# Towards sustainable mobility: prediction, modeling and assessment

Ganmin Yin[1,2]

[1]Singapore-MIT Alliance for Research and Technology (SMART), MIT

[2]Institute of Remote Sensing and Geographic Information System, Peking University

Dec 10, 2025

# Contents

# Introduction

# About me

**Ganmin Yin**

**Email:** ganmin@mit.edu

**Web:** www.sunshineyin.site

## *Education*

2025.09 – present: **Postdoc** at SMART Centre, MIT

2020.09 – 2025.07: **Ph.D.** in GIScience, Peking University

2016.09 – 2020.07: **B.S.** in GIScience, Peking University

## *Research interests*

- **Urban mobility & Human dynamics**
- **Urban growth & Urban economics**

} Research topics

- Big data analytics & Geospatial AI
- Social sensing & Remote sensing

} Methods & data

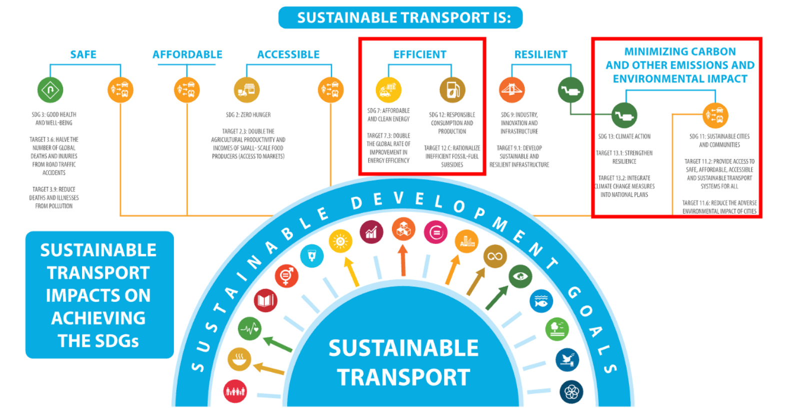
- Sustainable development goals (SDGs)

→ Research goals

# Background

***"Sustainable transport drives sustainable development"***

*— Ban Ki-moon, 2016*




The United Nations Secretary General's High-Level Advisory Group on Sustainable Transport, October 2016

# From small data to geospatial big data

## Small data

- Travel log, survey, census, etc.
- [Pros] demographics, travel mode/purpose, etc.
- [Cons] coverage, resolution

## Geospatial big data

- Remote sensing, social sensing
- Wide coverage, high resolution, more information, etc.
- Brought a new paradigm to geography and urban studies

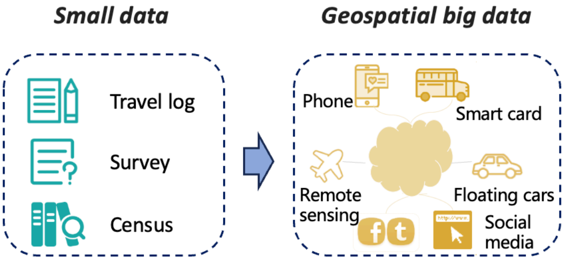



Figure: From small data to geospatial big data

# The rise of data-driven research paradigm

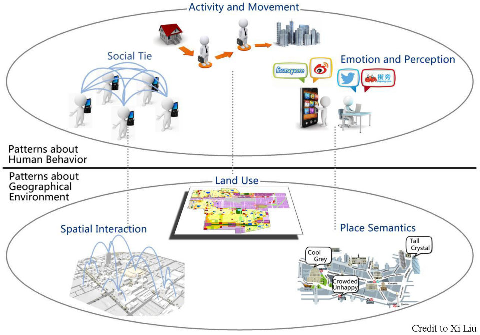



Figure: Social sensing framework for geography and urban studies [Liu et al., 2015]

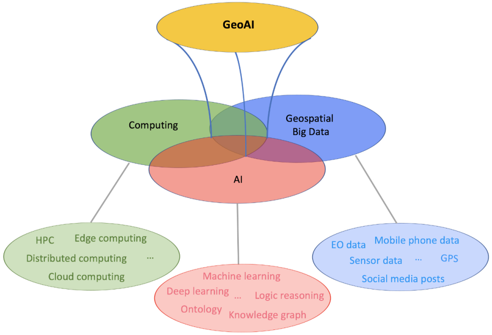
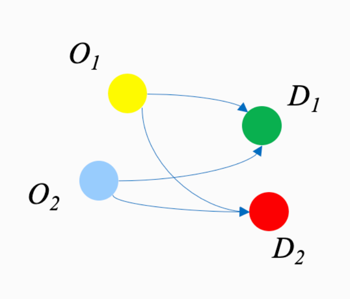
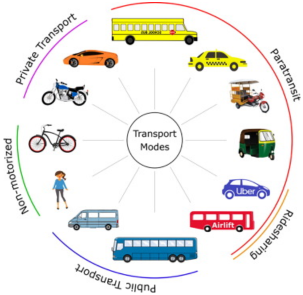



Figure: A three-pillar view of GeoAI: Geospatial big data, computing and AI [Li, 2020]

# Three major topics: prediction, modeling and assessment




**Travel demand prediction**




**Mode choice modeling**

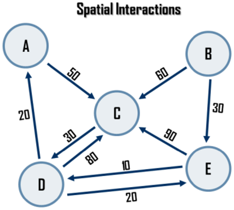

**Mode shift assessment**

# Topic 1: Travel demand prediction

# Topic 1: Travel demand prediction

**[Problem] How to estimate travel demand between two locations?**

- Commuting flow prediction
- Subway-bikesharing integration prediction




O/D Matrix

| | A | B | C | D | E | Ti |
|---|---|---|---|---|---|---|
| A | 0 | 0 | 50 | 0 | 0 | 50 |
| B | 0 | 0 | 60 | 0 | 30 | 90 |
| C | 0 | 0 | 0 | 30 | 0 | 30 |
| D | 20 | 0 | 80 | 0 | 20 | 120 |
| E | 0 | 0 | 90 | 10 | 0 | 100 |
| Tj | 20 | 0 | 280 | 40 | 50 | **390** |

Figure: Spatial interaction and OD travel demand prediction


[1] **Yin, G.**, Huang, Z.*, Bao, Y., Wang, H., Li, L., Ma, X., & Zhang, Y. (2023). ConvGCN-RF: A hybrid learning model for commuting flow prediction considering geographical semantics and neighborhood effects. ***GeoInformatica***, 27(2), 137-157. **(IF=2.6)**

[2] **Yin, G.**, Fu, C., Ren, S., Yan, X., Qi, J., Bao, Y., & Huang, Z.* (2025). Traffic prediction and road space optimization for the integration of dockless bike-sharing and subway. ***Sustainable Cities and Society***, 121, 106162. **(IF=12.0)**

# Topic 1.1: Commuting flow prediction

## Spatial interaction models

- Gravity model [Zipf, 1946]
- Radiation model [Simini et al., 2012]
- [Cons] Limited inputs, low accuracy

## Machine learning models

- Tree models: Random forest, XGBoost
- Neural networks: DeepGravity [Simini et al., 2021], SIGCN [Yao et al., 2020]
- [Cons] Spatial interaction, geographical semantics, spatial proximity, etc.

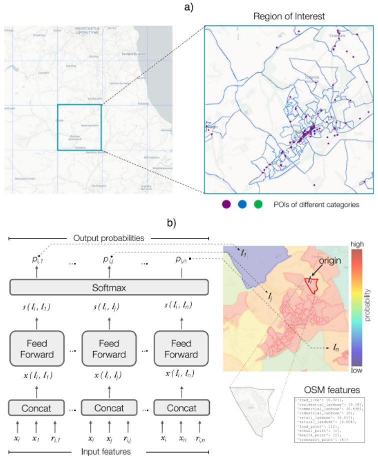



Figure: DeepGravity [Simini et al., 2021]

# Topic 1.1: Commuting flow prediction

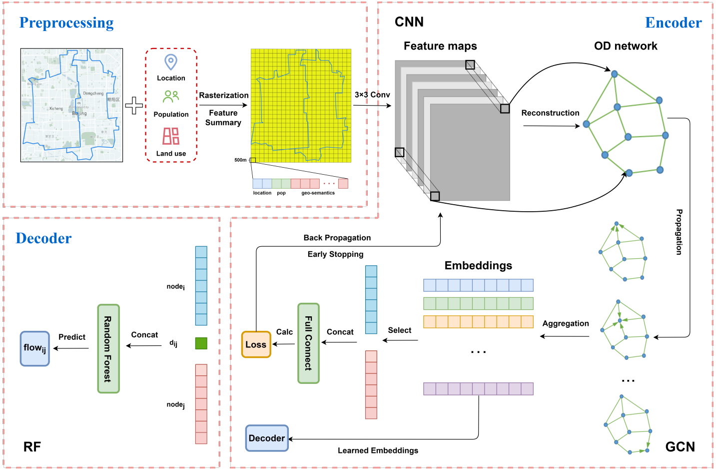



Figure: The three-layer framework: (1) Preprocessing, (2) Encoder, (3) Decoder.

# Topic 1.1: Commuting flow prediction

## [Geospatial data vs image data]

- Geographic region $\leftrightarrow$ image space
- Grid cell $\leftrightarrow$ pixel
- Grid features $\leftrightarrow$ channel
- Buffer $\leftrightarrow$ receptive field

## [Spatial proximity modeling]

$$Y_{i,j} = \sigma\left(\sum_{h=-p}^{p}\sum_{w=-p}^{p}\sum_{c=1}^{C} X_{i+h,j+w,c} \cdot K_{h+p,w+p,c} + b\right)$$

$\mathbf{Y}$: Output fusing nearby features

$\mathbf{X}$: Input 3D feature map

$\mathbf{K}$: Convolutional kernel

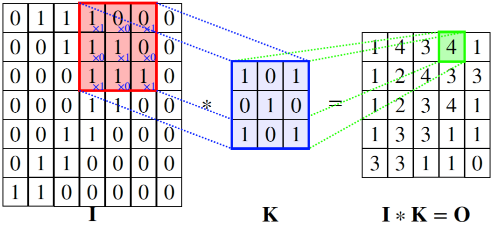



Figure: Convolution considers nearby features

# Topic 1.1: Commuting flow prediction

Table: Comparison of different models on a Beijing dataset (~50k ODs)

| Model | RMSE | MAPE | CPC |
|---|---|---|---|
| Gravity model | 16.064 | 0.751 | 0.707 |
| Radiation model | 19.048 | 0.933 | 0.337 |
| Random forest | 13.541 | 0.617 | 0.758 |
| XGBoost | 12.443 | 0.604 | 0.756 |
| **Node2vec-RF (ours)** | **11.087** | **0.550** | **0.771** |
| **GCN-RF (ours)** | **9.871** | **0.452** | **0.810** |
| **ConvGCN-RF (ours)** | **9.553** | **0.439** | **0.817** |
| Improved | $\downarrow$ 23.2% | $\downarrow$ 27.3% | $\uparrow$ 8.1% |

# Topic 1.1: Commuting flow prediction

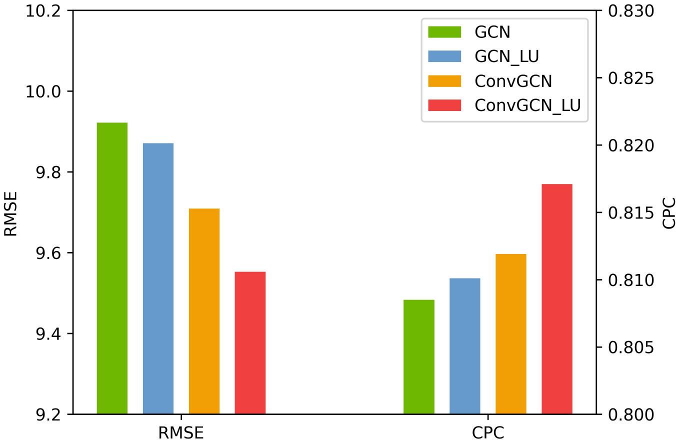



Figure: Stacking effects of land use and convolution: feature fusion and spatial discretization.

# Topic 1.2: Subway-bikesharing integration prediction

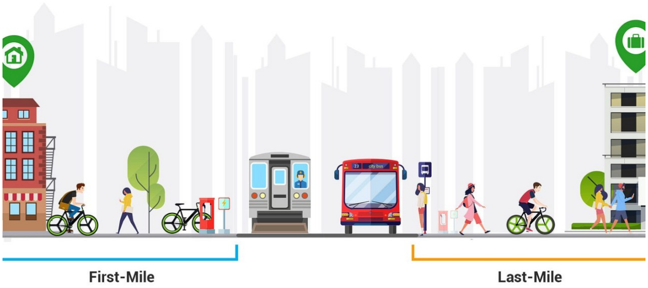



Figure: First/last mile connectivity in public transportation.

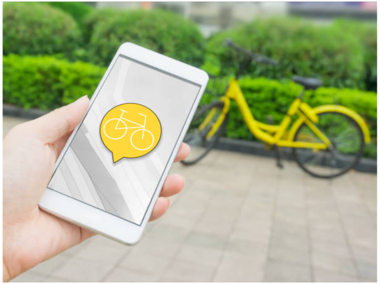


Figure: Shared bikes as a first/last mile solution.

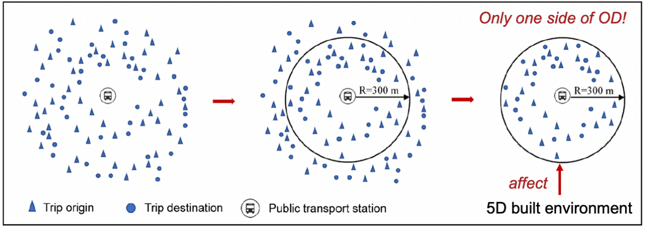
# Topic 1.2: Subway-bikesharing integration prediction

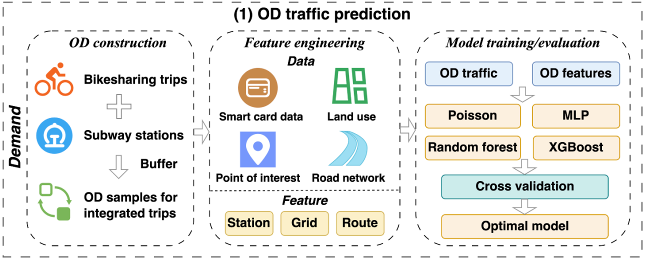
## Previous studies

- Definition of integration
- Buffer analysis
- 5D built environment
- Influencing factors analysis
- Only one side of OD

## Our study

- OD construction
- Multisource geo-data
- OD integration prediction
- More granular OD



# Topic 1.2: Subway-bikesharing integration prediction

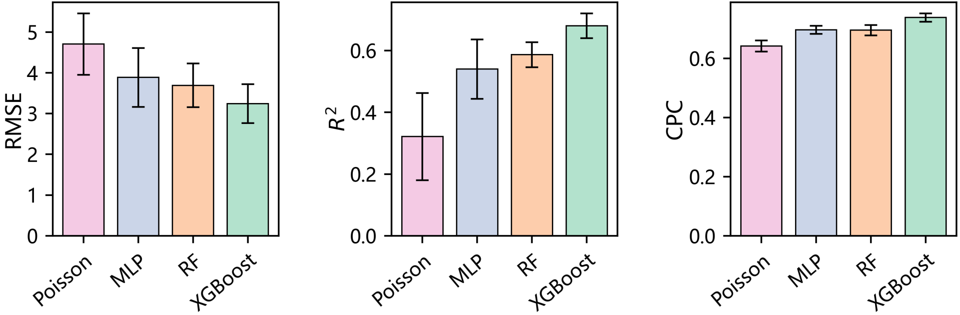



Figure: Comparison of different models on integration flow prediction.

# Topic 1.2: Subway-bikesharing integration prediction


Figure: Driving in Non-Motorised Vehicle lanes to attract Rs 2000 fine, New Delhi (TIMES, 2016)


Figure: Bike Lanes are NOT for Vehicles, Vancouver (Source: https://bikehub.ca/)

# Topic 1.2: Subway-bikesharing integration prediction

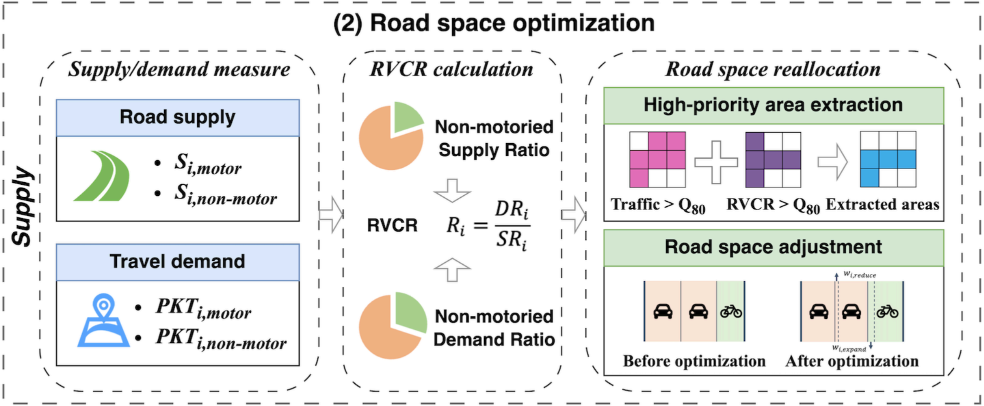



Figure: Reducing the width of motor lanes reasonably to increase non-motor road space.

# Topic 1.2: Subway-bikesharing integration prediction

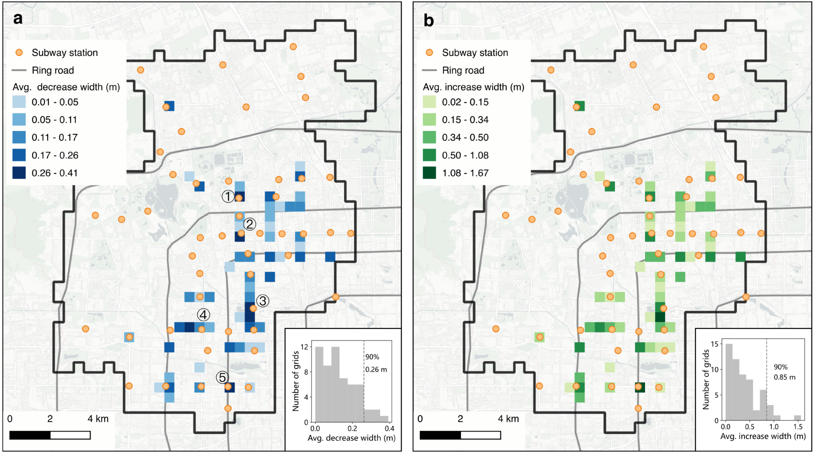



Figure: After optimization, the width of motor lanes mostly meets the safety standard.

# Topic 2: Mode choice modeling

# Topic 2: Mode choice modeling

**[Problem] Why do people choose certain modes of transportation?**

- Public transit share extraction
- Public transit mode choice modeling
- Active mobility mode choice modeling

Figure: Various modes of transportation


[1] **Yin, G.**, Huang, Z.*, Yang, L., ..., & Liu, Y. (2023). How to quantify the travel ratio of urban public transport at a high spatial resolution? A novel computational framework with geospatial big data. ***International Journal of Applied Earth Observation and Geoinformation***, 118, 103245. **(IF=8.6)**

[2] **Yin, G.**, Huang, Z.*, Lu, L., Qi, J., Liu, Y., Yan, X., Ren, S., & Bao, Y. An improved Beta-binomial model for public transit mode share: Incorporating nonlinear and interaction effects. ***Under Review***.

[3] **Yin, G.**, Huang, Z.*, Fu, C., Ren, S., Bao, Y., & Ma, X. (2024). Examining active travel behavior through explainable machine learning: Insights from Beijing, China. ***Transportation Research Part D: Transport and Environment***, 127, 104038. **(IF=7.7)**

# Topic 2.1: Public transit share extraction

**[Public transit share]**

$$\text{Public transit share} = \frac{\text{Public transit flow}}{\text{Total flow}} \quad \text{for each OD pair}$$

**[Smart card data]**

- Individual-level public transit usage
- Widely used in metro and bus systems
- Open system vs closed system
- [Station to station ODs]

**[Mobile phone location data]**

- Individual-level trajectory data
- Wide coverage, high resolution
- GPS signals vs cellular signals
- [Location to location ODs (grid)]

# Topic 2.1: Public transit share extraction

**Step 1: Route planning**

**Input:** Grid pairs with total flow, $GP$

**Output:** Planning routes for each grid pair, $gp.routes$ for all $gp \in GP$

**for** each $gp \in GP$ **do**<br>
    $R \leftarrow$ the fastest 3 planning routes of $gp$<br>
    $r' \leftarrow$ the fastest planning route of $gp$<br>
    **for** each $r \in R$ **do**<br>
        **if** $r.dura > r'.dura + 15$ **then**<br>
            remove $r$ from $R$<br>
        **end if**<br>
    **end for**<br>
    $gp.routes \leftarrow R$<br>
**end for**

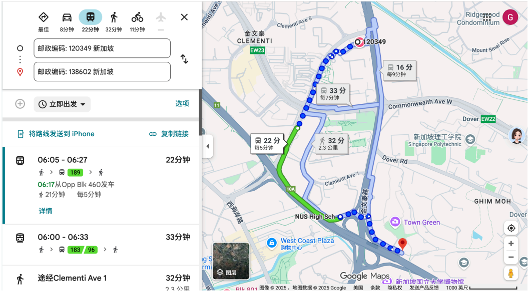



Figure: Route planning function in Google Maps

# Topic 2.1: Public transit share extraction

**Step 2: Route matching**

**Input:** Planning routes $gp.routes$, PT routes $PT$
**Output:** Potential source grid pairs of PT routes, $pt.gp$ for all $pt \in PT$

**for** each $pt \in PT$ **do**
  **for** each $gp \in GP$ **do**
    **for** each $r \in gp.routes$ **do**
      **if** $r.ori = pt.ori$ and $r.des = pt.des$ **then**
        add $gp$ to $pt.gp$
      **end if**
    **end for**
  **end for**
**end for**

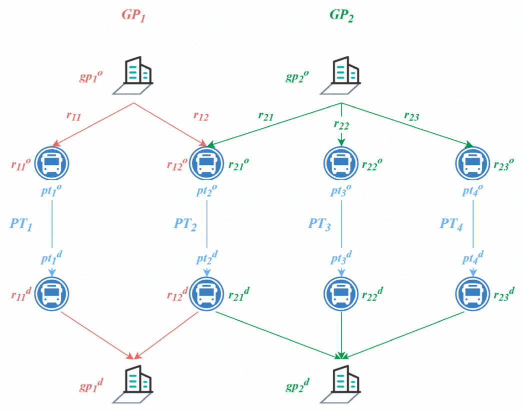



Figure: Matching the planned routes with PT routes

# Topic 2.1: Public transit share extraction

**Step 3: Flow assignment**

**Input:** PT routes $PT$, grid pairs $GP$

**Output:** Public transit share $PT\_Share$

```
PT_Share ← ∅
for each pt ∈ PT do
    assign pt.flow to pt.gp
end for
for each gp ∈ GP do
    gp.pt_share ← gp.pt_flow/gp.total_flow
    PT_Share ← PT_Share ∪ {gp.pt_share}
end for
return PT_Share
```

$$\min \quad \sum_{i=1}^{H}\sum_{j=1}^{G} dura_{ij} \cdot pt_{ij} \tag{1}$$

$$\text{s.t.} \quad \sum_{j=1}^{G} pt_{ij} = f_i, \quad \forall i \tag{2}$$

$$\sum_{i=1}^{H} pt_{ij} \leq t_j, \quad \forall j \tag{3}$$

$$dura_{ij} = \begin{cases} T(r_{jk}), & \text{if } PT_i \text{ matches } r_{jk} \\ +\infty, & \text{otherwise} \end{cases} \tag{4}$$

$$pt_{ij} \in [0, \min(f_i, t_j)], \ \text{integer} \tag{5}$$

# Topic 2.2: Public transit mode choice modeling

**[Bernoulli trial]**

Whether an individual chooses public transportation follows a Bernoulli distribution.

**Definition:**

- $y$: whether to choose public transportation (0 or 1)
- $p$: the probability

$$y \sim \text{Bernoulli}(p)$$

**[Binomial distribution]**

The traffic of public transportation among a group (under independence) follows a Binomial distribution.

**Definition:**

- $y$: the PT traffic
- $n$: the total traffic
- $p$: the probability

$$y \sim \text{Binomial}(n, p)$$

# Topic 2.2: Public transit mode choice modeling

**[Binomial model]**

The PT traffic of the i-th OD

The total traffic and PT share of the i-th OD

$$\boxed{Y_i} \sim \text{Binomial}(\boxed{n_i, p_i})$$

$$\boxed{\text{logit}(p_i)} = \log\left(\frac{p_i}{1-p_i}\right)$$

Link function in the Binomial model

$$= \boxed{\beta_0 + \sum_{j=1}^{m} \beta_j x_{ij}}$$

Linear combination of influencing factors

where $\beta_0$ is the intercept, $\beta_j$ are coefficients, and $x_{ij}$ are independent variables

# Topic 2.2: Public transit mode choice modeling

**[Two limitations]**

### 1.Independence Assumption Violation

Travelers within the same OD pair are correlated:

- Socio-demographic characteristics
- Travel preferences

### 2.Linear Relationship Limitation

As a type of GLMs, it can only model simple linear relationships:

- Ignores potential nonlinear effects
- Misses interaction effects

### Consequence

- Residual variance is larger than expected
- Leads to over-dispersion problem

### Examples

- Distance's nonlinear relationship
- Cost advantage's interaction effect with distance on PT share

# Topic 2.2: Public transit mode choice modeling

**[Improved Beta-Binomial model]**

Beta-Binomial model with beta prior distribution

Dispersion param. of beta distribution

$$Y_i \sim \text{Beta-Binomial}(n_i, p_i, \phi_i)$$

when $\phi_i \to \infty$, it degenerates into the Binomial model

$$\begin{aligned}\text{logit}(p_i) &= \log\left(\frac{p_i}{1-p_i}\right) \\ &= \beta_0 + \sum_{j\in\mathcal{S}} s_j(x_{ij}) + \sum_{j\in\mathcal{L}} \beta_j x_{ij} + \sum_{(j,k)\in\mathcal{I}} \beta_{jk} x_{ij} x_{ik}\end{aligned}$$

Link function

Non-linear term (penalized B-spline)

Linear term

Interaction term

when $\mathcal{S}, \mathcal{L}, \mathcal{I}$ denote the sets of smooth/non-linear, linear, and interaction terms

# Topic 2.2: Public transit mode choice modeling

**[Results]**

Table: Comparison of goodness-of-fit indicators

| Model | DEV | AIC | BIC |
|---|---|---|---|
| Binomial model | 478625.2 | 478651.2 | 478777.3 |
| **Beta-binomial model** | **454740.0** | **454768.0** | **454903.8** |
| **Nonlinear model** | **445583.8** | **445657.6** | **446015.6** |
| **Interaction model** | **445403.3** | **445483.5** | **445872.2** |
| Improved | ↓ 6.9% | ↓ 6.9% | ↓ 6.9% |

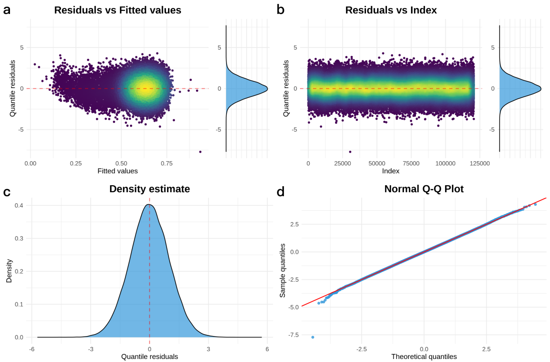



Figure: Residual diagnostics of the optimal model.

# Topic 2.2: Public transit mode choice modeling

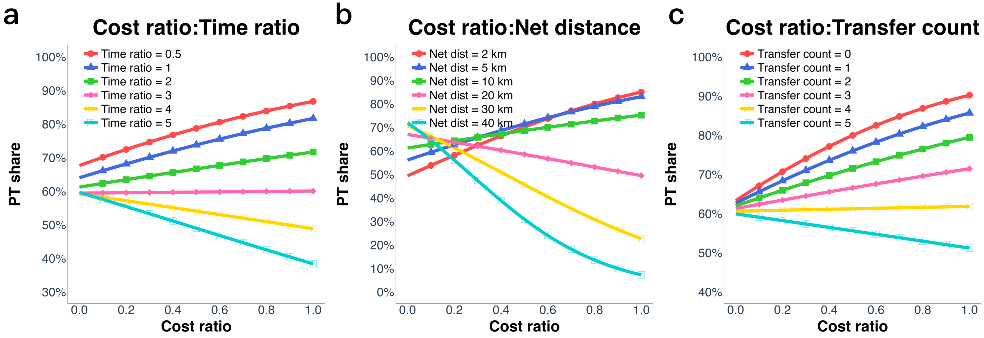



Figure: Interaction effects of transit-to-taxi cost ratio reveal a cost-based compensatory mechanism, challenging the price elasticity theory (monotonic negative relationship between price and demand).

# Topic 2.3: Active mobility mode choice modeling

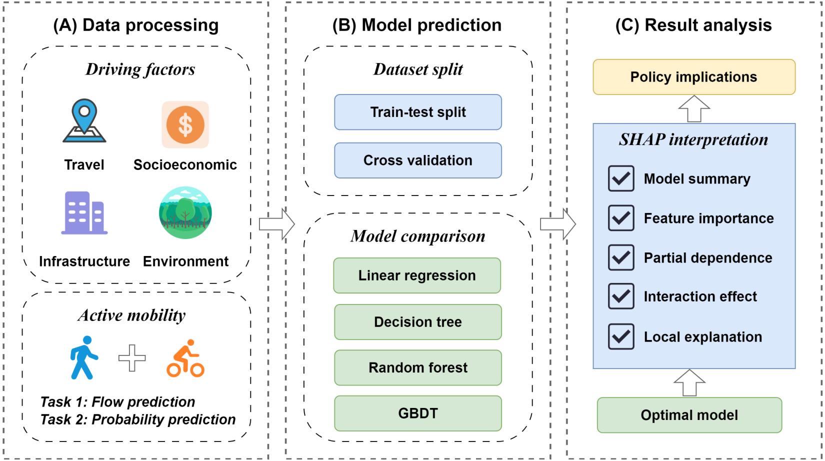



Figure: The modeling framework: (a) Data processing, (b) Model prediction, (c) Result analysis.

# Topic 2.3: Active mobility mode choice modeling

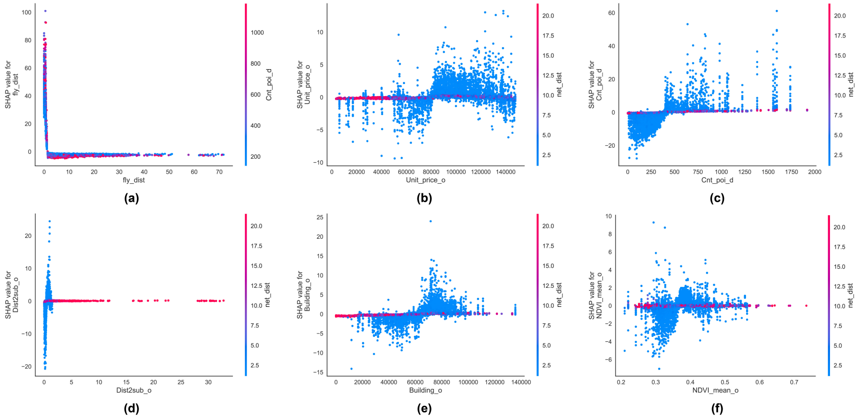



Figure: SHAP dependence plots of six representative features for flow prediction.

# Topic 2.3: Active mobility mode choice modeling

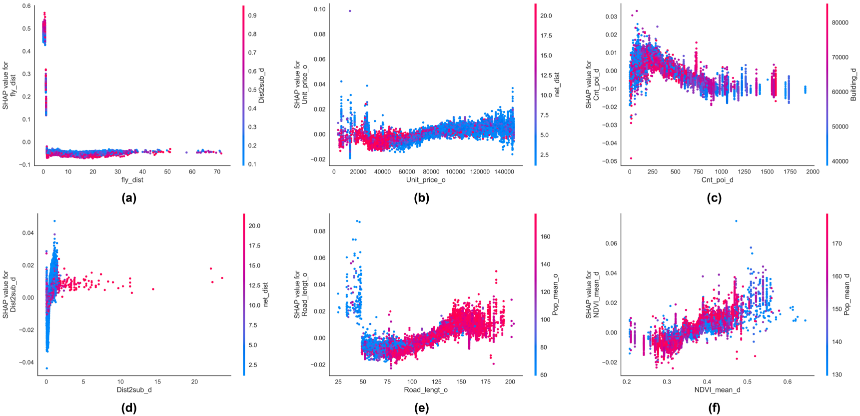



Figure: SHAP dependence plots of six representative features for probability prediction.

# Topic 2.3: Active mobility mode choice modeling

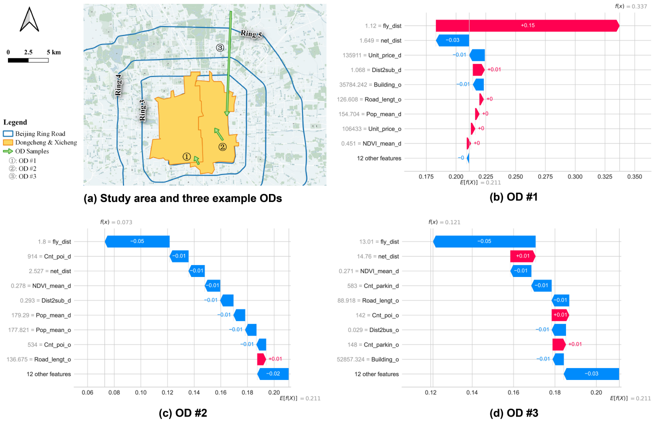



Figure: Explanation of three example ODs based on local SHAP analysis.

# Topic 3: Mode shift assessment

# Topic 3: Mode shift assessment

**[Problem] What impacts will travel mode shift bring?**

- Carbon reductions of bike-sharing
- Travel mode optimization for sustainability

Figure: Travel mode shift for sustainable mobility

[1] **Yin, G.**, Huang, Z.*, Wang, X., Tang, M., Ren, S., & Bao, Y. (2026). Unveiling the overestimated carbon reductions of dockless bike-sharing: A data-driven analysis. ***Transportation Research Part D: Transport and Environment***, 150, 105071. **(IF=7.7)**

[2] **Yin, G.**, Huang, Z.*, Ren, S., Tang, M., Yan, X., Zheng, J., Qi, J., Bao, Y. & Ma, X. (2026). Balancing efficiency and emissions through travel mode shifts: A multi-objective analysis from Beijing, China. ***Transportation Research Part A: Policy and Practice***, 204, 104782. **(IF=6.8)**

# Topic 3.1: Carbon reduction of bike-sharing

**[Assumption]**

- Emerging mobility (e.g., bike-sharing, EVs) greatly reshapes traffic systems
- Environmental benefits are attributed to the substitution of travel modes

**[For example]**

- If bike-sharing replaces private car trips, carbon emissions are reduced
- If bike-sharing replaces walking, no carbon reductions are observed

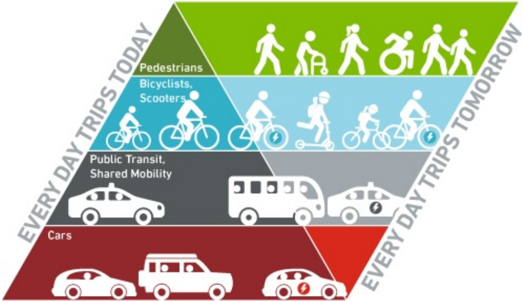



Figure: Wish a better travel mode structure for future.

# Topic 3.1: Carbon reduction of bike-sharing

**[Carbon reduction calculation model]**

$$\underbrace{CR}_{\text{Carbon reductions for a bike-sharing trip}} = \underbrace{f_{\text{substitute}}}_{\text{Emission factor of the substituted travel modes}} \times d \times l$$

$$f_{\text{substitute}} = \sum_i \underbrace{p_i}_{\text{Proportion of the } i\text{-th travel mode (from local authorities)}} \times \underbrace{f_i}_{\text{Emission factor of the } i\text{-th mode}}$$

where $d$ is the trip distance and $l$ is the load factor (set to 1 for bike-sharing).

# Topic 3.1: Carbon reduction of bike-sharing

**[Two limitations]**

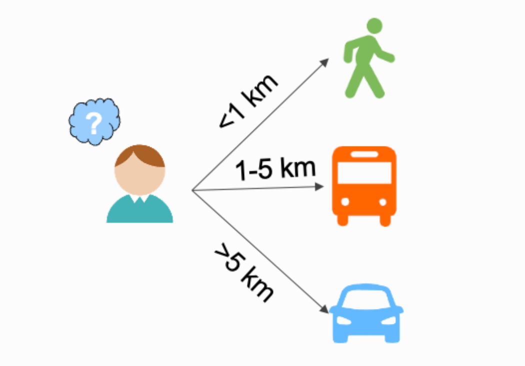



Figure: Travel mode share varies with distance.

Figure: Bike-sharing rebalancing increases unignorable operational emissions.

# Topic 3.1: Carbon reduction of bike-sharing

**[Improved calculation model]**

Emission factor of substituted travel modes

Emission factor of a shared bike using LCA

$$CR = (\boxed{f_{\text{substitute}}} - \boxed{f_{\text{bikesharing}}}) \times d \times l$$

$$f_{\text{substitute}} = \sum_i \boxed{p_i} \times \boxed{f_i}$$

Distance-sensitive proportions using Monte Carlo simulation

Emission factor of the $i$-th mode

where $d$ is the trip distance and $l$ is the load factor (set to 1 for bike-sharing).

# Topic 3.1: Carbon reduction of bike-sharing

Table: Multi-scenarios to quantify overestimation: S1 is previous studies, S4 is our study.

| Scenario | Distance | Lifecycle emissions |
|---|---|---|
| S1 | × | × |
| S2 | ✓ | × |
| S3 | × | ✓ |
| S4 | ✓ | ✓ |

**Previous studies would have overestimated by about 44.9%!**

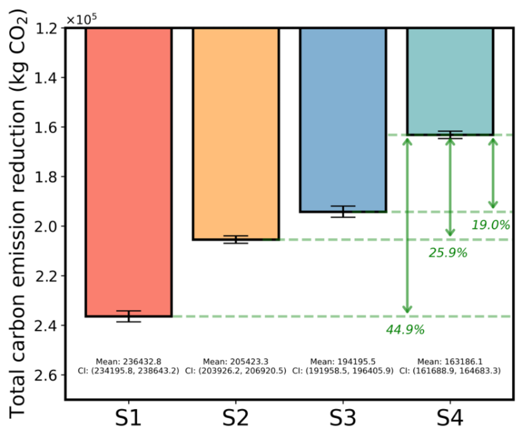



Figure: Overestimation using S4 as the baseline (Shenzhen, ∼1.42 million trips).

# Topic 3.1: Carbon reduction of bike-sharing

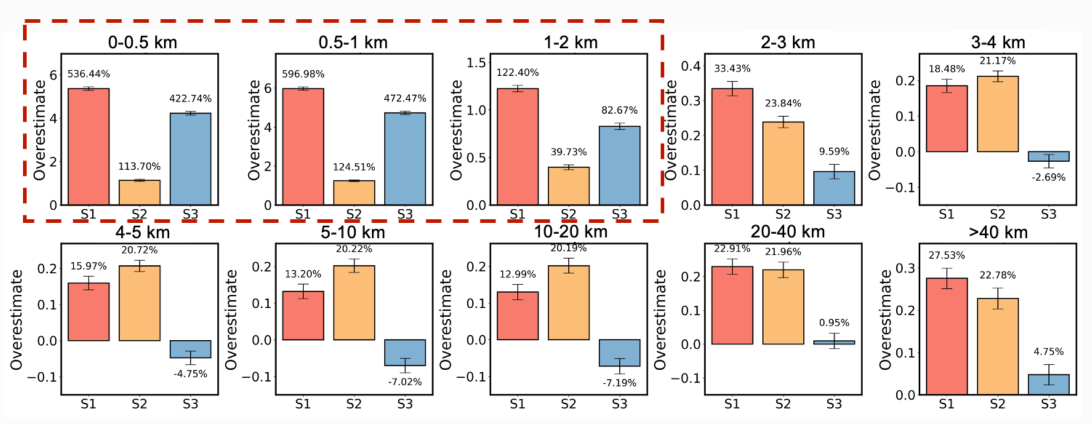
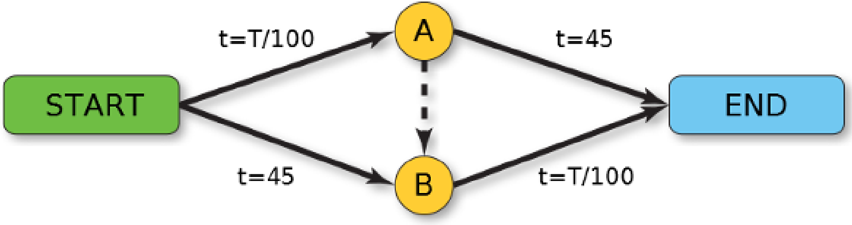



Figure: Trips within 2km are overestimated by more than 100%.

# Topic 3.2: Travel mode optimization for sustainability

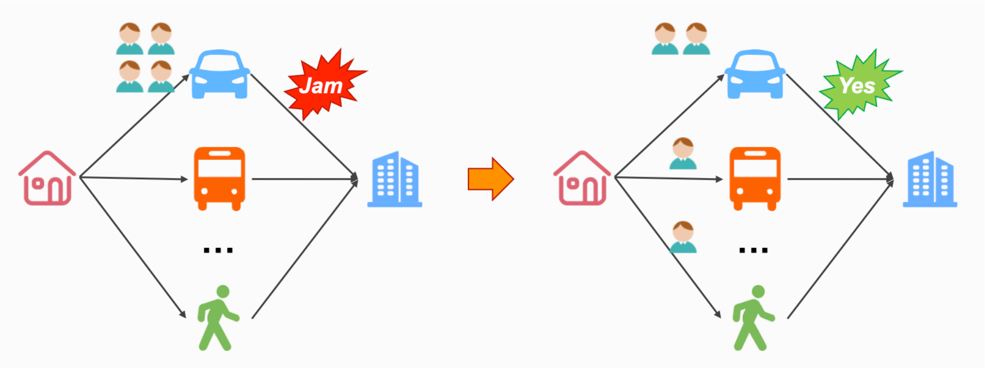
**[Braess's paradox]**




If 4000 cars need to take this route, what will happen when the super fast route (dashed line) is added?

# Topic 3.2: Travel mode optimization for sustainability

**[Maybe behavioral adjustment?]**




**Will travel mode shift both improve efficiency and reduce emissions?**

# Topic 3.2: Travel mode optimization for sustainability

**[Mathematical expression]**

Objective 1: total travel time of a given OD

Decision variable: $k$-th mode's traffic

$$\text{Minimize:} \quad \boxed{F_1(\mathbf{x})} = \sum_{k=1}^{m} \boxed{x_k} \cdot t_k$$

Objective 2: total carbon emissions of the OD

$$\text{Minimize:} \quad \boxed{F_2(\mathbf{x})} = \sum_{k=1}^{m} x_k \cdot d_{k,\text{motor}} \cdot f_k$$

$$\text{Subject to:} \quad \sum_{k=1}^{m} x_k = \boxed{N},$$

Constraint: Upper / lower bound

Constraint: Total traffic as $N$

$$\boxed{0 \leq x_k \leq N}, \quad k = 1, 2, \ldots, m$$

$$\boxed{x_k \in \mathbb{Z}^+}, \quad k = 1, 2, \ldots, m$$

Constraint: Integer

$\mathbf{x} = (x_1, x_2, \ldots, x_m)$: decision variable, $t_k$:travel time, $d_{k,\text{motor}}$: motor dist., $f_k$:emission factor of mode $k$.

# Topic 3.2: Travel mode optimization for sustainability

**[Multi-scenario settings]**

(1)

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathcal{P}} F_1(\mathbf{x})$$

$$\text{s.t. } F_2(\mathbf{x}) \le F_2(\mathbf{x}^r)$$

$$F_2(\mathbf{x}^r) = \sum_{k=1}^{m} x_k^r \cdot d_{k,\text{motor}} \cdot f_k$$

**S-I: Efficiency priority**

(2)

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathcal{P}} F_2(\mathbf{x})$$

$$\text{s.t. } F_1(\mathbf{x}) \le F_1(\mathbf{x}^r)$$

$$F_1(\mathbf{x}^r) = \sum_{k=1}^{m} x_k^r \cdot t_k$$

**S-II: Emission priority**

(3)

$$nF_j(\mathbf{x}) = \frac{F_j(\mathbf{x}) - \min_{\mathbf{x}\in\mathcal{P}} F_j(\mathbf{x})}{\max_{\mathbf{x}\in\mathcal{P}} F_j(\mathbf{x}) - \min_{\mathbf{x}\in\mathcal{P}} F_j(\mathbf{x})}$$

$$j = 1, 2$$

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathcal{P}} \max\left(nF_1(\mathbf{x}),\ nF_2(\mathbf{x})\right)$$

**S-III: Balanced**

# Topic 3.2: Travel mode optimization for sustainability

**[Results: single OD]**

Table: Information of a randomly selected OD.

| Mode | Traffic | Duration $(hour)$ | Distance $(km)$ |
|---|---|---|---|
| Car | 11 | 0.569 | 11.556 |
| Bus | 4 | 1.384 | 10.464 |
| Subway | 5 | 0.826 | 8.252 |
| Active travel | 0 | 0.746 | 11.187 |

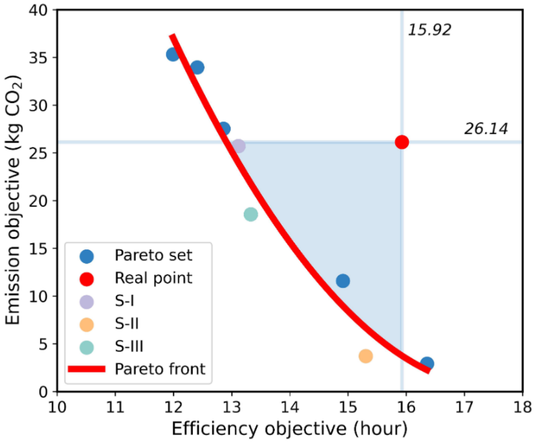


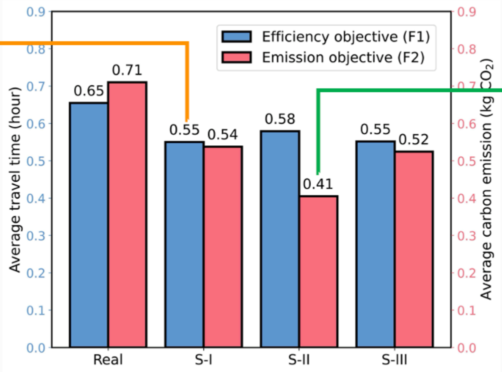

Figure: Objective function space after optimization.

# Topic 3.2: Travel mode optimization for sustainability

**[Results: experiment on Beijing]**

***Travel efficiency***

**1. Average travel time can be reduced by up to 6.3 mins/trip.**




***Carbon emissions***

**2. Carbon emissions can be reduced by up to 0.30 kgCO2-eq/trip.**

Beijing, ~0.86 million ODs

**Travel mode shift can both improve efficiency and reduce emissions!**

# Topic 3.2: Travel mode optimization for sustainability

**[A personal thought...]**

**Perhaps not infrastructure, but our behavior that really matters?**

# Conclusions

# Takeaways

## Travel demand prediction

- Developed a geospatial AI model for commuting flow prediction
- Designed a framework for subway-bikesharing integration prediction and optimization

## Mode choice modeling

- Developed a novel framework for public transit mode share extraction at the OD level
- Proposed an improved Beta-binomial model for public transit mode choice modeling
- Examined active travel behavior through explainable machine learning

## Mode shift assessment

- Uncovered the overestimated carbon reductions of dockless bike-sharing
- Proposed a multiobjective travel mode optimization framework for sustainability

# Welcome your submissions!

**Special Issue**

Advanced Information Systems: Data-Driven and Geospatial Approaches

**Guest Editors**

Dr. Yi Bao
Dr. Xiao Zhou
Dr. Ganmin Yin
et al.

**Deadline**

15 June 2026

*electronics*

IMPACT FACTOR 2.6

CITESCORE 6.1

# Thank you!

ganmin@mit.edu

https://www.sunshineyin.site/

# References I


Li, W. (2020).
Geoai: Where machine learning and big data converge in giscience.
*Journal of Spatial Information Science*, (20):71–77.

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., and Shi, L. (2015).
Social sensing: A new approach to understanding our socioeconomic environments.
*Annals of the Association of American Geographers*, 105(3):512–530.

Simini, F., Barlacchi, G., Luca, M., and Pappalardo, L. (2021).
A deep gravity model for mobility flows generation.
*Nature communications*, 12(1):6576.

# References II


Simini, F., González, M. C., Maritan, A., and Barabási, A.-L. (2012).
A universal model for mobility and migration patterns.
*Nature*, 484(7392):96–100.

Yao, X., Gao, Y., Zhu, D., Manley, E., Wang, J., and Liu, Y. (2020).
Spatial origin-destination flow imputation using graph convolutional networks.
*IEEE Transactions on Intelligent Transportation Systems*, 22(12):7474–7484.

Zipf, G. K. (1946).
The p 1 p 2/d hypothesis: on the intercity movement of persons.
*American sociological review*, 11(6):677–686.