Simulating Urban Flow with Geographically Explicit Synthetic Population

Boyu Wang, Andrew Crooks

AAG Annual Meeting, Honolulu, HI

April 20, 2024

University at Buffalo The State University of New York

Introduction

- Urban human mobility is an active research field that studies human mobility patterns in urban areas.
- Through individual's movement, higher level phenomena such as traffic congestion and disease outbreaks emerge.
- Understanding how and why people move around a city plays an important role in urban planning, traffic control, and public health.
- An agent-based model (ABM) is developed to simulate people's daily commute patterns.

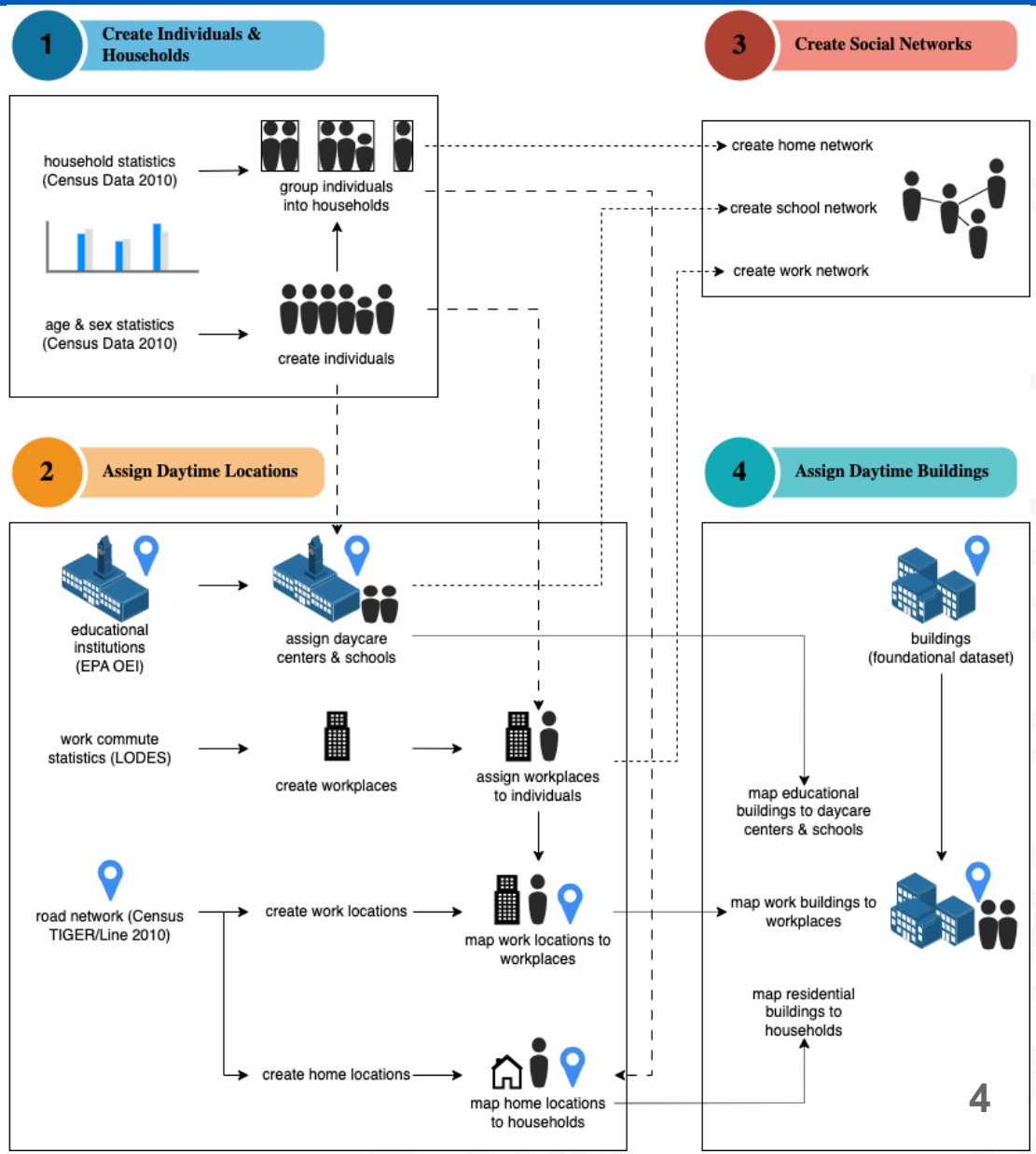
How can we utilize census data in ABM to uncover emerging urban commute patterns?

Approach

- Study area: Manhattan, New York City.
- Agents are initialized with a synthetic population created out of census data.
- Agents' home and work locations are derived from census data.

Approach: Synthetic Population

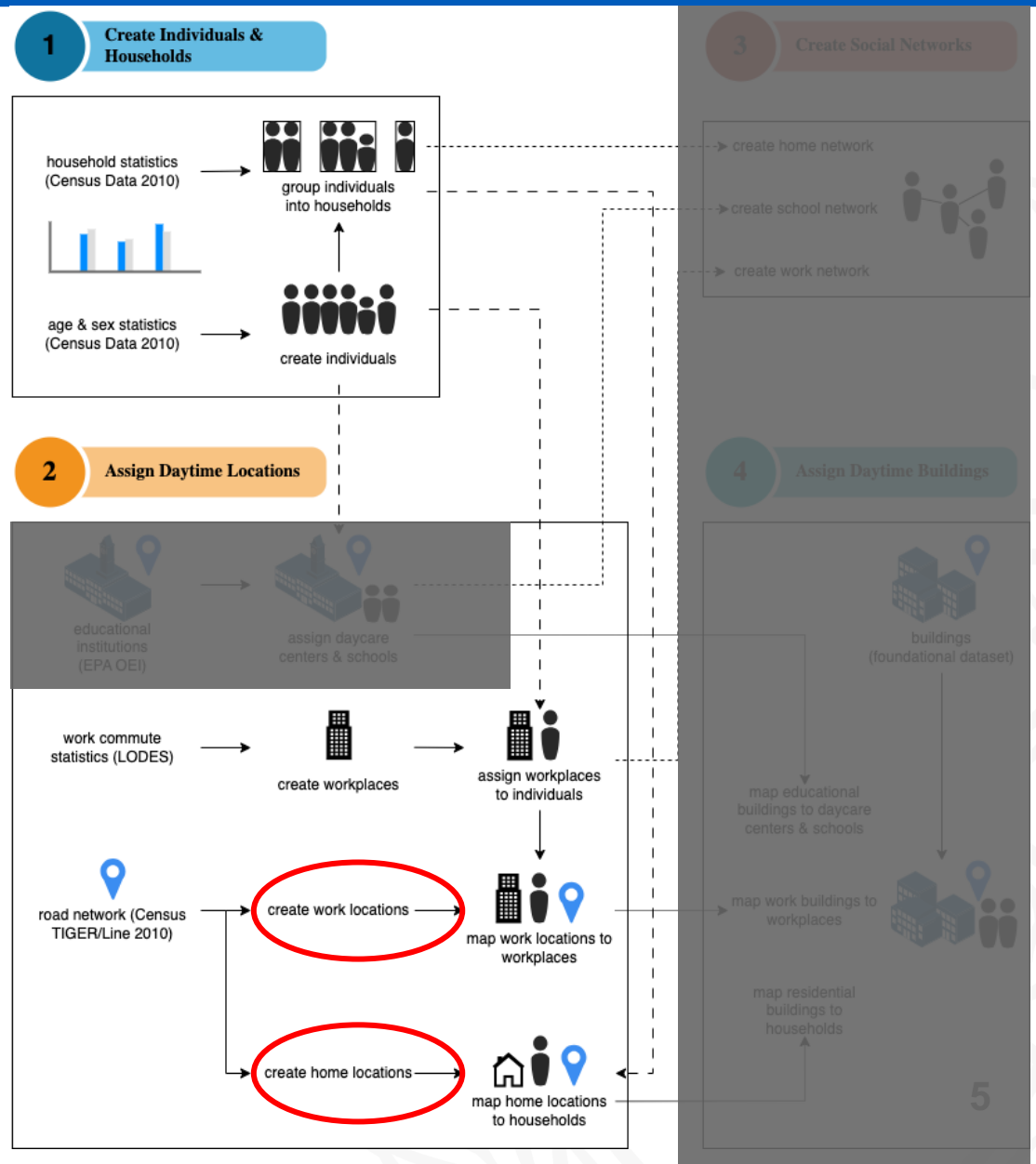
- Jiang, N., Crooks, A.T., Kavak, H. *et al*. A method to create a synthetic population with social networks for geographically-explicit agent-based models. *Comput.Urban Sci*. 2, 7 (2022).
- Source code available at: https://github.com/njiang8/Create_Synthetic_Population



Approach: Synthetic Population

The synthetic home and work locations are sampled from road network based on road types:

- **home locations** are sampled at 50m intervals on local neighborhood road, rural road, city street, and private road for service vehicles
- **work locations** are sampled at 20m intervals on secondary roads



Approach: Agent-Based Model

Model is developed using **Mesa** & **Mesa-Geo** in Python, both of which are open-source free software, with source code publicly available at https://github.com/projectmesa

- The only Python ABM framework with GIS support
- Easy integration with the Python ecosystem
- Expanding community

Come and join us!

- NumFOCUS Small Development Grant
- Google Summer of Code

SCAN ME

Approach: Agent-Based Model

GeoSpace

- Contains a road network layer, constructed from TIGER/Line shapefiles of roads in the study area.
- Implemented by two underlying data structures:
  - a topological network, for shortest path calculation
  - a k-d tree, for nearest node queries

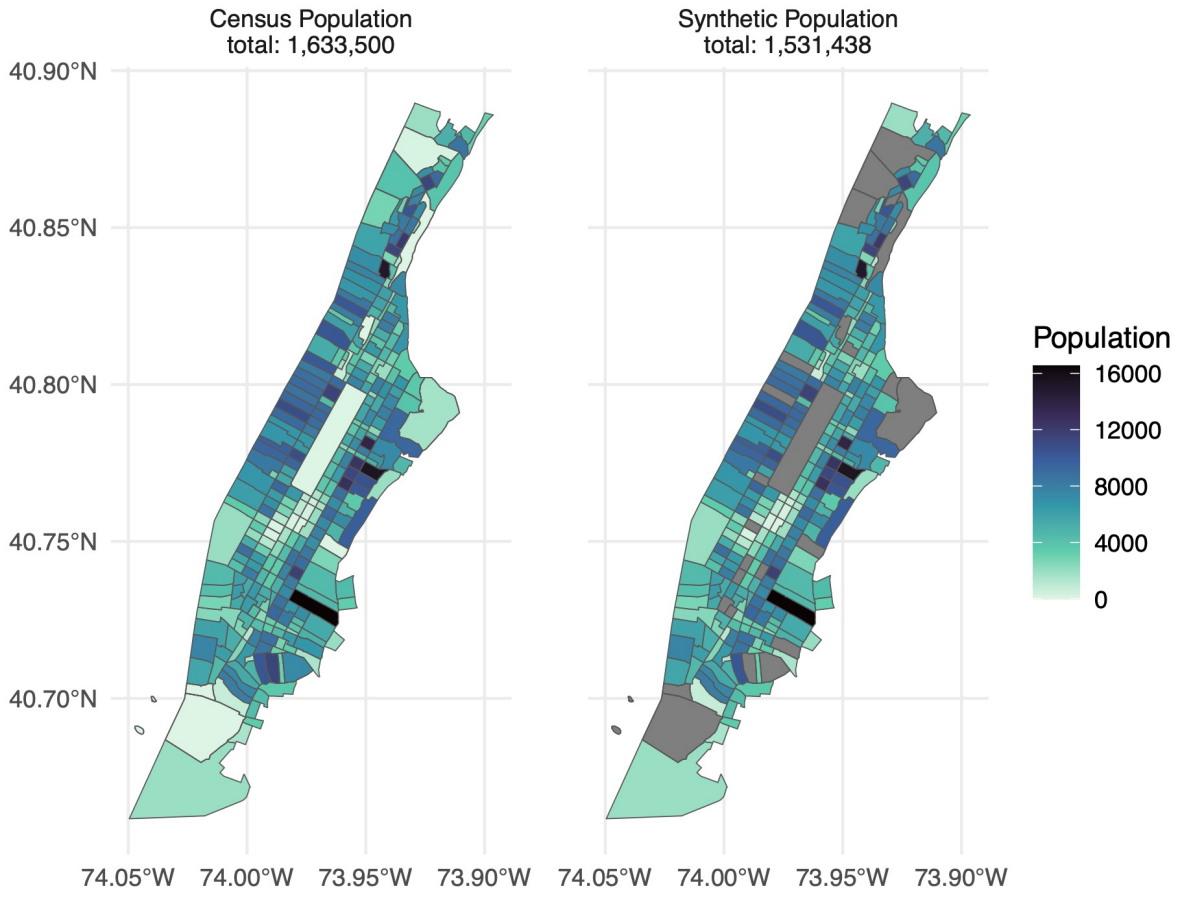
GeoAgents

- The commuters are the GeoAgents in the model, initialized with the synthetic population created earlier.
- Each agent has a unique id, home and work locations, randomly sampled on road networks.
- Every agent has a random start time to go to work (between 6am to 9am). After 8 hours, agents commute back home from workplaces.
- All agents travel at the same speed of 12km/h (subject to further calibration and validation).

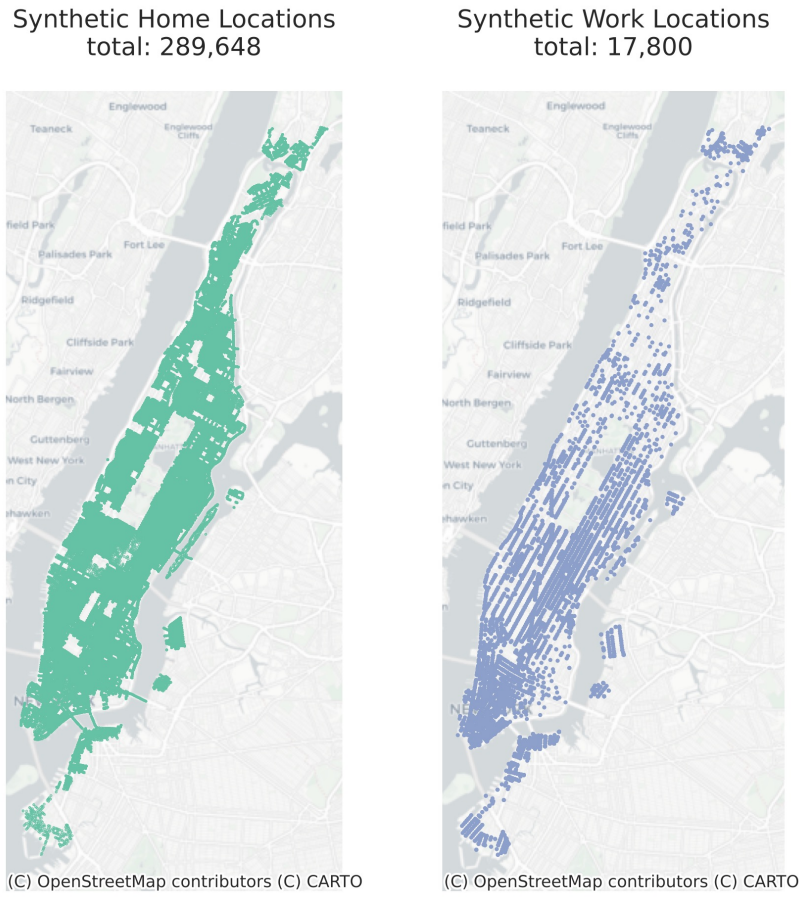
Results: Synthetic Population

- The total population from census data is 1,633,500. The synthetic population contains a total number of 1,531,438 individuals, which is 6.25% less than true population size.
- The amount of people who both live and work inside the study area is 443,893. This is the number of agents that are included in the commuting model (ideally).

Results: Synthetic Population

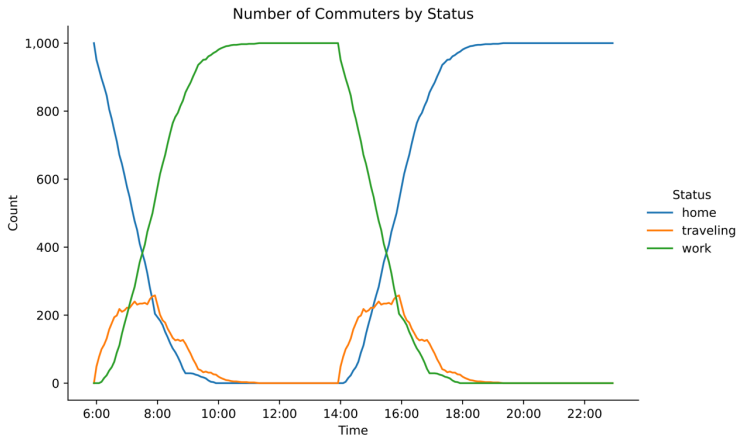


Synthetic population by census tracts in the study area.

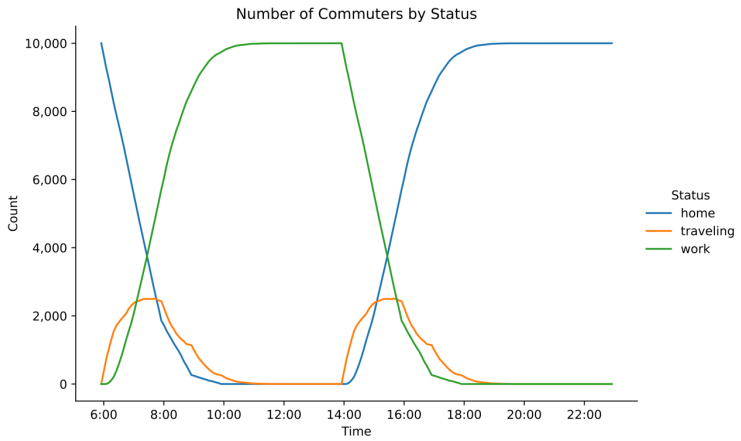


Synthetic home and work locations, sampled on road network based on road types.

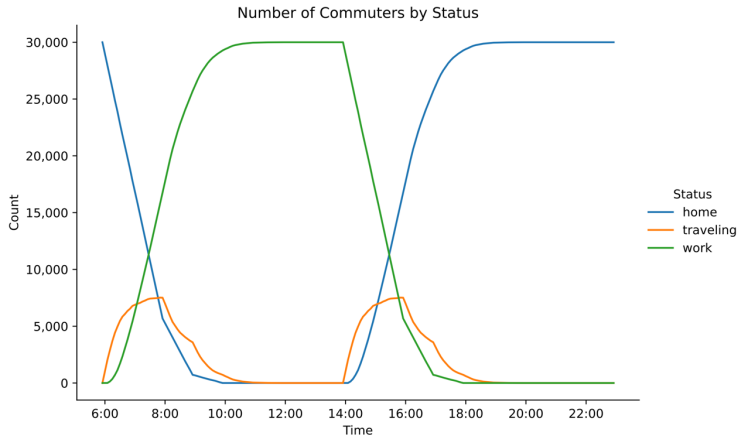
Results: Agent-Based Model



(a)

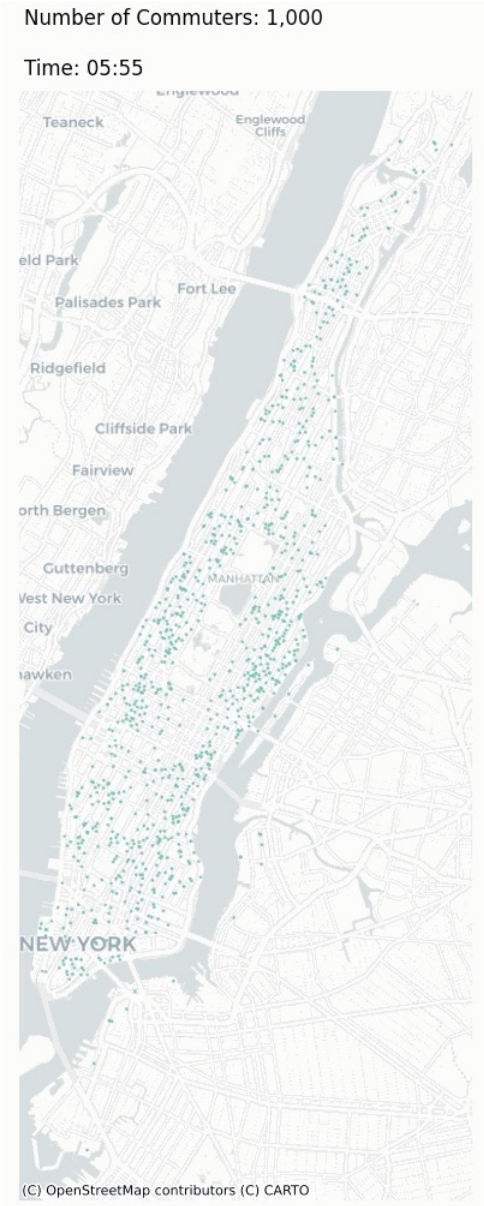


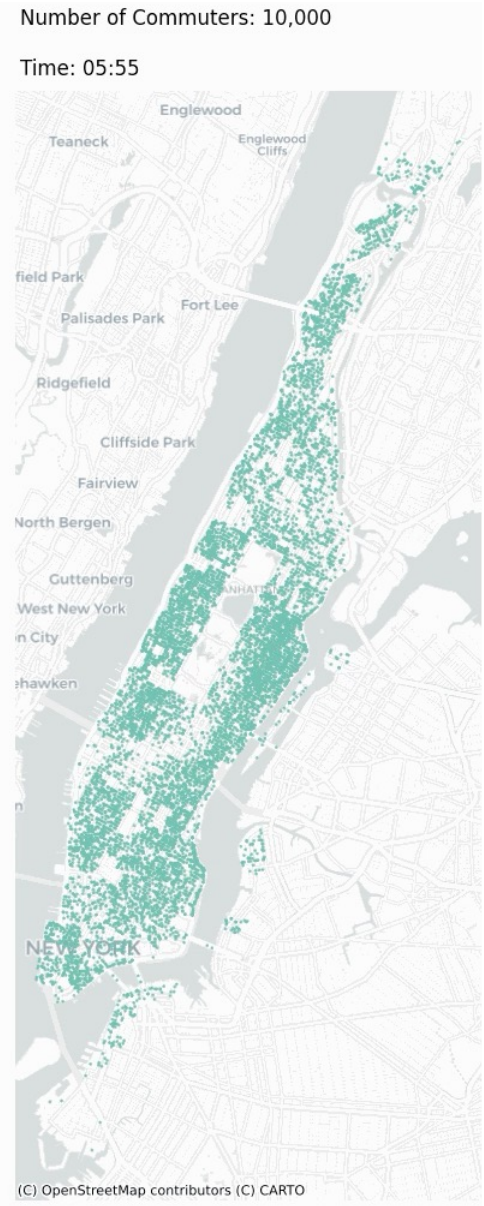
(b)

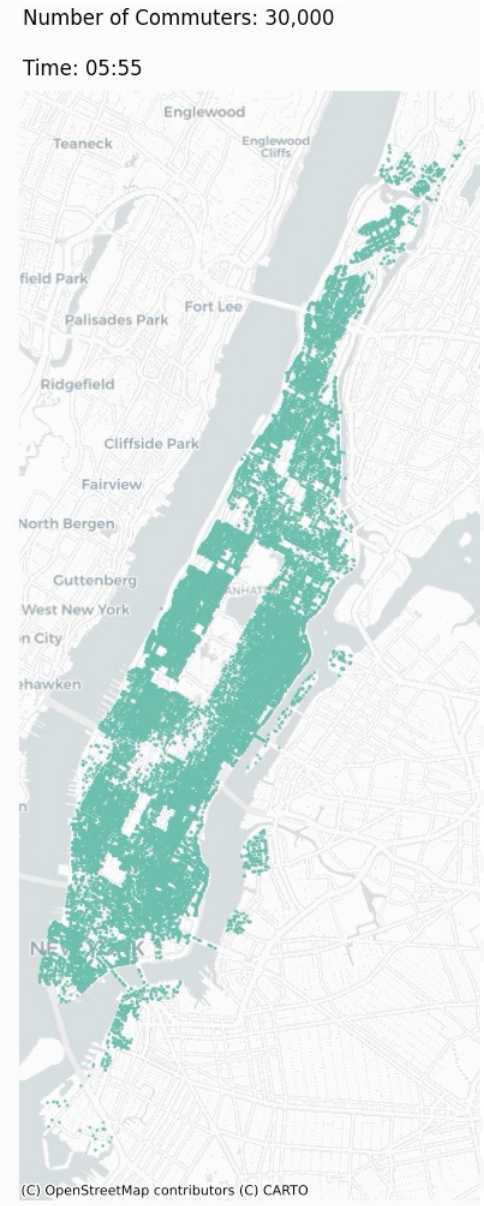


(c)

Number of commuters by status when total number of commuters is (a) 1,000, (b) 10,000, and (c) 30,000.


Number of Commuters: 1,000
Time: 05:55
Teaneck
Englewood Cliffs
Palisades Park
Fort Lee
Ridgefield
Cliffside Park
Fairview
Guttenberg
MANHATTAN
City
NEW YORK
(C) OpenStreetMap contributors (C) CARTO


Number of Commuters: 10,000
Time: 05:55
Englewood
Teaneck
Englewood Cliffs
Palisades Park
Fort Lee
Ridgefield
Cliffside Park
Fairview
Guttenberg
West New York
NEW YORK
(C) OpenStreetMap contributors (C) CARTO


Number of Commuters: 30,000
Time: 05:55
Englewood
Teaneck
Englewood Cliffs
Palisades Park
Fort Lee
Ridgefield
Cliffside Park
Fairview
Guttenberg
West New York
(C) OpenStreetMap contributors (C) CARTO

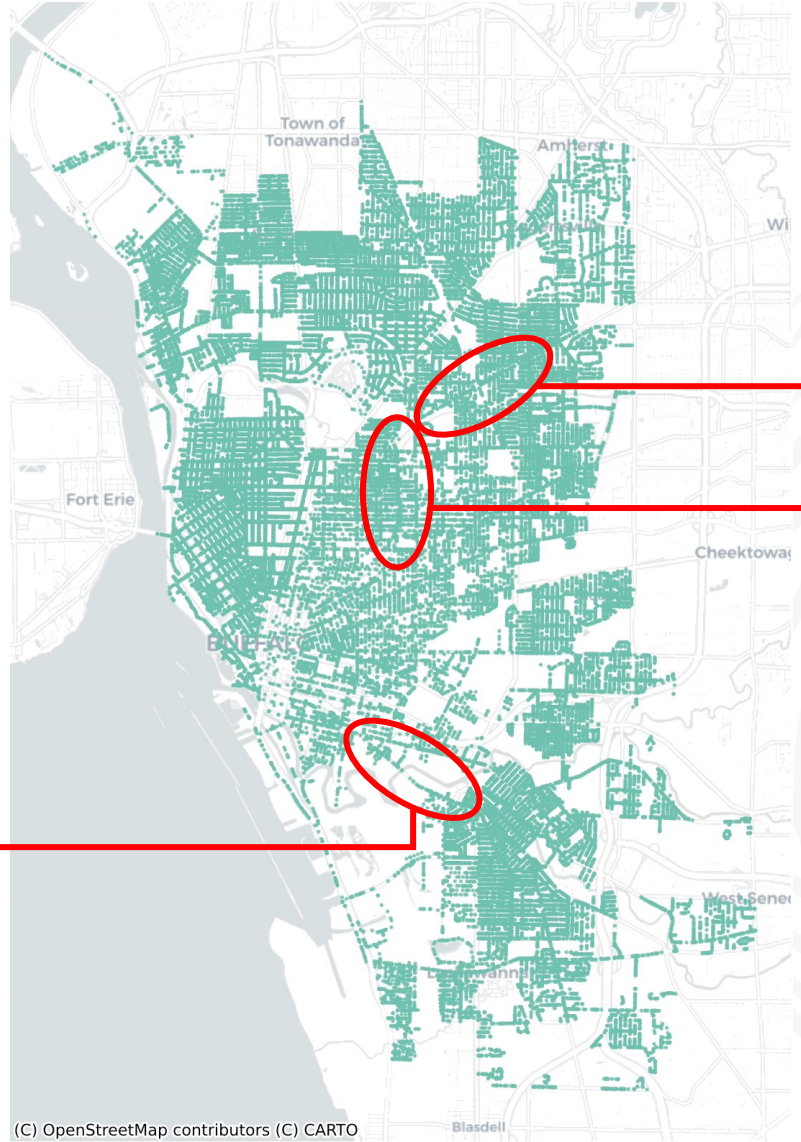
Emergence

Broadway

Emergence

Number of Commuters: 62,079

Time: 05:55



Kensington Avenue

Kensington Expressway

South Park Avenue

"**South Park Avenue** is the most important street in South Buffalo and is the primary alternative to the Buffalo Skyway (New York State Route 5) when traveling between the city and the southern suburbs."
https://en.wikipedia.org/wiki/South_Park_Avenue

Improving the Agent-Based Model

Calibrate model parameters (e.g., commuter speed) and validate results with travel surveys

- Average travel time, number of trips, etc.

Expand the geographic boundary

- To include more agents whose home and workplaces are outside the current study area

Expand the types of activities

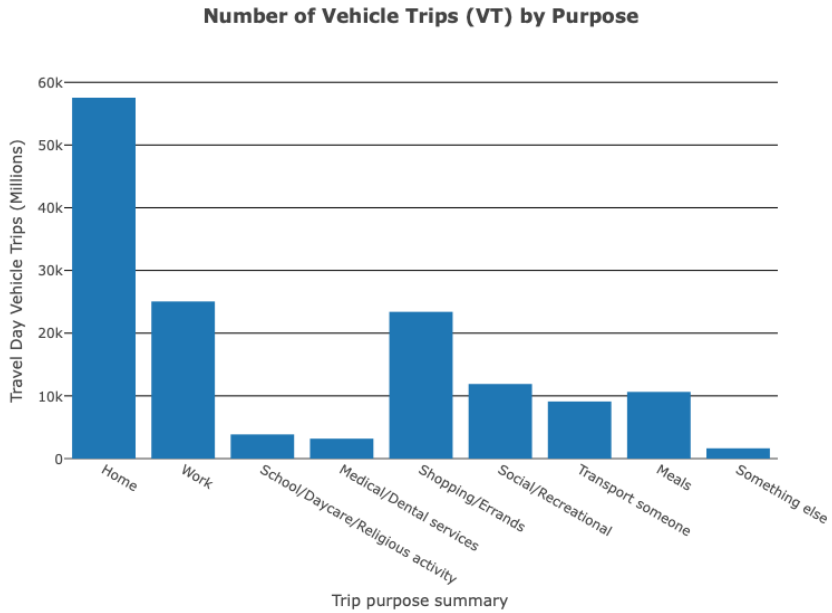
- To include travels between home, workplaces, shopping/errands places, schools, and recreational centers, etc.

Have heterogeneous agents

- People of different characteristics (e.g., age, gender) may behave differently

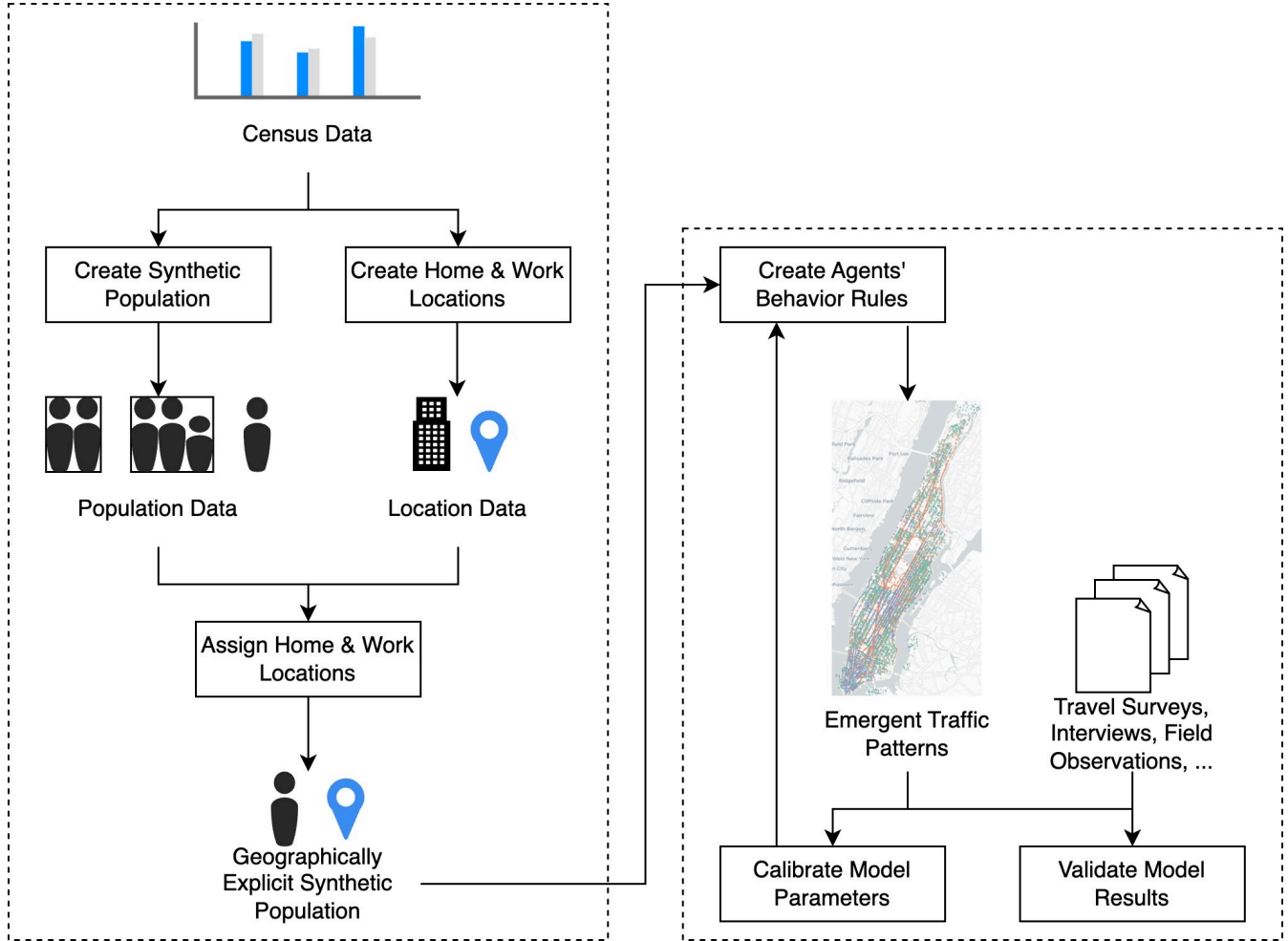
Include social behaviors (i.e., coordinated movements)

- Parents send kids to schools
- Friends go to the same restaurant from different home locations

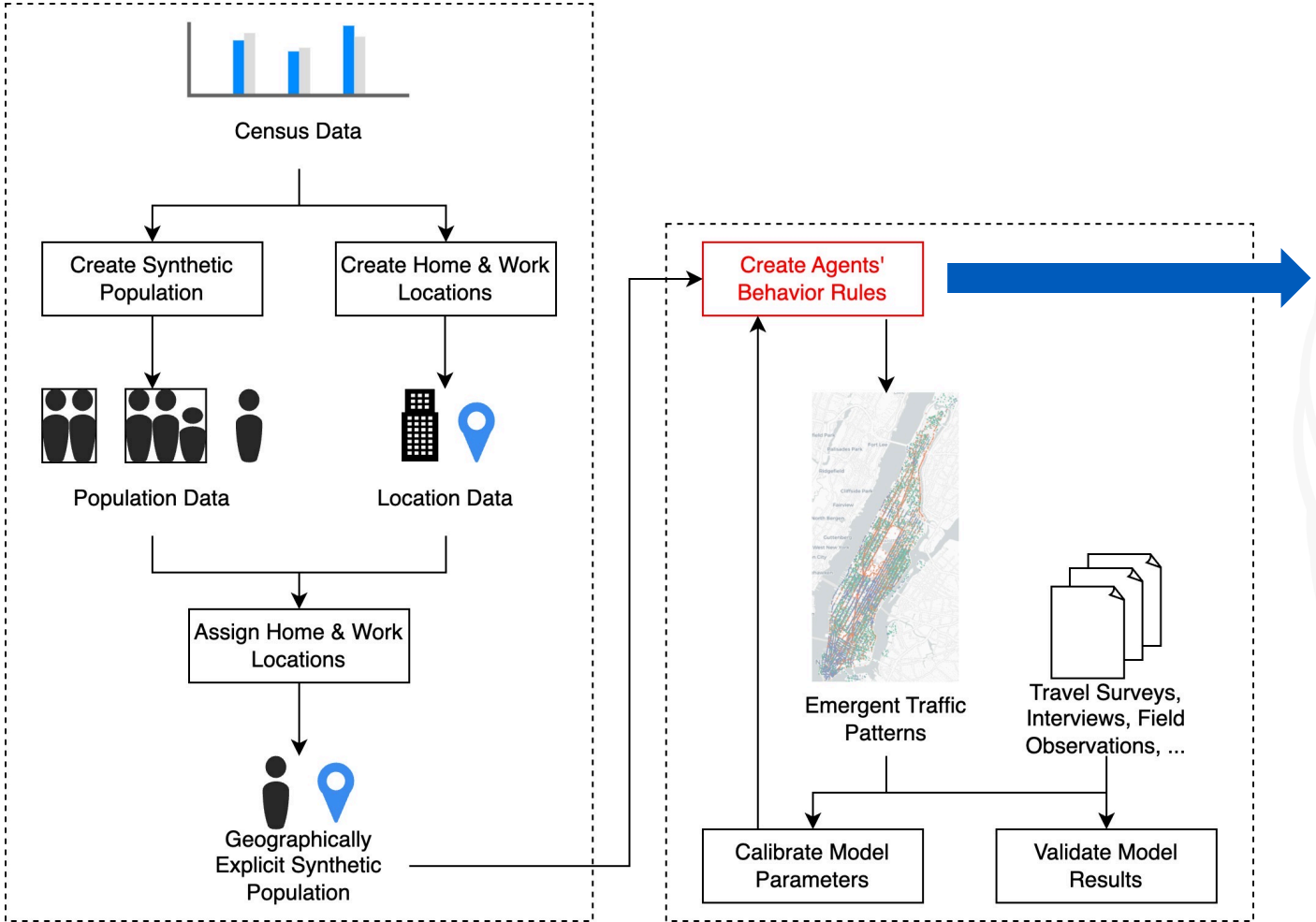


Source: Popular Vehicle Trips Statistics
https://nhts.ornl.gov/vehicle-trips

Improving the Agent-Based Model

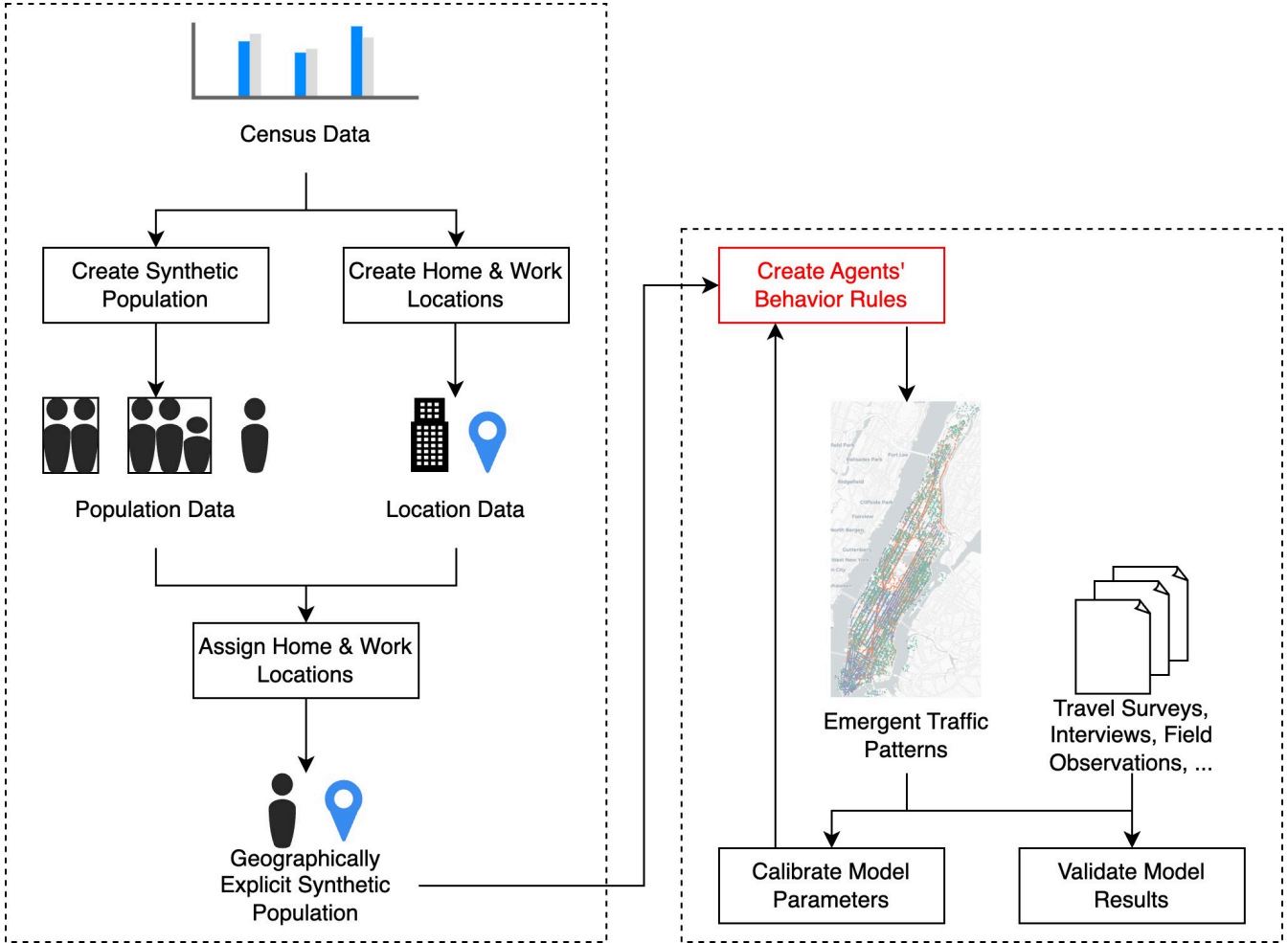


Improving the Agent-Based Model



- Go to work in the morning by shortest path between home and work locations.
- Travel at the same speed of 12km/h (adjustable).
- Stay at work for 8 hours.
- Travel back home by the same shortest path in the morning.

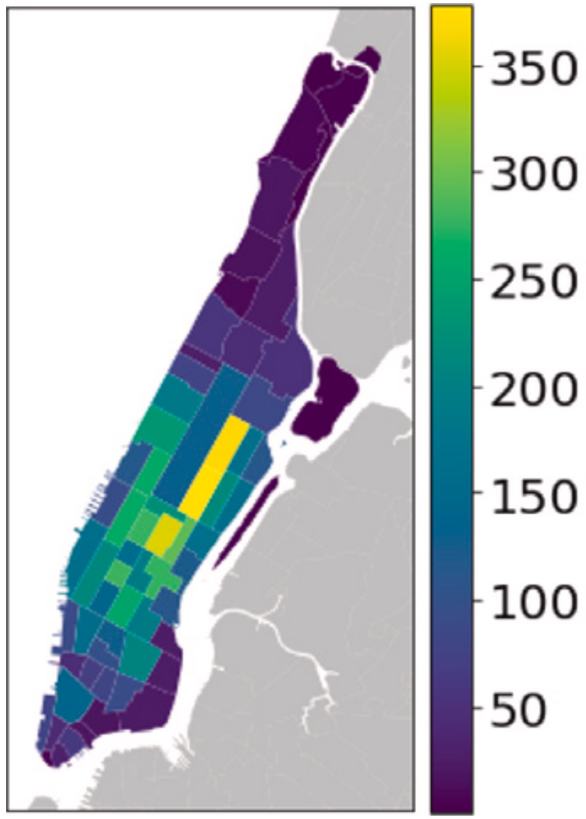
Improving the Agent-Based Model



A more data-driven approach?

Improving the Agent-Based Model

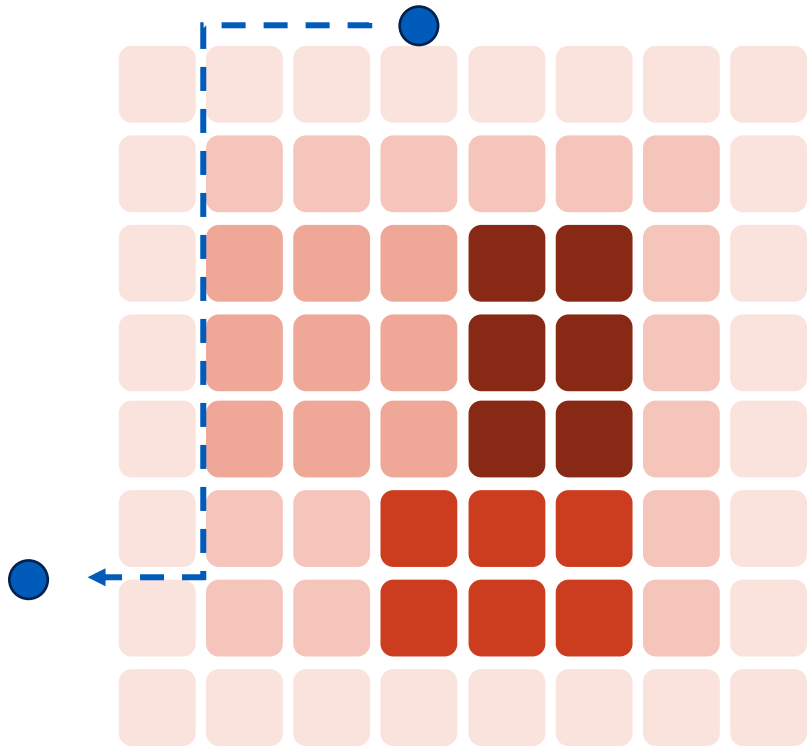
Taxi traffic as a proxy for urban human mobility



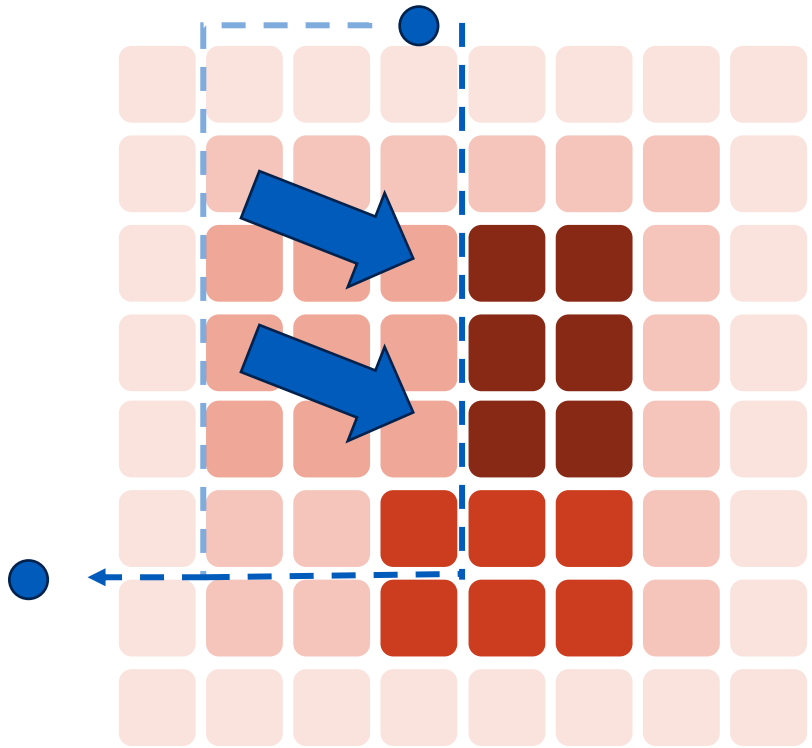
Taxi zones of Manhattan color-coded based on the average hourly activity from NYC Taxi data

Source: Tygesen, M. N., Pereira, F. C., & Rodrigues, F. (2023). Unboxing the graph: Towards interpretable graph neural networks for transport prediction through neural relational inference. *Transportation research part C: emerging technologies*, 146, 103946.

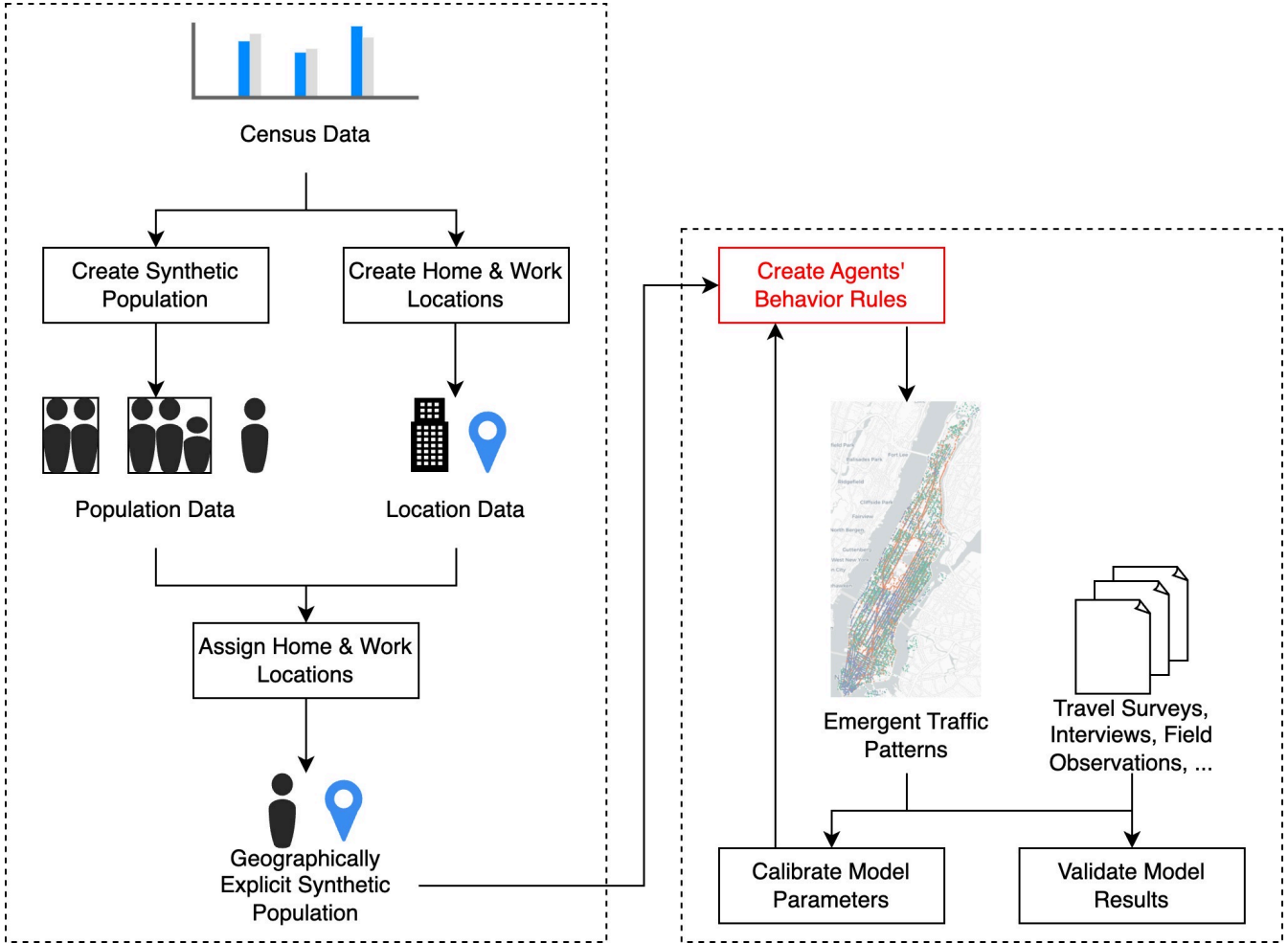
Most Probable Commute Path



Most Probable Commute Path

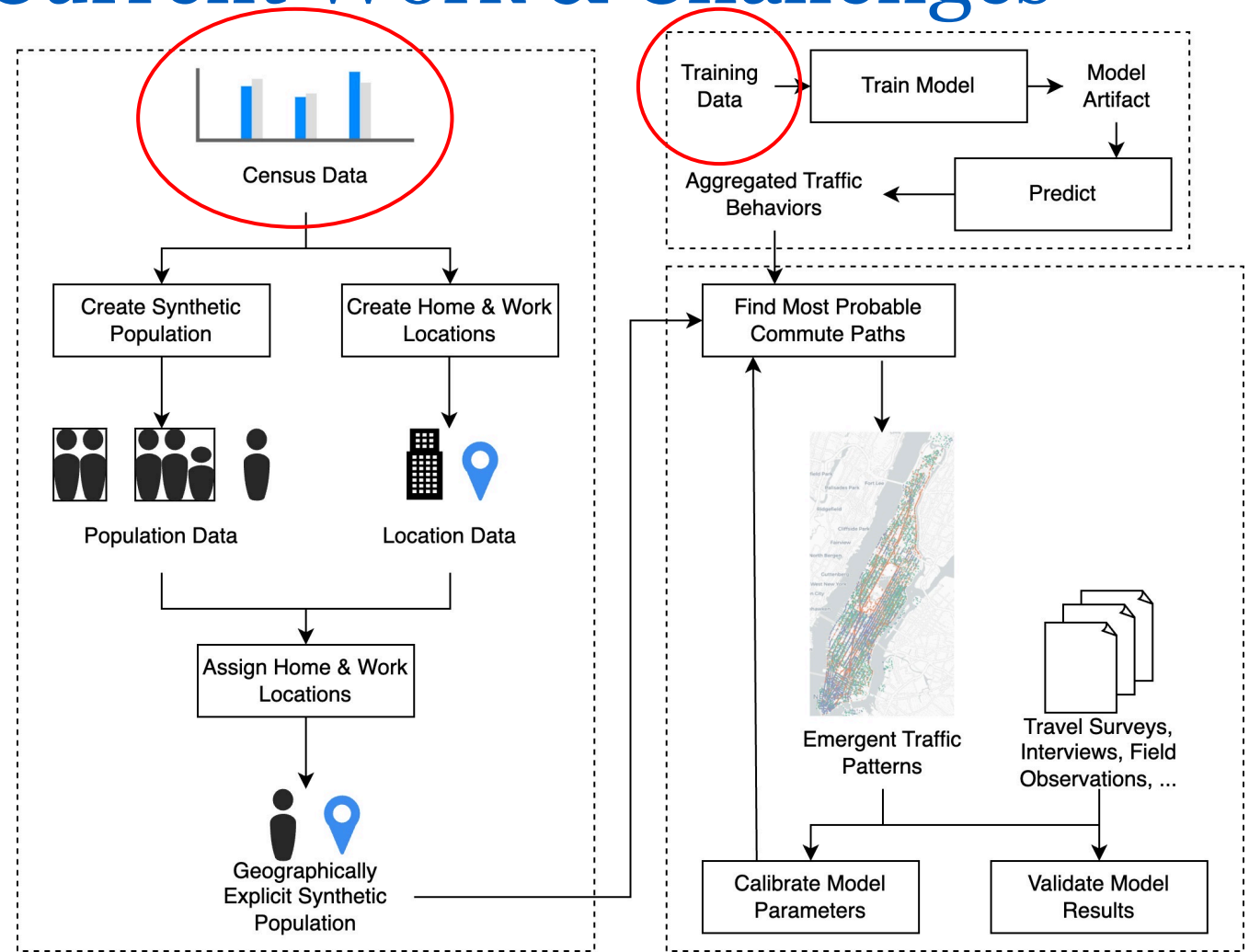


Improving the Agent-Based Model

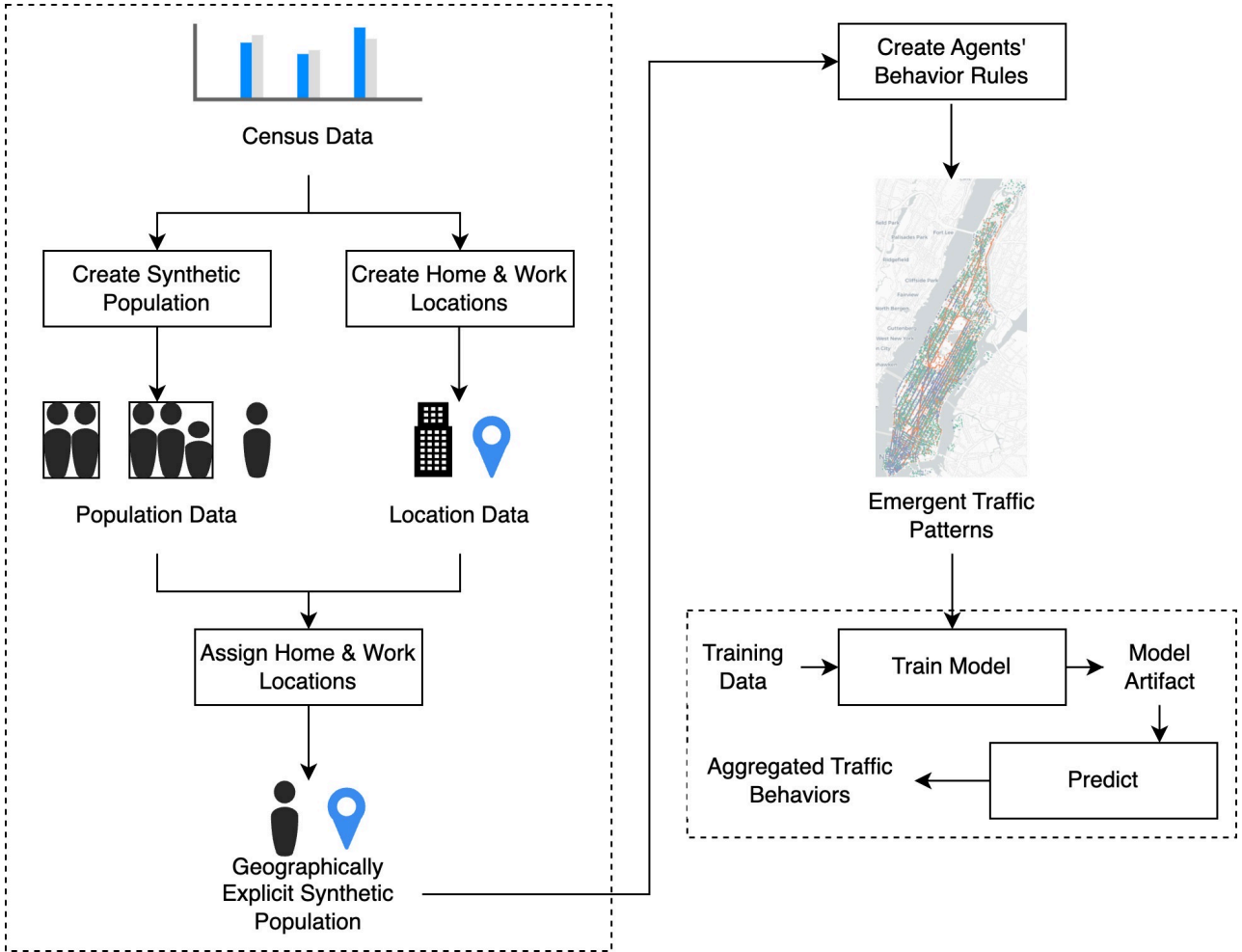


A more data-driven approach?

Current Work & Challenges



On A Side Note…



Conclusion

- An agent-based model of urban commuting in Manhattan is developed to simulate people's daily commute patterns, based on census survey data.
- Emerging traffic patterns can be observed through agents' simple behaviors.
- Such modeling technique can be complemented by recent advancements in GeoAI techniques.
- The model can be used as a basis to answer interesting "what-if" questions.
  - How many more minutes do people have to commute if certain roads are closed for maintenance?
  - Imagine a x% adoption rate of autonomous vehicles. Does is worth it to have one or more dedicated lanes for them?

University at Buffalo The State University of New York

Thank you for listening!
Welcome comments, questions and suggestions.

bwang44@buffalo.edu
atcrooks@buffalo.edu

@BoyuWang_
@AndyCrooks

https://wang-boyu.github.io
https://gisagents.org