Machine Learning for Climate Analysis

Tony Marziale

7/5/2025

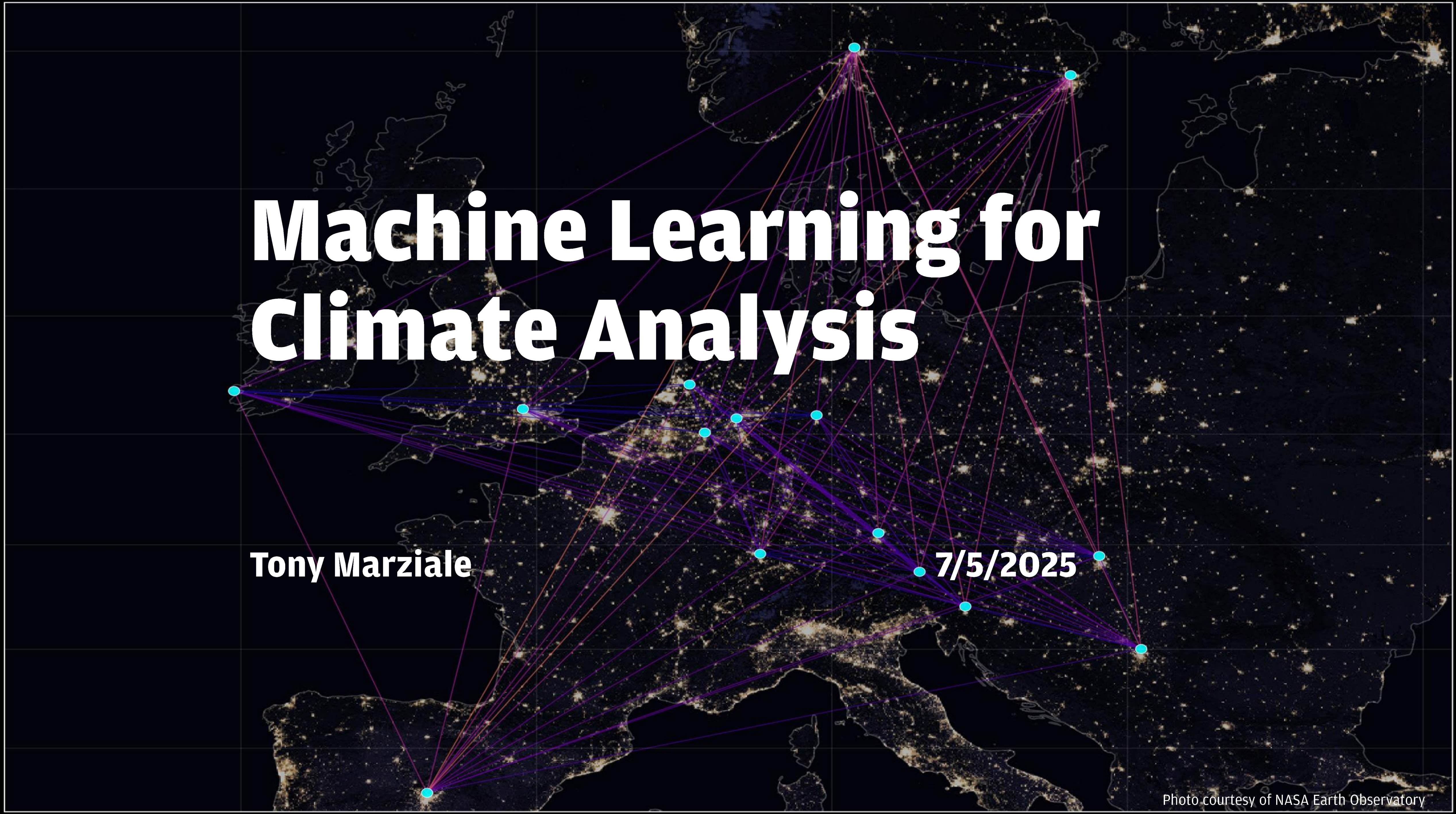


Photo courtesy of NASA Earth Observatory

Hypotheses

Seasonal Amplitude Compression

Though temperatures have risen overall, Winters have warmed faster than Summers resulting in smaller inter-seasonal differentials, with more dramatic effects in Northern Europe than Southern Europe.

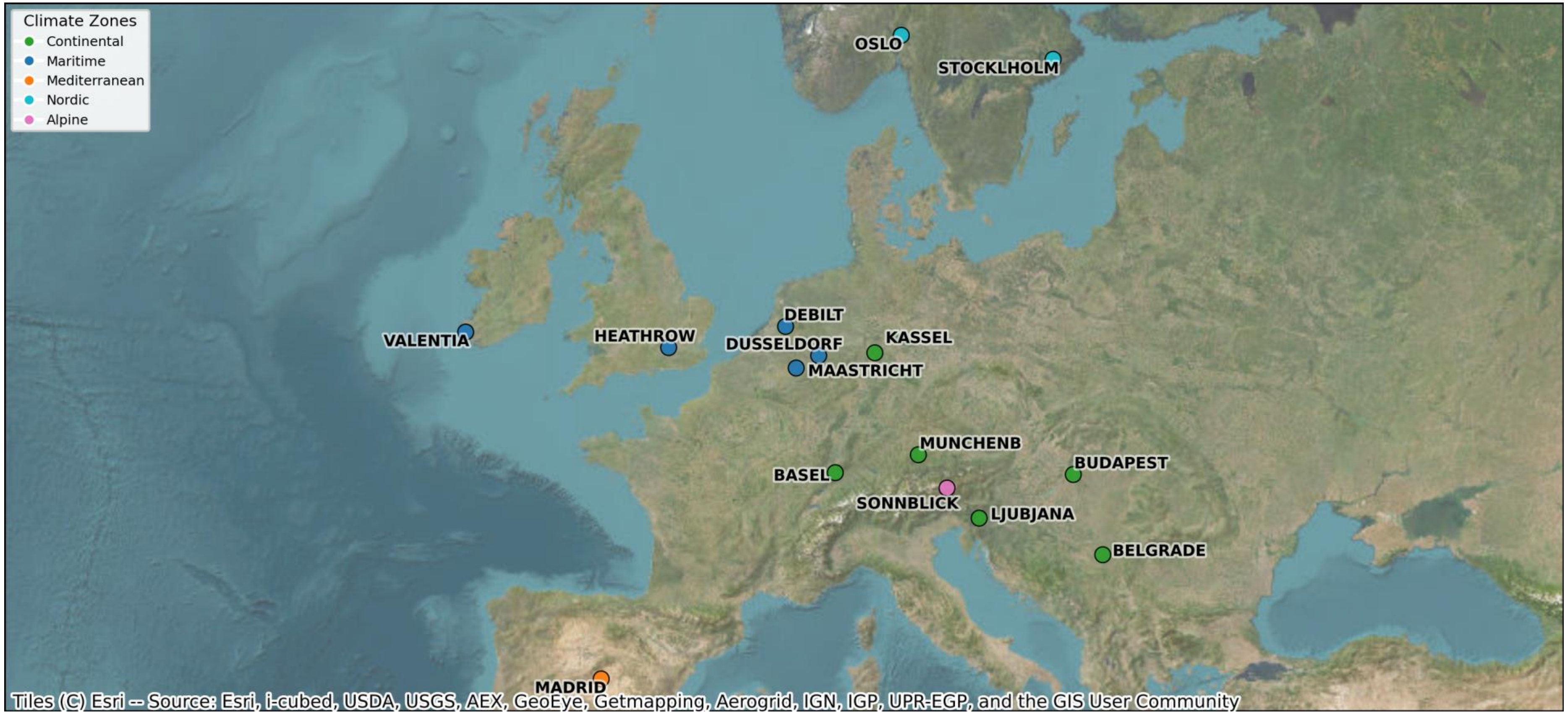
A Moveable Feast - Shifting Picnic Weather

The timing and quantity of days with pleasant weather have shifted as Spring begins earlier and Autumn stretches later into the year.

Compound Weather Event Frequency

Multi-variable weather extremes (such as days with very high heat, low humidity, and high solar radiation) have become more frequent over time.

Data Overview



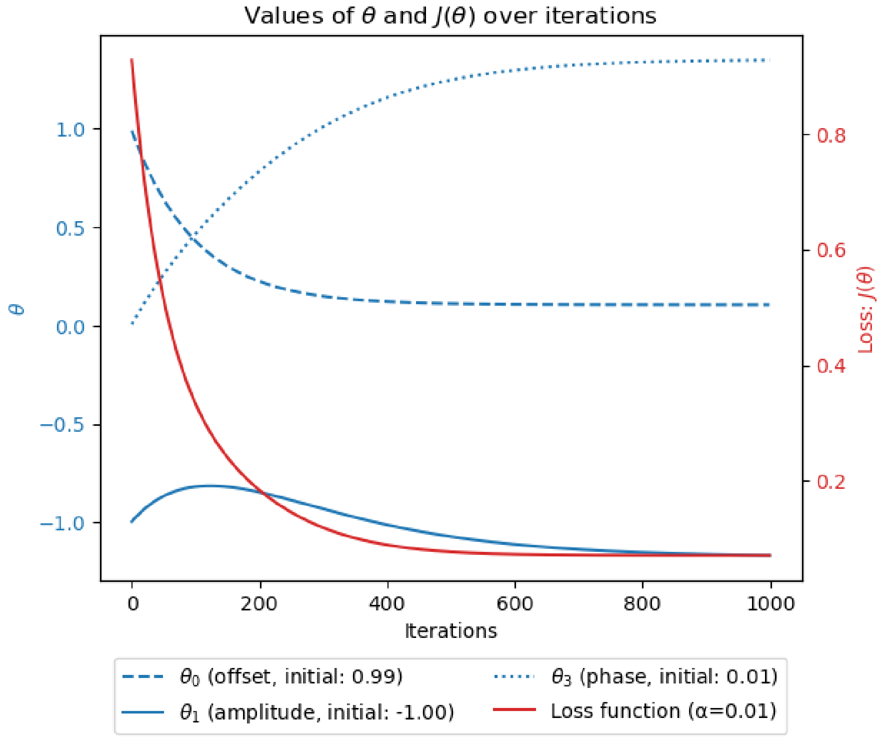
Contents: 344,250 daily weather observations across 15 European locations (1960-2022).
Features include seven weather measurements (temperature, humidity, precipitation, sunshine, radiation).

Limitations and Possible Bias:
Small sample of only 15 stations could introduce spatial bias.
Data quality can vary by location and time of collection.
Source: European Climate Assessment and Dataset

Gradient Descent Optimization

What is Gradient Descent? Gradient descent is an optimization algorithm that finds the best parameters for a model by iteratively adjusting them in the direction that most reduces prediction error. It powers diverse algorithms from linear regression to deep neural networks, making it fundamental to modern Machine Learning.

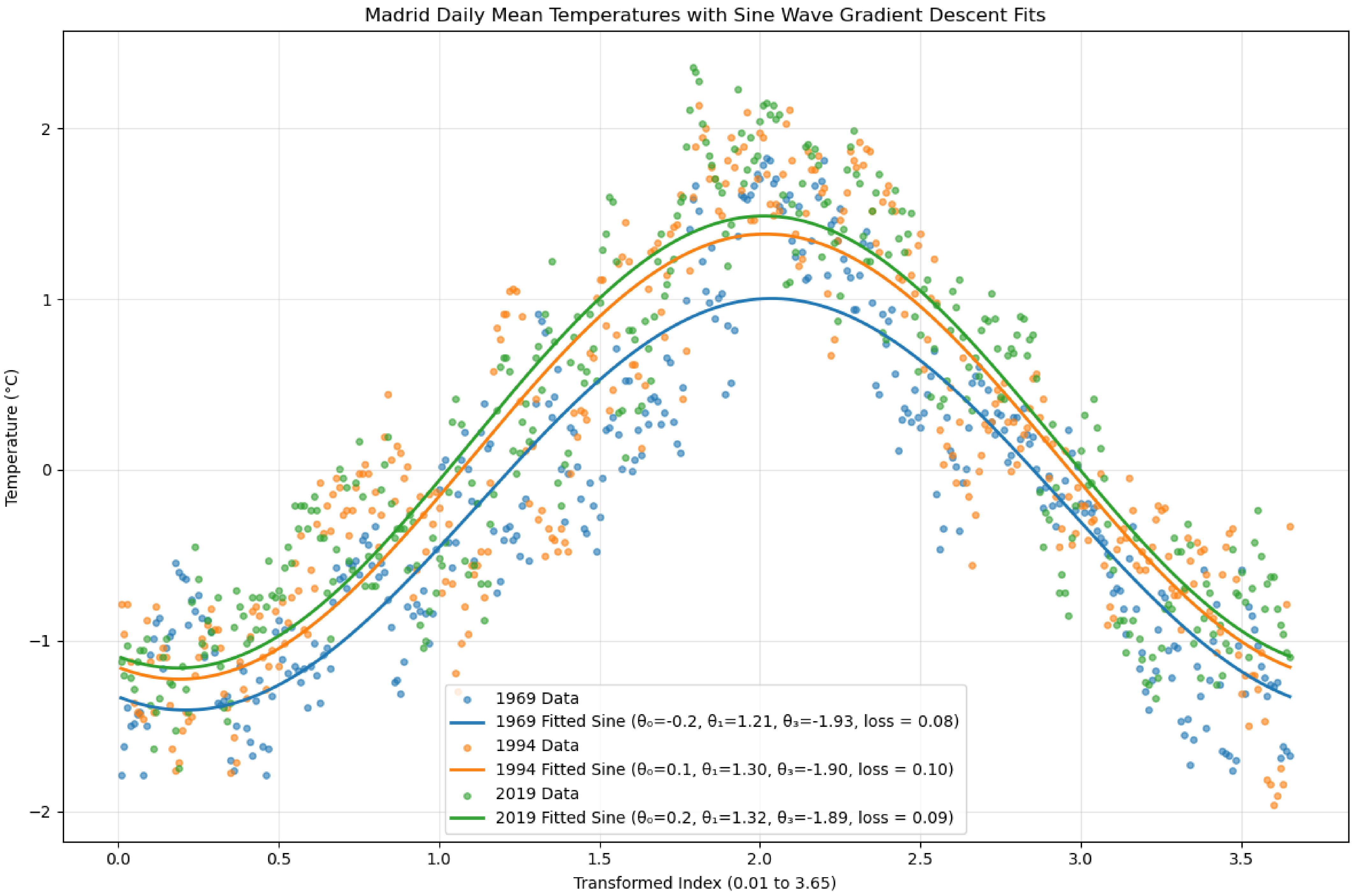
Sine Wave Fitting - Gradient Descent



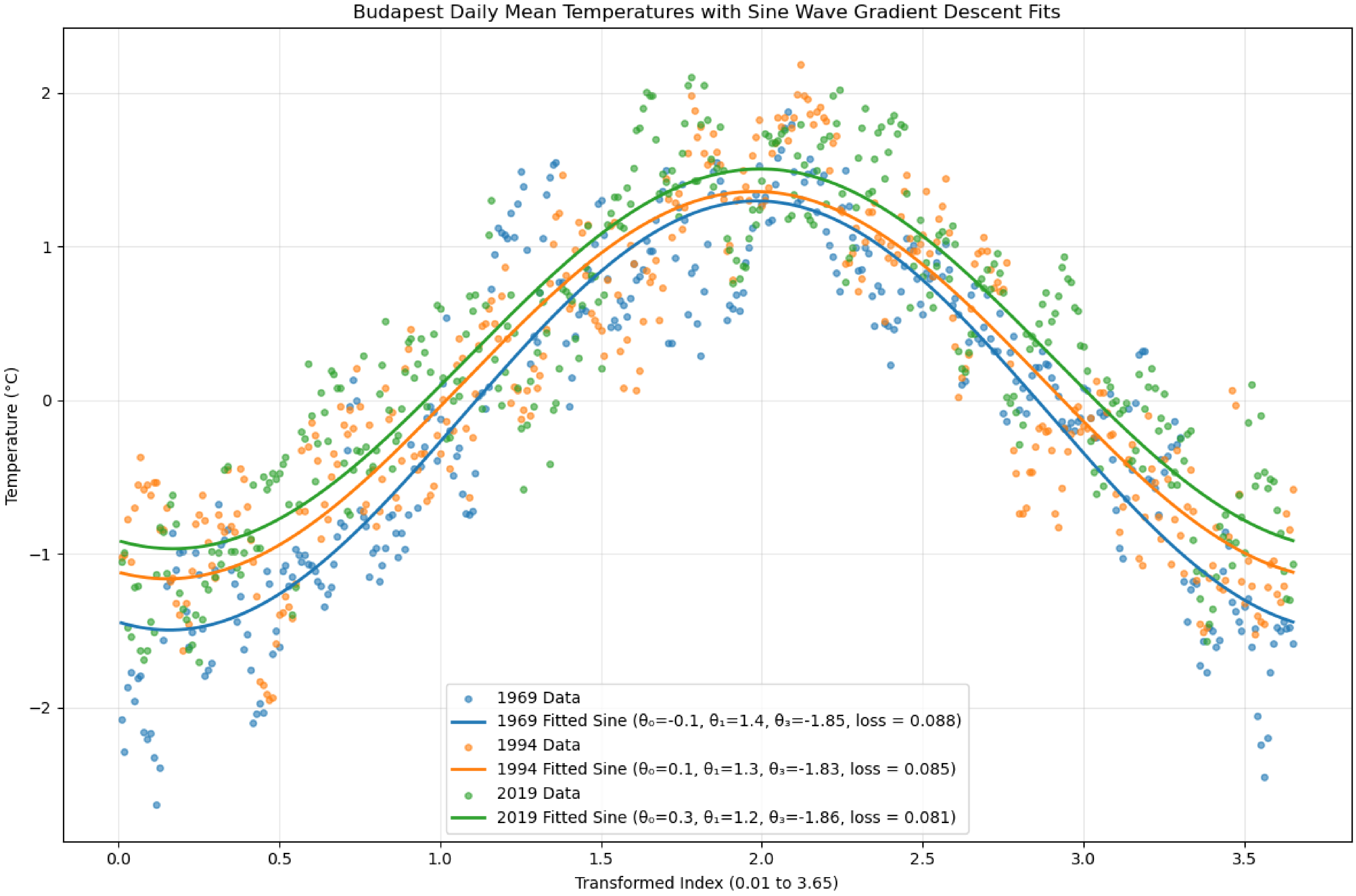
For our data, we can use gradient descent to fit a sine wave to our daily temperature data to discover the optimal seasonal temperature curve for each city and reveal climate patterns without manual parameter selection.

Gradient descent optimizes coefficients θ_0 (baseline temperature), θ_1 (seasonal amplitude), and θ_3 (phase shift, i.e. peak timing) by:

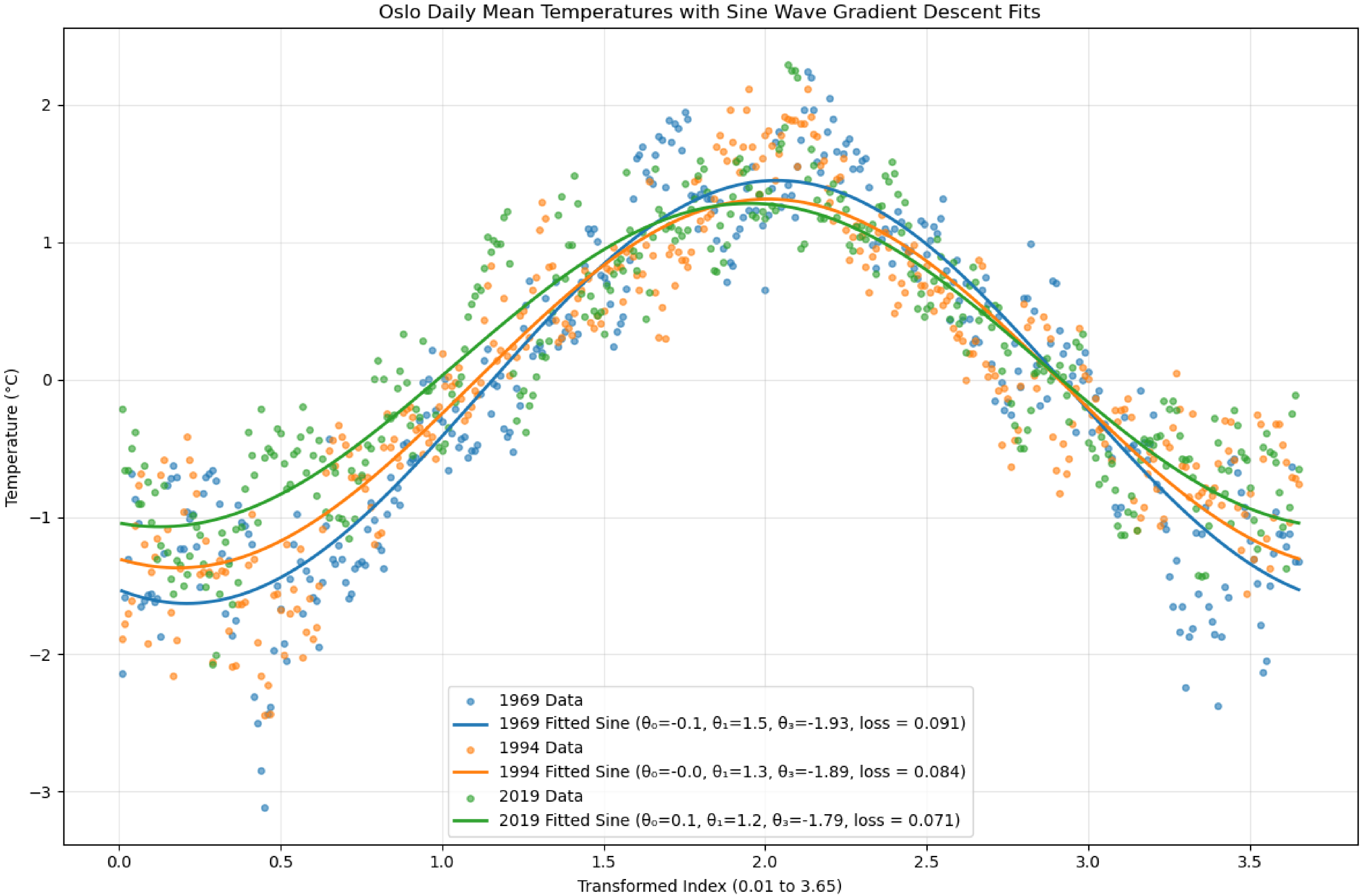
1. Starting with rough approximate (or random) parameter values
2. Calculating prediction error across all temperature measurements
3. Computing gradients showing how each parameter affects total error
4. Adjusting parameters in the direction that reduces error most rapidly
5. Repeating until the model accurately captures seasonal temperature cycles



Madrid Daily Mean Temperatures with Sine Wave Gradient Descent Fits
Temperature (°C)
Transformed Index (0.01 to 3.65)
1969 Data
1969 Fitted Sine (θ₀=-0.2, θ₁=1.21, θ₃=-1.93, loss = 0.08)
1994 Data
1994 Fitted Sine (θ₀=0.1, θ₁=1.30, θ₃=-1.90, loss = 0.10)
2019 Data
2019 Fitted Sine (θ₀=0.2, θ₁=1.32, θ₃=-1.89, loss = 0.09)



Budapest Daily Mean Temperatures with Sine Wave Gradient Descent Fits
Temperature (°C)
Transformed Index (0.01 to 3.65)
2
1
0
−1
−2
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
1969 Data
1969 Fitted Sine (θ₀=-0.1, θ₁=1.4, θ₃=-1.85, loss = 0.088)
1994 Data
1994 Fitted Sine (θ₀=0.1, θ₁=1.3, θ₃=-1.83, loss = 0.085)
2019 Data
2019 Fitted Sine (θ₀=0.3, θ₁=1.2, θ₃=-1.86, loss = 0.081)


Oslo Daily Mean Temperatures with Sine Wave Gradient Descent Fits
Temperature (°C)
Transformed Index (0.01 to 3.65)
2
1
0
−1
−2
−3
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
1969 Data
1969 Fitted Sine (θ₀=-0.1, θ₁=1.5, θ₃=-1.93, loss = 0.091)
1994 Data
1994 Fitted Sine (θ₀=-0.0, θ₁=1.3, θ₃=-1.89, loss = 0.084)
2019 Data
2019 Fitted Sine (θ₀=0.1, θ₁=1.2, θ₃=-1.79, loss = 0.071)

Gradient Descent Analysis

- All three cities show increasing θ_0 values over 50 years, indicating higher baseline temperatures. Budapest shows the most dramatic shift (0.37 increase from 1969-2019), while Oslo exhibits the most rapid recent warming acceleration (0.13 gain between 1994-2019 vs 0.06 from 1969-1994).

- Madrid shows **increasing seasonal amplitudes** (1.21 $\rightarrow$1.32), while Budapest shows **moderate compression** (1.40 $\rightarrow$ 1.24), and Oslo **extreme compression** (1.54 $\rightarrow$ 1.18).

- θ_3 represents the phase shift and determines the timing of the wave's peak. The model for Oslo suggests that Summer peaked a full **14 days earlier** there in 2019 than in 1969. Madrid's seasons may also be moving; its 2019 phase was shifted **4 days earlier** than in 1969. Budapest showed **no change** in trend.

- J(θ) reveals striking differences in climate stability. Budapest (0.088 $\rightarrow$ 0.081) and Oslo (0.091 $\rightarrow$ 0.071) show **improving predictability** while Madrid shows initial **decreasing predictability** (0.080 $\rightarrow$ 0.104), followed by stabilization.

- These results provide **strong evidence** for the Seasonal Amplitude Compression Hypothesis and shows that it is **geographically dependent**. The combination of θ_1 amplitudes and J(θ) values reveals that Northern cities are becoming more predictable with less defined seasons, while Southern cities show increases in both seasonality and volatility.

Supervised Learning Algorithms

What is Supervised Machine Learning? Supervised learning trains algorithms on labeled data to make predictions about new, unseen examples. The algorithm learns patterns by studying input features (weather variables) paired with known outcomes (pleasant/unpleasant weather classifications).

Our Classification Problem:

Target Variable: Binary classification of daily weather as "pleasant" or "unpleasant"
Goal: Build models that can accurately classify weather conditions as pleasant or not

Why Supervised Learning Fits This Problem:

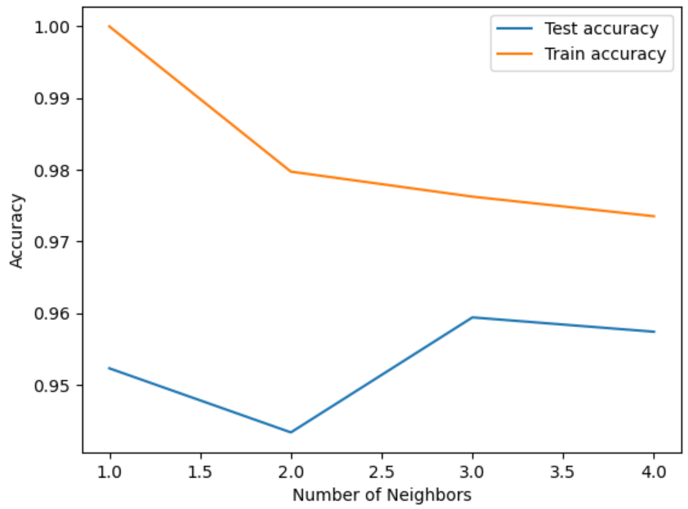
- We have decades of weather data with established pleasant/unpleasant classifications, providing the labeled examples needed for supervised learning.
- Multiple quantitative weather variables create a multi-dimensional feature space where algorithms can identify complex patterns distinguishing pleasant from unpleasant conditions.
- We can compare different supervised approaches (k-NN, Decision Trees, Artificial Neural Networks) to find the most effective method for weather-based classification in various geographic contexts.

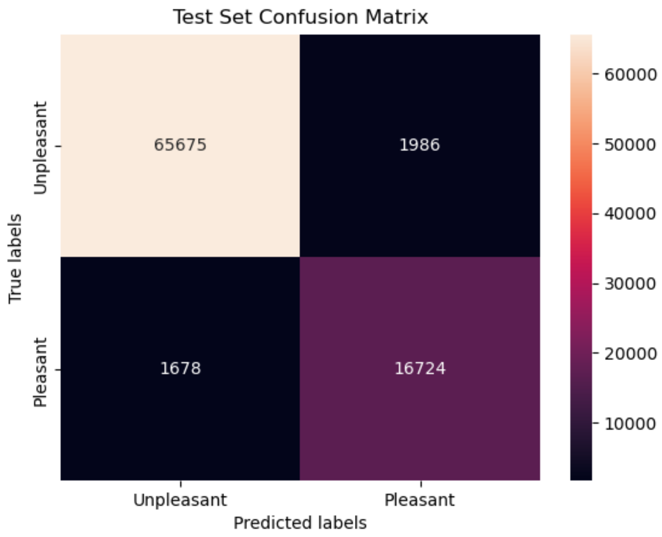
k-Nearest Neighbors

- Classifies data based on majority vote of nearest data points - number of neighbors (*k*) is set by the user.
- Higher *k* reduces overfitting to noise but may miss local patterns.

Optimal Neighbors : *k* = 3
Accuracy: 95.94%

Class	Precision	Recall	F1-Score
Pleasant	0.89	0.91	0.90
Unpleasant	0.98	0.97	0.97





The model achieved remarkable accuracy in classifying our data. As we would expect for unbalanced data, it performed better on the majority class (unpleasant days) than the minority class (pleasant days).

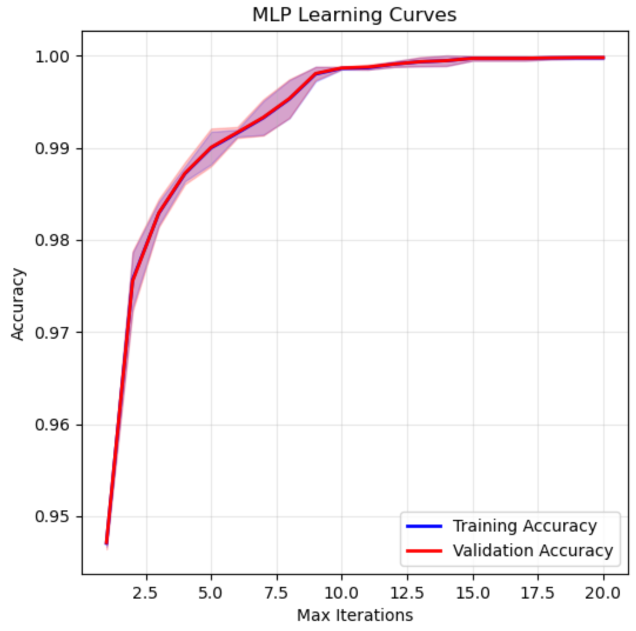
For both classes, precision and recall were about equal meaning the model is not biased towards misclassifying data as falsely negative or falsely positive.

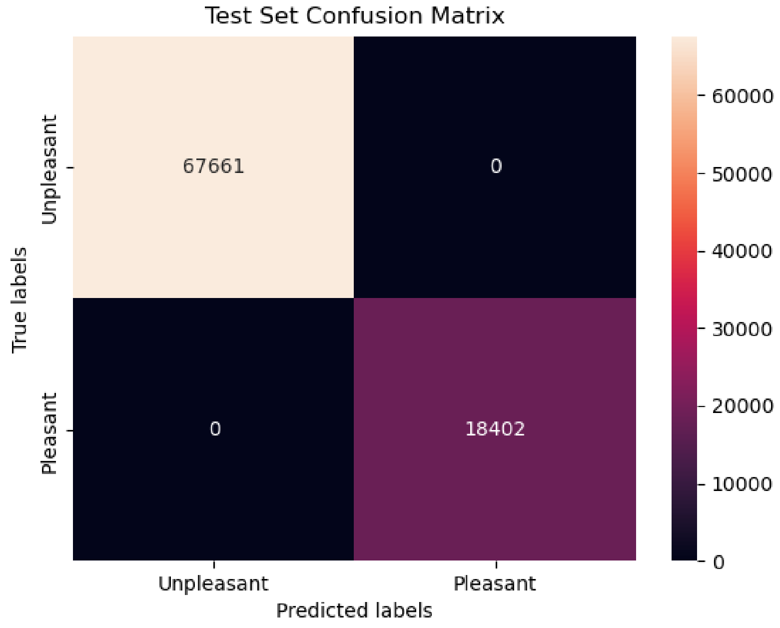
Multi-Layer Perceptron

- Artificial neural network with input, hidden, and output layers where each neuron applies weights to inputs, sums them, and passes the result through an activation function.
- Learns by using backpropagation to calculate gradients and adjust weights during training to minimize prediction error on labeled data.

Hidden Layers: 3
Neurons per Layer: [10, 5, 5]
Iterations: 50

Class	Precision	Recall	F1-Score
Pleasant	1.00	1.00	1.00
Unpleasant	1.00	1.00	1.00





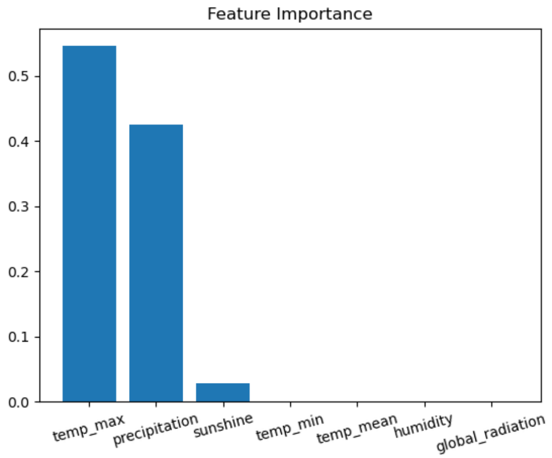
Although the model achieved perfect classification for our data, it did so at a very high computational cost and took a long time to converge despite the simplicity of this classification problem. For these reasons, I would not recommend this approach, especially as simpler models had nearly equal - or better - results.

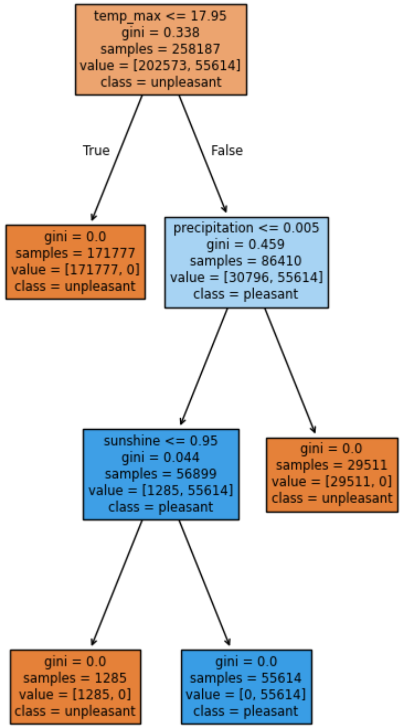
Decision Trees

- Partitions data by selecting optimal split points that minimize inaccurate classifications, creating a series of binary decisions until stopping criteria are met.
- Similar to human decision-making process, therefore usually more interpretable than other methods.

Tree Depth: 3 levels
Total Nodes: 4 leaves
Most Important Feature:
Maximum Temperature

Class	Precision	Recall	F1-Score
Pleasant	1.00	1.00	1.00
Unpleasant	1.00	1.00	1.00





Not only did the model achieve perfect classification, running quickly and efficiently on the dataset, but it also revealed the underlying logic behind the Pleasant/Unpleasant classification system:

Pleasant days are defined as those with maximum temperatures of at least 18° C (64.4° F), no precipitation, and at least 1 hour of sunshine.

Wrap-Up

A sincere and heartfelt thank you for your time and attention.

For questions, suggestions, or job offers ☺ I can be reached at:
TonyMarziale@gmail.com

More work available at:
Github.com/tonylevine
Tableau.com/tonymarz2408