

Projecting the Urban Heat Island Effect Using Historical Climate Trends and Land Use-Land Cover

Samhita Srivatsan (Monta Vista High School), Moneel Patel (Seven Lakes High School)

NASA STEM Enhancement in Earth Sciences

Abstract Summary—This research study investigates how developed land cover and weather trends can forecast the UHIE in the Greater Austin Region with two distinct modeling frameworks: (1) Keras sequential models to identify correlations between environmental factors closely linked to the UHIE and (2) developed softwares in QGIS Modules for Land Use Change Evaluation (MOLUSCE) in conjunction with high resolution satellite imagery provided by Multi-Resolution Land Characteristics land cover/land use data.

Keywords—Urban Heat Island, Prediction, Machine Learning Model, Satellite Imagery, Land Use-Land Cover, Greater Austin Region

I. INTRODUCTION

In recent years, historical highs have been recorded for land surface temperatures (LST) in urbanizing areas, and have been identified as closely related to a decrease in air quality as a measurement of Particulate Matter 2.5 and contamination of freshwater bodies as reflected by decreasing Freshwater Quality Index (WQI) scores.

The Urban Heat Island Effect (UHIE) is “widely recognized as a heat accumulation phenomenon, which is the most obvious characteristic of urban climate caused by urban construction and human activities” [1, 10]. In other words, the spatial pattern in which urban landscapes record higher temperatures within closer areas in contrast to surrounding rural, non-urbanized, areas, in some cases 10° C higher [2]. Urban areas are characterized by a lack of greenery and the presence of clustered buildings, roads, and infrastructure that create “islands” experiencing global warming differently than outlying, less densely populated areas. Researchers have recorded daytime temperatures up to 7° higher and nighttime temperatures up to 5° higher in these islands than in surrounding areas. [42].

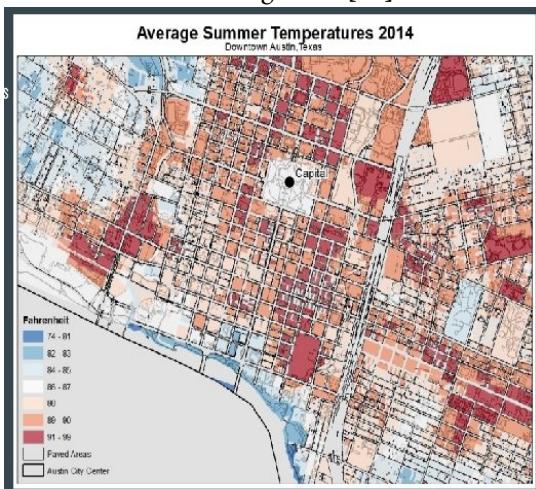


Figure 1A: Downtown Austin, TX (Hayhoe et al, 2014)

Possible explanations for this phenomenon rely on “rapidly heating urban surfaces consisting of buildings, asphalt, bare-soil, and short grasses” [2] and are typically quantified through air and surface temperature measurements [3]. Such Urban Heat Islands contribute to many problems, most notably, ones detrimental to human health such as increased exposure to heat waves which can potentially lead to respiratory and cardiovascular health diseases [4]. As closely related to UHIE, the factor of Land Surface Temperature (LST) is important to understand in measuring the UHIE.

LST is commonly derived from satellite data with remote sensing, most notably using Moderate Resolution Imaging Spectroradiometer (MODIS) thermal infrared bands [5] as a standard measurement. LST is widely considered a measurement of the UHIE and their relationship has been historically well-documented and widely explored by climate researchers across the world [21, 22]. Efforts to explain the increase in LST in urban areas have been done by analyzing the difference in “spectral albedos” [2] or measure of the reflectivity of a material in relation to thermal radiation. Researchers have recorded a trend in “loss of forest and cropland to urban use” [6] due to urbanization, noting a decrease in Normalized Vegetation Differentiation Index (NDVI), a standard measurement of vegetation health via near-infrared light. To generalize this factor for study areas, spatial land surface classification is broadly associated with varying levels of albedo such that an explanation is provided for higher surface temperature temperatures.

This two-part study is driven by the following questions: How effective is forecasting land use and land cover and land surface temperature (as indicators of the UHIE) using computer programs? Can we use these predictions to visualize the UHIE and implement countermeasures? How can developed land cover and weather trends be used to forecast the Urban Heat Island Effect?

In the scope of this geospatial temporal analysis, existing data on environmental factors such as soil moisture, relative humidity, air and surface temperature, and population growth will be compared to each other and to reported changes in the UHIE over time to identify patterns that could aid in forecasting the severity of the UHIE in various urban areas in the future. High resolution satellite imagery and data on the distance from roadways and inland water bodies are employed to predict the growth of the Greater Austin Region, and what that indicates about the expansion of the Urban Heat Island. With the consideration

of variables already examined in the past and the addition of urban sprawl analysis , a more general conclusion on the predictability of the UHIE can be derived based on developing a greater understanding of the behavior of metropolitan zones in Austin.

II. PROJECT OVERVIEW

A. Project Criteria

The following are Project Criteria decided at the beginning of the research process:

- Model's validation loss will be less than 2%
 - This criteria was not met. The validation losses were, on average, less than 4%.
- Model will yield predictions for average daily temperatures for summer 2025 (via colored map)
 - This criteria was not met due to data inconsistencies limiting the temporal range of predictions. There was not a strong enough foundation for either model to create a temperature map as far out as 2025. The time restrictions of the Synopsys Science Fair also prevented the achievement of this goal.
- Over 80% of output variance will be explained by model
 - This criteria was met by the Keras sequential model (Framework I).

B. Project Constraints

The following are Project Criteria decided at the beginning of the research process:

1. Will be entirely based on publicly available data (satellite imagery and CSV files)
2. Data will be analyzed using tools on ArcGis Online (less features than standard ArcGis)
3. Software must run in real-time, adapting to predictions as it is validated on more data
4. Regional & temporal data inconsistencies, narrow scope of factors and geographic region are limitations.

III. STUDY

A. Study Area

Its rapid industrialization and ongoing urban growth made the Greater Austin Region (30.267, -97.733) in Austin, Texas, an ideal site for which to forecast urban sprawl and analyze interactions of environmental factors with an emphasis on surface and land temperature.



Figure 1B: Greater Austin Region, Austin Texas, as defined by the coordinates: (30.267, -97.733).

According to the Köppen-Geiger Climate Classification [46, 47], this region is characterized by Humid Subtropical Climate [45], with evenly distributed precipitation exempting peaks at May, June, and October. This region experiences Southerly winds and low stratus clouds at night. The year with the highest recorded average temperatures was 2017, and the lowest recorded average temperatures were in 1899 [45]. A progressive increase in average temperatures is also evidenced by air and surface temperatures considered in Framework I.

The study focuses on this region for initial feasibility, but aims to extend these analyses to a national or global scale.

B. Framework I: Correlating Environmental Factors via Keras Sequential Model

The first technique employed was using machine learning to identify and confirm correlations between factors speculated to be closely linked to the Urban Heat Island Effect. This framework focused on the following

factors: air temperature (primary), land surface temperature (primary), relative humidity, soil moisture, population growth. Accurate and sufficient data for vegetation cover was initially intended to be used, but was unavailable.

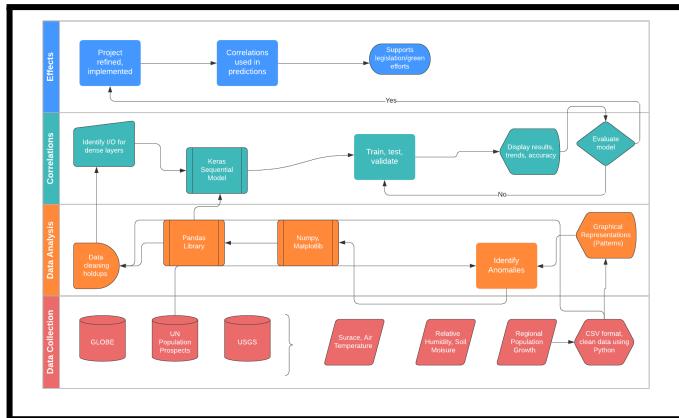


Figure 2: Conceptual Model of Framework I (from down to up).

(a) Data Collection & Constraints

The GLOBE Advanced Data Access Tools [49] and visualization system was used to retrieve data on air and surface temperatures, relative humidity, and soil moisture in CSV format.



Figure 3: GLOBE Data Visualization System.

The study also utilized global monthly air temperature and humidity data, courtesy of the United States Geological Survey [48]. The 2019 Revision of the United Nations World Population Prospects [50] database was used to access population growth trends and projections, and the 2018 Revision of the World Urbanization Prospects [51] database was used to access information on annual rate of change of percentage urban (global) and annual urban population at mid-year globally. This allowed for a cross-functional analysis to assist in future urban planning, environmental resource

management, and determining the health and habitability of an urban zone.

(b) Data Analysis

There were many inconsistencies and “Not a Number” values in the data. The data was manually and automatically imported into Jupyter Notebooks on Google Colaboratory in CSV file format, and was cleaned using the NumPy and Pandas software libraries. These helped manipulate dataframes and plot data. Visualizations were conducted using Matplotlib and Seaborn to perform initial analyses on the nature and scope of the data before making correlations and predictions.

```

atemp.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 571 entries, 1 to 571
Data columns (total 14 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   organization_id  571 non-null    object  
 1   org_name          571 non-null    object  
 2   site_id           571 non-null    object  
 3   site_name         571 non-null    object  
 4   latitude          571 non-null    object  
 5   longitude         571 non-null    object  
 6   elevation         571 non-null    object  
 7   measured_on      571 non-null    datetime64[ns]
 8   airtemps:userid   571 non-null    float64 
 9   airtemps:measuredat 571 non-null    object  
 10  airtemps:solarmeasuredat 571 non-null    object  
 11  airtemps:currenttemp(degC) 571 non-null    float64 
 12  airtemps:comments  342 non-null    object  
 13  airtemps:globeteams  0 non-null    float64 
dtypes: datetime64[ns](1), float64(3), object(10)
memory usage: 66.9 KB
  
```

Figure 4: Pandas Dataframe Summary.

(c) Correlations

Identifying correlations is a critical step in making predictions. Knowing the relationship between factors enables us to fill in gaps in data. For example: if there is a gap in air temperature data around February 2015, but we know the surface temperature and relative humidity at the time, and have an understanding of the link between these variables based on testing over granular data collected over many years, we may be able to accurately calculate an air temperature value that is close to the actual value.

```

Real Output <=> Predictions (First 10 values)
[10.2] <=> [10.19517]
[10.2] <=> [10.24102]
[9.4] <=> [10.206987]
[10.4] <=> [11.291314]
[10.4] <=> [11.104922]
[9.8] <=> [10.973987]
[13.7] <=> [11.557368]
[13.8] <=> [10.897681]
[11.9] <=> [11.373532]
[11.9] <=> [12.201732]
Model: "sequential"

```

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 32)	96
dense_1 (Dense)	(None, 10)	330
dense_2 (Dense)	(None, 1)	11

Total params: 437
Trainable params: 437
Non-trainable params: 0

Figure 5: Model's validation losses (discrepancy between verified data and predicted output) < 4%.

In deep learning, a layer is a container that receives weighted input and transforms it using a set of nonlinear functions, passing the output of that layer to the next. A sequential model uses a stack of layers, and the Keras library provides this functionality in Python, within a Jupyter Notebook.

A network of Keras sequential models was used to make preliminary predictions about the past and present, and compare them to actual data points. The relationship between different environmental factors was quantified in the form of correlation coefficients.

Figure 6: Loading Keras into Jupyter Notebook before building model

Once data gaps were filled in, this cleaned data, along with recognized correlations and patterns were employed in a prototype artificial neural network model (ANN) to predict future temperatures.

The second, simultaneous, part of the study was studying changes in past land cover land use to speculate changes in future urban sprawl in this region. This framework used various Geographic Information Systems such as QGIS, ArcGIS, NextGIS, and DivaGIS to better understand the geospatial implications of the Urban Heat Island Effect.

(a) Data Collection

For the purpose of this research study, we used QGIS, an open source satellite imagery software. In order to predict future land use changes, we used the Modules for Land Use Change Simulations (MOLUSCE) plugin developed by Asia Air Survey and NextGIS.

We imported data on inland water bodies, major roadways, and the distance from these water bodies and roadways in the study area. Data was sourced from DIVA-GIS and rasterization was conducted on ArcGIS.

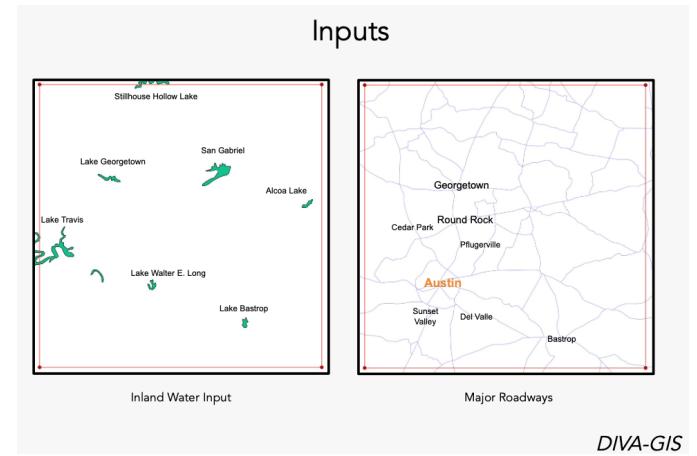


Figure 7: Input roadways and inland water body data from DIVA-GIS

We then rasterized and used the Euclidean distance function in ArcGIS on the input data. (Rasterization is taking scene geometry and converting it into pixels for display on a screen.)

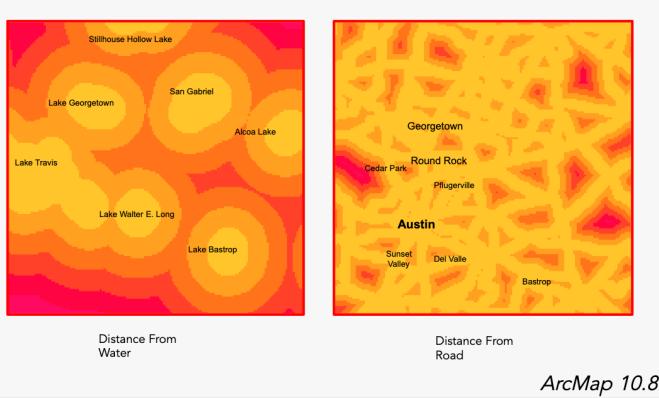


Figure 8: Using the Euclidean Distance form on input data in ArcGIS.

(b) Prediction Model

We used rasters of land use categories from two different time periods (and repeated over many different combinations of time periods) and rasters of explanatory variables: distance from roadways and distance from inland water bodies.

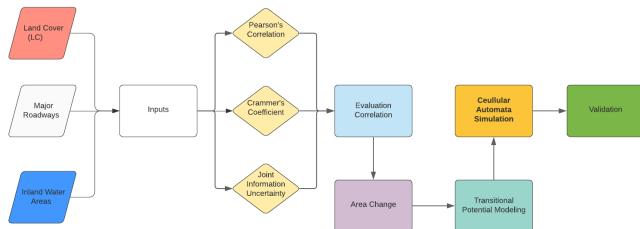


Figure 9: Conceptual Model of Framework II (from left to right).

Next, we input the data into an evaluation correlation. We used Cramer's Coefficient, to measure the strength of association between the different land cover classifications. The prediction model then created a transitional matrix and calculated the individual pixels in area-changes between the classifications. This was the basis for an ANN-based transition potential model finding confidence and probabilities of certain pixels changing.

We finally input the data into a cellular automata simulation with sets of pixels being classified into neighborhoods and land use change predictions based on these clusters.

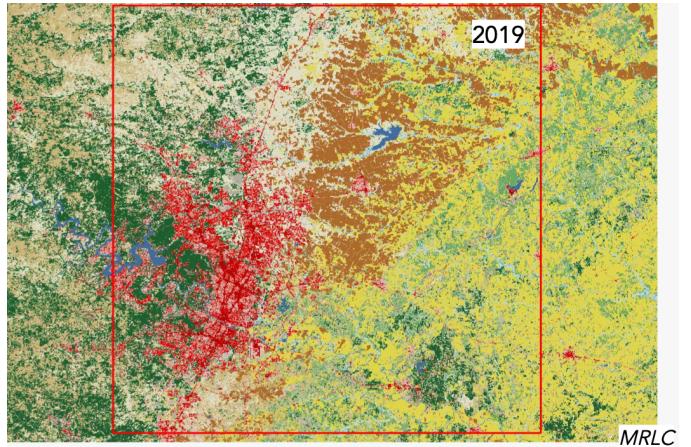


Figure 10: Sample 2019 map prior to simulation (sourced from Multi-resolution Land Characteristics Consortium).

IV. RESULTS & ANALYSES

A. Notable Outputs

The Keras model displayed a steady increase in the Austin air and surface temperatures over time, and provided us with a functional algorithm for calculating the temperature in the future. The model's validation losses were consistently below 4% or 0.04. Over 80% of output variance was explained by the model.

The QGIS visualization predicted the Urban Heat Island sprawl in 2022, based on the increase in temperatures and changes in other environmental factors found via Framework I.

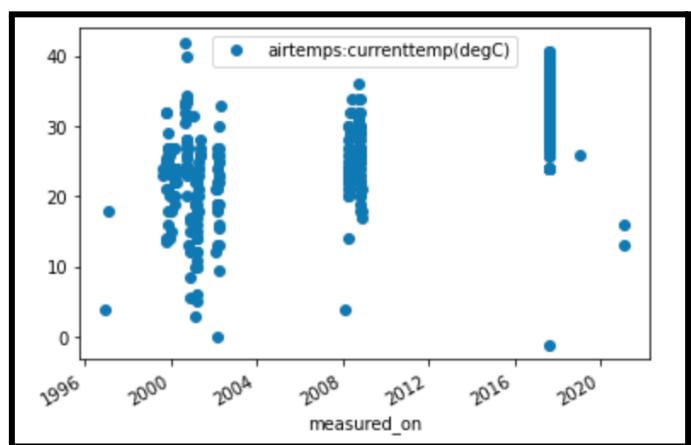


Figure 10: Steady increase in air temperature (and other environmental factors) in the Greater Austin Region.

B. Corroboration

By using the outputs of Framework I in conjunction with the prediction model in Framework II, it is evident that an

increase in certain environmental factors yields a sprawl of UHIE.

Though a direct causation has yet to be investigated, there exists a correlation between land cover-land use and air and surface temperature in the Greater Austin Region.

1. Limitations & Variation

- Lack of complete data (temporal inconsistencies, Not a Number values)
- No access to ArcGIS Desktop, used free student version of software
- Variation in temperature and precipitation changes over time
- Lack of knowledge about future climate change in study site in response to human activity
- Lack of understanding of region-specific policy (and implications) regarding response to global temperature change

V. DISCUSSION

A. Extension & Scaling

This technology can be adapted for use in cities across the United States, especially ones with established urban infrastructure. The availability of satellite imagery and city blueprints would be foundational tools to allow for the scaling of this software and analysis.

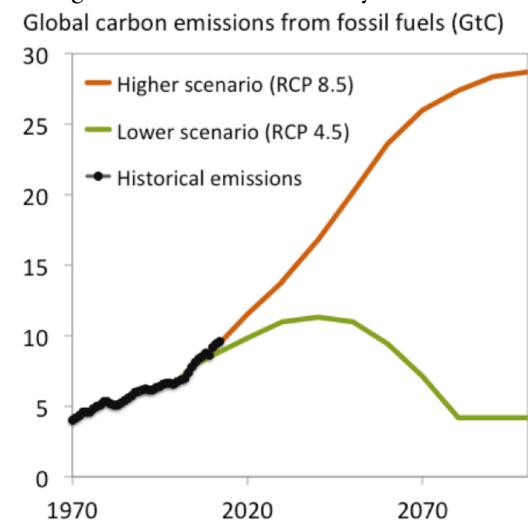


Figure 11: IPCC's Fourth Assessment Report (Hegerl et al 2007)

B. Uses & Impact

Science cannot work alone. With the aid of citizen scientists, policymakers, and ample time and legislation, there are various impactful implementations of this technology.

- Guide local city planners in counteracting the UHIE

- Pass legislation to decrease energy consumption and carbon emissions in metropolitan areas
- Create designated green spaces and preserve urban greenery
- Reduce disparity in relative urban temperatures
- Slow the sprawl and spread of urban canyons and urban heat islands

VI. ACKNOWLEDGEMENTS

We are immensely grateful for all of the support we have received from the NASA SEES UHIE intern team, the GLOBE Mission Earth team, and the following researchers and postdoctoral fellows at the University of Toledo: Dr. Kevin Czajkowski, Sara Mierzwiak, Orlarwale Oluwafemi, and Dr. Yitong Jiang.

VII. REFERENCES

- [1] <https://www.sciencedirect.com/science/article/pii/S1877705816332039>
- [2] <https://www.tandfonline.com/doi/abs/10.1080/01431169208904271>
- [3] <https://www.sciencedirect.com/science/article/abs/pii/S0034425703003390>
- [4] <https://ieeexplore.ieee.org/document/7079467>
- [5] <https://www.tandfonline.com/doi/abs/10.1080/0143116031000116417>
- [6] <https://www.ingentaconnect.com/content/asprs/pers/2003/00000069/00000009/art00011>
- [7] <https://www.epa.gov/pm-pollution/particulate-matter-pm-basics>
- [8] <https://www.sciencedirect.com/science/article/abs/pii/S1352231021000194>
- [9] <https://www.sciencedirect.com/science/article/abs/pii/S135223101930706X>
- [10] http://www.paternott.com/pdf/Oake1982_UHI.pdf

- [11] <https://www.sciencedirect.com/science/article/abs/pii/S0924271617300035>
- [12] <https://link.springer.com/article/10.1007%2Fs11356-018-697-0>
- [13] <https://www.sciencedirect.com/science/article/abs/pii/S1352231018304242>
- [14] <https://ieeexplore.ieee.org/abstract/document/7552814>
- [15] [https://www.mdpi.com/2072-4292/10/12/2034.htm](https://www.mdpi.com/2072-4292/10/12/2034)
- [16] <https://austintexas.gov/news/austins-population-continues-another-decade-growth-according-us-census-bureau-0>
- [17] <https://www.mdpi.com/2073-4433/11/9/926>
- [18] <https://www.nature.com/articles/srep12467>
- [19] <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2010RG000345>
- [20] <https://www.annualreviews.org/doi/abs/10.1146/annurev.es.23.110192.000351?journalCode=ecolsys.1>
- [21] <https://www.sciencedirect.com/science/article/pii/S1878029610000952>
- [22] <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2008JD010035>
- [23] <https://www.sciencedirect.com/science/article/pii/S2667010021001712>
- [24] <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.622.3649&rep=rep1&type=pdf>
- [25] <http://dx.doi.org/10.1016/j.isprsjprs.2017.03.014>
- [26] <https://doi.org/10.1175/2009JCLI2900.1>
- [27] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6950409/>
- [27] <https://storymaps.arcgis.com/stories/04ee5b88bca141b08abb7c5a746a52ec>
- [29] <https://www.mrlc.gov/data>
- [30] <https://www.weather.gov/media/ewx/climate/ClimateSummary-ewx-Austin.pdf>
- [31] <https://earthexplorer.usgs.gov/>
- [32] <https://www2.purpleair.com/>
- [33] <https://landsat.gsfc.nasa.gov/satellites/landsat-8/>
- [34] https://www.researchgate.net/publication/335176595_Land-UseLand-Cover_Changes_and_Their_Impact_on_Surface_Urban_Heat_Islands_Case_Study_of_Kandy_City_Sri_Lanka
- [35] https://earthobservatory.nasa.gov/features/MeasuringVegetation/measuring_vegetation_2.php#:~:text=NDVI%20is%20calculated%20from%20the,nd%20less%20near%2Dinfrared%20light.
- [36] <http://se.asee.org/proceedings/ASEE2000/Yeh01.pdf>
- [37] <https://www.mdpi.com/2072-4292/12/13/2138>
- [38] <https://doi.org/10.5194/hess-24-2577-2020>
- [39] <https://doi.org/10.3390/rs12091433>
- [40] <https://doi.org/10.1016/j.jhydrol.2020.124905>
- [41] <https://doi.org/10.1016/j.scitotenv.2019.135148>
- [42] <https://www.epa.gov/heatislands>
- [43] https://www.ipcc.ch/site/assets/uploads/2018/03/ar4_wg2_full_report.pdf
- [44] <https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-wg1-chapter9-1.pdf>
- [45] <https://www.weather.gov/media/ewx/climate/ClimateSummary-ewx-Austin.pdf>
- [46] <https://www.globe.gov/documents/358135/359681/Koppen-Geiger+Guide>
- [47] <https://www.nature.com/articles/sdata2018214>
- [48] <https://earthexplorer.usgs.gov>
- [49] <https://www.globe.gov/es/globe-data/retrieve-data>
- [50] <https://population.un.org/wpp/>
- [51] <https://population.un.org/wup/>