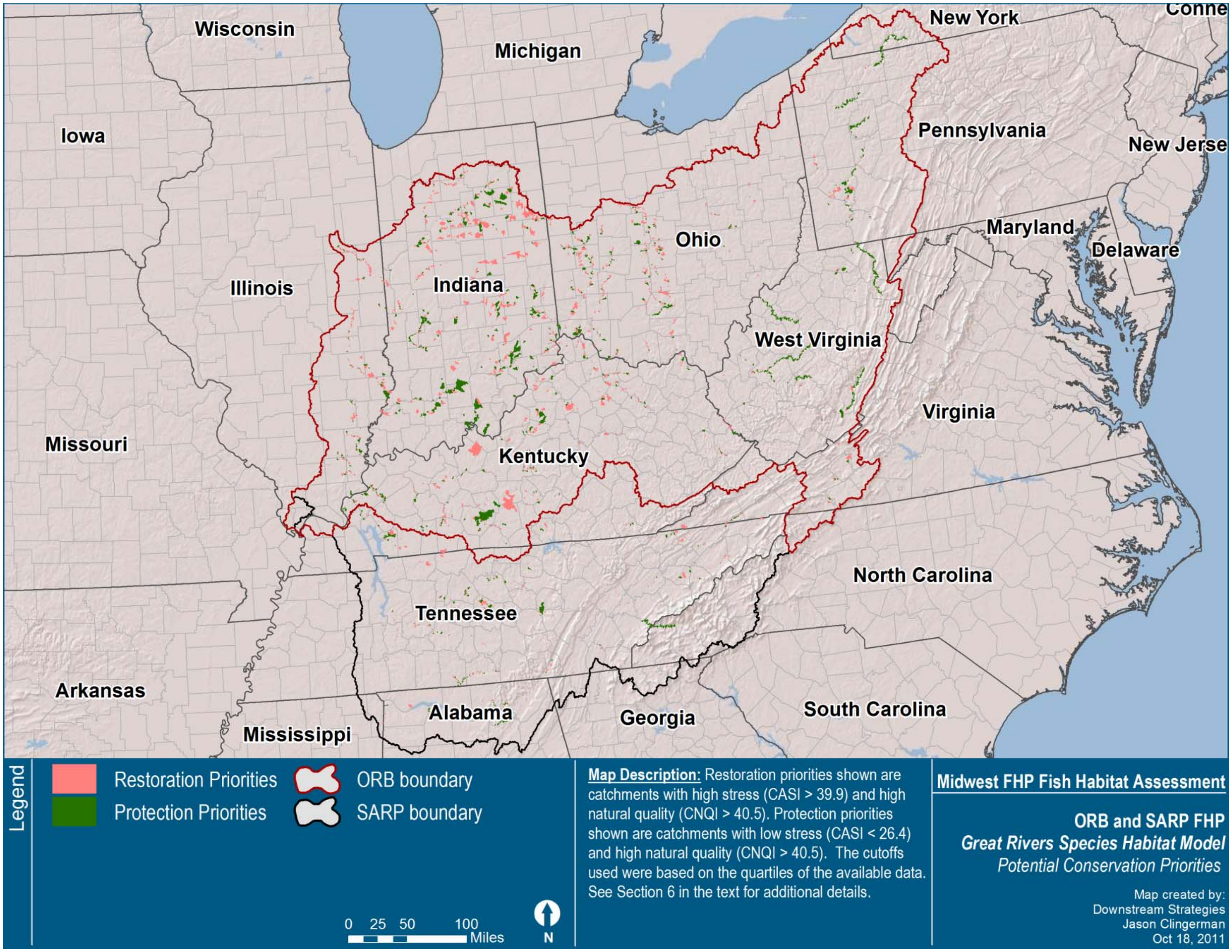
A plot of CNQI versus CASI values for all catchments in the study area (Figure 109) can be used as a reference when defining thresholds for categories of CNQI and CASI scores for use in the development of restoration and protection priorities. In the example shown (Figure 110), thresholds for restoration (high natural potential coupled with high anthropogenic stress) were set to CNQI greater than 40.5 and CASI greater than 39.9 (third quartiles). The thresholds used for protection (high natural potential and low anthropogenic stress) priorities were CNQI greater than 40.5 and CASI less than 26.4 (first quartile).

Figure 109: CNQI versus CASI values for all catchments for great rivers species



Note: Breakpoints for CNQI and CASI classes in this example are denoted by dashed lines. The arrows indicate the directions of increasing potential protection (green arrow) or restoration (red arrow) priority. The red box indicates catchments defined as restoration priorities under the example scenario. The green box indicates catchments defined as protection priorities under the same scenario.

Figure 110: Restoration and protection priorities for great rivers species



8. INTOLERANT MUSSELS

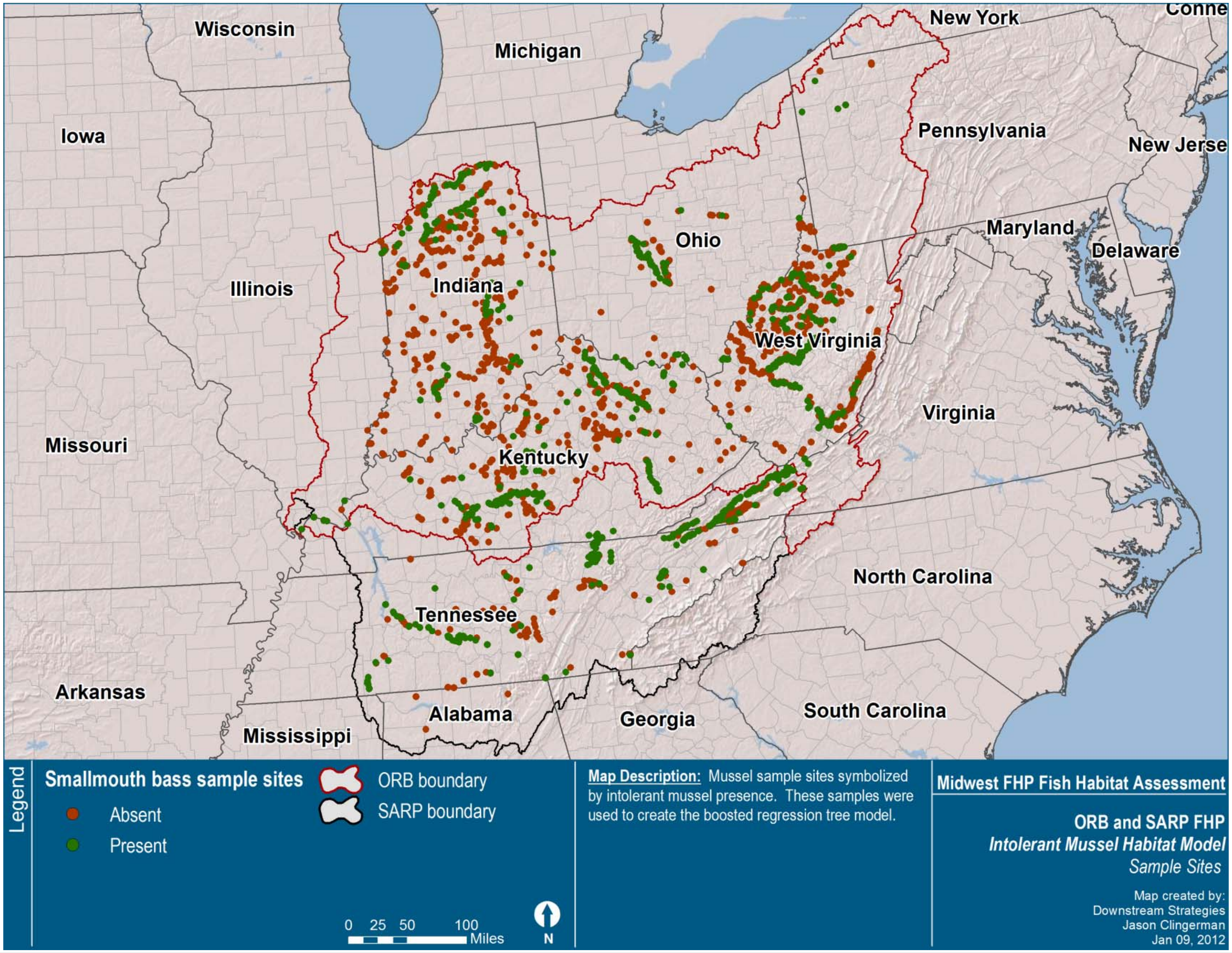
8.1 Modeling inputs

DS used a list of predictor variables selected by ORB and SARP to develop a ten-fold CV BRT model for intolerant mussels at the 1:100k catchment scale. The model was used to produce maps of expected current intolerant mussel distribution and maps of expected current natural habitat quality and anthropogenic stress at that scale throughout the extents of both FHPs.

DS cooperated with ORB and SARP to arrive at a list of landscape-based habitat variables used to predict the presence of intolerant mussels throughout the region; ultimately, those variables were also used for characterizing habitat quality and anthropogenic stress. From an initial suite of 372 catchment attributes, DS and the FHPs compiled a list of 92 predictors for evaluation. From that list, 47 variables were removed due to statistical redundancy ($r > 0.6$) or logical redundancy, resulting in a final list of 45 predictor variables for the BRT model and assessment. See Appendix A for a full data dictionary and the metadata document for variable processing notes.

ORB and SARP provided DS with a presence-absence dataset for intolerant mussels comprised of 3,341 observations collected in streams. Intolerant mussels in this analysis are defined as those species that are intolerant to human disturbance; a list of those species considered intolerant can be found in Table 25 (located in Appendix A).These samples were taken over a time frame spanning 1996 to 2010. Figure 111 maps all of the sampling sites that were used to construct the model and indicates the ORB and SARP boundaries to which the modeling outputs were applied.

Figure 111: Intolerant mussels modeling area and sampling sites



8.2 Modeling process

8.2.1 *Predictive performance*

The final selected model was comprised of 1,800 trees. The model had a CV correlation statistic of 0.605±0.007, and a CV ROC score of 0.845±0.004. The CV correlation statistic indicates the mean correlation resulting from each fold (ten in this case) of the cross-validation process.

8.2.2 *Variable influence*

The BRT output includes a list of the predictor variables used in the model ordered and scored by their relative importance. The relative importance values are based on the number of times a variable is selected for splitting, weighted by the squared improvement to the model as a result of each split, and averaged over all trees (Friedman and Meulman, 2003). The relative influence score is scaled so that the sum of the scores for all variables is 100, where higher numbers indicate greater influence. Of the 45 predictor variables used to develop the intolerant mussel model, all had a relative influence value greater than zero (Table 21). The five most influential predictors, which accounted for almost 40% of the total influence in the model, were:

- network drainage area,
- network baseflow index,
- mean annual precipitation,
- network density of dams, and
- network alluvium cover.

The five most influential anthropogenic stressors, which accounted for over 19% of the total influence, were:

- network dam density,
- network surface water consumption,
- network forested land cover,
- network density of road crossings, and
- local impervious surface cover.

The five most influential natural habitat variables, which contributed over 39% of the total influence, were:

- network drainage area,
- network baseflow index,
- mean annual precipitation,
- network alluvium cover, and
- network shale bedrock.

Network drainage area, the single most important variable in terms of relative influence, contributed over 12% of the total influence.

Table 21: Relative influence of all variables in the final intolerant mussel model

Variable code	Variable description	Relative influence
cumdrainag	Network drainage area	12.39
BFI_meanc	Network mean baseflow index	8.23
precip	Mean annual precipitation	8.12
damsc_den	Network density of dams	5.98
surf3pc	Network alluvium surficial geology cover	5.27
brock7pc	Network shale bedrock geology cover	5.17
water_swc	Network surface water consumption	5.02
soil4pc	Network soil type D cover	4.2
minelevraw	Minimum catchment elevation	4.02
forpc	Network forested land cover	2.97
roadcrc_den	Network density of road crossings	2.89
imp06	Local impervious surface cover	2.71
imp06c	Network impervious surface cover	2.65
grasspc	Network grassland land cover	2.64
soil2pc	Network soil type B, B/D cover	2.45
pastp	Network pasture land cover	2.43
cropsp	Network rowcrop land cover	2.17
tric_den	Network density of Toxic Release Inventory sites	2.16
pastpc	Network pasture land cover	1.96
wetlandp	Local wetland land cover	1.74
slope	Slope of catchment flowline	1.49
surf6pc	Network residuum surficial geology cover	1.45
popdens	Local population density	1.45
brock5pc	Network sand/gravel bedrock geology cover	1.2
soil2p	Local soil type B, B/D cover	1.01
water_gwc	Local groundwater consumption	0.98
ripdisp	Local riparian disturbance	0.97
roadcr_den	Local density of road crossings	0.97
cercc_den	Network density of Superfund sites	0.65
dams_den	Local density of dams	0.64
surf2pc	Network outwash surficial geology cover	0.64
grassp	Local grassland land cover	0.61
soil4p	Local soil type D cover	0.52
minesc_den	Network density of mines	0.46
soil1pc	Network soil type A, A/D cover	0.46
brock1p	Local carbonate bedrock geology cover	0.38
surf4pc	Network lacustrine surficial geology cover	0.29
brock2pc	Network felsic/igneous bedrock geology cover	0.2
surf8p	Network colluvium surficial geology cover	0.16
soil1p	Local soil type A, A/D cover	0.13
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.08
TRI_den	Local density of Toxic Release Inventory sites	0.05
surf5pc	Network loess surficial geology cover	0.04
surf1p	Local till surficial geology cover	0.03
surf6p	Local residuum surficial geology cover	0.01

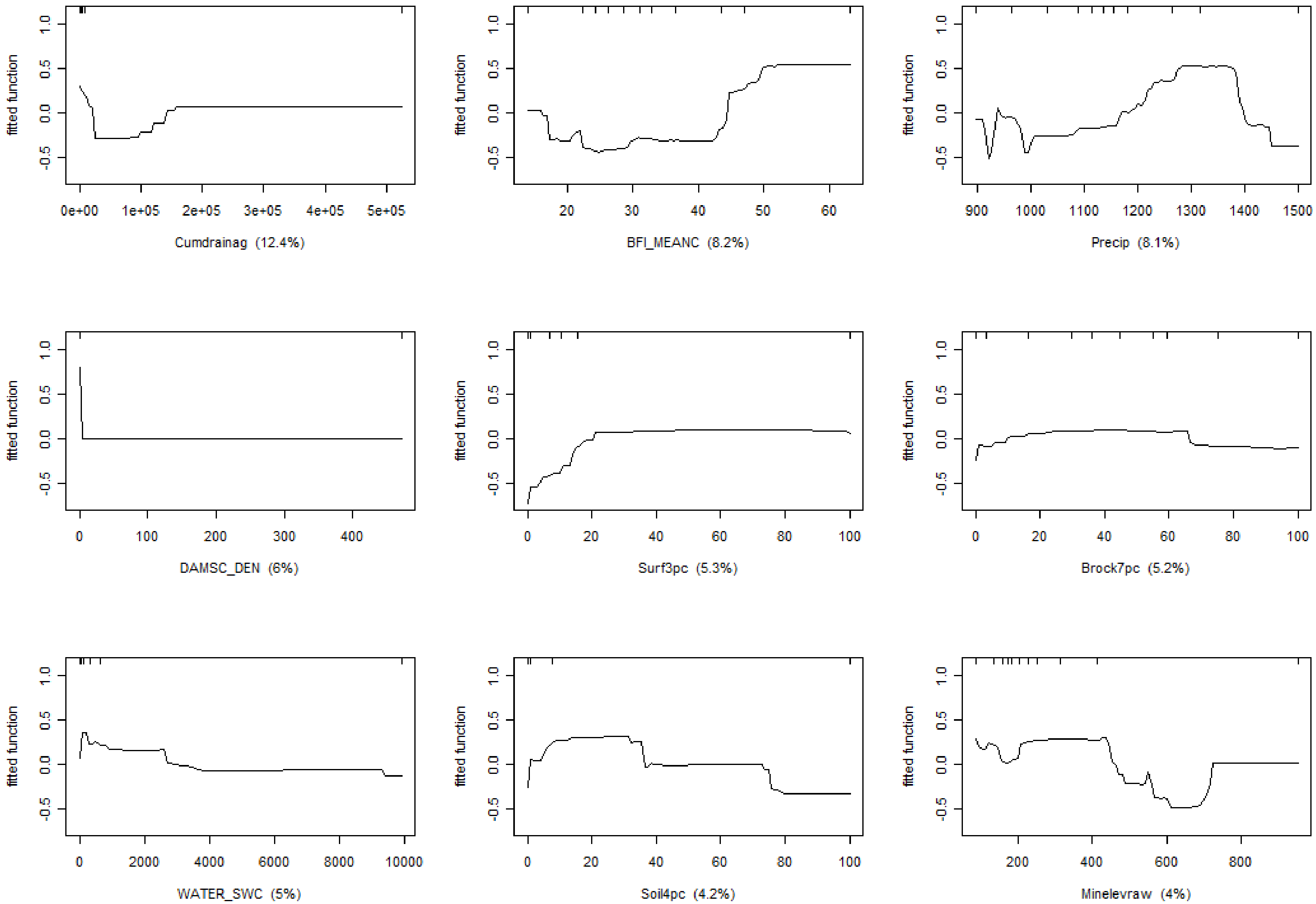
Note: Individual variables are highlighted according to whether they were determined to be anthropogenic in nature (red highlight) or natural (green highlight).

8.2.3 *Variable functions*

The BRT output also contains quantitative information on partial dependence functions that can be plotted to visualize the effect of each individual predictor variable on the response after accounting for all other variables in the model. Similar to the interpretation of traditional regression coefficients, the function plots are not always a perfect representation of the relationship for each variable, particularly if interactions are strong or predictors are strongly correlated. However, they do provide a useful and objective basis for interpretation (Friedman, 2001; Friedman and Meulman, 2003).

These plots show the trend of the response variable (y-axis) as the predictor variable (x-axis) changes. The response variable is transformed (usually to the logit scale) so that the magnitude of trends for each predictor variable's function plot can be accurately compared. The dash marks at the top of each function represent the deciles of the data used to build the model. The function plots for the nine most influential variables in the intolerant mussel model (see Table 21 for reference) are illustrated in Figure 112 below. The plots for all 46 variables are shown in Appendix B.

Figure 112: Functional responses of the dependent variable to individual predictors of intolerant mussels



Note: Only the top nine predictors, based on relative influence (shown in parentheses; see Appendix A for descriptions of variable codes), are shown here. See Appendix B for plots of remaining predictor variables.

8.3 Post-modeling

The variable importance table and partial dependence functions of the final BRT model were used to create the post-modeling indices of natural habitat quality and anthropogenic stress for Intolerant mussels. The CNQI was comprised of 23 variables with relative influence greater than zero that were classified as natural habitat features (Table 22). The CASI was comprised of 16 variables with relative influence greater than zero that were classified as anthropogenic habitat features (Table 23). To calculate the cumulative indices (i.e., CNQI and CASI), each of the individual natural or anthropogenic variables used in the two indices was converted to a metric by first applying the appropriate transformations, based on their function plots, and then rescaling the transformed measures to a 0 to 100 scale. To calculate the cumulative index from the individual metrics, the metrics were first multiplied by their appropriate weighting factors and then summed. The CNQI and CASI scores were a result of a rescaling of those weighted and summed metrics, again from 0 to 100.

8.3.1 *Variable weights*

Table 22 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CNQI. The five most influential factors in the CNQI were:

- network drainage area,
- network baseflow index,
- mean annual precipitation,
- network alluvium cover, and
- network shale bedrock.

Table 23 summarizes the relative influence values and the derived post-modeling weighting factors used in the construction of the CASI. The five most influential factors in the CASI were:

- network density of dams,
- network surface water consumption,
- network forested land cover,
- local impervious surface cover, and
- network impervious surface cover.

Table 22: Relative influence and weights for natural variables on intolerant mussels

Variable code	Variable description	Relative influence	Weighting factor
cumdrainag	Network drainage area	12.39	1
BFI_meanc	Network mean baseflow index	8.23	0.66
precip	Mean annual precipitation	8.12	0.66
surf3pc	Network alluvium surficial geology cover	5.27	0.43
brock7pc	Network shale bedrock geology cover	5.17	0.42
soil4pc	Network soil type D cover	4.2	0.34
minelevraw	Minimum catchment elevation	4.02	0.33
soil2pc	Network soil type B, B/D cover	2.45	0.2
slope	Slope of catchment flowline	1.49	0.12
surf6pc	Network residuum surficial geology cover	1.45	0.12
brock5pc	Network sand/gravel bedrock geology cover	1.2	0.1
soil2p	Local soil type B, B/D cover	1.01	0.08
surf2pc	Network outwash surficial geology cover	0.64	0.05
soil4p	Local soil type D cover	0.52	0.04
soil1pc	Network soil type A, A/D cover	0.46	0.04
brock1p	Local carbonate bedrock geology cover	0.38	0.03
surf4pc	Network lacustrine surficial geology cover	0.29	0.02
brock2pc	Network felsic/igneous bedrock geology cover	0.2	0.02
surf8p	Network colluvium surficial geology cover	0.16	0.01
soil1p	Local soil type A, A/D cover	0.13	0.01
surf5pc	Network loess surficial geology cover	0.04	0
surf1p	Local till surficial geology cover	0.03	0
surf6p	Local residuum surficial geology cover	0.01	0

Table 23: Relative influence and weights for anthropogenic variables on intolerant mussels

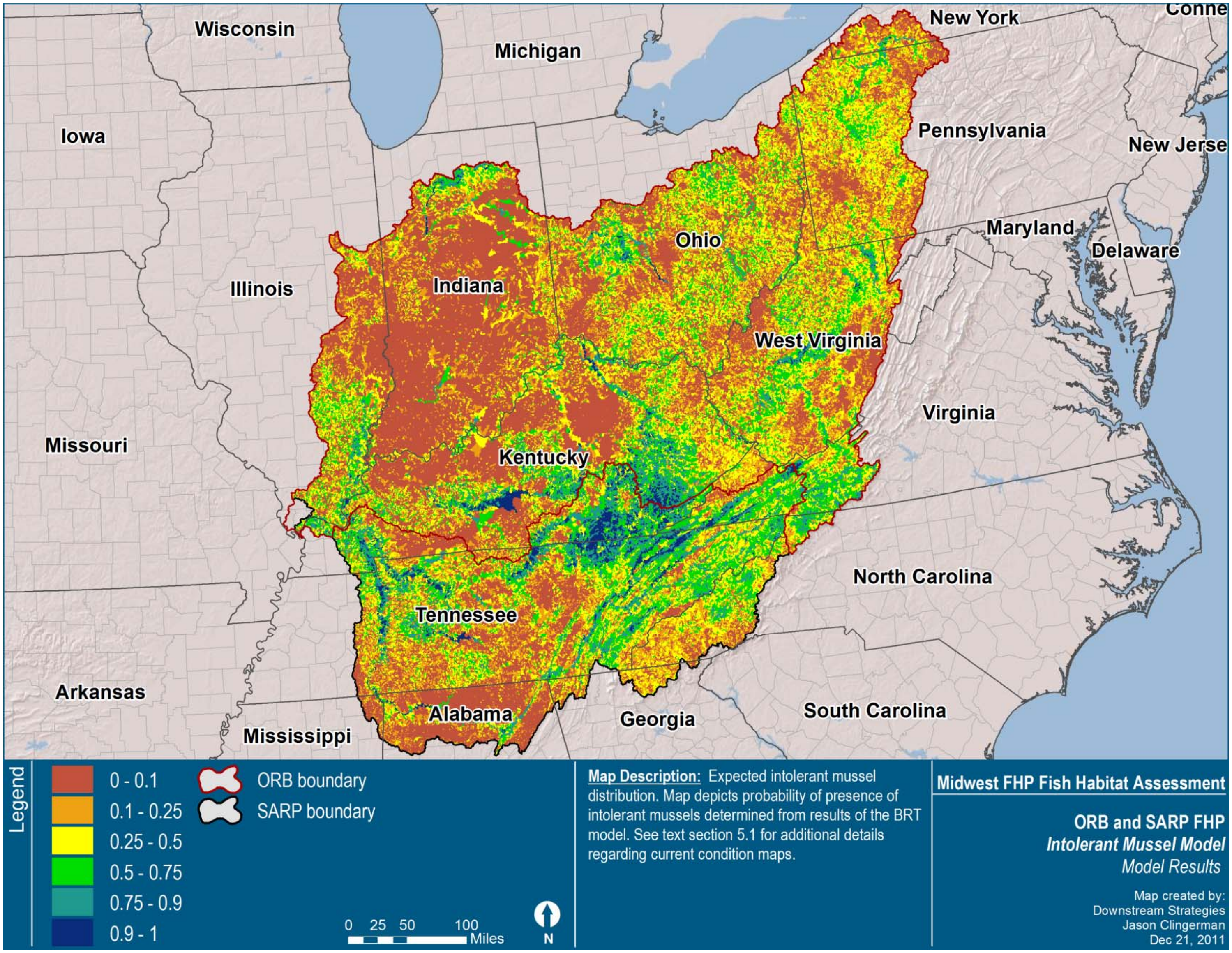
Variable code	Variable description	Relative influence	Weighting factor
damsc_den	Network density of dams	5.98	1
water_swc	Network surface water consumption	5.02	0.84
forpc	Network forested land cover	2.97	0.5
imp06	Local Impervious surface cover	2.71	0.45
imp06c	Network impervious surface cover	2.65	0.44
pastp	Network pasture land cover	2.43	0.41
cropsp	Network rowcrop land cover	2.17	0.36
pastpc	Network pasture land cover	1.96	0.33
wetlandp	Local wetland land cover	1.74	0.29
popdens	Local population density	1.45	0.24
water_gwc	Local groundwater consumption	0.98	0.16
ripdisp	Local riparian disturbance	0.97	0.16
cercc_den	Network density of Superfund sites	0.65	0.11
dams_den	Local density of dams	0.64	0.11
minesc_den	Network density of mines	0.46	0.08
NPDES_den	Local density of National Pollutant Discharge Elimination System permits	0.08	0.01

8.4 Mapped Results

8.4.1 *Expected current conditions*

Intolerant mussel probability of presence was calculated for all 1:100k stream catchments in the study area using the BRT model. The predicted probability values ranged from 0.002 to 1. The mean predicted probability value for the 226,919 total catchments was 0.375. There were 24,129 catchments with a predicted probability of presence greater than 0.75, and 42,830 catchments where the probability of presence was between 0.5 and 0.75. These results are mapped in Figure 113.

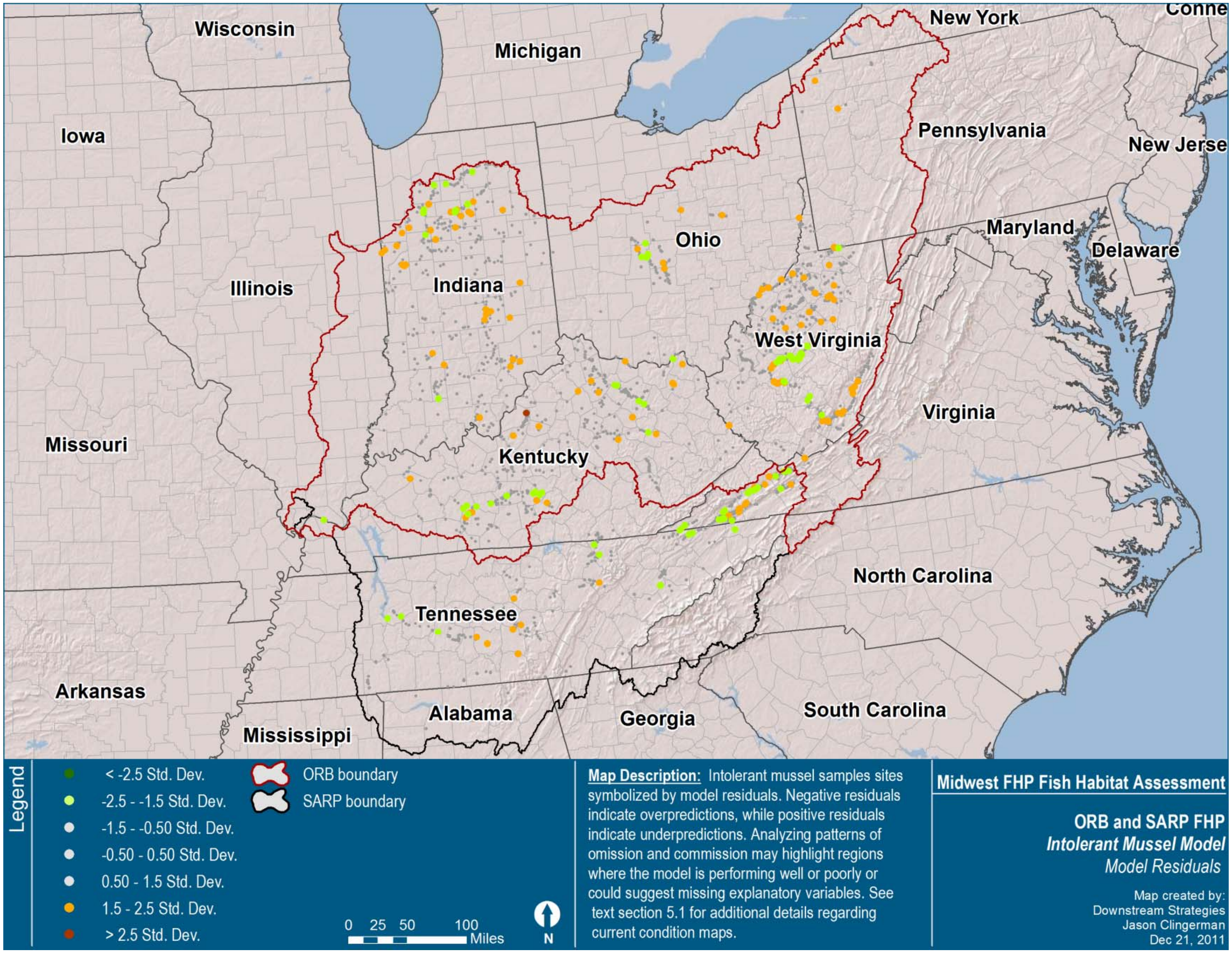
Figure 113: Expected intolerant mussels distribution



8.4.2 *Spatial variability in predictive performance*

Analyzing patterns of omission and commission may highlight regions where the model is performing well or poorly or could suggest missing explanatory variables (Figure 114). To assess omission and commission, residuals are also calculated by the BRT model. The residuals are a measure of the difference in the measured and modeled values (measured value *minus* modeled value). Negative residuals indicate overpredictions (predicting higher values than are true), while positive residuals indicate underpredictions (predicting lower values than are true).

Figure 114: Distribution of intolerant mussel model residuals by sampling site



8.4.3 *Indices of stress and natural quality*

Maps of CNQI and CASI illustrate the spatial distribution of natural habitat potential (i.e., CNQI score) and anthropogenic stress (i.e., CASI score) in ORB and SARP. CNQI and CASI scores are mapped in Figure 115 and Figure 116, respectively. The top five most influential variables toward the calculation of CNQI are shown in Figure 117-Figure 121. The top five variables contributing toward the calculation of CASI are mapped in Figure 122-Figure 126. CNQI, CASI, and their metrics are all scaled on a 0-100 scale (see Section 8.3 for more details on CNQI and CASI calculation). For CNQI, higher values indicate higher natural quality, while higher values for CASI indicate higher levels of anthropogenic stress.

Figure 115: Cumulative natural quality index for intolerant mussels

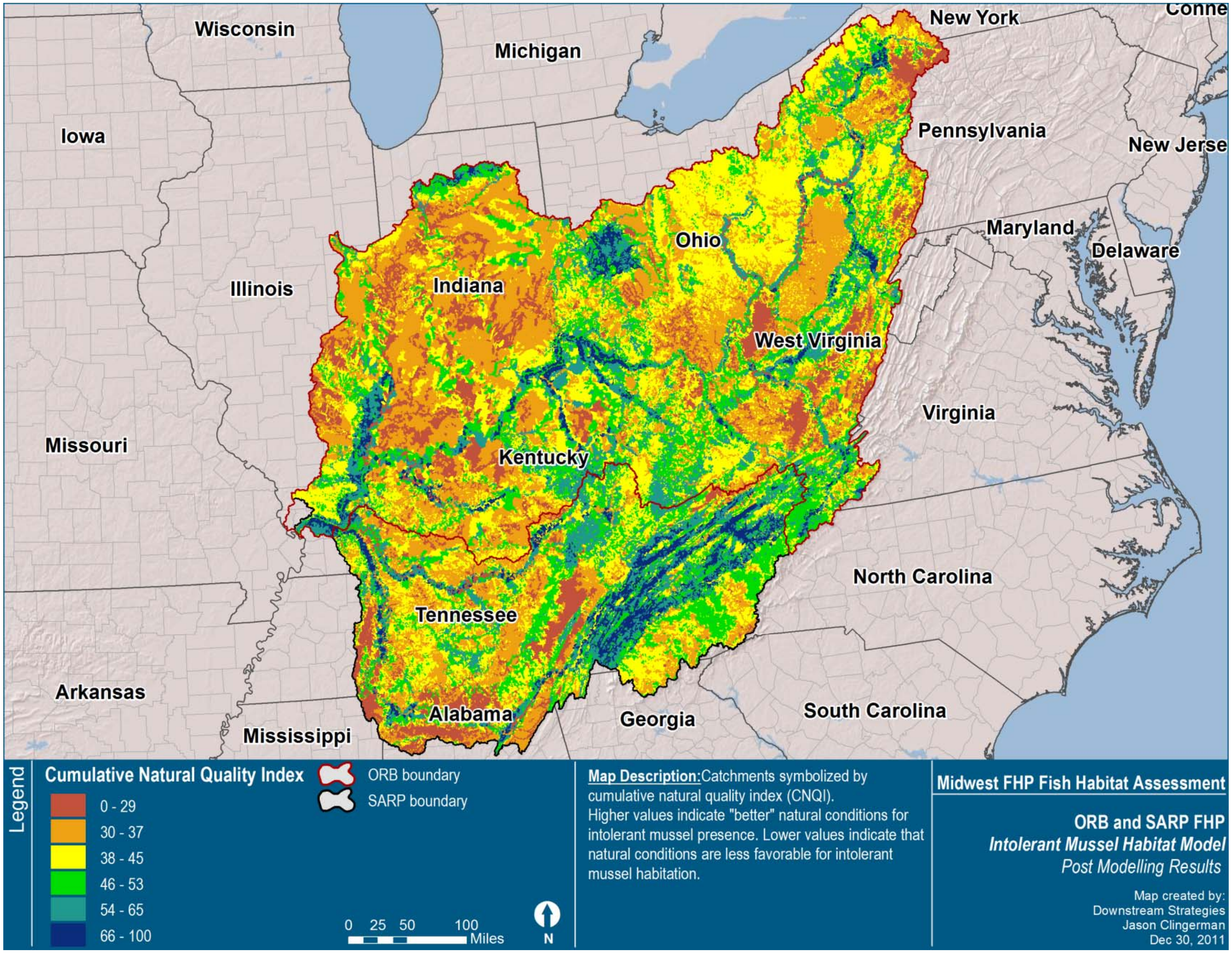


Figure 116: Cumulative anthropogenic stress index for intolerant mussels

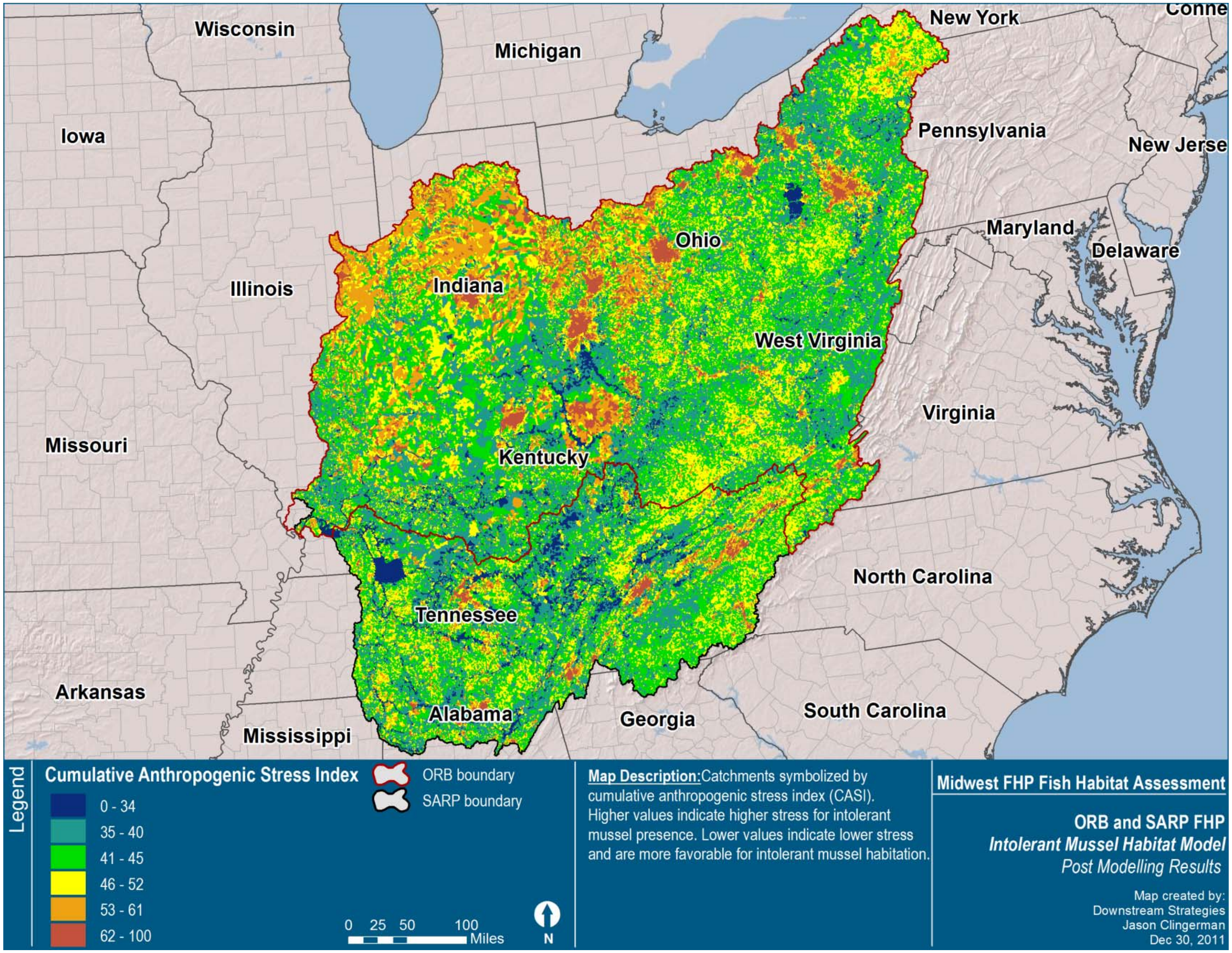


Figure 117: Most influential natural index metric for intolerant mussels

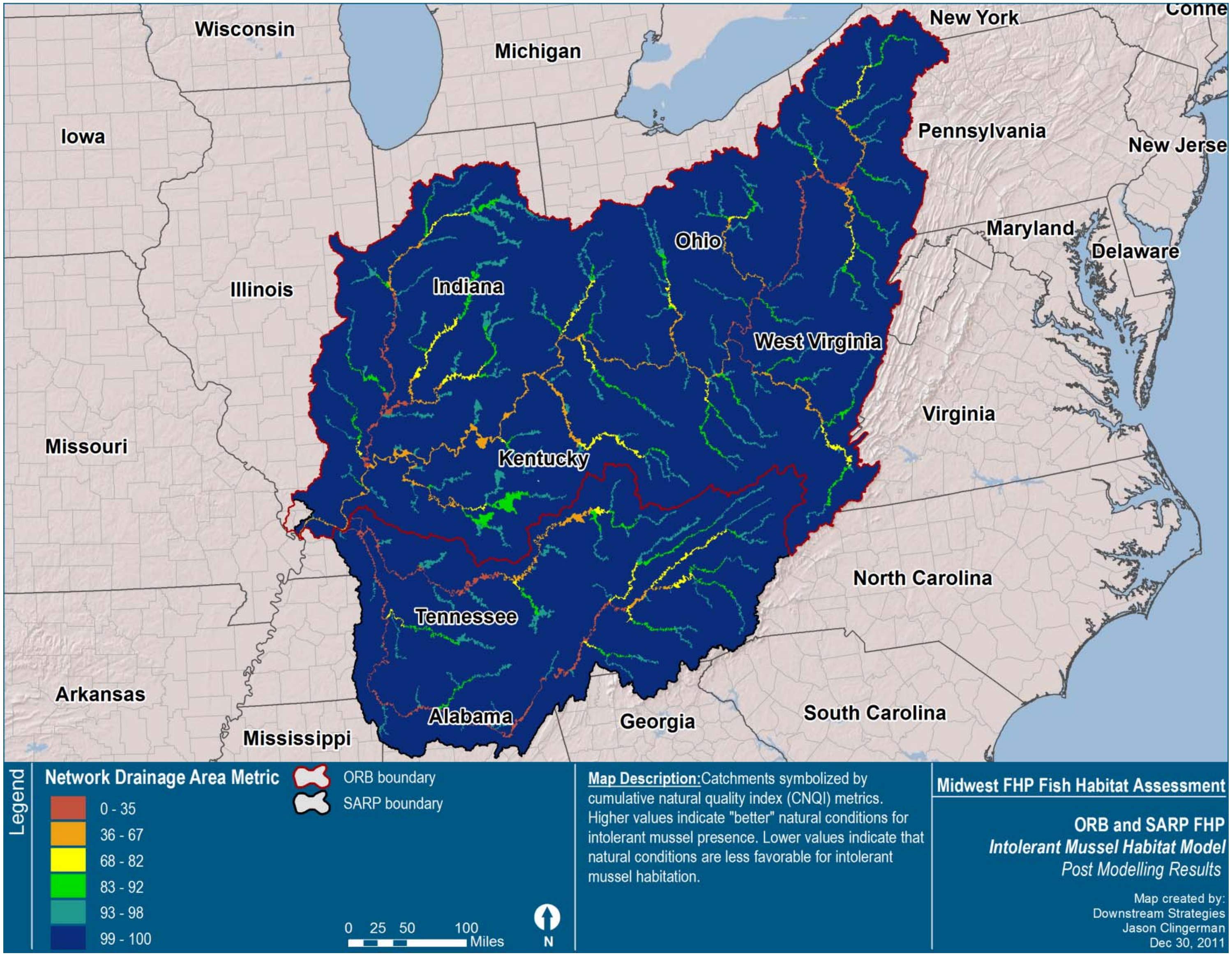


Figure 118: Second most influential natural index metric for intolerant mussels

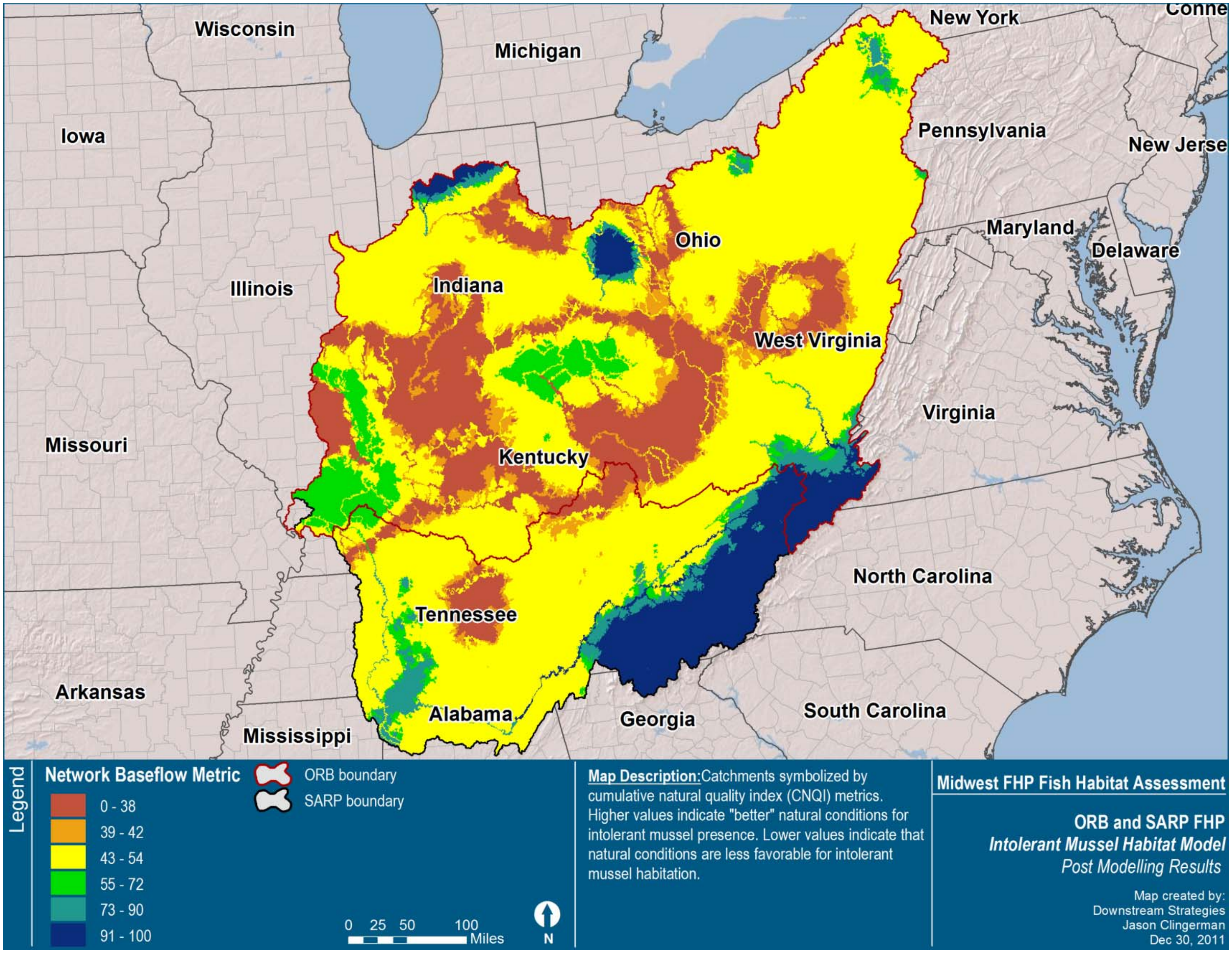


Figure 119: Third most influential natural index metric for intolerant mussels

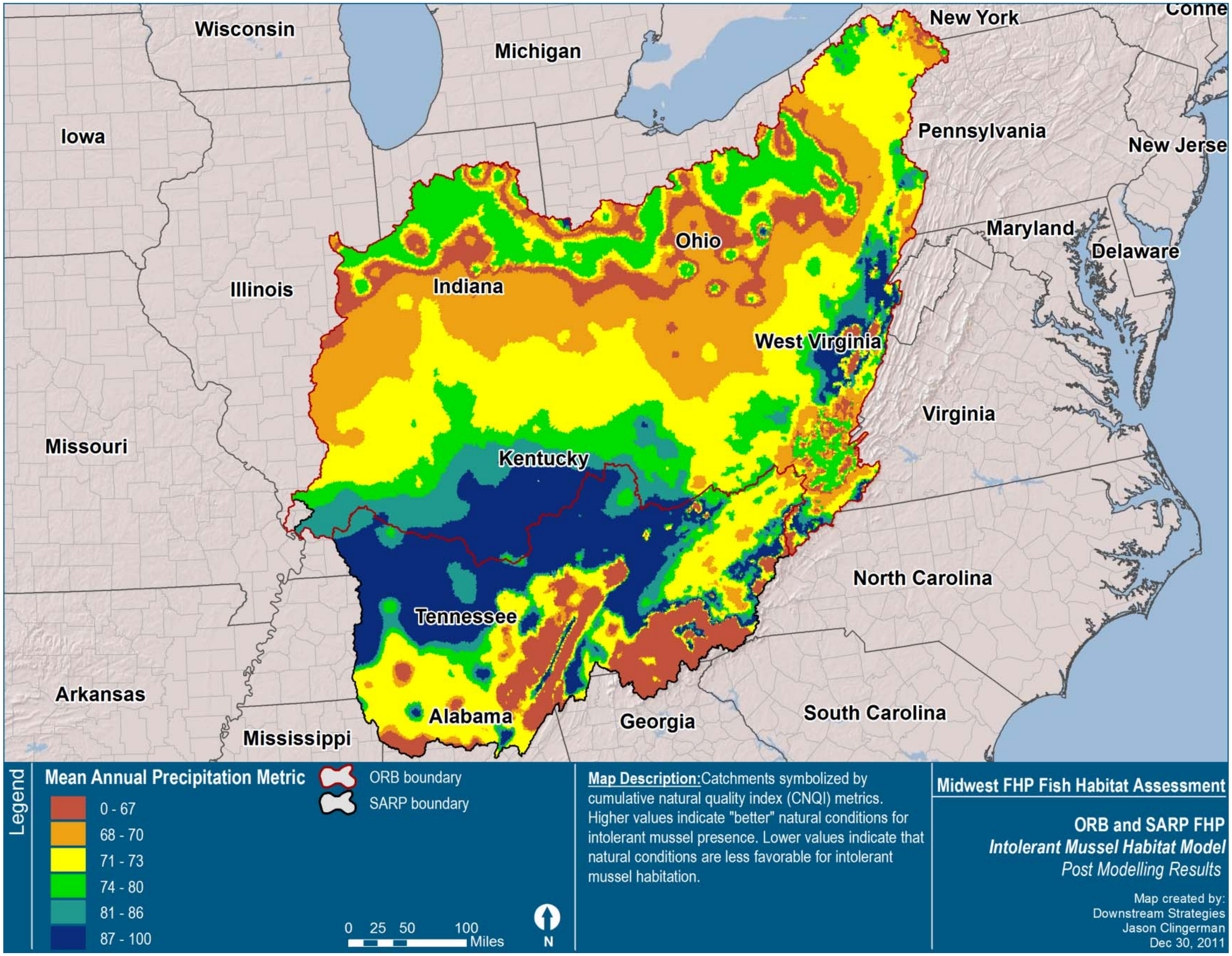


Figure 120: Fourth most influential natural index metric for intolerant mussels

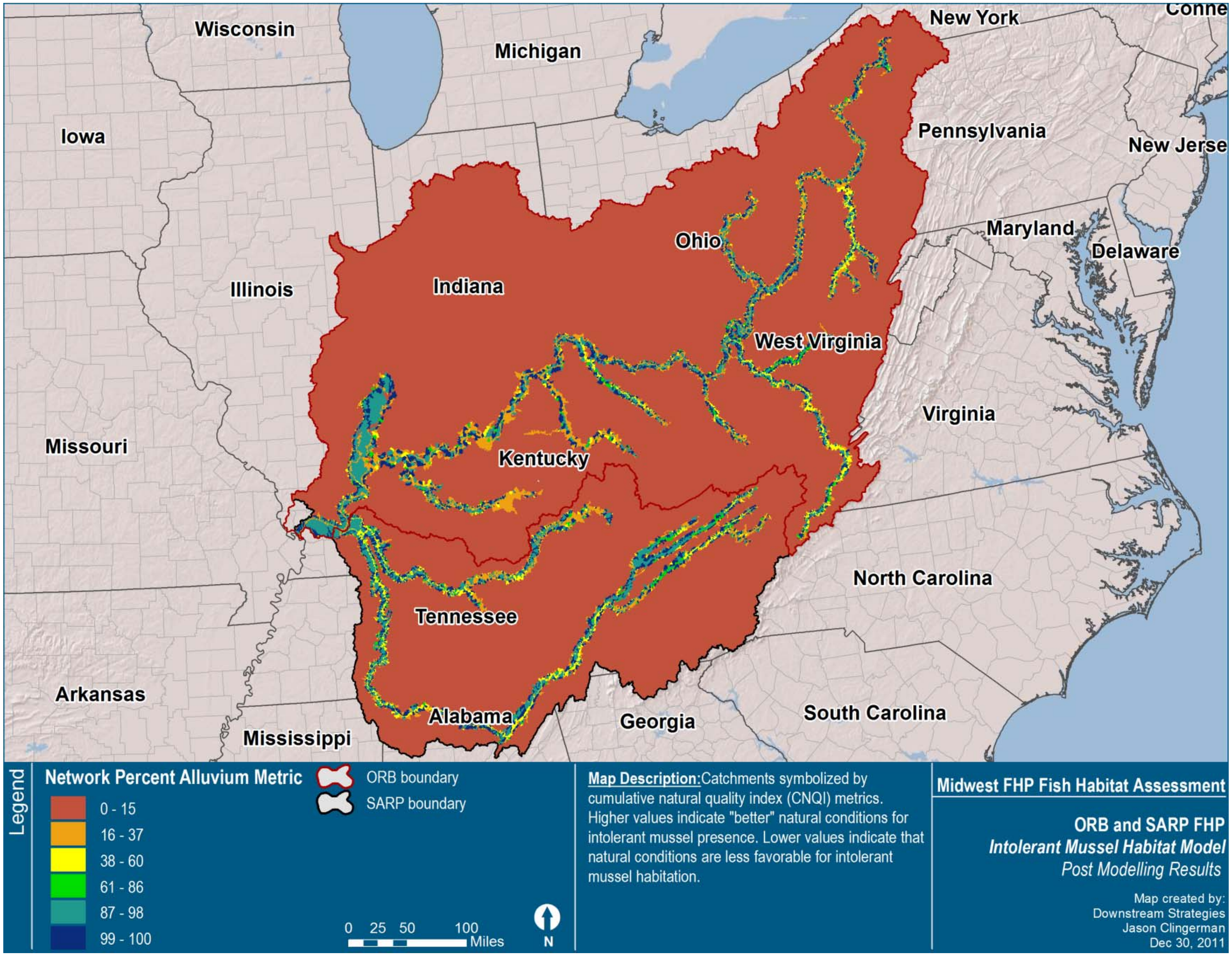


Figure 121: Fifth most influential natural index metric for intolerant mussels

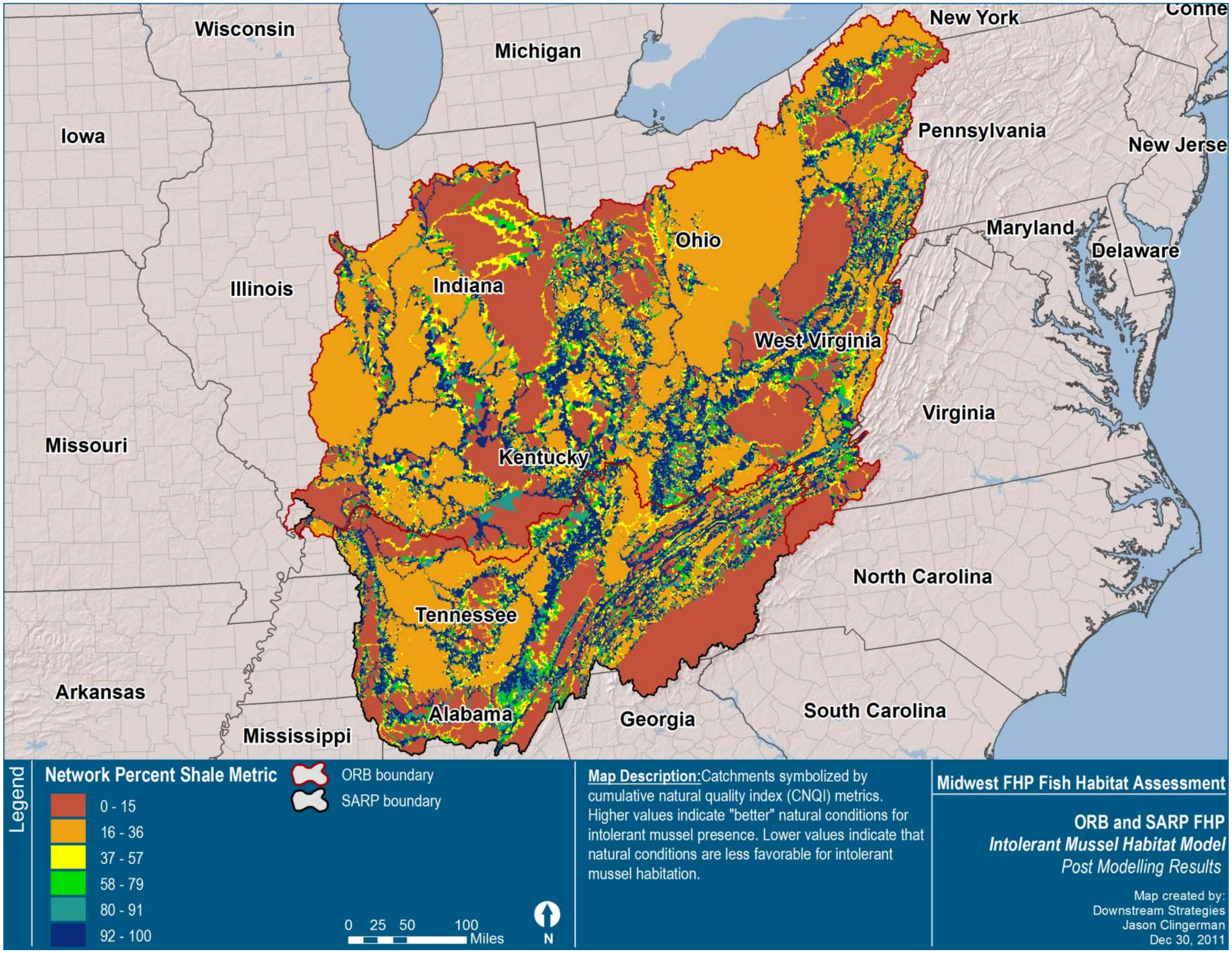


Figure 122: Most influential anthropogenic index metric for intolerant mussels

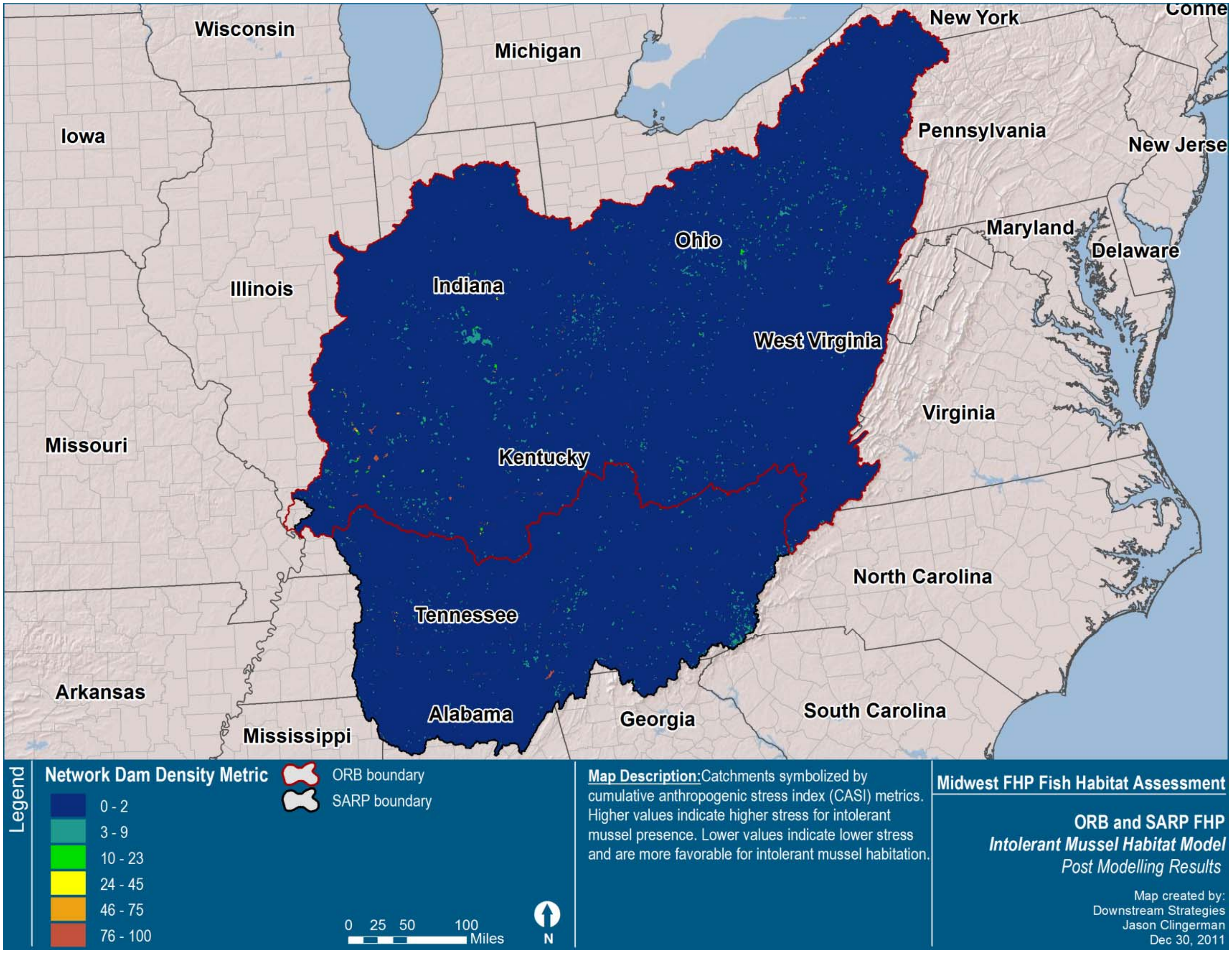


Figure 123: Second most influential anthropogenic index metric for intolerant mussels

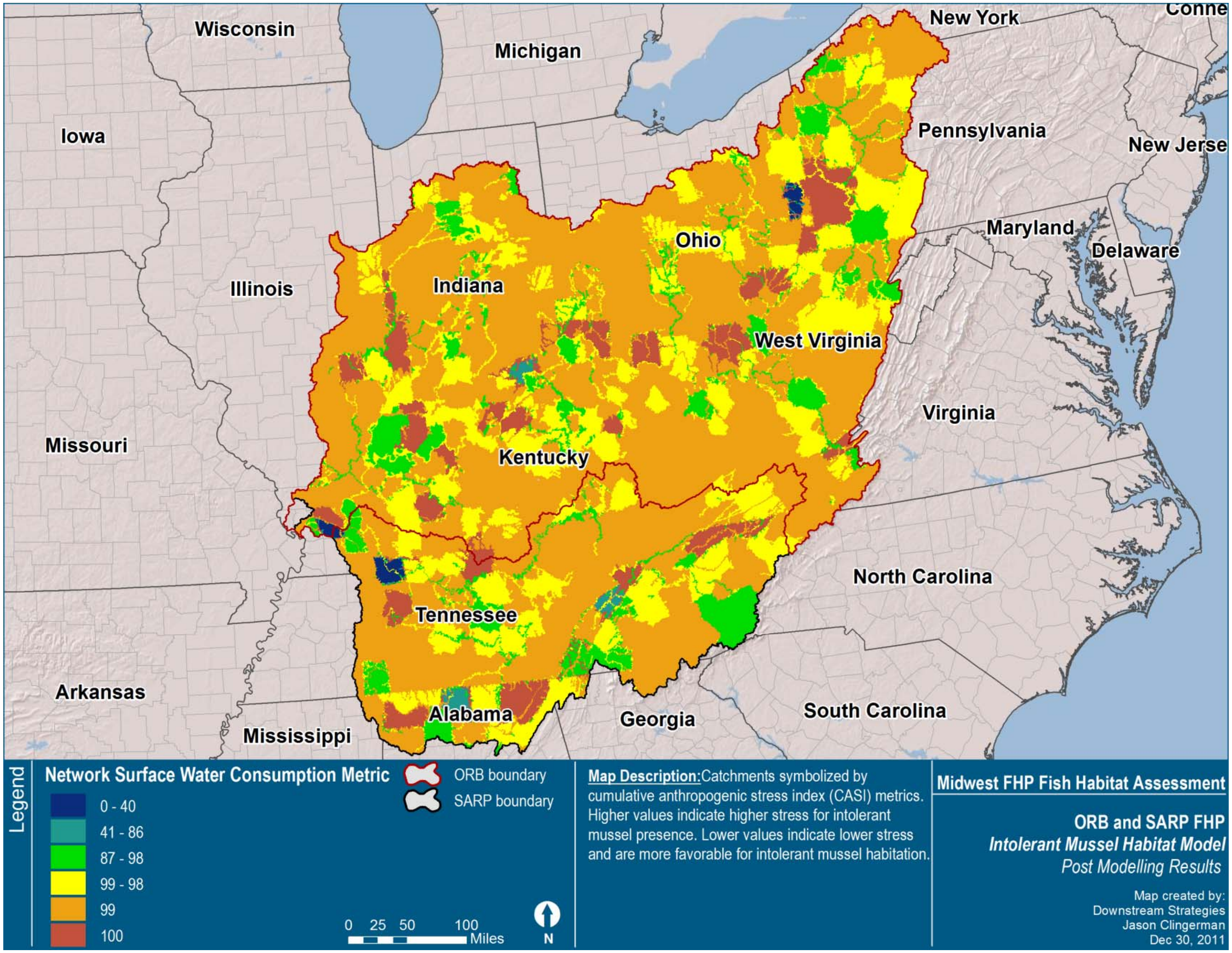


Figure 124: Third most influential anthropogenic index metric for intolerant mussels

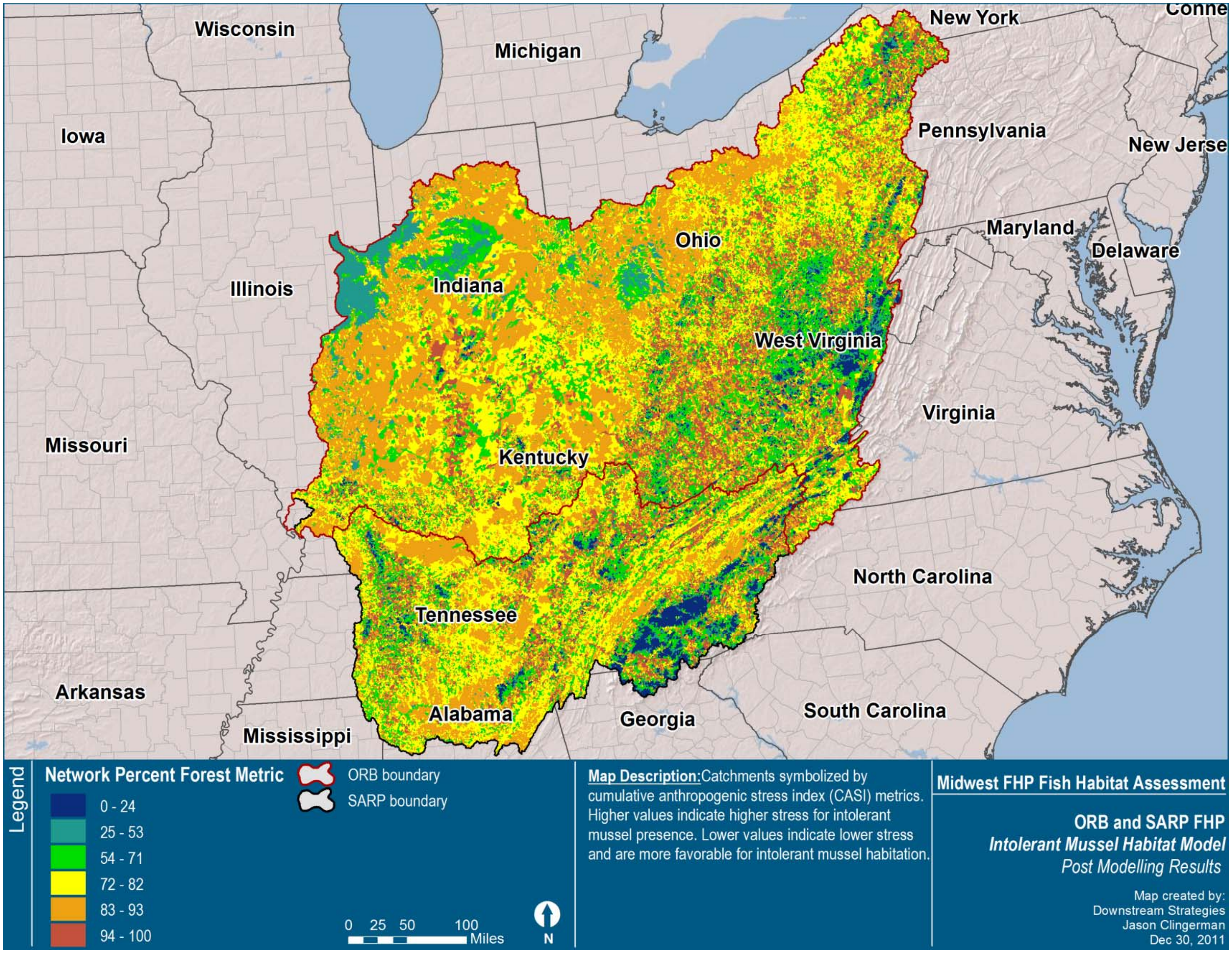


Figure 125: Fourth most influential anthropogenic index metric for intolerant mussels

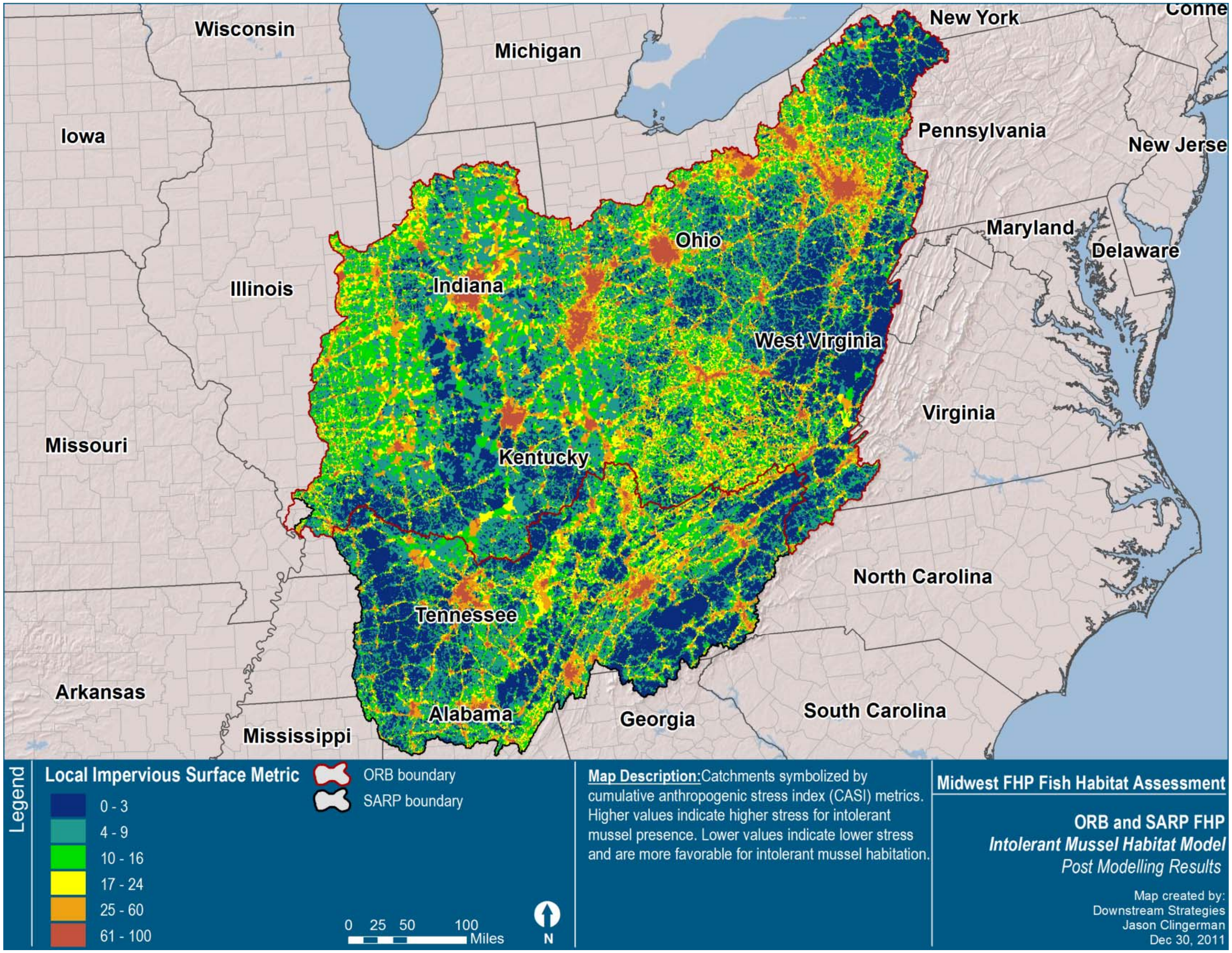
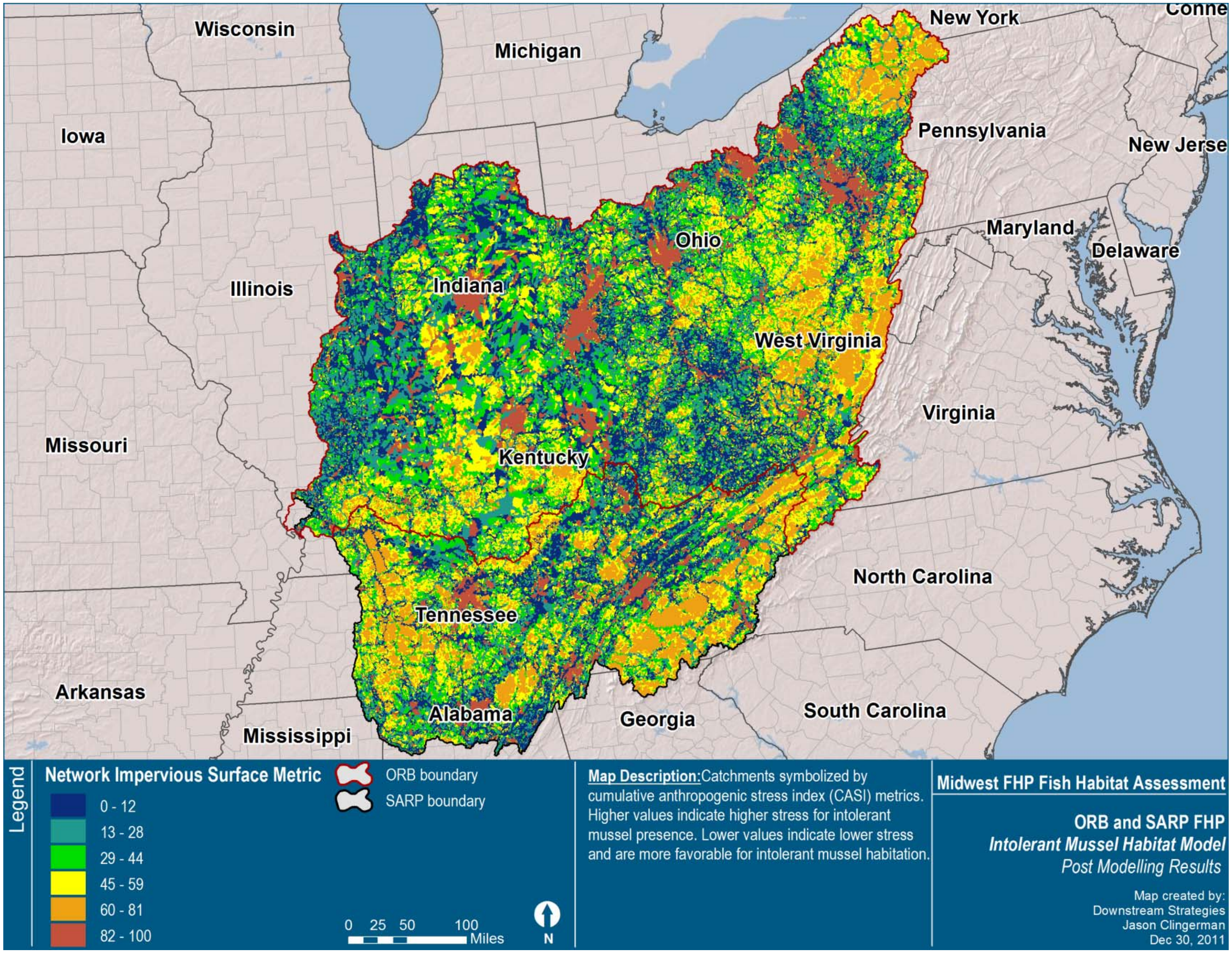


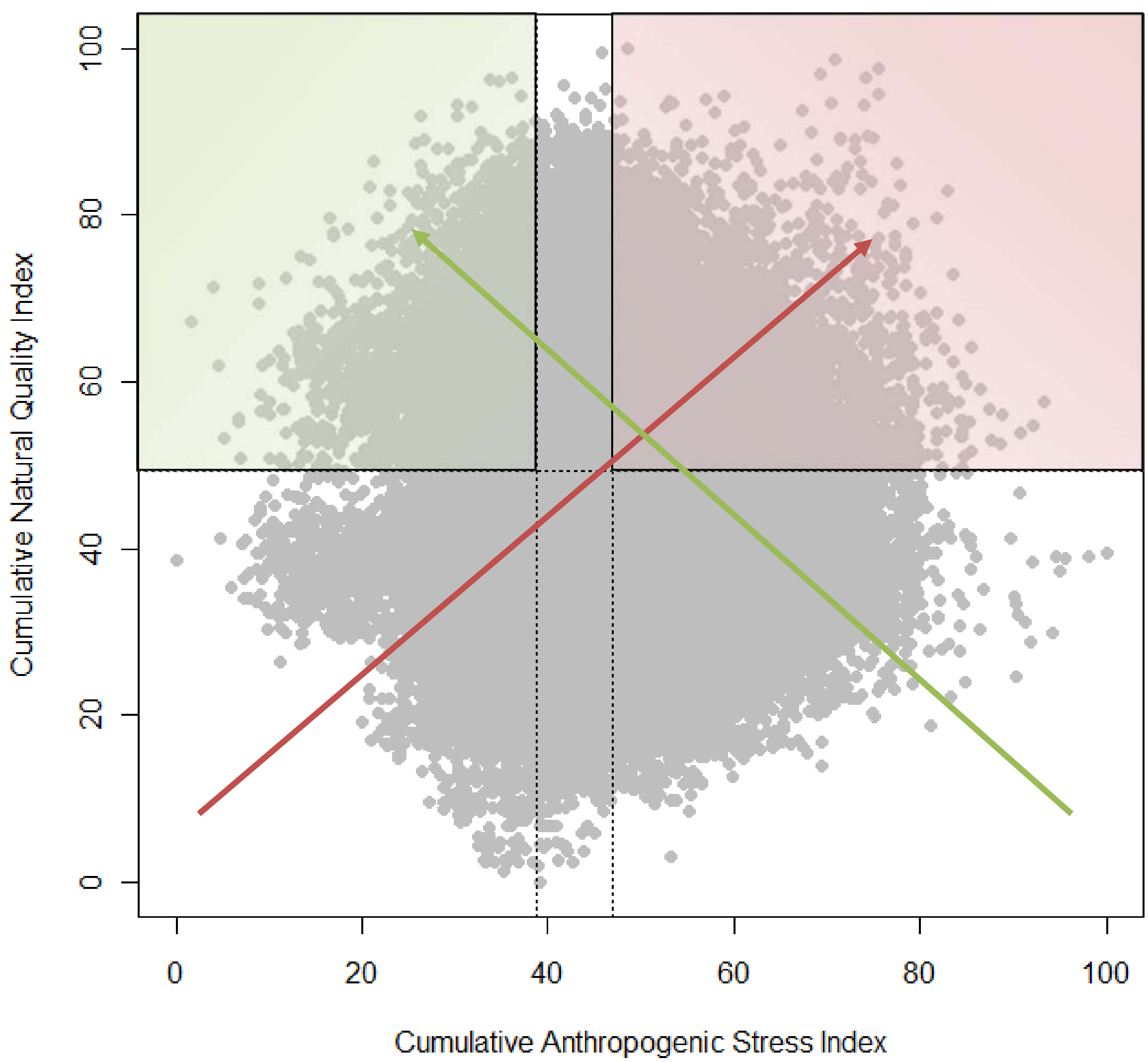
Figure 126: Fifth most influential anthropogenic index metric for intolerant mussels



8.4.4 *Restoration and protection priorities*

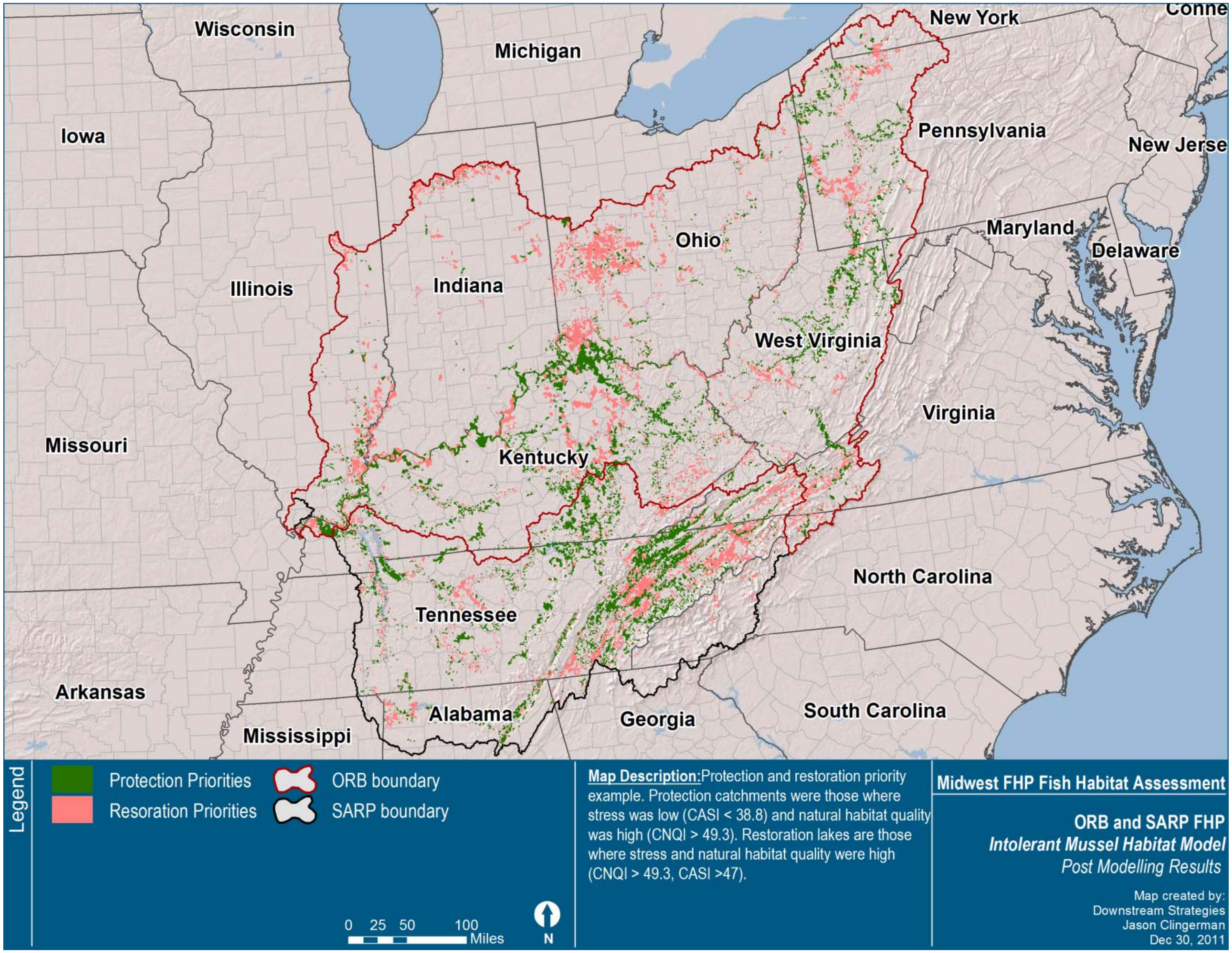
A plot of CNQI versus CASI values for all catchments in the study area (Figure 127) can be used as a reference when defining thresholds for categories of CNQI and CASI scores for use in the development of restoration and protection priorities. In the example shown (Figure 128), thresholds for restoration (high natural potential coupled with high anthropogenic stress) were set to CNQI greater than 49.3 and CASI greater than 47.0 (third quartiles). The thresholds used for protection (high natural potential and low anthropogenic stress) priorities were CNQI greater than 49.3 and CASI less than 38.8 (first quartile).

Figure 127: CNQI versus CASI values for all catchments for intolerant mussels



Note: Breakpoints for CNQI and CASI classes in this example are denoted by dashed lines. The arrows indicate the directions of increasing potential protection (green arrow) or restoration (red arrow) priority. The red box indicates catchments defined as restoration priorities under the example scenario. The green box indicates catchments defined as protection priorities under the same scenario.

Figure 128: Restoration and protection priorities for intolerant mussels



9. LIMITATIONS AND SUGGESTIONS FOR FUTURE WORK

In general, while the estimates of probability of presence, index scores, CNQI, and CASI generated through this assessment process represent a useful and objective means for assessing aquatic habitat and prioritizing habitats for restoration or protection, there are some limitations that are important to consider. Results generated through the modeling process are ultimately limited by the quality of data used to generate them. In the future, the model can be improved by improving the resolution and precision of the data. For example, some county-level data were used as predictor variables although the data likely generalize conditions at the catchment scale. In some cases, this resulted in generalizations in CASI or in the individual CASI metrics, which is evidenced by the visibly unnatural hard break lines at some county boundaries. Although these variables—such as network cattle density and network surface water consumption—were limited in spatial resolution, they still had high relative influence in the BRT model and were important to retain for predictive performance. In the future, refinement of these county-level variables or inclusion of higher resolution surrogates could improve both the precision of the BRT model predictions and post-modeling indices.

A second limitation is that the data and maps represent only a snapshot in time. Therefore, the models may not represent conditions before or after the data were collected or created. For example, any habitat lost or gained due to increased impervious surface cover since the 2006 National Land Cover Database (NLCD) was not considered in this assessment. Similarly, a portion of the uncertainty can be attributable to the temporal mismatches between the fish collection data and landscape data. As such, improving the temporal match between those datasets for future work would be beneficial.

Although the BRT statistical modeling algorithm automatically accounts for interactions between predictor variables, the post modeling process used to generate CNQI and CASI does not fully account for interactions. The post modeling outputs were derived from the relative influence and function plot outputs from BRT. Those outputs themselves are objective approximations useful for model interpretation and interrogation; however, they are not used in estimating the predicted values in BRT. Therefore, the individual function plots from which CASI and CNQI and their metrics were derived did not account for variable interactions because they are not accounted for in the BRT relative influence values or function plots. For example, while appropriately accounting for an interaction between ecoregion and stream size would require separate function plots for the effect of drainage area in each separate ecoregion, the function plot generated by BRT represents an average effect of drainage area across all ecoregions. In that example, the effect of drainage area is overestimated in some ecoregions and underestimated in others. This is not a limitation specific to the methodology used in this assessment, but common in other popular predictive modeling approaches. The advantage of the approach using BRT, however, is in the improved ability to predict current conditions (i.e., probability of presence) relative to other methods.

While continuous response variables can be modeled, binomial response variables can generally be modeled with greater precision in cases where the response data vary in collection method or date. Throughout this assessment, we have generally found that binomial (i.e., presence-absence) response variable models have performed better than continuous (i.e., abundance-based) variable models. In the future, basing diversity metrics on the presence-absence of targeted species, rather than relative abundance, may improve their precision.

There were also a few important issues that were beyond the scope of this project. Acid precipitation, biological interactions, and local habitat variation are all important in structuring fish communities. These variables were not directly used as predictor variables, although, when possible, surrogates were used to approximate variation in the model resulting from these processes.

Acid precipitation, for example, is prevalent throughout the Ohio River basin, especially in high-elevation headwater streams. Acid precipitation can cause declines in abundance, or shifts in fish communities. To attempt to account for acid precipitation, bedrock geology was used as a predictor variable. Bedrock geology is linked to soil acid-buffering capacity, and therefore the amount of acid reaching the stream.

Local habitat measures such as water quality (pH, alkalinity, instream temperature), physical habitat complexity, and substrate size are examples of local measures important to structuring fish communities. These measures could not be directly quantified in this analysis given the scope and scale of the project. However, since each catchment's land cover and geology was included in the analysis, some aspects of water quality were indirectly modeled. Likewise, habitat complexity and substrate size could be partially captured by the combination of stream slope and bedrock and surficial geology. Nonetheless, exclusion of detailed local measures likely accounts for some uncertainty in the model results. Thus, the results from this analysis should be combined with local expert knowledge and additional field data to arrive at the most accurate representation of habitat conditions.

In addition, inclusion of biological interactions in future models could improve the precision of the model and the ability to quantify its influence on the response variables. Specifically, important biological interactions in this system could include the negative interactions resulting from the introduction of non-native or other stocked fishes, such as brown trout or Asian carp.

One other specific limitation that could potentially explain additional variability in ORB models is the lack of a quality mining predictor variable. The only variable supplied, mines, is incomplete in nature. Surface mines were seemingly captured to an extent with the NLCD grassland variable but underground mines were apparently unmodeled as predictor variables. For a region where both surface and underground mining are so prevalent and pervasive, their inclusion would likely increase model accuracy significantly, at least throughout the coal mining regions.

Finally, another important consideration for managing aquatic habitat at this scale, which was not considered directly in this analysis, is climate change. Potential impacts from climate change include altered thermal regimes, stream flow regimes, and physical habitat. Particularly for coldwater fishes such as trout and sculpin, future warming could result in increased population isolation due to confinement to headwater habitats or more localized thermal refugia. Specifically, identifying catchments vulnerable to climate change and important to species of interest in this system—for example, those on the fringe of meeting the upper thermal criteria for a species—could represent an important and supplementary next step in the identification of restoration and protection priorities for targeted aquatic populations.

REFERENCES

Bear B (2006) Prioritization of eastern Wyoming prairie streams for conservation. Wyoming Game and Fish Department Administrative Report, Cheyenne, Wyo.

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. Journal of Animal Ecology 77: 802-13.

Friedman JH (2001) Greedy function approximation: a gradient boosting machine. Annals of Statistics 29: 1189-1232.

Friedman JH, Meulman JJ (2003) Multiple additive regression trees with application in epidemiology. Statistics in Medicine 22: 1365-81.

Patton T (2001) Application of a modified index of centers of density to four drainages within the Missouri River drainage, Wyoming. Wyoming Game and Fish Department Completion Report, Cheyenne, Wyo.

Appendix A: DATA DICTIONARY

Field	Description	Source	Data last modified
Comid	catchment comid (unique identifier)	NHDPlus	4/11/2011
Grid_code	grid code identifier	NHDPlus	4/11/2011
Grid_count	Number of cells in catchment grid, 30m	NHDPlus	4/11/2011
Prod_unit	NHDPlus production unit (subdivides the region)	NHDPlus	4/11/2011
Areasqkm	area of catchment, sq km	NHDPlus	4/11/2011
Huc8	8 digit Hydrologic Unit Code	Midwest FHAP	4/11/2011
Huc12	12 digit Hydrologic Unit Code (NRCS WBD)	Midwest FHAP	4/11/2011
Huc12_name	12 digit Hydrologic Unit Code Name (NRCS WBD)	Midwest FHAP	4/11/2011
Eco_code2	Ecoregion code (majority), Level II, catchment	USEPA Omernik Ecoregions for North America, Level II	4/11/2011
Eco_code3	Ecoregion code (majority), Level III, catchment	USEPA Omernik Ecoregions for North America, Level III	4/11/2011
Streamleve	Stream level	NHDFlowlineVAA.dbf	4/11/2011
Streamorde	Strahler stream order	NHDFlowlineVAA.dbf	4/11/2011
Fromnode	From node number (top of flowline)	NHDFlowlineVAA.dbf	4/11/2011
Tonode	To node number (bottom of flowline)	NHDFlowlineVAA.dbf	4/11/2011
Hydroseq	Hydrologic sequence number	NHDFlowlineVAA.dbf	4/11/2011
Levelpathi	Hydrologic sequence number of most downstream flowline in level path	NHDFlowlineVAA.dbf	4/11/2011
Pathlength	Distance to terminal flowline downstream along the mainpath (kilometers)	NHDFlowlineVAA.dbf	4/11/2011
Terminalpa	Hydrologic sequence number of terminal flowline	NHDFlowlineVAA.dbf	4/11/2011
Arbolatesu	An estimate of miles of stream upstream of a flowline. Always 0. (square kilometers	NHDFlowlineVAA.dbf	4/11/2011
Divergence	0 – not part of a divergence	NHDFlowlineVAA.dbf	4/11/2011
Startflag	1 – main path of a divergence	NHDFlowlineVAA.dbf	4/11/2011
Terminalfl	2 – minor path of a divergence	NHDFlowlineVAA.dbf	4/11/2011
Dnlevel	0 – not a headwater flowline	NHDFlowlineVAA.dbf	4/11/2011
Thinnercod	1 – a headwater flowline	NHDFlowlineVAA.dbf	4/11/2011
Uplevelpat	0 – not a terminal flowline	NHDFlowlineVAA.dbf	4/11/2011
Uphydroseq	1 – a terminal flowline	NHDFlowlineVAA.dbf	4/11/2011
Upminhydro	Streamlevel of mainstem downstream flowline	NHDFlowlineVAA.dbf	4/11/2011
Dnlevelpat	Ordinal value used to display various network densities	NHDFlowlineVAA.dbf	4/11/2011
Dnminhydro	Upstream mainstem level path identifier	NHDFlowlineVAA.dbf	4/11/2011

Dndraincou	Upstream mainstem hydrologic sequence number	NHDFlowlineVAA.dbf	4/11/2011
Cumdrainag	Cumulative drainage area in square kilometers	catchmentattributesflow.dbf	4/11/2011
Maflowu	Mean Annual Flow (cfs) at bottom of flowline as computed by Unit Runoff Method	catchmentattributesflow.dbf	4/11/2011
Maflowv	Mean Annual Flow (cfs) at bottom of flowline as computed by Vogel Method	catchmentattributesflow.dbf	4/11/2011
Mavelu	Mean Annual Velocity (fps) at bottom of flowline as computed by Unit Runoff Method	catchmentattributesflow.dbf	4/11/2011
Mavelv	Mean Annual Velocity (fps) at bottom of flowline as computed by Vogel Method	catchmentattributesflow.dbf	4/11/2011
Incrflowu	Incremental Flow (cfs) for Flowline as computed by the Unit Runoff Method	catchmentattributesflow.dbf	4/11/2011
Maxelevraw	Maximum elevation (unsmoothed) in meters	catchmentattributesflow.dbf	4/11/2011
Minelevraw	Minimum elevation (unsmoothed) in meters	catchmentattributesflow.dbf	4/11/2011
Maxelevsmo	Maximum elevation (smoothed) in meters	catchmentattributesflow.dbf	4/11/2011
Minelevsmo	Minimum elevation (smoothed) in meters	catchmentattributesflow.dbf	4/11/2011
Slope	Slope of flowline (cm/cm)	catchmentattributesflow.dbf	4/11/2011
Precip	Mean annual precipitation in mm	catchmentattributestempprecip.dbf	4/11/2011
Temp	Mean annual temperature in degrees centigrade * 10	catchmentattributestempprecip.dbf	4/11/2011
Mississ	Is catchment found along Mississippi River mainstem (1=yes, 0=no)	Jackie	4/11/2011
Warnings	warnings on catchment anomalies	based on NHDPlus attributes	4/11/2011
Wbareacomi	NHD waterbody/area feature COMID	NHD flowline attributes	4/11/2011
Ftype	NHD flowline feature type	NHD flowline attributes	4/11/2011
Fcode	NHD flowline feature code	NHD flowline attributes	4/11/2011
Wbftype	NHD waterbody feature type	NHD waterbody attributes	4/11/2011
Wbfcode	NHD waterbody feature code	NHD waterbody attributes	4/11/2011
Areaftype	NHD area feature type	NHD area attributes	4/11/2011
Areafcode	NHD area feature code	NHD area attributes	4/11/2011
Catchtyype	Catchment flowline feature type (flowline and waterbody/area combined)	based on NHD	4/11/2011
Nlcd0611a	NLCD 2006 open water, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0611p	NLCD 2006 open water, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0621a	NLCD 2006 developed, open space, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0621p	NLCD 2006 developed, open space, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0622a	NLCD 2006 developed, low intensity, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0622p	NLCD 2006 developed, low intensity, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0623a	NLCD 2006 developed, medium intensity, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0623p	NLCD 2006 developed, medium intensity, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0624a	NLCD 2006 developed, high intensity, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0624p	NLCD 2006 developed, high intensity, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0631a	NLCD 2006 barren land (rock/sand/clay), area (sq km), catchment	NLCD 2006	4/11/2011

Nlcd0631p	NLCD 2006 barren land (rock/sand/clay), area (%), catchment	NLCD 2006	4/11/2011
Nlcd0641a	NLCD 2006 deciduous forest, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0641p	NLCD 2006 deciduous forest, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0642a	NLCD 2006 evergreen forest, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0642p	NLCD 2006 evergreen forest, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0643a	NLCD 2006 mixed forest, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0643p	NLCD 2006 mixed forest, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0652a	NLCD 2006 shrub/scrub, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0652p	NLCD 2006 shrub/scrub, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0671a	NLCD 2006 grassland/herbaceous, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0671p	NLCD 2006 grassland/herbaceous, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0681a	NLCD 2006 pasture/hay, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0681p	NLCD 2006 pasture/hay, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0682a	NLCD 2006 cultivated crops, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0682p	NLCD 2006 cultivated crops, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0690a	NLCD 2006 woody wetlands, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0690p	NLCD 2006 woody wetlands, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0695a	NLCD 2006 emergent herbaceous wetlands, area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd0695p	NLCD 2006 emergent herbaceous wetlands, area (%), catchment	NLCD 2006	4/11/2011
Nlcd0611ac	NLCD 2006 open water, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0611pc	NLCD 2006 open water, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0621ac	NLCD 2006 developed, open space, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0621pc	NLCD 2006 developed, open space, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0622ac	NLCD 2006 developed, low intensity, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0622pc	NLCD 2006 developed, low intensity, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0623ac	NLCD 2006 developed, medium intensity, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0623pc	NLCD 2006 developed, medium intensity, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0624ac	NLCD 2006 developed, high intensity, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0624pc	NLCD 2006 developed, high intensity, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0631ac	NLCD 2006 barren land (rock/sand/clay), area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0631pc	NLCD 2006 barren land (rock/sand/clay), area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0641ac	NLCD 2006 deciduous forest, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0641pc	NLCD 2006 deciduous forest, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0642ac	NLCD 2006 evergreen forest, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0642pc	NLCD 2006 evergreen forest, area (%), upstream cumulative	NLCD 2006	4/11/2011

Nlcd0643ac	NLCD 2006 mixed forest, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0643pc	NLCD 2006 mixed forest, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0652ac	NLCD 2006 shrub/scrub, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0652pc	NLCD 2006 shrub/scrub, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0671ac	NLCD 2006 grassland/herbaceous, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0671pc	NLCD 2006 grassland/herbaceous, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0681ac	NLCD 2006 pasture/hay, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0681pc	NLCD 2006 pasture/hay, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0682ac	NLCD 2006 cultivated crops, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0682pc	NLCD 2006 cultivated crops, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0690ac	NLCD 2006 woody wetlands, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0690pc	NLCD 2006 woody wetlands, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0695ac	NLCD 2006 emergent herbaceous wetlands, area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd0695pc	NLCD 2006 emergent herbaceous wetlands, area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd06deva	NLCD 2006 Developed land cover classes (21, 22, 23, 24), area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd06devp	NLCD 2006 Developed land cover classes (21, 22, 23, 24), area (%), catchment	NLCD 2006	4/11/2011
Nlcd06fora	NLCD 2006 Forested land cover classes (41, 42, 43), area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd06forp	NLCD 2006 Forested land cover classes (41, 42, 43), area (%), catchment	NLCD 2006	4/11/2011
Nlcd06agac	NLCD 2006 Agriculture land cover classes (81, 82), area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd06agpc	NLCD 2006 Agriculture land cover classes (81, 82), area (%), catchment	NLCD 2006	4/11/2011
Nlcd06weta	NLCD 2006 Wetland land cover classes (90, 95), area (sq km), catchment	NLCD 2006	4/11/2011
Nlcd06wetp	NLCD 2006 Wetland land cover classes (90, 95), area (%), catchment	NLCD 2006	4/11/2011
Nlcd06devac	NLCD 2006 Developed land cover classes (21, 22, 23, 24), area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd06devpc	NLCD 2006 Developed land cover classes (21, 22, 23, 24), area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd06forac	NLCD 2006 Forested land cover classes (41, 42, 43), area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd06forpc	NLCD 2006 Forested land cover classes (41, 42, 43), area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd06agac	NLCD 2006 Agriculture land cover classes (81, 82), area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd06agpc	NLCD 2006 Agriculture land cover classes (81, 82), area (%), upstream cumulative	NLCD 2006	4/11/2011
Nlcd06wetac	NLCD 2006 Wetland land cover classes (90, 95), area (sq km), upstream cumulative	NLCD 2006	4/11/2011
Nlcd06wetpc	NLCD 2006 Wetland land cover classes (90, 95), area (%), upstream cumulative	NLCD 2006	4/11/2011
Impervs	Impervious surface area (allocation per segment): area (km2)	NLCD 2001 Impervious Surface Area	4/11/2011
Impervsc	Impervious surface area (accumulation of upstream segments): total upstream area (km2)	NLCD 2001 Impervious Surface Area	4/11/2011
Nlcd11a	Open Water/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd11p	Open Water/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd12a	Perennial Ice/Snow/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011

Nlcd12p	Perennial Ice/Snow/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd21a	Developed, Open Space/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd21p	Developed, Open Space/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd22a	Developed, Low Intensity/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd22p	Developed, Low Intensity/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd23a	Developed, Medium Intensity/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd23p	Developed, Medium Intensity/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd24a	Developed, High Intensity/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd24p	Developed, High Intensity/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd31a	Barren Land (Rock/Sand/Clay)/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd31p	Barren Land (Rock/Sand/Clay)/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd32a	Unconsolidated Shore/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd32p	Unconsolidated Shore/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd41a	Deciduous Forest/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd41p	Deciduous Forest/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd42a	Evergreen Forest/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd42p	Evergreen Forest/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd43a	Mixed Forest/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd43p	Mixed Forest/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd51a	Dwarf Scrub/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd51p	Dwarf Scrub/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd52a	Shrub/Scrub/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd52p	Shrub/Scrub/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd71a	Grassland/Herbaceous/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd71p	Grassland/Herbaceous/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd72a	Sedge/Herbaceous/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd72p	Sedge/Herbaceous/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd73a	Lichens/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd73p	Lichens/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd74a	Moss/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd74p	Moss/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd81a	Pasture/Hay/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd81p	Pasture/Hay/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd82a	Cultivated Crops/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd82p	Cultivated Crops/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011

Nlcd90a	Woody Wetlands/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd90p	Woody Wetlands/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd91a	Palustrine Forested Wetland/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd91p	Palustrine Forested Wetland/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd92a	Palustrine Scrub/Shrub Wetland/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd92p	Palustrine Scrub/Shrub Wetland/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd93a	Estuarine Forested Wetland/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd93p	Estuarine Forested Wetland/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd94a	Estuarine Scrub/Shrub Wetland/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd94p	Estuarine Scrub/Shrub Wetland/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd95a	Emergent Herbaceous Wetlands/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd95p	Emergent Herbaceous Wetlands/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd96a	Palustrine Emergent Wetland (Persistent)/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd96p	Palustrine Emergent Wetland (Persistent)/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd97a	Estuarine Emergent Wetland/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd97p	Estuarine Emergent Wetland/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd98a	Palustrine Aquatic Bed/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd98p	Palustrine Aquatic Bed/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd99a	Estuarine Aquatic Bed/allocation per segment: area in square kilometers	2001 NLCD Allocation tables	4/11/2011
Nlcd99p	Estuarine Aquatic Bed/allocation per segment: area-weighted percent	2001 NLCD Allocation tables	4/11/2011
Nlcd11ac	Open Water/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd11pc	Open Water/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd12ac	Perennial Ice/Snow/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd12pc	Perennial Ice/Snow/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd21ac	Developed, Open Space/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd21pc	Developed, Open Space/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd22ac	Developed, Low Intensity/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd22pc	Developed, Low Intensity/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd23ac	Developed, Medium Intensity/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd23pc	Developed, Medium Intensity/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd24ac	Developed, High Intensity/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd24pc	Developed, High Intensity/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd31ac	Barren Land (Rock/Sand/Clay)/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd31pc	Barren Land (Rock/Sand/Clay)/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd32ac	Unconsolidated Shore/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011

Nlcd32pc	Unconsolidated Shore/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd41ac	Deciduous Forest/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd41pc	Deciduous Forest/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd42ac	Evergreen Forest/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd42pc	Evergreen Forest/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd43ac	Mixed Forest/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd43pc	Mixed Forest/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd51ac	Dwarf Scrub/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd51pc	Dwarf Scrub/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd52ac	Shrub/Scrub/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd52pc	Shrub/Scrub/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd71ac	Grassland/Herbaceous/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd71pc	Grassland/Herbaceous/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd72ac	Sedge/Herbaceous/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd72pc	Sedge/Herbaceous/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd73ac	Lichens/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd73pc	Lichens/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd74ac	Moss/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd74pc	Moss/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd81ac	Pasture/Hay/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd81pc	Pasture/Hay/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd82ac	Cultivated Crops/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd82pc	Cultivated Crops/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd90ac	Woody Wetlands/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd90pc	Woody Wetlands/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd91ac	Palustrine Forested Wetland/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd91pc	Palustrine Forested Wetland/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd92ac	Palustrine Scrub/Shrub Wetland/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd92pc	Palustrine Scrub/Shrub Wetland/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd93ac	Estuarine Forested Wetland/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd93pc	Estuarine Forested Wetland/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd94ac	Estuarine Scrub/Shrub Wetland/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd94pc	Estuarine Scrub/Shrub Wetland/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd95ac	Emergent Herbaceous Wetlands/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd95pc	Emergent Herbaceous Wetlands/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011

Nlcd96ac	Palustrine Emergent Wetland (Persistent)/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd96pc	Palustrine Emergent Wetland (Persistent)/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd97ac	Estuarine Emergent Wetland/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd97pc	Estuarine Emergent Wetland/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd98ac	Palustrine Aquatic Bed/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd98pc	Palustrine Aquatic Bed/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Nlcd99ac	Estuarine Aquatic Bed/upstream accumulation: area in square kilometers	2001 NLCD Accumulation tables	4/11/2011
Nlcd99pc	Estuarine Aquatic Bed/upstream accumulation: area-weighted percent	2001 NLCD Accumulation tables	4/11/2011
Imp06avg	NLCD 2006 percent impervious, average, catchment	NLCD 2006	4/11/2011
Imp06avgc	NLCD 2006 percent impervious, average, upstream cumulative	NLCD 2006	4/11/2011
Soil0a	Revised soil hydrologic group code 0 (urban areas/water), area (sq km), catchment	STATSGO	4/11/2011
Soil0p	Revised soil hydrologic group code 0 (urban areas/water), area (%), catchment	STATSGO	4/11/2011
Soil1a	Revised soil hydrologic group code 1 (A), area (sq km), catchment	STATSGO	4/11/2011
Soil1p	Revised soil hydrologic group code 1 (A), area (%), catchment	STATSGO	4/11/2011
Soil2a	Revised soil hydrologic group code 2 (B), area (sq km), catchment	STATSGO	4/11/2011
Soil2p	Revised soil hydrologic group code 2 (B), area (%), catchment	STATSGO	4/11/2011
Soil3a	Revised soil hydrologic group code 3 (C), area (sq km), catchment	STATSGO	4/11/2011
Soil3p	Revised soil hydrologic group code 3 (C), area (%), catchment	STATSGO	4/11/2011
Soil4a	Revised soil hydrologic group code 4 (D), area (sq km), catchment	STATSGO	4/11/2011
Soil4p	Revised soil hydrologic group code 4 (D), area (%), catchment	STATSGO	4/11/2011
Soil0ac	Revised soil hydrologic group code 0 (urban areas/water), area (sq km), upstream cumulative	STATSGO	4/11/2011
Soil0pc	Revised soil hydrologic group code 0 (urban areas/water), area (%), upstream cumulative	STATSGO	4/11/2011
Soil1ac	Revised soil hydrologic group code 1 (A), area (sq km), upstream cumulative	STATSGO	4/11/2011
Soil1pc	Revised soil hydrologic group code 1 (A), area (%), upstream cumulative	STATSGO	4/11/2011
Soil2ac	Revised soil hydrologic group code 2 (B), area (sq km), upstream cumulative	STATSGO	4/11/2011
Soil2pc	Revised soil hydrologic group code 2 (B), area (%), upstream cumulative	STATSGO	4/11/2011
Soil3ac	Revised soil hydrologic group code 3 (C), area (sq km), upstream cumulative	STATSGO	4/11/2011
Soil3pc	Revised soil hydrologic group code 3 (C), area (%), upstream cumulative	STATSGO	4/11/2011
Soil4ac	Revised soil hydrologic group code 4 (D), area (sq km), upstream cumulative	STATSGO	4/11/2011
Soil4pc	Revised soil hydrologic group code 4 (D), area (%), upstream cumulative	STATSGO	4/11/2011
Bedr1a	Bedrock geology, group code 1 (Carbonate), area (sq km), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr1p	Bedrock geology, group code 1 (Carbonate), area (%), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr2a	Bedrock geology, group code 2 (felsic-igneous), area (sq km), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr2p	Bedrock geology, group code 2 (felsic-igneous), area (%), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr3a	Bedrock geology, group code 3 (mafic-igneous), area (sq km), catchment	USGS state geologic maps for Midwest	4/11/2011

Bedr3p	Bedrock geology, group code 3 (mafic-igneous), area (%), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr4a	Bedrock geology, group code 4 (metamorphic), area (sq km), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr4p	Bedrock geology, group code 4 (metamorphic), area (%), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr5a	Bedrock geology, group code 5 (sand and gravel), area (sq km), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr5p	Bedrock geology, group code 5 (sand and gravel), area (%), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr6a	Bedrock geology, group code 6 (sandstone), area (sq km), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr6p	Bedrock geology, group code 6 (sandstone), area (%), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr7a	Bedrock geology, group code 7 (shale), area (sq km), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr7p	Bedrock geology, group code 7 (shale), area (%), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr8a	Bedrock geology, group code 8 (water), area (sq km), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr8p	Bedrock geology, group code 8 (water), area (%), catchment	USGS state geologic maps for Midwest	4/11/2011
Bedr1ac	Bedrock geology, group code 1 (Carbonate), area (sq km), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr1pc	Bedrock geology, group code 1 (Carbonate), area (%), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr2ac	Bedrock geology, group code 2 (felsic-igneous), area (sq km), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr2pc	Bedrock geology, group code 2 (felsic-igneous), area (%), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr3ac	Bedrock geology, group code 3 (mafic-igneous), area (sq km), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr3pc	Bedrock geology, group code 3 (mafic-igneous), area (%), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr4ac	Bedrock geology, group code 4 (metamorphic), area (sq km), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr4pc	Bedrock geology, group code 4 (metamorphic), area (%), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr5ac	Bedrock geology, group code 5 (sand and gravel), area (sq km), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr5pc	Bedrock geology, group code 5 (sand and gravel), area (%), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr6ac	Bedrock geology, group code 6 (sandstone), area (sq km), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr6pc	Bedrock geology, group code 6 (sandstone), area (%), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr7ac	Bedrock geology, group code 7 (shale), area (sq km), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr7pc	Bedrock geology, group code 7 (shale), area (%), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr8ac	Bedrock geology, group code 8 (water), area (sq km), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Bedr8pc	Bedrock geology, group code 8 (water), area (%), upstream cumulative	USGS state geologic maps for Midwest	4/11/2011
Ethanol	Number of ethanol plants, catchment	Renewable Fuels Association website	4/11/2011
Wetlanda	wetland feature total area, sq km, catchment	NWI and GAP wetlands	4/11/2011
Wetlandp	wetland area, percent of catchment's area (o to 100%)	NWI and GAP wetlands	4/11/2011
Imp303	impaired stream (303d listed) present in catchment (0=no, 1=yes)	EPA 303(D) listings by state	4/11/2011
BFImean	Base Flow Index, mean (catchment)	USGS	4/11/2011
rechgMean	Recharge (estimated mean annual natural groundwater discharge, mm/yr), catchment mean	USGS Open File Report 03-311	4/11/2011
N_kg	total estimated Nitrogen inputs, kg/year, catchment	USGS	4/11/2011
P_kg	total estimated Phosphorous inputs, kg/year, catchment	USGS	4/11/2011

N_kgden	total estimated N inputs (kg/year), per sq km, catchment	USGS	4/11/2011
P_kgden	total estimated P inputs (kg/year), per sq km, catchment	USGS	4/11/2011
AreaSqkmC	area of catchment's upstream contributing area, sq km	calculated	4/11/2011
catchesC	number of upstream catchments	calculated	4/11/2011
ethanolC	number of ethanol plants, upstream contributing area	Renewable Fuels Association website	4/11/2011
wetlandAC	wetland area, sq km, upstream contributing area	NWI and GAP wetlands	4/11/2011
wetlandPC	wetland area, percent of upstream contributing area (0 to 100%)	NWI and GAP wetlands	4/11/2011
imp303C	number of upstream catchments within impairment (303D listings in catchment)	EPA 303(D) listings by state	4/11/2011
imp303PC	percent of upstream catchments within impairment (303D listings)	EPA 303(D) listings by state	4/11/2011
BFImeanC	Base Flow Index, mean (upstream contributing area)	USGS	4/11/2011
rechgMeanC	Recharge (estimated mean annual natural groundwater discharge, mm/year), mean value for all upstream catchments	USGS Open File Report 03-311	4/11/2011
N_kgC	total estimated Nitrogen inputs, kg/year, upstream cumulative	USGS	4/11/2011
P_kgC	total estimated Phosphorous inputs, kg/year, upstream cumulative	USGS	4/11/2011
N_kgdenC	total estimated N inputs (kg/year), per sq km, upstream cumulative	USGS	4/11/2011
P_kgdenC	total estimated P inputs (kg/year), per sq km, upstream cumulative	USGS	4/11/2011
Water_gw	LOCAL: USGS National Atlas of the US: Ground Water consumption by COUNTY 2000: Millions gallons per day/km2	local_disturbance_variables.dbf	4/11/2011
Water_sw	LOCAL: USGS National Atlas of the US: Surface Water consumption by COUNTY 2000: Millions gallons per day/km2	local_disturbance_variables.dbf	4/11/2011
Cattle	LOCAL: Agricultural Census 2002, 1:2M scale, INTEGER: average number of cattle/acre farmland	local_disturbance_variables.dbf	4/11/2011
Popdens	LOCAL: US Population Density 2000, NOAA, scale 1km, #/km2	local_disturbance_variables.dbf	4/11/2011
Roadcr	LOCAL: Census 2000 TIGER Roads, 1:100K scale, road crossings identified by INTERSECT, with points generated, #/km2	local_disturbance_variables.dbf	4/11/2011
Roadlen	LOCAL: Census 2000 TIGER Roads, 1:100K scale, units not given - m/km2	local_disturbance_variables.dbf	4/11/2011
Dams	LOCAL: National Inventory of Dams, 2002-2004, #/km2	local_disturbance_variables.dbf	4/11/2011
Mines	LOCAL: USGS Active Mines and Mineral Processing Plants, 2003, #/km2	local_disturbance_variables.dbf	4/11/2011
Tri	LOCAL: USEPA, 2007: #/km2 Toxics Release Inventory Program sites	local_disturbance_variables.dbf	4/11/2011
Npdes	LOCAL: USEPA, 2007: #/km2 National Pollutant Discharge Elimination System permits	local_disturbance_variables.dbf	4/11/2011
Cerc	LOCAL: USEPA, 2007: #/km2 Compensation and Liability Information System permits	local_disturbance_variables.dbf	4/11/2011
Water_gwc	NETWORK: USGS National Atlas of the US: Ground Water consumption by COUNTY 2000: Millions gallons per day/km2	network_disturbance_variables.dbf	4/11/2011
Water_swc	NETWORK: USGS National Atlas of the US: Surface Water consumption by COUNTY 2000: Millions gallons per day/km2	network_disturbance_variables.dbf	4/11/2011
Cattlec	NETWORK: Agricultural Census 2002, 1:2M scale, INTEGER: average number of cattle/acre farmland	network_disturbance_variables.dbf	4/11/2011
Popdensc	NETWORK: US Population Density 2000, NOAA, scale 1km, #/km2	network_disturbance_variables.dbf	4/11/2011

Roadcrc	NETWORK: Census 2000 TIGER Roads, 1:100K scale, road crossings identified by INTERSECT, with points generated, #/km2	network_disturbance_variables.dbf	4/11/2011
Roadlenc	NETWORK: Census 2000 TIGER Roads, 1:100K scale, units not given - m/km2	network_disturbance_variables.dbf	4/11/2011
Damsc	NETWORK: National Inventory of Dams, 2002-2004, #/km2	network_disturbance_variables.dbf	4/11/2011
Minesc	NETWORK: USGS Active Mines and Mineral Processing Plants, 2003, #/km2	network_disturbance_variables.dbf	4/11/2011
Tric	NETWORK: USEPA, 2007: #/km2 Toxics Release Inventory Program sites	network_disturbance_variables.dbf	4/11/2011
Npdesc	NETWORK: USEPA, 2007: #/km2 National Pollutant Discharge Elimination System permits	network_disturbance_variables.dbf	4/11/2011
Cercc	NETWORK: USEPA, 2007: #/km2 Compensation and Liability Information System permits	network_disturbance_variables.dbf	4/11/2011

Table 24: Intolerant fish species list

Scientific name	Common name
Ammocrypta clara	Western sand darter
Ammocrypta pellucida	Eastern sand darter
Ammocrypta vivax	Scaly sand darter
Carpiodes velifer	Highfin carpsucker
Clinostomus elongatus	Redside dace
Clinostomus funduloides	Rosyside dace
Cottus bairdii	Mottled sculpin
Crystallaria asprella	Crystal darter
Cycleptus elongatus	Blue sucker
Cyprinella analostana	Satinfin shiner
Cyprinella camura	Bluntface shiner
Cyprinella galactura	Whitetail shiner
Erimystax dissimilis	Streamline chub
Etheostoma baileyi	Emerald darter
Etheostoma barbouri	Teardrop darter
Etheostoma barrenense	Splendid darter
Etheostoma bellum	Orangefin darter
Etheostoma camurum	Bluebreast darter
Etheostoma cinereum	Ashy darter
Etheostoma flavum	Saffron darter
Etheostoma histrio	Harlequin darter
Etheostoma maculatum	Spotted darter
Etheostoma microlepidum	Smallscale darter
Etheostoma obeyense	Barcheek darter
Etheostoma parvipinne	Goldstripe darter
Etheostoma percnurum	Duskytail darter
Etheostoma proeliare	Cypress darter
Etheostoma pyrrhogaster	Firebelly darter
Etheostoma rafinesquei	Kentucky darter
Etheostoma rufilineatum	Redline darter
Etheostoma sagitta	Arrow darter
Etheostoma sanguifluum	Bloodfin darter
Etheostoma simoterum	Snubnose darter
Etheostoma smithi	Slabrock darter
Etheostoma stigmaeum	Speckled darter
Etheostoma swaini	Gulf darter
Etheostoma swannanoa	Swannanoa darter
Etheostoma tippecanoe	Tippecanoe darter
Etheostoma variatum	Variegate darter
Etheostoma virgatum	Striped darter
Etheostoma vitreum	Glassy darter
Etheostoma zonale	Banded darter
Etheostoma zonistium	Bandfin darter
Exoglossum laurae	Tonguetied minnow
Exoglossum maxillingua	Cutlip minnow
Fundulus catenatus	Northern studfish

Fundulus dispar	Starhead topminnow
Hiodon alosoides	Goldeye
Hiodon tergisus	Mooneye
Hybopsis amblops	Bigeye chub
Hybopsis amnis	Pallid shiner
Hypentelium etowanum	Alabama hog sucker
Hypentelium nigricans	Northern hog sucker
Ichthyomyzon bdellium	Ohio lamprey
Ichthyomyzon fossor	Northern brook lamprey
Ichthyomyzon gagei	Southern brook lamprey
Ichthyomyzon greeleyi	Mountain brook lamprey
Lampetra appendix	American brook lamprey
Lepomis miniatus	Redspotted sunfish
Luxilus coccogenis	Warpaint shiner
Lythrurus lirus	Mountain shiner
Macrhybopsis aestivalis	Speckled chub
Macrhybopsis gelida	Sturgeon chub
Macrhybopsis meeki	Sicklefin chub
Moxostoma carinatum	River redhorse
Moxostoma duquesnei	Black redhorse
Moxostoma lacerum	Harelip sucker
Moxostoma valenciennesi	Greater redhorse
Nocomis effusus	Redtail chub
Nocomis leptocephalus	Bluehead chub
Notropis albizonatus	Palezone shiner
Notropis anogenus	Pugnose shiner
Notropis ariommus	Popeye shiner
Notropis boops	Bigeye shiner
Notropis chalybaeus	Ironcolor shiner
Notropis heterodon	Blackchin shiner
Notropis heterolepis	Blacknose shiner
Notropis leuciodus	Tennessee shiner
Notropis photogenis	Silver shiner
Notropis procne	Swallowtail shiner
Notropis rubellus	Rosyface shiner
Notropis rubricroceus	Saffron shiner
Notropis sp1	Sawfin shiner
Notropis telescopus	Telescope shiner
Notropis texanus	Weed shiner
Notropis volucellus	Mimic shiner
Noturus elegans	Elegant madtom
Noturus eleutherus	Mountain madtom
Noturus exilis	Slender madtom
Noturus flavus	Stonecat
Noturus hildebrandi	Least madtom
Noturus miurus	Brindled madtom
Noturus phaeus	Brown madtom
Noturus stigmosus	Northern madtom
Noturus trautmani	Scioto madtom

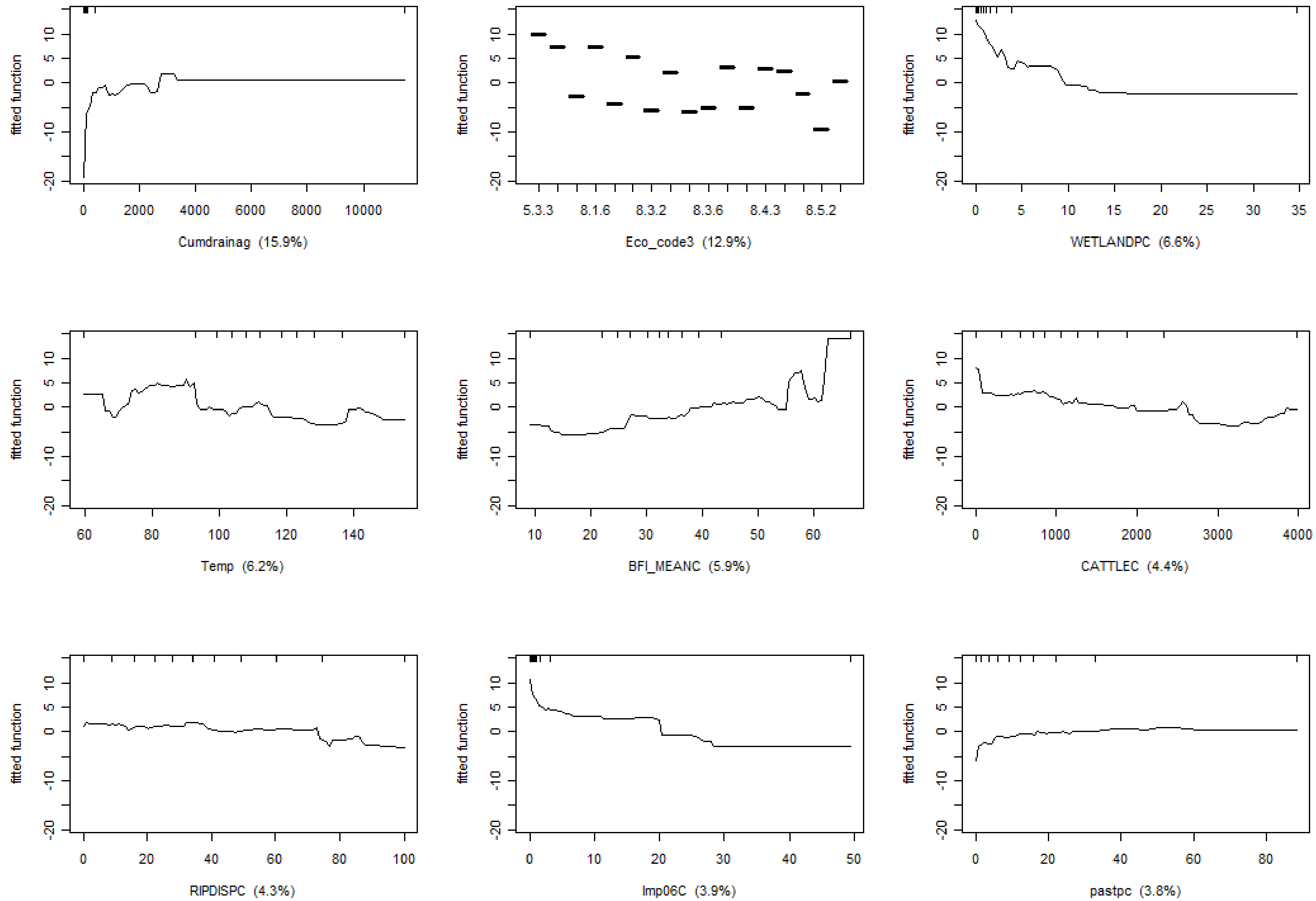
Opsopoeodus emiliae	Pugnose minnow
Percina burtoni	Blotchside logperch
Percina copelandi	Channel darter
Percina evides	Gilt darter
Percina macrocephala	Longhead darter
Percina oxyrhynchus	Sharpnose darter
Percina phoxocephala	Slenderhead darter
Percina squamata	Olive darter
Percina stictogaster	Frecklebelly darter
Phenacobius uranops	Stargazing minnow
Phoxinus cumberlandensis	Blackside dace
Phoxinus erythrogaster	Southern redbelly dace
Phoxinus saylori	Laurel dace
Phoxinus tennesseensis	Tennessee dace
Polyodon spathula	Paddlefish
Thoburnia atripinnis	Blackfin sucker

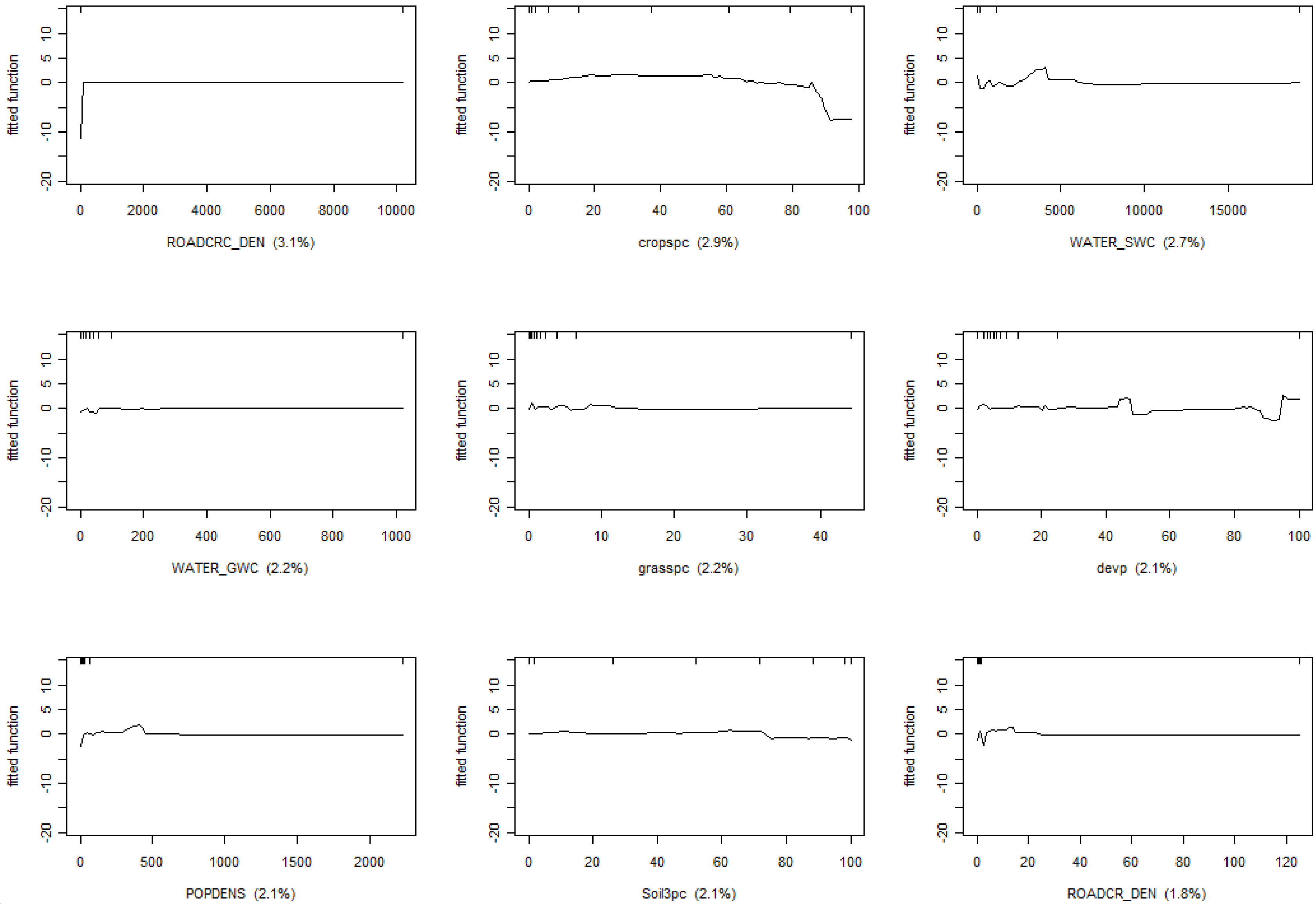
Table 25: Intolerant mussel species list

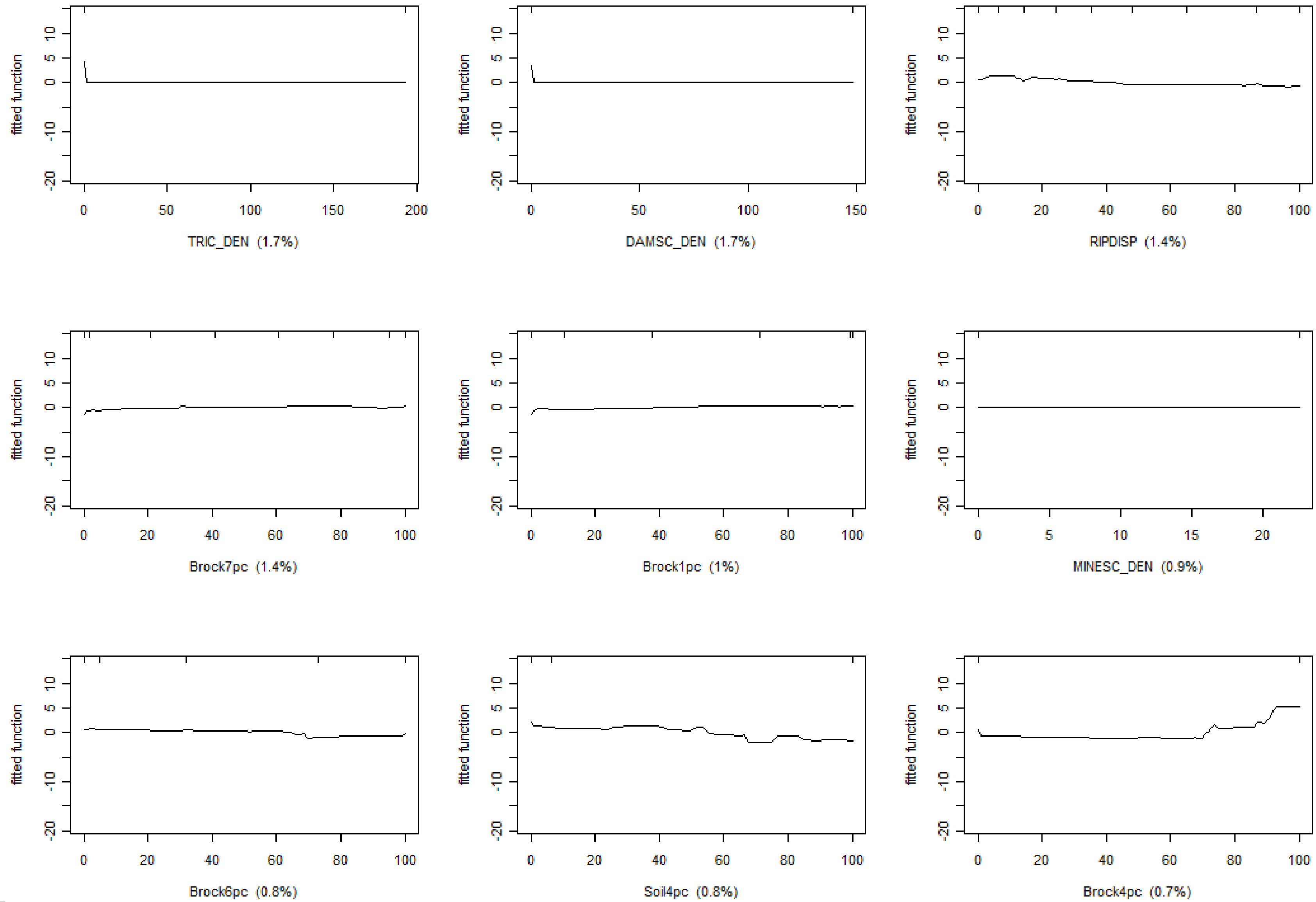
Scientific name	Common name
Actinonaias pectorosa	Pheasantshell
Alasmidonta heterodon	Dwarf wedgemussel
Alasmidonta triangulata	Southern elktoe
Alasmidonta atropurpurea	Cumberland elktoe
Alasmidonta varicosa	Brook floater
Anodontoides denigratus	Cumberland papershell
Cumberlandia monodonta	Spectaclecase
Cyclonaias tuberculata	Purple wartyback
Cyprogenia stegaria	Fanshell
Elliptio crassidens	Elephant-ear
Epioblasma spp	—
Fusconaia cor	Shiny pigtoe
Fusconaia cuneolus	Finerayed pigtoe
Fusconaia cordatum	Ohio pigtoe
Lampsilis abrupta	Pink mucket
Lampsilis virescens	Alabama lampmussel
Lasmigona holstonia	Tennessee heelsplitter
Lemiox rimosus	Birdwing pearlymussel
Medionidus acutissimus	Alabama moccasinshell
Medionidus conradicus	Cumberland moccasinshell
Obovaria retusa	Ring pink
Obovaria subrotunda	Round hickorynut
Pegias fabula	Littlewing pearlymussel
Plethobasus cicatricosus	White wartyback
Plethobasus cooperianus	Orangefoot pimpleback
Plethobasus cyphyus	Sheepnose
Pleuronaia barnesiana	Tennessee pigtoe
Pleuronaia dolabelloides	Slabside pearlymussel
Potamilus capax	Fat pocketbook
Ptychobranchus spp	—
Ptychobranchus fasciolaris	Kidneyshell
Quadrula cylindrica strigillata	Rough rabbitsfoot
Quadrula intermedia	Cumberland monkeyface
Quadrula sparsa	Appalachian monkeyface
Toxolasma pullus	Savannah lilliput
Toxolasma lividus	Purple lilliput

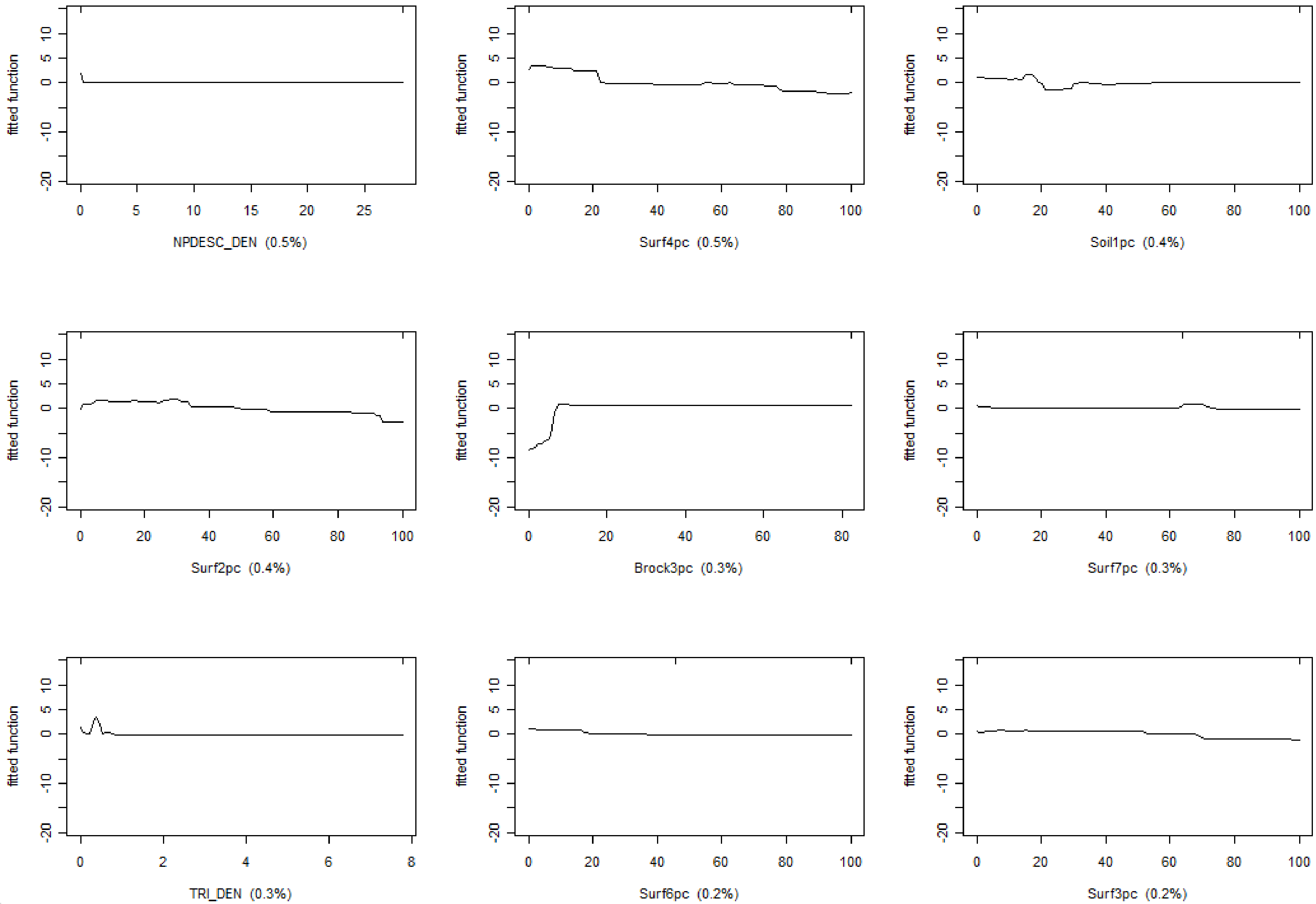
Appendix B: FUNCTIONAL RESPONSE PLOTS

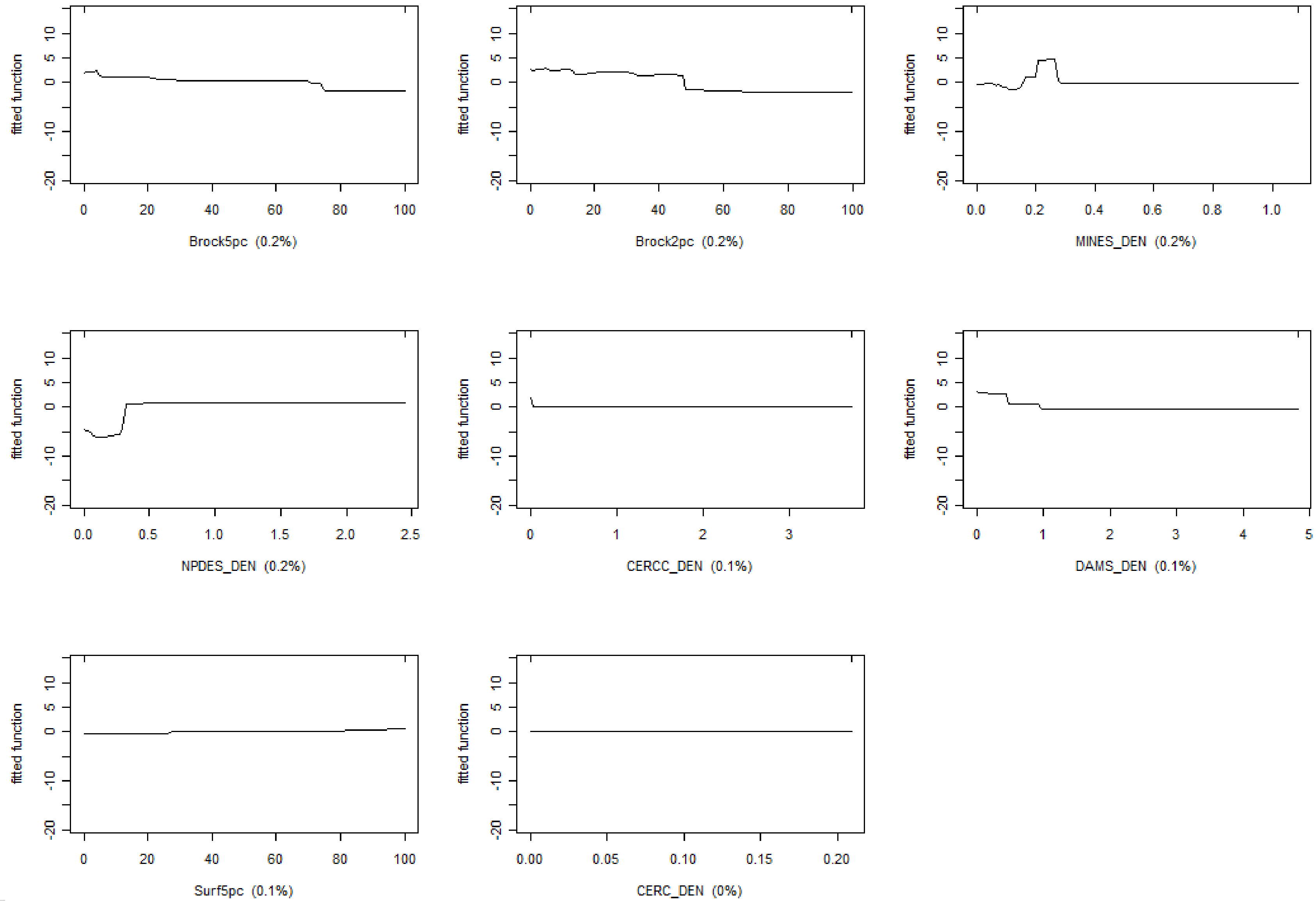
Small stream signature fish index



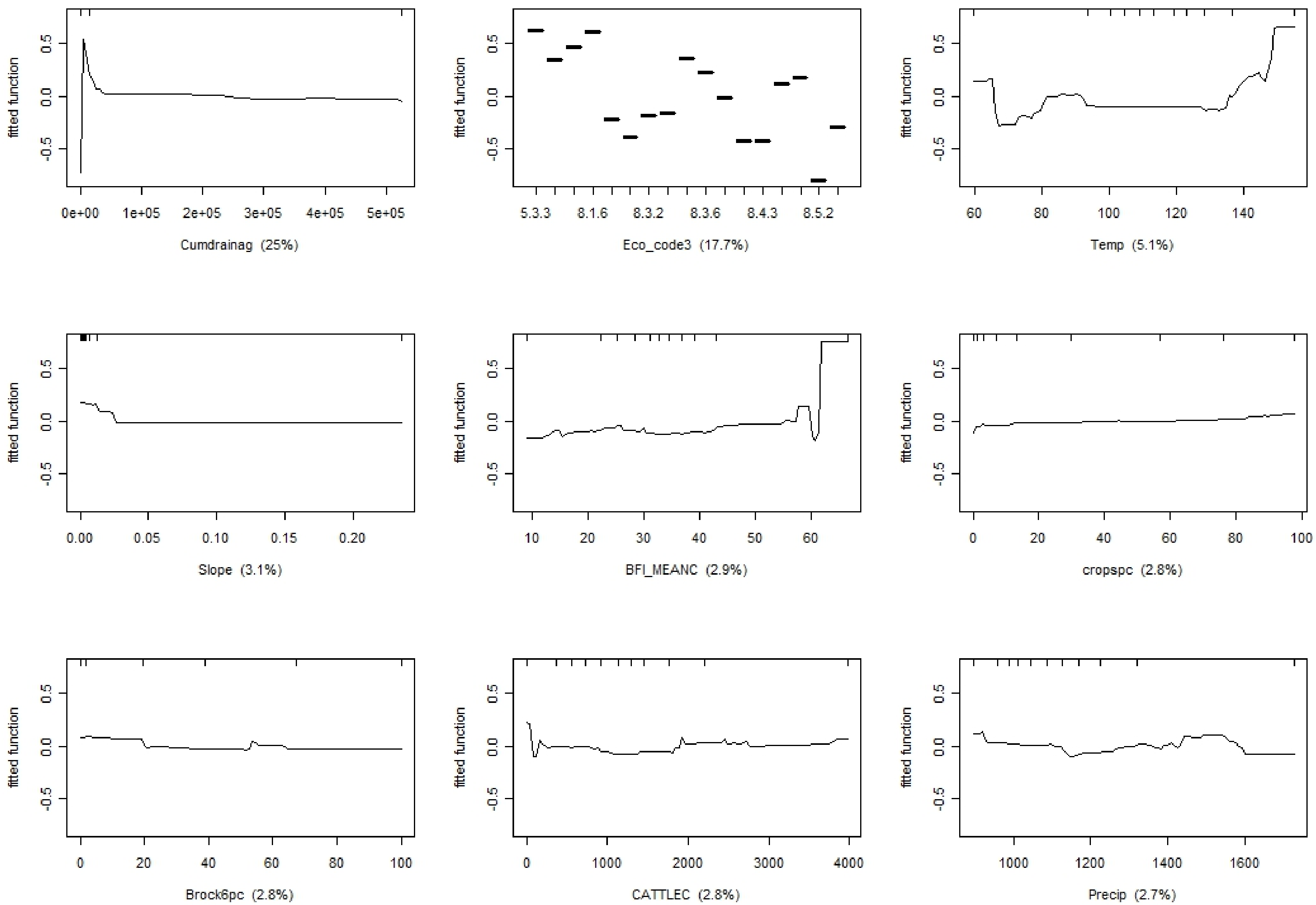

fitted function
ROADCRC_DEN (3.1%)
cropspc (2.9%)
WATER_SWC (2.7%)
WATER_GWC (2.2%)
grasspc (2.2%)
devp (2.1%)
POPDENS (2.1%)
Soil3pc (2.1%)
ROADCR_DEN (1.8%)

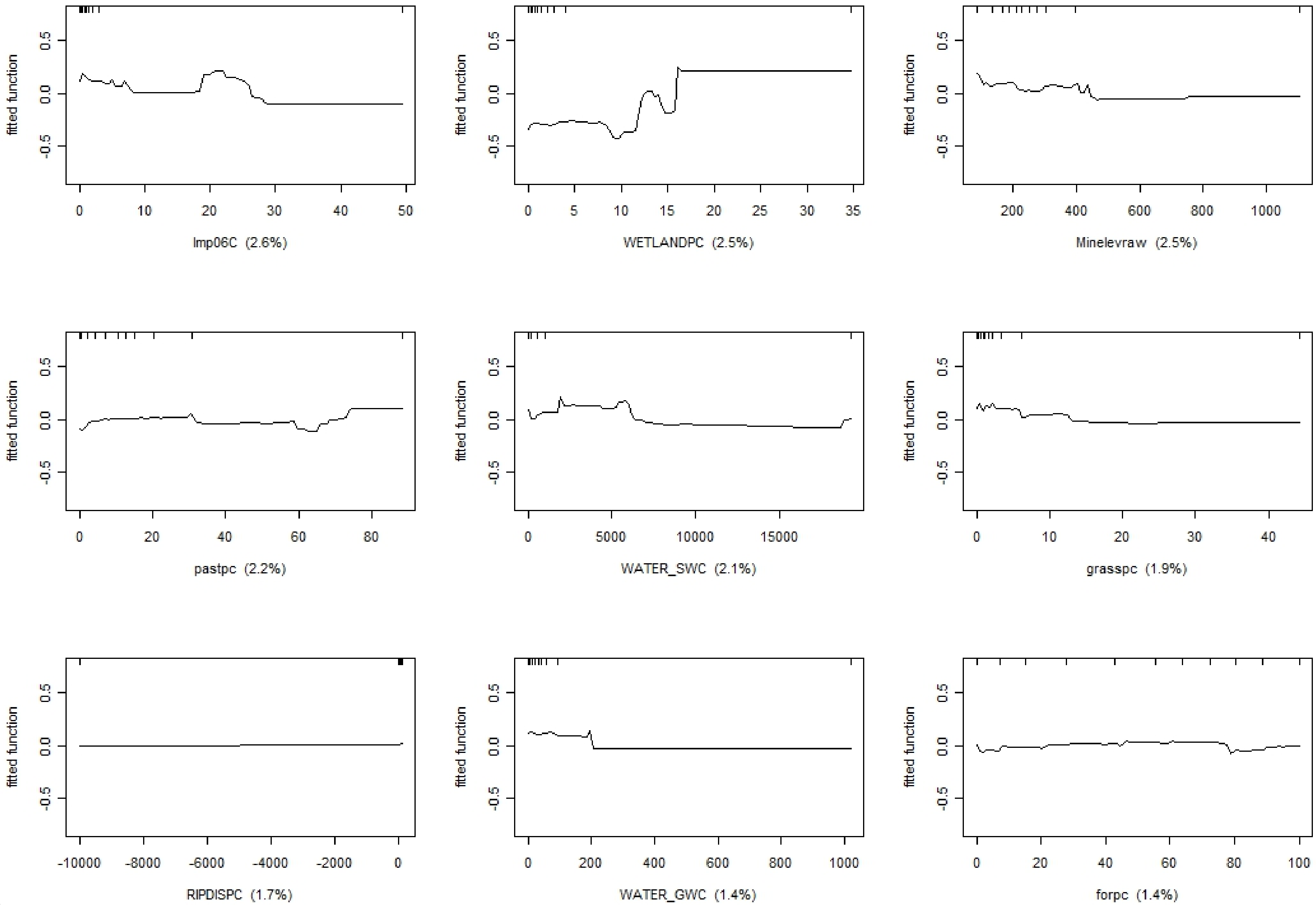

fitted function
TRIC_DEN (1.7%)
DAMSC_DEN (1.7%)
RIPDISP (1.4%)
Brock7pc (1.4%)
Brock1pc (1%)
MINESC_DEN (0.9%)
Brock6pc (0.8%)
Soil4pc (0.8%)
Brock4pc (0.7%)

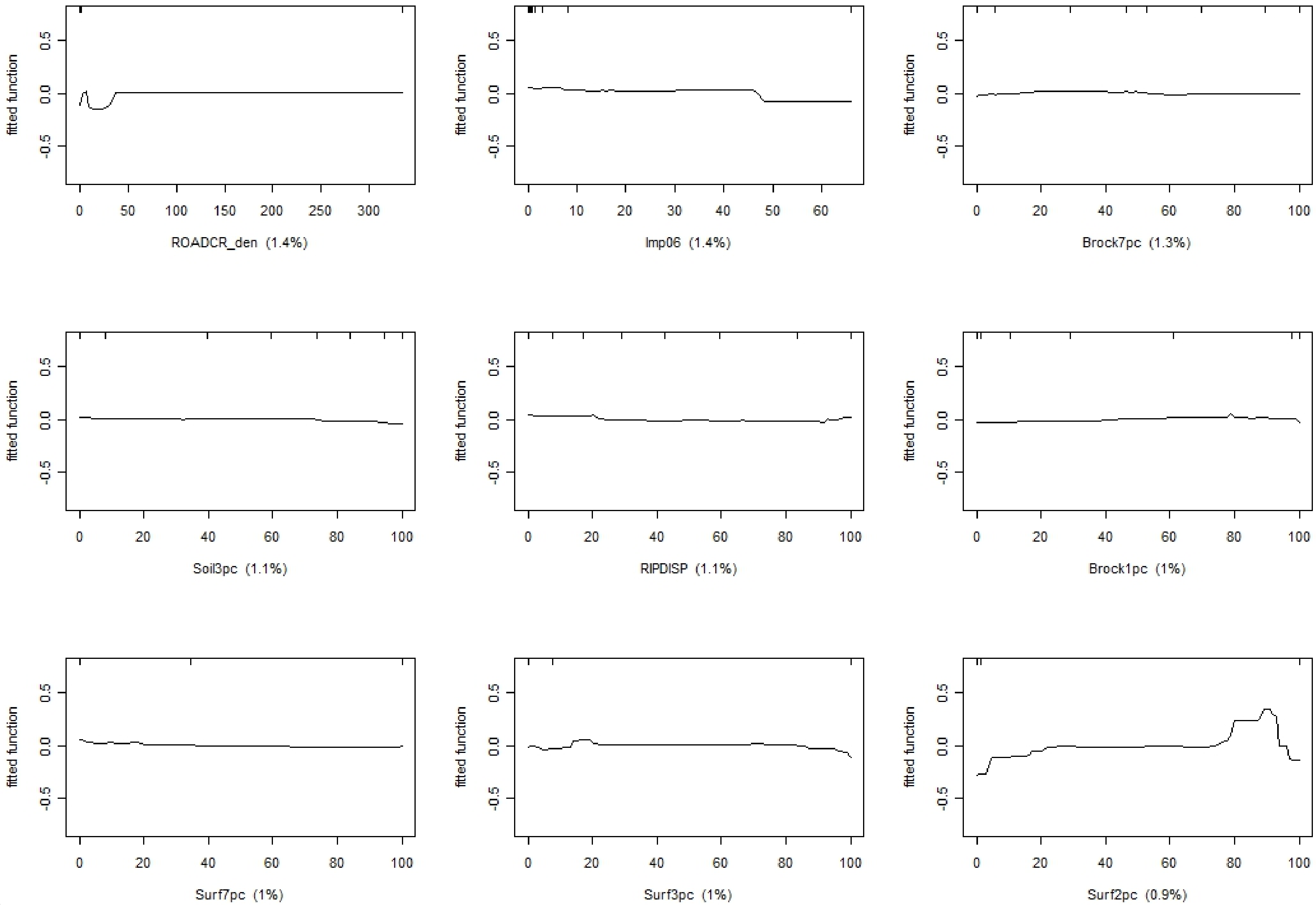

fitted function
NPDESC_DEN (0.5%)
Surf4pc (0.5%)
Soil1pc (0.4%)
Surf2pc (0.4%)
Brock3pc (0.3%)
Surf7pc (0.3%)
TRI_DEN (0.3%)
Surf6pc (0.2%)
Surf3pc (0.2%)

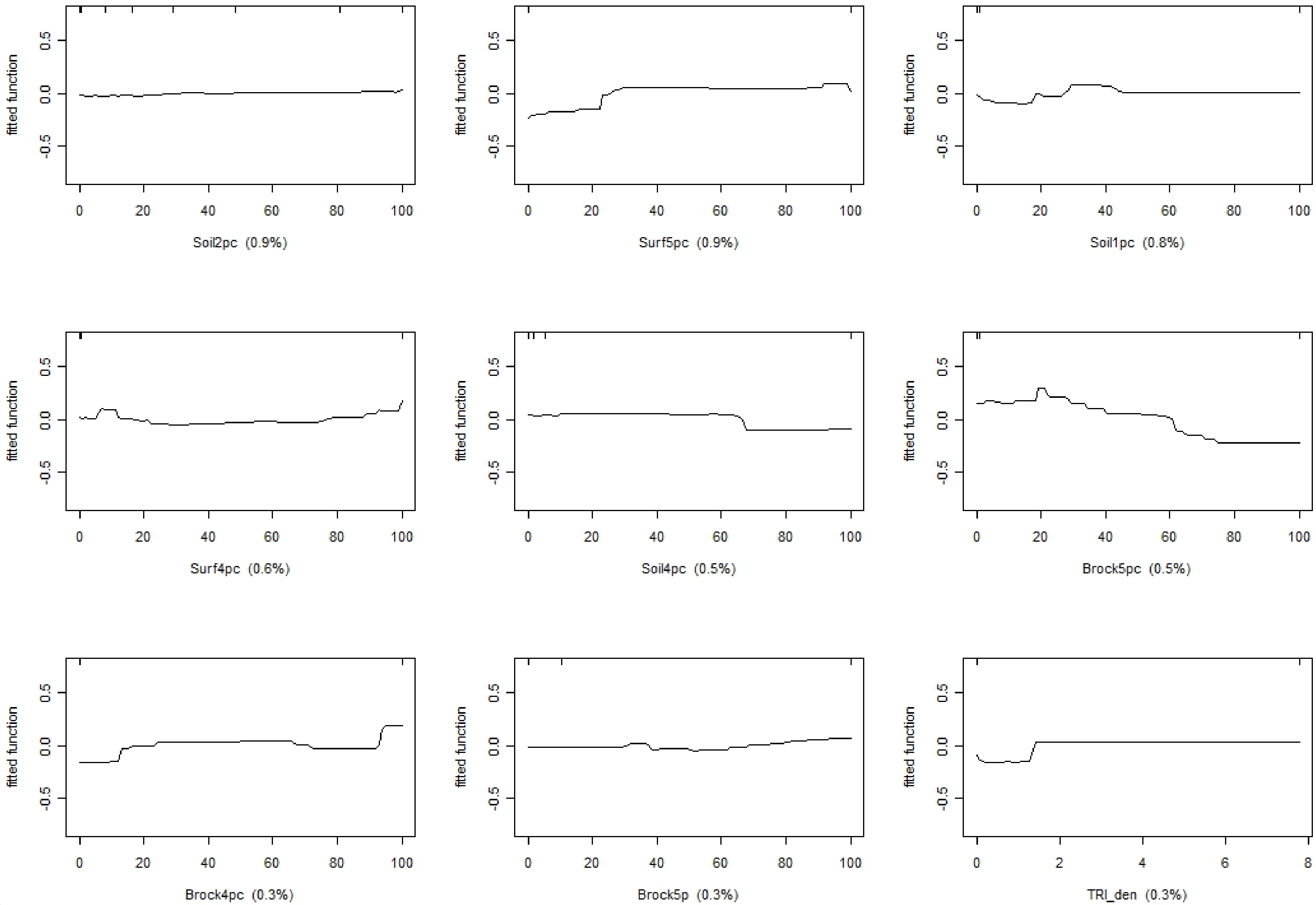

fitted function
-20
-10
0
5
10
0
20
40
60
80
100
Brock5pc (0.2%)
Brock2pc (0.2%)
0.0
0.2
0.4
0.6
0.8
1.0
MINES_DEN (0.2%)
0.5
1.0
1.5
2.0
2.5
NPDES_DEN (0.2%)
1
2
3
CERCC_DEN (0.1%)
4
5
DAMS_DEN (0.1%)
Surf5pc (0.1%)
0.00
0.05
0.10
0.15
0.20
CERC_DEN (0%)

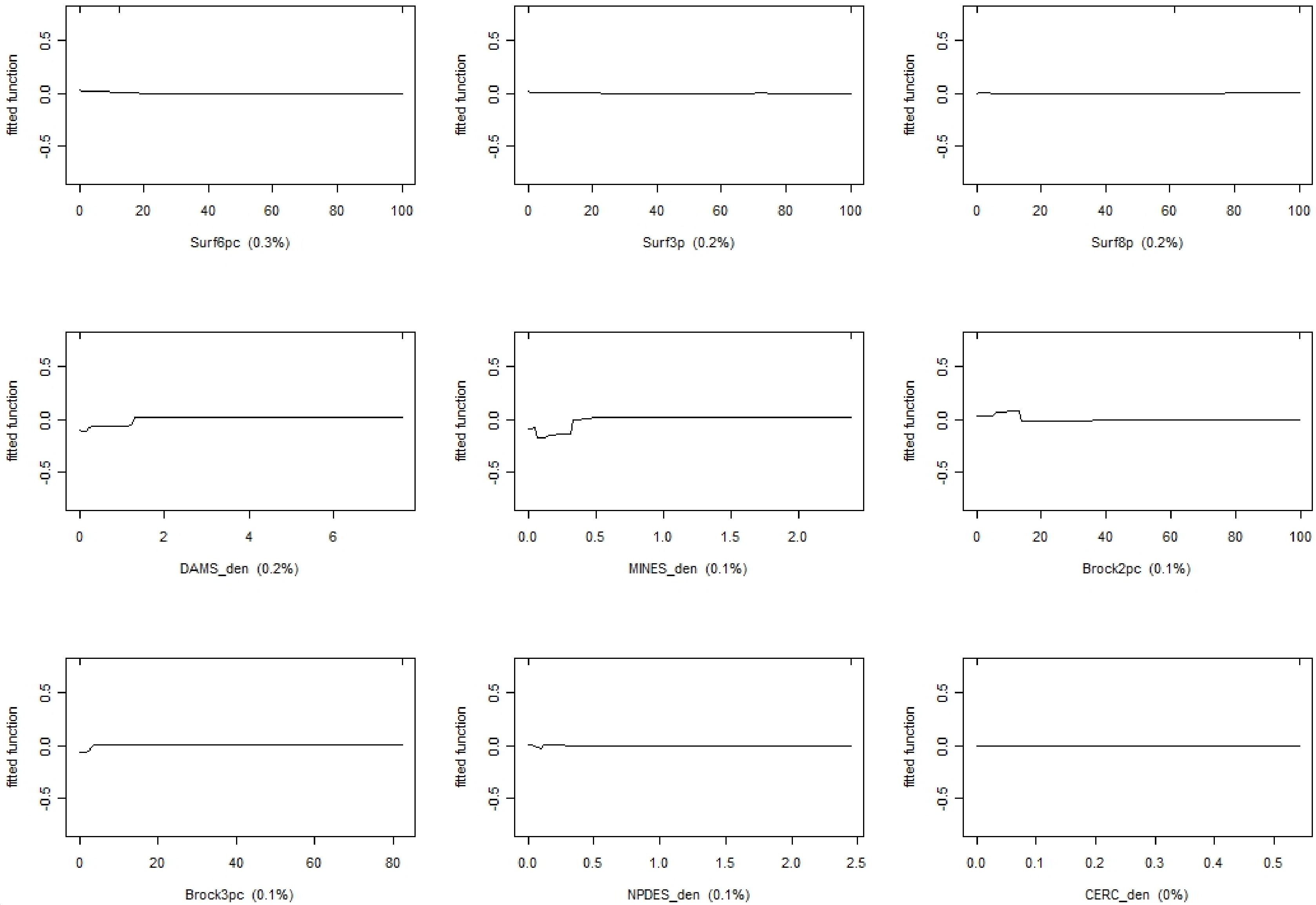
Modified index of centers of diversity



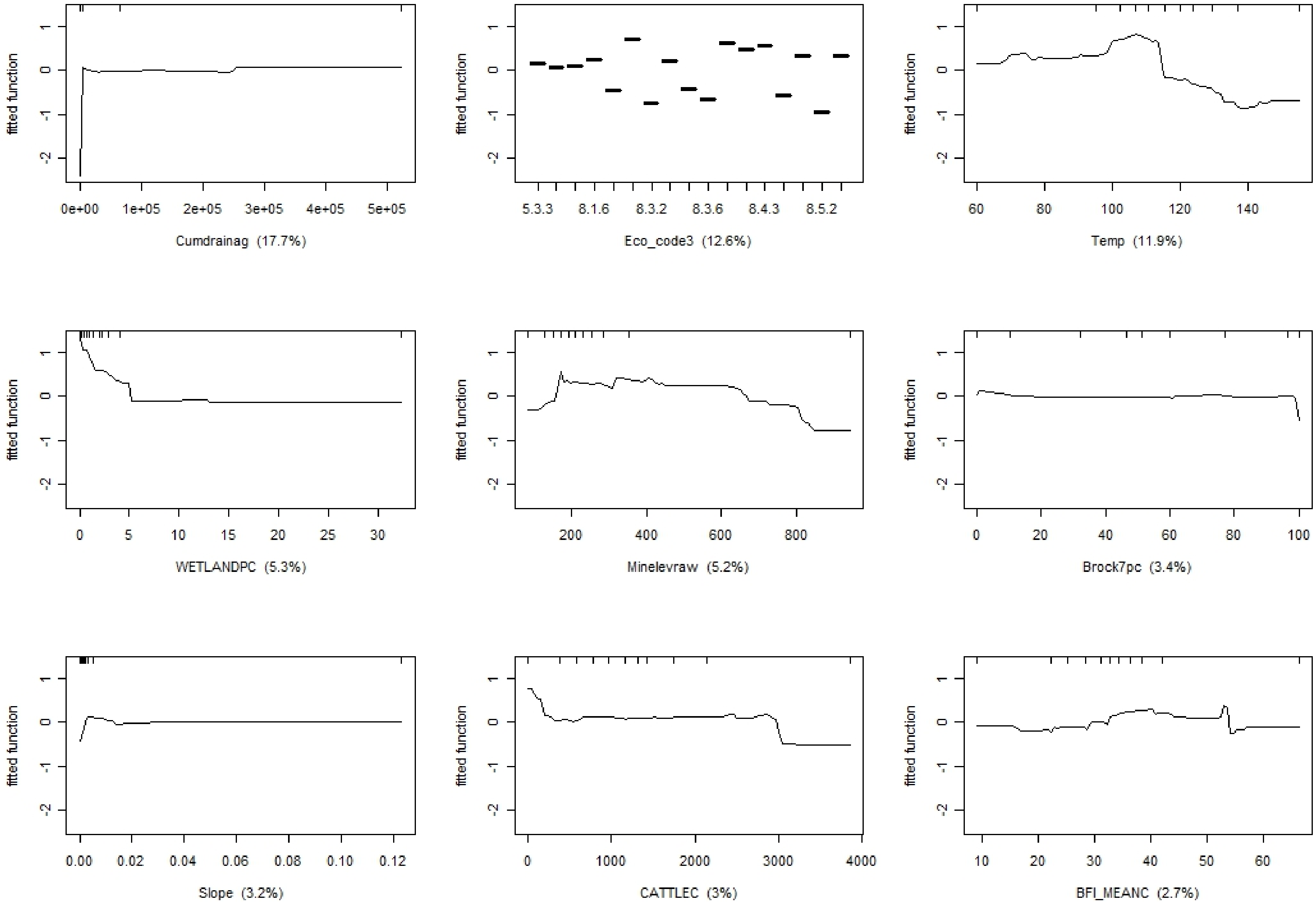

fitted function
0.5
0.0
-0.5
0 10 20 30 40 50
Imp06C (2.6%)
0 5 10 15 20 25 30 35
WETLANDPC (2.5%)
200 400 600 800 1000
Minelevraw (2.5%)
0 20 40 60 80
pastpc (2.2%)
0 5000 10000 15000
WATER_SWC (2.1%)
0 10 20 30 40
grasspc (1.9%)
-10000 -8000 -6000 -4000 -2000 0
RIPDISPC (1.7%)
0 200 400 600 800 1000
WATER_GWC (1.4%)
0 20 40 60 80 100
forpc (1.4%)

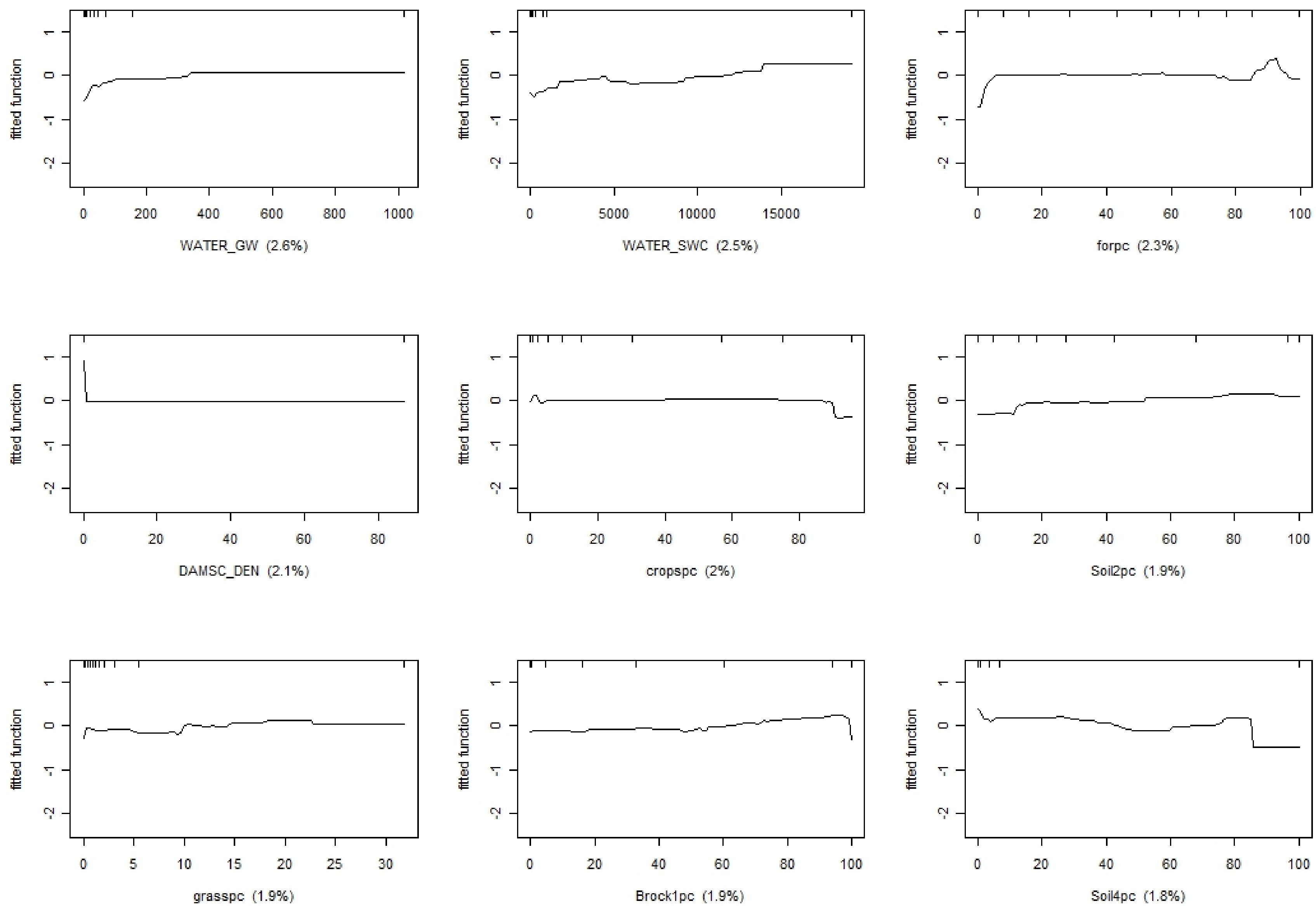

fitted function
ROADCR_den (1.4%)
Imp06 (1.4%)
Brock7pc (1.3%)
Soil3pc (1.1%)
RIPDISP (1.1%)
Brock1pc (1%)
Surf7pc (1%)
Surf3pc (1%)
Surf2pc (0.9%)

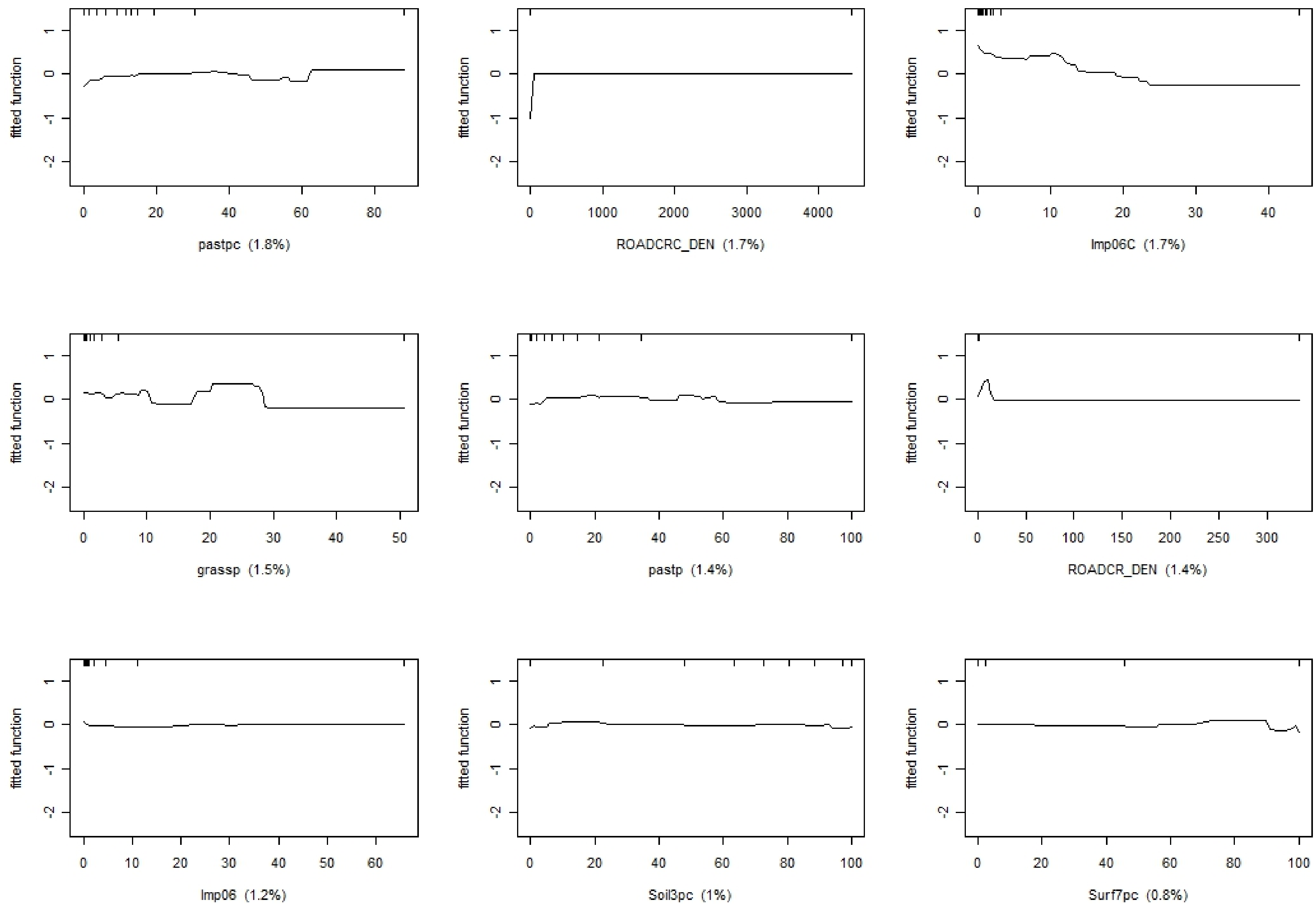

fitted function
0.5
0.0
-0.5
0
20
40
60
80
100
Soil2pc (0.9%)
Surf5pc (0.9%)
Soil1pc (0.8%)
Surf4pc (0.6%)
Soil4pc (0.5%)
Brock5pc (0.5%)
Brock4pc (0.3%)
Brock5p (0.3%)
2
4
6
8
TRI_den (0.3%)

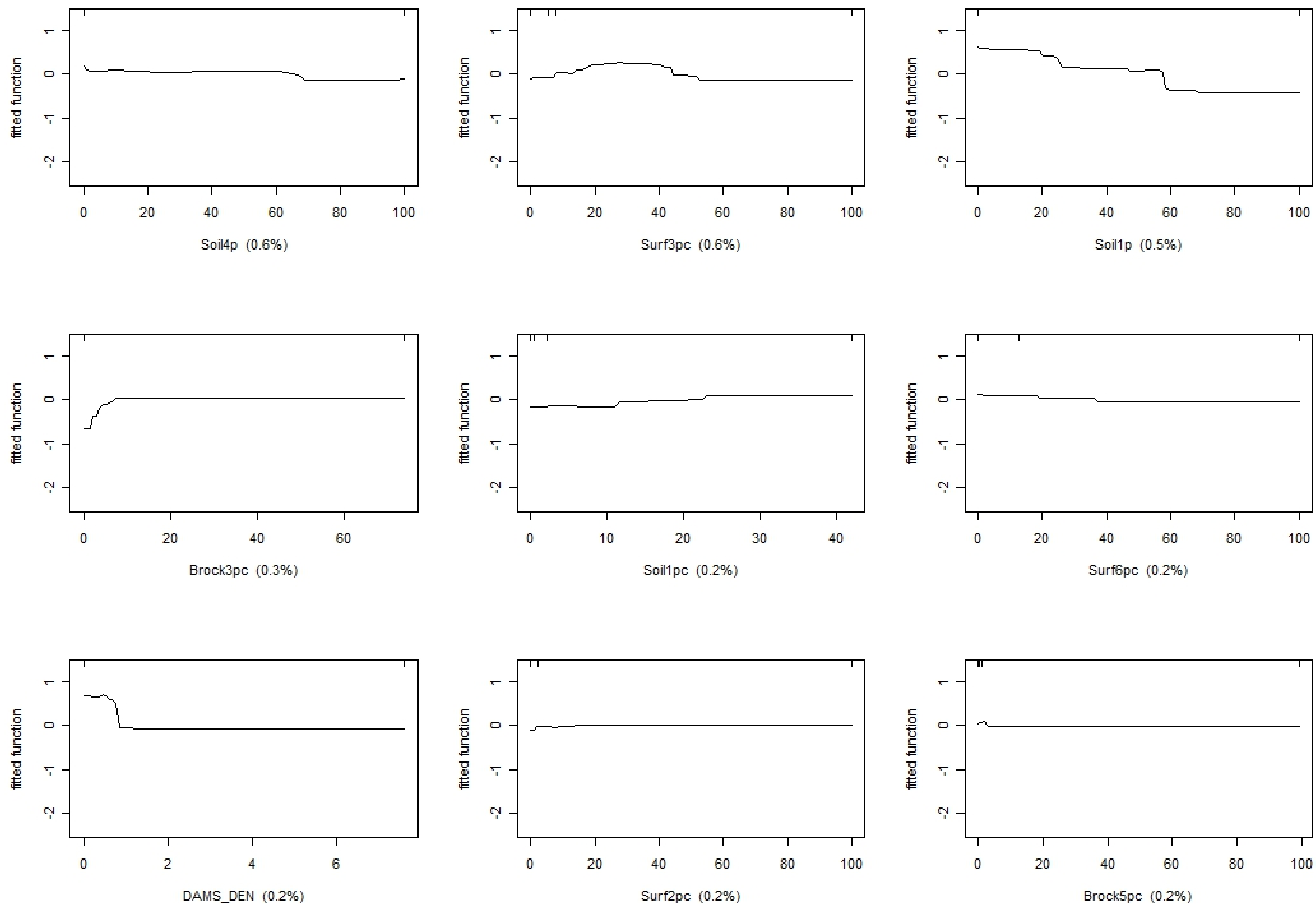

fitted function
Surf6pc (0.3%)
Surf3p (0.2%)
Surf8p (0.2%)
DAMS_den (0.2%)
MINES_den (0.1%)
Brock2pc (0.1%)
Brock3pc (0.1%)
NPDES_den (0.1%)
CERC_den (0%)

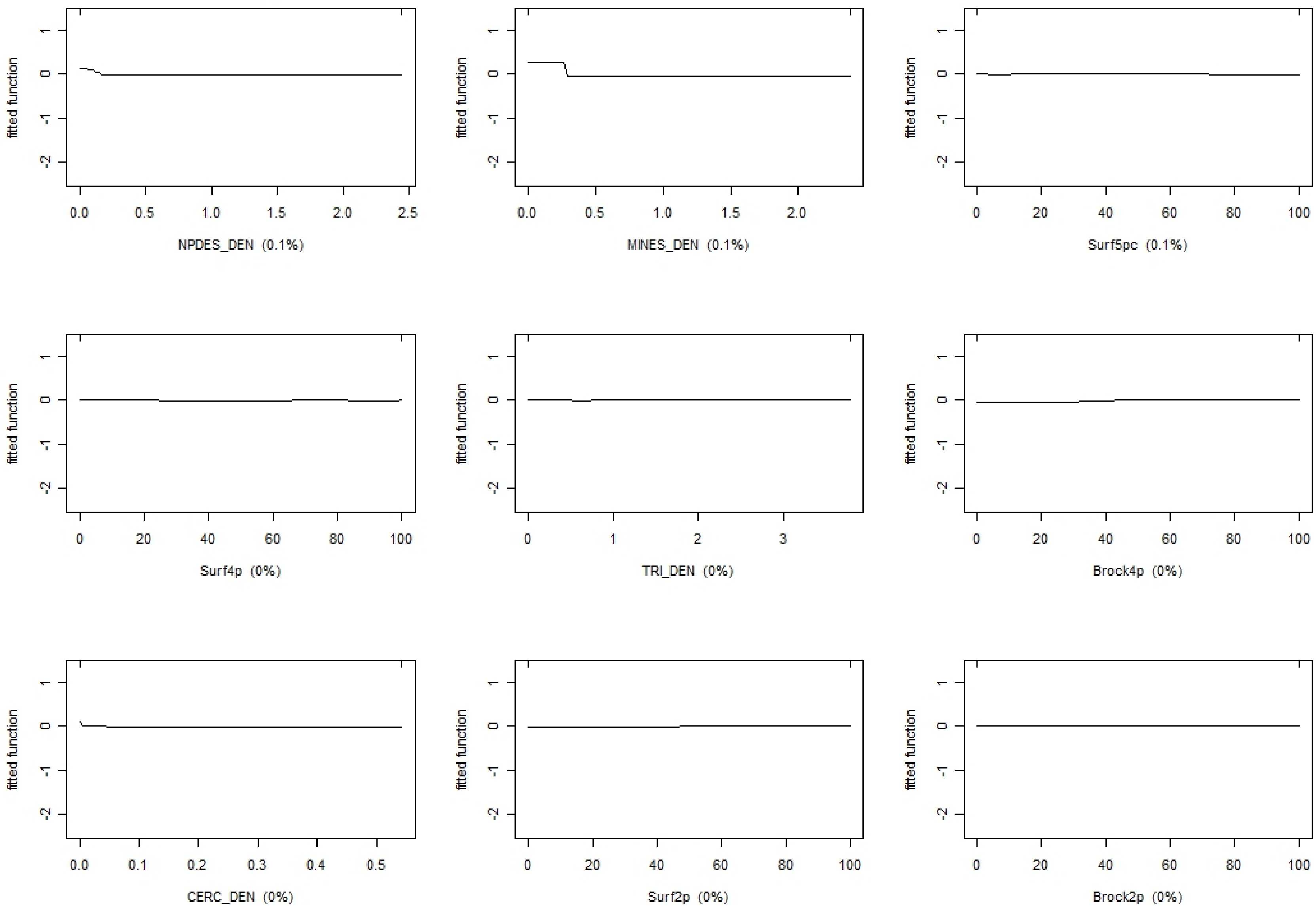
Smallmouth Bass

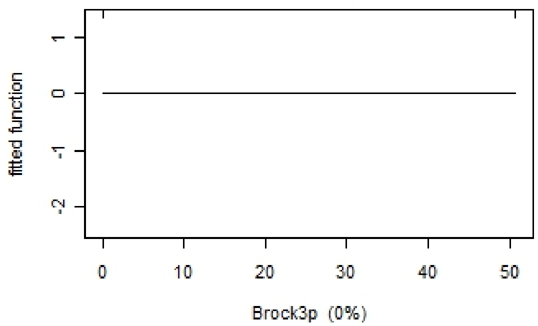



fitted function
1
0
-1
-2
0
200
400
600
800
1000
WATER_GW (2.6%)
0
5000
10000
15000
WATER_SWC (2.5%)
0
20
40
60
80
100
forpc (2.3%)
DAMSC_DEN (2.1%)
cropspc (2%)
Soil2pc (1.9%)
5
10
15
20
25
30
grasspc (1.9%)
Brock1pc (1.9%)
Soil4pc (1.8%)

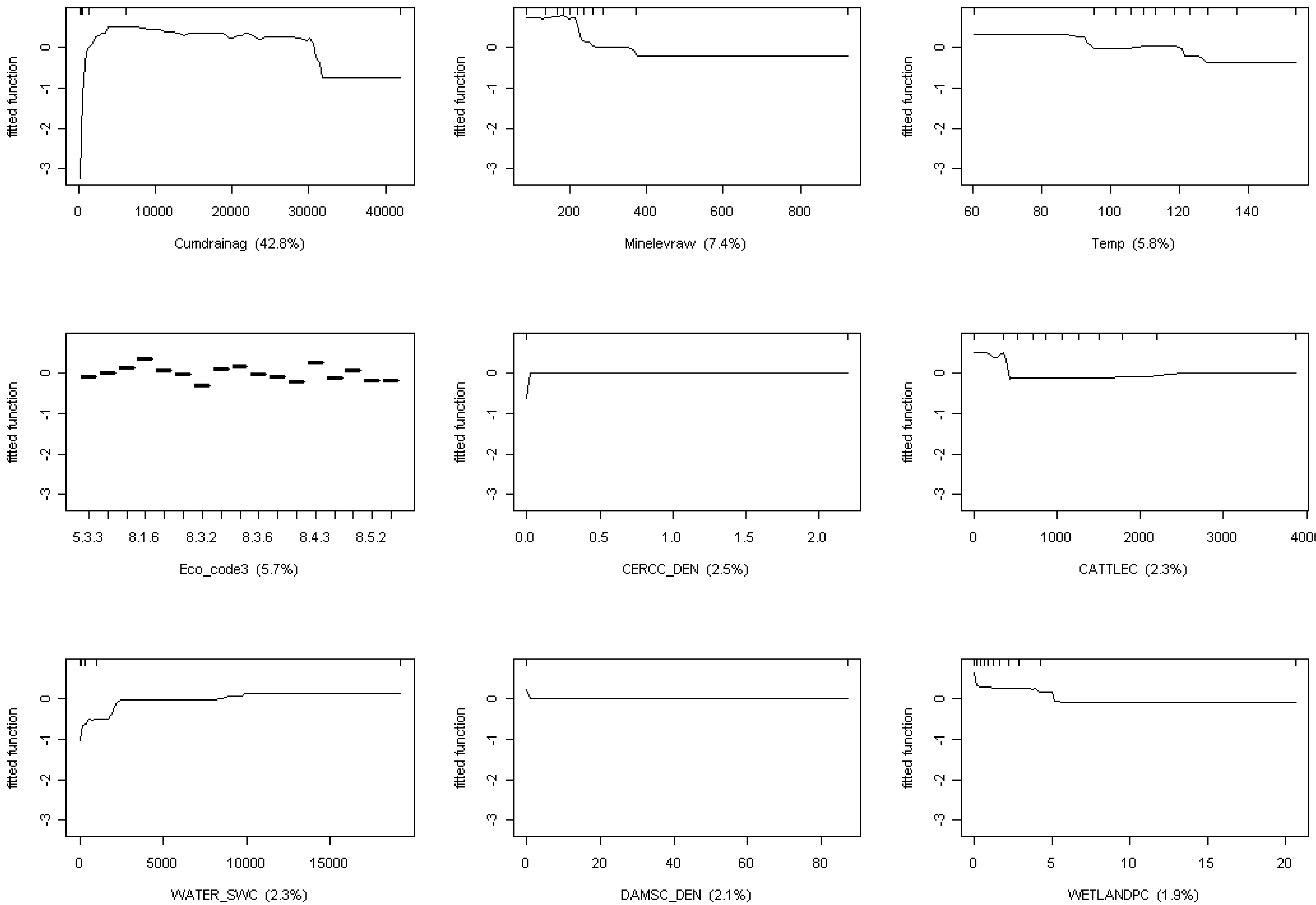

fitted function
pastpc (1.8%)
ROADCRC_DEN (1.7%)
Imp06C (1.7%)
grassp (1.5%)
pastp (1.4%)
ROADCR_DEN (1.4%)
Imp06 (1.2%)
Soil3pc (1%)
Surf7pc (0.8%)


fitted function
Soil4p (0.6%)
Surf3pc (0.6%)
Soil1p (0.5%)
Brock3pc (0.3%)
Soil1pc (0.2%)
Surf6pc (0.2%)
DAMS_DEN (0.2%)
Surf2pc (0.2%)
Brock5pc (0.2%)

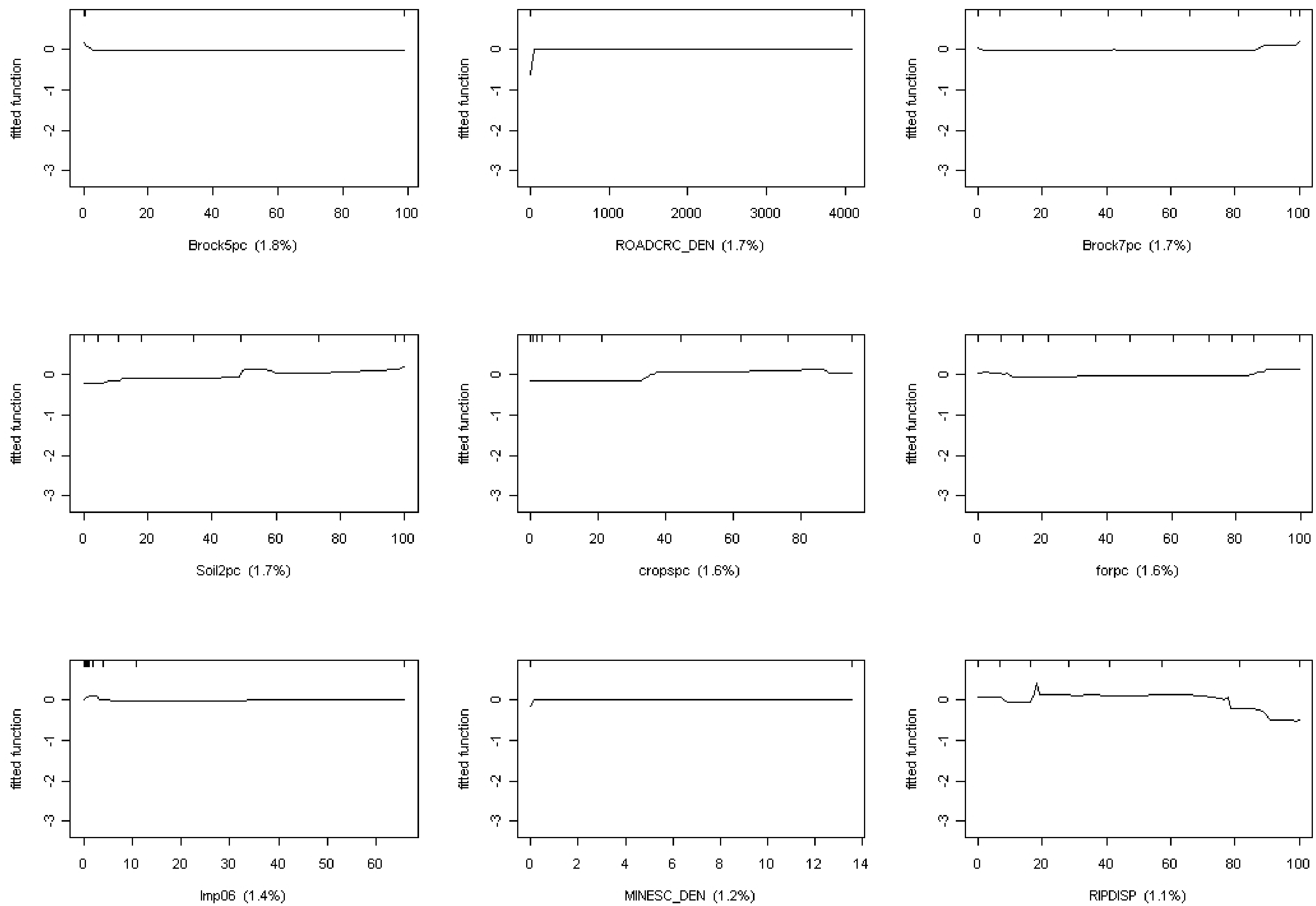

fitted function
NPDES_DEN (0.1%)
MINES_DEN (0.1%)
Surf5pc (0.1%)
Surf4p (0%)
TRI_DEN (0%)
Brock4p (0%)
CERC_DEN (0%)
Surf2p (0%)
Brock2p (0%)

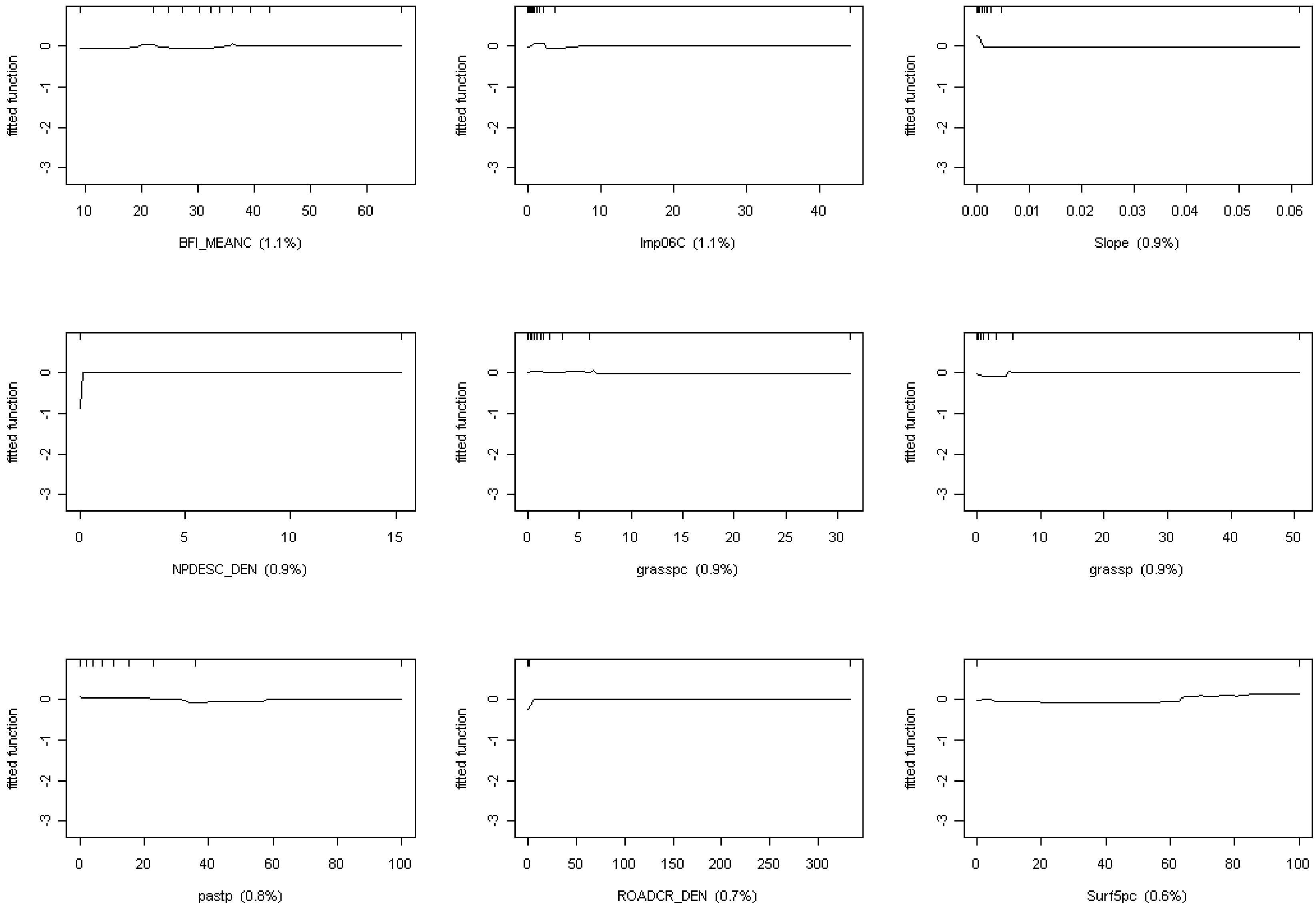

fitted function
1
0
-1
-2
0
10
20
30
40
50
Brock3p (0%)

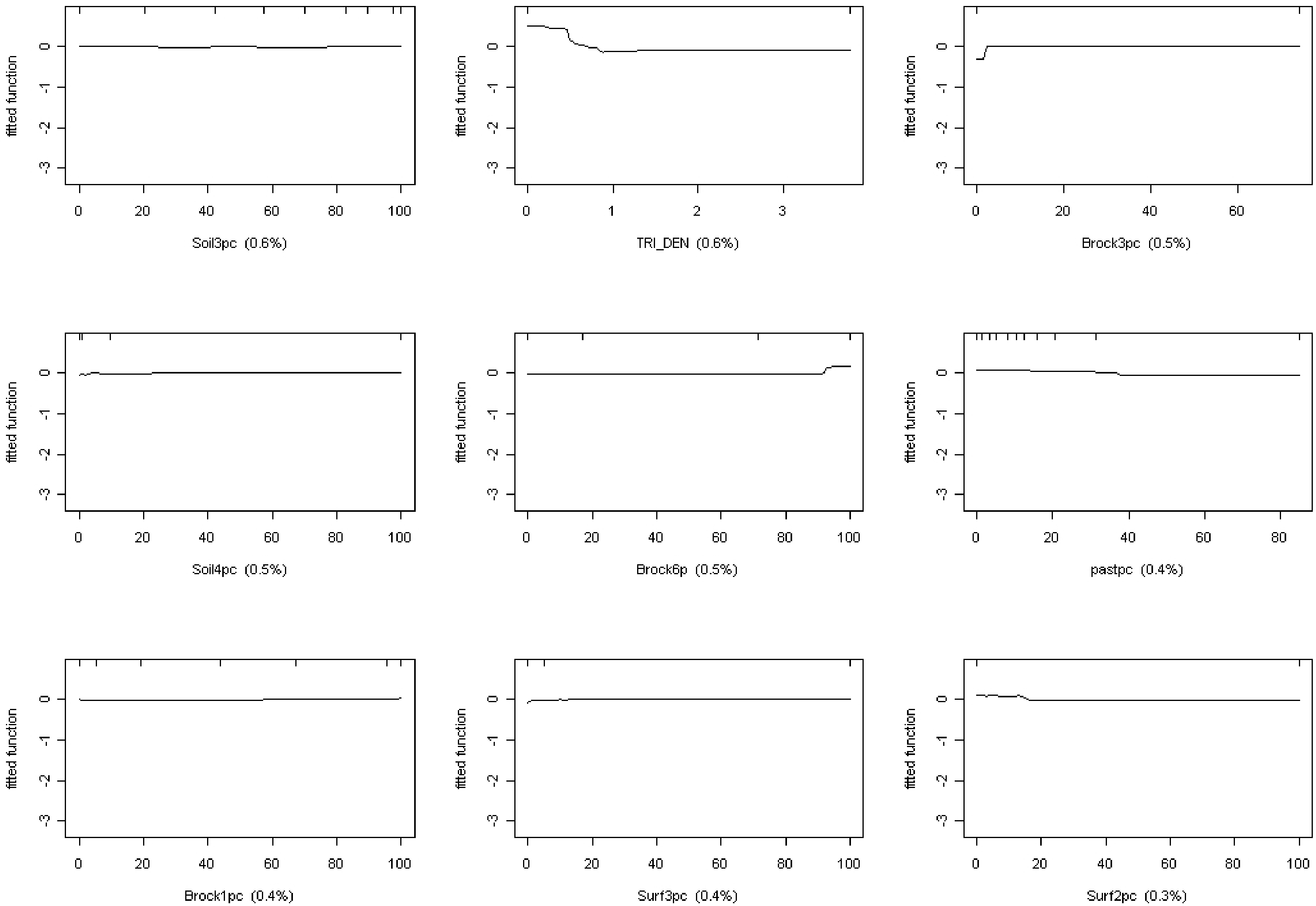
Redhorse

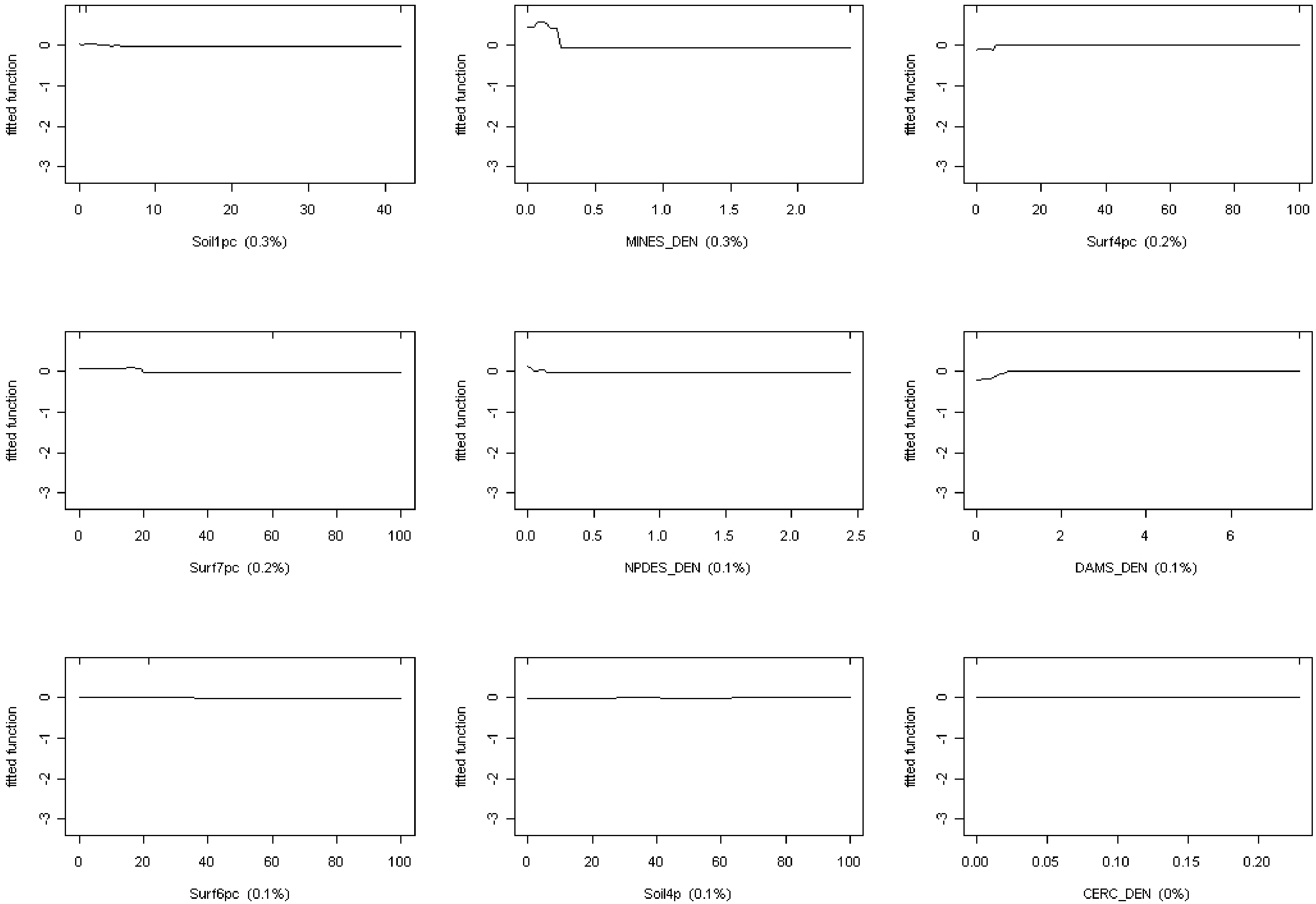


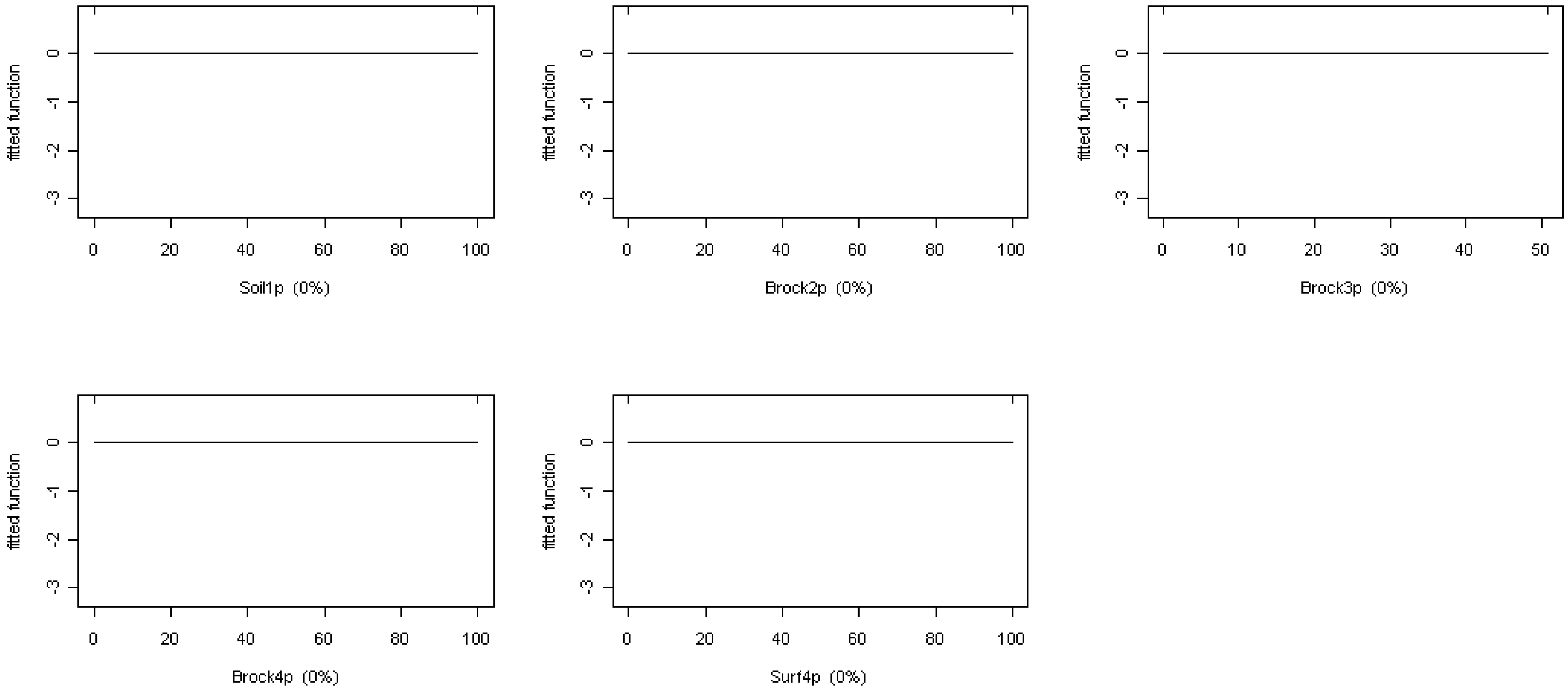
fitted function
0
-1
-2
-3
0 10000 20000 30000 40000
Cumdrainag (42.8%)
200 400 600 800
Minelevraw (7.4%)
60 80 100 120 140
Temp (5.8%)
5.3.3 8.1.6 8.3.2 8.3.6 8.4.3 8.5.2
Eco_code3 (5.7%)
0.0 0.5 1.0 1.5 2.0
CERCC_DEN (2.5%)
0 1000 2000 3000 4000
CATTLEC (2.3%)
0 5000 10000 15000
WATER_SWC (2.3%)
0 20 40 60 80
DAMSC_DEN (2.1%)
0 5 10 15 20
WETLANDPC (1.9%)


fitted function
Brock5pc (1.8%)
ROADCRC_DEN (1.7%)
Brock7pc (1.7%)
Soil2pc (1.7%)
cropspc (1.6%)
forpc (1.6%)
Imp06 (1.4%)
MINESC_DEN (1.2%)
RIPDISP (1.1%)

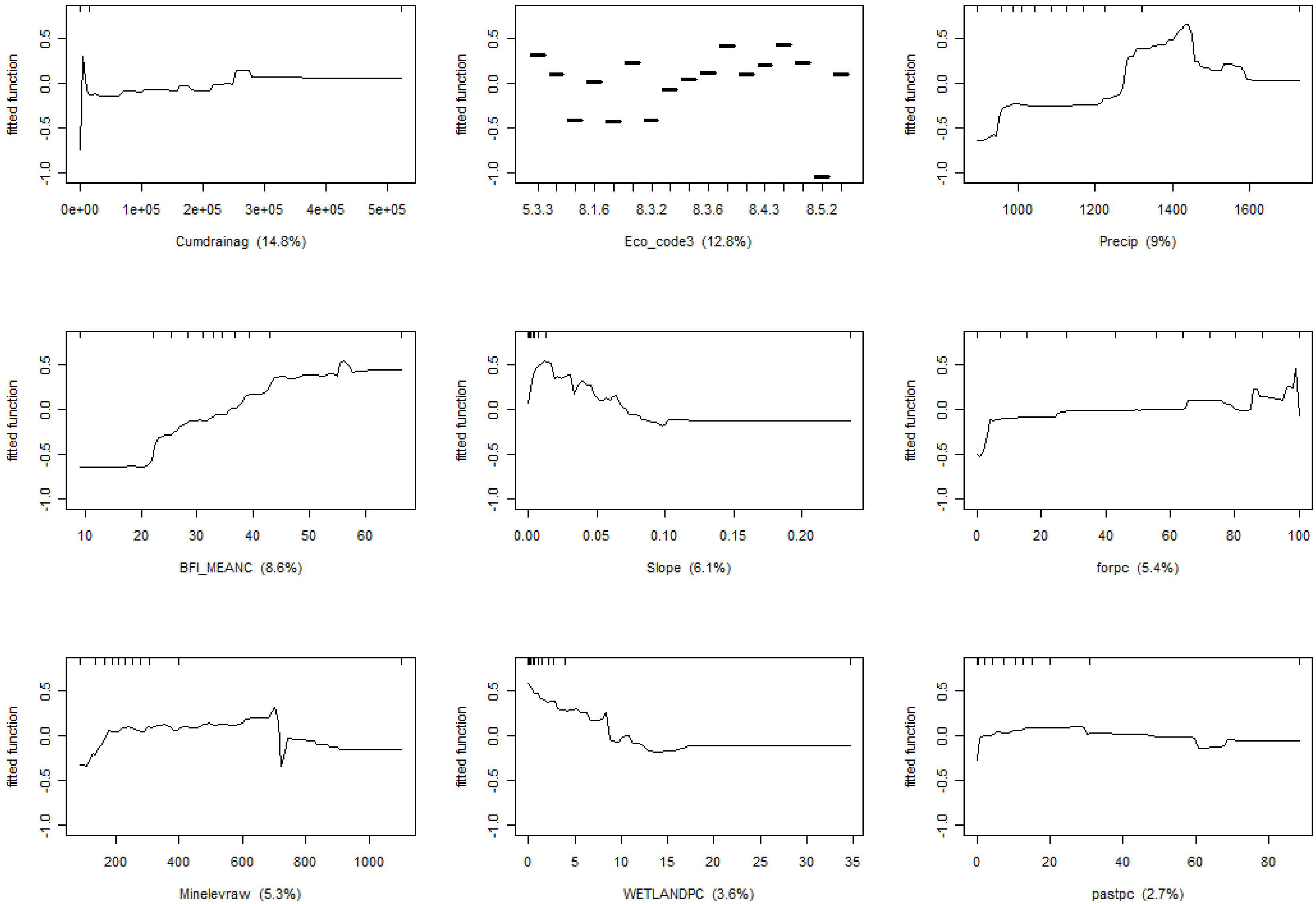

fitted function
0
-1
-2
-3
10 20 30 40 50 60
BFI_MEANC (1.1%)
0 10 20 30 40
Imp06C (1.1%)
0.00 0.01 0.02 0.03 0.04 0.05 0.06
Slope (0.9%)
0 5 10 15
NPDESC_DEN (0.9%)
0 5 10 15 20 25 30
grasspc (0.9%)
0 10 20 30 40 50
grassp (0.9%)
0 20 40 60 80 100
pastp (0.8%)
0 50 100 150 200 250 300
ROADCR_DEN (0.7%)
0 20 40 60 80 100
Surf5pc (0.6%)

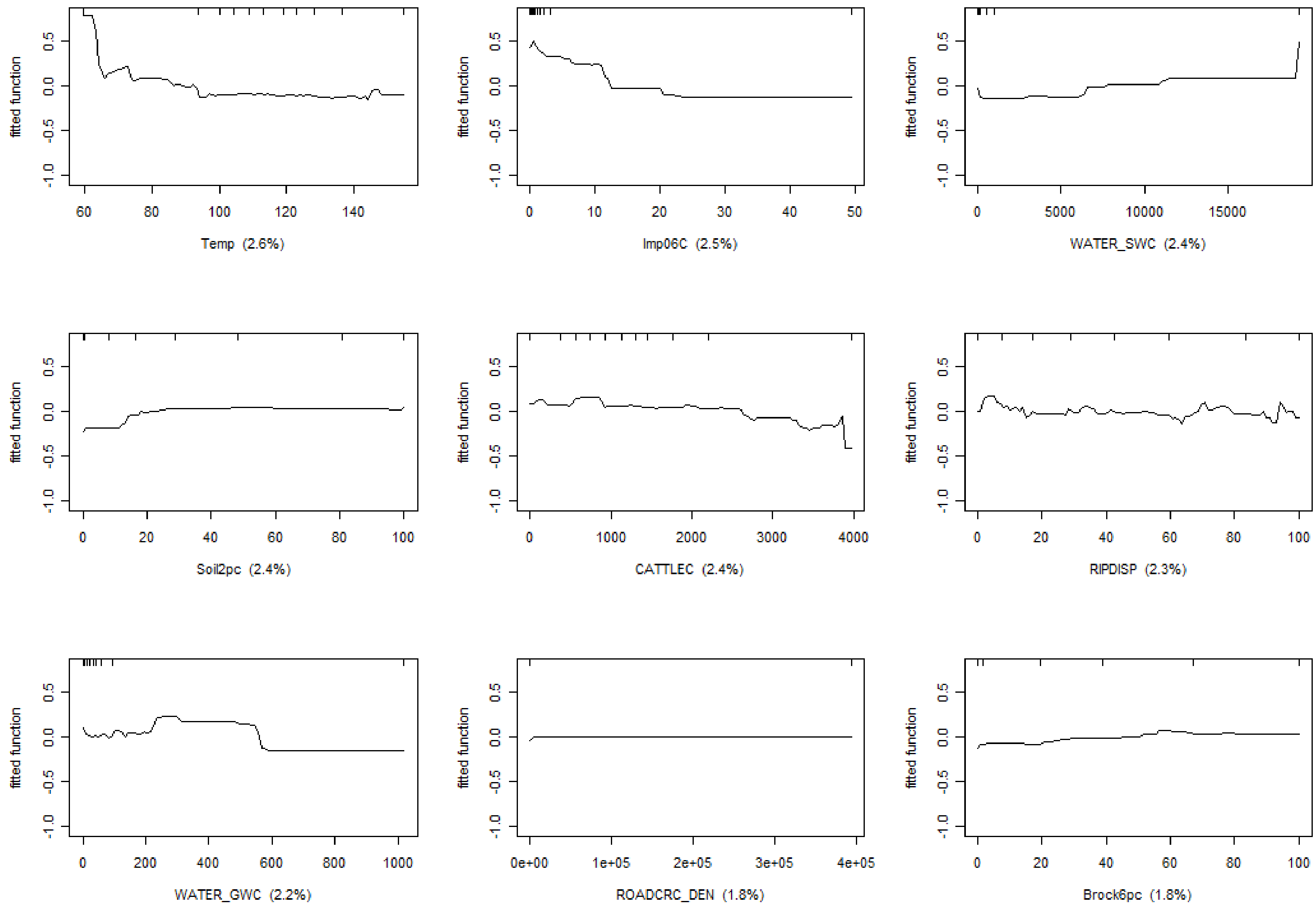

fitted function
Soil3pc (0.6%)
TRI_DEN (0.6%)
Brock3pc (0.5%)
Soil4pc (0.5%)
Brock6p (0.5%)
pastpc (0.4%)
Brock1pc (0.4%)
Surf3pc (0.4%)
Surf2pc (0.3%)

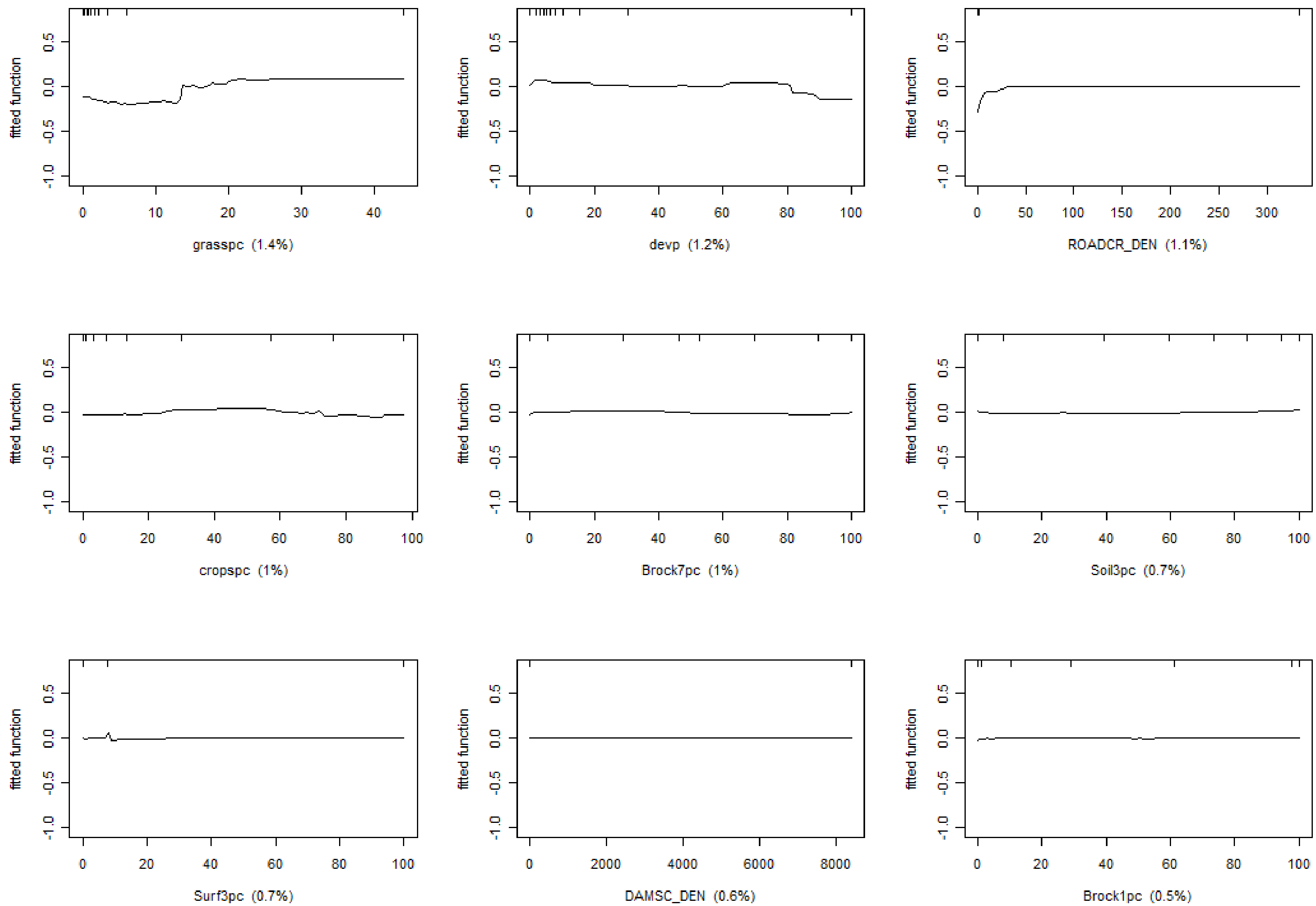

fitted function
Soil1pc (0.3%)
MINES_DEN (0.3%)
Surf4pc (0.2%)
Surf7pc (0.2%)
NPDES_DEN (0.1%)
DAMS_DEN (0.1%)
Surf6pc (0.1%)
Soil4p (0.1%)
CERC_DEN (0%)

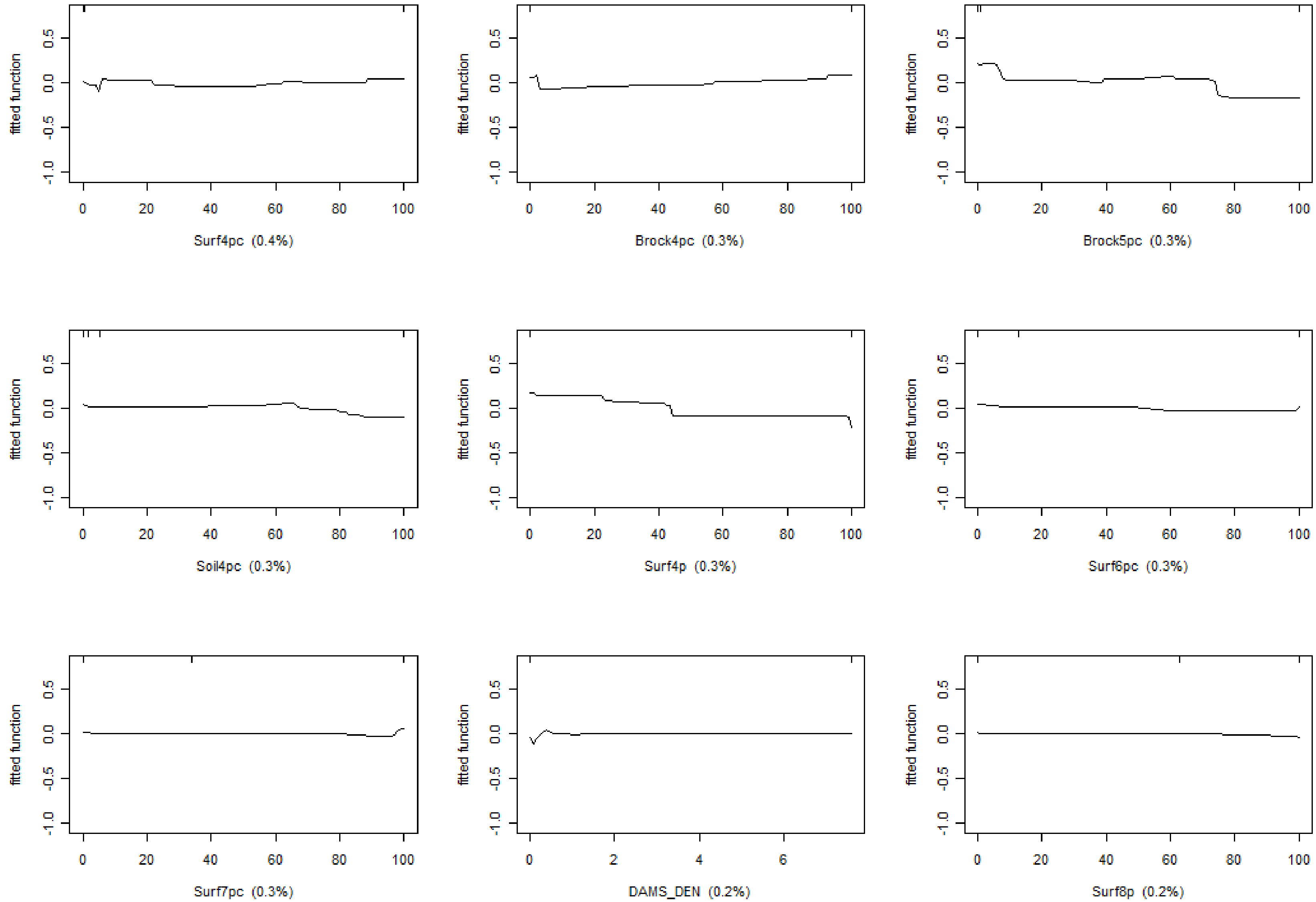

fitted function
0
-1
-2
-3
0 20 40 60 80 100
Soil1p (0%)
fitted function
0
-1
-2
-3
0 20 40 60 80 100
Brock2p (0%)
fitted function
0
-1
-2
-3
0 10 20 30 40 50
Brock3p (0%)
fitted function
0
-1
-2
-3
0 20 40 60 80 100
Brock4p (0%)
fitted function
0
-1
-2
-3
0 20 40 60 80 100
Surf4p (0%)

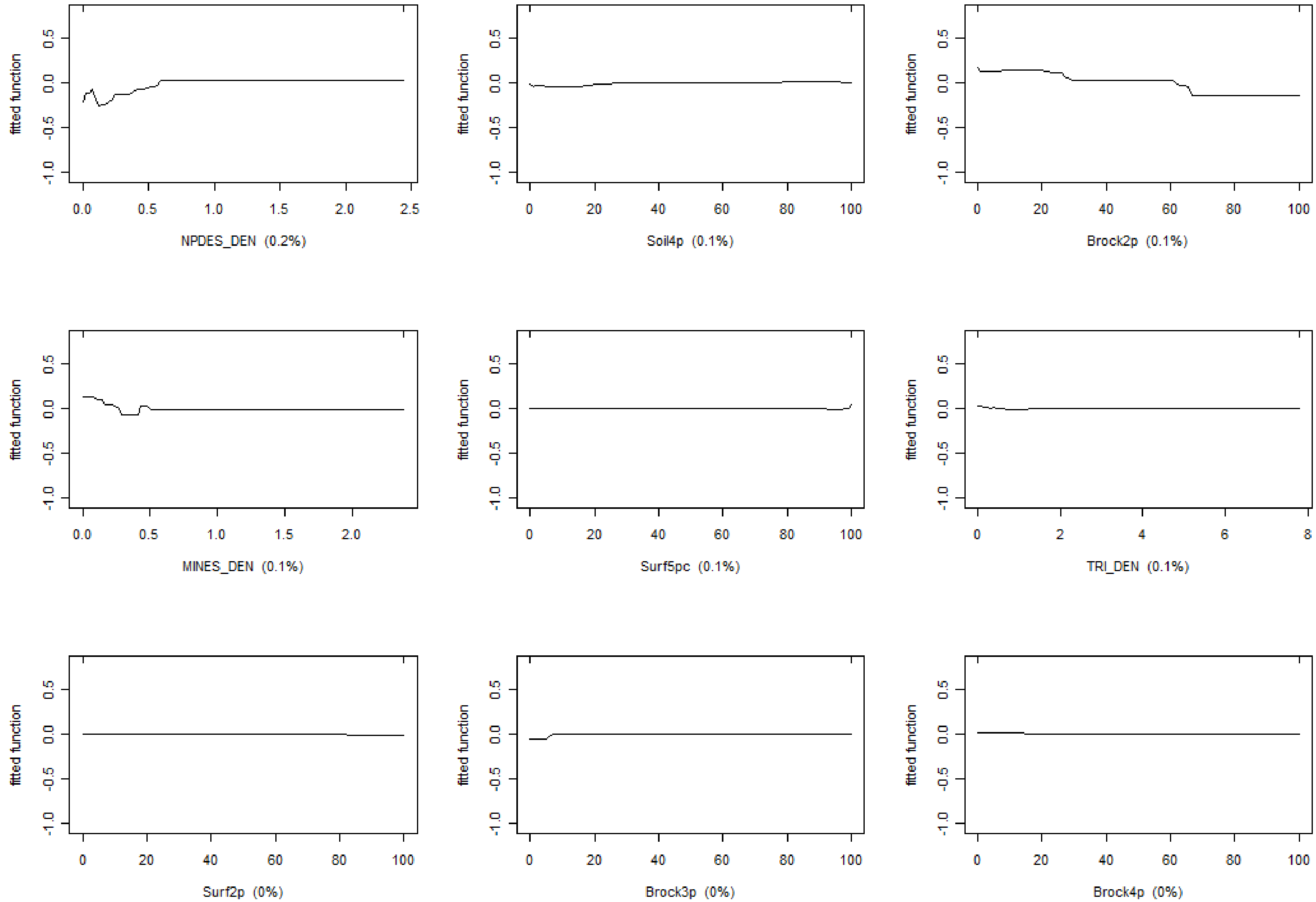
Percent Intolerant Fish

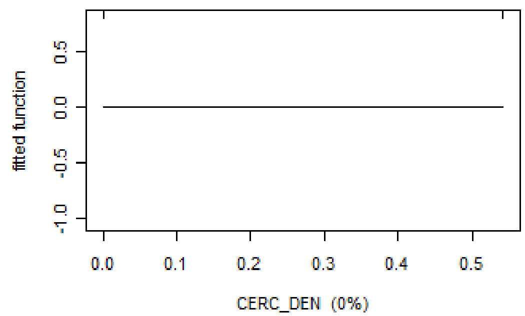



fitted function
0.5
0.0
-0.5
-1.0
60 80 100 120 140
Temp (2.6%)
0 10 20 30 40 50
Imp06C (2.5%)
0 5000 10000 15000
WATER_SWC (2.4%)
0 20 40 60 80 100
Soil2pc (2.4%)
0 1000 2000 3000 4000
CATTLEC (2.4%)
0 20 40 60 80 100
RIPDISP (2.3%)
0 200 400 600 800 1000
WATER_GWC (2.2%)
0e+00 1e+05 2e+05 3e+05 4e+05
ROADCRC_DEN (1.8%)
0 20 40 60 80 100
Brock6pc (1.8%)

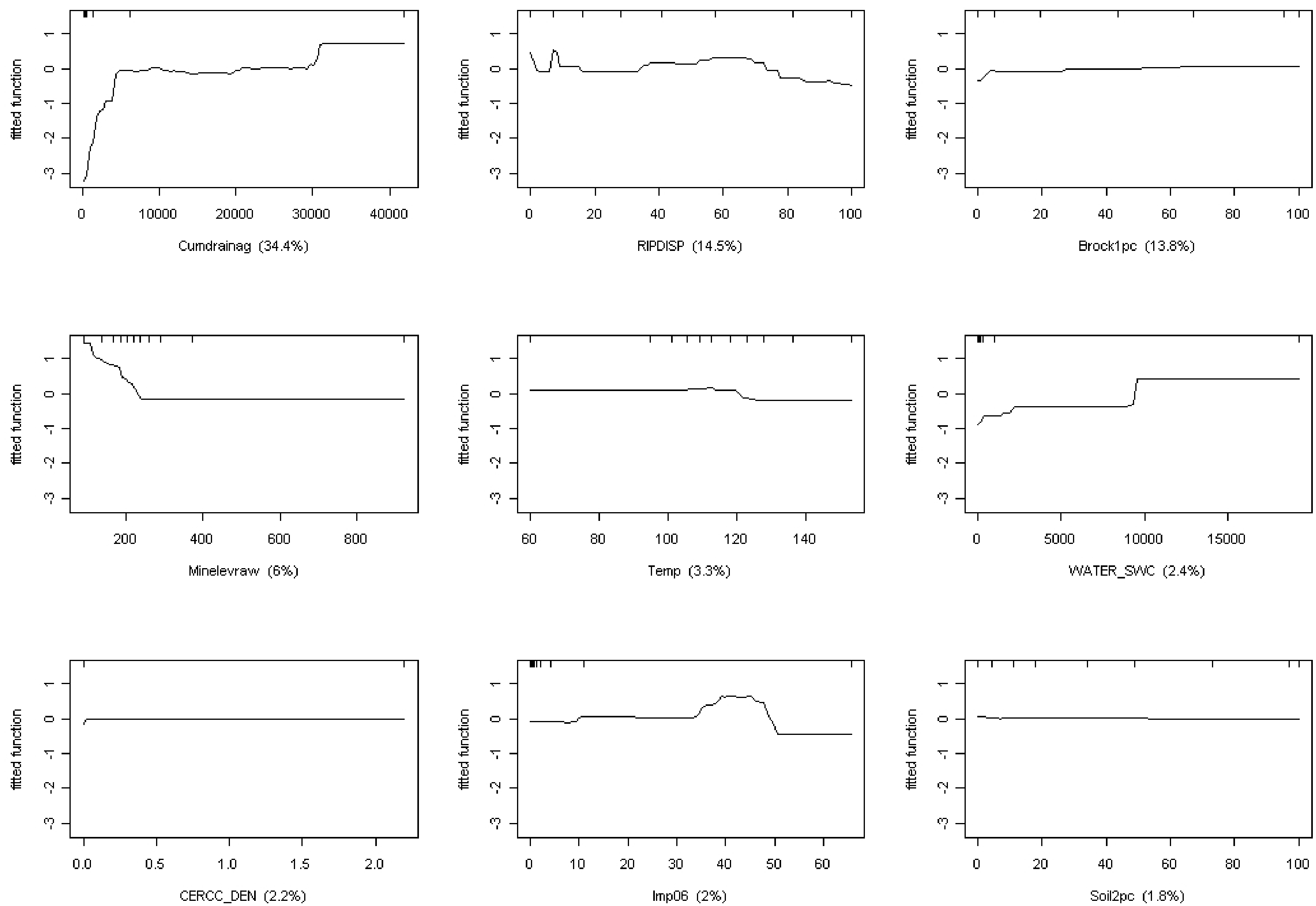

fitted function
-1.0
-0.5
0.0
0.5
0
10
20
30
40
grasspc (1.4%)
0
20
40
60
80
100
devp (1.2%)
0
50
100
150
200
250
300
ROADCR_DEN (1.1%)
cropspc (1%)
Brock7pc (1%)
Soil3pc (0.7%)
Surf3pc (0.7%)
0
2000
4000
6000
8000
DAMSC_DEN (0.6%)
Brock1pc (0.5%)

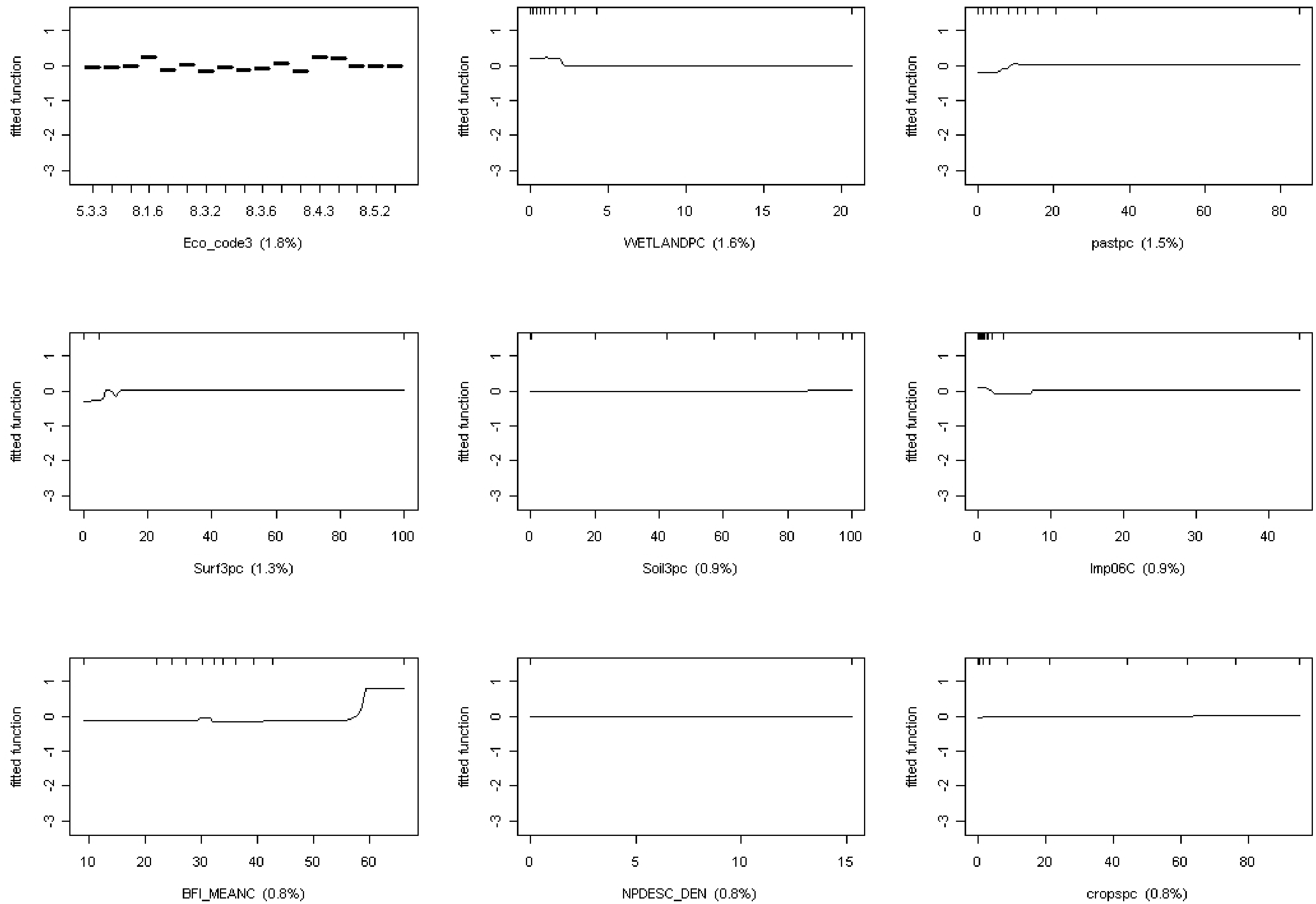

fitted function
Surf4pc (0.4%)
Brock4pc (0.3%)
Brock5pc (0.3%)
Soil4pc (0.3%)
Surf4p (0.3%)
Surf6pc (0.3%)
Surf7pc (0.3%)
DAMS_DEN (0.2%)
Surf8p (0.2%)

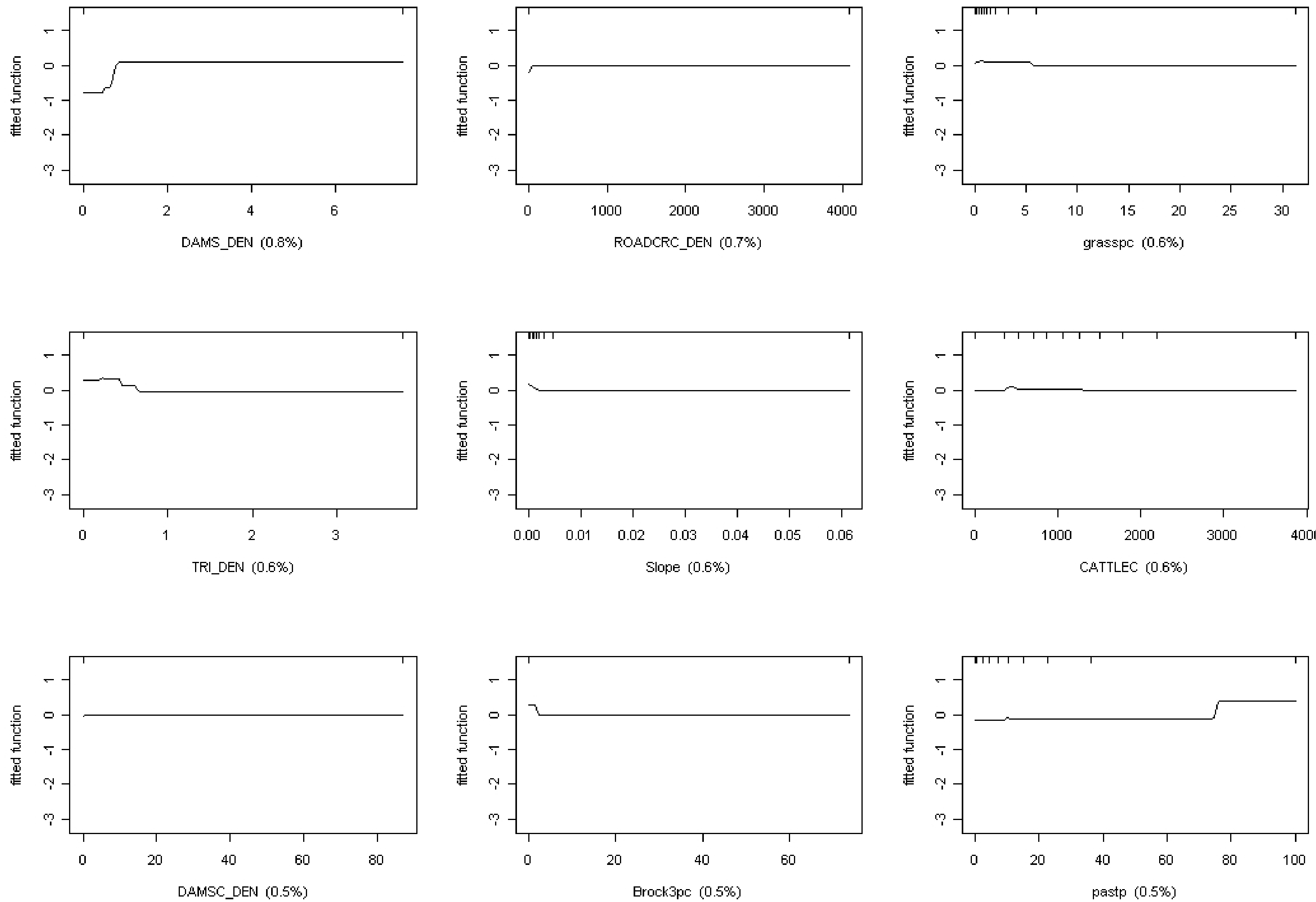

fitted function
0.5
0.0
-0.5
-1.0
0.0 0.5 1.0 1.5 2.0 2.5
NPDES_DEN (0.2%)
0 20 40 60 80 100
Soil4p (0.1%)
Brock2p (0.1%)
0.0 0.5 1.0 1.5 2.0
MINES_DEN (0.1%)
Surf5pc (0.1%)
0 2 4 6 8
TRI_DEN (0.1%)
Surf2p (0%)
Brock3p (0%)
Brock4p (0%)


fitted function
0.5
0.0
-0.5
-1.0
0.0
0.1
0.2
0.3
0.4
0.5
CERC_DEN (0%)

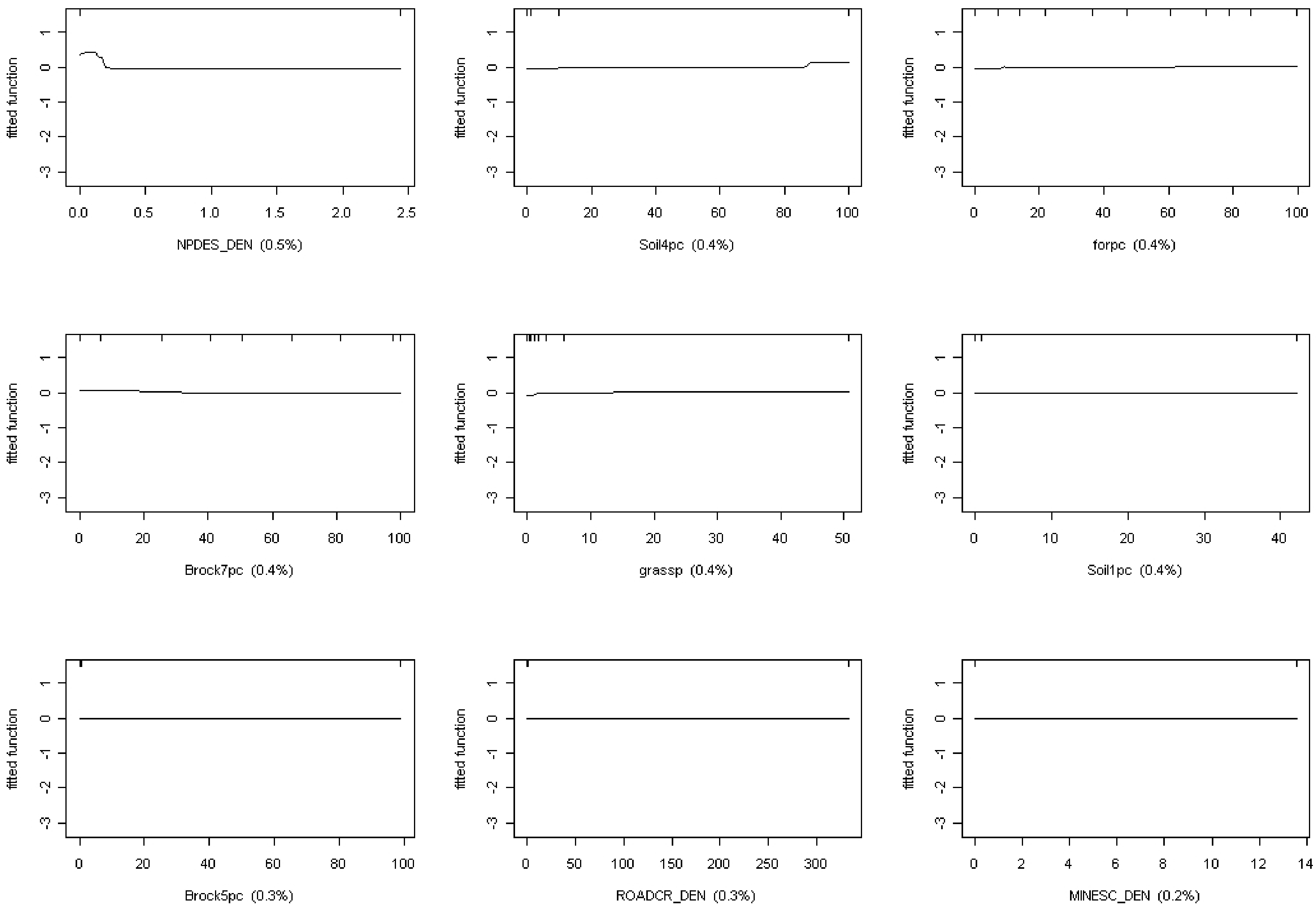
Great rivers species

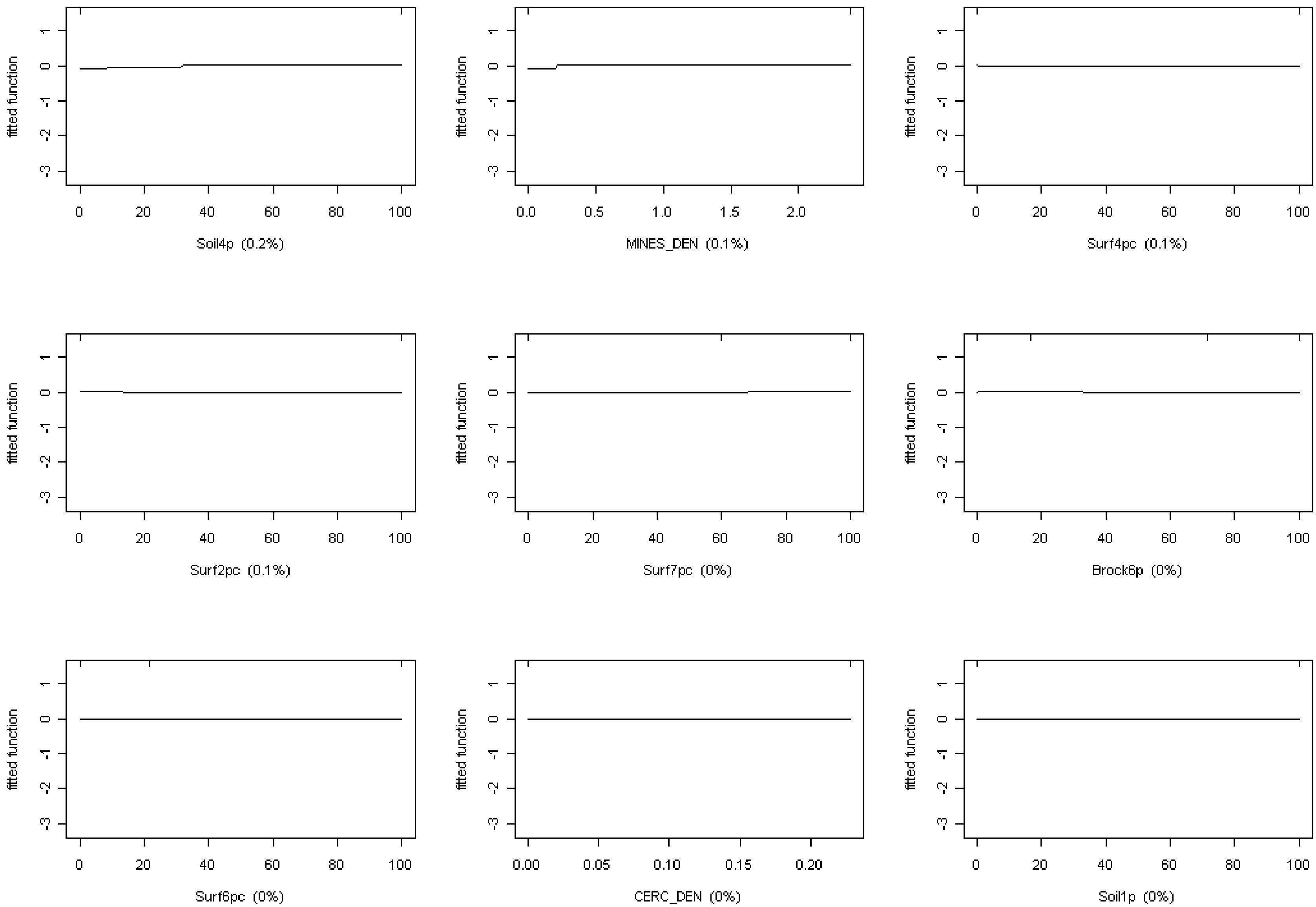


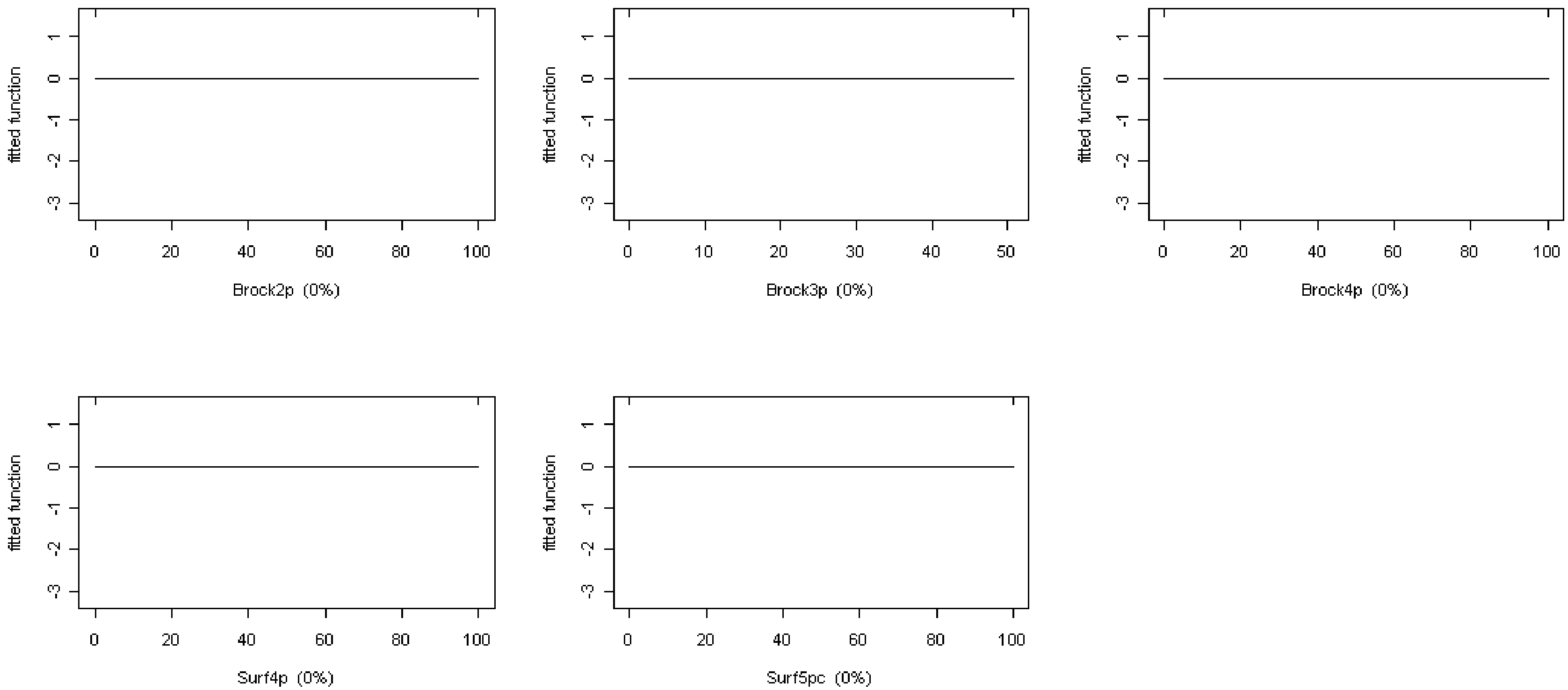

fitted function
1
0
-1
-2
-3
5.3.3 8.1.6 8.3.2 8.3.6 8.4.3 8.5.2
Eco_code3 (1.8%)
0 5 10 15 20
WETLANDPC (1.6%)
0 20 40 60 80
pastpc (1.5%)
0 20 40 60 80 100
Surf3pc (1.3%)
0 20 40 60 80 100
Soil3pc (0.9%)
0 10 20 30 40
Imp06C (0.9%)
10 20 30 40 50 60
BFI_MEANC (0.8%)
0 5 10 15
NPDESC_DEN (0.8%)
0 20 40 60 80
cropspc (0.8%)



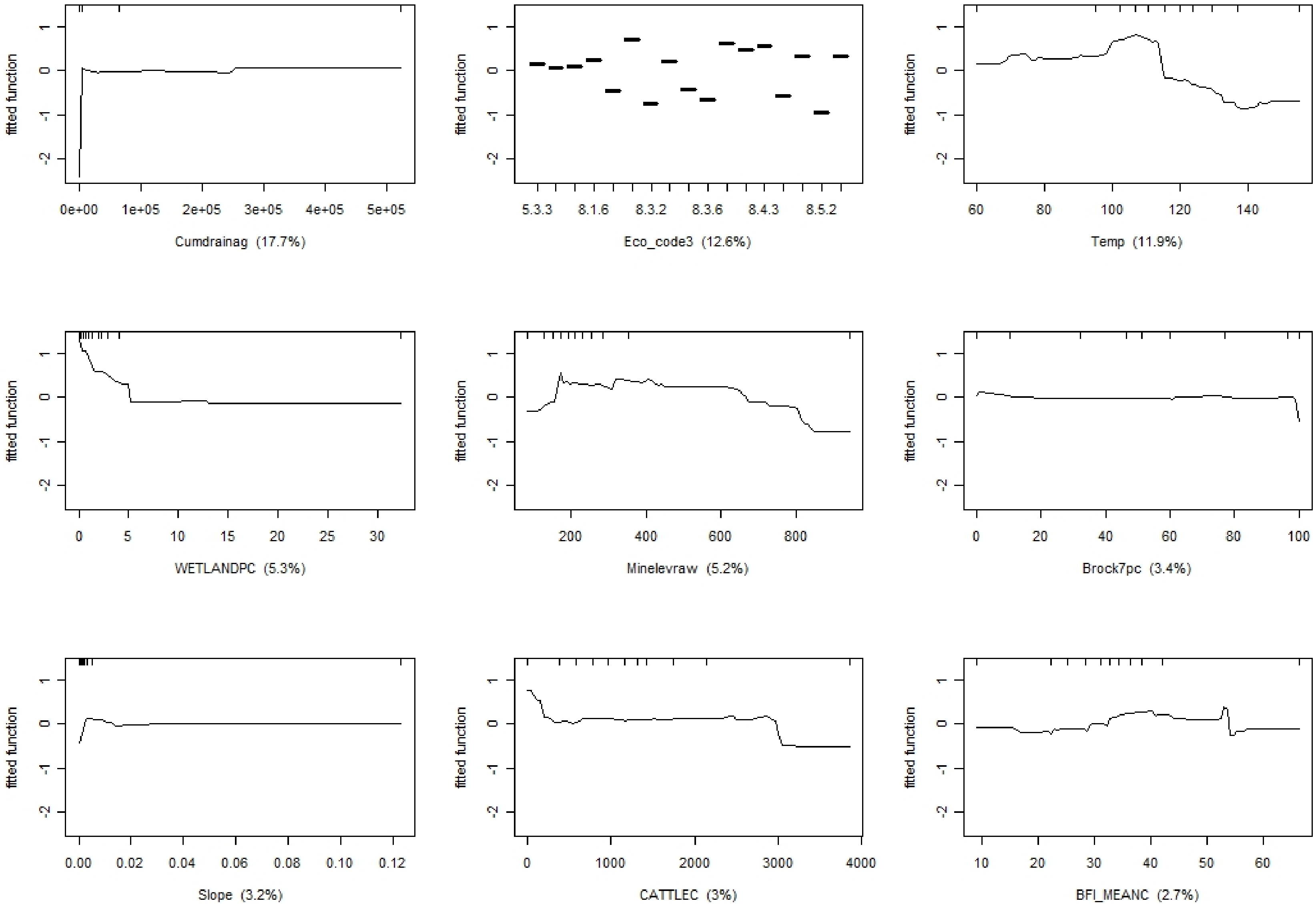
DAMS_DEN (0.8%)
ROADCRC_DEN (0.7%)
grasspc (0.6%)
TRI_DEN (0.6%)
Slope (0.6%)
CATTLEC (0.6%)
DAMSC_DEN (0.5%)
Brock3pc (0.5%)
pastp (0.5%)
fitted function

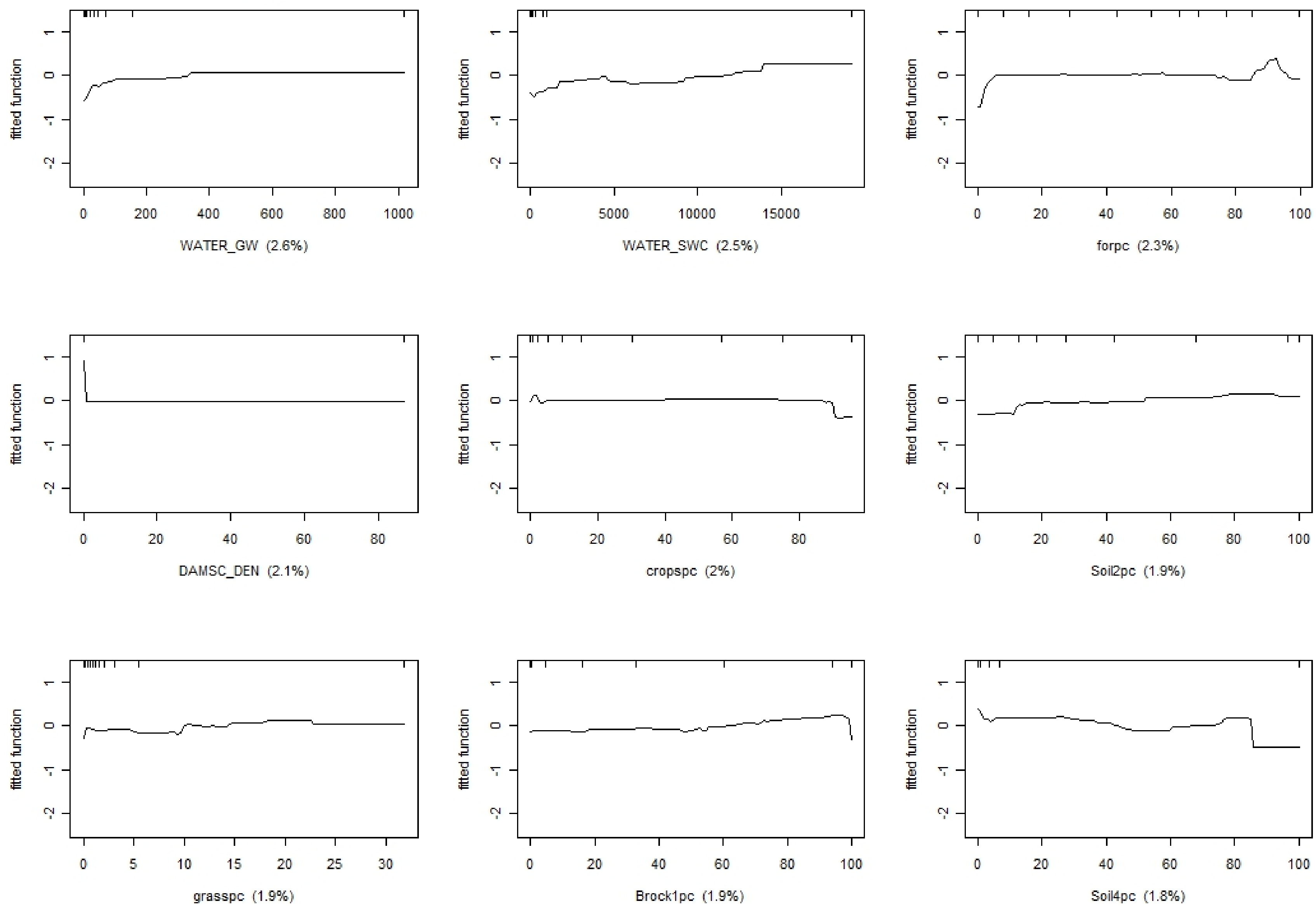

fitted function
NPDES_DEN (0.5%)
Soil4pc (0.4%)
forpc (0.4%)
Brock7pc (0.4%)
grassp (0.4%)
Soil1pc (0.4%)
Brock5pc (0.3%)
ROADCR_DEN (0.3%)
MINESC_DEN (0.2%)

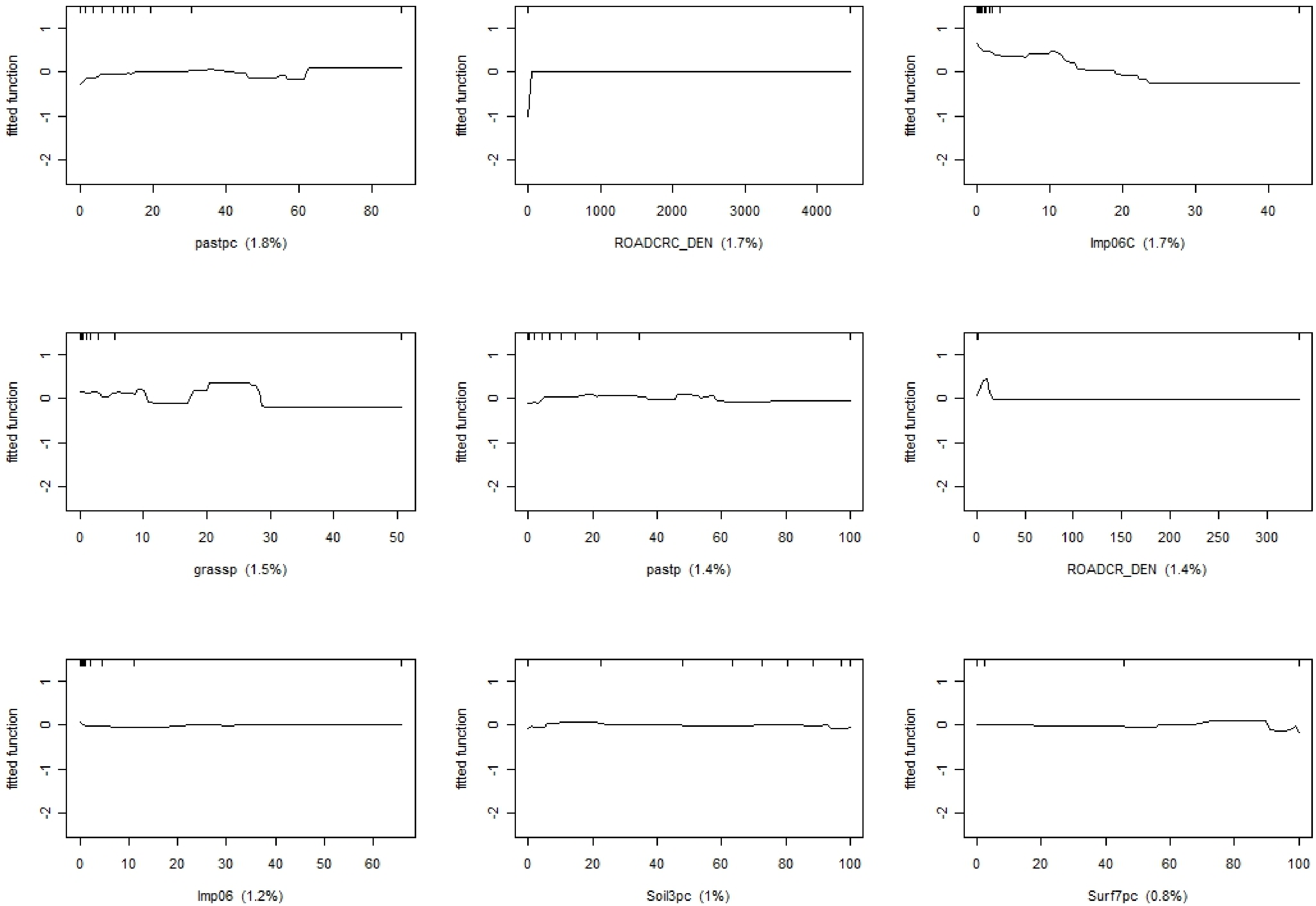

fitted function
Soil4p (0.2%)
MINES_DEN (0.1%)
Surf4pc (0.1%)
Surf2pc (0.1%)
Surf7pc (0%)
Brock6p (0%)
Surf6pc (0%)
CERC_DEN (0%)
Soil1p (0%)

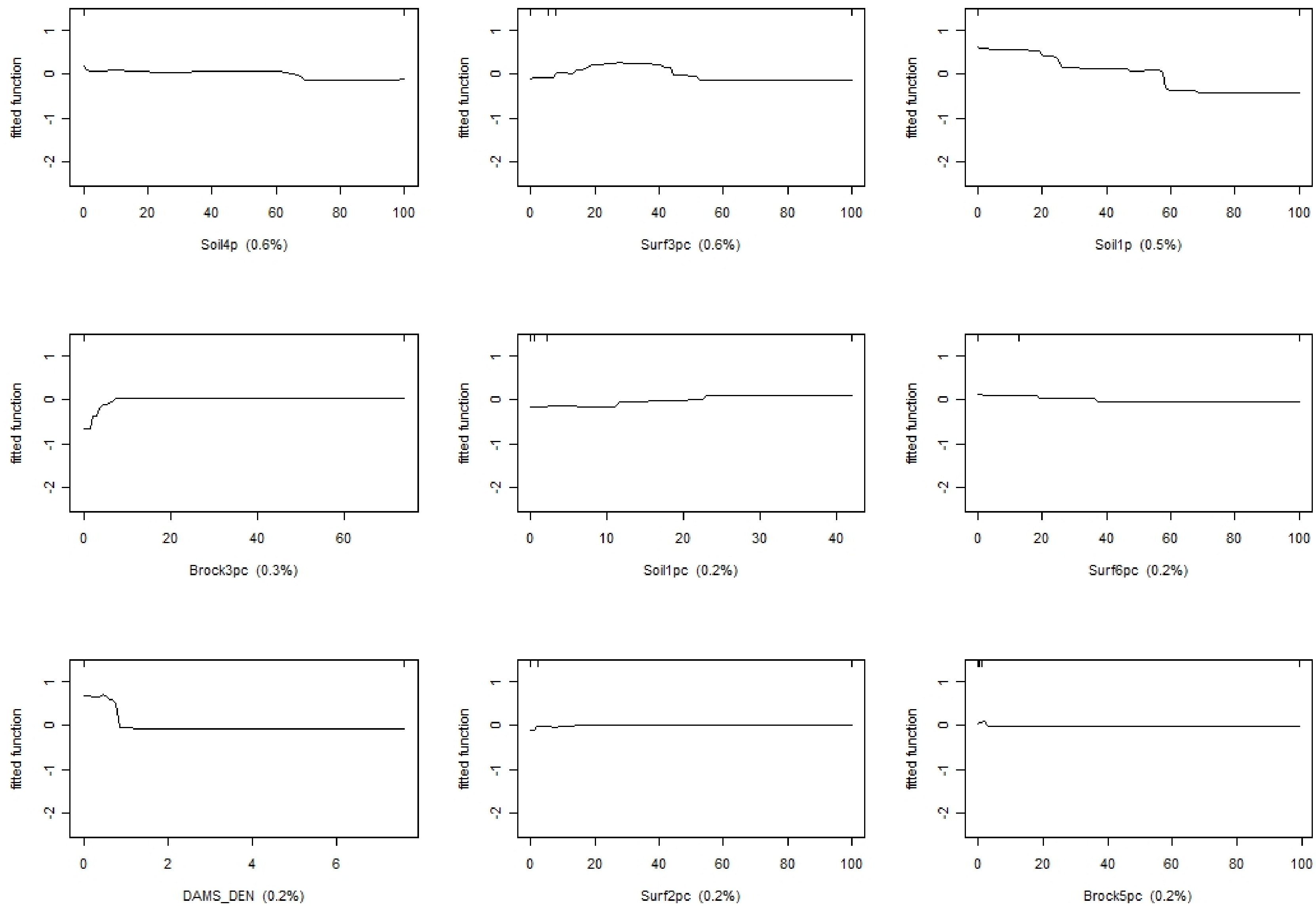

fitted function
Brock2p (0%)
Brock3p (0%)
Brock4p (0%)
Surf4p (0%)
Surf5pc (0%)

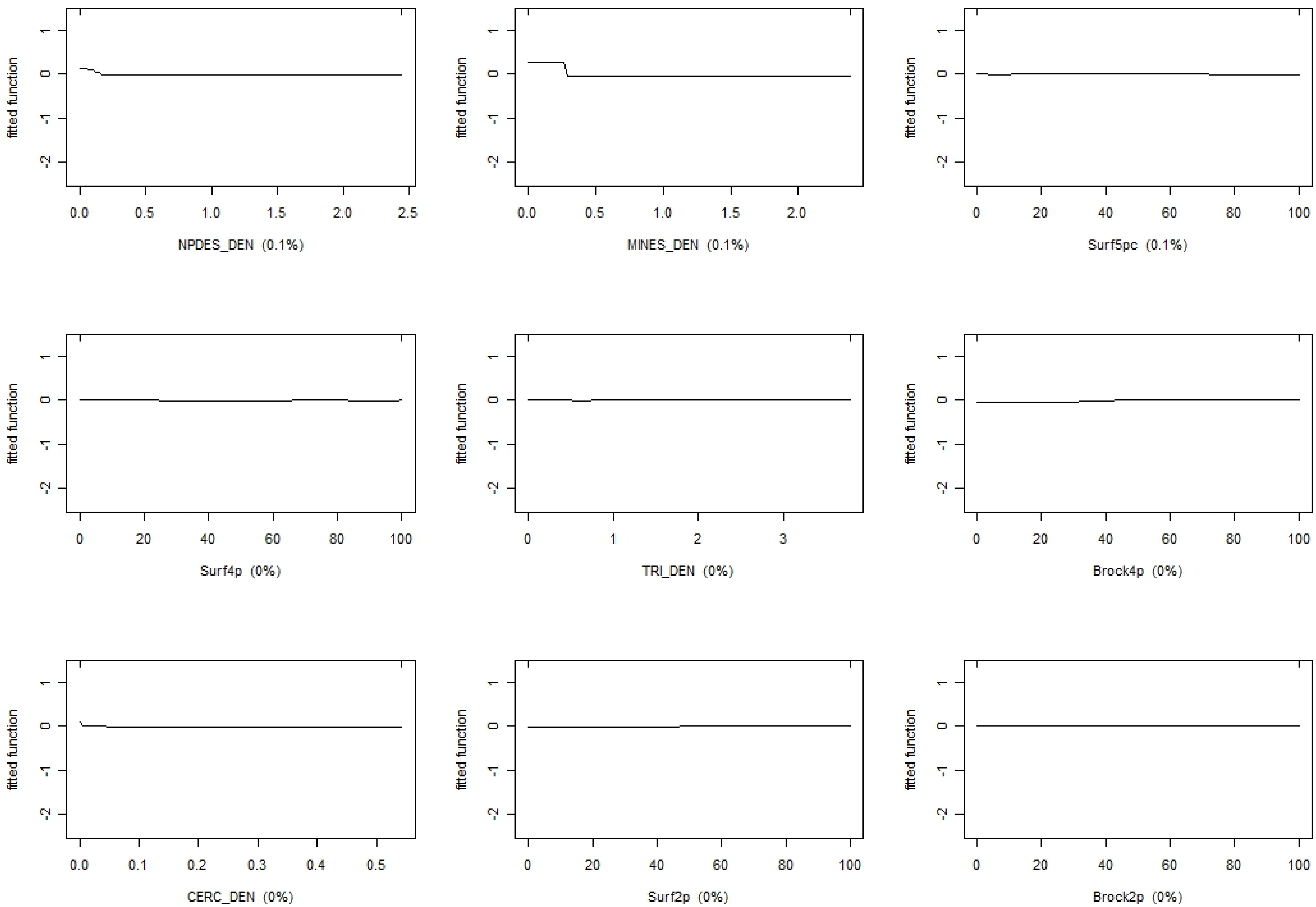
Intolerant mussel

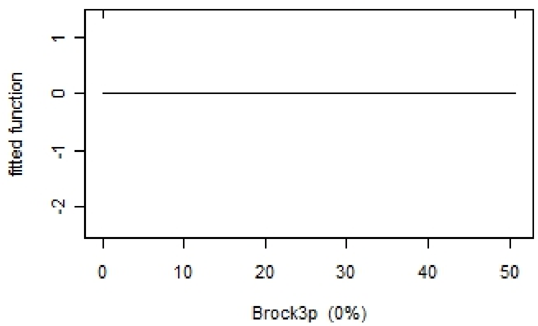



fitted function
1
0
-1
-2
0
200
400
600
800
1000
WATER_GW (2.6%)
0
5000
10000
15000
WATER_SWC (2.5%)
0
20
40
60
80
100
forpc (2.3%)
0
20
40
60
80
DAMSC_DEN (2.1%)
0
20
40
60
80
cropspc (2%)
0
20
40
60
80
100
Soil2pc (1.9%)
0
5
10
15
20
25
30
grasspc (1.9%)
0
20
40
60
80
100
Brock1pc (1.9%)
0
20
40
60
80
100
Soil4pc (1.8%)


fitted function
pastpc (1.8%)
ROADCRC_DEN (1.7%)
Imp06C (1.7%)
grassp (1.5%)
pastp (1.4%)
ROADCR_DEN (1.4%)
Imp06 (1.2%)
Soil3pc (1%)
Surf7pc (0.8%)


fitted function
Soil4p (0.6%)
Surf3pc (0.6%)
Soil1p (0.5%)
Brock3pc (0.3%)
Soil1pc (0.2%)
Surf6pc (0.2%)
DAMS_DEN (0.2%)
Surf2pc (0.2%)
Brock5pc (0.2%)


fitted function
NPDES_DEN (0.1%)
MINES_DEN (0.1%)
Surf5pc (0.1%)
Surf4p (0%)
TRI_DEN (0%)
Brock4p (0%)
CERC_DEN (0%)
Surf2p (0%)
Brock2p (0%)


fitted function
1
0
-1
-2
0
10
20
30
40
50
Brock3p (0%)