

Air Pollution Exposure Assessment for Epidemiological Health Studies

Rima Habre, ScD

habre@usc.edu

Assistant Professor, Division of Environmental Health, University of Southern California

Visiting Scientist, Harvard TH Chan School of Public Health

NASA Health and Air Quality Applied Sciences Team Meeting, Jul 16, 2018

LA DREAMERs, USC ECHO Center

Santa Barbara

“Lifecourse Approach to Developmental Repercussions of Environmental Agents on Metabolic and Respiratory health”

Director of Exposure Assessment, LA DREAMERs (PIs Breton and Gilliland)

Co-Chair Air Pollution Subcommittee, NIH ECHO Program



Pregnancy cohort study

Children's cohort study

USC CHILDREN'S HEALTH STUDY

UNIVERSITY OF SOUTHERN CALIFORNIA



ECHO

Environmental influences
on Child Health Outcomes

A program supported by the NIH

USC and partner institutions awarded \$6 million children's environmental health grant from NIH

Posted on Wednesday, September 21, 2016 by Scehsc UscEdu

Share

 Tweet

Research to look at prenatal and early life environmental influences on lifetime health related to asthma and obesity

Staff Report

Media Coverage: Press Enterprise, Daily News, The Washington Times, Redding.com, Daily Breeze, Pasadena Star News

LOS ANGELES – September 21, 2016 – Researchers at the Keck School of Medicine of USC have been awarded a 2 year \$6 million grant, as the first phase of a large seven-year National Institutes of Health, Environmental Influences on Child Health Outcomes (ECHO) initiative involving more than 30 research entities. The USC based research team will investigate health issues related to asthma and obesity.



**LA DREAMERS
USC ECHO
Cohorts**

USC CHILDREN'S HEALTH STUDY
UNIVERSITY OF SOUTHERN CALIFORNIA

CHS by the numbers

- 22 years
- 270 + studies published
- 11,689 Children
- 25 investigators
- 28 field team staff...

<https://healthstudy.usc.edu/>



ECHO

Environmental influences
on Child Health Outcomes

A program supported by the NIH

ECHO—A Nationwide Program

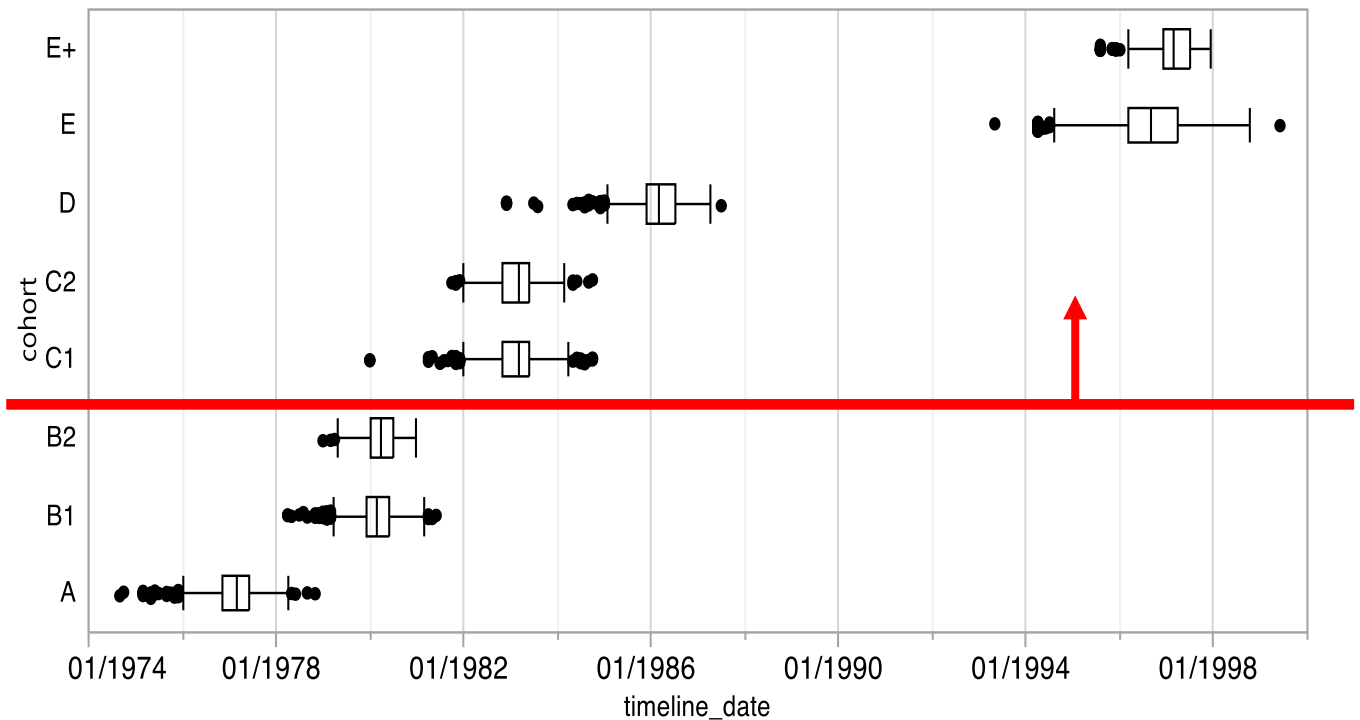


CHS Cohort(s), ~12,000 children

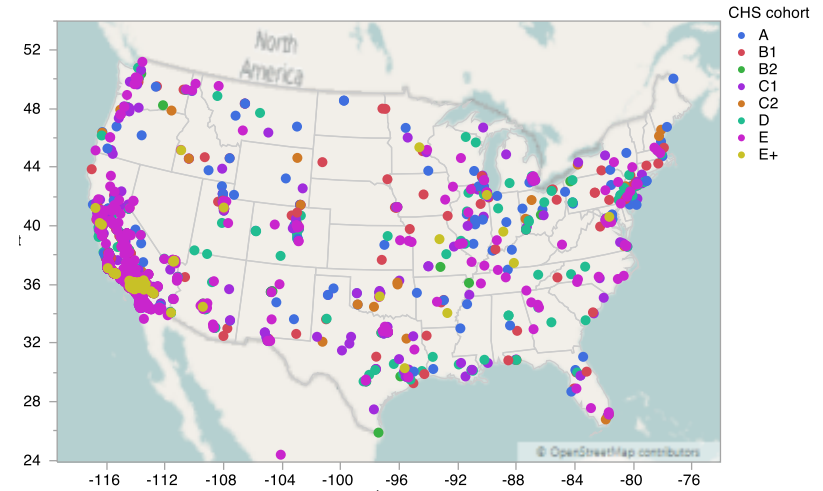
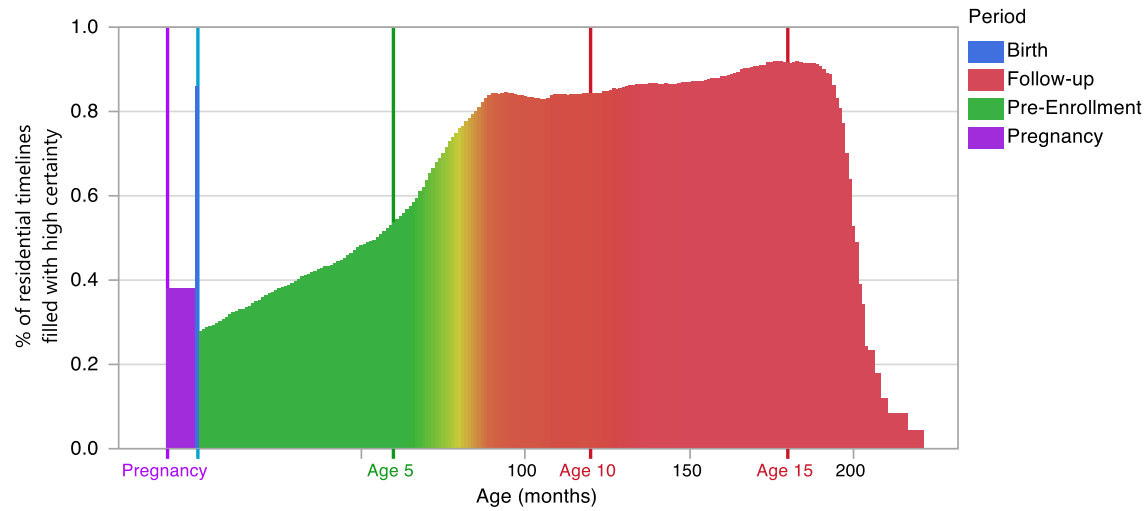


<https://healthstudy.usc.edu/>

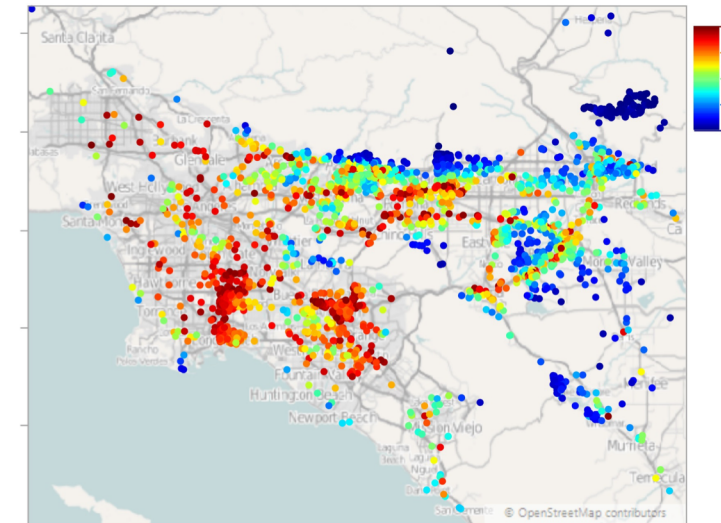
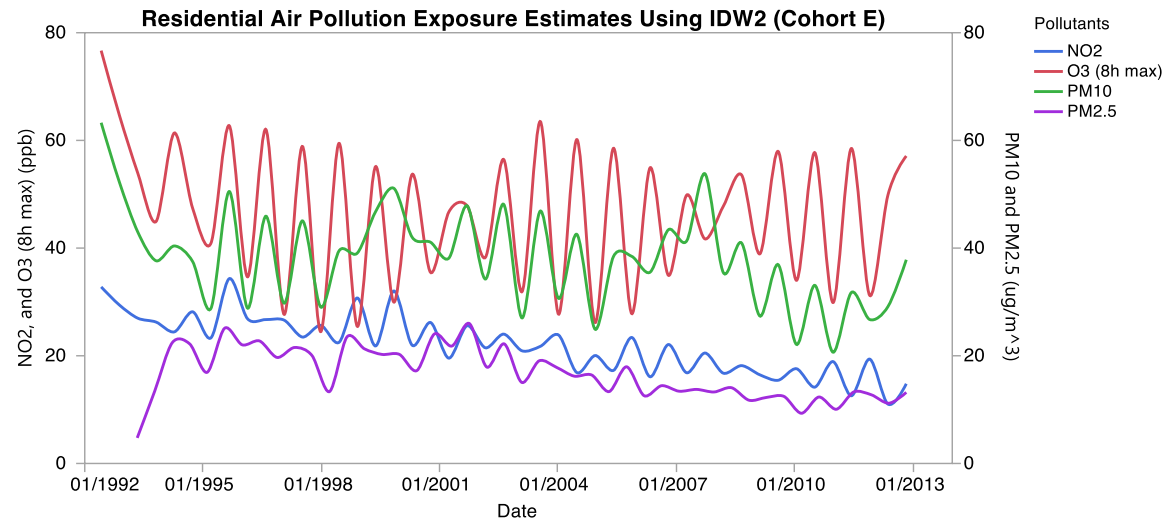
Pregnancy Periods in CHS Cohorts A-E



1) Lifetime residential histories assembled and geocoded in the Children's Health Study



2) Residential air pollution exposure estimated (regional lifetime estimates, left; traffic-related pregnancy estimates, right)

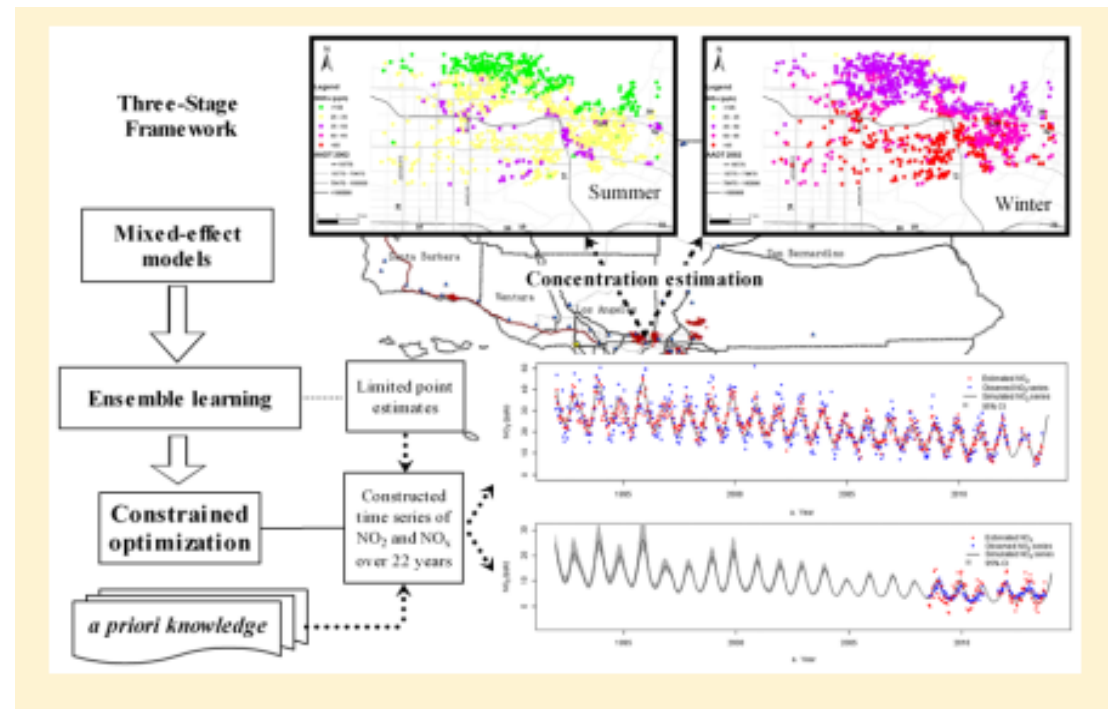


NO₂/NO_x Models

- Novel, three-stage spatiotemporal modeling framework to estimate short-term (two-week average) NO₂/NO_x concentrations and uncertainties in their prediction over 22 years (1992-2013)
- SoCal model CV R²: NO₂ 0.85 and NO_x 0.86 (Li et al. 2017)
- California model CV R²: NO₂ 0.88 and NO_x 0.88 (Li et al. 2018, in submission)

Constrained Mixed-Effect Models with Ensemble Learning for Prediction of Nitrogen Oxides Concentrations at High Spatiotemporal Resolution

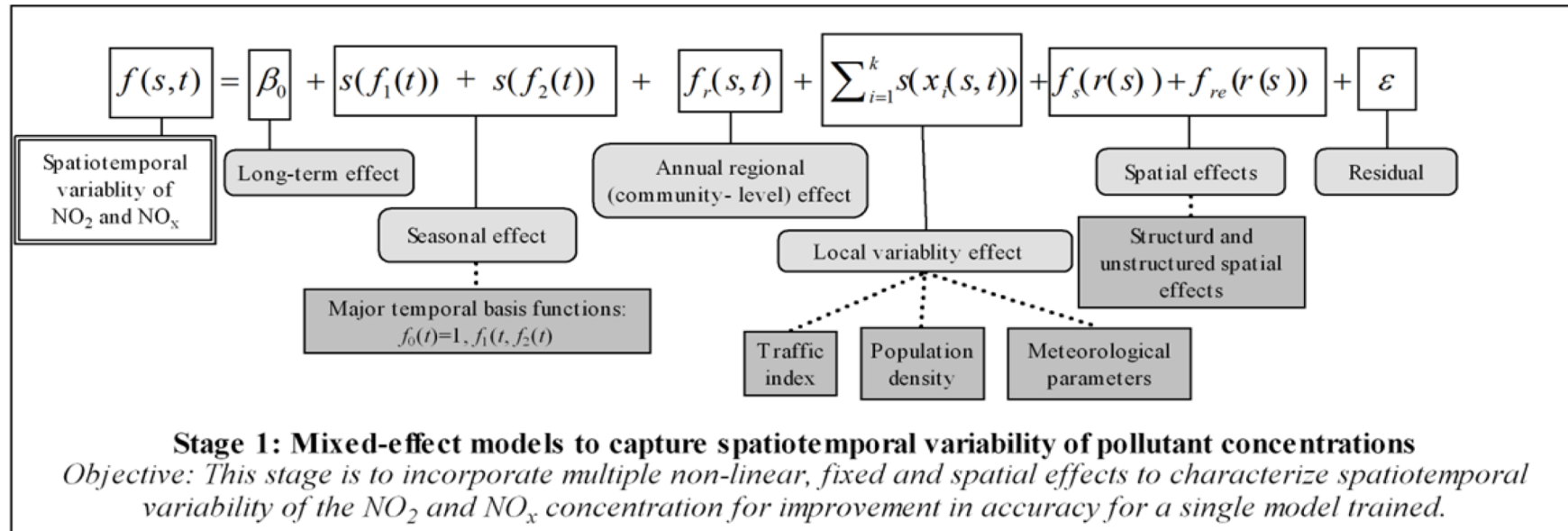
Lianfa Li,^{*,†,‡,§} Fred Lurmann,[§] Rima Habre,^{*,†} Robert Urman,[†] Edward Rappaport,[†] Beate Ritz,^{||} Jiu-Chiuan Chen,[†] Frank D. Gilliland,[†] and Jun Wu^{*,⊥}



Li L, Lurmann F, Habre R et al. Constrained mixed-effect models with ensemble learning for prediction of nitrogen oxides concentrations at high spatiotemporal resolution, Environ Sci Technol., 2017, 5:9920-9929

Model Framework

Stage 1



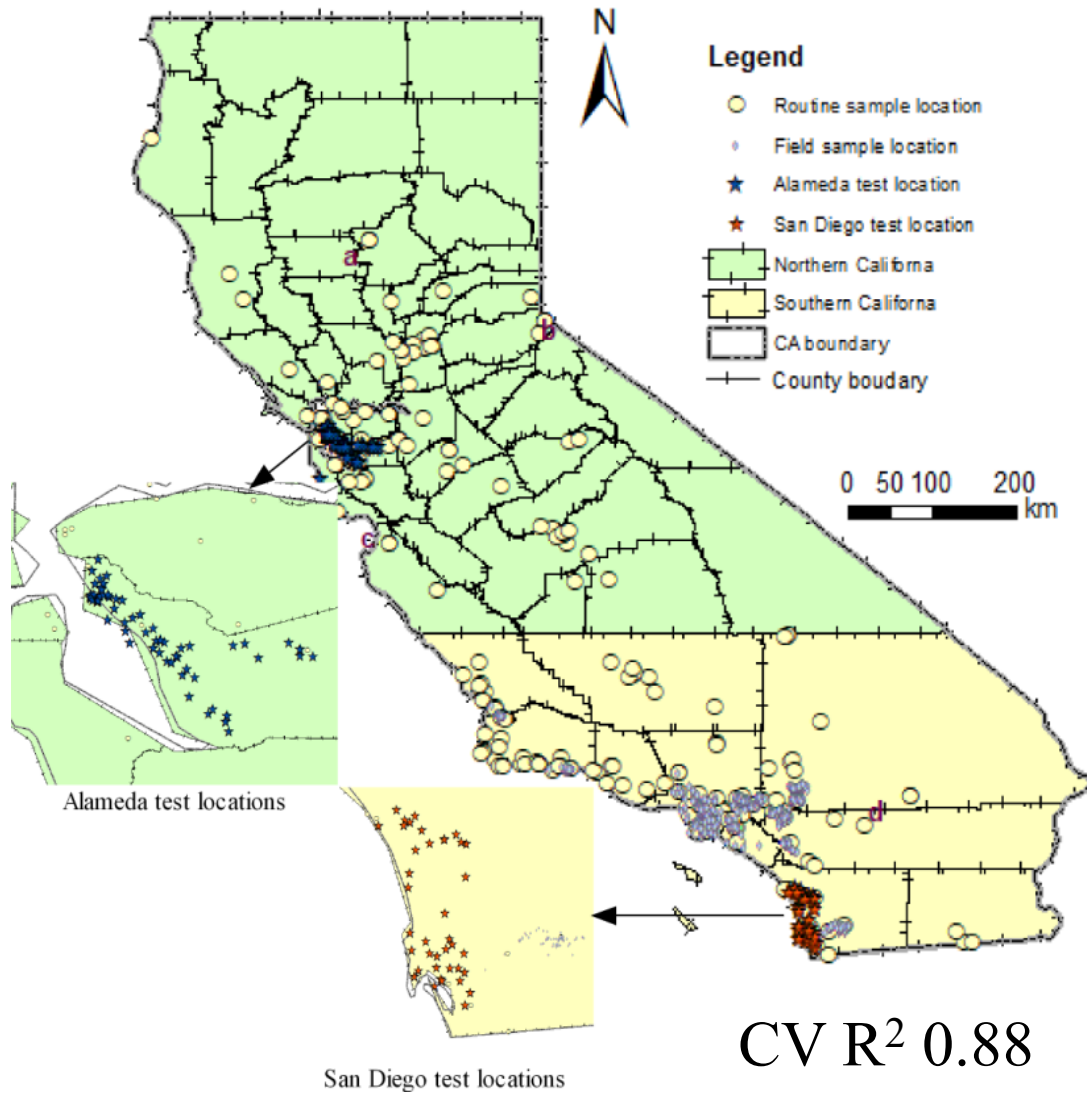
Stage 2

Stage 2: Ensemble learning to reduce uncertainty and variability in point prediction.
Objective: This stage is to leverage bootstrap aggregation to reduce variance in the prediction based on the models from Stage 1. The output includes the averaged estimates and its uncertainty metric (standard deviation).

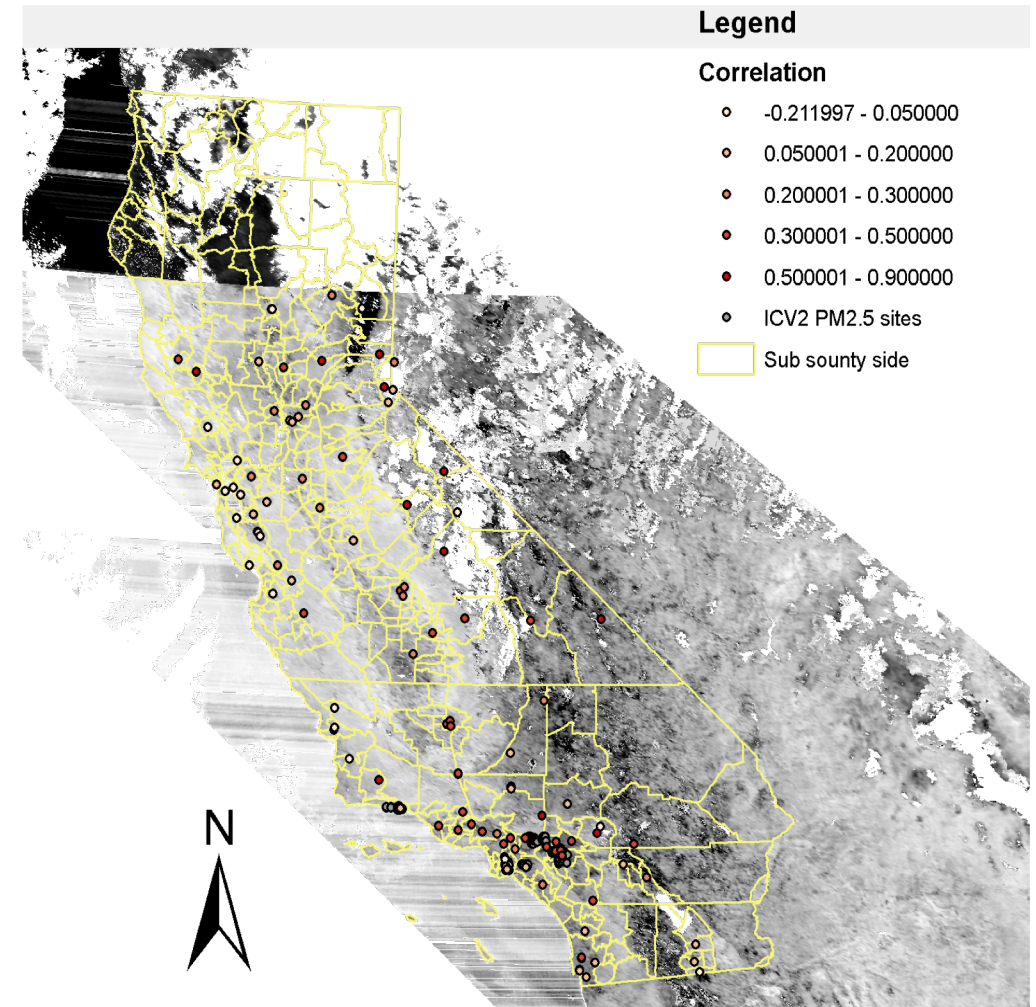
Stage 3

Stage 3: Constrained optimization to help with long-term continuous time series estimation
Objective: This stage is to leverage a priori knowledge, and limited aggregated estimates from Stage 2 to construct a stable, long-term, and continuous time series for the concentrations of a target location over the entire study period.

Li et. Al (2018, *in submission*) California Spatiotemporal NO_x Model



Li et. Al (2018, *in progress*) California Spatiotemporal PM_{2.5} Model

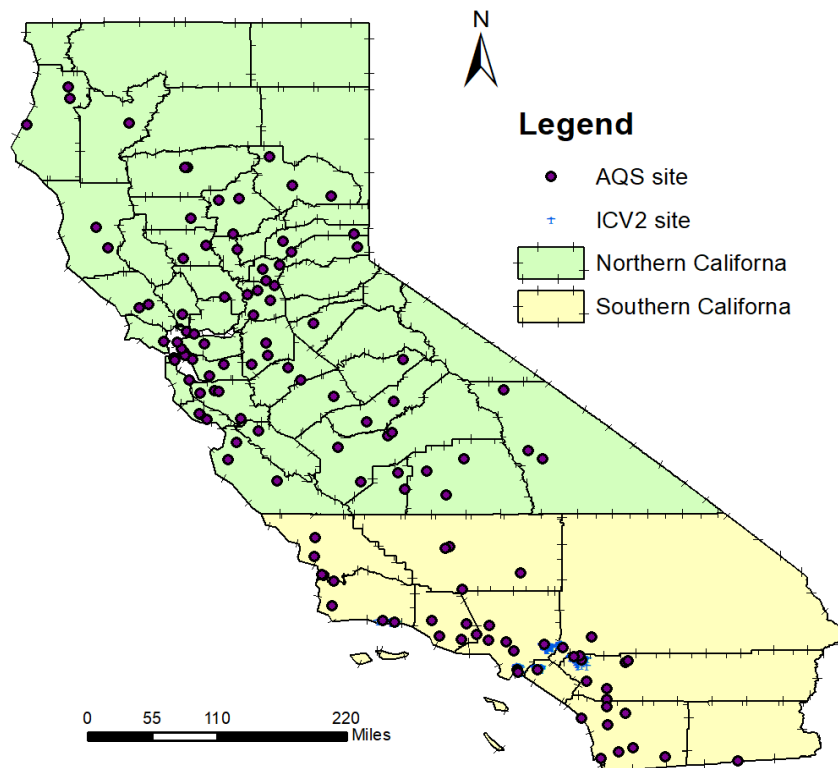


Li et. Al (2018, *in progress*) California Spatiotemporal PM_{2.5} Model

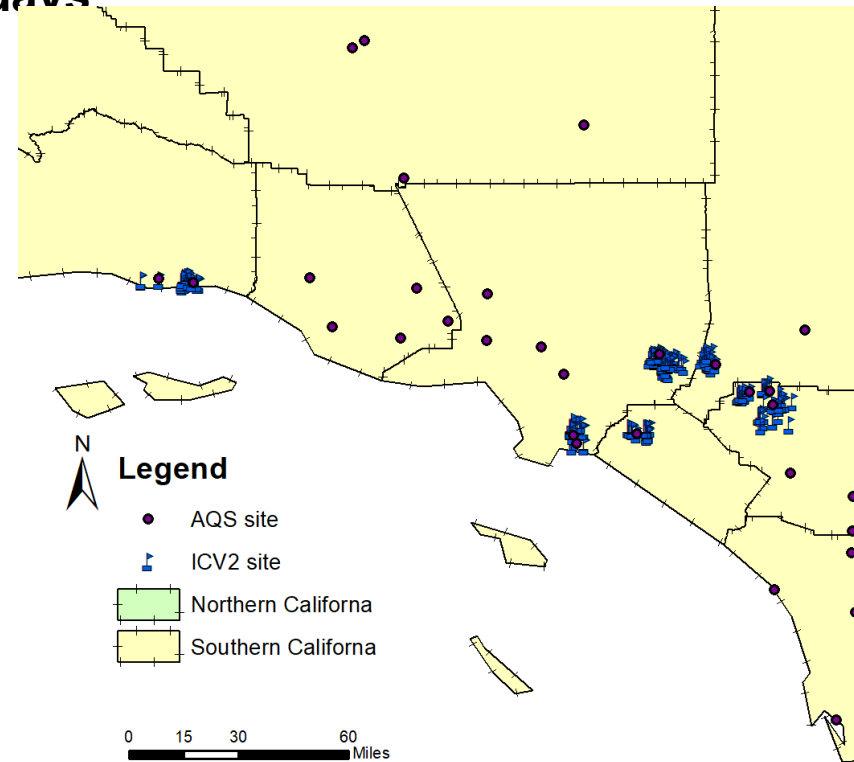
- Building two PM_{2.5} models, weekly
 - ~2000 onwards using MAIAC 1km x 1km data
 - ~1980 onwards using MERRA-2 50km x 50km data
- Fitting temporal basis functions to 50km x 50km MERRA-2 grid cells to reconstruct long term time series (1980's onwards)
- Using MERRA-2 AOD to help interpolate missing MAIAC AOD using machine learning approaches
- Incorporating MERRA-2 GMI Replay Simulation AOD/surface-level PM_{2.5}

PM_{2.5} Measurements (2000+ model)

127 monitoring AQS stations for
California with 160,652 daily averages



278 sampling locations for USC ICV2
with 937 measurements of 12 to 16
days



MAIAC AOD, 1km x 1km

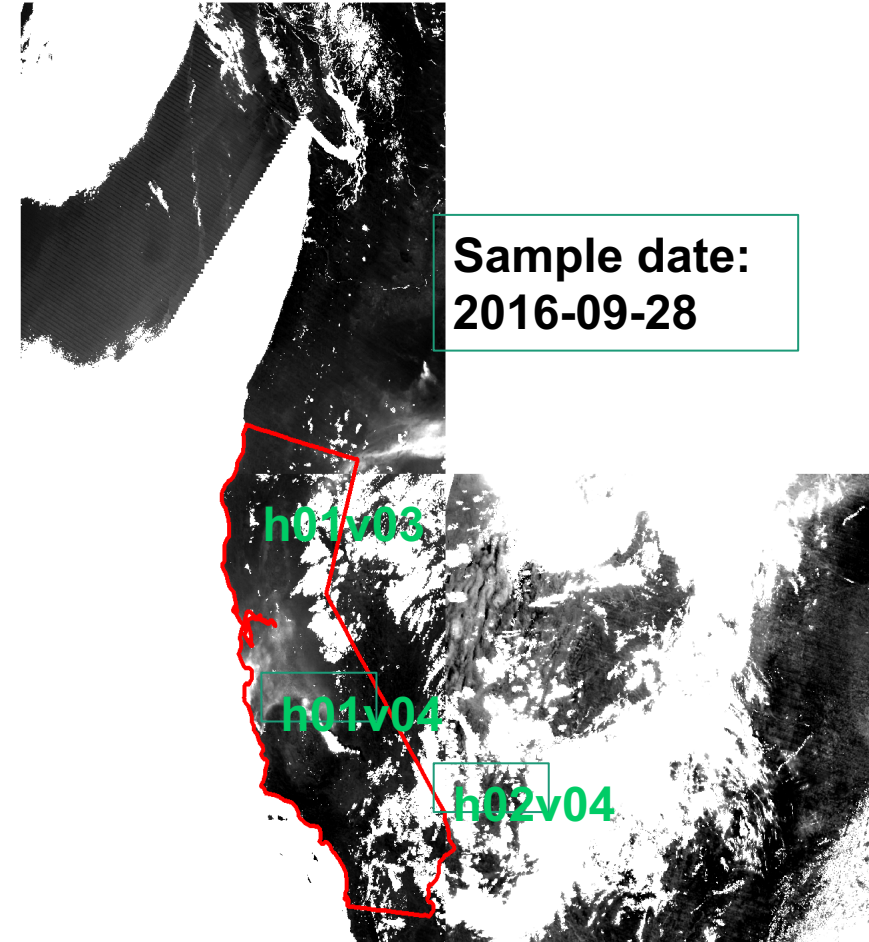


MAIAC 2000-2016 CA Tiles

H01V03 small northern region of CA

H01V04 large area of CA (southern CA)

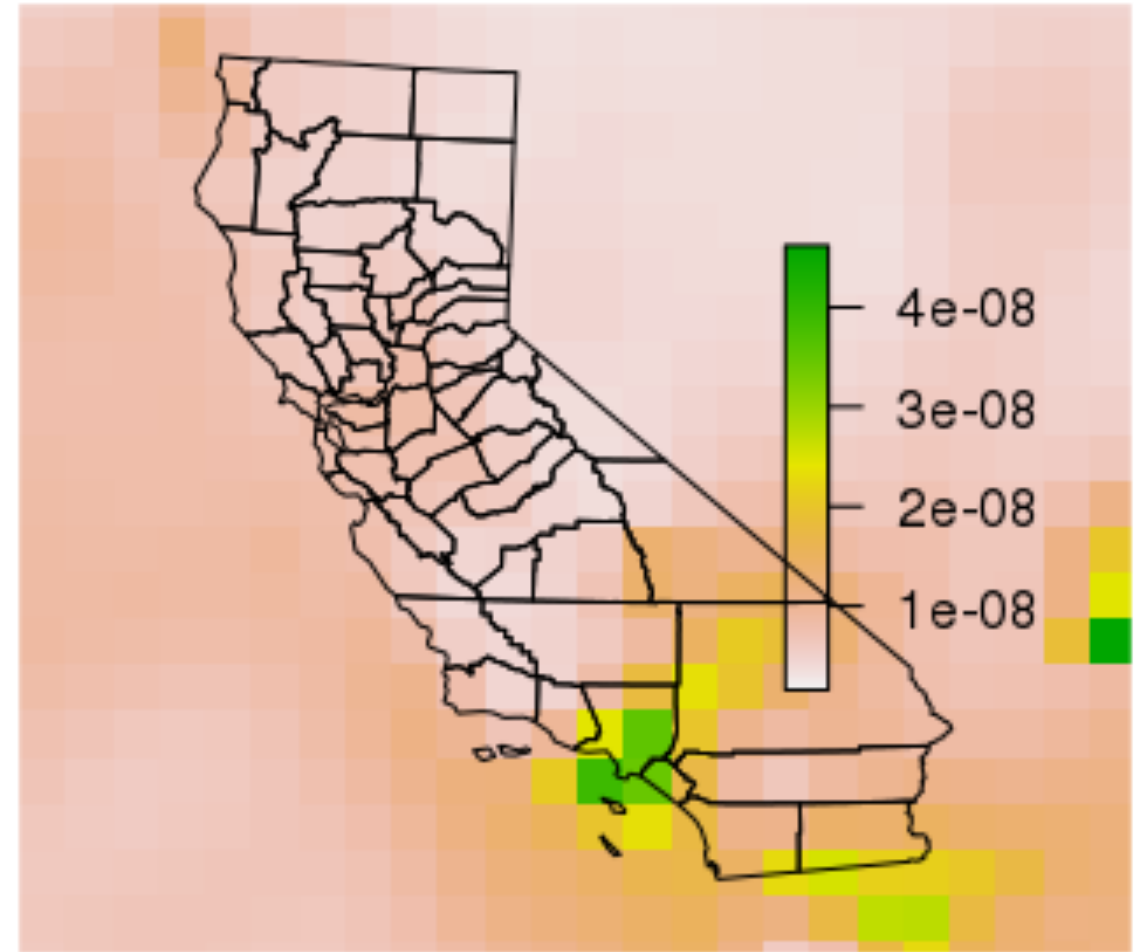
H02V04 covering a small corner of CA



MERRA-2 GMI Replay Simulation

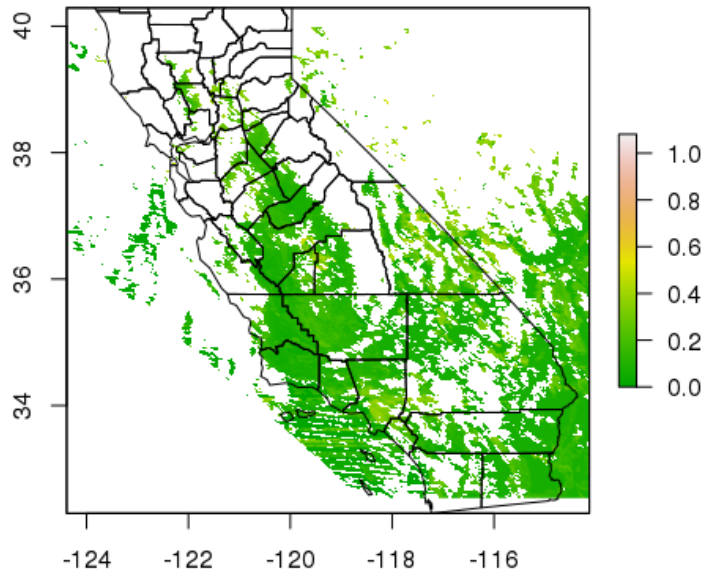
Surface-level $PM_{2.5}$

- 57 km x 57 km spatial
- Hourly temporal
- Sample date 2006-06-07
- $PM_{2.5_BOT}$ (kg/m^3): Total reconstructed $PM_{2.5}$

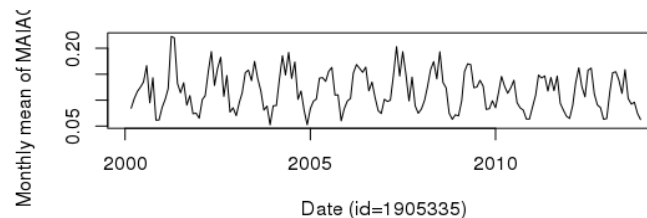
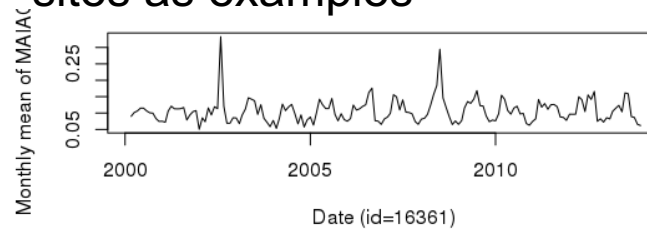


Using monthly mean MAIAC AOD, construct temporal basis functions

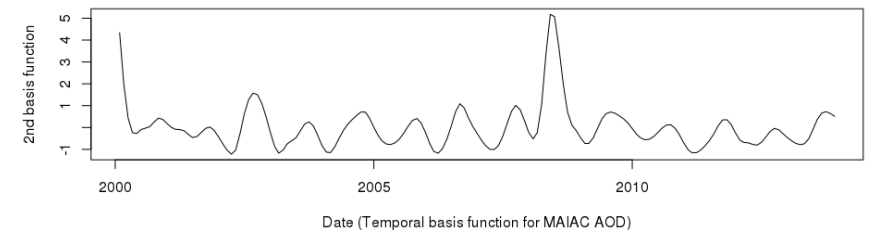
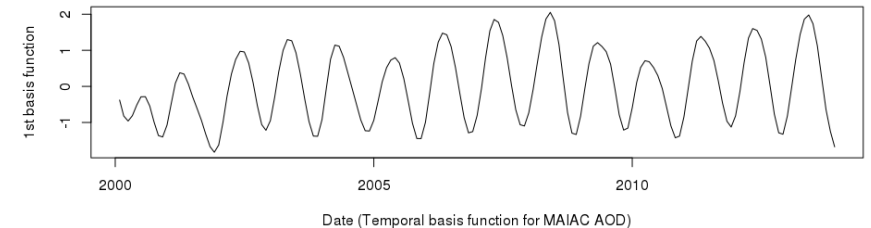
February 2000



Monthly means at different
sites as examples



Temporal basis functions generated
from random sample ($n=10,000$),
reflect regional long-term temporal
variation, we expect regional
differences



Imputation of missing daily MAIAC AOD

- Currently using coordinates (space), elevation, air temperature, wind speed, specific humidity, surface shortwave radiation flux, MERRA-2 daily AOD, and monthly MAIAC AOD, day of week index

$$\begin{aligned} AOD_{jt} = & s(X_j) + s(Y_j) + s(X_j^2) + s(Y_j^2) + s(X_j \times Y_j) + \beta_1 CF \\ & + \beta_2 CMTAOD_{jt} + \beta_3 + Temp_{jt} + \beta_4 RH_{jt} + \beta_5 SH_{jt} \\ & + \beta_6 Ele_j + D_t + \varepsilon_{jt} \end{aligned}$$

Coordinates: x , y , x^2 , y^2 and $x*y$ to capture spatial autocorrelation;

Air temperature: min and max air temperature, difference between minimum and maximum temperatures;

Specific humidity;

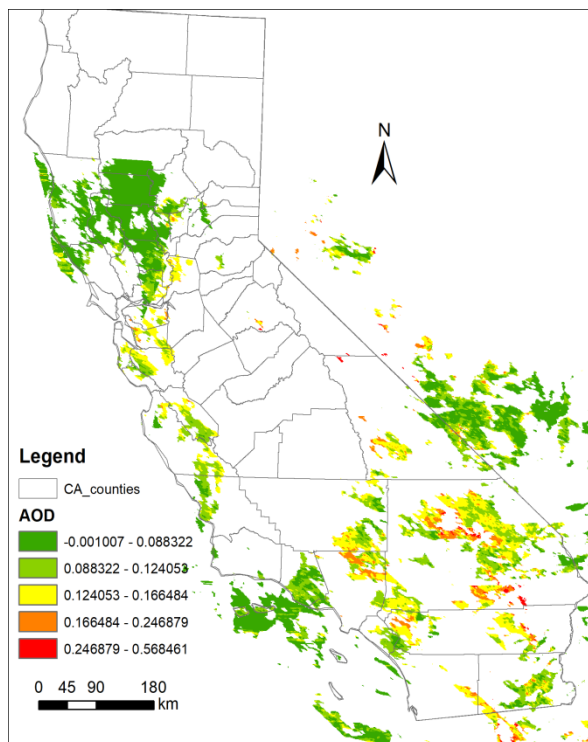
Surface downwelling shortwave flux in air;

Elevation;

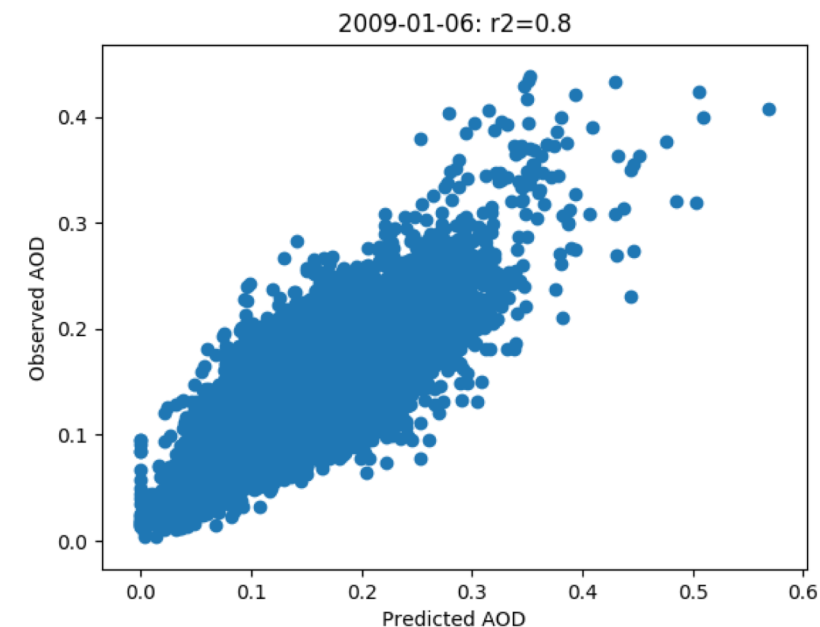
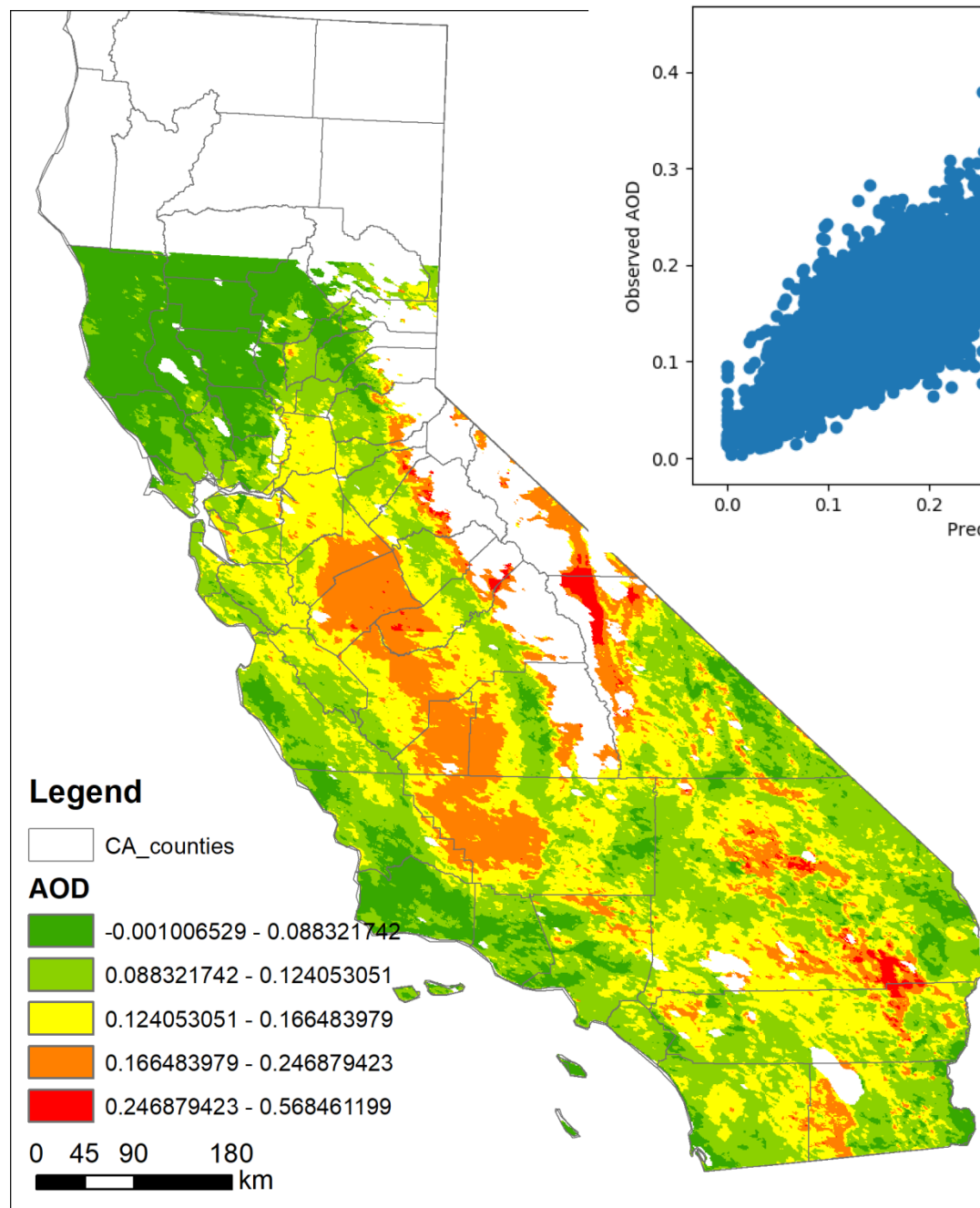
MERRA2 reanalysis AOD: background AOD predictor ;

Monthly mean AOD: to capture temporal trend and spatial difference;

Day index: dummy variable for day index (1-7)



Original image (2009-01-06)



Imputed image:
Several places
still not
imputed due to
snow or water

Preliminary findings, please do not cite or circulate.

Thank You

- Lianfa Li, postdoctoral research associate
- Mariam Girguis, postdoctoral research associate
- ECHO exposure assessment team
 - Fred Lurmann, Jun Wu, Meredith Franklin, Nathan Pavlovic
- LA DREAMERs PIs
 - Carrie Breton and Frank Gilliland