Thermal Inequity Landscapes in Dane County: An Analysis of Heat/Vegetation vs. Income/Race

Guy Weibel, Mihir Heda

April 29, 2025

Abstract :

Urban Heat Islands and the cooling effects of green space raise environmental equity concerns. Our project investigated the distribution of Land Surface Temperature (LST) and vegetation (NDVI) relative to income and race across Dane County, WI census tracts. Using Landsat imagery (Summer 2023) processed via Google Earth Engine and US Census data, we performed correlation and spatial autocorrelation analyses. Results revealed significantly higher LST and lower NDVI in tracts with lower median incomes and higher minority percentages. NDVI exhibited strong positive spatial clustering (Moran's I $\approx$ 0.71, $p < 0.001$). These findings indicate potential thermal inequities, suggesting geographically uneven distribution of environmental burdens and amenities within the county.

Project Proposal

Motivation & Topics of Interest

Urban areas often experience elevated temperatures, impacting health and quality of life. Green spaces (or vegetation) mitigate this effect via shading and evapotranspiration. We investigated whether the benefits of green spaces and the burden of excess heat might be unevenly distributed between different socioeconomic or racial groups within Dane County, Wisconsin. We aim to explore potential correlations between these environmental factors (Land Surface Temperature - LST, Normalized Difference Vegetation Index - NDVI) and key demographic indicators (median household income, percentage of minority population) and examine the spatial clustering of green space.

Required Data

- **Analysis Unit Boundaries:** Dane County Census Tracts (Obtained from US Census TIGER/Line).
- **Land Surface Temperature (LST):** Derived from Google Earth Engine (GEE) LANDSAT/LC08/C02/T1_L2 and LANDSAT/LC09/C02/T1_L2 collections using the ST_B10 band (Thermal Infrared Sensor), averaged for Summer 2023 (June 1st - August 31st).
- **Vegetation Greenness (NDVI):** Derived from GEE LANDSAT/LC08/C02/T1_L2 and LANDSAT/LC09/C02/T1_L2 collections using the standard formula (NIR-Red) / (NIR+Red) via SR_B5 (Near Infrared) and SR_B4 (Red) bands, averaged for Summer 2023.
- **Demographic Data:** Median Household Income (Table B19013) and Race counts (Table DP05, used to calculate percentage minority population) from data.census.gov (American Community Survey - ACS 5-Year Estimates 2018-2022) for Dane County census tracts.

Analysis Methods

1. **Acquire & Prepare Data:** Load Census Tract boundaries into a GeoDataFrame (using GeoPandas). Process Landsat imagery in GEE to calculate mean LST and mean NDVI per tract using zonal statistics. Load Median Household Income data and calculate the Percentage Minority Population per tract (using Pandas). Ensure consistent geographic identifiers (GEOID) for merging.
2. **Integrate Data:** Join the calculated mean LST, mean NDVI, Median Income, and Percentage Minority Population data to the Census Tract GeoDataFrame based on the common GEOID. Handle any missing data appropriately (e.g. due to cloud cover in satellite imagery).
3. **Visualize Distributions & Relationships:** Create choropleth maps to visualize the spatial distribution of LST, NDVI, income, and minority percentage across tracts. Generate scatter plots to visually examine pairwise relationships between these variables. (Note: Visualizations were left in the .ipynb and are not in this summary).
4. **Quantify Correlations:** Calculate Pearson correlation coefficients (r) to measure the strength and direction of linear associations between LST, NDVI, median income, and race percentage, considering only tracts with complete data.
5. **Analyze Spatial Clustering:** Compute Global Moran's I for mean NDVI to test for significant spatial autocorrelation (i.e., whether green space is clustered, dispersed, or randomly distributed across Dane County tracts). Assess the statistical significance (p-value) and strength of clustering.

Results and Interpretation

Spatial Distributions

Analysis revealed distinct spatial patterns. Lower median household incomes and higher percentages of minority populations are concentrated primarily within the central urban core of Madison and adjacent tracts. Higher incomes and lower minority percentages are generally found in the surrounding suburban and outlying areas (Maple Bluff, Middleton, Verona...). Mean Land Surface Temperature (LST) was significantly higher in the central urban areas, closely corresponding spatially with areas of lower income and higher minority population percentage, demonstrating a clear Urban Heat Island (UHI) effect. Conversely, Mean NDVI (indicating vegetation greenness) was lowest in these same central urban tracts and highest in the less developed, outlying areas. The visual overlap between high LST, low NDVI, lower income, and higher minority percentage tracts is notable based on map analysis (detailed in the .ipynb submission).

Correlations

Quantitative analysis confirmed the observed relationships based on tracts with complete data:

- **LST & NDVI:** A very strong negative correlation (r = -0.90) was found, confirming the expected inverse relationship: areas with less vegetation experience significantly higher surface temperatures.
- **Environmental & Socioeconomic Factors:** Moderate correlations were observed:
  - Higher LST is associated with lower Median Income (r = -0.54).
  - Higher LST is associated with higher Minority Percentage (r = 0.33).
  - Higher NDVI (more green space) is associated with higher Median Income (r = 0.65).
  - Higher NDVI is associated with lower Minority Percentage (r = -0.39).
- **Socioeconomic Factors:** A moderate negative correlation exists between Median Income and Minority Percentage (r = -0.51), indicating tracts with higher minority populations tend to have lower median incomes.

These correlations suggest systematic relationships between environmental conditions and socio-demographic characteristics at the census tract level.

Spatial Clustering

The analysis of spatial autocorrelation for Mean NDVI yielded a Global Moran's I statistic of **0.7098**. This result is statistically highly significant (p-value = 0.0010, Z-score = 13.2574), indicating **strong positive spatial autocorrelation**. This means that census tracts with high NDVI values tend to be clustered near other high-NDVI tracts, and low-NDVI tracts are clustered near other low-NDVI tracts. Green space is not randomly distributed across the county but is geographically clustered.

Interpretation and Discussion

The combined results strongly suggest the presence of **thermal inequity landscapes** within Dane County. The data indicates that census tracts characterized by lower median household incomes and higher percentages of minority residents are disproportionately exposed to higher summer surface temperatures and have less access to green space (lower NDVI) compared to wealthier, less diverse tracts. The strong negative correlation between LST and NDVI highlights the crucial role of vegetation in mitigating surface heat. The significant positive spatial clustering of NDVI suggests that the benefits of this cooling effect are not evenly distributed but are concentrated in specific geographic areas, which, based on the correlation results, tend to be the higher-income, lower-minority tracts. While correlation does not imply causation, these findings align with patterns observed in many US cities and provide quantitative evidence supporting the hypothesis that the burdens of urban heat and the benefits of green space are unevenly distributed along socioeconomic and racial lines in Dane County.

Limitations

- **Correlation vs. Causation:** This analysis identifies statistical associations, not cause. Other factors not measured (e.g., building density, impervious surface fraction, housing age, infrastructure quality) also influence LST and NDVI and may confound the observed relationships.
- **Scale:** Analysis was performed at the census tract level, which is subject to oversimplification of a space. Results might differ if we used different spatial units, or higher granularity, as aggregation may mask finer-scale variations within tracts.
- **Temporal Scope:** LST and NDVI represent average conditions during a single summer season (2023) and may not capture year-round variations, extreme heat event conditions, or long-term trends. Demographic data is based on 5-year estimates.
- **Data Gaps:** Cloud cover inherent in satellite imagery resulted in missing LST/NDVI data for some tracts, potentially biasing the analysis if missingness is not random.
- **LST Definition:** We used satellite-derived Land Surface Temperature, which differs from the air temperature typically reported in weather forecasts, although they are related.

Conclusion

Our analysis successfully demonstrated significant spatial patterns and statistical correlations between land surface temperature, vegetation greenness, median household income, and racial composition in Dane County census tracts. The findings indicate potential environmental inequities, where lower-income and higher-minority communities appear to experience greater heat exposure and possess less green space coverage. The strong spatial clustering of vegetation further emphasizes the uneven distribution of these environmental amenities. These results provide a valuable quantitative baseline for further investigation and could inform local planning and policy decisions aimed at mitigating the UHI effect, promoting equitable access to green infrastructure and its associated benefits.

Contributions:
Equal work between Guy and Mihir throughout the entire project.