Bat Bridge Habitat Probability - June 2022 - NC Department of Transportation

File Geodatabase Feature Class

Tags

Bats, Bridge, random forest model, machine learning, Transportation, NRTR, NCDOT, Environment, Location, North Carolina, ATLAS, North Carolina

Summary

This dataset was originally created in April 2020 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.

Severe population declines in numerous North American bat species due to white-nose syndrome have prompted increased concern for monitoring bat populations at various life-history stages. There are also increased legislative protections for species at the state and federal levels to help limit impacts. As part of its broader effort to avoid and minimize impacts on federally listed bat species, NCDOT is improving their internal processes regarding species surveys and monitoring efforts. This effort is a key component of NCDOT's ATLAS program. In particular, NCDOT is using machine-learning models to refine species maps for federally listed species. Bats were originally excluded from NCDOT's machine learning models because the federally listed species had too few records to develop an adequate model; bat habitat usage is highly variable through seasons, life-history stages, and across species; and approved programmatic habitat conservation plans (HCPs) are already in place for much of the state and for most priority bat species. To provide linkage, automated, and screening model tools relevant to bat research needs and NCDOT efforts to streamline bridge repair and maintenance, a probabilistic model was developed (i.e., generalized linear model [GLM]) for bat-use in bridges. The GLM was originally developed from a dataset of 709 NCDOT bridges that had undergone bat habitat assessments and surveys looking for roosting bats. It was then modified based on expert review and field testing. The results of the GLM can be applied to a larger, statewide dataset of NCDOT bridges, and the outcome of that application has been applied as final shapefiles.

Because environmental data available for many of the 709 bridges were not available for other bridges throughout the state, engineered specifications that may serve as proxies for bat selection parameters in bridges were considered. These GLM inputs were developed with bat experts from non-government organizations (NGOs), federal agencies, NCDOT, and other state agencies. A refined list of structural variables was used to assess which, if any, were relevant in predicting bat roosting in bridges, how they contributed to bat roosting, and which model structure was the most effective in predicting potential bat roost use.

Preliminary results were shared with NCDOT and field testing was performed to quality control the GLM's efficacy. Field testing feedback was incorporated into the GLM and the model was re-scored and re-evaluated.

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.

Description

The Bat Bridge Habitat Probability dataset is a point layer depicting probable locations for bridges that contain Bat habitats in NC.

A targeted element of project ATLAS is modeling threatened and endangered species throughout the state to better understand how they interact with NCDOT infrastructure including bridges. It was proposed to develop a generalized linear model (GLM) to model bat roost-use in bridges in North Carolina. This proposal, approved by NCDOT, benefits the ATLAS program by increasing the currently limited information on bat roost-use in bridges and by allowing NCDOT to better screen and prioritize bridge survey efforts. In addition, the GLM could leverage existing bat data, existing bridge-survey data, and the existing NCDOT engineering data for bridges (WIGINS data).

Increasingly bats are being found using bridges and culverts throughout various seasons. Current research is focusing on micro-habitat quantification to understand why specific roosts are selected within a bridge, and this has fundamental ecological relevance and importance, particularly if transportation departments can use this data to incorporate bat roosting spaces as an onsite mitigation tool. However, models with such targeted data collection cannot easily be expanded to statewide probabilistic models which can have wide ranging utility for project screening and risk assessment for bridge repair and maintenance.

Although it is widely accepted that bats use bridges as roosting sites, little attention has been given to understanding the combined characteristics of bridge and habitat associated with their use of bridges as roosting sites in North Carolina. Studies on how engineered bridge features corresponded to bridge roost selection by bats were conducted. Because surveyed bridges included any sign of bats as a "used" bridge, this model considers any roost including short-term night roosts. A major goal of this research was to better understand when bridge replacement, repair, or rehabilitation projects have the potential for "taking" (i.e., harassing, injuring, or killing) bat species listed as federally threatened or endangered under the US Endangered Species Act (ESA 1973). The federally endangered gray bat (Myotis grisescens), federally threatened northern long-eared bat (M. septentrionalis), and the little brown bat (M. lucifugus) that is currently under review as a candidate species for listing, are all found in bridge structures in North Carolina.

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). Annual 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	-84.401717	**East**	-75.445724
North	36.604896	**South**	33.769572

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, 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 2020-04-23 00:00:00
PUBLICATION DATE 2022-06-17 00:00:00

Hide Thesaurus ▲

THEME KEYWORDS Bats, Bridge, random forest model, machine learning, Transportation, NRTR, NCDOT, Environment, Location, North Carolina, ATLAS, North Carolina

THESAURUS ►
TITLE User
CREATION DATE 2020-04-23 00:00:00
PUBLICATION DATE 2022-06-17 00:00:00

Hide Thesaurus ▲

Hide Topics and Keywords ▲

Citation ►

TITLE Bat Bridge Habitat Probability - June 2022 - NC Department of Transportation
CREATION DATE 2020-04-23 00:00:00
PUBLICATION DATE 2022-06-17 00:00:00

PRESENTATION FORMATS digital map
FGDC GEOSPATIAL PRESENTATION FORMAT vector digital data

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 point of contact

CONTACT INFORMATION ►
PHONE
VOICE 919-707-6136

ADDRESS
TYPE physical
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 originator

CONTACT INFORMATION ▶
PHONE
VOICE 919-707-6136

ADDRESS
TYPE physical
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
TYPE physical
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). Annual 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
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 2020-04-23 00:00:00
ENDING DATE 2020-04-23 00:00:00

EXTENT
GEOGRAPHIC EXTENT
BOUNDING RECTANGLE
EXTENT TYPE Extent used for searching
* WEST LONGITUDE -84.401717
* EAST LONGITUDE -75.445724
* NORTH LATITUDE 36.604896
* SOUTH LATITUDE 33.769572
* EXTENT CONTAINS THE RESOURCE Yes

EXTENT IN THE ITEM'S COORDINATE SYSTEM
* WEST LONGITUDE 412591.002996

* EAST LONGITUDE 3043914.441126
* SOUTH LATITUDE 48905.798709
* NORTH LATITUDE 1039124.545629
* 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 originator

CONTACT INFORMATION ►
PHONE
VOICE 919-707-6136

ADDRESS
TYPE physical
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 physical

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 BatBridgeHabitatProbability
* OBJECT TYPE point
* OBJECT COUNT 13387

Hide Vector ▲

ARCGIS FEATURE CLASS PROPERTIES ►
FEATURE CLASS NAME BatBridgeHabitatProbability
* FEATURE TYPE Simple
* GEOMETRY TYPE Point
* HAS TOPOLOGY FALSE
* FEATURE COUNT 13387
* 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 2020-04-23 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 2020-04-23 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

First Model Revisions:

The statewide GLM shapefile was reviewed on ArcGIS Online by species experts at the US Fish and Wildlife Service (USFWS), NGOs, and NCDOT. Their comments indicated that overall the GLM was performing well. Reviewers recommended changes making the GLM less conservative (increasing specificity) and incorporating bridge height and deck materials which were being poorly reflected in the GLM. The model was comprised of several variables queried from the WIGINS database, but bridge height was not available during this initial model revision.

Reviewers were encouraged to comment on agreement or disagreement with the GLM, the reasons why, and whether the model should be changed or retained at specific bridges. The GLM was adjusted from these comments by including main structure material and by including beam type and material, then it was rerun. We reapplied the new probability scoring to bridges across the state and saved the updated shapefile for implementation into the digital data collection framework designed by HDR and for ease of reference and review. The new data collection framework was an automated data collection protocol that calculated surrounding habitat for each bridge structure and auto-populated digital data forms for ecological data collection. This initiative was intended to streamline field data collection, ensure completed data forms, and facilitate future data modeling efforts during bat habitat assessments at bridges. The automated system was beta tested in the fall of 2019 and included a field test for the probabilistic GLM's accuracy and applicability. Field testing was completed by teams from Copperhead and HDR and informed the second model revisions.

Second Model Revisions:

Second model revisions incorporated a field test of the model and expert investigation of modeled bridge scoring. Increasing the GLM's specificity continued to be particularly relevant to wooden decked bridges where the GLM predicted high probability of bat use, but the field circumstances and past research have ruled out that likelihood. Further, field testing documented that source data being used for deck structures was incorrect, and a new query was initiated from the WIGINs database.

The WIGINS database queries are highly detail specific. We encountered two primary issues with the queries: incomplete data and incorrect data columns. Incomplete data were an issue when regional or county data were separated during a query, i.e., a formatting error, or when data were still being entered and updated. Through the process improvements within NCDOT, staff are continuing to collect data to fill known information gaps and the lag between these data updates can result in unpredictable gaps in the data. This required careful screening prior to a variable's use in the model resulting in not all variables being used. For example, data gaps on bridge height prevented use of this variable during our initial model runs, but they were entered in fall 2019 and were subsequently incorporated. The later inclusion of bridge height data helped improve our model during the second round of revisions. Another example of missing data was guardrail materials which were inconsistently provided so were omitted from the model.

One primary data column – Structure Type – was uploaded incorrectly in the initial model. Because the WIGINS database is organized with coded data labels and because the labels are not ecologically related, we initially requested structure material type instead of deck type to model the bridge deck materials. Structure Type consists of a 3-digit code that includes the span materials and span design (FHWA 1995); however, this differs from the WIGINS coding system and definition for deck type. Deck type is a variable that is developed and coded within NCDOT and is not part of the FHWA (Federal Highway Administration) guidance document. It is, however, included in the WIGINS reports. After review and field verification, deck type better reflected the surface that the bats contacted under

the bridges. By altering our query request, we improved our model, particularly pertaining to predictions of false positives on smaller wooden bridges. Accuracy increased on the field verified models recommended for change by 7%.

The final GLM incorporates repaired deck source data, newly available bridge height data, past reviewer comments, and comments from the field test. The results below include model specificity and sensitivity and some summary statistics for model improvements associated with revisions. The final GLM was a global model (i.e., includes all parameters and available for analysis) that was dredged using the "MuMIn" package in the R software environment. The top two reduced models were compared for the final model and are presented as "Reduced 1" and "Reduced 2" below. The GLM with the highest sensitivity was considered the most appropriate and was applied to all bridges within the state to produce a GIS shapefile for bat expert review. During GLM reduction, AIC was considered to identify the most parsimonious and powerful model within the ROSE model framework, not across model frameworks. Because parameters differed in their units, scales, and values, we scaled data prior to model development.

CONFORMANCE TEST RESULTS
TEST PASSED Yes
RESULT EXPLANATION
Pass

PRODUCT SPECIFICATION ►
TITLE NCDOT Geospatial Data Specifications
CREATION DATE 2020-04-23 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
Initial Model Development:

Original model development was completed using two primary data sources: bat habitat assessments (ecological data) in North Carolina and bridge engineering specifications (WIGINS). The ecological data-list was developed from bridges surveyed for bat use by trained wildlife biologists and literature review. The WIGINS database contains the engineered specifications for each bridge, including measurements and materials. Ecological data forms were considered to help develop corollaries between necessary ecological needs for bats and engineered WIGINS information to develop a robust GLM framework that could be applied statewide.

Bridge habitat assessments are carried out according to the NCDOT Preliminary Bat Habitat Assessments (Structures, Caves, & Mines) (NCDOT 2019) and include quantifying land cover, identification of bridge features likely important to bat species (e.g., structure material, crevices present), and presence of bats. Originally, we hoped to use the ecological data almost exclusively, but both datasets (ecological and WIGINS) included data gaps, superfluous information that would not inform the model, and differing structure identification numbers. Consequently, we collected remotely sensed data to augment ecological data. However, we still needed to link surveyed bridges from the existing ecological data to engineered data and remotely sensed data. Because these datasets were not designed to be integrated with each other and integration was complex, the project team (Jessica Tisdale [HDR], Richard Borthwick [Copperhead], and Eric Wilson [NCDIT]) met with NCDOT staff

(Walt Tallman) to understand the definitions of the engineered features and which features (as defined by an engineer) correspond to bat use features (as defined by a biologist). These were supported through expert opinion and literature review. Two data gaps posed challenges in coordinating the two datasets: incomplete bridge structure numbers and inaccurate or missing GPS locations. The following are the specific sources of these issues:

- Missing county code: All WIGINS structure numbers (i.e., unique bridge identification number) consist of six-digit codes with the leading two digits corresponding to a county code. In many of the habitat assessment structure numbers, the county code had been omitted and the structure number (consequently less than 6 digits) did not correspond with the WIGINS database. Both datasets also include the county name, so we used this data (where complete) to augment the structure number from the habitat assessment data with the appropriate code.
- Missing structure number: Without a structure number, the structure was untraceable through the WIGINS database. Where this occurred, GPS data were used to find "nearest neighbor" bridges in the WIGINS database.
- Missing GPS data: The lack of GPS data resulted in no geo-referenced information for the assessed bridge. If the structure number was complete, the bridge was directly linked to WIGINS through the structure number. Two bridges were missing GPS data and a structure number. These sites were deleted from the GLM training source data. Although they could have been manually identified through aerial photos on a map based on location descriptions, neither site had bats roosting in them and were not likely to contribute significantly to the GLM.

In addition to these issues, the coordinates provided from both datasets were not being displayed correctly when imported into GIS software (i.e., ArcMap or QGIS) (Ryan Dugger, HDR, Pers. Comm.). Therefore, bridge GPS locations from both datasets were added to Google Earth Pro, each dataset was exported as a *.KML file, and each *.KML was imported into ArcMap to spatially join the two *.KML files into a single shapefile (Appendix 1).

Because the habitat assessment data is dependent on a field visit and our intention was to test models that could be sourced from existing data and applied to bridges that have not yet been surveyed, we used ArcMap and remotely available data to help determine proxies for ecological variables. The remotely sensed data included the latest iteration of National Land Cover Database (NLCD) (Yang et al. 2018) to determine presence of water at each site and surrounding land cover, and we used the mineral resources data systems (MRDS) embedded in the ATLAS layer to determine cave or mine presence within 0.5-miles of the crossing structure. Due to missingness, we were forced to remove sun exposure, an important consideration for bridge-roosting bats.

Our team interpreted the WIGINS engineering data from 13,910 bridges to provide proxies for ecologically relevant variables. For example, joins and seals are entered as total lengths in unique columns depending on the join or seal type. However, because we considered these data a proxy for crevice availability in a bridge, we combined varied seal and join data into one column used as total crevice potential. Crevices have been shown to be important bat roost selection but identifying remote parameters to define crevices is challenging. In addition to seals and joins, we considered the number of spans and age of the bridge to further refine proxies for crevices.

After consolidation was complete, the data had an imbalance in bat presence ("1") and absence ("0"). Consequently, we ran a series of GLMs with various techniques to address the imbalance. This included zero-inflated GLMs through the "pscl" package in the R software environment (R Core Team 2019) and a random over-sampling example (ROSE) GLM using the "ROSE" package in the R software environment. Because model generation requires knowledge of use, inflated zeroes in a dataset (i.e., one case is rare) can result in a poor model fit or an over-fit predicting negative or absent data. Because a model tends to focus on prevalent data, a data imbalance can compromise model training and (due to scarcity of a class) model accuracy testing.

Several methods exist to address data imbalances, but there are some conflicting opinions about how to best address the issue. We tested three model variations to explore which improved model predictive efficacy. First, we ran a conventional model with a reduced negative sample using a modified Wilson's editing. This method generally requires a reduction of both positive and negative occurrences weighted more heavily on negative samples. However, due to the large imbalance in this dataset, meaningfully reducing zero occurrences was only possible if extraction included only negative data. This method resulted in an unquantifiable loss of data and a relatively low-quality model. Second, we ran a zero-inflated negative binomial (ZINB) model which attempts to account for the imbalance within the model framework. This output focused on the prevalent (i.e., negative) data, a well-documented shortcoming. Third, we considered a smoothed bootstrap form of re-sampling from

the data to develop a balanced dataset for model training (ROSE). This was the only oversampling effort used and we deemed it appropriate as we found that when the data imbalance was severe, oversampling was effective.

A ROSE model consists of a random oversampling of sites with replacement for the rare class, and without replacement for the common class. These methods help avoid the issue of missingness that arises from undersampling (i.e., removing zeroes), but increase the risk of over-fitting as sample units of the minority class are duplicated. To offset the issue of overfitting, we proposed randomly generating new sample unit observations based on nearest-neighbor values from existing sample points. The generation of new artificial data has the benefit of reducing over-fitting risks. However, there are inherent concerns in randomly generated new data. This has been subsequently studied for applicability and efficacy. Due to the unknowns in bat-roost selection, we opted to use the ROSE model as opposed to generating new data. This resulted in better predictive power than either negative data removal or ZINB models.

The model was used to estimate the probability of bats using a bridge for a roost. To create a confusion matrix of successful predictions, a probability threshold was set to allow for the assignment of a binary “yes” (i.e., “batty”) or “no” (i.e., “non-batty”) based on the estimated probabilities. Thresholds were adjusted when running each GLM so that the probability value was optimized. To complete this assessment, we tabled various probability thresholds and their respective confusion matrices to maximize both specificity and sensitivity. We completed this trial-and-error contrasting table for each model iteration as it was possible that the optimal threshold would change.

We used threshold assignments to calculate model specificity and sensitivity. Specificity is the true negative rate (number of predicted negatives divided by the total number of actual negatives) while sensitivity is the true positive rate (number of predicted positives divided by the total number of actual positives). Specificity is our ability to accurately predict bat absence from bridges while sensitivity is our ability to accurately predict bat presence. For our purposes, sensitivity is more important as we do not want to miss bat bridges. As sensitivity increases, our rate of false positives also increases, resulting in over-prediction of bat presence. The optimal model was conservative in nature and low specificity meant that we were routinely predicting “non-batty” bridges as “batty” when using a 32% threshold. For both zero-inflated and ROSE model structures, specificity and sensitivity were considered to evaluate which method of addressing a zero-inflated dataset predicted most effectively. Akaike’s Information Criteria (AIC) or Bayesian Information Criteria (BIC) are generally used to compare model structures because, when the response variable remains the same, these criteria indicate improvements in model fit as explanatory variables are altered. However, our model iterations altered the response variable distribution, so an alternative model scoring framework was required. Predicted outputs and probability scores were calculated for the statewide data, saved as a shapefile, and uploaded to the HDR and NCDOT Sharepoint hub for file storage.

June 2022 Update: Binary probability was replaced with High/Low probability for Bat presence. Field was renamed to HabitatProb

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 originator

CONTACT INFORMATION ▶
PHONE
VOICE 919-707-6136

ADDRESS
TYPE physical

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
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 originator

CONTACT INFORMATION ►
PHONE
VOICE 919-707-6136

ADDRESS
TYPE physical
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
TYPE physical
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 BatBridgeHabitatProbability ►
* TYPE Feature Class
* ROW COUNT 13387
DEFINITION
Locations for Bats in Bridges in NC

DEFINITION SOURCE
NCDOT

FIELD OBJECTID ►
* ALIAS FID
* 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 StrNo ►
* ALIAS StructureNumber
* DATA TYPE String
* WIDTH 80
* PRECISION 0
* SCALE 0
FIELD DESCRIPTION
  Bridge Structure Number

DESCRIPTION SOURCE
  NCDOT

Hide Field StrNo ▲

FIELD Prob ►
* ALIAS Probability
* DATA TYPE Double
* WIDTH 8
* PRECISION 0
* SCALE 0
FIELD DESCRIPTION
  Bat presence probability

DESCRIPTION SOURCE
NCDOT

Hide Field Prob ▲

FIELD HabitatProb ►
* ALIAS HabitatProbability
* DATA TYPE String
* WIDTH 10
* PRECISION 0
* SCALE 0
FIELD DESCRIPTION
Probability of Bat Habitat

DESCRIPTION SOURCE
NCDOT

LIST OF VALUES
VALUE High
DESCRIPTION High Probability of Bat presence
ENUMERATED DOMAIN VALUE DEFINITION SOURCE NCDOT

VALUE Low
DESCRIPTION Low Probability of Bat presence
ENUMERATED DOMAIN VALUE DEFINITION SOURCE NCDOT

Hide Field HabitatProb ▲

Hide Details for object BatBridgeHabitatProbability ▲

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 14:11:33
LAST MODIFIED IN ARCGIS FOR THE ITEM 2024-01-29 15:33:43

AUTOMATIC UPDATES
HAVE BEEN PERFORMED Yes
LAST UPDATE 2024-01-29 15:33:43

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
TYPE physical
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
Annual 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
TYPE physical
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 ▲