

Summary of Urban Heat Island (UHI) Index Modeling - Team Shuanggao

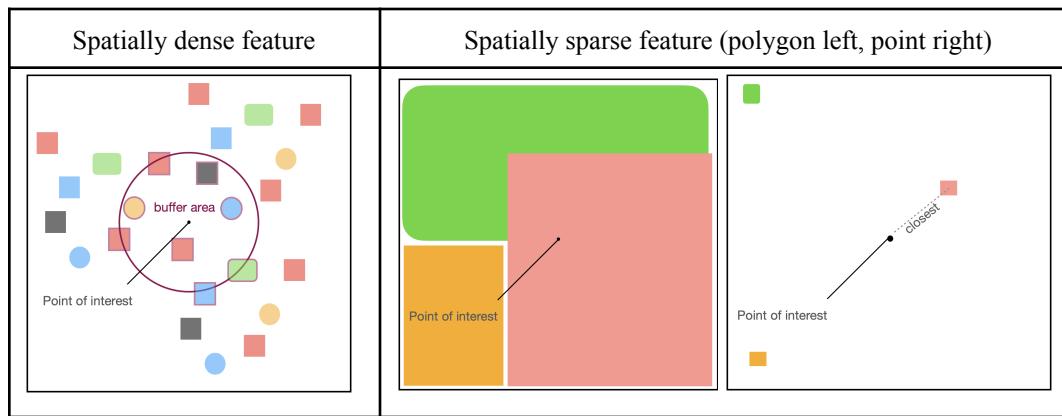
The model

We use ExtraTreesRegressor from scikit-learn for training.

Spatial feature deviation

We consider two types of feature construction based on the property of spatial features at hand

- **Spatially dense features (SD):** we construct buffer size within point of interest (lon, lat) and compute the statistics. For buffer-based methods, we may construct multiple buffer sizes to better understand the feature.
- **Spatially sparse feature (SP):** we perform data assignment, i.e., for each point of interest, we assign the feature based on the (1) which polygon it lies in (if the feature is polygon); (2) which point it is closest to (if the feature is a point).



Utilized features

We list the data used in the following table. For each data, we will specify data source, explanation, method (SD / SP), used statistics and other details if possible. (**More details can be found in the jupyter notebook!**)

POI: point of interest (lon,lat)

Building footprint	Explanation: Building footprint of the Bronx and Manhattan regions (the provided .kml file) Method: SD (polygon) Statistics: total area of the building within the buffer, counts of building overlapped with the buffer Other details: buffer size is in [0.06, 0.07, 0.08, 0.09, 0.1] for building count, [0.0005, 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1] for building area. All under projection 4326.
Weather	Explanation: The weather data from 2 weather stations (Bronx and Manhattan) Method: SP (point) Statistics: Weather characteristics (temperature, rainfall, humidity, etc.) in the region the POI belongs to. Other details: N/A
Building elevation	Explanation: The elevation of the building (not building height) Source: https://data.cityofnewyork.us/City-Government/Building-Elevation-and-Subgrade-BES-/bsin-59hv/about_data Method: SD (polygon)

	Statistics: Mean value of all building overlapped with the buffer Other details: buffer size is in [0.005, 0.01] under projection 4326
Building height	Explanation: The height of the building Source: https://data.cityofnewyork.us/City-Government/Building-Footprints/5zhs-2jue/about_data Method: SD (polygon) Statistics: Mean “Floor Area Ratio (FAR)” of all building overlapped with the buffer Other details: buffer size is in [500 meters] under projection 3857
Building energy	Explanation: The energy consumption of the building Source: https://energy.cusp.nyu.edu/#/ Method: SD (point) Statistics: Mean “Energy Use Intensity (EUI)” of all building overlapped with the buffer Other details: buffer size is in [0.005, 0.01, 0.15] under projection 4326
Population data	Explanation: Regional population data Source: https://www.kaggle.com/datasets/muonneutrino/new-york-city-census-data Method: SP (point) Statistics: Total population of the region the POI belongs to. Other details: N/A
Mapping Inequality	Explanation: Region type (by income as well as functionality) Source: https://dsl.richmond.edu/panorama/redlining/data/NY-Manhattan/ , https://dsl.richmond.edu/panorama/redlining/data/NY-Bronx Method: SP (polygon) Statistics: Income grade(1~10) / Area / Functionality(residential/commercial/industrial) of the region the POI belongs to. Other details: N/A
Commute data	Explanation: Commute type (car/walk/public/bike) of each region Source: https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/walking-driving-and-cycling/?id=2415#display=summary Method: SP (polygon) Statistics: #people who prefer public transportation in the region the POI belongs to. Other details: N/A
Air Quality	Explanation: Regional PM2.5 index Source: https://a816-dohbesp.nyc.gov/IndicatorPublic/data-explorer/air-quality/?id=2023#display=summary Method: SP (point) Statistics: PM2.5 value in the region the POI belongs to. Other details: N/A
Facilities	Explanation: Facilities geometry data Source: https://data.cityofnewyork.us/City-Government/Facilities-Database/ji82-xba5/about_data Method: SD (polygon) Statistics: number of “park and plaza” overlapped with the buffer Other details: buffer size is in [0.01] under projection 4326