COLUMBIA UNIVERSITY

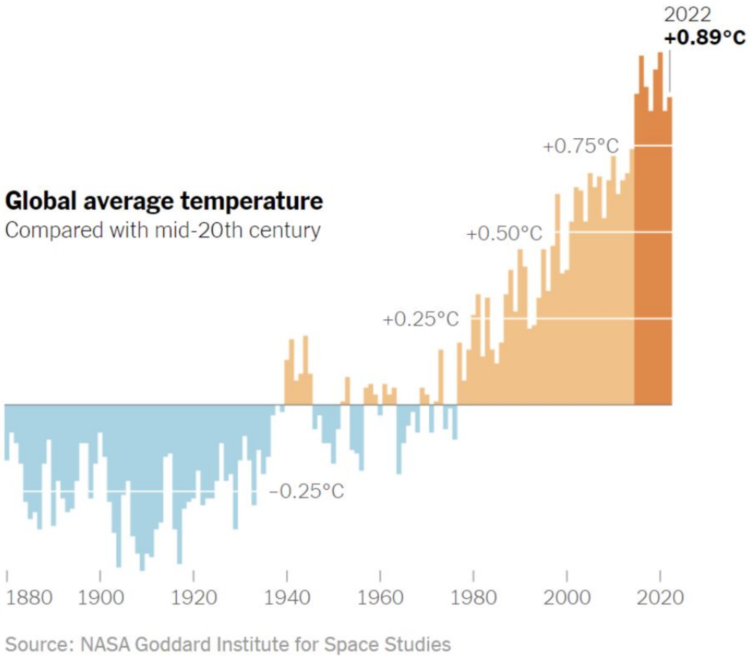
Climate Change Prediction with Time-Series Modeling

By: Adeel Arif, Boping Song, Malaikah Khan

Problem Statement & Motivation

What real-world problem does this project address?

- Global temperatures **continue to rise** due to human-driven climate change.[6]
- Impacts are uneven and regions like **Africa face higher vulnerability** due to exposure + low adaptive capacity.
- Temperature shifts affect heat stress, agriculture, water access, and public health.



Objectives & Research Questions

Research Questions:

1. To what extent can an LSTM-based model reliably forecast future temperatures at a global scale?

2. How does the performance of a global LSTM-based model compare to continent-specific LSTM models across different regions (e.g., Europe, Asia, Africa)?

Clear, measurable goals:

- Evaluate the accuracy of global temperature forecasts using LSTM (RMSE, MAE, directional accuracy).
- Compare continent-specific model performance to the global model across six continents.

Data & Preprocessing

- **Berkeley Earth Global Land Temperature Dataset**[1]
    - Contained earth surface temperature data from **1743-2013**
    - 532.83 MBs

- Our dataset contained **577,462 rows, 4 columns**
    - **243 total countries + islands/territories**

	dt	AverageTemperature	AverageTemperatureUncertainty	Country
577457	2013-05-01	19.059	1.022	Zimbabwe
577458	2013-06-01	17.613	0.473	Zimbabwe
577459	2013-07-01	17.000	0.453	Zimbabwe
577460	2013-08-01	19.759	0.717	Zimbabwe
577461	2013-09-01	NaN	NaN	Zimbabwe

Data Cleaning

Step 1: Country-Continent Mapping

pycountry-convert 0.7.2

`pip install pycountry-convert`

Manual Mapping

- We used the **pycountry-convert package** to map country to continent
- With the regions that the package did not have, we **manually mapped** them
    - (e.g. "Timor Leste": "Asia", "Virgin Islands": "North America, "Baker Island": "Oceania")

Step 2: Remove Duplicated Dates

- For duplicated dates, we took an **average of the temperature value**

Step 3: Remove Any Large Gaps With Dates

- **Removed large temporal gaps** between consecutive observations
- **Excluded all gaps >3 months** to ensure consistent spacing for reliable model training.

Step 4: Interpolation

- Filled missing temperature values using **interpolation**
- Applied Min-Max normalization to scale all temperature values.

Final Cleaned Dataset for Modeling

Continent	Count
Europe	157,688
Asia	114,516
Africa	105,230
North America	90,135
South America	29,509
Oceania	32,892

Total Count: 533,022

44,440 data points were removed

- **Antarctica** was removed given low count (3052)
- **Oceania** is a geographical region that **includes the continent of Australia** plus many Pacific nations

Architecture/Model Overview

- We implemented **three neural network architectures** and trained each in two settings:
  - **Global model** using all countries combined
  - **Six continent-specific models** (Europe, Asia, Africa, North America, Oceania, South America)

Architecture	Description	# Global Models	# Continent Models	Total
Baseline Attention LSTM	Single Feature Input (Average Temperature)	1	6	7
Enhanced Attention LSTM	Multi-Feature Inputs (Seasonality, Lags, Rolling Average)	1	6	7
Attention Bi-LSTM	Bidirectional + multi-feature	1	6	7

Columbia University

Main Output: Predicted temperature for the next month (t + 1)

Total: 21 models

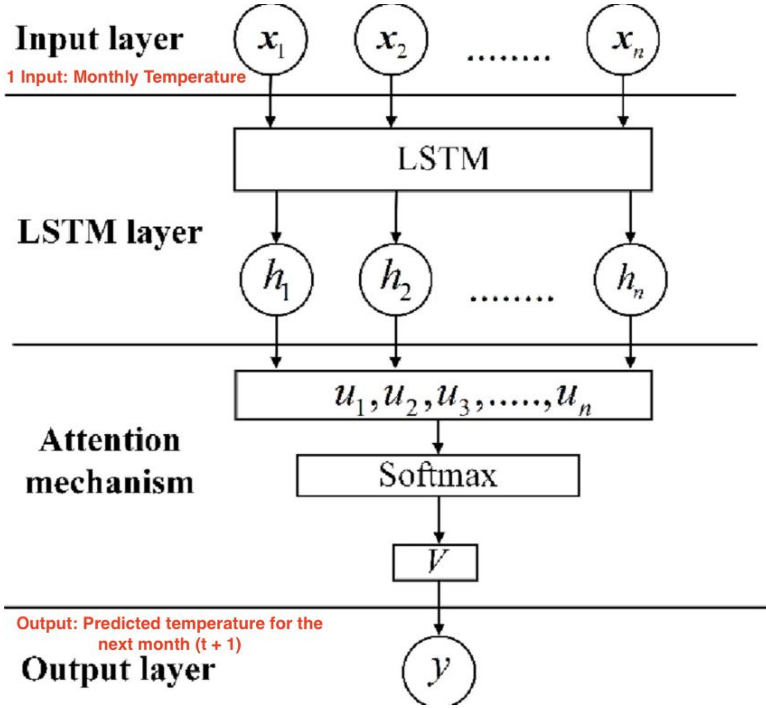
Model 1 – LSTM + Attention Basic

Architecture Hyperparameters

- **Lookback window: 36 months** — captures multi-year climate patterns
- **LSTM hidden size: 64** — enough capacity to learn trends without overfitting
- **LSTM layers: 1** — keeps model simple and stable for time-series forecasting

Training Hyperparameters

- **Learning rate: 0.001** — adjusts the weights in small, stable steps
- **Loss function: MSELoss** — standard regression loss for forecasting
- **Batch size: 64** — processes 64 samples per step, giving stable gradients
- **Epochs: 30** — sufficient for convergence without overfitting



Model 2: Advanced LSTM + Attention (6 Features)

We used the same hyperparameters as Model 1, but added **6 new inputs**:

1. Average Temperature (same as model 1)

2. Sine transformation (month_sin) — captures smooth cyclical seasonality

3. Cosine transformation (month_cos) — complements sine to represent full yearly cycle

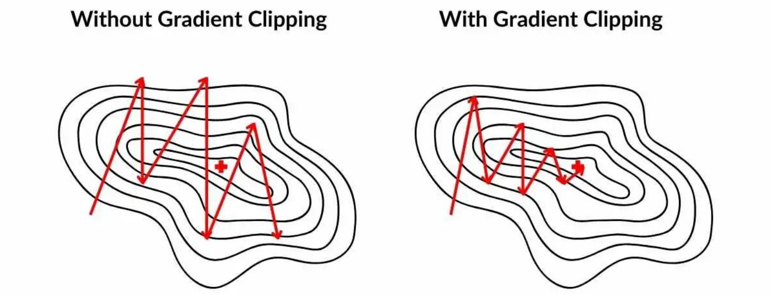
4. Lagging average (lag12) — temperature from 12 months ago → yearly pattern memory

5. lag24 — temperature from 24 months ago → multi-year trend awareness

6. Rolling average (roll3) — 3-month moving average → recent short-term trend smoothing

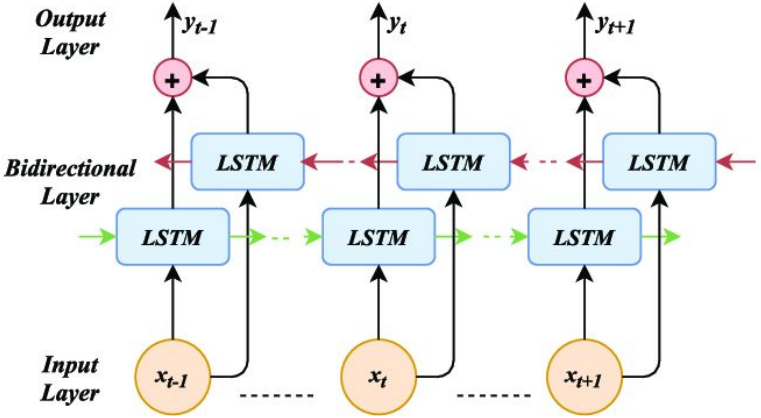
These extra inputs made the model **higher-dimensional** and more sensitive during training.

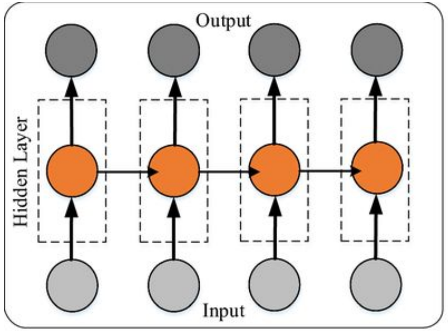
- This led the LSTM to occasionally produce **very large gradients**.
- Those large updates caused **training instability** and even **NaN losses** early on.
- **Gradient clipping** fixed this by **preventing exploding gradients** (when gradients become too large and cause unstable weight updates during training)[3]



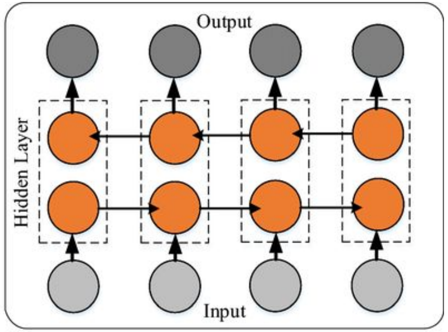
Model 3: Bidirectional-LSTM

We used the **same inputs, hyperparameters, and training setup as Model 2**





LSTM architecture
(Hochreiter and Schmidhuber, 1997)



Bi-LSTM architecture
(Graves and Schmidhuber, 2005)

- LSTM remembers information through a hidden state that moves forward in time **(only carries memory from past)**.[2]
- BiLSTM uses **two hidden states (one forward and one backward)** → so it learns patterns from both directions.
  - This gives the model **more context** and often **better accuracy** for cyclical climate data.

Model 3: Bi-LSTM Continued

Feature	LSTM	Bi-LSTM
Contextual Information	Has access only to past context when predicting the next state.	Uses both forward and backward context for a more complete view of the sequence.
Performance	Sufficient for tasks where past context is adequate.	Typically provides better performance in tasks where full sequence context (both past and future) is important.
Training Time	Faster due to single-direction processing.	Slower due to the processing of sequences in both directions.

We hypothesize that:

1) **Advanced LSTM will outperform basic LSTM**
2) **Bi-LSTM will outperform LSTM models (basic & advanced)**

Experiments & Results

Part 1 - Global Metrics

RMSE (Root Mean Squared Error) : Measures the average size of errors, treating all deviations equally

MAE (Mean Absolute Error): Measures of the average magnitude of errors between paired observations expressing the same phenomenon

Directional accuracy (DA): Measures how often a forecast correctly predicts the direction of change (up or down) of a variable, rather than its exact value

Model	RMSE	MAE	Directional Accuracy
Baseline Attention LSTM	0.0258	0.0155	0.842
Enhanced Attention LSTM	**0.0204**	**0.0133**	**0.875**
Attention Bi-LSTM	0.0211	0.0155	0.870

Among the global model, the Enhanced Attention LSTM achieves the lowest RMSE and MAE and the Directional Accuracy

Part 2 - Continent-Level Metrics

- Attention Bi-LSTM model shows substantial improvements in RMSE and MAE across all continents
- Directional accuracy remains relatively similar across models with only variations by region

MAE:

Continent	Baseline Attention LSTM	Enhanced Attention LSTM	Attention Bi-LSTM
Europe	3.693739	2.478485	**0.033690**
Asia	1.249834	1.138645	**0.023536**
Africa	0.839264	0.723852	**0.024564**
North America	0.735268	1.732746	**0.019073**
Oceania	0.489724	0.699873	**0.038547**
South America	0.722365	2.226599	**0.071998**

RMSE:

Continent	Baseline Attention LSTM	Enhanced Attention LSTM	Attention Bi-LSTM
Europe	4.391178	3.02192	**0.041422**
Asia	2.036672	1.65024	**0.031520**
Africa	1.204461	1.072081	**0.034632**
North America	1.41378	3.075578	**0.028855**
Oceania	0.6802	0.885159	**0.049648**
South America	1.082204	2.609257	**0.085169**

DA:

Continent	Baseline Attention LSTM	Enhanced Attention LSTM	Attention Bi-LSTM
Europe	0.869093	0.881826	0.883119
Asia	0.885417	0.904542	0.881022
Africa	0.830061	0.851309	0.853076
North America	0.84163	0.847692	0.850125
Oceania	0.651606	0.702714	0.699769
South America	0.792405	0.694011	0.687718

Discussion & Analysis

Insights / Patterns

- **Seasonality improves prediction** — lag12/24 + sin/cos capture yearly climate cycles.
- **Higher DA in Northern Hemisphere** — more stable seasonal structure.
- **Regional variation matters** — continents show different climate dynamics.

Limitations / Biases

- **Data quality varies** — some continents have sparse or noisy historical records.
- **Hemispheric bias** — opposite seasons confuse global models.
- **Global model challenge** — mixing climates weakens model strength.

What Worked

- **BiLSTM + features** — consistently lowest RMSE/MAE across regions.
- **Continent-level models** — align better with local seasonal behavior.
- **Seasonal features** — lag + sin/cos improve model performance.

What Didn't

- **Global-only model** — struggles with mixed or opposite-season regions.
- **Oceania prediction** — reversed seasons + limited data reduce DA.

Ethical Considerations

Bias & Fairness:

We examined the data structure across regions, since each continent has different time ranges recorded.

Transparency:

All modeling steps were documented, including sequence window (36 months), scaling, month sin/cos encoding, and area embeddings.

Privacy:

Data only come from public climate datasets (Berkeley Earth) and contain no individual-level or sensitive information.

Societal Impact:

Benefits: Clearer climate insights for researchers and policymakers

Risks: Over reliance on model outputs without domain expertise could distort public understanding of climate risk

Conclusion & Future Work

Key results/Takeaways

- Enhanced Attention LSTM achieved the lowest global RMSE and MAE and achieved the highest directional accuracy.
- At the continent level, Bi-LSTM dramatically reduced RMSE and MAE (for example Europe: 4.39 → 0.04) showing strong regional improvement.
- Directional accuracy remained high across all models with small variations depending on regional seasonality.[7]
- Bi-LSTM consistently outperformed other models on continent-specific data, showing it captures regional patterns more effectively.

How this project extends beyond class

- Addresses real-world climate challenges such as warming trends, extreme weather risk, and regional climate variability.
- Provides a transferable framework for environmental agencies, researchers, and policymakers to model continent-level climate behavior.
- Can lead to future research on anomaly detection, long-term climate forecasting, and multi-variable climate modeling.

Future work

- Add deeper architectures (more LSTM, Bi-LSTM, or attention layers) and perform broader hyperparameter tuning.
- Incorporate additional climate variables such as CO_2, precipitation, sea-level data, and emissions
- Test transformer-based time series models to compare performance and improve long-range climate forecasting

References

1. "Climate Change: Earth Surface Temperature Data." *Kaggle*, www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data.

2. GeeksforGeeks. "Difference between a Bidirectional LSTM and an LSTM." *GeeksforGeeks*, 26 Aug. 2024, www.geeksforgeeks.org/deep-learning/difference-between-a-bidirectional-lstm-and-an-lstm/.

3. GeeksforGeeks. "Understanding Gradient Clipping." *GeeksforGeeks*, 18 Jan. 2024, www.geeksforgeeks.org/deep-learning/understanding-gradient-clipping/.

4. GeeksforGeeks. "Understanding the Moving Average (MA) in Time Series Data." *GeeksforGeeks*, 6 Feb. 2024, www.geeksforgeeks.org/machine-learning/understanding-the-moving-average-ma-in-time-series-data/.

5. Inc, TUNE. "Pycountry-Convert: Extension of Python Package Pycountry Providing Conversion Functions." *PyPI*, pypi.org/project/pycountry-convert/.

6. IPCC. "Sixth Assessment Report — IPCC." *IPCC*, 2023, www.ipcc.ch/assessment-report/ar6/.

7. Spider Financial, info@numxl.com. "MDA - Mean Directional Accuracy." *Help Center*, 14 June 2019, support.numxl.com/hc/en-us/articles/360029220972-MDA-Mean-Directional-Accuracy.