YellowLance - Potential Habitat, July 2022 - NC Department of Transportation

File Geodatabase Feature Class

Tags

Yellow lance, Elliptio lanceolata, Elli_lanc, Freshwater Bivalve, Aquatic, threatened, mussel, USFWS Bridge/culvert programmatic agreement, random forest model, machine learning, Transportation, NRTR, NCDOT, Environment, Location, North Carolina, ATLAS

Summary

This dataset was originally created in February 2022 as part of the Project ATLAS initiative at NCDOT to support the Environmental Analysis Unit (EAU) Mitigation and Modeling Unit with project delivery in the development phase.

The Yellow lance model is a Random Forest (machine-learning) model. Random Forest models were used to predict the probability of habitat at the scale of USGS National Hydrography Data (NHD Plus V2) catchments and their associated stream segments (median area: 272 acres; 5th and 95th percentile area: 4 acres and 1541 acres).

This dataset supports the production of the Natural Resources Technical Report (NRTR). This dataset also contains information that may assist biologists in preparing background information for field surveys, in order to address protected species for Threatened & Endangered Species Survey Reports, and/or Biological Assessments. The model is intended to be used in project management tools to:
1. flag areas of higher versus lower risk of "May Affect" biological conclusions within a species range to improve project planning and management, and
2. add a landscape-scale perspective to improve biologists' field planning and site assessment

Description

The Yellow lance Potential Habitat dataset is a polygon layer depicting high, moderate and low potential habitat locations for Yellow lance in NC counties.

The Yellow lance (*Elliptio lanceolata*) is a sand-loving species often found buried deep in clean, coarse to medium sand, although it can sometimes be found in gravel substrates. Yellow lances often are moved with shifting sand and eventually settle in sand at the downstream end of stable sand and gravel bars. This species depends on clean, moderate flowing water with high dissolved oxygen. This species is found in medium-sized rivers to smaller streams. Historically, the yellow lance ranged from the Patuxent River Basin in Maryland, to the Potomac River Basin in Maryland/Virginia, the Rappahannock, York, James, and Chowan River basins in Virginia, and the Tar and Neuse River basins in North Carolina.

The three levels are: Low, Moderate, and High Probability of Potential Habitat (based on similarity of environmental conditions to those found at known occurrence locations). The category thresholds were set based on the distribution of predicted values for known habitat. The High-Moderate threshold is set at the level where 90% of the observed potential habitat (species presence and reviewer judgments) falls within the High category (Presence Percent Correctly Classified). The Moderate-Low threshold were set at the level where 8% of the observed potential habitat (species presence and reviewer judgments) falls within the Moderate category and 2% within the Low category. The final thresholds for this species are 0.452 and 0.754 for the Low-Moderate and Moderate-High thresholds, respectively. Lower thresholds result in more of the range labelled as High probability of habitat and greater misclassification of known non-habitat locations. Higher thresholds result in more of the range labelled as Low probability of habitat and greater misclassification of known habitat locations.

Given the larger spatial unit of ecological models and ecological characteristics of aquatic species, the landscape scale environmental attributes of potential habitat varied greatly among sites and could closely resembled non-habitat sites. It is important to remember that potential habitat (and non-habitat) can occur at any classification level within a catchment and must be verified by a qualified biologist.

For more information please click here https://xfer.services.ncdot.gov/gisdot/Metadata/Atlas/TechDocs/

Datasets developed under Project ATLAS do not replace any NRTR work for future projects and may not be used as a replacement for site visits / field surveys by qualified professionals and hence should be used only as a supporting platform for decision making. Use of this dataset for project scoping or screening is merely pre-decisional.

Credits

The Environmental Analysis Unit (EAU) Mitigation and Modeling Unit within NCDOT was tasked to create this dataset. This dataset supports the production of the Natural Resources Technical Report (NRTR). Maintenance of this dataset is handled by the EAU.

Support and maintenance of the enterprise spatial database where this data resides is handled by NCDIT's Transportation GIS Unit.

Use limitations

The North Carolina Department of Transportation shall not be held liable for any errors in this data. This includes errors of omission, commission, errors concerning the content of the data, and relative and positional accuracy of the data. This data cannot be construed to be a legal document. Primary sources from which this data was compiled must be consulted for verification of information contained in this data.

Datasets developed under Project ATLAS do not replace any NRTR work for future projects and may not be used as a replacement for site visits / field surveys by qualified professionals and hence should be used only as a supporting platform for decision making. Use of this dataset for project scoping or screening is merely pre-decisional.

Extent

West -78.872364 **East** -77.351467
North 36.453002 **South** 35.237112

Scale Range

Maximum (zoomed in) 1:5,000
Minimum (zoomed out) 1:625,000

ArcGIS Metadata ►

Topics and Keywords ►

THEMES OR CATEGORIES OF THE RESOURCE biota, geoscientificInformation, inlandWaters, location, transportation, environment

* CONTENT TYPE Downloadable Data
EXPORT TO FGDC CSDGM XML FORMAT AS RESOURCE DESCRIPTION No

PLACE KEYWORDS North Carolina

THESAURUS ►
TITLE User
CREATION DATE 2022-02-25 00:00:00
PUBLICATION DATE 2022-06-01 00:00:00

Hide Thesaurus ▲

THEME KEYWORDS Yellow lance, Elliptio lanceolata, Elli_lanc, Freshwater Bivalve, Aquatic, threatened, mussel, USFWS Bridge/culvert programmatic agreement, random forest model, machine learning,Transportation, NRTR, NCDOT, Environment, Location, North Carolina, ATLAS

THESAURUS ►
TITLE User
CREATION DATE 2022-02-25 00:00:00
PUBLICATION DATE 2022-06-01 00:00:00

Hide Thesaurus ▲

Hide Topics and Keywords ▲

Citation ►

TITLE YellowLance - Potential Habitat, July 2022 - NC Department of Transportation
CREATION DATE 2022-02-25 00:00:00
PUBLICATION DATE 2022-06-01 00:00:00

PRESENTATION FORMATS digital map

Hide Citation ▲

Citation Contacts ►

RESPONSIBLE PARTY
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE originator

CONTACT INFORMATION ►
PHONE
VOICE 919-707-6136

ADDRESS
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

RESPONSIBLE PARTY
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant

CONTACT'S ROLE resource provider

CONTACT INFORMATION ▶
PHONE
VOICE 919-707-6136

ADDRESS
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

RESPONSIBLE PARTY
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE point of contact

CONTACT INFORMATION ▶
PHONE
VOICE 919-707-6136

ADDRESS
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

Hide Citation Contacts ▲

Resource Details ▶

DATASET LANGUAGES English (UNITED STATES)
DATASET CHARACTER SET utf8 - 8 bit UCS Transfer Format

STATUS completed
SPATIAL REPRESENTATION TYPE vector

* PROCESSING ENVIRONMENT Version 6.2 (Build 9200) ; Esri ArcGIS 10.8.1.14362

CREDITS

The Environmental Analysis Unit (EAU) Mitigation and Modeling Unit within NCDOT was tasked to create this dataset. This dataset supports the production of the Natural Resources Technical Report (NRTR). Maintenance of this dataset is handled by the EAU.

Support and maintenance of the enterprise spatial database where this data resides is handled by NCDIT's Transportation GIS Unit.

Hide Resource Details ▲

Extents ►

EXTENT
GEOGRAPHIC EXTENT
BOUNDING RECTANGLE
WEST LONGITUDE -84.017454
EAST LONGITUDE -81.839144
SOUTH LATITUDE 35.022011
NORTH LATITUDE 36.167407
EXTENT CONTAINS THE RESOURCE Yes

EXTENT
DESCRIPTION
Data collection is complete.

GEOGRAPHIC EXTENT
BOUNDING RECTANGLE
WEST LONGITUDE -84.422111
EAST LONGITUDE -75.416034
SOUTH LATITUDE 33.730557
NORTH LATITUDE 36.617257
EXTENT CONTAINS THE RESOURCE Yes

TEMPORAL EXTENT
BEGINNING DATE 2022-06-01 00:00:00
ENDING DATE 2022-06-01 00:00:00

EXTENT
GEOGRAPHIC EXTENT
BOUNDING RECTANGLE
EXTENT TYPE Extent used for searching
* WEST LONGITUDE -78.872364
* EAST LONGITUDE -77.351467
* NORTH LATITUDE 36.453002
* SOUTH LATITUDE 35.237112
* EXTENT CONTAINS THE RESOURCE Yes

EXTENT IN THE ITEM'S COORDINATE SYSTEM
* WEST LONGITUDE 2038106.736538
* EAST LONGITUDE 2484940.686373
* SOUTH LATITUDE 545190.556465
* NORTH LATITUDE 983842.851403
* EXTENT CONTAINS THE RESOURCE Yes

Hide Extents ▲

Resource Points of Contact ►

POINT OF CONTACT
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE point of contact

CONTACT INFORMATION ►
PHONE
VOICE 919-707-6136

ADDRESS
TYPE
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

Hide Resource Points of Contact ▲

Resource Maintenance ►

RESOURCE MAINTENANCE
UPDATE FREQUENCY as needed

SCOPE OF THE UPDATES dataset

OTHER MAINTENANCE REQUIREMENTS
Maintenance of this dataset is handled by the Environmental Analysis Unit (EAU) Mitigation and Modeling Unit. Currently updating this dataset has not been planned. Support and maintenance of the enterprise spatial database where this data resides is handled by NCDIT's Transportation GIS Unit.

MAINTENANCE CONTACT
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE originator

CONTACT INFORMATION ►
PHONE
VOICE 919-707-6136

ADDRESS
TYPE
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh

ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

Hide Resource Maintenance ▲

Resource Constraints ►

LEGAL CONSTRAINTS
LIMITATIONS OF USE
The North Carolina Department of Transportation shall not be held liable for any errors in this data. This includes errors of omission, commission, errors concerning the content of the data, and relative and positional accuracy of the data. This data cannot be construed to be a legal document. Primary sources from which this data was compiled must be consulted for verification of information contained in this data.

SECURITY CONSTRAINTS
CLASSIFICATION unclassified
CLASSIFICATION SYSTEM None

LIMITATIONS OF USE
The North Carolina Department of Transportation shall not be held liable for any errors in this data. This includes errors of omission, commission, errors concerning the content of the data, and relative and positional accuracy of the data. This data cannot be construed to be a legal document. Primary sources from which this data was compiled must be consulted for verification of information contained in this data.

CONSTRAINTS
LIMITATIONS OF USE

The North Carolina Department of Transportation shall not be held liable for any errors in this data. This includes errors of omission, commission, errors concerning the content of the data, and relative and positional accuracy of the data. This data cannot be construed to be a legal document. Primary sources from which this data was compiled must be consulted for verification of information contained in this data.

Datasets developed under Project ATLAS do not replace any NRTR work for future projects and may not be used as a replacement for site visits / field surveys by qualified professionals and hence should be used only as a supporting platform for decision making. Use of this dataset for project scoping or screening is merely pre-decisional.

Hide Resource Constraints ▲

Spatial Reference ►

ARCGIS COORDINATE SYSTEM
* TYPE Projected
* GEOGRAPHIC COORDINATE REFERENCE GCS_North_American_1983

* PROJECTION NAD_1983_StatePlane_North_Carolina_FIPS_3200_Feet
* COORDINATE REFERENCE DETAILS
PROJECTED COORDINATE SYSTEM
WELL-KNOWN IDENTIFIER 102719
X ORIGIN -121841900
Y ORIGIN -93659000
XY SCALE 3048.0060960121918
Z ORIGIN -100000
Z SCALE 10000
M ORIGIN -100000
M SCALE 10000
XY TOLERANCE 0.0032808333333333331
Z TOLERANCE 0.001
M TOLERANCE 0.001
HIGH PRECISION true
LATEST WELL-KNOWN IDENTIFIER 2264
WELL-KNOWN TEXT
PROJCS["NAD_1983_StatePlane_North_Carolina_FIPS_3200_Feet",GEOGCS["GCS_North_American_1983",DATUM["D_North_American_1983",SPHEROID["GRS_1980",6378137.0,298.257222101]],PRIMEM["Greenwich",0.0],UNIT["Degree",0.0174532925199433]],PROJECTION["Lambert_Conformal_Conic"],PARAMETER["False_Easting",2000000.002616666],PARAMETER["False_Northing",0.0],PARAMETER["Central_Meridian",-79.0],PARAMETER["Standard_Parallel_1",34.33333333333334],PARAMETER["Standard_Parallel_2",36.16666666666666],PARAMETER["Latitude_Of_Origin",33.75],UNIT["Foot_US",0.3048006096012192],AUTHORITY["EPSG",2264]]

REFERENCE SYSTEM IDENTIFIER
VALUE 2264
* CODESPACE EPSG
* VERSION 6.12(9.0.0)

Hide Spatial Reference ▲

Spatial Data Properties ►

VECTOR ►
* LEVEL OF TOPOLOGY FOR THIS DATASET geometry only

GEOMETRIC OBJECTS
FEATURE CLASS NAME YellowLancePotentialHabitat
* OBJECT TYPE composite
* OBJECT COUNT 8282

Hide Vector ▲

ARCGIS FEATURE CLASS PROPERTIES ►
FEATURE CLASS NAME YellowLancePotentialHabitat
* FEATURE TYPE Simple
* GEOMETRY TYPE Polygon
* HAS TOPOLOGY FALSE
* FEATURE COUNT 8282
* SPATIAL INDEX TRUE
* LINEAR REFERENCING FALSE

Hide ArcGIS Feature Class Properties ▲

Hide Spatial Data Properties ▲

Data Quality ▶

SCOPE OF QUALITY INFORMATION ▶
RESOURCE LEVEL dataset

Hide Scope of quality information ▲

DATA QUALITY REPORT - COMPLETENESS OMISSION ▶
MEASURE DESCRIPTION
After processing, the dataset is checked for drawing display and number of records and file sizes compared with source materials.

CONFORMANCE TEST RESULTS
TEST PASSED Yes
RESULT EXPLANATION
Pass

PRODUCT SPECIFICATION ▶
TITLE NCDOT Geospatial Data Specifications
CREATION DATE 2019-10-01 00:00:00
PUBLICATION DATE 2021-06-17 00:00:00

Hide Product specification ▲

Hide Data quality report - Completeness omission ▲

DATA QUALITY REPORT - CONCEPTUAL CONSISTENCY ▶
MEASURE DESCRIPTION
The dataset is converted to file geodatabase (FGDB) format using tools in ArcGIS. The geometry is checked, and if needed repaired

CONFORMANCE TEST RESULTS
TEST PASSED Yes
RESULT EXPLANATION
Pass

PRODUCT SPECIFICATION ▶
TITLE NCDOT Geospatial Data Specifications
CREATION DATE 2019-10-01 00:00:00
PUBLICATION DATE 2021-06-17 00:00:00

Hide Product specification ▲

Hide Data quality report - Conceptual consistency ▲

DATA QUALITY REPORT - QUANTITATIVE ATTRIBUTE ACCURACY ►

MEASURE DESCRIPTION

In addition to direct species observations, expert judgments of habitat conditions (Potential Habitat versus Non-Habitat) were gathered through desktop and field rapid site assessments.

Species experts provided feedback at two stages of model development. An early draft model was reviewed in ArcGIS Online. Completed via a desktop application, experts could scan and zoom to provide feedback on model performance based on their regional knowledge and any data layers they chose to add and view. Experts entered desktop assessments as high precision points (e.g., "this location is habitat or non-habitat") and/or low precision polygons (e.g., "this region generally provides good habitat or nonhabitat). After further development, an improved model received field validation. Field validation consisted of a rapid visual site assessment to judge a stream segment as Potential Habitat versus Non-Habitat based on evidence of stream condition and, in some cases, observation of associated species. All field validation reviews were entered as highprecision line features denoting the stream section observed.

In both reviews, time, knowledge, and/or access limitations prevented a true randomized design, but experts attempted to maximize the distribution of feedback spatially and across performance categories (True and False Positives, True and False Negatives) and to target areas with high uncertainty (high variance among models). Judgments vary in their spatial precision, including both polygons (low precision) and points and lines (high precision). Experts' judgments identify some non-habitat locations, but these are in the minority.

To ensure a balanced sample, additional catchments were randomly drawn as needed from the background (No Data catchments) to serve as pseudoabsence. Pseudoabsence catchments are unsurveyed catchments modeled treated as absence sites for the purposes of defining absence (versus presence) habitat characteristics.

CONFORMANCE TEST RESULTS

TEST PASSED Yes

RESULT EXPLANATION

Pass

PRODUCT SPECIFICATION ►

TITLE NCDOT Geospatial Data Specifications

CREATION DATE 2020-10-01 00:00:00

PUBLICATION DATE 2021-06-17 00:00:00

Hide Product specification ▲

Hide Data quality report - Quantitative attribute accuracy ▲

Hide Data Quality ▲

Lineage ►

LINEAGE STATEMENT

All species' models use occurrence data from the Natural Heritage Program (NHP Tier 2 data) and the NC Wildlife Resource Commission (NCPAWS data). Species occurrence observations identified as historic (extirpated) or older than 2000 are excluded from models, as are data with low spatial accuracy (e.g., NHP Accuracy "4-Low" or "5-Very Low"). NHP "3-Medium" accuracy observations were retained and labelled as Low precision and NHP 2-High and 1-Very High accuracy observations as

High precision data. The occurrence data are presence-only and are not a randomized or representative sample from the species' range. Models score any catchment with a current, moderate to high precision occurrence as "Potential Habitat".

Often multiple data records fell within a single catchment. For the purposes of modeling, to avoid pseudo-replication, duplicates were removed; each catchment was scored as potential habitat or non-habitat. Where a single catchment contained records of both potential habitat and non-habitat (OBSERVATION SET = "Conflict"), it was scored as habitat. Based on these decisions, the final model predicts the probability of any habitat within the catchment rather than the abundance of habitat within the catchment. With the available observation data (including pseudoabsence) 618 of the 6762 catchments were classified within the species expert delineated model range.

The environmental data attributes included variables drawn from the US EPA StreamCat and the NHD Plus V2 data. From the available StreamCat data, 87 raw data variables plus 12 indices were extracted. The raw data variables primarily described local catchment characteristics across a spectrum of land cover, climatic, physiographic, hydrologic, chemical, geological, and disturbance metrics. A few of these variables provided data at finer (riparian only) and coarser (all upstream watershed area) scales. The 12 indices (six for the local catchment and six for the entire upstream watershed area) were calculated values summarizing multiple raw variables to represent the quality of aquatic habitat conditions.

PROCESS STEP ▶
DESCRIPTION
Geodatabase was forwarded on to the GIS Unit for publishing as part of data for project ATLAS.

PROCESS CONTACT
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE point of contact

CONTACT INFORMATION ▶
PHONE
VOICE 919-707-6136

ADDRESS
TYPE
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

Hide Process step ▲

PROCESS STEP ▶
DESCRIPTION
Data was reviewed in ESRI's Data Reviewer tool to verify geometry.

PROCESS CONTACT
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE point of contact

CONTACT INFORMATION ▶
PHONE
VOICE 919-707-6136

ADDRESS
TYPE
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

Hide Process step ▲

PROCESS STEP ▶
DESCRIPTION
Random Forest models were used to predict the probability of habitat at the scale of USGS National Hydrography Data (NHD Plus V2) catchments and their associated stream segments (median area: 272 acres; 5th and 95th percentile area: 4 acres and
1541 acres). Random Forest models generate predictions through repeated construction of decisiontree style models. At multiple points during model construction and assessment, the random forest process draws a random subset of habitat and non-habitat data and
selects a random subset of the environmental variables by which to compare them.

This randomization is beneficial to reduce overfitting of the models. The model procedure tracks
(1) how frequently sites are predicted to be habitat vs non-habitat,
(2) which variables contributed most to accurate classification of habitat vs non-habitat sites, and
(3) overall statistics about model performance.

The models were run in R primarily using functions from the randomForest and rfUtilities packages. Multiple random forest models were run, each initiated with a different suite of environmental (predictor) and habitat/non-habitat (response) data. Models meeting minimum performance criteria were averaged to produce a single predicted probability of habitat per catchment. These model suites included six core models, plus two additional models if associate species had been defined. “Full Set” refers to models initiated with the 87 raw environmental variables from StreamCat plus the NHD Plus environmental variables, while the “Index Set” refers to models initiated with StreamCat’s 12 calculated condition indices plus NHD Plus environmental variables. For the response variable, models

were run with all species observations and expert judgments (“All Habitat Data”), run with only the verified target species observations (“No Reviewer Habitat Data”), run with only the high precision observations (excluding large generalized polygons from reviewer and observation sets) (“No Low Precision Data”), and run with all habitat data for the target species plus the current, high precision observations of associate species (“Plus Associates” for species with expert identified associates).
• Full Set, All Habitat Data
• Index Set, All Habitat Data
• Full Set, No Reviewer Habitat Data
• Index Set, No Reviewer Habitat Data
• Full Set, No Low Precision Habitat Data
• Index Set, No Low Precision Habitat Data
• Full Set, Plus Associate Species Data
• Index Set, Plus Associate Species Data

The inclusion of reviewer data, low precision data, and/or associate species data did not always improve the models. Model statistical scores could be high, indicating good fit to available data, but fail to match experts’ expectations based on knowledge of regional habitat variability. Also, low precision comments (large polygons) often included mixed habitat, introducing more noise than signal to the training data. Throughout the analysis, aattemp was made to balance consideration of both statistical and expert assessment.

Statistics to evaluate model performance:
• Sensitivity (SENS): The proportion of known habitat correctly identified as Potential Habitat (true positive rate).
• Specificity (SPEC): The proportion of known non-habitat (including pseudoabsence) correctly identified as Non-Habitat (true negative rate).
• Area Under the Curve (AUC): A summary of overall model performance based on both sensitivity and specificity.
• Cohen’s Kappa (KAPPA): A summary of overall model performance based on performance relative to a random classification.

Individual models were accepted and carried forward if they achieved values of 0.6 or greater for sensitivity, specificity, and area under the curve. For this species, the models accepted and averaged to produce the final map and statistics were: FullSet,
FullSetNoRev,
FullSetWithAssoc,
Indices,
IndicesNoRev,
IndicesWithAssoc

All statistical values are to be interpreted with caution as each has unique limitations. The most informative statistic depends on the intended application and the nature of the underlying data.

PROCESS CONTACT
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE point of contact

CONTACT INFORMATION ▶
PHONE
VOICE 919-707-6136

ADDRESS
TYPE
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US

E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

Hide Process step ▲

Hide Lineage ▲

Distribution ►

DISTRIBUTOR ►
CONTACT INFORMATION
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE distributor

CONTACT INFORMATION ►
PHONE
VOICE 919-707-6136

ADDRESS
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

Hide Distributor ▲

DISTRIBUTION FORMAT
* NAME File Geodatabase Feature Class
VERSION 10.5

Hide Distribution ▲

Fields ►

DETAILS FOR OBJECT YellowLancePotentialHabitat ►

* TYPE Feature Class
* ROW COUNT 8282

DEFINITION

High, moderate and low potential habitat locations for Yellow Lance in NC counties.

DEFINITION SOURCE

NCDOT

FIELD OBJECTID ►

ALIAS OBJECTID

* DATA TYPE OID
* WIDTH 4
* PRECISION 0
* SCALE 0
* FIELD DESCRIPTION

Internal feature number.

* DESCRIPTION SOURCE

Esri

* DESCRIPTION OF VALUES

Sequential unique whole numbers that are automatically generated.

Hide Field OBJECTID ▲

FIELD Shape ►

* ALIAS Shape
* DATA TYPE Geometry
* WIDTH 0
* PRECISION 0
* SCALE 0
* FIELD DESCRIPTION

Feature geometry.

* DESCRIPTION SOURCE

Esri

* DESCRIPTION OF VALUES

Coordinates defining the features.

Hide Field Shape ▲

FIELD Mean ►

* ALIAS Mean
* DATA TYPE Double
* WIDTH 8
* PRECISION 0
* SCALE 0

FIELD DESCRIPTION

Multiple random forest models were created using different data selection criteria (eg include or exclude low precision points, include or exclude reviewer desktop judgments, use raw environmental variables or composite indices, etc). Each model outputs a probability of suitable habitat value. Using all models passing statistical tests (>.6 for sensitivity, specificity, and Area under ROC Curve) Mean and Standard Deviation were calculated.

Based on the mean, the thresholds for the bins (Low, Moderate, High probability) were calculated.

DESCRIPTION SOURCE
NCDOT

Hide Field Mean ▲

FIELD SD ►
ALIAS StandardDeviation
* DATA TYPE Double
* WIDTH 8
* PRECISION 0
* SCALE 0
FIELD DESCRIPTION
Multiple random forest models were created using different data selection criteria (eg include or exclude low precision points, include or exclude reviewer desktop judgments, use raw environmental variables or composite indices, etc). Each model outputs a probability of suitable habitat value. Using all models passing statistical tests (>.6 for sensitivity, specificity, and Area under ROC Curve) Mean and Standard Deviation were calculated.

DESCRIPTION SOURCE
NCDOT

Hide Field SD ▲

FIELD Shape_Length ►
* ALIAS Shape_Length
* DATA TYPE Double
* WIDTH 8
* PRECISION 0
* SCALE 0
* FIELD DESCRIPTION
Length of feature in internal units.

* DESCRIPTION SOURCE
Esri

* DESCRIPTION OF VALUES
Positive real numbers that are automatically generated.

Hide Field Shape_Length ▲

FIELD Shape_Area ►
* ALIAS Shape_Area
* DATA TYPE Double
* WIDTH 8
* PRECISION 0
* SCALE 0

* FIELD DESCRIPTION
Area of feature in internal units squared.

* DESCRIPTION SOURCE
Esri

* DESCRIPTION OF VALUES
Positive real numbers that are automatically generated.

Hide Field Shape_Area ▲

FIELD PotHabitat ►
* ALIAS PotHabitat
* DATA TYPE String
* WIDTH 25
* PRECISION 0
* SCALE 0
FIELD DESCRIPTION
Model Output - Low, Moderate or High potential habitat

DESCRIPTION SOURCE
NCDOT

LIST OF VALUES
VALUE Low
DESCRIPTION Low probability of Habitat
ENUMERATED DOMAIN VALUE DEFINITION SOURCE NCDOT

VALUE Mod
DESCRIPTION Moderate probability of Habitat
ENUMERATED DOMAIN VALUE DEFINITION SOURCE NCDOT

VALUE High
DESCRIPTION High probability of Habitat
ENUMERATED DOMAIN VALUE DEFINITION SOURCE NCDOT

VALUE No data
DESCRIPTION No data available within USFWS area for species
ENUMERATED DOMAIN VALUE DEFINITION SOURCE NCDOT

Hide Field PotHabitat ▲

Hide Details for object YellowLancePotentialHabitat ▲

Hide Fields ▲

Metadata Details ►

METADATA LANGUAGE English (UNITED STATES)
METADATA CHARACTER SET utf8 - 8 bit UCS Transfer Format

SCOPE OF THE DATA DESCRIBED BY THE METADATA dataset
SCOPE NAME * dataset

* LAST UPDATE 2024-01-29

ARCGIS METADATA PROPERTIES
METADATA FORMAT ArcGIS 1.0
STANDARD OR PROFILE USED TO EDIT METADATA ISO19139
METADATA STYLE ISO 19139 Metadata Implementation Specification

CREATED IN ARCGIS FOR THE ITEM 2024-02-01 14:59:07
LAST MODIFIED IN ARCGIS FOR THE ITEM 2024-01-29 16:13:50

AUTOMATIC UPDATES
HAVE BEEN PERFORMED Yes
LAST UPDATE 2024-01-29 16:13:50

Hide Metadata Details ▲

Metadata Contacts ►

METADATA CONTACT
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit
CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE point of contact

CONTACT INFORMATION ►
PHONE
VOICE 919-707-6136

ADDRESS
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

Hide Metadata Contacts ▲

Metadata Maintenance ►

MAINTENANCE
UPDATE FREQUENCY as needed

OTHER MAINTENANCE REQUIREMENTS
Maintenance of this dataset is handled by the Environmental Analysis Unit (EAU) Mitigation and Modeling Unit. Support and maintenance of the enterprise spatial database where this data resides is handled by NCDIT's Transportation GIS Unit.

MAINTENANCE CONTACT
ORGANIZATION'S NAME North Carolina Department of Transportation - EAU Mitigation and Modeling Unit

CONTACT'S POSITION Environmental Program Consultant
CONTACT'S ROLE originator

CONTACT INFORMATION ►
PHONE
VOICE 919-707-6136

ADDRESS
DELIVERY POINT Century Center Building B, 1020 Birch Ridge Drive
CITY Raleigh
ADMINISTRATIVE AREA NC
POSTAL CODE 27610
COUNTRY US
E-MAIL ADDRESS ATLAS@ncdot.gov

HOURS OF SERVICE
9:00am – 5:00pm Monday - Friday

CONTACT INSTRUCTIONS
Please send an email with any issues, questions or comments regarding the ATLAS Data Search Tool, ATLAS Screening Tool or ATLAS Workbench. If it is an immediate need, please call the contact number or indicate as such in the subject line in an email.

Hide Contact information ▲

Hide Metadata Maintenance ▲

Metadata Constraints ►

SECURITY CONSTRAINTS
CLASSIFICATION unclassified
CLASSIFICATION SYSTEM None

LIMITATIONS OF USE
The North Carolina Department of Transportation shall not be held liable for any errors in this metadata. This includes errors of omission, commission, errors concerning the content of the data, and relative and positional accuracy of the data. This data cannot be construed to be a legal document. Primary sources from which this data was compiled must be consulted for verification of information contained in this data. Datasets developed under Project ATLAS do not replace any NRTR work for future projects and may not be used as a replacement for site visits / field surveys by qualified professionals and hence should be used only as a supporting platform for decision making. Use of this dataset for project scoping or screening is merely pre-decisional.

CONSTRAINTS
LIMITATIONS OF USE
The North Carolina Department of Transportation shall not be held liable for any errors in this metadata. This includes errors of omission, commission, errors concerning the content of the data, and relative and positional accuracy of the data. This data cannot be construed to be a legal document. Primary sources from which this data was compiled must be consulted for verification of information contained in this data.Datasets developed under Project ATLAS do not replace any NRTR work for future projects and may not be used as a replacement for site visits / field surveys by qualified professionals and hence should be used only as a supporting platform for decision making. Use of this dataset for project scoping or screening is merely pre-decisional.

Hide Metadata Constraints ▲