Environmental Exposome

August 2023

Jinkook Lee

University of Southern California

GATEWAY TO
GLOBAL
AGING
DATA

g2aging.org

Enhance the Gateway with contextual information to understand environmental risk factors of dementia and cognitive health
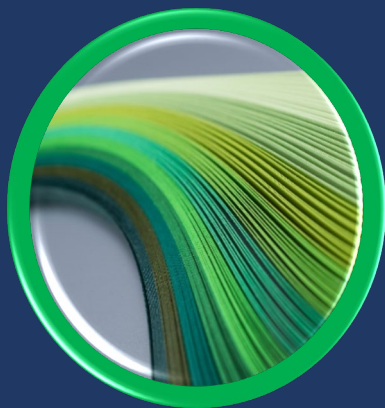
Goal

Specific Aims

Dr. Sara Adar

PROVIDE SCIENCE-READY DATA

Harmonized, high quality information for linkage with study participants based on location

DEVELOP DATA MASKING STRATEGY

Optimize data integrity while protecting participant privacy

CREATE STATISTICAL GUIDANCE

Inform modeling choices when evaluating relationships of exposome with health

From Data Generation to Publication

Data Generation

Quality Assurance and Checks

Transfer of Data

Documentation| Publication
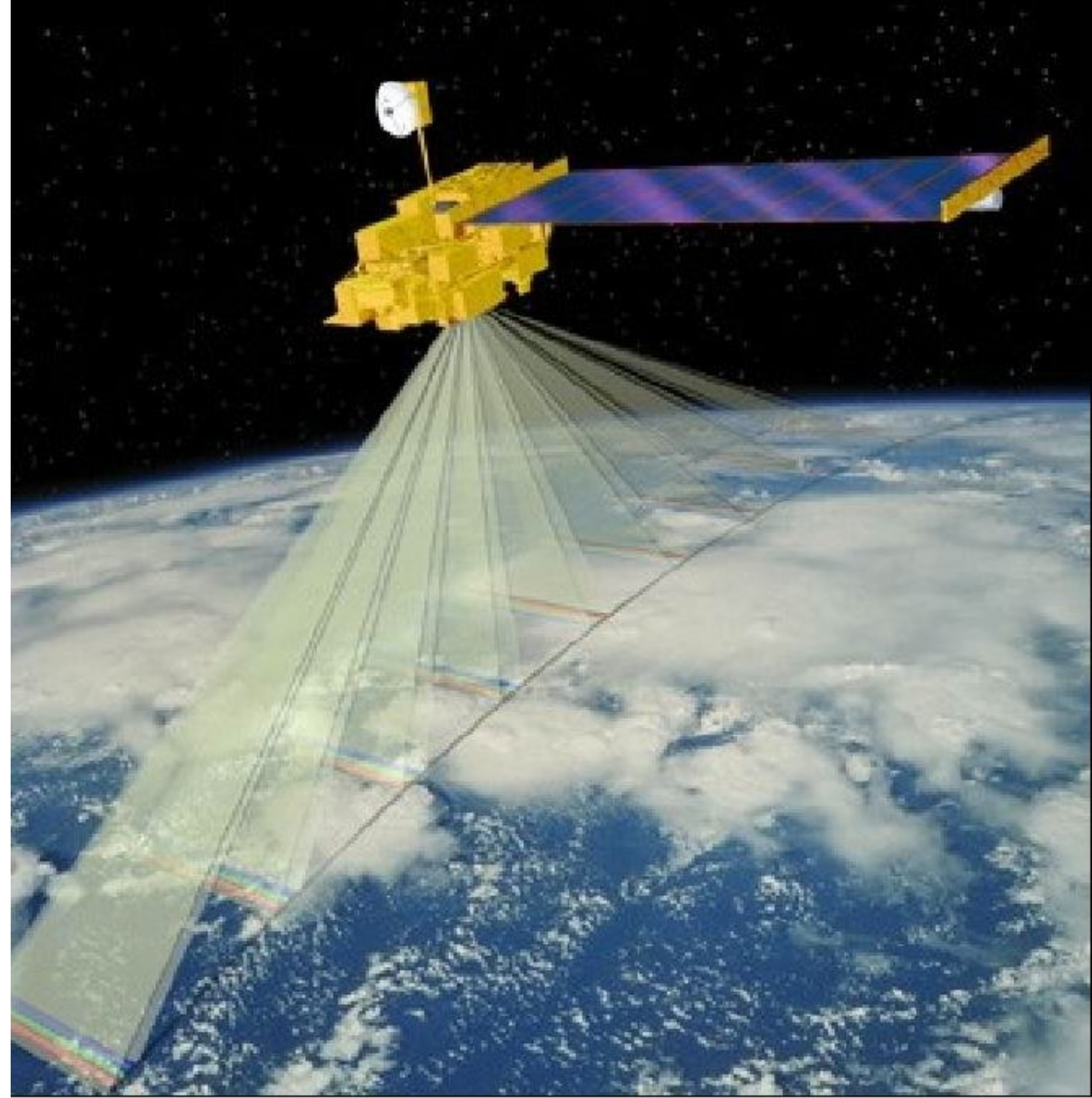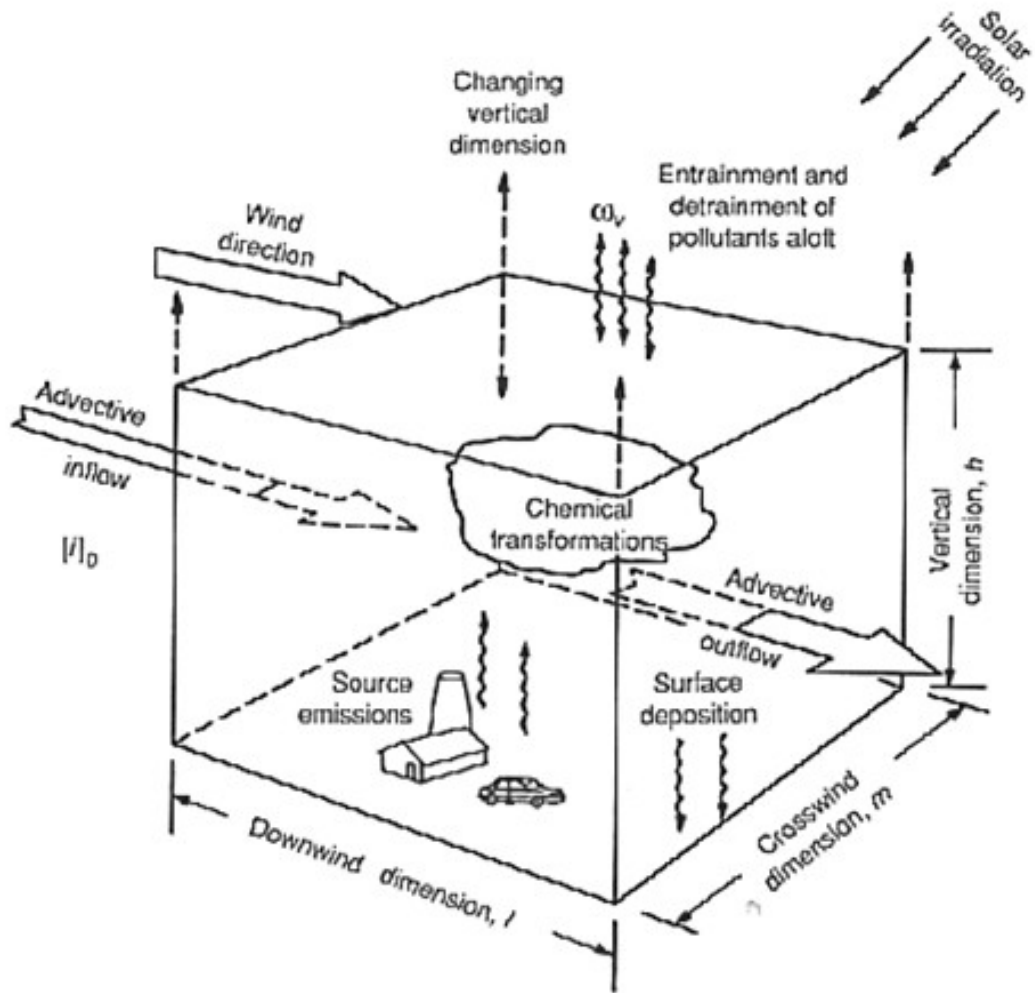
Data Generation to Linkage

Step 1. Generated Country-Specific Datasets

- $PM_{2.5}$ (2010-2019 at 1km resolution), $PM_{2.5}$ source fraction information, NO_2, and O_3 for

- US, India, Ireland, N. Ireland, England, Mexico, and Chile.

- Generated population-weighted descriptive statistics for all datasets

Data Integration Model for Air Quality (DIMAQ)
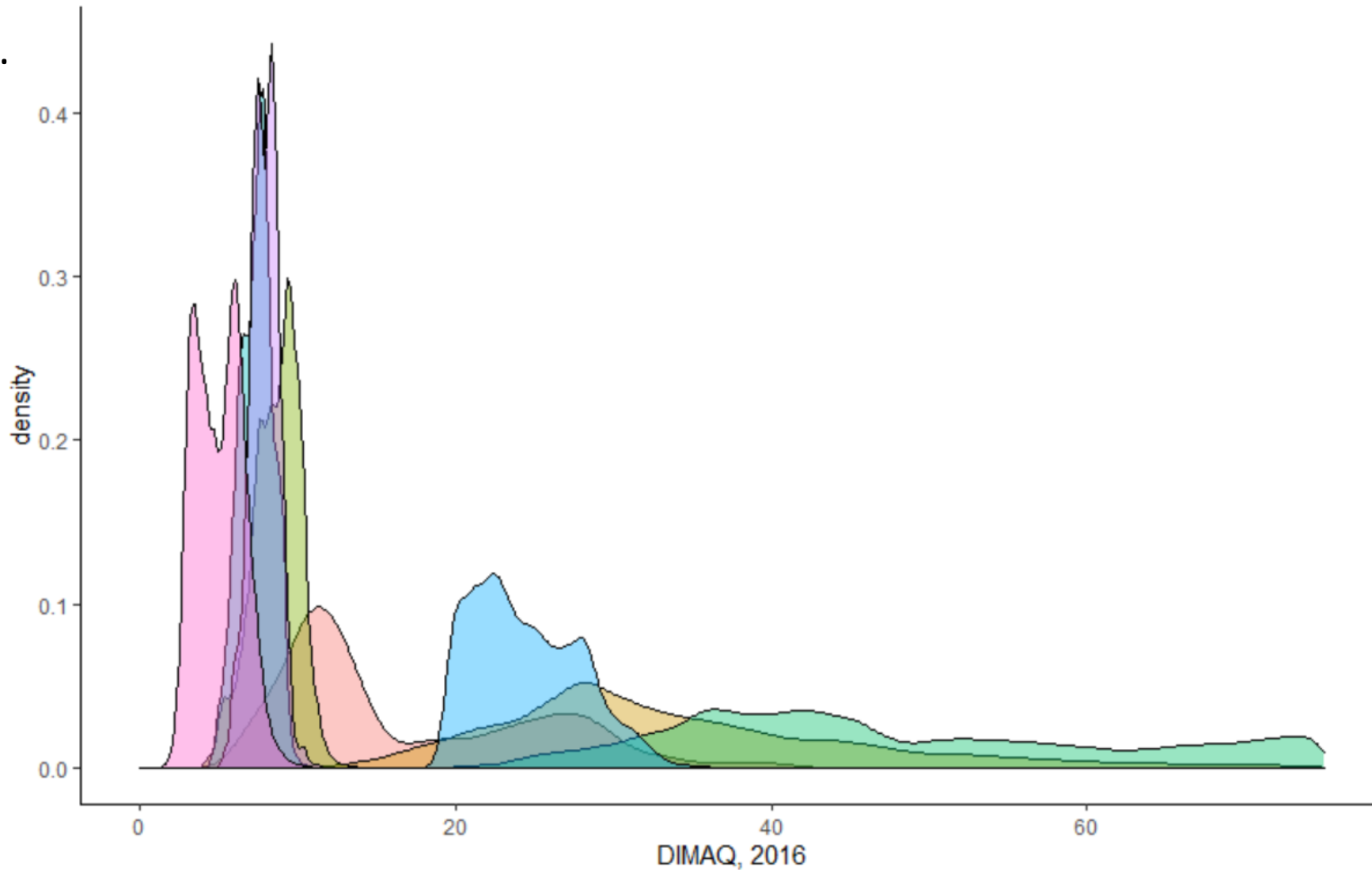
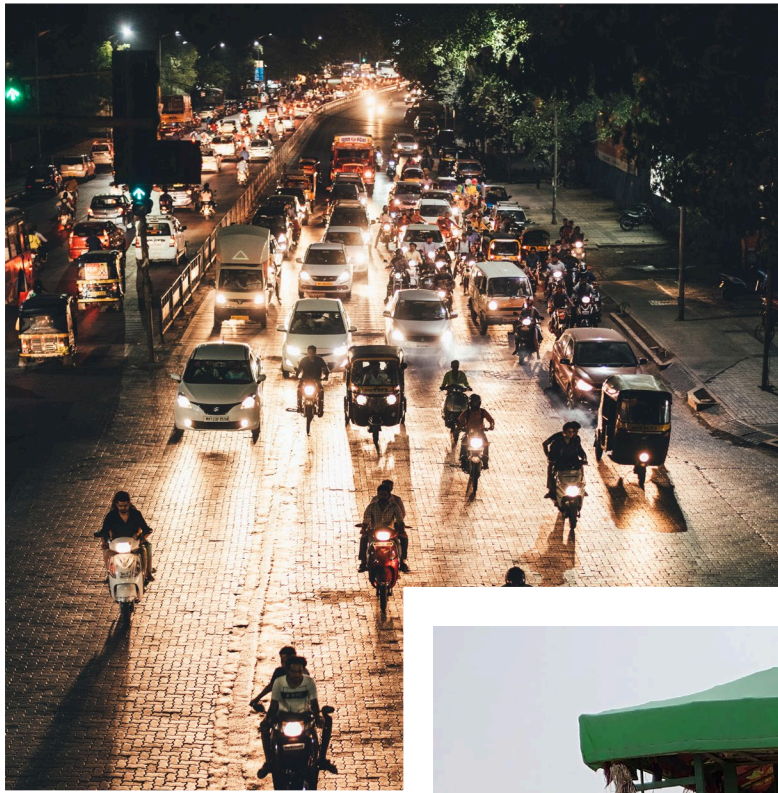Dr. Gavin Shaddick

Harmonized Air Pollution Data, PM$_{2.}$

Dr. Matt Thomas

Initially produced at a 10 km^2 scale globally, but has been refined to 1km^2 for our project (2000-2019)

Source-specific PM$_{2.5}$

(2017)
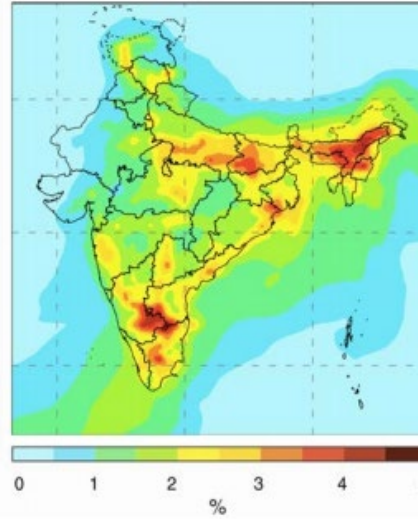
Dr. Michael Brauer
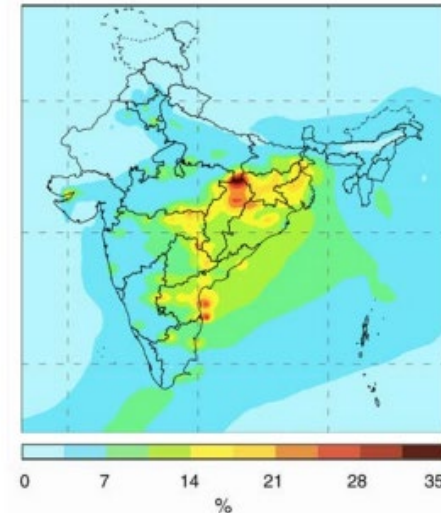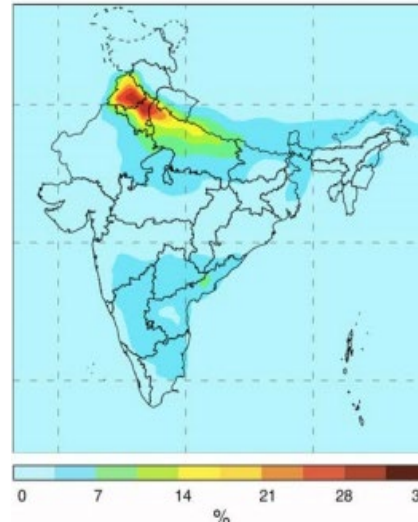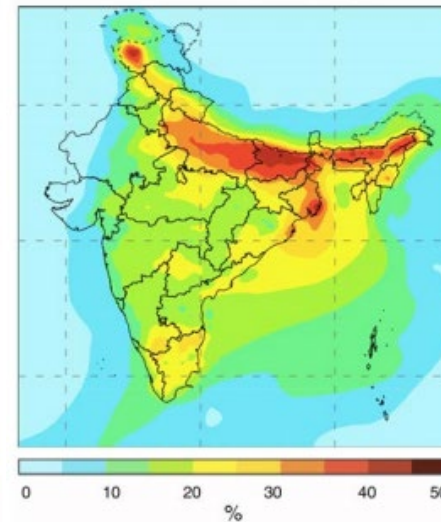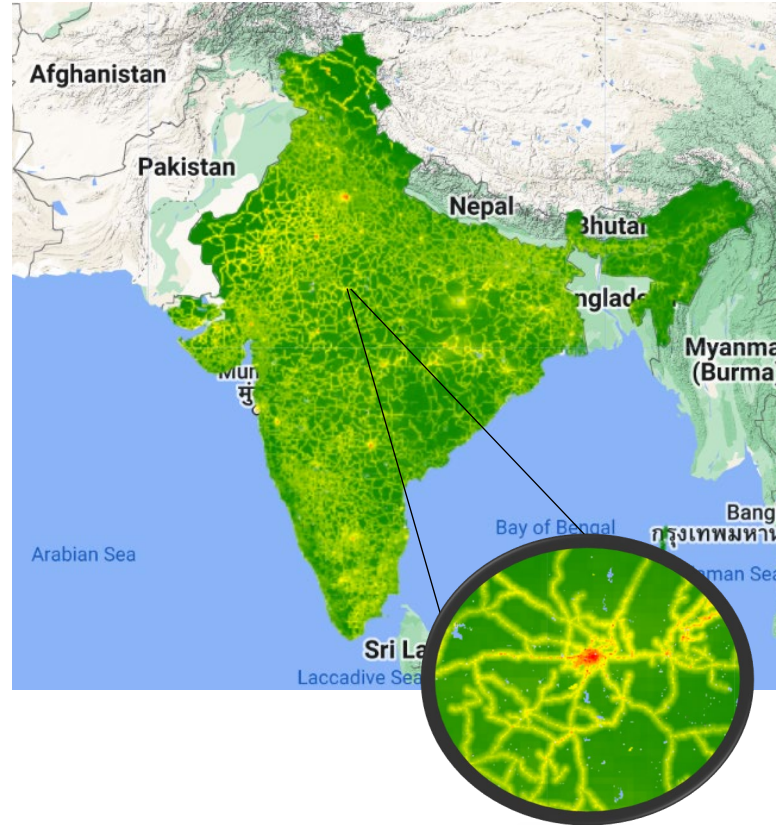
Mohegh & Anenberg. 2020. Figshare.
DOI:10.6084/m9.figshare.12968114.v4.

Dr. Susan Annenberg

Outdoor NO$_2$ Estimation (2005-2020)

Outdoor O$_3$ Estimation (1990-2017)
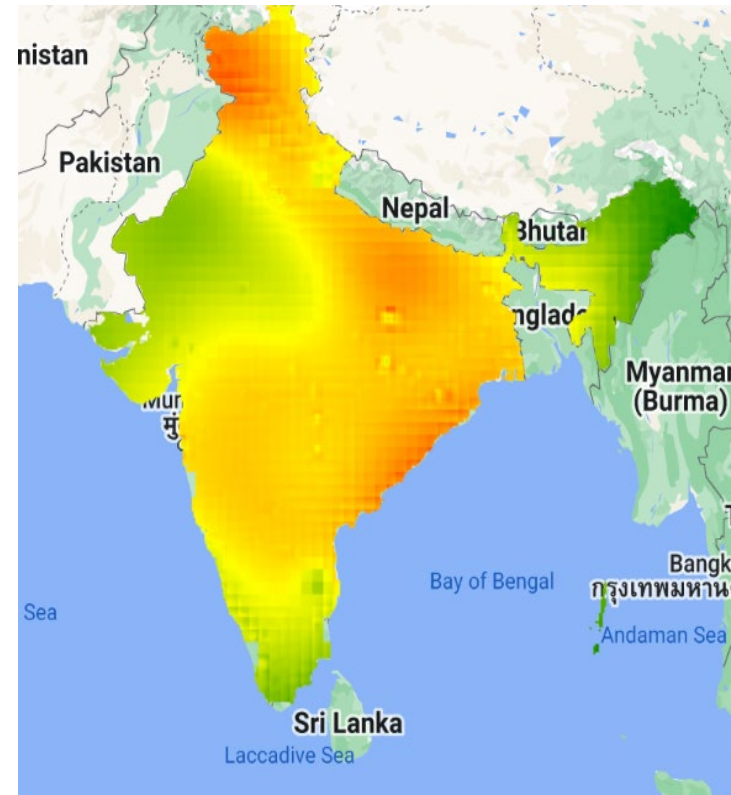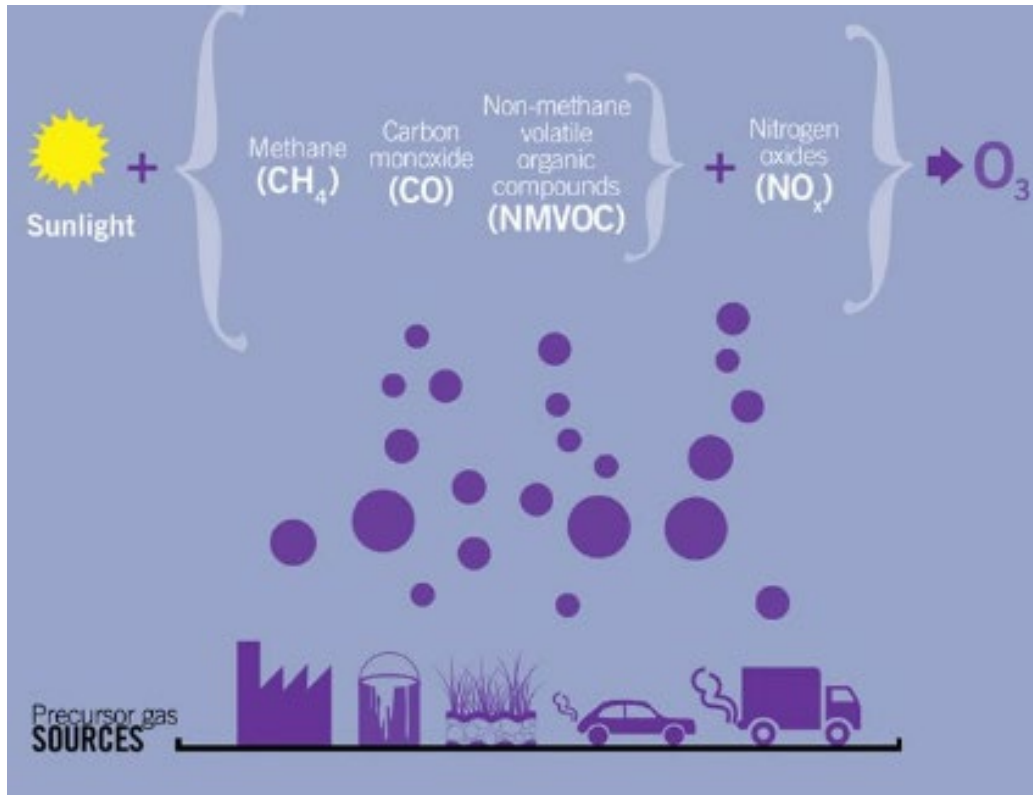
Dr. Jason West

DeLang et al. 2021. Envrionmental Science & Technology. DOI:10.1021/acs.est.0c07742

Data Generation to Linkage

Step 1. Generated Country-Specific Datasets

Step 2. Reviewed and Revised Datasets with Issues

- Identified artifacts in India, Ireland, and N. Ireland PM$_{2.5}$ estimates that have been resolved by Dr. Thomas. Currently investigating unusual features in PM$_{2.5}$ estimates for Mexico.

- Discovered missingness in global O$_3$ data for coastal cities. Initiated a new collaboration and contract with Dr. Jason West at UNC to resolve this issue with remodeling of the coasts underway.

Data Generation to Linkage

Step 1. Generated Country-Specific Datasets

Step 2. Reviewed and Revised Datasets with Issues

Step 3. Linked Exposome Data to Surveys

- HRS, LASI, TILDA, and NICOLA successfully linked cohorts with exposome measures
- In the process of geocoding for ELSA, MHAS, and SPS/Chile-Cog

Team Activities

- *Monthly Conference Calls* – We hosted monthly calls to facilitate harmonization across projects and consensus building around scientific choices for variables and analyses.

- *Individual Survey Support* – We hosted calls with each survey to support their unique challenges with geocoding and data linkage.

- *Centralized Resource Sharing* – Created a project Google Drive for all meeting notes, group decisions, data dissemination, and data documentation.

- *Writing Groups* – Formed writing groups to utilize the Gateway Exposome estimates within the various surveys and across surveys.

- *Capacity Building* – Gateway team hosted trainings for the ELSA and LASI teams regarding geospatial mapping and geocoding.

- *Workshop Planning* – Initiated discussions with MELODEM leadership about an exposome workshop for the Gateway project in the Summer of 2024.

Data Documentation & Website Development
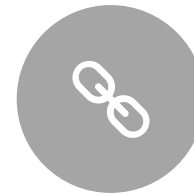
Sample Methods Section

Citation

Specifications

Description

User Notes

Code

Expanding Environmental Exposome

Dr. Ruth Hunter

Pollution:
Nighttime light, noise

Physical environments: greenspace, blue space, walkability

Climate:
temperature, precipitation

Spatial and Temporal Resolutions of Environmental Exposome Data

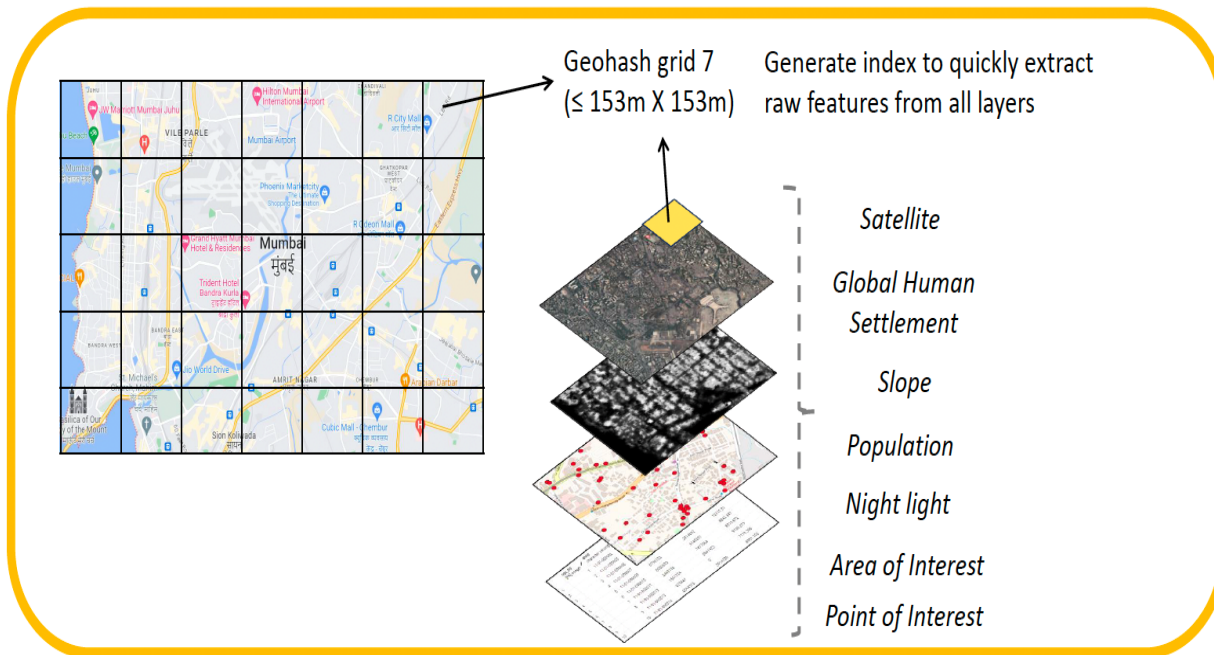| Exposure	Temporal Resolution	Spatial Resolution	Years	Source
Nighttime Light	annual median	~500m	2012-2020	Elvidge et al. 2021, Remote sensing. DOI:10.3390/rs13050922
Greenspace	annual mean, minimum, maximum	250m	2001-2020	Didan. 2015. NASA EOSDIS Land Processes. DOI:10.5067/MODIS/MOD13Q1.006
Bluespace	static	150m	2000-2012	Lamarche et al. 2017, Remote Sensing.DOI:10.3390/rs9010036
Temperature	daily	0.5degree	1981-2020	CPC Global Unified Temperature data, https://psl.noaa.gov
Precipitation	daily	0.5degree	1981-2020	CPC Global Unified Temperature data, https://psl.noaa.gov

Innovations:
Multimodal AI to capture physical environment

Dr. Yao-Yi Chiang.   Min Namgung

Overall Approach – 1. Build Region Foundation Model

Geohash grid 7
(≤ 153m X 153m)   Generate index to quickly extract
raw features from all layers

Satellite

Global Human Settlement

Slope

Population

Night light

Area of Interest

Point of Interest

Goal: Develop region foundation model to fuse geographically related data

Overall Approach – 2 Build Regional Analytic Model

Generate embedding vector
for each data modal

Region Foundation Model

Regional Analytic Model

- **Scene prediction**

- **Socio-economic indicator prediction (e.g., walkability)**
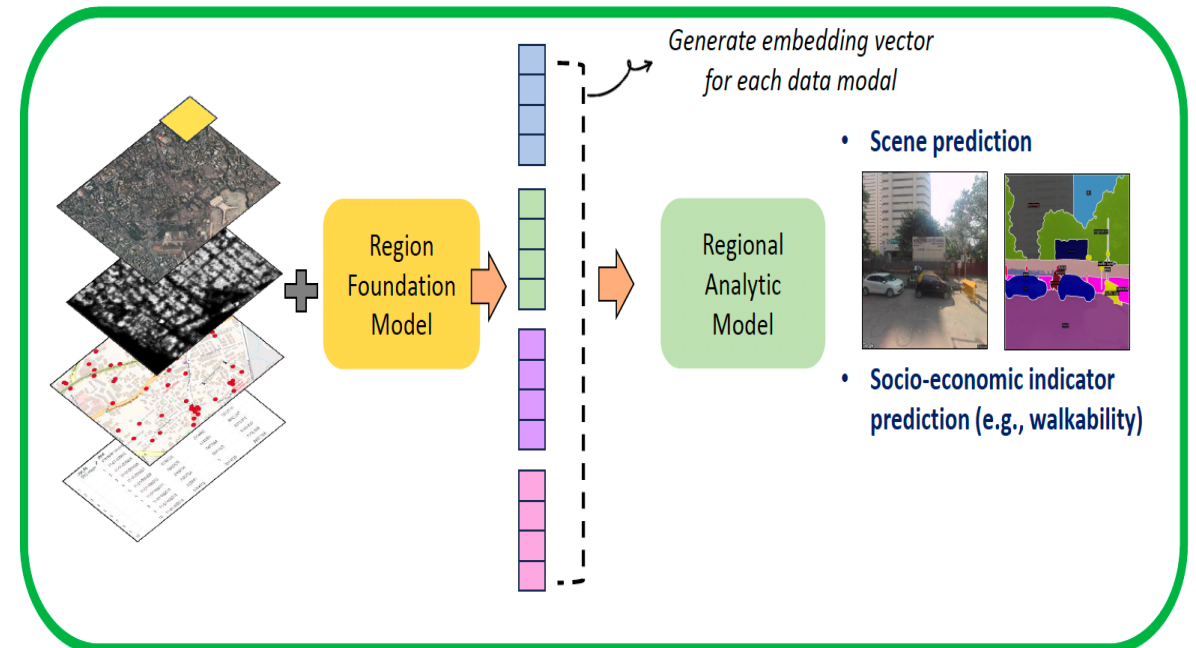
Goal: Build Regional analytic model by adapting region foundation model

CrOss mOdal Knowledge Injected Embedding (COOKIE🍪)
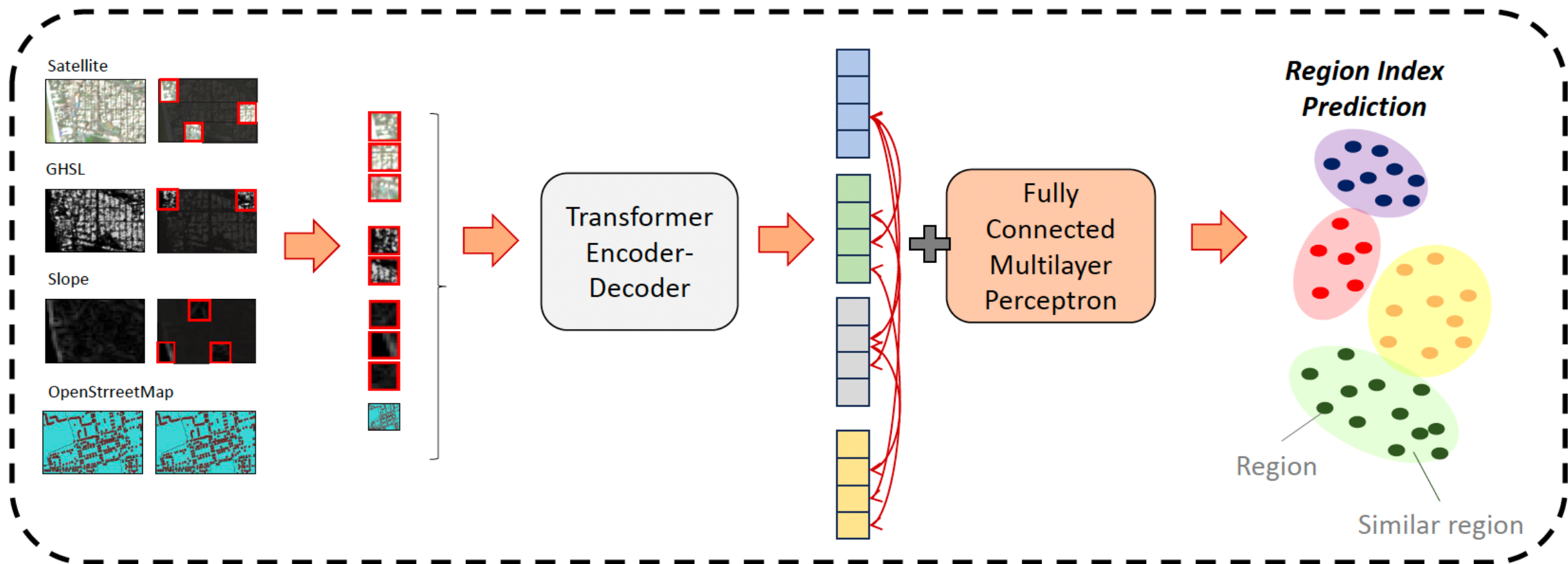
Purpose: Initial Representation Model

Measuring
Indoor Pollution Levels

Measuring Indoor Temperature, Humidity, and CO_2

Publications | Presentations

- PM2.5, NO2, O3 and disability in HRS, under review
- PM2.5 and cognition in LASI, ready for submission
- PM2.5 and cognition in ELSA, presented at the *International Society of Environmental Epidemiology*
- PM2.5 and mental health in TILDA, presented at the *European Health Economics Association*
- Integrating multimodal data and spatial dependence for LASI, to be presented at the *Association for the Advances in Artificial Intelligence*

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Household use of polluting cooking fuels and late-life cognitive function: A harmonized analysis of India, Mexico, and China

Joseph L. Saenz [a,*], Sara D. Adar [b], Yuan S. Zhang [c], Jenny Wilkens [d], Aparajita Chattopadhyay [e], Jinkook Lee [d,f], Rebeca Wong [g]

[a] Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, United States
[b] Department of Epidemiology, University of Michigan, Ann Arbor, MI, United States
[c] Carolina Population Center, University of North Carolina, Chapel Hill, NC, United States
[d] Center for Economic and Social Research, University of Southern California, Los Angeles, CA, United States
[e] Department of Development Studies, International Institute for Population Sciences, Mumbai, India
[f] Department of Economics, University of Southern California, Los Angeles, CA, United States
[g] Sealy Center on Aging, University of Texas Medical Branch, Galveston, TX, United States

JAMA Network

JAMA Internal Medicine | Search All | Enter Search Term

Original Investigation

ONLINE F

August 14, 2023

Comparison of Particulate Air Pollution From Different Emission Sources and Incident Dementia in the US

Boya Zhang, PhD[1]; Jennifer Weuve, ScD[2]; Kenneth M. Langa, MD, PhD[3,4,5,6]; et al

» Author Affiliations

Method Development

- *Geomasking* – Conducted simulation studies to compare three different geomasking options for maintaining respondent confidentiality and evaluating the impacts on statistical inference.

- *Spatial confounding* – Conducting simulation studies to assess the impacts of adjustment for space in the US, England, India, Ireland, and N. Ireland.

Dr. Kayleigh Keller

GATEWAY TO
GLOBAL
AGING
DATA

g2aging.org

NIA/NIH 2R01 AG030153

Thank You!