The Impact of Tree Cover on the Urban Heat Islands Effect

Abstract

An analysis of the relationship between land surface temperature and tree cover for the City of San Jose, Santa Clara County, USA. Land surface temperatures are derived from Landsat 8 imagery. Tree cover is obtained from the Multi-Resolution Land Characteristics Consortium. Presents results in maps, charts, and tables. Draws conclusions, identifies limitation of this study, and suggests areas for further study.

Christopher Prendergast

December 12th, 2023

Table of Contents

The Impact of Tree Cover

on the Urban Heat Islands Effect

1 Project Description

The goal of this project is to explore the urban heat island effect by analyzing the relationship between tree cover and temperature. This project looks at the urban areas of the City of San Jose, California.

1.1 Null Hypothesis

The null hypothesis is that the pattern of temperature is unrelated to tree cover and could have occurred at random.

2 Analysis Steps

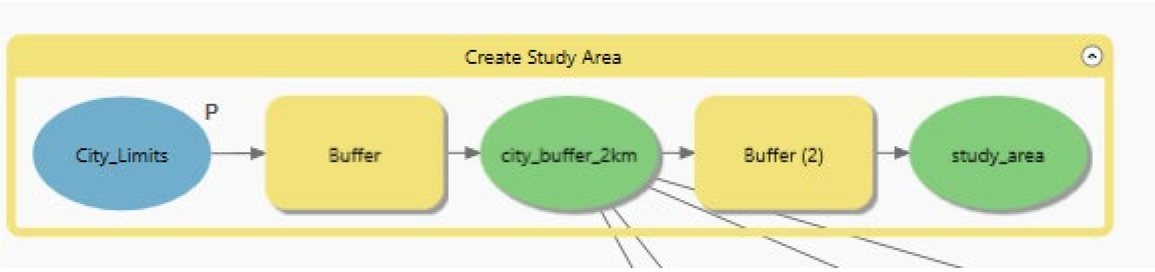
Document and present the analysis steps in model builder. Break the problem down into solvable components that can be modeled using model builder. Quantify and evaluate the spatial questions. Use model builder to document required functions. Include screen capture of model builder.

2.1 Step 1 – Prepare raster datasets

The steps to download and prepare the temperature dataset from Landsat imagery are detailed in Appendix A. The tree cover dataset was obtained from the ESRI Living Atlas.

2.2 Step 2 – Prepare the study area polygon

A 2 km buffer is added around the City of San Jose limits to eliminate edge effects in the later processing. The results are subsequently clipped to the actual city limits in a later step.

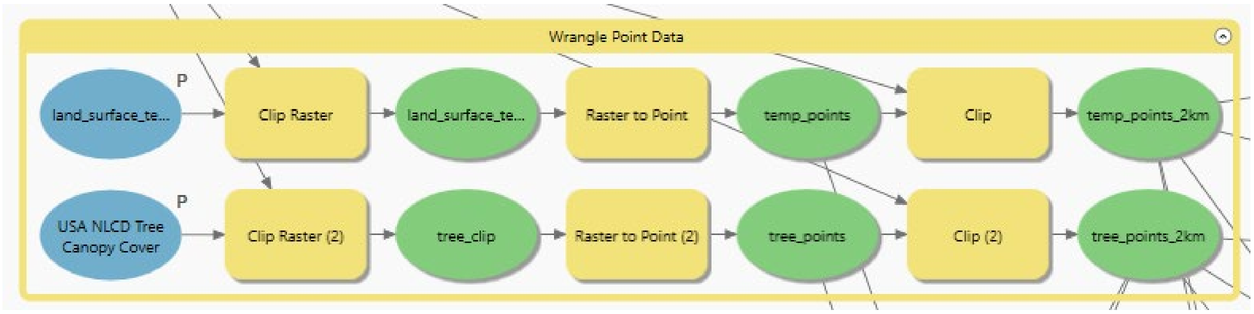


2.3 Step 3 – Wrangle the point data

This step transforms the raster data into a format that can be statistically analyzed by tools which require vector data. It also selects the subset of data points within a 2km buffer around the study area to focus and streamline the analysis.

- Clip the two raster datasets for tree cover and temperature to the extent of the 2 km buffer surrounding the City of San Jose limits.
- Extract point data from both raster datasets. Run the *Raster to Point (Conversion)* tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/conversion/raster-to-point.htm

- Clip the point data to using the polygons for the 2km buffer surrounding the City of San Jose limits.



2.4 Step 4 – Look for spatial autocorrelation.

This step looks for spatial autocorrelation in the two input datasets. Given a set of features and an associated attribute, this tool evaluates whether the pattern expressed is clustered, dispersed, or random. When the z-score or p-value indicates statistical significance, a positive Moran's I index value indicates tendency toward clustering, while a negative Moran's I index value indicates tendency toward dispersion.

- Run the *Spatial Autocorrelation (Global Moran's I) (Spatial Statistics)* tool.
- https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/spatial-autocorrelation.htm

2.5 Step 5 – Optimized hotspot, and outlier analyses

This step looks for statistically significant hot spots and outliers in the two input datasets. Both tools evaluate the characteristics of the input feature class to produce optimal results.

The optimized hot spot analysis tool creates a map of statistically significant hot and cold spots using the Getis-Ord Gi* statistic.

The optimized outlier analysis tool creates a map of statistically significant hot spots, cold spots, and spatial outliers using the Anselin Local Moran's I statistic.

- Run the *Optimized Hot Spot Analysis (Spatial Statistics)* tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/optimized-hot-spot-analysis.htm
- Run the *Optimized Outlier Analysis (Spatial Statistics)* tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/optimizedoutlieranalysis.htm

2.6 Step 6 – Hotspot (Getis-Ord Gi*), and cluster and outlier (Anselin Local Moran’s I) analyses

This step looks for hot spots, clusters, and outliers using a defined neighborhood around each datapoint using the distance specified.

The hot spot analysis tool identifies statistically significant hot spots and cold spots using the Getis-Ord Gi* statistic. The z-scores and p-values are measures of statistical significance that tell you whether or not to reject the null hypothesis, feature by feature. In effect, they indicate whether the observed spatial clustering of high or low values is more pronounced than one would expect in a random distribution of those same values.

The cluster and outlier analysis tool identifies statistically significant hot spots, cold spots, and spatial outliers using the Anselin Local Moran's I statistic. The z-scores and p-values are measures of statistical significance which tell you whether or not to reject the null hypothesis, feature by feature. In effect, they indicate whether the apparent similarity (a spatial clustering of either high or low values) or dissimilarity (a spatial outlier) is more pronounced than one would expect in a random distribution.

- Run the *Hot Spot Analysis (Getis-Ord Gi*) (Spatial Statistics)* tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/hot-spot-analysis.htm
- Run the *Cluster and Outlier Analysis (Anselin Local Moran's I) (Spatial Statistics)* tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/cluster-and-outlier-analysis-anselin-local-moran-s.htm

2.7 Step 7 – Generalized Linear Regression

This step performs a generalized linear regression to generate predictions or to model a dependent variable in terms of its relationship to a set of explanatory variables. In this analysis the temperature is the dependent variable, and the tree cover is the explanatory variable. This tool is used to fit a continuous ordinary least square (OLS) model.

- Perform a spatial join between the tree cover and temperature point feature datasets to generate a single dataset with both tree cover and temperature attributes.
- Run the *Generalized Linear Regression (Spatial Statistics)* tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/generalized-linear-regression.htm

2.8 Step 8 – Geographically Weighted Regression

This step attempts to run a geographically weighted regression to find a better fit compared to the ordinary least squares regression performed by the generalized linear regression tool above. It does this by including the location in the regression analysis, performing a local form of regression used to model spatially varying relationships. The GWR tool provides a local model of the variable or process you are trying to understand or predict by fitting a regression equation to every feature in the dataset. The GWR tool constructs these separate equations by incorporating the dependent and explanatory variables of features within the neighborhood of each target feature.

- Run the *Geographically Weighted Regression (GWR) (Spatial Statistics)* tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/geographically-weighted-regression.htm

2.9 Step 9 – Compare hotspots

This step compares the hotspots found for tree cover and temperature within the study area to identify any relationship between their patterns.

Compares two hot spot analysis result layers and measures their similarity and association.

The similarity and association between the hot spot result layers is determined by comparing the significance level categories between corresponding features in both input layers. The similarity measures how closely the hot spots, cold spots, and nonsignificant areas of both hot spot results spatially align. The association (or dependence) measures the strength of the underlying statistical relationship between the hot spot variables (similar to correlation for continuous variables).

TODO: Need to have same points in both datasets (spatially joined) to run this tool!

- Run the *Hot Spot Analysis Comparison (Spatial Statistics)* tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/hot-spot-comparison.htm

2.10 Step 8 – Prepare data for presentation

This step prepares the analysis results for presentation by converting the tree cover and temperature point feature data to rasters for easier and quicker visualization.

- Run the IDW (Spatial Analyst) tool. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/idw.htm

3 Evaluation

Interpret the results. Provide a slide that evaluates and analyzes the results in the context of the question posed, data limitations, accuracy, and other implications. Present this to the class.

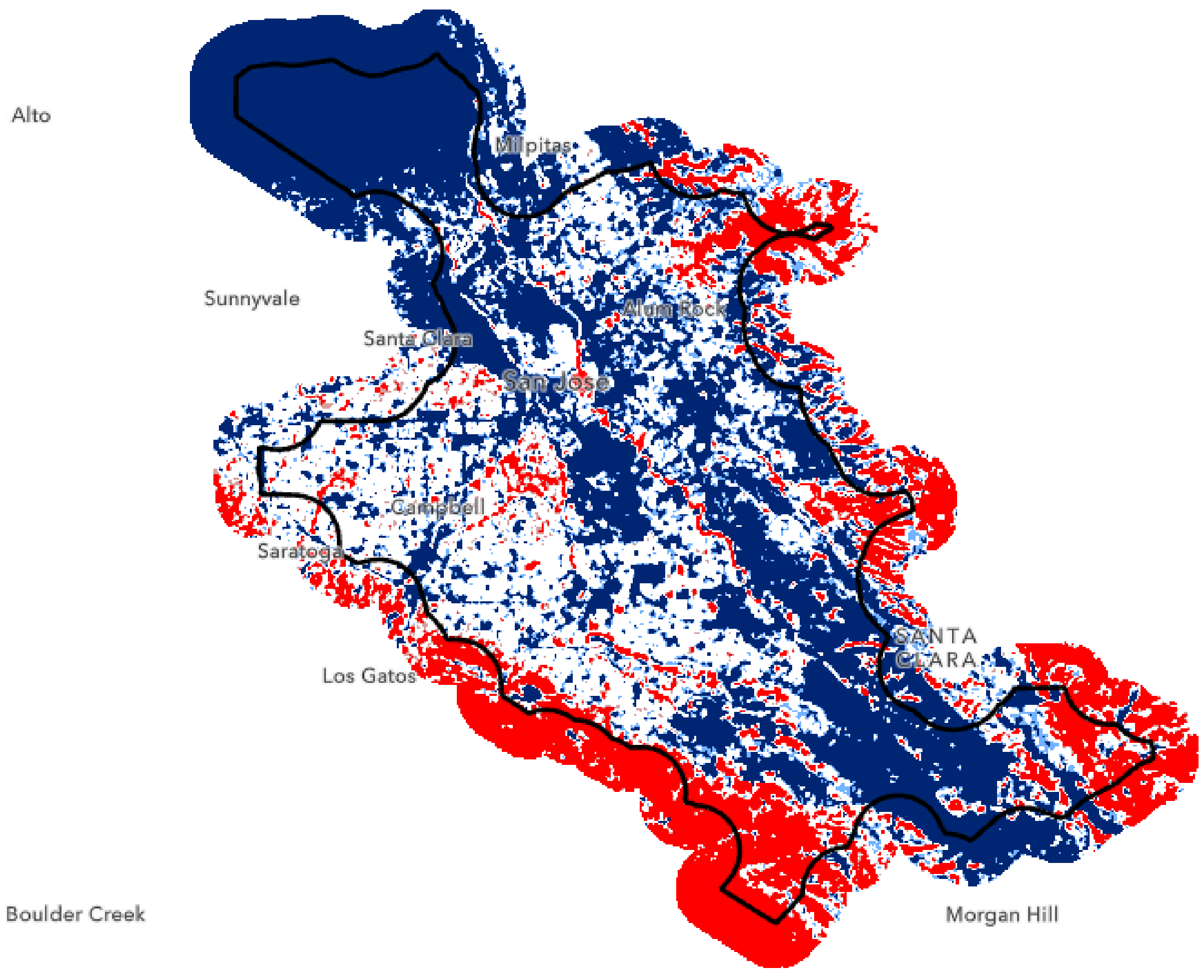
3.1 Spatial Autocorrelation

Both the tree cover and the temperature datasets display statistically significant spatial autocorrelation.

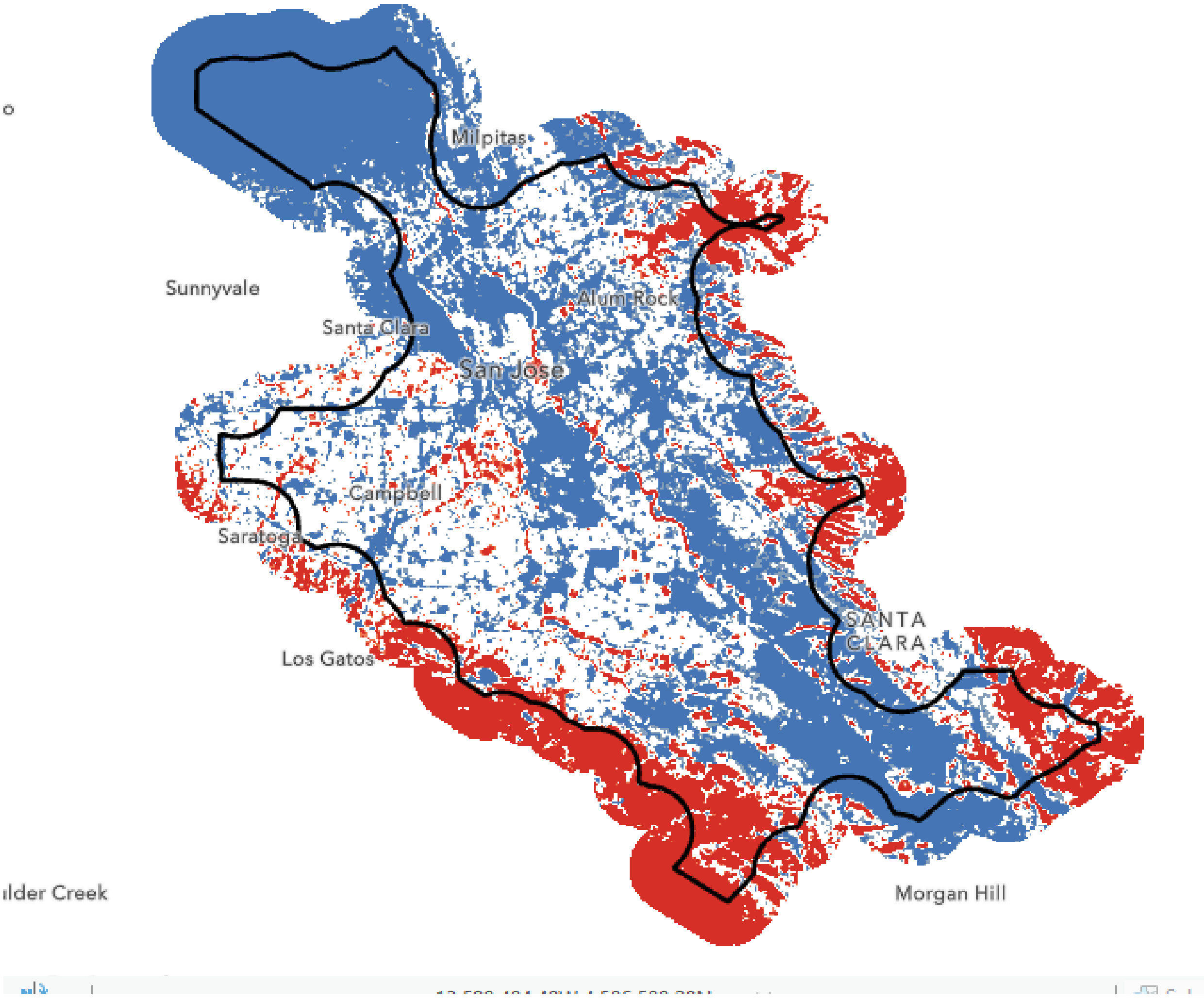
3.2 Hot Spots

Statistically significant hot spots exist in both the temperature and tree cover datasets.

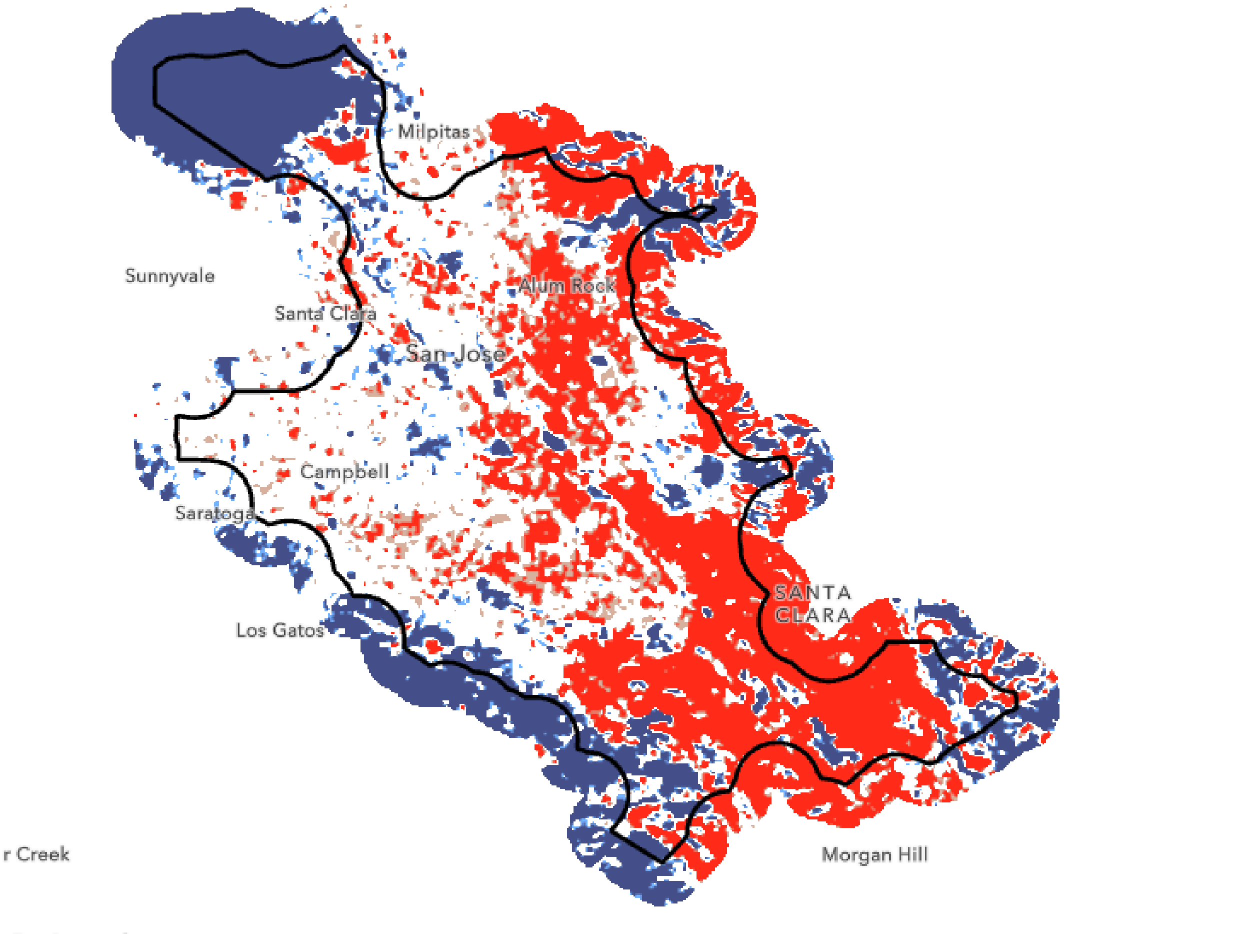
3.2.1 Percent Tree Cover Hot Spots



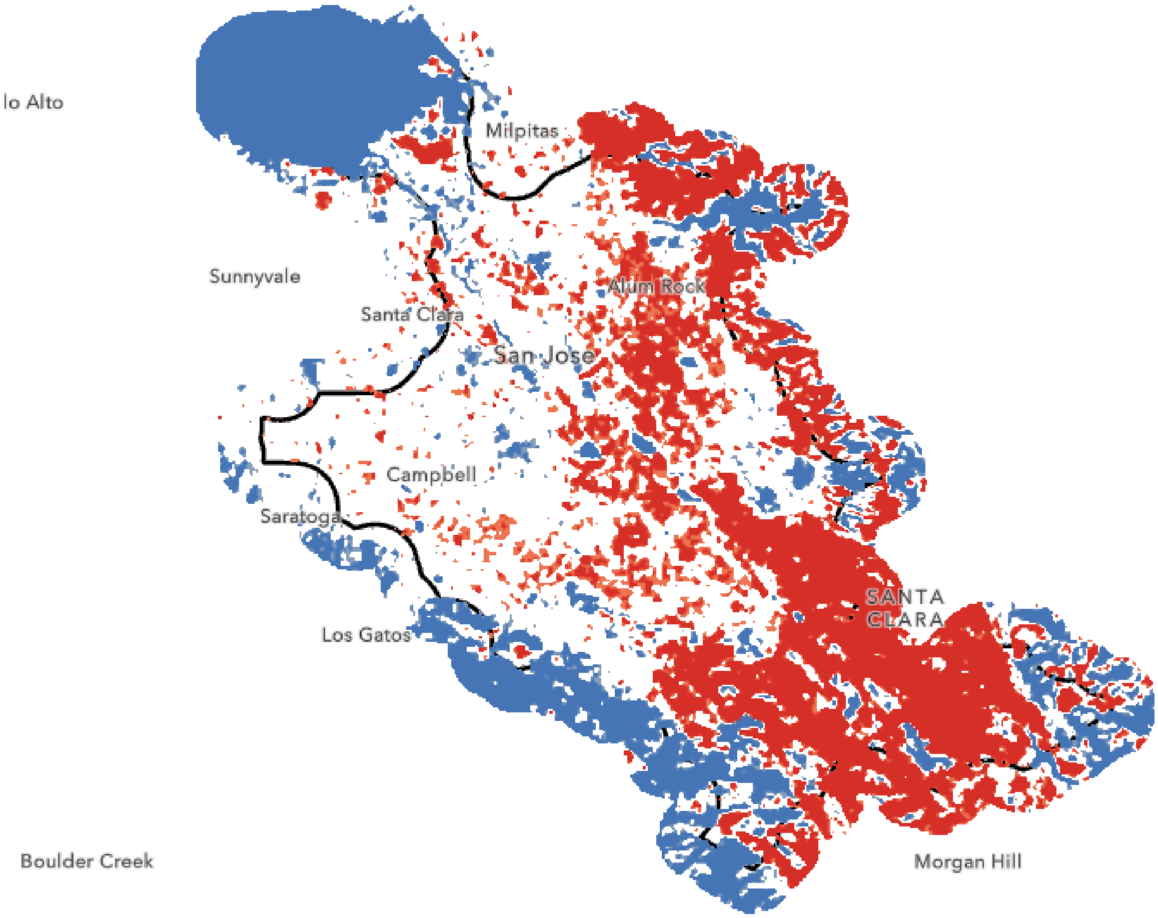
3.2.2 Percent Tree Cover Optimized Hot Spots



3.2.3 Temperature Hot Spots



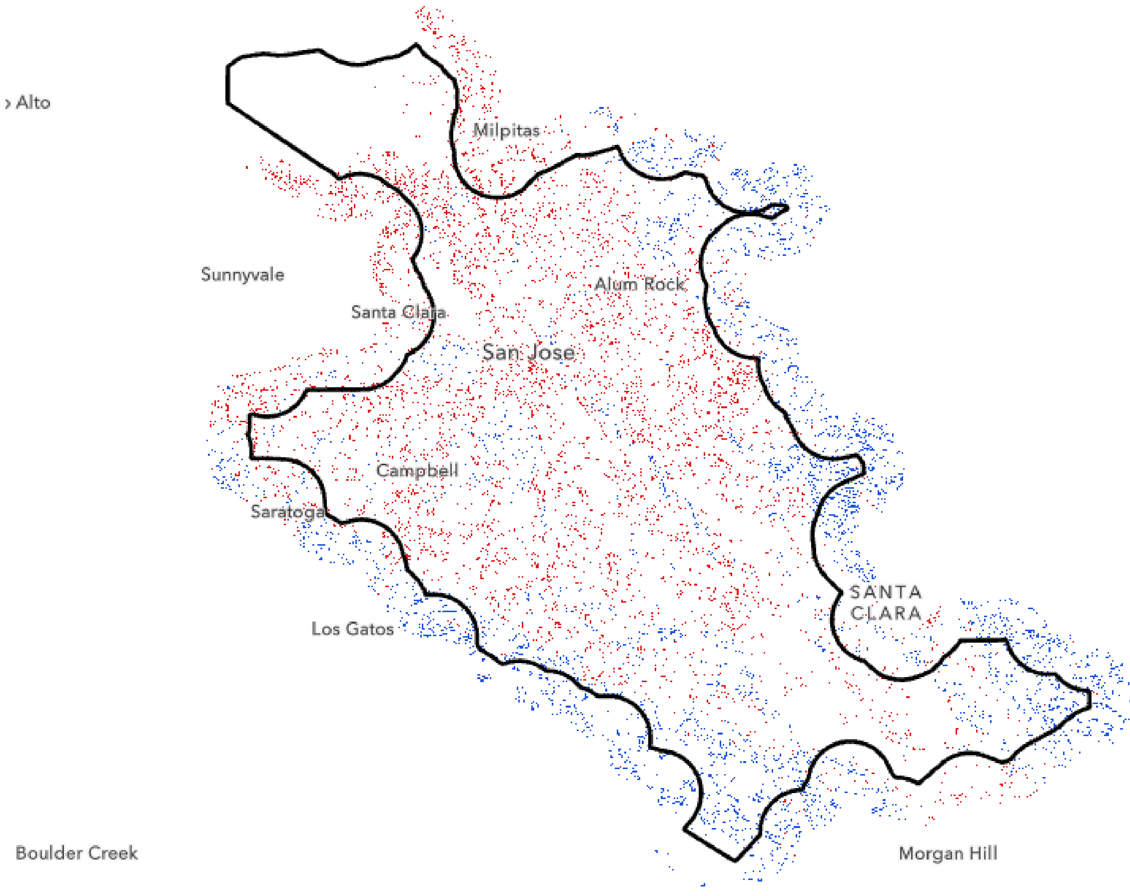
3.3 Temperature Optimized Hot Spots



3.4 Clusters and Outliers

Statistically significant clusters and outliers exist in both the tree cover and temperature datasets.

3.5 Percent Tree Cover Clusters and Outliers

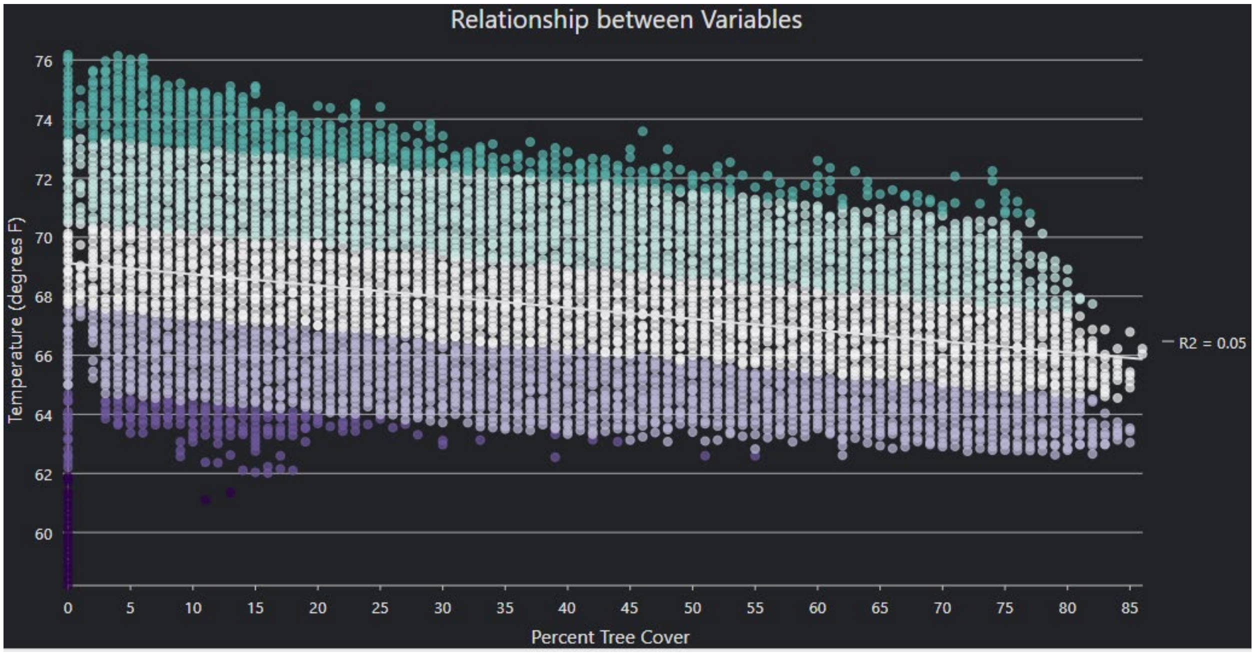

› Alto
Milpitas
Sunnyvale
Alum Rock
Santa Clara
San Jose
Campbell
Saratoga
SANTA
CLARA
Los Gatos
Boulder Creek
Morgan Hill

3.6 Temperature Clusters and Outliers

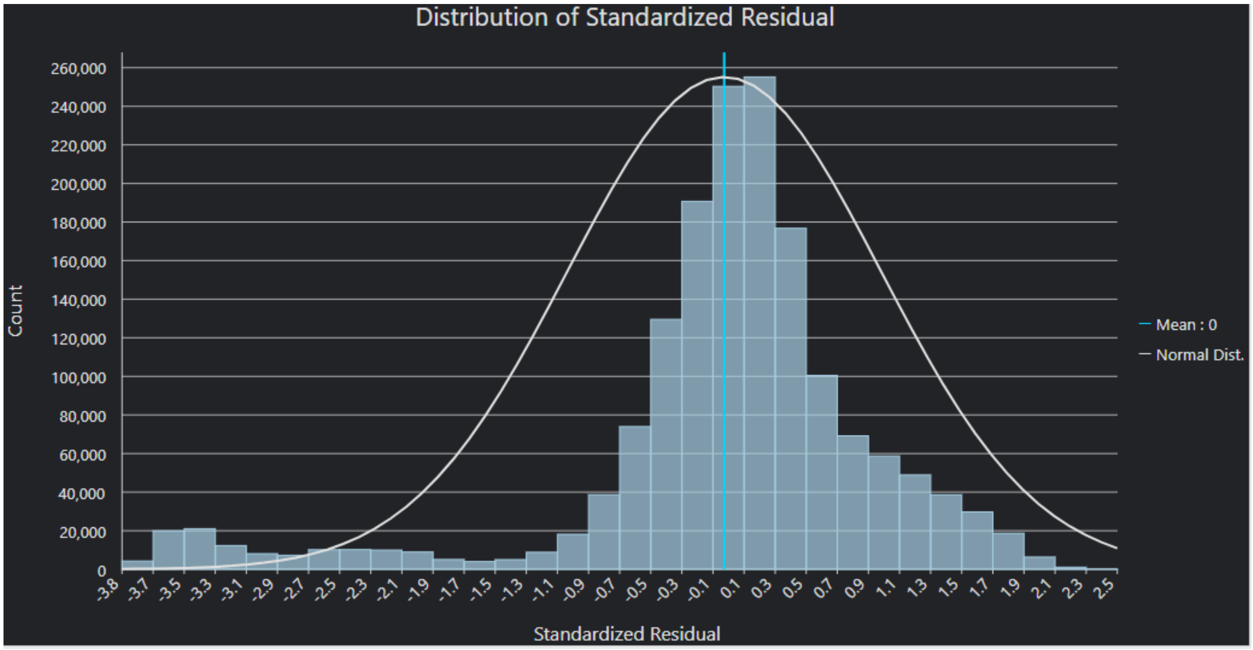
3.7 Linear Regression

There is a statistically significant strong correlation between the temperature and the tree cover.

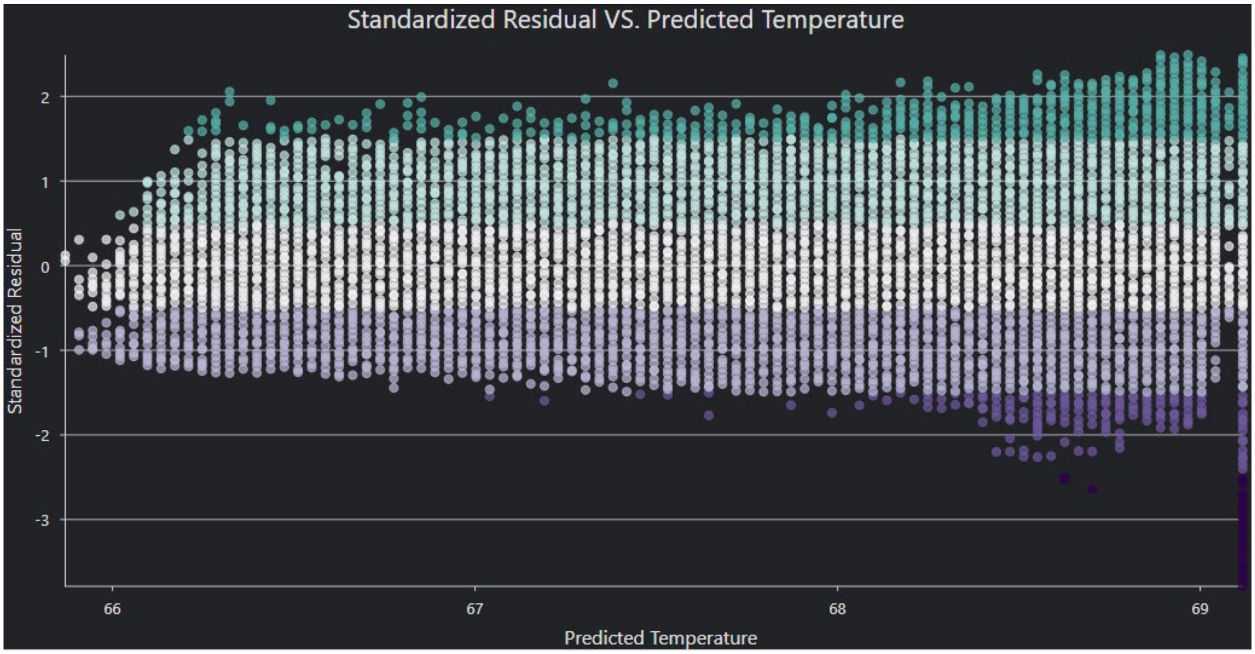
3.8 Relationship Between Variables



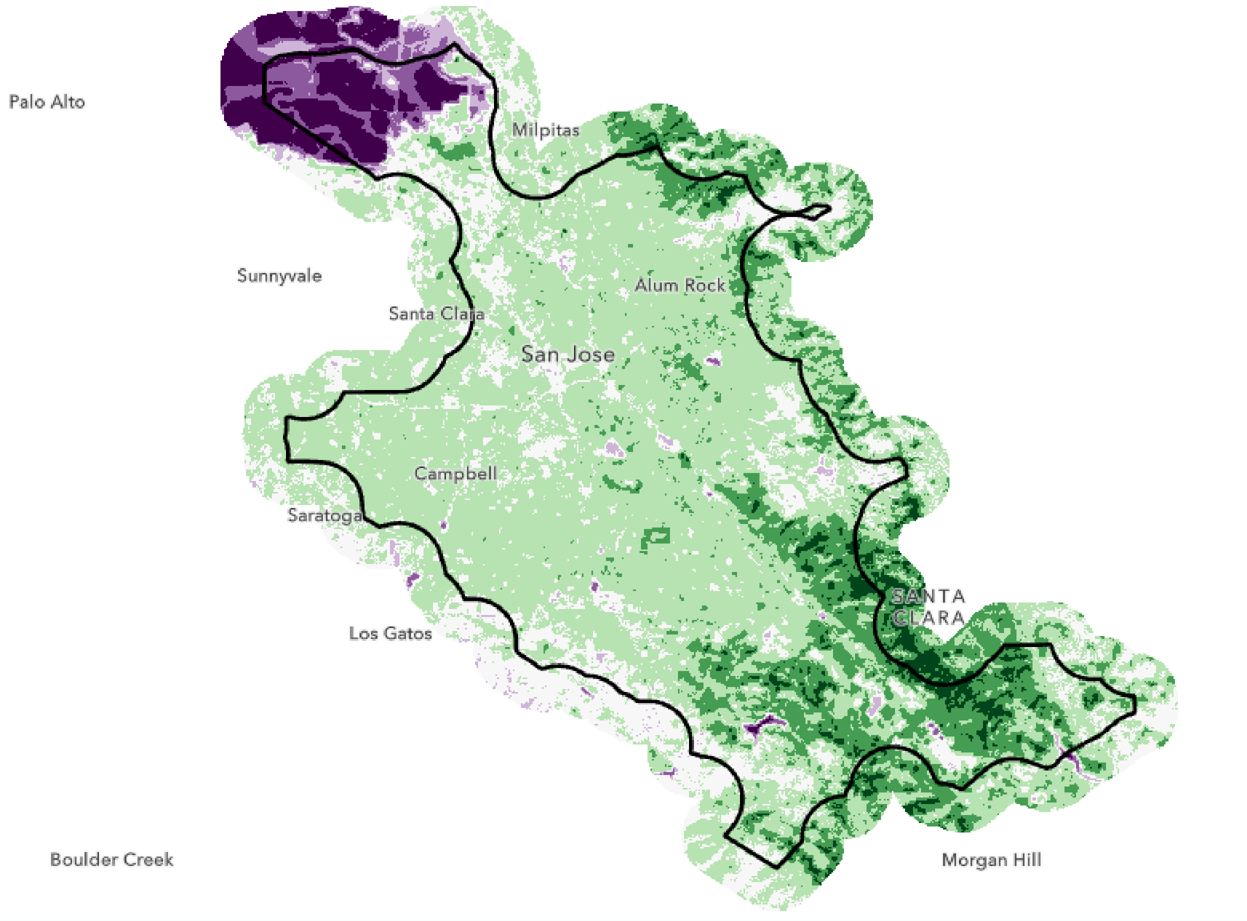
3.9 Distribution of Standardized Residuals



3.10 Standardized Residual vs Predicted Plot



3.11 Standardize Residual Map



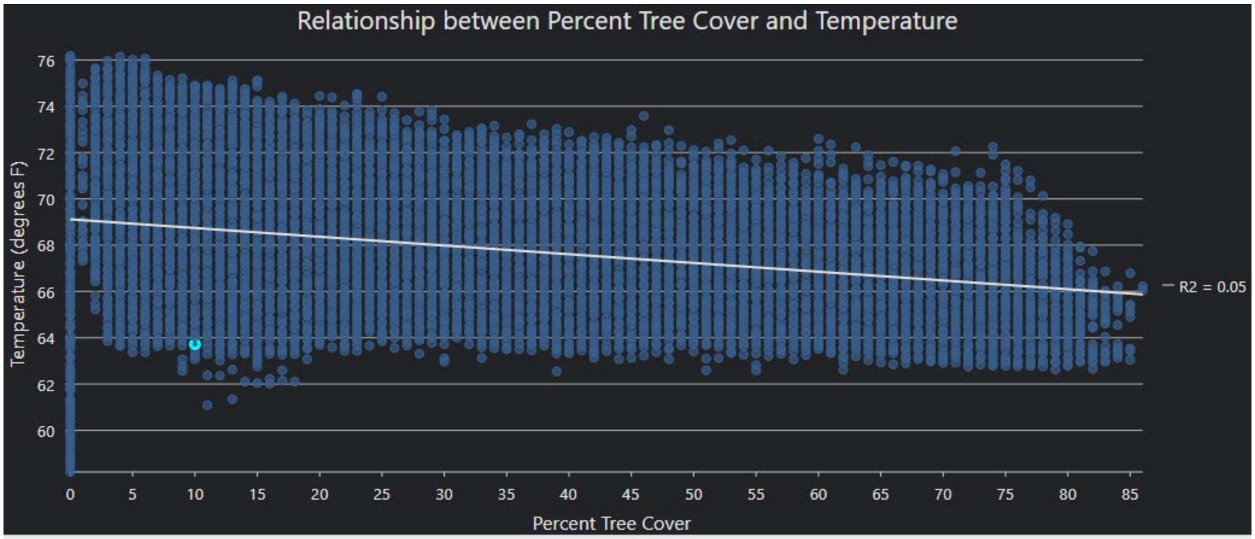
3.12 Geographically Weighted Regression

It was not possible to run a geographically weighted regression between the tree cover and temperature datasets due to the lack of variation in the relationship between these variables across the study area. Within the study area the relationship between temperature and tree cover appears to be independent of location.

```
ERROR 110222: Unable to estimate at least one local model due to multicollinearity (data redundancy).
```

At least one local model could not be estimated because some of the Explanatory Variable(s) parameter values are highly correlated with each other.

3.13 Scatter Plot



4 Current Challenges and Directions for Future Work

Describe and issues or problems you encountered. Spatial analysis is a continuous and iterative process that often leads to further questions and refinements. Provide a summary of the challenges you had or new directions that were identified during the analysis.

Because the input raster datasets came from different sources the grid of pixels does not align exactly. This issue was resolved when the data was converted to points and a spatial join using nearest neighbor was performed. An alternative approach would be to resample one raster dataset to align with the other before the analysis begins. Other approaches are possible using the Extraction toolset. Both approaches increase the level of uncertainty and reduce the accuracy of the analysis results.

In this case, the Optimized Hotspot Analysis tool does not work well for the regular grid of data points produced from raster datasets. It does not find an optimal scale for the analysis through assessing the intensity of clustering by calculating Moran's I at increasing distance. So, the tool falls back on using a neighborhood defined by the 30 nearest neighbors to each point. Unfortunately, this process consumes a large amount of CPU resources and takes a long time to run for no real added value.

For similar reasons, the Optimized Cluster and Outlier tool does not find any outliers.

It was also not possible to run the Geographically Weighted Linear Regression tool on this dataset. The tool raises an error due to multicollinearity of the datasets because it is unable to find any statistically significant variation across the geography of the study area. This issue could be addressed by increasing the size of the study area to include more non-urban areas outside the City of San Jose such as parts of the San Francisco Bay.

Another challenge is the size of the dataset created by raster imagery with a 30m-by-30m pixel size.

These issues could be addressed in future work by resampling the raster datasets to a larger pixel size, trading off some precision for faster processing. Alternatively, data could be aggregated to a generated polygonal grid or preexisting administrative polygons such as census blocks.

5 Results

Present the results. The best information and analysis becomes increasingly valuable when it can be effectively presented and shared with a larger audience. Present any charts, graphs, maps, story maps, or apps that help support the results of your analysis.

6 Conclusion

Make a decision / accept or reject null hypothesis. Spatial analysis and GIS are used to support the decision-making process. A successful spatial analysis helps you to accept or reject the null hypothesis which can lead to the understanding necessary to drive decisions and action. Inform the class on the results of the analysis.

- The ordinary least squares regression found a very strong correlation between tree cover and temperature in the study area. The null hypothesis can be rejected with greater than 99% certainty (p-value ?).
- The analysis found no significant spatial autocorrelation in either of the input data sets.
- Although this analysis only shows a correlation between tree cover and temperature, it can be concluded that trees cover in urban areas make a significant contribution to reducing the urban heat island effect. This knowledge should be used to drive policy decision and focus action to preserve and expand the urban forest.

7 Appendix A – Data Sources

This analysis used the following data source.

- USA NLCD Tree Canopy Cover
  https://www.arcgis.com/home/item.html?id=f2d114f071904e1fa11b4bb215dc08f3
- Landsat 8 satellite imagery scene LC08_L2SP_044034_20210828_20, from August 28, 2021, obtained from the USGS Earth Explorer website: https://earthexplorer.usgs.gov/

8 Appendix B – Raster dataset preparation

8.1 Obtain Landsat Imagery for Study Area

Obtain Landsat 8 imagery covering the study area from the USGS EarthExplorer website. Select the scene LC08_L2SP_044034_20210828_20 from August 28, 2021, for this study.

- Download all bands.
- Load all bands into a composite raster.

The model builder diagram below illustrates this process.

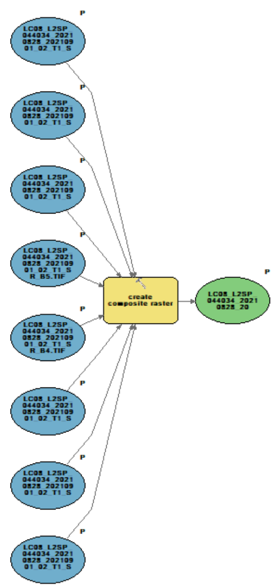


Figure 1 - Create a composite raster.

8.2 Create Land Surface Temperature Raster

Process bands 4, 5, and 10 of the composite raster created above to estimate the land surface temperature for each pixel in the study area. For a fuller explanation of the algorithm see “How to Use Arcgis pro to Map Urban Heat Islands” (Oppong). The model builder diagram below shows the process to calculate the land surface temperature.

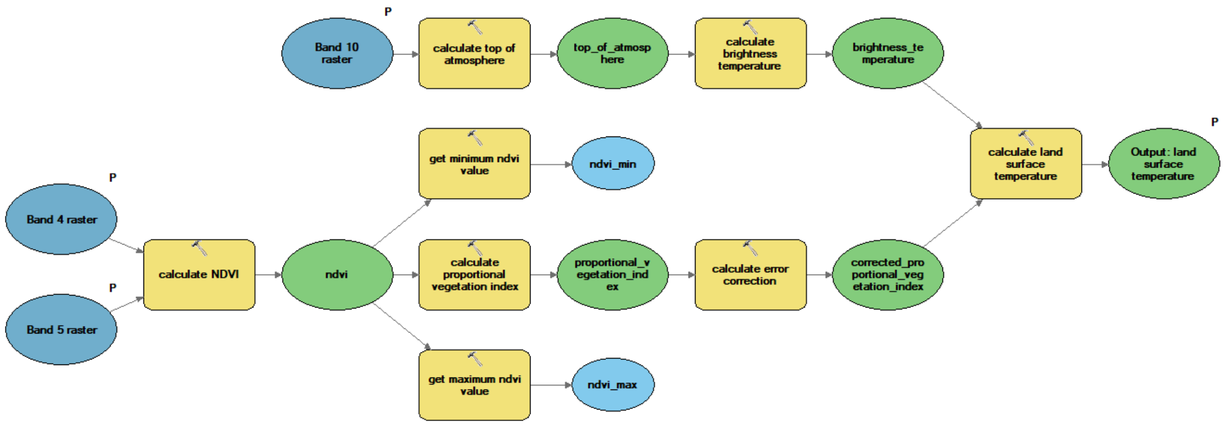


Figure 2 - Model to calculate land surface temperature.

The steps followed were as follows:

1) Generate an ***nvdi*** raster using the rater calculator expression:

```
Float( "%Band 5 raster%" - "%Band 4 raster%") / Float( "%Band 5 raster%" + "%Band 4 raster%")
```

2) Extract minimum and maximum pixel values from the ***ndvi*** raster.
3) Generate a ***proportional_vegetation_index*** raster from the ***nvdi*** raster using the raster calculator expression:

```
Square(Float("%ndvi%" - "%ndvi%".minimum) / Float("%ndvi%".maximum - "%ndvi%".minimum))
```

4) Generate a ***corrected_proportional_vegetation_index*** raster from the ***proportional_vegetation_index*** raster using the raster calculator expression:

```
(4E-3 * "%proportional_vegetation_index%") + 9.86E-1
```

5) Calculate a ***top_of_atmosphere*** raster from band 10 using the raster calculator expression:

```
(3.342E-4 * "%Band 10 raster%") + 1E-1
```

6) Generate a ***brightness_temperature*** raster from the ***top_of_atmosphere*** raster using the raster calculator expression:

```
(1.3210789E3 / Ln((7.748853E2 / "%top_of_atmosphere%") + 1.0)) - 273.15
```

7) Calculate the final ***land_surface_temperature*** raster from the ***brightness_temperature*** and ***corrected_proportional_vegetation_index*** rasters using the raster calculator

```
"%brightness_temperature%" / (1.0 + (1.15E-3 * "%brightness_temperature%" / 1.4388) *

Ln("%corrected_proportional_vegetation_index%"))
```

expression:

8) Save the ***land_surface_temperature*** raster in the project geodatabase.
9) Manually set the symbology to display layer including equal interval classification and color ramp.

8.3 Tree Cover layer

To select just one year in the series, first turn the time series off on the time slider, then create a definition query on the layer which selects only the desired year.

9 Appendix C – Works Cited

"City Limits." *San Jose CA GIS Open Data*, City of San Jose, 7 Feb. 2021, gisdata-csj.opendata.arcgis.com/datasets/CSJ::city-limits/explore.

Oppong, Jeff. "How to Use Arcgis pro to Map Urban Heat Islands." *Geography Realm*, 5 July 2021, www.gislounge.com/how-to-use-arcgis-pro-to-map-urban-heat-islands/.

"USA NLCD Tree Canopy Cover." *Arcgis.Com*, ESRI, 1 Apr. 2023, www.arcgis.com/home/item.html?id=f2d114f071904e1fa11b4bb215dc08f3.

10 Appendix D-- Output from Groprocessing Tools

```
Executing (Spatial Join): SpatialJoin
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\tree_points_2k
m
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\temp_points_2k
m C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points
"Join one to one" KEEP_ALL "grid_code_Mean "grid_code_Mean" true true false 255 Long 0
0,Mean,#,C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\tree_
points_2km,grid_code,-1,-1;grid_code_Mean_1 "grid_code_Mean_1" true true false 255
Float 0
0,Mean,#,C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\temp_
points_2km,grid_code,-1,-1" Closest "50 Meters" # #
Start Time: Sunday, December 10, 2023 8:31:17 PM
Succeeded at Sunday, December 10, 2023 8:32:11 PM (Elapsed Time: 53.65 seconds)
Executing (Generalized Linear Regression (2)): GeneralizedLinearRegression
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points
grid_code_Mean_1 "Continuous (Gaussian)"
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\glr
grid_code_Mean # # # #
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\glr_predicted
C:\ArcGIS_local_projects\urban_heat_islands_pro\glr_model.ssm
Start Time: Sunday, December 10, 2023 8:32:11 PM
```

10.1.1.1.1 Summary of GLR Results [Model Type: Continuous (Gaussian/OLS)]

Variable	Coefficient[a]	StdError	t-Statistic	Probability[b]	Robust_SE	Rob
Intercept	69.117029	0.002758	25056.840037	0.000000*	0.003213	21510.1
GRID_CODE_MEAN	-0.037778	0.000126	-299.510652	0.000000*	0.000097	-388.7

10.1.1.1.2 GLR Diagnostics

Input Features	join_points	Dependent Variable
Number of Observations	1640132	Akaike's Information Criterion (AICc)[d]
Multiple R-Squared[d]	0.051858	Adjusted R-Squared[d]
Joint F-Statistic[e]	89706.630856	Prob(>F), (1,1640130) degrees of freedom
Joint Wald Statistic[e]	151141.214023	Prob(>chi-squared), (1) degrees of freedom
Koenker (BP) Statistic[f]	78206.533980	Prob(>chi-squared), (1) degrees of freedom
Jarque-Bera Statistic[g]	1635107.129984	Prob(>chi-squared), (2) degrees of freedom

10.1.1.1.3 Notes on Interpretation

*	An asterisk next to a number indicates a statistically significant p-value (p < 0.01).
a	Coefficient: Represents the strength and type of relationship between each explanatory variable and variable.
b	Probability and Robust Probability (Robust_Pr): Asterisk (*) indicates a coefficient is statistical < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Robust Probability (Robust_Pr) to determine coefficient significance.

c	Variance Inflation Factor (VIF): Large Variance Inflation Factor (VIF) values (> 7.5) indicate redu explanatory variables.
d	R-Squared and Akaike's Information Criterion (AICc): Measures of model fit/performance.
e	Joint F and Wald Statistics: Asterisk (*) indicates overall model significance (p < 0.01); if the K Statistic [f] is statistically significant, use the Wald Statistic to determine overall model signi
f	Koenker (BP) Statistic: When this test is statistically significant (p < 0.01), the relationships m consistent (either due to non-stationarity or heteroskedasticity). You should rely on the Robust Pr (Robust_Pr) to determine coefficient significance and on the Wald Statistic to determine overall mo
g	Jarque-Bera Statistic: When this test is statistically significant (p < 0.01) model predictions are residuals are not normally distributed).

Succeeded at Sunday, December 10, 2023 8:34:23 PM (Elapsed Time: 2 minutes 11 seconds)

Executing (Optimized Hot Spot Analysis): OptimizedHotSpotAnalysis C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\temp_opt_hotspot_2km grid_code_Mean_1 "Count incidents within fishnet grid" # # C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\temp_density # "120 Meters"

Start Time: Sunday, December 10, 2023 8:34:26 PM

10.1.1.1.4 Initial Data Assessment

Making sure there are enough weighted features for analysis....

- There are 1640132 valid input features.

Evaluating the Analysis Field values....

10.1.1.1.5 GRID_CODE_MEAN_1 Properties:

Min	58.0330
Max	76.3523
Mean	68.6385
Std. Dev.	2.9573

Looking for locational outliers....

- There were no outlier locations found.

Checking to see if the Environment Settings include a raster analysis mask....

- Raster analysis mask not set; constructing convex hull....

10.1.1.1.6 Scale of Analysis

- The Neighborhood Distance used was 120 meters.

10.1.1.1.7 Hot Spot Analysis

Finding statistically significant clusters of high and low grid_code_Mean_1 values....

- There are 1066241 output features statistically significant based on an FDR correction for multiple testing and spatial dependence.
- 0% of features had less than 8 neighbors based on the distance band of 120 meters

10.1.1.1.8 Output

Creating output feature class:
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\temp_opt_hotspot_2km

1. Red output features represent hot spots where high grid_code_Mean_1 values cluster.
2. Blue output features represent cold spots where low grid_code_Mean_1 values cluster.

Creating output raster layer:
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\temp_density

- Using optimal fixed distance band (120 meters) for the kernel density search radius.
- The surface will be clipped to a convex hull of the input points.

Succeeded at Sunday, December 10, 2023 8:41:50 PM (Elapsed Time: 7 minutes 23 seconds)

Executing (Optimized Outlier Analysis): OptimizedOutlierAnalysis C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\temp_opt_cluster_2km grid_code_Mean_1 "Count incidents within fishnet grid" # # "Balanced (499 permutations)" # "120 Meters"

Start Time: Sunday, December 10, 2023 8:41:51 PM

10.1.1.1.9 Initial Data Assessment

Making sure there are enough weighted features for analysis....

- There are 1640132 valid input features.

Evaluating the Analysis Field values....

10.1.1.1.10 GRID_CODE_MEAN_1 Properties:

Min	58.0330
Max	76.3523
Mean	68.6385
Std. Dev.	2.9573

Looking for locational outliers....

- There were no outlier locations found.

10.1.1.1.11 Scale of Analysis

- The Neighborhood Distance used was 120 meters.

10.1.1.1.12 Optimized Outlier Analysis

- Creating the random reference distribution with 499 permutations....
- Finding statistically significant outliers of high and low grid_code_Mean_1 values....
- There are 1063186 output features statistically significant based on an FDR correction for multiple testing and spatial dependence.
- There are 150 statistically significant high outlier features.
- There are 407 statistically significant low outlier features.
- There are 413712 features part of statistically significant low clusters.
- There are 648917 features part of statistically significant high clusters
- 0% of features had less than 8 neighbors based on the distance band of 120 meters

10.1.1.1.13 Output

- Pink output features are part of a cluster of high grid_code_Mean_1 values.
- Light Blue output features are part of a cluster of low grid_code_Mean_1 values.
- Red output features represent high outliers within a cluster of low grid_code_Mean_1 values..
- Blue output features represent low outliers within a cluster of high grid_code_Mean_1 values.

Succeeded at Sunday, December 10, 2023 9:29:27 PM (Elapsed Time: 47 minutes 35 seconds)

Executing (Optimized Hot Spot Analysis (2)): OptimizedHotSpotAnalysis C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\tree_opt_hotspot_2km grid_code_Mean "Count incidents within fishnet grid" # # C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\tree_density # "120 Meters"

Start Time: Sunday, December 10, 2023 9:29:28 PM

10.1.1.1.14 Initial Data Assessment

Making sure there are enough weighted features for analysis....

- There are 1640132 valid input features.

Evaluating the Analysis Field values....

10.1.1.1.15 GRID_CODE_MEAN Properties:

Min	0.0000
Max	86.0000
Mean	12.6681
Std. Dev.	17.8264

Looking for locational outliers....

- There were no outlier locations found.

Checking to see if the Environment Settings include a raster analysis mask....

- Raster analysis mask not set; constructing convex hull....

10.1.1.1.16 Scale of Analysis

- The Neighborhood Distance used was 120 meters.

10.1.1.1.17 Hot Spot Analysis

Finding statistically significant clusters of high and low grid_code_Mean values....

- There are 1196434 output features statistically significant based on an FDR correction for multiple testing and spatial dependence.
- 0% of features had less than 8 neighbors based on the distance band of 120 meters

10.1.1.1.18 Output

Creating output feature class:
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\tree_opt_hotspot_2km

1. Red output features represent hot spots where high grid_code_Mean values cluster.
2. Blue output features represent cold spots where low grid_code_Mean values cluster.

Creating output raster layer:
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\tree_density

- Using optimal fixed distance band (120 meters) for the kernel density search radius.
- The surface will be clipped to a convex hull of the input points.

Succeeded at Sunday, December 10, 2023 10:10:15 PM (Elapsed Time: 40 minutes 47 seconds)

Executing (Optimized Outlier Analysis (2)): OptimizedOutlierAnalysis C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\tree_opt_cluster_2km grid_code_Mean "Count incidents within fishnet grid" # # "Balanced (499 permutations)" # "120 Meters"

Start Time: Sunday, December 10, 2023 10:10:16 PM

10.1.1.1.19 Initial Data Assessment

Making sure there are enough weighted features for analysis....

- There are 1640132 valid input features.

Evaluating the Analysis Field values....

10.1.1.1.20 GRID_CODE_MEAN Properties:

Min	0.0000
Max	86.0000
Mean	12.6681

Std. Dev.	17.8264

Looking for locational outliers....

- There were no outlier locations found.

10.1.1.1.21 Scale of Analysis

- The Neighborhood Distance used was 120 meters.

10.1.1.1.22 Optimized Outlier Analysis

- Creating the random reference distribution with 499 permutations....
- Finding statistically significant outliers of high and low grid_code_Mean values....
- There are 1199309 output features statistically significant based on an FDR correction for multiple testing and spatial dependence.
- There are 44869 statistically significant high outlier features.
- There are 33879 statistically significant low outlier features.
- There are 804091 features part of statistically significant low clusters.
- There are 316470 features part of statistically significant high clusters
- 0% of features had less than 8 neighbors based on the distance band of 120 meters

10.1.1.1.23 Output

- Pink output features are part of a cluster of high grid_code_Mean values.
- Light Blue output features are part of a cluster of low grid_code_Mean values.
- Red output features represent high outliers within a cluster of low grid_code_Mean values..
- Blue output features represent low outliers within a cluster of high grid_code_Mean values.

Succeeded at Sunday, December 10, 2023 11:05:29 PM (Elapsed Time: 55 minutes 12 seconds)

Executing (Hot Spot Analysis (Getis-Ord Gi*) (3)): HotSpots C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points grid_code_Mean C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_tree_hotspot FIXED_DISTANCE_BAND Euclidean Row 120 # # APPLY_FDR #

Start Time: Sunday, December 10, 2023 11:05:29 PM

Succeeded at Sunday, December 10, 2023 11:08:53 PM (Elapsed Time: 3 minutes 24 seconds)

Executing (Cluster and Outlier Analysis (Anselin Local Moran's I) (3)): ClustersOutliers C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points grid_code_Mean C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_tree_cluster FIXED_DISTANCE_BAND Euclidean Row 120 # APPLY_FDR 499 #

Start Time: Sunday, December 10, 2023 11:08:54 PM

Succeeded at Monday, December 11, 2023 9:06:31 AM (Elapsed Time: 9 hours 57 minutes 36 seconds)

Executing (Hot Spot Analysis (Getis-Ord Gi*) (4)): HotSpots C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points grid_code_Mean_1 C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_temp_hotspots FIXED_DISTANCE_BAND Euclidean Row 120 # # APPLY_FDR #

Start Time: Monday, December 11, 2023 9:06:32 AM

Succeeded at Monday, December 11, 2023 9:09:53 AM (Elapsed Time: 3 minutes 20 seconds)

Executing (Cluster and Outlier Analysis (Anselin Local Moran's I) (4)): ClustersOutliers C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points grid_code_Mean_1 C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_temp_cluster FIXED_DISTANCE_BAND Euclidean Row 120 # APPLY_FDR 499 #

Start Time: Monday, December 11, 2023 9:09:53 AM

```
Succeeded at Monday, December 11, 2023 10:04:44 AM (Elapsed Time: 54 minutes 51
seconds)
Executing (Point to Raster): PointToRaster
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_temp_hots
pots Gi_Bin
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_temp_hots
pots_ras "Most frequent" NONE 30 BUILD
Start Time: Monday, December 11, 2023 10:04:45 AM
Succeeded at Monday, December 11, 2023 10:04:49 AM (Elapsed Time: 3.90 seconds)
Executing (Point to Raster (2)): PointToRaster
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_tree_hots
pot Gi_Bin
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_tree_hots
pot_ras "Most frequent" NONE 30 BUILD
Start Time: Monday, December 11, 2023 10:04:51 AM
Succeeded at Monday, December 11, 2023 10:04:54 AM (Elapsed Time: 3.54 seconds)
Executing (Point to Raster (3)): PointToRaster
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_tree_clus
ter COType
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_tree_clus
ter_ras "Most frequent" NONE 30 BUILD
Start Time: Monday, December 11, 2023 10:04:55 AM
Succeeded at Monday, December 11, 2023 10:04:59 AM (Elapsed Time: 3.59 seconds)
Executing (Hot Spot Analysis Comparison): HotSpotAnalysisComparison
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_tree_hots
pot
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_temp_hots
pots
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_tree_hots
pot_comparison 8 499 "Fuzzy weights" "-3 -3 1;3 3 1;-3 -2 0.71;3 2 0.71;-3 -1 0.55;3 1
0.55;-2 -2 1;2 2 1;-2 -1 0.78;2 1 0.78;-1 -1 1;1 1 1;0 0 1" # NO_EXCLUDE
Start Time: Monday, December 11, 2023 10:05:00 AM
```

10.1.1.1.24 Global Hot Spot Analysis Comparison Results

Similarity Value	0.2609
Expected Similarity Value	0.2917
Spatial Fuzzy Kappa	-0.0434
Number of Non-Significant Features	228562 (13.94%)

10.1.1.1.25 Categorical Weights Table

	Cold 99%	Cold 95%	Cold 90%	Not Significant	Hot 90%
Cold 99%	1.00	0.71	0.55	0.00	0.00
Cold 95%	0.71	1.00	0.78	0.00	0.00
Cold 90%	0.55	0.78	1.00	0.00	0.00
Not Significant	0.00	0.00	0.00	1.00	0.00
Hot 90%	0.00	0.00	0.00	0.00	1.00
Hot 95%	0.00	0.00	0.00	0.00	0.78
Hot 99%	0.00	0.00	0.00	0.00	0.55

10.1.1.1.26 Hot Spot Significance Level Pair (Counts)

	Hot Spot 2 Significance Level

Hot Spot 1 Significance Level	Cold 99%	Cold 95%	Cold 90%	Not Significant	Hot 90%	Hot 95%
Cold 99%	145435	9563	6281	172147	21562	40125
Cold 95%	3909	2089	1349	40368	4819	8361
Cold 90%	2019	999	680	21271	2539	4392
Not Significant	19571	10140	7049	228562	25396	38431
Hot 90%	1794	977	604	12076	586	827
Hot 95%	3704	1913	1216	17951	838	1339
Hot 99%	172033	21478	10951	81516	3329	5161
Total	348465	47159	28130	573891	59069	98636

10.1.1.1.27 Hot Spot Significance Level Pair Counts (Percentages)

	Hot Spot 2 Significance Level				
Hot Spot 1 Significance Level	Cold 99%	Cold 95%	Cold 90%	Not Significant	Hot 90%
Cold 99%	21.69	1.43	0.94	25.68	3.22
Cold 95%	3.50	1.87	1.21	36.17	4.32
Cold 90%	3.65	1.80	1.23	38.41	4.58
Not Significant	4.41	2.29	1.59	51.51	5.72
Hot 90%	9.05	4.93	3.05	60.89	2.95
Hot 95%	11.77	6.08	3.87	57.06	2.66
Hot 99%	55.91	6.98	3.56	26.49	1.08

Succeeded at Monday, December 11, 2023 10:14:59 AM (Elapsed Time: 9 minutes 58 seconds)

Executing (Spatial Autocorrelation (Global Moran's I)): SpatialAutocorrelation C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points grid_code_Mean GENERATE_REPORT "Fixed distance band" Euclidean Row 120 # #

Start Time: Monday, December 11, 2023 10:15:03 AM

10.1.1.1.28 Global Moran's I Summary

Moran's Index	0.796732
Expected Index	-0.000001
Variance	0.000000
z-score	4911.233990
p-value	0.000000

Distance measured in meters

Writing html report....
C:\ArcGIS_local_projects\urban_heat_islands_pro\MoransI_Result_23160_28472_2.html

Succeeded at Monday, December 11, 2023 10:20:13 AM (Elapsed Time: 5 minutes 9 seconds)

```
Executing (Spatial Autocorrelation (Global Moran's I) (2)): SpatialAutocorrelation
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points
grid_code_Mean_1 GENERATE_REPORT "Fixed distance band" Euclidean Row 120 # #
Start Time: Monday, December 11, 2023 10:20:13 AM
```

10.1.1.1.29 Global Moran's I Summary

Moran's Index	0.978406
Expected Index	-0.000001
Variance	0.000000
z-score	6031.110776
p-value	0.000000

```
Distance measured in meters

Writing html report....
C:\ArcGIS_local_projects\urban_heat_islands_pro\MoransI_Result_23160_28472_3.html

Succeeded at Monday, December 11, 2023 10:25:19 AM (Elapsed Time: 5 minutes 5 seconds)
Executing (Point to Raster (4)): PointToRaster
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_temp_clus
ter COType
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_temp_clus
ter_ras "Most frequent" NONE 30 BUILD
Start Time: Monday, December 11, 2023 10:25:19 AM
Succeeded at Monday, December 11, 2023 10:25:22 AM (Elapsed Time: 3.27 seconds)
Executing (Point to Raster (5)): PointToRaster
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\temp_points_2k
m grid_code
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\temp_points_2k
m_ras "Most frequent" NONE 30 BUILD
Start Time: Monday, December 11, 2023 10:25:23 AM
Succeeded at Monday, December 11, 2023 10:25:26 AM (Elapsed Time: 2.35 seconds)
Executing (Point to Raster (6)): PointToRaster
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\tree_points_2k
m grid_code
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\tree_points_2k
m_ras "Most frequent" NONE 30 BUILD
Start Time: Monday, December 11, 2023 10:25:27 AM
Succeeded at Monday, December 11, 2023 10:25:30 AM (Elapsed Time: 3.20 seconds)
```

GLR

```
Executing (Generalized Linear Regression (2)): GeneralizedLinearRegression
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\join_points
grid_code_Mean_1 "Continuous (Gaussian)"
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\glr
grid_code_Mean # # # #
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\glr_predicted
C:\ArcGIS_local_projects\urban_heat_islands_pro\glr_model.ssm
Start Time: Monday, December 11, 2023 3:26:53 PM
```

10.1.1.1.30 Summary of GLR Results [Model Type: Continuous (Gaussian/OLS)]

Variable	Coefficient[a]	StdError	t-Statistic	Probability[b]	Robust_SE	Robust_t	Robust_P
Intercept	69.117245	0.002758	25058.090143	0.000000*	0.003213	21511.166074	0.00000

GRID_CODE_MEAN	-0.037781	0.000126	-299.554917	0.000000*	0.000097	-388.825568	0.00000

10.1.1.1.31 GLR Diagnostics

Input Features	join_points	Dependent Variable	GRID_CODE_MEAN_1
Number of Observations	1640144	Akaike's Information Criterion (AICc)[d]	8123742.014277
Multiple R-Squared[d]	0.051873	Adjusted R-Squared[d]	0.051872
Joint F-Statistic[e]	89733.148461	Prob(>F), (1,1640142) degrees of freedom	0.000000*
Joint Wald Statistic[e]	151185.322609	Prob(>chi-squared), (1) degrees of freedom	0.000000*
Koenker (BP) Statistic[f]	78200.992240	Prob(>chi-squared), (1) degrees of freedom	0.000000*
Jarque-Bera Statistic[g]	1635391.396415	Prob(>chi-squared), (2) degrees of freedom	0.000000*

10.1.1.1.32 Notes on Interpretation

*	An asterisk next to a number indicates a statistically significant p-value (p < 0.01).
a	Coefficient: Represents the strength and type of relationship between each explanatory variable and the dependent variable.
b	Probability and Robust Probability (Robust_Pr): Asterisk (*) indicates a coefficient is statistically significant (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use the Robust Probability column (Robust_Pr) to determine coefficient significance.
c	Variance Inflation Factor (VIF): Large Variance Inflation Factor (VIF) values (> 7.5) indicate redundancy among explanatory variables.
d	R-Squared and Akaike's Information Criterion (AICc): Measures of model fit/performance.
e	Joint F and Wald Statistics: Asterisk (*) indicates overall model significance (p < 0.01); if the Koenker (BP) Statistic [f] is statistically significant, use

	the Wald Statistic to determine overall model significance.
f	Koenker (BP) Statistic: When this test is statistically significant (p < 0.01), the relationships modeled are not consistent (either due to non-stationarity or heteroskedasticity). You should rely on the Robust Probabilities (Robust_Pr) to determine coefficient significance and on the Wald Statistic to determine overall model significance.
g	Jarque-Bera Statistic: When this test is statistically significant (p < 0.01) model predictions are biased (the residuals are not normally distributed).

```
Succeeded at Monday, December 11, 2023 3:29:03 PM (Elapsed Time: 2 minutes 10 seconds)
Executing (Point to Raster (10)): PointToRaster
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\glr STDRESID
C:\ArcGIS_local_projects\urban_heat_islands_pro\geospatial_analysis.gdb\glr_std_residu
al_ras "Most frequent" NONE 30 BUILD
Start Time: Monday, December 11, 2023 3:29:04 PM
Succeeded at Monday, December 11, 2023 3:29:08 PM (Elapsed Time: 3.97 seconds)
```