



Summer 2024

Cali Urban Development II
Investigating the Impacts of Land Use Change on Urban Heat and Social
Vulnerability in Cali, Colombia

DEVELOP Technical Report

August 9th, 2024

Gabi Davidson-Gomez (Project Lead)
Brenna Bruffey
Deya Rodriguez
Isabelle Solórzano

Advisors:

Dr. Morgan Gilmour, NASA Ames Research Center (Science Advisor)
Dr. Alexandra Christensen, NASA Jet Propulsion Laboratory (Science Advisor)
Dr. Xia Cai, NASA Langley Research Center (Science Advisor)
Lisa Tanh, Esri (Science Advisor)

Previous Contributors:

Tallis Monteiro
Gabi Davidson-Gomez
Nathan Tesfayi
Raquel Trejo

Lead:

Lauren Webster (California – Ames)

1. Abstract

The surface urban heat island (SUHI) effect is an environmental phenomenon resulting in cities with higher temperatures than rural areas due to increased pavement and decreased cooling from vegetation. The city of Santiago de Cali in Colombia faces SUHI challenges exacerbated by land use change. The Cali municipal government agency, Departamento Administrativo de Gestión del Medio Ambiente, and the community organization Fundacion Dinamizadores Ambientales partnered with NASA DEVELOP to evaluate communities in Cali most vulnerable to urban heat. This project illustrated the utility of using NASA Earth observations to evaluate the relationship between land use, temperature, and social factors in Cali, Colombia between 2013 and 2023. The team used Landsat 7 Enhanced Thematic Mapper Plus (ETM+), Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS), and Landsat 9 OLI-2/TIRS-2 to generate land surface temperature (LST) maps in Google Earth Engine through NASA DEVELOP's Urban Heat Exposure Assessment Tempe 2.0 tool. Cloud cover limited the project feasibility, but it improved with Landsat 9 data. In ArcGIS Pro, the team found that LST was significantly higher in urban areas than in wetlands or forests. Using R Studio, the team ran a principal component analysis and found that health care and green space access were negatively correlated, and Afro-Colombian ethnicity was positively correlated with LST. With knowledge of the most impacted and vulnerable regions, the partner organizations can prioritize establishing healthcare facilities and green spaces in those areas to reduce the impacts of urban heat.

Key Terms

Colombia, land surface temperature (LST), land use/land cover change (LULC), remote sensing, socioeconomic vulnerability, surface urban heat islands (SUHIs), urban development, vegetation loss

2. Introduction

The surface urban heat island (SUHI) effect is an environmental phenomenon where urban areas experience higher temperatures than their rural surroundings due to high levels of urbanization (Li et al., 2021). Paved surfaces and buildings absorb high levels of energy from the sun and then radiate that energy as heat to the atmosphere, making urban areas hotter than their rural counterparts. The SUHI effect is a prominent issue in Santiago de Cali, Colombia, colloquially known as Cali, due to decades of deforestation, wetland degradation, and development. This investigation aims to examine the impacts of land use change on the SUHI effect and social vulnerability. Identifying spatial relationships between SUHIs and land cover change is important information for the local population to understand the extent of heat disparity and its impacts on different communities in Cali.

Colombia's rapid urbanization over the past three decades caused significant changes in land use, including the purposeful draining of wetlands for agriculture, which led to the shrinking of wetland area by 99% in Cali (Ocampo-Marulanda et al., 2021). Understanding the environmental ramifications of this expansion is crucial for future sustainable urban planning. A recent study (Salazar Tamayo and Julio Estrada, 2022) highlighted the lack of effective and sustainable management of urban growth in Cali by analyzing urban expansion patterns using Landsat 5 and census data. This study revealed that local authorities often underestimate rates of population growth and land requirements. However, urban development that emphasizes the preservation of blue and green spaces (e.g. wetlands and forests) can mitigate SUHI impacts (Li et al., 2021). One study examined the impact of land use and land cover changes on SUHI patterns, highlighting how wetland fragmentation leads to increased land surface temperatures (LST; Cai et al., 2016). In Colombia, the loss and fragmentation of natural lands coupled with the increase in impervious surfaces created unknown impacts from SUHI that warrant further exploration.

Research assessing SUHIs often incorporates remote sensing as a critical tool to determine how heat is distributed. The application of Landsat 8 imagery can facilitate the identification of extreme heat areas and the development of a vulnerability index, guiding urban planning for climate resilience (Dialesandro et al., 2021). Moreover, scientists have demonstrated the efficacy of Google Earth Engine (GEE) for processing extensive Landsat data to monitor long-term SUHI trends and advocated for the platform's use in global SUHI studies

(Ravanelli et al., 2018). Other studies used remote sensing techniques to investigate SUHI effects in Cali, with findings indicating significant temperature increases in highly urbanized areas, particularly in the southwestern and eastern parts of the city (Preciado Vargas & Aldana Olave, 2011). This DEVELOP project is built on prior research, using remote sensing tools to map SUHI intensity and analyze its spatial correlation with land use and land cover changes and socioeconomic vulnerability.

The first term of this project, conducted in spring 2024, investigated wetland declination due to urban and agricultural development in Cali from 2002 to 2023. Cali is located within the Valle del Cauca department in Colombia and is divided into an urban zone of twenty-two comunas, which are administrative districts further divided into neighborhoods, and a rural zone of fifteen corregimientos (Figure 1; Equipo del Sistema de Indicadores Sociales, Alcaldía de Santiago de Cali. (n.d.)).



Figure 1. Study Area of Cali, Colombia (left), Cali within Colombia (top right), and Colombia within South America (bottom right).

Two of Cali’s environmentally focused organizations, the Departamento Administrativo de Gestión del Medio Ambiente (DAGMA) and Fundación Dinamizadores Ambientales, partnered with the spring 2024 team to achieve their shared interest in developing community conservation initiatives and educational materials to increase citizen participation in wetland management. DAGMA is a part of Cali’s municipal government focused on environmental stewardship, while the Fundación Dinamizadores Ambientales is a Cali-based environmental justice non-profit. The spring team mapped land use and land cover change, delineated wetland extent as of 2023, and assessed the potential for wetland presence. The team found high wetland potential in the southeastern part of Cali now dedicated to agriculture, indicating that these low-elevation areas were previously wetlands. The results also suggested an extensive network of riparian wetlands in Cali. Notably, the team found that the El Pondaje and Charco Azul urban wetlands declined in size from 2002 to 2023, and that urbanized areas increased by around 2000 hectares.

The summer 2024 Cali Urban Development II team continued the partnership with DAGMA and the Fundación Dinamizadores Ambientales to identify areas most impacted by the SUHI effect from January 2013 through December 2023 and identify comunas where cooling interventions could be implemented. The team assessed the feasibility of using NASA Earth observations and remote sensing methods to investigate urban heat in Cali, while providing the partner organizations with end products that can assist their decision-

making and aid in their objective to increase community understanding of urban heat. The team aimed to identify the areas of the city that experienced the highest temperatures, evaluate how land use has changed, and locate areas experiencing high temperatures and social vulnerability.

3. Methodology

3.1 Data Acquisition

3.1.1 Urban Heat

A spring 2022 DEVELOP team created the Urban Heat Exposure Assessment Tempe (UHEAT) 2.0 tool, which uses census data, Earth observations, and other open-source data to generate maps of environmental variables such as land surface temperature (LST; Agrawal et al.). The team utilized the geoprocessing portion of code from the UHEAT tool, which relied on imagery from Landsat 8 and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) for calculated environmental variables. The team reworked the UHEAT 2.0 code by incorporating data from Landsat 7 through 9 to generate the maximum number of images per selected year (U.S. Geological Survey Earth Resources Observation and Science Center, 2022). Team members added the study area shapefile as an asset to GEE to clip each variable to this area. The team generated LST maps for five years: 2013, 2015, 2018, 2020, and 2023, and a composite of the entire study period. The team selected the years 2013, 2018, and 2023 to visualize changes over the beginning, middle, and end of the study period. Additionally, the year 2015 served as a case study due to intense weather events caused by El Niño. The team also included the year 2020 to draw meaningful connections with the socioeconomic data, much of which was from 2020. For all years, the team generated composite images using the entire year to produce the most accurate and cloud-free aggregate images possible with the available data (Table A1).

3.1.2 Landcover

For the land use/land cover map (LULC), the team used Collection 2 Level 2 data from Landsat 8 Operational Land Imager (OLI; U.S. Geological Survey Earth Resources Observation and Science Center, 2023–2024). Using the USGS Earth Explorer website, the team filtered the data to only include images from January 1, 2013 – December 31, 2023, with a cloud cover less than or equal to 47% (Table B1). It is important to note that the team increased this threshold to 56% for the 2021 image due to the unfavorable distribution of cloud cover over the study area at lower cloud cover filters. Then, team members loaded Surface Reflectance metadata and Quality Assessment (QA) bands of the filtered images into ArcGIS Pro 3.2.

3.1.3 Social Vulnerability

To analyze social vulnerability to urban heat, the team first gathered data from the partners and the “Cali en Cifras” report, which makes statistical data for Cali available at citywide and comuna-specific levels (Morales & Perilla Galvis, 2021). The metrics chosen to indicate vulnerability included census data about health, age, ethnicity, socioeconomic status, and access to both public and ecosystem services (Table C1; Alcaldía de Santiago de Cali., 2022). The team acquired all indicators on a comuna level to enable direct comparison and demonstrate how factors determining social vulnerability are distributed spatially in Cali. The team then organized the selected indicators into an Excel table for processing.

3.2 Data Processing

3.2.1 Urban Heat

The UHEAT 2.0 tool code was originally in Python syntax, so team members modified it to function within GEE’s JavaScript API. The team utilized the tool to create maps of daytime LST, which relied on Landsat data, for the years and time periods selected, using QA bits to mask out clouds. For the Landsat-derived variables, the team merged collections from Landsat 7, 8 and 9, using all the satellites available for each year. For example, in 2013, only two Landsat satellites were operational, so the team utilized data from these two satellites. In contrast, the team used data from the three satellites which were operational in 2023. The team clipped each environmental variable to the study area asset and converted all temperatures from Fahrenheit to Celsius.

3.2.2 Landcover

The team ran a raster containing the surface reflectance bands from each year individually through Esri's Landsat 8 Deep Learning Landcover Classification model using the Classify Pixels using Deep Learning geoprocessing tool (Esri, 2023). The deep learning model classified land use into 16 different categories (broken down into three water/wetland types, four urbanization levels, three forest types, two agriculture types, and four miscellaneous categories). The Extract by Mask tool limited the output to the urban region comprising the study area. Due to Cali's proximity to the equator, persistent cloud cover introduced uncertainty into the landcover classification. Therefore, team members created a cloud mask from raster functions by using the QA band to transpose bits and applying the Boolean Not function to clip from the LULC raster pixels having high confidence cloud, cloud shadow, or cirrus (Xu, 2023; Figure B1; Table B2).

3.2.3 Social Vulnerability

After acquiring demographic data, the team prepared the data for analyses in RStudio (Table C1). First, the team converted data categories listing numbers of people or households to percentages of the population. These categories included the Afro-Colombian and Indigenous populations, the populations under the age of 5 and over the age of 65, and the number of households in the two most economically disadvantaged social strata. Team members ensured that the population data used to calculate percentages for each indicator matched the year the data were collected. Second, to determine the proportions of the population that lacked access to public utilities, the team subtracted the percentages describing access to public utilities from 100. The team applied the same method to determine the percentage of the population that lacked a secondary education. Then, the team calculated sums from figures that were initially separated in Cali demographic data, such as the total number of health facilities, and the percentage of the population living with disabilities. Finally, the team calculated the area of green space per comuna in QGIS using the Intersect geoprocessing tool on a partner-provided green spaces shapefile and comuna boundary shapefiles. For each comuna, team members divided the green space by the population to produce the area of green space per capita.

3.3 Data Analysis

3.3.1 Urban Heat

The team exported TIF files of LST for each selected year to ArcGIS Pro to generate a breakdown of the mean LSTs per comuna using the zonal statistics tool. Furthermore, the team calculated the change in LST between the years 2013 to 2023. Team members then checked the validity of the Landsat-derived LST results. First, the team generated LST results for 2013 and 2023 using Aqua MODIS satellite data to visually verify if the same temperature trends were demonstrated across the period (Wan, Hook, & Hulley, 2013). The team then quantified temperature uncertainty in Cali using Landsat's quality assessment band to run an uncertainty mask. The team then used the verified results to run analyses on LST's relationship with LULC and social vulnerability.

3.3.2 Landcover

To determine the extent of urban expansion during five- and ten-year intervals, the team used the Change Detection Wizard in ArcGIS Pro with a categorical change method to detect differences in the landcover. The team compared 2013 to 2018, 2018 to 2023, 2013 to 2022, and 2013 to 2023 to assess what period experienced the greatest change, along with the overall urban increase. The outputs were symbolized by Class_To. The team used the 2013 to 2022 change results to compare with the 2013 to 2023 changes to ensure that the 2023 landcover did not contain major errors or outliers. The team used the Change Detection Wizard to quantify the difference between the two change rasters and found it to be minimal, so the 2013-2022 change raster is not included in the results.

To facilitate further comparison, the team used the Zonal Histogram tool within ArcGIS Pro to create a table containing the total count of pixels classified under each of the LULC classes per comuna. The team then calculated the number of pixels classified as low to high-intensity development as a percentage of the total number of pixels in each comuna to measure the degree of urbanization. For further comparison, the team

calculated the difference in development between the start and end years as a percentage. The team members also ran the Zonal Statistics to Table tool to calculate the average temperature each year for each type of landcover classification.

3.3.3 Social Vulnerability

The team utilized R Studio to conduct a Principal Component Analysis (PCA) to understand the relationships between the socioeconomic indicators chosen. Team members selected PCA as a relevant method to analyze socioeconomic data because of its effectiveness in reducing data with many dimensions. With 15 variables measuring Cali residents' various demographic characteristics and ability to access resources, the team wanted to determine which factors would most influence residents' vulnerability to environmental burdens: in this case the high temperatures symptomatic of the urban heat island effect. Additionally, the script within the UHEAT 2.0 code for generating PCA-based heat vulnerability, heat exposure, and heat priority scores was not adaptable for this international project because it was designed to draw from U.S. Census data. To facilitate a PCA in R Studio, team members first installed the packages "corr," "corrplot," "ggcorrplot," "FactoMineR," and "factoextra" for correlation analysis, correlation matrix creation, multivariate data analysis, and PCA output visualization. Second, the team normalized the data and ran the PCA, once with solely the 15 socioeconomic variables, and a second time with the added variable of 2023 LST. Third, team members visualized the results for each PCA as biplots and correlation matrices to explore how the variables related to one another. The team then returned to the Excel spreadsheet and aggregated each row of socioeconomic variables to calculate each comuna's social vulnerability.

4. Results & Discussion

4.1 Analysis of Results

4.1.1 Urban Heat

The team found that, when analyzing LST across the entire study period, urban heat in Cali tended to be concentrated around central comunas with the periphery experiencing lower surface temperatures, especially in the south and west (Figure 2).

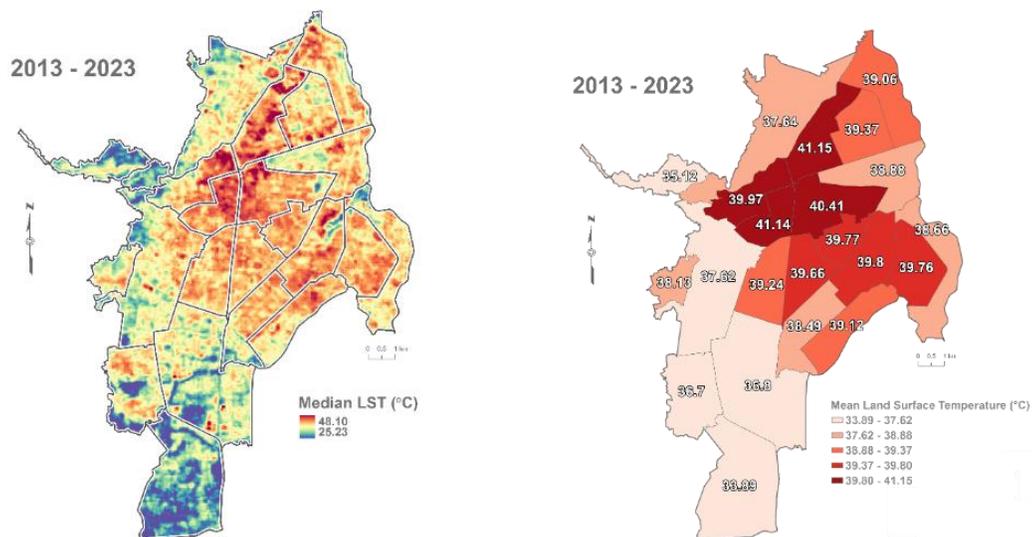


Figure 2. Median LST (°C) (left) and mean LST (right) in Cali from 2013 through 2023. Grey boundaries separate individual comunas.

Over time, the team found that LST increased from 2013 to 2015, mostly plateaued in 2018, and decreased in 2020 and 2023, with an overall decrease across the decade (Figure 3; Table 1).

Table 1
Yearly LST Ranges

Year	Minimum (°C)	Median (°C)	Maximum (°C)
2013	21.20	39.63	51.01
2015	21.47	40.77	51.94
2018	24.26	40.63	52.90
2020	21.40	36.06	48.96
2023	19.10	35.66	45.50

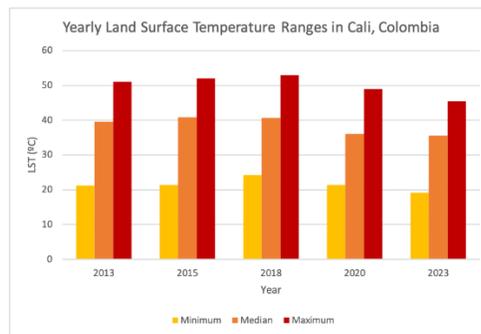


Figure 3. Yearly LST Ranges.

Despite the decrease in temperatures during the study period, the intensity of the SUHI effect that the comunas experienced increased. In a 2015 study conducted by the municipal government of Cali, including one of this project's partners, DAGMA, researchers classified SUHIs based on the deviation from the median surface temperature calculated across the study area, guided by the thresholds listed below (Table 2; Corporación Autónoma Regional del Valle del Cauca [CVC] et al., 2015). To visualize Cali's present heat islands, the team reclassified the image composite from the entire study period according to these thresholds (Figure 4).

Table 2
SUHI Classification Thresholds

SUHI Classification	Temperature (°C)
Weak	< 2
Moderate	2 - 4
Strong	4 - 6
Very Strong	> 6

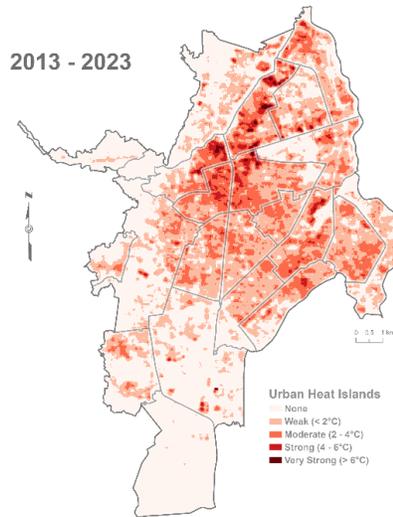


Figure 4. SUHIs in Cali from median temperatures across 2013 through 2023.

In that study, only one comuna (comuna 4, in the north central region of Cali) had a strong SUHI effect intensity while most others were considered moderate (CVC et al., 2015). However, when evaluating SUHIs at the comuna level across the entire study period, the team found that there were ten comunas with a strong intensity and eight with a very strong intensity (Table D1). As such, there are currently more residents within Cali that are affected by the SUHI effect than estimated by the 2015 study.

4.1.2 Landcover

Cali's municipality was divided into four general categories of land use: city, forests, wetlands, and agriculture which can all be further distinguished into more specialized groups using the deep learning classification model (Figure 5). Other land classification tools, such as the supervised classifier used in term I of this project, previously struggled to differentiate wetlands from forests; however, the deep learning model was successful in identifying 81% of known wetlands in 2023 (Monteiro et al., 2024). For the 19% that were unidentified, many were mistakenly classified as open water. To check the validity of the land classification tool, the team generated a confusion matrix using Sentinel-2 10-m landcover data as a reference (Table E1). The team found a kappa coefficient of 0.80 which suggests the model is moderately accurate (Mao & Wang, 2012; Table E2).

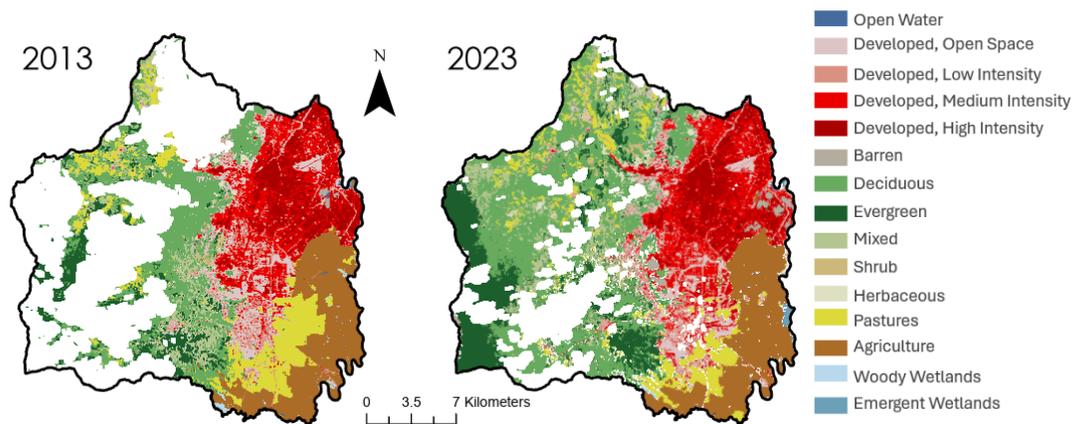


Figure 5. Landcover classification in 2013 and 2023. The white sections are due to the cloud mask.

Many landcover alterations occurred between 2013 and 2023 (Figure 6). The urban areas in Cali expanded, particularly in the southeast, and in many areas, the intensity of development increased to encompass an area

of 52.8 km². It is important to note that while development intensity increased within the city, deforestation and the rural to urban land conversion were mainly at the south end of the city. During the study period, Cali experienced a 15.6% population increase which likely led to increased development and suburban sprawl (United Nations, 2024). The change from medium to high developed intensity was the largest. Additionally, 39% of the pastures in 2013 became crops in 2023, likely due to an increased need to feed people as the population within the city grew. Similarly, 8.4 km² of deciduous forests were converted, with 34% of the forest lost becoming pastures and 31% becoming developed open space which is likely to be developed more in the future.

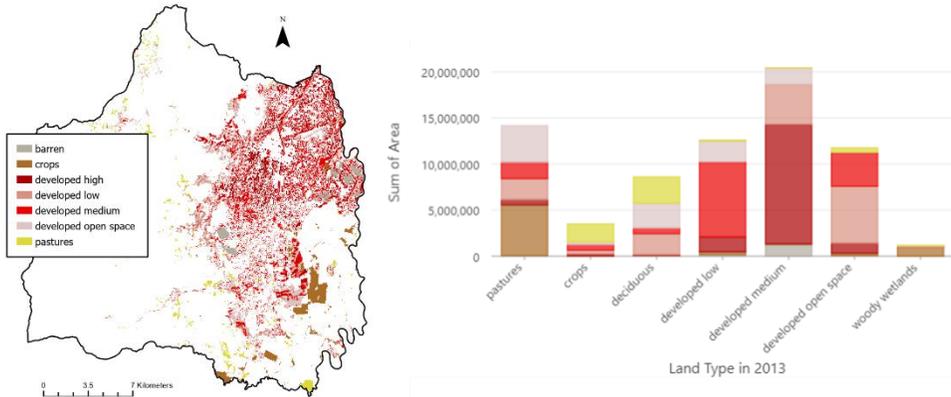


Figure 6. Human alterations to landcover from 2013 to 2023. The stacked column colors correspond to the land type classification in 2023. Area is measured in m².

Land use patterns also influence the temperature of each area. The team consolidated land classes into simplified groups and found that while the actual temperature values varied between years, the annual temperature differences between land types remained the same (Figure 7; Table E1). Team members conducted a Student’s *t*-test and found that the difference in temperature between urban land and forested land was significant ($p = 0.0013$), with urban land yielding the highest temperatures in most years. Conversely, the forested lands consistently had the lowest temperatures at approximately 31 °C on average. The wetlands also had cooler temperatures with an average of 33 °C. One explanation for the wetlands being warmer than forested areas is their size. Many of the wetlands in Cali are small and easily heated, whereas the larger patches maintain cooler temperatures better (Cai et al., 2016). It is interesting to note that the hottest land cover types, developed urban areas and croplands, are the result of human modifications to the land. The Valle de Cauca region produces several crops, such as sugarcane, that may contribute to these high temperatures (U.S. Department of Commerce, 2021).

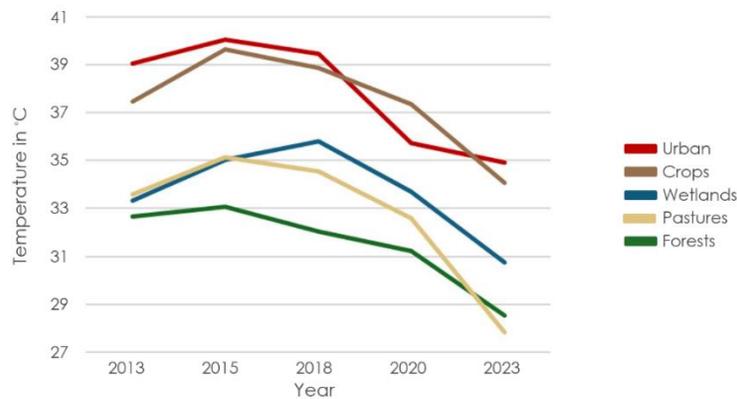


Figure 7. Mean LST of simplified land cover classes.

After running a zonal analysis, the team found that comunas in the center of the city had the highest intensity of urban development, with many having more than 97% of their area classified as developed low to high intensity. Comunas on the periphery, and especially those to the south, tended to have more sparse development (Figure 8). However, while the outskirts of the city were less developed, they experienced the largest changes in development over the study period (Figure 8).

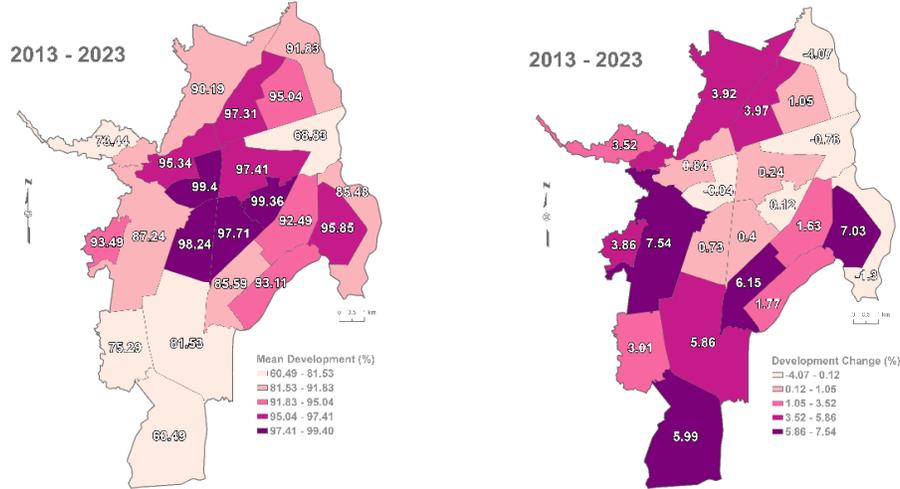


Figure 8. Urbanization at the comuna level across the study period, illustrated by mean development percentage (left) and change in development (right).

To better understand the relationship between land use changes and urban heat, the team plotted the percentage of urban development per comuna against the comuna’s mean land surface temperature and found that the two variables had a moderately strong positive correlation, with an R-squared value of 0.7 (Figure 9). Moreover, the team also visualized this relationship with the bivariate choropleth map shown below that once again showed the core-periphery divergence (Figure 9). This means that, generally, comunas that are more developed tend to have higher mean land surface temperatures. No significant correlation was found between changes to development and changes to land surface temperatures across the study period.

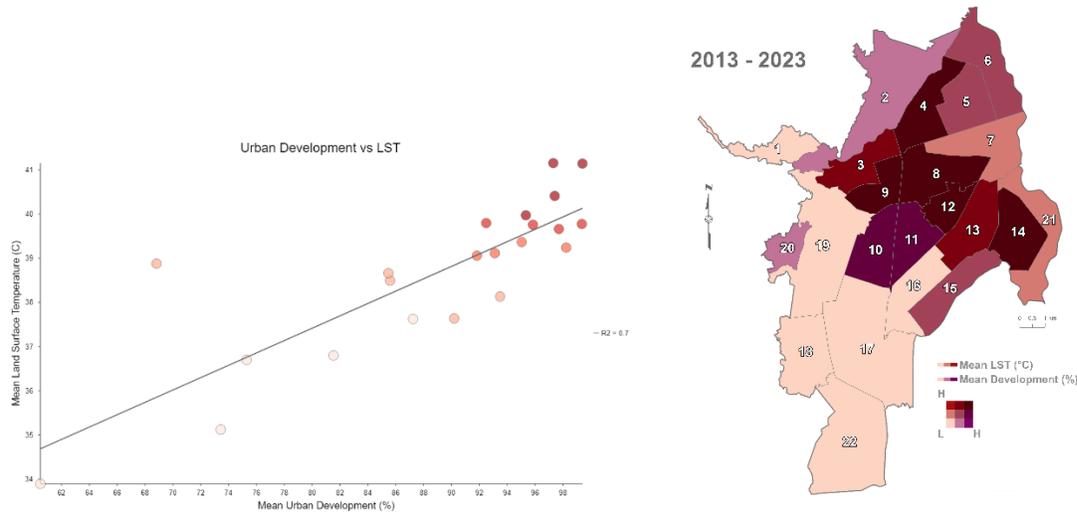


Figure 9. Visualizations of the relationship between development and land surface temperatures.

4.1.3 Social Vulnerability

By examining the correlation matrices produced with the PCA results, the team found that the socioeconomic variables most indicative of social vulnerability were the adult illiteracy rate, percentage of households in the most socioeconomically disadvantaged social stratum, percentage of households enrolled in the SISBEN program (Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales) receiving social and health benefits designed to serve low-income residents, and access to green space (Figure C1). Team members found that the areas with the hottest temperatures correlated to higher Afro-Colombian populations (correlation coefficient = 0.57) and less access to green space (correlation coefficient = -0.39; Figure C2). On a comuna level, comunas 1, 14, 15, 20, and 21, all on the outskirts of Cali, had the highest aggregate social vulnerability scores. Comparing social vulnerability scores with land surface temperature to determine heat risk demonstrated that comunas 13, 14, 15, and 21, on the eastern side of the city, are most at-risk, experiencing both high temperatures and high social vulnerability (Figure 10). The team also created a map displaying urban green spaces and wetlands overlaid over an average land surface temperature raster along with a line graph comparing green space area with LST per comuna, revealing that although there is much temperature variation surrounding urban green spaces throughout the study area, a comuna-based analysis shows an inverse relationship between green space and LST, with higher temperatures in comunas with less green space and lower temperatures experienced in comunas with more green space (Figure C3).

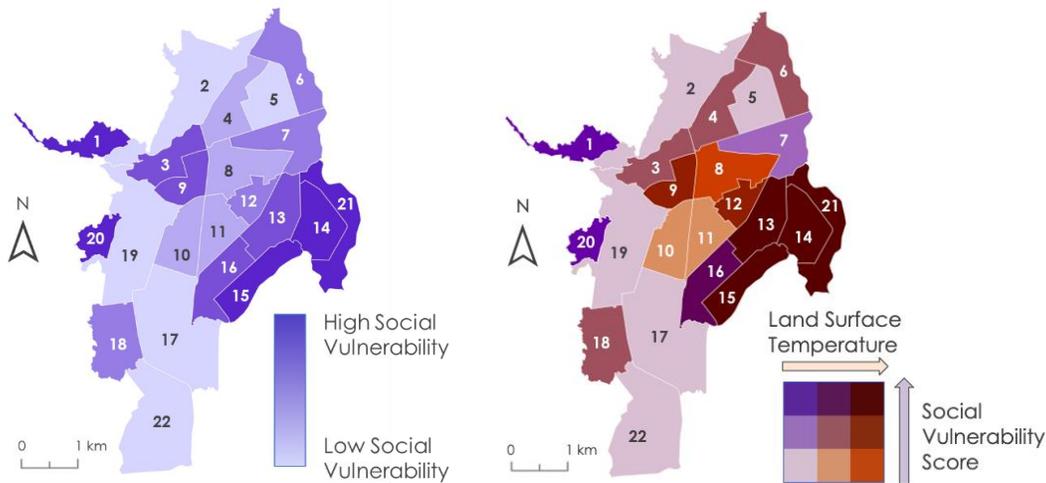


Figure 10. Social vulnerability by comuna (*left*) and heat risk assessment calculated by comparing aggregate social vulnerability scores with land surface temperature (*right*).

4.2 Errors & Uncertainties

The most prominent limitation throughout the project was cloud cover, which posed a challenge to landcover classification and LST outputs. Cloud cover is a persistent obstacle to using Earth observations in Cali because of the region’s proximity to the equator and tropical climate. In fact, all key years in the study period had a mean cloud cover percentage above 60% (Table 3). Originally, the team planned to use only Landsat 8 to generate LST data, however, the clouds limited the number of satellite images available to use in our analysis to less than 20 per year. The team adapted by incorporating all the current operational Landsat satellites in their analysis. Still, the number of images generated for years relying on 2-3 Landsat satellites was between 34-63 images. The team employed a cloud mask which was able to remove the clouds but left large areas of no data in some datasets, which heavily affected the landcover maps.

Table 3

Cloud cover percents across the study period

Year	Minimum (%)	Mean (%)	Maximum (%)
2013	30	73	96

2015	16	68	92
2018	30	69	95
2020	22	71	94
2023	16	64	95

The deep learning model for landcover classification was generally accurate at differentiating land types. However, the model frequently misidentified urban land as barren and, occasionally, wetlands as open water. The model was trained using the National Land Cover Database, which is run by the USGS and based on landcover imagery in the United States. The ecosystems and construction materials in Colombia are likely different and could impact the model’s accuracy (Caro, 2012).

Finally, some of the data used for the social vulnerability analysis was only available for 2013, the very first year of the study period, potentially making some of the figures calculated for socioeconomic variables outdated while other data were more recent (Table C1). Due to this limitation, the social vulnerability PCA and maps may not paint the most accurate picture of social vulnerability and heat risk in Cali at present day.

4.3 Feasibility & Partner Implementation

The feasibility of this study relies heavily on the team’s ability to meet the image requirements of 35 images or more used for accurate calculations, such as for land surface temperature. Acquiring enough Landsat-derived images to accurately depict changes in land surface temperature over time will prove difficult for years prior to the launch of Landsat 9 in 2021. Moreover, even in years with three satellites, the team found that, on a pixel-by-pixel basis, most pixels had less than 30 data inputs for calculation after applying the cloud mask (Table A2). To mitigate these limitations, the partners can utilize enhanced cloud masking techniques to improve cloud detection and retain more usable data. Another major limiting factor is the uncertainty in the surface temperature bands from Landsat ETM+/TIRS/TIRS-2. Previous research made use of a quality assessment mask that excluded pixels where the uncertainty was above 4 °C (Nuñez et al., 2023). However, close to 60% of the study area was above this threshold with an uncertainty ranging from 3.5 - 5.5 °C (Figure 11).

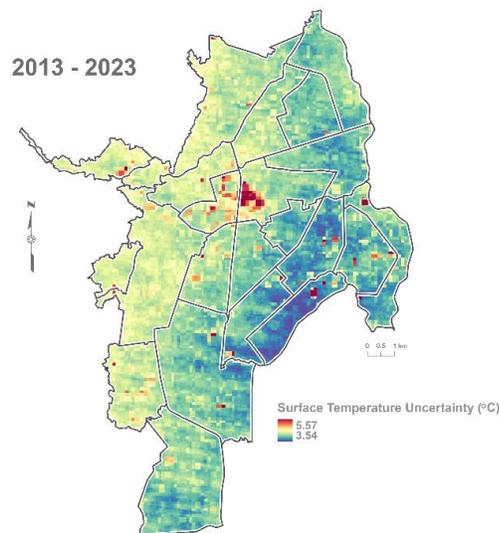


Figure 11. Land surface temperature uncertainty in Cali from median temperatures across 2013 through 2023.

In fact, the earlier years in the study period tended to have higher uncertainties than later years which could have impacted the results (Table 4).

Table 4

Land surface temperature quality assessment (ST_{QA}) values per year

Year	Minimum (°C)	Mean (°C)	Maximum (°C)
2013	3.37	4.17	6.30
2015	3.16	4.02	6.02
2018	3.60	4.23	5.81
2020	3.28	3.97	5.46
2023	3.01	3.54	4.84

5. Conclusions

The team found that temperature varies significantly between urban and forested lands, which shows that landcover strongly impacts surface temperatures. This finding indicates that the proportion of developed land area per comuna significantly increases the temperatures residents experience. The high temperatures in crop lands signify that the types of vegetation grown are important, with trees as the most effective at reducing heat. The temperatures were hottest in the center of the city, where there is less green space and less vegetation cooling. The team also found less green space in areas of higher social vulnerability, on the outskirts of the city where the greatest rates of development during the study period occurred. Specifically, the team identified Comunas 13, 14, 15 and 21 as the most heat-vulnerable and likely to receive prioritized implementation of green spaces and health facilities. These takeaways support the partner organizations by identifying communities at the greatest risk to heat, along with factors that will influence future risk. The relationship between urban heat, landcover use, and the comunas' social demographics will be useful for future city planning.

Overall, this project demonstrated that using remote sensing methods to understand the impacts of land use change, urban heat islands (or SUHI?), and social vulnerability in Cali is feasible for future planning and implementation, including future urban heat studies. Despite the difficulty in obtaining suitable images due to persistent cloud cover, improvements were made upon past studies of urban heat in Cali, the most notable of which depended on a single satellite image from 2015 to identify urban heat islands (CVC et al., 2015; Table 1), while this project utilized an average of 43.2 images per year for five select years throughout the study period. Furthermore, the Esri deep learning model for landcover proved effective for the study area, enabling a detailed examination of development intensity and land cover change which could be combined with temperature data for a comprehensive analysis. Compared to term I of this project, which used a supervised classifier dependent on team members selecting training points for each land cover type and yielded a kappa coefficient of 0.20, indicating low accuracy, the deep learning model yielded a kappa coefficient of 0.80. Employing a deep learning model could save partner organizations valuable time in comparison to the current frequently-used strategy of digitizing aerial and satellite images to create land use and land cover maps. Finally, the availability of comuna-specific socioeconomic datasets on Cali municipal websites allowed for a comuna-level understanding of how social factors increase heat risk for communities in different areas of the city. As demographic data from Cali's 2018 census and upcoming 2024 census are released and the municipality updates its online resources accordingly, partners could conduct another social vulnerability analyses that incorporates more recent data.

6. Acknowledgements

The team would like to thank:

Partners: Fundación Dinamizadores Ambientales and Departamento Administrativo de Gestión del Medio Ambiente (DAGMA)

Advisors: Dr. Morgan Gilmour (NASA – ARC), Dr. Alexandra Christensen (NASA – JPL), Dr. Juan Torres-Pérez (NASA – ARC), Dr. Xia Cai (NASA – LaRC), Lisa Tanh (Esri)

Fellow: Maya Hall (DEVELOP ARC)

Lead: Lauren Webster (DEVELOP ARC)

Special thanks to the Spring 2024 Cali Urban Development I Team for their previous work on this project and the UHEAT Urban Development (Spring 2022) DEVELOP team for developing the UHEAT 2.0 Tool.

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

Blue Space – visible water features including wetlands, pools, and rivers

Composite Images – an image made from combining several other images

Comunas – city divisions in Cali, similar to districts in the U.S.

Correlation Matrix – a statistical method to evaluate the relationship between variables

Earth Observations – satellites and sensors that collect information about the Earth’s physical, chemical, and biological systems over space and time

Enhanced Thematic Mapper Plus (ETM+) – the sensor on Landsat 7

Google Earth Engine (GEE) – a cloud-based platform used to access and analyze satellite imagery

Green Space – areas of grass, trees, or vegetation used for recreational purposes in urban environments

Heat Disparity – unequal distribution of heat in urban areas and neighborhoods which causes disproportionate impacts to residents

Kappa Coefficient – a statistic that represents the amount of association between continuous variables

Land Surface Temperature (LST) – describes how hot a material of the earth’s surface would feel to touch at a given location

Land use/Landcover (LULC) – describes the type of land and its purpose

Moderate Resolution Imaging Spectroradiometer (MODIS) – the sensor on the satellite Aqua

Operational Land Imager (OLI) – a sensor on satellites Landsat 8 and 9

Principal Component Analysis (PCA) – a statistical analysis method used to reduce the dimensionality of data with many variables in order to understand which few components explain the variation among most of the data

Quality Assessment Band (QA) – satellite data that allows users to filter the pixels to remove clouds

Thermal Infrared Sensor (TIRS) – a sensor on satellites Landsat 8 and 9

Urban Heat Island Effect (UHI) – an environmental phenomenon where urban areas experience higher temperatures than their rural surroundings due to high levels of urbanization

8. References

- Agrawal, A., Sik Cho, M., Farid, Z., & Machuca, V. (2022). *Increasing capabilities and updating the Urban Heat Exposure Assessment for Tempe (UHEAT) Tool*. NASA DEVELOP National Program, Virtual – Pop-Up Project. <https://ntrs.nasa.gov/citations/20220004583>
- Alcaldía de Santiago de Cali. (2022). *Informacion Censo 2018 Barrio-comuna.xls* [Data set]. Demografía de Santiago de Cali. Retrieved from <https://www.cali.gov.co/planeacion/publicaciones/144497/demografia-de-santiago-de-cali/genPagdoc2187=1>
- Cai, Y., Zhang, H., Zheng, P., & Pan, W. (2016). Quantifying the impact of land use/land cover changes on the urban heat island: A case study of the natural wetland distribution area of Fuzhou City, China. *Wetlands*, 36(2), 285–298. <https://doi.org/10.1007/s13157-016-0738-7>
- Caro, S. (2012). Asphalt specifications and research challenges in Colombia. *Asphalt Paving Technology-Proceedings Association of Asphalt Technologists*, 81, 797.
- Corporación Autónoma Regional del Valle del Cauca, Centro Internacional de Agricultura Tropical, & Departamento Administrativo de Gestión del Medio Ambiente. (2015). *Identificación de Zonas y Formulación de Propuestas para el Tratamiento de Islas de Calor: Convenio CVC–CIAT–DAGMA No. 110 de 2015*. <https://cgspace.cgiar.org/items/7fd0a6ced-b7ba-4f4e-a2ad-63610a250cf3>
- Dialesandro, J., Kruskopf, M., Marvin, C., & Vargas, M. (2023). *San Diego urban development: Utilizing NASA Earth observations to identify drivers of extreme urban heat and generate a high-resolution vulnerability index for urban planning and climate resiliency in San Diego, California*. NASA DEVELOP National Program, Arizona – Tempe. <https://ntrs.nasa.gov/citations/20210015142>
- Esri. (2023). Deep learning model to perform land cover classification on Landsat 8 imagery. <https://www.arcgis.com/home/item.html?id=e732ee81a9c14c238a14df554a8e3225>
- Equipo del Sistema de Indicadores Sociales, Alcaldía de Santiago de Cali. (n.d.). *Consulta de perfiles por comunas*. Sistema de Indicadores Sociales. <https://indicadores.cali.gov.co/consulta-indicadores/dimensiones-sis-comunas/perfiles>
- Li, X., Stringer, L. C., Chapman, S., & Dallimer, M. (2021). How urbanization alters the intensity of the urban heat island in a tropical African city. *PLoS One*, 16(7). <https://doi.org/10.1371/journal.pone.0254371>
- Monteiro, T., Tesfayi, N., Davidson-Gómez, G., & Trejo, R. (2024, March 21). *Cali Urban Development: Using NASA earth observations to assess wetlands and land reclamation in Cali, Colombia*. [PowerPoint slides]. NTRS – NASA Technical Reports Server. <https://ntrs.nasa.gov/citations/20240003207>
- Mao, W., & Wang, F.-Y. (2012). New advances in intelligence and security informatics. In Elsevier eBooks. <https://doi.org/10.1016/c2011-0-07828-6>
- Morales, G. E., & Perilla Galvis, D. M. (2021). *Cali en Cifras 2021*. Departamento Administrativo de Planeación, Santiago de Cali, Colombia. <https://www.cali.gov.co/documentos/1705/documentos-de-cali-en-cifras/u>
- Núñez, Y., Hoyos, N., & Arellana, J. (2023). High land surface temperatures (LSTs) disproportionately affect vulnerable socioeconomic groups in Barranquilla, Colombia. *Urban Climate*, 52, <https://doi.org/10.1016/j.uclim.2023.101757>

- Ocampo-Marulanda, C., Carvajal-Escobar, Y., Perafán-Cabrera, A., & Restrepo-Jiménez, L.M. Desiccation of Wetlands and Their Influence on the Regional Climate. Case Study: Ciénaga de Aguablanca, Cali, Colombia. *Tropical Conservation Science*. 2021;14. doi:10.1177/19400829211007075
- Preciado Vargas, M., & Aldana Olave, A. (2011). Análisis de presencia de islas de calor en Santiago de Cali empleando técnicas de teledetección. *Ventana Informatica*, 24. <https://doi.org/10.30554/ventanainform.24.162.2011>
- Ravanelli, R., Nascetti, A., Cirigliano, R. V., Di Rico, C., Leuzzi, G., Monti, P., & Crespi, M. (2018). Monitoring the impact of land cover change on surface urban heat island through Google Earth Engine: Proposal of a global methodology, first applications and problems. *Remote Sensing*, 10(9), Article 1488. <https://doi.org/10.3390/rs10091488>
- Salazar Tamayo, M. M., & Julio Estrada, J. D. (2022). Planning gaps: Unexpected urban expansion in five Colombian metropolitan areas. *Buildings and Cities*, 3(1), 725–744. <https://doi.org/10.5334/bc.2400>
- United Nations. (2024). World Population Prospects 2024. Department of Economic and Social Affairs Population Division. <https://population.un.org/wpp/>
- U.S. Department of Commerce. (2021). Santiago de Cali Profile. International Trade Administration. https://www.trade.gov/sites/default/files/2021-03/Cali%20Profile_0.pdf
- U.S. Geological Survey Earth Resources Observation and Science Center. (1999-2022). *Landsat 7 ETM+ Collection 2 Level-2 Science Products* [Data set]. U.S. Geological Survey. <https://doi.org/10.5066/F7Q52MNK>
- U.S. Geological Survey Earth Resources Observation and Science Center. (2021-2023). *Landsat 8-9 OLI/TIRS Collection 2 Level-2 Science Products* [Data set]. U.S. Geological Survey. <https://doi.org/10.5066/p9ogbgm6>
- U.S. Geological Survey. (2024). *Landsat 8-9 Collection 2 (C2) Level 2 Science Product (L2SP) Guide* (Report No. LSDS-1619 Version 6.0). Author.
- Wan, Z., Hook, S., Hulley, G. (2013-2023). *MYD11A1 MODIS/Aqua Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006* [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center. <https://doi.org/10.5067/MODIS/MYD11A1.006>
- Xu, H. (2023, February 5). *Create a cloud-free image composite from satellite imagery*. Esri. <https://www.esri.com/arcgis-blog/products/arcgis-pro/imagery/cloud-free-image-composite/>

9. Appendices

Appendix A: LST calculation limitations

Table A1

Image availability per year

Year	Images from Landsat 7	Images from Landsat 8	Images from Landsat 9	Total Images
2013	19	15	N/A	34
2015	18	22	N/A	40
2018	20	20	N/A	40
2020	20	19	N/A	39
2023	23	19	21	63
Composite	199	210	39	448

**N/A entries are a result of Landsat 9 not having been launched*

Table A2

Ranges of image data inputs left, for LST analysis, after applying cloud mask

Year	Smallest number of data inputs	Mean number of data inputs	Largest number of data inputs
2013	6	11	17
2015	7	20	22
2018	6	14	21
2020	4	19	25
2023	16	24	30
All	87	200	215

Appendix B: LULC Image Cloud Cover

Table B1

Image inputs for the LULC deep learning model

Year	Product ID	Date Acquired	Cloud Cover %
2013	LC08_L2SP_009058_20130902_20200913_02_T1	9/2/2013	44.18
2015	LC08_L2SP_009058_20151229_20200908_02_T1	12/29/2015	41.44
2018	LC08_L2SP_009058_20180511_20200901_02_T1	5/11/2018	37
2020	LC08_L2SP_009058_20200109_20200823_02_T1	1/9/2020	42.9
2023	LC08_L2SP_009058_20230829_20230906_02_T1	8/29/2022	31.17

Table B2

Transpose bits raster function input (U.S. Geological Survey, 2024)

Output Bit*	Input Bit	Bit Value – 0**	Bit Value - 1
0	0	Image data	Fill data
1	1	No cloud dilation	Cloud dilation
2	2	None to low confidence cirrus	High confidence cirrus
3	3	None to low confidence cloud	High confidence cloud
4	4	None to low confidence cloud shadow	High confidence cloud shadow

*Bits numbered 5-15 set to constant fill value of 0

**Hence, a cloud free pixel will have all 16- bit values equal 0

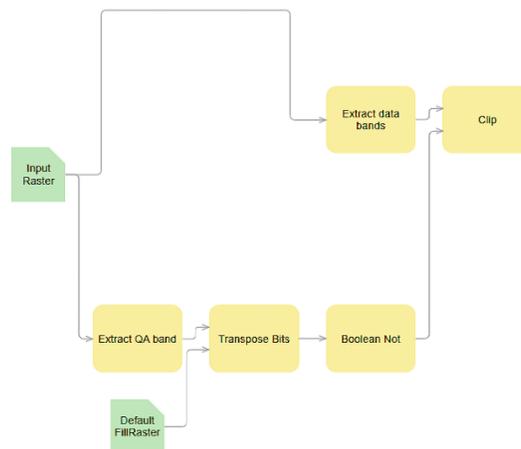


Figure B1. Cloud mask workflow within ArcGIS Pro's raster function editor (Xu, 2023).

Appendix C: *Social Vulnerability Data & Results*

Table C1

Social vulnerability indicators and Cali demographic data sources

Indicator	Year	Source	Methodology (if applicable)
Percentage of households in social stratum 1*	2020	<i>Cali en Cifras 2021</i> p.180	--
Percentage of households in social stratum 2*	2020	<i>Cali en Cifras 2021</i> p.180	--
Percentage of the population lacking a basic secondary education	2013	<i>Consulta de perfiles por comunas</i> webpage - PDFs downloaded for each comuna	100 - comuna "Tasa de escolaridad neta básica secundaria" (TENS figure). e.g. for Comuna 1, 100 - 64.0 = 36% of the population does not have a basic secondary education
Adult illiteracy rate	2013	<i>Consulta de perfiles por comunas</i> webpage - PDFs downloaded for each comuna	--
Percentage of the population with disabilities	2013	<i>Consulta de perfiles por comunas</i> webpage - PDFs downloaded for each comuna	% "blindness" + % "deafness" + % "muteness" figures from comuna profiles
Percentage of households on low-income healthcare benefits	2013	<i>Consulta de perfiles por comunas</i> webpage - PDFs downloaded for each comuna	--
Total number of health facilities	2021	<i>Cali en Cifras 2021</i> p.171-173	#"Puestos de salud" + #"Puestos de salud y CAB" + #"Centro hospital" + #"Hospitales y clinicas"
Households without electricity, water, and sewage coverage	2013	<i>Consulta de perfiles por comunas</i> webpage - PDFs downloaded for each comuna	100 - energy/water/sewage coverage percentage = % not covered
Households without natural gas coverage	2013	<i>Consulta de perfiles por comunas</i> webpage - PDFs downloaded for each comuna	100 - natural gas percentage = % not covered
Households without garbage collection coverage	2013	<i>Consulta de perfiles por comunas</i> webpage - PDFs downloaded for each comuna	100 - garbage collection percentage = % not covered
Percentage of the population that is Afro-Colombian	2018	"Barrio Etnico 2018" tab of <i>Informacion Censo</i>	--

		2018 Barrio-comuna data set	
Percentage of the population that is Indigenous	2018	"Barrio Etnico 2018" tab of <i>Informacion Censo 2018 Barrio-comuna</i> data set	--
Percentage of the population under the age of 5	2018	"Barrio Quinquenales 2018" tab of <i>Informacion Censo 2018 Barrio-comuna</i> data set	--
Percentage of the population age 65 and over	2018	"Barrio Quinquenales 2018" tab of <i>Informacion Censo 2018 Barrio-comuna</i> data set	--
Access to green space (square meters per capita)	2022	"Zonas Verdes" partner-provided shapefile and "Cali en Cifras" 2022 population data	Area of green space within each comuna divided by population of the comuna

*The two social strata refer to how households are classified by the Colombian government in order of access to resources and as a tool to determine which households would benefit most from public services. They are a general approximation of the hierarchy of poverty and wealth present among Colombian residents. For the purposes of this study’s social vulnerability analysis, we chose to use data from the two most disadvantaged strata, categorized in Colombia as “low-low” (stratum 1) and “low” (stratum 2) status.

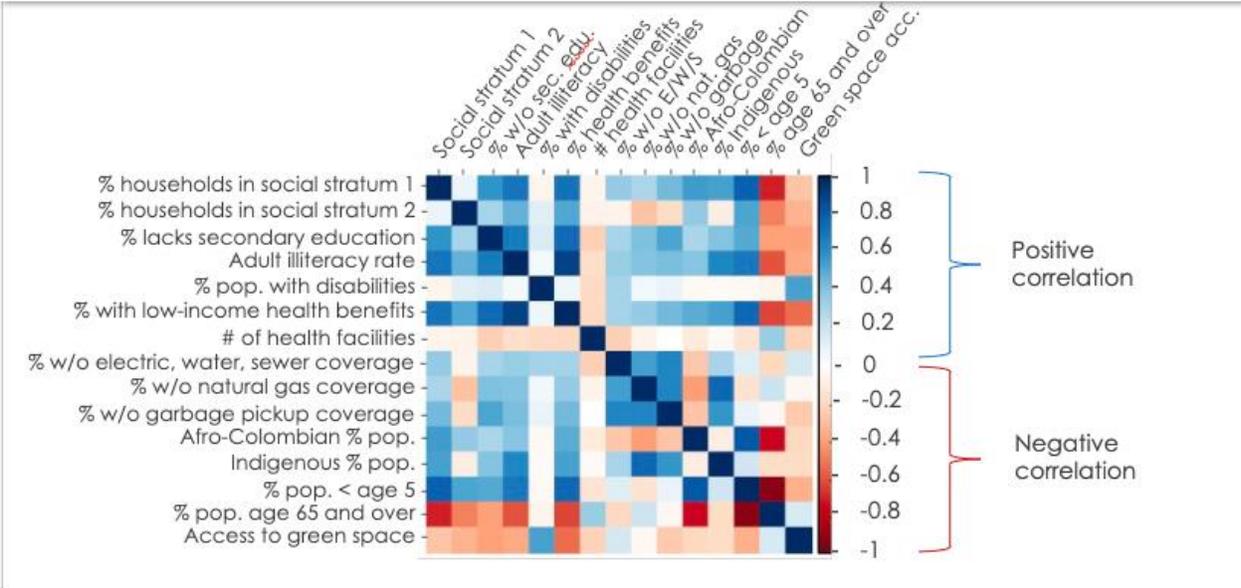


Figure C1. Correlation matrix showing the strength of correlation between the 15 selected socioeconomic variables.

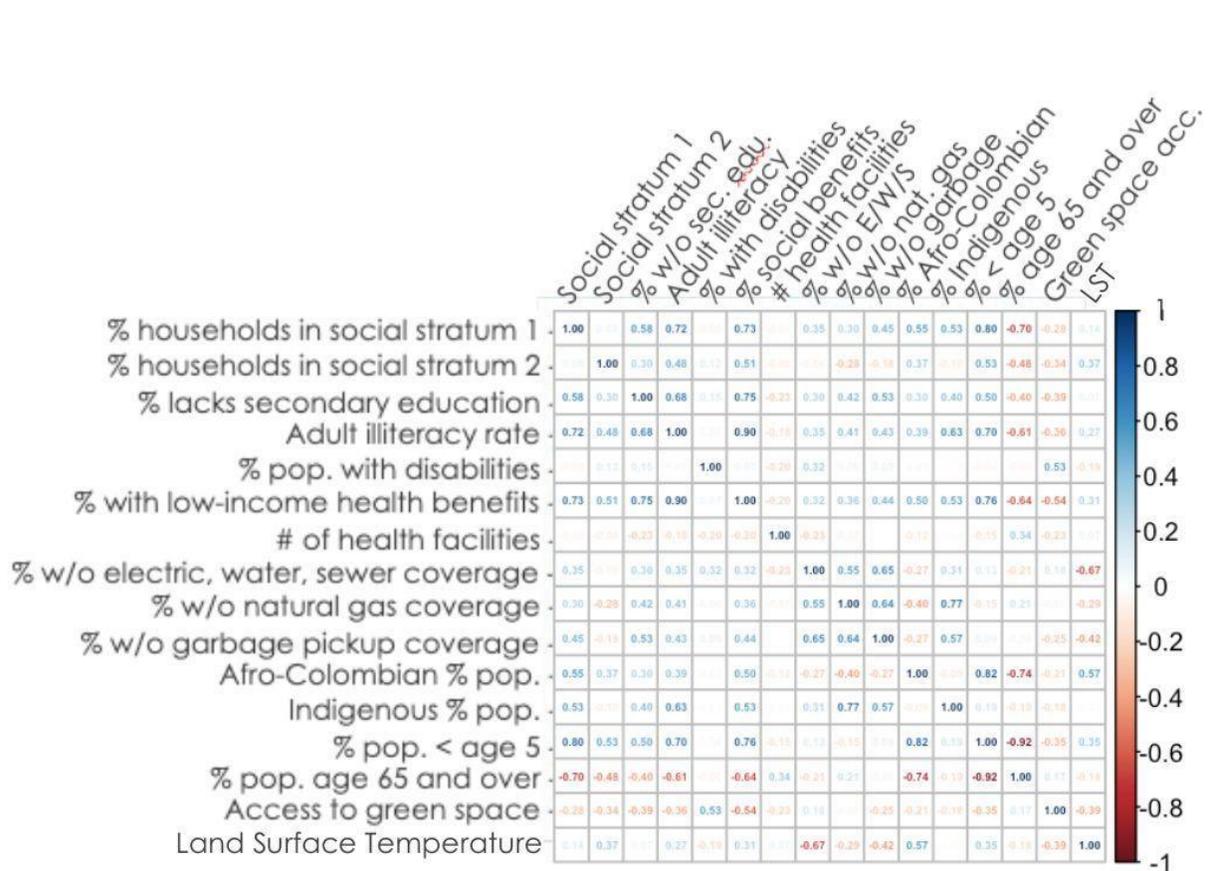


Figure C2. Correlation matrix with correlation coefficients showing the strength of correlation between the 15 selected socioeconomic variables and land surface temperature.

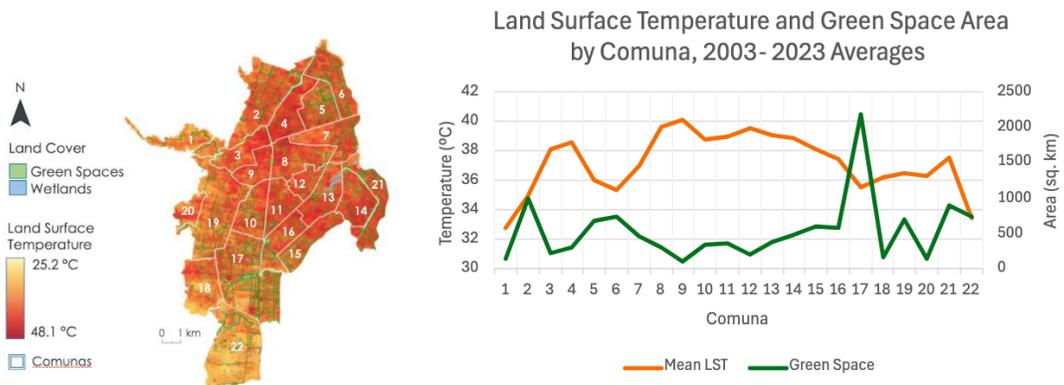


Figure C3. Map of land surface temperature averages throughout the entire study period, overlaid with partner-provided shapefiles of urban green spaces and wetlands (left), and line graph depicting the relationship between average LST over the entire study period and green space area per comuna (right).

Appendix D: *SUHI by Comuna*

Table D1

SUHI Classification by Comuna across the Study Period

Name	Maximum LST (°C)	Difference (39.04°C)	SUHI Classification
Comuna 1	40.58	1.54	Weak
Comuna 2	45.23	6.19	Very Strong
Comuna 3	47.15	8.11	Very Strong
Comuna 4	48.10	9.06	Very Strong
Comuna 5	46.04	7	Very Strong
Comuna 6	45.03	5.99	Strong
Comuna 7	45.20	6.16	Very Strong
Comuna 8	46.67	7.63	Very Strong
Comuna 9	44.79	5.75	Strong
Comuna 10	43.48	4.44	Strong
Comuna 11	43.13	4.09	Strong
Comuna 12	42.74	3.7	Moderate
Comuna 13	45.53	6.49	Very Strong
Comuna 14	43.74	4.7	Strong
Comuna 15	44.24	5.2	Strong
Comuna 16	43.14	4.1	Strong
Comuna 17	46.76	7.72	Very Strong
Comuna 18	43.11	4.07	Strong
Comuna 19	44.64	5.6	Strong
Comuna 20	42.29	3.25	Moderate
Comuna 21	43.27	4.23	Strong
Comuna 22	41.26	2.22	Moderate

Appendix E: *Confusion Matrix*

Table E1

Reclassification of Landsat and Sentinel LULC data

Landsat 8 Classes	Sentinel-2 Classes	Reclassified Value
Deciduous Forest	Trees	1 - Forest
Evergreen Forest		
Mixed Forest		
Open Water	Water	2 - Wetlands
Woody Wetlands	Flooded Vegetation	
Emergent Herbaceous Wetlands		
Cultivated Crops	Crops	3 - Crops
Developed, Open Space	Built	4 - Urban
Developed, Low Intensity		
Developed, Medium Intensity		
Developed, High Intensity		
Barren Land*		
Hay/Pasture	Rangeland	5 - Pastures
Shrub/Scrub		
Herbaceous		

*Barren land included under “Urban” due to the model’s misclassification of developed land

Table E2

Confusion Matrix for LULC Deep Learning Classification Tool

Class	1	2	3	4	5	Total	U_Accuracy*	Kappa
1 - Forest	167	0	3	8	10	188	0.89	0
2 - Wetlands	0	6	1	0	0	7	0.86	0
3 - Crops	1	0	52	1	0	54	0.96	0
4 - Urban	3	0	4	127	3	137	0.93	0
5 - Pastures	14	0	7	4	15	40	0.38	0
Total	185	6	67	140	28	426	0	0
P_Accuracy*	0.90	1	0.78	0.91	0.54	0	0.86	0
Kappa	0	0	0	0	0	0	0	0.80

*U_Accuracy or "User's accuracy" shows what percentage of pixels were classified accurately, but also includes false positives, also known as type I error. Meanwhile, P_Accuracy or "Producer's accuracy" shows what percentage of pixels were classified correctly, including false negatives, also known as type II error.