Predicting the Urban Heat Island Effect on San Diego's Athletic Fields Using Machine Learning

Anna Wang and Leo Meng
Nasa SEES Urban Heat Island Effect Team
Mentored by Dr. Jiang Yitong, Dr. Kevin Czajkowski, Dr. Parisa Derakhshesh

Background

The **urban heat island effect (UHIE)** describes human-caused discrepancies in the temperature of urban and rural areas. Factors: material type (absorption, emittance), surrounding vegetation.

San Diego: diverse microclimates, 8th largest city in U.S. by population

Methodology

Selected for consideration were **25 outdoors San Diego athletic fields** used as primary competition venue for a National Collegiate Athletic Association collegiate level team or professional team. Each field was assigned a 300m surrounding buffer to compare the variations between the field and its surroundings.

Data Extraction:
USGS images taken by Landsat 8/9 and Sentinel 2 Level 2A were analyzed on ArcGIS. These images have <10% cloud cover and are 30m in resolution.

At each site, and for each predictive variable, the August values for each year of **2015-2026** were averaged.

Predictor variables extracted and averaged:

- **NDVI (Normalized Difference Vegetation Index),** the measure of vegetation cover
  - within the entire area of a site
- **Albedo,** the reflectivity of a surface
  - within the entire area of a site
- **NDMI (Normalized Difference Moisture Index),** the measure of soil water content
  - within the field itself **AND** its isolated buffer, to analyze how real grass and turf vary in moisture
- **NDBI (Normalized Difference Build-Up Index),** the measure of urban, above surface infrastructure
  - within the field itself **AND** its isolated buffer, to highlight the influence of surrounding infrastructure on any given field

And one **response variable: LST (Land Surface Temperature),** extracted at the fields and buffers to account for the cooling influence of the Pacific Ocean.

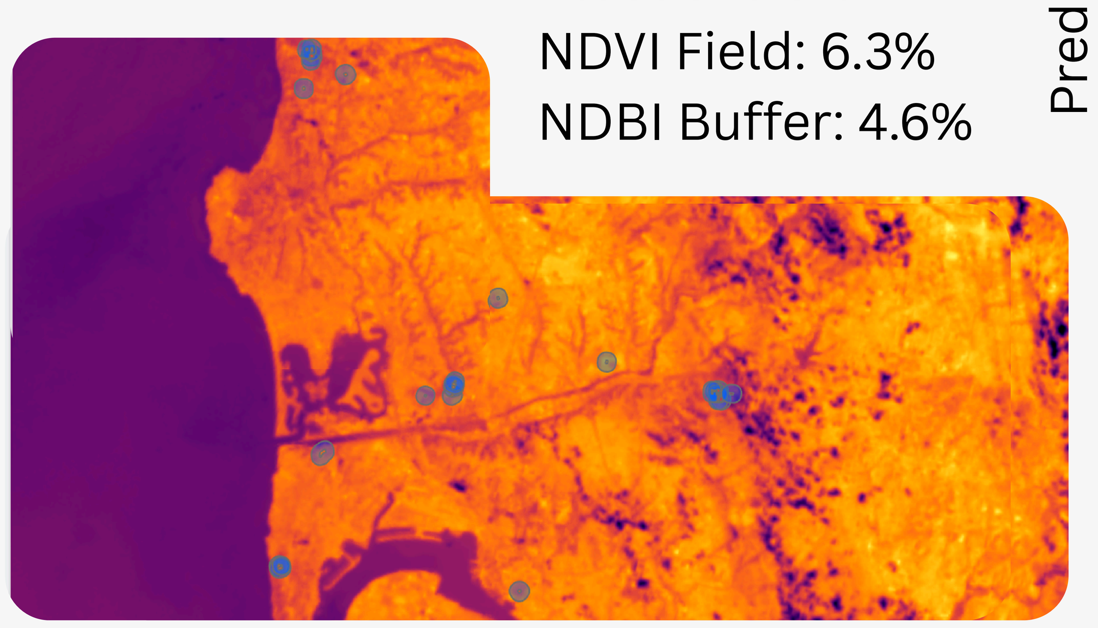
Results

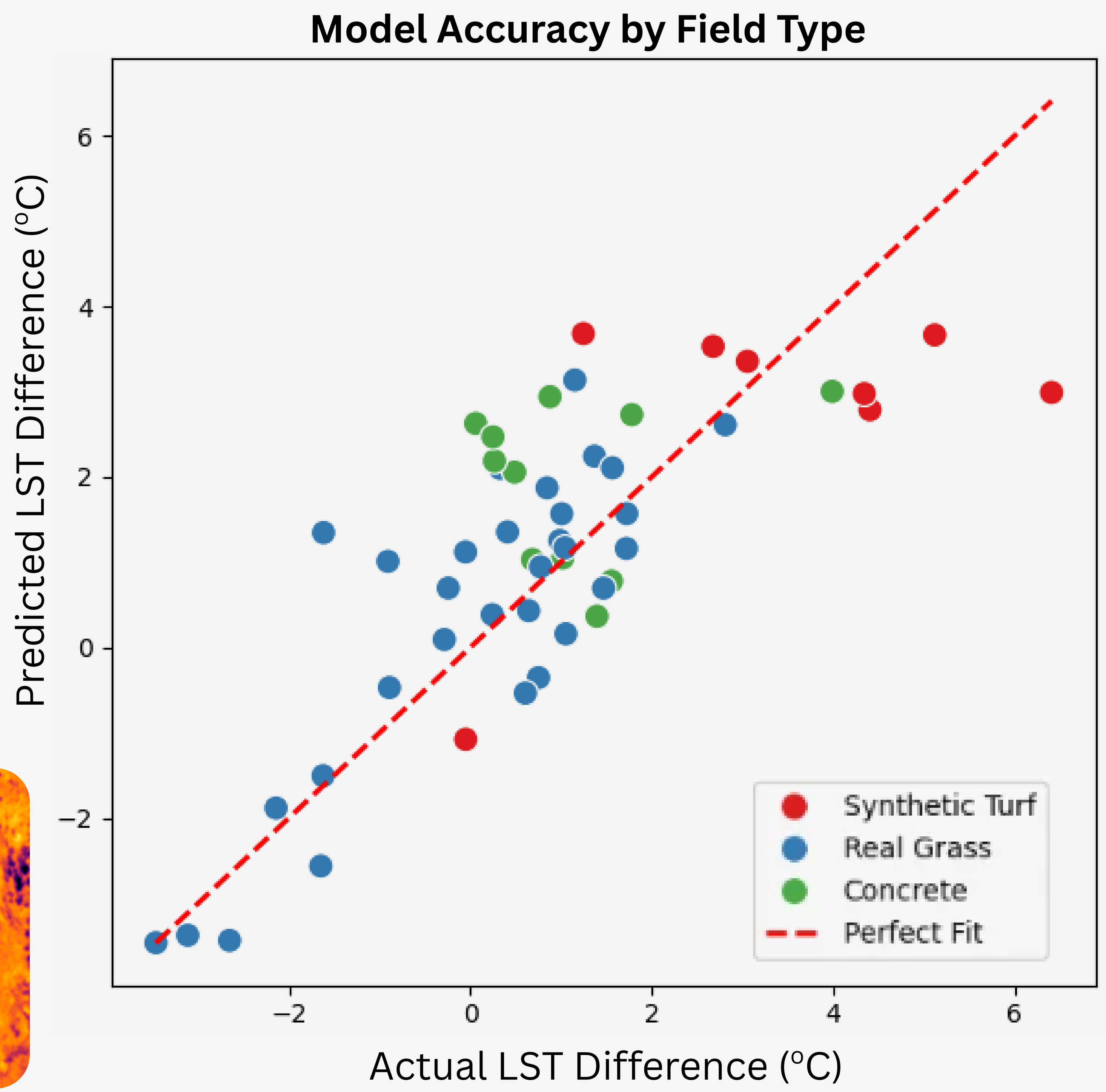
A Random Forest Model was trained using a **80/20 ratio.**

The model achieved an **R-squared Score** of **0.56** and **Average Error (RMSF): 1.29 °C Difference**

Error by field type:
Synthetic: 1.81 °C
Real Grass: 1.01 °C
Concrete: 1.53 °C

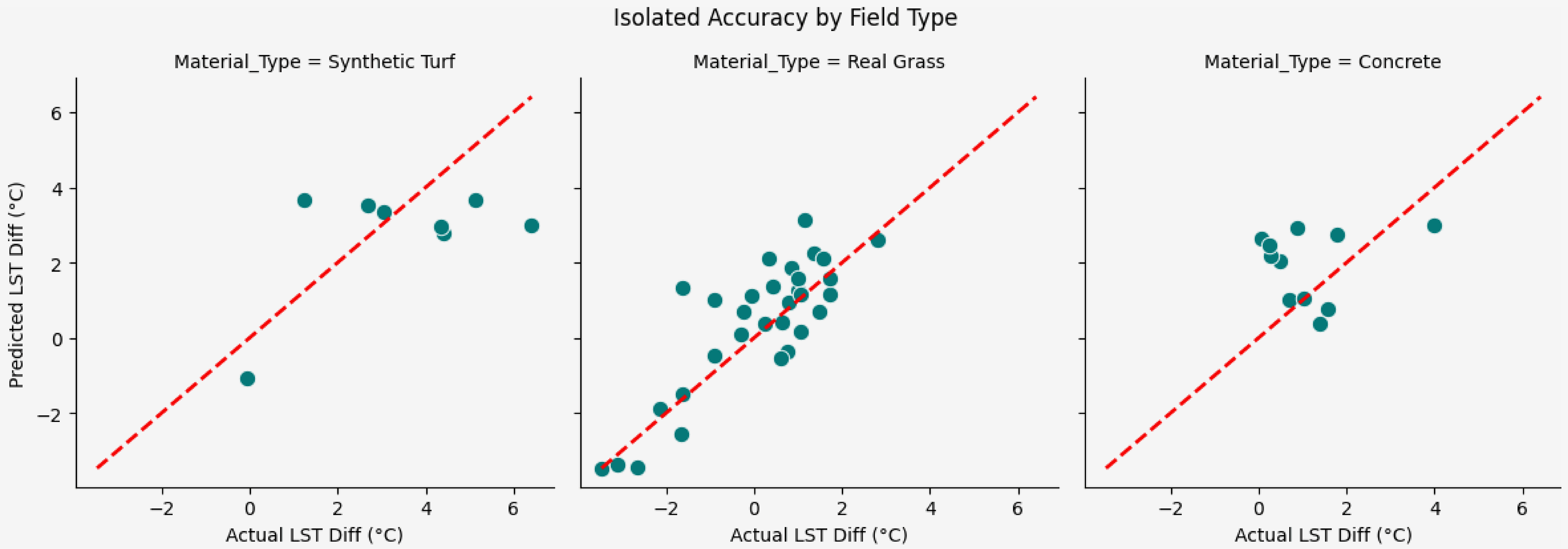
Predictive power:
NDMI Field: 57.6%
NDVI: 16.8%
NDBI Buffer: 9.1%
Albedo: 6.3%
NDVI Field: 6.3%
NDBI Buffer: 4.6%





How does the model fair at predicting different surface types?

The model was most accurate in predicting LST at real grass locations and least accurate in predicting LST at synthetic turf fields. Synthetic turf matches closely with the albedo and NDVI of real grass but contains less moisture, and thus, cannot regulate temperature as easily. Both grass and concrete raise the field's temperature relative to its surroundings, while real grass maintains the field temperature and can even decrease it.



Discussion

Field moisture content is the most informative predictor for LST difference, emphasizing that **evaporative cooling and latent heat** flux keep fields cool, not just the presence of greenery.

Despite the limitations of our data collection, our model achieved a robust R^2 score of 0.56. **Macro-level** satellite **remote sensing** is a highly viable and accessible tool for screening regional microclimate risks.

Future Work

Supplement with new data: The study was limited by the short timeframe of the SEES program. To improve the accuracy of our model, new parameters and new fields of a more even material distribution will be incorporated.

Expand the scope: The ultimate purpose of this project, as inspired by the World Cup, is to predict the UHIE and LST in athletic fields to prevent **athlete heat illnesses.**

Citations

Cardenas, Catherine; Aronow, Jacob; Beltran, Lorenzo. Predicting Urban Heat Island Growth in North Carolina's Research Triangle Using Machine Learning. NASA SEES UHIE 2025, presented at AGU Winter 2025.

Cheng, W., Spengler, J. O., & Brown, R. D. (2020). A Comprehensive Model for Estimating Heat Vulnerability of Young Athletes. International Journal of Environmental Research and Public Health, 17(17), 6156. https://doi.org/10.3390/ijerph17176156

Crowson, M., Williams, J., Sharma, J., & Pettorelli, N. (2024). Challenges for monitoring artificial turf expansion with satellite remote sensing. Environmental Monitoring and Assessment, 196(6), 580. https://doi.org/10.1007/s10661-024-12724-w

Mantas, V., Czajkowski, K., Xian, G., Jiang, Y., Namjooyan, F., Oluawfemi, O.. (2024). Characterizing temperature hotspots within Urban Heat Islands in the United States. Sports fields as a case-study. AGU Winter 2024.

Thoms, A., Brosan, J., Zidek, J., Sorochan, J. (2014). Models for Predicting Surface Temperatures on Synthetic Turf Playing Surfaces. International Sports Engineering Association Conference. https://www.researchgate.net/publication/275536365_Models_for_Predicting_Surface_Temperatures_on_Synthetic_Turf_Playing_Surfaces

Wald, A., & Demorest, S. (2022). Race to Beat the Heat: Climate Change Impacts Physical Activity. The Journal for Nurse Practitioners, 18(4), 388–394. https://doi.org/10.1016/j.nurpra.2021.12.005

Zhang, Y., Zhang, X., Yan, D., Wang, Y., Zhang, M., Li, Y., & Wu, L. (2025). Quantifying the cooling effect of green spaces on urban heat island effect. Geomatics, Natural Hazards and Risk, 17(1). https://doi.org/10.1080/19475705.2025.2605714