Identifying Important Roadway Corridor Factors Which Affect Crashes Influenced by Unsafe Driving Behaviors to Help Law Enforcement Proactively Reduce Crashes

John McCombs, University of Central Florida (UCF) Ph.D. Candidate

Haitham Al-Deek, Ph.D., P.E., UCF Professor of Engineering

Adrian Sandt, Ph.D., UCF Research Associate

Grady Carrick, Ph.D., Enforcement Engineering, Inc.

TRBAM Paper 24-01040

Introduction

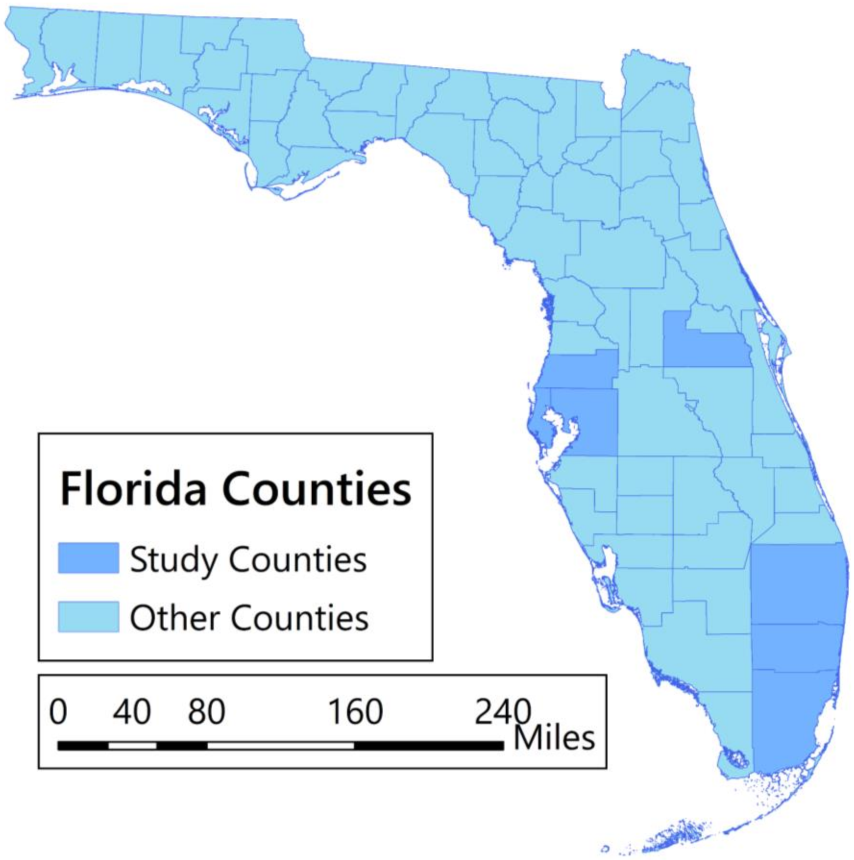
- Limited traffic enforcement staffing in recent years.
- Rise in fatal and injury (FI) crashes and unsafe driving behaviors since COVID-19.
- Proactive enforcement can help prevent crashes, but knowing the best locations for enforcement can be challenging.
- Identifying roadway factors which affect crashes involving unsafe driving behaviors can help law enforcement allocate their officers effectively.

Research Problem and Goal

- Main goal: Identify important roadway corridor factors which affect crash rates for unsafe driving behaviors.
- Factors identified via corridor-level random forest regression models using data from Florida arterials.
- This research can help law enforcement agencies prioritize locations for adjustments to enforcement to prevent crashes influenced by unsafe driving behaviors.

Data Scope

- Data retrieved for urban and suburban arterial roadways in seven Florida counties.
  - Orange, Hillsborough, Pinellas, Pasco, Palm Beach, Broward, and Miami-Dade.
  - Previously identified by the Florida Department of Transportation (FDOT) as high-risk with many FI crashes.
  - Likely have the most potential for crash reduction through proactive enforcement.



Roadway Data

- Geographic information system (GIS) files from FDOT.
  - Inventoried roadway data: lane count, traffic signal locations, functional classification, context classification.
  - Operational data: historical annual average daily traffic (AADT), speed limits.
- Google Maps satellite and Street View imagery.
  - Intersection counts, primary median type.
  - Presence of school zones, railroad crossings, limited access facility exits or entrances, bus stops, lighting, bicycle lanes, sidewalks, crosswalks, midblock crossings.

Citation Data

- Law enforcement citation data from 2017-2021 for seven unsafe driving behaviors.
- Citations related to non-driver behaviors or complex behaviors were not considered.

1. Failure to obey a traffic control device
2. Failure to stop at a steady red signal
3. Failure to drive in a single lane
4. Failure to yield while turning left
5. Driving under the influence
6. Reckless driving
7. Careless driving

Corridor Identification

- Corridors identified using geoprocessing tools in ArcMap.
  - Primary features (cannot change within a corridor): context classification, lane count.
  - Additional features: at least four years of AADT data, at least one lane in each direction, even number of lanes, contain at least one signalized intersection, at least 0.5 miles long.
- 548 corridors identified using this corridor definition.
- After filtering for corridors with known enforcement and removing outliers, 406 corridors covering over 800 centerline miles of arterial roadway were retained.

Corridor Summary Statistics

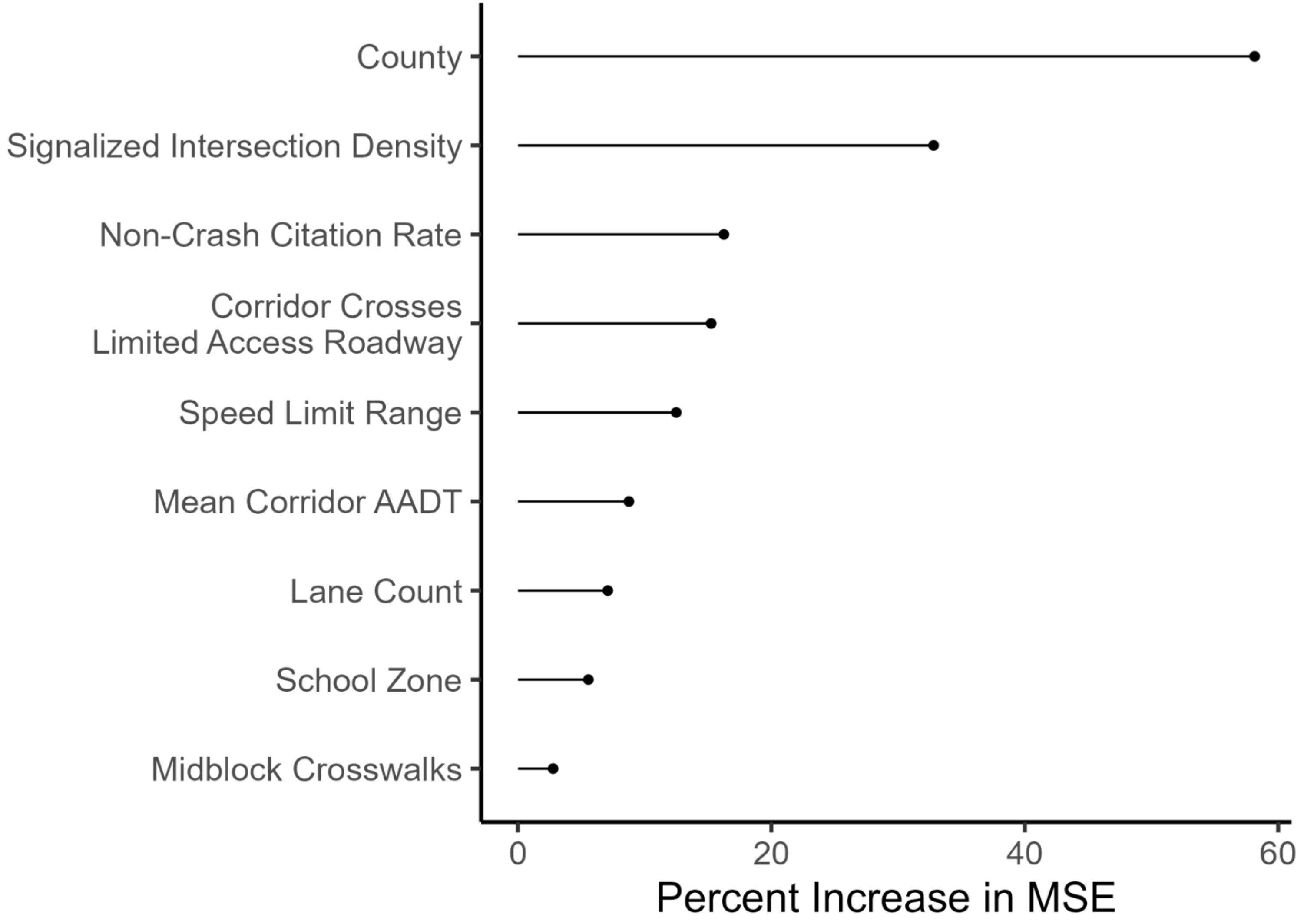
- 406 corridors contained 21,876 citations for the seven considered violations during the study period.
- Over two-thirds of the total citations were issued for careless driving.
- 85% of citations (18,518) were related to a crash.
  - 27% of these crashes (5,053) were FI crashes.
- Mean corridor traffic volume of about 37,500 vehicles/day.

Random Forest Models

- Variable importance plots (VIPs) show the order of important variables in a model.
  - Higher percent increase in mean squared error (MSE) means greater importance for prediction.
- Partial dependence plots (PDPs) show the marginal effect of one variable on the predicted value of the response.
- Three models developed:
  - Total crash rate (model 1).
  - FI crash rate (model 2).
  - FI crash rate for careless and reckless driving (model 3).

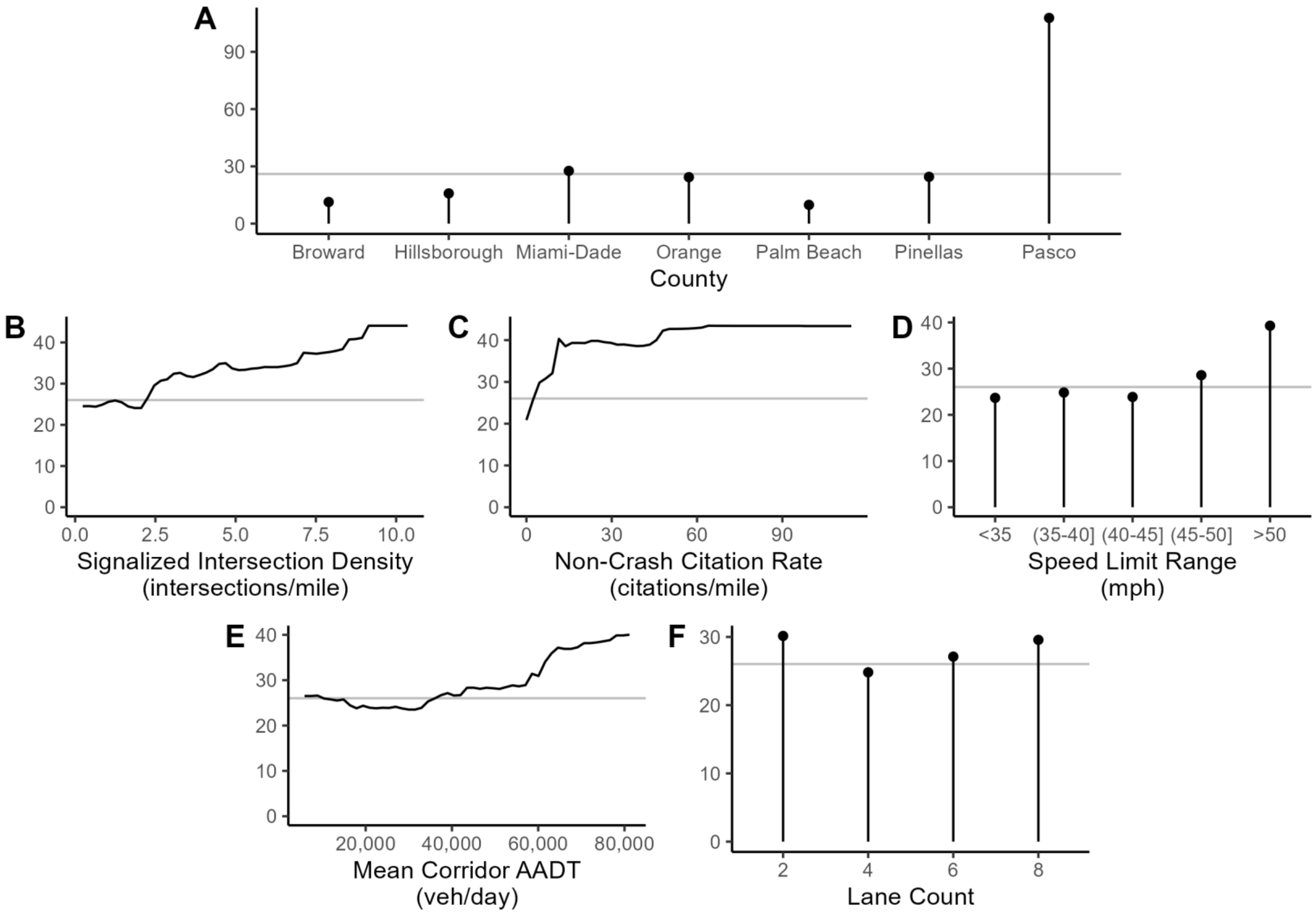
Model 1 Results & VIP

- 406 corridors containing 18,518 crashes.
- Pseudo R^2 = 45.96%.
- Four variables used at each tree split.
- Most important variables: County, Signalized Intersection Density.



Model 1 PDPs

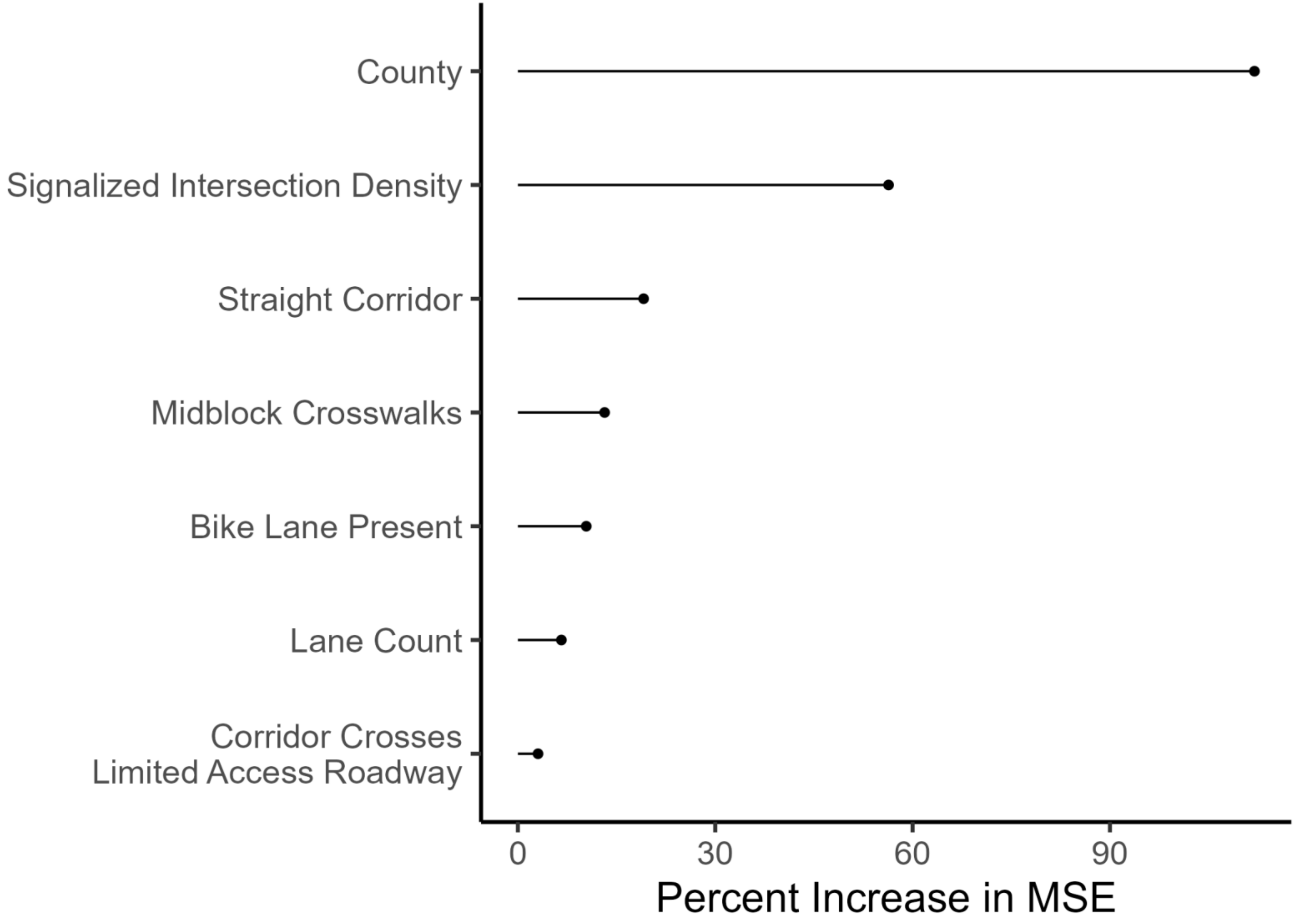
- Attributes of corridors likely to benefit from proactive enforcement:
  - In Pasco County.
  - High signalized intersection densities.
  - Contain limited access exit or entrance.
  - Speed limits greater than 45 mph.
  - Mean corridor traffic volumes greater than 60,000 vehicles/day.



Civil, Environmental and Construction Engineering
UCF
UNIVERSITY OF CENTRAL FLORIDA

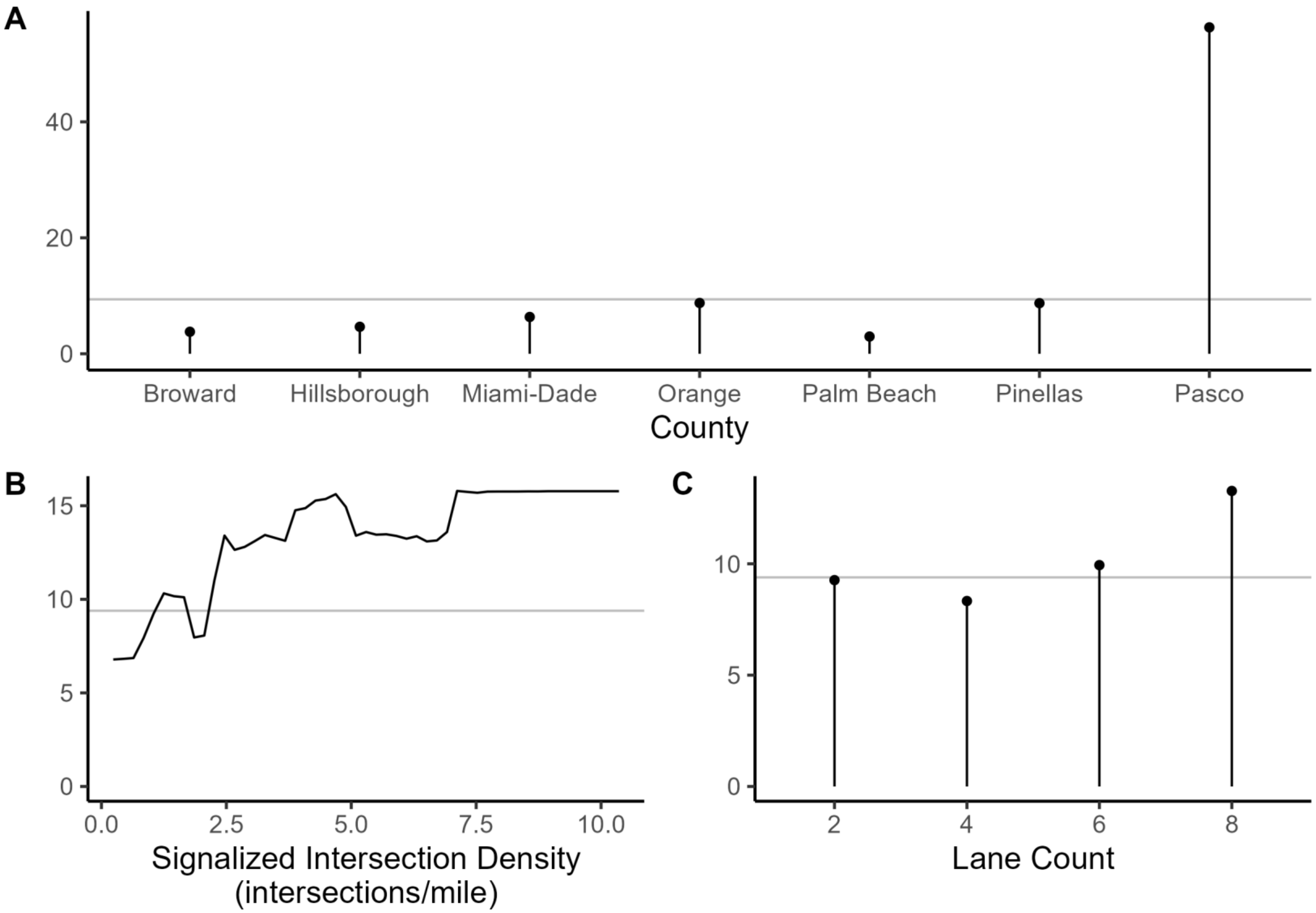
Model 2 Results & VIP

- 314 corridors containing 5,053 FI crashes.
- Pseudo R^2 = 57.02%.
- Seven variables used at each tree split.
- Straight corridors and corridors with bicycle lanes now identified as important.



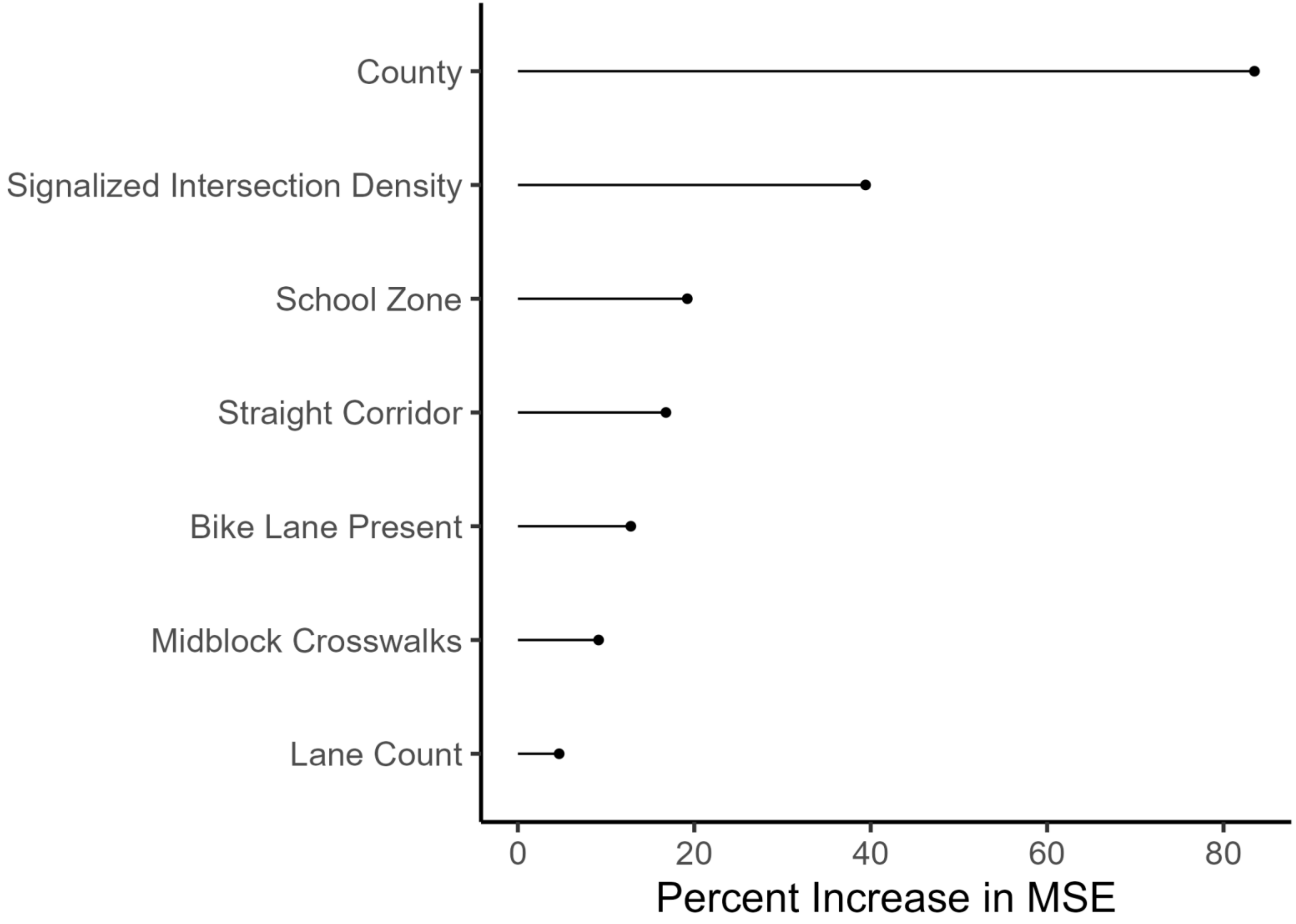
Model 2 PDPs

- Attributes of corridors likely to benefit from proactive enforcement:
  - In Pasco County.
  - High signalized intersection densities.
  - Limited horizontal curvature.
  - Contain midblock crossings.
  - Contain bicycle lanes.
  - Have six or more lanes.



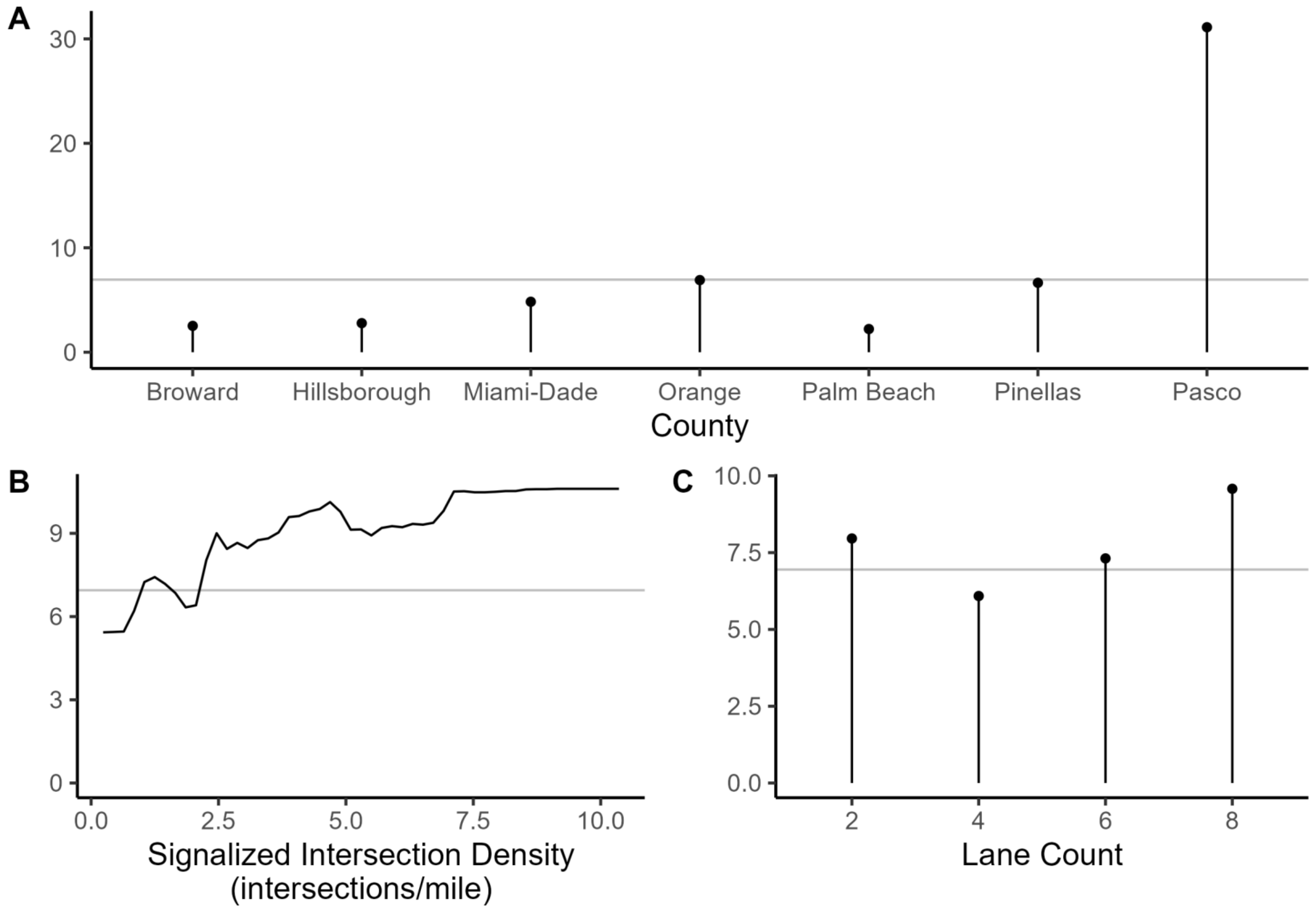
Model 3 Results & VIP

- 289 corridors containing 3,467 FI crashes that involved careless or reckless driving.
- Pseudo R^2 = 59.38%.
- Four variables used at each tree split.
- School zone presence identified as more important in Model 3 than in Model 1.



Model 3 PDPs

- Attributes of corridors likely to benefit from proactive enforcement:
  - In Pasco County.
  - High signalized intersection densities.
  - Do not contain a school zone.
  - Limited horizontal curvature.
  - Contain bicycle lanes.
  - Contain midblock crossings.
  - Have either two lanes or more than four lanes.



Civil, Environmental and Construction Engineering
UCF
UNIVERSITY OF CENTRAL FLORIDA

Summary and Conclusions

- Using random forests with law enforcement data can help agencies with limited staffing prioritize locations for proactive enforcement.
  - Important variables show which features can affect crash frequency due to unsafe driving behaviors.
  - Random forests can capture non-linear effects typical models cannot.
- County and signalized intersection density were consistently identified as important factors affecting crashes.
  - Pasco County corridors and corridors with greater than 7.5 signalized intersections/mile would likely benefit the most from proactive law enforcement.

Future Research

- Non-crash citation rate was important in model 1, but the relationship between crashes and non-crash citations needs further study.
- Future research can incorporate data on existing law enforcement patrols to recommend specific adjustments or explore other analysis methods, such as neural networks.

Questions?

- John McCombs: John.McCombs@ucf.edu
- Haitham Al-Deek, Ph.D., P.E. (Corresponding Author): Haitham.Al-Deek@ucf.edu