Global Building Emissions Estimation

Climate TRACE
MIDS Capstone 2024

Duke University

Our team

Kyle Bradbury, Ph.D.

Mentor & Client

Barbara Flores

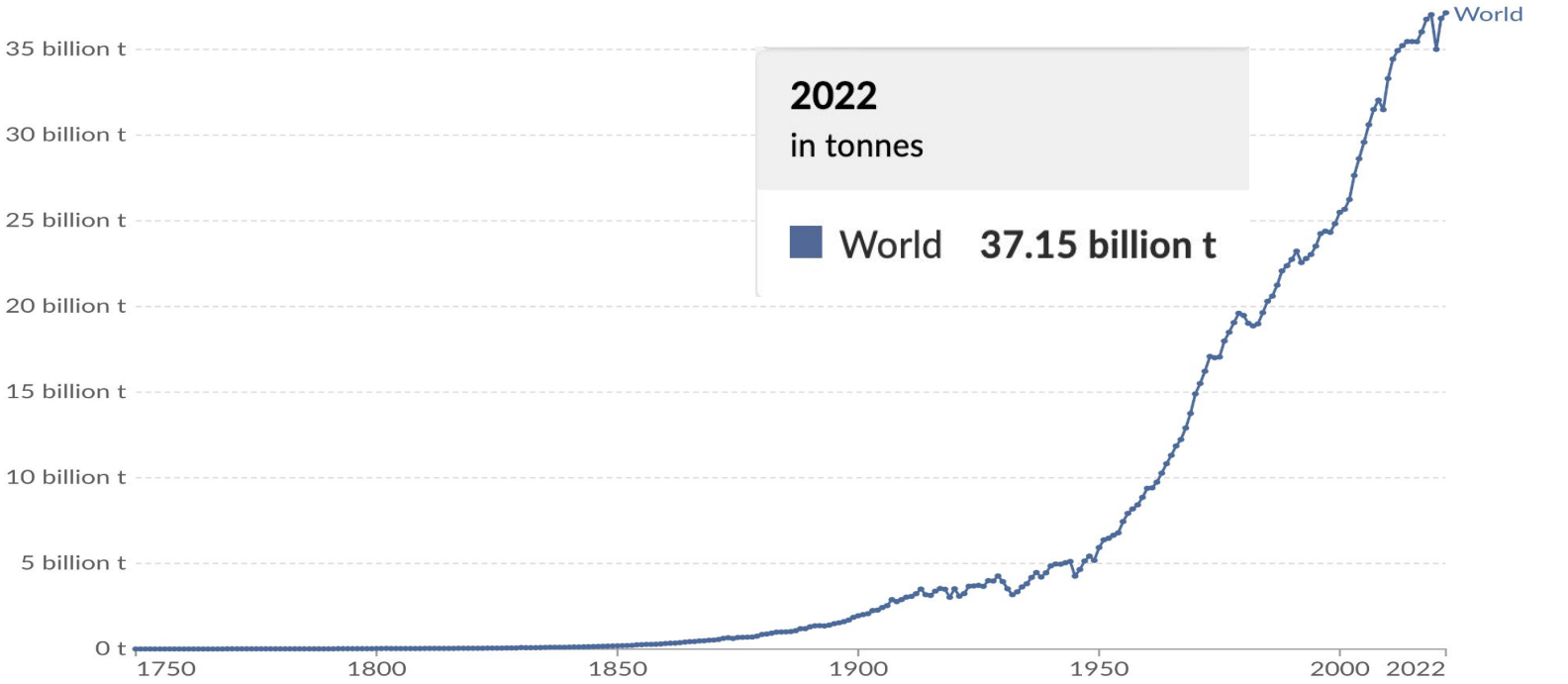
Meixiang Du

Jiechen Li

Yulei (Alicia) Xia

Annual CO_2 emissions

Carbon dioxide (CO_2) emissions from fossil fuels and industry[1]. Land-use change is not included.



Data source: Global Carbon Budget (2023)

OurWorldinData.org/co2-and-greenhouse-gas-emissions | CC BY

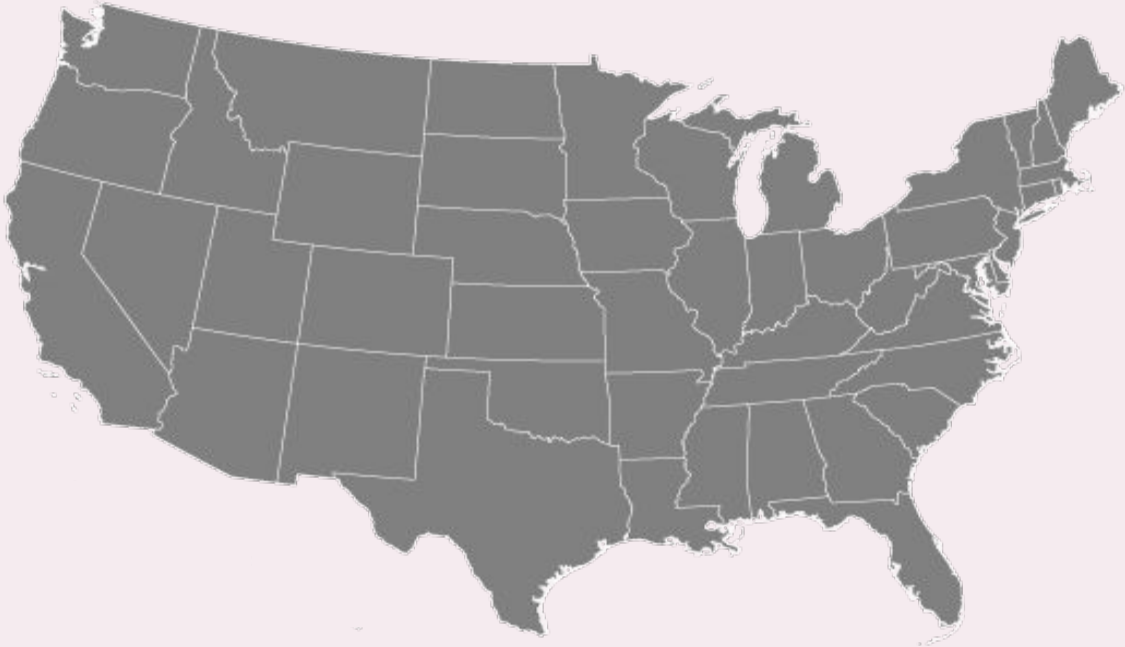
1. Fossil emissions: Fossil emissions measure the quantity of carbon dioxide (CO_2) emitted from the burning of fossil fuels, and directly from industrial processes such as cement and steel production. Fossil CO_2 includes emissions from coal, oil, gas, flaring, cement, steel, and other industrial processes. Fossil emissions do not include land use change, deforestation, soils, or vegetation.

Hannah Ritchie and Max Roser (2020) - "CO_2 emissions" Published online at OurWorldinData.org. Retrieved from: 'https://ourworldindata.org/co2-emissions' [Online Resource]

If we use 37.15 billion tons of CO_2 to cover the entire surface of continental U.S. ...

Hannah Ritchie and Max Roser (2020) - "CO_2 emissions" Published online at OurWorldinData.org. Retrieved from: 'https://ourworldindata.org/co2-emissions' [Online Resource]

...it would form a layer of CO_2 about 2.58 meters thick.



Hannah Ritchie and Max Roser (2020) - "CO_2 emissions" Published online at OurWorldinData.org. Retrieved from: 'https://ourworldindata.org/co2-emissions' [Online Resource]

Challenges

Outdated

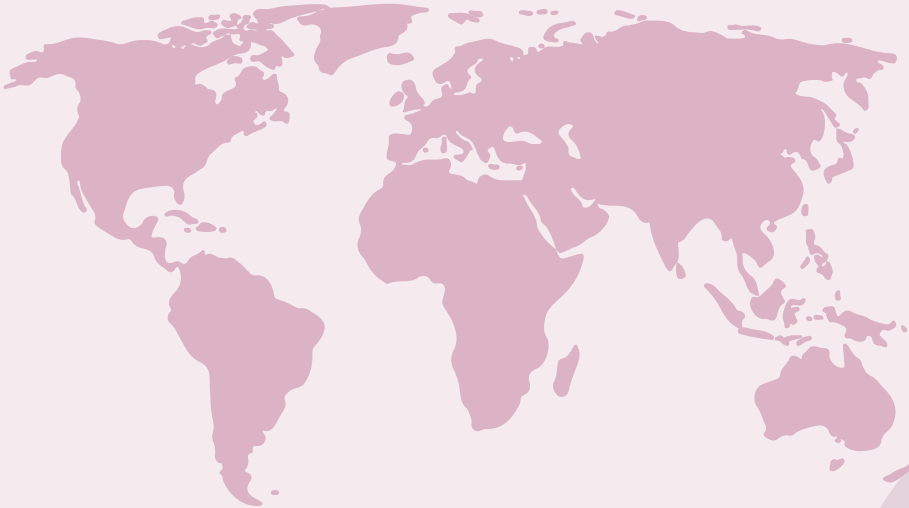
Historically limited and incomplete inventories in country level

Latency

The latest emissions report to UN, but still many countries struggle to report emissions

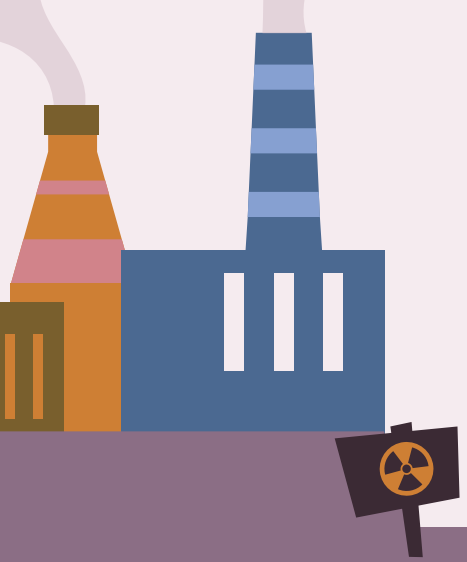
Behind Paywalls

Many Granular and detailed data, when available – are only available to paid subscribers

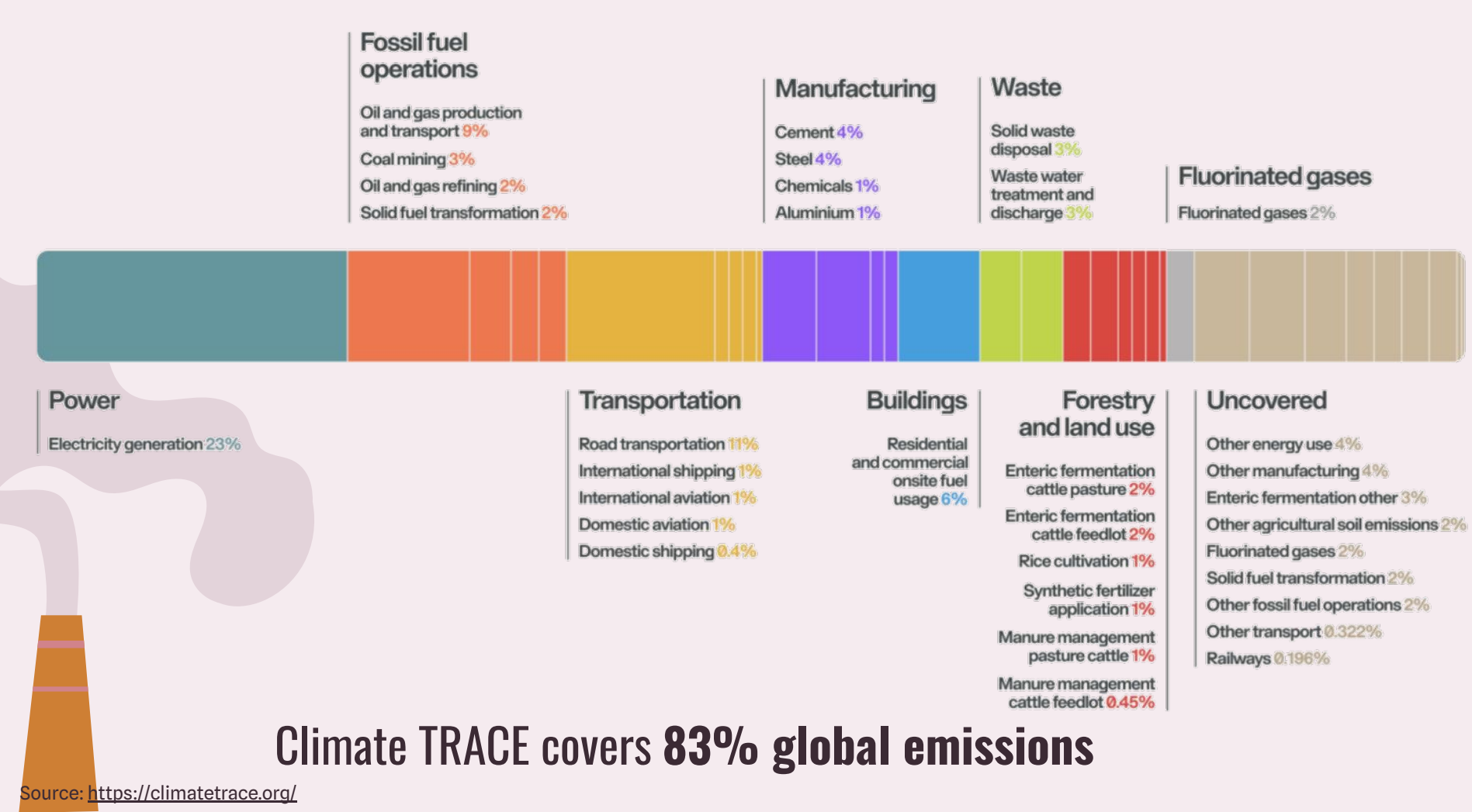


Climate TRACE

- A global non-profit coalition;
- Estimate and track greenhouse gas (GHG) emissions;
- Provide fast, easy and free data access to the public.

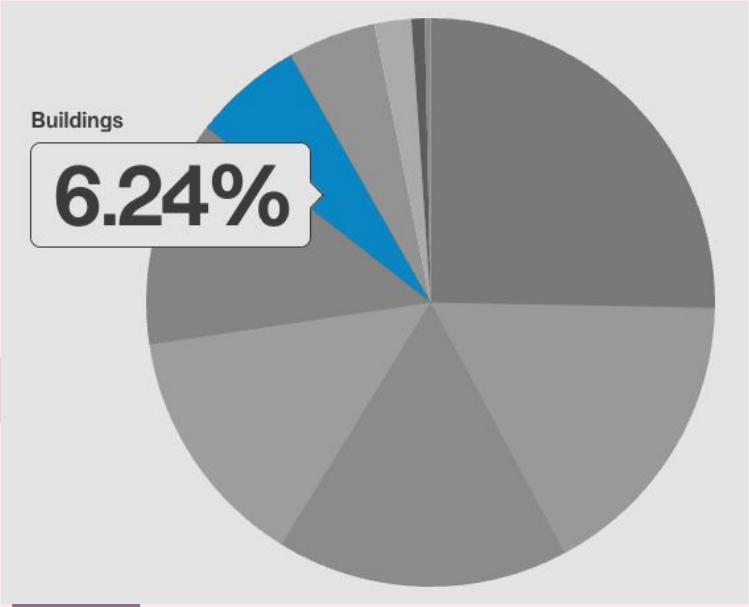


Source: https://climatetrace.org/



Climate TRACE covers **83% global emissions**

Source: https://climatetrace.org/



Duke University

Leads modeling **direct emissions** from fuels combusted within **buildings** based on satellite-derived data.

Source: https://climatetrace.org/

Buildings Emissions Estimation

A × EUI × EF = GHG

A	EUI	EF	GHG
Residential Building Characteristics [Area] Non-residential Building Characteristics [Area]	Energy Use Intensity Factors [Joules/Area]	Emissions Intensity Factors [tons GHG/Joule]	Onsite emissions from buildings [tons GHG]

Cooking

House heating

Natural Gas

Oil

Coal

Electricity/District Heating excluded (not on-site)

Source:https://docs.google.com/presentation/d/1YFDsizQTxMvClI0R2rf00jJoY4TSoPFx473tJnmVijA/edit#slide=id.g2f1c4a00eba_0_123

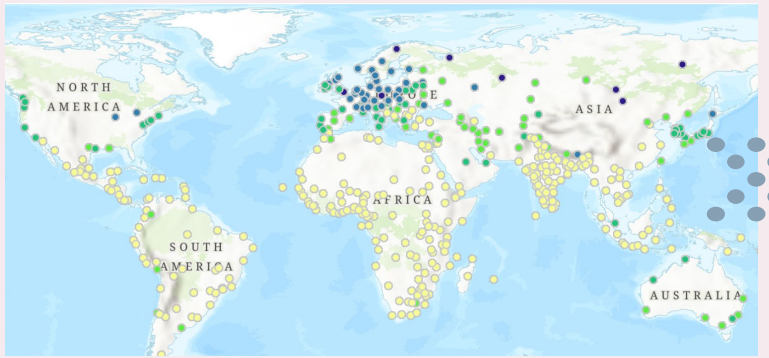
Problem Statement

Develop an approach to estimate **global onsite building emissions in high resolution grid.**

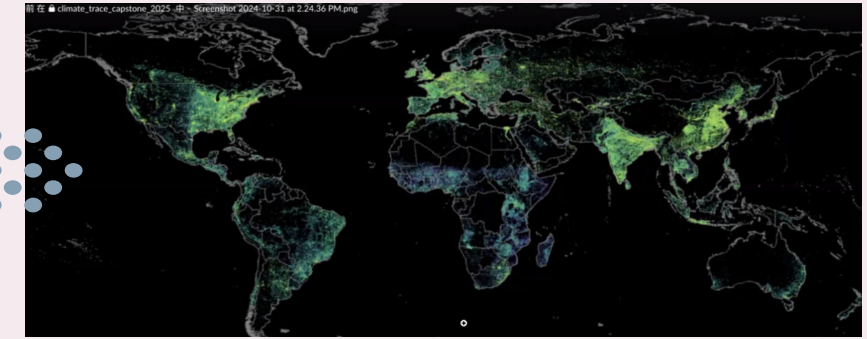
- Share **open-source methodologies** for replicability
- **Validate models** for uncertainty assessment

Potential solution

We will investigate machine-learning-based techniques to estimate global emissions for both residential and commercial buildings, scaling from **483 data points to the whole world at resolution of 1 km²**.



483 EUI data distribution



schematics

Our Plan

MVP

Goal to Reach

Preparation → Feature Selection → Feature Engineering → Modelling → Evaluation → Estimate energy consumption → Validate energy consumption

- Understand problem-kick off meeting, and related material review
- Learn domain tools and library(ArcPy, QGIS, Geopandas, Rasterio, rasterstats, Osgeo (GDAL))
- Literature review, start from assigned paper and explore

- From Previous study, including behavior, building, social economic, weather data.
- Literature review on EUI prediction
- From domain specific models, base models,
- From Expert and life experience.

- Conventional feature engineering tactic to improve representative power of selected data.
- Specific scale down engineer from Country or other top down scale data to grid level data.

- Select model options that may fit the data and predict EUI.
- Fit the models to training data, and optimize performance.
- Integrate with satellite image(or graph) model to further improve performance.

- Metrics design(MAPE, WAPE).
- Cross validation across different geo space.
- Develop uncertainty estimates on the prediction results.

- Once the model is defined, estimate EUI consumption some points

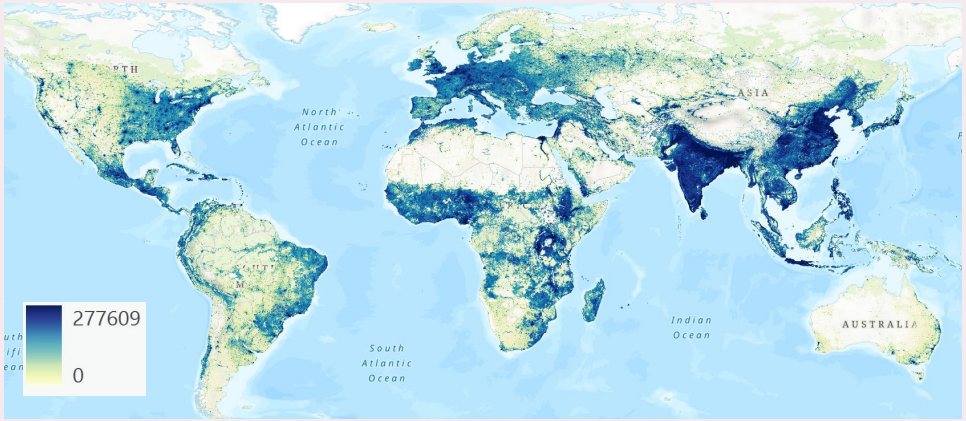
- Validate the estimated consumption from the previous step

Feature Map–work in progress

From previous study, literature review and further interpretation of the project.

Building geometry data	Weather	Socioeconomics	Policy/Law	Behavior
To be continued	Temperature, Heating Degree Days(HDD), Humidity, Comfort Index	Human Development Index(HDI), GDP, Population, Educational Index, Income Index, Urbanization	Paris Agreement	To be continued

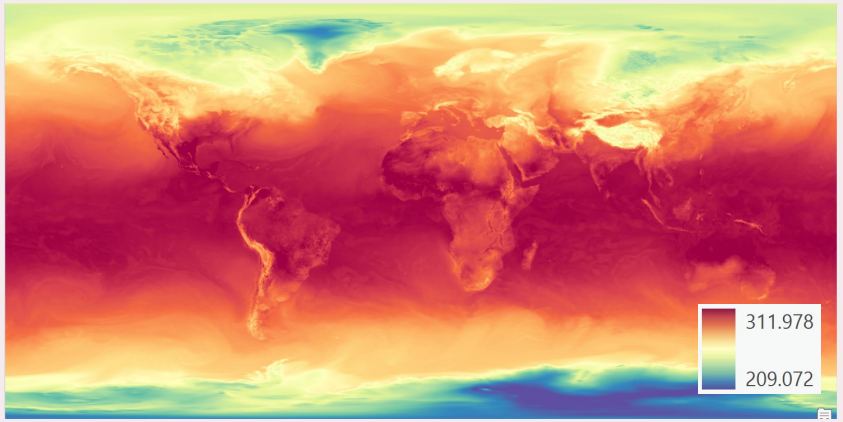
Population 2020



Scale: 1:207,421,922

Resolution: 30 arcsec/pixel
Dara source: European Commission and the Group on Earth Observations

Temperature (1 h)



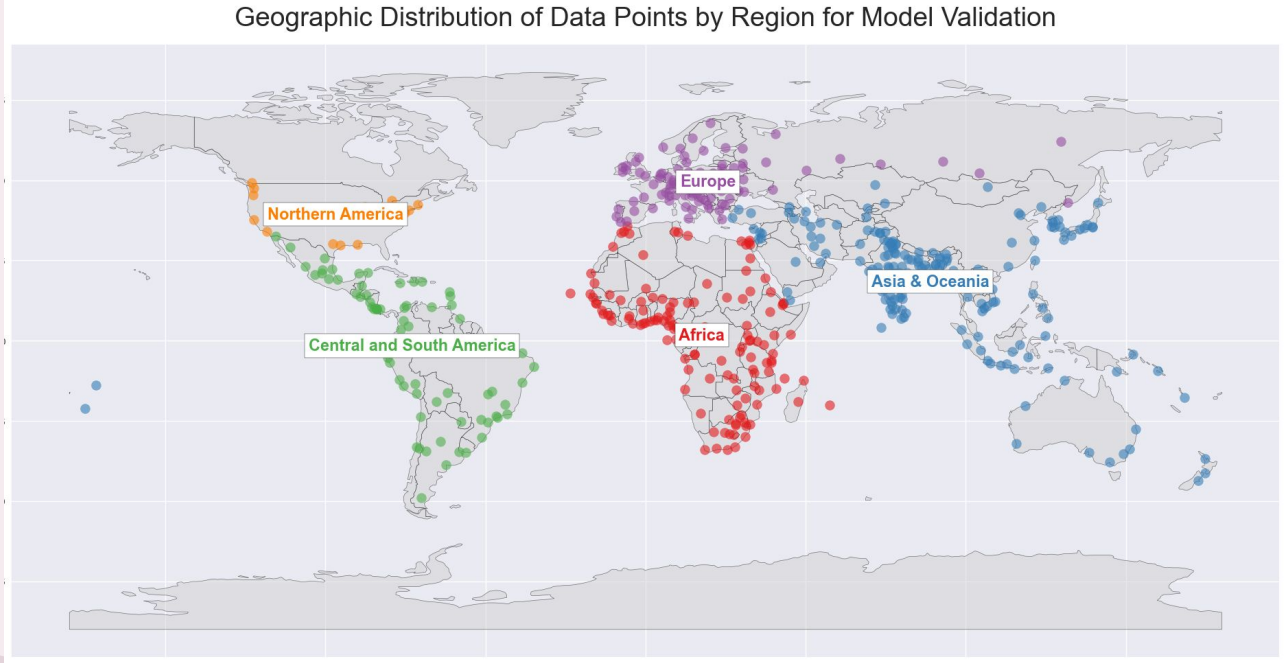
Scale: 1:173,915,198

Dara source: European Commission and the Group on Earth Observations

Modeling

- **Predict EUI**
- **Using ML Models**
- **483 Points**
- **Experimental Design based in Regions for validation**

Experimental Design



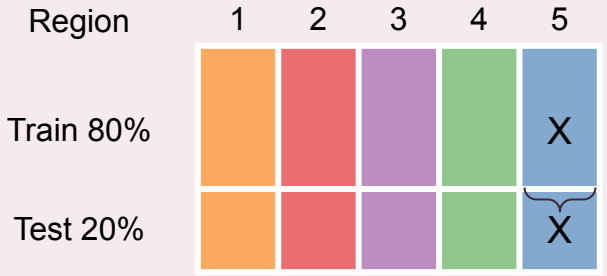
- One challenge with global data is the difficulty of extrapolating results due to regional variations.

- To address this, we will validate our predictions at the regional level and estimate Energy Use Intensity while considering these effects. This approach helps us identify biases and enhances the model's robustness for better extrapolation.

- We de are considering 5 regions

Experimental Design

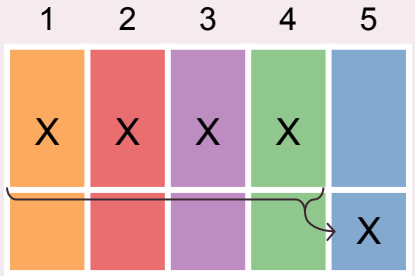
Within-Domain



We train our model on 80% of the data from each region and test it on the remaining 20% of the same region.

To obtain a global result, we calculate the average of our evaluation metrics across the 5 regions.

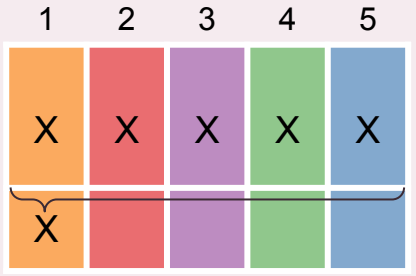
Cross-Domain



We train our model using 80% of the data from 4 regions and test it on the 20% of the 5th region. We repeat this to evaluate the model's extrapolation.

The global result is the average of all outcomes.

All Domains



We use 80% of the data from all 5 regions for training and test on the 20% of the first region. This process repeats for each region,

The global result is the average across all tests.

Potential Impact & Opportunities

Transforming Data into Action

Provide accurate, up-to-date building emissions data on a global scale with 1km^2 resolution.

Democratizing Information

Empower communities that currently lack access to crucial emissions information.

Enabling Policy Innovation

Allow municipalities to design and implement targeted climate policies.

Thank you!

Do you have any questions?

Source: https://climatetrace.org/