Data Science Training Webinars

Spatiotemporal Simulation

Monte Carlo Method and Application in Urban Traffic Simulation

Yuxiang Zhong

Wuhan University

July 10, 2022

Center for Geographic Analysis

Harvard University

Geo-computation Center for Social Sciences

Wuhan University

China Data Institute

Future Data Lab

Contents

Spatial Data Disaggregation	空间数据拆分
Monte Carlo Simulation	蒙特卡洛模拟
Travel Demand Modeling	交通需求建模
Application of Monte Carlo method in the traffic simulation model	蒙特卡洛方法在交通需求建模中的应用
Case study	案例分析-基于蒙特卡洛的交通建模-以巴吞鲁日为例
Flow Chart	流程图
Workflow Implementation	KNIME工作流
The Outputs from the Workflow	KNIME工作流输出
Step by Step Workflow Execution	KNIME逐步执行
Discussion	讨论

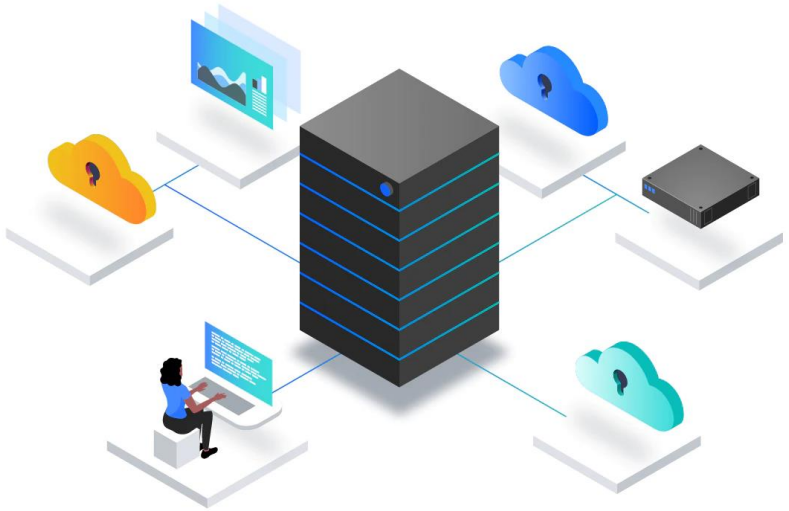
Spatial Data Disaggregation

Background:

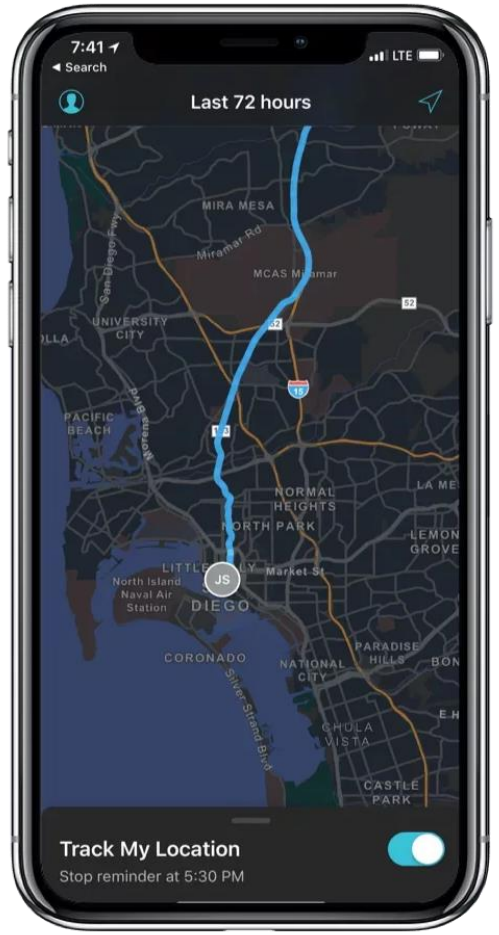
The scale at which spatial analysis is conducted tends to affect the result .

Spatial data often come as aggregated data in various areal units (sometime large areas) for various reasons:

- geo-privacy
- administrative convenience
- integration of various data sources
- limited data storage space



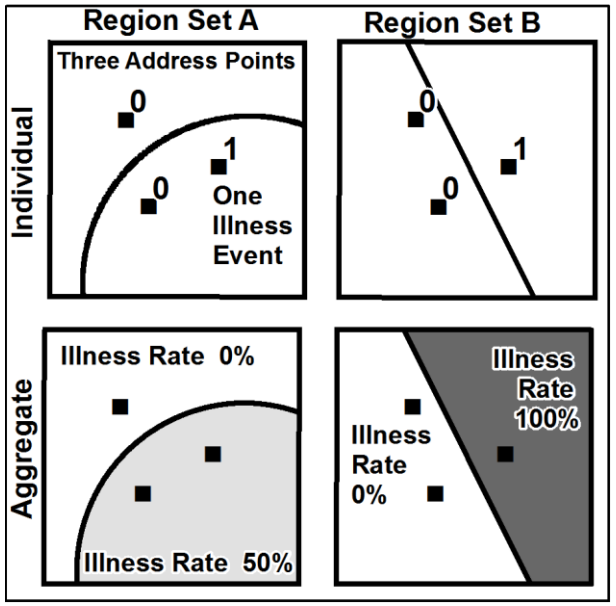
limited data storage space



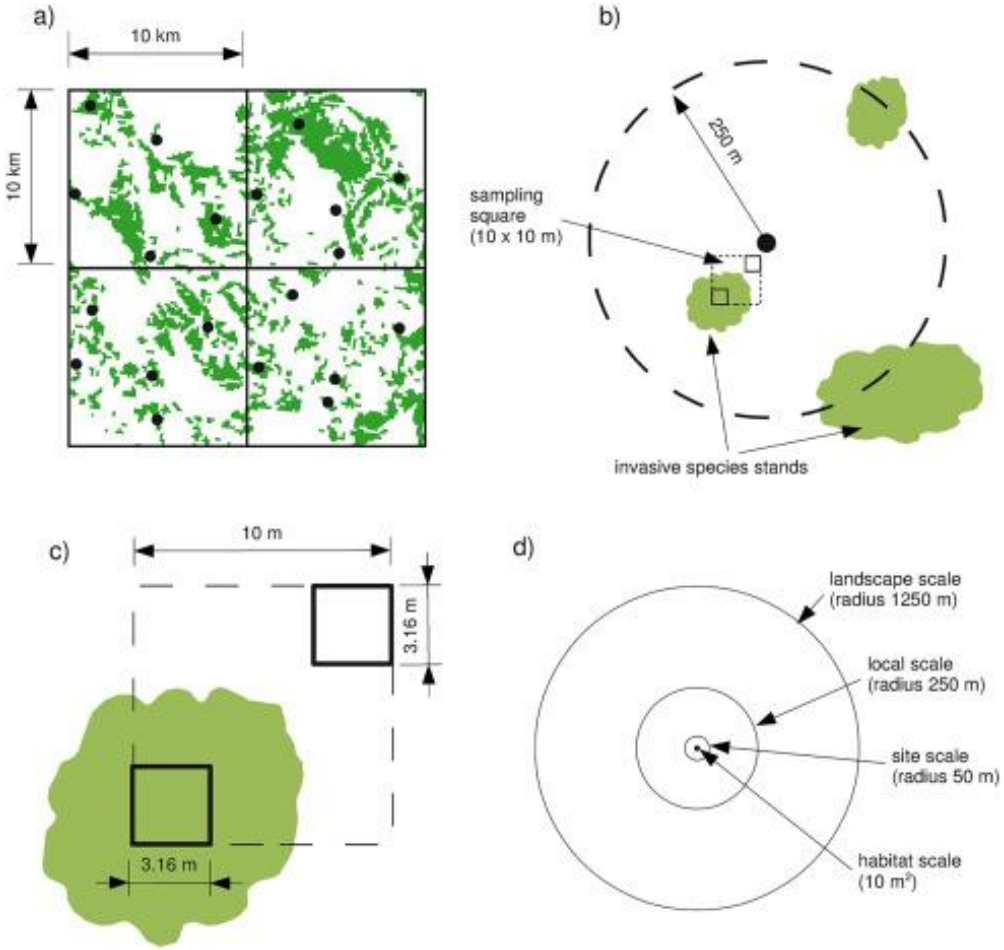
personal location information are shared

Spatial Data Disaggregation

- At broad scales, the high level of aggregation of data obscures the variability of geographic situations and, thus, 'dilutes' causal relationships.
- urban and regional planning, transport planning, environment, climate or geology often require high resolution socio-economic data.

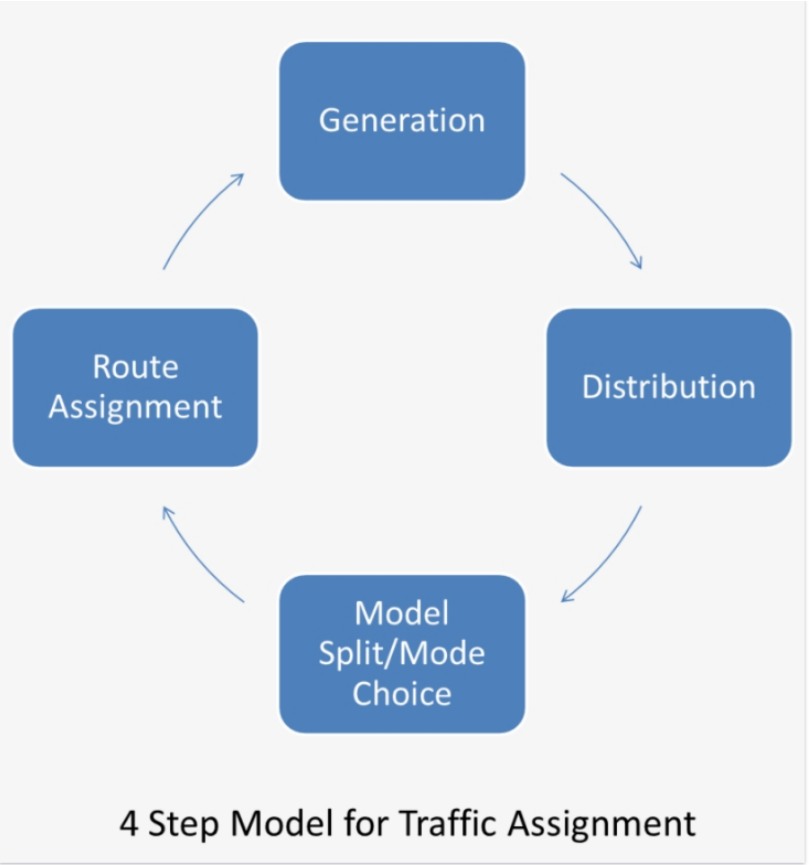


MAUP



Application of Four Step Model

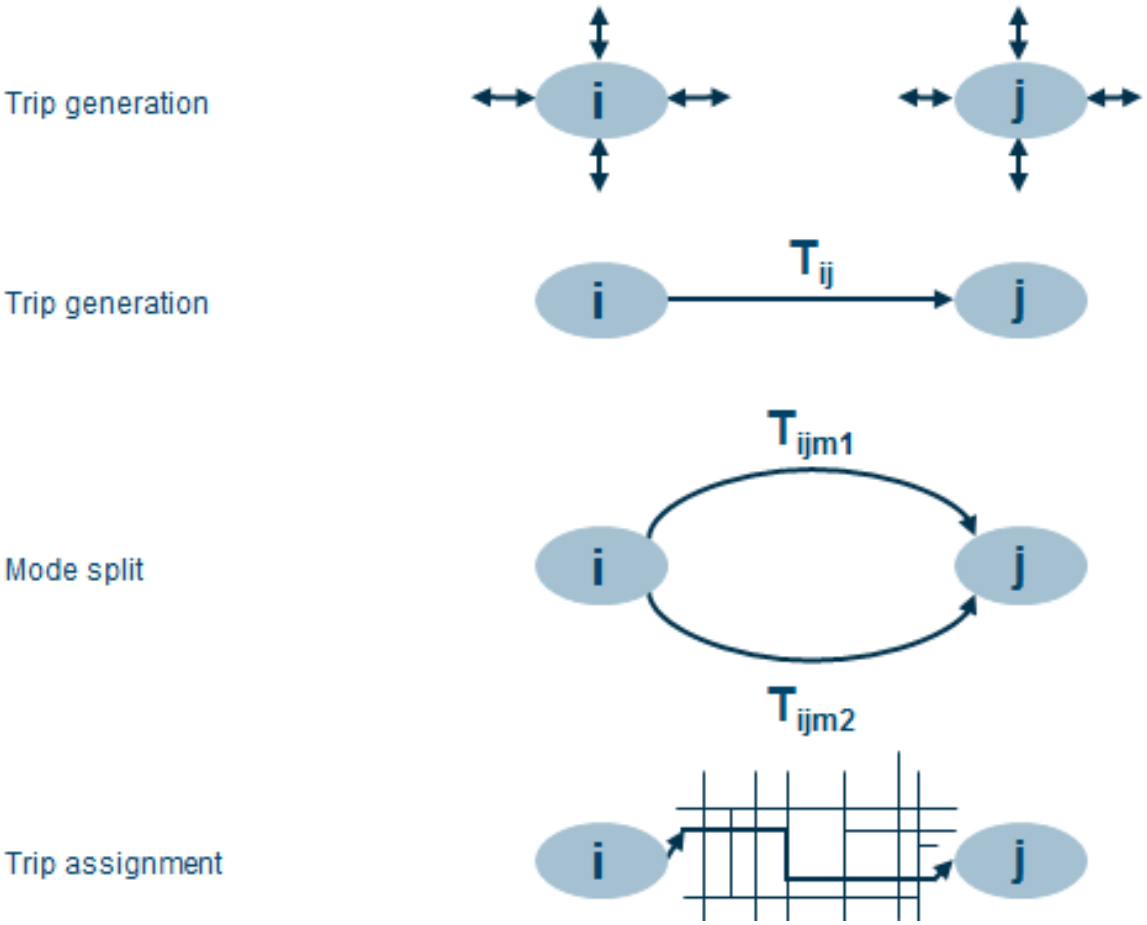
- ✓ Travel demand modeling(McNally, 2007)
- ✓ Excess/wasteful commuting measurement(Hu et al. 2015)
- ✓ Route Planning(Horváth et al. 2017)
- ✓ Freight Transportation and Distribution Systems(Hensher et al. 2007)



Travel Demand Modeling： four step model

The history of travel demand modeling, especially in the U.S., has been dominated by the four-step model (McNally, 2007)

- ✓ trip generation
- ✓ trip distribution
- ✓ mode choice
- ✓ trip assignment



Travel Demand Modeling： four step model

Trip generation

Traffic originated from areas or going to the same units is collectively referred to as trip generation

- **trip production**
  - ✓ estimated by trip purposes
  - ✓ examples : population, labor participation ratio, vehicle ownership, income

- **trip attraction**
  - ✓ different types of land use
  - ✓ examples : port, factory, hotel, hospital, school, office, store, service, park
  - ✓ various regression models

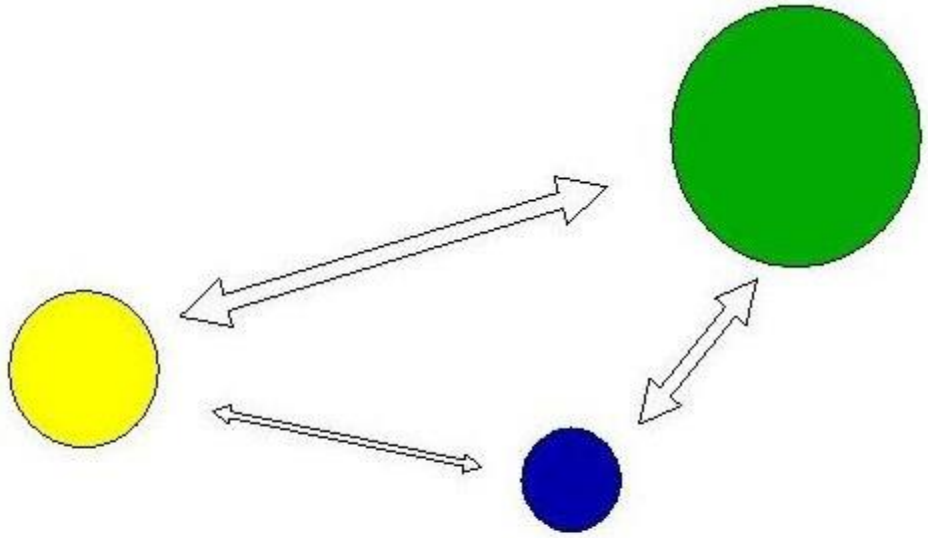
Travel Demand Modeling： four step model

Trip distribution

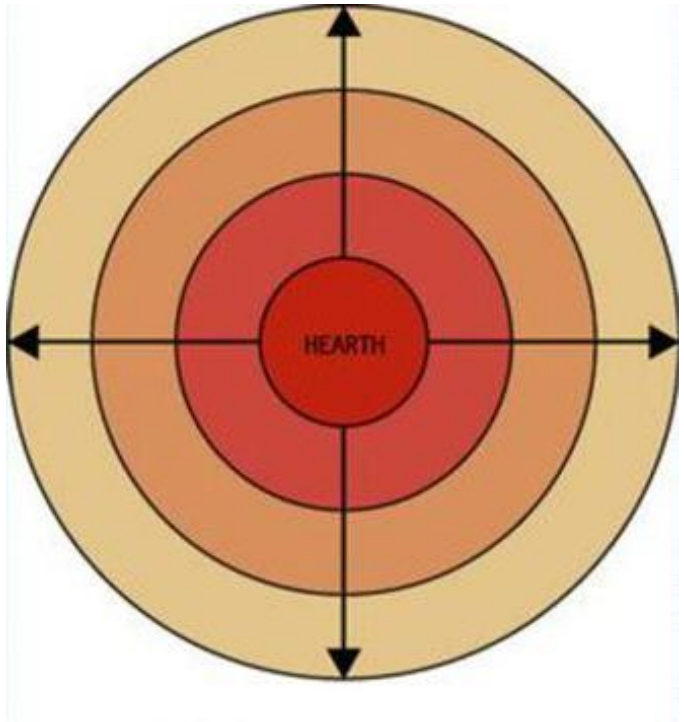
distributes the total number of trips from an origin (or to a destination) to specific O-D pairs

✓ **the gravity model**

$$T_{ij} = \alpha O_i D_j d_{ij}^{-\beta}$$



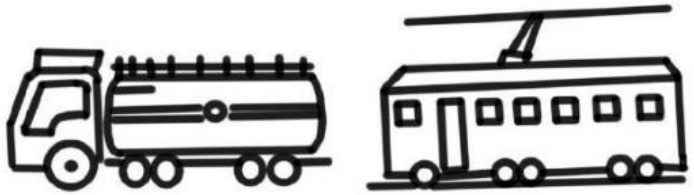
The shorter the distance between two objects, and the greater the mass of either (or both) objects, the greater the gravitational pull between the objects.



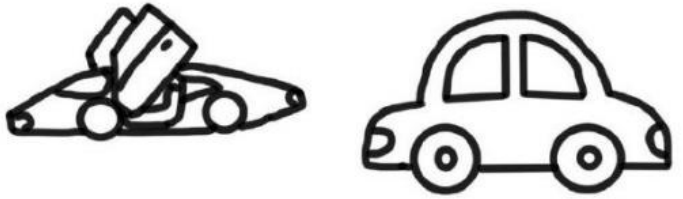
Travel Demand Modeling： four step model

Mode choice analysis

splits the traffic among available modes such as drive-alone, carpool, public transit, bicycle and others



- **attributes associated with a mode**



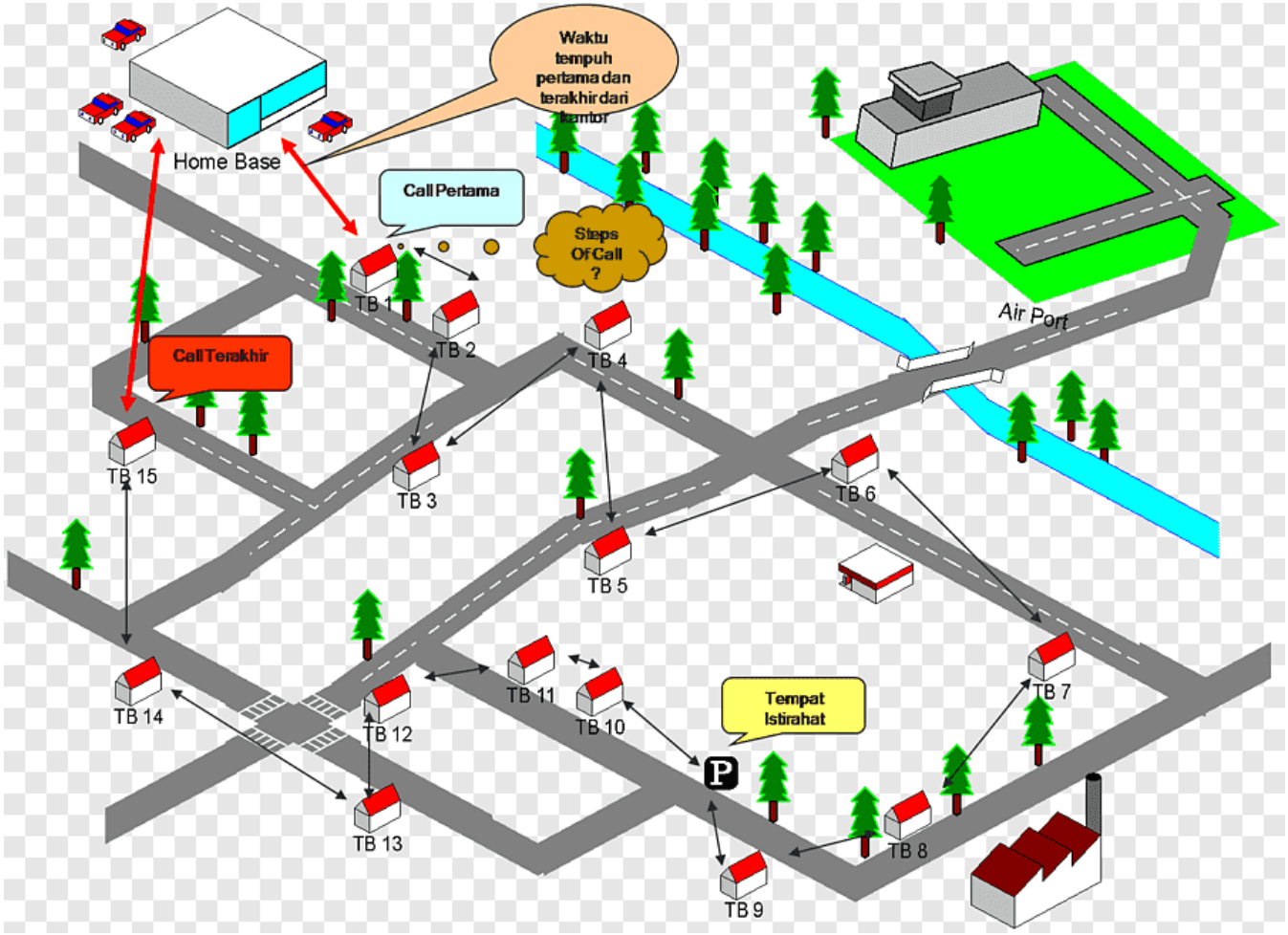
  - ✓ travel time
  - ✓ travel cost
  - ✓ convenience
  - ✓ comfort
  - ✓ trip purpose
  - ✓ automobile availability and reliability

Travel Demand Modeling： four step model

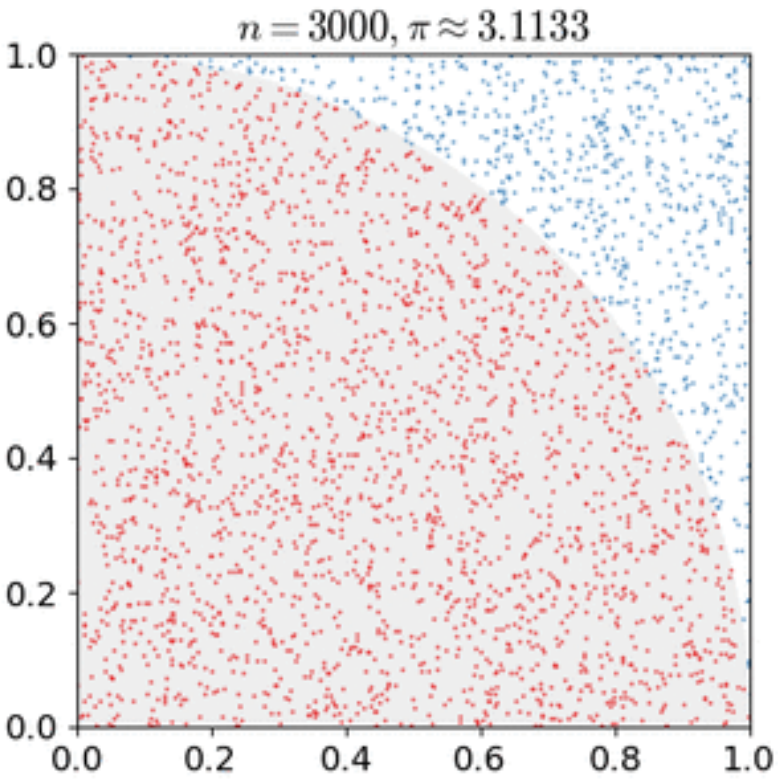
Trip assignment

adjusting traffic assignments on various routes in response to the change of speeds dependent on traffic flow until an equilibrium is reached.

equilibrium



Monte Carlo Simulation



Monte Carlo : A numerical analysis technique that uses repeated random sampling to generates suitable random numbers of parameters or inputs to explore the behavior of a complex system or process.

- ✓ Markov Chain Monte-Carlo(Van et al. 2018)
- ✓ Monte-Carlo based uncertainty analysis(Janssen et al. 2013)
- ✓ approximate Bayesian computation(Del et al. 2012)

The random numbers generated follow a certain probability distribution function (PDF) that describes the occurrence probability of an event. Some common PDFs include:

- uniform distribution
- normal distribution
- lognormal
- discrete distribution

Application of Monte Carlo method in the traffic simulation model

- **disaggregate area data to individual points within the areas**

  The first task is to disaggregate areal data randomly to individual points so that the density pattern of simulated points reflects the pattern at the zonal level.

- **disaggregae inter-zonal flow data to individual flows between points**

  The second is to randomly connect the points between zones, and the volume of resulting connections is consistent with a predefined (observed) inter-zonal flow pattern.

Case Study

Title Monte Carlo Method and Application in Urban Traffic Simulation

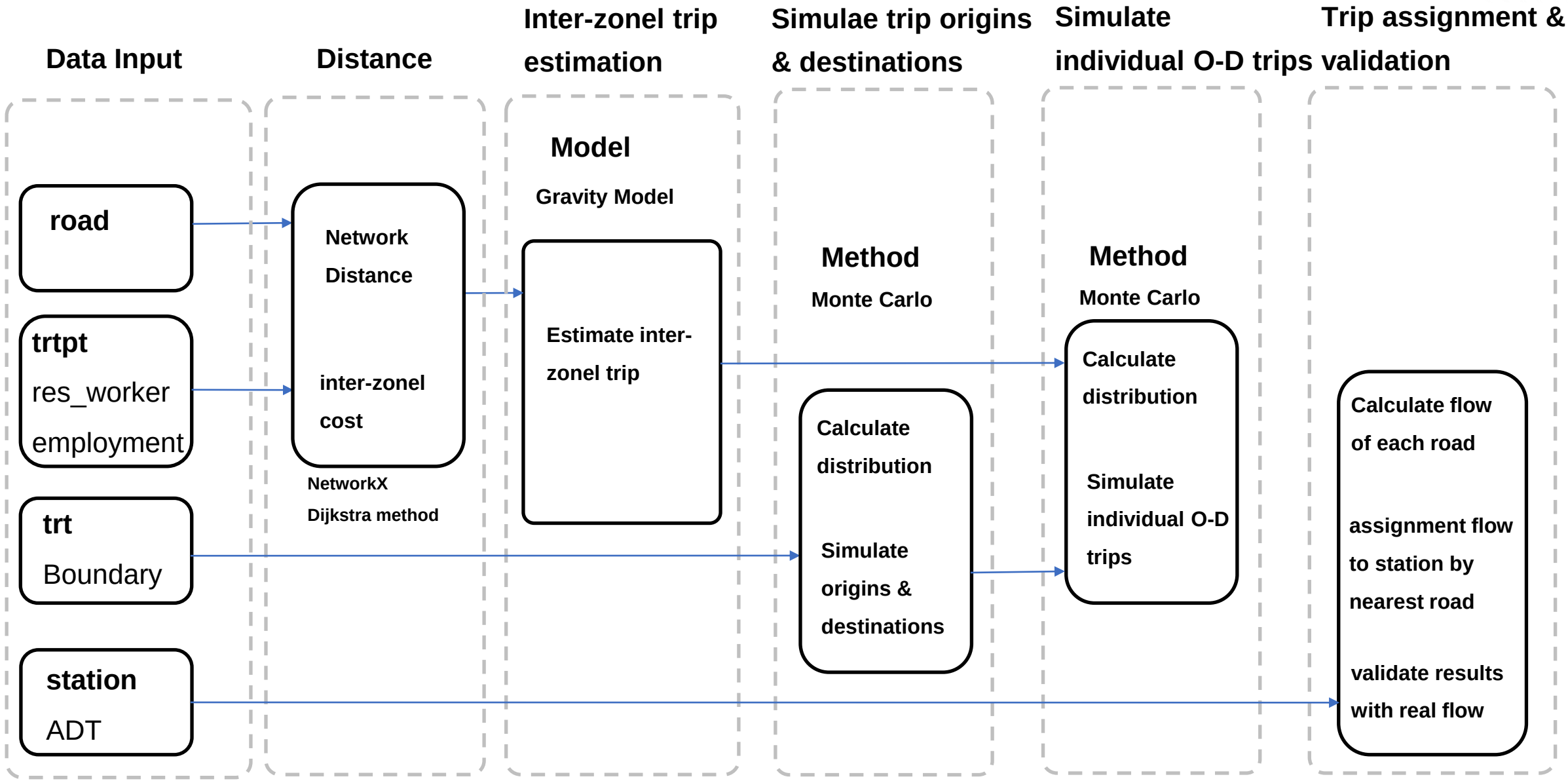
Objectives Spatial data disaggregation

Methodology Monte Carlo Method

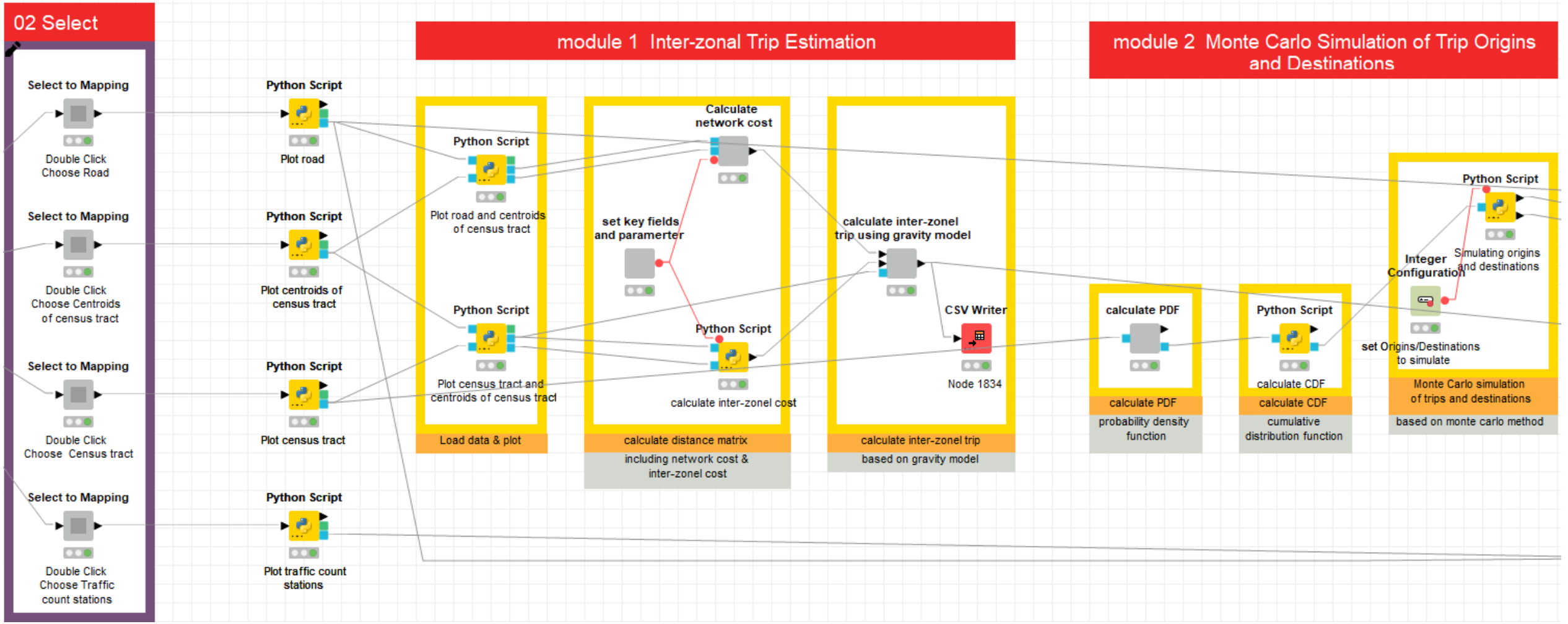
Data

Shapefile	Type	File Detail	Key Fields	Key Variable	Data Source
trt	Polygon	census tract	**res_worker**	resident workers in each census tract	2005 East Baton Rouge Parish
			employment	jobs in each census tract	
trtpt	Point	census tract centroids			2005 East Baton Rouge Parish
road	Polyline	road network dataset including the edges			road shapefile in Baton Rouge Parish 2005
station	Point	816 traffic count stations	**ADT**	annual average daily traffic	AADT in 2005

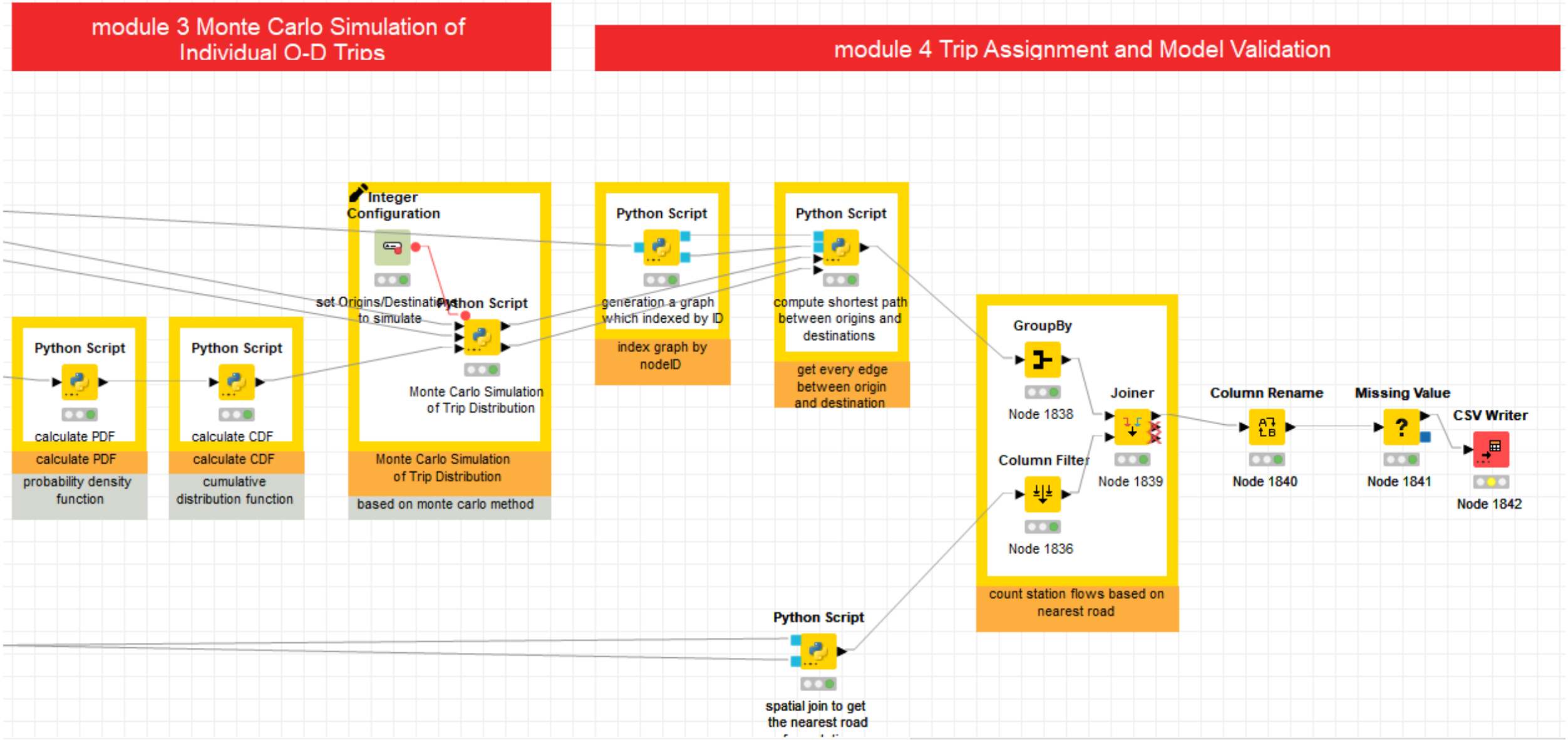
Flowchart for Case Study:Module 4



The Workflow Implementation

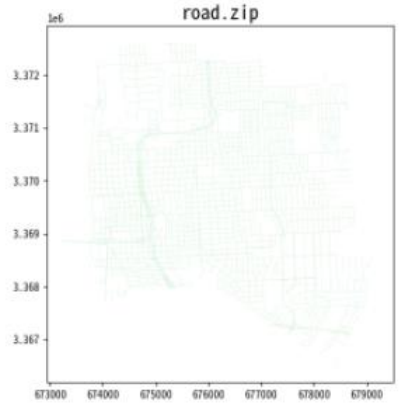


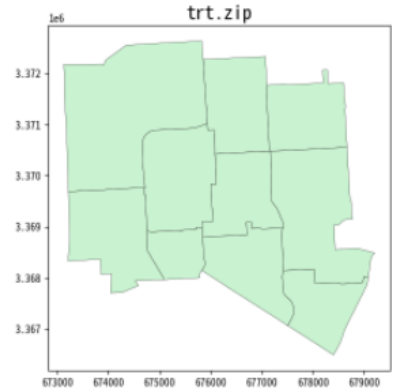
The Workflow Implementation

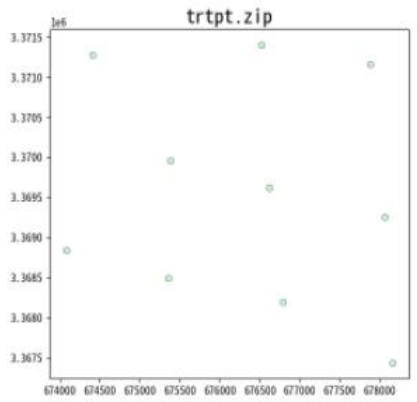


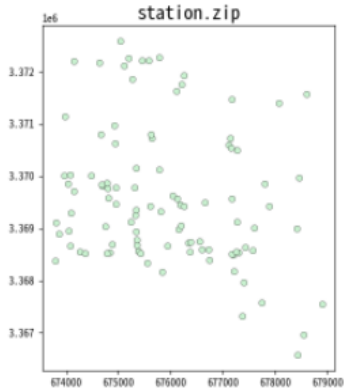
The Workflow Implementation

Preparation









Upload GeoData zip

Select file road.zip

Upload GeoData zip

Select file trtpt.zip

Upload GeoData zip

Select file trt.zip

Upload GeoData zip

Select file station.zip

Road Shapefile Download

Shapefile (zip)

TractCentroid Shapefile Download

Shapefile (zip)

Tract Shapefile Download

Shapefile (zip)

Station Shapefile Download

Shapefile (zip)

The Workflow Implementation

Step 1:Define UID Field & Speed Parameter to Calculate Network cost

explanation for key fields

Field	meaning
UID	the unique identification for cencus tract
Speed	the speed on the road network

set UID field

trtID

speed

670.56

Show 10 entries

Search:

Row Index	RowID	24	25	29	30	34	35	40	103
0	24	4.5803825408091745	4.555266936989233	4.724656112771416	5.8365777246368715	3.745789809618081	4.768000139435696	4.496705782730107	6.304030303190617
1	25	4.555266936989233	4.530151333169292	4.699540508951475	5.8114621208169295	3.7206742057981392	4.7428845356157545	4.471590178910165	6.278914699370675
2	29	4.724656112771416	4.699540508951475	4.868929684733656	5.980851296599112	3.8900633815803216	4.912273711397936	4.640979354692346	6.448303875152858
3	30	5.8365777246368715	5.8114621208169295	5.980851296599112	7.092772908464568	5.001984993445777	6.024195323263392	5.752900966557803	7.560225487018314
4	34	3.745789809618081	3.7206742057981392	3.8900633815803216	5.001984993445777	2.9111970784269867	3.933407408244602	3.662113051539013	5.469437571999523
5	35	4.768000139435696	4.7428845356157545	4.912273711397936	6.024195323263392	3.933407408244602	4.955617738062217	4.684323381356627	6.491647901817138
6	40	4.496705782730107	4.471590178910165	4.640979354692346	5.752900966557803	3.662113051539013	4.684323381356627	4.413029024651038	6.220353545111549
7	103	6.3040303031906175	6.2789146993706755	6.448303875152858	7.560225487018314	5.469437571999523	6.491647901817138	6.220353545111549	8.02767806557206
8	104	5.087487590845711	5.062371987025769	5.231761162807952	6.343682774673408	4.252894859654617	5.2751051894722325	5.003810832766644	6.811135353227154
9	105	4.817567303820688	4.792451700000747	4.961840875782928	6.073762487648384	3.982974572629594	5.005184902447208	4.733890545741619	6.54121506620213

Showing 1 to 10 of 10 entries

Previous 1 Next

The Workflow Implementation

Step 2:Estimate Inter-zonal Trip Based on Gravity Model

set scalar factor a

0.001

set distance friction coefficient b

2

set size_origin field

res_worker

set size_destinaton field

mployment

Row Index	RowID	O_ID	D_ID	trips	O_ID (Iter #1)	D_ID (Iter #1)	trips (Iter #1)
0	Row0	24	24	322	24	24	28
1	Row1	24	25	89	24	25	3
2	Row2	24	29	124	24	29	4
3	Row3	24	30	172	24	30	3
4	Row4	24	34	13	24	34	0
5	Row5	24	35	55	24	35	1
6	Row6	24	40	65	24	40	1
7	Row7	24	103	325	24	103	5
8	Row8	24	104	104	24	104	1
9	Row9	24	105	220	24	105	6

		Search O_ID	Search D_ID	Search trips	Search O_ID (Iter #1)	Search D_ID (Iter #1)	Search trips (Ite

Showing 1 to 10 of 100 entries

Previous 1 2 3 4 5 … 10 Next

The Workflow Implementation

Step 3:Monte Carlo Simulation of Trip Origins and Destinations

set size_destinaton field

mployment

set size_origin field

res_worker

simulateThreshold

100

O_X	O_Y	ZoneID
702478.676975415	3354864.14641448	1
704696.181908568	3356447.85309883	1
703150.575862824	3353305.14599347	1
702616.937091235	3356425.65539643	1
700390.308359557	3358023.67571273	1
703965.754284229	3357388.38437156	1
701614.892089973	3354112.02507886	1
704128.799248364	3354751.5320685	1
706279.470542048	3356671.26154444	1
705344.364730125	3355840.40383864	1

D_X	D_Y	ZoneID
703925.064336033	3352788.91663881	1
701996.466588677	3355405.19818109	1
702617.815818116	3353400.53498993	1
704494.29587378	3353617.53444397	1
702463.868745075	3356910.79300024	1
703133.228899591	3354510.080814	1
702592.013540929	3355638.88608011	1
705097.355685674	3358105.60964829	1
703250.215380562	3352950.56814698	1
703621.956392952	3354612.41286826	1

The Workflow Implementation

Step 4:Monte Carlo Simulation of Trip Distribution

Define the LoopCount

3

tripsThreshold

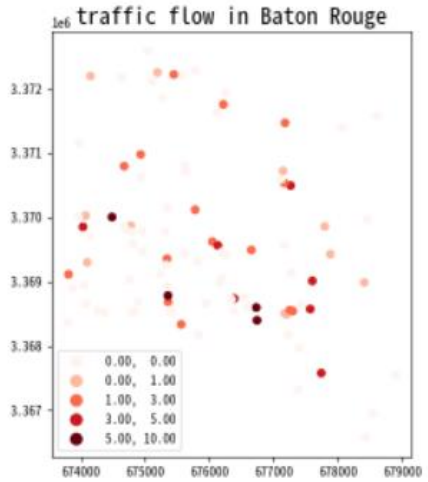
100

O_x	O_y	D_x	D_y	RouteID
675373.199978953	3365792.95753157	674201.4503061579	3372610.8395192	1
675273.738295522	3365473.97753462	686728.018113684	3366554.05191123	2
689900.130726309	3373456.45116396	696849.205588825	3371836.5604318	3
684119.033402859	3372355.26618587	686469.3059347	3365686.15180851	4
675617.285730249	3371378.63735532	689304.53822283	3370248.67640257	5
692250.410372493	3361498.38176478	674742.564185607	3370896.8550695	6
681778.4611917561	3363247.76060866	680095.487324084	3368598.23726603	7
679109.8504362471	3363472.94971678	678862.251640553	3376854.59760377	8
686686.730337439	3367838.6102334	672391.585363013	3365093.29007341	9
675385.709210098	3372587.23641693	684852.148533471	3363078.01928482	10

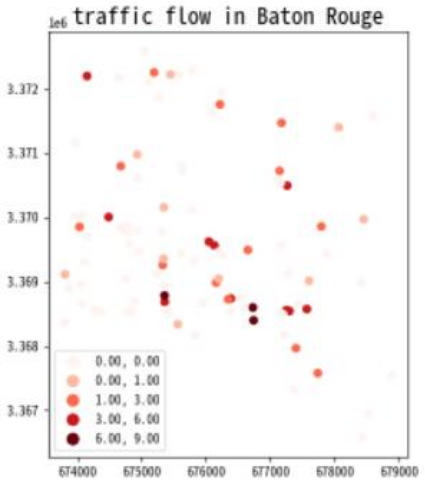
The Workflow Implementation

Result :Trip Assignment

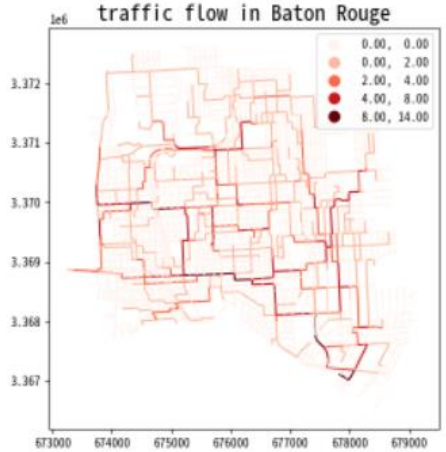
station flow(a=0.001,b=2)



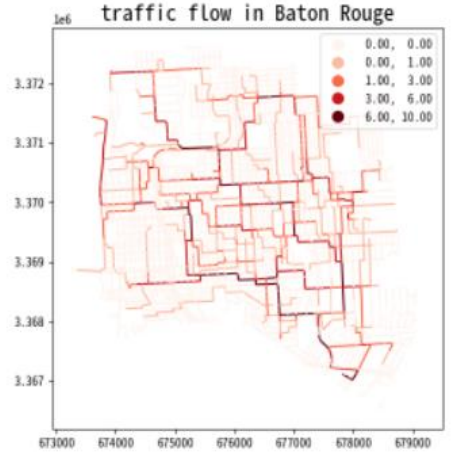
station flow(a=0.002,b=4)



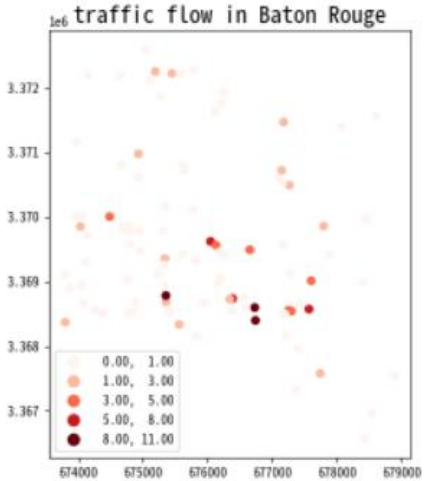
road net flow(a=0.001,b=2)



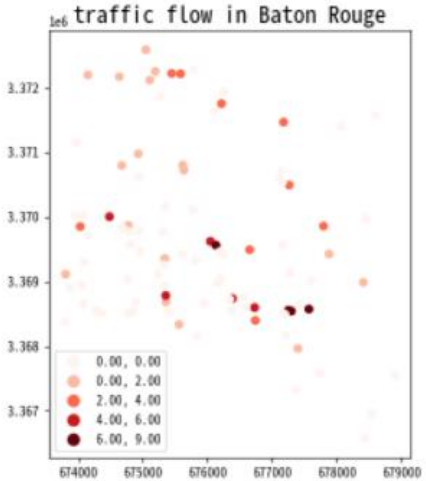
road net flow(a=0.002,b=4)



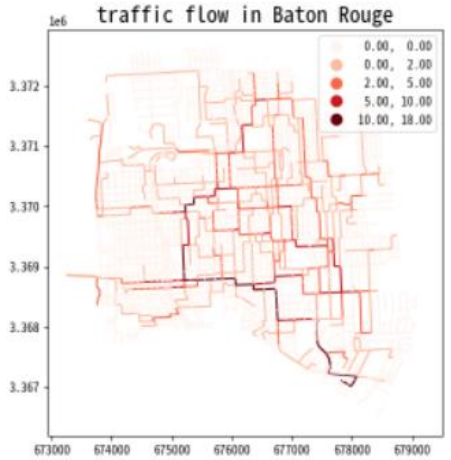
station flow(a=0.003,b=6)



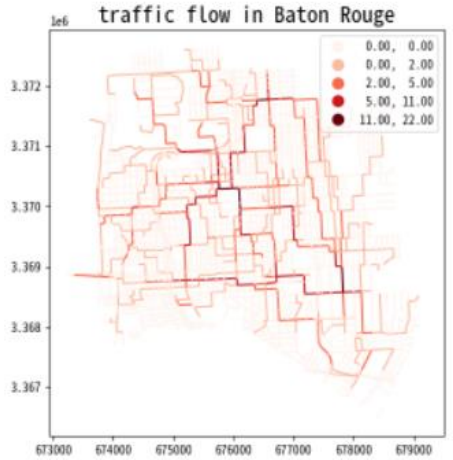
station flow(a=0.004,b=8)



road net flow(a=0.003,b=6)

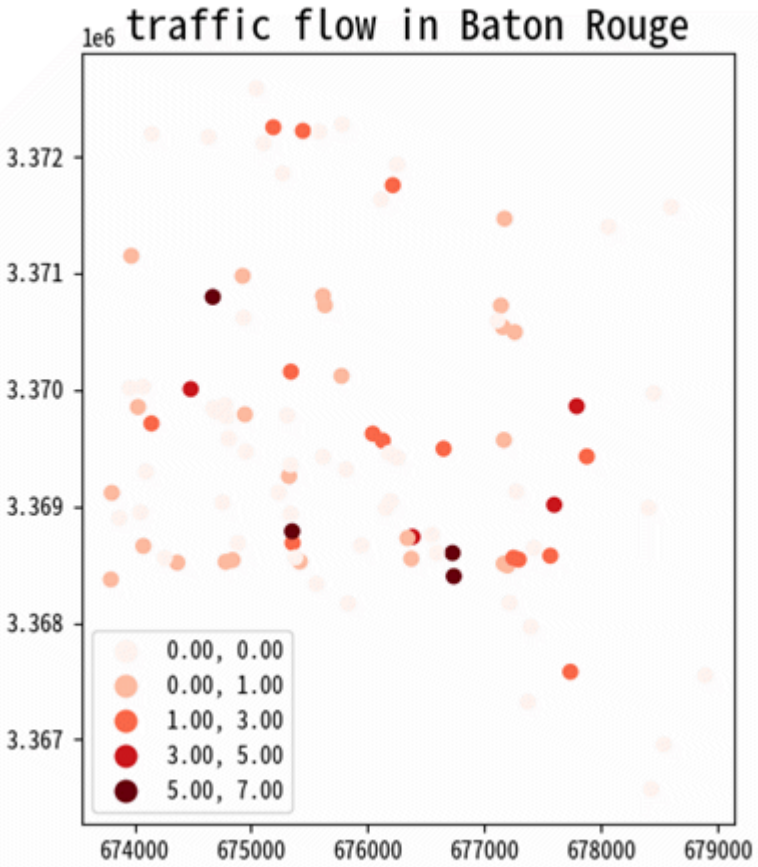


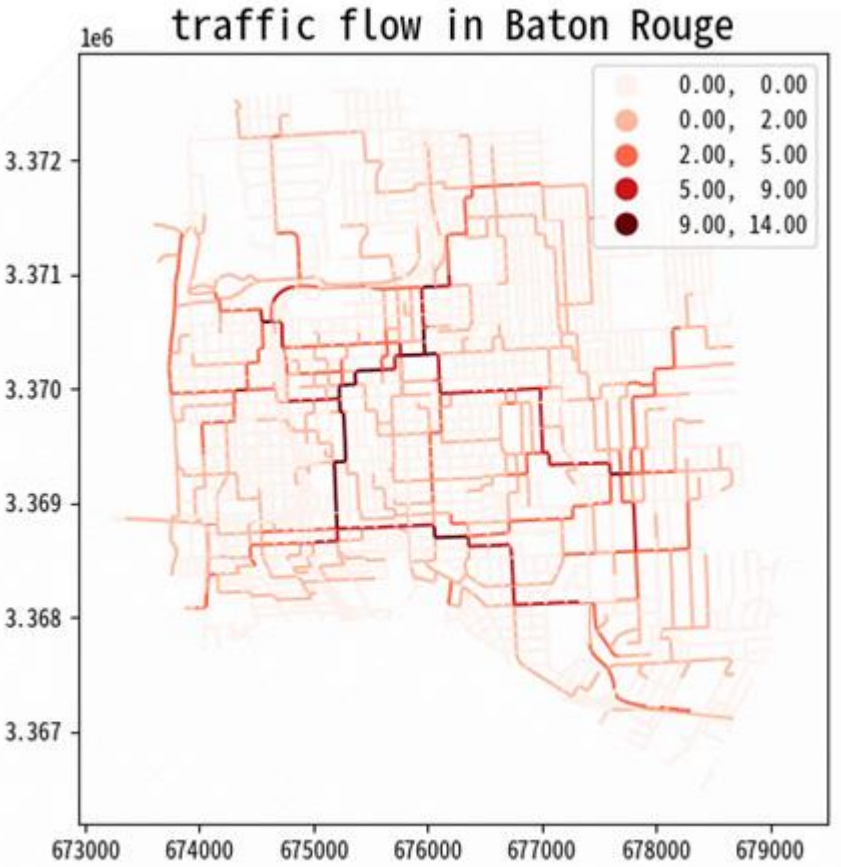
road net flow(a=0.004,b=8)



The Workflow Implementation

Result :Trip Assignment





Future Work

✓ Consider multimodal transportation networks

In this case, travel speed is set as a constant value. Indeed, individual intend to take different transport(e.g., carpool, transit, walking or bicycling).

✓ consider better trip assignment method based on better traffic equilibrium model

In this case, we adjust traffic assignments on various routes based on minimum travel cost principle. In future work, we could consider better traffic equilibrium model.

✓ Consider temporal variation of travel demand & network state

Indeed, in different period of the day, travel demand of individuals vary. The network status also changes (e.g., traffic congestion, reversible lanes). These factors could be taken into consideration.

References

Fahui Wang, 2015. Quantitative Methods and Socio-Economic Applications in GIS. https://www.routledge.com/Quantitative-Methods-and-Socio-Economic-Applications-in-GIS/Wang/p/book/9781138843622.

Hu, Y., & Wang, F. (2015). Decomposing excess commuting: A Monte Carlo simulation approach. Journal of Transport Geography, 44, 43-52.

Web Site and Contacts

Project web site: http://spatialdatalab.org

Co-mentors:

Yu, Hanchen, hanchenyu@fas.harvard.edu
Lingbo Liu, lingbo.liu@whu.edu.cn

Advisors:

Guan, Wendy, wguan@cga.harvard.edu
Shuming Bao, sbao@umich.edu

Files on Google Drive:
https://drive.google.com/drive/u/0/folders/1q7rpxck1BPx1yDzENOLW7XH9RKEZMri8

THANKS

Spatial Data Lab http://spatialdatalab.org

Contact Spatialdatalab@lists.fas.harvard.edu

zyxiang@whu.edu.cn

Center for Geographic Analysis

Harvard University

Geo-computation Center for Social Sciences

Wuhan University

China Data Institute

Future Data Lab