Analysis of Animal-Vehicle Collisions on Florida Roads Using the Getis-Ord Gi* Spatial Statistic Tool in ArcGIS

Large Animal Hotspot Analysis

Final

10/19/21
revised 3/4/22

Deliverable 3b --- Final Model Results

Project No. BDV24-TWO985-04

Daniel J. Smith, Ph.D., A.I.C.P., Principal Investigator
Department of Biology

Crystal Gagne, M.Sc. Student
Computer Science Department

Jennifer Rote, Biological Technician
Department of Biology

University of Central Florida

Analysis of Animal-Vehicle Collisions on Florida Roads Using the Getis-Ord Gi* Spatial Statistic Tool in ArcGIS

Introduction

The Florida Department of Transportation (FDOT), along with resource agencies and non-profit organizations (NGOs) have expressed an interest in furthering conservation initiatives to protect Florida wildlife by identifying and prioritizing habitat corridors where wildlife-highway conflicts are likely to occur. FDOT strives to improve highway safety and one measure that would facilitate this objective would be to identify wildlife-highway conflicts in advance so that wildlife crossing considerations can be included in early transportation project planning. A decision-support model using geographic information systems (GIS) was created for this purpose. Kernel density estimate and hotspots analysis techniques performed on animal-vehicle-collision (AVC) and species observational data along with updated target species habitat models and landscape data were used to identify and prioritize locations for potential mitigation (such as wildlife crossing structures) to improve highway safety and to facilitate conservation objectives.

Data and Analytical Methods

Large Animal Collision Hotspots. We evaluated available crash data for 4 large animal species included as part of the target species list assembled in collaboration with a Technical Advisory Group comprised of resource and transportation agency and NGO representatives (see Appendix for target species list and TAG representation). The four species included Florida panther, Florida black bear, white-tail deer, and American alligator.

<u>Available Data</u>:

A search of available data revealed high quality datasets for Florida panther and Florida black bear maintained by the Florida Fish and Wildlife Conservation Commission (FWC). These two datasets include crash locations in the state from the 1970s to the present. We were unable to obtain similar data for the American alligator or white-tail deer. Road mortality data was available for Florida Key deer and American crocodile maintained by FWC and the US Fish and Wildlife Service (FWS). These are both isolated populations and were not included in the analysis at this time, though they could be analyzed separately to identify significant crash hotspots using the same method in this study.

Alternatively, we considered observation data from sources like iNaturalist, Bison, FNAI and FWC, however the number of recorded observations were minimal for white-tail deer and American alligator and not representative of the population size and widespread distribution of these species in the state. This is likely because they are common, and observers do not readily or consistently record or submit observations of these species to these catalogs. Therefore, we considered this data unreliable for identifying crash hotspots.

As a substitute we turned to the FDOT-maintained Crash Analysis Reporting System. This is a digital database of all highway crash reports submitted by law enforcement (referred to in this document as “crash data”). The crash data is available in GIS formats, however digital records only date back to 2012. We identified applicable collision data using two filters. One filter was listed under “Environmental Condition 1, 2, or 3” and labeled “Animal(s) in Roadway”. This allowed us to download all collisions that were, at least partially, caused by an animal being in the roadway. The second filter was listed under “Harmful Event” and labeled “Animal”. This pulled out all collisions due to a vehicle striking an animal. These filters contained a significant amount of overlap and duplicates were removed prior to inclusion in our final AVC database. Unfortunately, species or type of animal is not included in the database. This information may be available in written reports associated with each record, but it was not feasible for us to explore this as many thousands of records are involved.

Since species or animal type is not specified, certain assumptions were made regarding suitability of the crash data as a substitute data source with the absence of available white-tail deer and American alligator collision data in the large animal road hotspots analysis. For example, we suggest that most reported crashes (in the two selected categories above) were associated with significant damage to vehicles and/or resulted in human injury. Therefore, it is probable that most of these incidents involved a large animal such as white-tail deer, wild pig, American alligator, Florida panther or Florida black bear. Another assumption was that a certain component of the crash data may be due to collisions with livestock, particularly in rural areas. We predict this may be a small percentage, but nonetheless, are also likely to be large animals and are thus relevant for inclusion in the large animal hotspots model.

<u>Data Preparation:</u>

Several steps were followed to adapt the data for use in the hotspots model. This included a few housekeeping tasks required to eliminate duplication and to refine accuracy/suitability of the data.

- First, we removed all locations in the crash data that occurred on urban roads (per FDOT classification). This was necessary to remove the high number of records that are likely tied to collisions with domestic animals, e.g., dogs and cats. These records would disproportionately skew the results of the analysis to urban settings and is not consistent with the objectives of the project—targeting wildlife collision hotspots and opportunities where wildlife crossing structure solutions are applicable. Limiting the crash data to rural roads only eliminated this confounding factor.

- Second, all duplicates in the crash data corresponding to panther or bear collisions were removed.

- Third, the panther and bear roadkill locations as well as the crash data were further filtered in relation to the presence of existing wildlife crossings and wildlife fencing. Any AVCs that occurred prior to the date that wildlife crossings or fencing was installed (and located within the fenced section) were removed from the analysis. Those that occurred after the date wildlife crossings or fencing were installed were retained.

The next step was to create a segmented roads layer to merge with the AVC location point data. This step was necessary to remove unwanted influence of nearest neighbors from nearby or intersecting roads (a typical effect of Kernel Density or Hotspots Analysis with point data). The segmented road layer also provides a standard length of roadway for defining hotspots (we chose one linear mile intervals). Using the most recent FDOT data layers of paved federal, state and county roads with traffic counts (Annual Average Daily Traffic, AADT 2019), we performed a dissolve to eliminate previous segmentation and then created points at one-mile intervals. Using the 1 mi interval points, we split the dissolved roads layer creating 1 mi segments. Next, using the buffer tool, we created a 200 m wide buffer layer. Special treatments (using multiple operations in ArcGIS) were required to remove overlaps and recreate road intersections. A spatial join to the original roads layer was then required to repopulate the original roads attribute data. The result was a suitable base layer of buffered 1 mi road segments.

To add the collision data, we created a new field in the attribute table of each of the AVC point data layers (panther, bear, crash data [all current through 2019]). For each corresponding dataset, the new *integer* field was labeled as either “FP_count”, “BB_Count”, or “CD_Count”, respectively and the value for each record was set to “1”. All three AVC layers were then merged. The merged AVC data layer was then spatially joined to the one-mile segments. Using the join function in “Joins and Relates”, the sum of collisions (by type) per dataset was calculated, resulting in a layer that contained the number of AVCs per mile segment. We also created a sum field (SUM_All3) and calculated the total combined AVCs for each segment. The “Count” fields by AVC type provide the ability to assign different weighting factors to the panther, bear and crash data and perform calculations to adjust relative level of importance (could be associated with risk to listed species and/or potential crash severity).

Hotspots Analysis Using the Getis-Ord Gi* Statistic:

The Getis-Ord Gi* tool in ArcGIS (ESRI version 10.6), given a set of weighted features, identifies statistically significant hotspots. It uses the number of incidents (in our case, animal-vehicle collisions per segment), the distance from the mean, and the value of neighboring segments in calculating relative significance. There are three parameters of the tool that require input to run this analysis (below are the settings we applied):

- Conceptualization of Spatial Relationships – Describes how the features interact with one another in space. We selected the option INVERSE DISTANCE (nearby neighboring features have a larger influence on the computations for a target feature than features more distant). This was important because it ensured that the routine prioritized adjoining or near clusters of AVCs.
- Distance Calculation Method – We chose the standard setting, EUCLIDEAN DISTANCE (the straight-line distance between two features).
- Threshold Distance Band – Specifies a “cut-off distance” for which features outside the threshold are not considered for that feature. We chose the DEFAULT, which is the calculated *Euclidean distance* that ensures every feature has at least one neighbor.

In combination with the inverse distance parameter, using the default distance band ensured that all data was accounted for in the analysis and that no features were left out. Because of this, the default distance band produced the best results.

We performed the Getis Ord Gi* analysis on the three applicable sets of AVC data we obtained (Florida panther, Florida black, FDOT crash data). We ran models on the *merged, spatially joined AVC data of one-mile segments* at two separate scales, statewide and by FDOT District. For all datasets, the analysis was limited to road segments that contained at least one mortality.

At both the statewide and district level, the analysis was performed using decimal weights of 1.4 for panthers, 1.2 for bears, and 1.0 for the crash data. The weighting factors were used to elevate the significance of the Florida panther as a federally endangered species and the Florida black bear as an important and at-risk management species (both populations vulnerable in part due to an exceptionally high rate of vehicle-related deaths). The other reason the weightings were necessary was to control for the much greater abundance of crash data occurrences. We considered several weighting factors before choosing 1.4, 1.2, and 1.0 as the most balanced (other weighting factors considered: integer weight ratios -- 5:3:1, 4:2:1, 3:2:1 and decimal weight ratios -- 1.3, 1.2, and 1.0). The 40% (1.4) and 20% (1.2) increase in value of panthers and bears, respectively resulted in no significant panther or bear hotspots from independent panther and bear model runs from being lost in the combined results. In addition, several crash data hotspots were added to the total number of significant segments.

As referred to above, we also performed the hotspots analysis on independent panther and bear data in FDOT districts where the number of collisions for each was substantial. For District 1 we analyzed the unweighted panther mortalities. For Districts 2, 3, and 5 the hotspot analysis was performed on unweighted bear mortalities. In districts where panther and bear road mortalities were too few to perform the Getis Ord Gi* statistical analysis, panther, bear and crash data were combined.

Results

Large Animal Collision Hotspots. The Getis-Ord Gi* analysis results are presented in three tiers of significance: 99%, 95% and 90%. The table below provides the p value and Z score differences. It is a cluster analysis technique that evaluates the probability of spatial randomness. In relative terms, the 99% tier is 10 times more likely to be significant than the 90% tier. Road segments within the 90% level have a 1 in 10 chance of being random, and thus are only probable as hotspot clusters, rather than being statistically significant. All road segments below the 90% confidence tier were deemed statistically insignificant.

z-score (Standard Deviations)	p-value (Probability)	Confidence level
< -1.65 or > +1.65	< 0.10	90%
< -1.96 or > +1.96	< 0.05	95%
< -2.58 or > +2.58	< 0.01	99%

from ESRI - How Hot Spot Analysis (Getis-Ord Gi) works—ArcGIS Pro | Documentation

The data is presented in multiple forms. First, this package includes ArcGIS geodatabases at the state and district scales. Second is a set of the same spatial results in KMZ format for use with Google Earth. Third, is a set of pdf summary data tables with map images.

The data tables contain several fields (see Appendix) that represent the computations of the hotspot analysis; in addition are several fields that provide context to- and evaluation of- these locations such as conservation land status, Florida greenways network designation, panther and black bear habitat and range, and FDOT road projects (in early planning stages).

<u>Statewide Model Summary</u>

At the statewide scale, there were 7,401 1-mi segments that had at least one AVC (see Appendix C folder – excel tables and pdf files: statewide). Of these, 424 road segments had values above 90% confidence level; 195 of these were in the 99% category, 110 were within the 95% category and the remaining 119 were in the 90% category. All 99% tier road segment hotspots were found within Districts 1, 2, 3, 5, and 7. These districts strongly reflect the presence of weighted black bear or panther data as well as highly significant crash data locations. At the statewide scale, only five other significant hotspots were identified outside of these districts, all five in District 7.

As might be expected, clusters of hotspots were most notable in Ocala National Forest and other black bear population concentration areas and within Collier and Hendry Counties where most panthers are found.

<u>District One Summary</u>

Two versions of the hotspot analysis were performed for District One, a combined, weighted dataset, and a panther only dataset. For the combined, weighted dataset, a total of 86 significant hotspots were identified, 42 in the 99% category, 34 in the 95% category and 10 in the 90% category (Appendix C – pdf). The hotspots, indicative of panther data, are mostly associated with roads in Collier and Hendry counties. Another notable cluster of hotspots is within Highlands and Glades Counties on the US 27 and SR 70 corridors where numerous bear fatalities have occurred as well as a few panther fatalities.

The panther-only analysis identified hotspots on mostly the same roads as the PRIT (Panther recovery Implementation Team) Transportation SubTeam Hotspots Analysis but does exhibit notable differences in the level of significance of various hotspot locations. It includes 8 and 5 hotspots at the 99% and 95% confidence levels, respectively.

<u>District Two Summary</u>

We performed two hotspot analyses in District Two, a combined, weighted dataset, and a bear only dataset. For the combined, weighted dataset, there were 130 hotspots (75 – 99%, 48 – 95%, 7 – 90%). Although bears contributed to this high number of significant hotspot segments, when compared to the bear only analysis results it is apparent that FDOT crash data also had an impact (Appendix C – pdf maps). Most concentrations are within existing or proposed conservation lands and priority one or two greenways and in abundant or common bear range and high-quality bear habitat areas. Several hotspot locations appear to coincide with agricultural and edge habitats, which would also translate to the common occurrence of white-tail deer.

In the bear model for District Two, 20 total hotspots were identified (8 – 99% confidence tier, 3 – 95% confidence tier, and 9 – 90% confidence tier). These were concentrated on roads near the Aucilla WMA, near the Suwannee River and Osceola NF, and the landscape linkage from Ocala NF to Camp Blanding (a composite of different existing and proposed conservation lands).

<u>District Three Summary</u>

As in District Two, the same two datasets were analyzed for District Three. In the combined, weighted data hotspots model, 129 AVC hotspots were identified (55 – 99%, 38 – 95%, 36 – 90%). In the 99% confidence tier, most hotspot locations were within abundant or common bear range, 85.5% were within Priority 1 or 2 greenways, and 56.4% were within existing conservation lands (Appendix C – pdf). Contextual characteristics were similar for hotspot locations, most were found within large conservation areas or priority 1 or 2 greenways. Hotspot concentrations are notable on roads within Eglin and Tindal AFBs, an expansive, commercially owned, forested area east of Panama City, US 98 within Apalachicola NF, Tates Hell SF and St. Marks NWR, several roads in public lands and natural areas to the west, south and east of Tallahassee, and US 27 and US 98 within the Aucilla WMA area.

The District Three black bear hotspots model identified 38 road segments, 23 in the 99% confidence tier and 15 in the 95% confidence tier. Their locations are in the same areas as discussed above. A comparison of the two models for District Three shows that the FDOT crash data enhances and expands the hotspots identified in the bear model. This is consistent for both bear and deer presence, given that all areas are primarily forested natural areas.

District Four Summary

Due to the minimal numbers of panther and bear collisions in District Four, we only conducted the hotspots analysis on a combined, weighted dataset. The results included a total of 16 hotspots, 7 in the 99% category and 9 in the 95% category. Sample size (total n=322 AVCs) was small when compared to other districts. Notable locations included I-75 near the Broward-Collier county line (included one panther and multiple bear collisions), SR 70 and the Turnpike near the northwest corner of St. Lucie County, respectively (see Appendix C, pdf maps).

District Five Summary

In District Five, we completed the hotspot analyses on the combined, weighted dataset and the bear only dataset. For the combined, weighted dataset, there was a total of 84 hotspots (43 – 99%, 21 – 95%, 20 – 90%). Although bears contributed to this high number of significant hotspot segments, when compared to the bear only analysis results it is apparent that FDOT crash data enhanced the results (Appendix C – pdf maps). All hotspots were concentrated on roads within the greater Ocala National Forest ecosystem, and nearly all were within existing or proposed conservation lands, priority one or two greenways and in abundant or common bear range and high-quality bear habitat areas. Concentrations of white-tail deer are also high within this area.

In the bear model for District Five, 44 total hotspots were identified (24 – 99% confidence tier, 14 – 95% confidence tier, 6 – 90% confidence tier). Locations are essentially the same as described above. It is noteworthy to point out that segments of two roads that were identified as hotspots in the results are part of the Wekiva Parkway Project currently under construction (SR 49 and CR 46a) and are being mitigated for AVC impacts.

District Six Summary

As with District Four, sample size was quite small in District Six (total n = 47 AVCs). The hotspot analysis is questionable in this case with a mean of slightly above one collision/segment. Even though the model ran and produced results, only four hotspots were identified (at the 99% confidence level) with an n=2 AVCs for each hotspot (the location, on US 41 just east of SR 997, included one panther collision). The hotspot on US 1 on Big Pine Key is located outside the fenced area associated with existing wildlife crossings constructed for the Florida key deer.

District Seven Summary

Because of minimal numbers of panther and bear collisions in District Seven, we only conducted the hotspots analysis on a combined, weighted dataset. The results included a total of 24 hotspots, 12 in the 99% category and 12 in the 95% category. Sample size (total n=586 AVCs) was smaller in comparison to most other districts. There were 15 bear collisions included in the analysis in District Seven. Three of the hotspot segments involved multiple bear mortalities. Most hotspots were located within existing conservation lands, however coincidence with greenways and bear range/habitat is more variable (Appendix C – pdf). Notable locations included SR 44 near Withlacoochee River and Half Moon WMA, various roads within Weekiwatchee Preserve/Chassahowitzka NWR, and roads bordering Starkey Wilderness Preserve, Cross Bar Ranch and Conner Preserve (see Appendix C, pdf maps).

Appendix

A key to the data layer attribute tables (in alphabetical order):

- AADT_2019 – Annual Average Daily Traffic in 2019; if 0, indicates no value available
- Abundant_or_Common_Bear_Range - Is the segment in the abundant or common black bear range?
- Black_Bear_AVC_Count - Number of collisions from the black bear roadkill dataset, per individual segment
- Census_2019 – Census Block designation in 2019 where road segment is located
- Confidence_Level - Confidence Level
- County - Florida Counties
- Description_of_Functional_Class - Description of FDOT functional roadway classification
- District - FDOT Management District #s
- Existing_Conservation_Land - Is the segment in an existing conservation area?
- FID - Record number
- Florida_Panther_AVC_Count - Number of collisions from the panther roadkill dataset, per individual segment
- Florida_Panther_Habitat - Is the segment in panther habitat?
- Functional_Classification_Code – FDOT functional roadway classification codes
- Gi_Bin - Confidence level tiers used in the results (99%, 95%, 90%, <90%)
- Gi_PValue - Getis-Ord Gi generated P value
- Gi_ZScore - Getis-Ord Gi generated Z score
- Jurisdiction - Agency responsible for construction/maintenance
- Name - Road name and/or number
- Nearest_Neighbors - Number of nearby neighbors used in the analysis
- Number_of_Lanes - Number of road lanes

- Occasional_Black_Bear_Range - Is the segment in the occasional black bear range?
- Primary_Florida_Panther_Range - Is the segment in primary panther range?
- Priority_1_or_2_Greenway - Is the segment in a priority 1 or 2 greenway?
- Priority_3_to_6_Greenways - Is the segment in a priority 3 (or more) greenway?
- Proposed_Conservation_Land - Is the segment in a proposed conservation area?
- Road_Description - Road name/number and description
- Road_Project - Is the segment in an FDOT planned road project (early or pre-process)?
- Road_Status - Status under the State Highway System
- Roadway - FDOT roadway identifier
- Route_Number – State/county road number designation
- Rural_Crash_Data_AVC_Count - Number of collisions/segment from the FDOT crash dataset (rural areas only)
- Secondary_Florida_Panther_Range - Is the segment in secondary (or other) panther range?
- Speed_Limit – posted speed limit; if 0, indicates no value available
- State_and_BMU_Bear_Habitat - Is the segment in BOTH the state and BMU black bear habitat?
- State_or_BMU_Bear_Habitat - Is the segment in EITHER the state or BMU black bear habitat (not both)?
- Total_AVC_Count - Combined number of animal vehicle collisions, per individual segment from all three data sources (crash data-FDOT, black bear-FWC, panther-FWC)
- US_Route - US highway number
- Weighted_AVC_Count – Weighted count (crash data wgt = 1, black bear wgt = 1.2, panther wgt = 1.4)