Estimating Near-Surface Air Temperature from Satellite Land Surface Temperature Using Machine Learning

Felix Ozpaker
Mentor: Yuling Liu

cisess
Cooperative Institute for Satellite Earth System Studies

Objectives

Develop a machine learning framework to estimate daily maximum and minimum near-surface air temperature by integrating satellite LST with GFS Meteorological Fields.

Data & Methodology

Built the training dataset combining:

- Global Historical Climatology Network (GHCNd)
- Global Forecast System (GFS)
- L3 Daily Global VIIRS LST

Developed an XGBoost machine learning model trained on 2021 - 2024 data and tested on 2025 data.

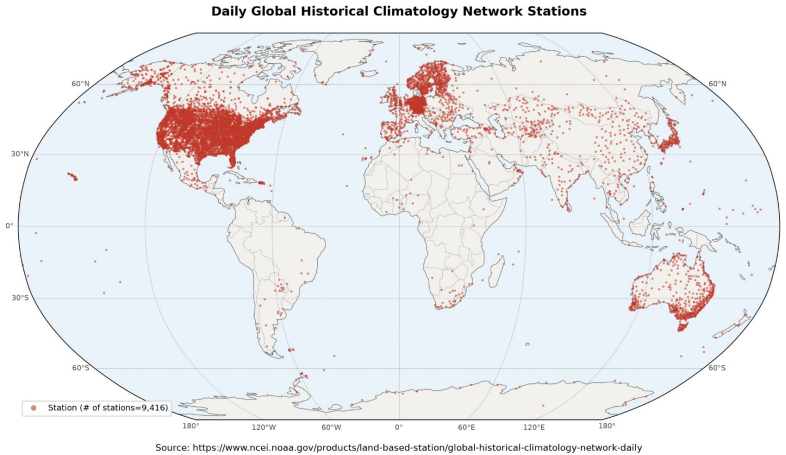


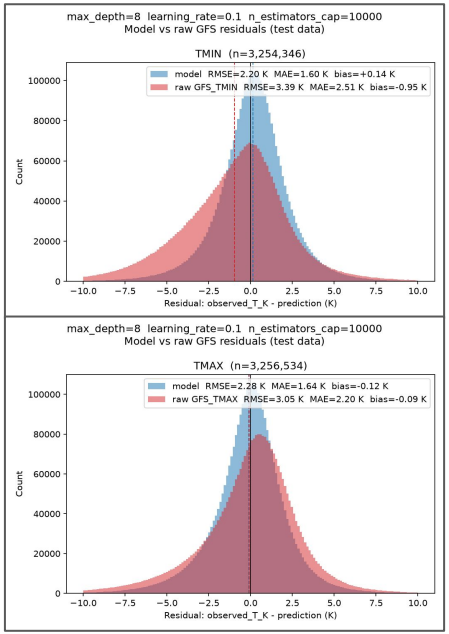
Figure: Geographic distribution of the 9,416 GHCN-Daily stations used as ground-truth observations for model training

EARTH SYSTEM SCIENCE INTERDISCIPLINARY CENTER

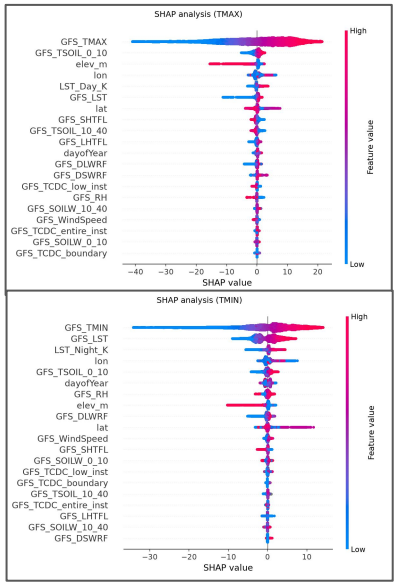
cisess
Cooperative Institute for Satellite Earth System Studies

Results

- Using the independent 2025 test dataset, the models achieved RMSEs of 2.28 K and 2.20 K for daily maximum and minimum air temperature, respectively.
- The models achieved R^2 values of 0.965 and 0.961 and MAEs of 1.64 K and 1.60 K for daily maximum and minimum air temperature, respectively.
- The minimum temperature model reduced the GFS baseline's bias from -0.95 K to -0.14 K.



Figures: Residual plot distribution for the TMAX (top) and TMIN (bottom) models



Figures: SHAP summary plots showing feature impact and ranking for the TMAX (top) and TMIN (bottom) model predictions

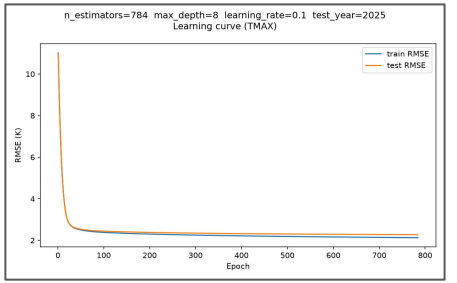


Figure: Train/test RMSE (K) vs. epoch, showing convergence and the point where early stopping parameters stop training

TMAX	Early Stopping Rounds: 100	Early Stopping Tolerance: 0.01								
max_depth	learning_rate	n_estimators_cap	best_iteration	train_rmse	test_rmse	mae	test_r2	test_bias	overfit_percent	delta_vs_gfs
8	0.1	10000	683.00	2.14	2.28	1.64	0.965	0.12	6.95%	-0.77
7	0.1	10000	829.00	2.19	2.30	1.66	0.964	0.11	4.80%	-0.75
8	0.07	10000	665.00	2.19	2.30	1.66	0.964	0.12	4.99%	-0.74
8	0.05	10000	810.00	2.21	2.32	1.67	0.963	0.12	4.60%	-0.73
7	0.07	10000	766.00	2.24	2.33	1.68	0.963	0.12	3.91%	-0.72
6	0.1	10000	781.00	2.27	2.35	1.70	0.962	0.11	3.52%	-0.70
8	0.03	10000	848.00	2.27	2.35	1.70	0.962	0.12	3.68%	-0.69
7	0.05	10000	827.00	2.27	2.36	1.70	0.962	0.12	3.51%	-0.69
6	0.07	10000	851.00	2.29	2.37	1.71	0.962	0.11	3.16%	-0.68
7	0.03	10000	891.00	2.32	2.39	1.73	0.961	0.12	2.91%	-0.66
8	0.02	10000	797.00	2.33	2.40	1.74	0.961	0.12	2.97%	-0.65
6	0.05	10000	748.00	2.35	2.41	1.75	0.960	0.12	2.57%	-0.64
7	0.02	10000	934.00	2.36	2.42	1.76	0.960	0.12	2.49%	-0.63
8	0.01	10000	942.00	2.39	2.44	1.77	0.959	0.13	2.29%	-0.61
6	0.03	10000	778.00	2.40	2.45	1.78	0.959	0.12	2.08%	-0.60
6	0.02	10000	896.00	2.43	2.47	1.80	0.958	0.12	1.80%	-0.58
7	0.01	10000	893	2.45	2.49	1.81	0.958	0.13	1.67%	-0.56
6	0.01	10000	1064	2.49	2.52	1.83	0.957	0.13	1.32%	-0.53

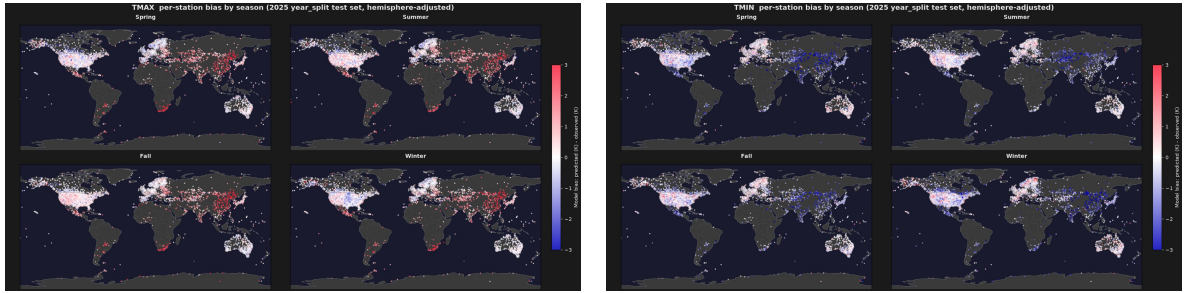
Figure: Parameter sweep over max_depth and learning rate, ranked by test RMSE

EARTH SYSTEM SCIENCE INTERDISCIPLINARY CENTER

cisess
Cooperative Institute for Satellite Earth System Studies

Analysis

- TMAX model overestimates across Asia, South Africa, and Mexico.
- TMIN model underestimates across Asia, South Africa, and Mexico.
- Spatial bias is heavily correlated with density of ground stations.
- Both models have various spatial biases dependent on season (e.g. TMAX underestimates the Midwest in the Winter and overestimates in the Fall).
- At the station level, both models perform best during the summer.



Figures: Station-level bias by season, hemisphere-adjusted (2025 test set) for TMAX (left) and TMIN (right)

Model bias by 15° latitude zone (2025 year_split test set)

TMAX

Latitude zone	Model bias (K)	Number of observations	Number of stations
75° to 90°	+0.23	3,757	11
60° to 75°	-0.08	213,844	606
45° to 60°	+0.07	803,192	2,287
30° to 45°	+0.11	1,832,632	5,268
15° to 30°	+0.53	149,653	451
0° to 15°	+0.76	22,697	66
-15° to 0°	-0.03	12,757	38
-30° to -15°	+0.19	69,820	214
-45° to -30°	+0.09	134,712	398
-60° to -45°	+0.73	3,047	9
-75° to -60°	+1.14	8,710	29
-90° to -75°	+1.44	1,713	5

TMIN

Latitude zone	Model bias (K)	Number of observations	Number of stations
75° to 90°	-0.45	3,756	11
60° to 75°	+0.06	213,806	606
45° to 60°	-0.17	802,031	2,287
30° to 45°	-0.11	1,830,973	5,267
15° to 30°	-0.51	150,280	452
0° to 15°	-0.68	22,751	66
-15° to 0°	-0.05	12,755	38
-30° to -15°	-0.10	69,868	214
-45° to -30°	-0.10	134,643	398
-60° to -45°	-0.65	3,047	9
-75° to -60°	-0.97	8,724	29
-90° to -75°	-1.20	1,712	5

Figure: Mean model bias (predicted K - observed K) by 15° latitude zones for TMAX (left) and TMIN (right) on 2025 test data

EARTH SYSTEM SCIENCE INTERDISCIPLINARY CENTER