NASA DEVELOP National Program
California - Ames

Spring 2024

San Joaquin Valley Health & Air Quality II

Assessing Urban Heat Island Distribution and its Intersections with Air Quality to Understand Converging Vulnerabilities

DEVELOP Technical Report

March 29th, 2024

Shea Rousseau (Project Lead)
Gabriela Colombo
Tyler C. Padua
Carlo Gomez

Advisors:
Dr. Juan Torres-Pérez, NASA Ames Research Center (Science Advisor)
Dr. Kenton Ross, NASA Langley Research Center (Science Advisor)
Dr. Morgan Gilmour, NASA Ames Research Center (Science Advisor)
Dr. Xia Cai, NASA Langley Research Center (Science Advisor)
Dr. Joshua Dimon, University of California Berkeley, Disasters Laboratory (Science Advisor)
Lisa Tanh, Esri (Science Advisor)

Previous Contributors:
Jonathan Szeto
Jasper Beardslee
Piper Christian
Bethany MacCarter
Alma Quintero

Lead:
Lauren Webster (ARC)

1. Abstract

The city of Stockton, California, located within the San Joaquin Valley (SJV), is a major hub for agricultural production and has endured the continuous threat to community health from nitrogen dioxide (NO_2) and increasing temperatures. The convergence of these issues occurs within historically segregated communities that are disproportionately facing health risks related to heat and air quality. Little Manila Rising (LMR), a social and environmental justice (EJ) advocacy non-profit, partnered with NASA DEVELOP for a second term project to evaluate county wide urban heat islands, sociodemographic vulnerability, landcover classification, and the convergence of these variables. We utilized Landsat 8 Thermal Infrared Sensor (TIRS) and Operational Land Imager (OLI) data to produce a land surface temperature (LST) and Normalized Difference Vegetation Index (NDVI) map. They added Centers for Disease Control (CDC) socioeconomic data from 2020 to identify which communities in Stockton were more susceptible to these environmental factors. Additionally, we used imagery from National Agriculture Imagery Program (NAIP) to create a landcover map differentiating developed infrastructure from tree canopy cover. We discovered that south Stockton, where LMR resides, had the worst convergence of urban heat, air pollution, low canopy coverage and sociodemographic vulnerability compared to northern and rural parts of the city. This was further substantiated by statistical analysis showing a strong positive relationship between areas of high LST and low vegetation. The results provided LMR with compelling evidence to support in their EJ advocacy, and in their efforts to inform state officials of the discriminatory issues they face.

Key Terms

Urban heat islands, environmental justice, air pollution, sociodemographic vulnerability, land surface temperature, nitrogen dioxide

2. Introduction

2.1 Background Information

The city of Stockton lies within the greater San Joaquin Valley (SJV) of California and is a community of over 320,000 inhabitants with a culturally and economically diverse population of residents (US Census Bureau, 2022). In 2018, Stockton was named the most culturally diverse city in the United States, with a population of 11.6% identifying as Black or African American, 45.2% Hispanic or Latino, 20.9% Asian, and 17.6% white, not including Hispanic or Latino (Galvin, 2020, U.S. Census Bureau, n.d.). The city also includes a total population of 54,000 unauthorized individuals, 71% of which traveled from Mexico and 11% from the Philippines within SJV (Profile of the Unauthorized Population - County Data, n.d.).

Increased urban development has resulted in extensive impervious surfaces and limited vegetation throughout Stockton, which causes the city to experience abnormal temperatures and climate anomalies (Jumari et al., 2023). Impervious surfaces are artificial surfaces that do not allow water infiltration and contribute to heat by modifying the surface albedo (Ma and Peng, 2021). They alter the absorption and retention of heat in cities and contribute to a phenomenon known as the Urban Heat Island (UHI) effect (Jumari et al., 2023). Higher temperatures derived from UHIs pose risks to public health such as heatstroke, heat exhaustion, air quality degradation and energy consumption (Tong et al., 2021). In addition, UHIs often disproportionately affect communities that have historically been affected by redlining (Chakraborty et al., 2019). The combination of extreme heat and poor air quality has been proven to have compounding effects on health and is shown to increase chances of heart related mortality in women and older adults (Xu, 2023).

2.2 Scientific Basis

We utilized remote sensing tools and techniques to examine differences in heat, air quality, vegetation, and land cover to calculate trends and relationships between environmental factors. Calculating land surface temperature (LST) works best as an indicator of overall heat in an area (Plott et al., 2023). Previous studies have used remote sensing from space-borne satellites to demonstrate UHI analysis and its relationship to LST

(Jumari et al., 2023). LST as a measure of heat does not include ambient heat, which is the temperature felt within the air. Large spatial and temporal variations of UHI effects are influenced by factors such as impervious surface area, vegetation cover, landscape structure, and climate (Zhou et al., 2019). Vegetation indices, such as the normalized difference vegetation index (NDVI), produced on 16-day intervals and at multiple spatial resolutions, provide consistent spatial and temporal monitoring of vegetation canopy greenness, a composite property of leaf area, and chlorophyll (MODIS Web, n.d.). Land cover classification data from National Agriculture Imagery Program (NAIP) are available for the continental US and include multi-decadal imagery collected at high spatial resolution with very low cloud coverage (Maxwell et al., 2017). Analyzing NDVI data, along with land cover classification of landscape structure, proved to be essential techniques when understanding the broader representation of UHIs in the San Joaquin County. By also including sociodemographic data such as level of education, race, and income, we identified locations with a compounding effect of both social and environmental vulnerability.

2.3 Study Area and Period

The study area encompassed all of San Joaquin County (SJC), which includes the cities of Stockton, Lodi, Tracy, and Manteca (Figure 1*)*. Much of the study focused on Stockton, specifically south Stockton where our partners Little Manila Rising (LMR) are located. The study period spanned from 2012 to 2023 and focused on summertime months when extreme heat waves were more prevalent and air pollution congregated within the SJV. Investigating this decadal stretch of time allowed us to analyze patterns of abnormal heat experiences.

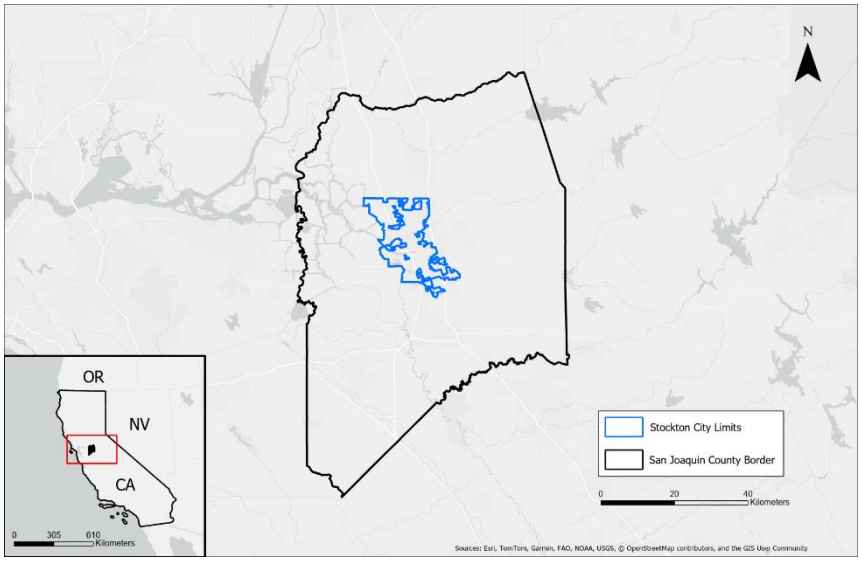

Figure 1. Study Area of the San Joaquin County border (black lines) and the Stockton city limits (blue lines). Inset: Location of study area within the state of California.

2.4 Term I Results

In the summer of 2023, DEVELOP participants analyzed the relationship between air quality and social disparities across the SJV (Szeto et al., 2023). In partnership with LMR, they found that air pollution is not equitably distributed throughout the SJV, with NO_2 specifically concentrated around urban areas and particulate matter (PM) enhanced in locations with high agricultural burning. This contributed to the accumulation of fine particulate matter 2.5 (PM 2.5), which is a mixture of solids and aerosols and is defined as particles that are 2.5 microns in diameter or less (California Air Resources Board, n.d.). Long term exposure to PM 2.5 can severely impact a person's lungs and heart (US EPA, n.d.).

2.5 Project Partners & Objectives

LMR is a social and EJ advocacy non-profit located in southern Stockton that emerged from the mass migration of Filipinos after the Philippine-American war in 1902 (Milestones: 1899–1913 - Office of the Historian, n.d.). LMR supports marginalized communities by developing solutions to harmful public policy and institutionalized racism through focusing its initiatives on programs such as urban tree planting, air quality control, and community outreach. Our project will provide LMR with a detailed understanding of urban heat and its convergences with air quality concerns. The results will be shared within LMR and its Urban Forestry program, community members, and stakeholders to incorporate into their mission to achieve multifaceted equity in Stockton.

Our objectives were to produce a social vulnerability index (SVI) based on sociodemographic variables and identify areas experiencing extreme heat and poor air quality within SJC. We also aimed to map socially vulnerable regions against neighborhoods experiencing high concentrations of NO_2 and increased LST, noting any overlap between the converging layers.

3. Methodology

3.1 Data Acquisition

3.1.1 Earth Observations Datasets for Land Surface Temperature, Vegetation, and Landcover Classification

We collected open-source data for each variable that we examined, listed in Table 1. This included raster images from Landsat 8 Thermal Infrared Sensor (TIRS), Landsat 8 Operational Land Imager (OLI), Sentinel 5 TROPOMI, and NAIP. We acquired Landsat 8 TIRS from Google Earth Engine (GEE) for our analysis of LST within the date range of 2012 – 2023 overlooking San Joaquin County. The TIRS sensor provided high resolution imagery of the thermal radiation reflecting off land surfaces such as asphalt and rooftops.

We replicated this process in GEE to collect Landsat 8 OLI satellite imagery within the date range of 2012 – 2022. The OLI sensor provided us with a vegetation coverage map of San Joaquin County. We collected term I's Sentinel-5 TROPOMI imagery that supplied the NO_2 measurements to calculate our index for air pollution exposure. Lastly, we obtained National Agricultural Imagery Program (NAIP) data from the USGS EarthExplorer online program to observe different land cover types with a resolution of one-by-one meter squared, filtering for cloud free images from 2022.

Table 1. *List of Sensors and Data Products*

Platform & Sensor	Parameter(s)	Date Ranges	Purpose	Source
Landsat 8 OLI	Surface reflectance (albedo)	2012 - 2022	OLI provided thermal analysis data for vegetation greenness.	GEE
Landsat 8 TIRS	Land surface temperature (temperature anomaly)	2012 - 2023	TIRS provided thermal analysis data for daytime land temperatures.	GEE
Sentinel 5 TROPOMI	Average NO_2	2019, 2021, 2022	TROPOMI provided NO_2 data for the study area.	GEE

NAIP	Aerial Imagery	2022	NAIP provided high resolution raster images utilized for landcover classification	USGS EarthExplorer

3.1.2 Ancillary Data

In addition to the satellite data, we acquired demographic datasets from term I of the San Joaquin Health and Air Quality Project. This included datasets from EPA's Environmental Justice Screening and Mapping tool, and U.S. Census Bureau California 2020 census tracts (Table 2). Next, we isolated the variables that aligned with our partners' biggest concerns which included: percent low income, percent people of color, population in poverty, percent less than high school education, percent under age 5, and percent over age 64.

Table 2. *Social, Environmental, and Validation Datasets*

Data Source	Data Name	Data Type	Year(s)
EPA	Environmental Justice Screening Tool (EJScreen)	CSV text file of community vulnerability	2016–2020
U.S. Census Bureau	California Census Tracts 2020	California census tract shapefile	2020

3.2 Data Processing

3.2.1 Land Surface Temperature Processing

Once we imported the Landsat 8 TIRS data, we set filter dates for the image collection to July 1 through August 31, during the 2012 – 2023 timeframe, to potentially capture the hottest days and avoid cloud coverage. To obtain accurate surface temperature data, we also filtered the image collection with a cloud mask so that each image contained < 20% cloud cover. Finally, we clipped the satellite image to only cover San Joaquin County and applied the Landsat 8 TIRS to portray the average LST of our study area. We also used a scale factor so that the image could be interpreted in terms of degrees Celsius. After we finalized the average LST map within GEE, we exported the map to ArcGIS Pro as a TIFF file. We used the zonal statistics tool to combine the LST map with the county census tract shapefile to create average LST values per San Joaquin County census tract.

3.2.2 Processing for Normalized Difference Vegetation Index

We used GEE to extract vegetation within the San Joaquin County from Landsat 8 OLIdata. We began by filtering images that correlated to our study period of 2012 to 2022 from July 1 to August 31 Next, images with cloud coverage > 1% were removed using GEE. We then clipped the raster image to our study bounds of San Joaquin County using a shapefile from the Stockton GIS data catalogue. Once these steps were completed, we used *Equation 1* to extract band values associated with vegetation. The final output produced a raster image with values ranging from –1, little to no vegetation, to 1, dense vegetation.

$$NDVI = \frac{RED - NIR}{RED + NIR} \qquad \text{Equation 1}$$

Equation 1. NDVI equation using red and near infrared bands.

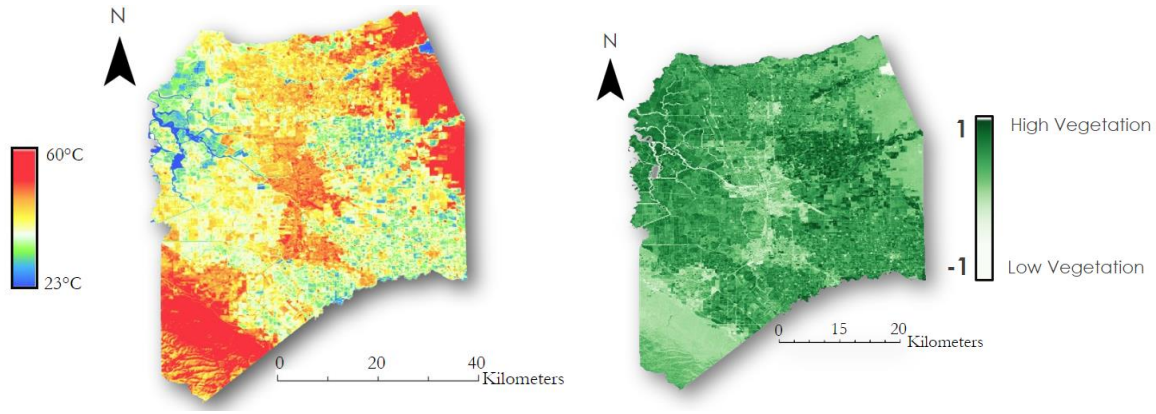

Figure 2. This is a cloud-masked mean LST (left) and NDVI (right) image clipped to the San Joaquin County study boundary for the period July 1 – August 31.

3.2.3 Processing for Social Vulnerability Index

We uploaded the social vulnerability variables from EPA's EJScreen into Microsoft Excel. In another Excel sheet, we added our census data containing the GEOIDs for all the census tracts in SJC. We joined the demographic data with the geographic data using the GEOIDs as a common attribute field. Since some census tracts did not have any data and our calculation for the SVI required a complete data set, we filled all the empty fields with the average of each variable.

3.2.4 Processing for Land Cover Classification

To assess the dominant land cover type of south Stockton, we performed a supervised land cover classification in ArcGIS Pro. We sourced 7 images from 2022 by the NAIP of demogr aphically distinct areas within Stockton and clipped them together to create a mosaic of images within the city boundary. Before classifying the imagery, we grouped pixels into segments to reduce errors and inaccuracies. To distinguish features more clearly, we changed the band combination to assign the red band to near infrared (Band 4), the green band to red (Band 1), and the blue to blue (Band 3). This allowed for training data to be more accurately assigned to their corresponding classes when segmented. We created seven distinct classes representing water, wetlands, developed, canopy, agriculture, barren, and grass. Using the Training Samples Manager tool within ArcGIS Pro, we assigned a minimum of 10 training samples to each class before performing the classification. We used the K-Nearest Neighbor classification method, which uses proximity to other classes to determine classification (Noi and Kappas, 2017).

Once Stockton was classified, we extracted both the developed and canopy classes from the land cover raster to assess the individual land cover types. This was done by using the Summarize Raster tool in ArcGIS Pro to display the pixel counts for each class within individual census tracts. We created a new column within the attribute table and divided the pixels of each class by the total number of pixels in each census tract to normalize the counts for each class. Finally, we visualized these values as a gradient of percentage of each land cover type within census tracts.

3.3 Data Analysis

3.3.1 Social Vulnerability Analysis

Vulnerability is defined as a population's exposure and sensitivity plus/minus their adaptive capacity (ARSET, n.d.). We used EJScreen data to find indicators of vulnerability, the same sociodemographic from the ancillary data, and quantified them through the creation of vulnerability indices. First, we acquired data sorted by census tract ID for our specific vulnerability indicators, and then ranked each value within a variable from lowest unit to highest. Using these ranked variables, we manipulated them into percentile ranks which were summed to create our final social vulnerability index (SVI). The higher the final SVI value is, the more susceptible an area is to environmental hazards. We visualized areas of low and high exposure by joining our vulnerability index to the SJC census tract shapefile and re-symbolized it using a choropleth map.

3.3.2 Vegetation and Temperature Correlation Analysis

We uploaded both mean LST and NDVI raster images to ArcGIS Pro 3.2 along with the study area shapefile. To graph the correlation between LST and NDVI, we used the Fishnet tool to create a 100 by 100-point grid over the study area. All point values layered over the raster images were extracted using the Extract Multi Values to Point tool. We extracted the associated value for each point data to display in the attribute table of the previous output file. Next, we filtered out values that were associated with water to not skew the regression line, using the Clip tool to extract only the points within the study area. We then graphed the data points onto a scatter plot in ArcGIS Pro and conducted a linear regression to determine the correlation between these two variables.

3.3.3 Air Quality Analysis

We recycled the previous team's raster imagery, which was taken by Sentinel-5 TROPOMI, of NO_2. The raster displayed 95th percentile and median NO_2 moles/meter^2. For our purposes, we used the 95th percentile NO_2 raster. First, we calculated the average NO_2 in SJC using the Raster Calculator tool in QGIS for the years 2019, 2021, and 2022. This resulted in a single, consolidated raster with each pixel representing the average NO_2 moles/meter^2. Next, we used our shapefile of SJC, which contains all the census tract areas, the average NO_2 raster, and the Zonal Statistics tool to find the average NO_2 moles per square meter at the census tract level. With our updated shapefile with NO_2 data, we visualized the census tracts in SJC that have the most exposure to high NO_2 concentrations using a choropleth map. Furthermore, we noted the locations of major roads and highways to identify any spatial correlation between distance to transportation corridors and NO_2 concentrations.

3.3.4 Heat and Air Quality Analysis

To combine the LST and air quality maps, we created a new LST and air quality index (LSTAQI) that we mapped against our SVI. LSTAQI was created using the same method as the SVI but with one additional step. Once we converted the ranks into percentiles for NO_2 and LST, we multiplied the NO_2 percentiles by 0.33 and the LST percentiles by 0.66. Since we wanted to emphasize the importance of LST in the index, we decided that it would take up 2/3 or 0.66 of the total index, and NO_2 would take up 1/3 or 0.33 of the total index. We added the two weighted percentiles together to create the LSTAQI. Next, we joined the new LSTAQI and the SVI with the SJC shapefile to capture social vulnerability, NO_2, and LST with census tract data. In ArcGIS Pro, we symbolized the two variables using a bivariate choropleth map. The map displays socially vulnerable areas as well as combined LST and NO_2 hazards. Areas that overlap between these two variables are symbolized as a blended color.

3.3.5 Principal Component Analysis

We carried out our principal component analysis (PCA) to reduce the dimensionality of our data while preserving most of the variance when creating our social vulnerability index. The variables used in the PCA are population in poverty, percent people of color, percent low income, percent less than high school education, percent less than age 5, percent over age 64, and mean LST. The PCA was created in Python using these libraries: pandas, NumPy, matplotlib.pyplot, and PCA from sklearn.decomposition. First, we read and

saved our social vulnerability variables into an object called 'data' and normalized it. Next, we created a correlation matrix and plotted it so that we had a table matrix and plot matrix. Third, we computed the eigenvectors and eigenvalues of the covariance matrix. Since we had seven variables, we created seven principal components. The output data frame displayed the variance of each principal component per census tract. With this data frame, we visualized the variance that each principal component contains using a scree plot (*Figure A3)*.

3.3.6 Temperature Variation Analysis

We calculated the urban heat intensity by finding the difference between average temperatures in an urban location and its rural counterpart. First, we created polygons within the different areas using the original LST map and produced histograms of the temperature data, showing the total average temperatures. South Stockton experienced an 8.4 °C (15.12 °F) higher temperature difference compared to rural locations as shown in *Figure A7* and *Figure A9*. These histograms allowed us to compare the two distinct spreads of temperature between rural and urban locations and showed us if the data had a normal distribution. The data were normally distributed, and we ran a two sample Wilcoxon T-test within R 4.3.2 to confirm that there was a statistically significant difference. We exported the histogram data from GEE to Excel to expand the "count" column so that our data table could be read within R. Once the data were cleaned, we ran a Wilcoxon T-test comparing rural and south Stockton temperatures and received a p-value of < 0.001. This result meant that we could reject the null hypothesis that stated there was no difference between the temperatures in the rural and urban regions. Using the same procedure, we compared south and north Stockton and found that south Stockton had a 3 °C (5.4 °F) higher temperature difference compared to north Stockton, as seen in *Figure A8* and *Figure A9*. We validated this conclusion by running the same T-test within R and established that the temperature difference was statistically significant because of a resulting p-value of < 0.001.

4. Results & Discussion

4.1 Analysis of Results

4.1.1 Social Vulnerability Results

The results of the PCA show that the strongest signs of social vulnerability are the percentage of people of color, percentage population of low income, and percentage of population with less than a high school education. Additionally, five out of seven principal components contain 90% of the total variance that our social vulnerability variables portray. We produced a choropleth map that displays census tracts of SJC with lowest to highest social vulnerability (Figure 3).

The choropleth map results align with the hypothesis that the social vulnerability index decreases in the census tracts furthest from highly populated areas such as Stockton. Our team also identified higher index values in south Stockton, while northern Stockton has lower index values. This observation aligns with LMR's principle that south Stockton faces unequitable living situations because of redlining.

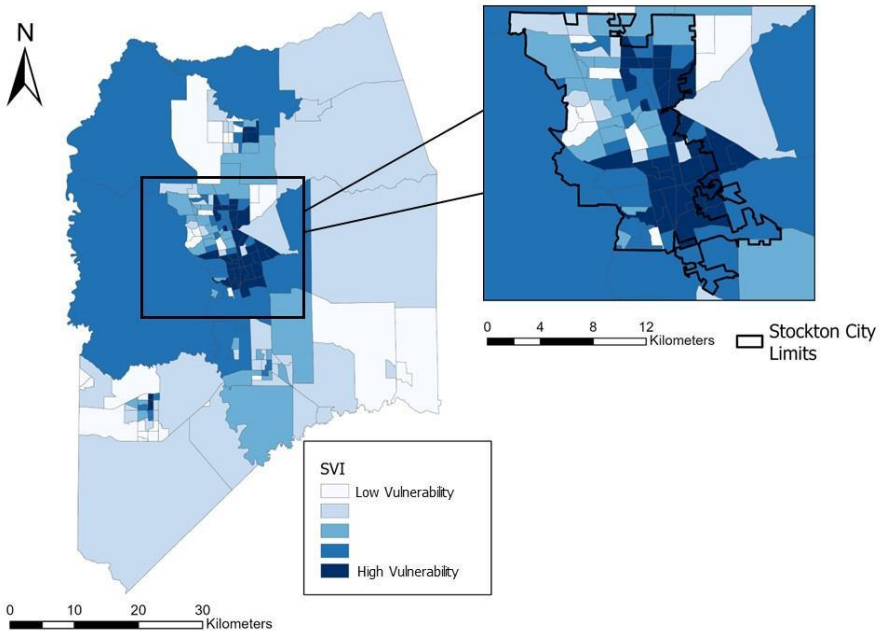

Figure 3. Social Vulnerability Choropleth Map - Social Vulnerability Index calculated for San Joaquin County, CA. Inset: The city of Stockton.

4.1.2 Vegetation and Temperature Correlation Results

The distribution of points plotted in our regression line graph show a strong negative correlation and an R-squared value of 0.65, indicating that 65% of the variance in temperature is explained by vegetation. This is considered a strong correlation and explains why communities with low vegetation will experience higher temperatures compared to those that have denser surrounding vegetation. Additionally, most of the clustered vegetation in Figure 2 are around areas of agricultural cultivation.

4.1.3 Nitrogen Dioxide Concentrations

Further reinforcing the results of term I, we found that high NO_2 levels are seen around urban areas while non-urban areas have less NO_2. We suspect that NO_2 levels are higher in urban areas because of the high number of vehicles driving along highways and major roads that cut through Stockton. We detected larger levels of NO_2 in south Stockton compared to north Stockton. The higher NO_2 levels may be due to densely industrialized urban land cover in the area. Overall, census tracts within the Stockton city limits contain higher concentrations of NO_2, with south Stockton facing the highest in the county, which reinforces the trend where south Stockton experiences higher frequencies of environmental hazards.

One factor that we did not account for until after creating this map was that the units of moles/meter^2 are not good measurements for understanding experienced air quality. For example, we don't know if 2.64×10^{-4} moles/meter^2 is considered hazardous air quality conditions. The United States Air Quality Index (AQI) has six categories from good to hazardous and measures NO_2 by parts per billion (ppb). Making the conversion from mol/meter^2 to ppb is doable but not trivial. If we had this idea sooner rather than later, it would have been feasible to make this conversion.

4.1.4 Air Quality Vulnerability

We combined the social vulnerability, air quality, and LST maps to create bivariate choropleth maps to determine which census tracts are experiencing high LST, poor air quality, and high levels of social vulnerability. Some regions may be highly susceptible to injustices but lack strong NO_2 concentrations.

Conversely, some regions may contain excess amounts of NO_2 emissions, but are not as susceptible to injustices. We were primarily interested in areas that are vulnerable and have excess NO_2 emissions because these locations are where mitigation is needed the most. We first examined areas with only high social vulnerability. We suspect that these areas are highly vulnerable but hold low NO_2 concentrations because those census tracts lack infrastructure such as highways, roads, and industrial development that creates NO_2. Next, we looked at areas with only high NO_2 concentrations. We believe that these regions are highly urbanized but lack social vulnerability due to being more wealthy, educated, or lacking diversity. Last, we identified communities that overlap with moderate to high social vulnerability and NO_2 concentrations. We observed that south Stockton has the worst conditions.

4.1.5 Heat Vulnerability Results

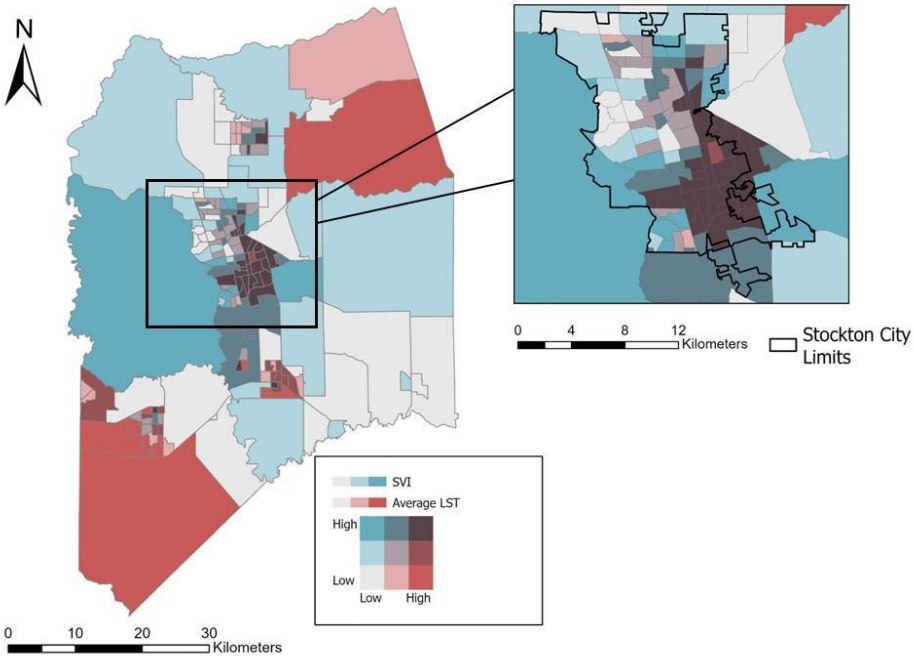

Figure 4. Heat and Social Vulnerability Bivariate Choropleth Map

After we created a new shapefile with the mean LST per county census tract as a field, we joined attribute tables with the SVI so that we could map both variables together (Figure 4). This map allows viewers to identify where there is a convergence of locations with high heat and high social vulnerability, while still distinguishing areas that only experience one or the other. We see in this figure that barren areas in the upper right (northeast) and lower left (southwest) corners face extreme heat with low social vulnerability because the population density is lower. However, it is apparent that south Stockton is experiencing both increased LST and social vulnerability. This is a significant issue because extreme heat is not being felt throughout the entire city and it is only severely impacting communities that have been disenfranchised and forced to live in these environments.

4.1.6 Final Heat and Air Quality Vulnerability Results

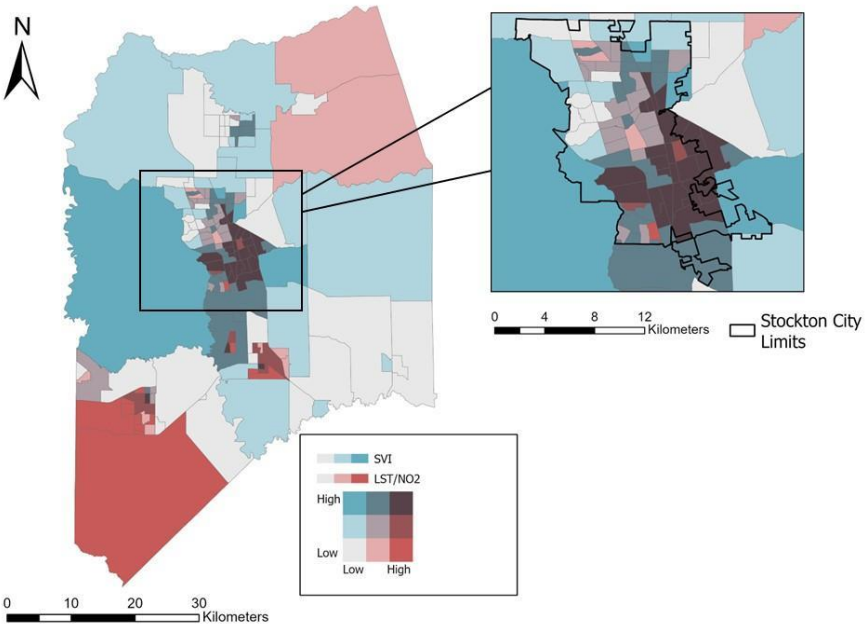

Figure 5. Final Bivariate Choropleth Map of heat, air quality, and social vulnerability

Through combining the three main variables of this study (LST, air quality, and social vulnerability), we showed that there are regions within SJC, and specifically Stockton, that experience extreme convergence of urban heat and poor air quality (Figure 5). Since the previous term provided our partners with a thorough air quality analysis, we weighed land surface temperature more heavily within this new LSTAQI. Our final air quality and LST vulnerability map looks like the previous maps as it portrays very comparable aspects with minor changes. Again, we observed that south Stockton is extremely exposed to these environmental hazards, posing risks for unsafe living conditions and health threats.

4.1.7 Land Cover Classification Visualization

We visualized Developed and Canopy land cover as maps along a gradient based on pixel counts for each land cover type by census tract, ranging from roughly 300 to 14,0000 pixels (Figure 6).

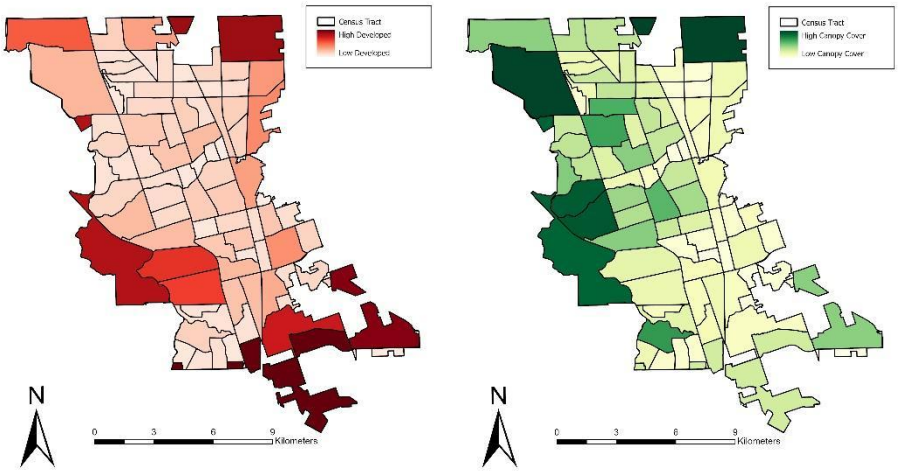

Figure 6. Land cover classes; developed (left) and canopy cover (right) within Stockton.

The developed map shows that the highest pixel count for this class lies in the south Stockton area, which is mainly composed of industrial zones, such as the Port of Stockton and the Stockton Metropolitan Airport. The canopy cover class shows the census tracts with the highest pixel counts are mainly concentrated throughout the northwest portion of Stockton. These areas are more affluent than south Stockton and contain a greater density of parks and natural features, including golf courses and bodies of water. In contrast, south Stockton experiences some of the lowest canopy cover in the city. This confirms our partners' concerns that residents in these areas have less access to shade in their neighborhoods and can also help to explain the increase in land surface temperature when compared to other parts of Stockton.

4.1.8 Principal Component Analysis Results

The correlation matrix in *Table A1* displays the degrees to which our variable pairs move in relation to each other. Values above zero show a positive correlation between variables where one variable increases the other variable increases too. Values below zero show a negative correlation between variables where one variable increases the other variable decreases. Values that are around 0 show that the variables have no correlation. Some of the variables that show high positive correlation with each other are percent people of color, percent less than high school education, percent low income, and average land surface temperature. The pair percent people of color and percent over age 64 have a strong negative correlation.

Figure A3 shows the percentage of variance that each principal component contains. In this case, principal component 1 represents 53% of the data and principal component 2 represents 22% of the data. The sum of all the principal components will be equal to 100%. To maximize the most variance while using the minimum number of principal components, our team selected the first four principal components for our analysis because these four make up at least 90% of the variance for our data set.

4.2 Feasibility for Partner Use and Future Recommendations

We found remote sensing tools such as ArcGIS, QGIS, and GEE to be vital in examining the relationship between land use and landcover, vegetation, temperature, and air quality within our study area. Aside from ArcGIS, the tools and data we utilized in our research are free access and public domain data that are available for anyone to download. With these end products, Little Manila Rising can focus their efforts on strategically placing cooling centers throughout Stockton to limit exposure to extreme heat events. In

addition, LMR can focus their urban forestry efforts on areas that contain the lowest canopy cover throughout south Stockton, shown in *Figure A4*, to continue to provide shade to residents. Finally, LMR can also use the creative communication brochure to continue to inform community members of the environmental hazards they are subjected to.

Little Manila Rising can continue to build on these end products by changing the Social Vulnerability Index to incorporate demographic factors they find more relevant in the future. They can also continue to monitor their urban forestry program objectives by assessing the change in canopy cover once trees have grown sufficiently enough to provide cooling effects. In addition, the air quality vulnerability map can be further improved by incorporating transit data to find the relationship between NO_2 concentration and proximity to major roads and highways. The Nitrogen Dioxide Choropleth Map can be improved by standardizing the legend to fit with the United States Air Quality Index. That way the numbers represented can communicate to the viewer the severity of the concentration of NO_2 from good, moderate, unhealthy, and hazardous. As previously recommended from Term I, TROPOMI can also measure carbon monoxide (CO), and formaldehyde (HCHO), which are other pollutants of concern that directly affect human health. We did not incorporate these other pollutants into our air quality analysis but echo this recommendation as well.

5. Conclusions

While exploring urban heat islands and air pollution within the San Joaquin County, we observed that southern Stockton, where our partners for this project reside, has a higher exposure rate to environmental issues compared its northern and rural counterparts. We discovered through our analysis of combining demographic data with our heat analysis map that southern communities disproportionately experienced increased urban temperatures. Additionally, through our analysis of developed landcover, we saw a large presence of poor air quality possibly due to major developments such as an airport, major highways, and shipment ports. This adds to the overall concern of increased health risks, such as high asthma rates, because of Stockton's history of redlining minority based low-income communities. Our vegetation and landcover analysis illustrated the lack of tree canopy cover in southern communities, which can intensify the effects of urban heat islands in these areas. Furthermore, after running statistical analysis comparing the differences in heat between north, south, and rural areas of Stockton we concluded that southern communities are at a much higher risk to increased temperatures. These results will help LMR provide crucial information to the nearby community members and aid them in their efforts to continue their Urban Forestry Program. They can also be used in urging the city to provide more funding that is directed toward providing cooling centers and heat mitigation strategies.

6. Acknowledgements

We would like to thank our partner, Little Manila Rising, for their collaboration. Thank you to everyone that helped and guided us in this project. Thank you to our advisors Dr. Morgan Gilmour, Dr. Juan Torres-Perez, Dr. Xia Cai, Dr. Kenton Ross, Dr. Joshua Dimon, and Lisa Tanh who showed up every week to answer any questions we had and guided us in this process. Big thank you to our Center Lead Lauren Webster and Impact Analysis Fellow Maya Hall who showed us the ropes in this stressful process with many big learning curves.

This material contains modified Copernicus Sentinel data (2023), processed by ESA. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration. This material is based upon work supported by NASA through contract 80LARC23FA024.

7. Glossary

Aerosol Optical Depth (AOD) – A measure of aerosols distributed within a column of air from the measuring instrument to either the ground or top of the atmosphere.
ArcGIS Pro – Geographic Information System (GIS) Software provided by Esri.
Canopy Cover – The floor area covered by a tree's leaves, branches, and stems when viewed from above.
Earth Observations (EOs) – satellites and sensors that collect information about the Earth's physical, chemical, and biological systems over space and time.
Geographic Information System (GIS) – computer-based tools used to store, visualize, analyze, and interpret geographic data.
Google Earth Engine (GEE) –Earth observations processing software
Land Surface Temperature (LST) – A proxy for calculating the temperature emitted of man-made objects and materials.
Land Surface Temperature Air Quality Index (LSTAQI) – The LSTAQI is an index made by weighing the land surface temperature and NO_2 values and adding them together. This index is compared with the social vulnerability index to find converging vulnerabilities.
National Agriculture Imagery Program (NAIP) – NAIP is aerial imagery taken aboard an airplane rather than a satellite.
Nitrogen Dioxide (NO_2) – Nitrogen Dioxide used as a proxy for vehicle emissions.
Operational Land Manger (OLI) – sensor on Landsat-8 satellite.
Principal Component Analysis (PCA) – A statistical test used when wanting to find the correlation between multiple related variables.
Social Vulnerability Index (SVI) - A way of measuring the "susceptibility" within a community using demographic data
Thermal Infrared Sensor (TIRS) – sensor on Landsat-8 satellite.
Tropospheric Monitoring Instrument (TROPOMI) – sensor on the Sentinel-5P satellite that measures NO_2

8. References

ARSET - Satellite Remote Sensing for Measuring Urban Heat Islands and Constructing Heat Vulnerability Indices. (2022). NASA Applied Remote Sensing Training Program (ARSET). http://appliedsciences.nasa.gov/get-involved/training/english/arset-satellite-remote-sensing-measuring-urban-heat-islands-and

California Air Resources Board. (2019, January 24). *Clean-air plan for San Joaquin Valley first to meet all federal standards for fine particle pollution | California Air Resources Board*. https://ww2.arb.ca.gov/news/clean-air-plan-san-joaquin-valley-first-meet-all-federal-standards-fine-particle-pollution

Centre for Geoinformatics & Planetary Studies, Dept. of Geology, Periyar University, Salem, Tamilnadu, India. (2016). Statistical Correlation between Land Surface Temperature (LST) and Vegetation Index (NDVI) using Multi-Temporal Landsat TM Data. *International Journal of Advanced Earth Science and Engineering*, *5*(1), 333–346. https://doi.org/10.23953/cloud.ijaese.204

Chakraborty, T., Hsu, A., Manya, D., and Sheriff, G. (2019). Disproportionately higher exposure to urban heat in lower-income neighborhoods: A multi-city perspective - IOPscience. *Environmental Research Letters*, *14*(10). https://doi.org/10.1088/1748-9326/ab3b99

Collins, J., and Dronova, I. (2019). Urban Landscape Change Analysis Using Local Climate Zones and Object-Based Classification in the Salt Lake Metro Region, Utah, USA. *Remote Sensing*, *11*(13), 1615. https://doi.org/10.3390/rs11131615

Dronova, I., Friedman, M., McRae, I., Kong, F., and Yin, H. (2018). Spatio-temporal non-uniformity of urban park greenness and thermal characteristics in a semi-arid region. *Urban Forestry & Urban Greening*, *34*, 44–54. https://doi.org/10.1016/j.ufug.2018.05.009

Flanagan, B. E., Gregory, E. W., Hallisey, E. J., Heitgerd, J. L., and Lewis, B. (2011). A Social Vulnerability Index for Disaster Management. *Journal of Homeland Security and Emergency Management*, *8*(1). https://doi.org/10.2202/1547-7355.1792

Galvin, G. (2020, January 22). America's Most Diverse City Is Still Scarred by Its Past [Review of America's Most Diverse City Is Still Scarred by Its Past]. US News. https://www.usnews.com/news/cities/articles/2020-01-22/stockton-california-americas-most-diverse-city-is-still-scarred-by-its-past#:~:text=In%202018%2C%20Stockton%20%E2%80%93%20whose%20roughly,based%20on%20recent%20census%20data.

Gunawardena, K. R., Wells, M. J., and Kershaw, T. (2017). Utilising green and bluespace to mitigate urban heat island intensity. *Science of The Total Environment*, *584–585*, 1040–1055. https://doi.org/10.1016/j.scitotenv.2017.01.158

Huang H, Yang H, Deng X, Zeng P, Li Y, Zhang L, Zhu L. Influencing Mechanisms of Urban Heat Island on Respiratory Diseases. Iran J Public Health. 2019 Sep;48(9):1636-1646. PMID: 31700819; PMCID: PMC6825676.

Ju, Y., Dronova, I., Rodriguez, D. A., Bakhtsiyarava, M., and Farah, I. (2023). Recent greening may curb urban warming in Latin American cities of better economic conditions. *Landscape and Urban Planning*, *240*, 104896. https://doi.org/10.1016/j.landurbplan.2023.104896

Kasniza Jumari, N. A. S., Ahmed, A. N., Huang, Y. F., Ng, J. L., Koo, C. H., Chong, K. L., Sherif, M., and Elshafie, A. (2023). Analysis of urban heat islands with landsat satellite images and GIS in Kuala Lumpur Metropolitan City. *Heliyon*, *9*(8), e18424. https://doi.org/10.1016/j.heliyon.2023.e18424

Kong, F., Sun, C., Liu, F., Yin, H., Jiang, F., Pu, Y., Cavan, G., Skelhorn, C., Middel, A., and Dronova, I. (2016). Energy saving potential of fragmented green spaces due to their temperature regulating ecosystem services in the summer. *Applied Energy*, *183*, 1428–1440.

Ma, X., and Peng, S. (2021). Assessing the quantitative relationships between the impervious surface area and surface heat island effect during urban expansion. PeerJ, 9, e11854. https://doi.org/10.1016/j.apenergy.2016.09.070

Maxwell, A.E., Warner, T. A., Vanderbilt, B. C., Ramezan, C. A. (2017). Land cover classification and feature extraction from National Agriculture Imagery Program (NAIP) orthoimagery: a review. *Photogrammetric Engineering & Remote Sensing,* 237-247.

Milestones: *1899–1913—Office of the Historian.* (n.d.). Retrieved February 29, 2024, from https://history.state.gov/milestones/1899-1913/war

MODIS Web. (n.d.). . https://modis.gsfc.nasa.gov/.

More Than History. (n.d.). Little Manila Rising. Retrieved February 29, 2024, from https://littlemanila.org/more-than-history

Noi, P. T., & Kappas, M. (2017). Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors*, *18*(1). https://doi.org/10.3390/s18010018

Plott, L., Pazmino, M., Hall, M., & Villalobos-Heredia, C. (2023). Remote Sensing of Urban Heat: DEVELOP White Paper. *NASA DEVELOP Program*, 1-8.

Profile of the Unauthorized Population—County Data (6077). (n.d.). Migrationpolicy.Org. Retrieved April 15, 2024, from https://www.migrationpolicy.org/data/unauthorized-immigrant-population/county/6077

Szeto, J., Beardslee, J., Christian, P., MacCarter, B., & Quintero, A. (2023). Evaluating the Overlap of Social vulnerabilities and Air Quality in the San Joaquin Valley Air Pollution Control District. *NASA DEVELOP Program*, 1–33.

Tong, S., Prior, J., McGregor, G., Shi, X., & Kinney, P. (2021). Urban heat: An increasing threat to global health. *BMJ (Clinical Research Ed.)*, *375*, n2467. https://doi.org/10.1136/bmj.n2467

U.S. Census Bureau QuickFacts: Stockton city, California. (n.d.). Retrieved February 27, 2024, from https://www.census.gov/quickfacts/fact/table/stocktoncitycalifornia/PST045222

US EPA, O. (2015, February 13). *Learn About Environmental Justice* [Overviews and Factsheets]. https://www.epa.gov/environmentaljustice/learn-about-environmental-justice

Xu, R. (2023). Extreme temperature events, fine particulate matter, and myocardial infarction mortality. *Circulation.* https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.122.063504

Yin, H., Kong, F., Middel, A., Dronova, I., Xu, H., and James, P. (2017). Cooling effect of direct green façades during hot summer days: An observational study in Nanjing, China using TIR and 3DPC data. *Building and Environment*, *116*, 195–206. https://doi.org/10.1016/j.buildenv.2017.02.020

Zhou, D., Xiao, J., Bonafoni, S., Berger, C., Deilami, K., Zhou, Y., Frolking, S., Yao, R., Qiao, Z., and Sobrino, J. A. (2019). Satellite Remote Sensing of Surface Urban Heat Islands: Progress, Challenges, and Perspectives. Remote Sensing, 11(1), 48. https://doi.org/10.3390/rs11010048

9. Appendices

Appendix A.

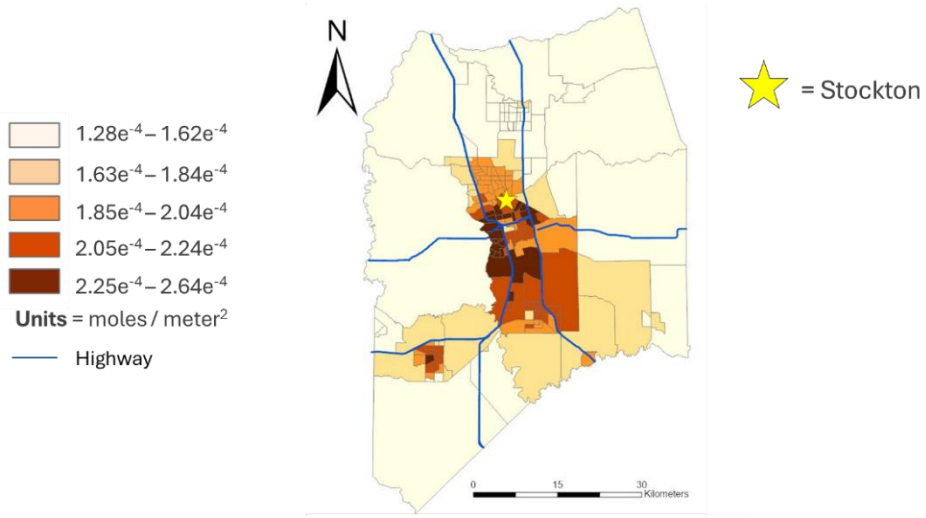

Figure A1. Nitrogen Dioxide Choropleth Map showing the distribution of NO_2 throughout the county with major roadways

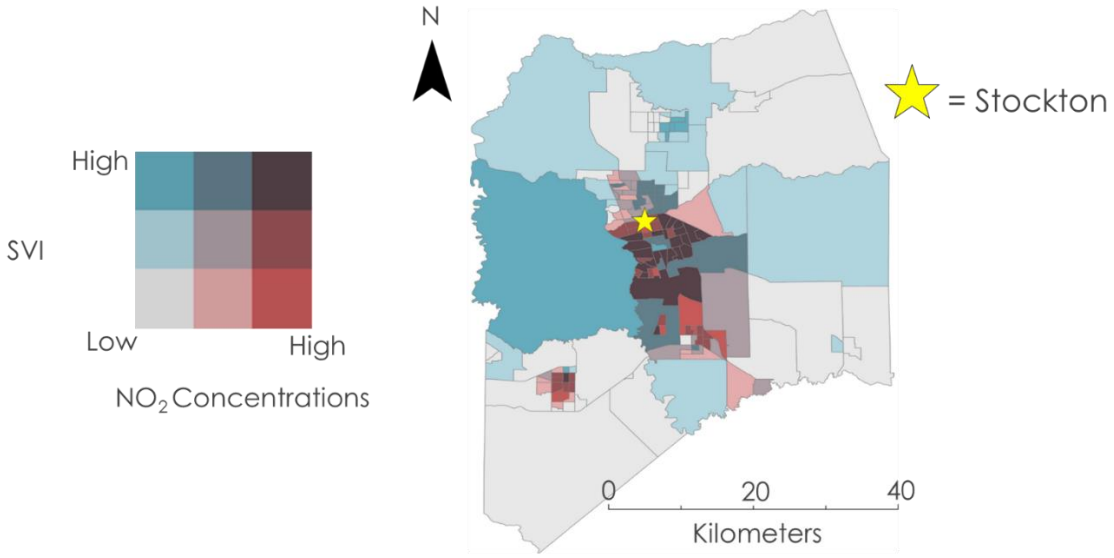

Figure A2. Air Quality Vulnerability Bivariate Choropleth Map

	population in poverty	% people of color	% low income	% <HS Education	% < age 5	% > age 64	lst_mean
population in poverty	1.000						
% people of color	0.052	1.000					
% low income	-0.328	0.555	1.000				
% <HS Education	-0.242	0.648	0.781	1.000			
% < age 5	0.031	0.365	0.459	0.379	1.000		
% > age 64	-0.285	-0.553	-0.233	-0.399	-0.378	1.000	
lst_mean	-0.021	0.507	0.273	0.336	0.277	-0.464	1.000

Table A1. Correlation matrix with values.

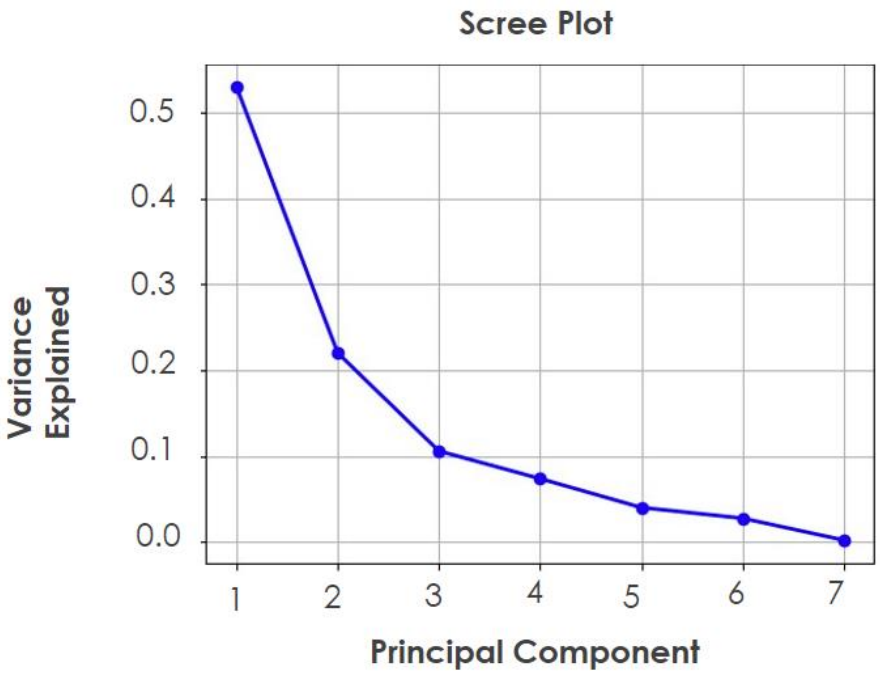

Figure A3. Scree Plot of principal components variance

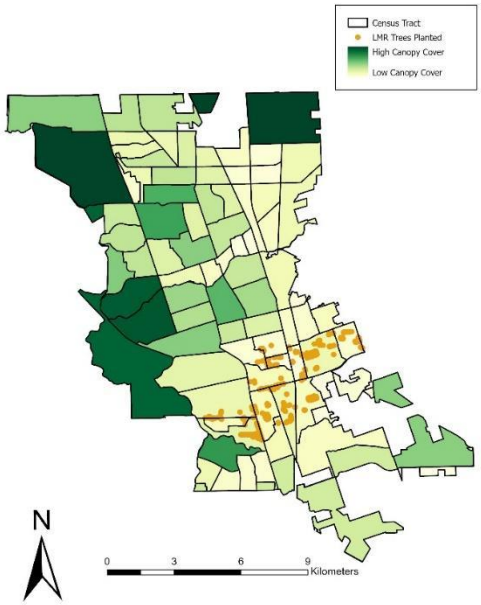

Figure A4. Canopy cover classification map with trees planted by Little Manila Rising

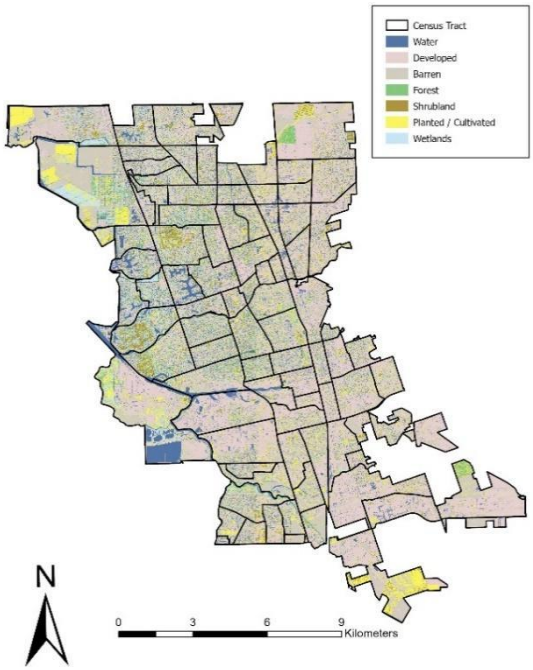

Figure A5. Land cover classification Stockton Map

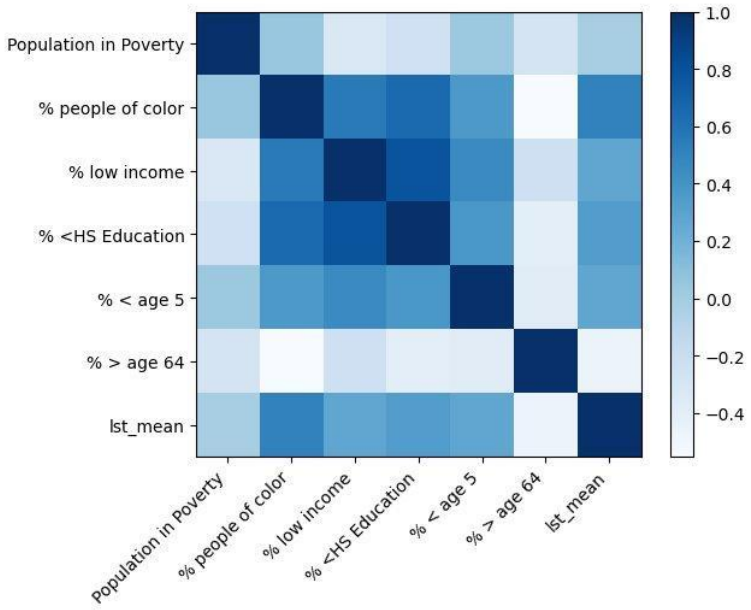

Figure A6. Social vulnerability index variables correlation matrix

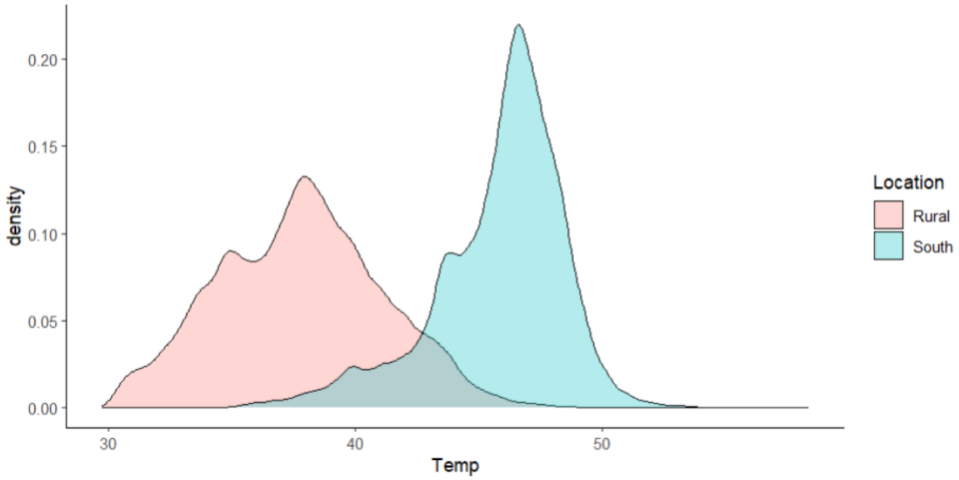

Figure A7. Histogram of average LST values comparing south and rural Stockton

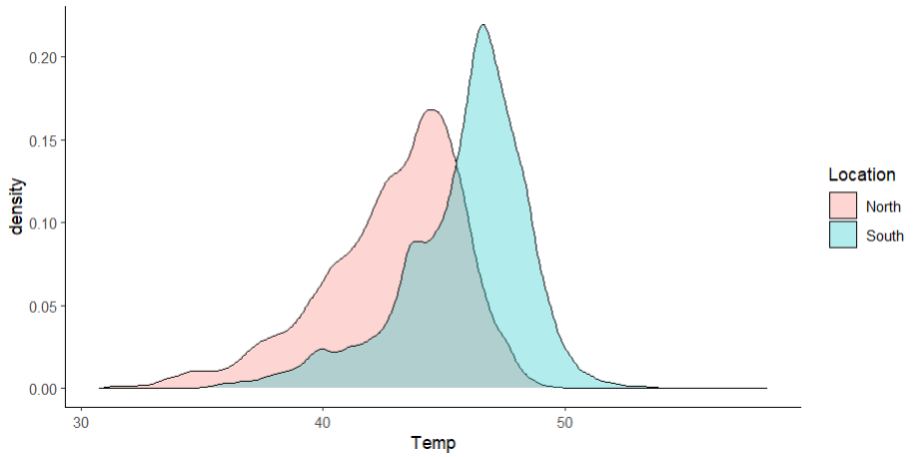

Figure A8. Histogram of average LST values comparing north and south Stockton

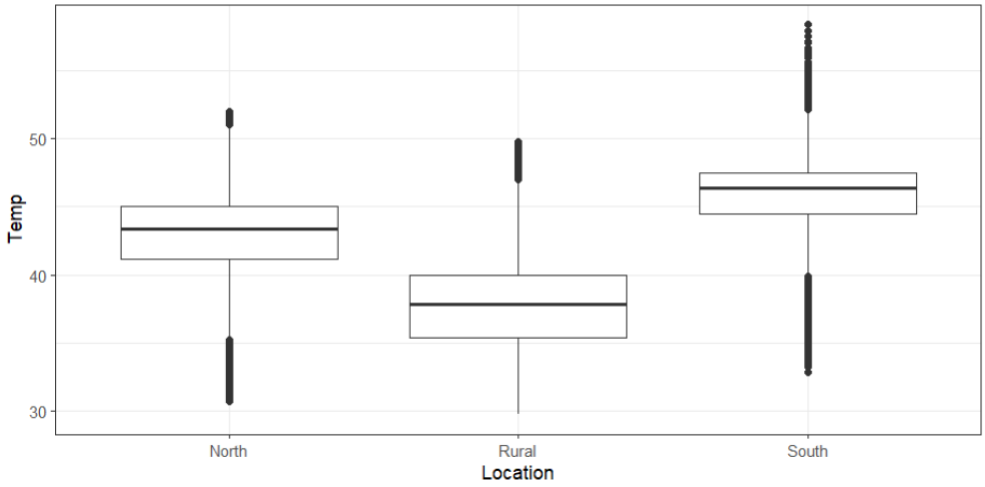

Figure A9. Box plot comparison of average LST values for north, rural, and south Stockton

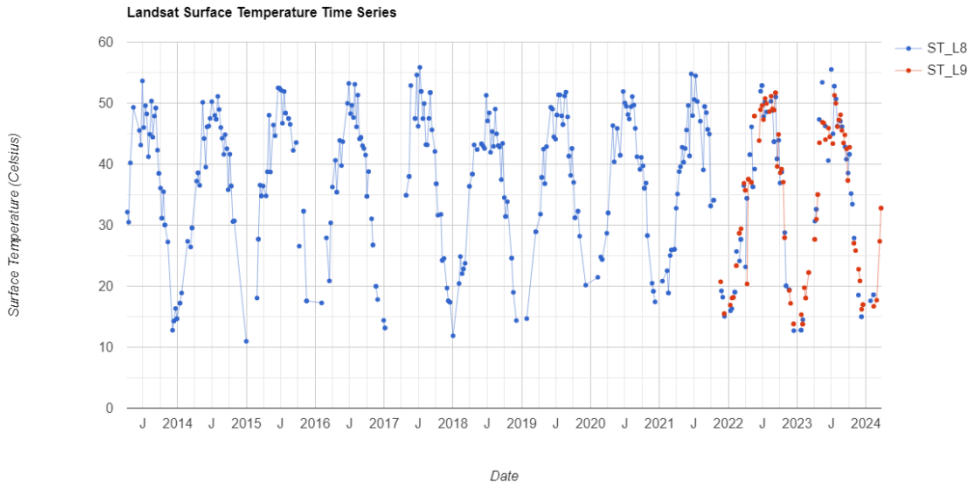

Figure A10. Time series of average LST values of a single pixel within a southern Stockton neighborhood