North Atlantic LCC Aquatic Habitat Assessment 2015

Chesapeake Bay Watershed Brook Trout Habitat and Climate Change Vulnerability Assessment

Final Report submitted to the North Atlantic Landscape Conservation Cooperative (NALCC) Assessment Project

Authors:
Jason Clingerman, Downstream Strategies
Todd Petty, West Virginia University
Fritz Boettner, Downstream Strategies

Abstract

Recently, the Chesapeake Bay Watershed Agreement established a management outcome focused on restoring and sustaining naturally reproducing brook trout populations in the Chesapeake Bay's headwater streams. Partners and stakeholders desired a statistically-valid predictive model that captured underlying cause and effect relationship between habitat characteristics and brook trout within this watershed. The model would ultimately be used to guide establishment of conservation priorities laid out within the Chesapeake Bay's Brook Trout Management Strategy.

Downstream Strategies, funded by the North Atlantic Landscape Conservation Cooperative, created a predictive model for brook trout within the Chesapeake Bay watershed. The analytical approach used for this assessment was boosted regression trees (BRT), a machine learning statistical method. The modeling process resulted in a series of quantitative outcomes, including: 1- predictions of expected current brook trout occupancy in all catchments within the modeling area, 2- a cross-validated measurement of model accuracy, 3- a measure of each predictor variable's relative influence on brook trout distribution, and 4- a measure of the functional relationship between each predictor variable and brook trout response. A post-modeling process was then used to quantify anthropogenic stress and natural habitat quality for all catchments within the study area based on BRT model outputs. In addition, we quantified how climate change may impact natural habitat quality and brook trout distributions. Finally, we detailed a case study that uses a hierarchical approach to establishing protection and restoration priorities at multiple scales using information on current brook trout distributions, levels of anthropogenic stress, underlying natural habitat quality, and expected impacts of future climate change.

Ultimately, all of the data and modeling results from this assessment will be incorporated into a web-based decision support tool. This tool will enable users to visualize and download data and model outputs, establish conservation priorities based on user-defined ranking criteria, calculate spatially-explicit predictions of brook trout response under various conservation scenarios, and assess conservation success within the context of future climate regimes (water temperature and rainfall). Combined, the modeling results contained within this report along with the publically accessible web application will improve public awareness of conditions and vulnerabilities of the Chesapeake Bay's headwater streams and empower resource managers to implement scientifically-defensible conservation actions. The web tool can be accessed at: www.fishhabitattool.org.

Acknowledgements

We offer a special thanks to Scott Schwenk, North Atlantic Landscape Conservation Cooperative, and Callie McMunigal, U.S. Fish and Wildlife Service for leading this project and coordinating partners.

This project would not have been possible without the invaluable technical assistance and review provided by the following: Nathaniel Hitt, U.S. Geological Survey; Mark Hudy, Eastern Brook Trout Joint Venture; Stephen Perry, Eastern Brook Trout Joint Venture; and Tyler Wagner, U.S. Geological Survey.

We would like to offer an additional thanks to Tyler Wagner, who provided datasets that were invaluable to the success of this effort.

Lastly, we appreciate all who participated in the process by providing data or feedback at various steps of the assessment. This group included:

Meredith Bartron, U.S. Fish and Wildlife Service; Doug Besler, North Carolina Wildlife Resources Commission; Jason Coombs, Eastern Brook Trout Joint Venture/University of Massachusetts; Sandra Davis, U.S. Fish and Wildlife Service; Jason Detar, Pennsylvania Fish and Boat Commission; Brad Fink, Virginia Depart of Game and Inland Fisheries; Nat Gillespie, U.S. Forest Service; Jennifer Greiner, U.S. Fish and Wildlife Service; Alan Heft, Maryland Department of Natural Resources; Fred Henson, New York State Department of Environmental Conservation; David Lawrence, National Fish and Wildlife Foundation; Ben Letcher, U.S. Geological Survey; Maritza Mallek, U.S. Fish and Wildlife Service; Rachel Muir, U.S. Geological Survey; Keith Nislow, U.S. Forest Service; Jake Rash, North Carolina Wildlife Resources Commission; Steve Reeser, Virginia Depart of Game and Inland Fisheries; Mike Shingleton, West Virginia Division of Natural Resources; Michael Slattery, U.S. Fish and Wildlife Service; Rich Starr, U.S. Fish and Wildlife Service; David Thorne, West Virginia Division of Natural Resources; Megan Tyrrell, U.S. Fish and Wildlife Service

Contents

Abbreviations

ACFHP	Atlantic Coastal Fish Habitat Partnership
ASI	anthropogenic stress index
BRT	boosted regression tree
DS	Downstream Strategies
DST	Decision Support Tool
HQI	natural quality index
HUC 8	Hydrologic Unit Code 8-Digit watershed
CART	classification and regression trees
CV	cross-validation
DS	Downstream Strategies
EBTJV	Eastern Brook Trout Joint Venture
FHP	Fish Habitat Partnership
GIS	geographic information systems
IPCC	Intergovernmental Panel on Climate Change
km	kilometer
NALCC	North Atlantic Landscape Conservation Cooperative
NHD	National Hydrography Dataset
NLCD	National Land Cover Database
ROC	receiver operating characteristic
USFWS	United States Fish and Wildlife Service
USGS	United States Geologic Survey
WVU	West Virginia University

Introduction

Project Background

NALCC Funding

Downstream Strategies, LLC (DS) was contracted by the North Atlantic Landscape Conservation Cooperative (NALCC) to perform aquatic assessments across the extent of the NALCC, which ranges from Maine to Virginia. These assessments were to be based off of previous work DS performed to assess habitats for numerous fish species for the Midwest Fish Habitat Partnerships (FHP). Brook trout were identified as a potential species of interest for an assessment because of the interest and availability of partners from the Chesapeake Bay Program. Consequently, a brook trout model for the entire Chesapeake Bay watershed was developed. The NALCC supported and coordinated the project in collaboration with a review team that included members from United States Fish and Wildlife Service (USFWS), the United States Geologic Survey (USGS), the Eastern Brook Trout Joint Venture (EBTJV), Atlantic Coastal Fish Habitat Partnership (ACFHP), West Virginia University (WVU), and numerous state fish and wildlife agencies.

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 and human influences on aquatic habitats. Specifically, the outcomes can be used to conduct fish habitat condition assessments based on a range of stakeholder-specified metrics and modeling endpoints to help determine natural drivers of aquatic conditions, as well as primary stressors to brook trout within the Chesapeake Bay watershed. The ultimate goal is to improve understanding of how local (e.g., stream water temperature) and network(e.g., upstream agriculture) processes influence stream conditions in the region and to provide additional knowledge, data, and tools to help prioritize and inform conservation and restoration actions throughout the Chesapeake Bay watershed.

Context with other assessments in North Atlantic

This report summarizes the predictive models and tools for brook trout that are part of the inland portion of the NALCC funded aquatic assessment project. Additional models of winter flounder and river herring are currently being constructed to complement this brook trout model for the North Atlantic region. Details of marine and diadromous fish habitat assessments will be described in separate reports.

Previous applications in Midwest Region

DS's inland fish species aquatic habitat modeling approach was developed for several FHPs in the Midwest region and was funded by the USFWS. These assessments served as the basis for the analysis detailed in this report. The Midwest assessments utilized the existing National Hydrology Dataset (NHD) and the NHD Plus (Horizon Systems, 2012) supplemental information on hydrology networks. Data included discrete catchment polygons that delineated the local drainage area for each specific stream segment. These catchments were utilized as our modeling unit, and predictor data were summarized within each distinct catchment. Response data were likewise summarized within catchments where available in order to create our predictive models, the results of which were also extrapolated to all catchments within the defined study areas. In total over 30 distinct models were created for the six FHPs within the Midwest. These model results were distributed as stand-alone geodatabases and within a desktop decision support tool which ran using desktop ArcGIS environment. Currently the decision support tool is being developed as a web-based application to provide improved accessibility to partners and stakeholders. Ultimately, the data and results from the Chesapeake Bay brook trout model will be incorporated into this web-based platform as well.

Review of previous brook trout assessments in Chesapeake Bay

Two important assessments of brook trout distributions within the EBTJV geographic boundary, which includes the Chesapeake Bay watershed, have recently been conducted by other researchers (EBTJV, 2015; Deweber and Wagner, 2015). The EBTJV (2015) assessment utilized data on known locations of brook, brown and rainbow trout to classify catchments 1:100K catchments according to the population types found within them. Population types included: brook trout present (exotic trout present) and brook trout present (exotic trout absent). They then used a set of criteria to extrapolate classifications of brook trout and exotic trout presence upstream. Remaining areas were classified as absences. Patches were then defined as clusters of interconnected catchments where brook trout are present. This assessment produced a categorical assignment of brook trout populations, both at the catchment and patch scale.

DeWeber and Wagner (2015) created a predictive model for brook trout at the extent of the entire EBTJV boundary. They used data on brook trout presences and absences within a Hierarchical Bayesian modeling framework to produce a predicted probability of brook trout presence for each 1:100K catchment. This predictive model utilized a modeled stream temperature variable (DeWeber and Wagner, 2014) along with other land use characteristics as predictor variables.

Justification of DS assessment

Stakeholders from within the Chesapeake Bay watershed desired a statistically valid predictive model that captured underlying cause and effect relationships between habitat characteristics and brook trout within this watershed. By building our model for only the Chesapeake Bay watershed (as opposed to a larger, regional extent), it produced results that are not influenced by data, processes, or relationships outside of the Chesapeake Bay watershed. During a case study of scale (Downstream Strategies, 2013), we found that model accuracy declines as the scale of the modeling effort increases. For example, we found that predictive accuracy of stream conditions within a particular hydrologic unit code 8-digit watershed (HUC 8) are maximized when HUC 8 scale-specific models are constructed, rather than regional scale models that include the HUC 8 of interest. Likewise, it is believed that a model constructed specifically for the Chesapeake Bay watershed will maximize predictive accuracy relative to a regional scale model that includes the Chesapeake Bay.

Furthermore, a key objective of the NALCC project is to produce a decision support tool with the capacity to run scenarios of how management actions may benefit brook trout habitat. To do this, we need a modeling structure that can be run efficiently within a web-based application. The statistical method we use (boosted regression trees (BRT), described in the subsequent BRT section) enables us to quickly produce quantitative measures of probability of presence, natural habitat quality, and stress. This allows for the creation of an "on-the-fly" scenario-based decision support tool (See Futuring Tool section for more description of this tool). The other existing brook trout models cannot be used in this manner.

Objectives

The over-riding objective of this project was to construct a useful Decision Support Tool (DST) built upon a validated predictive model of brook trout distributions both under current climate regimes as well as a variety of potential future climate regimes.

Specifically, we:

1 – Constructed and validated a BRT model that could reliably estimate the probability of brook trout occurrence in 1:100K scale catchments throughout the Chesapeake Bay watershed;

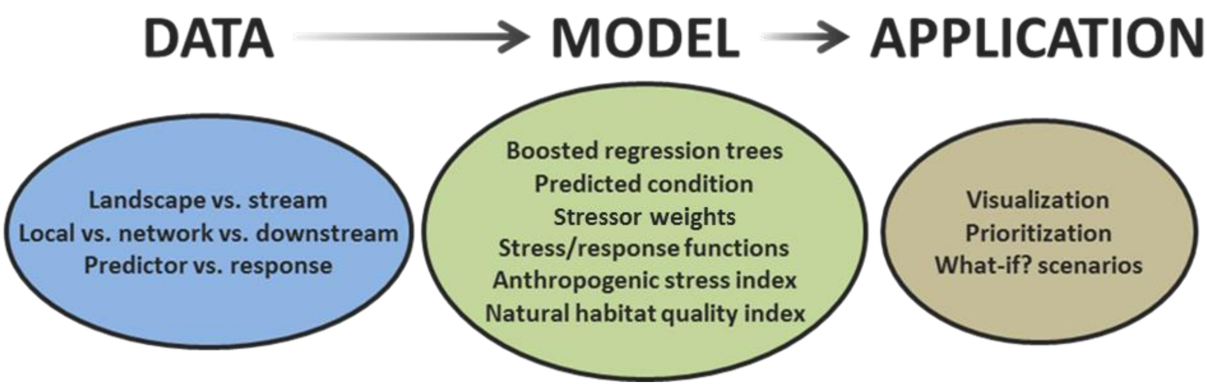
2 – Used BRT model outputs to calculate measures of underlying natural habitat quality and anthropogenic stress;

3 – Assessed future climate scenarios and the potential impact to brook trout populations (change in occupancy, stress, and natural habitat quality); and

4 – Created analytical tools to facilitate visualization of data and model results, prioritization of conservation actions, and estimation of brook trout habitat response to specific restoration actions under current and future climate scenarios.

A diagram of the general assessment process is outlined in Figure 1. DS acquired landscape and aquatic data from multiple sources to develop models and tools for visualizing expected current and potential future conditions and for prioritizing management actions.

FIGURE 1: DIAGRAM OF THE HABITAT ASSESSMENT PROCESS



Overview of Assessment Methodology

Data

Predictor Variables

The predictor variables were typically measures of land use or land cover derived from geographic information systems (GIS), such as percent impervious surface area or road crossing density. The predictor variables were compiled at multiple scales, including the local scale (e.g., single 1:100k NHD stream catchment), or the network scale (e.g., all upstream catchments and the local catchment). Predictor data consisted of both natural variables (such as geology or elevation) and data we classified as anthropogenic in nature. Anthropogenic predictors included predictors such as agriculture, impervious surfaces, and mining.

Response Variable

The response variable in a BRT model can be count data, continuous data, or binary data. The response variable for this project was the binary presence-absence of brook trout. DS compiled fish sample data from several state fish and wildlife agencies and then utilized the most recent sample within each catchment to create the final presence-absence response for modeling. This resulted in a single value of presence-absence for each catchment. Although the predictor variables were compiled from datasets created at multiple scales, the response variable was always measured at the local scale (e.g., individual sample site on a stream).

BRT

The statistical methodology utilized within the Midwest FHP assessments was BRT, a machine learning statistical method. This method was selected after careful review of many statistical methodologies. DS staff and partners, along with the stakeholders for the FHP assessments, decided upon BRT over competing methodologies after comparing and contrasting the strengths and weaknesses of each. BRT models combine decision trees and boosting methodologies, which often result in better cross-validated models than other methods (Elith et al., 2006), including classification and regression trees (CART). Decision trees are advantageous because (1) they can incorporate any type of predictor data (binary, numeric, categorical); (2) model outcomes are unaffected by differing scales of predictors; (3) irrelevant predictors are rarely selected; (4) they are insensitive to outliers and non-normalized data; (5) they can accommodate missing predictor data; (6) they can accommodate co-varying predictor variables; and (7) they can automatically handle interactions between predictors (Elith et al., 2008). The boosting algorithm used by BRT improves upon the accuracy of a basic regression tree approach by following the idea that averaging many models offers efficiency over finding a single prediction rule that is highly accurate (Elith et al., 2008). The software used to create the BRT models was R utilizing the 'gbm' package and source code from Elith et. al 2008 supplemental materials.

The modeling process results in a series of quantitative outcomes, including: predictions of expected current conditions of all catchments in the modeling area, measurement of prediction accuracy, a measure of each predictor'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. The predictions of current conditions were created by extrapolating the BRT model to all catchments within the modeling area. The unit of the predicted current condition for this assessment is the probability of brook trout presence. These current conditions are useful for assessing habitats and mapping the expected range of species.

Predictive accuracy was 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 and the CV correlation coefficient, are actually averages of ten separate ROC or correlation measurements. A standard error for the ten estimates is also provided. CV measures are designed to estimate how well the model will perform using independent data (i.e., data not used to build the model).

Additionally we evaluated predictive performance on a fully independent dataset. Ten percent of the available response data was held out to perform this test. We assessed the misclassification rates on this dataset utilizing five thresholds to indicate presence and absence. Thresholds to determine presence or absence were calculated using the following information: (1) training data prevalence, (2) where sensitivity equaled specificity, (3) maximum Kappa, (4) maximum percent correctly classified, and (5) the average of (2) and (3).

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.

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).

Residual Analysis

Analyzing patterns of omission and commission may highlight regions where the model is performing well or poorly or could suggest missing explanatory variables. Residuals are calculated by the BRT model and are used to assess Type I and Type II errors. The residuals are a measure of the difference in the measured and modeled values (measured value *minus* modeled value). Negative residuals indicate over-predictions (predicting higher values than are true), while positive residuals indicate under-predictions (predicting lower values than are true).

In order to assess spatial structure in the residuals, we used the 'ade4' package within R to run a Mantel test. This test utilized distance matrices of both the station locations and the residual value to determine if there is spatial structure in the residuals. We used 9999 permutations of this method to estimate a precise p-value.

Derivation of Anthropogenic Stress Index and Habitat Quality Index

Characterizing anthropogenic stress and natural habitat quality of aquatic habitats is a necessary process for helping natural resource managers identify place-based conservation and restoration strategies. A post-modeling process was used to characterize anthropogenic stress and natural habitat quality for all catchments within the study area. Stress and natural habitat quality indices and metrics were identified and calculated based on BRT model outputs, and details of those calculations are below.

Once developed, these indices of stress and habitat quality can be used to generate and visualize restoration and protection priorities by analyzing how stress reduction can increase the probability of brook trout presence. For example, areas of high natural quality and low stress could represent protection priorities, whereas areas of high natural quality and high stress may represent restoration priorities. In addition, we can quantify how climate change may affect brook trout distributions through an effect on underlying natural habitat quality over time.

Anthropogenic stress

Stress indices are useful for evaluating anthropogenic landscape drivers that structure aquatic responses. Natural resource managers can use stress indices and metrics to assess how anthropogenic processes are impacting aquatic responses and can utilize this information to site restoration projects in order to maximize efficiency. Individual stressors were identified by examining BRT model outputs, both the variable influence table and the functional relationship between predictor variables and response variable. Any predictor variable significantly affected by anthropogenic disturbance was included as a potential stressor.

Individual stress metrics were calculated by determining the increase in probability of presence for each catchment when the statistical effect of that predictor variable was removed. A new predictor variable dataset was produced to calculate each individual stressor metric. The new predictor dataset contained the same values as the original predictor dataset except for a single anthropogenic variable for which a stress metric was calculated. For this variable, the values were all set to reflect "no stress." This provided a hypothetical baseline that represented the removal of all stress from that predictor variable. The existing BRT model was then applied to the new hypothetical landscape data to provide an extrapolation of the current model assuming zero stress for that stressor. All the stressors used had examples of "no stress" in the training dataset used to build the model, which ensures that calculations of stress were not derived by extrapolating the model beyond the range of the data. The difference between the current predicted probability of presence and the probability of presence under this "no stress" situation indicated the change that could be attributable to stress. This process was repeated for each stressor to generate individual metrics of stress on a potential scale of 0-1. Higher stress values indicated a larger change in predicted probability of presence after removing stress, and lower stress values indicated that the catchment was relatively unaffected by removing stress (Table 1).

For each catchment, the individual stress metrics (e.g. agriculture stress, impervious surface stress, mining stress) were summed to produce an overall stress metric, the anthropogenic stress index (ASI). The generalized formulas for calculating individual stress metrics and ASI are as follows:

$$individual\ stress\ metric = probability\ of\ presence_{no\ stress} - probability\ of\ presence_{current}$$

$$anthropogenic\ stress\ index\ (ASI) = \sum individual\ stress\ metrics$$

TABLE 1: EXAMPLE OF STRESS CALCULATIONS

Comid	Current Condition Predictions	Stressor 1 Predictions	Stressor 1 Metric	Stressor 2 Predictions	Stressor 2 Metric	Anthro. Stress Index (ASI)
Catchment ID	**Predicted probability of occurrence using current landscape data**	**Predicted probability of occurrence when stressor 1 removed**	**(Stressor 1 pred – Current Pred)**	**Predicted probability of occurrence when stressor 2 removed**	**(Stressor 2 pred – Current Pred)**	**Stressor 1 Metric + Stressor 2 Metric**
1234567	0.80	0.90	0.10	0.80	0	0.10
1234568	0.25	0.50	0.25	0.35	.10	0.35
1234569	0.5	0.7	0.2	0.55	.05	0.25

Natural habitat quality

Natural habitat quality metrics provide baseline information on the optimal potential condition of a catchment. We defined natural quality as the maximum probability of presence under a zero-stress situation; essentially, the highest attainable condition in the catchment. These metrics allow natural resource managers to further classify each catchment and target specific land-based conservation or restoration actions.

The natural habitat quality index (HQI) was calculated directly from the BRT output. Metrics for 'natural' predictor variables were calculated using a different approach than for the stressor calculations detailed above. A single hypothetical 'no stress' dataset was created where all stressors were removed. The existing BRT model was then applied to this hypothetical predictor dataset, and the resulting probability of presence indicated the maximum condition attainable by removing all stress. This hypothetical situation where all stressors were zero was also represented in the training dataset, which ensures that these extrapolations are not outside of the range of the data used to build the model. The probability of presence calculated by the BRT model for this hypothetical 'no stress' dataset is the HQI and this value indicates the maximum condition expected in each catchment.

Applications

Hierarchical Visualization

This visualization tool is used to examine all the datasets used in the assessment process. Datasets include current conditions, stress and natural quality variables, socioeconomic information, and model results. Two scales of visualization are available: regional and local. The regional scale maps data by HUC12 for entire study areas, while the local scale maps catchment-level data within a single HUC8. All data from each scale can be mapped and exported.

Ranking/Prioritization

Users can rank catchments within a selected HUC8 watershed by selecting and weighting data. Variables can include modeling results and additional socioeconomic factors. The tool will produce a new output that displays catchments ranked by user criteria. All data can be exported and mapped.

Futuring Tool

The web-based futuring tool allows the user to examine brook trout habitat stressors for specific catchments. The user can then modify existing conditions and predict changes in overall condition, both locally and downstream for brook trout.

Supplemental information on how to use these tools, and a case study detailing example scenarios can be found in the separate report "Chesapeake Bay Brook Trout Assessment 2015: Using decision support tools to develop priorities".

Chesapeake Bay Brook Trout Model

Predictor Data

DS, in cooperation with the project's Technical Review Team, arrived at a list of landscape-based habitat variables (Appendix A) used to predict brook trout throughout the region; those variables were also used to characterize habitat quality and anthropogenic stress. DS and the Review Team compiled a list of 45 predictors for evaluation. From that list, 35 variables were removed due to statistical redundancy ($r > 0.6$), logical redundancy, or because of a lack of model influence, resulting in a final list of 10 predictor variables for the BRT model and assessment. Most predictor variables were gathered from public sources, but modeled stream temperature was acquired from Tyler Wagner, USGS, PA Cooperative Fish and Wildlife Research Unit. A detailed description of the modeled stream temperature variable can be found in DeWeber and Wagner (2014).

Response Data

DS compiled 16,261 unique stream fish collection records from 1995 to 2013, the details of which can be found in Table 2. A large portion of this data was also provided by Tyler Wagner. Other data were acquired by DS from state fish and wildlife agencies, or provided by coauthor Todd Petty. DS processed those data to create a presence-absence dataset for brook trout, which was comprised of data for 3,284 catchments. Figure 2 illustrates all of the sampling sites that were used to construct the model.

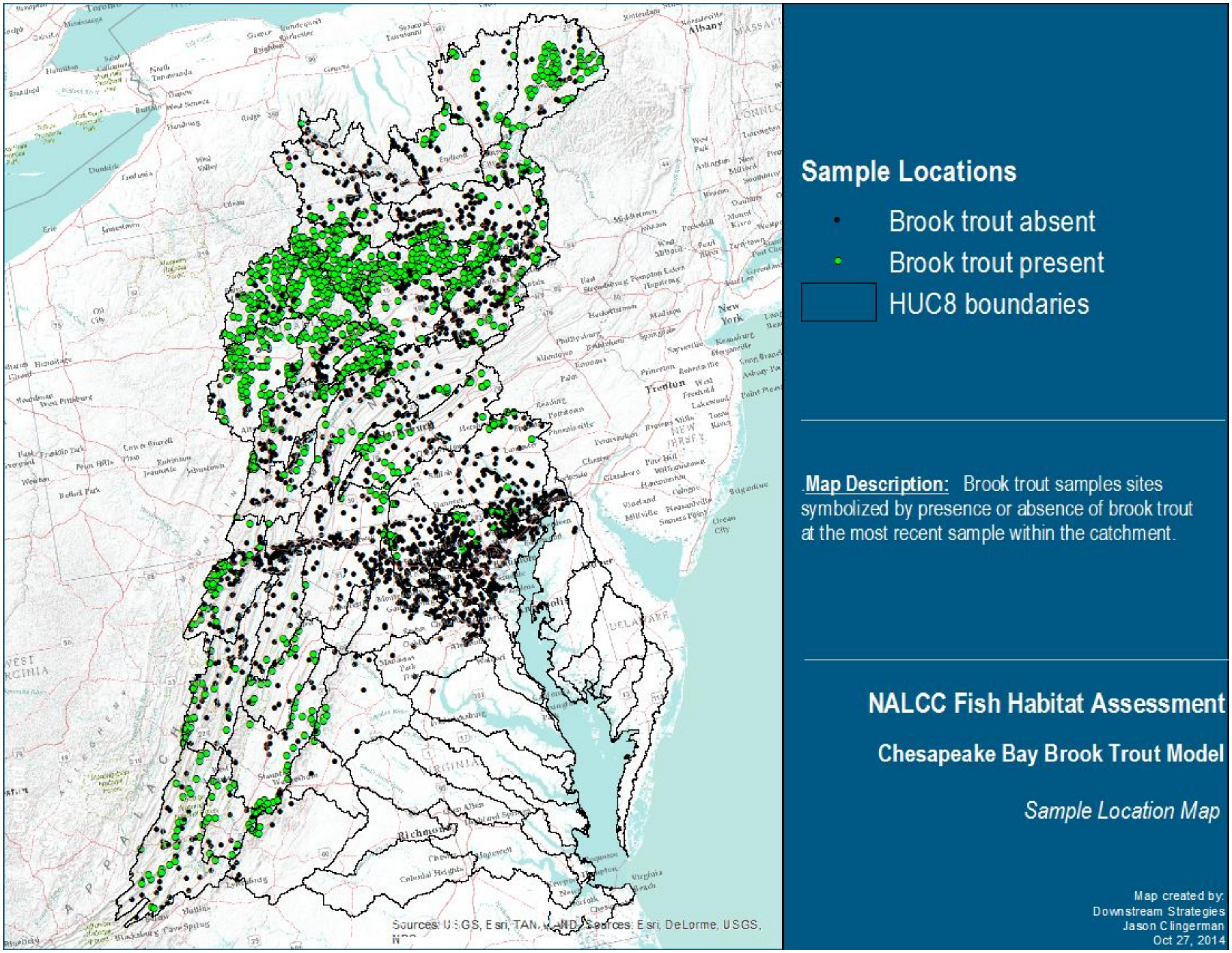
TABLE 2. FISH DATA SOURCES

Data Source	Data Provider	Date range	# samples
Pennsylvania Fish and Boat Commission (PAFBC)	Tyler Wagner	1995 – 2013	7,203
New York State Department of Environmental Conservation	Tyler Wagner	1995 – 2007	4,565

(NYDEC)			
Maryland Biological Stream Survey (MBSS)	N/A	1995 – 2001	3,081
Virginia Department of Game and Inland Fish (VADGIF)	N/A	2001 – 2010	611
Virginia Department of Environmental Quality (VADEQ)	Tyler Wagner	1995 – 2012	454
West Virginia Department of Environmental Protection(WVDEP)	Tyler Wagner	1997 – 2010	245
WVDEP	Petty	2006 – 2012	43
West Virginia Division of Natural Resources (WVDNR)	Petty	2001 – 2010	25
West Virginia Stream Classification Survey	Petty	2004 – 2009	21
Regional Environmental Monitoring and Assessment Program (REMAP)	Petty	2001	5
Lara Hedrick	Petty	2002	3
Mid-Atlantic Integrated Assessment (MAIA)	Petty	1997	3
Environmental Monitoring and Assessment Program (EMAP)	Petty	1997	1
National Rivers and Stream Assessment (NRSA)	Petty	2009	1

Note – Data source indicates the original collector of the data, and provider indicates the party that provided data directly to DS. Permission for data use was secured with collecting agencies when data was not publically-accessible.

FIGURE 2: BROOK TROUT MODELING AREA AND SAMPLING SITES



Final BRT Model

Model Details

As described previously, we used BRT to develop a predictive statistical model for brook trout at the 1:100k catchment scale throughout the historic range of brook trout within the Chesapeake Bay watershed. Portions of the Chesapeake Bay watershed outside of the species historic range were not included in our analysis. We utilized the default settings for model building for most options, including using a 10-fold cross validation procedure and bag fraction = 0.75. Tree complexity (interaction depth) was set to 1 (this setting is necessary to ensure proper stress and natural quality calculations) and learning rate was set at 0.01. Learning rate was chosen after examination of holdout deviance plot produced from the BRT model, and ensuring the model did not come to resolution too quickly or too slowly. The final selected model was comprised of 4,450 trees. This model was created using 90% of the available response data (n = 2,949), with the remaining 10% held out for later independent model validation.

Modeled stream temperature (DeWeber and Wagner 2014), which represents a natural habitat quality variable, was the single most important predictor variable in the model with a relative influence of 43% (Table 3). Stream temperature was also among the most important predictor variables in the DeWeber and Wagner (2015) model predicting brook trout occupancy range-wide. While some factors influencing stream temperature could be considered anthropogenic in nature (e.g. riparian forest cover), air temperature was the dominant factor influencing stream temperature (DeWeber and Wagner, 2014). Given this, we chose to consider stream temperature a natural habitat quality variable. In the subsequent Climate Assessment section, we analyze changes in stream temperature due to forecasted changes in air temperature.

The next most important predictors of brook trout occupancy were both anthropogenic stressors: mean network imperviousness and network percent agricultural land cover, which had relative influences of 22% and 9.7%, respectively. Agricultural land cover was among the most important variables for the DeWeber and Wagner (2015) model predicting brook trout occupancy and for the model predicting brook trout population status for subwatersheds in Hudy et al. (2008). Predictive models created for brook trout in the Great Lakes (Downstream Strategies, 2015) and Driftless (Downstream Strategies, 2012) regions indicated that agriculture, development, and impervious surfaces all acted as stressors as well.

The final anthropogenic predictor variable in the model was percent of upstream network mined. Other indices and metrics of mining intensity have been shown previously to impact stream ecosystems, and have been used in predictive models or to predict future land use changes (Petty et al. 2010, Merovich et al. 2013, Merriam et al. 2013). Petty et al. (2010) found that landscape indices related to mining had negative relationships with water quality, habitat quality, and macroinvertebrate communities in streams. Merriam et al. (2013) found that deep and surface mining intensity had impacts on water quality and macroinvertebrates, and used projected future changes in mining intensity to predict the impact from future land use scenarios on the stream ecosystem. Merovich et al. (2013) used mining indices as predictor variables in a BRT model to successfully predict stream water chemistry and macroinvertebrate biotic integrity. From examination of the function plot in Figure 13, it seems that the mining metric used in our model acts as a stressor at lower levels (i.e. brook trout likelihood of occurrence is highest at zero mined area and reduces as mined area increases). This relationship breaks down at very high percentages of mined area, which is likely caused by a relatively small sample from areas with high percentages of mined area.

TABLE 3: RELATIVE INFLUENCE OF ALL VARIABLES IN THE FINAL BROOK TROUT MODEL

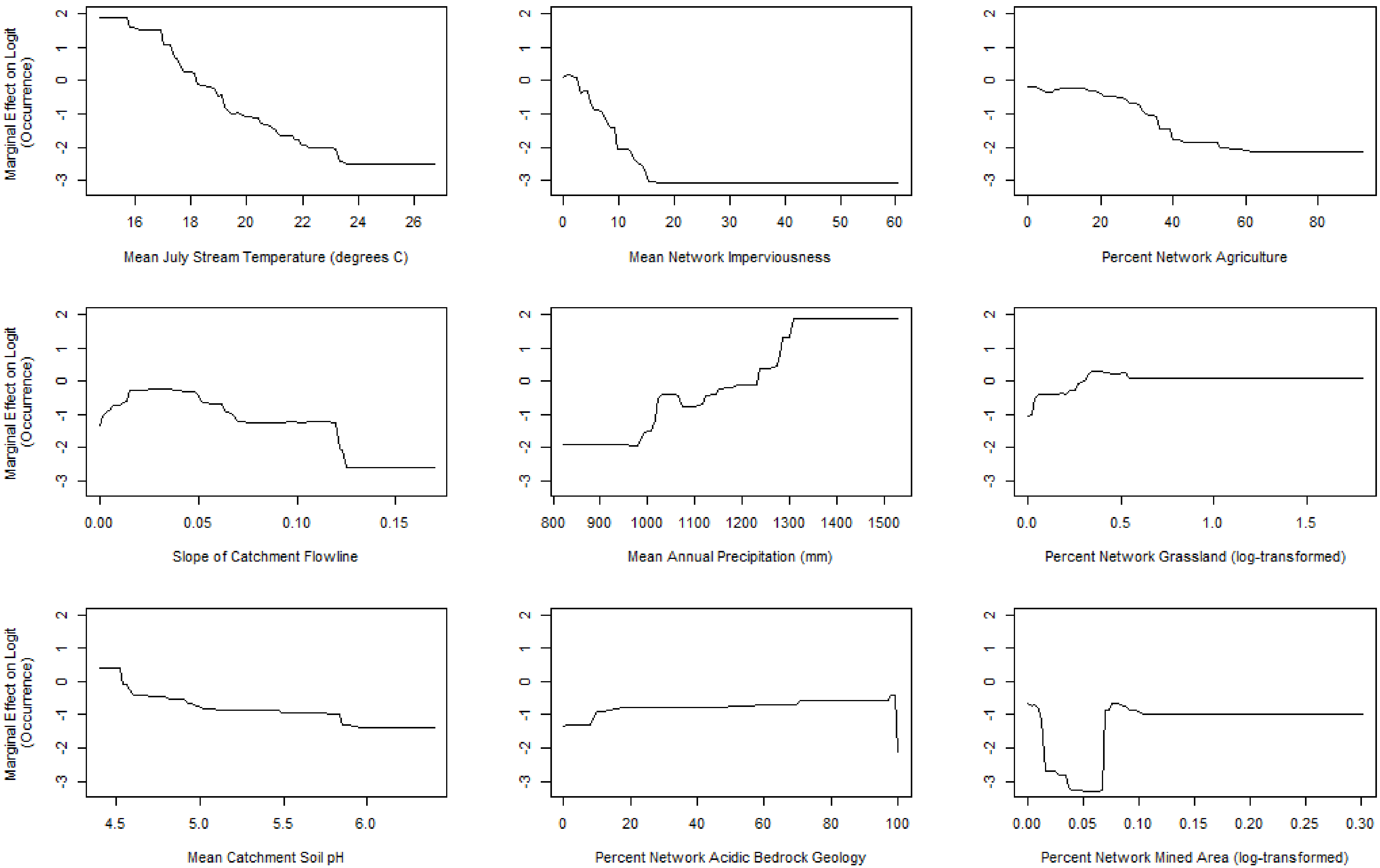
Variable Name	Variable Description	Relative Influence	Type of relationship
mnjuly	Mean July Stream Temperature (predicted)	42.67	Negative
IMP06C	Mean network imperviousness	21.59	Negative
Ag_pc	Network percent agricultural landcover	9.71	Negative
SLOPE_fix	Slope of catchment flowline	7.49	Variable
Precip	Mean annual precipitation	6.58	Positive

Log_Grass_pc	Log of network percent grassland cover	2.57	Positive
SoilpH	Catchment soil pH	2.53	Negative
Acid_geol_pc	Network percent acidic geology	2.51	Variable
Log_past_minepc	Log of network percent past mining areas	2.28	Negative
Log_Wet_pc	Log of network percent wetland cover	2.08	Variable

Note: Individual variables are highlighted according to whether they were determined to be anthropogenic (gray shading) or natural (no shading). Negative relationships indicate that general trends show that as the predictor increases, the likelihood of brook trout decrease. Positive relationships indicate the general trend is that likelihood of brook trout increases as the predictor variable value increases.

The function plots for the model, which show the marginal effect on the response variable (logit(p)) (y-axis) as the predictor variable (x-axis) changes, are shown in Figure 3 for the nine most influential variables in the brook trout model (Table 3). The dash marks at the top of each function represent the deciles of the data used to build the model. The plots for all 10 variables are shown in Appendix B.

Figure 3: Functional responses of the dependent variable to individual predictors of brook trout



Note: Only the top nine predictors, based on relative influence are shown here. See Appendix B for plots of all predictor variables.

Model Validation

The model had a CV correlation statistic of 0.759±0.008 and a CV ROC score of 0.929±0.005 and it explained 55% of the deviance in the response data. The remaining 10% of the available response data (n = 333) was held out to perform independent testing. We assessed the misclassification rates on this dataset utilizing five thresholds to indicate presence and absence. These thresholds represent several commonly used thresholds, and also one value we used to provide further summary (average of sensitivity = specificity and maximum kappa). Total misclassification rates ranged from 18.0% – 18.9%, commission error ranged from 7.5% - 10.2%, and omission error ranged from 8.1% - 11.4% (Table 4).

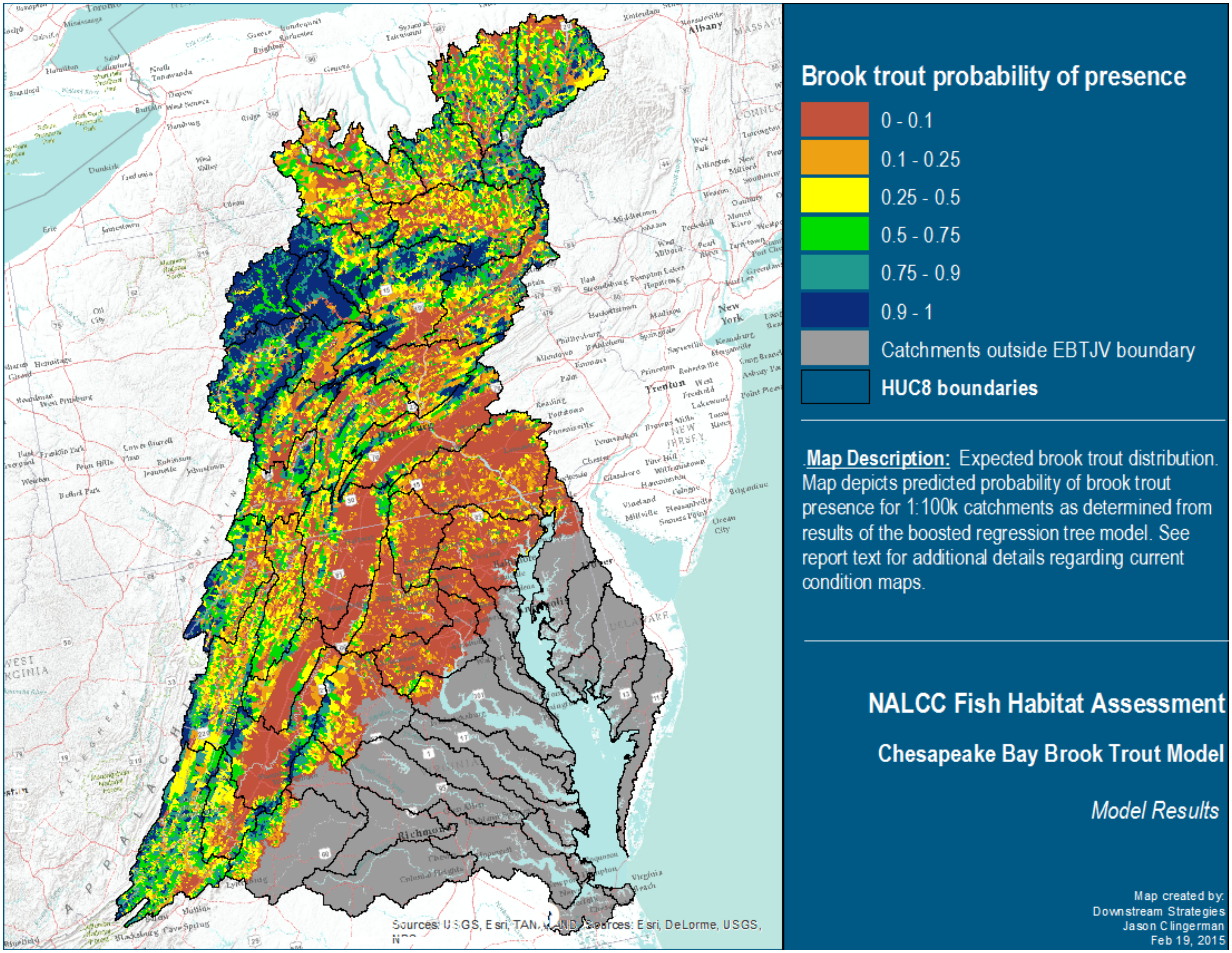
TABLE 4. MODEL MISCLASSIFICATION RATES

Threshold	Threshold Justification	False Positive Rate	False Negative Rate	Total Error Rate
0.38	Sensitivity = Specificity	10%	8%	18%
0.4	Training data prevalence	10%	9%	19%
0.42	Average of Sens=Spec & MaxKappa	9%	9%	18%
0.44	Maximum Kappa	9%	9%	18%
0.5	Maximum Percent Correction Classified (PCC)	8%	11%	19%

Map of current brook trout occupancy

Brook trout probability of presence was calculated for all 1:100k stream catchments in the study area using the BRT model. The predicted probability of presence ranged from 0 to 1, where 0 = absent and 1 = 100% probability of presence. The mean predicted probability was 0.33. Of the total 51,474 catchments in the Chesapeake Bay watershed and also within the historic brook trout distribution range, there were 9,605 catchments with a predicted probability of presence greater than 0.75 and 6,279 catchments where the probability of presence was between 0.5 and 0.75. These results are mapped in Figure 4.

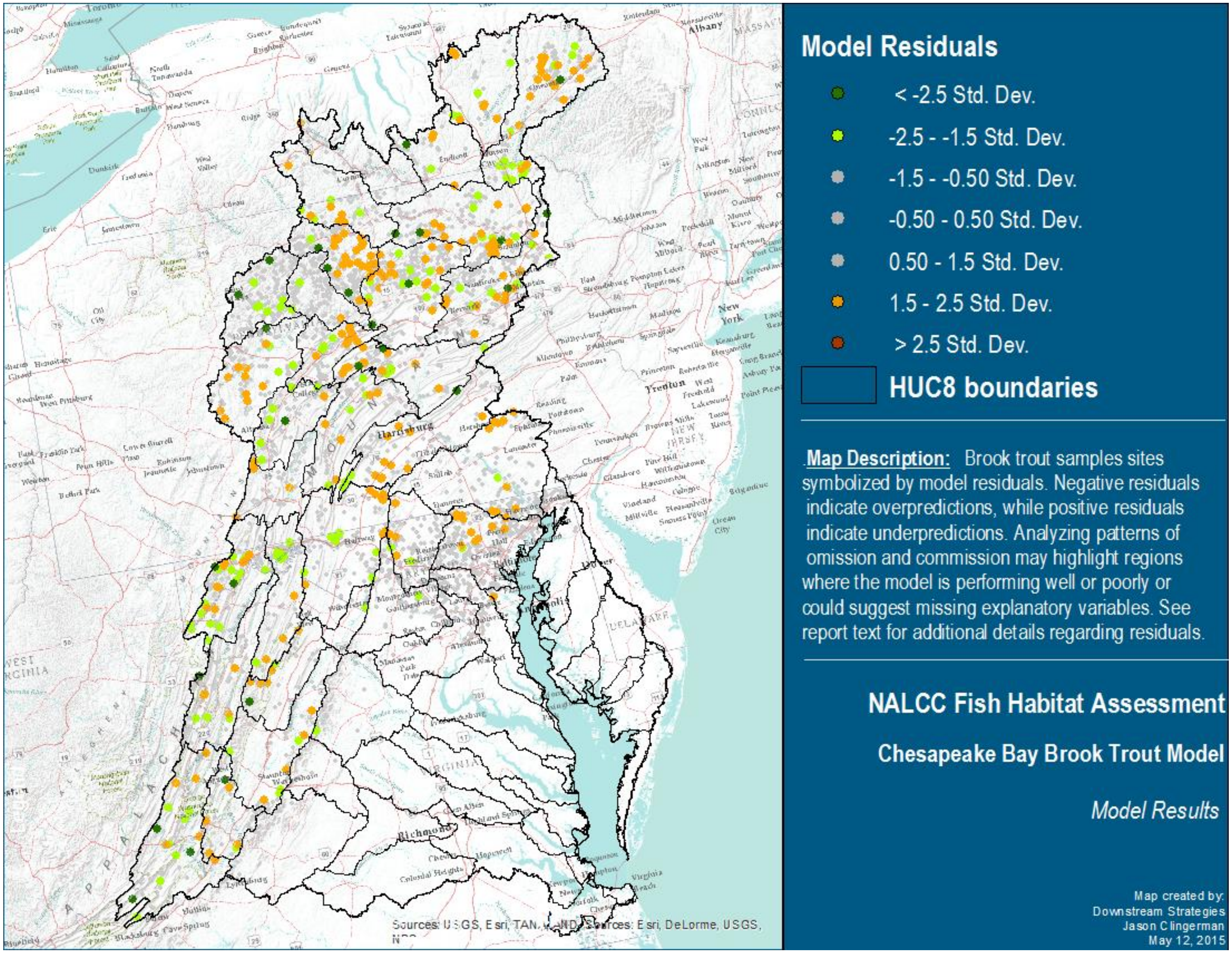
FIGURE 4: EXPECTED BROOK TROUT DISTRIBUTION



Spatial Residuals

The spatial distribution of residuals is shown in Figure 5. The results from the Mantel test, showed a simulated p-value of 0.001 and an r=0.103. This indicates that the residual values are spatially autocorrelated, which means that sample locations near one another are more likely to have more similar residual values than sample locations with a larger geographic distance between them. The p-value and r-value indicate that this relationship is statistically significant at the $p = 0.001$ level. The clustering of residuals suggests that there may be additional landscape scale information that could be used to improve the predictive power of the BRT model.

FIGURE 5: DISTRIBUTION OF BROOK TROUT MODEL RESIDUALS BY SAMPLING SITE



Anthropogenic Stress and Natural Habitat Quality

The variable importance table and partial dependence functions of the final BRT model were used to assess the potential stressors for the brook trout model. Within the model, there were three variables considered anthropogenic in nature (Table 3). These three stressors, network mean imperviousness (IMP06C), network agriculture land cover (Ag_pc) and log-transformed network percent past mining (Log_past_minepc), were used to calculate ASI for the brook trout model. See the 'Overview' section for details on how ASI and HQI were calculated for each model.

Maps of HQI and ASI illustrate the spatial distribution of natural habitat potential (i.e., HQI score) and anthropogenic stress (i.e., ASI score) in the Chesapeake Bay watershed. HQI and ASI scores are mapped in Figure 6 and Figure 7, respectively. The three metrics contributing toward the calculation of ASI are mapped in Figure 8, Figure 9, and Figure 10. HQI, ASI, and their metrics are all scaled on a 0-1 scale (see Overview section for more details on HQI and ASI calculation). For HQI, higher values indicate higher natural quality, while higher values for ASI indicate higher levels of anthropogenic stress.

Note that the stress values are not simply a measure of anthropogenic changes to the watershed, but also how much those changes are impacting brook trout. If an area was naturally unsuitable for brook trout (i.e. low natural quality index score), the stress index will also be low even if stressors are present in the area. In other words, stress can only be high if the natural habitat quality index is high. If natural habitat quality is so low that brook trout would likely be absent independent of stress, then the stress index is necessarily low as well. It is likely that stress on aquatic systems in general is much more widespread than is indicated in this model created specifically for brook trout. For all stress and natural quality indices, all catchments are shown, even in areas where the probability of presence is low. This is necessary and useful to consider areas outside of the current expected range where stress could have caused a historic population to be extirpated.

Figure 6: Habitat quality index for brook trout

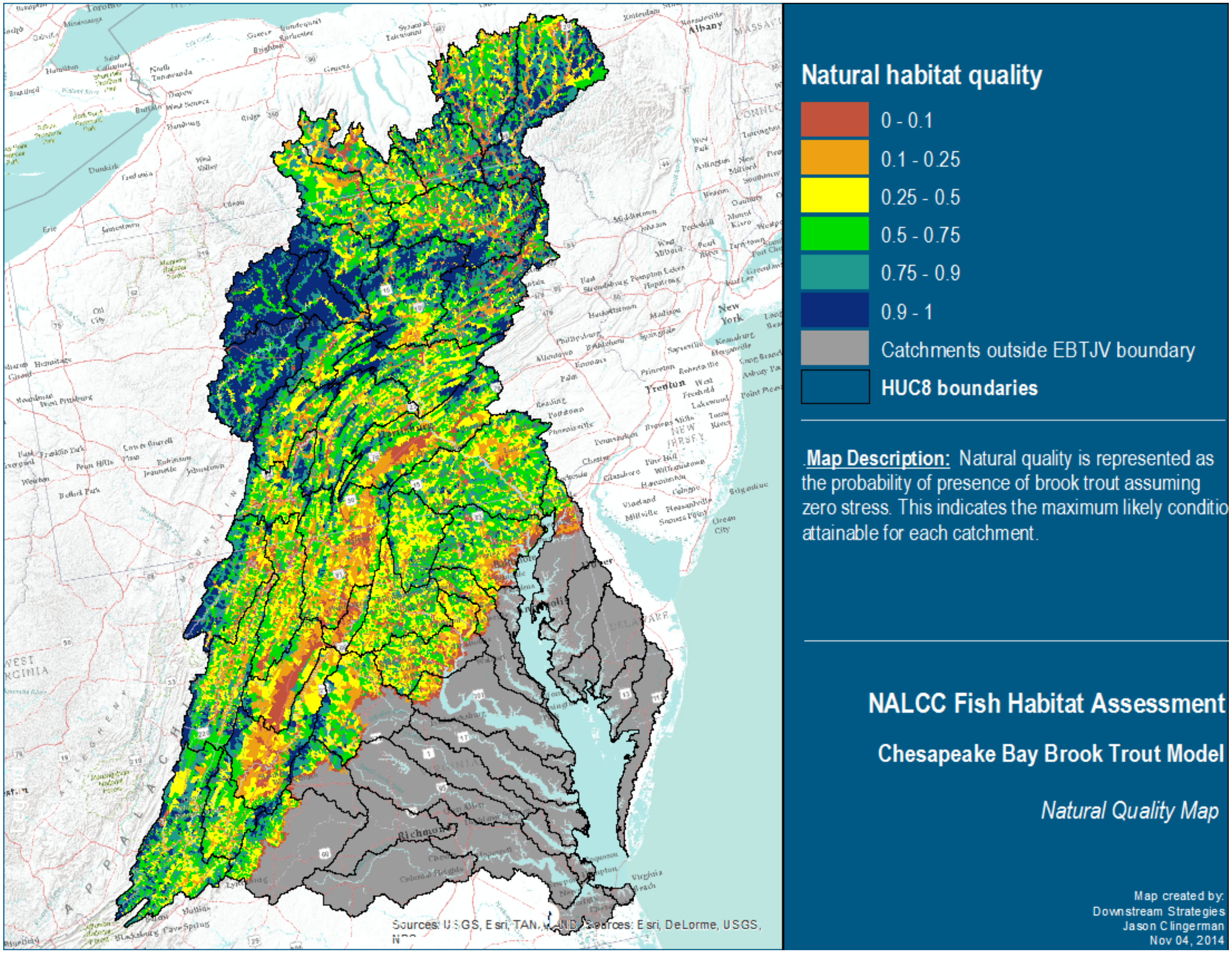


FIGURE 7: TOTAL ANTHROPOGENIC STRESS INDEX FOR BROOK TROUT

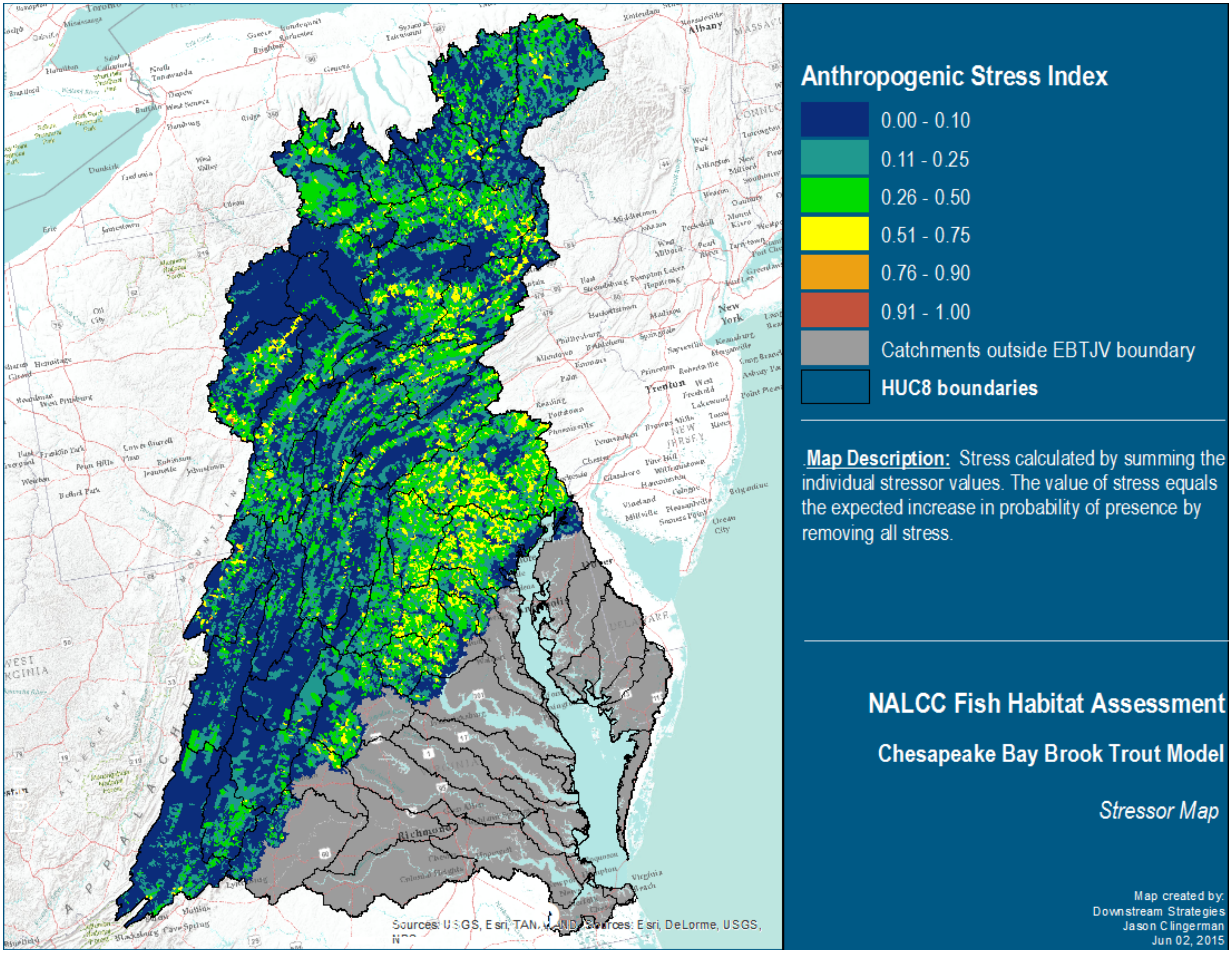
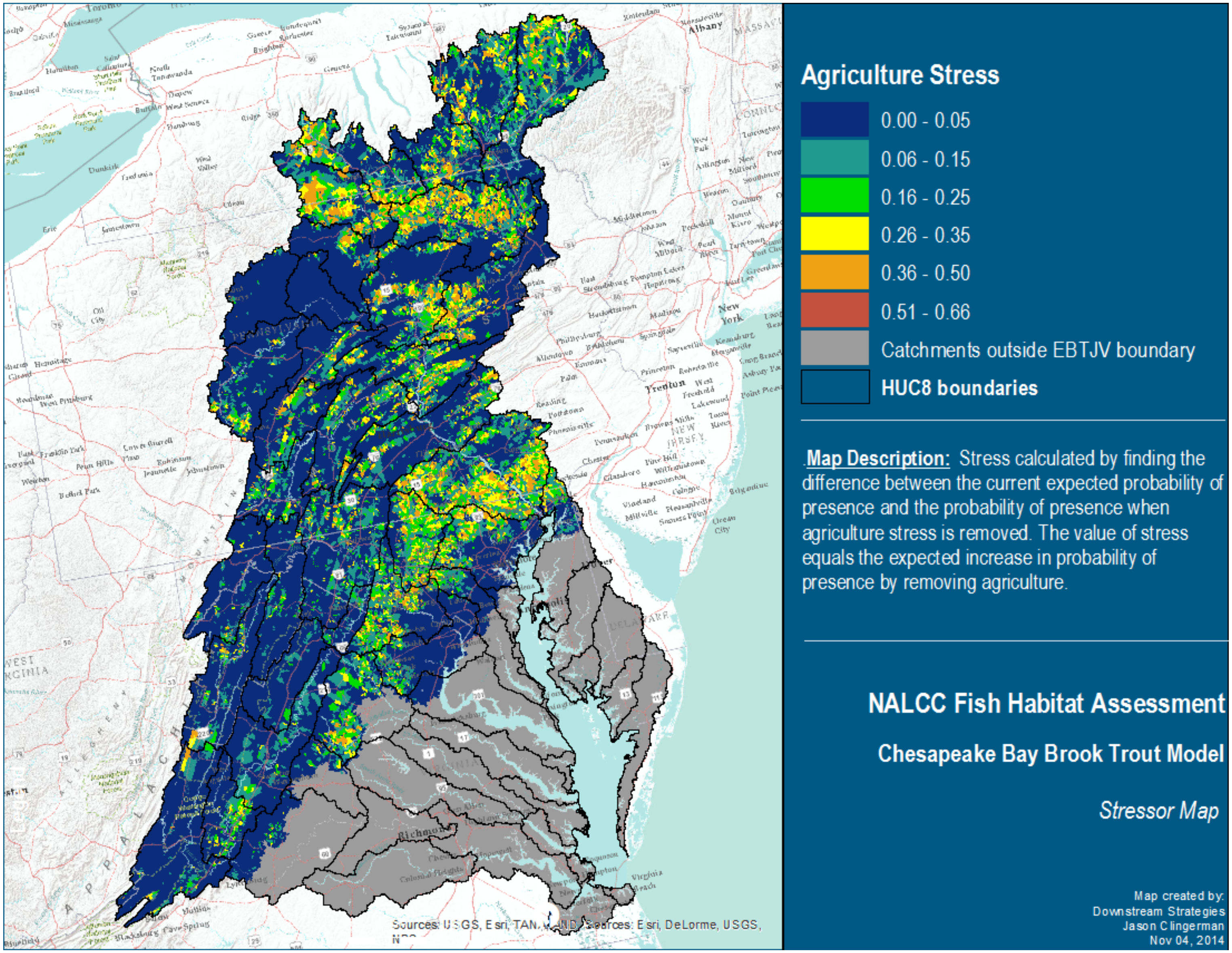
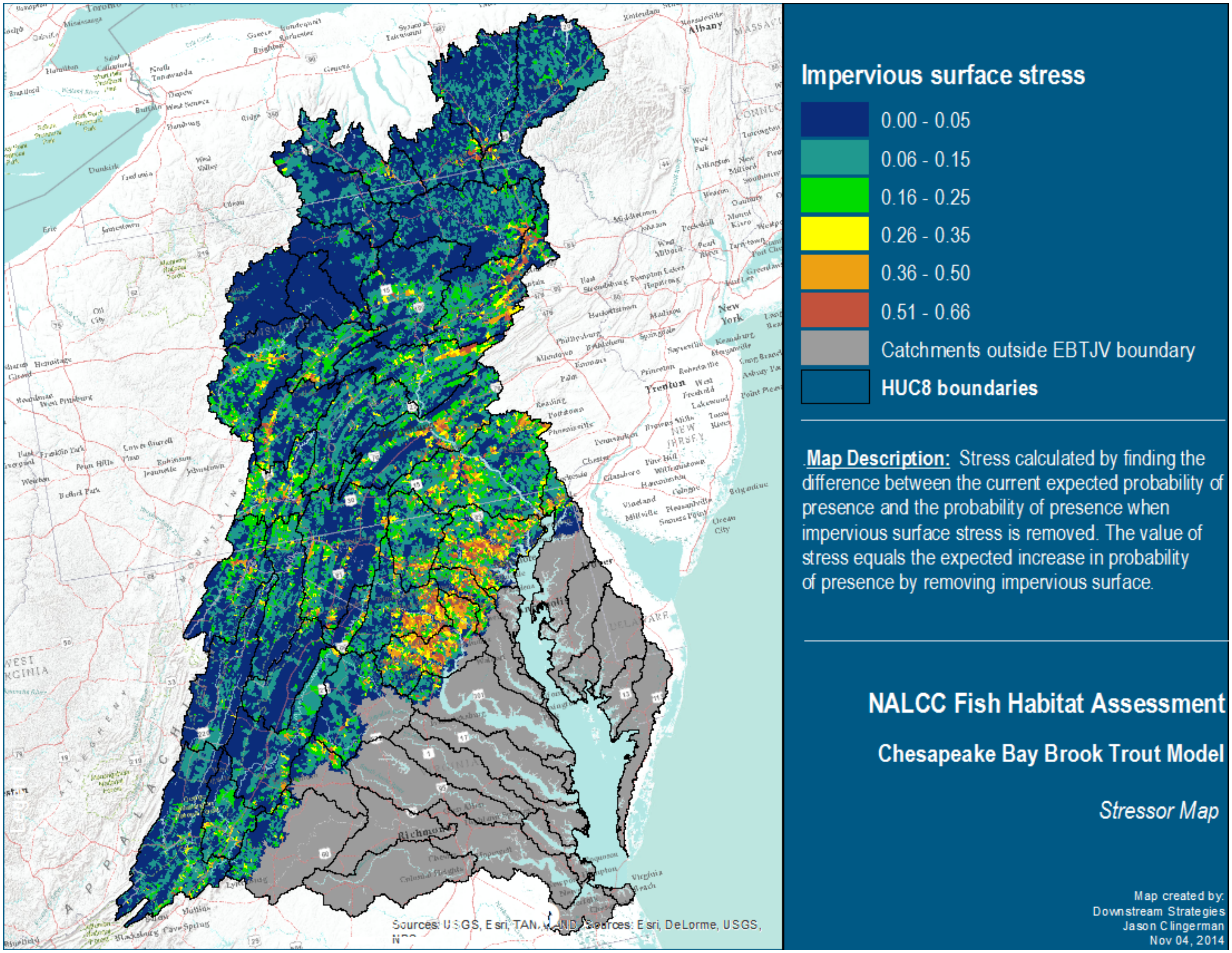


FIGURE 8: MOST INFLUENTIAL ANTHROPOGENIC STRESS INDEX METRIC FOR BROOK TROUT



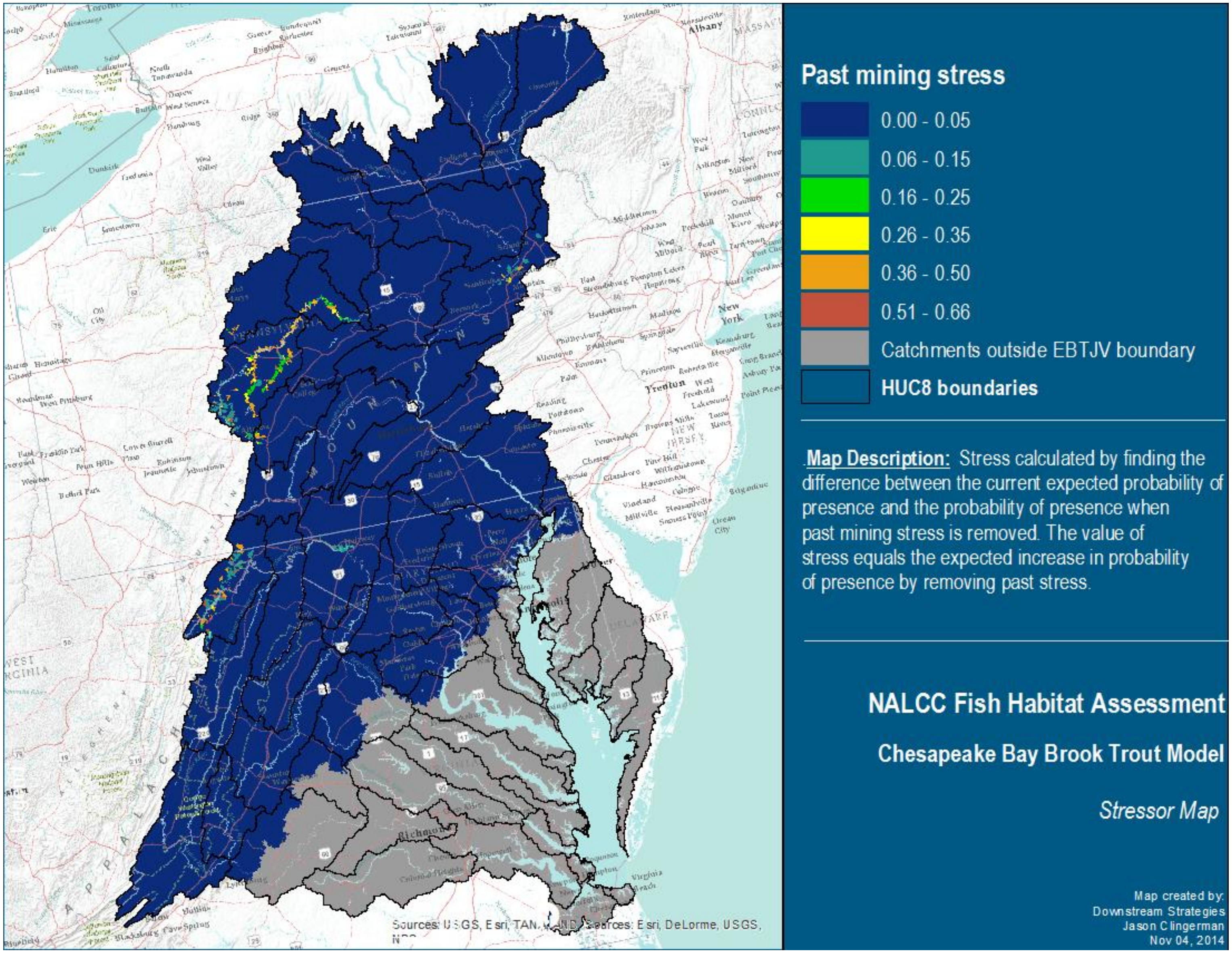
NOTE: "MOST INFLUENTIAL" REFERENCES THE RELATIVE INFLUENCE SCORES FROM THE BRT MODEL OUTPUT.

FIGURE 9: SECOND MOST INFLUENTIAL ANTHROPOGENIC INDEX METRIC FOR BROOK TROUT



NOTE: "MOST INFLUENTIAL" REFERENCES THE RELATIVE INFLUENCE SCORES FROM THE BRT MODEL OUTPUT.

FIGURE 10: THIRD MOST INFLUENTIAL ANTHROPOGENIC INDEX METRIC FOR BROOK TROUT



NOTE: "MOST INFLUENTIAL" REFERENCES THE RELATIVE INFLUENCE SCORES FROM THE BRT MODEL OUTPUT.

Comparison of occupancy predictions with EBTJV and Wagner model

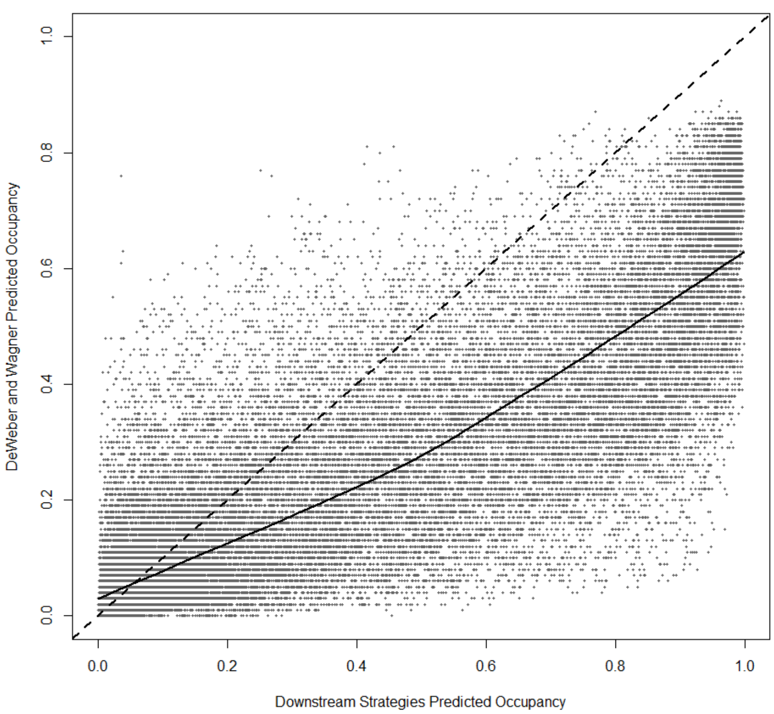
Our predicted occupancy values are comparable to the DeWeber and Wagner (2015) predictions for the same region. Figure 11 shows a scatter plot of predicted occupancy of each methodology plotted against one another, where R^2=0.71 when fitting a linear trendline (not shown). Within Figure 11, the points are plotted along with a smoothed loess line to fit the points, and dashed line where y=x, which would indicate the trend expected if both methodologies predicted equal occupancies for each catchment. The trends are quite similar, but the DeWeber and Wagner (2015) predictions are generally less than our predicted occupancies.

When evaluating model accuracy for catchments with predictions for both models, we found that our model's total error rate was lower than the DeWeber and Wagner (2015) total error rate when using a presence-absence threshold equal to training data prevalence for each model. The DeWeber and Wagner (2015) model's false-positive rate was very good, at approximately 1.7%, but the false-negative rate was 22.6%. Our model showed a more balanced error distribution, with 8.6% and 6.1% false-positive rates and false-negative, respectively. This indicates that for the Chesapeake Bay watershed, the DeWeber and Wagner (2015) model is very conservative and seems to consistently under-predict brook trout occurrence rates.

Some of this variation between values and predictive capacity is likely due to the different predictor variables utilized or differences in statistical structure between BRT and hierarchical bayesian regression. A recent study comparing the predictive power of differing modeling approaches found BRT models to be powerful in similar applications (Fleishman et al. 2014).

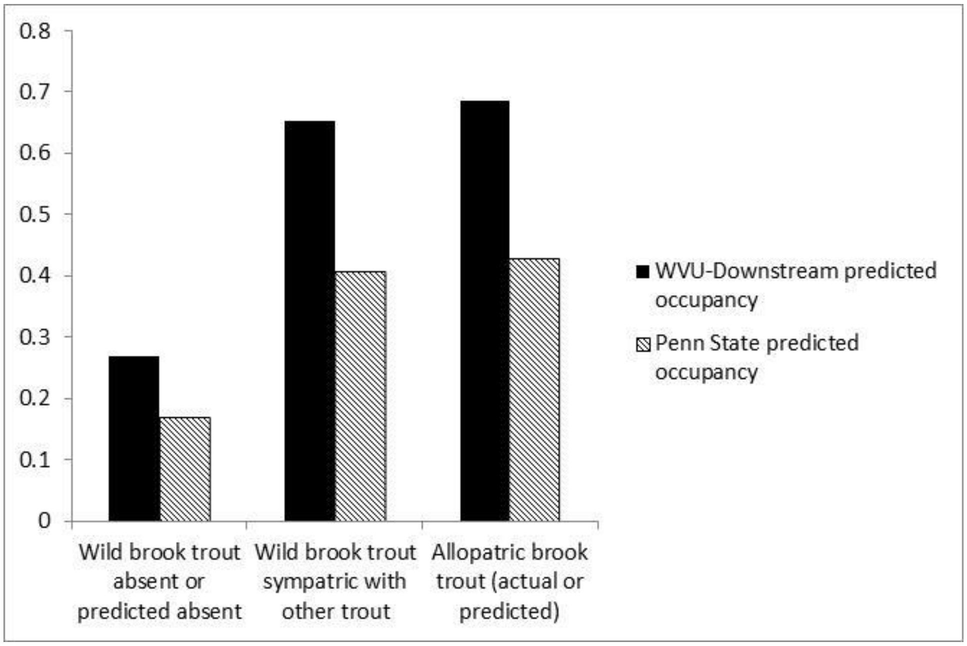
A portion of the remainder in the variation is likely do to the differing extents to which each model was created. The DS model extent was the Chesapeake Bay watershed, whereas the DeWeber and Wagner model extent was the entire EBTJV geographic range. It seems likely that influences outside of the focal Chesapeake Bay watershed may be causing higher variation and lower predictive accuracy within the Chesapeake Bay watershed for the DeWeber and Wagner model.

FIGURE 11. COMPARISON WITH PSU PREDICTIONS



Likewise, our predictions show reasonable agreement with EBTJV (2015) classifications. Our predictions of occupancy average 0.7 in catchments classified as allopatric brook trout populations, 0.65 in sympatric brook trout populations, and 0.25 in catchments classified as absent (Figure 12).

FIGURE 12. COMPARISON TO EBTJV CLASSIFICATIONS



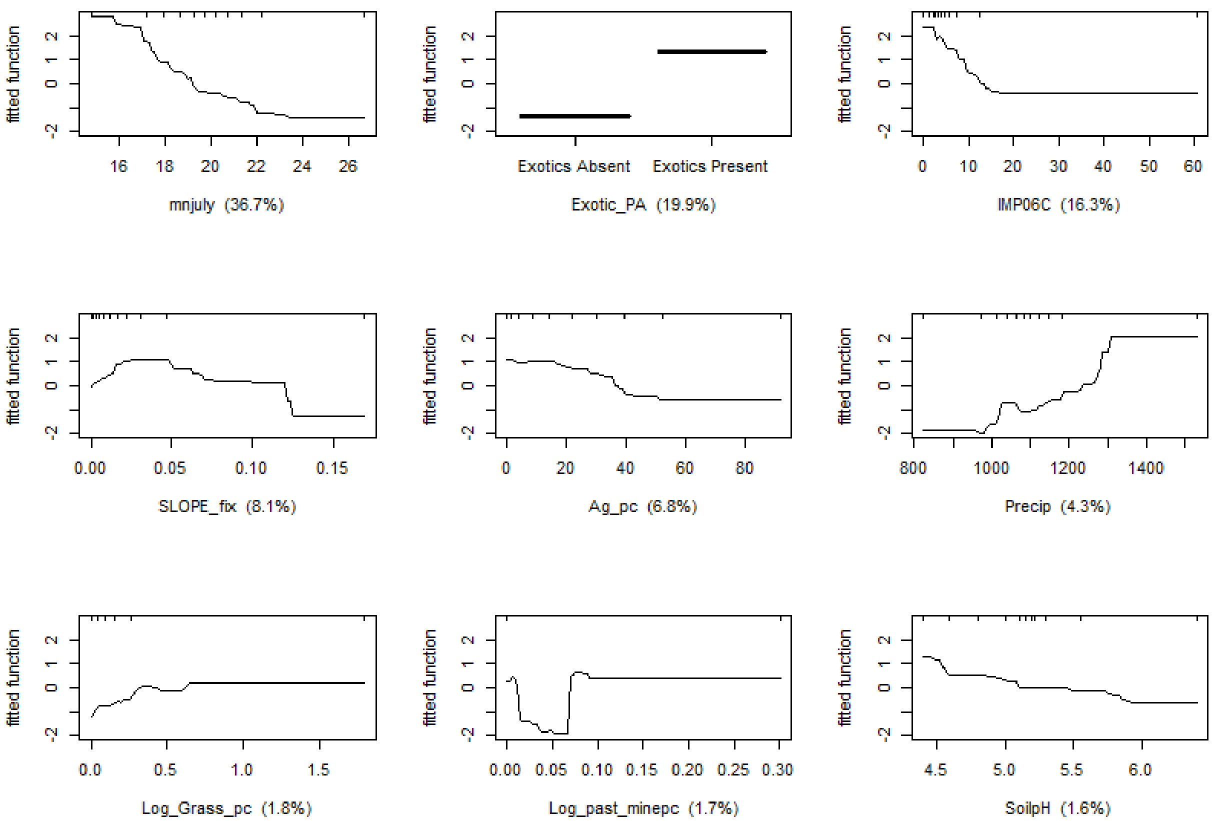
Influence of exotic trout on brook trout occurrence

The presence of exotic trout species has been indicated by experts to be a major threat to brook trout across their eastern range (EBTJV, 2006). The EBTJV (2006) report indicates that professionals deemed that exotic trout were a major stressor in Pennsylvania and New York, but were not identified in Maryland, Virginia or West Virginia as a major threat. In developing our final BRT occupancy model for brook trout, we decided against using exotic trout as a predictor variable for several reasons. First, continuous information on exotic trout throughout the Chesapeake Bay watershed does not exist. Consequently, a model that includes exotic trout cannot be used to predict brook trout occurrence continuously across the Bay watershed. Second, trout biologists from across the region disagree on the extent to which the presence of exotic trout actually serves as a stressor to brook trout and/or influences their current distributions (EBTJV, 2006). Third, because all trout share similar habitat characteristics, it is likely that brook trout and exotic trout distributions are highly correlated with similar landscape attributes (e.g., water temperature, forest cover, land use). Consequently, a model that includes exotic trout may influence underlying relationships between brook trout and natural habitat variables.

Nevertheless, in order to quantify the potential effects of exotic trout presence on brook trout occupancy, we conducted a sequence of post-modeling analyses. First, we used the EBTJV (2015) classification strategy to indicate presence or absence of exotic trout species, and constructed another BRT model using this information as a predictor variable. Although the resulting model could not be extrapolated to all unsampled catchments, it would provide a quantitative measure of the influence of exotic species on brook trout distributions at the scale of the Chesapeake Bay watershed.

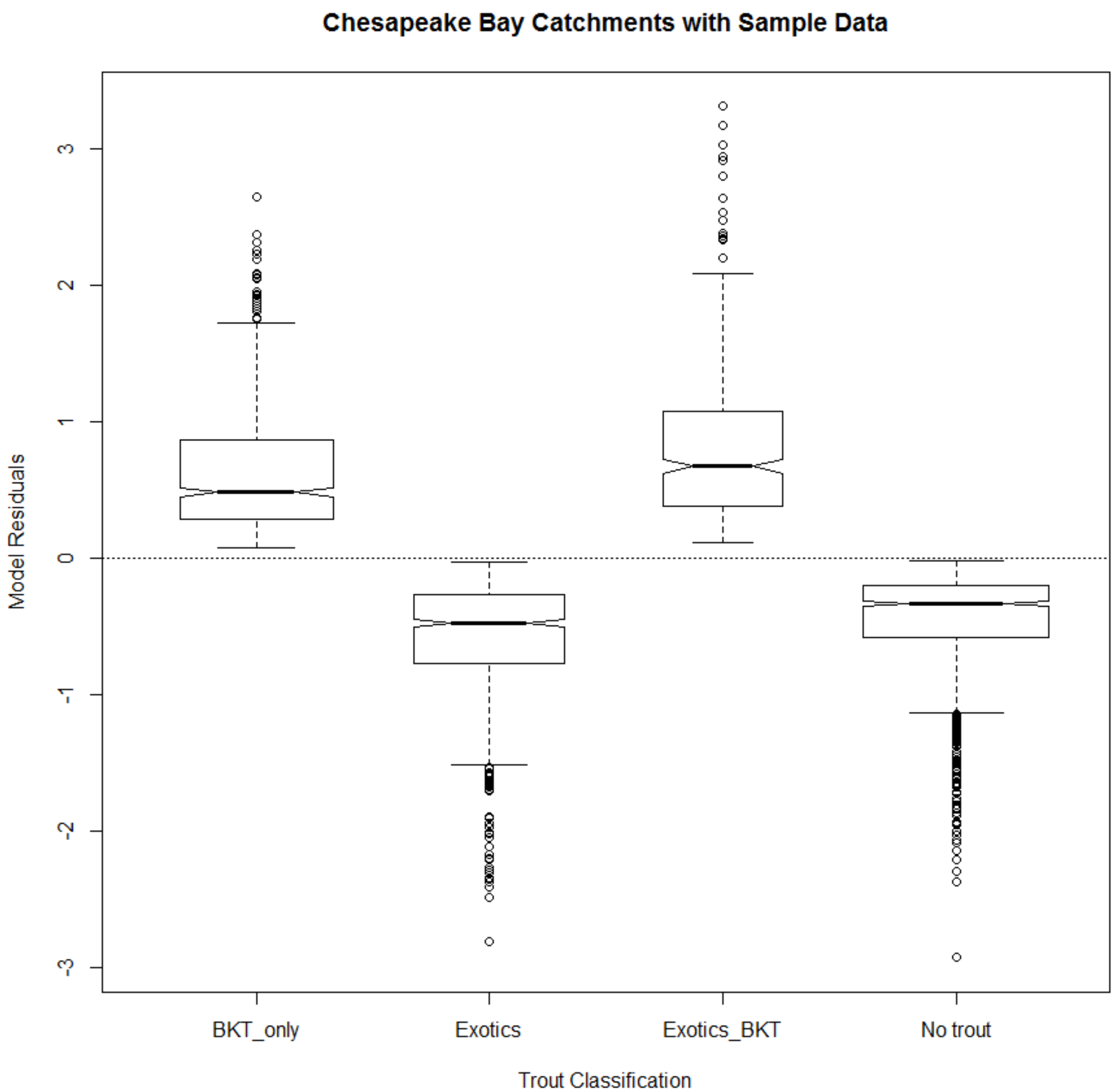
This new BRT model indicated that information on exotic trout does provide substantial additional explanatory power (Figure 13). However, as expected, we found that the statistical effect of exotics is an increased likelihood of brook trout presence. This is likely due to the fact that all trout species have shared habitat requirements. The curves of the other predictor variable's function plots remained relatively unchanged.

FIGURE 13. FUNCTION PLOTS FOR TOP NINE PREDICTOR VARIABLES WHEN INCLUDING EXOTIC TROUT PRESENCE AS A PREDICTOR VARIABLE



As a second analysis, we examined the relationship between the occurrence of exotic trout and the residuals from our original BRT model (i.e., the model that did not utilize exotic trout presence as a predictor). Highly negative residuals are indicative of areas where the model is "over-predicting" brook trout occurrence (i.e., predicting a high probability of occurrence but brook trout were absent). In contrast, large positive residuals are indicative of areas where the model is "under-predicting" brook trout occurrence (i.e., predicting a low probability of occurrence, but brook trout are present). If exotic trout are having a significant negative effect on brook trout occurrence, then we would expect sites classified as exotic trout only to have strongly negative residuals and for those residuals to be substantially more negative than sites classified as "no trout."

FIGURE 14. BOXPLOT OF MODEL RESIDUALS BY TROUT PRESENCE/ABSENCE



The results of this analysis indicate that areas classified as "exotic trout only" have residual patterns that are similar to those from areas classified as "no trout" (Figure 14), but when using ANOVA with a post-hoc Tukey test to determine difference in means, we find that there is a small, but statistically significant (p-value < 0.01) difference in the mean residual between the "exotic trout only" and "no trout" groups. This suggests a measurable negative effect of exotic trout on brook trout occupancy. Nevertheless, the actual difference between no trout and exotic only sites is very small, which suggests that numerous factors, including exotic trout, may be responsible for brook trout absences in high quality habitats. Other factors include: 1- undetected or unmapped stressors (e.g., abandoned underground mines), 2- isolation from core brook trout populations, 3- dispersal barriers, 4- brook trout detection errors, or 5- localized errors in predictions of water quality. An important goal of future work should be to identify and quantify additional factors that may be affecting the predictive power of our current model, including exotic trout, and incorporating this information into the next model iteration.

Our combined analysis (secondary BRT model plus residual analysis) suggests that the absence of information on exotic trout likely is not systematically affecting the explanatory / predictive power of our original BRT model. This is not to say that exotic trout cannot, or are not, having negative effects brook trout populations. There may be effects of exotic trout on brook trout abundance, or there may be localized effects on brook trout occurrence. Nevertheless, there is little evidence that the effect on brook trout occurrence is so widespread as to undermine the application of the original BRT model at the scale of the Chesapeake Bay watershed.

The results of this analysis further suggest that the best way for stakeholders to use information on exotic trout is within the context of the ranking tool. Stakeholders may very well want to prioritize conservation areas based on the presence or absence of exotic trout populations. This can easily be done within the ranking tool that we will provide. It is unnecessary for this information to be included in the occupancy model to be of value within the decision support tool.

Climate Assessment

Objectives/Introduction

For a coldwater obligate species such as brook trout, the impact of potential climate change is expected to shift and alter their distribution across the landscape (Comte et al., 2013; Hickling et al., 2006). In this assessment we quantified the anticipated resiliency and vulnerability to climate change for brook trout in the Chesapeake Bay watershed. These analyses will aid future restoration and protection priorities for brook trout, and can be considered alongside other factors such as current stress and current habitat quality to provide expectations of brook trout populations into the future.

Our assessment is based on large-scale climatic factors, including mean annual precipitation and mean annual or seasonal temperatures and assumes the current relationships between habitat and brook trout occurrence will persist into the future. Impacts resulting from changes in frequency or severity of individual storm events are beyond the scope of this assessment. It is also important to note the population parameter of interest in this analysis, brook trout occupancy. Impacts to population structure and dynamics resulting from climatic changes are also beyond the scope of this effort. An example of such an impact would be a severe localized flood event that does not change brook trout occupancy, but causes a shift in population structure because of high juvenile mortality.

Methods

Data

In our analysis, two predictor variables from the model described above were altered to capture potential future changes in climate: mean July stream temperature and mean annual precipitation. These two variables were available for several future climate scenarios. All projected future climate information was based on the Intergovernmental Panel on Climate Change (IPCC) A2 emissions scenario (Intergovernmental Panel on Climate Change. 2007), but several downscaled regional climate models (Hostestler et al., 2011) were used to represent more or less optimistic potential future climate conditions. All climate scenarios used for projections and their details are presented below in Table 5.

The mean July stream temperature used as a predictor variable in the original BRT model described above was produced as a result of the work of DeWeber and Wagner (2014). The authors of that study also produced future stream temperature predictions for a collection of climate models and future time frames (DeWeber and Wagner, unpublished data), and those data were used here to as a predictor variable for future scenarios. The stream temperature and annual precipitation data for each timeframe was averaged across a five year period that centered on the focal year.

Mean annual precipitation used as a predictor variable for the predictive model was compiled as part of the NHD plus datasets (Horizon Systems, 2012) for each catchment. This data was originally sourced from the Parameter-elevation Regressions on Independent Slopes Model (PRISM). For future mean annual precipitation projections, we utilized predicted mean annual precipitation from downscaled climate projections (Thrasher et al. 2013) from the time frames that matched those used for the projected stream temperatures. The Thrasher et al. (2013) data aggregated 39 global climate models to produce a daily average precipitation rate for each 800 meter grid cell, and we averaged the data annually for each grid cell across the five year periods used in the future stream temperature projections (DeWeber and Wagner, unpublished data).

TABLE 5. CLIMATE SCENARIO DETAILS

Timeframe	Predicted Stream Temperature			Predicted Annual Precipitation	
	Emissions Scenario	Climate Model	Mean July Temp (°C)	Data Source	Mean Annual Precip (mm)
Current	N/A	N/A	19.5	PRISM/NHD+	1069
2042	A2	EH5	19.7	NASA 39-model Ensemble	1147
2042	A2	GFDL	19.8	NASA 39-model Ensemble	1147
2042	A2	GENMOM	19.5	NASA 39-model Ensemble	1147
2062	A2	EH5	20.6	NASA 39-model Ensemble	1144
2062	A2	GFDL	20.2	NASA 39-model Ensemble	1144
2062	A2	GENMOM	20.3	NASA 39-model Ensemble	1144
2087	A2	EH5	21.7	NASA 39-model Ensemble	1159
2087	A2	GENMOM	20.6	NASA 39-model Ensemble	1159

Future status, habitat quality and stress

Predicted probability of presence for brook trout for future climate scenarios was calculated in a manner similar to the post-modeling methodology described above where the predictor variables used in the model were manipulated – this time to replace current climate data with projected future climate data. Probability of presence was calculated for each of the eight climate scenarios identified above.

Using the methodology described above in Derivation of Anthropogenic Stress Index and Habitat Quality Index, we also recalculated stress and natural quality under each potential climate scenario. This allowed us to calculate the differences between stress and natural quality between current and future conditions. For each climate scenario, the difference between current values and values calculated using future climate predictions was a way to indicate the potential effects of future climate scenarios.

Defining resilience and vulnerability

Climate resiliency and vulnerability were determined by analyzing predicted losses or gains in natural quality resulting from climate change. Underlying natural quality is directly impacted by changes in climate such as water temperatures and precipitation. Analyzing changes in modeled natural quality indicates the anticipated impacts on brook trout occupancy. Areas anticipated to have reduced natural habitat quality index scores were determined to be vulnerable to future climate change scenarios, while resilient areas were expected to remain unchanged or increase in natural quality under future climate scenarios.

Results/Discussion

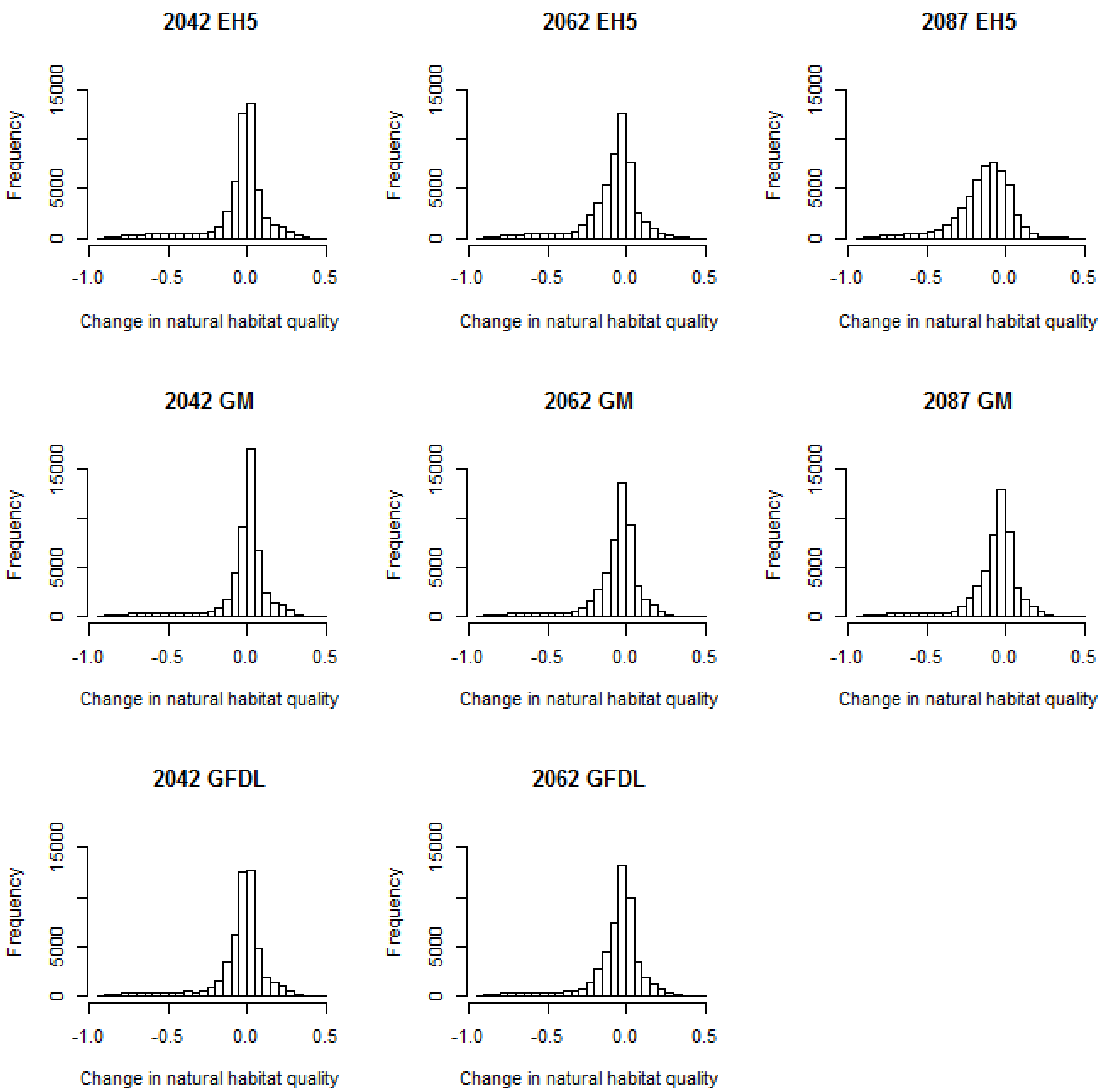
Watershed-wide results

This assessment produced a large amount of data, not all of which can be shown in a meaningful way within this report. All data produced will be available within the web-based decision support tool that will be part of the deliverables for this project, and also in stand-alone tables.

For this report, we have focused results and representative maps on two climate scenarios as examples of the data produced, both of which are from the 2062 time frame. This time frame is an actionable time frame (approximately 50 years), but is one far enough into the future that projected climate changes become more significant than projections at the 2042 time frame. The two downscaled regional climate models we will be presenting in this report are ECHAM5 (EH5) and GFDL. These were the two climate models that provided the greatest contrast in predicted temperature increases, with GFDL being more optimistic and EH5 being less optimistic. The third climate model analyzed here, GM, had results intermediate to EH5 and GM for the 2062 timeframe and are not shown in detail in this report.

A summary of the climate effects from each scenario is shown in the histogram in Figure 15. The effect of climate projections can be seen in the histograms of change in natural habitat quality for the EH5 model, as higher magnitude reductions in habitat quality are realized further out into the future. This trend is evident, but to a lesser extent in the GM model, while the GFDL model doesn't show much differentiation between the 2042 and 2062 scenarios (projections for this model were not available for the 2087 time frame).

FIGURE 15. CLIMATE EFFECT HISTOGRAMS FOR EACH CLIMATE SCENARIO.



A summary of mean values of the model results and stressors for each scenario and the current model outputs is shown below in

Table 6. This table shows that the most optimistic climate scenario varies depending on the timeframe assessed, at least when considering the entire study area. For the 2042 and 2087 timeframes, the GENMOM (GM) climate model is the most optimistic, but for the 2062 timeframe, GFDL is the most optimistic. Likewise, the least optimistic model for each timeframe also varied. The GFDL model was the least optimistic model for the 2042 time frame, while the EH5 model was the least optimistic for the 2062 and 2087 timeframes. For all future scenarios, predicted probability of presence and the natural quality index for brook trout was reduced compared to current predictions when analyzing the entire Chesapeake Bay watershed. Generally, stress values were also reduced, which would be anticipated with a reduction in natural quality and probability of presence since only areas with brook trout can be stressed utilizing our methodologies.

TABLE 6. CLIMATE SCENARIO RESULT SUMMARIES

Timeframe	Emissions Scenario	Climate Model	Mean prob. of presence	Mean natural quality	Mean impervious stress	Mean ag stress	Mean mining stress	Mean total stress
Current	n/a	n/a	0.336	0.512	0.081	0.069	0.003	0.153
2042	A2	EH5	0.307	0.471	0.077	0.065	0.003	0.144
2042	A2	GFDL	0.301	0.465	0.077	0.064	0.003	0.144
2042	A2	GENMOM	0.32	0.484	0.077	0.065	0.003	0.145
2062	A2	EH5	0.266	0.422	0.091	0.06	0.002	0.134
2062	A2	GFDL	0.283	0.444	0.074	0.062	0.003	0.139
2062	A2	GENMOM	0.28	0.439	0.073	0.062	0.003	0.138
2087	A2	EH5	0.227	0.37	0.065	0.052	0.002	0.12
2087	A2	GENMOM	0.276	0.434	0.072	0.061	0.003	0.136

The maps below (Figure 16 and Figure 17) show the change in natural habitat quality, which is the measure of climate effect for every catchment within the Chesapeake Bay watershed for the 2062 EH5 and GFDL models. For visualization purposes only, we classified changes into five categories. These categories were major decrease, minor decrease, no change, minor increase, and major increase. The minor categories were defined as a change between 0 +/- 0.20, and the major categories were defined as a change greater than +/- 0.20.

While the overall effect of future climate scenarios is negative for brook trout, there are specific regions that will be more resilient, and some are projected to have improved natural habitat quality. The areas identified to be resilient or improve show this result because of an increase in projected precipitation rates, which may moderate higher projected air temperatures or ameliorate effects from season low flow mortality. Figure 3**Error! Reference source not found.** (function lots from the original BRT model) illustrates the functional relationship between predicted probability of presence of brook trout and these two climatic factors, showing how increases in precipitation result in higher probabilities of occurrence.

While the overall patterns of increase and decrease are similar for both the EH5 and GFDL 2062 scenarios, it is apparent that the GFDL model is more optimistic of the two scenarios presented here. The areas projected to be resilient or improve are larger under the GFDL, and the most vulnerable areas are of higher intensity under the EH5 scenario.

FIGURE 16. CLIMATE CHANGE EFFECT FOR 2062 EH5 SCENARIO.

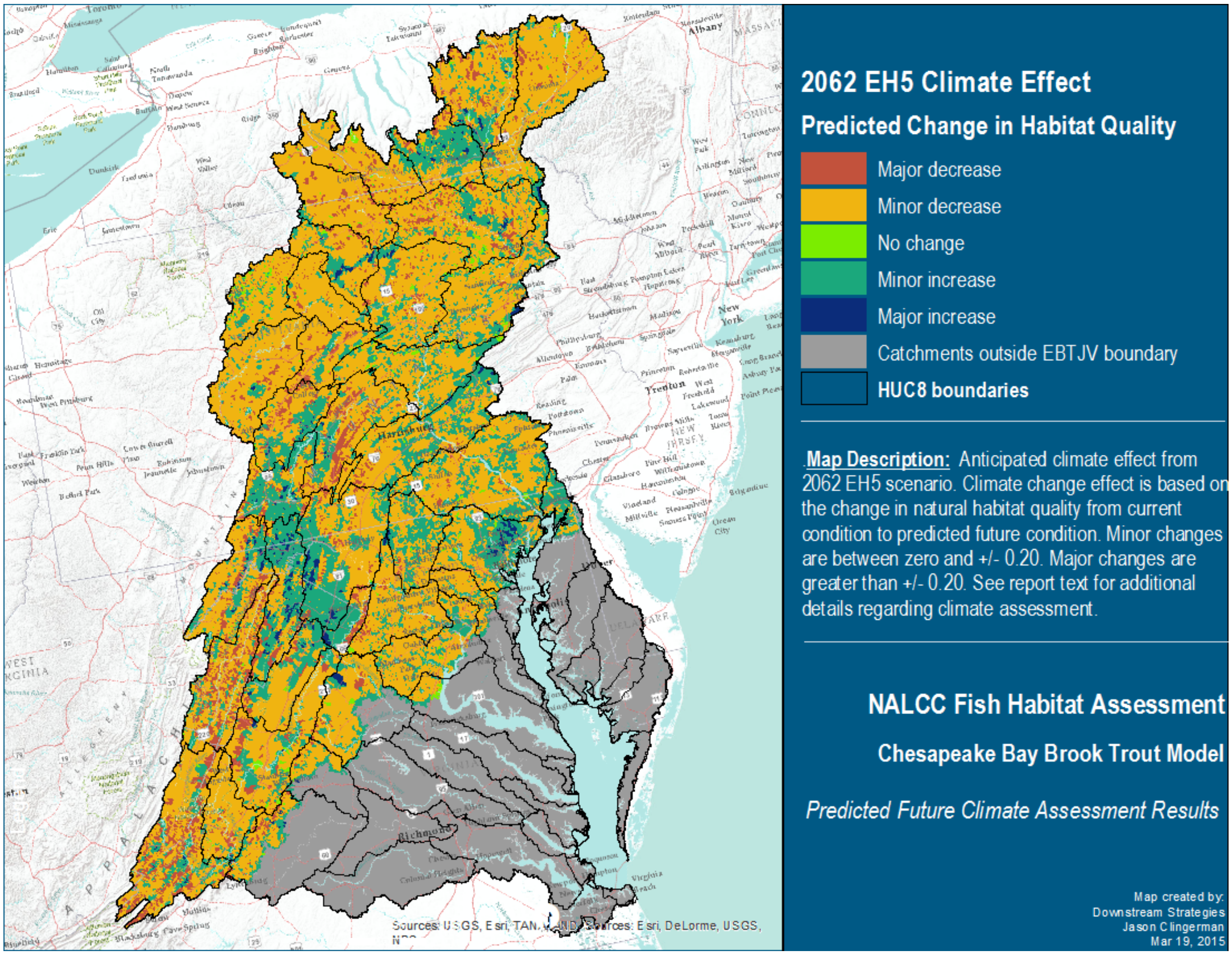
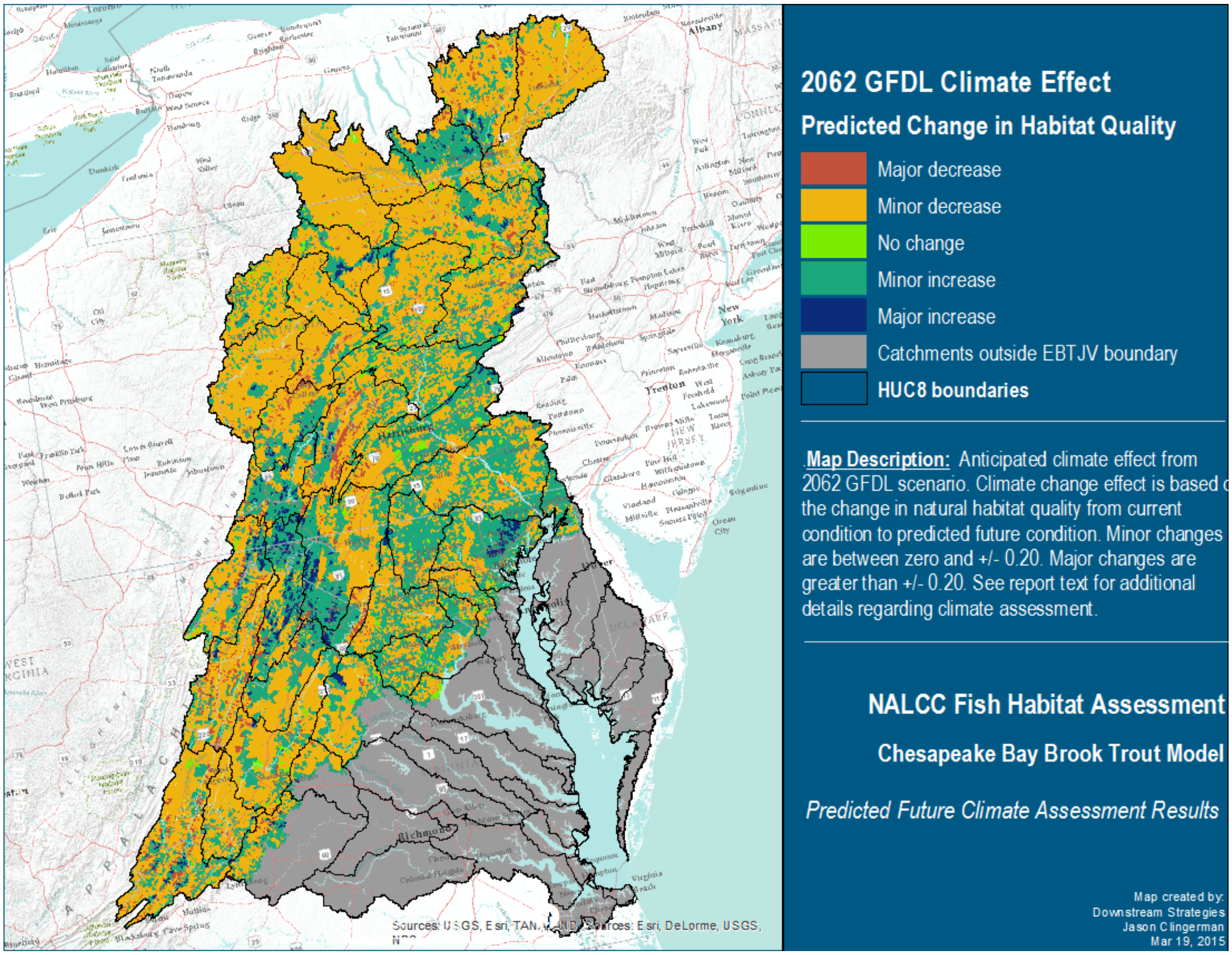


FIGURE 17. CLIMATE CHANGE EFFECT FOR 2062 GFDL SCENARIO.



Using climate vulnerability and resilience to inform priority establishment

Results from the future climate scenarios can allow natural resource managers to evaluate not only current conditions for restoration, but to also incorporate resiliency to climate change in decision making processes. This will allow for priorities for restoration to be established in areas where the work is expected to persist. Conversely, areas indicated as vulnerable to future climate scenarios can be identified and prioritized for actions that may ameliorate the impacts of warmer water and/or less precipitation.

Hierarchical Establishment of Priorities: Case Study

This case study shows an example of how we utilized a hierarchical process to establish restoration and protection priorities using the results from this assessment using the framework described in Merovich et al. (2013). This case study not only utilizes the information from the current model, but also uses future climate scenario predictions, which can allow natural resource managers to establish restoration priorities in areas where brook trout are expected to persist under future climate scenarios. Conversely, areas that are vulnerable to future climate scenarios could be identified and prioritized for actions that may ameliorate the impacts of warmer temperatures.

While this case study is reasonable and potentially useful at a watershed-wide scale, it is provided only as an example. Resource managers developing priorities for brook trout could incorporate data from other assessments, such as the EBTJV priority catchment information into a similar decision making process to further inform any priorities developed.

Fishery Value Calculation

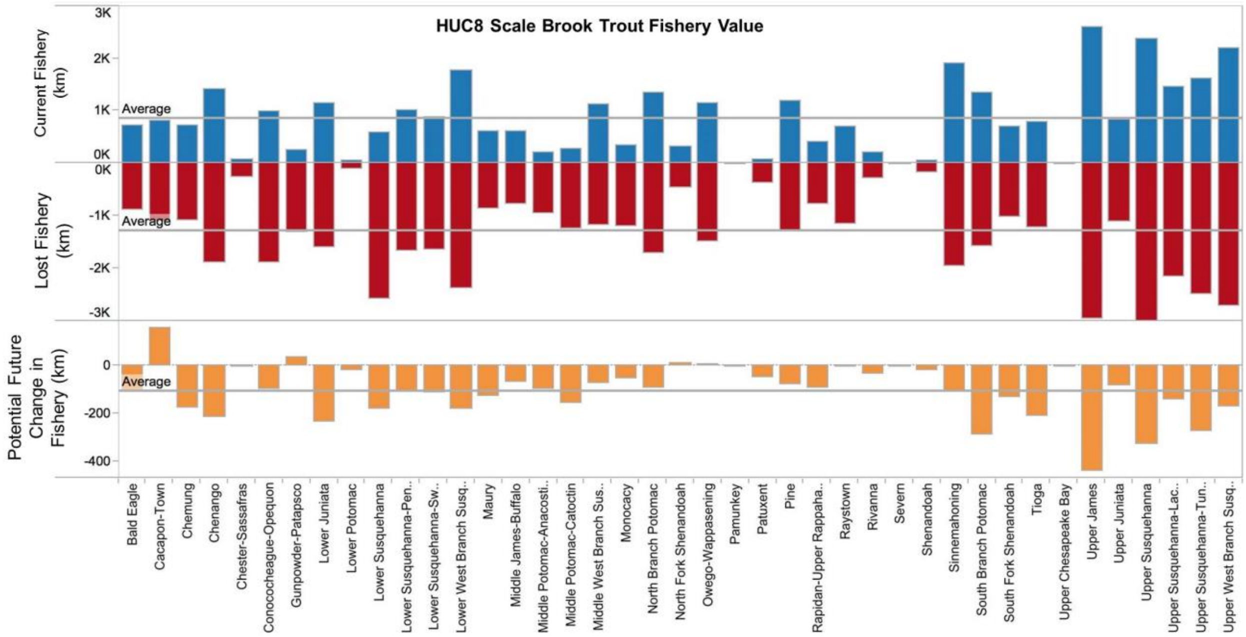
By evaluating the current status of the brook trout fishery, estimating lost fishery value due to landscape stressors, and estimating the potential change in future fishery condition due to climate; we can begin to develop broad scale conservation priorities at the HUC8 scale. Here we apply a process for estimating functional stream length that was developed for identifying acid remediation priorities (Petty and Thorne, 2005) and culvert replacement priorities (Poplar-Jeffers et al., 2009). To estimate the current functional value of a given stream segment as brook trout habitat, we multiplied the length of the segment (kilometer, km) by the current occupancy measure (varies from 0-1). The final value can be interpreted as an ecological function weighted length of habitat. This value can then be summed across all stream segments within a HUC12 or HUC8 watershed to provide a relative measure of current fishery value in units of stream segment length (km) at the larger scales. We can get a similar estimate of lost value for each segment by multiplying anthropogenic stress by the stream length. Once summed across segments within a watershed, this gives us a measure of the fishery value that has been lost due to anthropogenic stress on the landscape. Finally, we can multiply the change in natural habitat quality expected due to climate change by the stream length to get a measure of the potential lost fishery value that may result from climate change. The combination of these three measures (current value, value lost due to stress, and potential value loss due to climate) provides important information for setting conservation priorities at hierarchical scales (e.g., segment, HUC12, HUC8).

HUC8-scale priority establishment

Protection example: One priority could be to protect remaining brook trout populations within highly degraded HUC8 watersheds, especially when those areas are projected to remain resilient to future climate perturbations. Using Figure 18, we can see that the two HUC8 watersheds that stand out as resilient to climate change (positive orange bar) in the figure below are Cacapon-Town and Gunpowder-Patapsco. Of these two watersheds, the Gunpowder-Patapsco has a very small amount of current fishery remaining (blue bar) and has lost quite a large amount of habitat due to stress (red bar). This watershed will be the focus of our first scenario, where protection of remaining populations should be a priority.

Restoration example: From the same graphic in Figure 18, we can also identify those HUC8 watersheds best suited for restoration. Both the Upper James and Upper Susquehanna HUC8's possess relatively strong current fishery values (blue bar) and have also lost considerable value due to anthropogenic stress (red bar). This indicates ample opportunity to reduce stressors and build from strong remaining populations. Since the Upper Susquehanna has lower overall vulnerability to future climate change compared to the Upper James, it will be spotlighted for a priority restoration scenario.

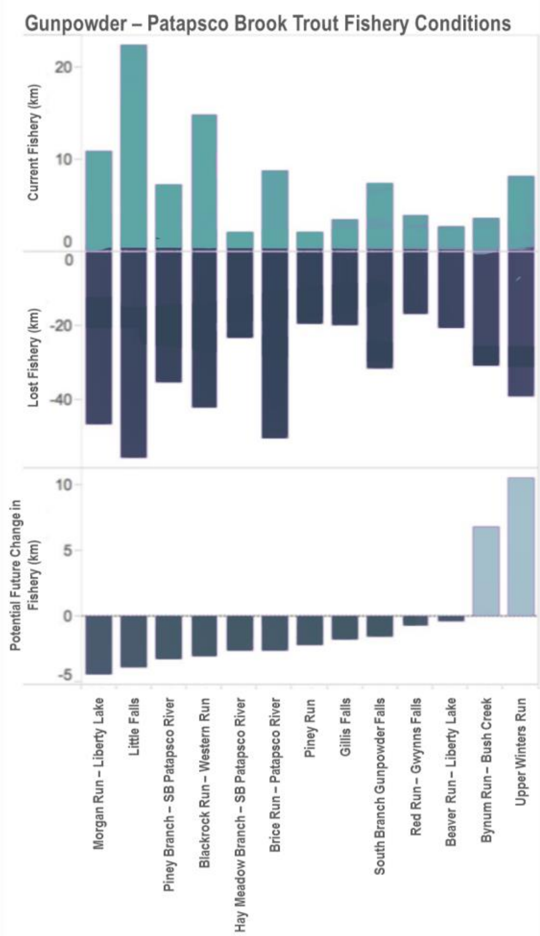
FIGURE 18. EVALUATION OF HUC8 CURRENT, LOST, AND FUTURE BROOK TROUT HABITAT.



HUC12 scale priority establishment

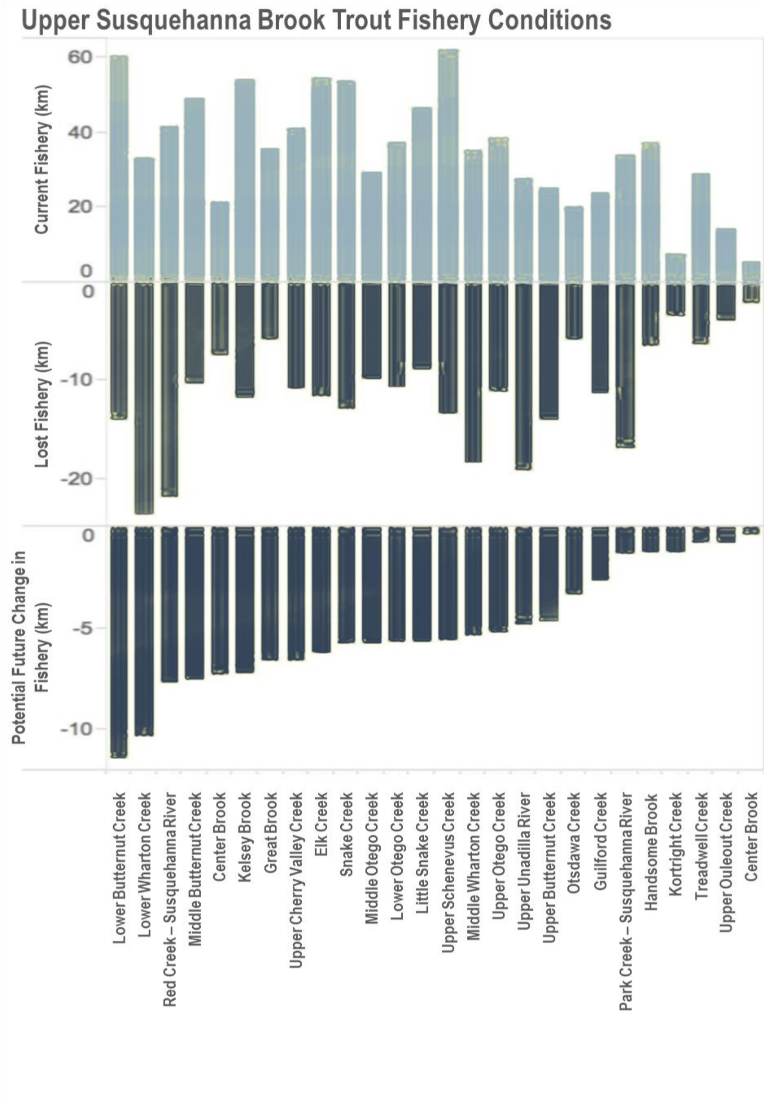
Protection example: Within our first example in the Gunpowder-Patapsco, an analysis of the same factors as above (current, lost, and future brook trout habitat value) within each HUC12 can further direct the establishment of protection priorities. For directing protection of remnant populations, focusing on those areas most resilient to climate change would be beneficial to ensure protections are not undermined by future climate conditions. Given that, the two HUC12 watersheds on the far right of Figure 19 (Bynum Run-Bush Creek and Upper Winters Run), would be watersheds to examine further for protection priorities. Areas with the highest overall remaining fishery would be other targets for this type of protection, so HUC12 watersheds Little Falls (second from left) and Blackrock Run-Western Run (fourth from the left) would also fall into this type of protection priority.

FIGURE 19. GUNPOWDER-PATAPSCO EVALUATION OF HUC12 CURRENT, LOST AND FUTURE BROOK TROUT HABITAT.



Restoration example: Within the Upper Susquehanna HUC8, where brook trout populations are currently predicted to be strong, protection may still be applicable for the HUC12s with the best conditions, but to evaluate restoration priorities, identifying HUC12s with moderate to high current condition, moderate to high lost fishery, and with the lowest detrimental impacts from future climate scenarios would be appropriate. Figure 20 shows the HUC12s that would most likely match those conditions within the Upper Susquehanna HUC8 would be Upper Schenevus Creek (highest current fishery and moderate lost fishery, near middle of chart) and Park Creek-Susquehanna River (sixth bar from the right: relatively high current fishery, high lost fishery, and very low climate vulnerability).

FIGURE 20. UPPER SUSQUEHANNA EVALUATION OF HUC12 CURRENT, LOST AND FUTURE BROOK TROUT FISHERIES.



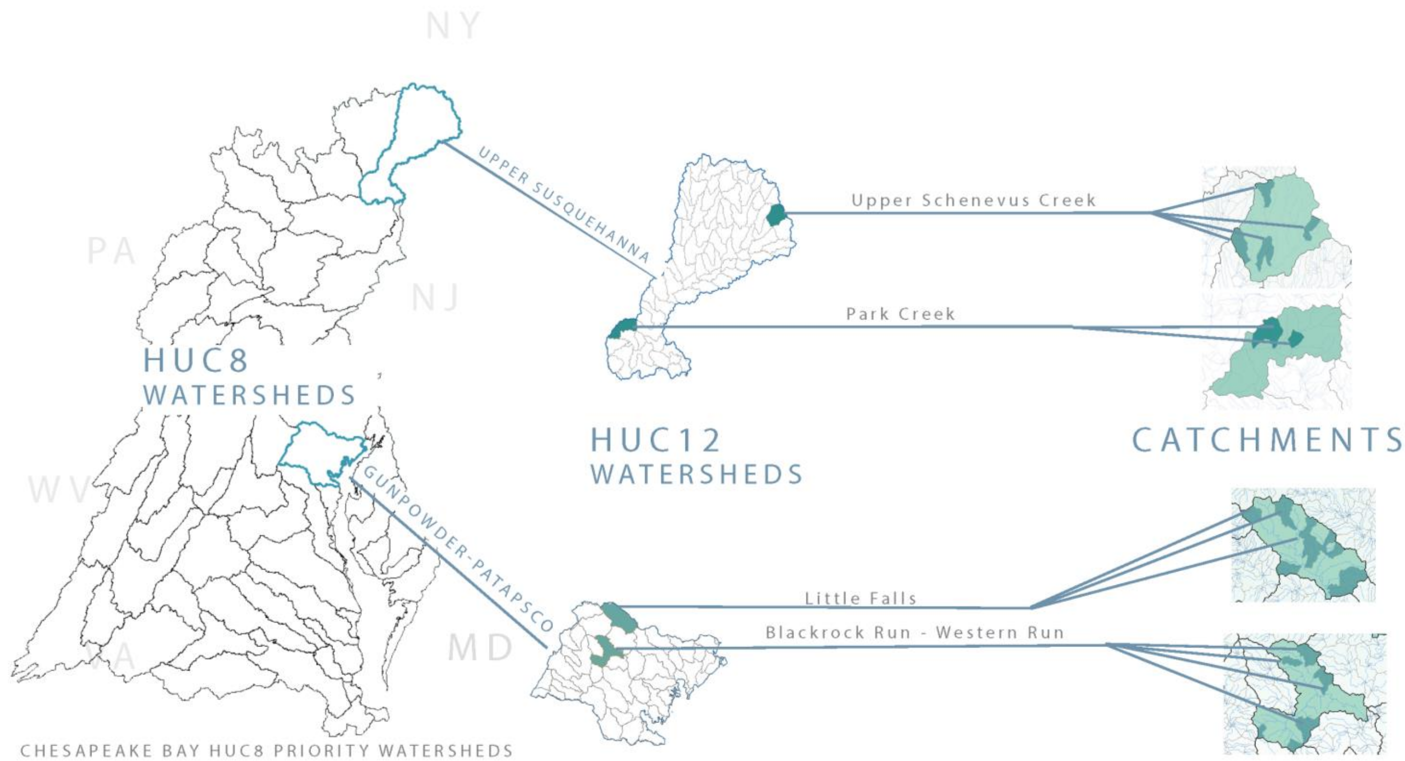
Segment (Catchment)-Scale Priority Establishment

Ultimately, all on-the-ground actions need to happen at the stream segment level. The analyses of data at the HUC8 and HUC12 can help to prioritize the best larger watershed for specific actions, but regardless of the broader priorities, catchment-level priorities are what natural resource managers will use to site specific actions. At the segment level, we can analyze several factors simultaneously to assess the most ideal stream segments for protection or restoration. Additional review of habitat conditions is also more possible within a relatively small number of focal catchments.

Protection example: For this example, we focused on the Little Falls and Blackrock Run-Western Run HUC12s identified in the HUC12 priority section. Catchment values were queried to show only those segments with high natural quality (>0.75) and high future natural quality (>0.75). The identified catchments (Figure 21) have high current fishery value and are anticipated be resilient to future climate scenarios. Upon further analysis of data for these catchments, we found these catchments to be highly agricultural (approximately 35% of land area) and relatively developed (7% mean imperviousness), so protection for these areas may include ensuring proper agricultural practices continue and that runoff from impervious areas is captured before entering streams.

Restoration example: For the two HUC12s selected as restoration priorities within the Upper Susquehanna HUC8 ('Upper Schenevus Creek' and 'Park Creek'), catchments were selected that have high natural quality (> 0.75), a current occupancy of 0.25 – 0.5, and high future natural quality score (> 0.75). This indicates segments (Figure 21) which have high underlying potential, slightly lower occupancy rates because of stress, a high future climate resiliency. These would be streams with good potential as brook trout habitat is restored. From further analysis, the main stressor for these 10 segments identified was agriculture, which averaged about 30% of the total land area. Likely restoration efforts for these areas may include exclusion fencing and implementation of other best agricultural practices.

FIGURE 21. HIERARCHICAL PRIORITIZATION SCENARIO.



Prioritization Summary

The above scenarios provide only a few examples of how these data can be used to establish priority areas for brook trout conservation actions in the Chesapeake Bay watershed. Applying a hierarchical approach that utilizes information summarized at various scales may provide the best watershed-wide priorities, but this approach could start at any level and continue to the segment level. For example, a state agency or watershed organization may only be concerned with the priorities within their work area. In such instances, groups could begin prioritizations at the HUC12 level, and establish priorities that matched their mission and conservation goals.

Discussion

Added value of this model

There are several key outcomes from this modeling effort that should be of considerable value for brook trout conservation in the Chesapeake Bay watershed. First, our new model provides substantially improved predictive power within Chesapeake Bay watershed, when compared to previous modeling efforts. Improved accuracy of this model is likely the result of a combination of factors including: a larger fishery dataset, the inclusion of additional stressor variables such as past mining intensity, the application of an extremely powerful machine-learning statistical approach, and constraint of the model extent to the Chesapeake Bay watershed.

Second, in addition to providing predictions of brook trout occupancy, we quantified the impacts that each anthropogenic stressor had on brook trout occurrence rates as well as the underlying potential of brook trout habitat in the absence of stress. These indices are critical in evaluating habitat restoration and protection priorities, and are additional products our effort produced that have not been fully accounted for in other efforts. In addition, they allow resource managers to identify areas within the broader watershed that are influenced by specific stressors. For example, although mining impacts may not be widespread threats or stressors to brook trout across its entire range, there are significant portions of the Chesapeake Bay watershed impacted by current and historical mining. By incorporating mining as a predictor variable in our model, we were able to quantify these impacts that were unaccounted for in other efforts.

Third, the statistical model that we produced, in addition to having high predictive accuracy, is extremely efficient analytically. Whereas most high level statistical modeling techniques require multiple hours to run at the scale of the Chesapeake Bay watershed, our use of BRT allows models to be run in seconds. The important aspect of this is that we are then able to incorporate the BRT occupancy model into an alternative futures modeling application. When users alter the landscape hypothetically, they are able to obtain immediate feedback on the potential effects of landscape change on brook trout populations. No other modeling approach combines analytical efficiency with predictive power like BRT does.

Finally, our analysis of future climate scenarios provides spatially explicit predictions of the potential impacts of future changes in precipitation and water temperature on brook trout habitats in the Chesapeake Bay watershed. These predictions will allow natural resource managers to assess future conditions as well as current brook trout conditions when making decisions about restoration or management efforts. The degree of error in these predictions is unclear. Nevertheless, they have value in identifying areas where brook trout populations may be at high risk in changing climate.

All of the above components will be embedded into a web-based decision support tool that will provide all interested stakeholders with access to all data and tools compiled as part of this effort. This tool will also utilize relevant data from other related efforts to allow a very thorough repository for all data pertaining to brook trout within the Chesapeake Bay watershed.

Limitations and suggestion for future work

In general, while the estimates of probability of presence, index scores, HQI, and ASI generated through this assessment 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.

While this model has been created for, and is highly accurate within the Chesapeake Bay watershed, its use is limited to only that geographic region. Results and habitat relationships cannot be applied to areas outside the study area, which ultimately restricts widespread use of this assessment. One suggestion for future work regarding the impact of model extent and scale is the need to examine the balance between statistically valid, region-wide models (e.g., DeWeber and Wagner model) and within-region specific models such as our assessment. Each model has applicability, and a detailed analysis of the tradeoffs and benefits of each type of assessment would be useful for future efforts. Furthermore, a line of future research should involve direct comparisons of BRT and hierarchical logistic regression approaches in the Chesapeake Bay watershed.

Competitive interactions with brown trout have been shown to decrease the occupancy of brook trout (Wagner et al., 2013), so the inclusion of interactions with exotic trout in future models could improve the precision of the model and the ability to quantify its influence on the response variable, given the proper scope and scale of assessment. For this assessment, data relating to non-native salmonids was examined to use as a predictor variable. After examination, it was determined that because of the similar habitat requirement of brook trout and non-native salmonids, that the presence of non-native salmonids would not be a useful predictor variable at the scale of this assessment. The modeling done by Wagner et al. (2013) only assessed conditions within streams that could support trout (brown or brook trout were sampled and the watershed size was less than 1,000 km^2), but when assessing all stream reaches the relationship between exotic trout and brook trout typically is positive. By excluding non-trout streams, Wagner et al. (2013) were better able to isolate differences in brook trout occupancy related to changes in brown trout presence, rather than finding relationships in brook trout occupancy across a wider range of habitat conditions. In the latter situation, as in this assessment, the influence of exotic trout is muddled by other habitat factors. Nevertheless, the biological interactions between non-native salmonids may account for some local variability in model results that were beyond the scope of this project: according to Elith and Leathwick (2009), this is a complex and difficult solution to implement in predictive models.

All results generated through the modeling process are ultimately limited by the quality and scale of data used in the model. In the future, the model can be improved by utilizing refined or higher quality predictor data. For example, many of the datasets used for predictor variables were based on a 30 meter grid cell (precipitation, land cover, impervious surfaces), and if resolution of those publically-available datasets improves to 10- or 1-meter grid cells that data would likely result in more accurate results. Data such as the mining predictor variable were based on data collected from multiple sources across states, and as such only the data that were similar across states could be utilized here while higher quality data only available in a certain state had to be omitted.

Also adding additional predictor variables that are deemed appropriate at structuring brook trout populations could be beneficial. While we feel confident that the major factors influencing brook trout in the Chesapeake Bay watershed have been included in this analysis, if future study indicates additional variables of importance, those should also be included. In the future, inclusion of more refined predictor variables or additional relevant predictor variables could improve both the precision of the BRT model predictions and post-modeling indices.

Another 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.

There were also a few additional issues that were beyond the scope of this project. Acid precipitation 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.

Local habitat measures such as water quality (pH, alkalinity, and conductivity), 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.

References

Comte, L., L. Buisson, M. Daufresne, and G. Grenouillet . 2013. Climate-induced changes in the distribution of freshwater fish: observed and predicted trends. Freshwater Biology 58:625-639.

DeWeber, J. and T. Wagner. 2015. Predicting brook trout occurrence in stream reaches throughout their native range in the eastern United States. Transactions of the American Fisheries Society 144: 11-24.

_____ . 2014. A regional neural network ensemble for predicting mean daily river water temperature. Journal of Hydrology 517: 187-200.

______ . Future Stream Temperature Projections for Eastern Brook Trout. Unpublished Data.

Downstream Strategies. 2015. Midwest fish habitat partnership fish habitat modeling results: great lakes basin fish habitat partnership. http://downstreamstrategies.com/documents/reports_publication/glb-results-summary_final.pdf

______ . 2013. Case Study: analysis of scale on boosted regression tree fish habitat models. Unpublished report. http://www.downstreamstrategies.com/documents/reports_publication/huc8_casestudy_final.pdf

______ . 2012. Midwest fish habitat partnership fish habitat modeling results: driftless area restoration effort (DARE). http://midwestfishhabitats.org/sites/default/files/dare.pdf

Eastern Brook Trout Joint Venture (EBTJV). 2015. Catchment assessment and habitat patch spatial layers. Version 04-23-15.

______ . 2006. Eastern brook trout: status and threats. Trout Unlimited, Arlington, VA. http://easternbrooktrout.org/reports/eastern-brook-trout-status-and-threats/at_download/file

Elith, J., C. Graham, R. Anderson, M. Dudı´k, S. Ferrier, A. Guisan, R. Hijmans, F. Huettmann, J. Leathwick, A. Lehmann, J. Li, L. Lohmann, B. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. Overton, A. Peterson, S. Phillips, K. Richardson, R. Scachetti-Pereira, R. Schapire, J. Sobero´n, S. Williams, M. Wisz, and N. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129-151.

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

Elith, J. and J. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution and Systematics 40:677-697.

Fleishman, E., J. Thomson, E. Kalies, B. Dickson, D. Dobkin, and M. Leu. 2014. Projecting current and future location, quality, and connectivity of habitat for breeding birds in the Great Basin. Ecosphere, 5(7), art82.

Friedman, J. 2001. Greedy function approximation: a gradient boosting machine. Annals of Statistics 29: 1189-1232.

Friedman, J. and J. Meulman. 2003. Multiple additive regression trees with application in epidemiology. Statistics in Medicine 22: 1365-81.

Horizon Systems. 2012. National Hydrography Dataset Plus – NHDPlus. http://www.horizon-systems.com/NHDPlus/index.php

Hickling, R., D. Roy, J. Hill, R. Fox, and C. Thomas. 2006. The distributions of a wide range of taxonomic groups are expanding polewards. Global Change Biology 12:450–455.

Hostetler, S., J. Alder, and A. Allan. 2011. Dynamically downscaled climate simulations over North America: methods, evaluation, and supporting documentation for users. Technical report, U.S. Geological Survey, Reston, Virginia.

Hudy, M., T. Thieling, N. Gillespie, and E. Smith. 2008. Distribution, status, and land use characteristics of subwatersheds within the native range of brook trout in the eastern united states. North American Journal of Fisheries Management 28:1069-1085.

Intergovernmental Panel on Climate Change. 2007. Climate change 2007: synthesis report. in: contribution of working groups I, II and III to the fourth assessment report of the Intergovernmental Panel on Climate Change. Technical report, IPCC, Geneva, Switzerland.

Merovich Jr., G. T., J. T. Petty, M. P. Strager, and J. B. Fulton. 2013. Hierarchical classification of stream condition: a house-neighborhood framework for establishing conservation priorities in complex riverscapes. Freshwater Science 32(2)874-891.

Merriam, E. R., J. T. Petty, M. P. Strager, A. E. Maxwell, and P. F. Ziemkiewicz. 2013. Scenario analysis predicts context-dependent stream response to landuse change in a heavily mined central Appalachian watershed. Freshwater Science 32(4):1246-1259.

Petty, J. T. and D. Thorne. 2005. An ecologically based approach to identifying restoration priorities in an acid-impacted watershed. Restoration Ecology 13(2):348-357.

Petty, J. T., J. B. Fulton, M. P. Strager, G. T. Merovich Jr., J. M. Stiles, and P. F. Ziemkiewicz. 2010. Landscape indicators and thresholds of stream ecological impairment in an intensively mined Appalachian watershed. Journal of North American Benthological Society 29(4):1292-1309.

Poplar-Jeffers, I., J. Petty, J. Anderson, S. Kite, M. Strager, and R. Fortney. 2009. Culvert replacement and stream habitat restoration: implications from brook trout management in an Appalachian watershed, USA. Restoration Ecology 17(3):404-413.

Thrasher B., J. Xiong, W. Wang, F. Melton, A. Michaelis, and R. Nemani. 2013. Downscaled climate projections suitable for resource management. Eos, Transactions American Geophysical Union 94(37):321-32.

Wagner T., J. Deweber, J. Detar, and J. Sweka. 2013. Landscape-scale evaluation of asymmetric interactions between brown trout and brook trout using two-species occupancy models. Transactions of the American Fisheries Society 142:353-361.

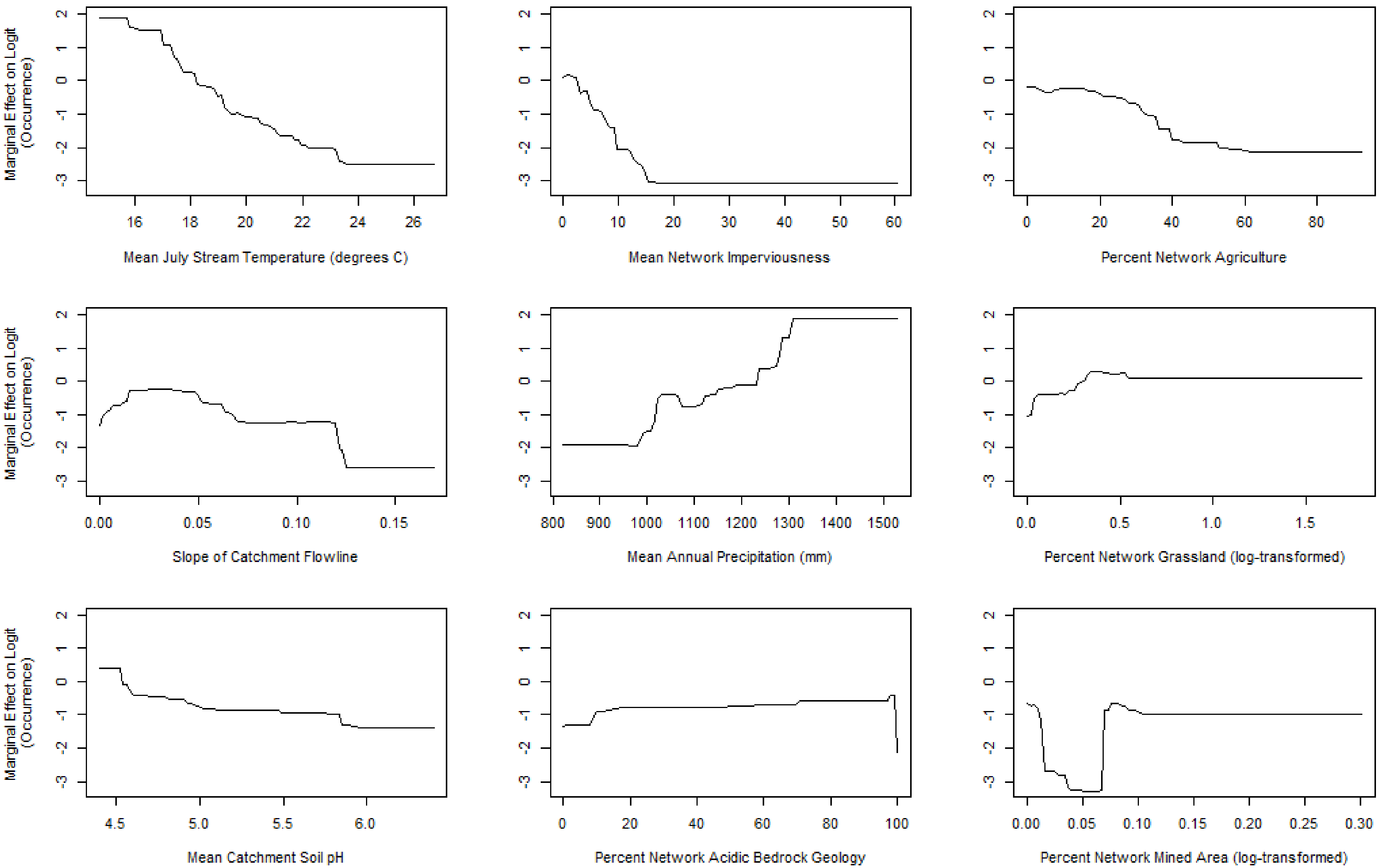
Appendix A: DATA DICTIONARY

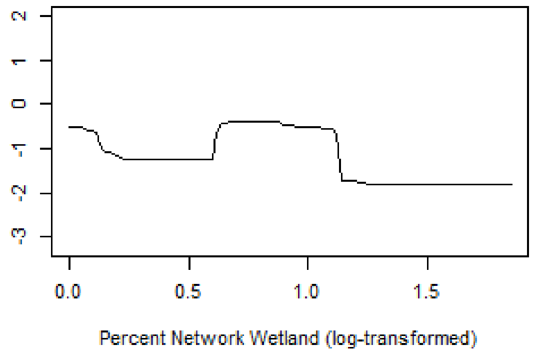
This table lists all the predictor variables compiled and considered as part of the model building process. Many variables were removed from model consideration because of high correlation with other variables, lack of variability (i.e. all values equal zero), or lack of relative influence in preliminary models. Predictor variables noted with a preceding asterisk (*) were included in the final predictive model.

Predictor Variable	Description	Source
FeatureID	Catchment identifier	NHDPlus
AreaSqKM	Cumulative drainage area (km2)	NHDPlus
Cu_AreaSq	Cumulative drainage area (km2)	NHDPlus
***Precip**	Mean annual precipitation (mm*100)	NHDPlus
Temp	Mean annual temperature (degrees centigrade * 100)	NHDPlus
RunOffV	Mean runoff (mm)	NHDPlus
Q0001A	Flow from runoff (cfs) using EROM method	NHDPlus
V0001A_fix	Velocity for Q0001A (fps)	NHDPlus
MAFLOWV_fix	Flow from runoff (cfs) using Vogel method	NHDPlus
MAVELV_fix	Mean annual velocity (fps) using Jobson method with MAFlowV	NHDPlus
MINELEVRAW_M	Minimum elevation (m)	NHDPlus
***SLOPE_fix**	Slope of flowline (m/m)	NHDPlus
Dev_p	Developed land cover (% catchment area)	NLCD 2006
For_p	Forested land cover (% catchment area)	NLCD 2006
Wet_p	Wetland land cover (% catchment area)	NLCD 2006
Bar_p	Barren land cover (% catchment area)	NLCD 2006
Grass_p	Grassland/herbaceous land cover (% catchment area)	NLCD 2006
Past_p	Pasture/hay land cover (% catchment area)	NLCD 2006
Crops_p	Cultivated crops land cover (% catchment area)	NLCD 2006
Ag_p	Agriculture land cover (% catchment area)	NLCD 2006
Dev_pc	Developed land cover (% cumulative upstream area)	NLCD 2006
For_pc	Forested land cover (% cumulative upstream area)	NLCD 2006
Wet_pc	Wetland land cover (% cumulative upstream area)	NLCD 2006
Bar_pc	Barren land cover (% cumulative upstream area)	NLCD 2006

Predictor Variable	Description	Source
Grass_pc	Grassland/herbaceous land cover (% cumulative upstream area)	NLCD 2006
***Log_Grass_pc**	Grassland/herbaceous land cover (log(% cumulative upstream area))	NLCD 2006
Past_pc	Pasture/hay land cover (% cumulative upstream area)	NLCD 2006
Crops_pc	Cultivated crops land cover (% cumulative upstream area)	NLCD 2006
***Ag_pc**	Agriculture land cover (% cumulative upstream area)	NLCD 2006
IMP06	NLCD 2006 percent impervious (catchment average)	NLCD 2006
***IMP06C**	NLCD 2006 percent impervious (cumulative upstream average)	NLCD 2006
Acid_geol_p	Acidic bedrock geology - V100P, V200P, V500P (% catchment area)	TNC Aquatic Classification, reclassified
Calc_geol_p	Calcareous bedrock geology - classes V300P, V400P (% catchment area)	TNC Aquatic Classification, reclassified
Other_geol_p	Other bedrock geology - classes V600P, V700P, V800P, V900P (% catchment area)	TNC Aquatic Classification, reclassified
***Acid_geol_pc**	Acidic bedrock geology - V100P, V200P, V500P(% cumulative upstream area)	TNC Aquatic Classification, reclassified
Calc_geol_pc	Calcareous bedrock geology - classes V300P, V400P(% cumulative upstream area)	TNC Aquatic Classification, reclassified
Other_geol_pc	Other bedrock geology - classes V600P, V700P, V800P, V900P (% cumulative upstream area)	TNC Aquatic Classification, reclassified
***SoilpH**	Soil pH (average within catchment)	UMASS
SO4_dep	Annual sulfate deposition (NTN), year 2011 (mean within catchment)	National Atmospheric Deposition Program
Invasive_cat	Invasive trout status, categorical (catchment)	Eastern Brook Trout Joint Venture
AML_dens	Abandoned Mine Lands sites (#/km2, catchment area)	OSM
AML_densc	Abandoned Mine Lands sites (#/km2, cumulative upstream area)	OSM
Current_surfminep	Currently permitted surface mines (% catchment area)	MDE, PADEP, WVDEP
Current_surfminepc	Currently permitted surface mines (% cumulative upstream area)	MDE, PADEP, WVDEP
past_minep	Surface mined areas, past (% catchment area)	MDE, PADEP, WVGES
past_minepc	Surface mined areas, past (% cumulative upstream area)	MDE, PADEP, WVGES
***Log_past_minepc**	Surface mined areas, past (log(% cumulative upstream area))	MDE, PADEP, WVGES
***mnjuly**	Mean predicted july water temperature (catchment)	DeWeber and Wagner, 2014

Appendix B: FUNCTIONAL RESPONSE PLOTS





2
1
0
-1
-2
-3
0.0
0.5
1.0
1.5
Percent Network Wetland (log-transformed)