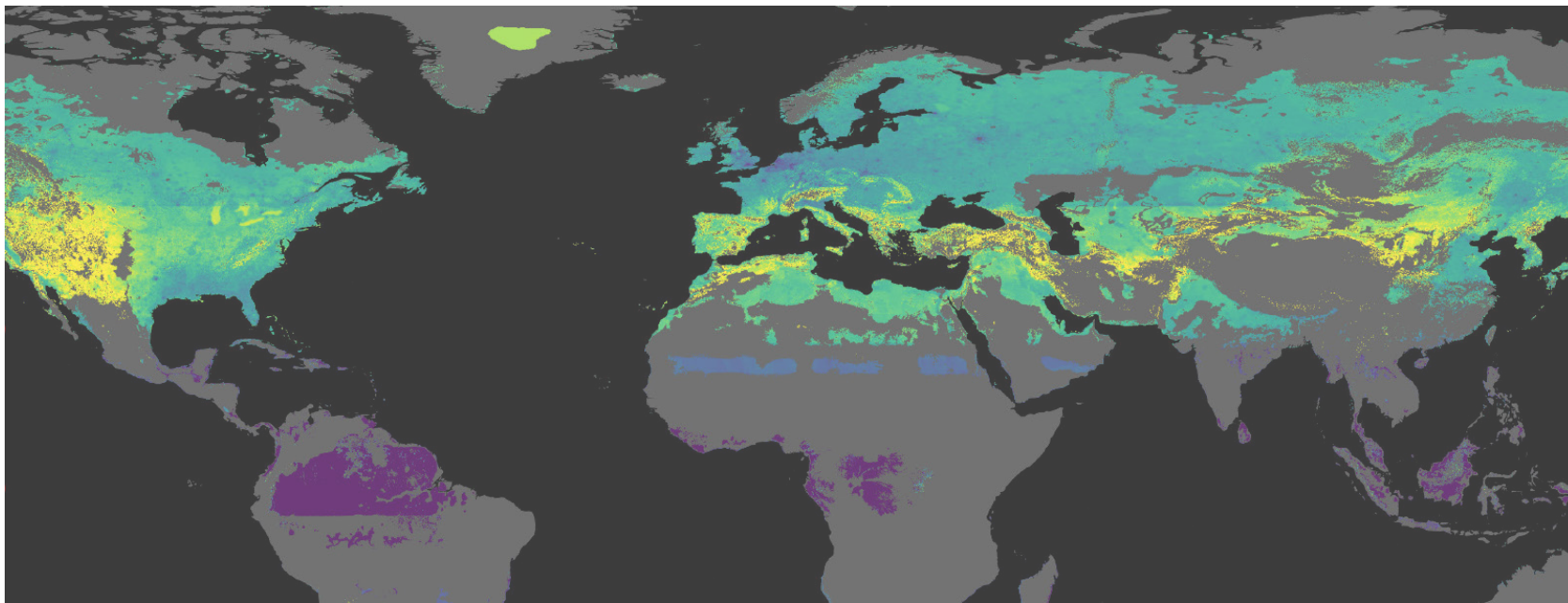




Global air quality mapping with explainable machine learning

JSC's End-of-Year Colloquium

8 DECEMBER 2022 | CLARA BETANCOURT | clara.betancourt@fz-juelich.de

IntelliAQ

K I : S T E

 **JÜLICH**
Forschungszentrum

Member of the Helmholtz Association



© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

OUTLINE

Global air quality mapping with explainable machine learning

- Ozone – a toxic trace gas
- Ozone mapping with machine learning
- Explainable Machine Learning
- The ozone map

THE EARTH'S ATMOSPHERE

Pollutants and trace gases make a difference!



The Earth's atmosphere

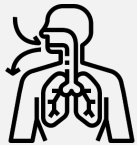


Air pollution

→ [Images by Pixabay](#)

TROPOSPHERIC OZONE

A toxic greenhouse gas



Lung damage

→ [Global Burden of Disease Study 2017](#)



Vegetation damage

→ [Emberson et al, 2018](#)

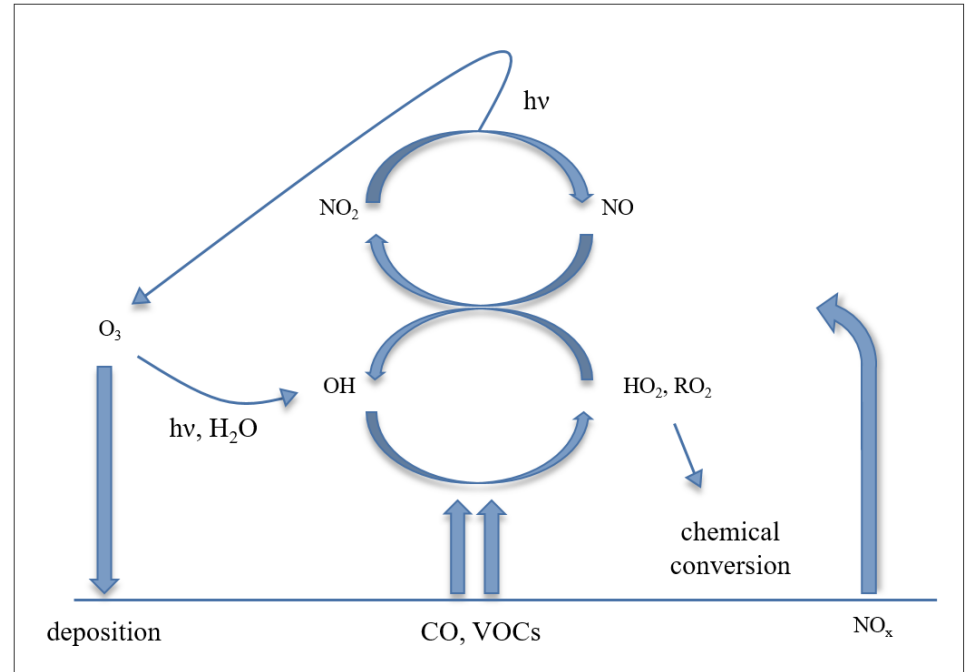


Short lived climate forcer

→ [IPCC, 2018](#)

Ozone impacts

→ [Icons by Flaticon](#)



Ozone processes

→ [Betancourt et al. 2021](#)

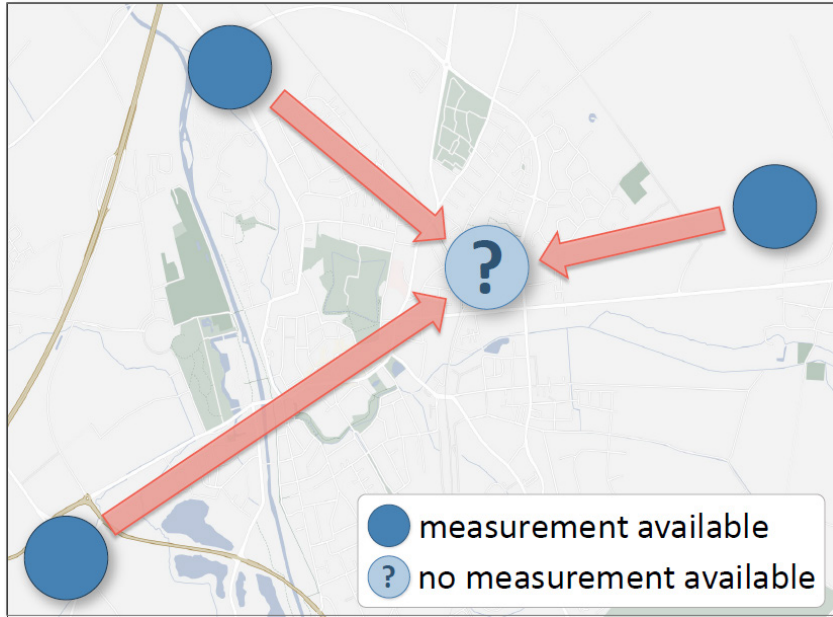
OUTLINE

Global air quality mapping with explainable machine learning

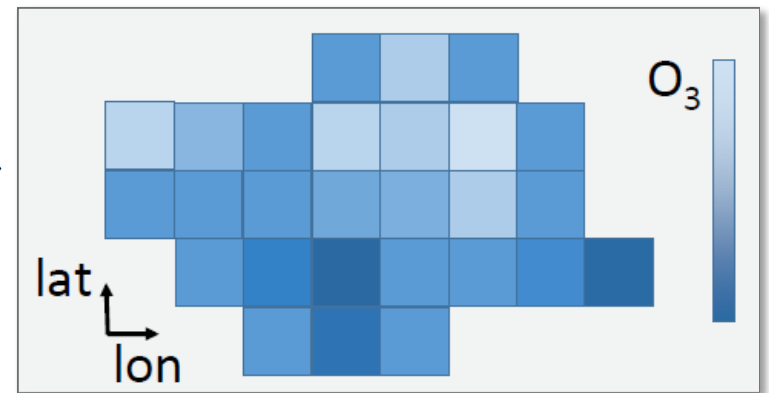
- Ozone – a toxic trace gas
- Ozone mapping with machine learning
- Explainable Machine Learning
- The ozone map

WHAT IS MAPPING?

From irregularly placed measurements to gridded data



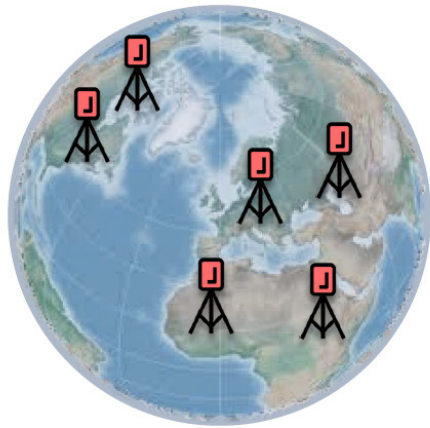
Ozone measurements



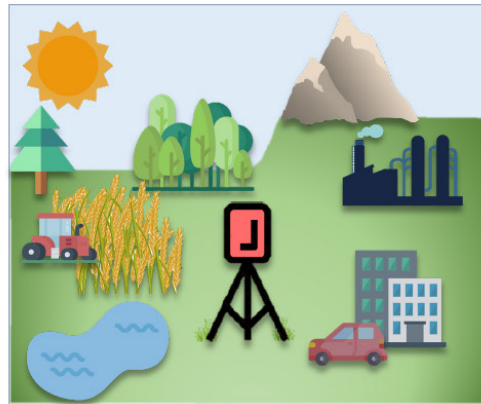
Ozone mapping

THE BENEFITS OF MACHINE LEARNING

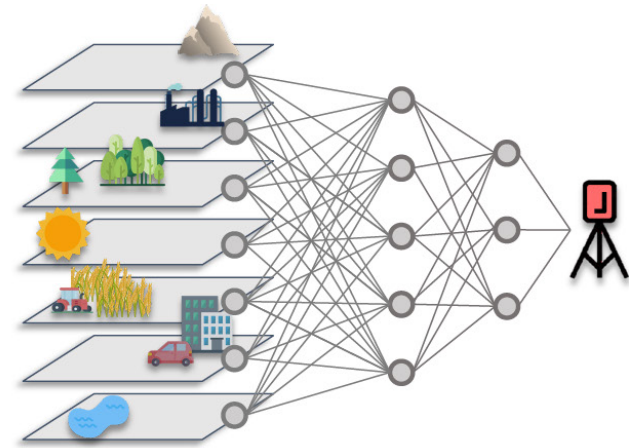
AQ-Bench dataset → [Betancourt et al. 2021](#)



The AQ-Bench dataset contains long-term air quality metrics and metadata at sites around the globe.

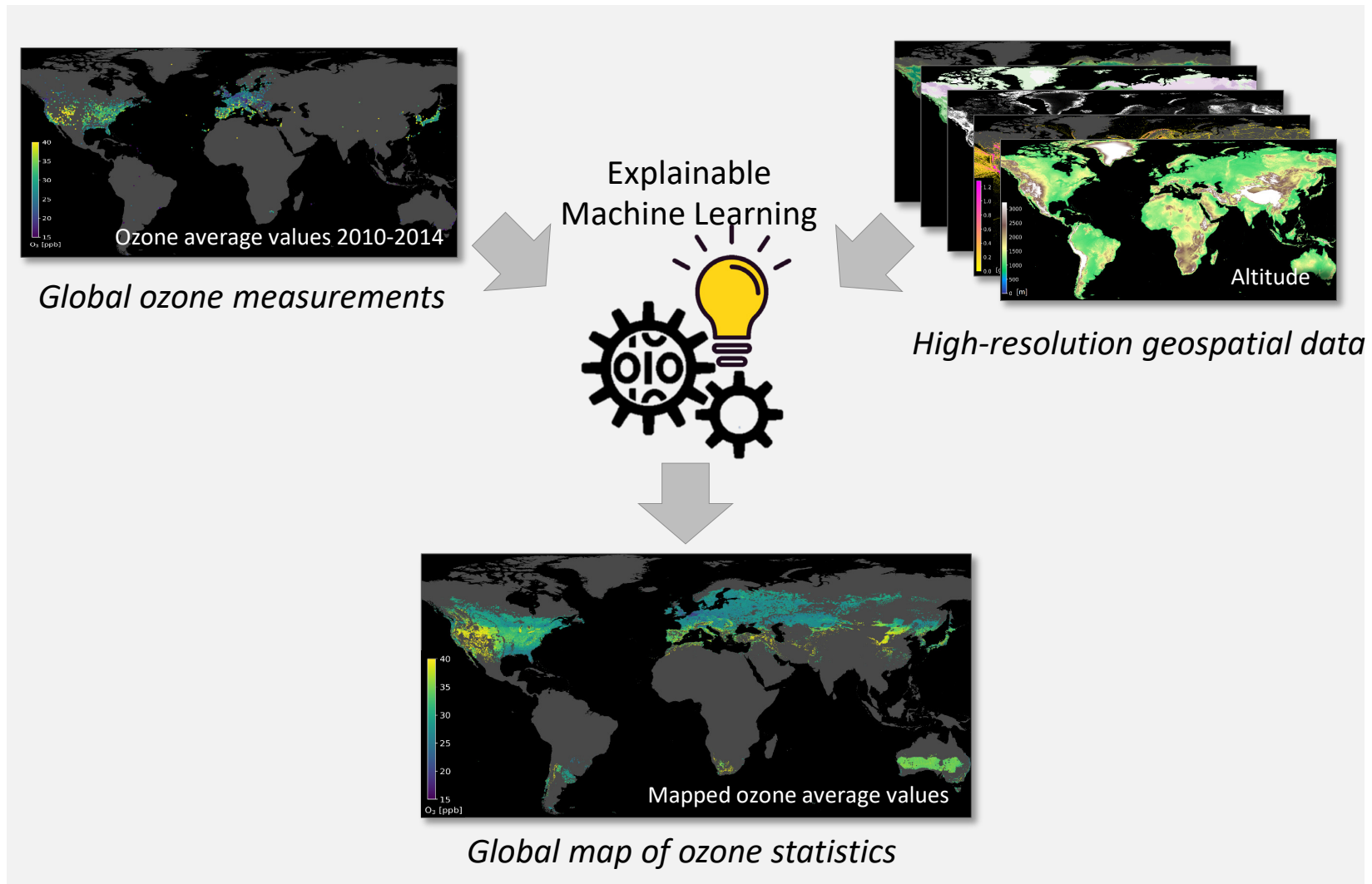


The air quality at a site is influenced by its surroundings.

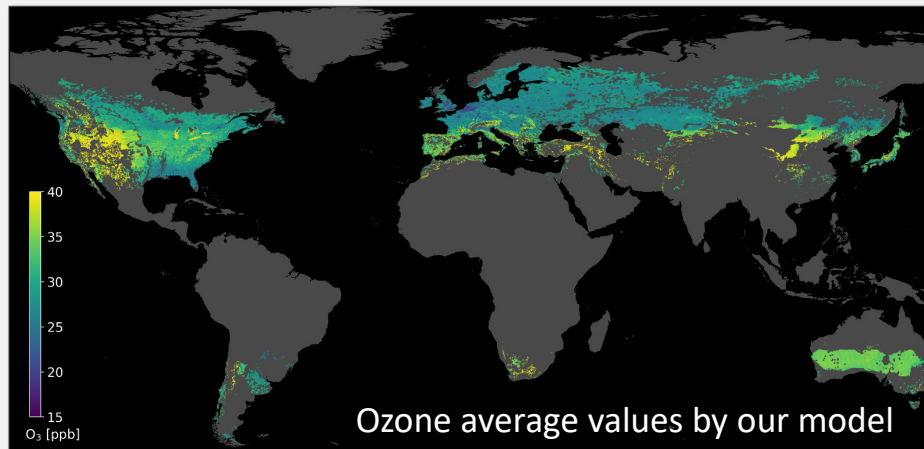
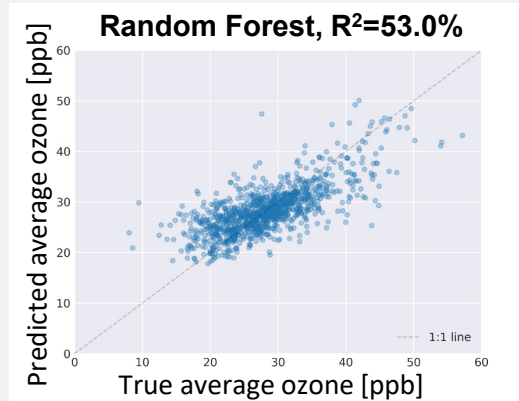


The proposed machine learning task is to train a machine learning algorithm which maps from metadata to long-term air quality metrics at measurement sites.

OZONE MAPPING



WHAT DO WE WANT TO EXPLAIN?



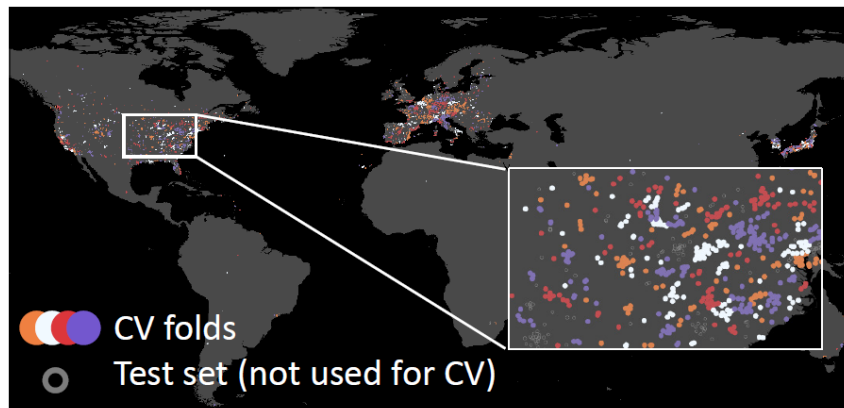
OUTLINE

Global air quality mapping with explainable machine learning

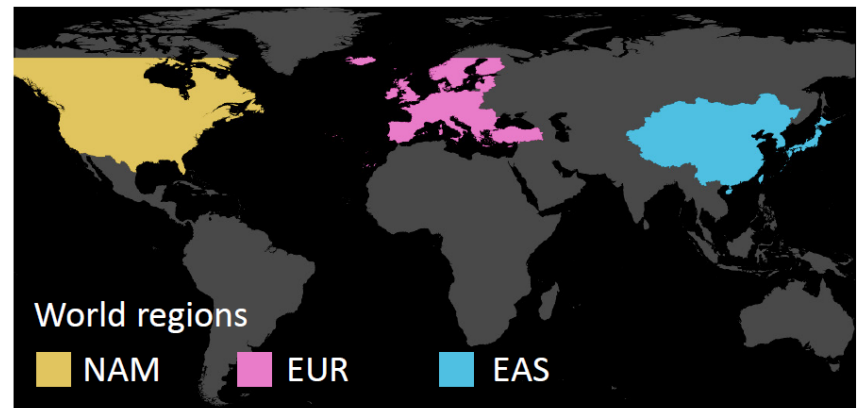
- Ozone – a toxic trace gas
- Ozone mapping with machine learning
- **Explainable Machine Learning**
- The ozone map

GLOBAL GENERALIZABILITY

Cross validation (CV) on different spatial scales



RMSE = 4 ppb



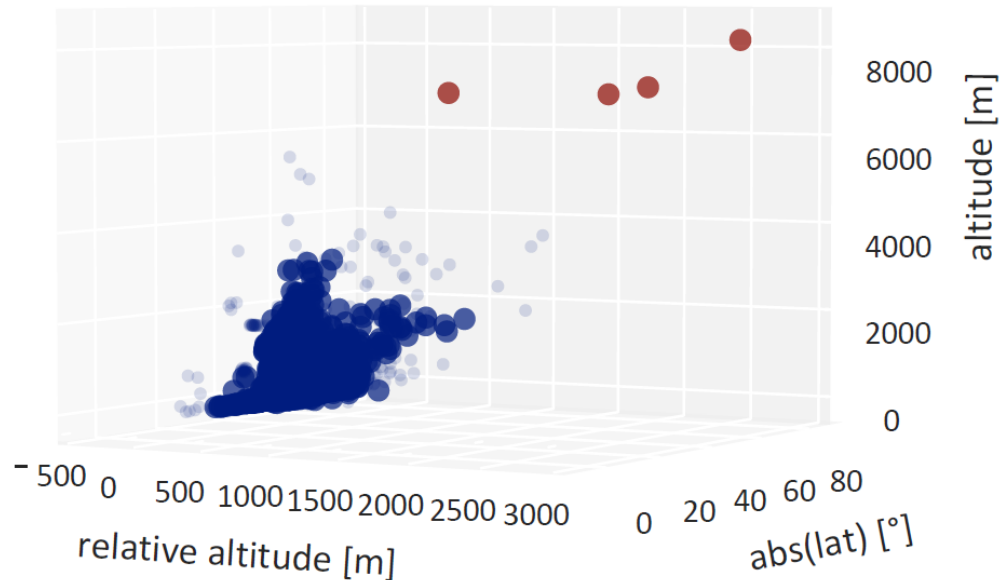
4 ppb < RMSE < ~ 5 ppb

APPLICABILITY

A look at the feature space

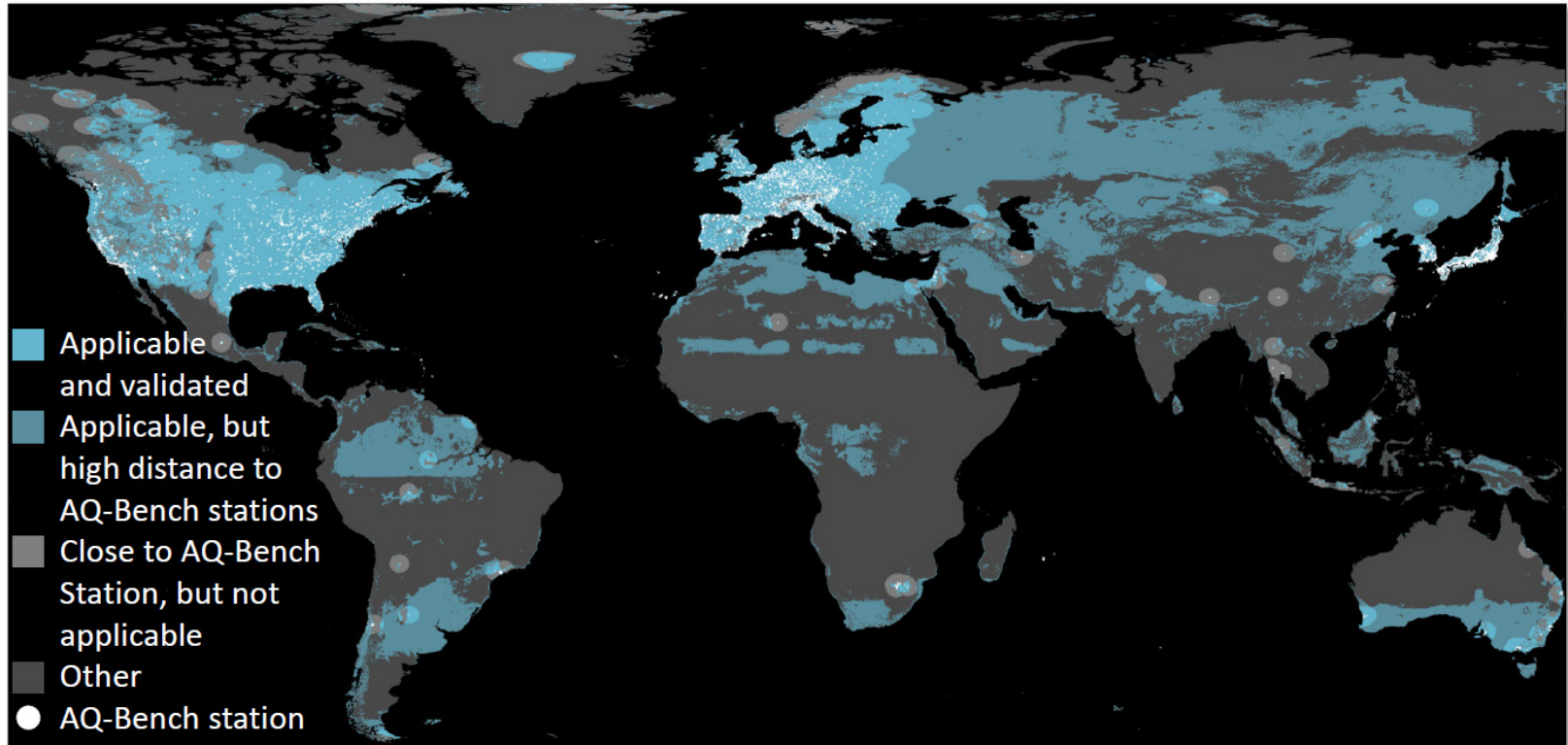
(method by Meyer et Pebesma, 2021)

- AQ-Bench dataset main cluster
- AQ-Bench dataset outliers
- Example gridded data points outside area of applicability



COMBINING THE FINDINGS

Generalizability and applicability

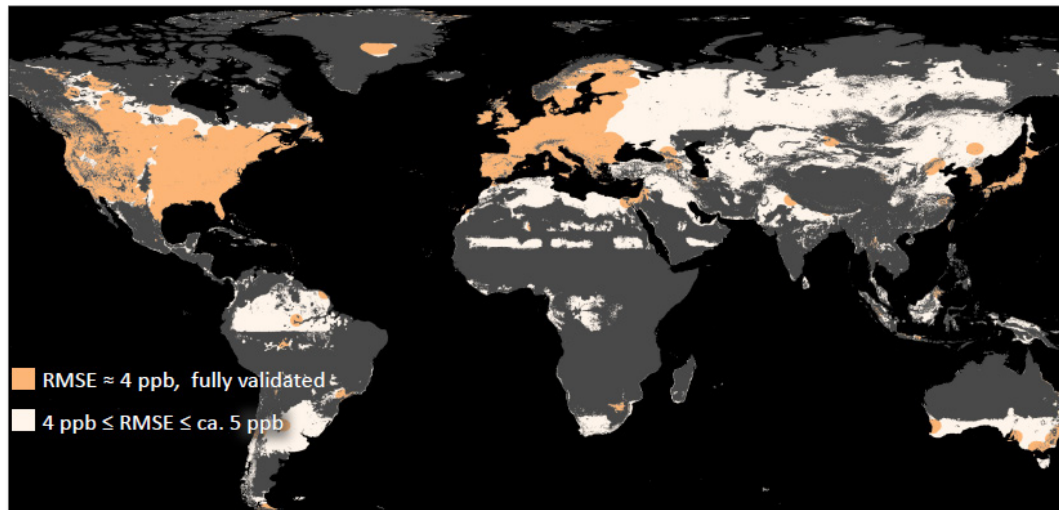
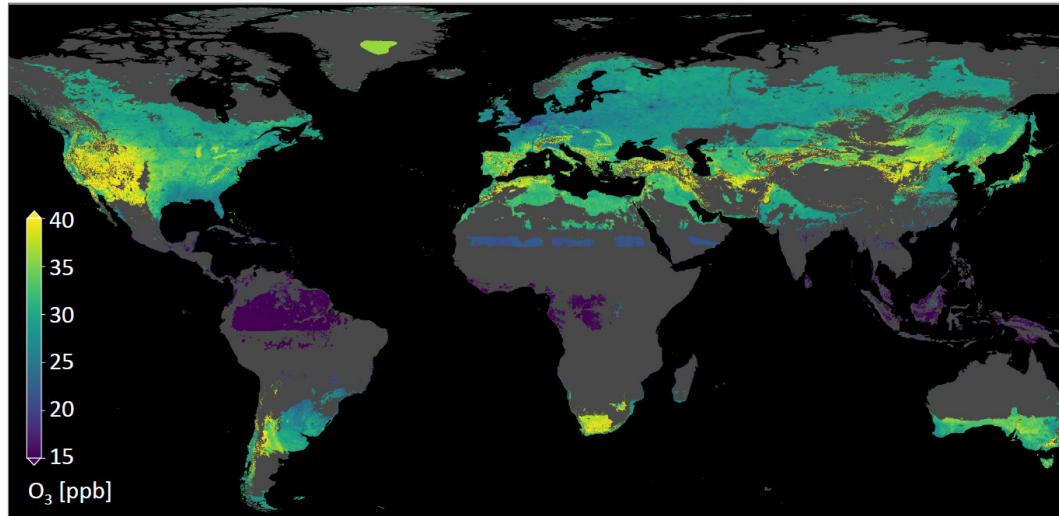


OUTLINE

Global air quality mapping with explainable machine learning

- Ozone – a toxic trace gas
- Ozone mapping with machine learning
- Explainable Machine Learning
- The ozone map

THE FINAL MAP → [Betancourt et al. 2022](#)



CONCLUSIONS

- Use of geospatial data for data driven approach
- Explainable machine learning makes map trustworthy

FUTURE RESEARCH

- Method is suitable for other ecological variables
- Dynamical mapping, with transformers

CONCLUSIONS

- Use of geospatial data for data driven approach

- Explanatory variables makes map trustworthy



→ [Adobe stock image](#)

- Method is suitable for other ecological variables
- Dynamical mapping, with transformers