Urban Sensing and Computing

Introduction and Application

王江浩

State Key Laboratory of Resources and Environmental Information System,
Institute of Geographic Sciences and Natural Resources Research,
Chinese Academy of Sciences

Dec 28, 2017 @ THU

Main Content

URBAN SENSING & COMPUTING

1. INTRODUCTION
2. BIG DATA IN URBAN STUDIES
3. LESSONS FROM BIG DATA

1 Introduction

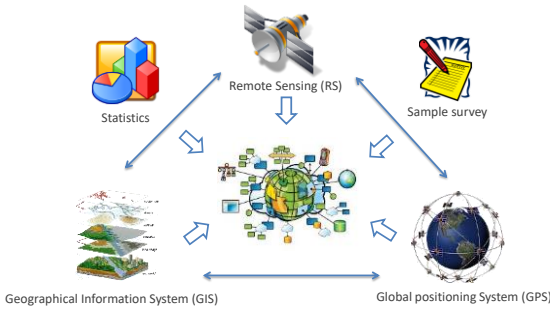
Introduction

- Urban is a complex giant system;
- Problem we faced: traffic congestion, air pollution; sustainable development
- Big data, from earth observation system to our daily lift, providing a new way to better understanding our world. Data is the new oil of the digital world.

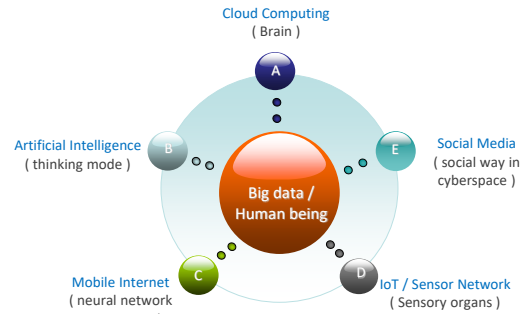
Goals for Sustainable Development

Sustainable Cities and Communities

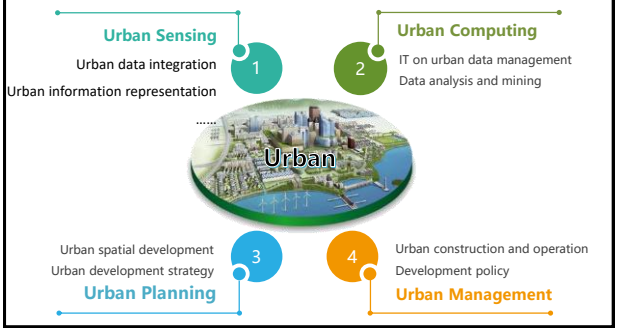
Traditional Way to Model Our World



New Technologies to Understand Our World



Urban Informatics

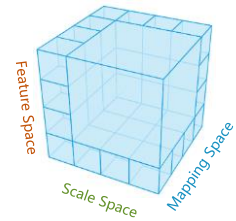


Pan-space Urban Sensing

cubic Urban Sensing Framework

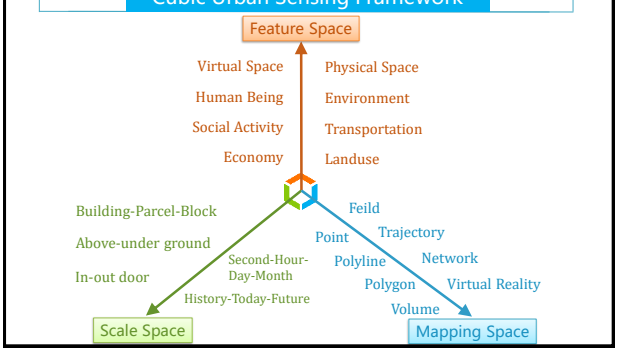
Pan-spatial GIS

—Academician Zhou

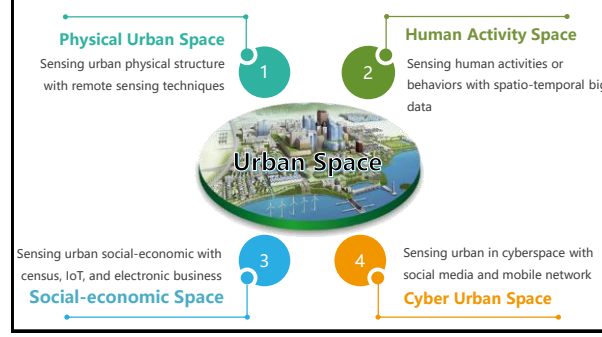


Pan-space Urban Sensing is crossover study investigating the use of digital sensors (including remote sensing, internet of things, wireless sensor network, citizen-sensing, et al.) to obtain, compute, and analyze urban dynamic environments to achieve more sustainable and efficient cities.

Cubic Urban Sensing Framework



Pan-Space Urban Sensing

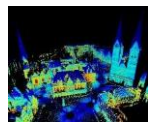


Pan-Space Urban Sensing

Pan-space Urban Sensing Technology

- **Physical Urban Space**
- Human Activity Space
- Social-economic Space
- Cyber Urban Space


Remote sensing


Lidar photogrammetry


Street View mapping system

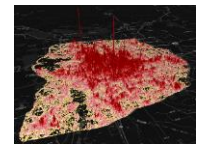
Sensing urban physical structure with remote sensing techniques

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

Pan-Space Urban Sensing

Pan-space Urban Sensing Technology

- Physical Urban Space
- **Human Activity Space**
- Social-economic Space
- Cyber Urban Space


Human activities


Human mobility

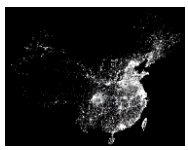
Sensing human activities or behaviors with spatio-temporal big data

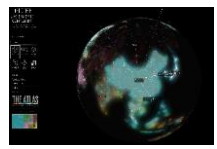
Understanding human dynamic distribution, mobility, and living behaviors with mobile phone or smart card data.

Pan-Space Urban Sensing

Pan-space Urban Sensing Technology

- Physical Urban Space
- Human Activity Space
- **Social-economic Space**
- Cyber Urban Space


China economic census


World economic trade with China

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

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

Pan-Space Urban Sensing

Pan-space Urban Sensing Technology

- Physical Urban Space
- Human Activity Space
- Social-economic Space
- **Cyber Urban Space**


Internet social network

Human emotions in Twitter

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

2 Big Data in Urban Studies

Research on Urban Sensing

Pan-Space Urban Sensing Research

Physical Urban Space	Model Dynamic Urban Structure
Human Activity Space	Sensing Chinese Dynamics
Social-economic Space	Social Sensing Urban Vibrancy
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 > **50 TB**.

Parcel level | AOI level | Building level | POI level

Data Integration / Fusion

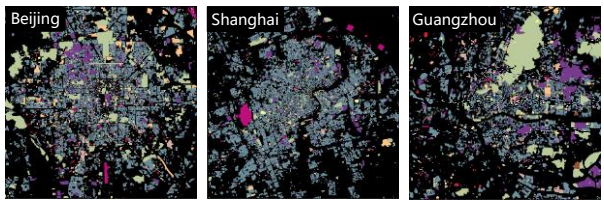
Data covers main cities in China. Pan-space (physical, human, social-economic, and cyber space) data were integrated or fused, including but not limited to,

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

Detailed Urban Land Use Classification

Problem: Remotely sensed imagery hard to tell the land use functions in urban area.

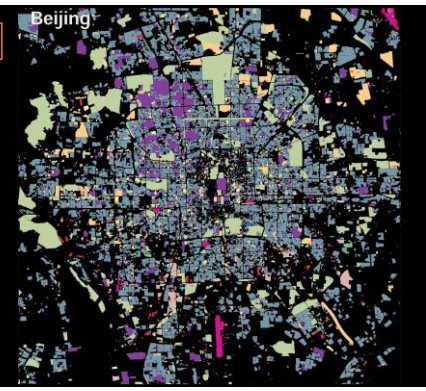
- Inferring urban detailed land use functions through pan-space urban information with **Tensor-based classification**



Urban Organs

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

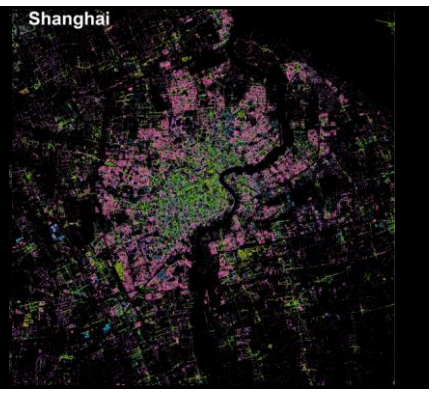
- 20 classes
- Overall accuracy: 91%



Urban Cells

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

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



Urban Pulse

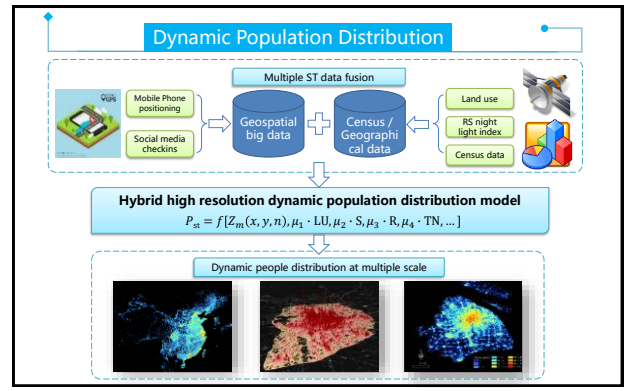
Sensing and forecasting urban traffic congestion with GPS equipped floating vehicles data in urban arteries.

Algorithm: Bayesian deep learning with spatio-temporal correlations

Forecasting accuracy: 85%

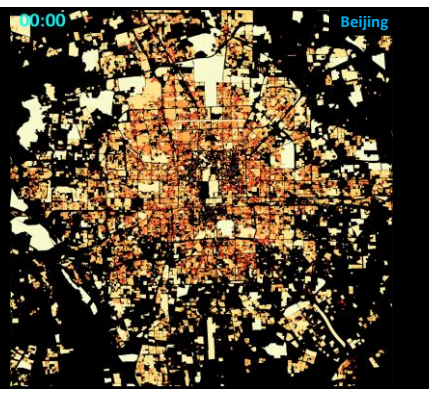
Human Activity Space

Sensing Chinese Dynamics



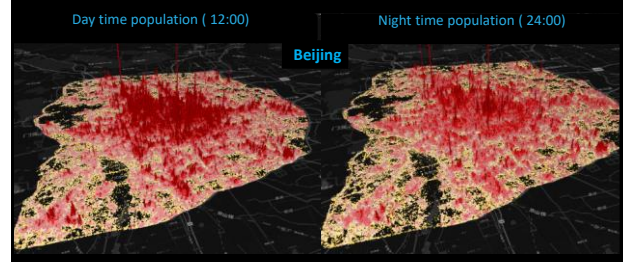
Urban Rhythm

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

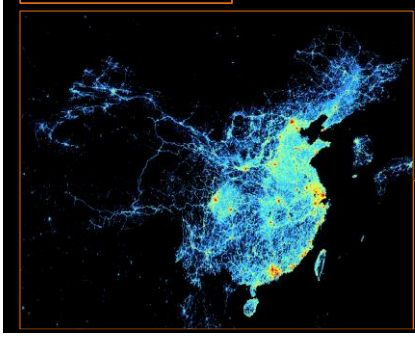


Urban Rhythm

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



Chinese Dynamics

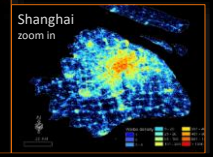


腾讯大数据 Tencent Big Data

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

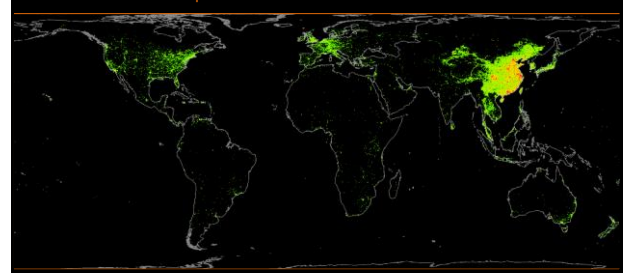
800 million users / day

50 billion location request / day



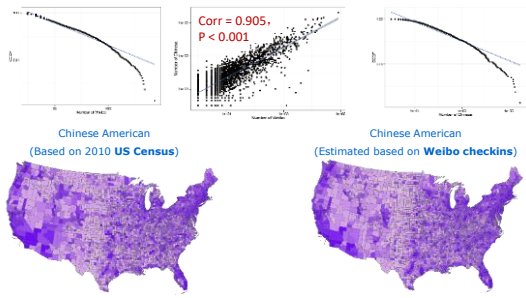
Chinese Footprint

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



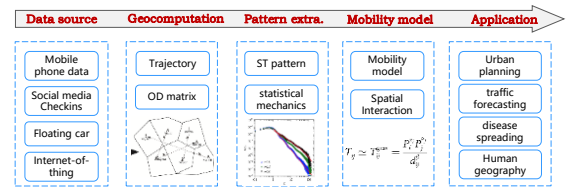
Social media data representativeness?

The Chinese American estimation based on Weibo checkins was validated with US Census in county level.



Model Chinese Mobility with Big Data

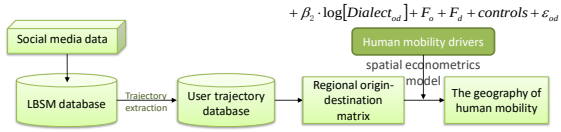
- Understanding individual human mobility is of fundamental importance for many applications from urban planning to disease spreading and traffic forecasting
- We propose social media (Weibo) checkins as a proxy for human mobility, as it relies on publicly available data and provides high resolution 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 discoveryed 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.

$$\log(T_{od}/L_o) = \beta_1 \cdot \log[Commuting_{od}] + \beta_2 \cdot \log[Dialect_{od}] + F_o + F_d + controls + \varepsilon_{od}$$

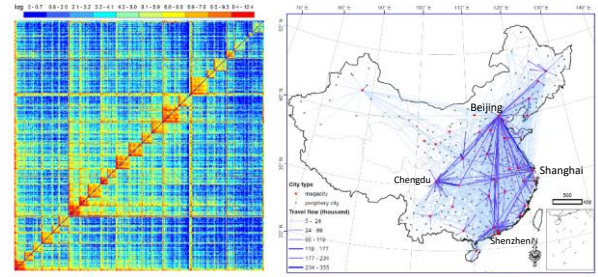


The Geography of Cultural Ties and Human Mobility: Big Data in Urban Contexts, *Annals of AAG*, 2016
The Geographical Legacies of Mountains: Impacts on Cultural Difference Landscapes, *Annals of AAG*, 2017

Understanding Chinese Mobility from Weibo

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

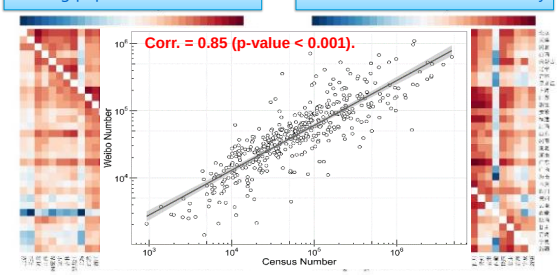
City-level Chinese mobility in geographical space



Cross-validation

Floating population in 6th Census

Weibo estimated human mobility



Model Chinese Mobility

- Cross-sectional OLS regression (in logs), city pairs

$$\log(T_{od}/L_o) = \beta_1 \cdot \log[Commuting_{od}] + \beta_2 \cdot \log[Dialect_{od}] + F_o + F_d + controls + \varepsilon_{od}$$

T_{od}/L_o : mobility flows between city pairs/total Weibo user flows of the origin city

- [Commuting]: commuting distance&time-(pecuniary mobility costs)
- [Dialect]: dialect distance- (non-pecuniary mobility costs)
- Fr: origin city fixed effect
- Fs: destination city fixed effect
- IV: historical dialect distance

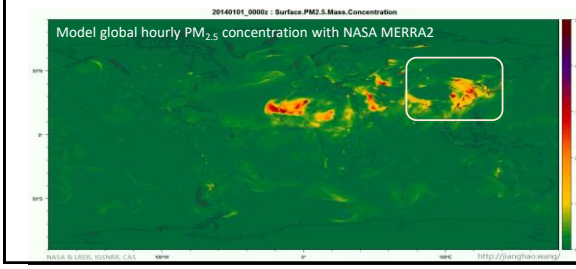
Cyber Urban Space

Citizen Emotion Sensing and $PM_{2.5}$

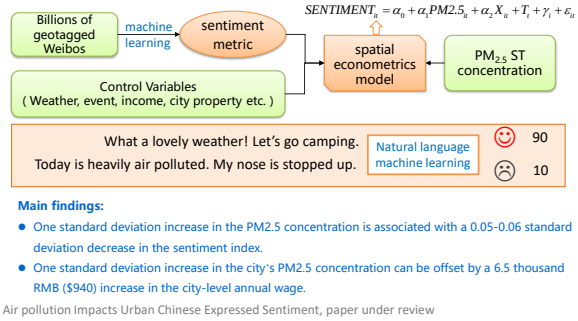
$PM_{2.5}$ Air Pollution in China

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

Study health impact ---> Measure social cost ?



Citizen Emotion Sensing and $PM_{2.5}$



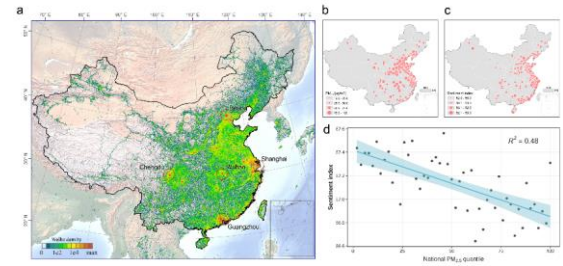
Main findings:

- One standard deviation increase in the PM2.5 concentration is associated with a 0.05-0.06 standard deviation decrease in the sentiment index.
- One standard deviation increase in the city's PM2.5 concentration can be offset by a 6.5 thousand RMB ($940) increase in the city-level annual wage.

Air pollution Impacts Urban Chinese Expressed Sentiment, paper under review

Emotion Sensing and Air Pollution

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



3 What can Urban Researcher Learn?

Lessons from Big Data

Urban Data Scientist in Action

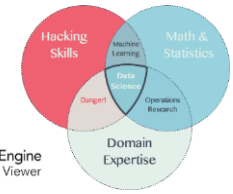
Acquisition python php Java

Database PostGIS PostgreSQL CouchDB Redis Cassandra mongoDB membase riak

Data Analysis OSGeo R

Visualization MapBox JavaScript CityEngine Web Viewer

- Domain Expertise
- Math & Statistics
- Computing skill

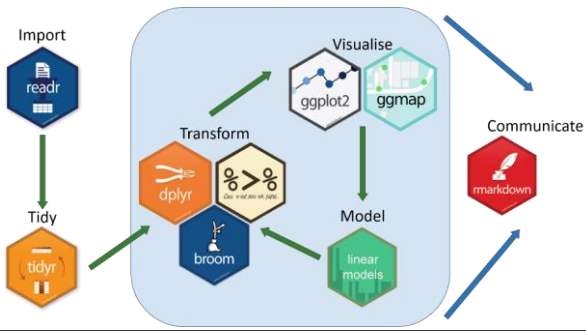


Urban Data Scientist

- Fundamentals
- Statistics
- Programming
- Machine Learning
- Text Mining
- Visualization
- Big Data
- Data Ingestion
- Data Munging
- Toolbox

Urban Data Scientist in Action

R in Action

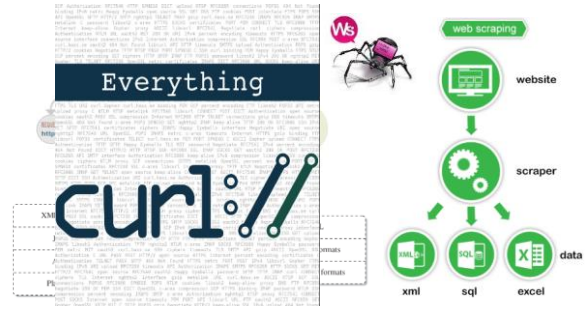


Python in Action

Data Collection

Acquisition and Management

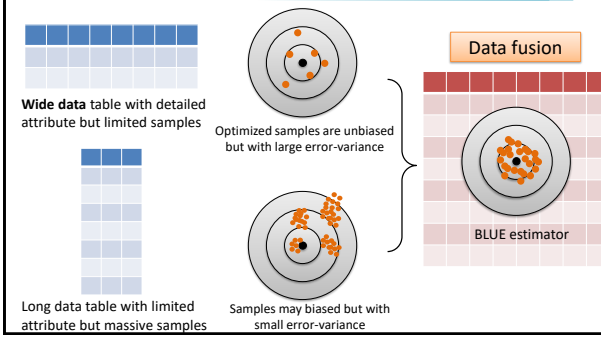
Web data Scraping



Remote, Social and Urban Sensing

	Remote Sensing	Social Sensing	Urban Sensing
Organization	Top-down ↓	↓ & ↑	Bottom-up ↑
Data accessibility	☺	☹	😐
Data representation	😐	☺	😐
Spatio-temporal granularity	☹	😐	☺
Model independence	😐	☹	☺

Small data and Big data

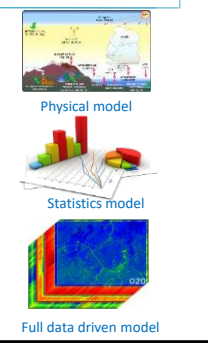


Data Analysis

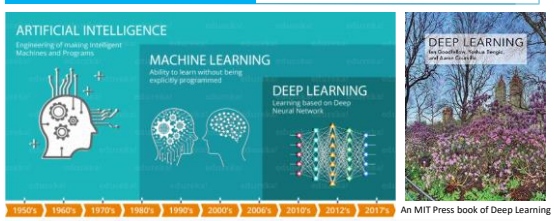
Model, Computing and Analyst

Data Driven Modelling

- Separate model
  - Spatial analysis
  - Time series analysis
- Spatio-temporal model
  - Statistical model
  - Physical model for specific field
  - Statistical + physical model
- Spatio-temporal data driven model
  - Machine learning
  - Statistical deep learning
  - Artificial intelligence algorithms



AI, ML, DL



An MIT Press book of Deep Learning

Caffe Chainer DL4J KERAS CNTK MatConvNet MINERVA mxnet Purine TensorFlow theano torch

Big Data in Earth Observation

WSN Wireless Sensor Networks

EOS RS, UAV, Model

Geospatial Analysis in the Cloud

Google Earth Engine

https://earthengine.google.com/

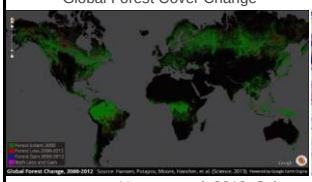
> 5 Petabytes of Earth observation data (imagery, weather, etc.)

Landsat
Sentinel
MODIS
ASTER
EO-1
DMSP-OLS

Topography
Landcover
Weather
Atmosphere
Population

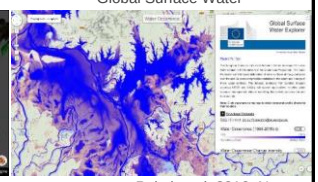
Geospatial Analysis in the Cloud

Global Forest Cover Change



Hansen et al. 2013. Science

Global Surface Water



Pekel et al. 2016. Nature

This is the first map of forest change that is globally consistent and locally relevant. What would have taken a single computer 15 years to perform was completed in a matter of days using Google Earth Engine computing.

- Professor Matt Hansen, University of Maryland

Analytic

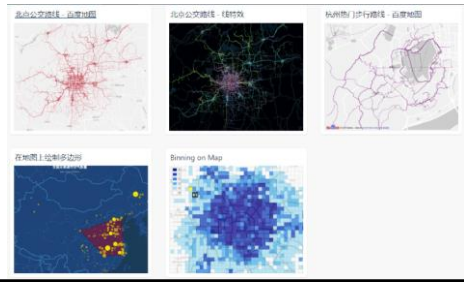
Mapping and Analysis

D3.js

Data-Driven Documents https://d3js.org/ https://www.jasondavies.com/

Echarts 3

ECharts 3 http://echarts.baidu.com/



Uber Visual

DECK.GL REACT-MAP-GL REACT-VIS

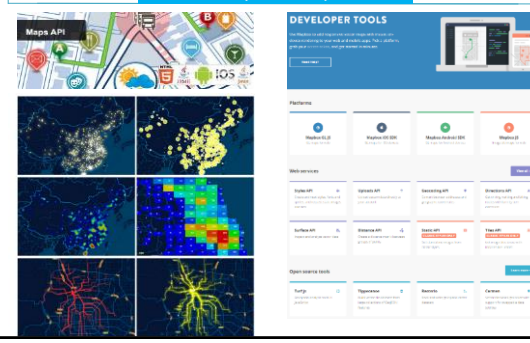
Cloud Cartography

MapBox https://www.mapbox.com/

CARTODB https://cartodb.com/

GEOHEY https://geohey.com/

Web Map Development



Deep Thinking in Action

Big data + Big model → Big unsolved question

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

Collaborators: Yong Ge, Siqi Zheng, Wenjie Wu, Ying Long, Xingjian Liu, *et al.*

Acknowledgement: The National Natural Science Foundation of China

http://jianghao.wang

wangjh@lreis.ac.cn