Applying Machine Learning to Model Freight Vehicle Type Choice

Usman Ahmed

Matthew Roorda

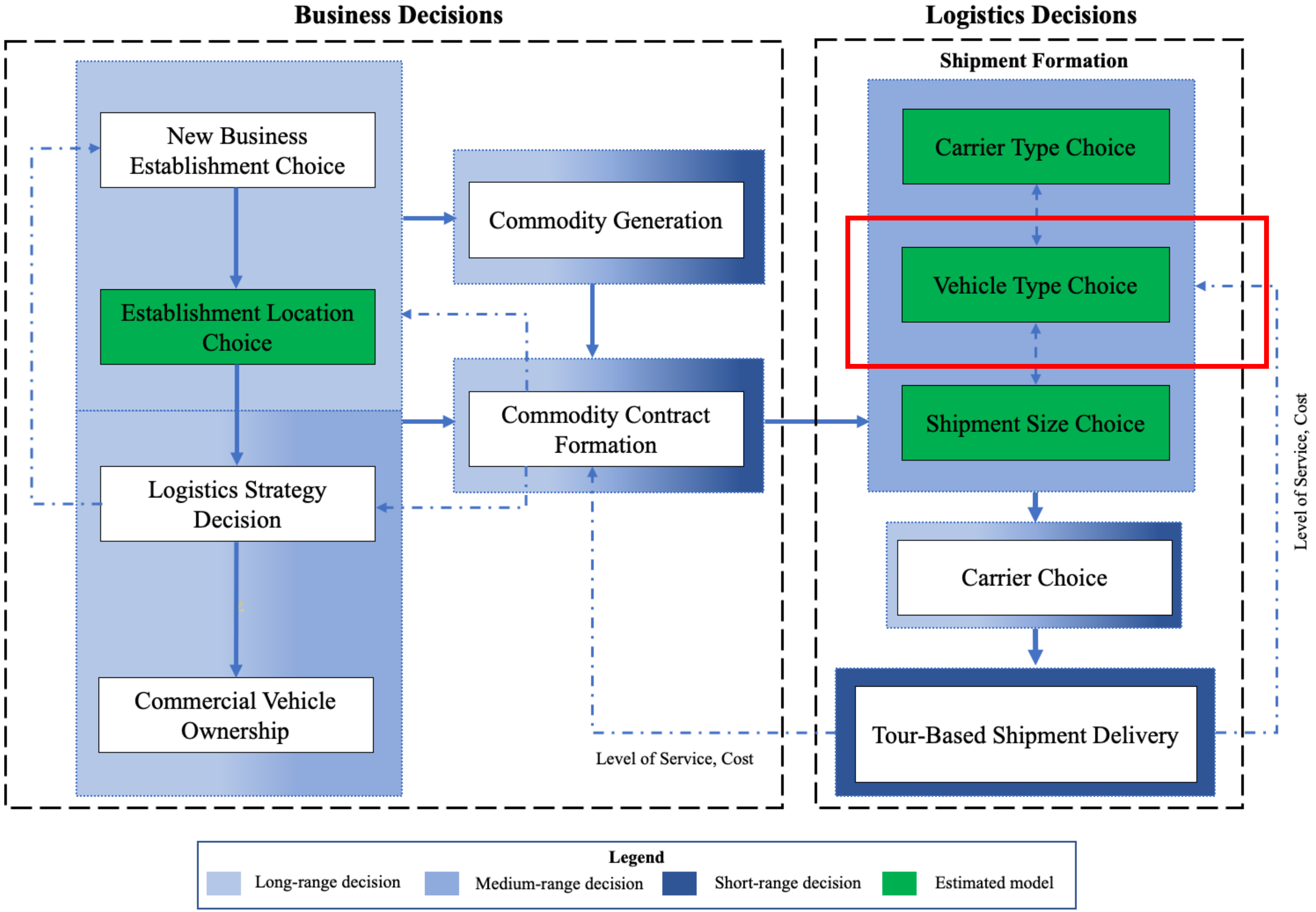
9th METRANS International Urban Freight Conference, Long Beach
May 25, 2022

City Logistics for the Urban Economy

CLUE

SMART FREIGHT CENTRE

Freight Business and Logistics Decisions Simulation Framework



Freight Mode v/s Vehicle Type

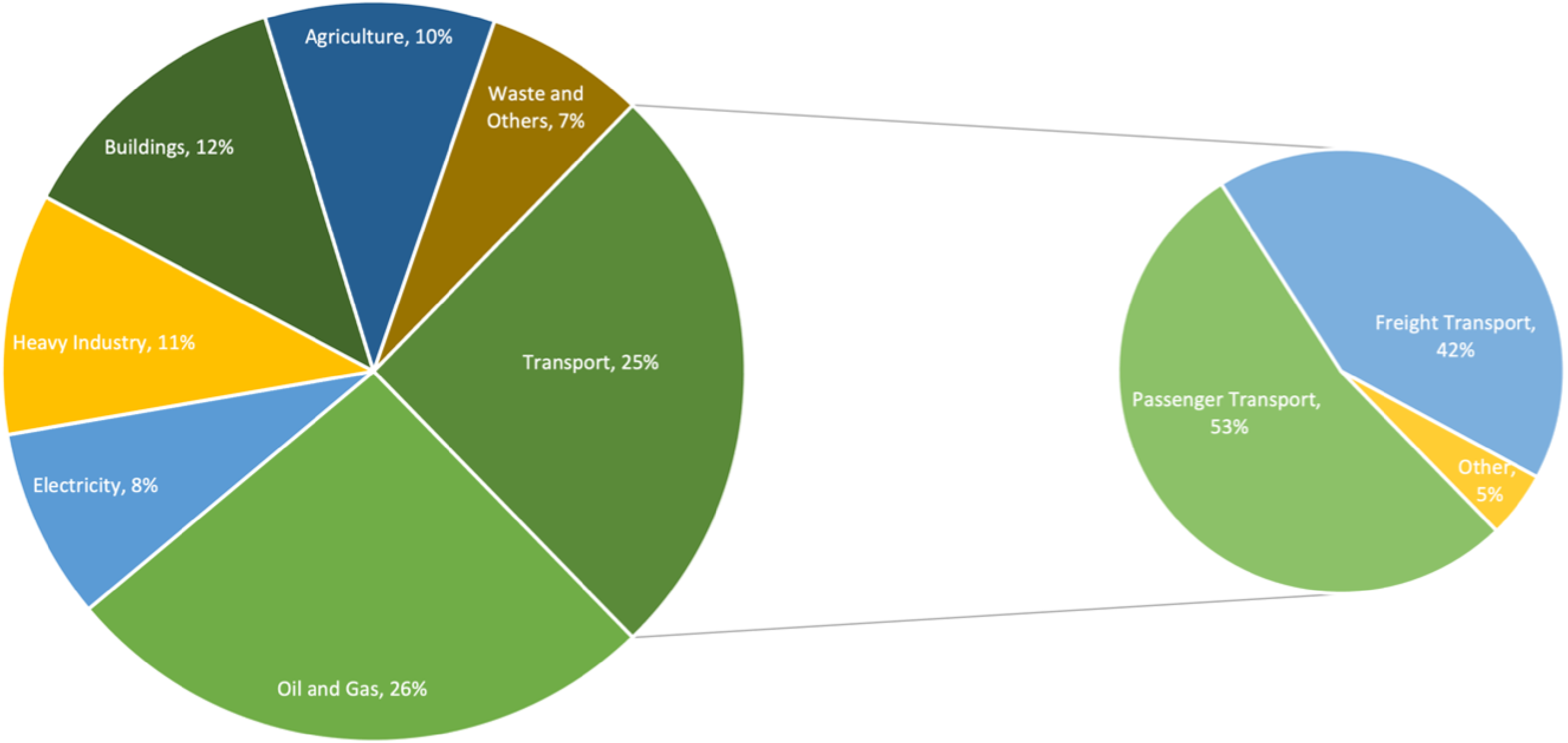
- Mode Choice:
  - Road, rail, air, water
  - Most relevant for inter-city, statewide, and national level studies

- Vehicle Type Choice:
  - Road-based mode: Passenger car, trucks, vans, etc.
  - Most relevant for city or metropolitan area level studies

Background and Motivation

- Freight flows have been increasing in Canada.
  - 16.7% increase in freight shipments from 2011 to 2017 (Statistics Canada 2020)
- Economic development of regions
- Global competitiveness of industries
- Changing trends in supply chain and logistics
- Major contribution to greenhouse gas emissions!

Background and Motivation



Background and Motivation

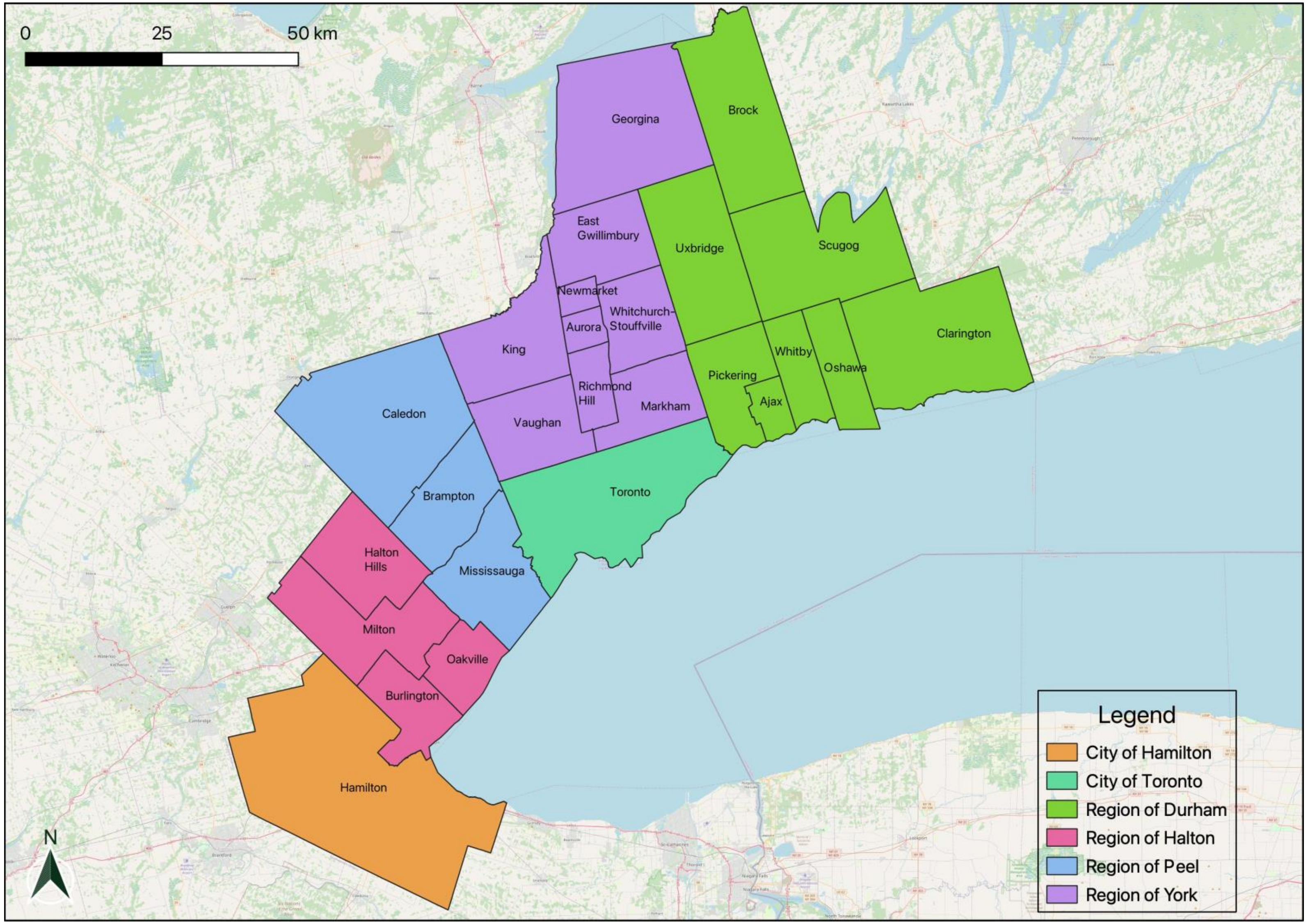
- Implications on quality of life of urban residents
  - Noise pollution
  - Traffic congestion
  - Safety impacts
  - Parking problems
  - Pavement damage

Study Objectives

- Study the factors behind freight vehicle type choice

- Comparison of discrete choice with machine learning methods
  - Discrete choice: Multinomial and mixed logit model
  - Machine learning: Random Forest

Study Area



0
25
50 km
Georgina
Brock
East Gwillimbury
Uxbridge
Scugog
Newmarket
Whitchurch-Stouffville
Aurora
King
Clarington
Whitby
Oshawa
Pickering
Richmond Hill
Ajax
Markham
Caledon
Vaughan
Brampton
Toronto
Halton Hills
Mississauga
Milton
Oakville
Burlington
Hamilton
N
Legend
City of Hamilton
City of Toronto
Region of Durham
Region of Halton
Region of Peel
Region of York

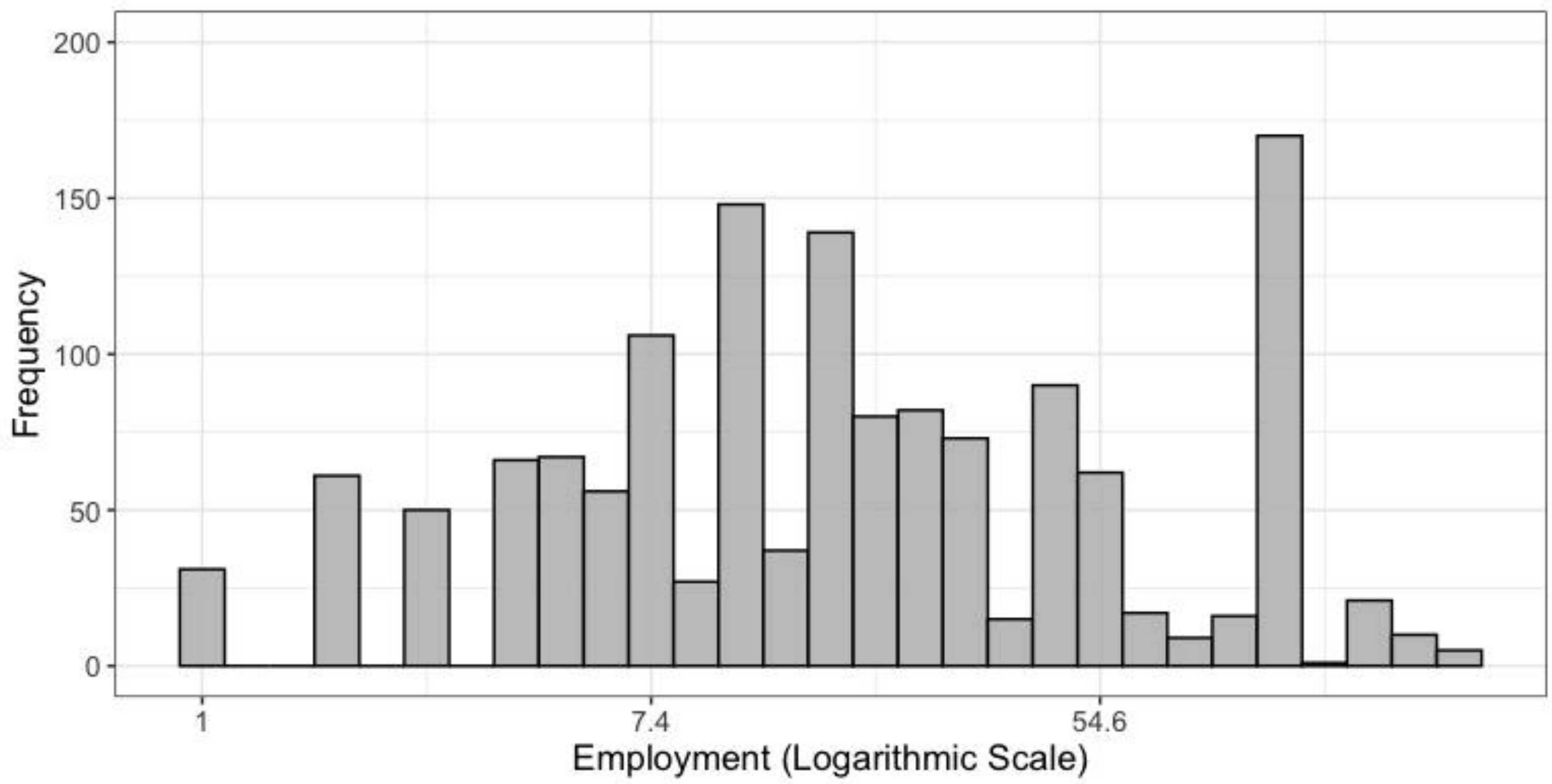
Data Source

- Commercial Travel Survey
  - Region of Peel (2006/07), Region of Durham (2010), Toronto Area (2012)

- Outbound Shipments
  - 1,439 shipments
  - 385 firms

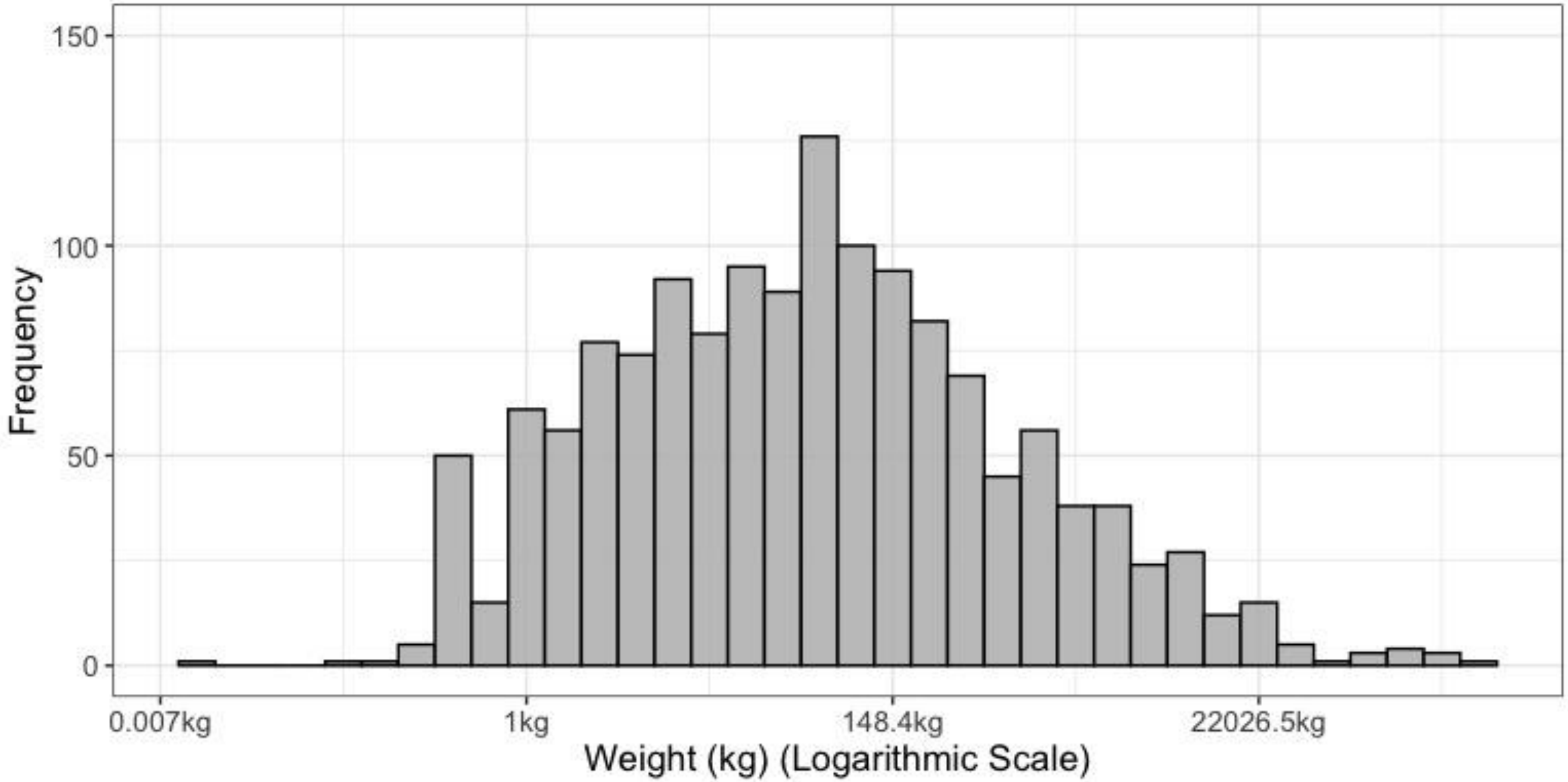
- Explanatory Variables
  - Industry type, commodity type
  - Shipment origin and destination (cities)
  - Employment and shipment weight

Data Source



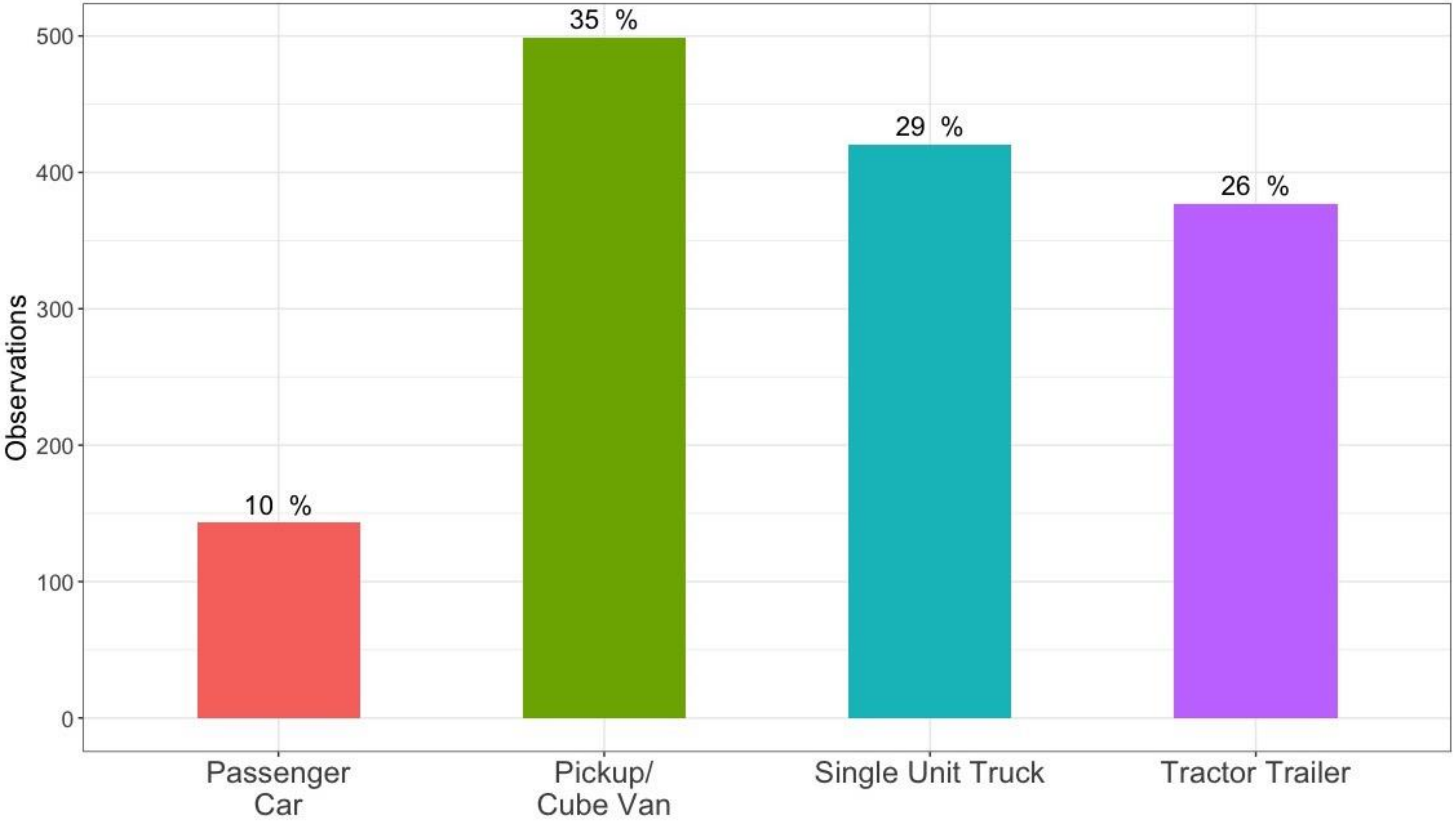
200
150
100
50
0
Frequency
1
7.4
54.6
Employment (Logarithmic Scale)

Data Source

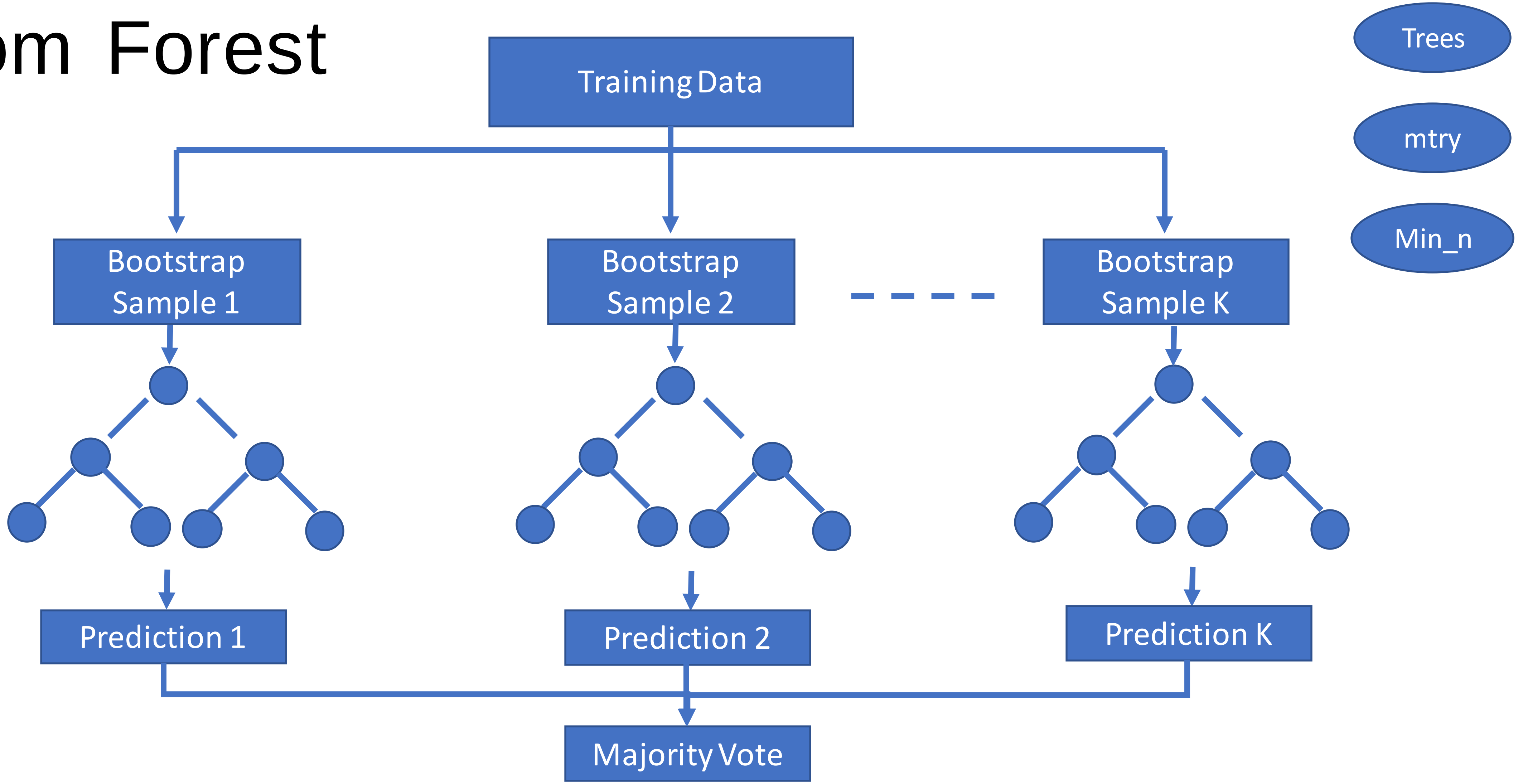


150
100
50
0
Frequency
0.007kg
1kg
148.4kg
22026.5kg
Weight (kg) (Logarithmic Scale)

Data Source – Vehicle Types



Random Forest



Random Forest – Variable Importance

- Shapley Additive Explanation (SHAP)
  - To assess the impact of explanatory variables on the model output
  - Comparison of model prediction with and without the variable
  - SHAP value is calculated for every observation
  - Variables are sorted based on the impact
  - Color of the point shows its value
    - Red: high value
    - Blue: low value

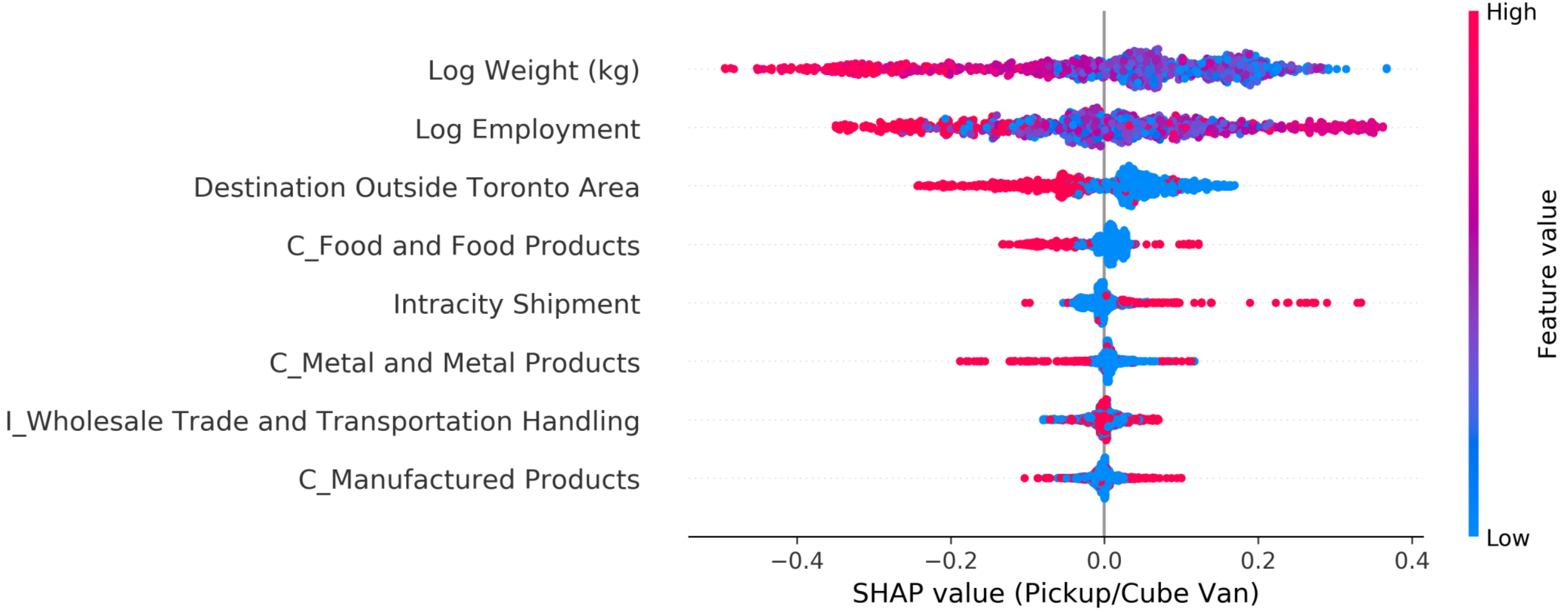
Training v/s Testing Data

- Models are developed on training data
  - RF: 10-fold cross validation
- Model prediction accuracy is calculated on testing data

- Training and testing data are divided based on firms
  - Training data: 269 firms with 1114 shipments (70%)
  - Testing data: 116 firms with 325 shipments (30%)

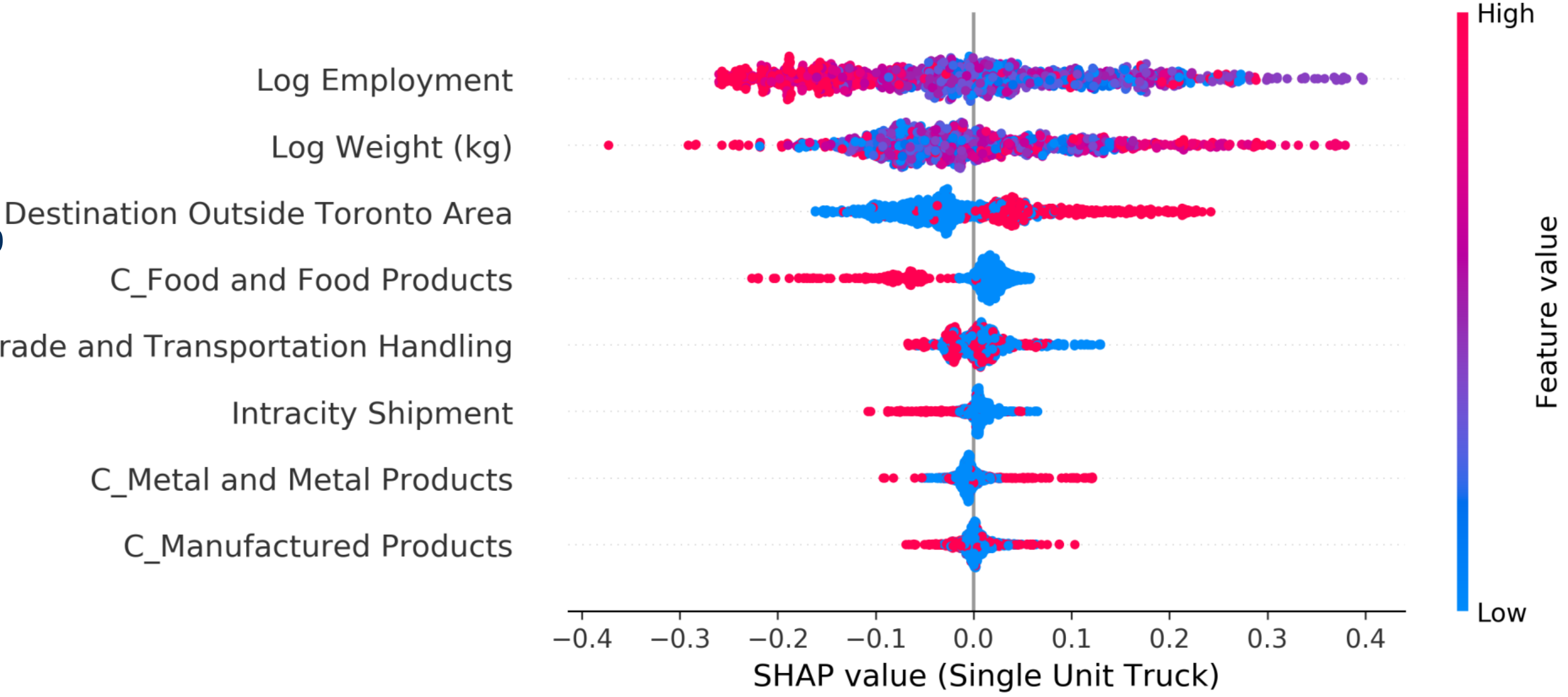
Results

Variable Importance



Variable Importance

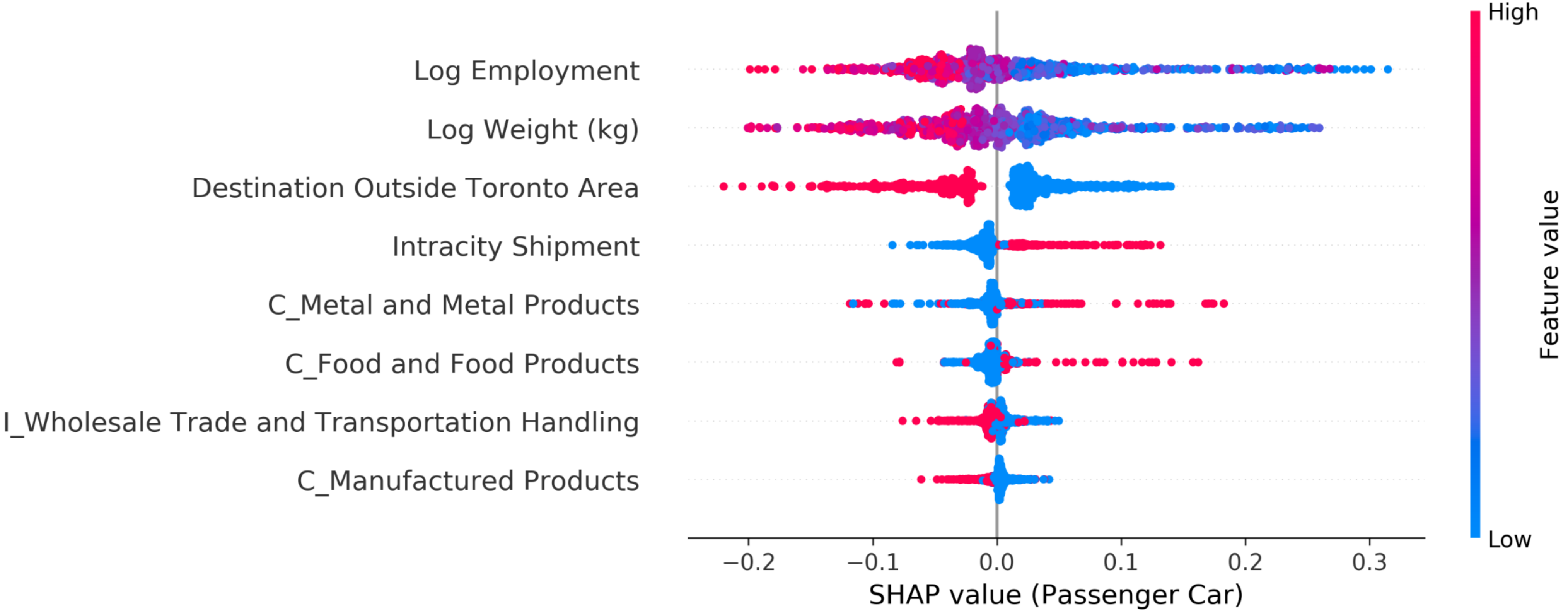
MNL: 1.64
Mix-MNL: 1.79



Variable Importance

MNL: 0.88

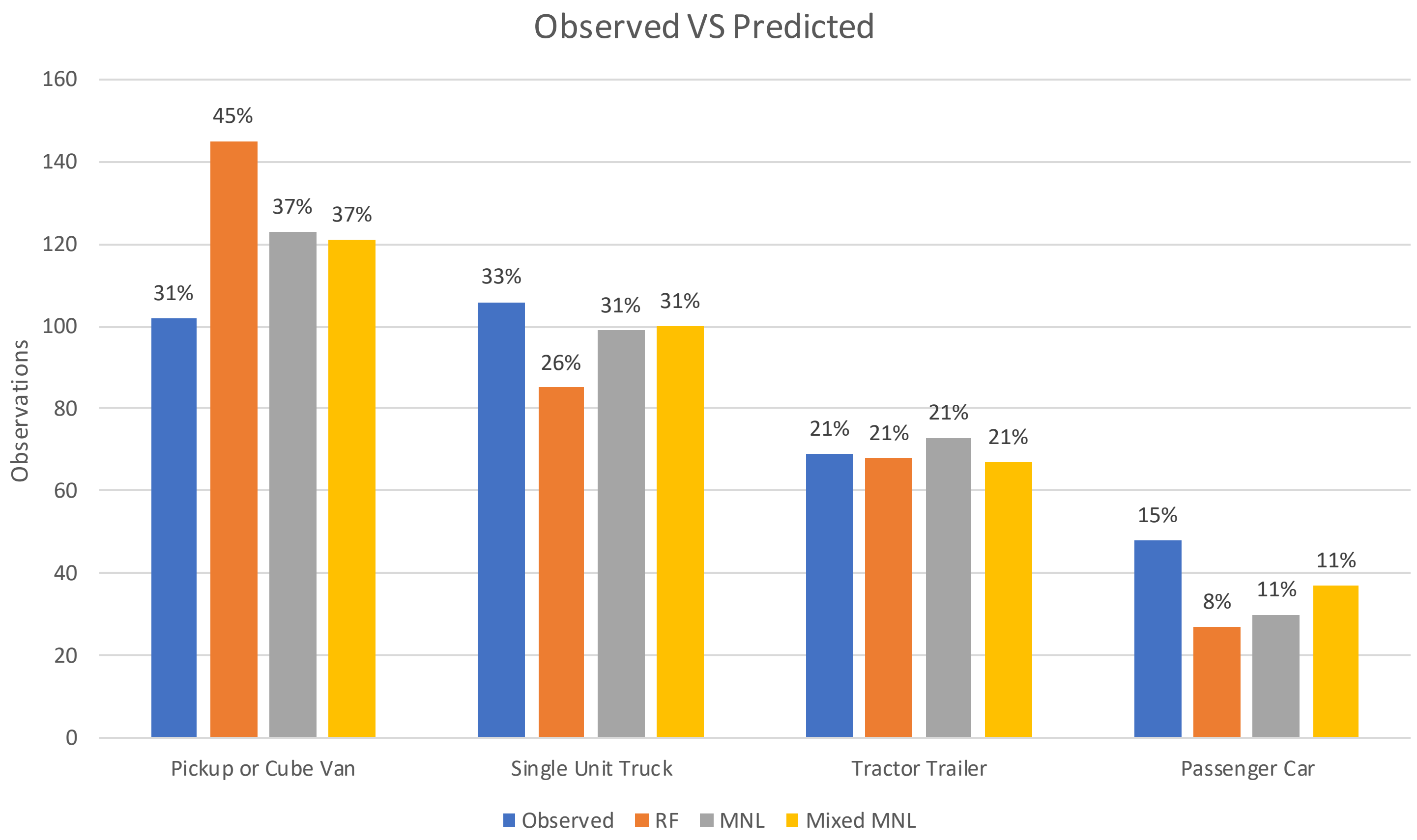
Variable Importance



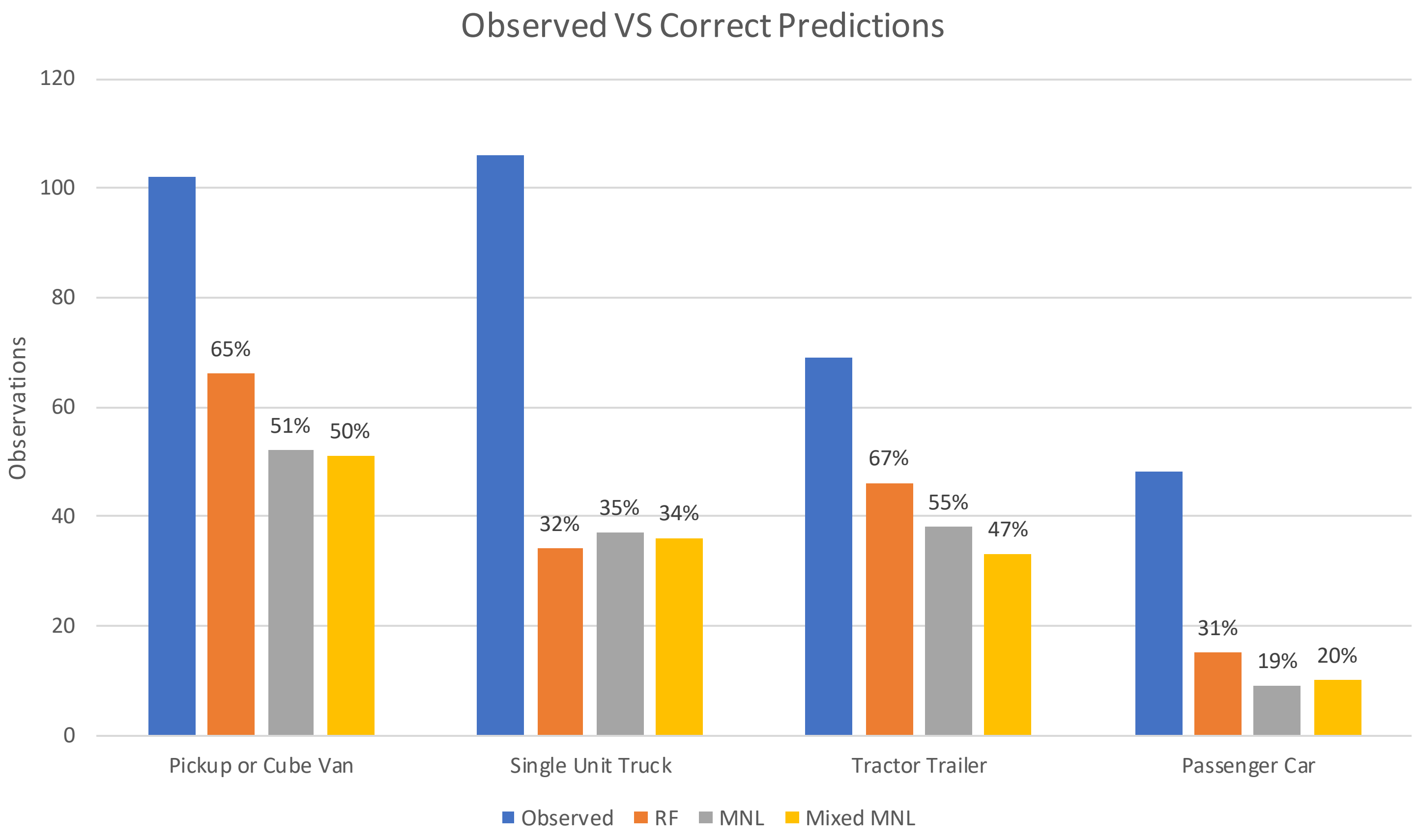
Discrete Choice Methods

- Larger firms are more likely to use larger vehicles
- Larger vehicles are more likely to be used for heavier shipments
- Intracity shipments are more likely to be transported using smaller vehicles
- Larger vehicles are more likely to be used for shipments destined outside of Toronto Area

Model Predictions



Prediction Accuracy



Results Summary

- Overall prediction accuracy
  - Random Forest: 50%
  - MNL: 42%
  - Mix-MNL: 40%

Conclusion

- Applications in policy analysis
  - Demand for parking facilities
  - Greenhouse gas emissions
  - E-commerce, same-day deliveries

Conclusion

- Freight vehicle type choice is studied using discrete choice and RF methods
- Commercial travel survey data are used to develop models for the Toronto Area
- RF results are interpreted using SHAP based variable importance
- RF model has higher prediction accuracy than DCM

More about this work!

- More details about models and results can be found in:

Ahmed, U., & Roorda, M. J. (2022). Modeling freight vehicle type choice using machine learning and discrete choice methods. Transportation Research Record, 2676(2), 541-552.