ClimateWins: Predicting Weather Variations with Machine Learning

Elena Svirko, PhD

August 2025

Introduction: Climatewins and their Goals

About ClimateWins

- European nonprofit addressing rising extreme weather events
- Focus on prediction and prevention through advanced analytics
- Believes machine learning can improve disaster preparedness

Long-term Goals

- Identify weather patterns outside the regional norm in Europe.
- Determine if unusual weather patterns are increasing.
- Generate possibilities for future weather conditions over the next 25 to 50 years based on current trends.
- Determine the safest places for people to live in Europe over the next 25 to 50 years.

Weather Classification Project Overview

Objective

Classify days as **pleasant** or **unpleasant** using a minimal number of predictors

- **Data**
  - Data Source: European Climate Assessment & Dataset Project
  - 18 weather stations across mainland Europe
  - Daily observations from 1960 through 2022
  - Measures include temperature, precipitation, humidity, cloud cover, etc.
- **Optimization Method**
  - Gradient Descent

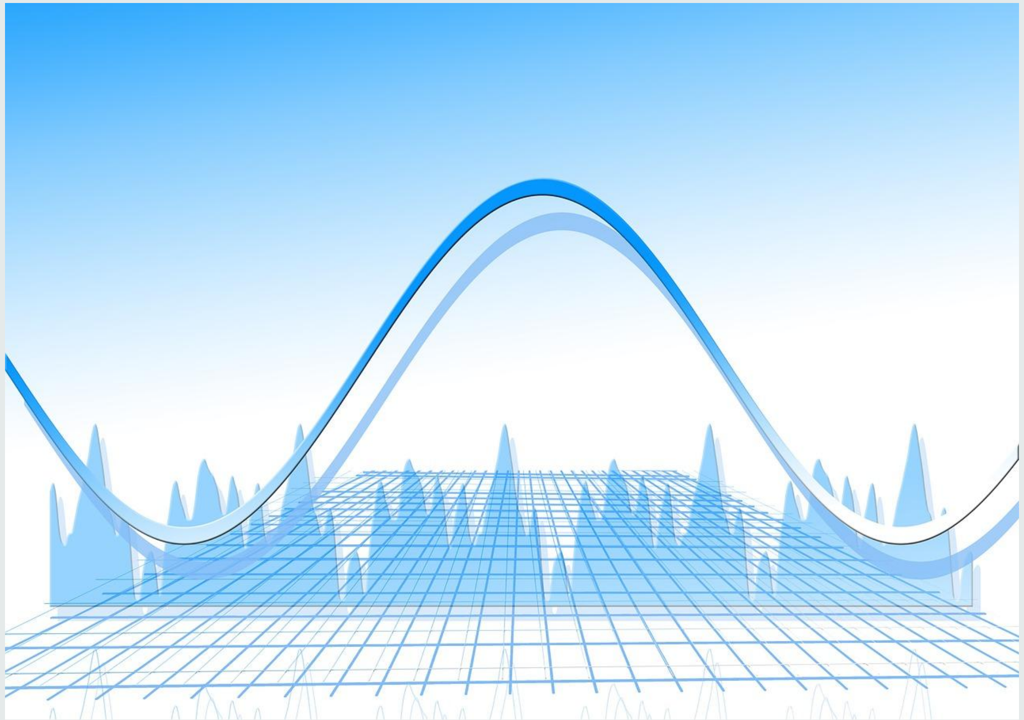
Evaluation & Performance Metrics

- **Data Challenge**

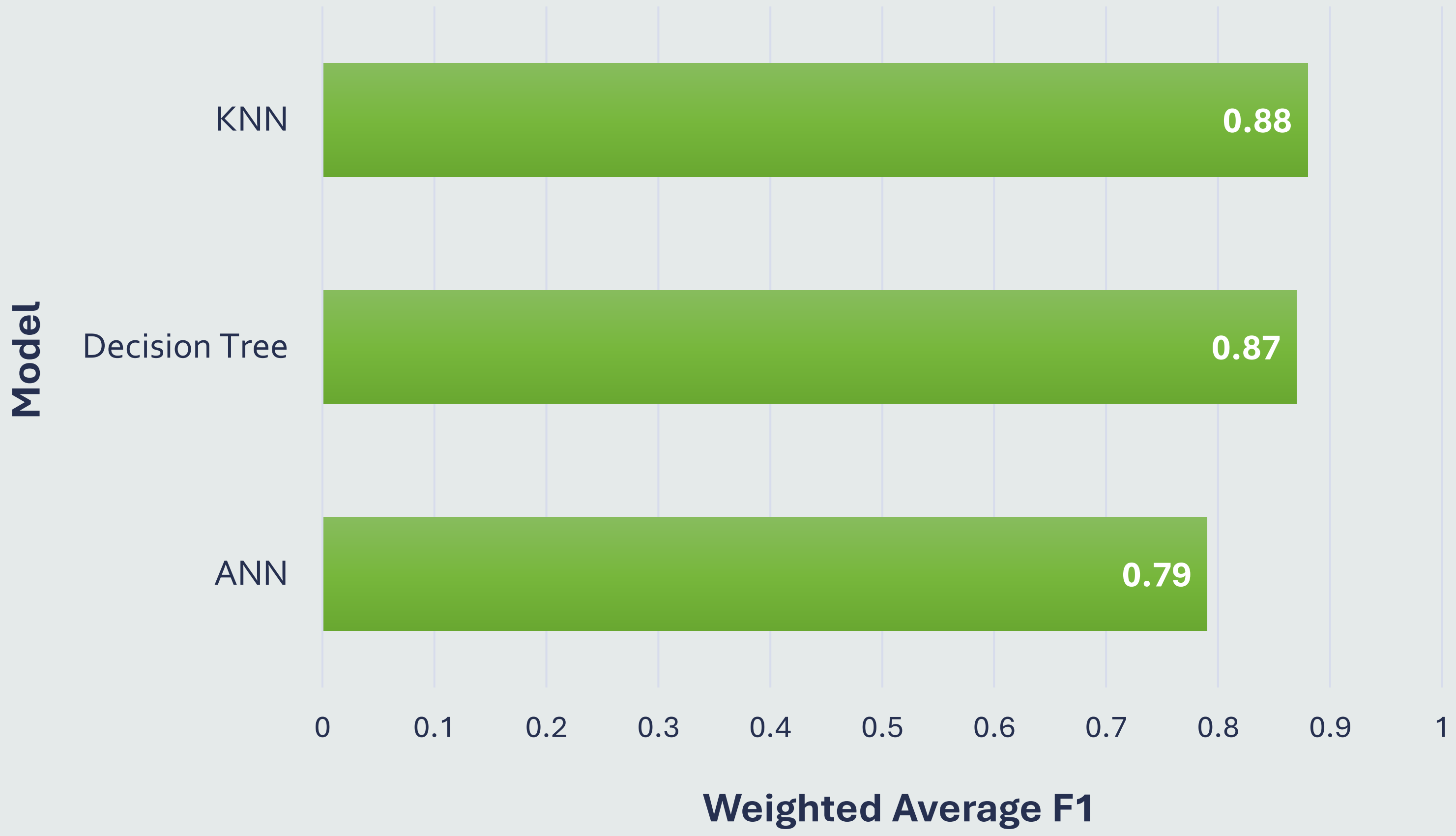
- Class imbalance: Far fewer pleasant days than unpleasant
- Risk: Model could predict "unpleasant" for all days and still show high accuracy

- **Evaluation Metric: Weighted F1 Score**

- Accounts for precision and recall across both classes
- Weighted to reflect class imbalance
- More informative than simple accuracy



Results: KNN, Decision Tree and ANN



K-Nearest Neighbours (KNN)

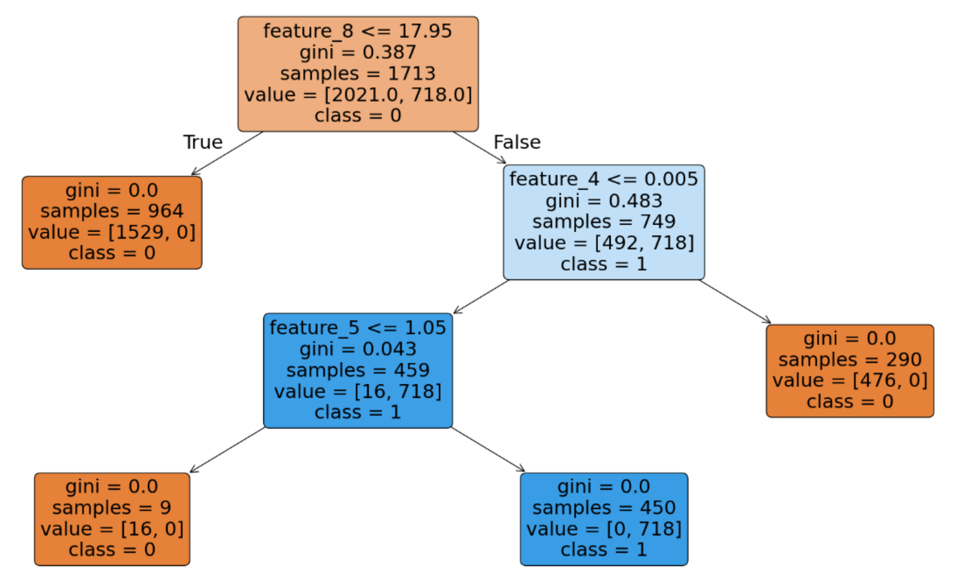
Decision Tree

Artificial Neural Network (ANN)

Random Forest Analysis

- Optimization with Random Search → weighted F1 scores **0.99–1.0**
- Most influential stations: **Madrid, Belgrade, Budapest**
- Only **3 features required**: maximum temperature, precipitation, sunshine

Fig 1. Decision Tree from Random Forest



Why this matters:

- Accurate classification with fewer inputs
- Lower computational costs
- Robust even when some data is missing
- Scalable to real-world, resource-limited settings

Convolutional Neural Network (CNN)

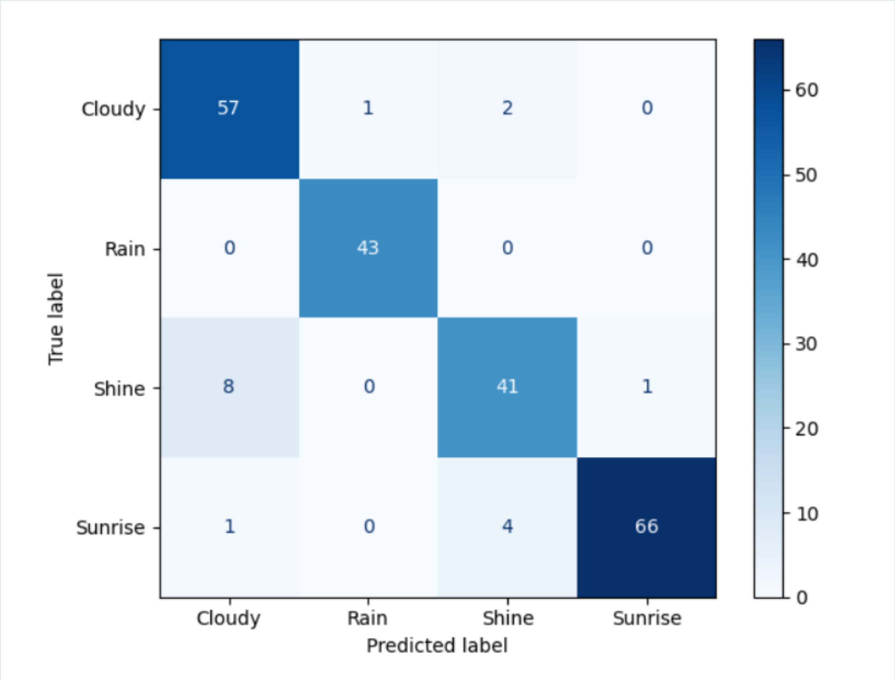
- **Weather classification:** Able to classify *pleasant vs. unpleasant* weather
  - Weighted F1 = **0.99**
- **Image classification:** Classified images into **4 categories:**

  ☁️ Cloudy, 🌧️ Rainy, ☀️ Sunny, 🌅 Sunrise
  - Weighted F1 = **0.92***

👉 Proof of concept: CNN can classify **radar/weather imagery**

👉 Potential for **operational use** in weather prediction

Fig 2. Confusion Matrix for Image Classification CNN



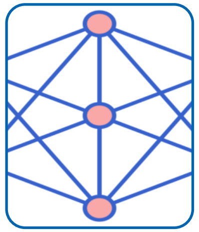
* Achieved using EfficientNetB0

Future Directions

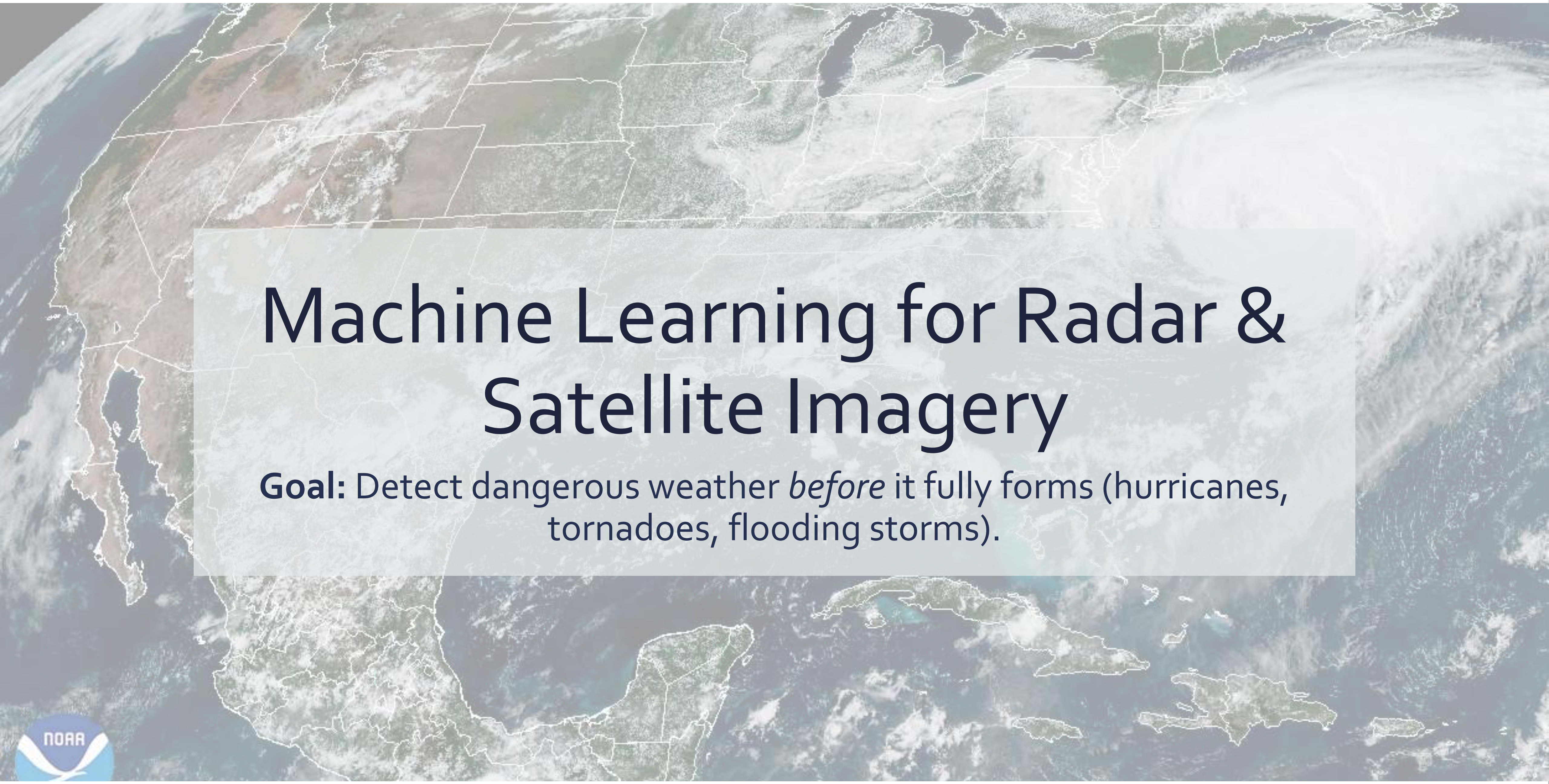
Application of machine learning for Climatewins goals

Radar & Satellite Imagery

Climate “Twin” Cities



Hybrid Models with Climate Projections



Machine Learning for Radar & Satellite Imagery

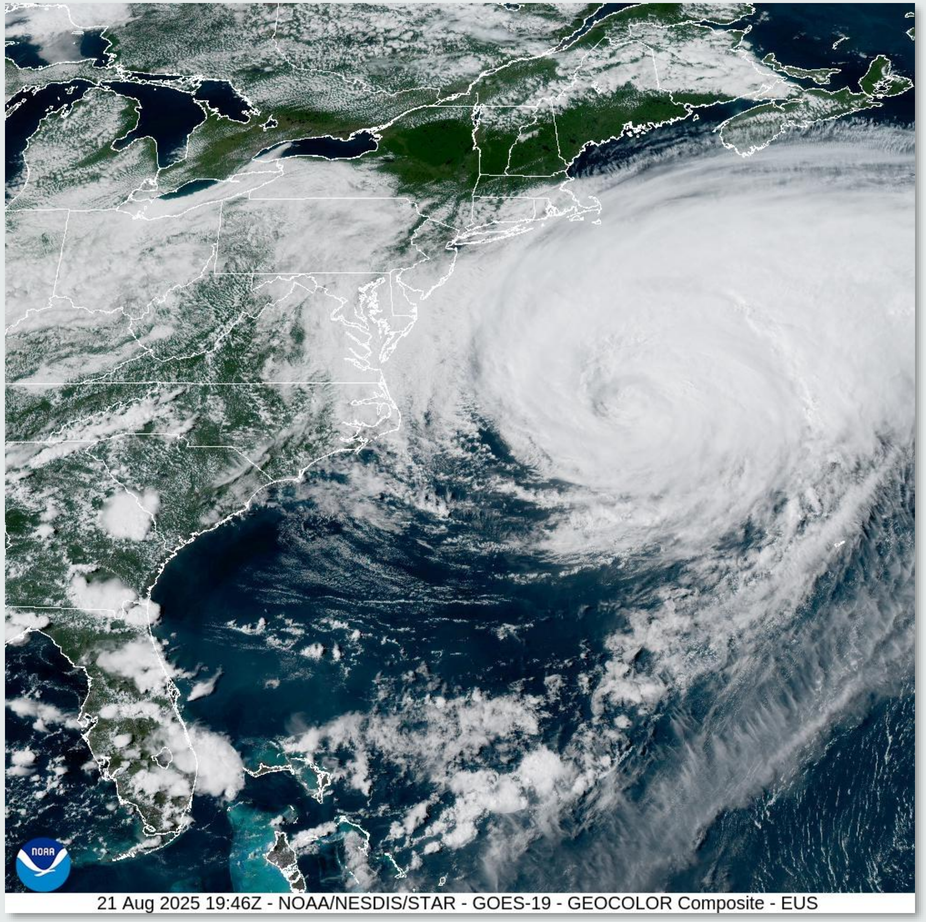
Goal: Detect dangerous weather *before* it fully forms (hurricanes, tornadoes, flooding storms).

GeoColor satellite imagery courtesy of CIRA/NOAA

Machine Learning for Radar & Satellite Imagery

- **CNN**
- Good at static image classification
- **LSTMs (Long Short-Term Memory networks)**
- Good at capturing sequences over time
- **CNN + LSTM Hybrid Models (ConvLSTM)**
- Capture changes over time in image sequences.
- ConvLSTM = combines CNN (space) + LSTM (time).
- **Best for storm evolution prediction**.

Fig 3. Hurricane Erin moving along the Atlantic coast



Machine Learning for Radar & Satellite Imagery

- **Generative Adversarial Networks (GANs)**

- **What they do:** GANs can **simulate future weather images** by learning how storms evolve.
- **Use case here:** GANs might be used for **nowcasting (0–6 hour predictions)**, generating what the next radar/satellite image might look like, and flagging when dangerous structures (like hurricane eyewalls or tornadic supercells) are about to emerge.

- **Transformers (Vision Transformers / Spatio-Temporal Transformers)**

- **What they do:** Newer models that can handle **both images and sequences**, learning complex relationships across space and time.
- **Use case here:** Applied to satellite and radar datasets, transformers could identify **long-range dependencies**, such as a chain of storm systems that increases flooding risk. This is cutting-edge but very promising.

Finding “Climate Twin Cities”

Goal: Identify European cities with historically similar weather patterns, then track how those relationships change under climate change.

Finding "Climate Twin Cities"

◆ **Why it matters:**

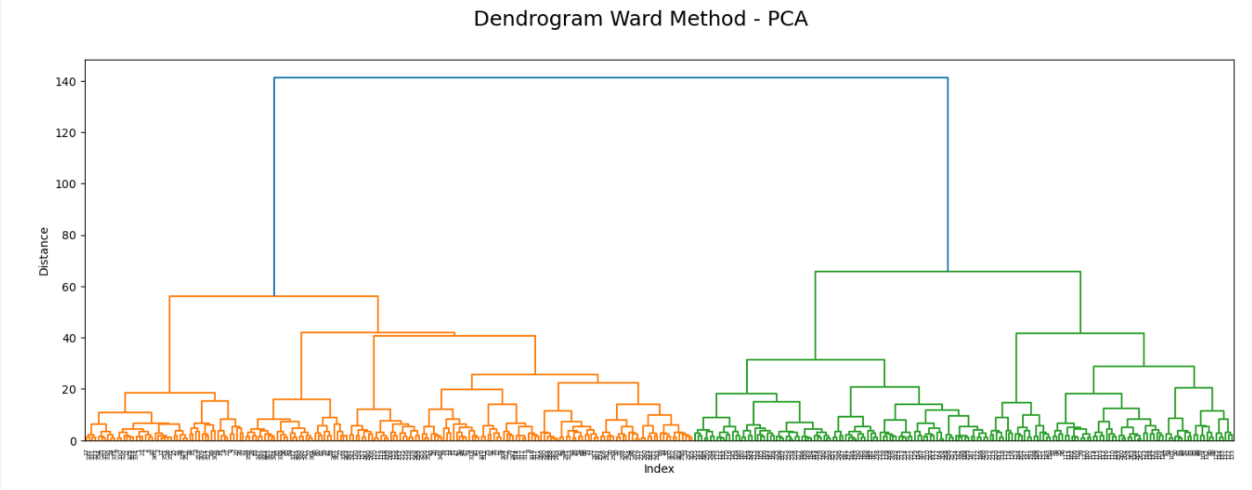
✓ Cities can **learn from their "climate twins"** (e.g., flood response, heatwave planning)

✓ Policymakers can **anticipate local risks** by studying other regions' experiences

Methods:

- **PCA (Principal Component Analysis)**
  - reduces climate data into key factors
- **Clustering (K-means / Hierarchical)**
  - groups cities into "climate families"

Fig 4. Example of Cluster Analysis of European Weather Data



Finding "Climate Twin Cities"

◆ **Graph Neural Networks (GNNs)**

- Map cities as networks of climate similarity

◆ **Divergence Tracker**

- Highlights cities whose climates are shifting fastest

💡 **Action Insights**

- Learn from cities that already face similar conditions to prepare early for new climate challenges.

Hybrid Model with Climate Projections

Goal: Predict how local patterns will evolve under global warming scenarios.

Hybrid Model with Climate Projections

The **IPCC** (Intergovernmental Panel on Climate Change) regularly publishes scenarios showing how the climate may change under different levels of global warming and emissions.

Combine **historical weather data** with **IPCC climate model outputs**

Machine learning could detect how **local patterns will evolve** under global warming scenarios

Hybrid Model with Climate Projections

- **Insights for Action:**
  - Safer regions for long-term settlement
  - Early-warning for unusual weather
  - Evidence-based policy guidance
  - Supports ClimateWins' mission of resilience

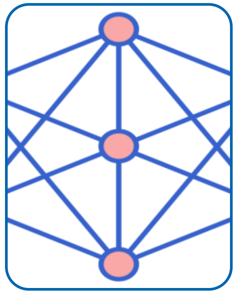
Data Needed Beyond Historical Weather

Radar & Satellite Imagery

- National radar mosaics
- Satellite bands (e.g. GeoColor)
- Severe event reports (floods, storms, tornadoes)

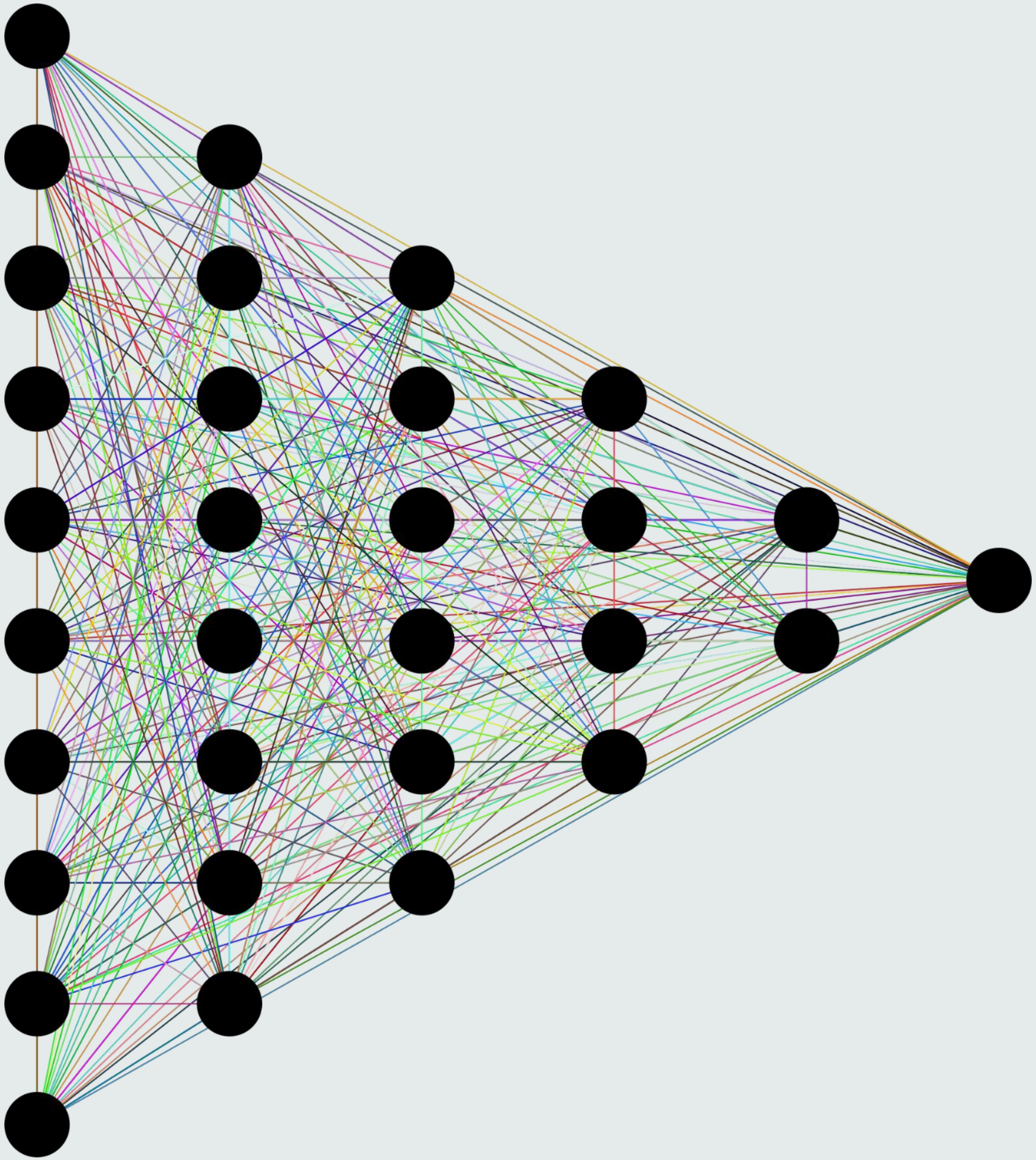
Climate "Twin" Cities

- Climate extremes (heatwaves, floods)
- Geographic context: elevation, distance to coast
- Socio-demographics



Hybrid Models with Climate Projections

- IPCC climate model outputs
- Vulnerability layers: floodplains, infrastructure, health data



Final Thoughts

- ✅ **Machine learning** offers powerful new tools for climate resilience
- ✅ My analysis shows **even simple models achieve strong accuracy**
- ✅ **Future direction**: expand datasets, explore hybrid models, connect directly to policy goals
- 🌍 **ClimateWins** has the potential to become a leader in **AI-driven weather prediction**

Thank you!

Questions?

📩 Feel free to reach out:
elena.y.svirko@gmail.com

🔗 Explore the code on GitHub:
https://github.com/ElenaSvirko/Climatewins_Machine_Learning_Project