

Urban Sensing and Computing

Big Data Analytic with Open Source Software

Jianghao Wang @ IGSNRR, CAS & MIT CFC Lab

2018-8-21 @ Prague GeoSTAT Summer School

URBAN COMPUTATION + OSGEO

1. INTRODUCTION & FRAMEWORK
2. BIG DATA IN URBAN STUDIES
3. GEOCOMPUTATION WITH OSGEO

1

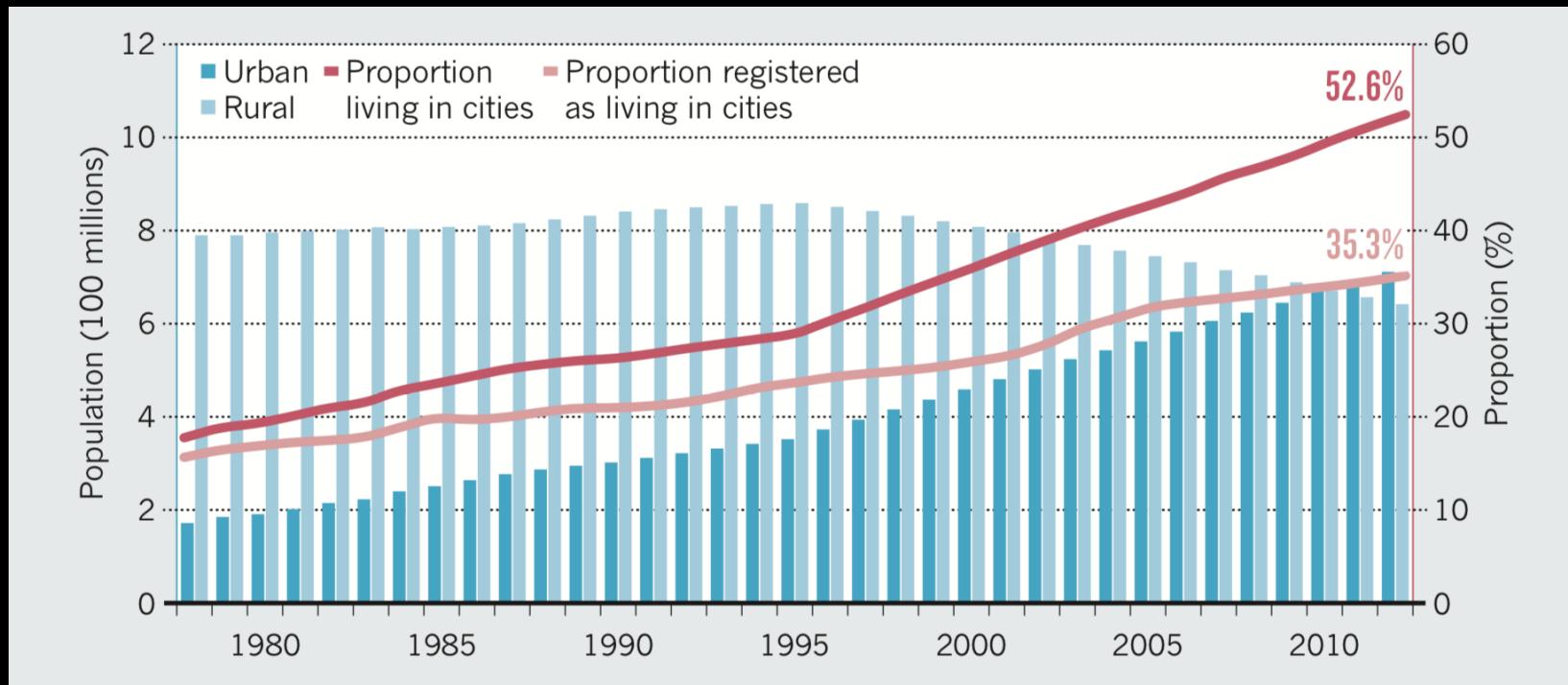


Introduction



Urbanization in China

In **1980**, only **20%** of Chinese lived in cities. However, the number reaches **56%** in **2017**, with **800 million** urban population; it's projected to reach **60%** by **2030**.



Problems with Rapid Urbanization

- Problem we faced: traffic congestion, air pollution; sustainable development
- City is a complex giant system; Smart / Intelligence management
- Big data, providing a new way to understanding our city.

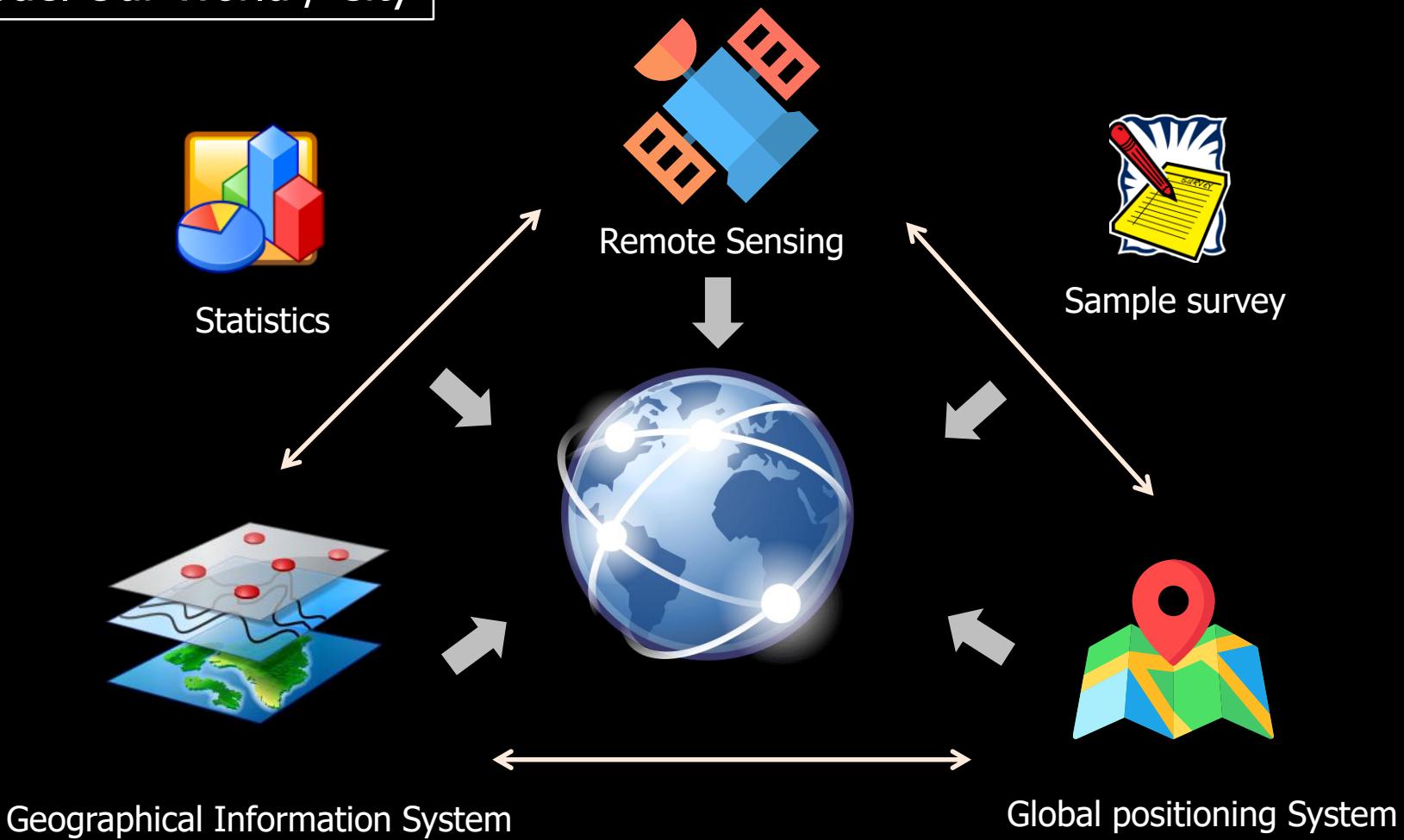


Problem we faced



2010 Shanghai Expo theme:
"Better City, Better Life."

Traditional Way to Model Our World / City

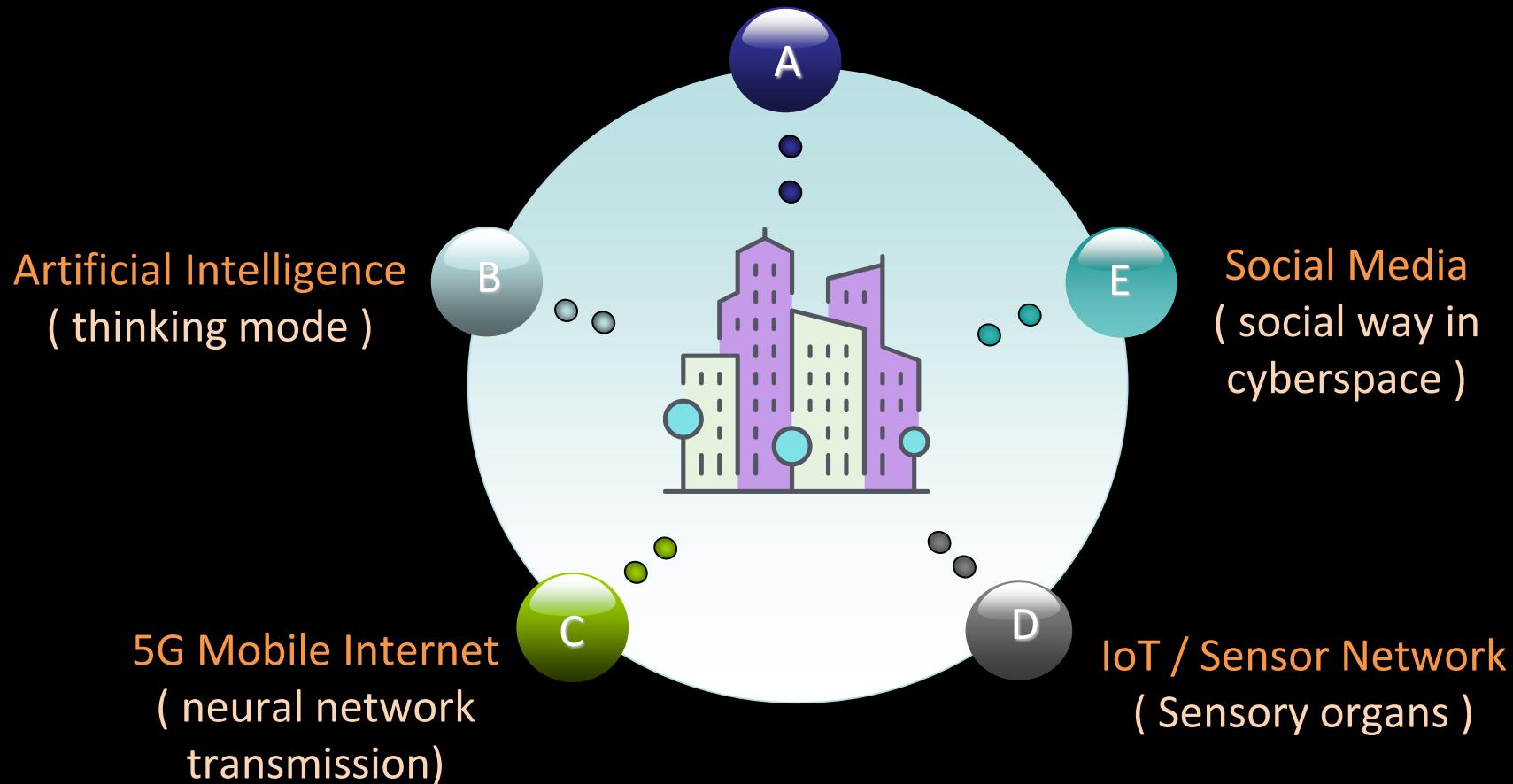


Senseable City



New Technologies to Understand Our City

Cloud Computing (Brain)





SUSTAINABLE DEVELOPMENT GOALS

1 NO POVERTY



2 ZERO HUNGER



3 GOOD HEALTH AND WELL-BEING



4 QUALITY EDUCATION



5 GENDER EQUALITY



6 CLEAN WATER AND SANITATION



7 AFFORDABLE AND CLEAN ENERGY



8 DECENT WORK AND ECONOMIC GROWTH



9 INDUSTRY, INNOVATION AND INFRASTRUCTURE



10 REDUCED INEQUALITIES



11 SUSTAINABLE CITIES AND COMMUNITIES



12 RESPONSIBLE CONSUMPTION AND PRODUCTION



13 CLIMATE ACTION



14 LIFE BELOW WATER



15 LIFE ON LAND



16 PEACE, JUSTICE AND STRONG INSTITUTIONS



17 PARTNERSHIPS FOR THE GOALS



SUSTAINABLE
DEVELOPMENT
GOALS

11 SUSTAINABLE CITIES AND COMMUNITIES



Make cities and human
settlements inclusive, safe,
resilient and sustainable

New Technologies in Urban Studies



Achieve a Sustainable City
Improve the Quality of Urban Life

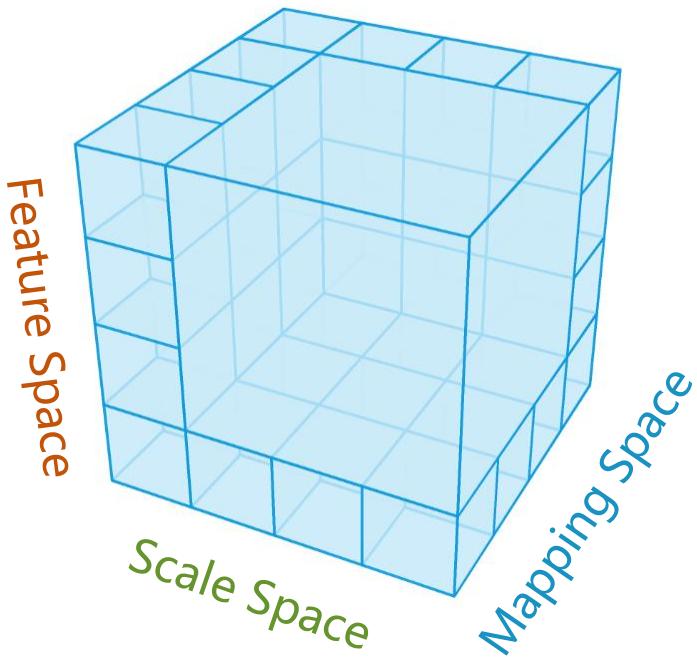
2



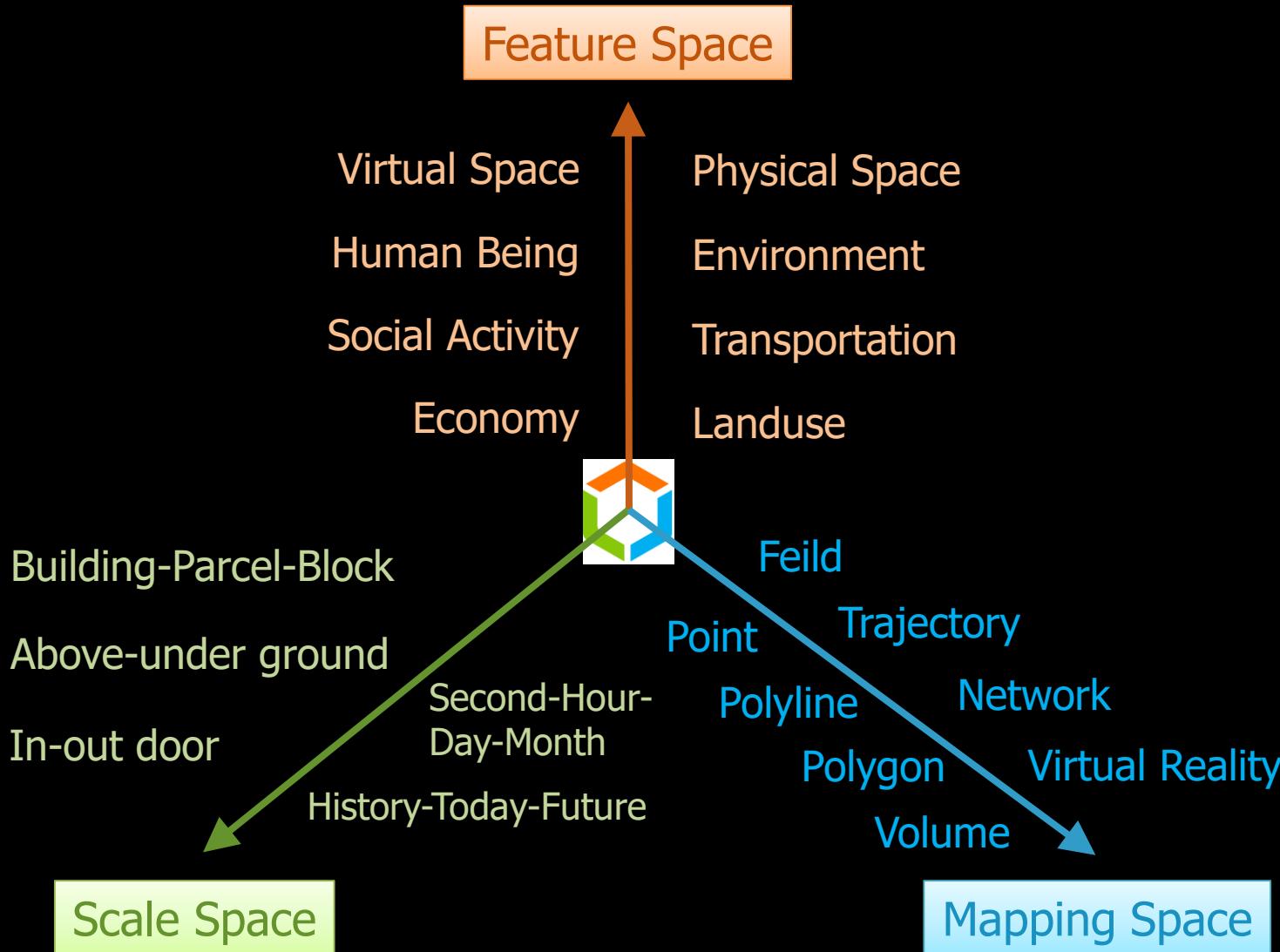
Framework



Urban Sensing and Computing Framework



Urban Sensing and Computing is crossover studies investigating the use of digital sensors (including remote sensing, internet of things, wireless sensor network, citizen-sensing, etc.) to obtain, compute, and analyze urban dynamic environments in order to achieve a more sustainable and efficient cities.



Physical Urban Space

Sensing urban physical structure with remote sensing techniques

1

Human Activity Space

Sensing human activities or behaviors with spatio-temporal big data

2



Urban Pan-space

Sensing urban social-economic with census, IoT, and electronic business

Social-economic Space

3

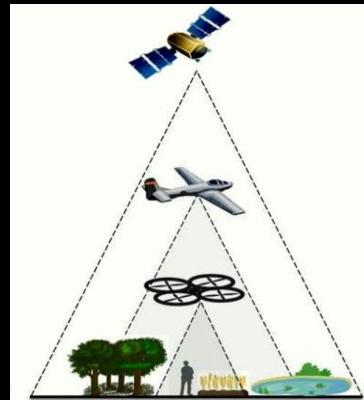
Sensing urban in cyberspace with social media and mobile network

4

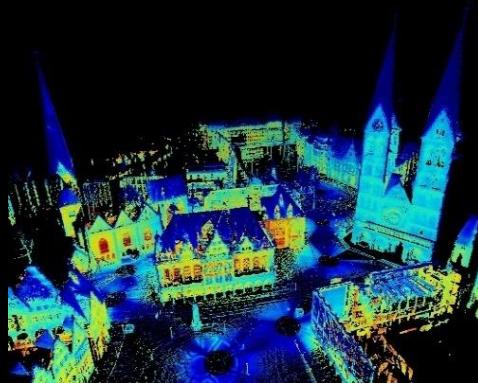
Cyber Urban Space



Pan-space Urban Sensing Technology



Remote sensing



Lidar photogrammetry



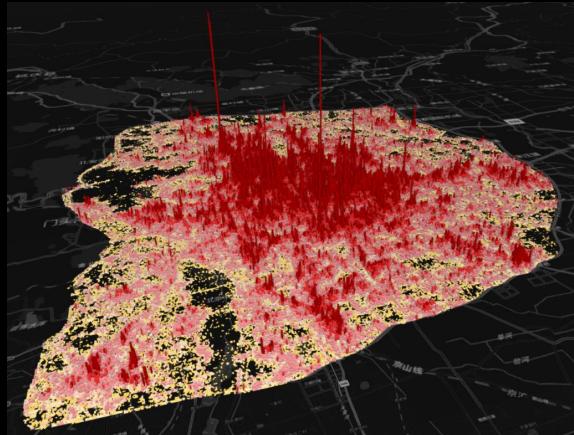
Street View mapping system

Sensing urban physical structure with (remote) sensing techniques

To get the basic information about our city, like land use/land cover, building, roads, greens, etc.



Pan-space Urban Sensing Technology



Human activities



Human Mobility

Sensing human activities or behaviors with spatio-temporal big data

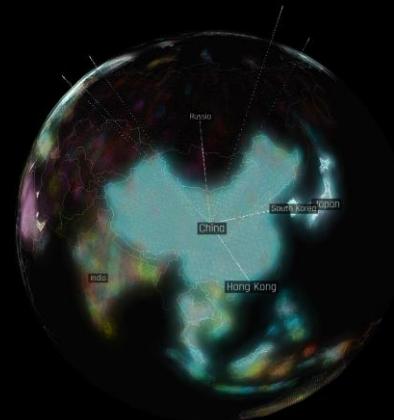
Understanding human dynamic distribution, mobility, and living behaviors with mobile phone or smart card data.



Pan-space Urban Sensing Technology



China economic census



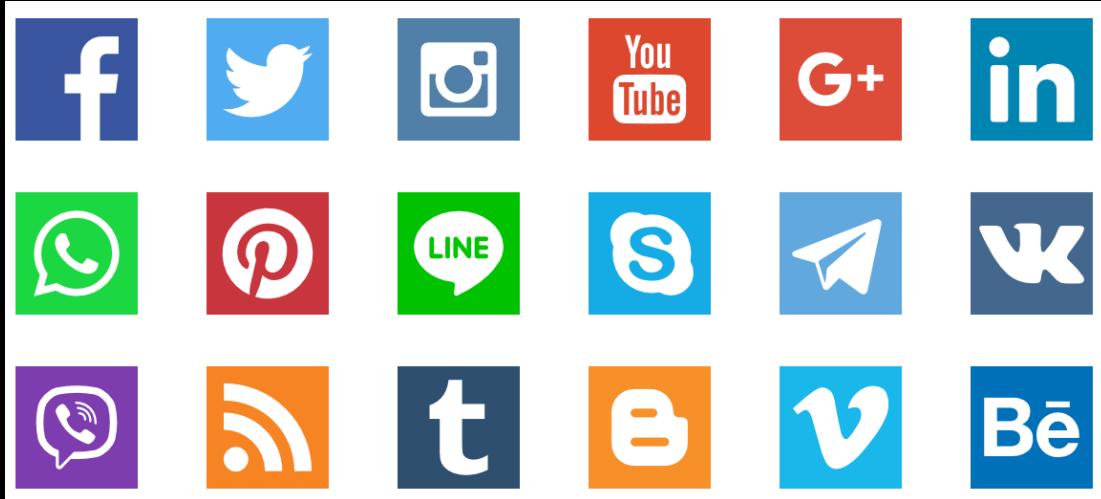
World economic trade with China

Sensing urban social-economic with IoT, and electronic business data

Using social-economic related geospatial big data to help better understanding our urban social and economic system.



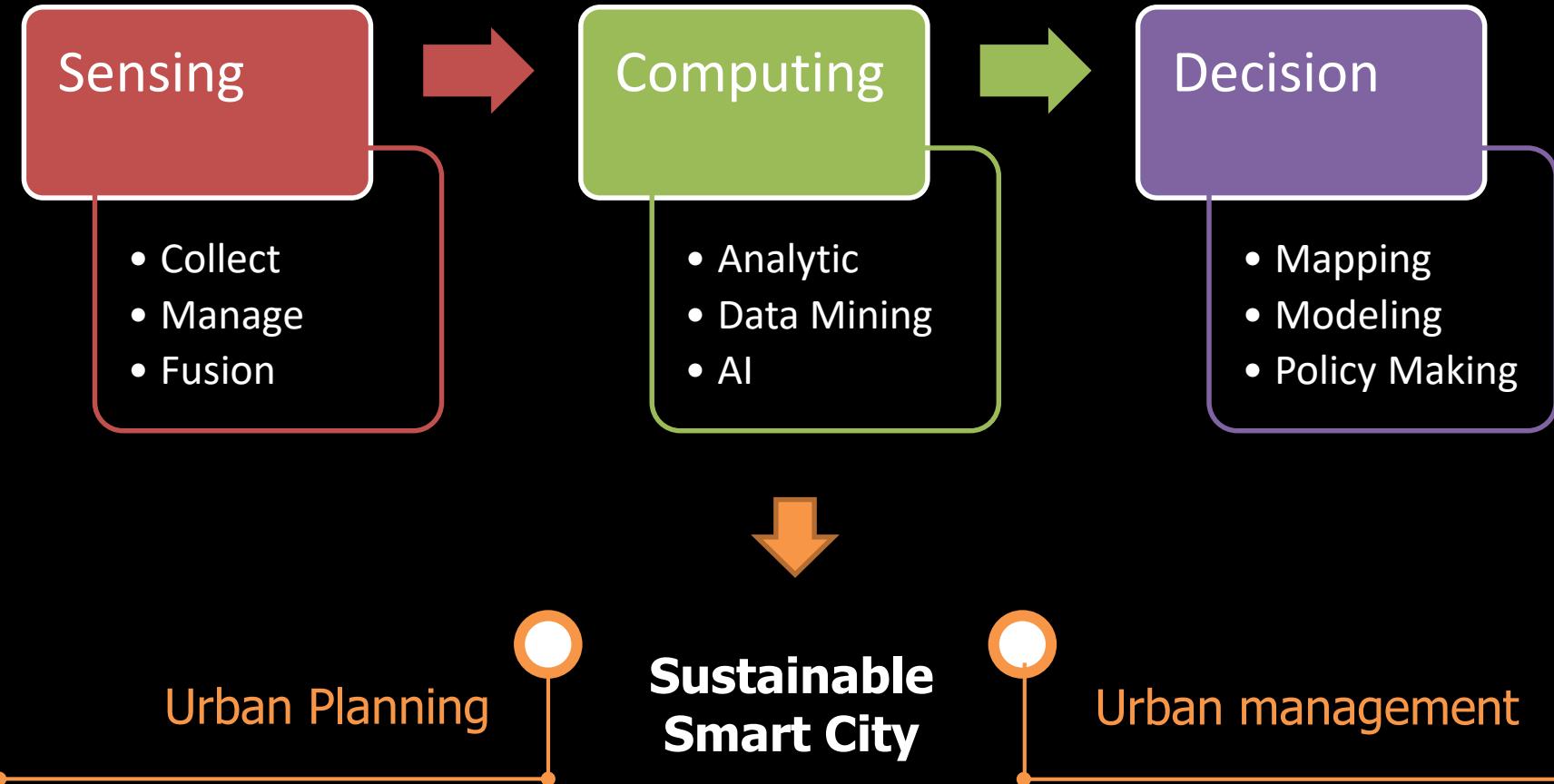
Pan-space Urban Sensing Technology



Sensing urban in cyberspace with social media and mobile network

Understanding citizen behavior or urban environment through online social media and study the relationship between physical and cyber urban space

Urban Data Science



3

Big Data in Urban Studies

Urban Studies with Big Data



Pan-Space Urban Research

Physical Urban Space

Modeling Dynamic Urban Structure

Human Activity Space

Sensing Chinese Dynamics

Social-economic Space

Mapping Urban Sustainability

Cyber Urban Space

Sensing Citizen Emotion and PM_{2.5}

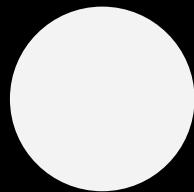
Physical Urban Space

Model Dynamic Urban Structure

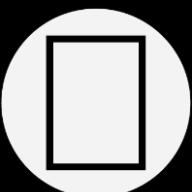
Urban Big Data Center

CAS-MIT China Urban
Big Data Center

A **pan-space, multi-scale, dynamic** China urban big database. Data volume > **100 TB**.



Block level



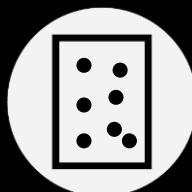
Parcel level



AOI level



Building level



POI level

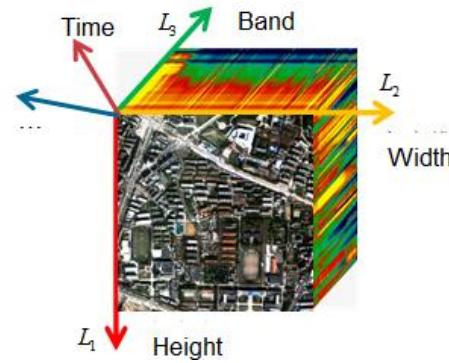


Data Integration / Fusion

Data covers main cities in China. Pan-space (physical, human, social-economic, and cyber space) data were integrated or fused, including but not limited to,

- Urban fundamental GIS data
- Census, Economic Census
- 365 * 24 hours MP positioning data
- Social media checkins

Tensor-based Land Use Classification



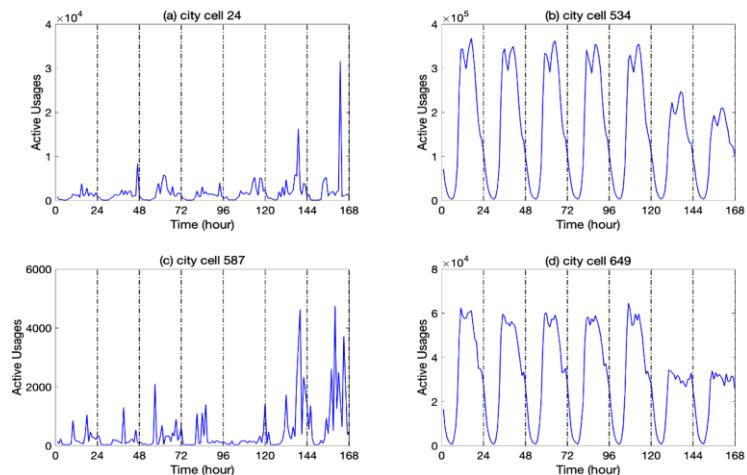
Tensor

Main advantages:

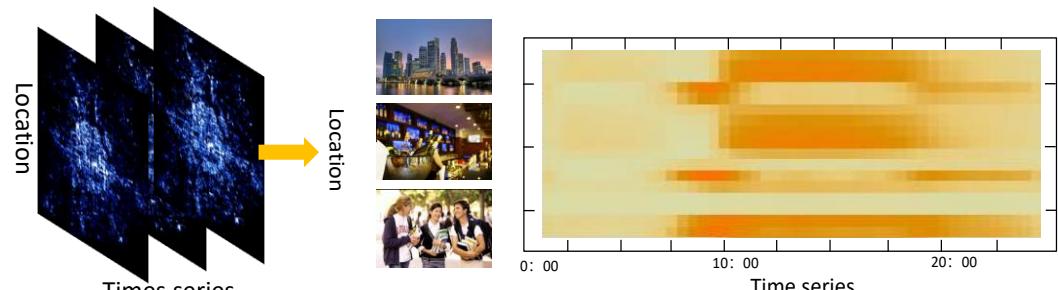
- ✓ Directly present the multiway info.
- ✓ Preserve the multi-dimensional correlations.

Objective function:

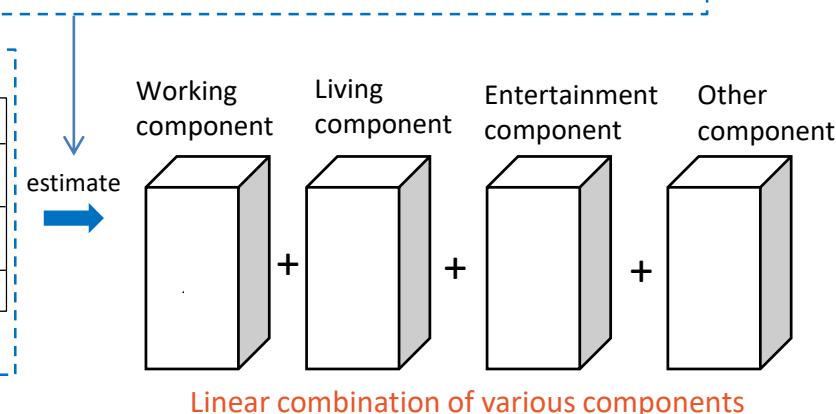
$$L(U, V, W) = \left\| Y - \sum_{k=1}^K u^{(k)} \circ v^{(k)} \circ w^{(k)} \right\|_F^2 + \Phi(U, V, W)$$



Respect to: $U \geq 0, V \geq 0, W \geq 0$



Mixed high order spatial-temporal data

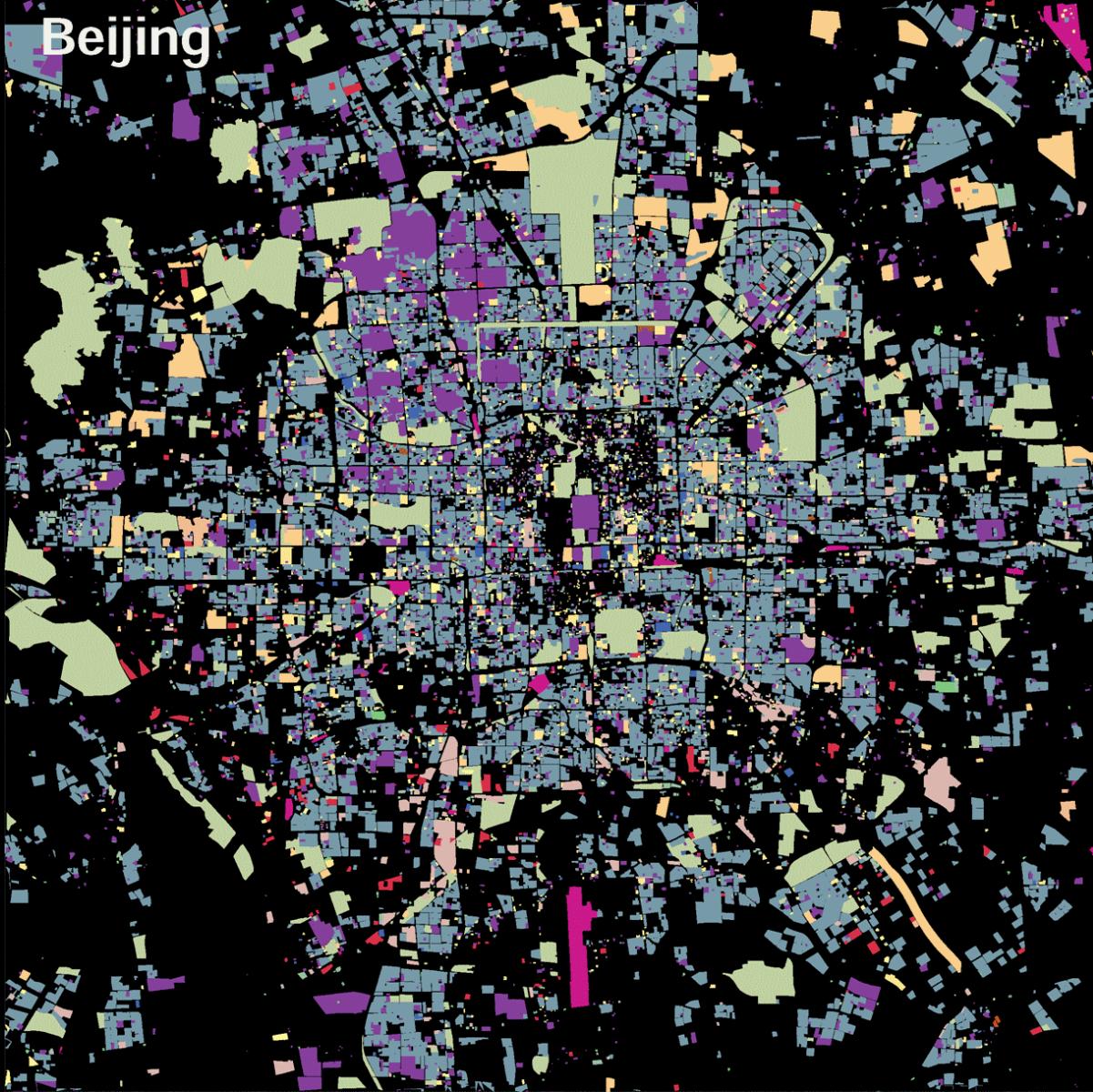


Urban Organs

Inferring urban detailed land use functions through pan-space urban information with Tensor-based artificial intelligence algorithm.

- 20 classes
- Overall accuracy: 91%

Beijing



Urban Cells

Understanding detailed urban structure and behave through the point-of-interest (POI) data.

Each point with different colors represent a urban unit (shop, school, hospital, or factory, etc.)

Shanghai

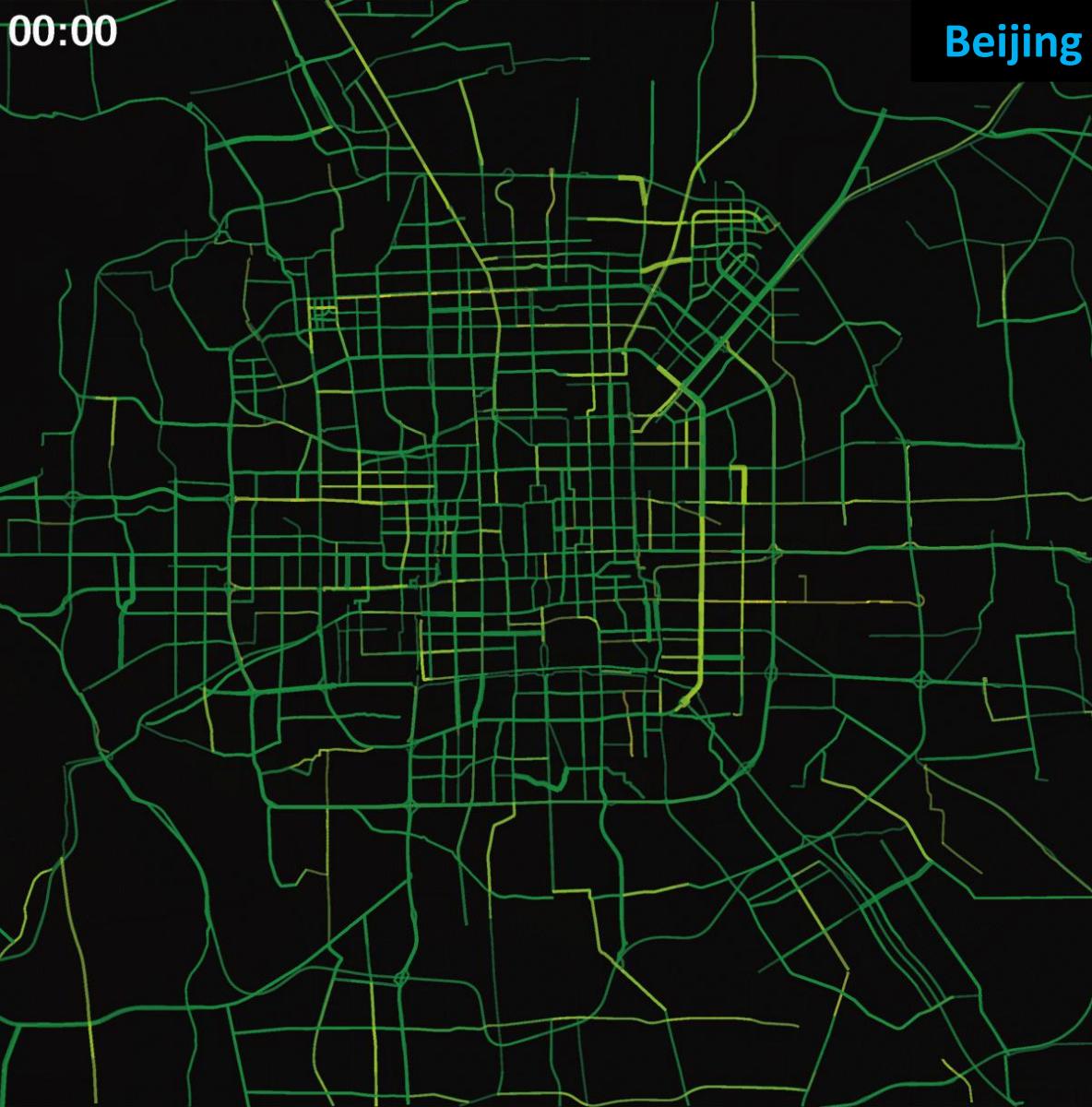


Urban Pulse

Sensing and forecasting
urban traffic congestion
with GPS equipped
floating vehicles data in
urban arteries.

Algorithm: Bayesian deep
learning with spatio-
temporal correlations

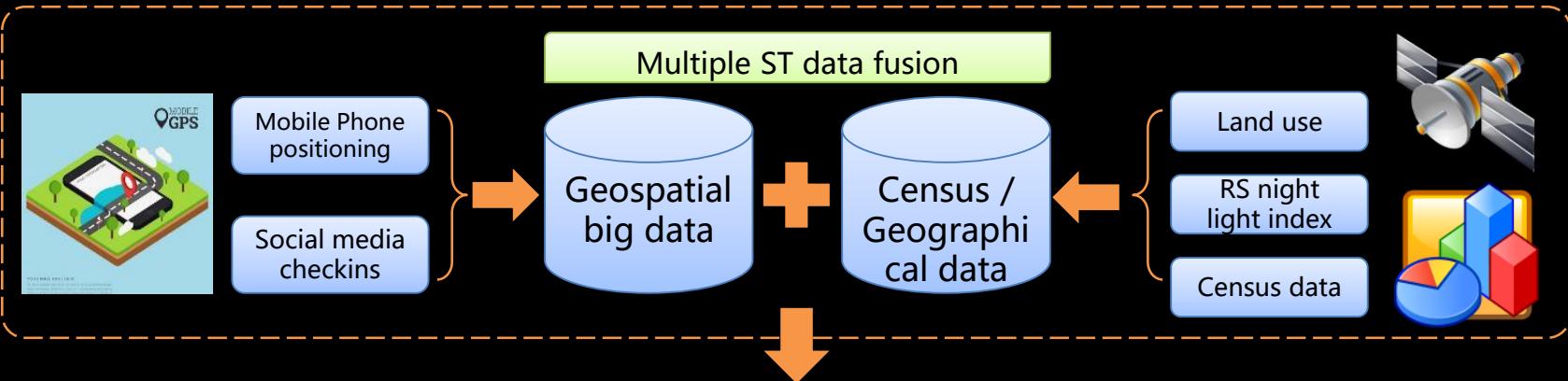
Forecasting accuracy: 85%



Human Activity Space

Sensing Chinese Dynamics

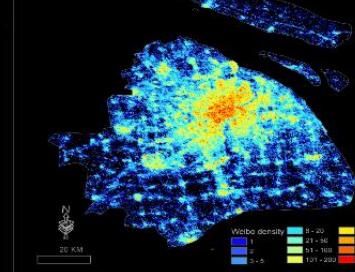
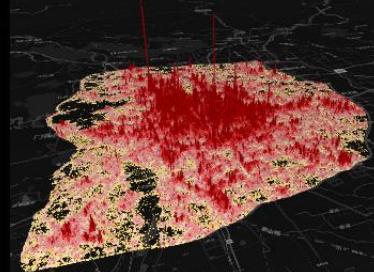
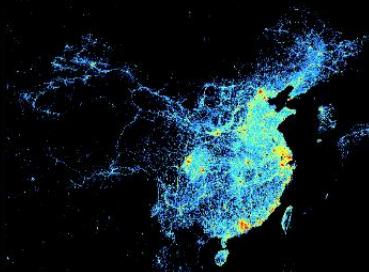
Dynamic Population Mapping



Hybrid high resolution dynamic population distribution model

$$P_{st} = f[Z_m(x, y, n), \mu_1 \cdot LU, \mu_2 \cdot S, \mu_3 \cdot R, \mu_4 \cdot TN, \dots]$$

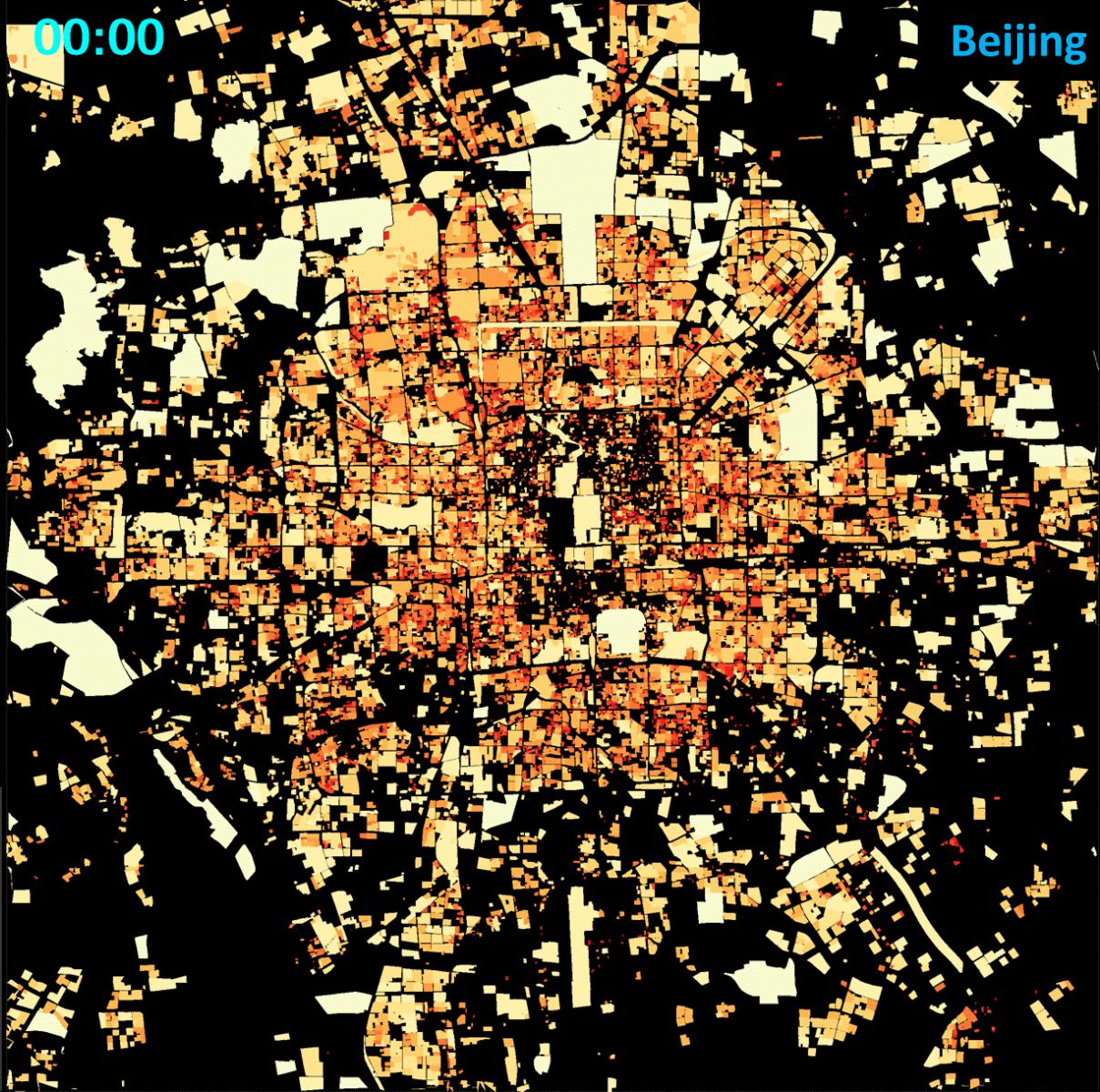
Dynamic people distribution at multiple scale



Urban Rhythm

Using mobile phone positioning data, we map urban human dynamic at very fine spatio-temporal scales

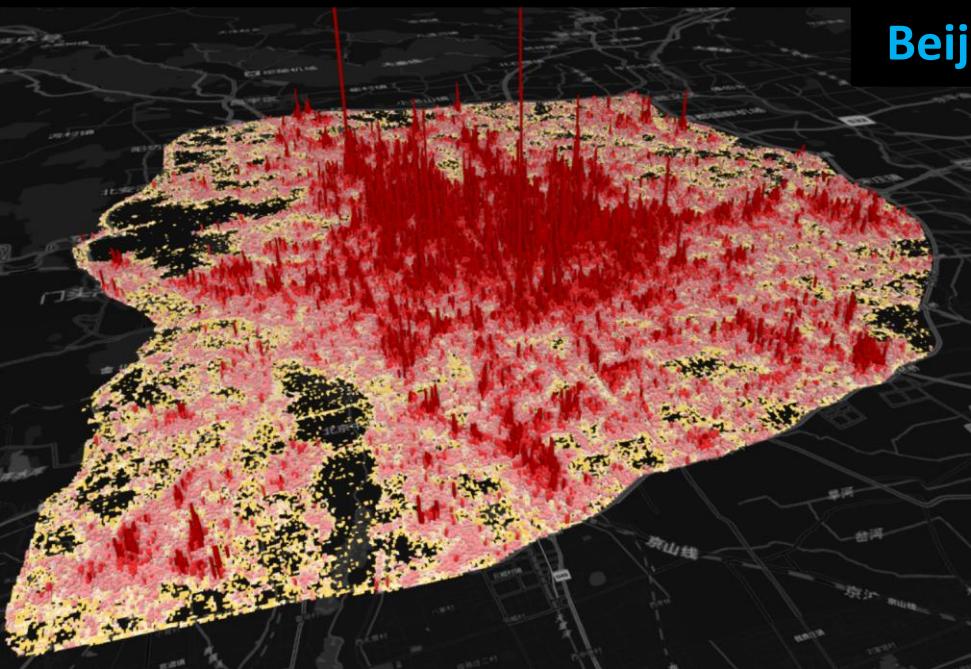
(Smart phone positioning requires)



Urban Rhythm

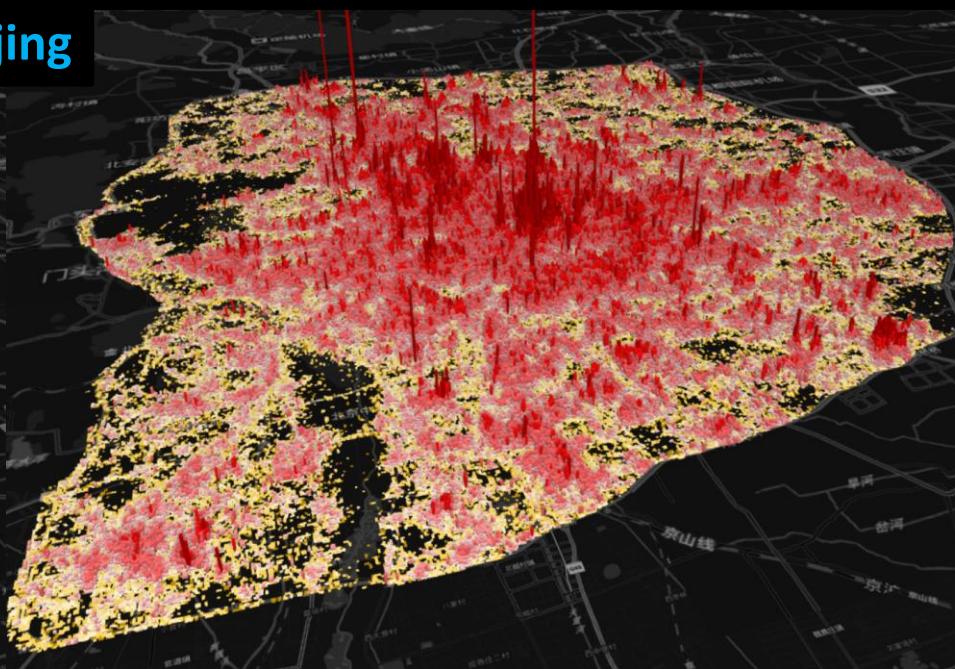
Dynamic population mapping at 100-m resolution was estimated by using census and multiple geospatial big data with hybrid area-to-point geostatistical technique.

Day time population (12:00)

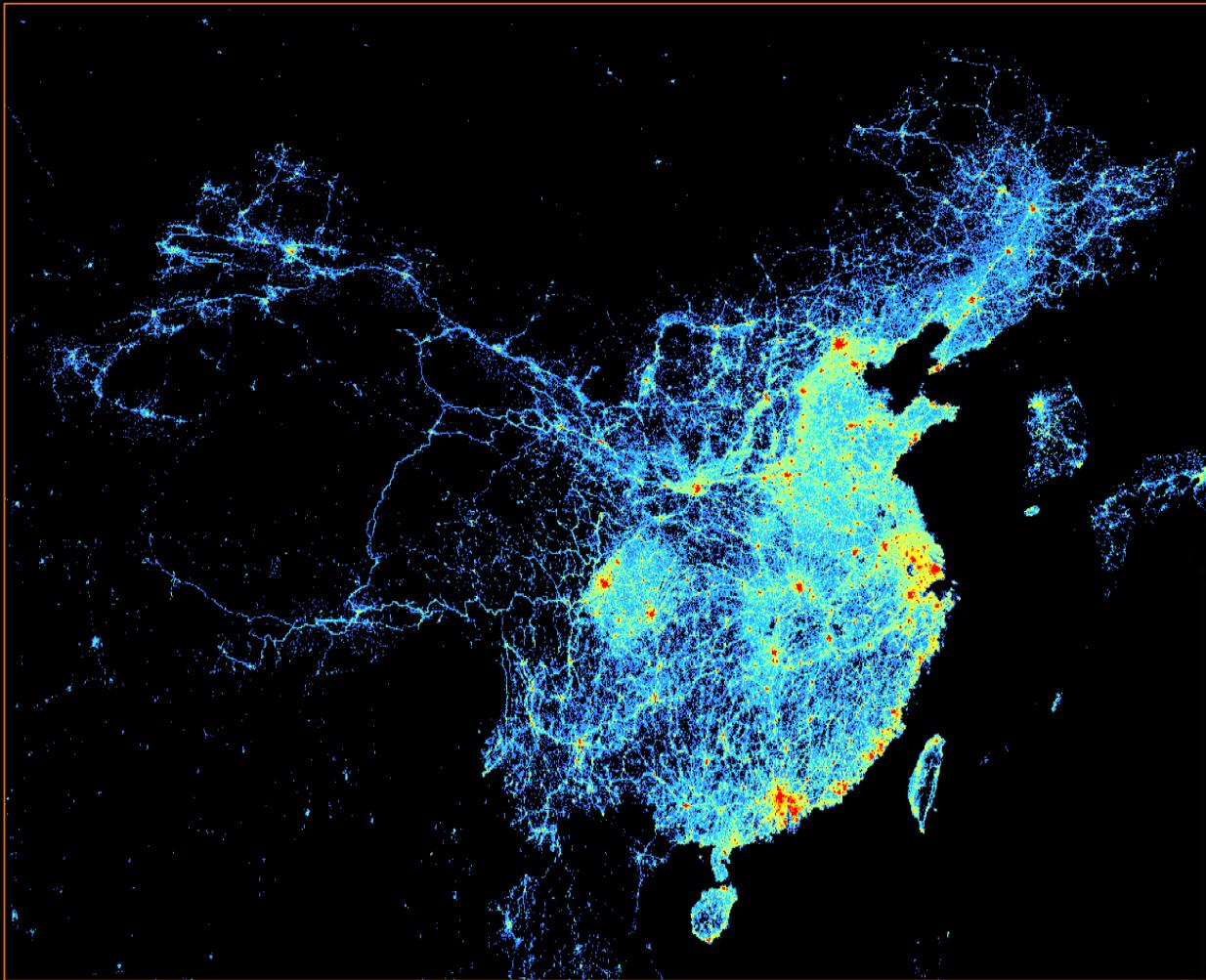


Beijing

Night time population (24:00)



Chinese Dynamics



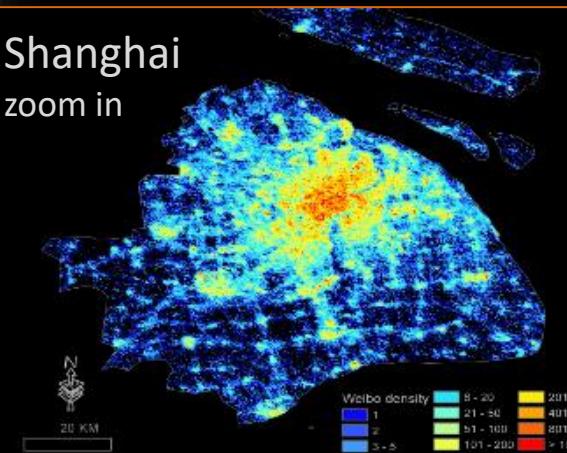
Location Based Service (LBS) data from *Tencent* Big Data center.

900 million users / day



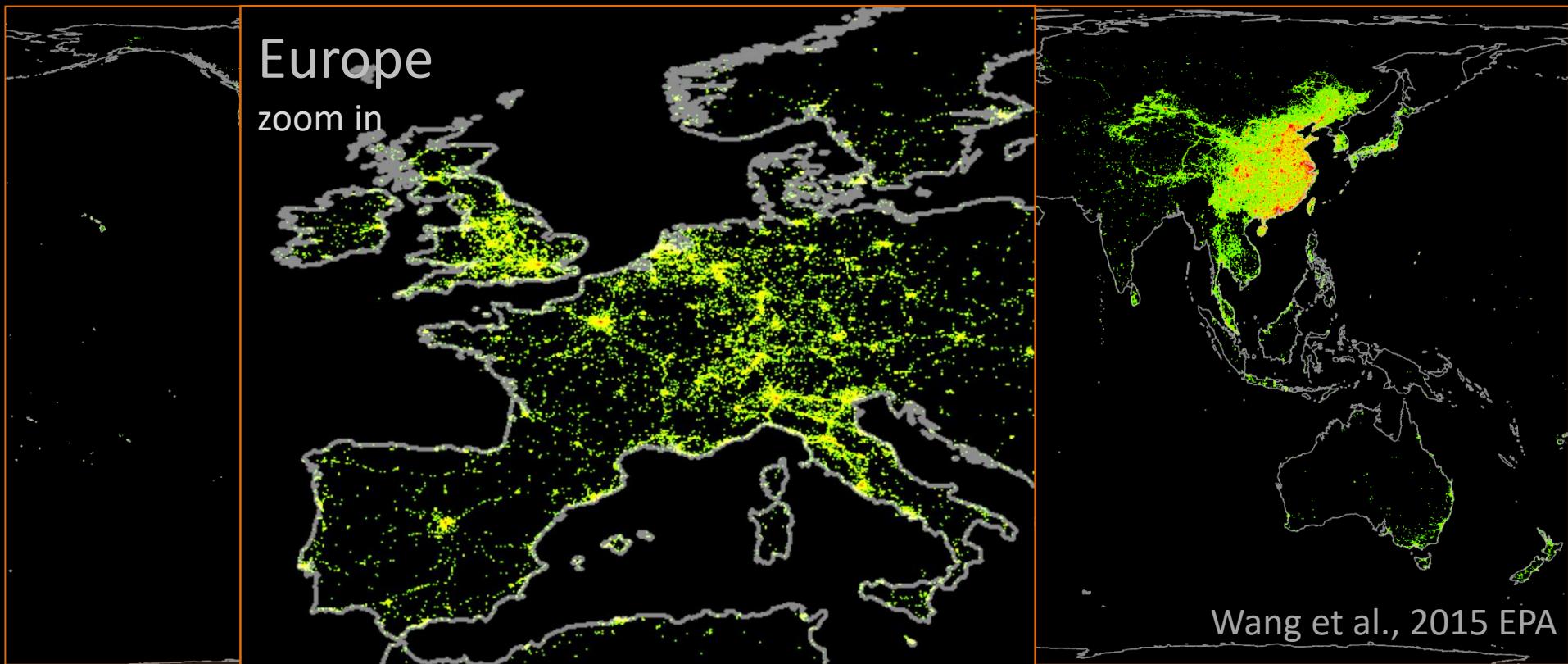
60 billion location request / day

Shanghai
zoom in



Chinese Footprint

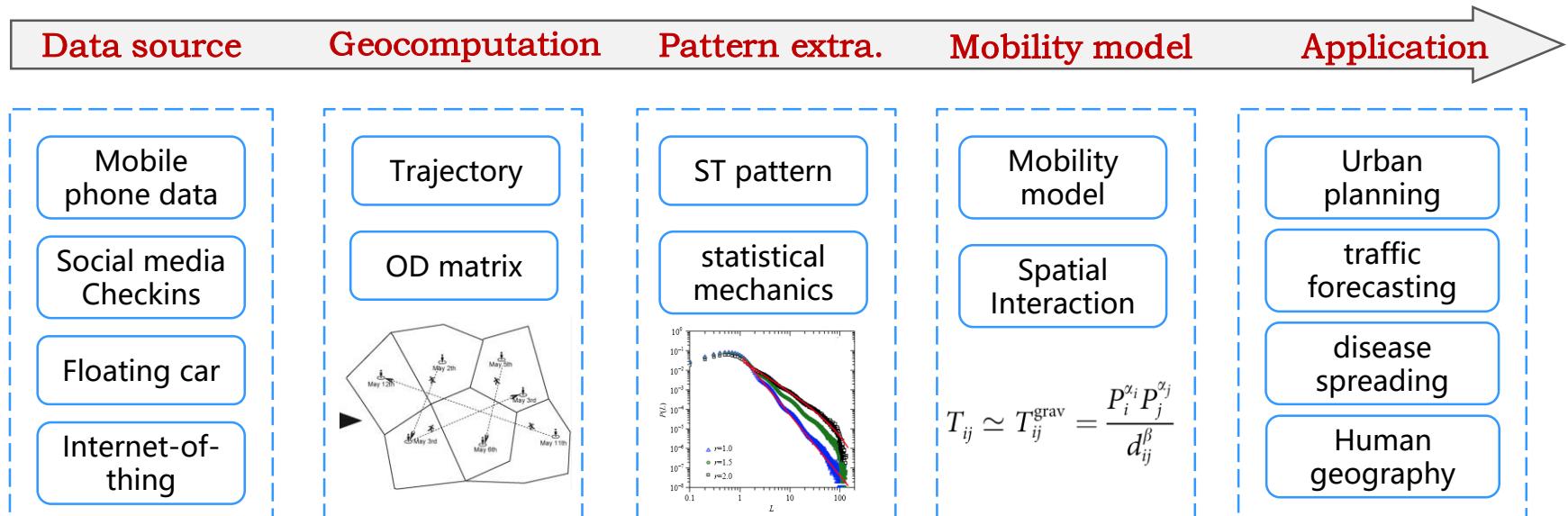
Mapping Chinese footprint worldwide with 400 millions Weibo (equate to Twitter) checkins to answer the question of “Where are the Chinese?”



Wang et al., 2015 EPA

Model Chinese Mobility

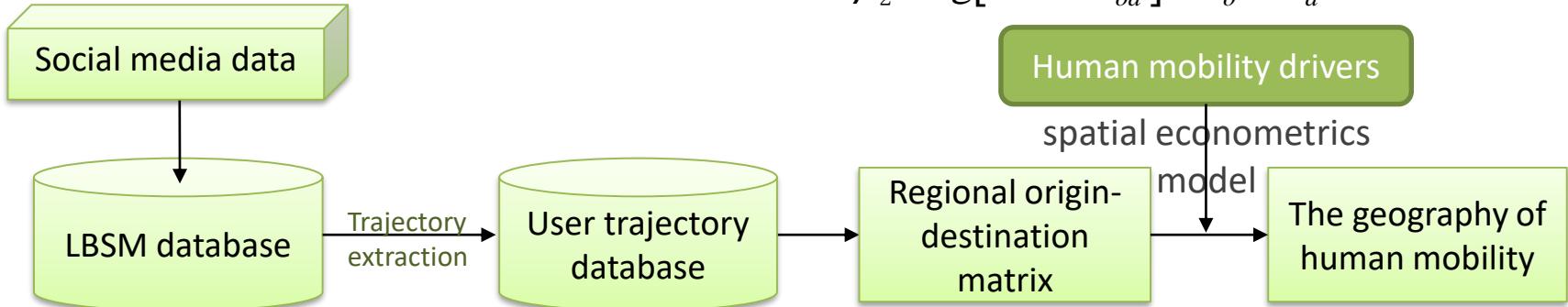
- Understanding individual human mobility is of fundamental importance for many applications from urban planning to disease spreading and traffic forecasting
- We propose social media (Weibo) checkins as a proxy for human mobility, as it relies on publicly available data and provides high resolution positioning when users opt to geotag their posts with their current location.



Understanding Chinese Mobility from Weibo

- Chinese mobility patterns is measured by big data computational strategy for identifying hundreds of millions of individuals' space–time footprint trajectories;
- We discovered dialect-based culture ties control the Chinese mobility pattern;
- Our study provides solid evidence that Weibo checkins can indeed be a useful proxy for tracking and predicting human movement.

$$\log(T_{od} / L_o) = \beta_1 \cdot \log[\text{Commuting}_{od}] \\ + \beta_2 \cdot \log[\text{Dialect}_{od}] + F_o + F_d + \text{controls} + \varepsilon_{od}$$

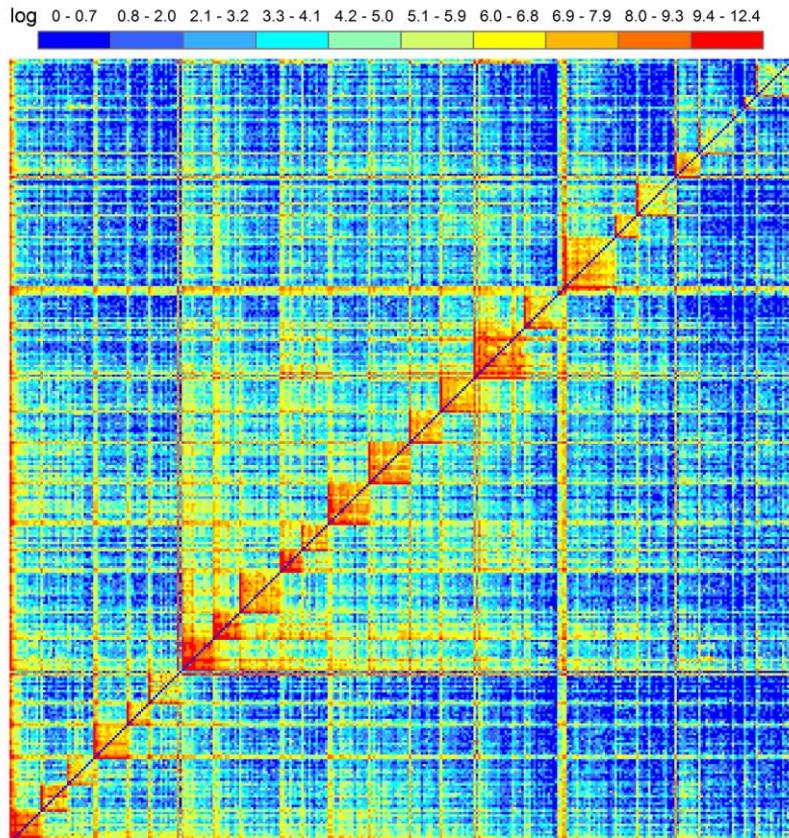


The Geography of Cultural Ties and Human Mobility: Big Data in Urban Contexts, *Annals of AAG*, 2016

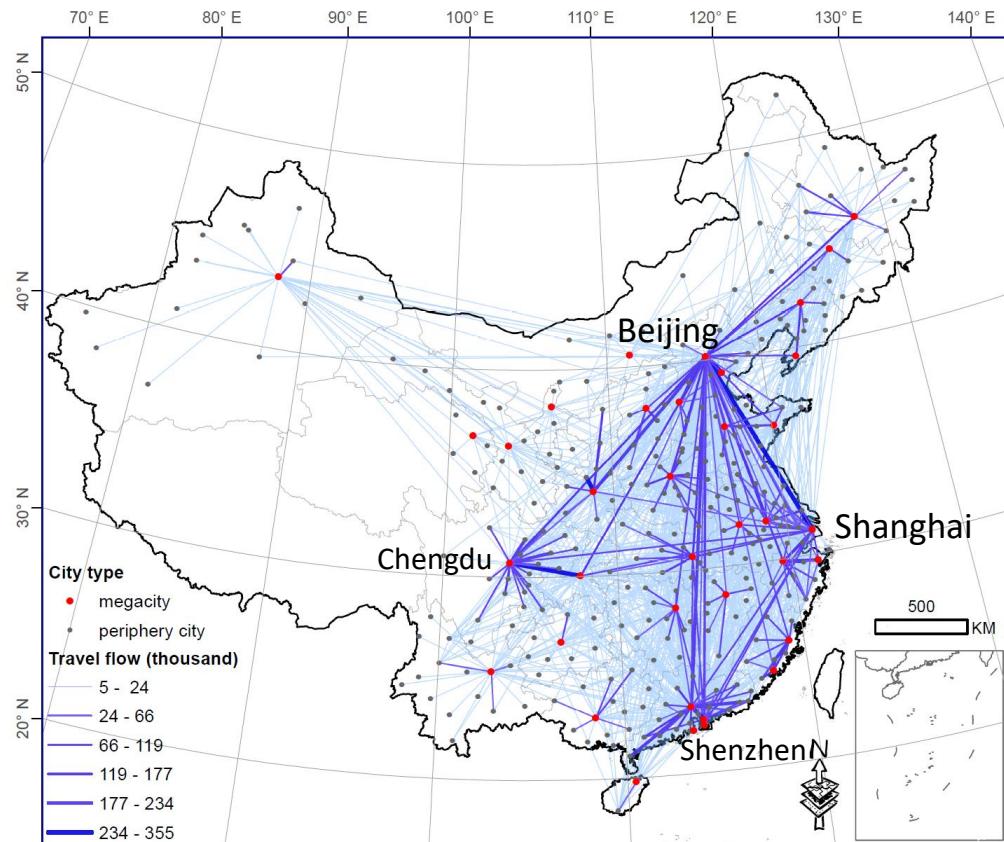
The Geographical Legacies of Mountains: Impacts on Cultural Difference Landscapes, *Annals of AAG*, 2017

Understanding Chinese Mobility from Weibo

City-level Chinese mobility
origin-destination (OD) matrix

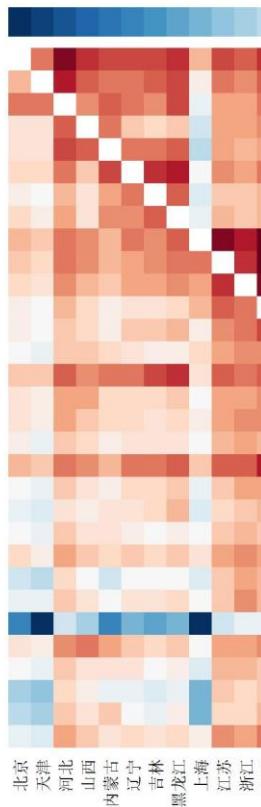


City-level Chinese mobility
in geographical space

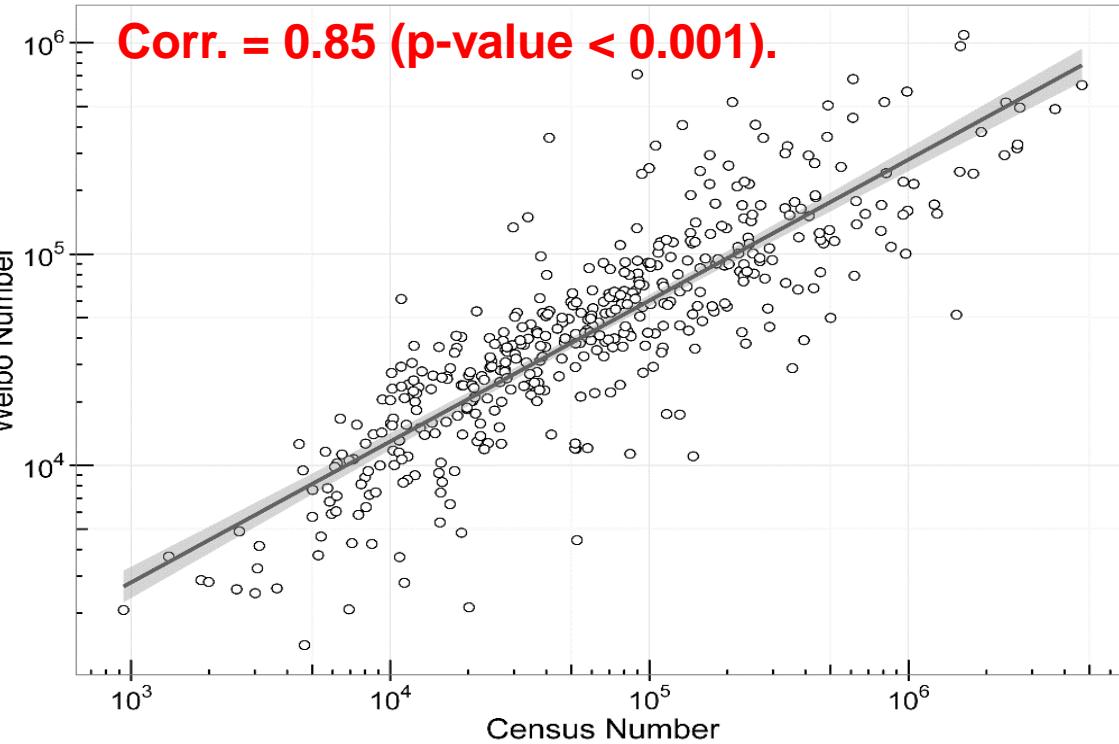


Cross-validation

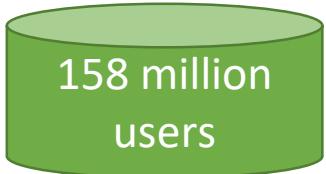
Floating population in 6th Census



Weibo estimated human mobility

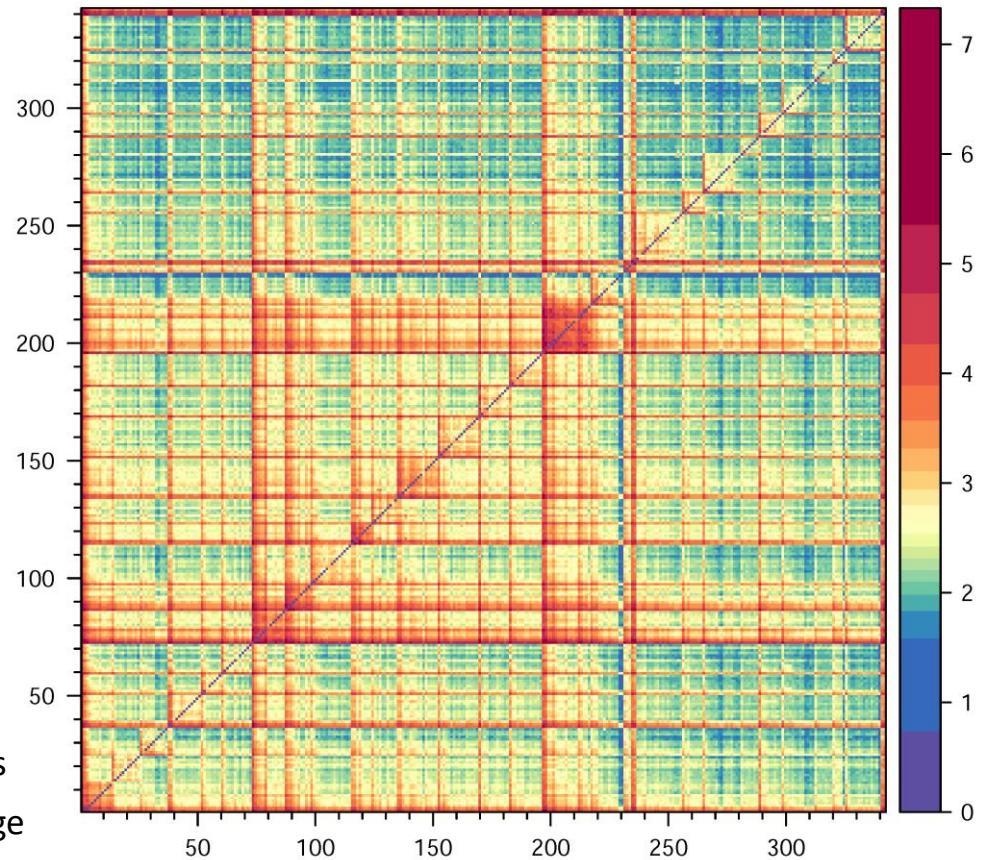


City Friendship



16.6 million users' friendship networks nodes, covers
2.20 billion edges (linkages), ~**140** friends on average

City-level Chinese Social Media
Friendship OD matrix



Social-Economic Urban Space

Mapping Urban Sustainability



SDG CHINA
美丽中国



<http://sdgcn.org>



2030年可持续发展议程

Sustainable Development Goals (SDG)

联合国SDG可持续发展目标和指标体系解析

地球大数据驱动的“美丽中国”全景评价

关于SDG

About SDG



The target of Urban Sustainability

11 SUSTAINABLE CITIES AND COMMUNITIES

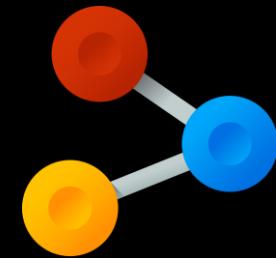


Make cities and human settlements inclusive, safe, resilient and sustainable



10

Target



15

Indicator

Public Transportation

<http://jianghao.wang/bus/>

Chinese Bus System

Home

Beijing 北京

Shanghai 上海

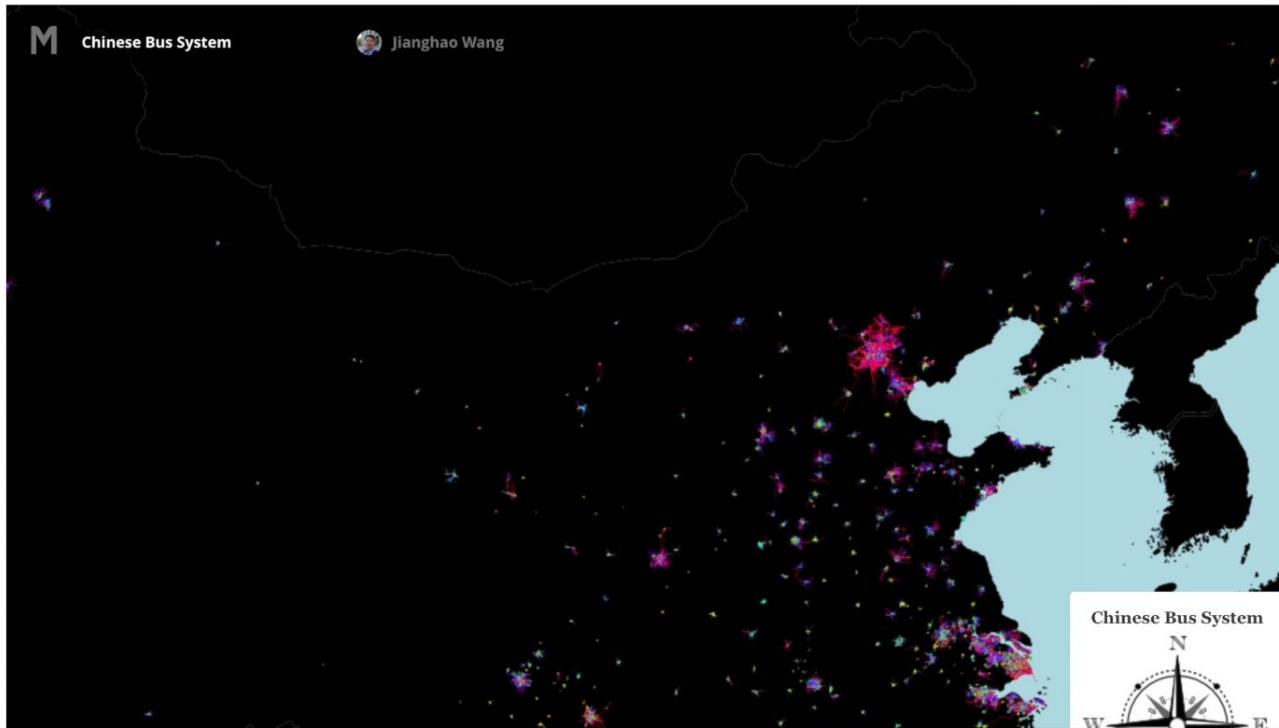
Guangzhou 广州

Shenzhen 深圳

About

Chinese bus system visualization

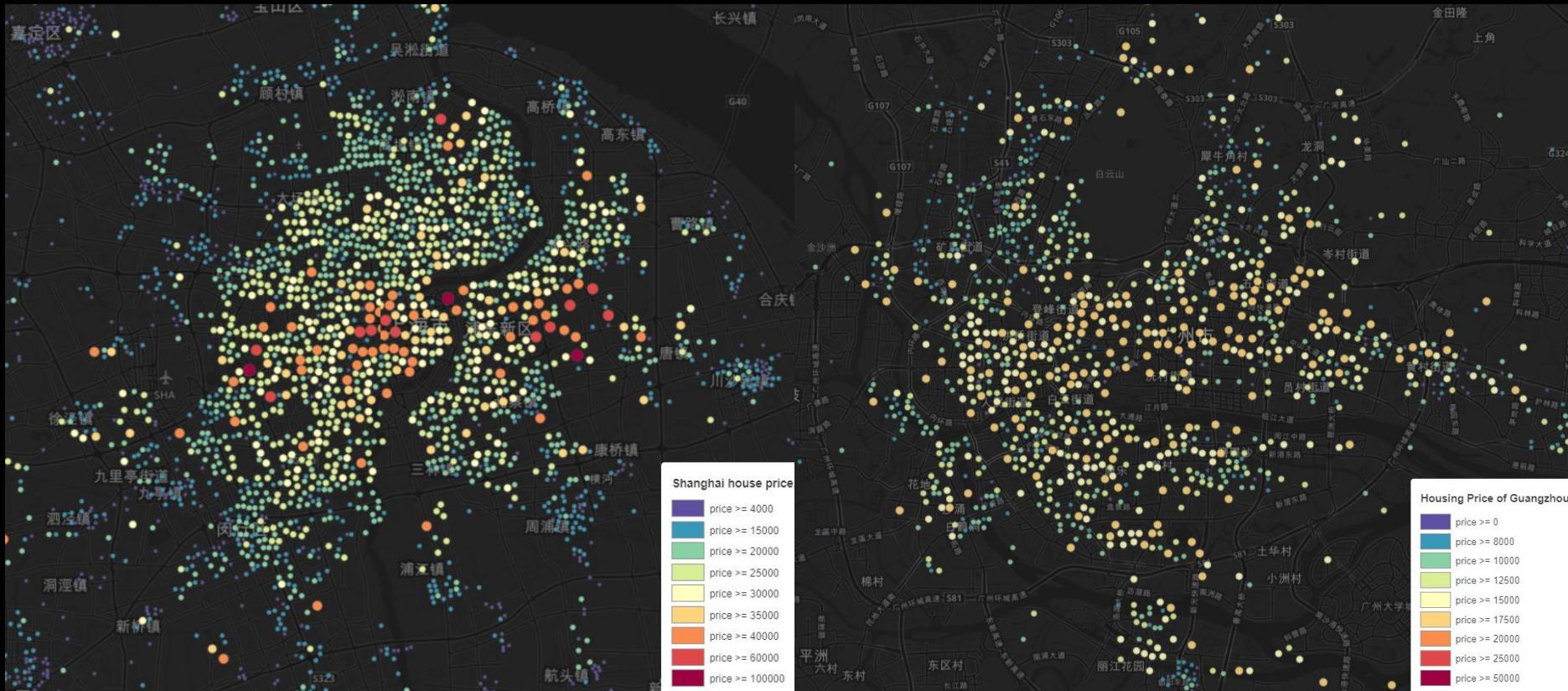
Visualize national wide prefecture-level cities bus system (exclude Taiwai province). There are totally 24,500 bus lines in China. [Full screen](#)



Urban Real Estate

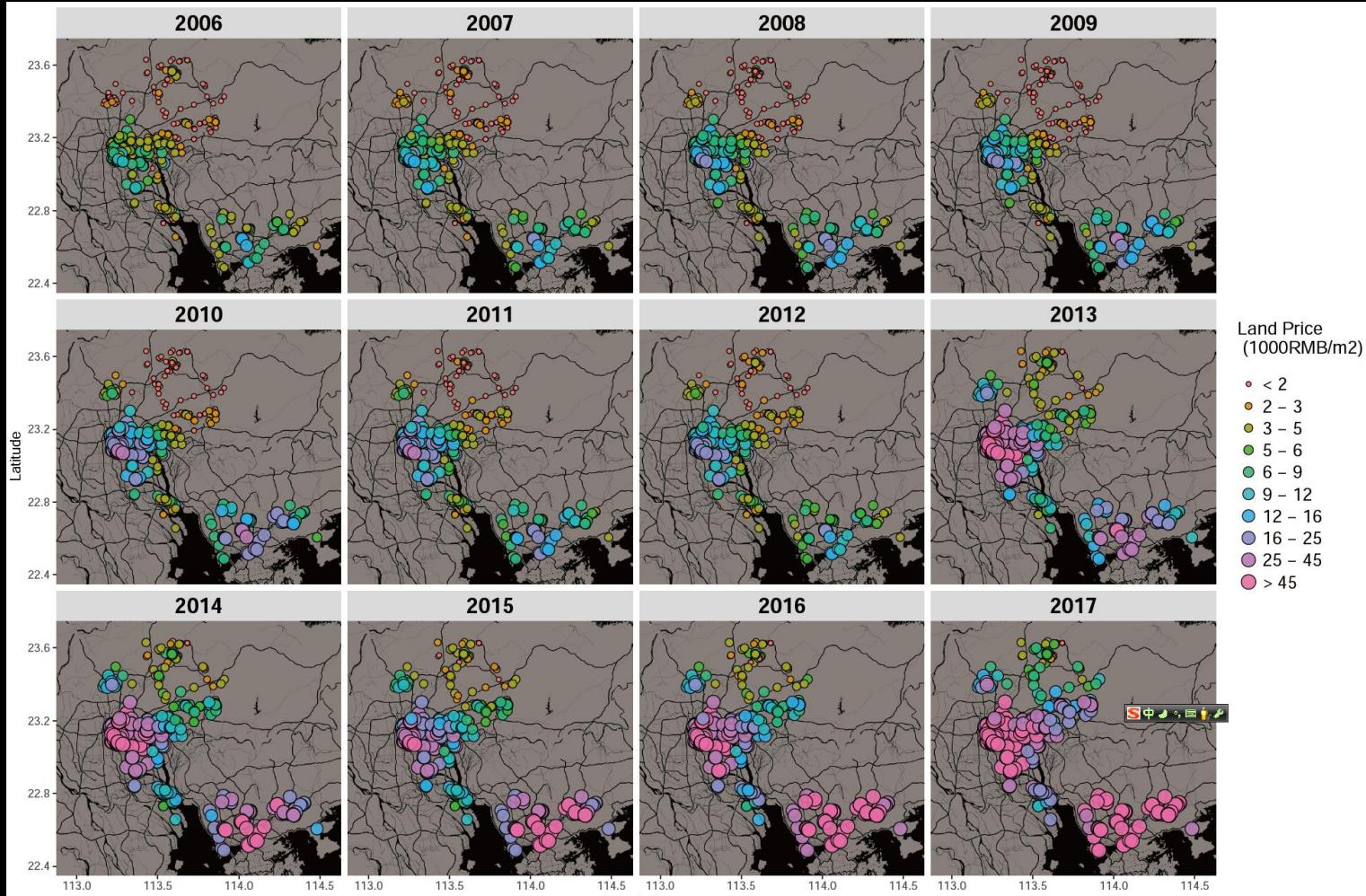
Land and housing panel prices

<http://jianghao.wang/HousePrice/>

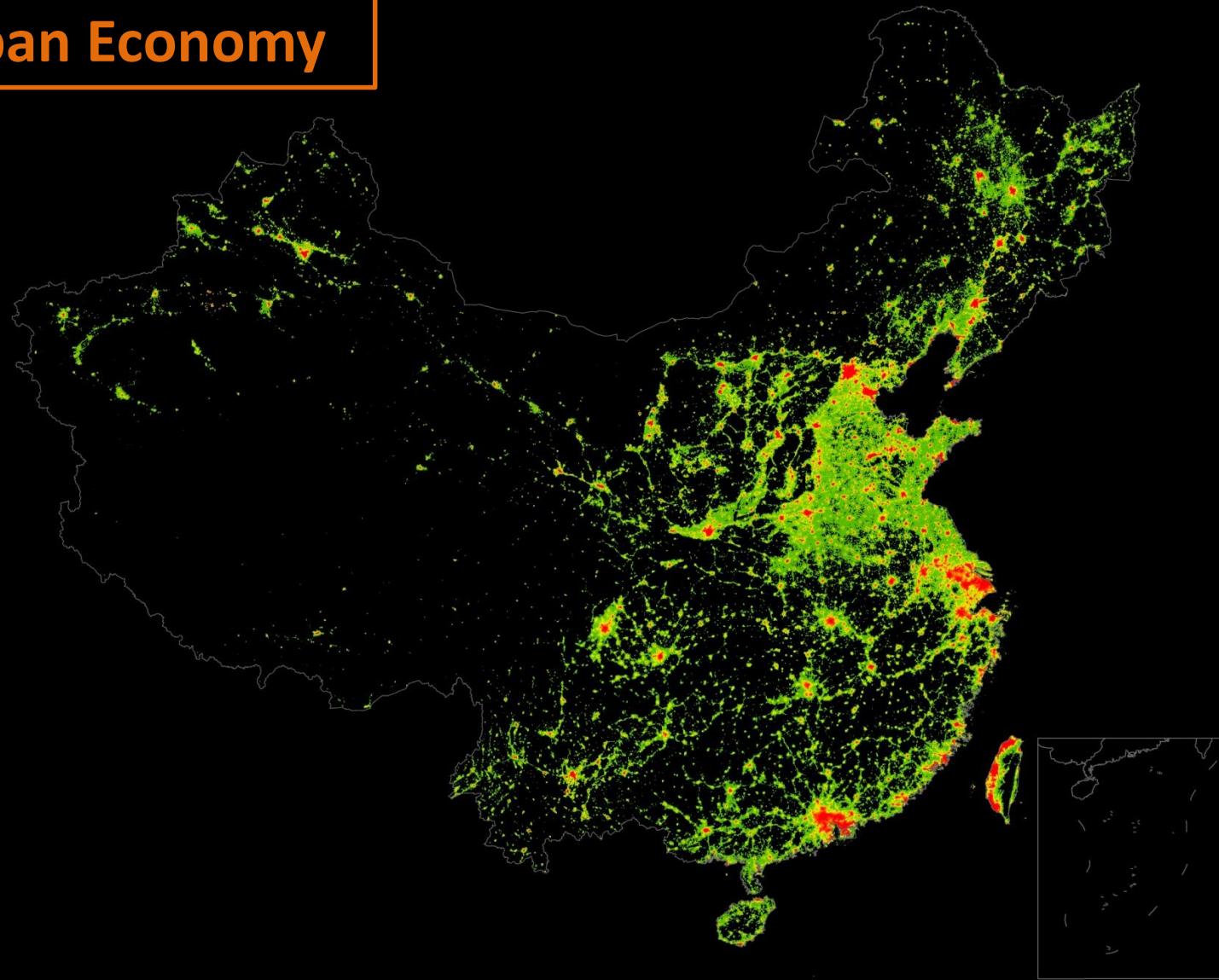


Urban Real Estate

Land transaction dynamic in PRD



Urban Economy



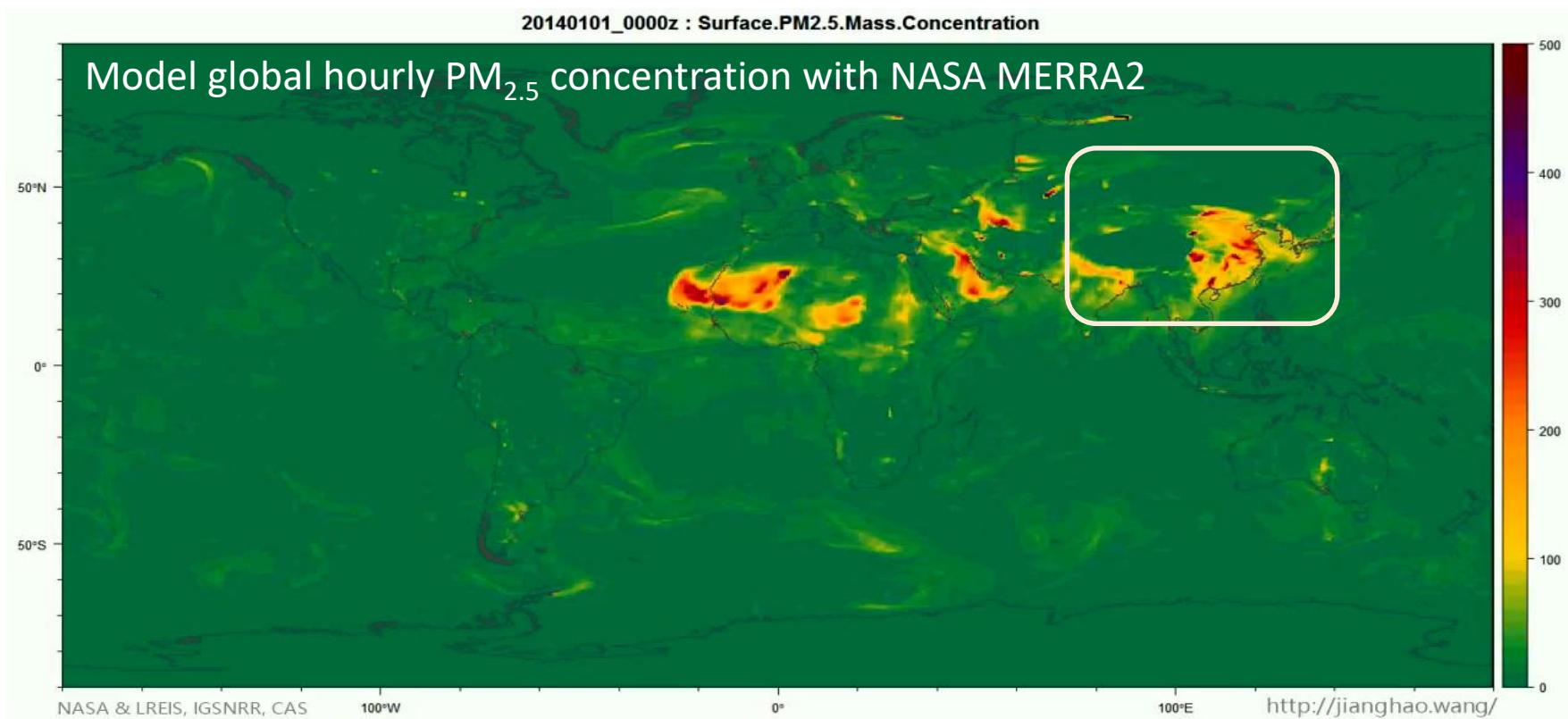
Cyber Urban Space

Citizen Emotion Sensing and PM_{2.5}

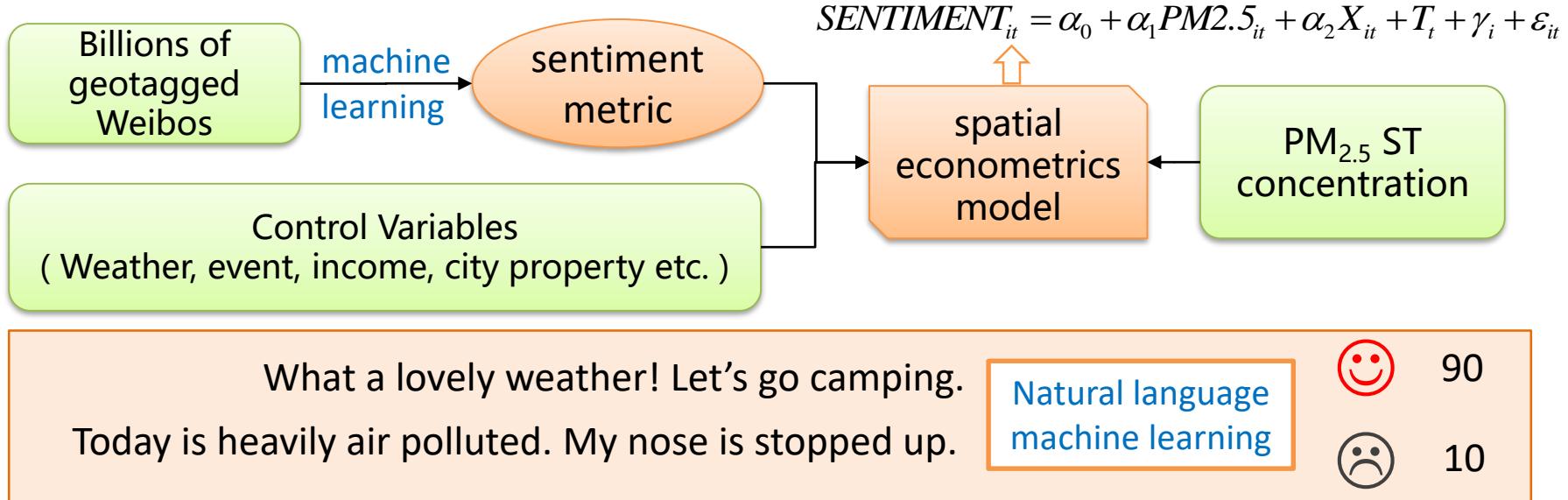
PM_{2.5} Air Pollution in China

China's high level of ambient air pollution causes sickness, excess mortality risk.

Study health impact ---> Measure social cost ?



Citizen Emotion Sensing and PM_{2.5}

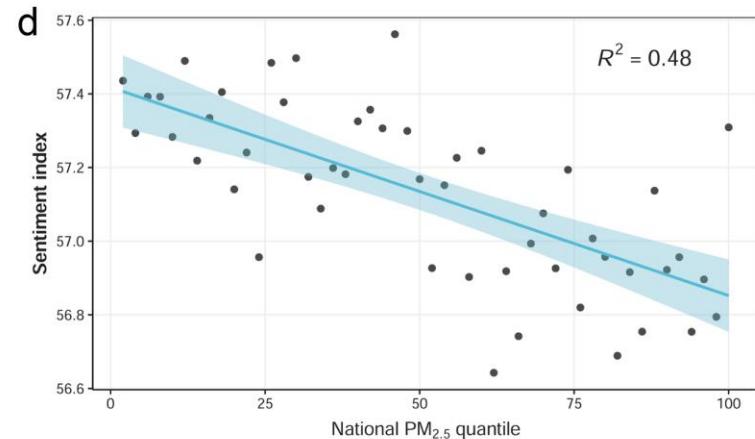
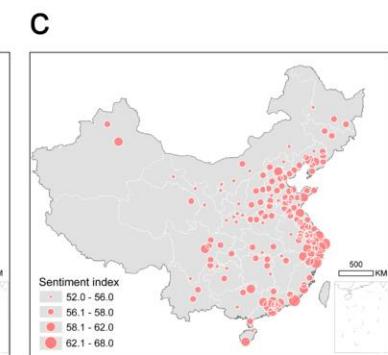
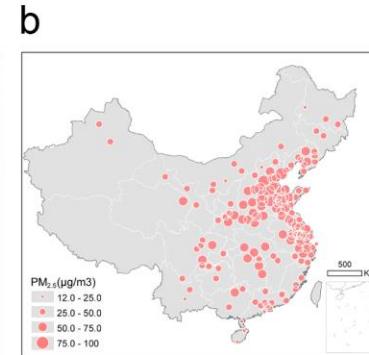
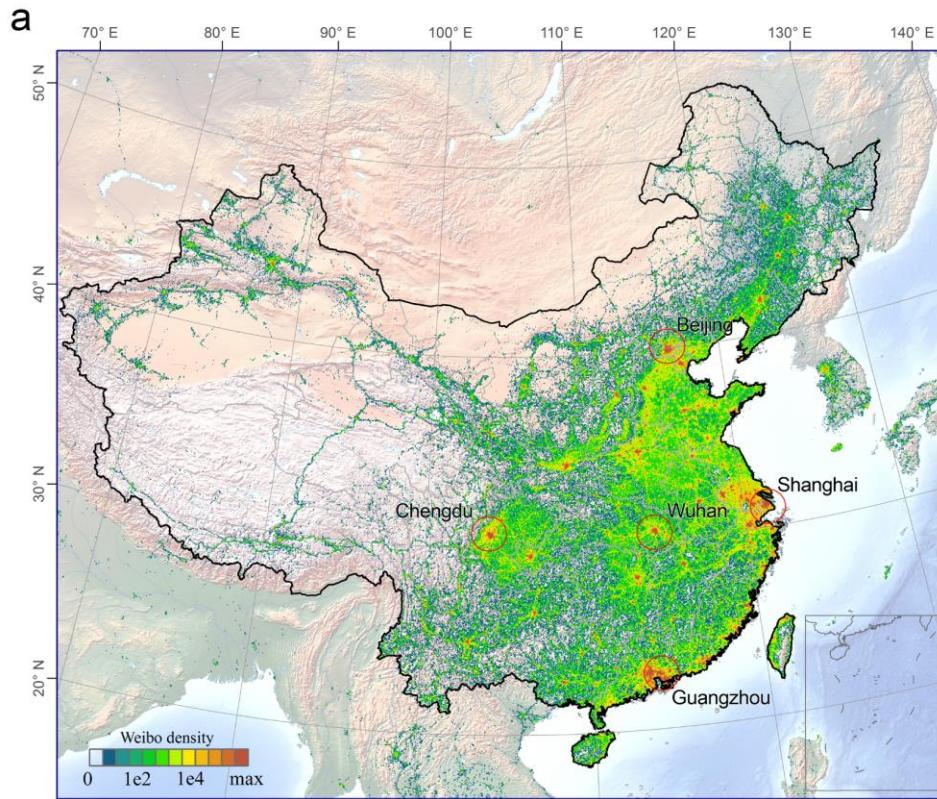


Main findings:

- One standard deviation increase in the PM_{2.5} concentration is associated with a 0.05-0.06 standard deviation decrease in the sentiment index.
- One standard deviation increase in the city's PM_{2.5} concentration can be offset by a 6.5 thousand RMB (\$940) increase in the city-level annual wage.

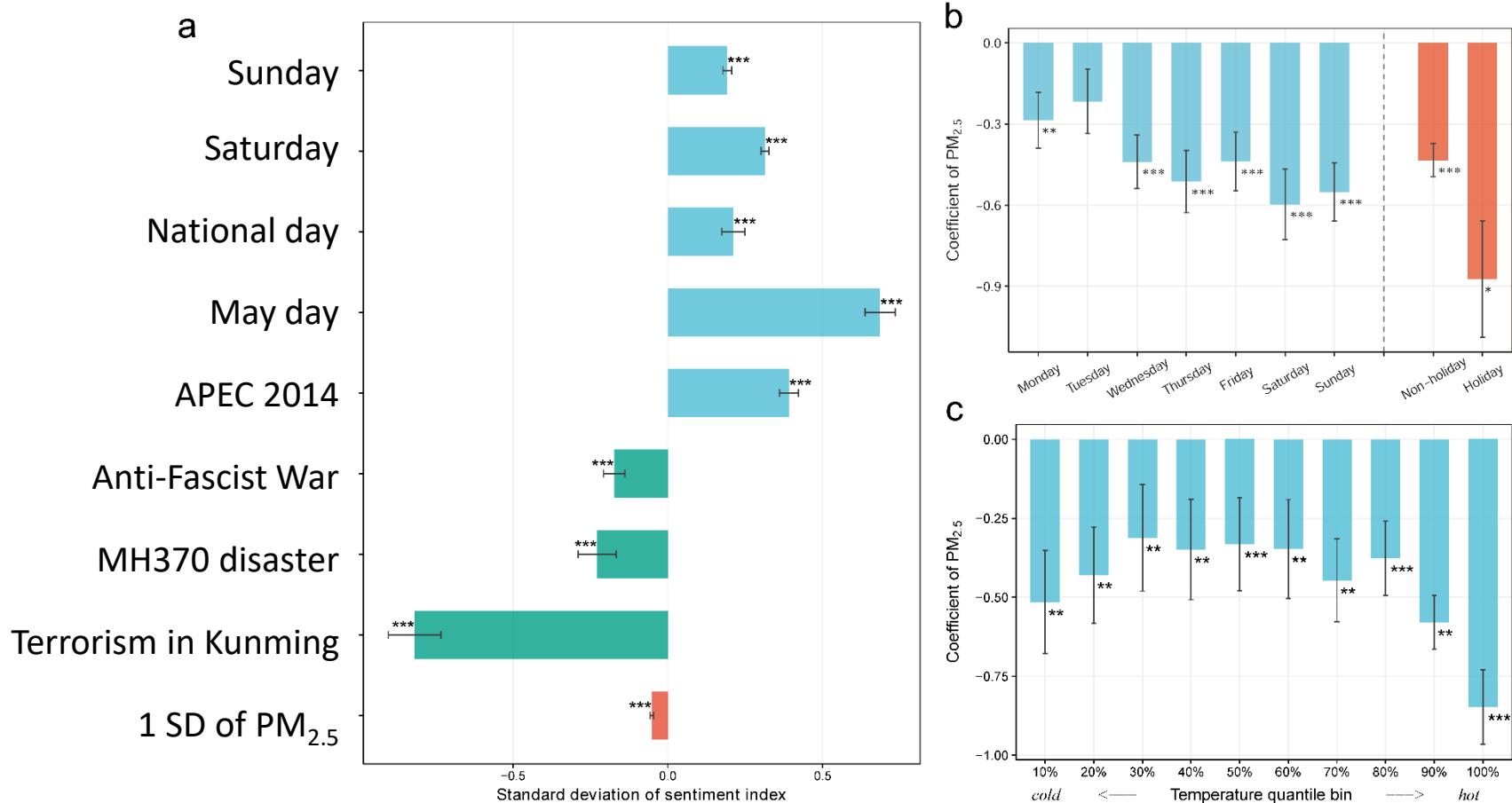
Emotion Sensing and Air Pollution

The Geography of Weibo posts, PM_{2.5} concentration and sentiment index, and their national relationship.



Emotion Sensing and Air Pollution

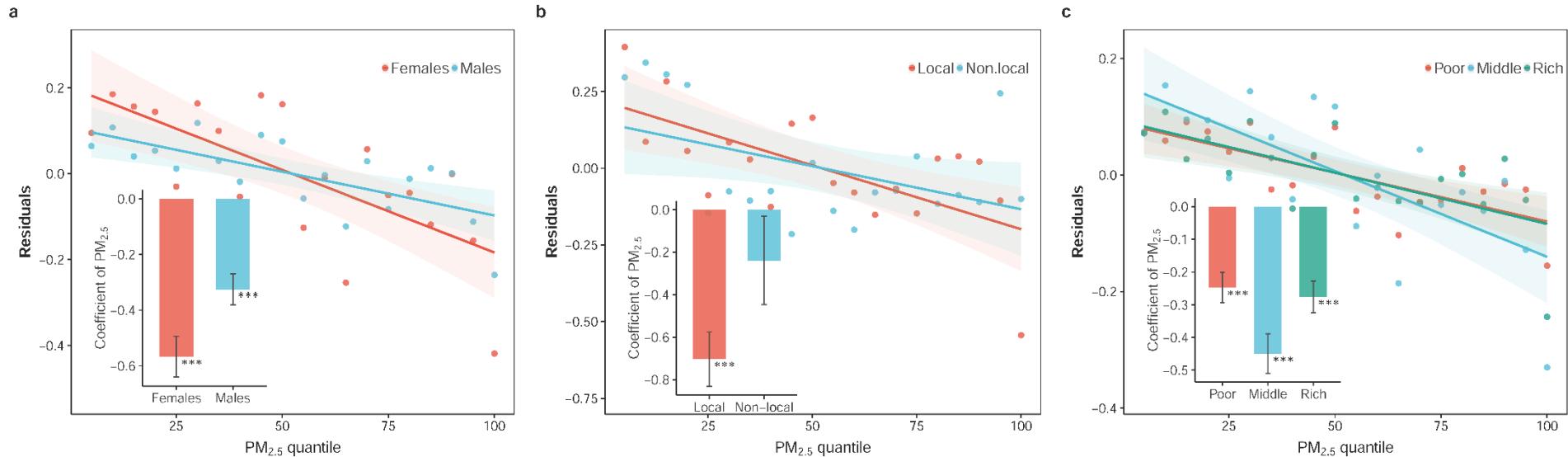
The heterogeneous effect of air pollution on sentiment on different days



Emotion Sensing and Air Pollution

The heterogeneous effect of air pollution on sentiment across population groups

- a) **males and females:** Females are more sensitive than males when $PM_{2.5}$ increase
- b) **local and non-local microbloggers:** Local people are more likely to be upset when $PM_{2.5}$ increases compared to non-local, indicating a long-term pollution exposure effect.
- c) **three income groups:** Both the rich and poor are less sensitive to air pollution than the middle class.

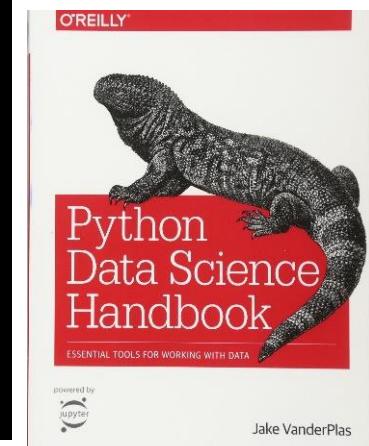
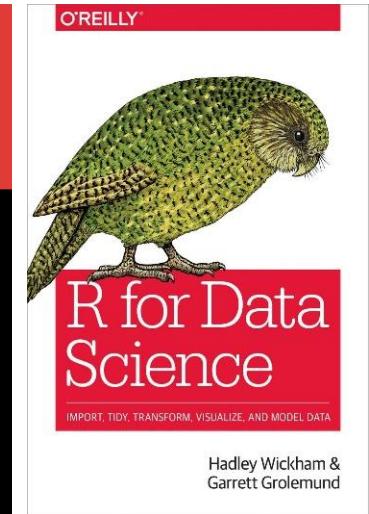


#

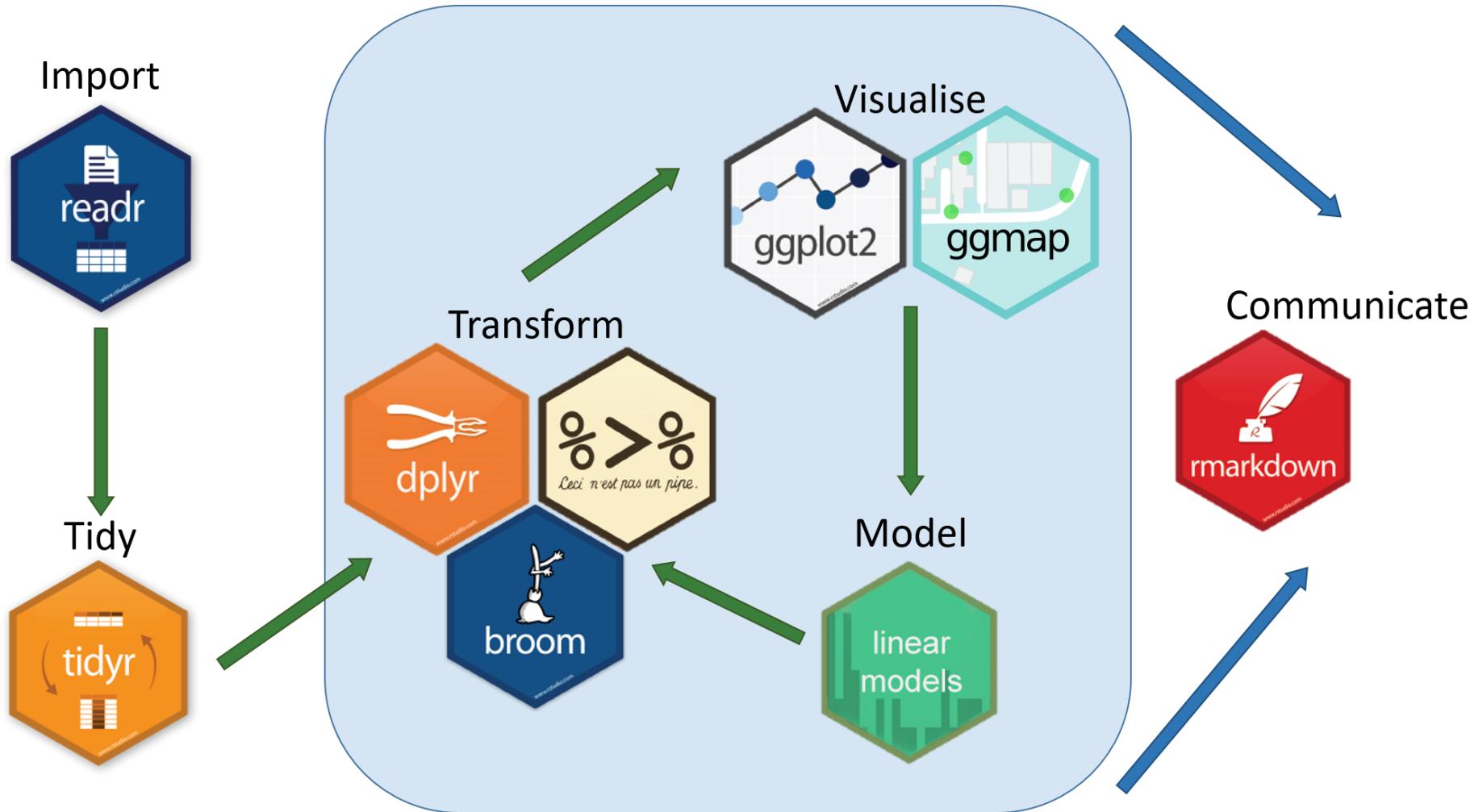
Geocomputation with OSGeo

Data Scientist in Action

Source Data
Store Data
Convert & ETL
Transform Data
Exploratory Analysis
Model Build & Generate Insights
Visualisation
Model Execution in Production



R Ecosystem for Data Science





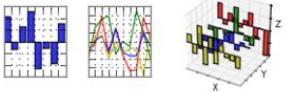
Python in Action



IP[y]: IPython
Interactive Computing

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



OSGeo Ecosystem



GEOS
Geometry
Engine
Open
Source

PROJ.4

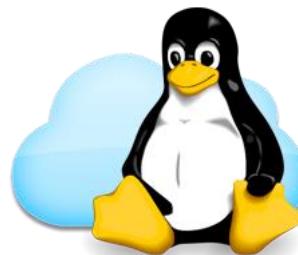


Data Management

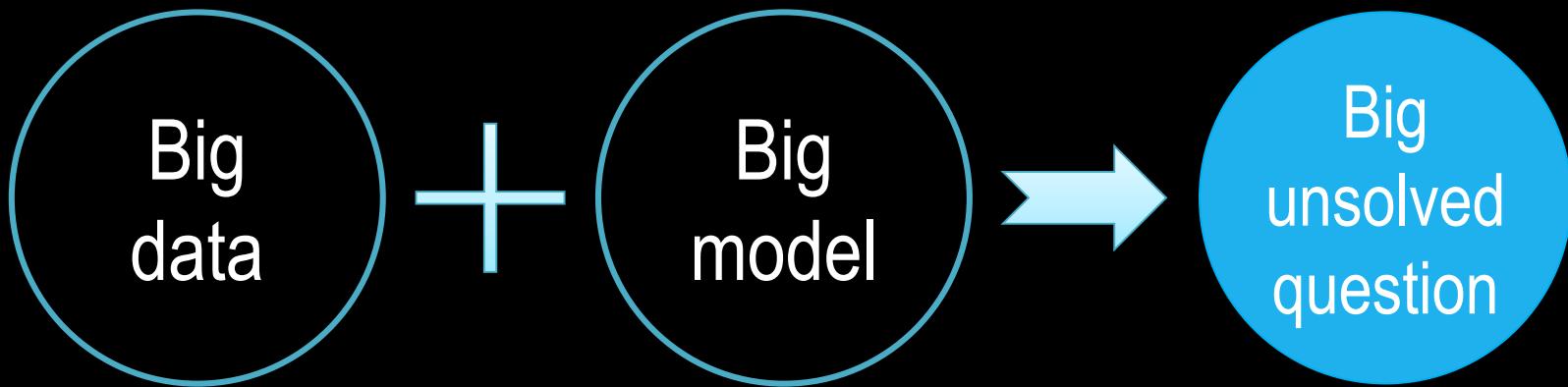
Computation

Visualization

My Geocomputation Workflow



Deep Thinking in Action



Cross-disciplinary cooperation !

Goodchild et al. (2012) PNAS:

“The supply of geographic information from satellite-based and ground-based sensors has expanded rapidly, encouraging belief in a new, fourth, or “big data,” paradigm of science that emphasizes **international collaboration, data-intensive analysis, huge computing resources, and high-end visualization**.”

Thanks

Q & A

<http://jianghao.wang>