

(a). Data Process

Covariates Process

Socio-demographic variables

- Population
- Median income
- Sex ration
- ...



Output Process

Traffic crash rates

- Overall crash rate
- Fatal crash rate

Traffic & crash records

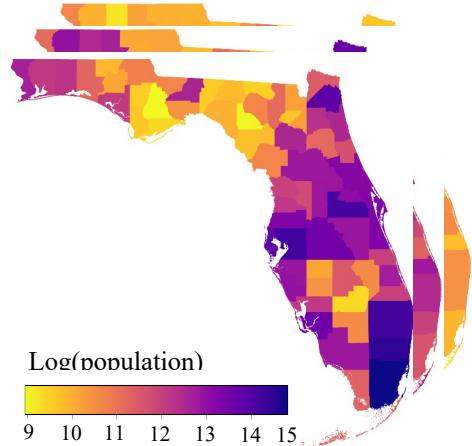
- Traffic crash records (2018-2022)
- Intersection densities
- AADT, VMT
- ...

Sources from FDOT

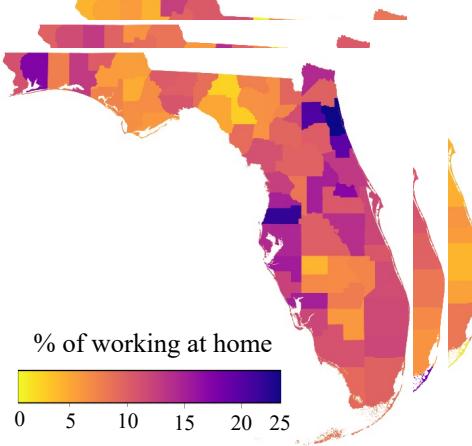
Sources from U.S. Census Bureau

(b). Spatial Inequality Modelling

Socio-demographic variables



Commute pattern variables



Variable Tests

Spatial lag

Endogeneity

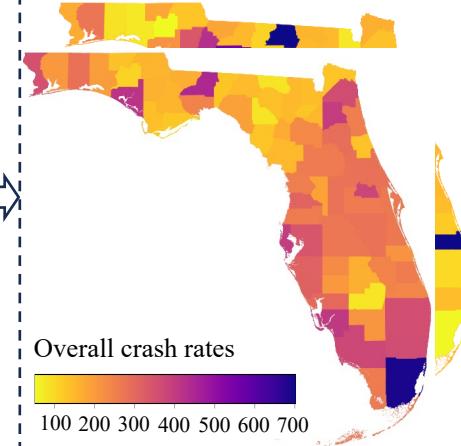
Modelling

SLM

MGWR

XGBoost

Overall & Fatal crash rates



(c). Simulated Comparison of Machine Learning with Spatial Statistical Models

Simulated Data Design

Spatial lag effects

Non-linear effects

Spatial heterogeneity

Interaction effects

Simulated Results Analysis

Quantitative Evidence

(d). Empirical Results and Discussion in Florida

Exploratory analysis of traffic crash rates

Results from statistical models (SLM, MGWR)

Results from XGBoost with SHAP

Discussion and Conclusions