Global, high-resolution statistical modelling of NO_2

Meng Lu[1], Oliver Schmitz[1], Kees de Hoogh[2,3], Perry Hystad[4], Luke Knibbs[5], Derek Karssenberg[1]

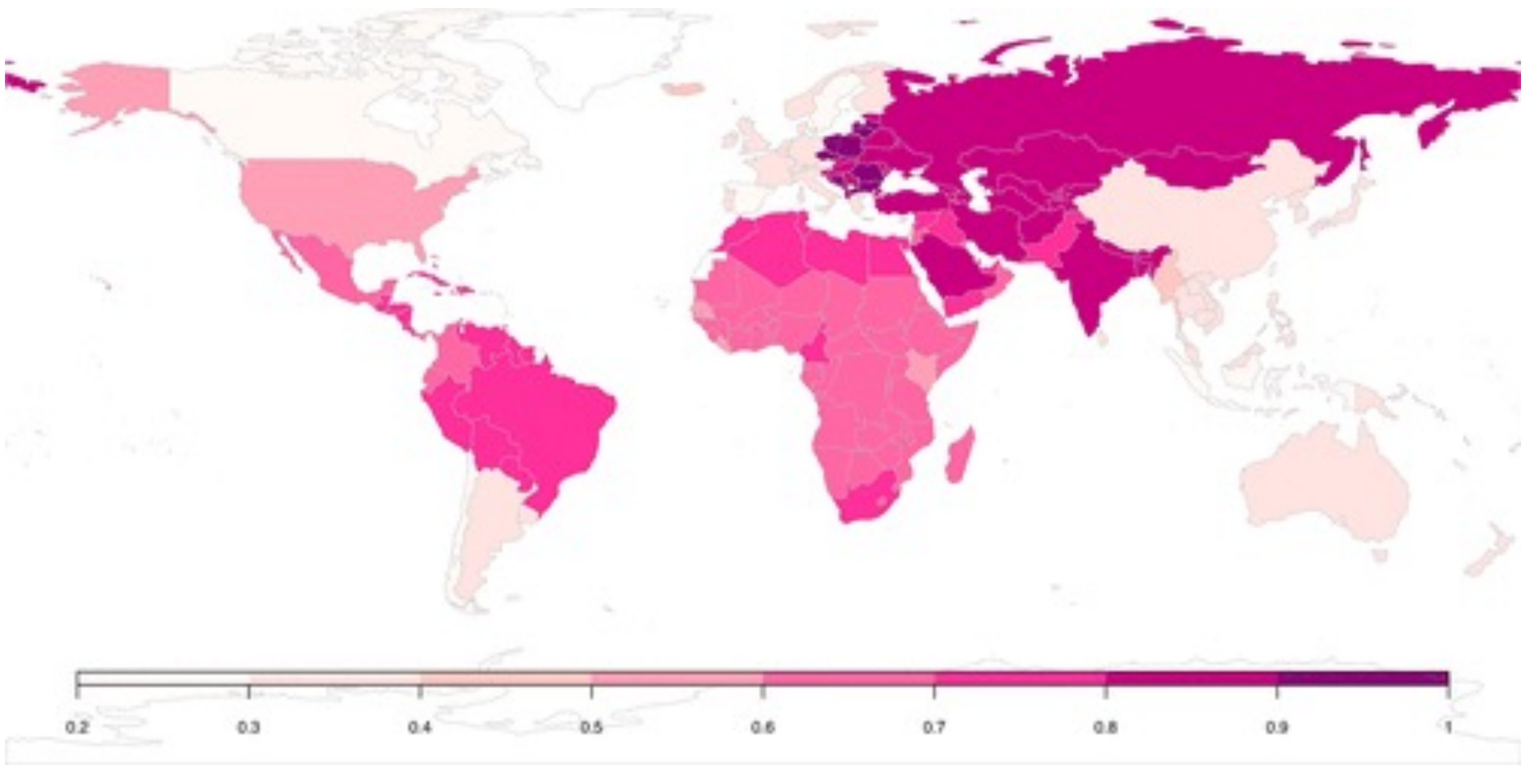
1 Department of Physical Geography, Utrecht University, the Netherlands
2 Swiss Tropical and Public Health Institute, Basel, Switzerland
3 University of Basel, Basel, Switzerland
4 Oregon state University, USA
5 The University of Queensland, Australia

Goal:
Global, high spatiotemporal resolution NO_2 mapping

NO_2: is part of a group of gaseous air pollutants produced as a result of road traffic and other fossil fuel combustion processes.

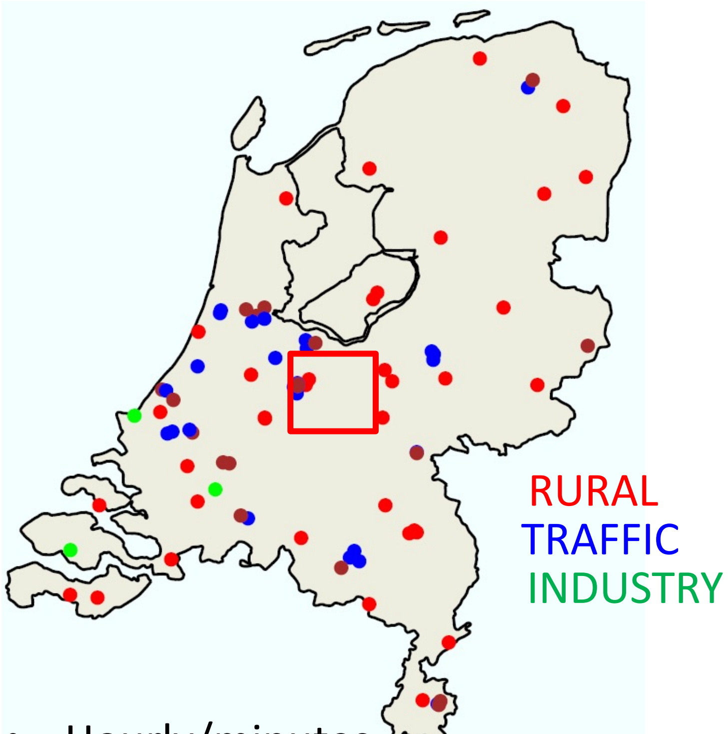
Why global?

- Lacking of ground monitoring stations in many countries.
- Consistent epidemiological studies world-wide
  Inequality in health and air pollution monitors in middle and low income countries

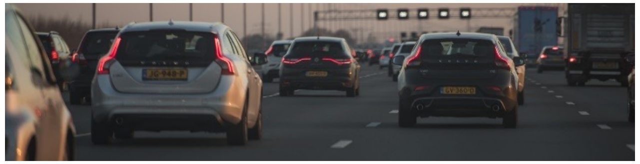


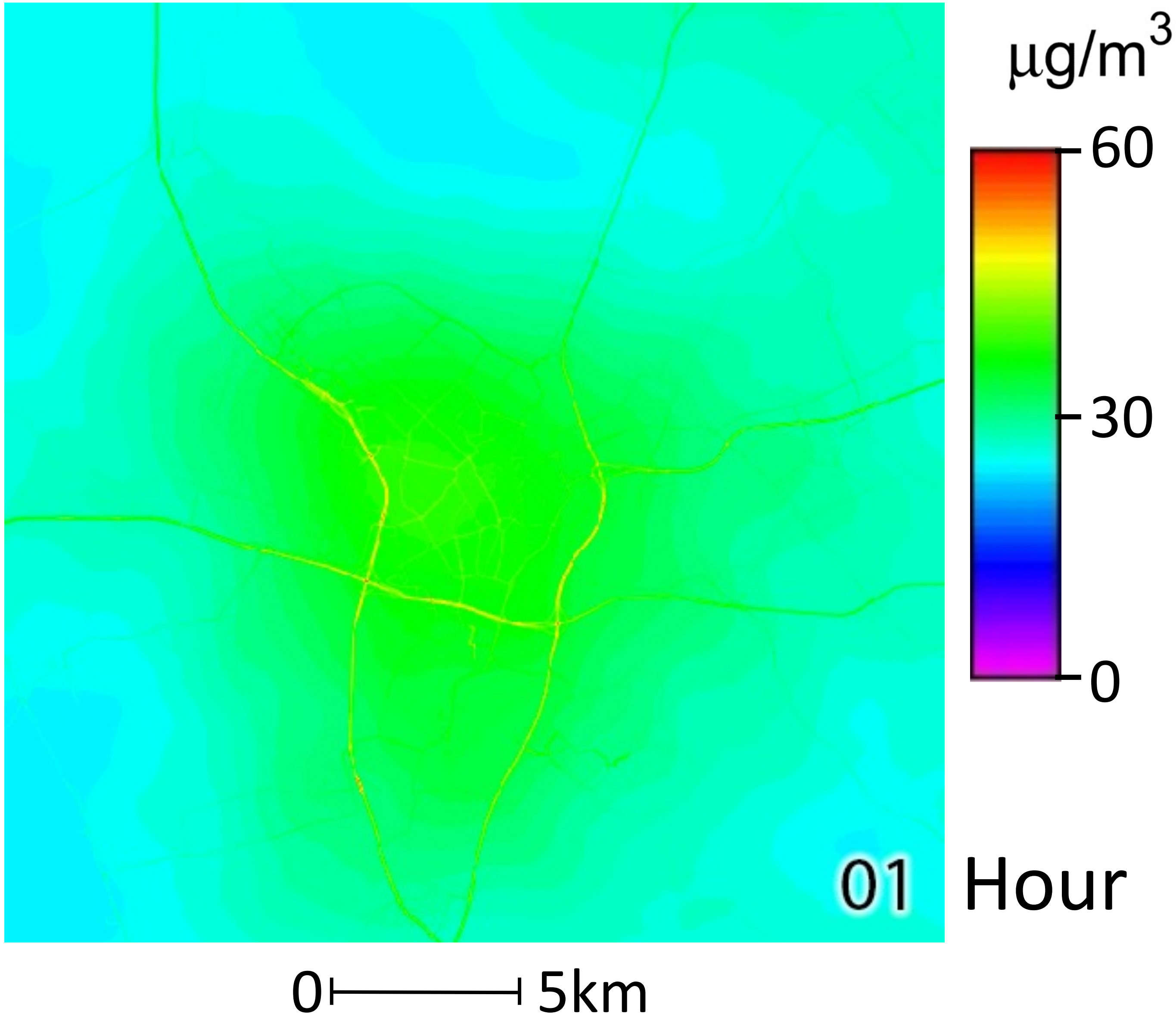
Shaddock et al., 2018

Why high spatial and temporal resolutions?



- Hourly/minutes
- LML data (landelijk meetnet luchtkwaliteit)

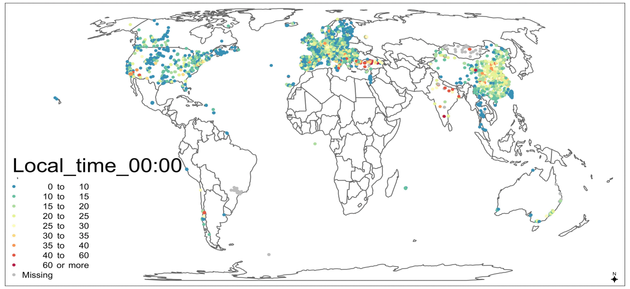




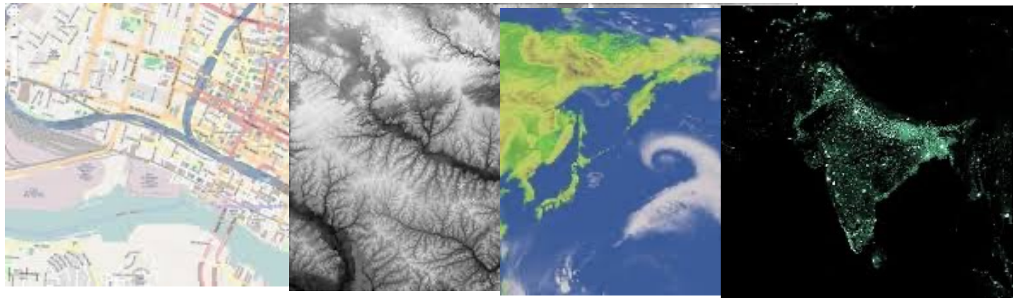
Statistical modeling

Based on relationships between NO_2 and Geospatial predictors and modeling the spatiotemporal dependence.

- Process-based models:
  Air dispersion models
  Chemical transportation models

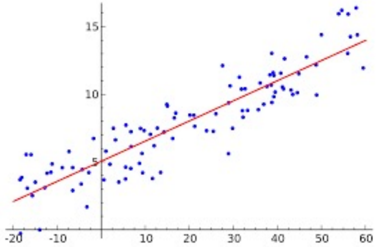


Station measurements



Predictors

Machine learning vs. Linear regression models



Global models Vs. National models

Daytimg models Vs. Nighttime models

Ensemble tree
- Random forest
- Gradient boosting
- XGboost

Linear regression
- Standard linear regression
- With regularisation

Lasso
Ridge
Elastic-Net

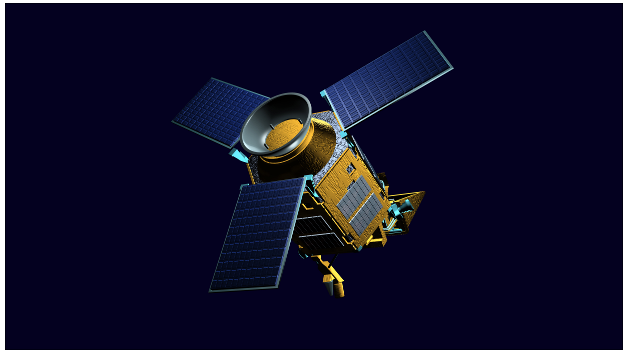
What we have done:

Lu, M. Schmitz, O., Van de Hoogh, K., Qin, K., & Karssenberg, D. Evaluation of different methods and data sources to optimise modelling of $NO2$ at a global scale. Environmental international

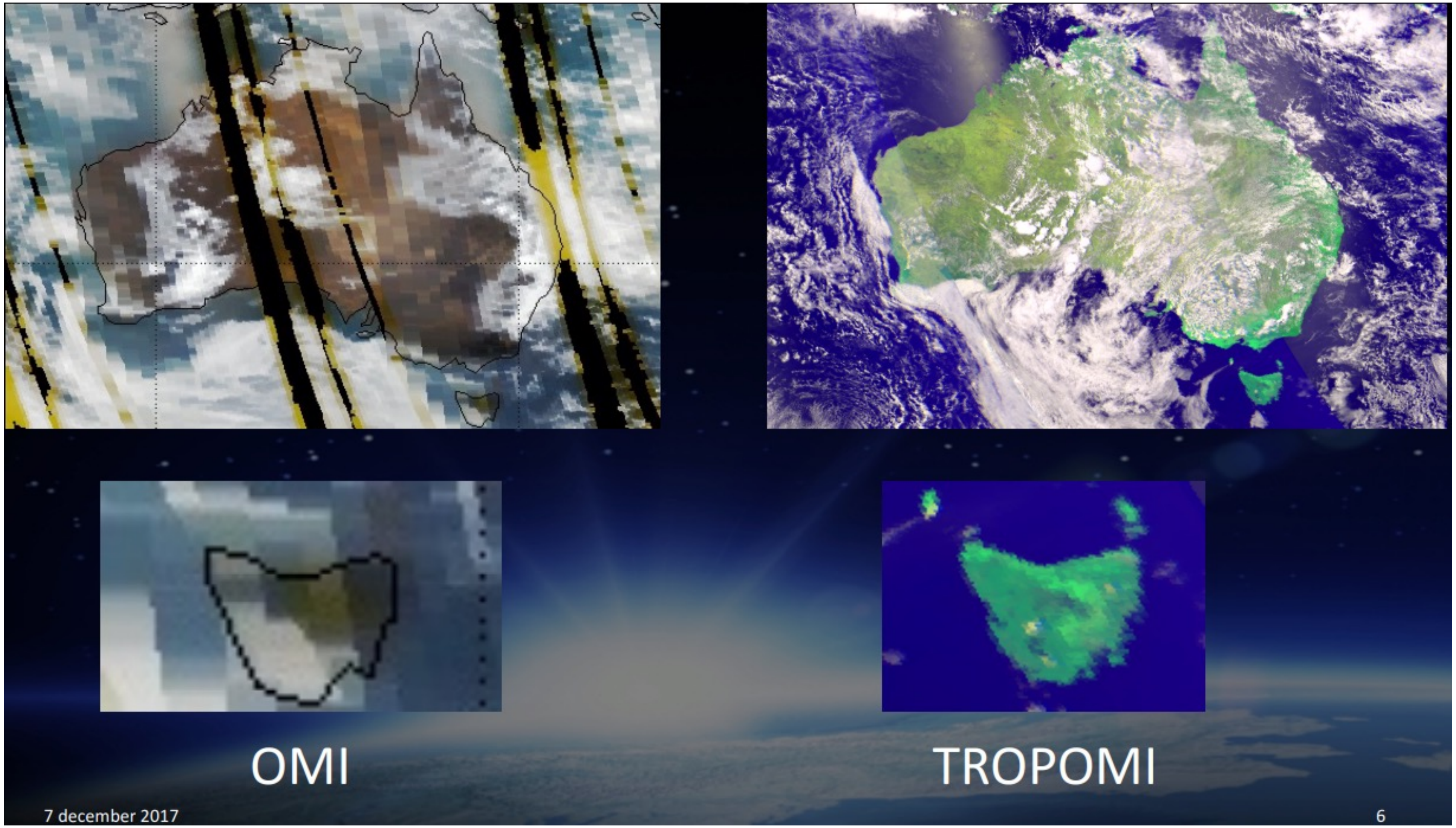
Tropomi: Sentinel 5p
(TROPOspheric Monitoring Instrument)

Launched Oct. 2017,

Resolution (NO2)
2018-2019: 7 km
Since Oct. 2019: 5.5 km



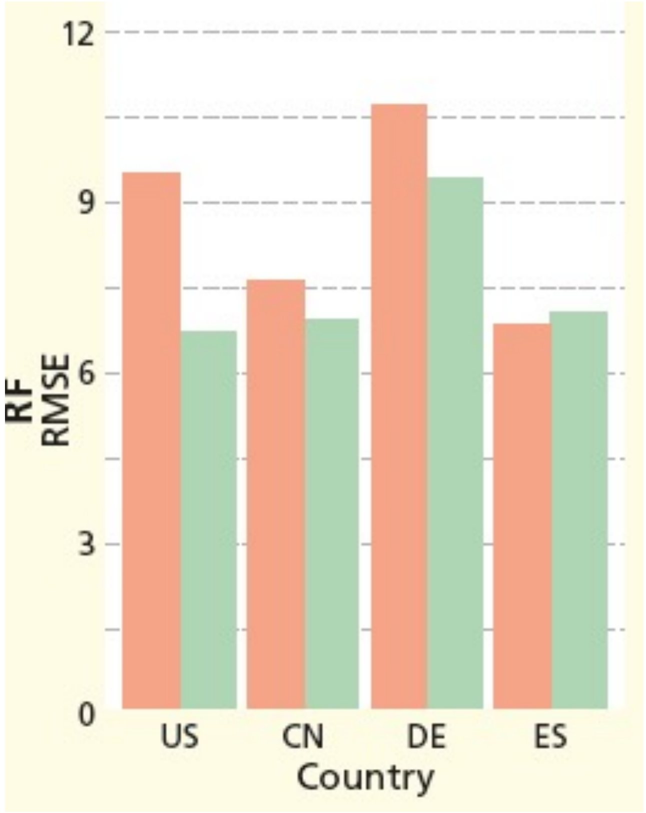
Measurements: NO_2, O_3 (7km × 28km), SO_2, methane and CO



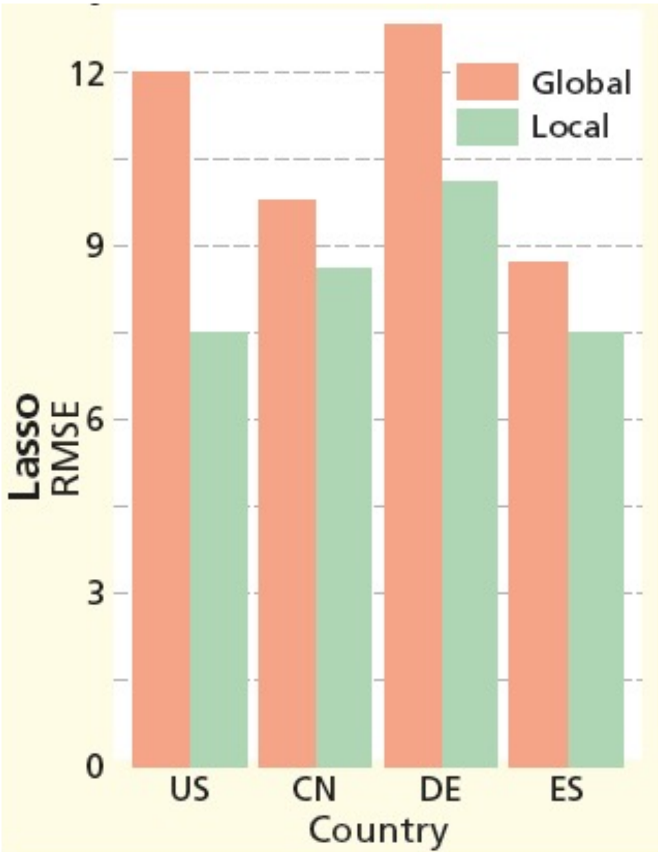
Souce [3]

- With the tree-based models, we obtained global modelling results that are comparible to local models

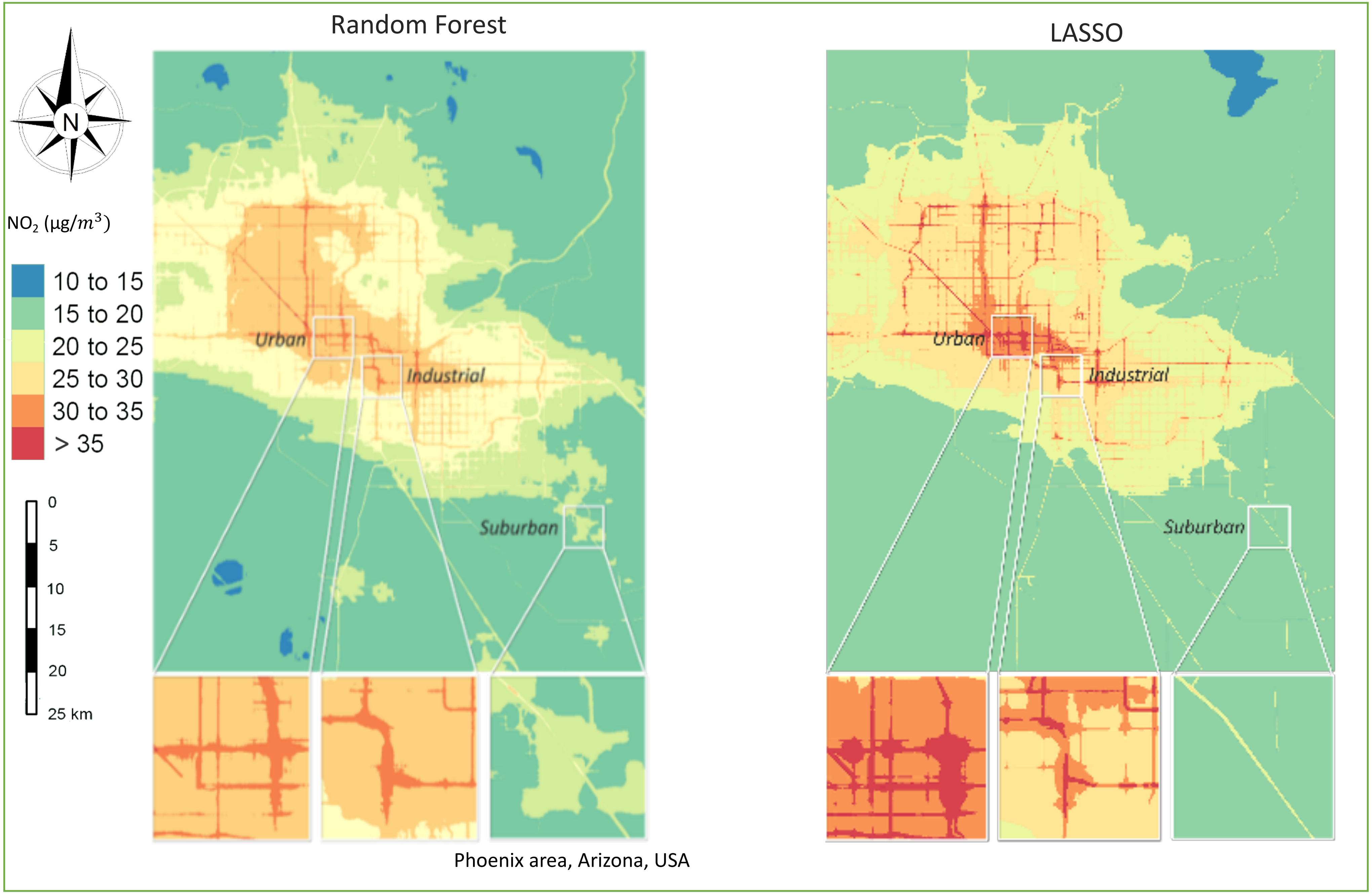
Random Forest

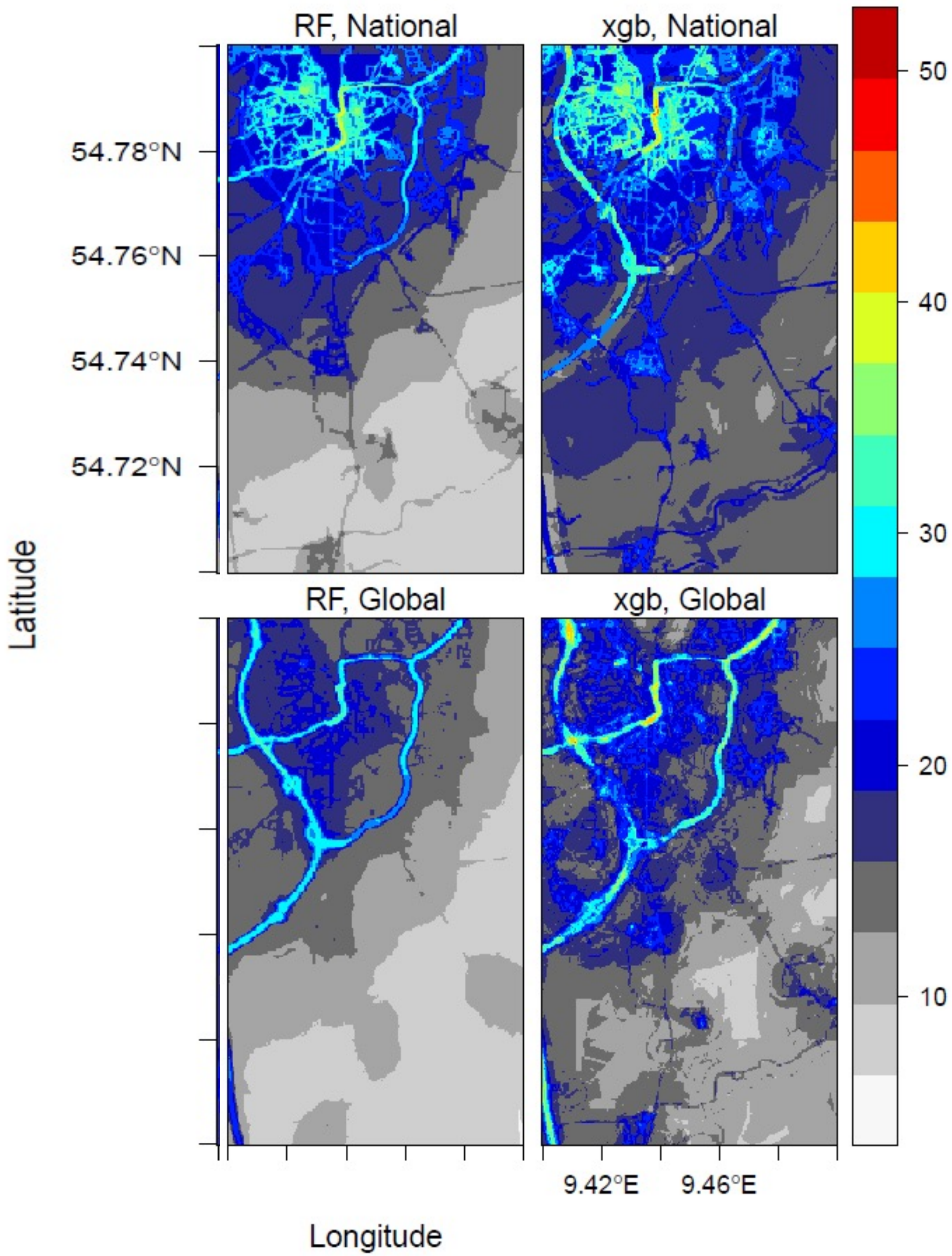


Lasso

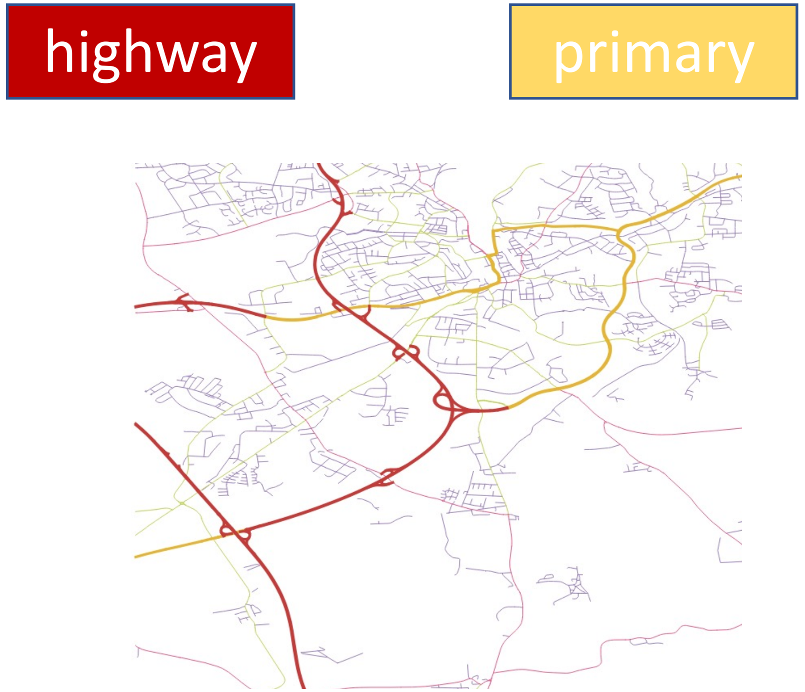


RF: random forest
CN: China
DE: Germany
ES: Spain


Random Forest
LASSO
N
NO$_2$ (μg/m^3)
10 to 15
15 to 20
20 to 25
25 to 30
30 to 35
> 35
0
5
10
15
20
25 km
Urban
Industrial
Suburban
Urban
Industrial
Suburban
Phoenix area, Arizona, USA

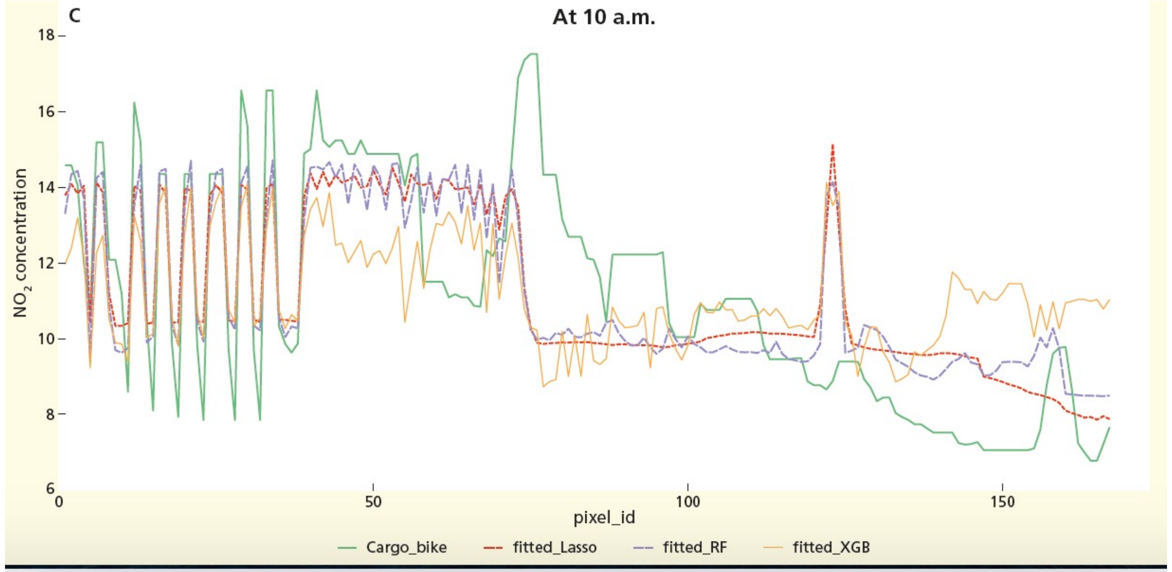


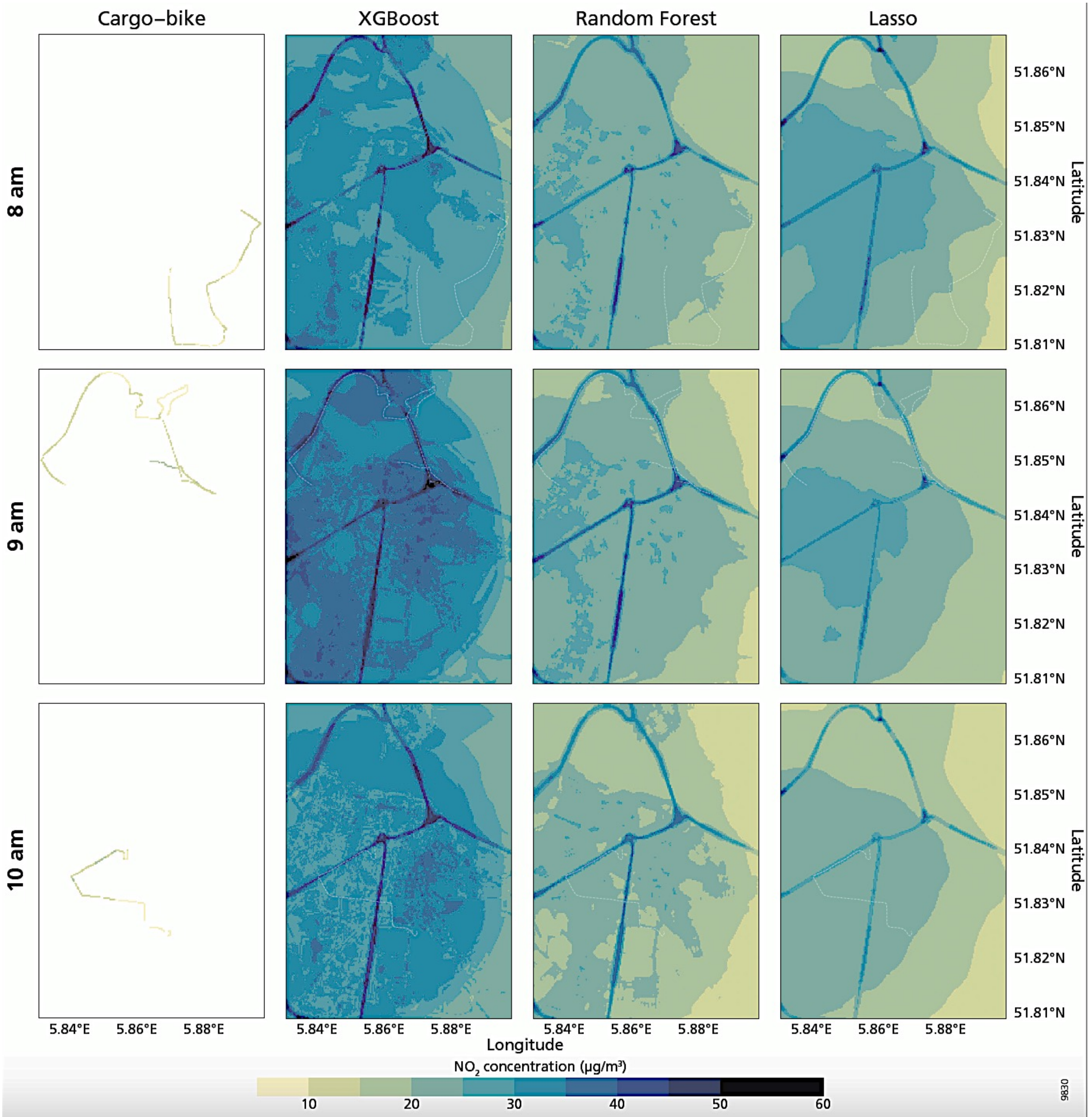
- Distinct prediction patterns despite the similar cross-validation accuracy between models.



Germany

Further validation: cargo-bike measurements

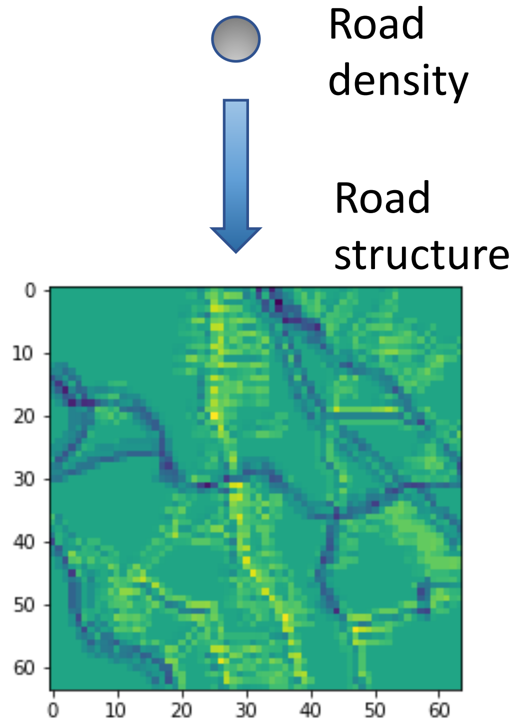




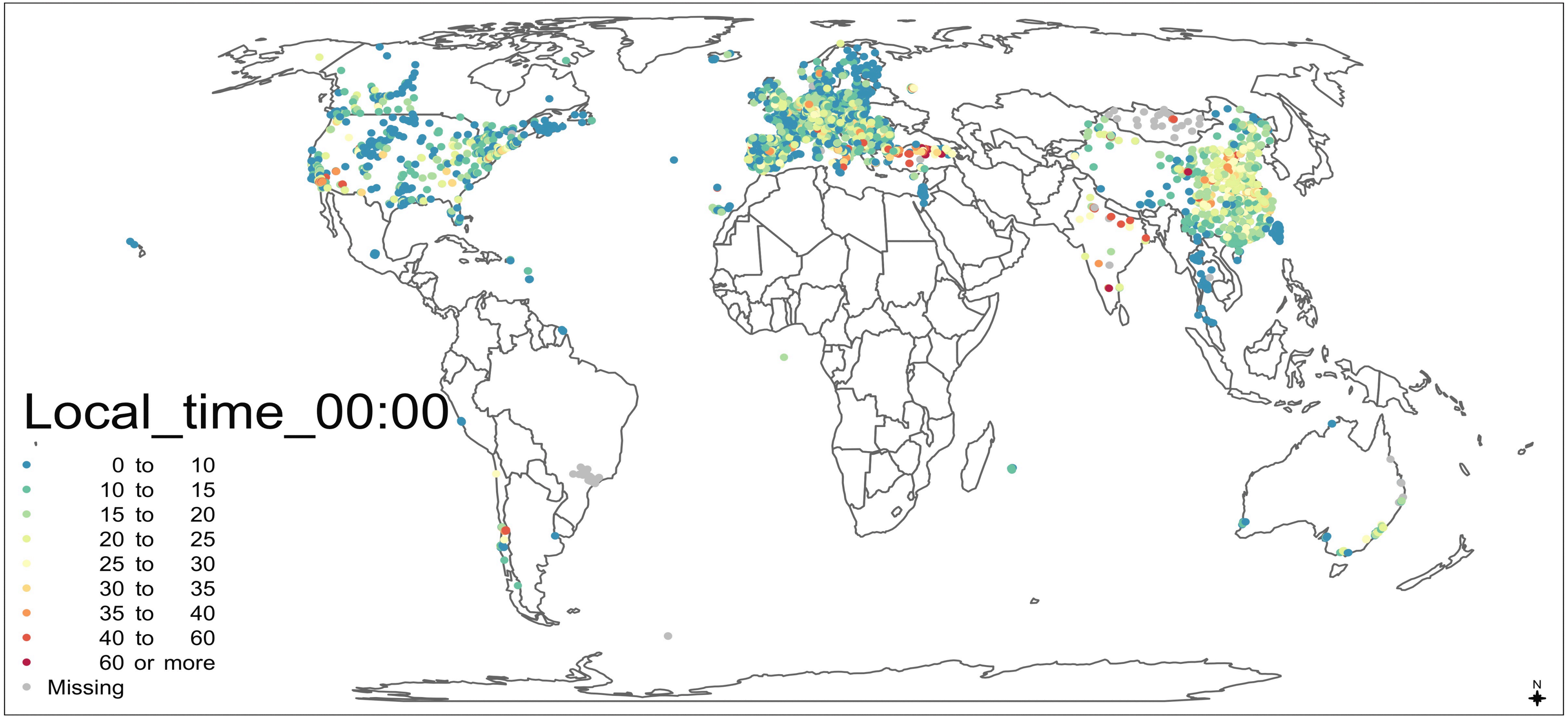
Challenges in statistical modeling and

Challenges:

- Spatiotemporal variability
- Spatially heterogeneous predictor-response relationships
- Integrate more spatial information



A new collection of ground monitor measurements:

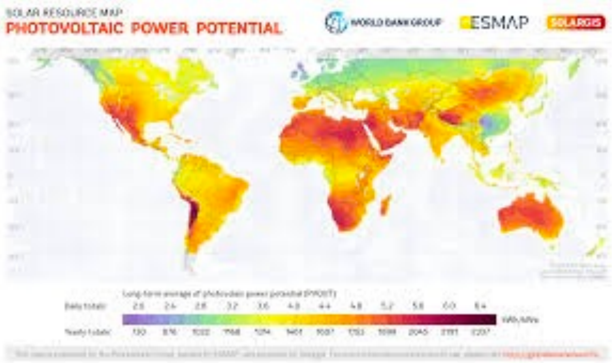


OpenAQ: an open science air quality community
Through collaboration: China, Australia, Canada, US, EU

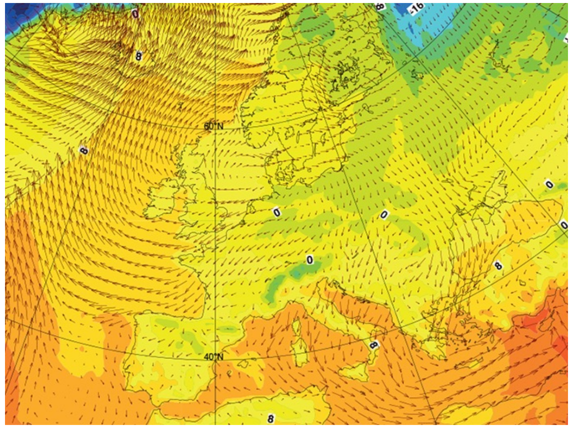
Updated Geospatial predictors and Tropomi measurements

Road and industrial area from OpenStreetMap (April 2020)

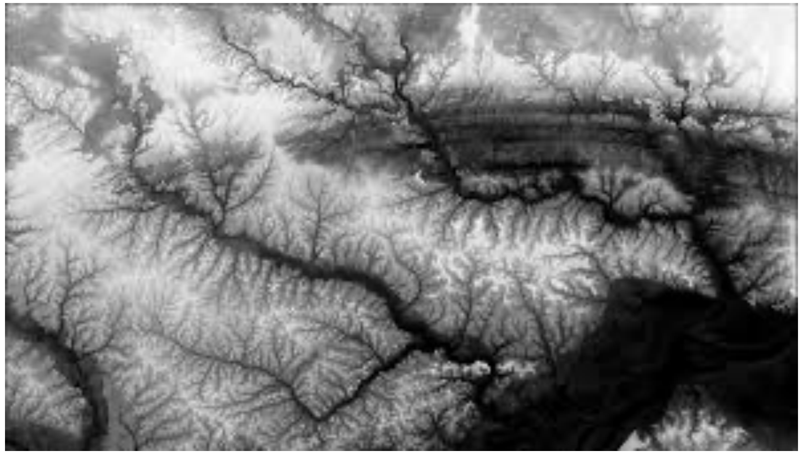
Earth nightlight: VIIRS (500 m)



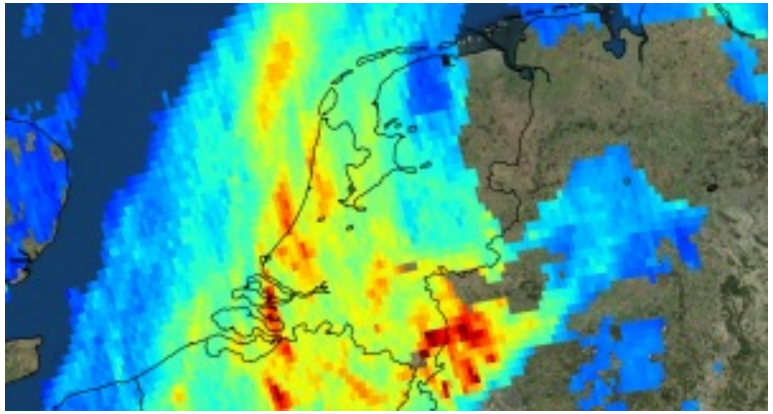
Solar radiation
Numerical model (Solargis)



Wind speed and temperature: ECMWF climate reanalysis, ERA5-land (9km)



Elevation: MERIT DEM (90 m)



Tropomi in 2019 (5.5 km resolution)

On-going work

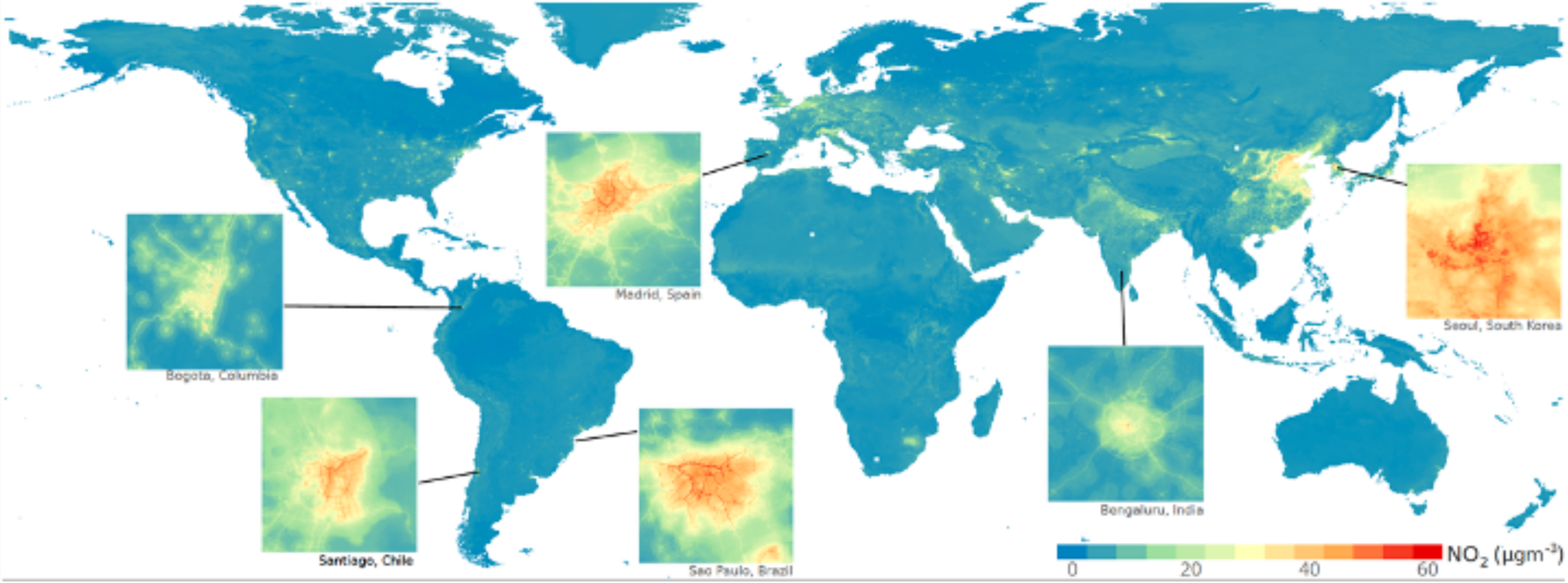
- Global map product separating between weekdays and weekends additional to day and night, using the XGBoost model.

- Hourly spatiotemporal model. (challenge 1)

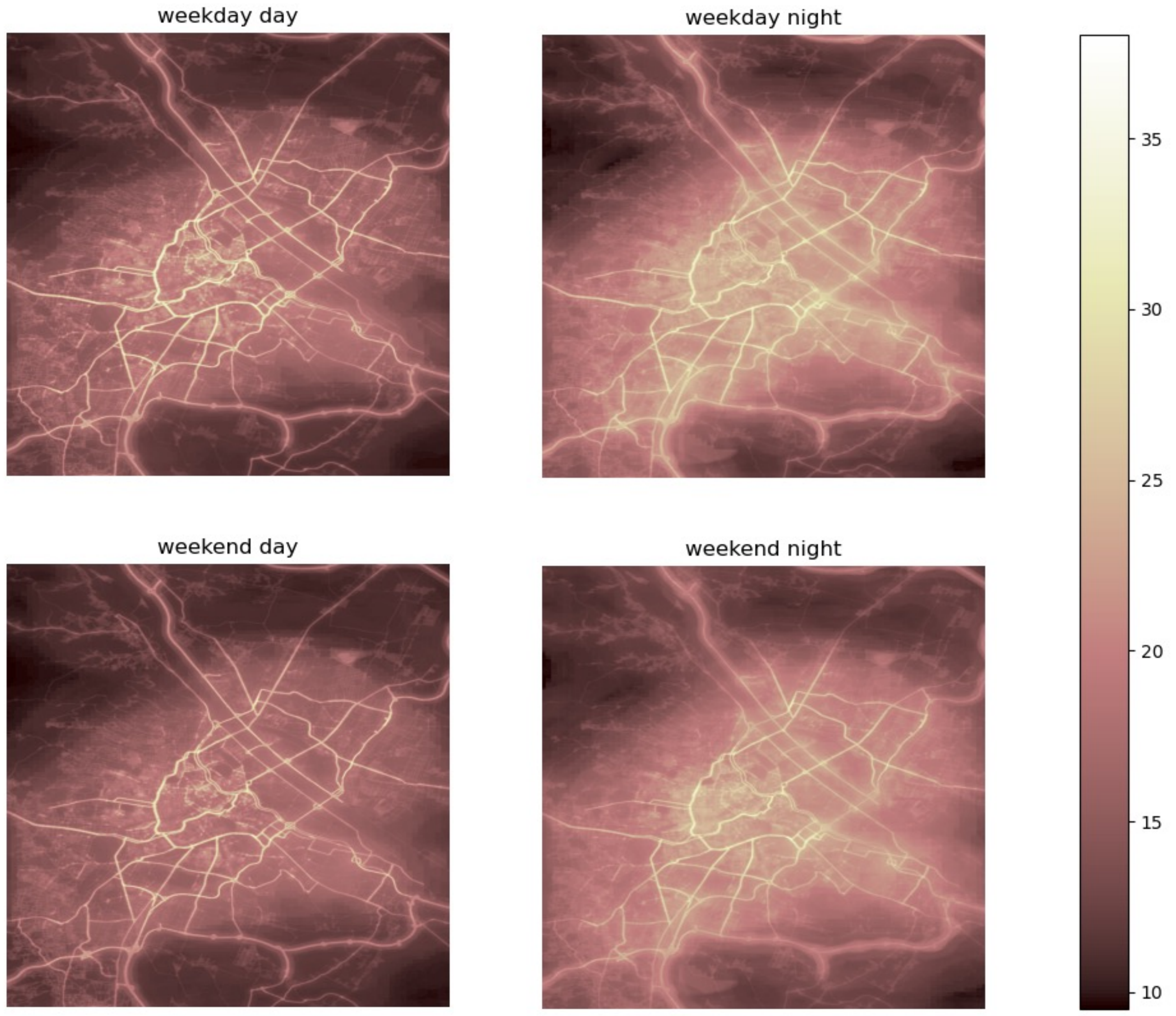
- Advanced hierarchical modelling:

1. Mixed-effect modeling. (challenge 2)
2. Deep convolutional neural networks over spatial and spectral-spatial domain. (challenge 3)

Global map



Madrid, Spain
Bogota, Columbia
Santiago, Chile
Sao Paulo, Brazil
Bengaluru, India
Seoul, South Korea
0
20
40
60
NO_2 (μgm^{-3})



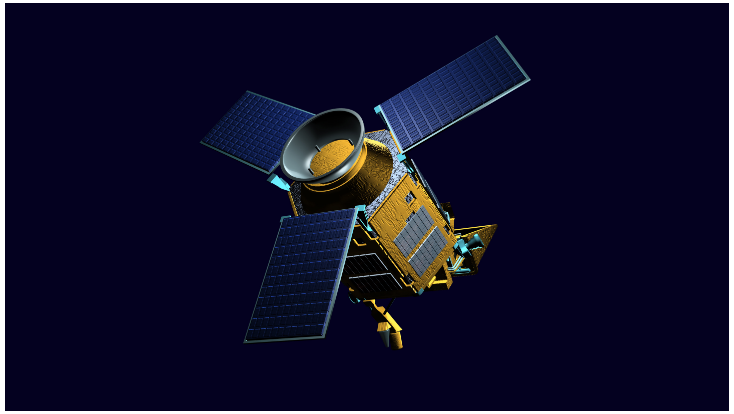
Example maps of Vienna, Austria

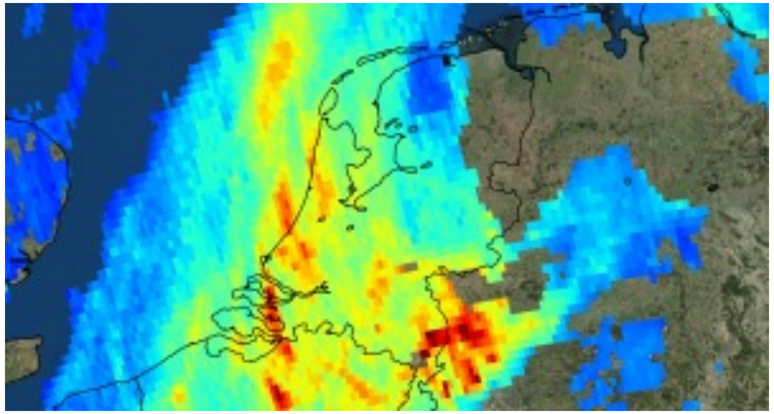
separating between day, night, weekday, weekends

Thank you

https://github.com/mengluchu/OpenGeoHub2020

Tropomi





NO_2, O_3 (7km × 28km), SO_2, methane and CO

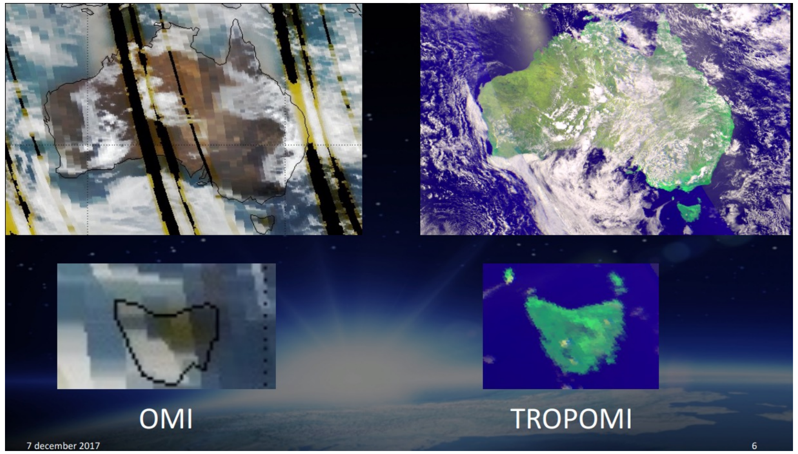
Electromagnetic field:
ultraviolet and visible (270–500 nm), near-infrared (675–775 nm), shortwave infrared (2305–2385 nm).

zoom in mode: 7 km × 3.5 km 5.5 km x 3.5 km since 2019

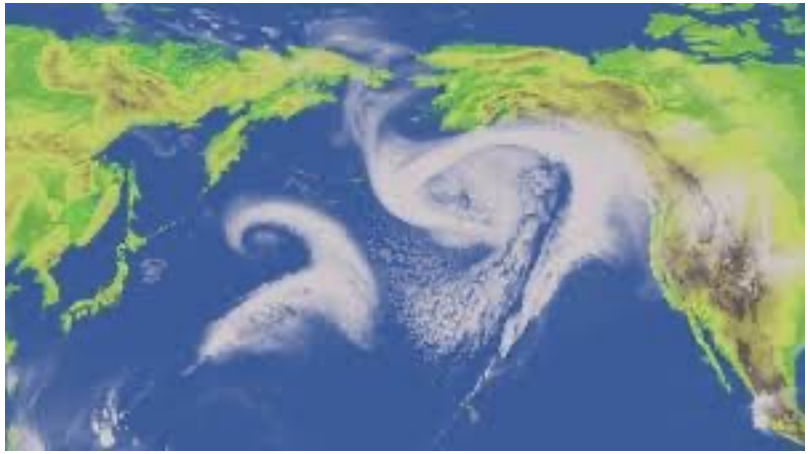
Geospatial predictors of the 2020 paper

Road and industrial area from
OpenStreetMap (vector)

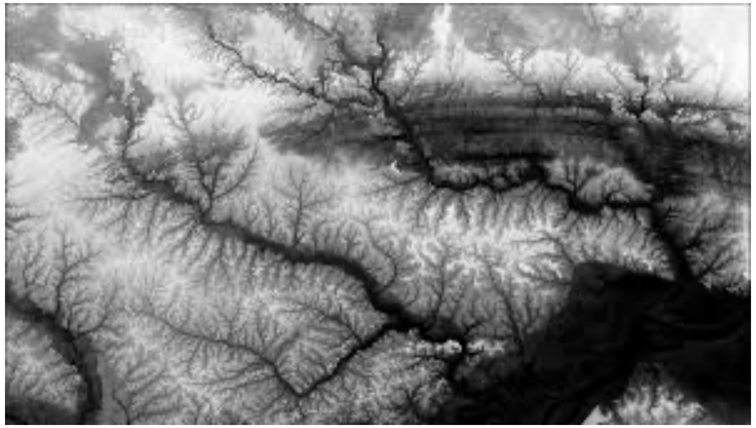
Population: GHS-POP R2019A(250 m)



Sentinel 5p (2018)
OMI (2017)
GEOS-CHEM (2012)



Wind speed and temperature



Elevation: SRTM (90 m)