GenHack 2025 – Week 1

Data Exploration and Urban Climate Adjustment

Team: 19 - UrbanCoolers

GenHack 2025 – Week 1: Data Exploration

Subject: Urban Heat Islands & Climate Adjustment

Team: [19 - UrbanCoolers]

Date: November 17, 2025

Period: Warm-up - Data Exploration and Understanding

Challenge Overview: The Fundamental Question

How does urbanization affect the accuracy of climate model data?

Our Mission:

- Analyze the Urban Heat Island (UHI) effect using multiple data sources.
- Understand the discrepancies between model data and ground truth.
- Use vegetation (NDVI) as an indicator of urbanization.

Data Sources:

- 🌡️ **ERA5-Land:** Temperature reanalysis (~9km resolution)
- 📍 **Ground stations:** Real temperature measurements (863 in France)
- 🌿 **Sentinel-2 NDVI:** Vegetation density maps (80m resolution)

Data Integration Strategy
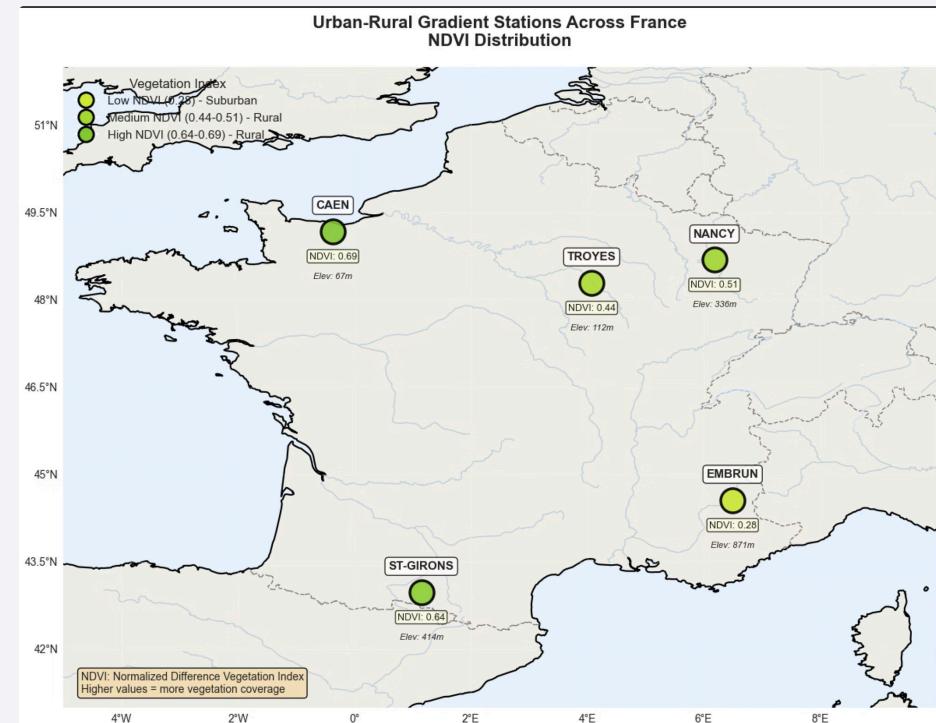
Strategic Station Selection:

5 stations representing an urban-rural gradient across France.
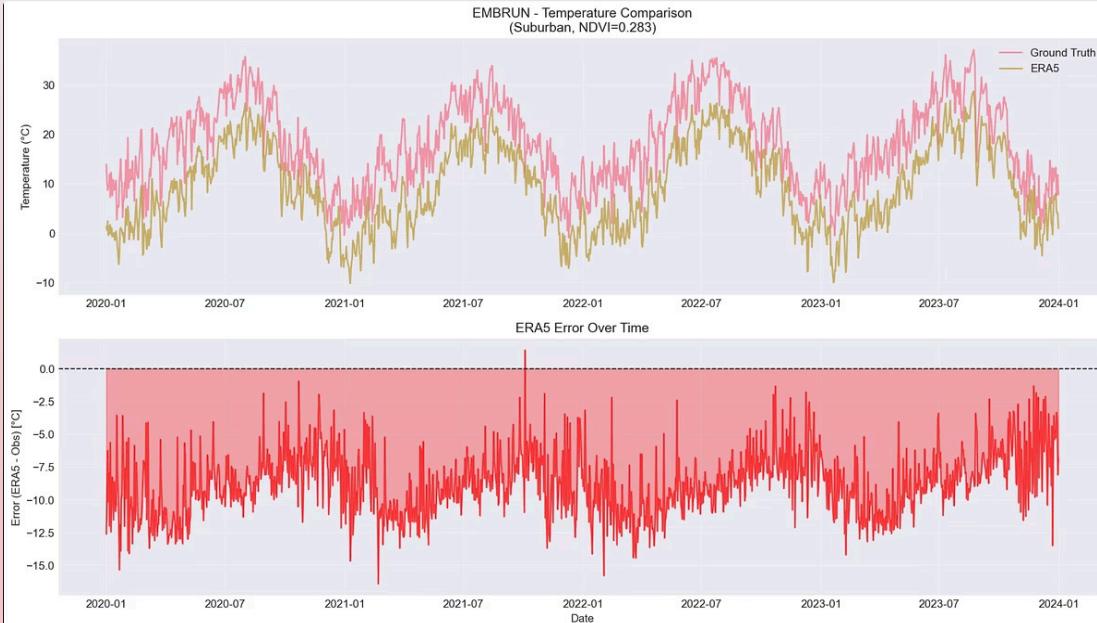
Station	NDVI	Category	Altitude
EMBRUN	0.28	Suburban	871m
TROYES	0.44	Rural	112m
NANCY	0.51	Rural	336m
ST-GIRONS	0.64	Rural	414m
CAEN	0.69	Rural	67m

Methodology:

- Extraction of ERA5 temperatures at station locations.
- Calculation of NDVI values from satellite data.
- Temporal alignment of all data sets (2020-2023).

Urban-Rural Gradient Stations Across France
NDVI Distribution

EMBRUN - Temperature Comparison
(Suburban, NDVI=0.283)

ERA5 Error Over Time

The Irrefutable Evidence: The Urban Cold Bias

EMBRUN Station

(Suburban, NDVI=0.28)

- **Bias: -8.8°C**
  (ERA5 significantly colder than reality)

- **RMSE: 9.12°C**
  (Very significant error)

- **Correlation: 0.960**
  (Good pattern match, but erroneous values)

This is the signature of the Urban Heat Island:

- ERA5's coarse resolution misses local urban warming.

- Systematic underestimation in built-up areas.

- More pronounced in areas with low vegetation.

Made with GAMMA

Rural Areas Show High Accuracy

Performance of Rural Stations (Temperature Accuracy by Station):

Station	Category	NDVI	Bias (°C)	RMSE (°C)	MAE (°C)	Corr
EMBRUN	Suburban	0.283	-8.80	9.12	8.80	0.960
TROYES-BARBEREY	Rural	0.441	-1.38	1.88	1.62	0.989
NANCY-OCHEY	Rural	0.512	-1.02	1.73	1.42	0.989
ST-GIRONS	Rural	0.638	-2.76	3.23	2.93	0.974
CAEN-CARPIQUET	Rural	0.685	-0.66	1.29	1.05	0.985

Key Observation:

- Rural areas: **Excellent accuracy** (RMSE < 2°C)
- Suburban areas: **Low accuracy** (RMSE > 9°C)
- **10 times the difference** in error magnitude!

Quantitative Relationship Established

A strong correlation between NDVI and model error has been identified, highlighting the critical role of vegetation in the accuracy of climate forecasts.

NDVI vs Bias
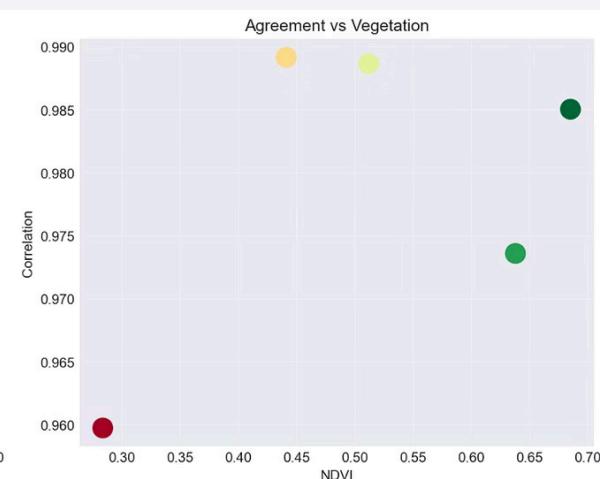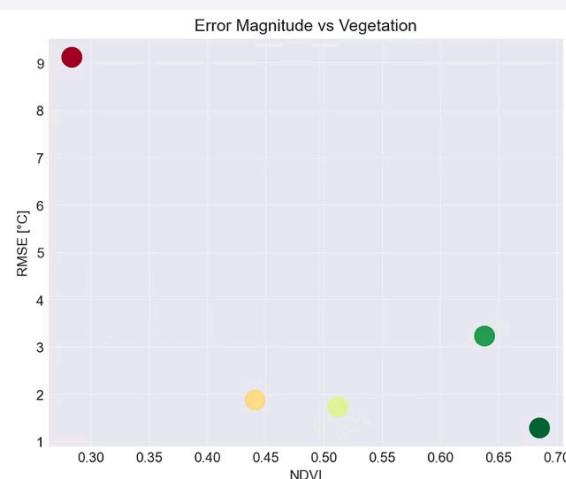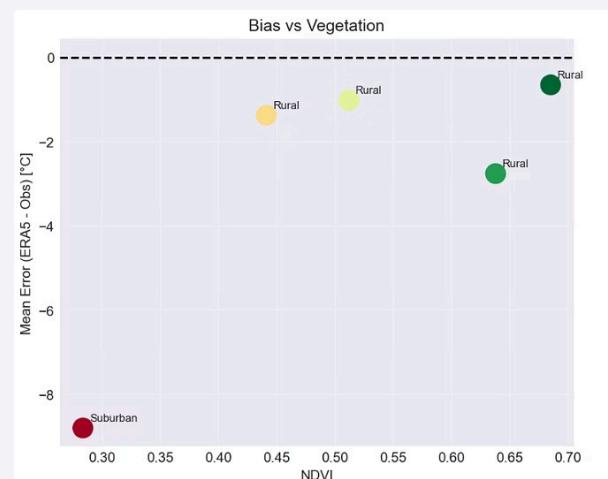
Correlation of **+0.758**

NDVI vs RMSE

Correlation of **-0.760**

- 📈 **Higher NDVI** (more vegetation) → **Lower Error.**

📉 **Lower NDVI** (urban areas) → **Higher Error**
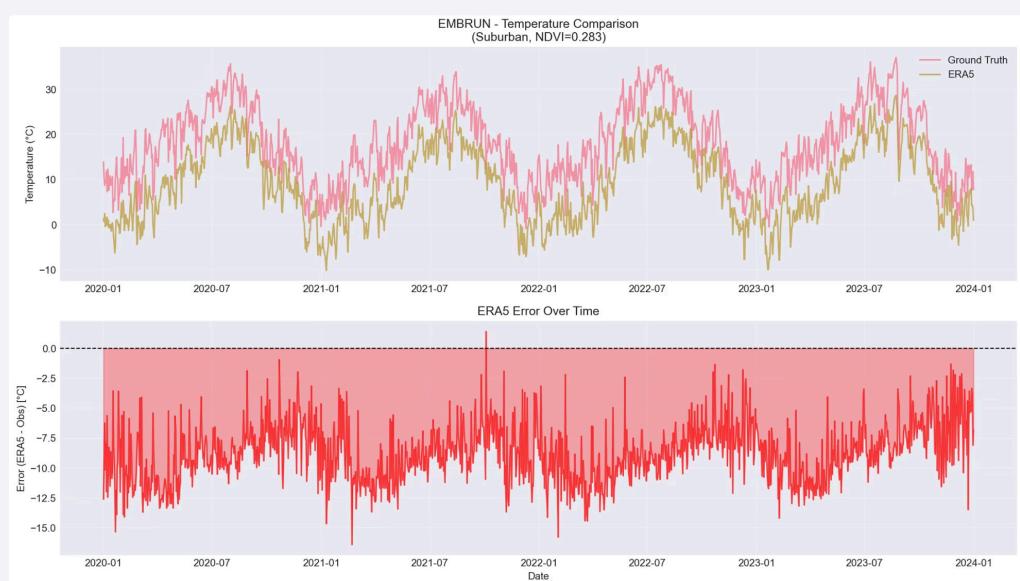
Visual Evidence:

Each point represents a weather station. A clear trend: less vegetation = greater cold bias.

# Visual Evidence: The Urban-Rural Divide
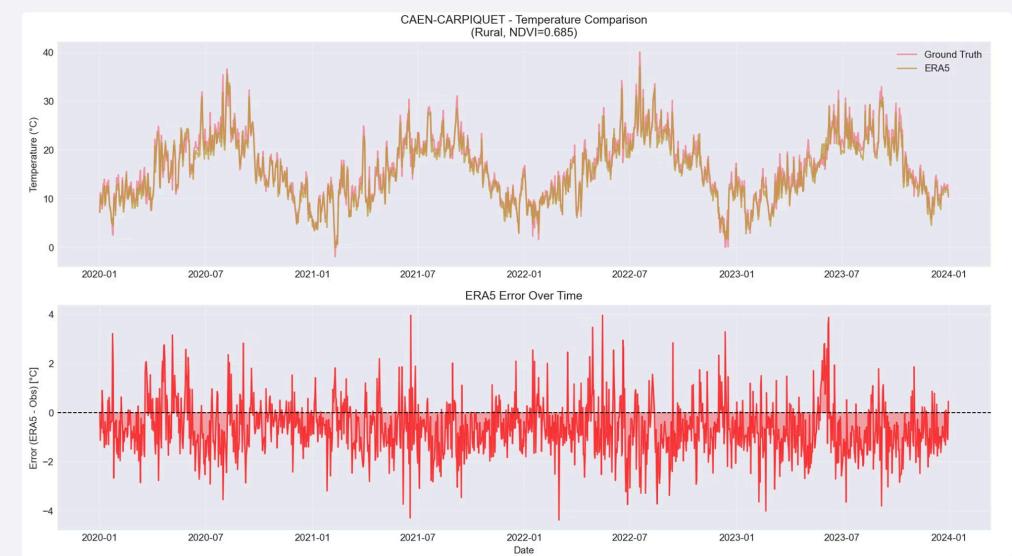
EMBRUN (Urban):

- Significant and persistent discrepancy between ERA5 and observations.

- Constant cold bias across all seasons.
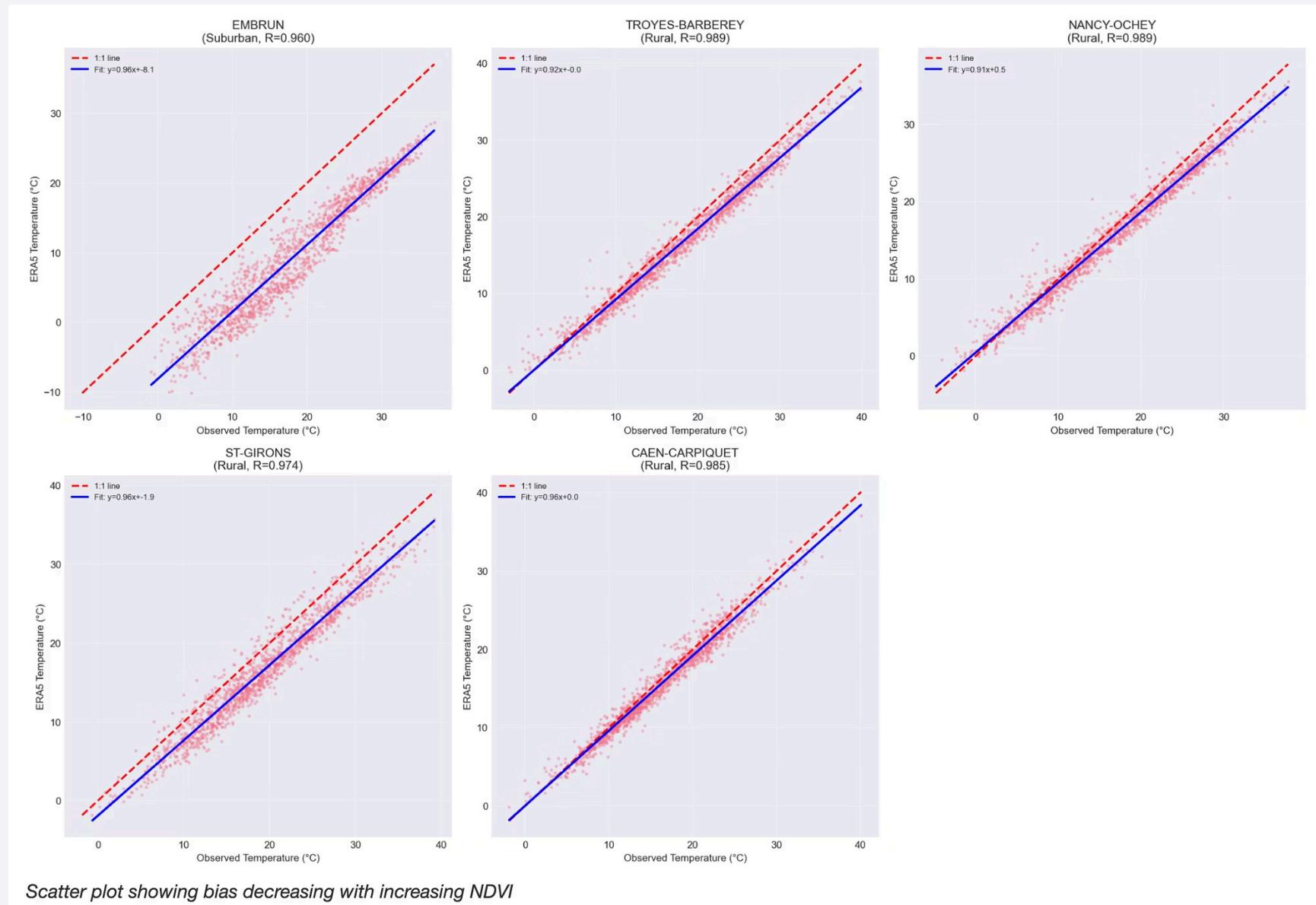
CAEN (Rural):

- Lines follow closely.

- Minor discrepancies, excellent overall fit.

The Story in a Single Graph: Urban heat islands create localized warming that ERA5 cannot capture, illustrating the importance of downscaling.

Visual Evidence: The Urban-Rural Divide

Scatter plot showing bias decreasing with increasing NDVI

Limitations & Data Notes

1

Current Constraints

- ✅ **Complete:** ERA5 (4 years), Ground stations (43 French)
- ⚠️ **Partial:** NDVI (1 file out of 19 analyzed)
- 📅 **Temporal:** Only one NDVI snapshot from Spring 2022

2

Why This Matters

- Vegetation changes seasonally → UHI intensity varies.
- Need for a comprehensive dataset for robust seasonal analysis.

3

Next Steps for Data
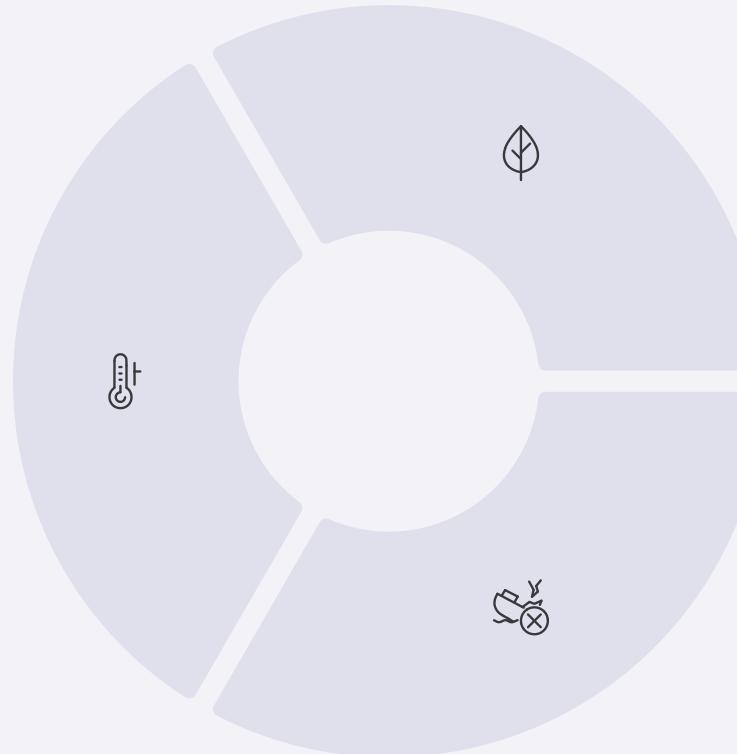
- Implement memory-efficient processing for the full NDVI dataset.
- Analyze seasonal patterns of the UHI effect.

Key Takeaways and Implications

Urban heat islands have a significant and underestimated impact on climate modeling, with concrete repercussions.

NDVI as Predictor

Strong indicator of model accuracy.

ERA5 Cold Bias

Systematic bias in urban areas due to UHIs.

Substantial Error

Up to 8.8°C difference.

Real Impact:

- 🏙️ **Urban Planning:** Underestimated heat risk.
- 🏥 **Public Health:** Misjudged heatwave impacts.
- ⚡ **Energy Demand:** Underestimated cooling needs.
- 📊 **Climate Science:** Attenuated urban warming signals.

Next Steps – Week 2 Preview

Week 2 Focus: Visualization & Communication

Planned Deliverables:

1. **Interactive UHI maps** of France and Europe
2. **Seasonal analysis** of UHI intensity variations
3. **Diurnal cycle** examination (if hourly data available)
4. **Full NDVI dataset** integration
5. **Compelling visual narrative** for policymakers