The content is Center Proprietary and the Terms of the Center membership agreement apply

Towards Spatiotemporal Science: A Preliminary Investigation

Chaowei Yang

NSF Spatiotemporal Innovation Center

George Mason University

杨超伟

美国国家基金会时空协同创新中心

乔治梅森大学

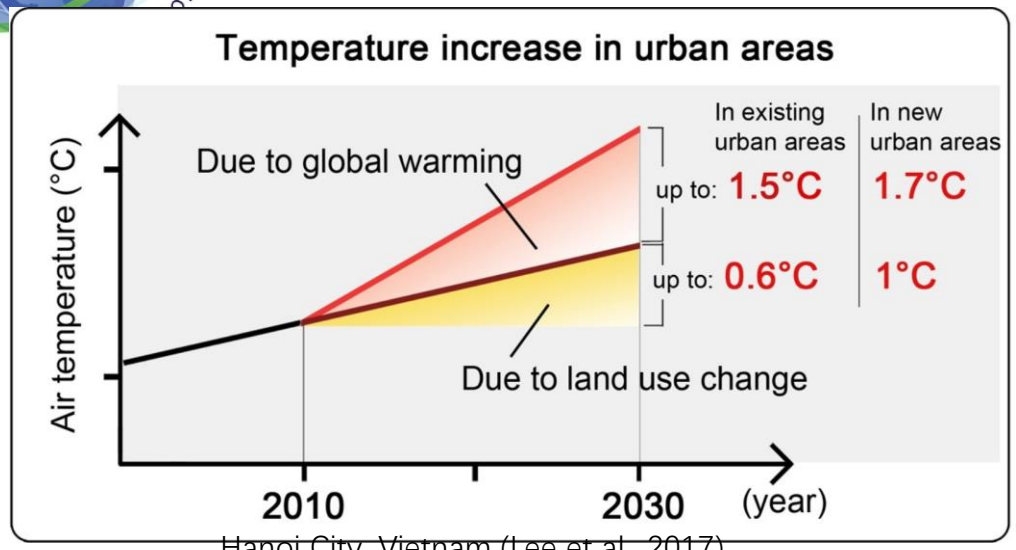
https://www.stcenter.net/

Agenda

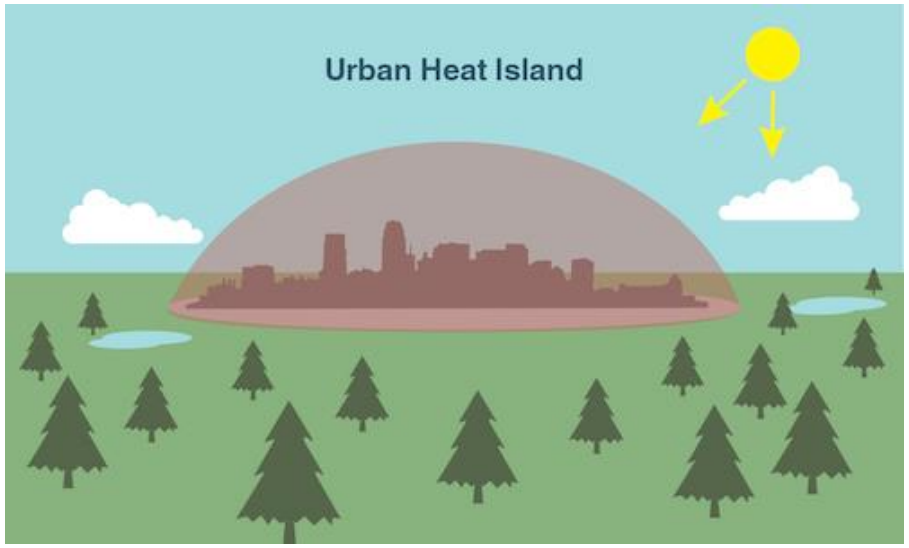
1. Global grand challenges
   1.1 Climate change, urban heat island, and air quality
   1.2 The COVID-19 Global pandemic
   1.3 Asteroids and man-made space objects
2. A formal and sustainable approach
3. Towards a science domain
4. Questions, Comments, Suggestions

1.1 Urban Heat Island

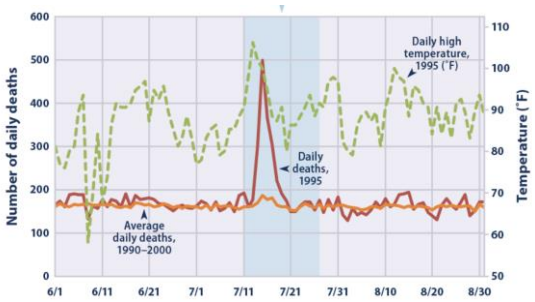
Jingchao Yang



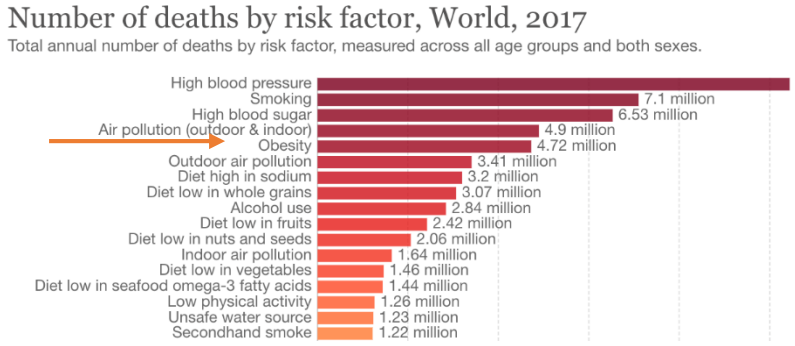
Hanoi City, Vietnam (Lee et al., 2017)



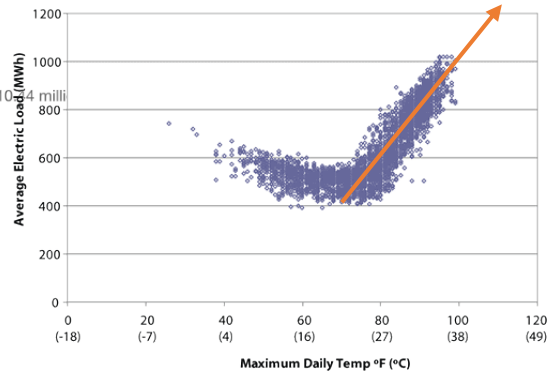
What Is an Urban Heat Island? | NASA Climate Kids



Heat wave in Cook County , Chicago
(*EPA, Climate Change Indicators: Heat-Related Deaths)*



https://ourworldindata.org/air-pollution



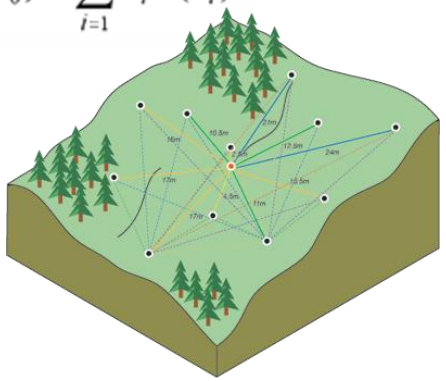
Li et al., 2018

New Orleans
(EPA, Heat Island Impacts)

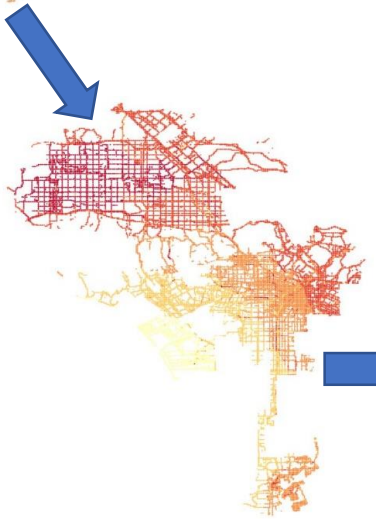
A spatiotemporal science approach

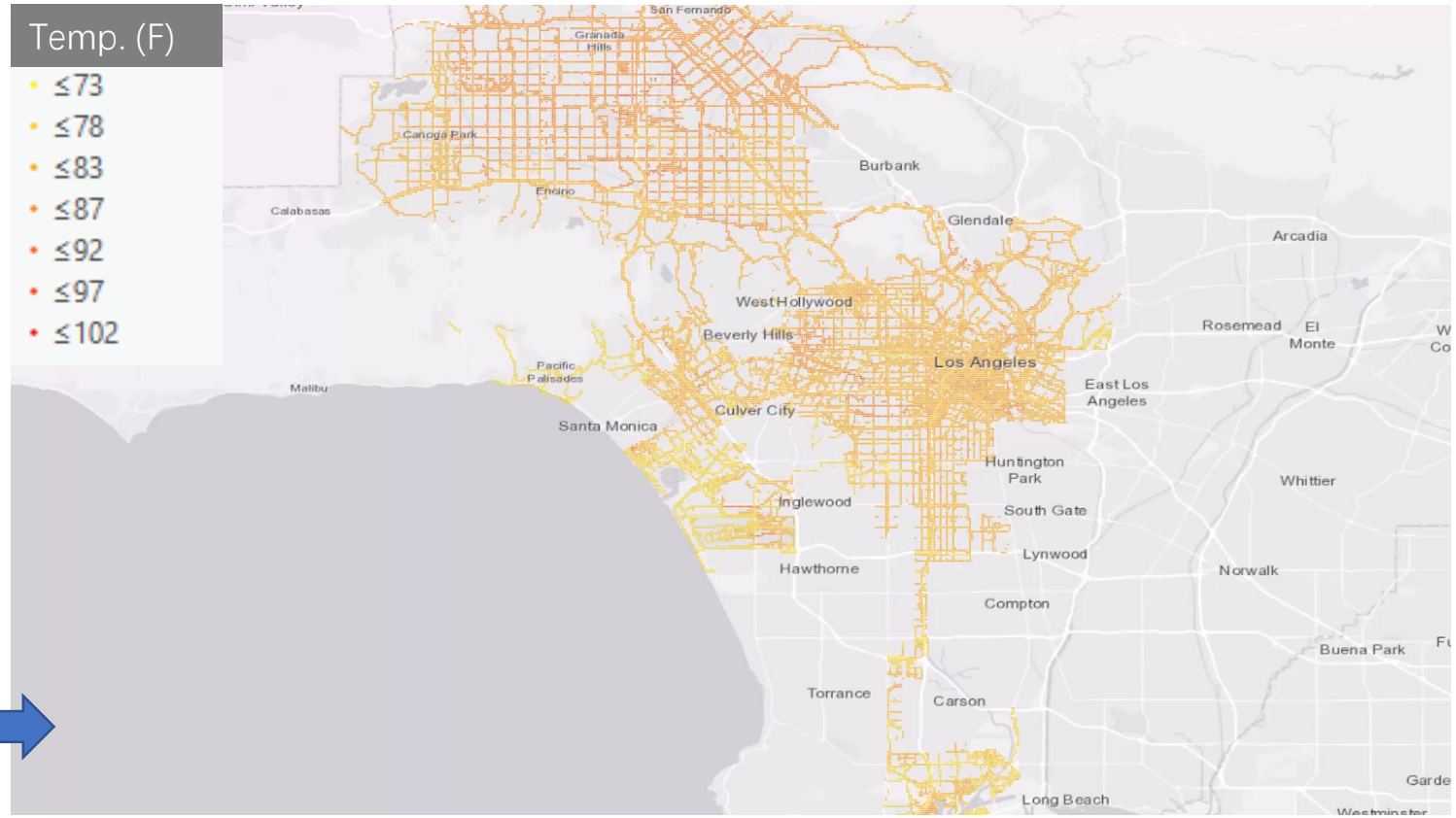
ST Kriging & many others

$$Z(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i)$$



Weights the surrounding measured values to derive a prediction for an unmeasured location





Jul 25-28, 2019 LA, CA

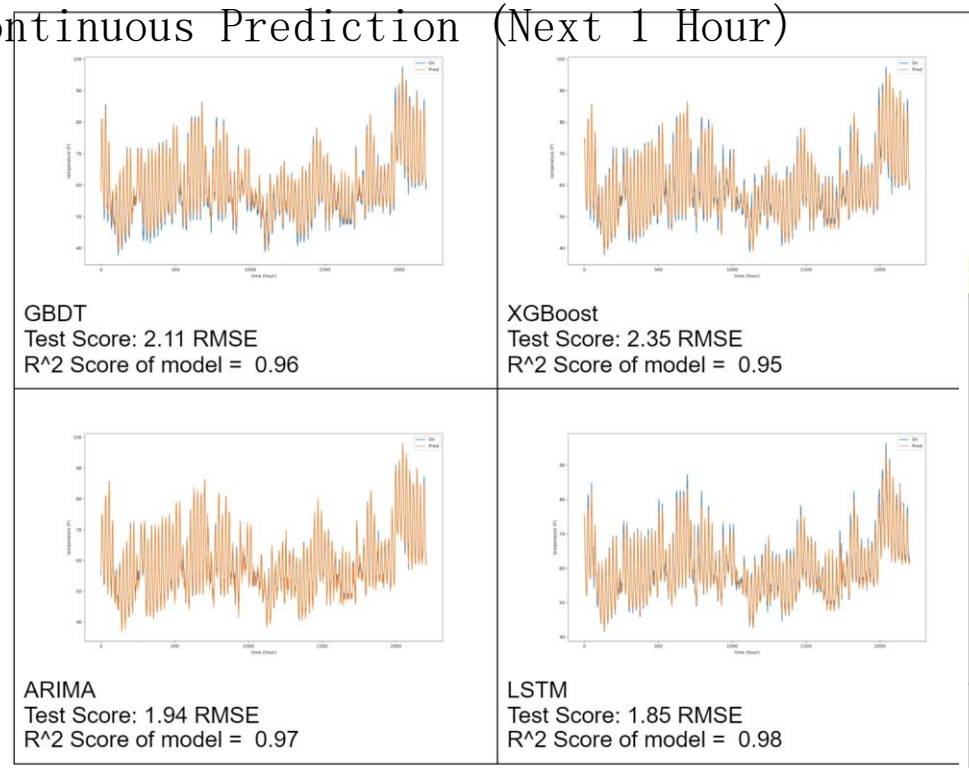
Initial Results: Model Comparison

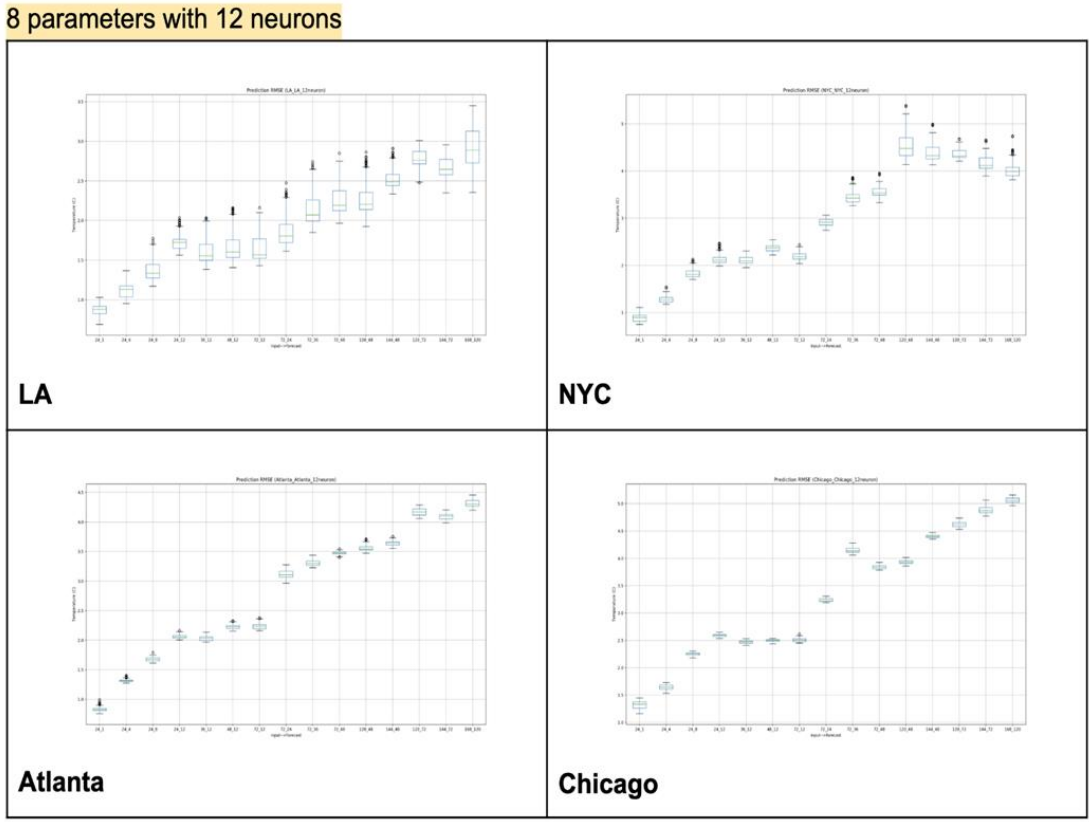
- Single-point Continuous Prediction (Next 1 Hour)

LSTM yields the best prediction result

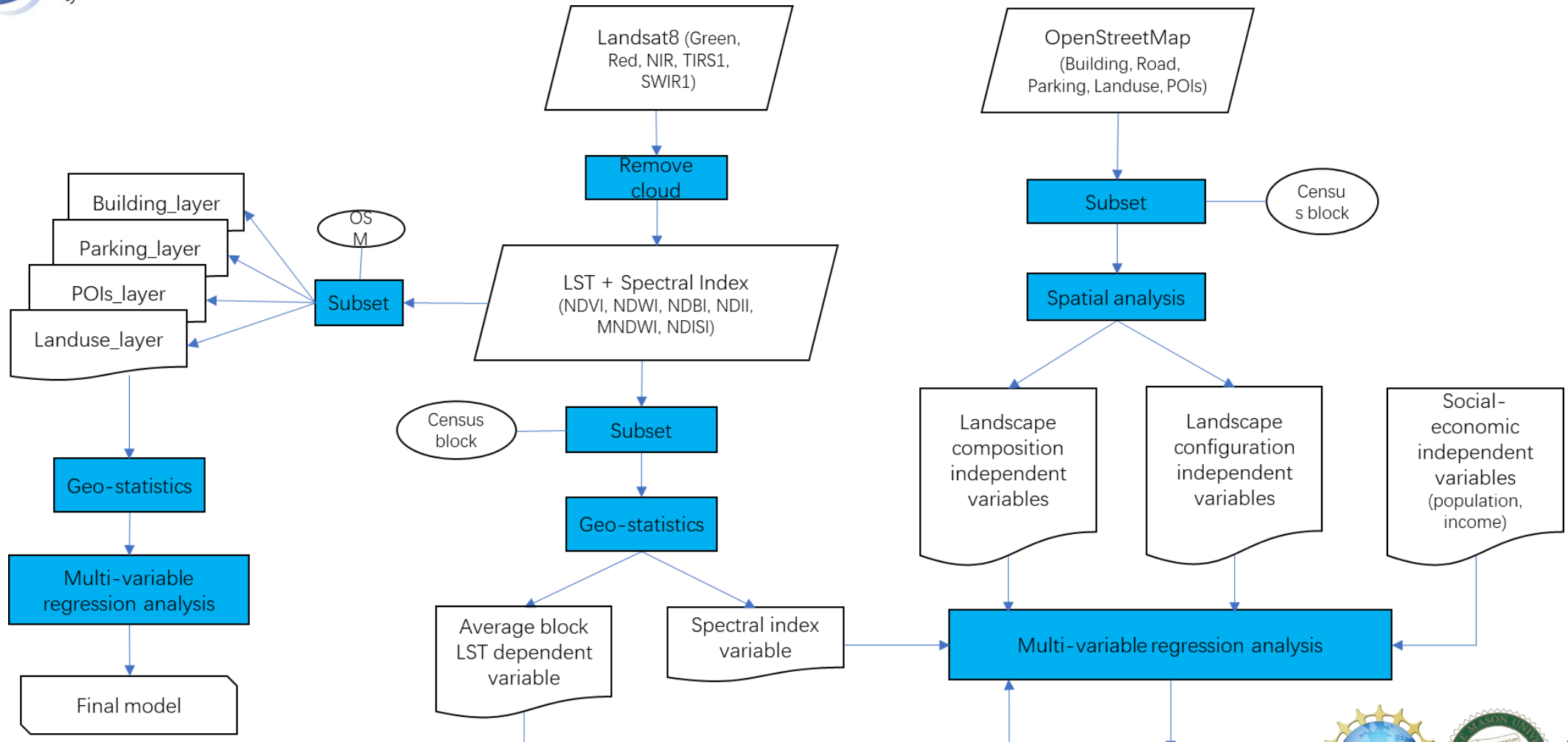
- minimum RMSE
- highest R^2

Selected for the further multi-step forecasting algorithm with multivariable.

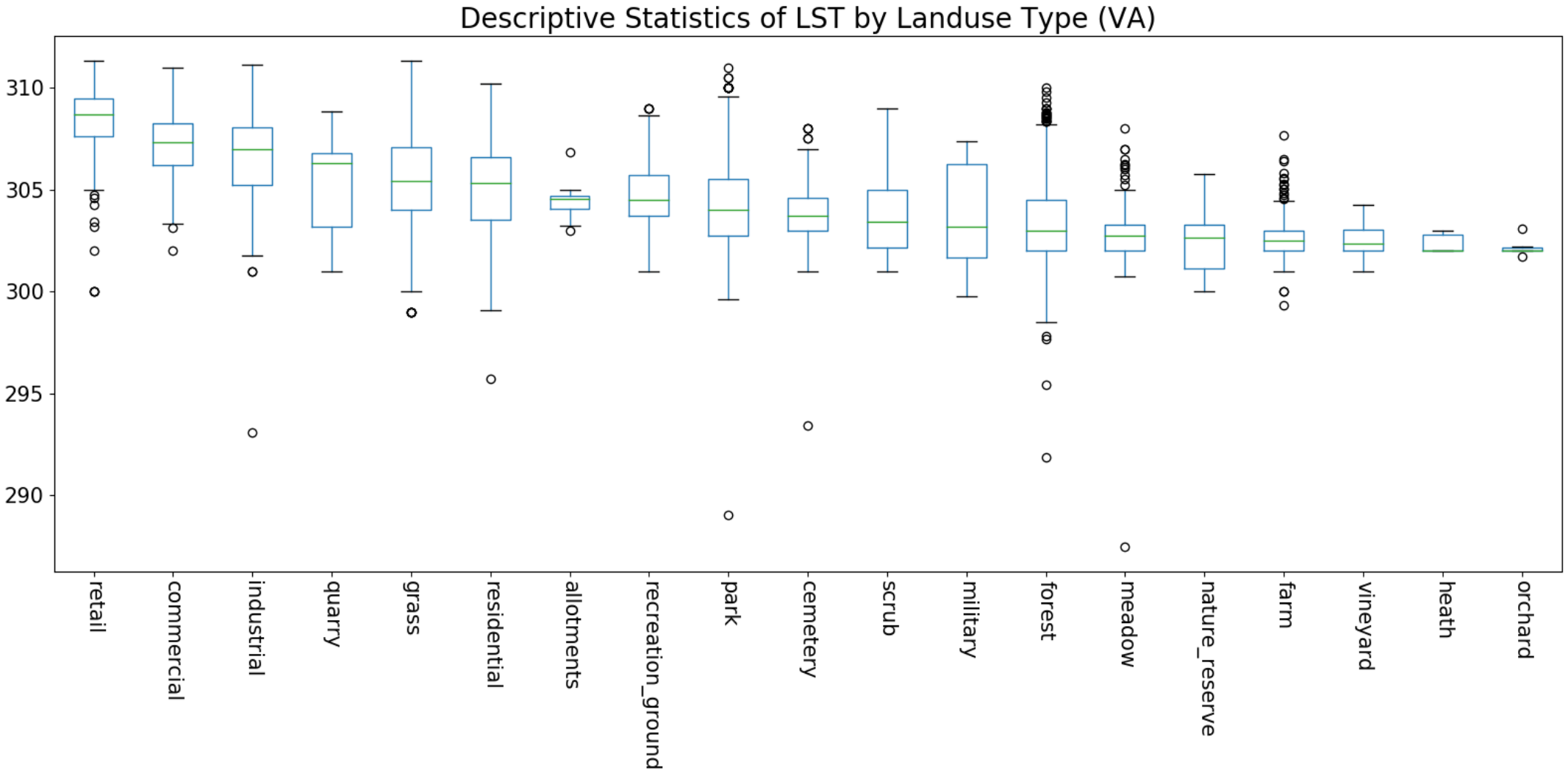




A Spatiotemporal Science Approach



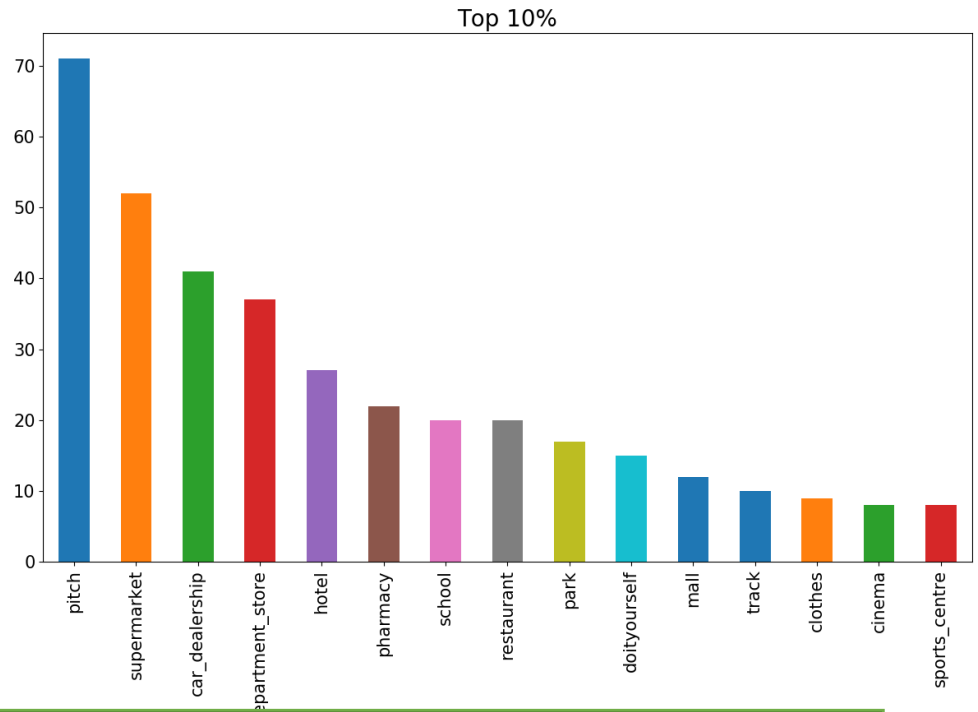
Result analysis

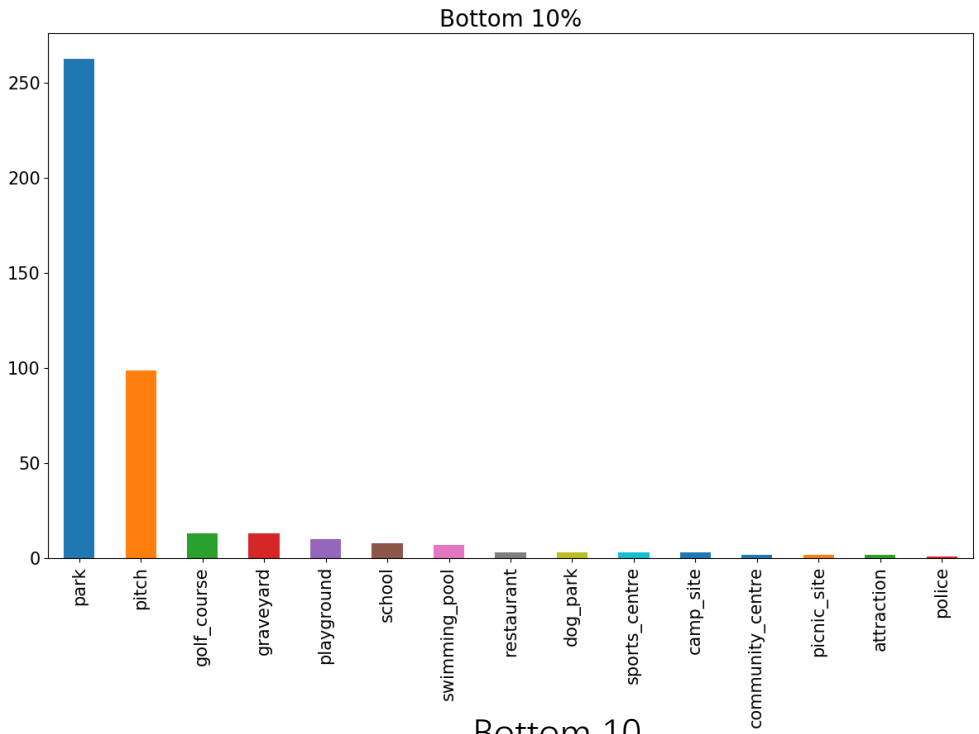


Taken VA as an example (K)

POI analysis (area > 900 m², number of POIs : 5901/11909)

Pitch,
Supermarket,
Car_dealership
Department_store
Hotel,
Pharmacy
School,
Restaurant
Park,
Doityourself





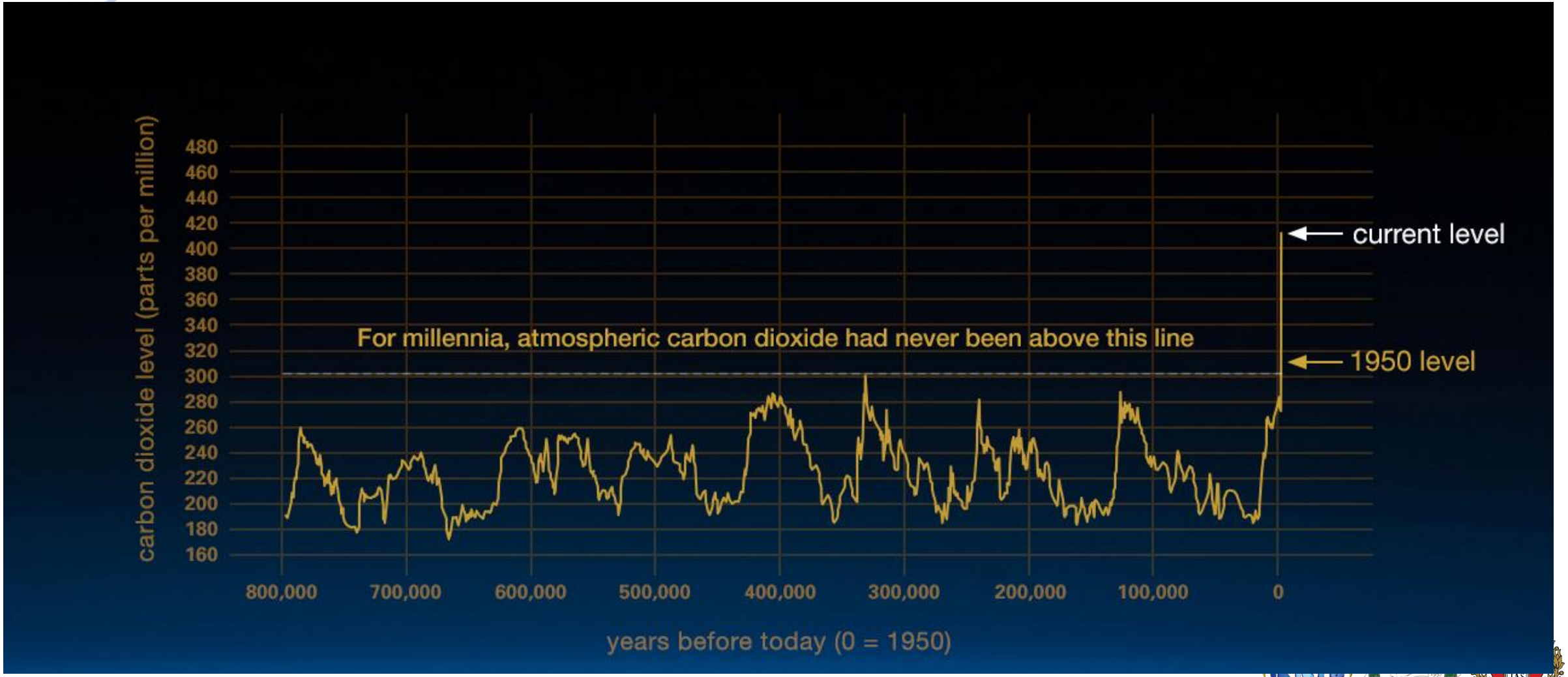
Park,
Pitch,
Golf_course,
Gaveyard
Playground
School,
Swimming pool
restaurant,
Dog_park
Sports centre
Camp_site

Solutions at the Building Level

Bottom 10

fclass	name	lst
pitch		300
park	Belmont Park	300.3
park	Meadowood Recreation Area	300.310183
park	Old Colchester Park & Preserve	300.425926
park	Pohick Bay Regional Park	300.550067
park	Huntley Meadows Park	300.730204
camp_site	Burke Lake Park Family Campground	300.893805
community_centre	Acadia Apartments	301
graveyard		301
park	Popes Head Stream Valley Park	301

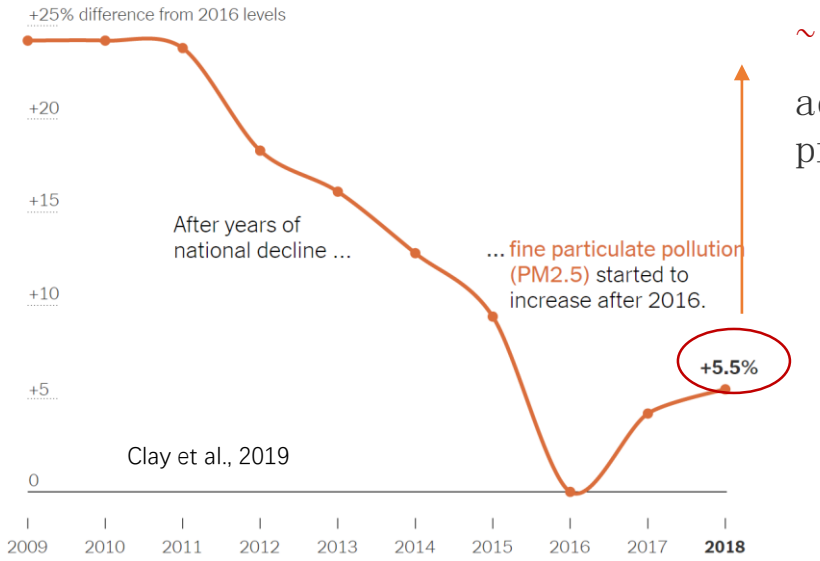
1.1.2 Climate Change & Air Quality



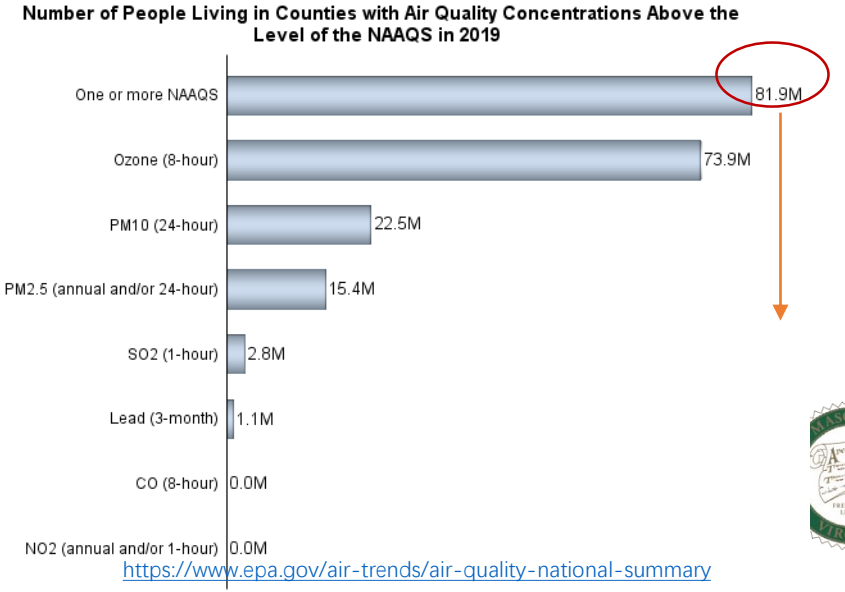
Evidence | Facts – Climate Change: Vital Signs of the Planet (nasa.gov)

Air Quality Problem: An Example of PM2.5

- PM2.5 - a major public health concern (Beelen et al., 2014)
  - 10 μg/m^3 up in $PM_{2.5}$ → 9% up in risk for lung cancer (IARC, 2013; Turner et al., 2020)
  - Traffic related air pollution levels up → lower fertility rates (Nieuwenhuijsen et al., 2014; Conforti et al, 2018).
- ~4.2 million premature deaths worldwide in 2016 (WHO)
  - ~200,000 deaths per year in the U.S. alone (Woodward and Levine, 2016)



~10,000 additional premature deaths

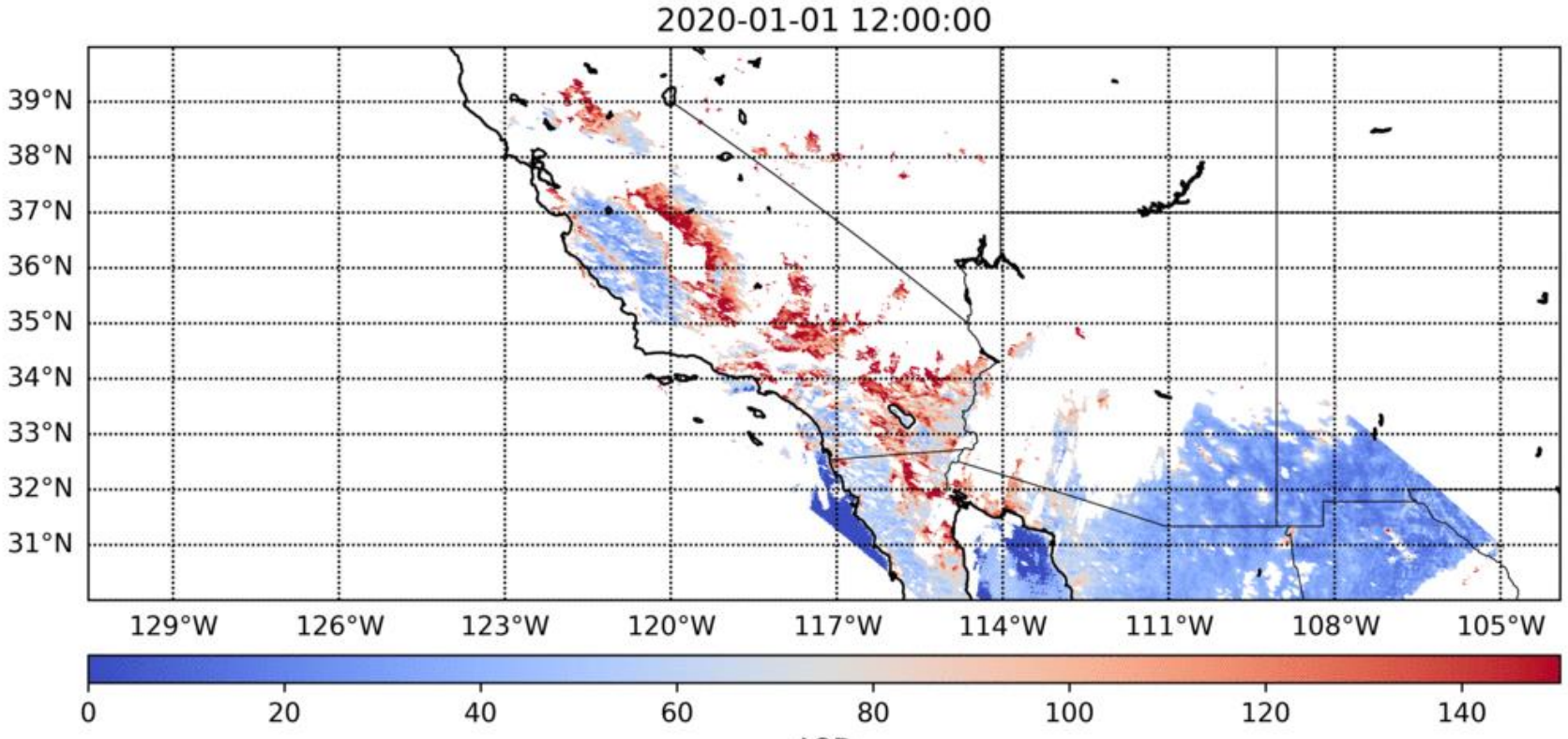


~82 million people nationwide lived in counties with pollution levels above the primary NAAQS in 2019

https://www.epa.gov/air-trends/air-quality-national-summary

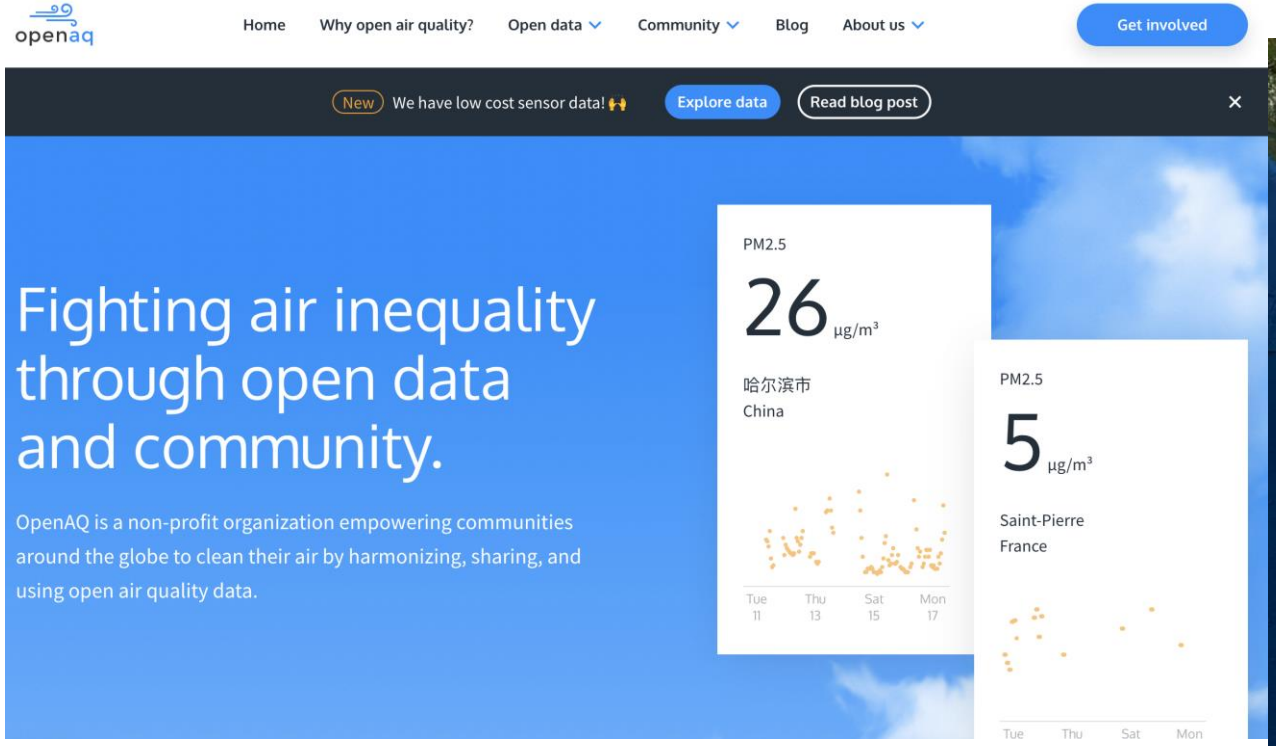
Problems

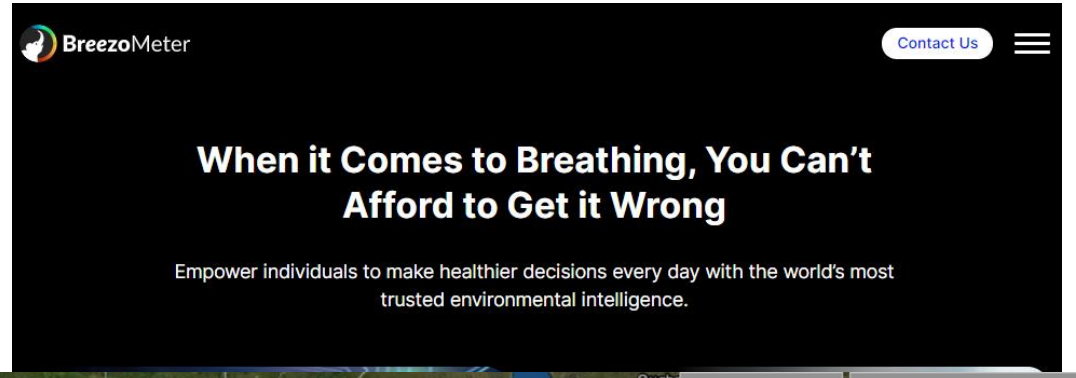
- 1. No global coverage gridded PM2.5 product (most satellite-based product are transferred from AOD and without ocean);
- 2. No international standard PM2.5 products collected from different countries and sensors.
- 3. No high spatiotemporal observations in key regions such as LA

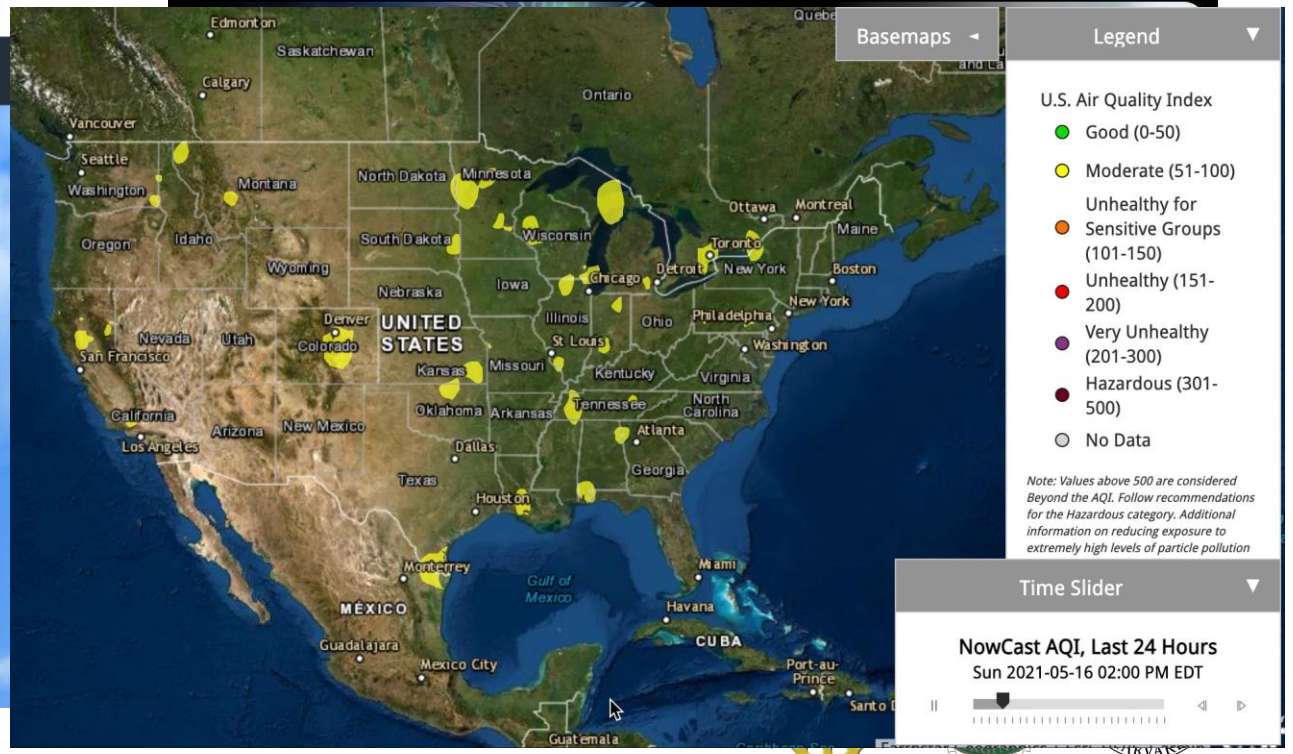


Public & Industrial Interest

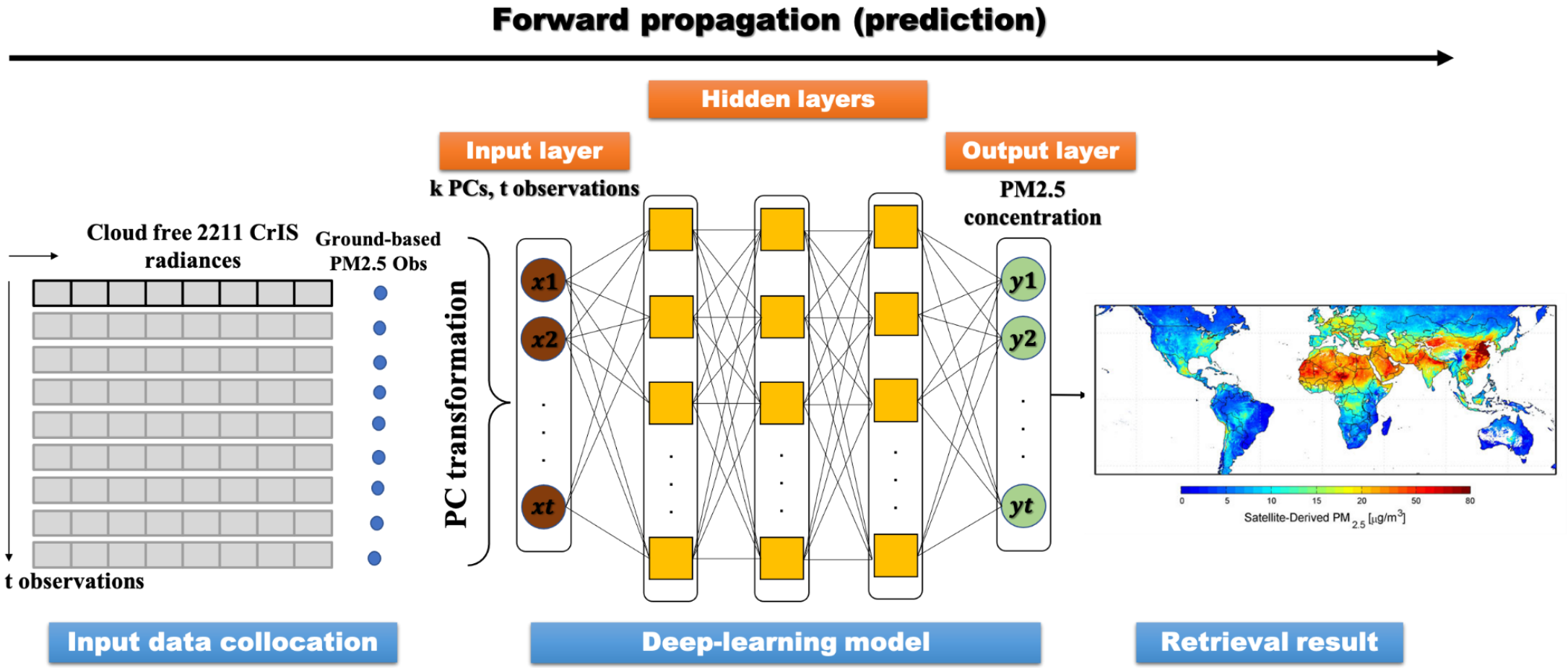
- Commercial companies leveraging air quality monitors and machine learning for real-time air quality monitors and short-term predictions



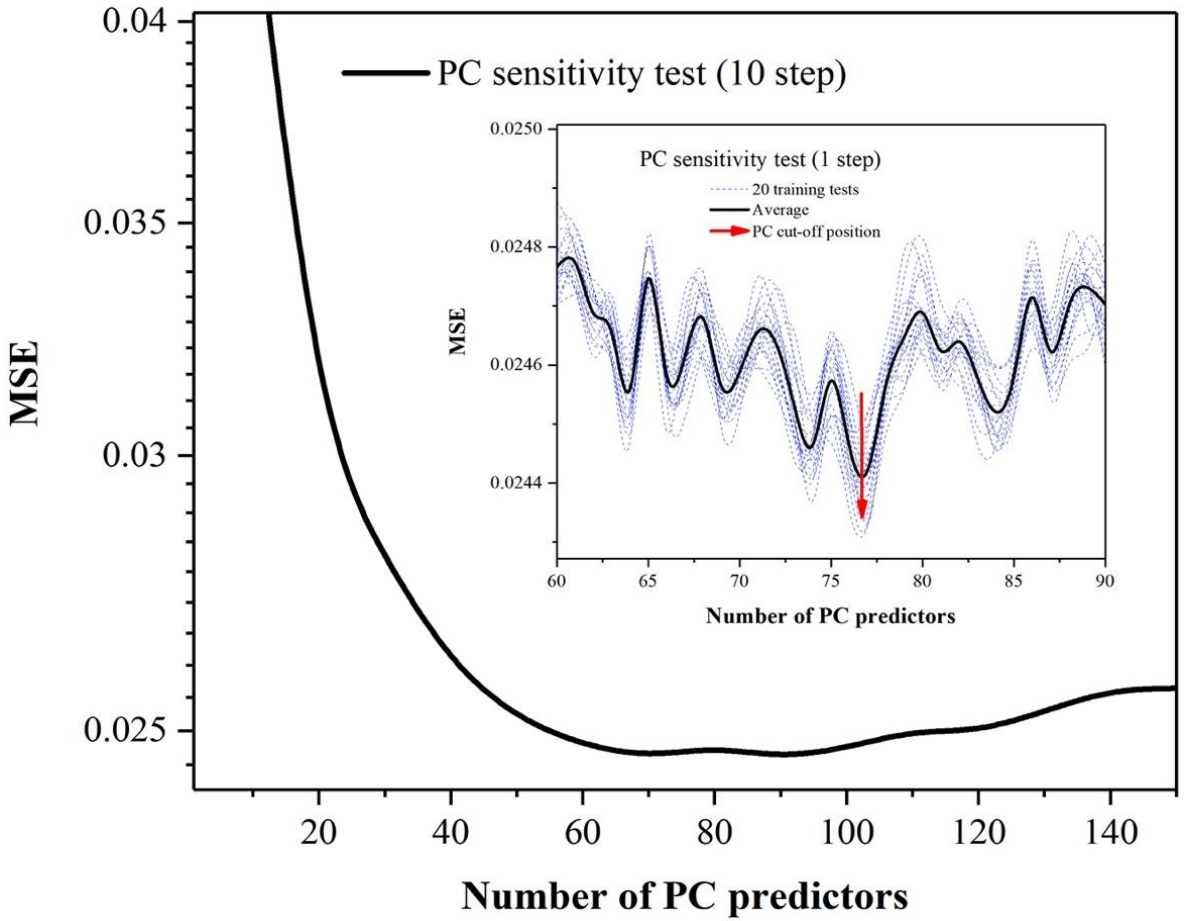




A Spatiotemporal Science Approach



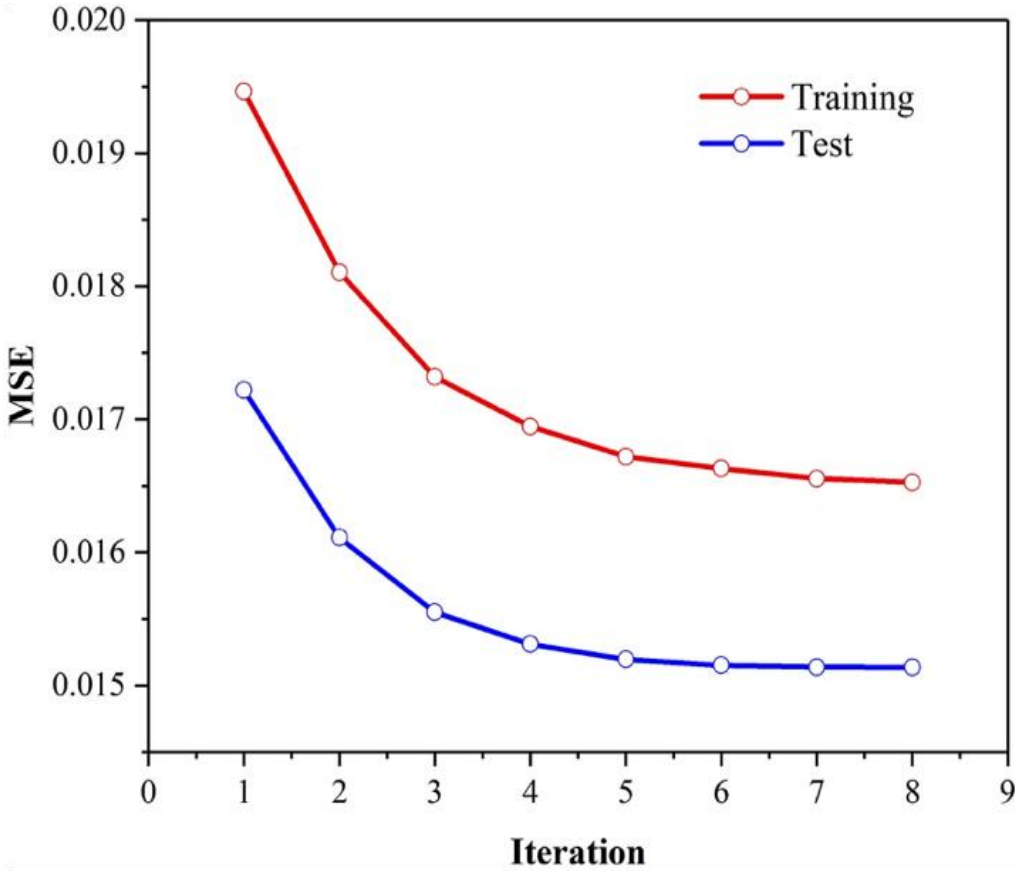
Sensitivity test: Input predictor



- Step 1: add 10 PCs each time from 10 to 150 PCs. Lowest MSE appears when PC number is between 60 and 90.
- Step 2: add 1 PC each time from 60 to 90, repeat 20 times. Average MSE become lowest when PC number is 77.

The top 77 Principal Components (PC) are used as predictors (instead of the totally 2211 channels)

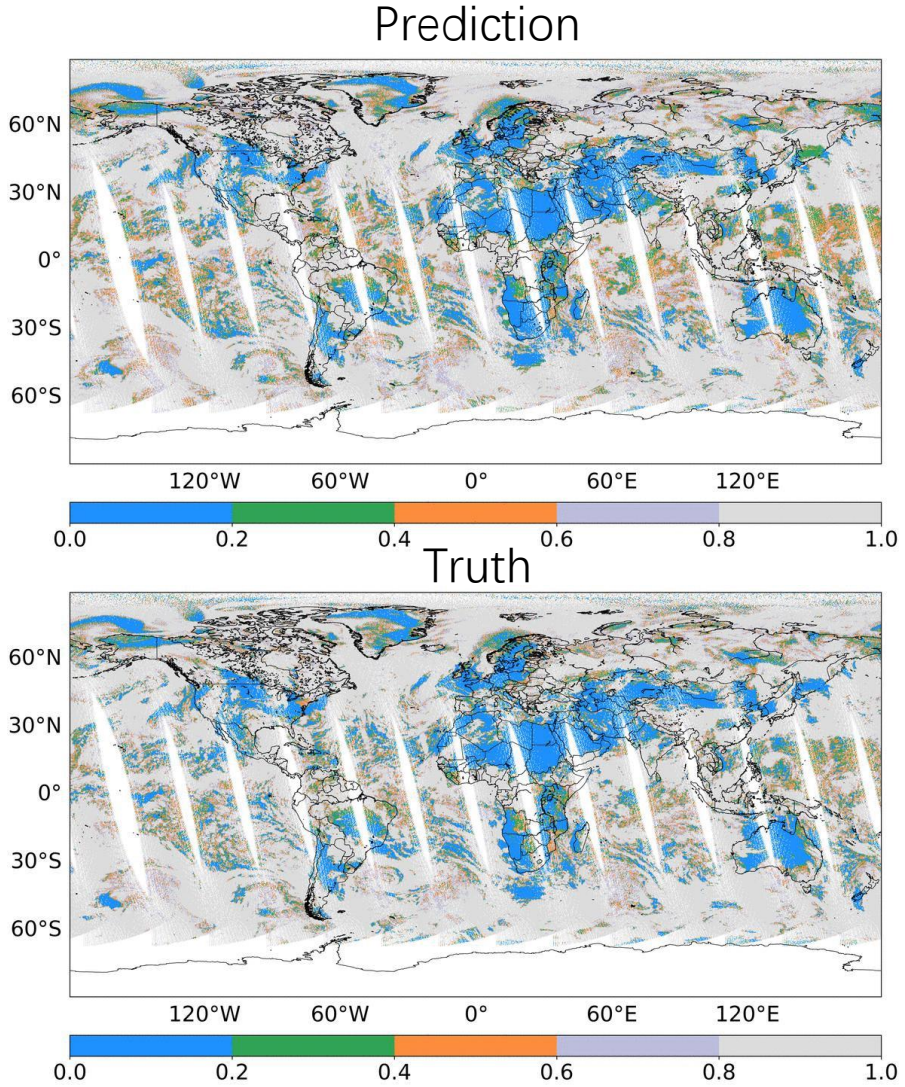
Sensitivity test: Iteration/epoch number

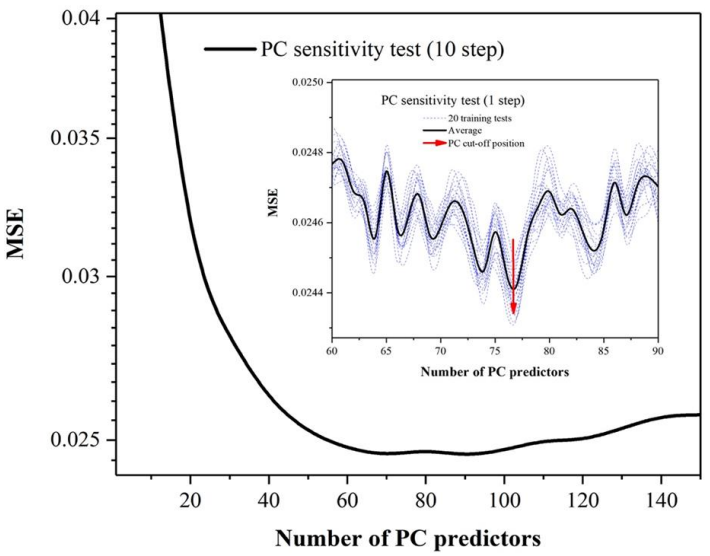


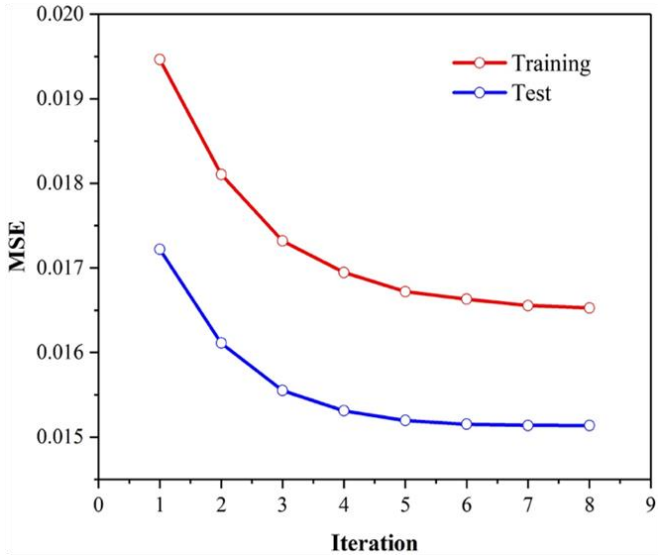
- The test accuracy does not change after 8th iteration, the training accuracy still increases.
- To avoid overfitting, training is stopped at the 60th epoch of 8th iteration.
- The training took 4-5 days, producing results is relatively fast.

Global Level Variable Retrieval Jun. 1~7 2020

Test accuracies of use cases from Jun. 01, 2020 to Jun. 07, 2020.



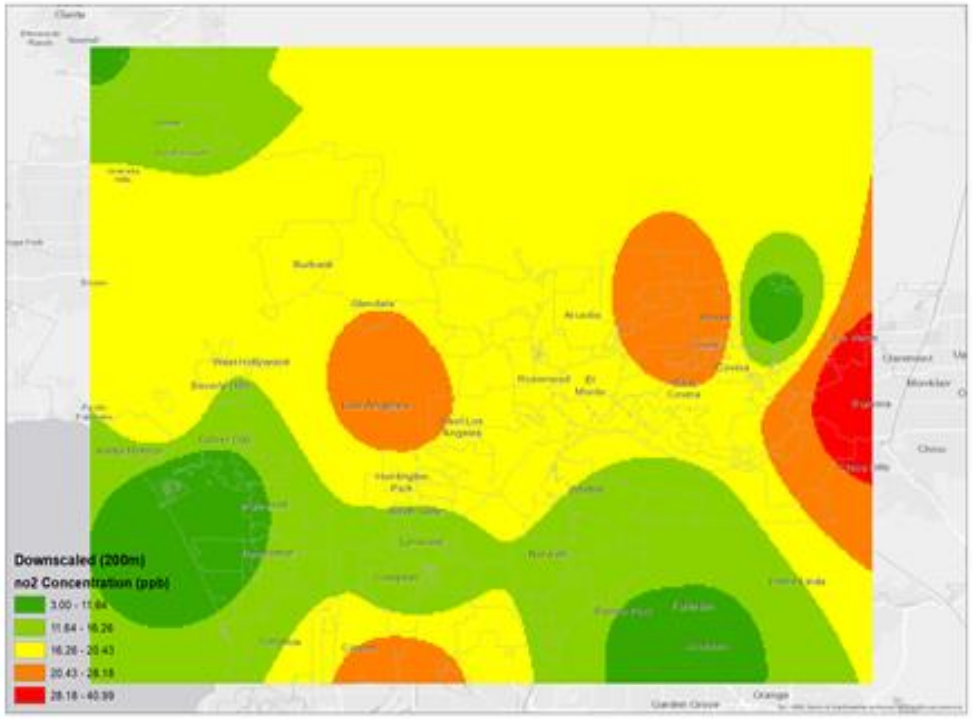




- Performance is stable and reliable for all use cases: **MSE~0.02, Pearson's R~0.92.**

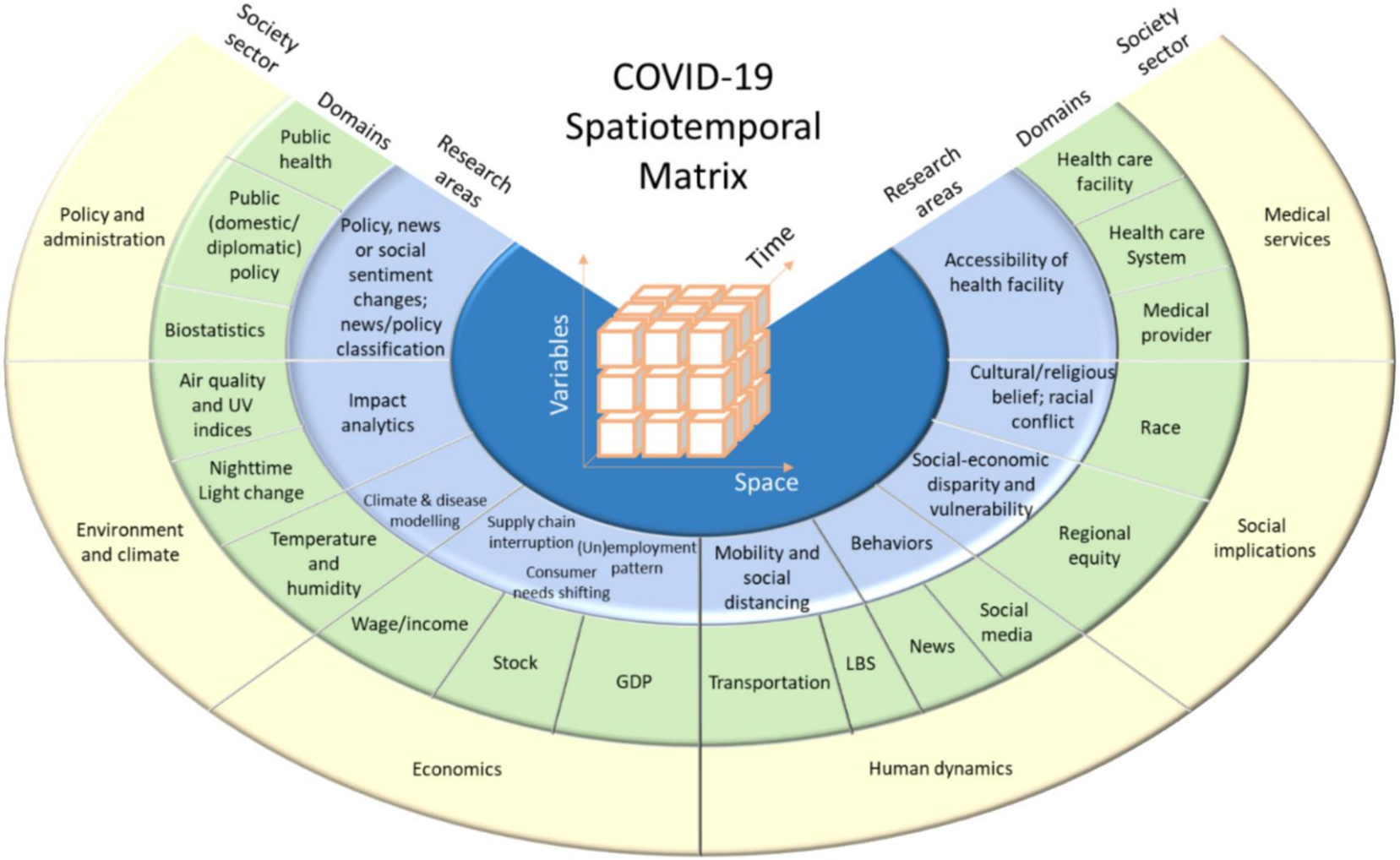
Get down to street level

(a) Original ground AQ observation



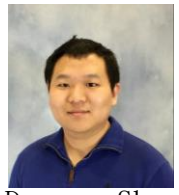
(b) Downscaled AQ

1.2 The Spatiotemporal Dimensions of the Global Pandemic

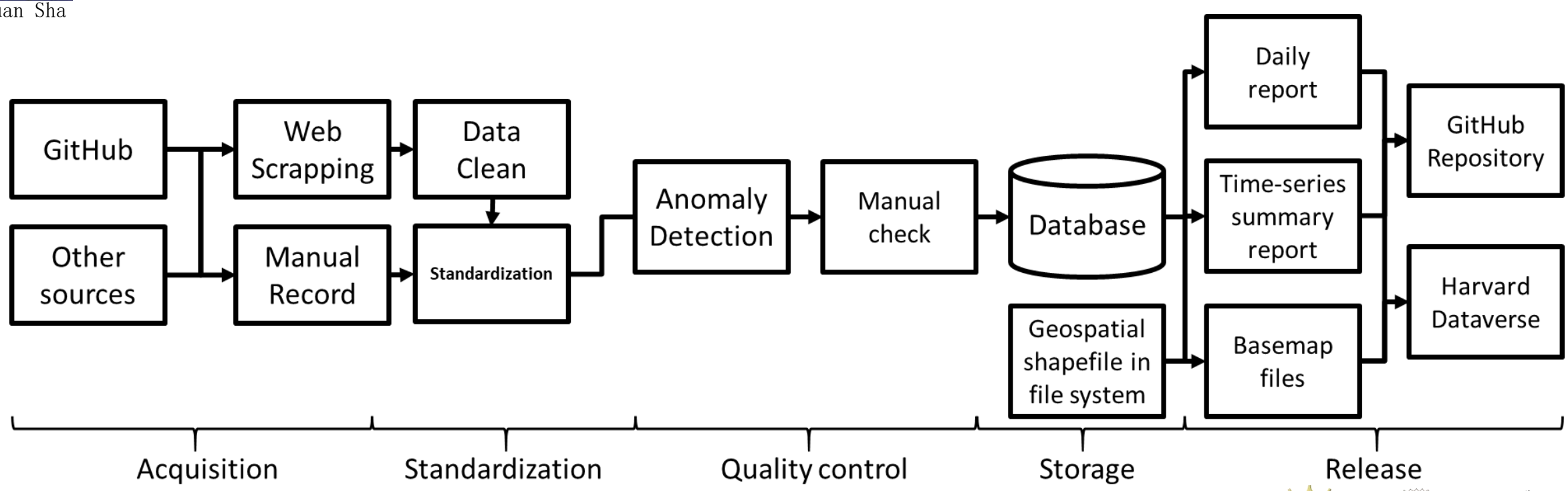


Yang et al., 2020

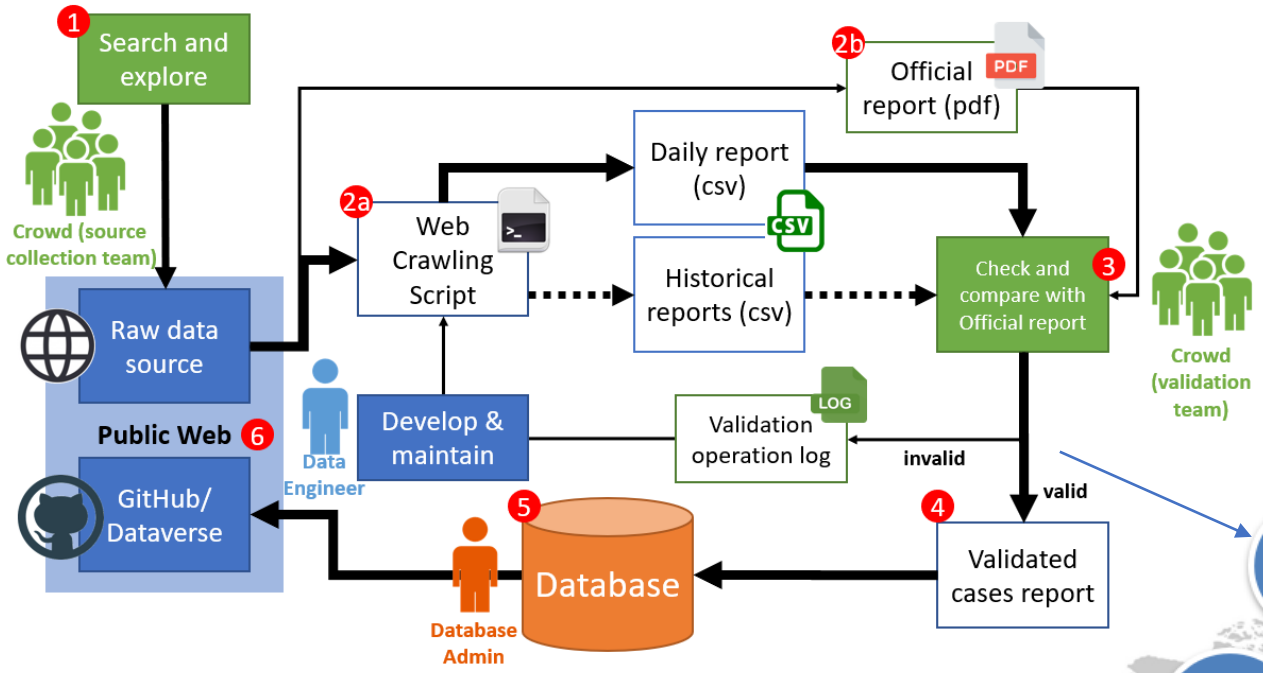
1.2.1 Spatiotemporal Data Collection, Collocation, and Visual Analytics

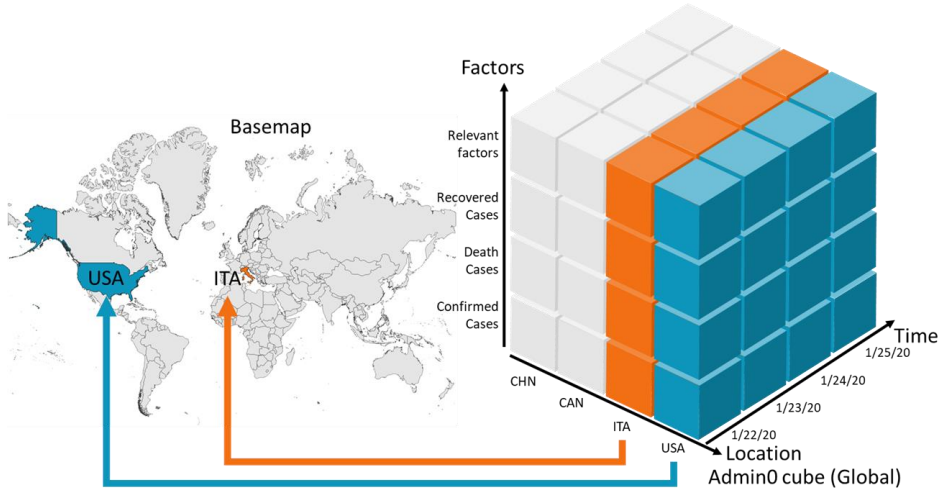


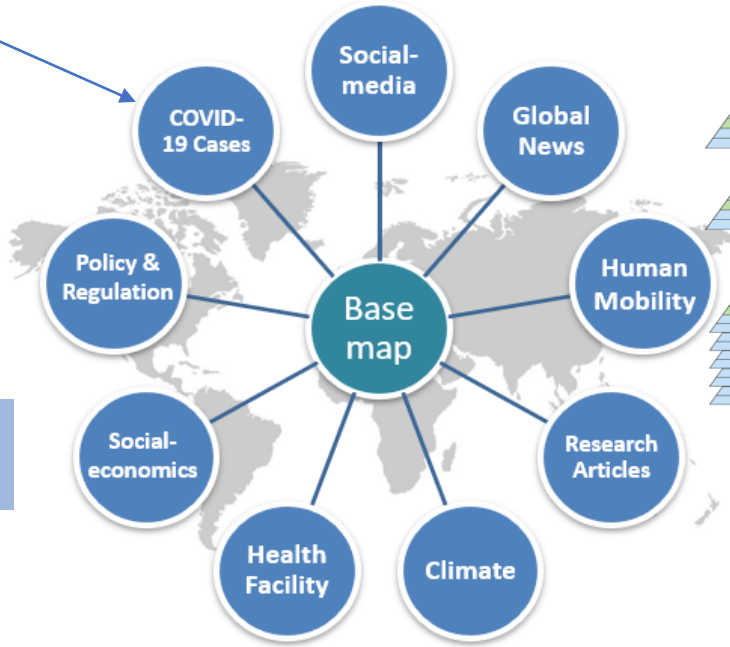
Dexuan Sha

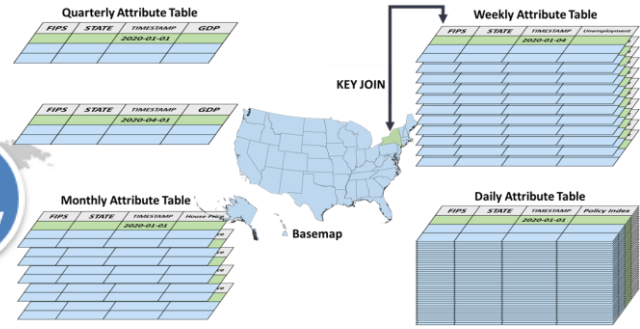


A Spatiotemporal Science Approach





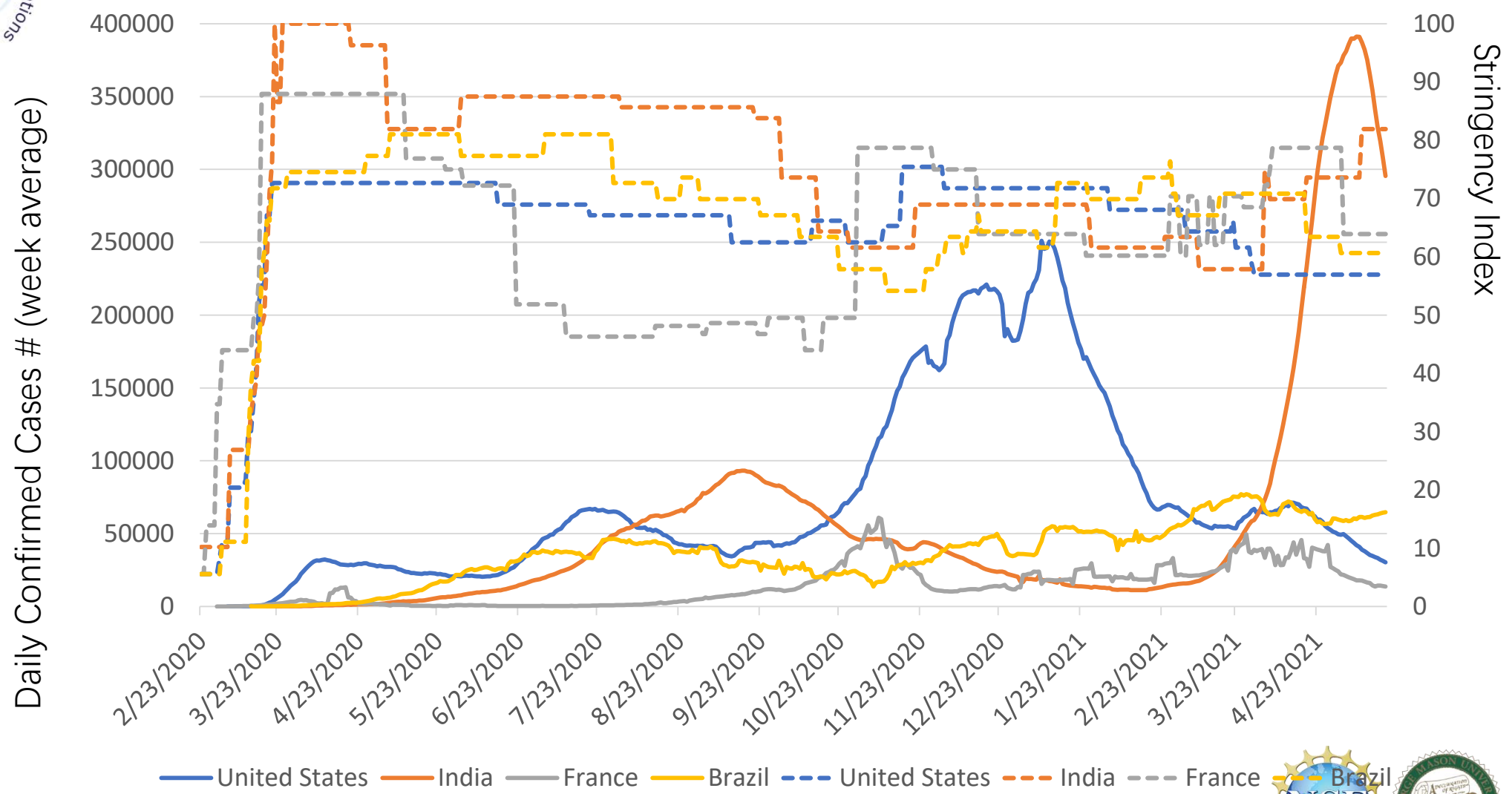




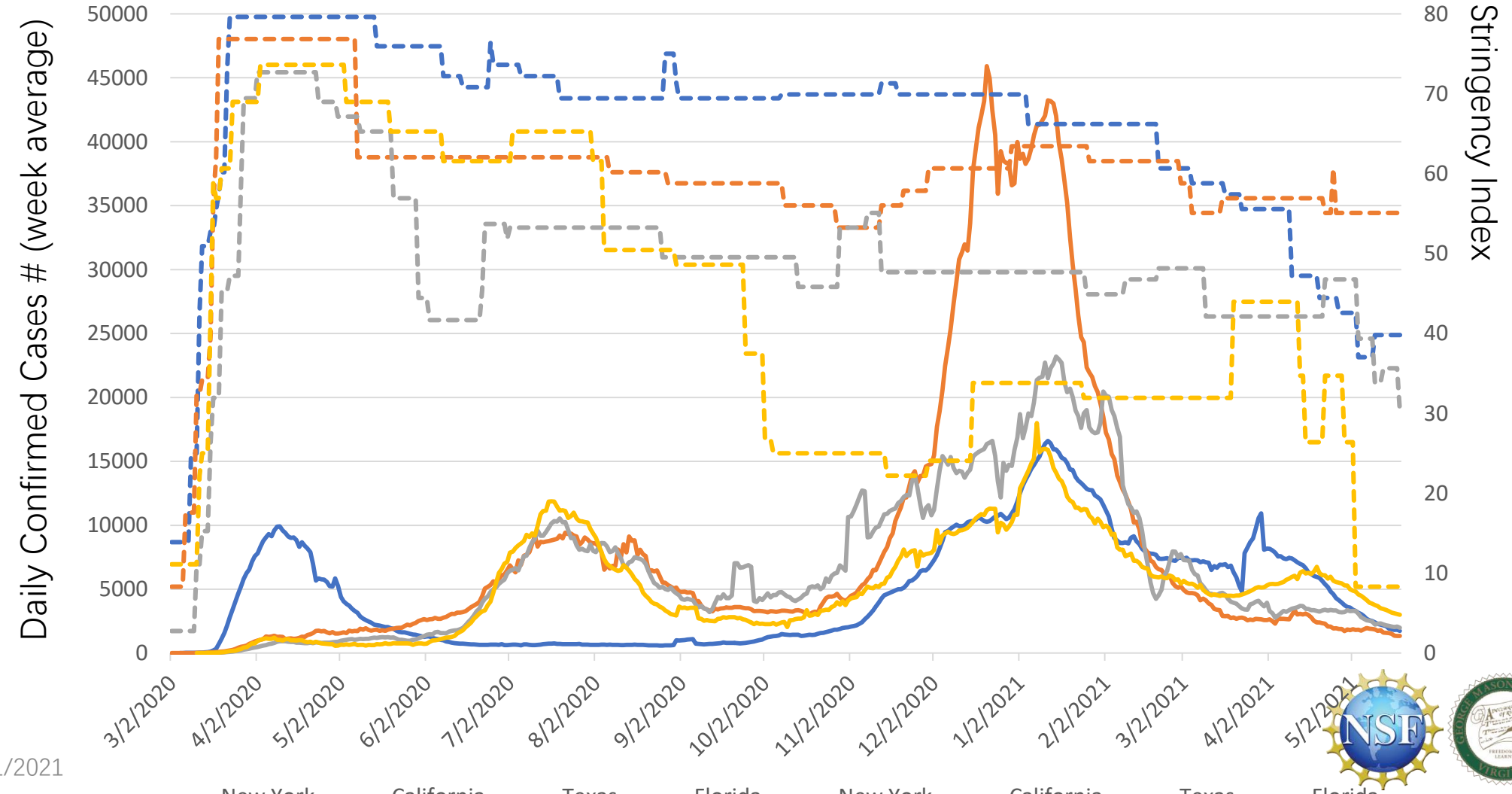
Map-based data and standard association

More reference for data description could be found here: Section 2.2. in https://www.tandfonline.com/doi/abs/10.1080/20964471.2020.1844934

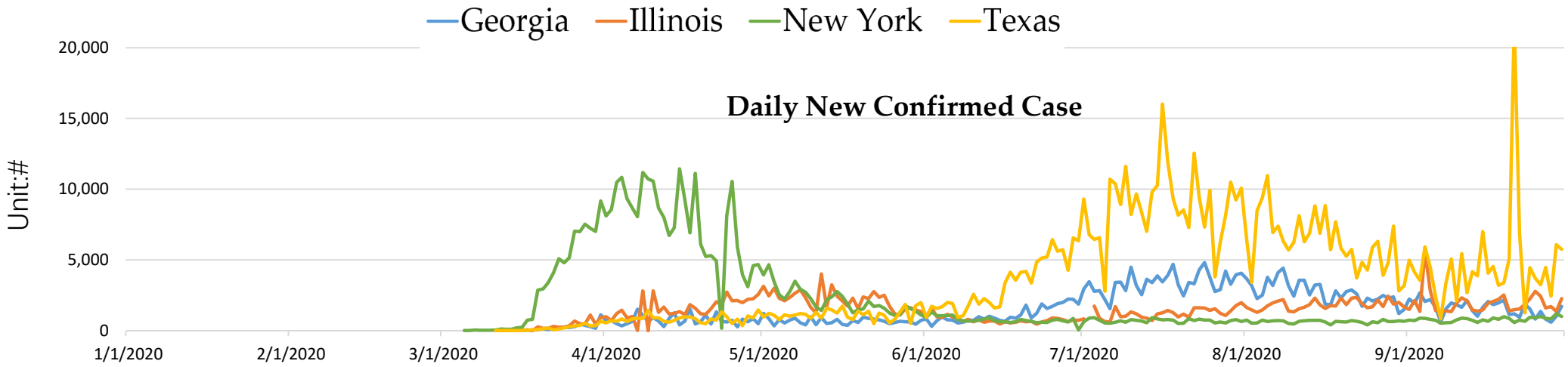
The Global Pandemic and Stringency Index?

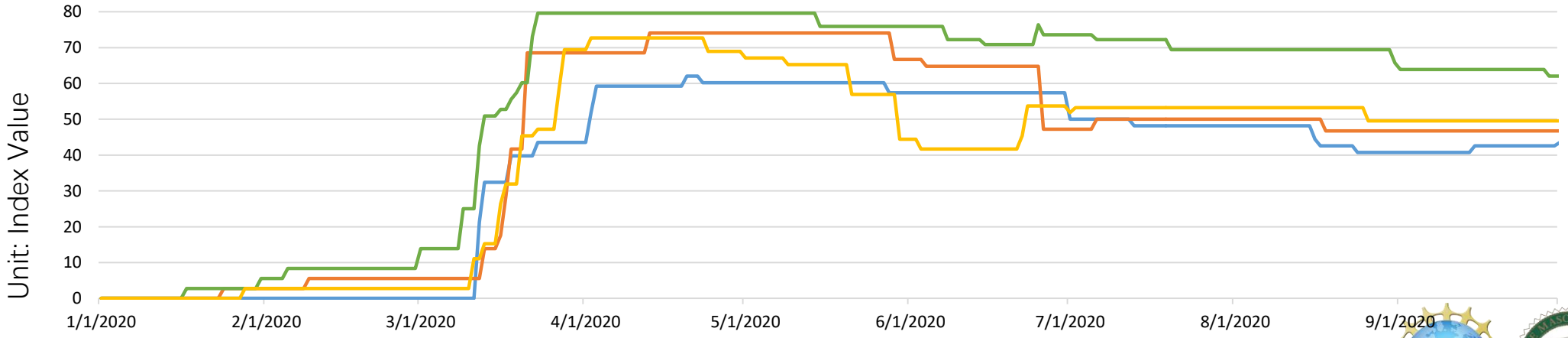


The US Pandemic and the Stringency Index?



Use Case - Trend of Economy

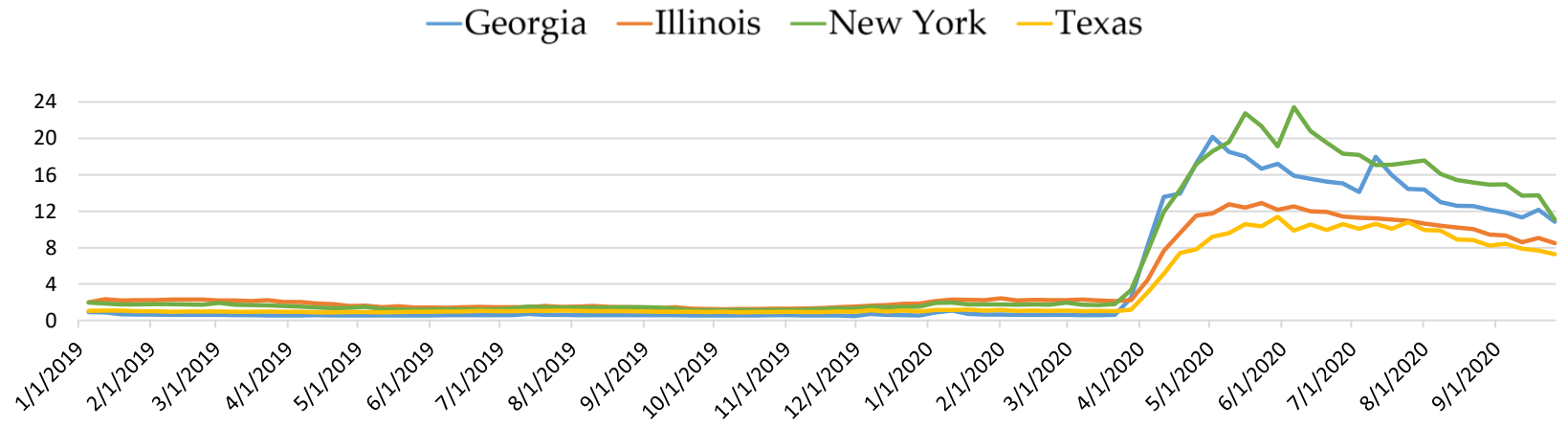




Policy Stringency Index

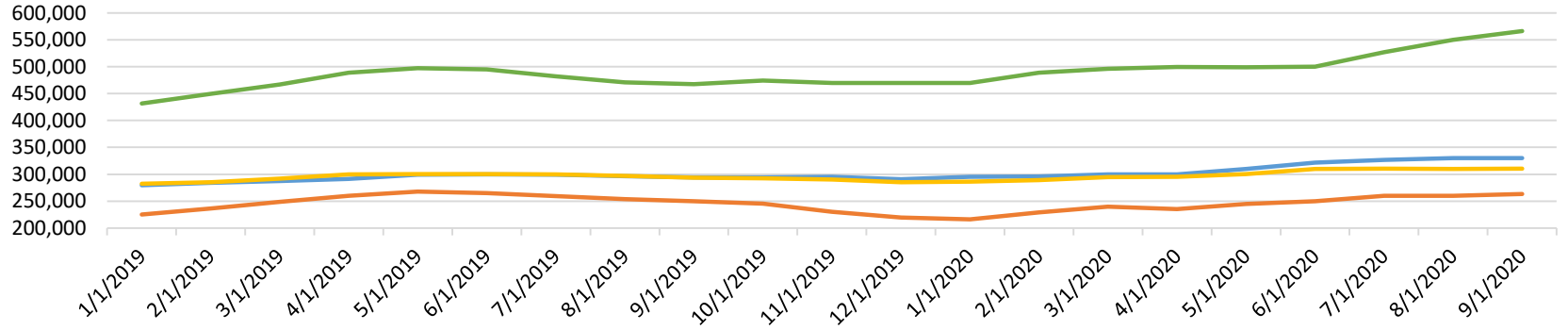
Use Case - Trend of Economy

Unit:%



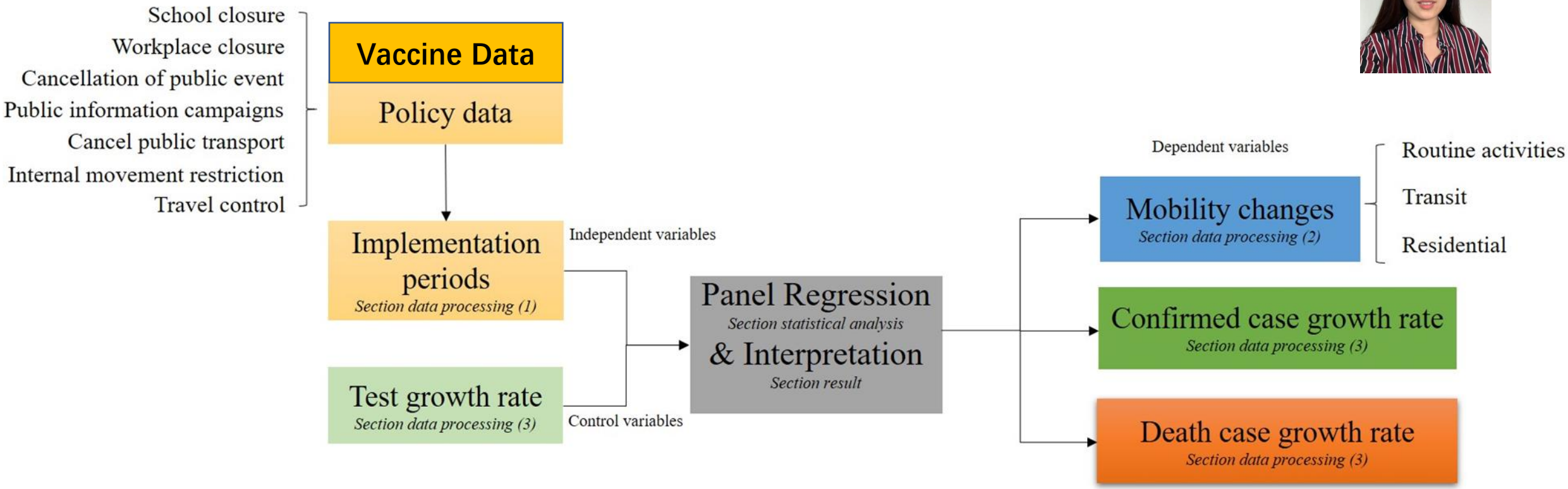
Unemployment Rate

Unit:$



Housing - Median Listing Price

1.2.2 How effective are different policies and vaccine?



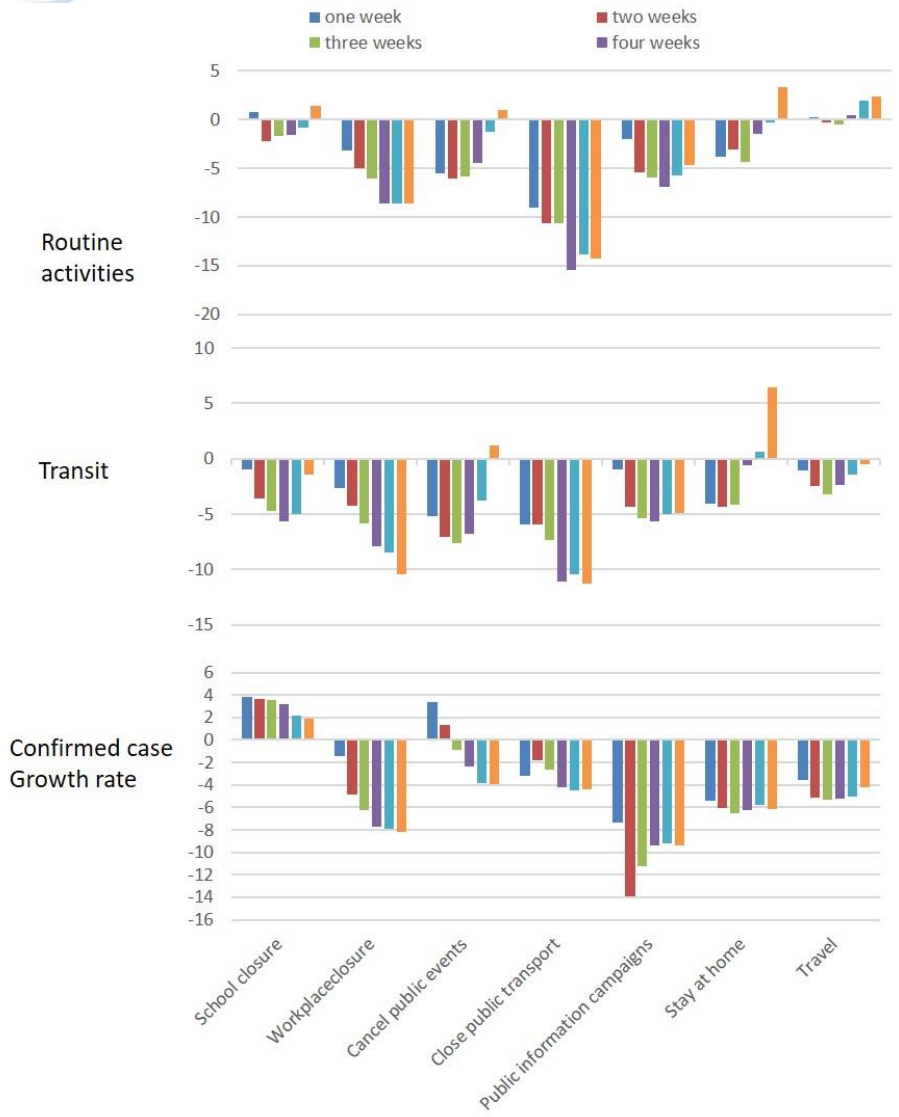
A Spatiotemporal Science Approach

- The seven policies in the main study of interest were denoted by *school*, *work*, *event*, *transport*, *campaign*, *home*, and *travel*, *vaccine*, *and air quality*.

$$ConfirmedGR_{sd} = \alpha + state_s + \left(\sum_{i\neq 0} \beta_{1i} school_{sdi} + \sum_{i\neq 0} \beta_{2i} work_{sdi} + \sum_{i\neq 0} \beta_{3i} event_{sdi} + \sum_{i\neq 0} \beta_{4i} transport_{sdi} + \sum_{i\neq 0} \beta_{5i} campaign_{sdi}\right.$$

…

Impacts of policies/vaccine on mobility, outbreak, and mortality



Stay-at-home orders:

- significantly decrease mobility in the first three weeks
- significantly decrease confirmed case growth rate

Workplace closures:

- significantly decrease in mobility
- non-consistent changes in mobility

Public event cancellations:

- significant changes in mobility
- significant results in the confirmed case growth rate after one month

Public transport closures

- significant changes in mobility
- no significant impacts on case growth rate

School closures and international/national travel

- impact on mobility were not significant
- impact on confirmed case growth rate were not significant

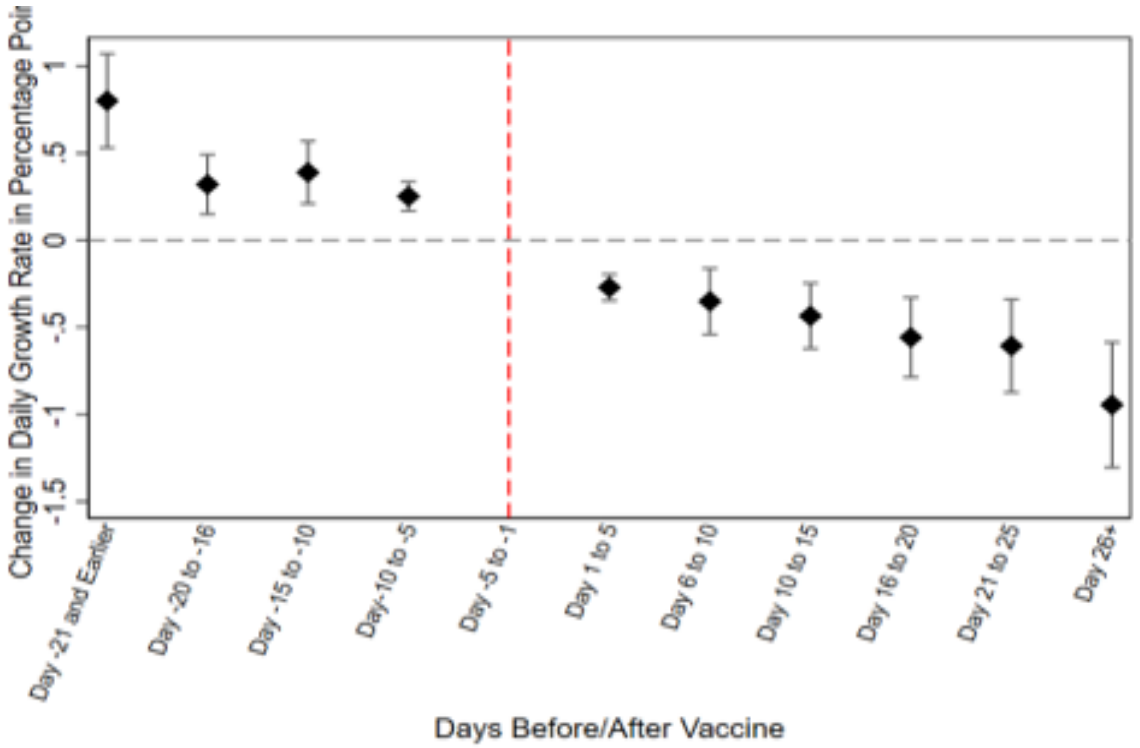
Public information campaigns:

- had little to no effect on mobility
- effectively reduce the COVID-19 confirmed case growth rate
- It indicates that this policy can significantly impact COVID-19 trends through channels other than changes in mobility

Counterfactual analysis: state-specific time trends for the confirmed case growth rate

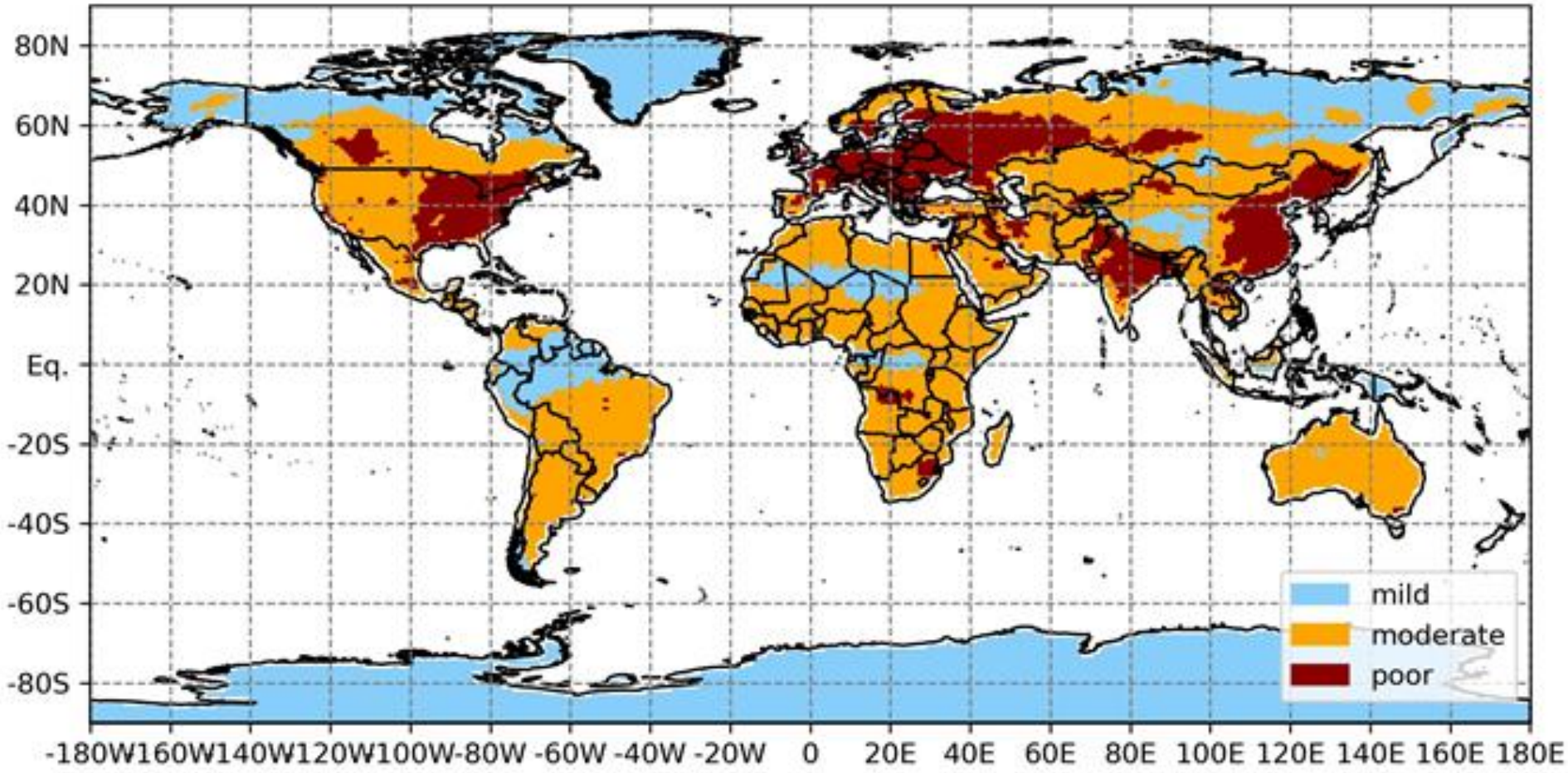
- Good fits between the predicted confirmed case growth rate and the observations.
- Large gaps between the predicted curves with and without policies

Effectiveness of vaccination



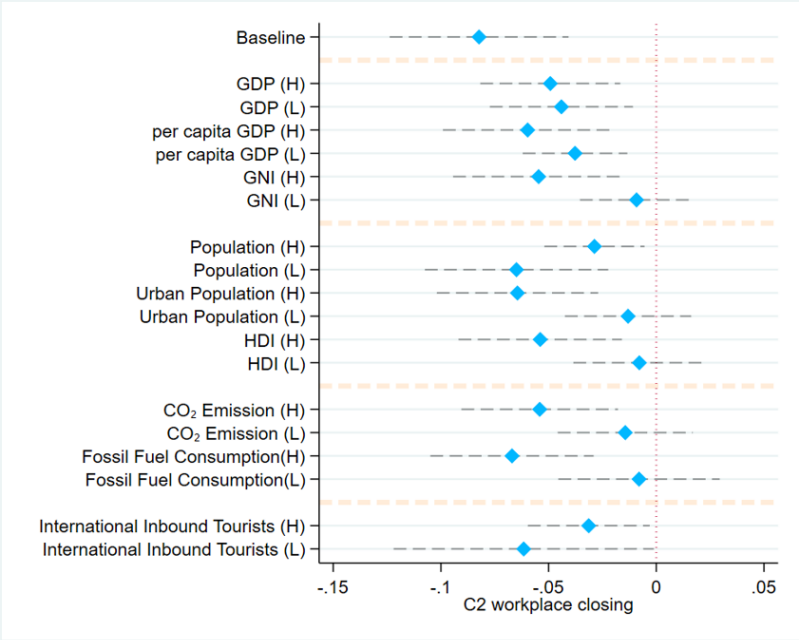
- Red vertical line represents the reference period.
- Black square symbols describe the point estimates (along with their 95% confidence intervals).
- shows that there were significant decreases in the daily COVID-19 case growth rate after the initial shots of the vaccine, with 0.124, 0.347, 0.345, 0.464, 0.490, and 0.756 percentage point declines in the 1-5, 6-10, 11-15, 16-20, 21-25, and 26 or more day post-periods after the start of vaccination

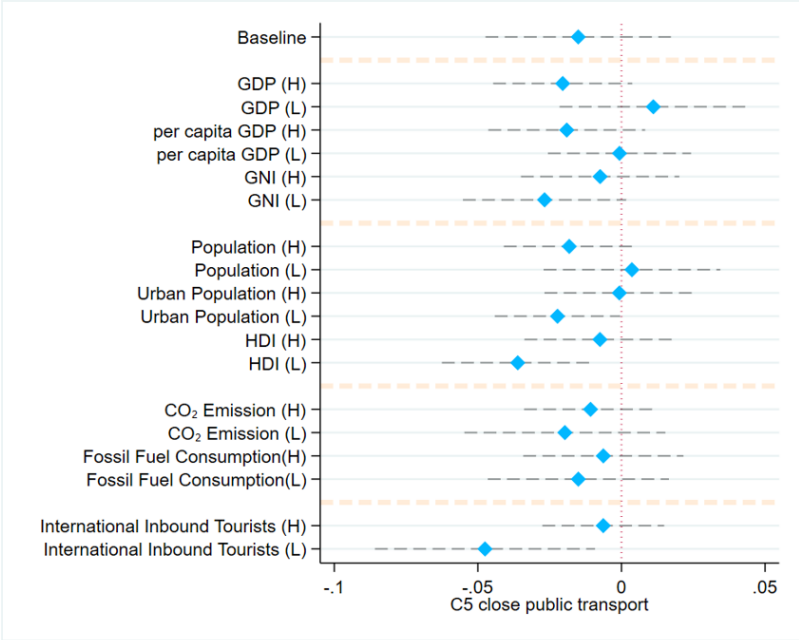
Analytical results: three NO_2 clusters in the global scale

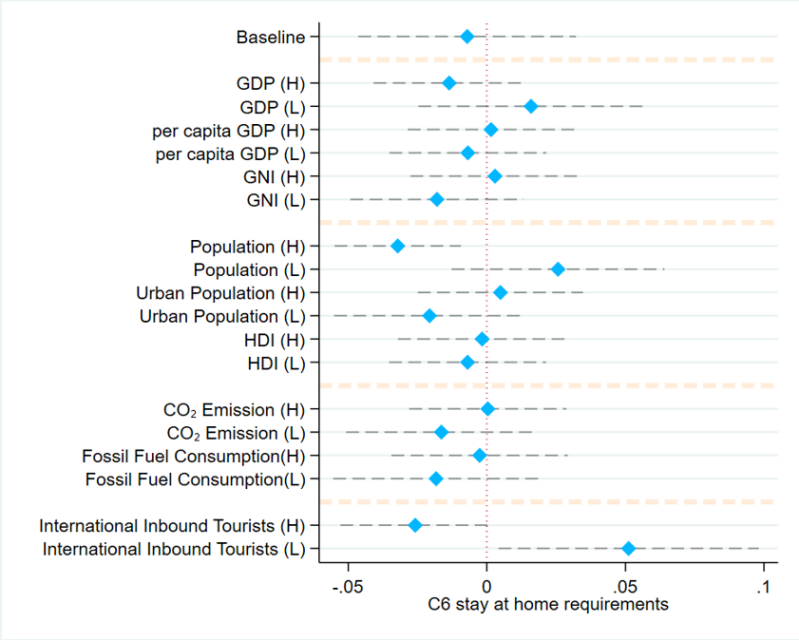


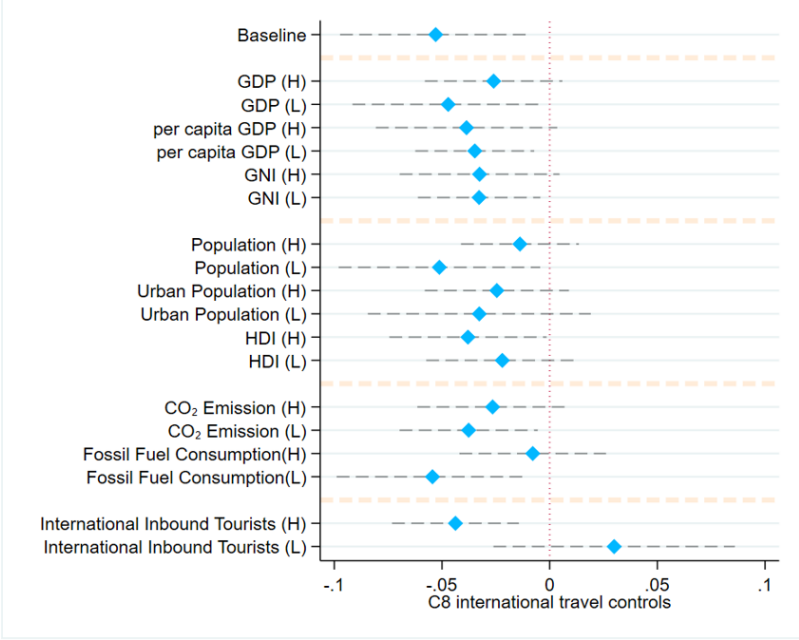
Spatial distribution of three clusters derived from long-term monthly mean NO2 TVCD

Heterogeneous impacts of policies in the poor cluster







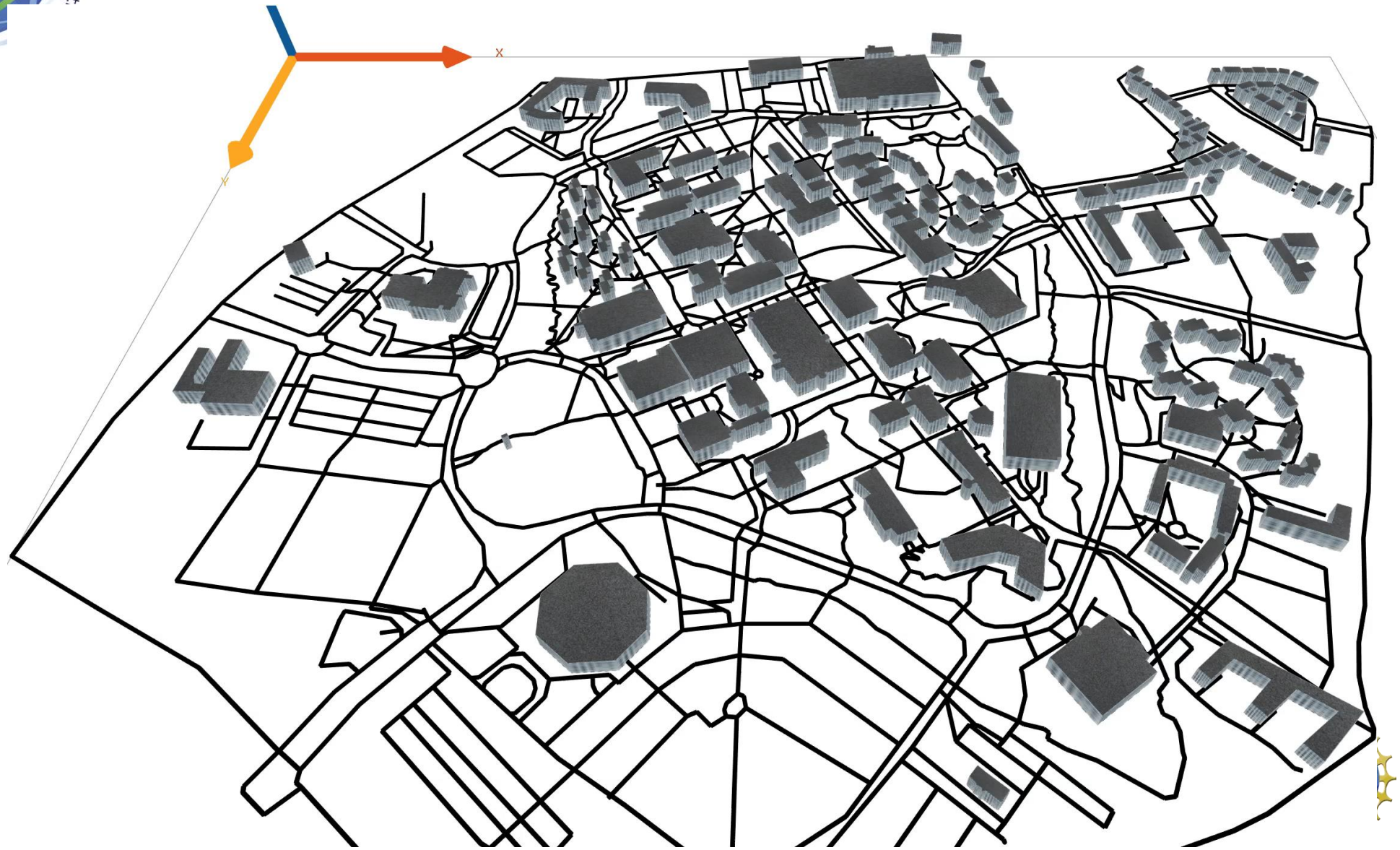


1.2.3 School Reopen

Hai Lan

- **Reopening is complicated** by the known high infectivity of the virus, complicating our ability to ensure the health and safety of students, employees, and families.
- **Lack of timely and scientific information** to support schools reopen decision making and control strategies design
- It is crucial to **predict** COVID-19 **spread trajectories** based on specific school/campus thus support decision making with risk level assessment and policy recommendations

GMU Simulation



A Spatiotemporal Science Approach

- Predicted COVID-19 cases trajectory C_p on studied campus/school :

$$C_p = \sum_t C_t \cdot I_{p,t} - \sum_t N_{quarantine,t} + \sum_t N_{recovered,t}$$

- Where each individual I

$$I = \{I_{id}, I_p, I_{inc_pd}, I_{age}, I_{symp}, I_{x,y}(t), I_{target}(t), I_{vac\ date}, I_{infdate}\}$$

on campus follow the daily activity setup in the simulation.

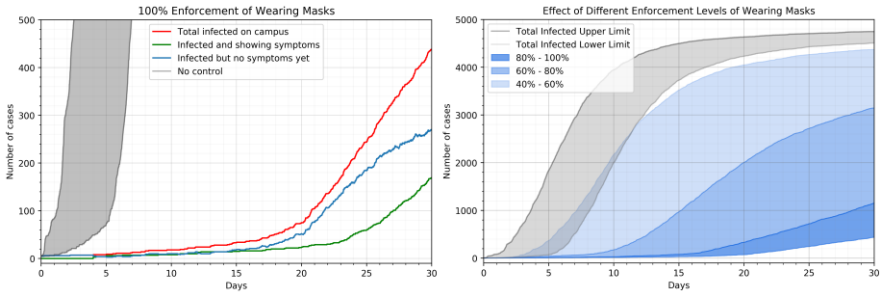
- While the probability of infection of each individual I_p for each contact is calculated by local R_0, corresponded contact matrices K, age group I_{age} and also impacted by enforced control policies. Hence,

$$I_p(t+1) = \begin{cases} 0.75 \cdot I_p & if \text{ asymptomatic} \\ I_p(t) \cdot e \cdot p_b & if\ building\ capacity\ ctrl \\ \gamma \cdot e \cdot I_p(t) & if\ social\ distancing \\ I_p(t) \cdot e \cdot p_{mask} & if\ masking \\ I_p(t) \cdot e \cdot p_s & if\ Hygiene \\ \gamma_v(t) \cdot I_p(t) & if\ vaccinated \end{cases}$$

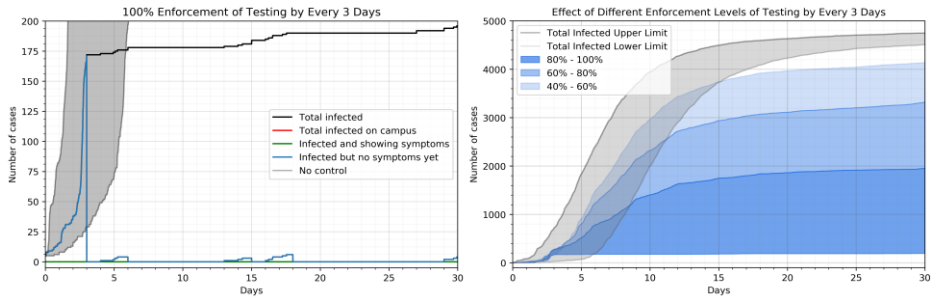
- The individual status $I_n(t)$ updated as:

$$I_n(t) = \begin{cases} \text{"susceptible"} & initial \\ \text{"exposed"} & incubation \\ \text{"infected "} & \\ \text{"quarantine"} & if\ symptom\ screening/viral\ testing/contract\ tracing \\ \text{"dead"} & if\ P(\gamma_{mr}) \\ \text{"recovered"} & if\ 1 - P(\gamma_{mr}) \end{cases}$$

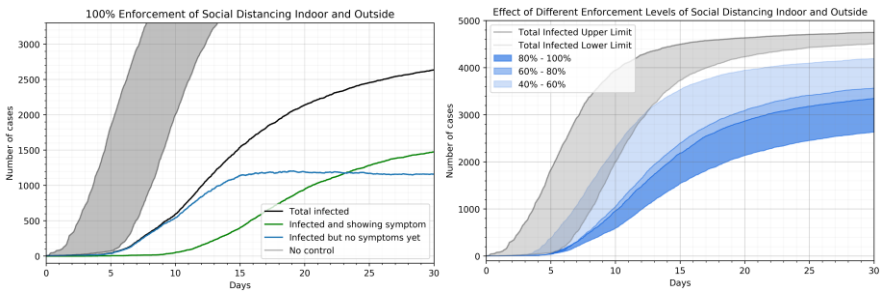
Policy Controlling Results



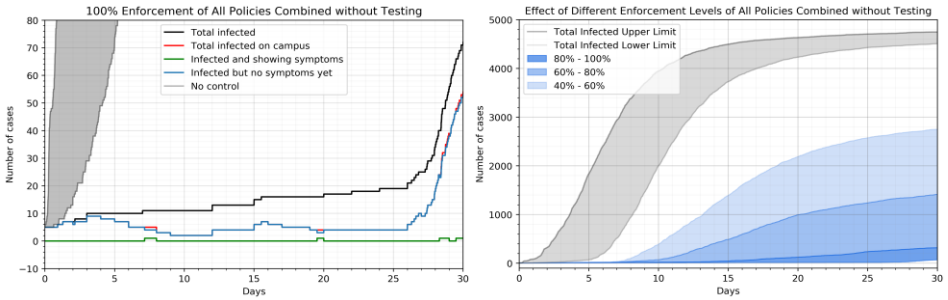
Masks Only



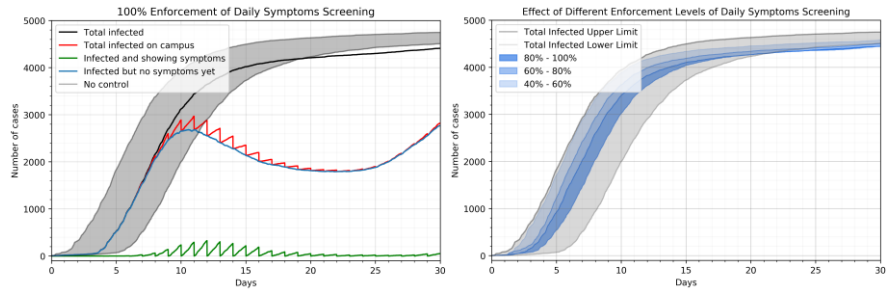
Viral Testing Only



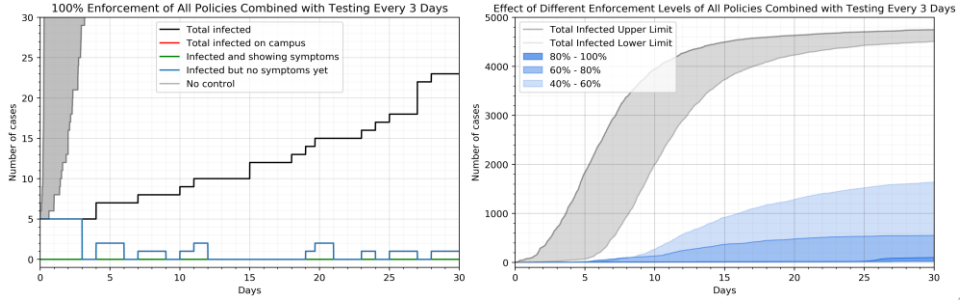
Social Distancing Only



All combined without testing

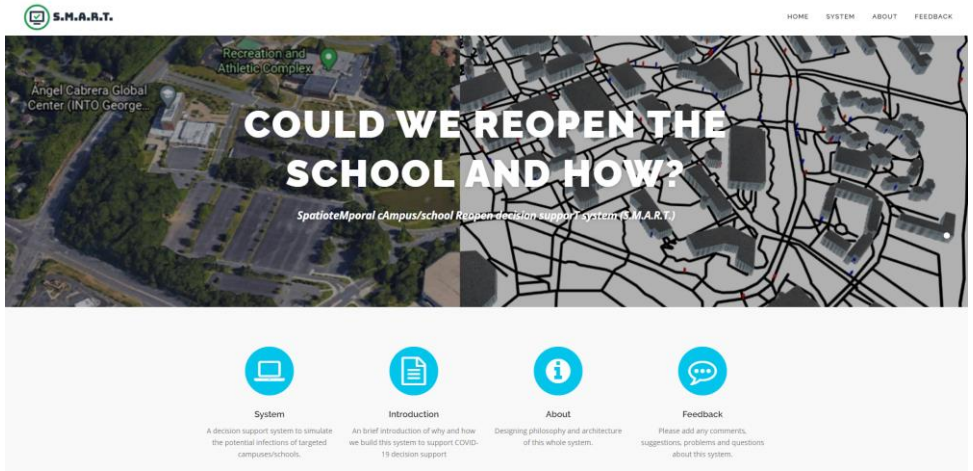


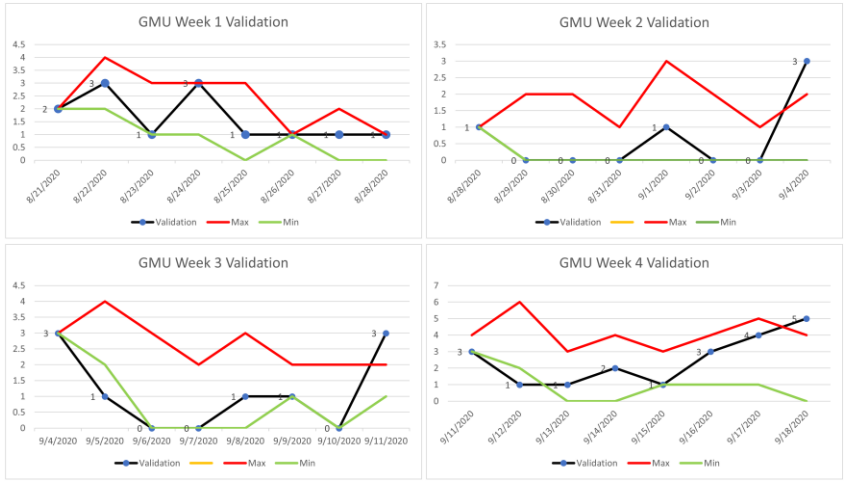
Symptoms Screening Only

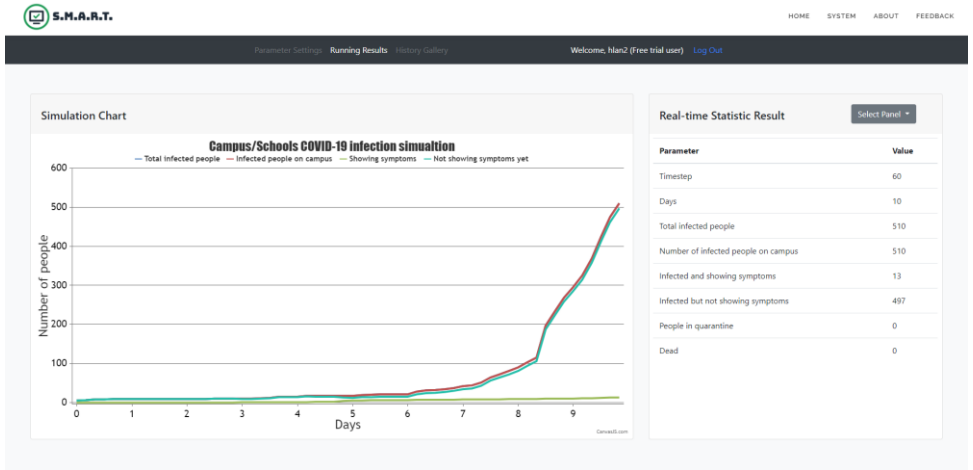


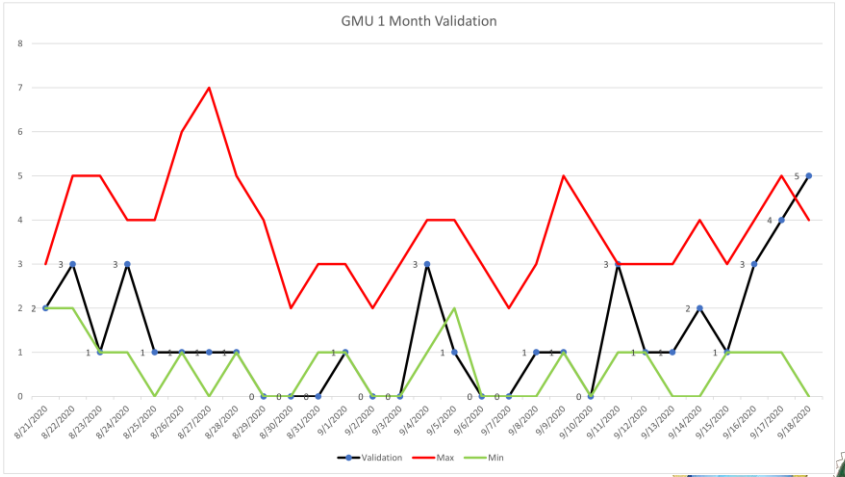
All combined

System and Result Evaluation





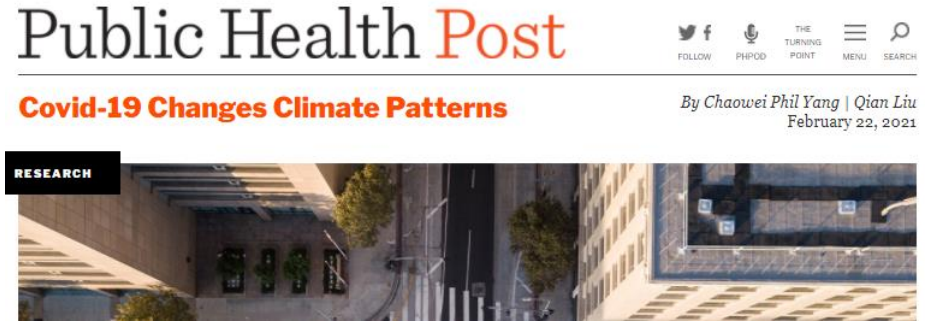




System link: https://smartgateway.stcenter.net

1.2.4 How was the environment impacted?

- **Challenges:**
- How different environmental factors are changed by COVID-19?
- Do they have similar responses in different regions?
- How to explain the changes?

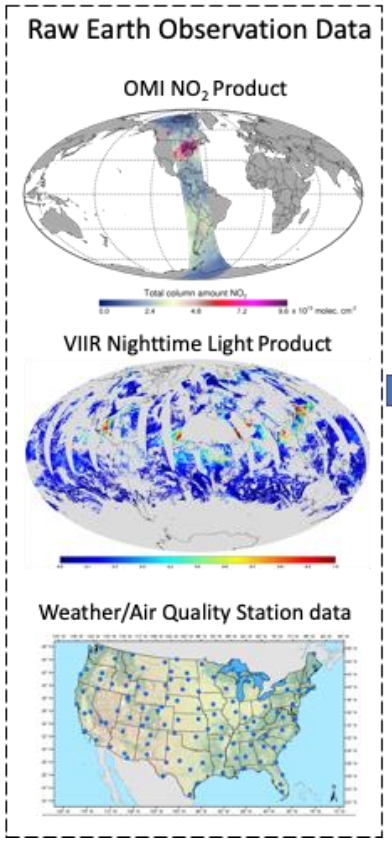


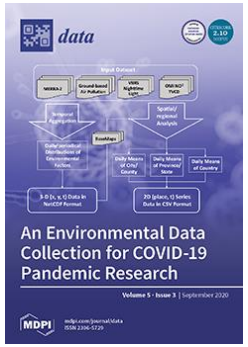
- **Motivation:**

➢ The factors are essential indicators of industrial production and global economic;

➢ Results can offer vital and practical basis for loss assessment and economic reopening and planning.

Workflow



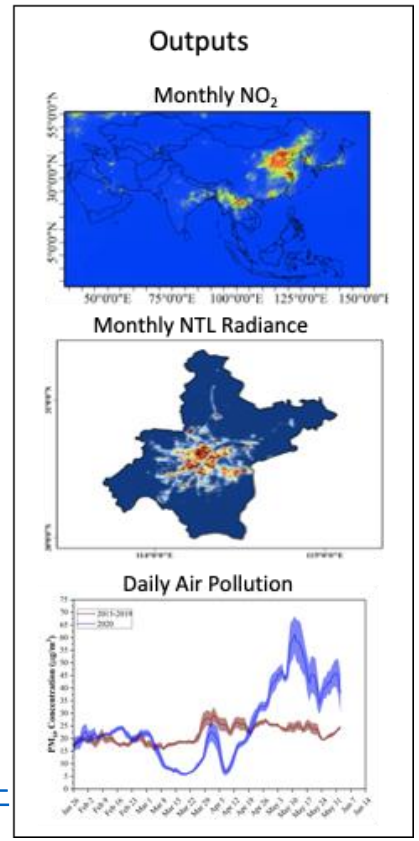


Spatiotemporal Environmental Data Analytical Tool

python data analysis and processing packages: numpy, netCDF4, math, scipy

Daily/Monthly/ Periodically Average, Season cycles

Visualization



OMI NO_2: https://github.com/stccenter/COVID-19-Data/tree/master/Environmental%20factors

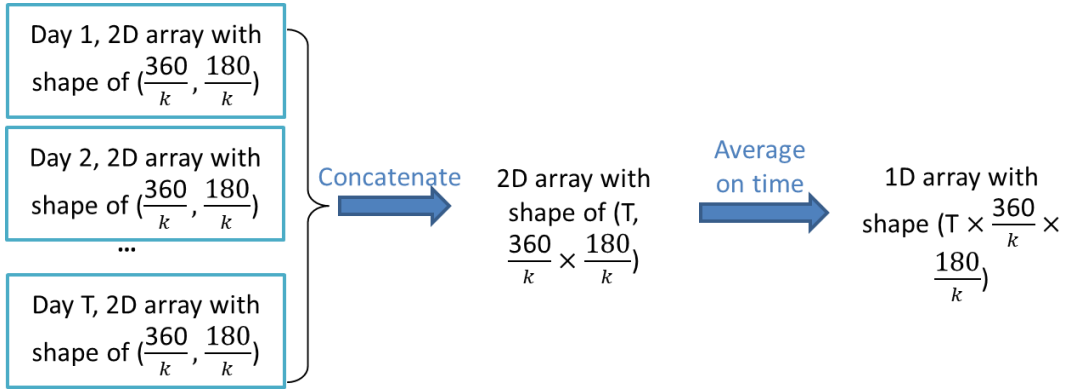
VIIR Nighttime Light: https://github.com/stccenter/COVID-19/tree/master/analysis/nightlight

Ground-based air pollution: https://github.com/stccenter/COVID-19/tree/master/analysis/nightlight

Liu, Q., Liu, W., Sha, D., Kumar, S., Chang, E., Arora, V., Lan, H., Li, Y., Wang, Z., Zhang, Y., Zhang, Z., Harris, J., Chinala, S. and Yang, C., 2020. An Environmental Data Collection for COVID-19 Pandemic Research. Data, 5(3), p.68.

A spatiotemporal science approach

a. Spatiotemporal dimension generalization



b. Spatiotemporal average

- $\overline{EF_t} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} EF_{i,j,t}}{m \times n}, t = 1,2,3,\ldots,T_{study}$

c. Time period spatiotemporal average

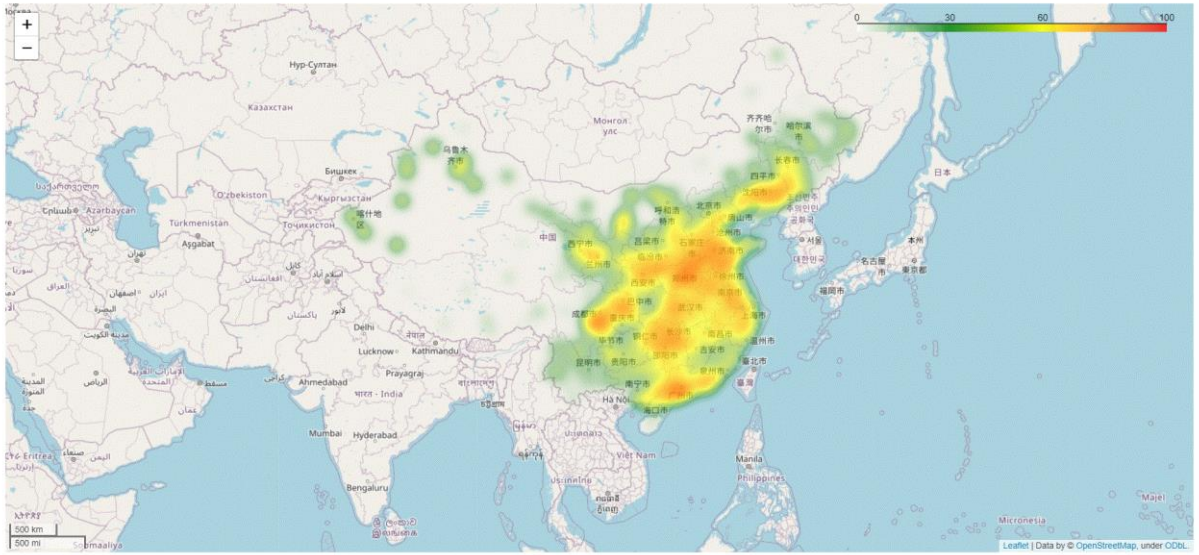
- $\overline{\mathrm{EF}}_{i,j} = \frac{\sum_{t=T_{period\ start}}^{T_{period\ end}} EF_{\mathrm{i,j,t}}}{T_{period}}, i \in [0, m-1], j \in [0, n-1]$

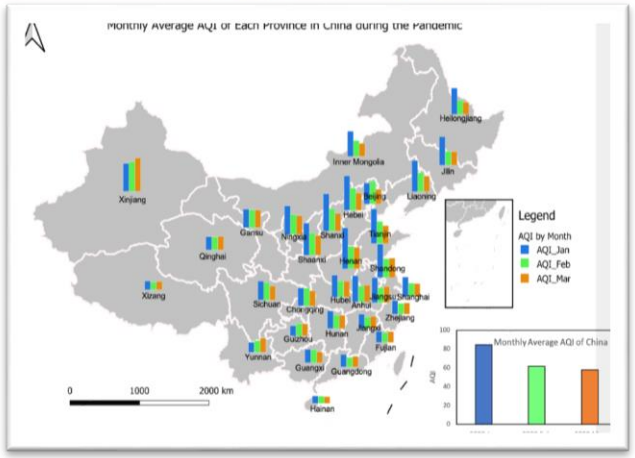
d. Spatiotemporal anomalies

$$A_{i,j} = \overline{EF(i,j)}_{p,y} - \overline{EF(i,j)}_{p,2010-2019},$$

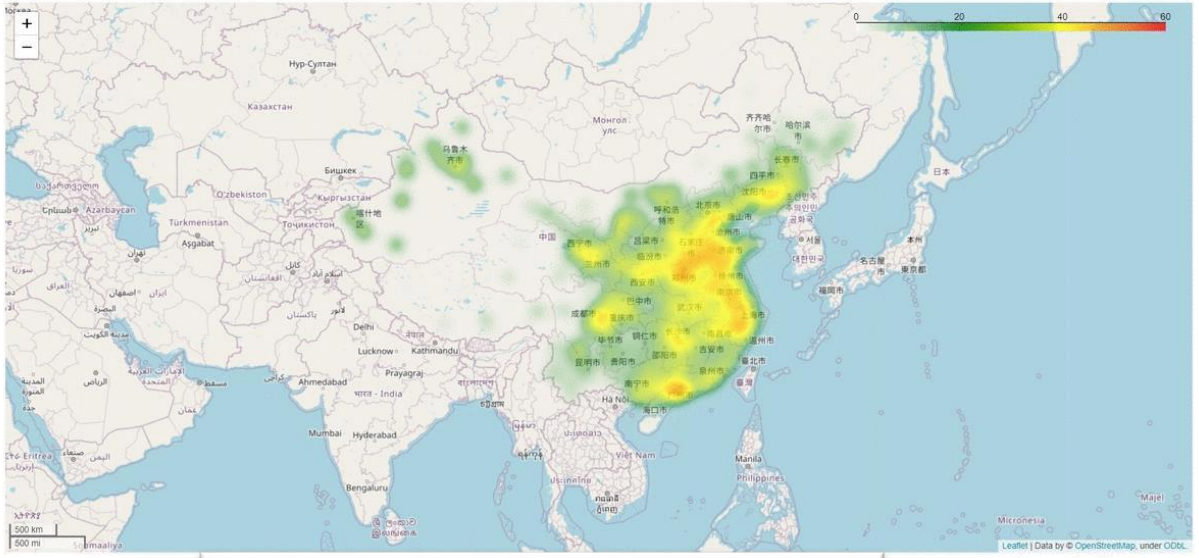
China air quality impact

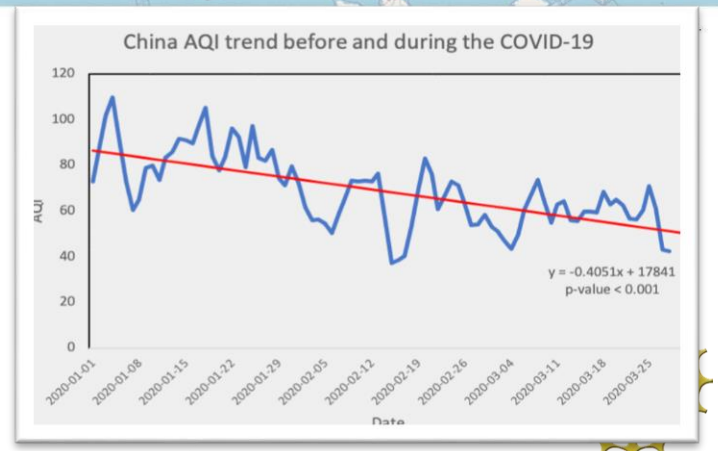
Heatmap of the Air Quality Index on Jan. 01, 2020





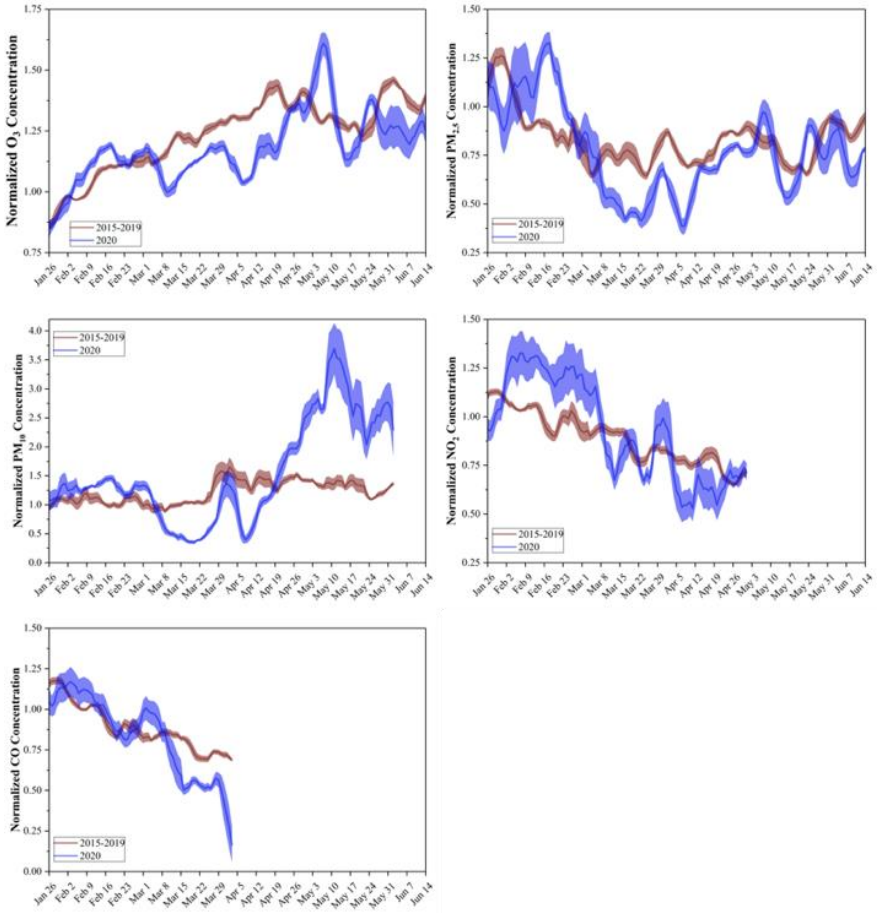
Heatmap of the NO2 Emission on Jan. 01, 2020

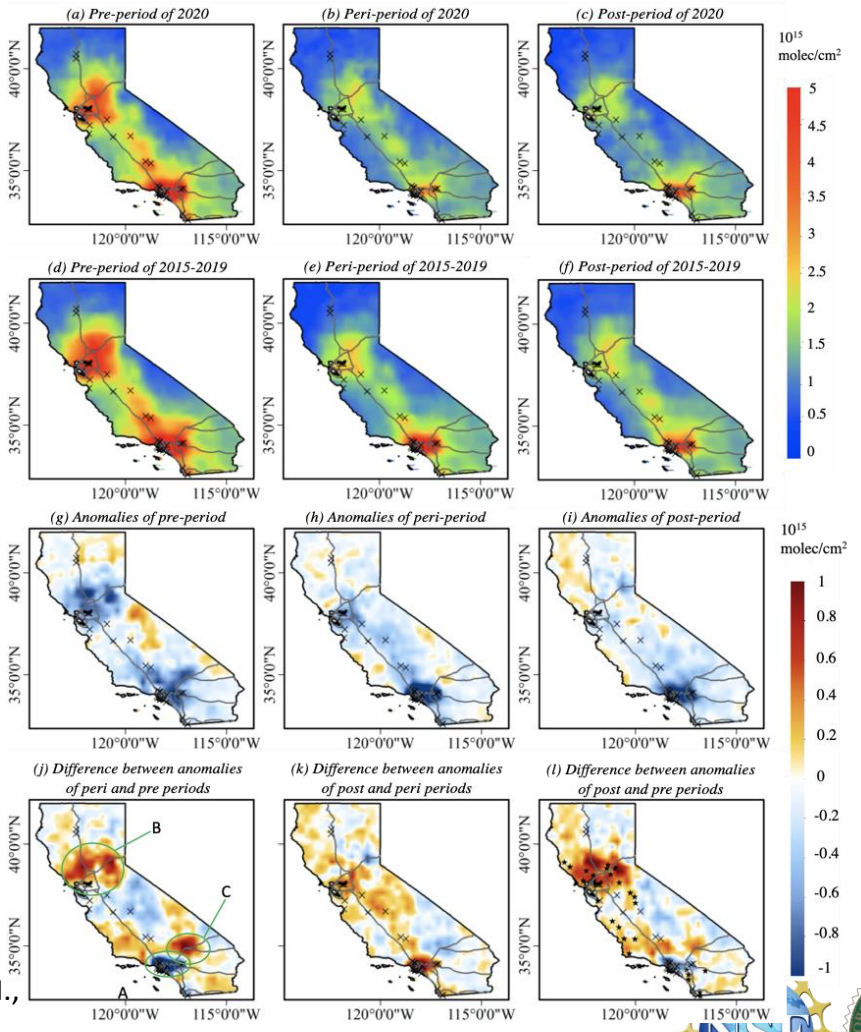




CA AQ Impact

- Impact on air quality in the U.S.



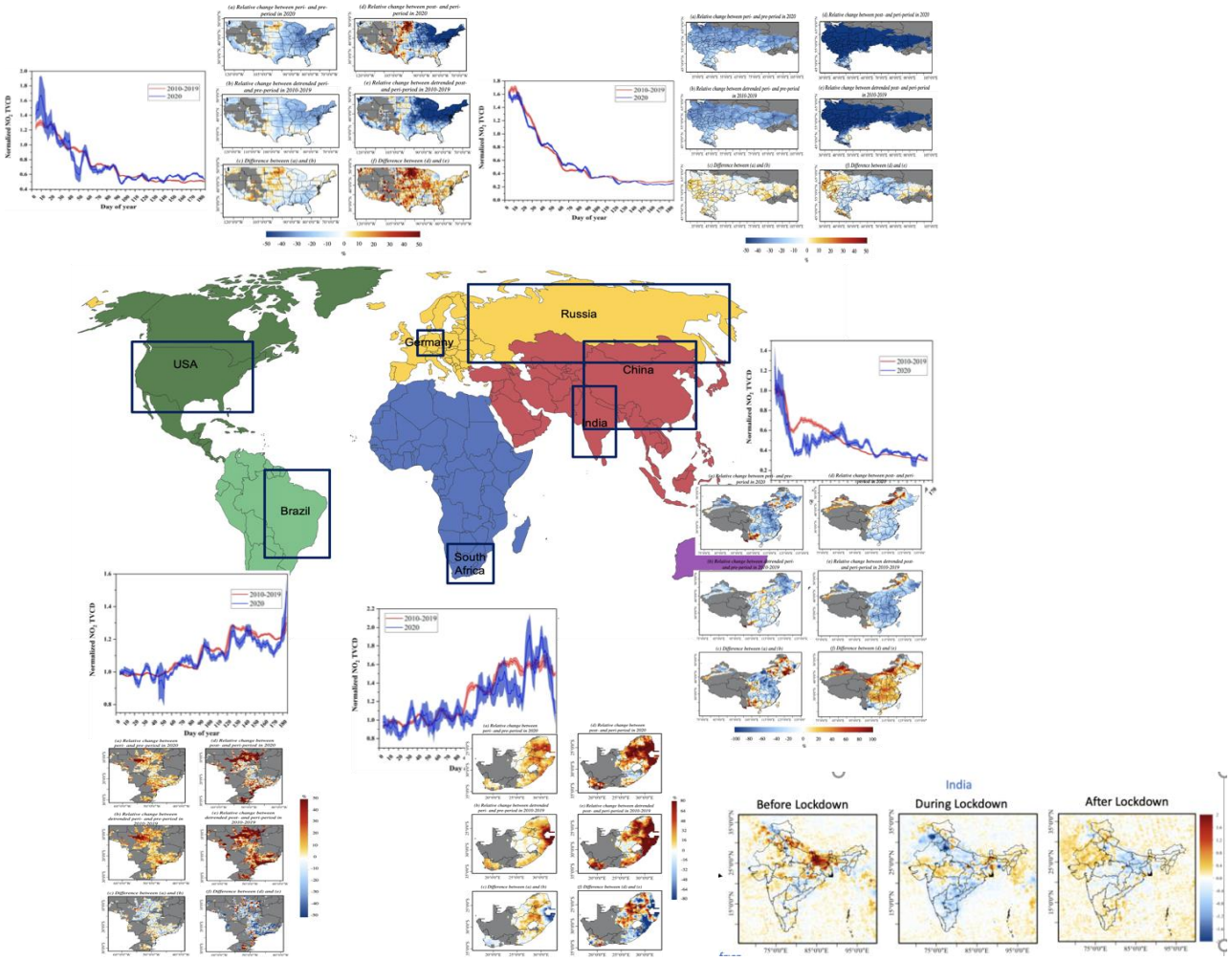


Liu, Q., Harris, J.T., Chiu, L.S., Sun, D., Houser, P., Yu, M., Duffy, D., Little, M., and Yang, C., 2020. Spatiotemporal Impacts of COVID-19 on Air Pollution in California, USA. Science of the Total Environment, p141592

Global Impact on NO2?

	Relative difference between peri- and pre-period (%)		Relative difference between post- and peri-period (%)	
	2020	2010-2019	2020	2010-2019
China	-50%	-38%	-23%	-41%
USA	-43%	-44%	4%	-8%
South Africa	33%	45%	6%	3%
Brazil	14%	17%	3%	4%
India	-9%	3%	-5%	-10%
Germany	-60%	-50%	-39%	-44%
Russia	-67%	-66%	-30%	-20%

- Mean tropospheric NO2 of most target countries significantly decreased (<-5%) during the lockdown, such as China, South Africa, India and Germany.
- Brazil and Russia slightly declined by 3% and 1% during lockdown.
- The USA increased by 1% during the lockdown.
- After the reopening, the NO2 rebounded in most target countries (>5%) except South Africa, Brazil and Russia.

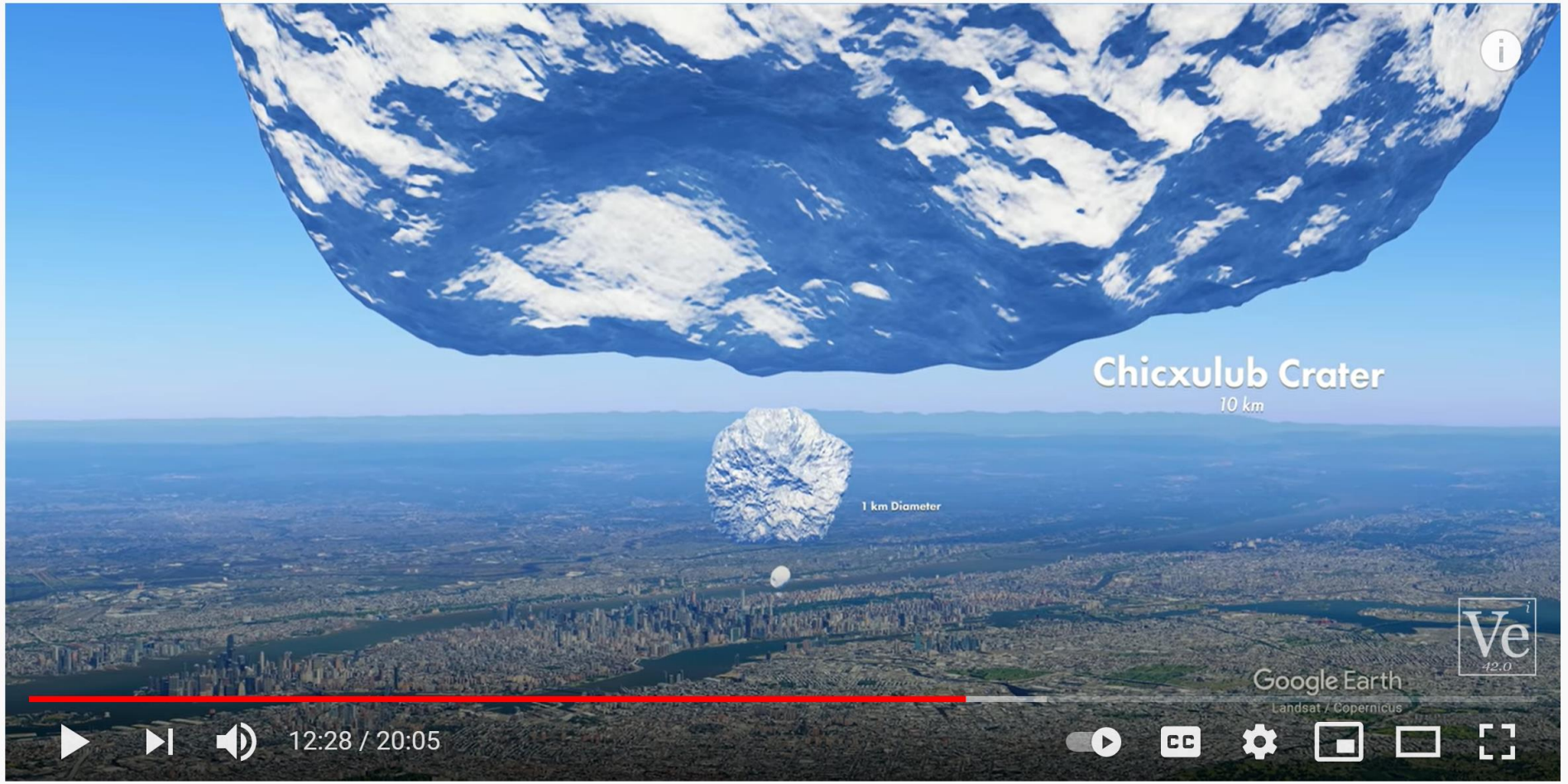


Liu, Q., et al., 2020. Spatiotemporal Changes in Global NO2 Emission Due to COVID-19 Mitigation Policies. Environment International, STOTEN.

1.3 Protecting our home planet

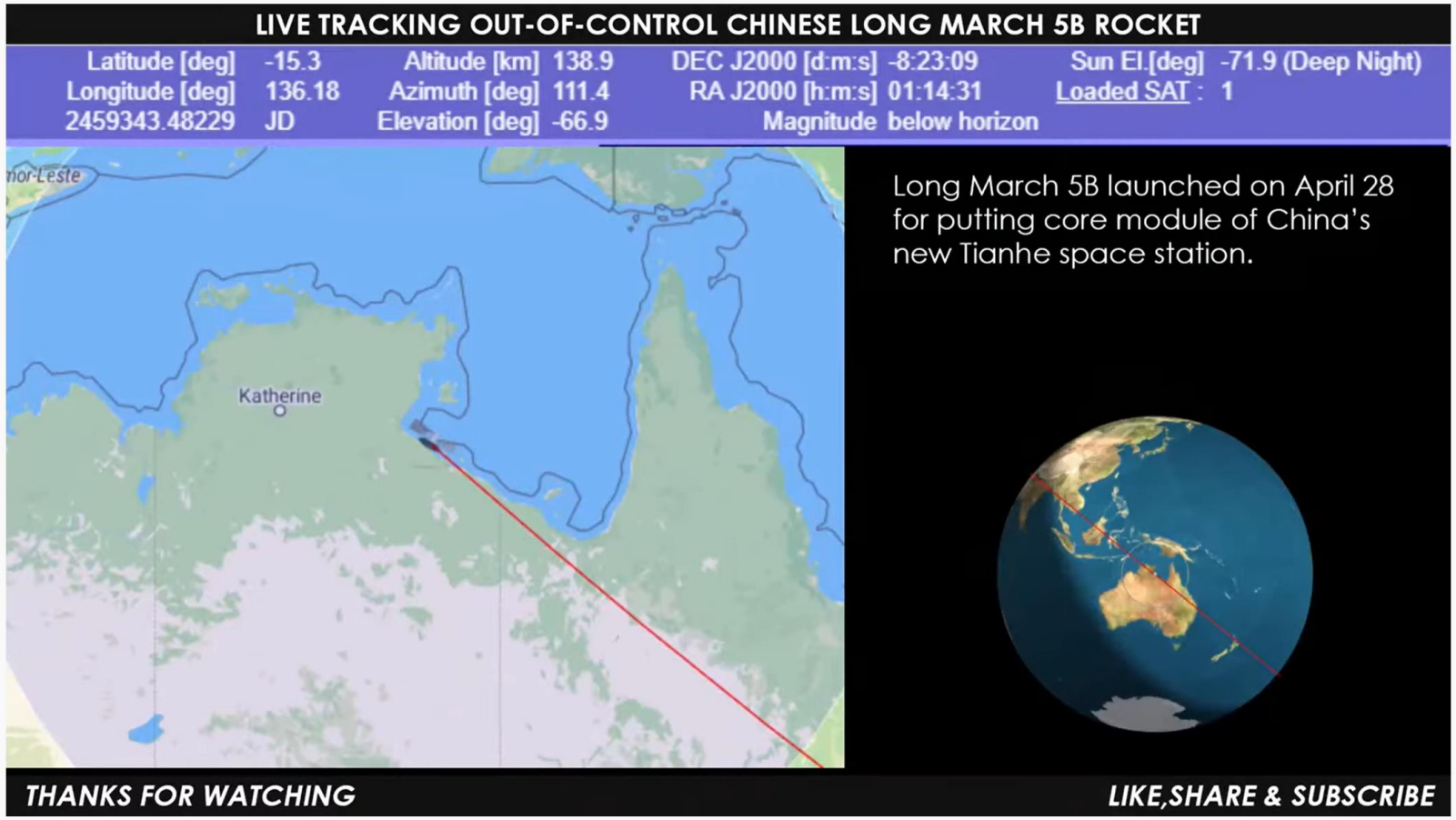
1) Asteroids
2) Man-made objects
3) Our understanding
4) The dilemma

Asteroids



[These are the asteroids to worry about - YouTube]

Man-made Objects

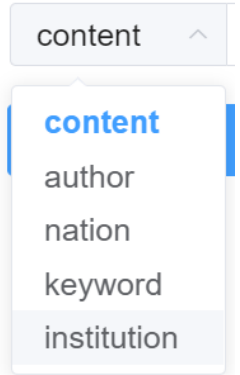


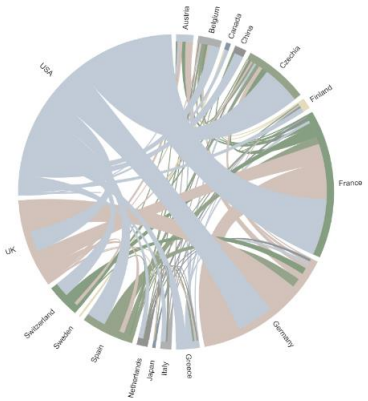
Out of control China rocket falling to earth live tracking Long March 5B Rocket debris

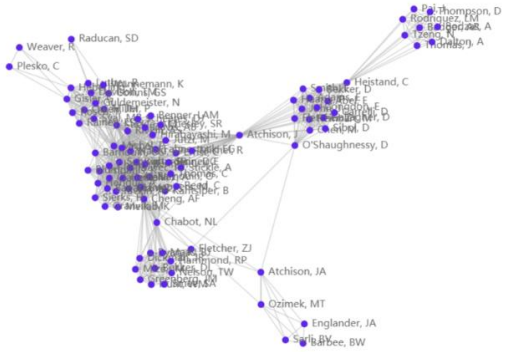
https://www.youtube.com/watch?v=rPw6wPgtCb8

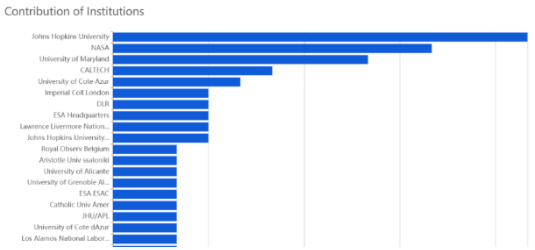
What we know?

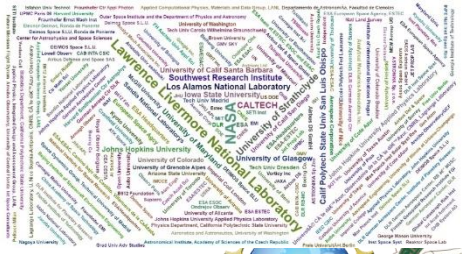
- Design
  - Database schema for PDC publication analysis based on ISI format
  - Extraction workflow from PDC pdf files
- Data Process
  - Manually extract and clean PDC abstract info for 2013, 2015 and 2017
  - Gateway contains over 1,400 PD publications, including PDC abstracts
  - Transform extracted data into JSON file
- Develop
  - Develop Analysis Interface
  - Integrate real data of 2013, 2015 and 2017 (JSON) to analysis module
- PD Publication Analysis Tool: allows options to search from content, author, nation, keyword or institution and see trends, relationships, ranking, and found references.











[Planetary Defense – PD Framework Gateway (stcenter.net)]

What we know?

- PD instruments: ~ 660 instruments
- Missions: Over 300 missions
- **Including instrument types such as:** Visible Cameras, IR Cameras, LIDAR, Magnetometers, Spectrometers, Analyzers, Multispectral Imagers, Detectors, Satellites, Monitors, and other instruments.

Missions	Mission Types
Deep Impact (EPOXI)	PD
Mars Odyssey	Mars
AIM	Cloud
AlSat-2 (Algeria Satellite-2)	Earth Observation
OSIRIS-Rex	Asteroid
AstroSat	Space
CartoSat-2A	Earth Observation

- Includes parameters such as: instrument type, instrument name, mission sponsor, mission flew, instrument provider, mission launch date, mass, power, dimensions, roles/capabilities, instrument composition, mission type, characteristics, comments, and the information source.
- Future work: Add the Instruments and Missions content to the PD Knowledge Base and continue work with Neo4J

MAGNETOMETERS

								Characteristics
Instrument Type	**Instrument Name**	**Provider**	**Mission Flew**	**Year**	**Mass**	**Power**	**Dimensions**	**Role/Capabilities**
Magnetometer	AMR Magnetometer	developed at LusoSpace, Portugal [1]	Aeolus / formerly ADM (Atmospheric Dynamics Mission)	launched on 22 August 2018 (21:20 GMT) [1],	300 g [1]	1 W [1]	85 mm x 53 mm x 60 mm (excl. connectors) [1]	"The magnetometer is a small (credit card surface dimension) and robust unit that can be used for several LEO missions" [1].

Characteristics	Characteristics	Characteristics			
Instrument Affiliation	**Composition Makeup**	**Characteristics**	**Cost**	**Comments**	**Information Source**
	"Possibility to generate external magnetic field in the chip by mean of built in coils" [1]	Bandwidth: up to 40 Hz (adjustable) [1]		Includes bandwidth from source 1. Mission type: Earth observation/wind.	[1] https://directory.eoportal.org/web/eoportal/satellite-missions/a/aeolus

Shams, I., Li, Y., Yang, J., Yu, M., Yang, C., Bambacus, M., Lewis, R., Nuth, J.A., Oman, L., Leung, R. and Seery, B.D., 2019. Planetary Defense Mitigation Gateway: A One-Stop Gateway for Pertinent PD-Related Contents. *Data*, *4*(2), p.47.

The Dilemma

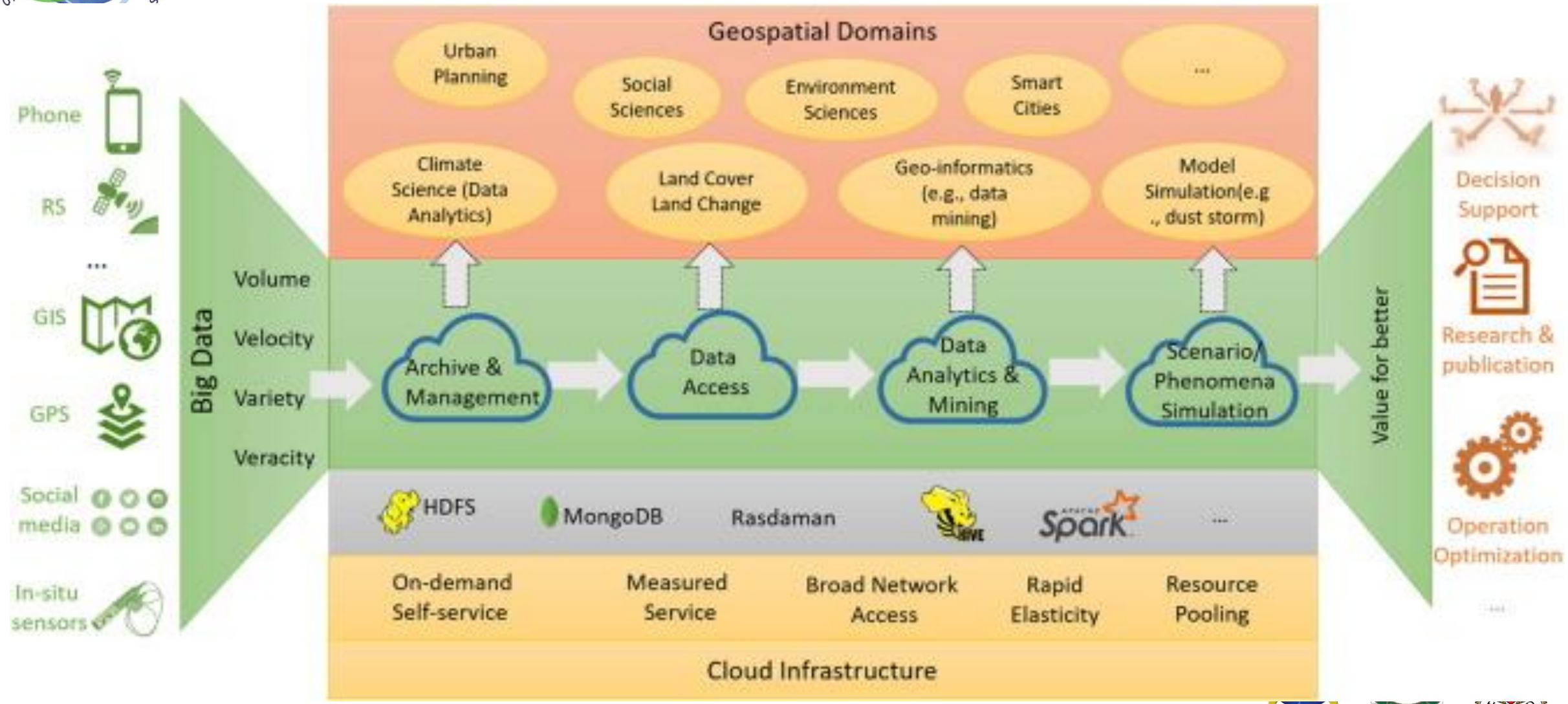
- We still miss many asteroids.
- We can't predict accurately where they will hit.
  - Good chance in ocean 7/10, but
  - *1% (half population),* **0.3% (urban area)**

How much of the earth is covered by city?

Population distribution

- We don't have an efficient mechanism (tools & diplomatic channels) to mitigate.
- A deeper **spatiotemporal science approach** is needed to obtain a comprehensive understanding.

2. A formal and sustainable approach



Yang, C., Yu, M., Hu, F., Jiang, Y. and Li, Y., 2017. Utilizing cloud computing to address big geospatial data challenges. Computers, environment and urban systems, 61, pp.120-128.

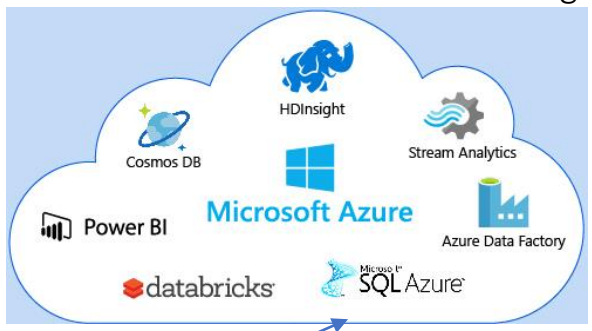
A Spatiotemporal Computing Infrastructure

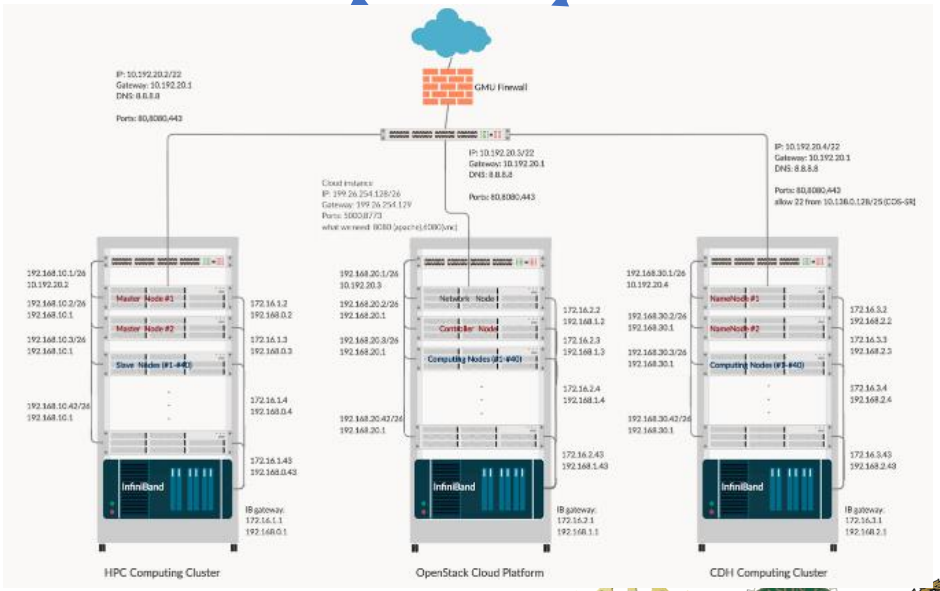
Zifu Wang

GMU Data Center

Three Computing Clusters designed for different purposes in GMU Datacenter contain *3024 cores*, *8Tb memory* and *220Tb Storage* in total:

- Hadoop Distributed Filesystem based **Cloud Computing** environment (CDH)
- Message Passing Interface powered **High Performance Computing** cluster (HPC)
- OpenStack based **Cloud Sharing** Platform (Cloud)





AWS

- Amazon has sponsored credits to our center to use AWS modules such as EMR, EC2, P3, and S3.
- Resources are utilized for publishing data, providing application services for public users.

Microsoft Azure

- Azure modules such as VM and Database, are utilized for publishing data, providing application services for public

Perform Research

Chaowei Yang
Director of I/UCRC at George Mason University

S V Subramanian
Co-Director of I/UCRC

Wendy Guan
Executive Director

Gordon Hanson
the Peter Wertheim Professor in Urban Policy at Harvard Kennedy School

Donglian Sun
Professor at GMU college of Science

Merce Crosas
University Research Data Management Officer at Harvard

Ben Lewis
Geospatial Technology Manager

Douglas Richardson
Distinguished Researcher

Ruixin Yang
Associate Professor at GMU

Liang Zhao
Emory Univ.

Xue Liu
Lead Researcher

Matthew Rice
Associate Professor at GMU college of Science

Devika Kakkar
Geospatial Engineer

Songqing Chen
Professor at GMU college of Engineering

Tao Hu
Postdoctoral Fellow

Josh Lieberman
Senior Researcher

Hai Lan
Postdoctoral Fellow

Sanmei Li
Associate Research Professor at GMU

Jingchao Yang
Graduate Research Assistant

Zifu Wang
Graduate Research Assistant

Yun Li
Graduate Research Assistant

Qian Liu
Graduate Research Assistant

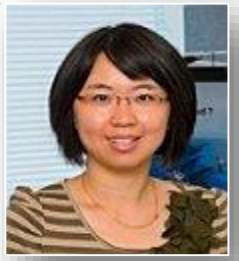
Dexuan Sha
Graduate Research Assistant

Anusha S. Malarvizhi
Graduate Research Assistant

Ishan Shams
Graduate Research Assistant

Cong Peng
Postdoctoral Fellow

Academic Participation



Wenwen Li
Tenured Associate Professor, Arizona State University, CAREER Awardee

Jing Li
Tenured Associate Professor, University of Denver

Qunying Huang
Tenured Associate Professor, Univ. of Wisconsin-Madison

Rezgui
Assistant Professor, New Mexico Tech

Zhenlong Li
Tenured Associate Professor, Univ. of South Carolina

Placed 10+ Tenure-line Faculty in phase I by GMU & UCSB, five of them obtained tenure including one early tenure and CAREER awardee. Many in top geospatial departments.

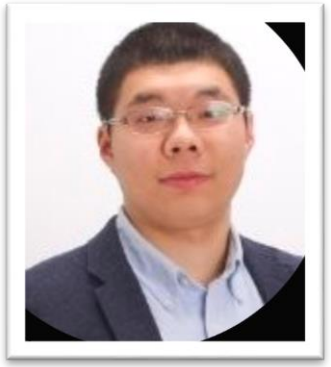
Starting Phase II

Manzhu Yu
Assistant Professor, Penn State

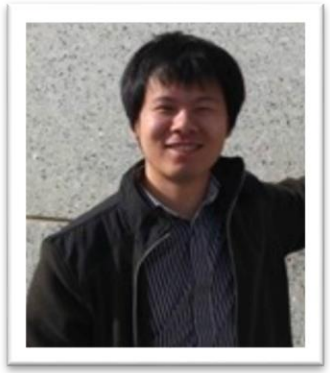
...

Industry Engagement

Fei Hu
Microsoft



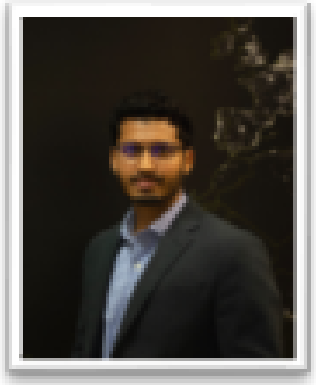
Yongyao Jiang
Amazon



Kai Liu
Facebook

...

Mengchao Xu
Torc Robotics



Ishan Shams
TechTrend, Inc

Han Qin
Ankura

Undergraduate Engagement

Jackson T. Harris
Dartmouth College
Cloud Classification/ AQ Impact
Co-Author of Data and STOTEN paper

Nadine Meister
Harvard University
COVID-19 ABM campus reopening simulation

Melanie Alfonzo
Johns Hopkins University
Fatality Prediction & Social Media Analytics

Megan A. Rice
Carnegie Mellon
COVID-19 variable correlation

Samantha Carr
McGill University
GPU-based multi-million level Analytics

Erin Abbott
University of Chicago
Spatiotemporal Patterns Analytics

Sophia O'Neill
Cornell University
COVID-19 data clean/ ArcCI prototype design

Oliver Chartock
Dartmouth College
Computing Infrastructure/ COVID-19 Visualization

Alma Topete
UCSB
Data Search, Collection, & Twitter Data Analyses

Community Building

- 12 Training Programs Since 2012, Latest Nov. and Dec., 2019
  - 200 professionals
  - 100 Speakers
  - Engaged 3 countries & UN
  - Approximately 20 Organizations: GMU, Harvard, UCSB, USC, ESRI, USGS, Microsoft, etc.
- Symposiums on Spatiotemporal Thinking, Computing and Applications
  - 9 time at AAG annual meeting
  - Average 10 sessions per year
  - Engaged total 7000 person times
- International Symposium on Spatiotemporal Computing
  - Two times, with this cancelled for the Pandemic
  - Approximately 120 participants each year
  - Publish with ISPRS and indexed by Web of Science

3. Towards Spatiotemporal Science

- Research & Development
  - **Theories such as ST-Texts**
  - **Technologies such as Smart Cities**
  - Applications such as Pandemic
  - Papers
  - Books & Texts
- Education & Curriculum
  - Graduates
  - Undergraduates
  - K-12 & Teachers

Community Building

Expert Meeting
ISSC series
ST symposium
Webinars

- Patents & Licenses
  - **Open Source**
  - **Commercial**
  - **Results adopted by members**
- Create business
  - Students
  - Collaboratives
  - **I-Corps**
  - **STTR/SBIR**
- Entrepreneur
  - **GOALI**
  - **PIRE**
  - INTERN

- Building and strengthening the spatiotemporal infrastructure
- Innovation towards solution & commercialization
- Address grand challenges, e.g., public health, natural hazards, geospatial intelligence

Acknowledgements

- Dr. Wendy Guan/Harvard and Dr. Shuming Bao/CDI
- We greatly thank NSF's I/UCRC program, and the support from our current and past members, NASA Goddard, NCCS, USGS, NASG, NGCC, Harris, Northrop Grumman, Microsoft, USDA, NOAA, UN, NGA, State Dept., Eastview Geospatial, OminiSci, CDI, RMDS Inc., FDL and the institutional support from GMU, Harvard and UCSB.
- We give our special thanks to our NSF program project directors, Rita Rodriguez, Dmitri Perkins, Behrooz Shirazi, our evaluators Donald Price & David Meyer, IAB chairs Daniel Q. Duffy, Myra Bambacus (past chair), and Lynn Usery (past chair).
- We Look forward to **collaborating with you to advance and formalize spatiotemporal sciences**.

4. Questions, Comments, Suggestions, & Discussions

cyang3@gmu.edu

Backup Slides