Measuring and examining the spatiotemporally varying transmissibility of COVID-19

Hao Yang and Angela Yao

University of Georgia

Outline

- **Background**
- **Research Questions**
- **Methods**
- **Results**
- **Conclusions**

Background

--Measuring transmissibility

R_0 : Basic Reproduction Number

R^e: Effective Reproduction Number

--Complex relationships between contextual factors and the transmissibility

- There are multiple contributing factors affecting the transmissibility
- The relationships vary over time and space.
- Machine learning methods provide insights into the complex relationships

Research Questions

- **Question-1**
  - From a longitudinal perspective, for a specific place (e.g., county), how do policy and behavior responses to the pandemic affect the changing R^e over time?
- **Question-2**
  - For a cross-sectional view, at a specific time, how is the spatial variability of R^e associated with the demographic, socioeconomic, environmental, and health behavioral characteristics of places?

Methods

1. Estimation of transmissibility R^e

- Major steps for the estimation of space- and time-specific effective reproduction numbers ($R^e_{t,m}$)

Reported Cases → $I_{t,m}$ → $\lambda_{t,m}$ → $R^e_{t,m}$

$I_{t,m}$: the time series of active cases at time t in county m

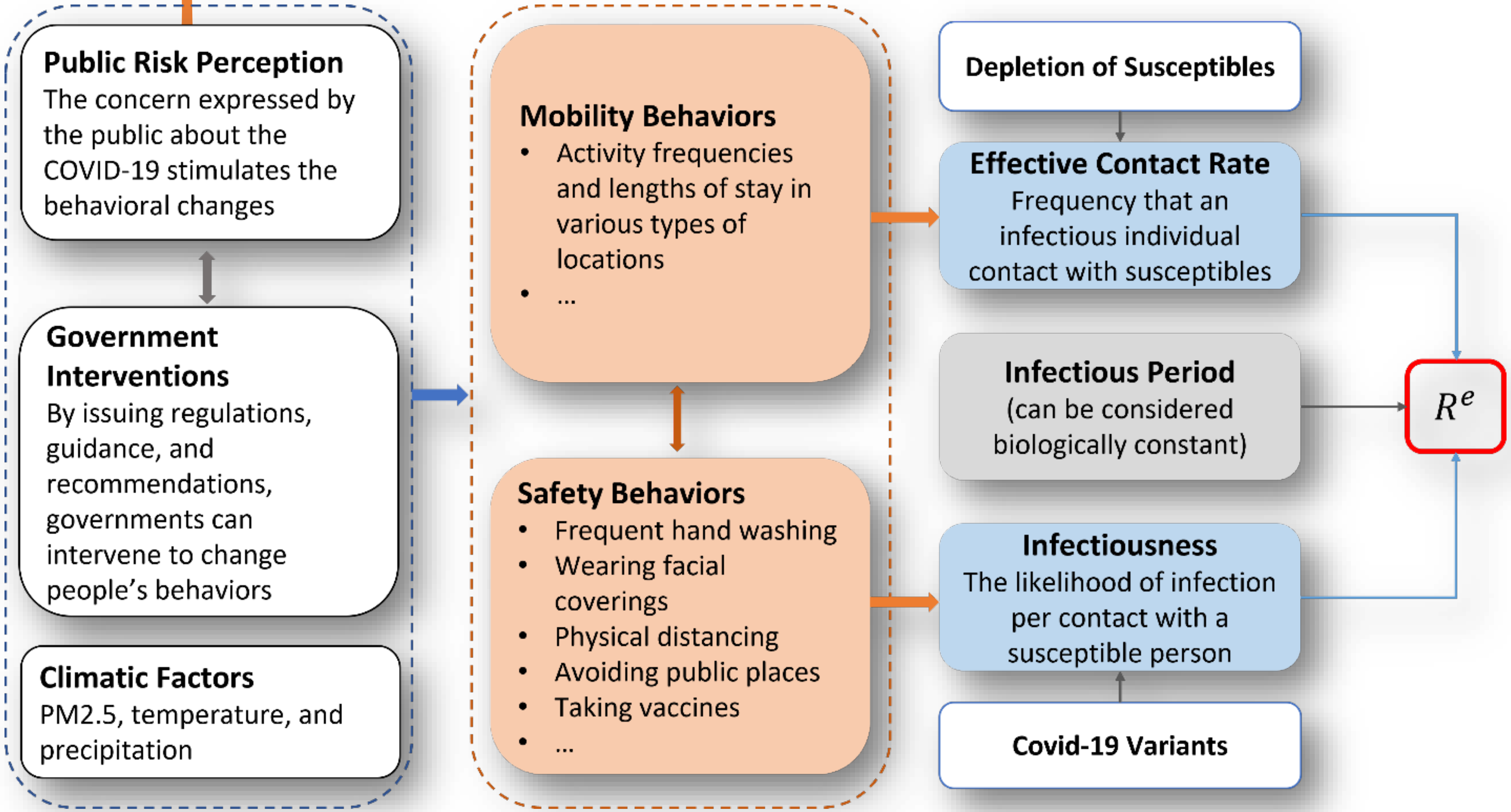
$\lambda_{t,m}$: the exponential growth rate of the time series of $I_{t,m}$

$R^e_{t,m}$: the estimation of the effective reproduction number

Methods

2. Relationship between contributing context factors and R^e

The flow of influences



Methods

3. Machine learning models

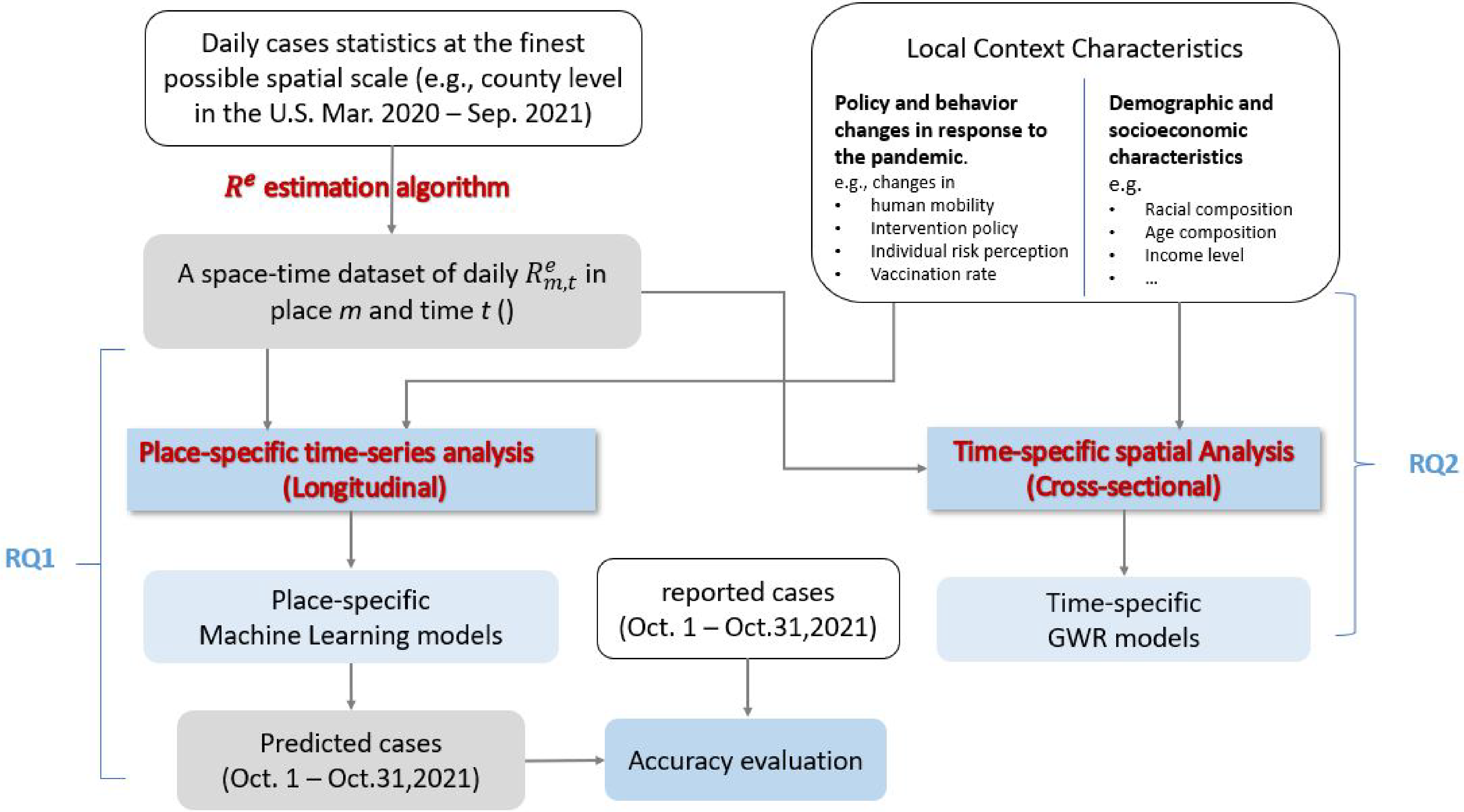
- Two regression models
  - Linear Regression
  - Exponential Regression

$$R_{m,t}^{e} = R_{m,0}^{e} e^{\sum_{i=1}^{n} \alpha_{m,i} * x_{m,i}}$$

- Two Decision Tree Ensemble methods
  - XGBoost algorithm
  - Random Forest (RF)

- Artificial Neural Nework (ANN)

Research Design



Results

1. **Estimation of R^e** at the county level in the U.S. Mar. 2020 – Sep. 2021

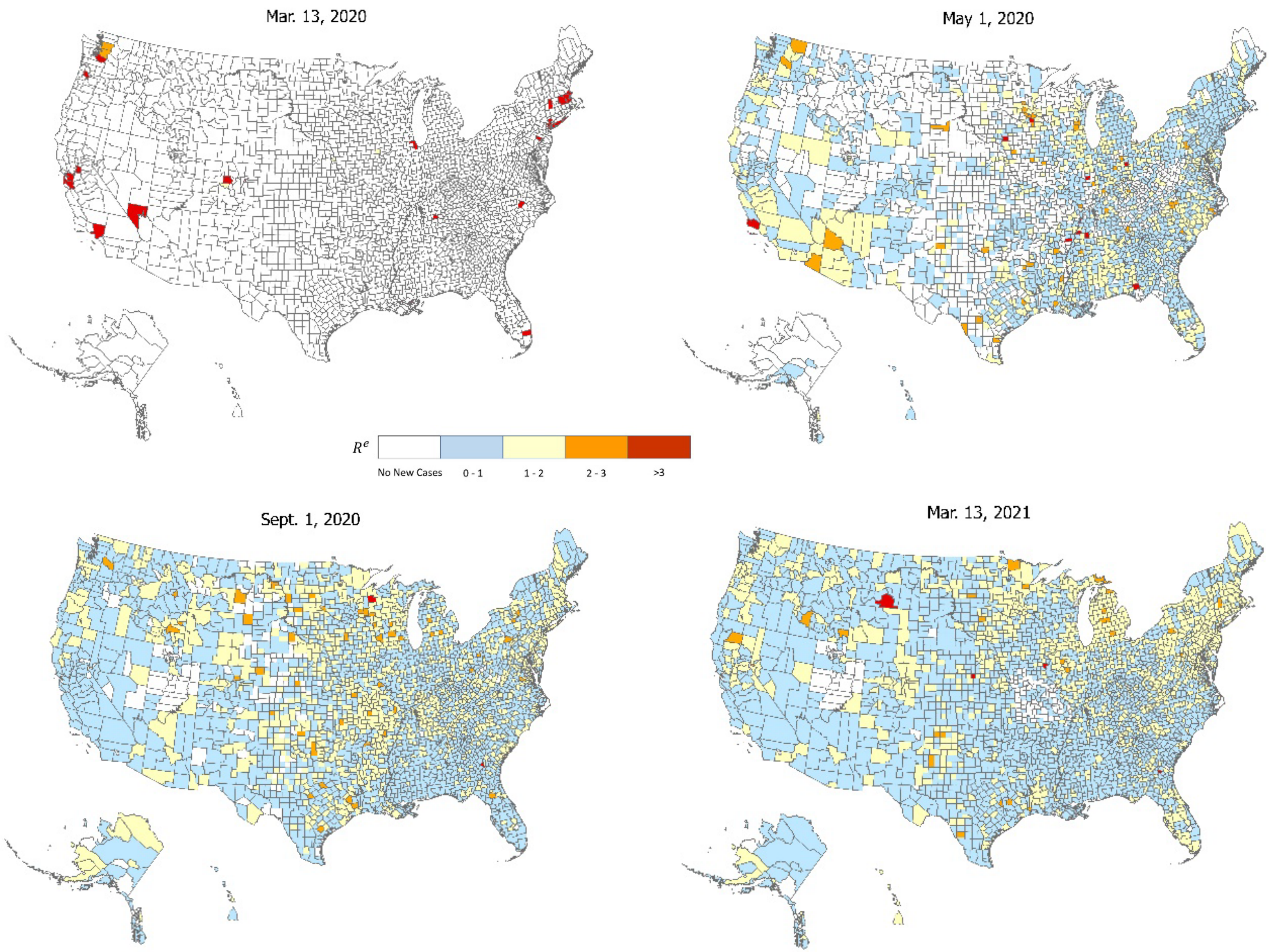
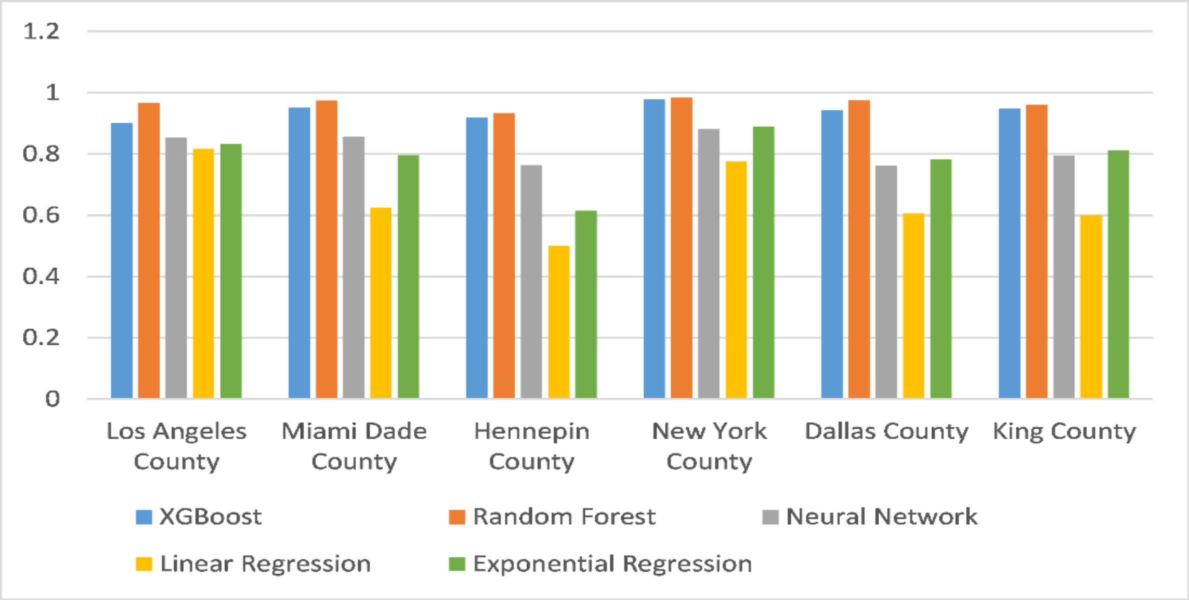


Fig. The variations of Covid-19 R^e in the United States across space and time

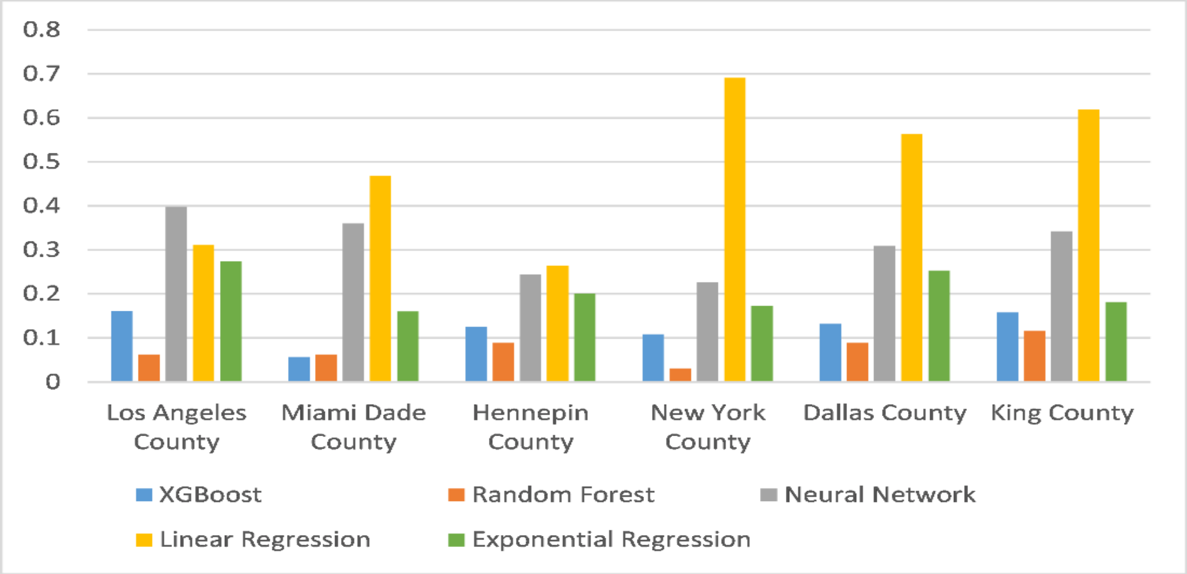
Results

2. **Machine models** – Training and validation with training data

➢ To validate the models and assess their performances, the study applies the 10-fold cross-validation for each model.



(a) The R^2 values of the five machine learning models for six selected counties

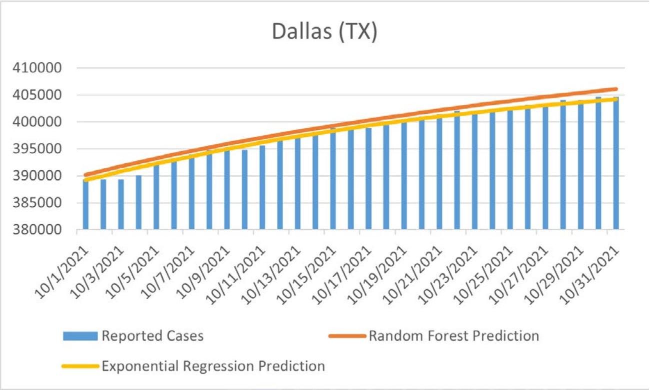


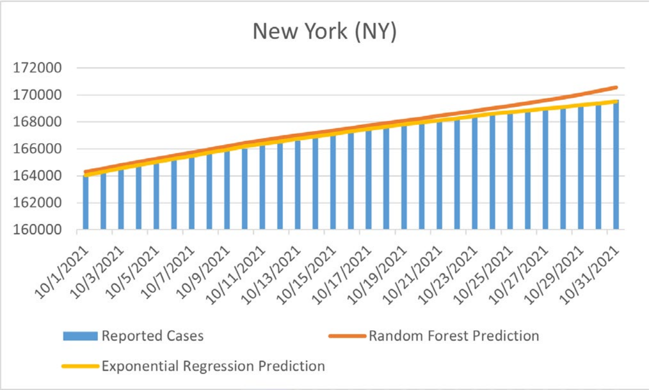
(b) The RMSE values of the five machine learning models for six selected counties

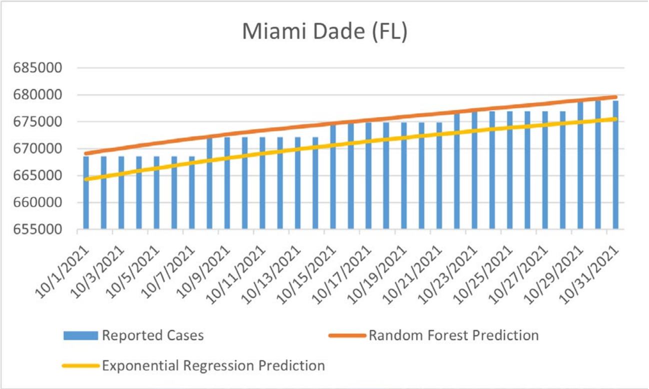
Results

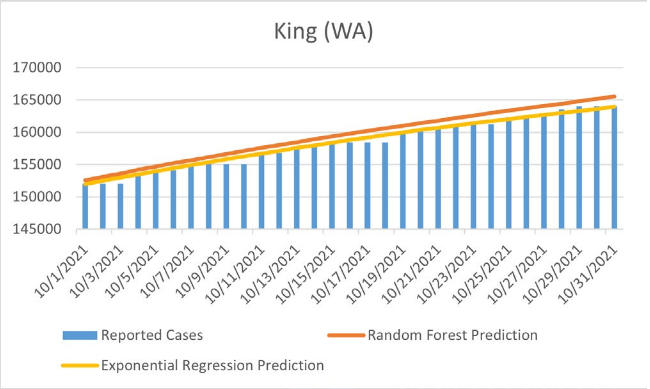
2. **Machine models** –
Evaluation with testing data

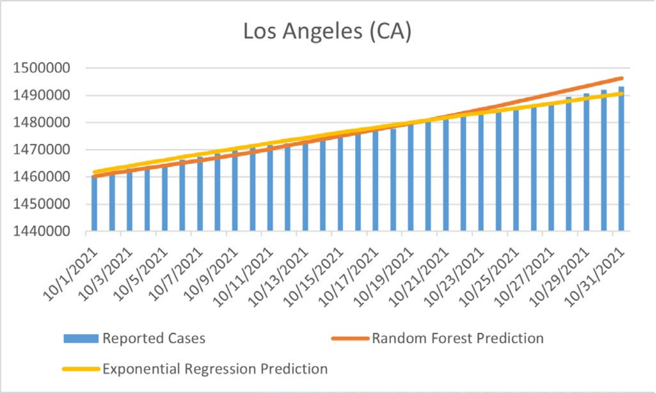
- The figure shows prediction results of the **RF** and the **exponential regression** model, compared against reported cases for the six selected counties in October 2021.

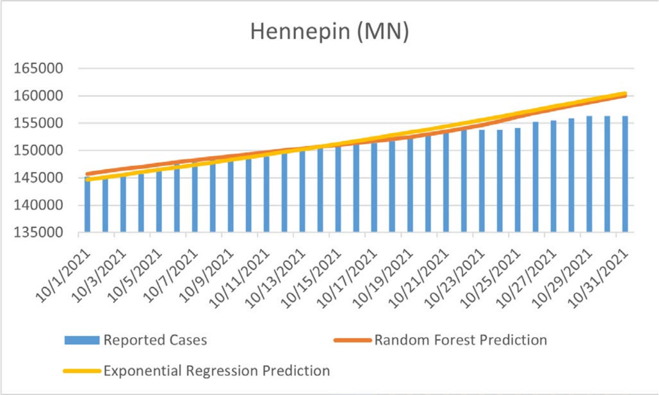










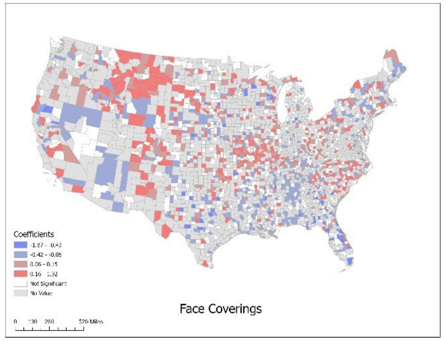


Results

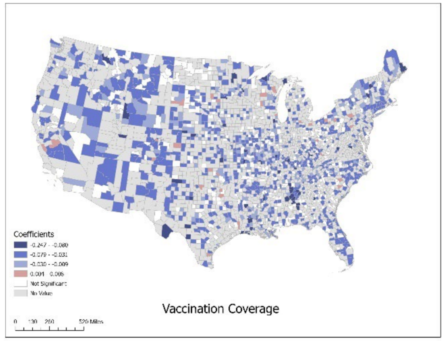
Spatial patterns of variable coefficients of the local exponential models

Exponential regression

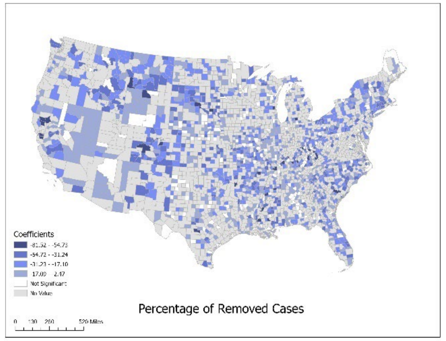
- Each local model has a set of coefficients corresponding to the respective explanatory variables.
- The variable coefficients vary across space.



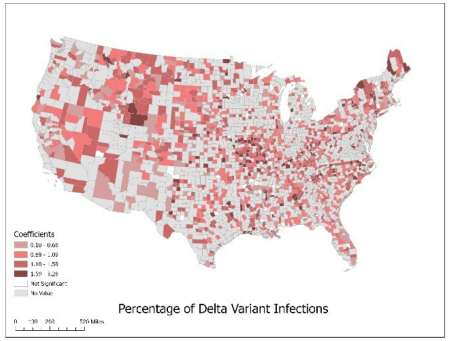
(a) Face Coverings



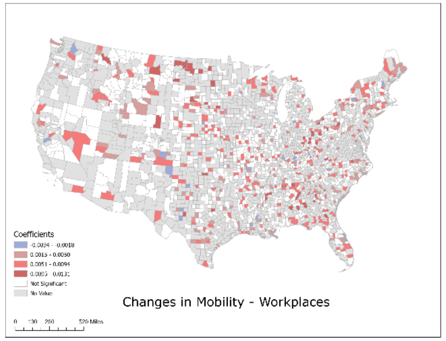
(b) Vaccination Coverage



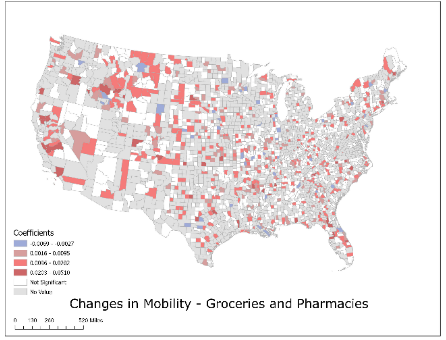
(c) Removed Cases



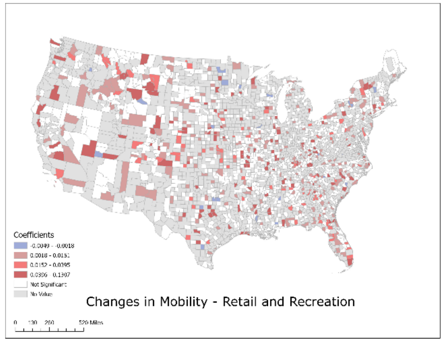
(d) Delta Variant



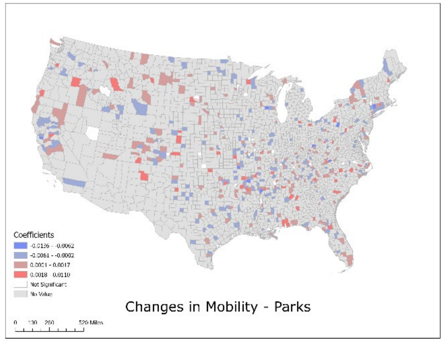
(e) Changes in mobility to Workplaces



(f) Changes in mobility to Groceries and Pharmacies



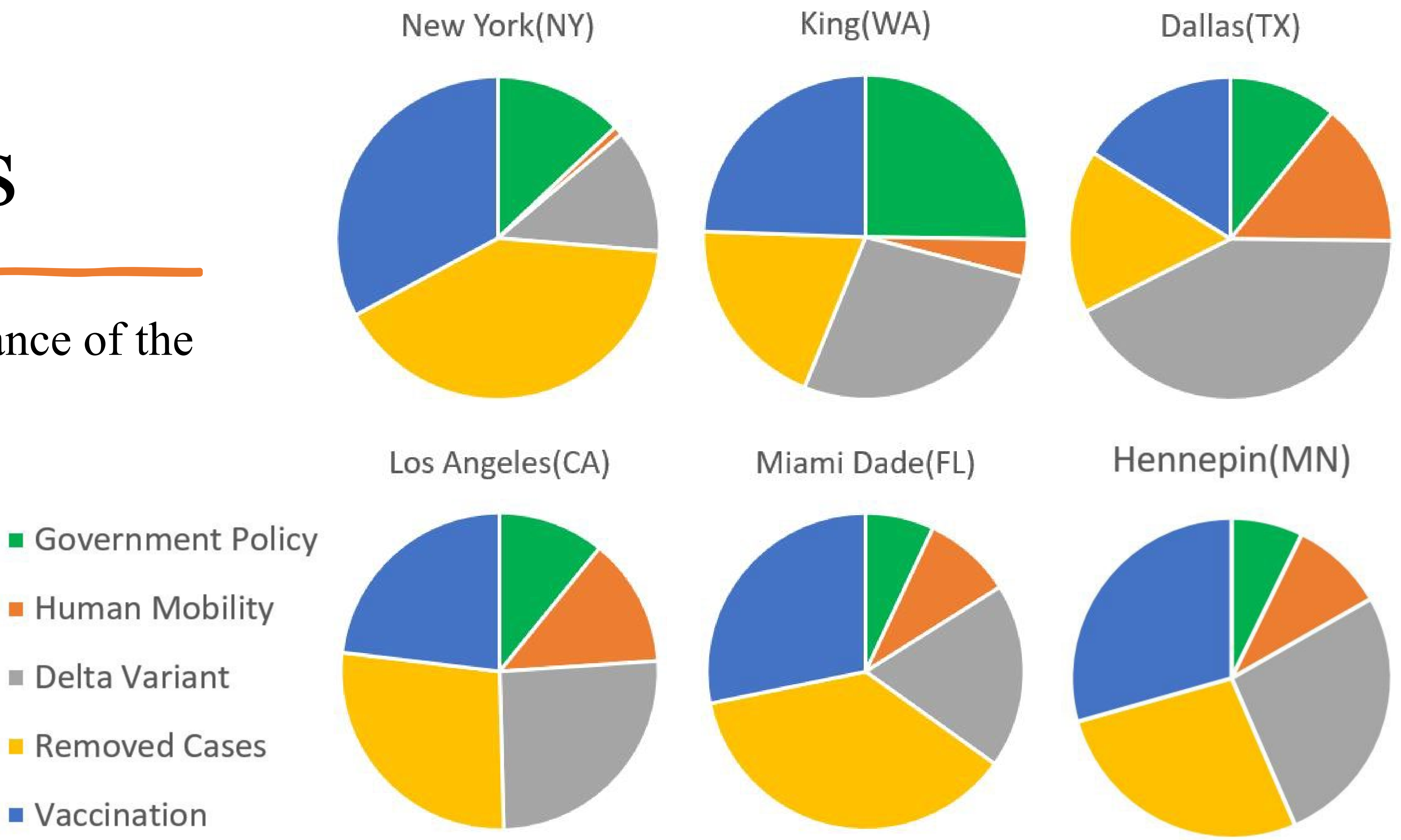
(g) Changes in mobility to Retail and Recreation



(h) Changes in mobility to Parks

Results

Feature importance of the RF model



Results

Random Forest model

- Random forest (RF) is the best-performing model. The figure shows the feature importance in each of the local RF models for the 1362 selected counties.

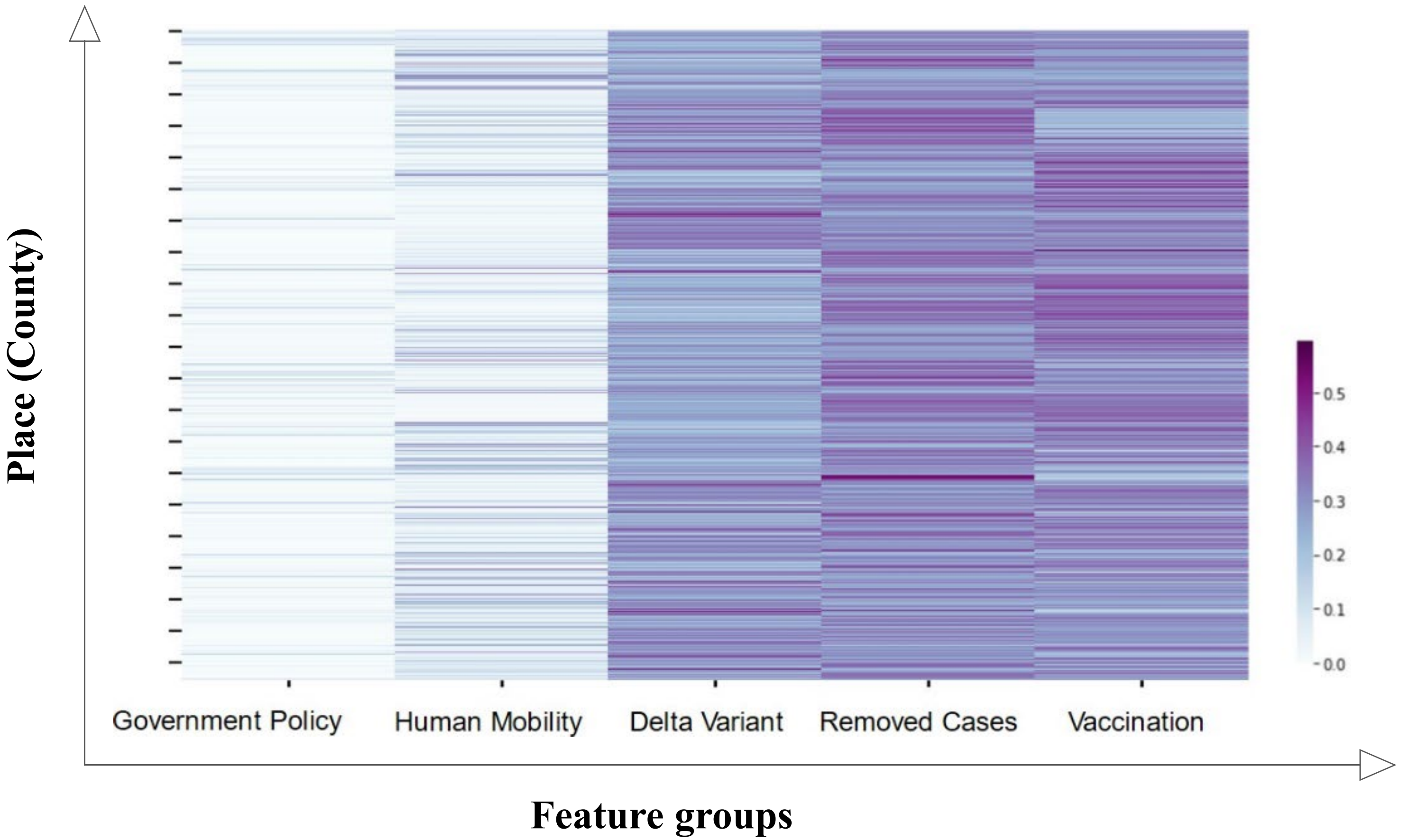


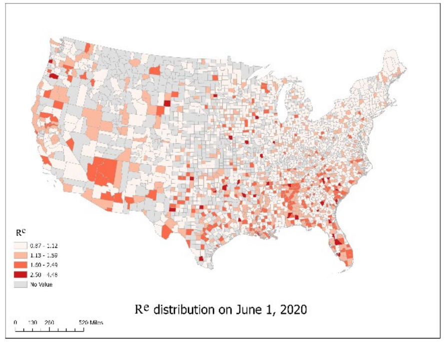
Fig. Feature importance of different predictive variables in the random forest models

Results

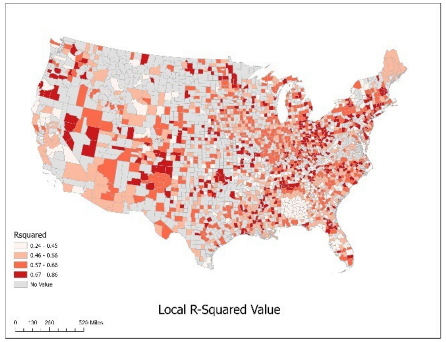
GWR model

- The GWR modeling is a cross-sectional study to investigate the relationship. It can be performed on data of any time slice during the study period.

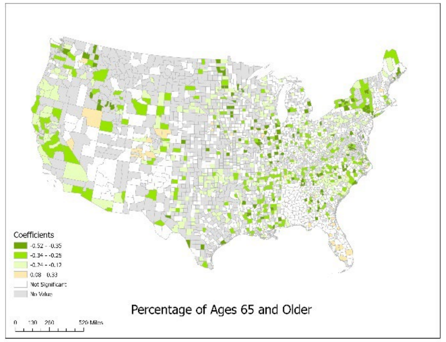
- Reported here are the results for June 1, 2020.



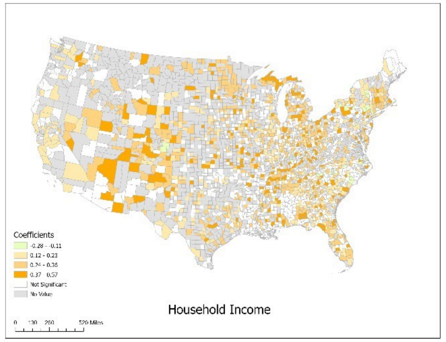
(a) R^e Distribution



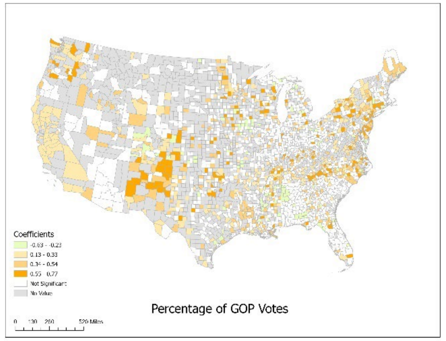
(b) Local R-Squared



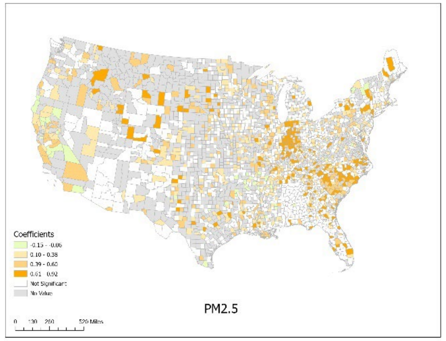
(c) Age 65 and over



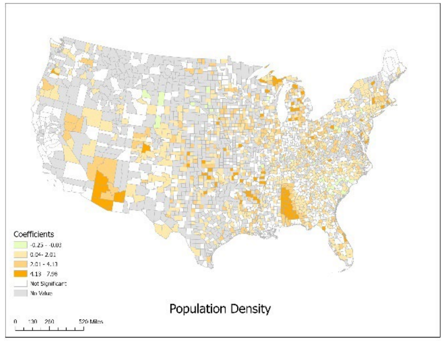
(d) Household Income



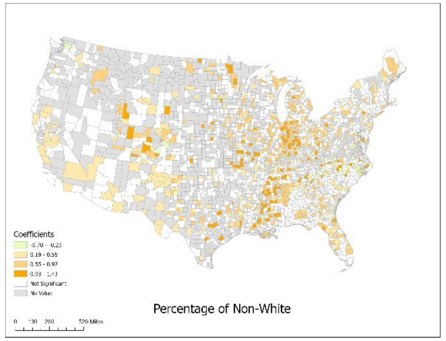
(e) GOP Votes



(f) PM 2.5



(g) Population Density



(h) Non-White

Spatial distribution of coefficients of the GWR model

Conclusions

Findings and contributions

- The temporal change of R^e in a place is positively associated with changes in human mobility and negatively associated with government intervention restrictions.

- Most importantly, R^e is significantly impacted by the depletion of the susceptible population.

- By modeling the relationships between R^e and local context factors, local policymakers and practitioners can better understand the effectiveness of various intervention policies.

Conclusions

Limitations and future work

- It is unknown whether different spatial scales may lead to other findings.
- The current research examines cross-sectional and temporal changes of R^e separately.

Thank You

Questions, comments, and suggestions are welcome