

NASA DEVELOP National Program
Arizona - Tempe
Spring 2020



Philadelphia Health & Air Quality
Assessing Land Surface Temperature, Vegetation Cover, and
Compounding Vulnerability Factors to Identify High Priority
Areas for Cooling Initiatives in Philadelphia, Pennsylvania

DEVELOP Technical Report
Final Draft - April 2nd, 2020

Brandy Nisbet-Wilcox (Project Lead)

Samuel Meltzer

Spencer Nelson

Charlotte Wagner

Dr. David Hondula, Arizona State University (Science Advisor)

1. Abstract

Heat is the leading cause of weather-related deaths in the US, with heat-related hospitalizations increasing by 2-5% between 2001-2010. In Philadelphia alone, 137 heat-related deaths were recorded between 2010-2018, while a total of 18 daily temperature records have been set since 2010. Temperature is relatively higher in cities compared to rural areas, a phenomenon known as the urban heat island effect. This effect exaggerates daytime maximum temperatures and nighttime heat retention in urban areas, which increases heat exposure in urban environments and especially impacts vulnerable populations. Vulnerability to heat-related illnesses is determined by a combination of risk factors, such as demographics, socioeconomic status, and preexisting health conditions. This project supported the Philadelphia Department of Public Health and Office of Sustainability by identifying priority areas for cooling interventions, such as heat danger educational outreach and urban tree planting. The team developed heat vulnerability scores for each census tract within Philadelphia. Remotely sensed land surface temperature, normalized difference vegetation index, normalized difference built-up index, normalized difference water index, and albedo data were calculated from Aqua Moderate Resolution Imaging Spectroradiometer and Landsat 8 Operational Land Imager/Thermal Infrared Sensor instruments. These variables were weighted against socioeconomic variables and preexisting health conditions using a principal component analysis. A total of 74 census tracts clustered were identified as high-risk areas for heat-related illnesses. 15 of these census tracts also had very low tree density (lower 20th percentile) and should be targeted for tree planting initiatives. The findings of this project will help target interventions to mitigate heat-related health issues and improve the overall wellness of Philadelphia residents.

Keywords

urban heat, heat vulnerability, land surface temperature, NDVI, NDBI, NDWI, albedo, remote sensing, principal component analysis

2. Introduction

2.1 Background Information

Heat-illnesses are a leading cause of weather-related deaths in the US. Heat-related hospitalizations increased by 2-5% between 2001 and 2010 (Centers for Disease Control and Prevention, 2019). Heat-related illnesses range from minor symptoms like heat cramps and heat exhaustion, to death from severe heat stroke, ischemic heart disease, diabetes, stroke, and respiratory diseases (Lugo-Amador, Rothenhaus, & Moyer, 2004; McGeehin & Mirabelli, 2001). In Philadelphia alone, 137 heat-related deaths were recorded between 2010 and 2018. A total of 18 daily temperature records have been set in Philadelphia since 2010 (Philadelphia Office of Sustainability, 2019).

Cities in the northeastern US are expected to experience the most severe increase in heat-related morbidity and mortality due to increasing frequency and severity of heatwaves (McGeehin & Mirabelli, 2001). The relatively greater temperatures in cities compared to rural areas, known as the urban heat island (UHI) effect, aggravate daytime maximum temperatures as well as nighttime heat retention (Oke, 1982). This increases the overall heat exposure of urban populations. Certain population groups are more susceptible to heat-related health effects. Risk factors

include demographic and socioeconomic factors as well as preexisting health conditions (Bao, Li, & Yu, 2015; Green et al., 2019).

Air temperature in cities may differ widely across neighborhoods as a result of the geographic distribution of the built environment and greenness (Harlan, Declet-Barreto, Stefanov, & Petitti, 2013). Relative humidity and wind speed also affect physiological heat exposure (Urban & Kysely, 2014; Steadman, 1984). However, these latter two factors cannot be directly measured with satellites at the spatial resolution needed to identify intraurban variability and ground-based measurements are also too sparse. Many studies of heat-related morbidity and mortality have used satellite-based images of land surface temperature (LST) as a proxy for air temperature since it can be obtained at high spatial resolution (Hondula et al., 2014). Satellite-retrieved normalized difference vegetation index (NDVI), normalized difference built-up index (NDBI), and albedo are sometimes added to analyses to better represent urban temperatures (Davis, Jung, Pijanowski, & Minor, 2016). More recently, normalized difference water index (NDWI) has been included to provide a more robust measure of vegetation health and heat exposure (Mushore et al., 2018).

Open green spaces such as parks and mixed urban/forested areas in cities reduce urban air temperatures (Davis, Jung, Pijanowski, & Minor, 2016; Hondula et al., 2018). Increasing urban green spaces is considered an important policy tool to reduce urban temperatures and increase heat resilience. Heat vulnerability indices integrate a combination of the above variables. They can be used to identify the areas that experience the highest temperatures and are home to population groups that are most vulnerable to heat-related illnesses.

2.2 Project Partners & Objectives

The Philadelphia Department of Health and the City of Philadelphia, Office of Sustainability, are concerned with the effects of extreme heat events in Philadelphia. Their objective is to protect vulnerable populations in Philadelphia which are disproportionately affected by urban heat and heat-related illness. They are also concerned that green spaces and street trees are unequally distributed throughout the city, further aggravating heat exposure among sensitive populations. The Philadelphia Department of Health currently provides a publicly available heat vulnerability index map using ArcGIS Online. The map displays priority areas and satellite retrieved LST covering 2013-2015, based on previous work at Arizona State University (Hondula, 2012). As the Department continues to develop its climate preparedness plans, it is important that they have access to current data covering different determinants of heat exposure. This project updates the socioeconomic, health, and environmental variables in the vulnerability index, and adds additional environmental factors that provide more detailed information on how heat exposure varies throughout the city. This data will allow the City of Philadelphia to identify priority areas for public health outreach, target vulnerable communities during policy development, and prioritize areas for cooling initiatives. In addition, this project identifies vulnerable census tracts with especially low tree density. Street tree planting is a cooling initiative of particular interest to the Office of Sustainability. They have an expanding tree planting program and are looking for key areas in the city that may be adversely

subjected to extreme heat-related health effects and would benefit the most from new trees (City of Philadelphia, Office of Sustainability, 2019).

3. Methodology

3.1 Data Acquisition

Census tract-level socioeconomic data for Philadelphia County were retrieved from the 5-year datasets produced by the 2018 American Community Survey using the Tidycensus package in R (Appendix Table A1). Prior public health research in Philadelphia and other urban contexts identified twelve variables as key indicators for heat sensitivity (Hondula et al., 2012; Johnson & Wilson, 2009; Uejio, Wilhelmi, Golden, Mills, Gulino, Samenow, 2011; Weber et al., 2015). These socioeconomic variables include percent of population age 65 and older, percent of population age 25 and older with less than a high school education, percent of population living in limited English-speaking households, percent of non-white population (including Hispanic/Latino), percent of population 65 and older living alone, and percent of population below the federal poverty level.

Data on population health status were downloaded from the 2018 Center for Disease Control and Prevention (CDC) 500-cities data set. Variables included hypertension, asthma, coronary heart disease, chronic obstructive pulmonary disease (COPD), diabetes, and obesity, which are correlated with heat-related illness and mortality (Harlan, Declet-Barreto, Stefanov, & Petitti, 2013). GIS-friendly data at the census tract level were downloaded from the CDC website for the whole US and subsetted to Philadelphia. Satellite data were retrieved using Google Earth Engine. LANDSAT/LC08/C01/T1_SR USGS Landsat 8 Surface Reflectance Tier 1 images, MYD09GA.006 MODIS/Aqua Surface Reflectance Daily L2G Global 1km and 500m images, and MYD11A1.006 MODIS/Aqua Land Surface Temperature and Emissivity Daily Global 1km images were downloaded for 2017-2019. A detailed overview of the datasets used in this analysis is shown in Table A1.

3.2 Data Processing

Each socioeconomic variable was calculated as a percent population based on the total number of respondents per question. The age, education, and language variables required pooling several sub-categories (age categories by sex, education by grade level, language by language family) before normalization. Socioeconomic and health status data sets were then joined in R by their unique geographical identifiers to create one dataset with all 12 socioeconomic and health indicators.

In addition to socioeconomic, the six environmental variables of LST, NDVI, NDBI, NDWI, daytime and nighttime LST, and albedo were included in the analysis. Images spanning a three-year period from 2017-2019 were included to increase data availability and confidence in results in a similar manner to the multi-year averages used in the American Community Survey (ACS) data set. The image collection was further filtered by season. Only images captured during the northern hemisphere summer (June 1 - September 1) were included to reflect the period of concern for heat-related health effects and maximize intra-urban temperature heterogeneity.

The most recent available 3-year summer averages of LST, NDVI, and NDBI were retrieved from Landsat 8 Surface Reflectance data collected using the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). These data are available at sufficiently high resolution (30 m and 100 m) for intra-urban analyses (Alhawiti & Mitsova, 2016; Kaplan, U. Avdan, & Z. Avdan, 2018). Cloud and cloud shadow, specified in the quality assurance band of the Landsat 8 Surface Reflectance dataset, were excluded from the image collection. The final Landsat 8 image collection of images from June 1-September 1, 2017-2019, included 16 images and about nine pixels per location on average.

Landsat 8 does not collect nighttime images. Thus, Aqua MODIS images were used to collect nighttime temperature in Philadelphia. Aqua MODIS provides a Land Surface Temperature and Emissivity Daily product with LST in Kelvin at 1km resolution, which was restricted to night images only for this work. Pixels where LST was not produced due to cloud effects or other reasons than clouds, as well as pixels with an error of >3 degrees Kelvin (per QC Night quality indicators), were removed from the image collection. The final image collection included 278 images with about 80 pixels per location. NDWI was derived from the Aqua MODIS Surface Reflectance Daily L2G Global dataset with a resolution of 500 meters. Data was limited to pixels produced at ideal quality across all bands (per 500-meter Reflectance Band Quality). The final image collection used to calculate NDWI comprised of 279 images with ~260 pixels per location.

3.3 Data Analysis

Daytime LST, NDVI, NDBI, NDWI, and albedo were calculated from thermal, near-infrared, and infrared wavelengths. LST is a function of brightness temperature (band 10 in the Landsat 8 Surface Reflectance product) and emissivity, a measurement of an object's ability to emit infrared energy (Jin & Liang, 2005; Zhang, Wang, Li, 2006). LST can be described by the following, Equation 1:

$$LST = \frac{BT}{(1 + (0.0000109 * (BT/0.01438) * \ln(E)))}$$

where BT is brightness temperature in Kelvin and E is dimensionless emissivity (Kumar, Bhaskar, & Padmakumari, 2012).

Albedo is the fraction of incident radiation that is reflected by a surface. It can be represented by Equation 2:

(1)

$$\text{Albedo} = 0.246 \text{ Blue} + 0.146 \text{ Green} + 0.191 \text{ Red} + 0.304 \text{ NIR} + 0.105 \text{ SWIR1} + 0.008 \text{ SWIR2}$$

where Blue, Green, Red, NIR, SWIR1 and SWIR2 are bands 2, 3, 4, 5, 6, and 7, respectively, in Landsat 8 Surface Reflectance images and the coefficients are empirically derived weighting coefficients (Olmedo, Ortega-Farías, de la Fuente-Sáiz, Fonseca- Luego, & Fuentes-Peñailillo, 2016).

(2)

NDVI is the normalized difference between near-infrared (NIR) and red reflectance (Red) following Equation 3:

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}}$$

where NIR and Red correspond to band 5 and band 4 in the Landsat 8 surface reflectance product, respectively (Orusa & Mondino, 2019). (3)

NDBI is the normalized difference between shortwave infrared (SWIR) and near-infrared (NIR) which are stored in bands 6 and 5, respectively and can be represented by Equation 4:

$$\text{NDBI} = \frac{\text{SWIR} - \text{NIR}}{\text{SWIR} + \text{NIR}}$$

NDWI is a measure of surface water content in plants and represents vegetation health when combined with NDVI. It is the normalized difference between the high plant reflectance wavelengths of 0.86- μm and 1.24- μm (Gao et al., 1996). In Aqua MODIS surface reflectance images these wavelengths are contained in the near-infrared band 2 and the short-wave infrared band 5, and NDWI is calculated using Equation 5:

$$\text{NDWI} = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWIR}} \quad (5)$$

For all six variables, we calculated median values for each location, as the number of pixels available for most locations was too low to assume a normal distribution of the data and calculate a sample mean. Mean values were then calculated for each census tract. Nighttime LST did not require a calculation as the band is already available in the Aqua MODIS product. The NDVI, NDBI, NDWI, day and nighttime LST, and albedo variables were added to the image collection and exported from GEE as a shapefile.

Based on the above variables, we developed three indexes to inform heat management policies: 1) a Heat Sensitivity Index (HSI, health and socioeconomic variables only), 2) a Heat Exposure Index (HEI, environmental variables only) and 3) a Heat Vulnerability Index (HVI, all variables) using Principal Component Analysis (PCA). PCA condenses complex patterns across variables into main drivers of data variability. These linearly uncorrelated variables are called principal components. As such, PCA can be used to identify spatial clusters of key factors driving heat vulnerability. The HSI includes the 12 socioeconomic and health variables described previously, and the HEI is derived from the six environmental variables described above. All 18 variables were included in the overall HVI.

PCA was performed on the input variables using R. The same PCA procedure was followed for all three indices. Prior to the PCA, input variables were transformed to z-scores (mean of 0, standard deviation of 1) to improve numerical stability and comparability between variables with widely differing units and ranges. PCA was done using the principal-function in base R with a varimax rotation. Only principal

components that had eigenvalues greater than one were retained (Kaiser, 1960). Components also were required to cumulatively explain approximately 80% of data variance and have eigenvalues higher than those that would have been generated by random data sets with the same number of variables and observations (Horn's method) using the paran-package in R (Glorfeld, 1995). Components were further assessed for their scientific consistency. For example, certain chronic diseases are expected to co-vary spatially and these known patterns should be reflected in the components. We further evaluated whether the sign, positive or negative, of each component produced by the PCA represented the best current scientific understanding of the real-world relationship between that variable and heat vulnerability. For example, the sign on the component related to LST would be expected to be positive, since an increase in temperature would produce an increase in heat vulnerability. However, since the PCA only detects patterns among input variables and is ignorant of the outcome variable of interest, this is not accounted for in the raw PCA output. We thus used scientific knowledge to determine whether the clustering of variables and sign on components identified by the PCA was plausible.

We then calculated scores for each component and census tract. Factor scores for each component and census tract were calculated by weighing the z-score of each variable by the 18x4 matrix product of the inverse correlation matrix of the data and the loadings resulting from the PCA. The resulting factor scores (number of census tracts times number of components) were summed to a single score. To create priority areas, the overall sensitivity index scores were then ranked by quintile, and classified as very low, low, medium, high and very high priority areas.

We then calculated the average street tree density for each census tract. To select priority areas for street trees in Philadelphia, which have the potential to reduce temperature and thus improve health and quality of life, areas with very high HVI (top 20%) were compared to street tree density. Street tree density by census tract was calculated by converting street segment vertices to center points. Length and number of trees as identified by the Department of Parks and Recreation were then summed by census tracts. The average number of street trees per 100 yards of street was calculated for each census tract, and binned into quintiles. The resulting street tree density map was overlaid onto the areas with very high heat vulnerability in ArcGIS Desktop 10.6.1. Areas that were in the upper quintile of the HVI and the bottom quintile of tree density were identified as tree initiative priority areas.

4. Results & Discussion

4.1 Principal Component Analysis Results

The 18 risk factors included in the Principal Component Analysis for the Heat Vulnerability Index (which includes socioeconomic, health, and environmental variables) clustered into four principal components which together explained 78 percent of variance observed in the data (*Figure 1*). The first component summarizes various socioeconomic and chronic health variables and has strong loadings on percent population below the federal poverty level, percent of non-white population as well as prevalence of asthma, COPD, hypertension, asthma, diabetes, and obesity. The second component is representative of heat exposure

and contains strong loadings for daytime LST, NDBI, NDVI, and NDWI. Following a scientific understanding of environmental risk factors of heat vulnerability, daytime LST and NDBI are anti-correlated with NDWI and NDVI. These two principal components alone explain 62 percent of overall variance observed in the data. The third and fourth principal components are representative of age-related health conditions and education, respectively, and explain an additional 16 percent of the observed variance.

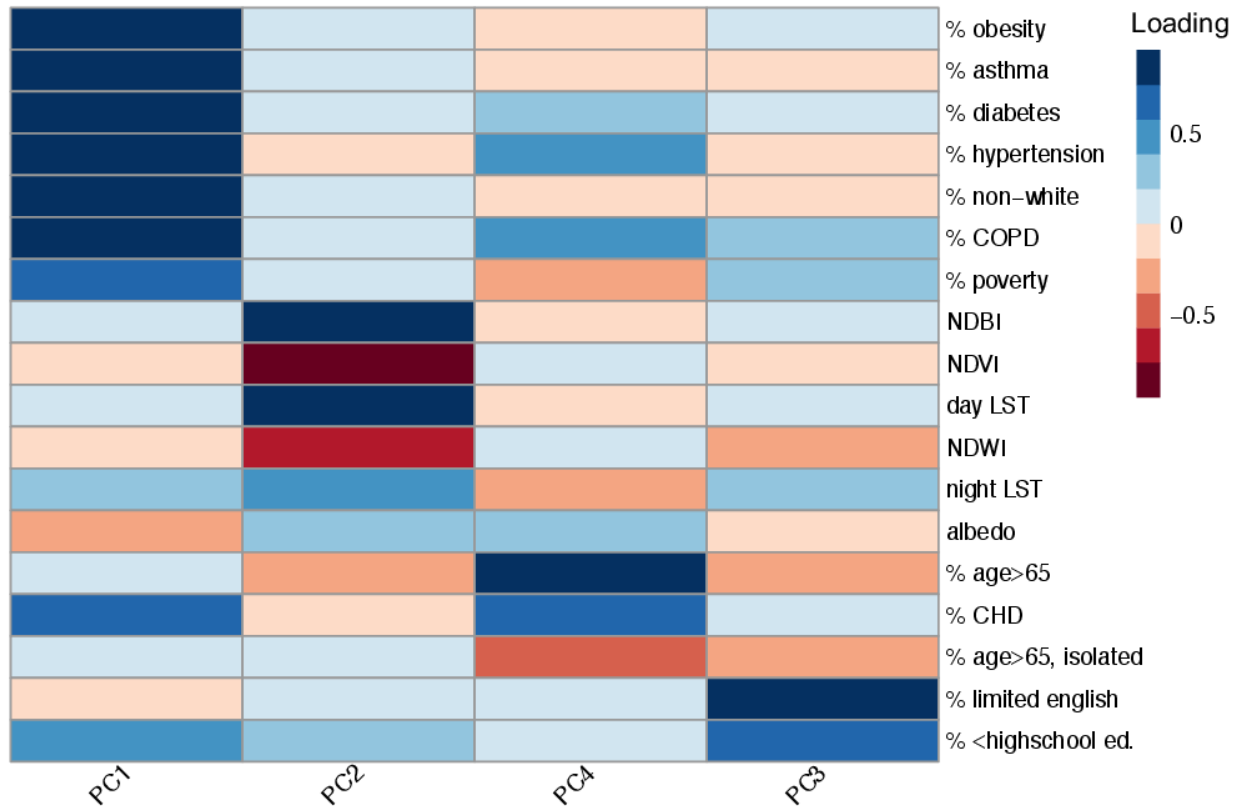


Figure 1. Loadings of Principal Component Analysis for 18 variables included in Heat Vulnerability Index. The principal component 1 (PC1) is driven by socioeconomic factors and chronic health conditions, PC2 is representative of heat exposure and driven by satellite-derived environmental variables. PC3 is representative of old age and age-related illnesses and PC4 represent socio-cultural drivers of heat vulnerability. Together the 4 components explain 78 percent of data variance.

The Heat Sensitivity Index and Heat Exposure Index allow for an in-depth and separate examination of the relationships of variables included within the two variable subgroups included in this analysis, namely the socioeconomic and health variables subgroup and the environmental variables subgroup. Accordingly, the resulting components in each of these indexes mirror the four components described in *Figure 1* but show additional weaker components that were not included. The Heat Sensitivity Index consists of three principal components, which explain 83 percent of data variance (*Figure A1*). The first component mirrors the socioeconomic and chronic health component in the Heat Vulnerability Index closely and explains 51 percent of data variance. The second component reflects

the third component in the heat vulnerability index, and comprises age and age-related health conditions, namely hypertension and coronary heart disease. The third component can be described as socio-cultural and has strong loadings on limited English-speaking households, less than high school education, and percent population below the federal poverty level. The Heat Exposure Index consists of two principal components, which together explain 81 percent of data variance (Figure A1). The first component consists of loadings for daytime LST, NDBI, NDVI, and NDWI of greater than absolute (0.8) and a loading of 0.6 on nighttime LST. The second component relates to heat retention and is driven by nighttime temperature and albedo.

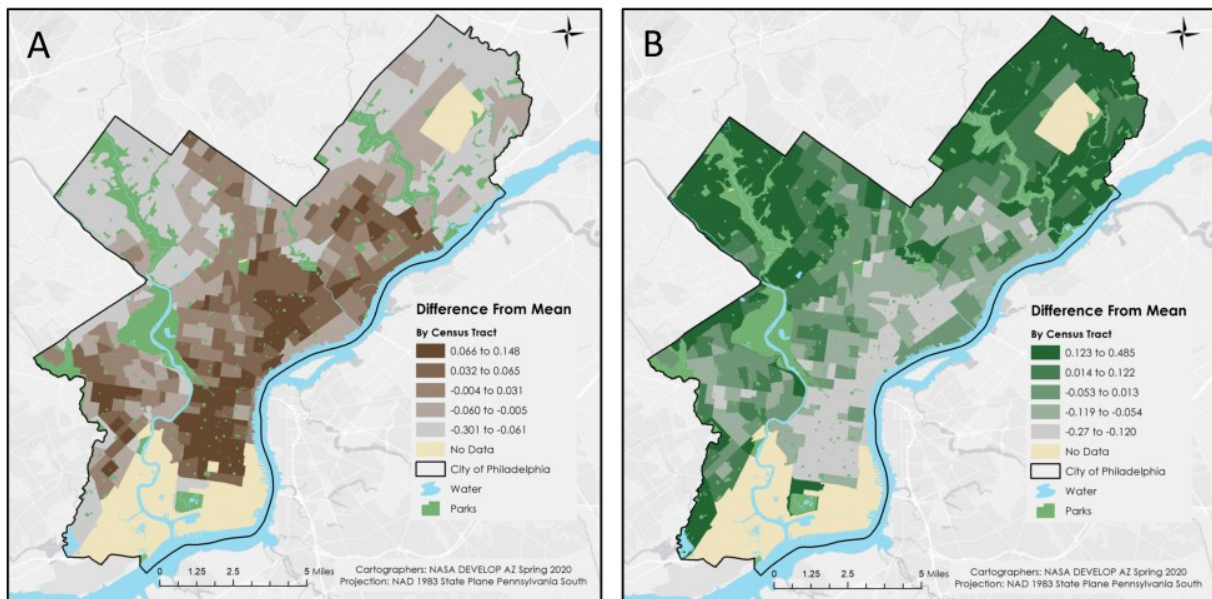


Figure 2. NDBI and NDVI in Philadelphia. Map A (left) shows NDBI in Philadelphia by census tract as a difference from the city mean. Map B (right) shows NDVI in Philadelphia by census tract as a difference from the city mean. Both maps are displayed in quintiles.

The strong inverse relationship between NDBI and NDVI indicated by the loadings from the PCA are confirmed in the raw data (see *Figure 1*). In *Figure 2*, it is evident that the spatial concentrations of high NDBI are correlated to areas of low NDVI, such as in Downtown and South Philadelphia. This can be explained by the high concentration of impervious surfaces in Downtown and South Philadelphia, such as buildings, and also the lack of yards and trees. Moreover, areas of particularly low NDBI are associated with areas of high NDVI, such as in the Far Northeast and the Germantown/Chestnut Hill areas in western Philadelphia. This can be explained by the low density of buildings and houses, as well as greater green spaces from parks and yards.

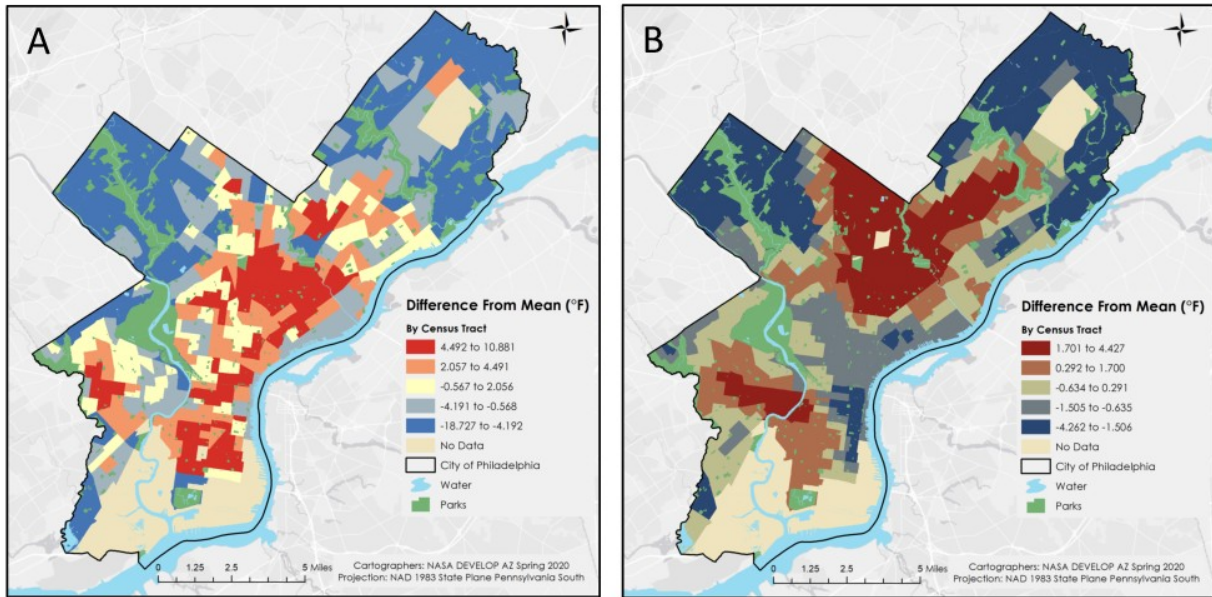


Figure 3. Daytime LST and Nighttime LST in Philadelphia. Map A (left) shows the difference from Philadelphia’s mean daytime LST in degrees Fahrenheit. Map B (right) displays nighttime LST as a difference from the city mean. Both maps are displayed in quintiles. Note the variance of temperature is much larger on the daytime LST map.

There were also some more unexpected results from the analysis. In the PCA of all 18 variables (heat vulnerability index), nighttime LST only showed a weak loading on the second component. In Figure 3, the noticeable spatial correlations and differences between daytime and nighttime LST are apparent. In both maps, it is clear that North Philadelphia has much higher day and night temperatures than most of the city, while those in Germantown/Chestnut Hill in western Philadelphia and the Far Northeast appear to have lower temperatures for both day and night. There are also several differences between the maps, as Downtown and the area north of it appear to be cooler on average during the night, while it is average to hotter than average during the day. Areas close to Philadelphia’s northwestern boundary appear to be much hotter than average at night and vary greatly during the day. In the PCA of the environmental variables only (Heat Exposure Index, *Figure A1*), night LST and albedo are identified as a separate principal component, indicating that their spatial variance differs statistically from other environmental variables (in particular NDVI, NDWI, and NDBI). This could indicate that places with a low albedo, and thus low reflectivity, retain heat more successfully throughout the day, than places with high albedo such as Downtown. It is also to be noted that the range of daytime temperature difference from the city mean is more than 3 times higher than that of nighttime temperature differences from the city mean (10.881 to -18.727 vs 4.427 to -4.262 degrees Fahrenheit). While this indicates that nighttime does provide some leveling of temperatures across the city, the higher nighttime temperatures in certain areas may nonetheless preclude relief from the day’s heat in certain areas of the city.

4.2 Priority Maps

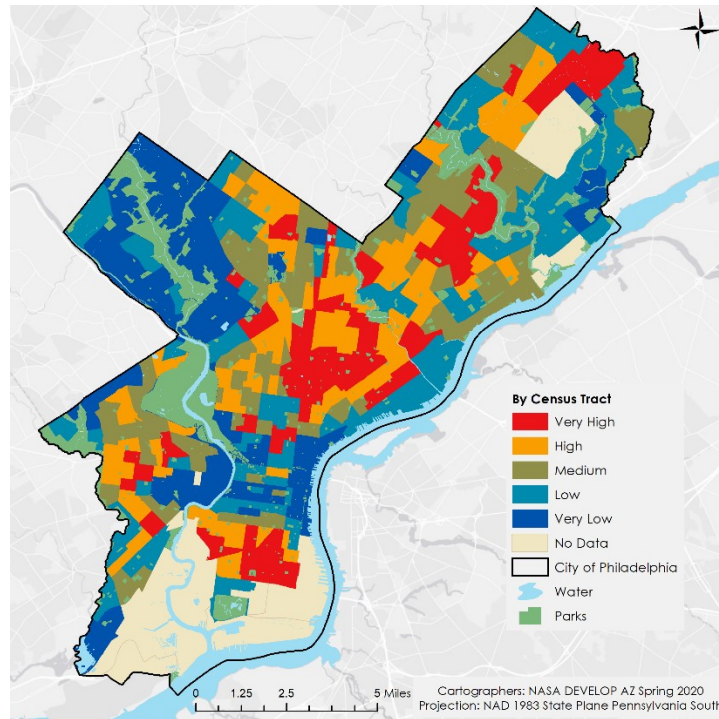


Figure 4. 2018 Heat Vulnerability Index for Philadelphia by census tract. The scores are categorized in quintiles. The Heat Vulnerability Index scores are based on all 18 socioeconomic, health and environmental variables.

Priority maps are categorized in quintiles as this will maintain the same number of census tracts in each category, making targeted efforts by policymakers more precise, especially as values and distributions change in future studies, the number of target areas will remain approximately the same (i.e. each category will always have approximately 75 census tracts). 74 census tracts were identified as very high heat vulnerability areas (Figure 4). They clustered in four areas of Philadelphia, the South, the Northeast, the Central Northeast, and the Far Northeast districts. The census tract with the highest heat vulnerability score (>2.5 standard deviations) was located in the Far Northeast district, but this area has very low population density and would likely not be a priority area for public outreach and cooling initiatives.

The South and Northeast districts, however, are some of the most densely populated areas in the US. The team's analysis indicates that these areas are subject to the highest concentration of socioeconomic and health risk factors resulting in very high heat vulnerability. Census tracts with very low heat vulnerability cluster in downtown and western Philadelphia. The Heat Sensitivity Index and the Heat Exposure Index identify the drivers of heat vulnerability clusters. Very high heat vulnerability in the South district is driven by high heat sensitivity (*Figure A2*). Heat exposure is medium to high in this area. Downtown, which has low to very low heat vulnerability due to its low very low to low heat exposure, nonetheless shows elevated heat sensitivity in the Heat Sensitivity Index. The Northeast district's high to very high heat vulnerability is a result of medium to very high heat sensitivity, and high to very high heat exposure. The cluster in the far northeast is primarily driven by elevated heat exposure (*Figure A3*). Some

census tracts in the southwest of Philadelphia also showed very high heat sensitivity, but had mixed scores in the heat exposure index.

The team also created a street tree density map of the census tracts deemed to have very high heat vulnerability. The map showed that the lowest quintile of census tracts for street tree density (bottom 20%) had on average less than 1.347 street trees per 100 yards of street. Among the 74 census tracts with very high heat vulnerability, 15 census tracts were also in the lowest street tree density category (*Figure 5*). Three of these 15 census tracts were located in the Central Northeast district and three were located in the South district, suggesting that street tree planting initiatives targeted at these areas may be particularly beneficial to local residents. In contrast, several areas with high heat vulnerability also have quite high street trees, such as in areas just north of Downtown, as well as a couple in the Far Northeast. These areas would likely benefit less from street tree planting to reduce vulnerability, and other options should be explored by policymakers to address their high vulnerability.

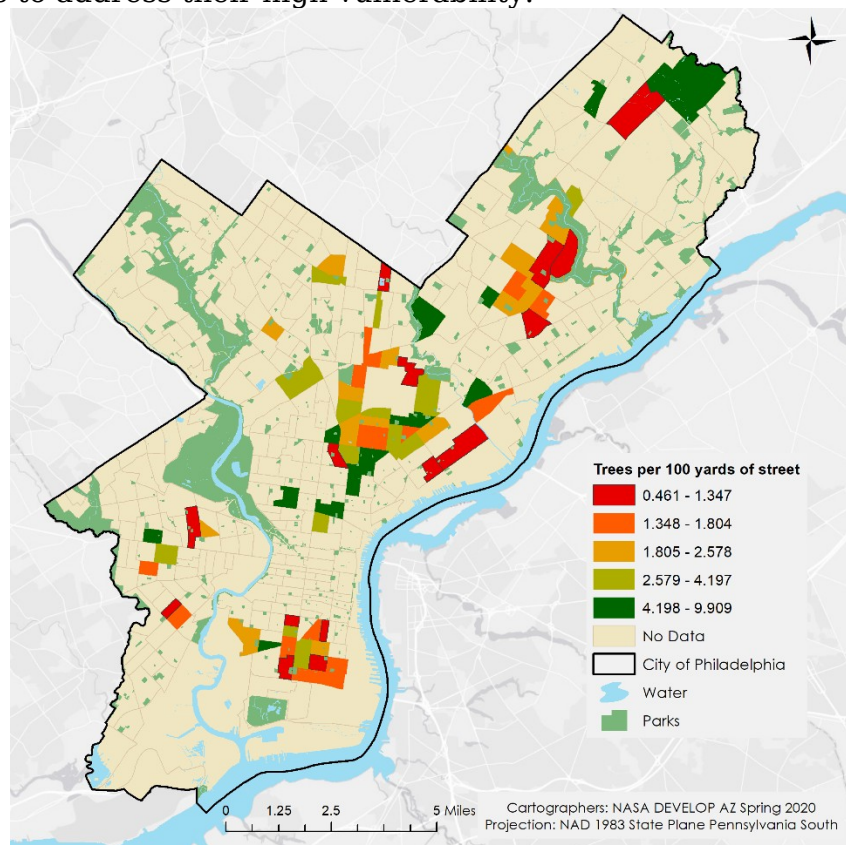


Figure 5. Density of street trees in census tracts with Very High Heat Vulnerability (top 20 percent). Street tree density is categorized in quintiles. Census tracts in the top category are in the lowest twenty percent of census tracts by street tree density and in the top 20 percent of heat vulnerability, and therefore priority areas for street tree planting.

4.3 Error Analysis/ Limitations

The heat vulnerability index that the team created was based on the twelve socioeconomic variables indicated by the partners at the City of Philadelphia as

being of most concern. There were many additional factors that could have been considered in the creation of this index, which could have impacted the strength and results of the analysis and potentially influenced which areas were selected as the highest priority areas for cooling interventions. Some of these factors could include additional diseases, younger age groups, or different measures of financial attainment such as income. Additionally, land surface temperature measures the temperature on the surface of the Earth. A different measurement, such as apparent temperature, would provide a better measure for what temperature people in the city experience by taking into account humidity and wind speed. Unfortunately, wind data collected at ground level across the city was not available for this calculation.

Another factor to be considered is that the ACS dataset leaves room for error because the spatial resolution is coarse at the census tract level and the data were a few years old. The CDC data were scaled down from block group resolution to census tract resolution to match the ACS dataset. Decennial Census data from 2010 would have been more informative in our analysis of priority areas by providing a finer spatial resolution for our analysis, but would have been much less current than the available CDC and ACS datasets. Another limitation is that this analysis does not represent true heat exposure, but rather assumes individuals remain near their homes. In reality, individuals move throughout the city during the day for work, school, or other activities. Whereas one person may be in an air-conditioned building at work and have a lower true heat exposure, a neighboring individual may have a higher true heat exposure by working outside. Heat exposure as represented in this work, however, will likely be more accurate for sensitive populations, like elderly or children, since these population groups are more likely to only move within a census tract throughout the day.

4.4 Future Work

This project could benefit from using 2020 Decennial Census data as Decennial Census data provides more accurate spatial representation of socioeconomic characteristics by census data. This would also allow for greater confidence in our estimate when determining priority areas based on the vulnerability index. Air quality data, including ozone and $PM_{2.5}$, could be added to the study if it were available at a finer resolution. The City is planning to do a large-scale air quality monitoring campaign in the near future covering several pollutants of interest. Once it is completed, these data could be included as an exposure variable to the vulnerability study. Temperature readings from ground stations throughout Philadelphia would help confirm temperature differences between neighborhoods which we estimated through satellite retrieved LST. Additionally, monitoring health outcomes, such as mortality rates, would help confirm whether the cooling initiatives based on this work are having a positive impact on resident wellness. They would also allow weighting our input variables in accordance with their importance for heat-related health outcomes, rather than weighting them equally. This would allow estimating actual health risk from heat, rather than just an assessment of spatial clustering in the variability of heat vulnerability indicator as provided by a PCA. All of these additional study components could improve confidence in the selection of priority areas for targeting public outreach, urban planning decisions, and implementing cooling initiatives, such as street tree planting initiatives.

5. Conclusions

This project demonstrates that NASA Earth observations can be used to assist a city by indicating which areas of the city would benefit most from public outreach, urban planning strategies, and cooling initiatives. The City of Philadelphia is concerned that vulnerable populations are disproportionately at risk for heat-related illness and knows that measures must be taken to curb the human health cost of increased intensity and frequency of extreme heat events in urban environments. They believe that those areas of the city that are inhabited by vulnerable populations and have the highest temperatures should be prioritized for public outreach and urban planning strategies and that the results of the NASA Earth observations analysis in this project will improve their ability to identify and prioritize such areas.

The results of the project indicate that the most densely populated residential areas of Philadelphia have the highest heat exposure. Average census tract summer daytime surface temperatures ranged from 75 to 105 degrees Fahrenheit. Four components resulting from the PCA explained 80 percent of the observed variance in the socioeconomic, health, and environmental variables included in this study. Based on the total scores derived from the PCA, seventy-four census tracts in Philadelphia were determined to have very high heat vulnerability and should be targeted by city planners for public outreach, urban planning strategies, and heat reduction measures. These census tracts are clustered in the South, Northeast, Central Northeast, and Far Northeast districts. A street tree planting program would be particularly beneficial in 15 census tracts which are home to communities with very high heat vulnerability and have few street trees.

The city previously created a vulnerability index web application to show which census tracts were most sensitive to extreme heat based on a variety of socioeconomic and health risk indicators, but the ability to assess heat exposure was limited to LST. The DEVELOP team updated this app using more recent data for socioeconomic and health indicators in addition to extending the data on heat exposure to up-to-date measurements of albedo, day LST, night LST, NDVI, NDBI, and NDWI data from NASA Earth observations. These observations together with the three priority indexes produced in this study will allow the partners to identify priority areas of heat sensitivity, heat exposure and overall vulnerability in Philadelphia, as well as provide them with access to detailed maps of the underlying variables. This will allow the partners at the City of Philadelphia to target public outreach at communities most affected by heat-related health issues, mitigate extreme heat in urban areas, and eventually decrease heat morbidity and mortality rates. NASA Earth observations are a useful and innovative tool to identify urban heat patterns for urban planning strategies such as cooling initiatives. The utilization of this technology by cities for decision making will result in improved welfare of residents and an overall healthier environment.

6. Acknowledgments

A special thanks to everyone at the Arizona - Tempe DEVELOP location for their support and guidance in the creation of this project.
Dr. David Hondula (Arizona State University)

Crystal Wespestad (NASA DEVELOP, Fellow/Lead)

Thanks also to our project partners who provided us with data and their expertise throughout this project.

Philadelphia Department of Public Health: Jason Hammer, Jessica Caum, and Alexandra Skula

City of Philadelphia, Office of Sustainability: Saleem Chapman

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors 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 NNL16AA05C.

7. Glossary

Centers for Disease Control and Prevention (CDC) – Federal public health agency in the United States.

Earth observations – Satellites and sensors that collect information about the Earth’s physical, chemical, and biological systems over space and time.

Land Surface Temperature (LST) – A measurement of how hot a surface on Earth would feel to the touch. A satellite measures the temperature of the first object that comes into its view when looking at the Earth

MODerate resolution Imaging Spectroradiometer (MODIS) – Imaging instrument aboard NASA’s Terra and Aqua satellites. MODIS has 36 spectral bands and views the entire Earth’s surface every 1 to 2 days.

Normalized Difference Built-Up Index (NDBI) – The quantifiable presence of urban land cover measured by the difference between shortwave infrared and infrared reflectance.

Normalized Difference Vegetation Index (NDVI) – The quantifiable presence of living vegetation measured using near-infrared and red-light reflectance.

Normalized Difference Water Index (NDWI) – An index estimating the surface water content of vegetation using near-infrared and shortwave infrared reflectance.

Operational Land Imager (OLI) – Imaging instrument aboard the Landsat 8 satellite. OLI can measure visible, near-infrared, and shortwave infrared wavelengths.

Thermal Infrared Sensor (TIRS) – Instrument aboard the Landsat 8 satellite. TIRS uses two thermal infrared wavelength bands in order to differentiate between the temperature of Earth’s surface and that of Earth’s atmosphere.

Urban Heat Island (UHI) effect – The phenomenon of relatively high temperatures in an urban area compared to the temperatures of surrounding rural areas.

8. References

- Alhawiti, R. H., & Mitsova, D. (2016). Urban Landsat-8 data to explore the correlation between urban heat island and urban land uses. *IJRET: International Journal of Research in Engineering and Technology*, 5(3), 457-466.
- Bao, J., Li, X., & Yu, C. (2015). The construction and validation of the heat vulnerability index, a review. *International Journal of Environmental Research and Public Health*, 12(7), 7220-7234. <https://doi.org/10.3390/ijerph120707220>
- Centers for Disease Control and Prevention (2019, February 2019). CDC's Tracking Network in Action: Extreme Heat. Retrieved February 20, 2019 from <https://www.cdc.gov/features/trackingheat/index.html>.
- Davis, A. Y., Jung, J., Pijanowski, B. C., & Minor, E. S. (2016). Combined vegetation volume and "greenness" affect urban air temperature. *Applied Geography*, 71, 106-114. <https://doi.org/10.1016/j.apgeog.2016.04.010>
- Gao, B.-C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257-266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)
- Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, 55(3), 377-393. <https://doi.org/10.1177/0013164495055003002>
- Green, H., Bailey, J., Schwarz, L., Vanos, J., Ebi, K., & Benmarhnia, T. (2019). Impact of heat on mortality and morbidity in low and middle income countries: A review of the epidemiological evidence and considerations for future research. *Environmental Research*, 171, 80-91. <https://doi.org/10.1016/j.envres.2019.01.010>
- Harlan, S. L., Deplet-Barreto, J. H., Stefanov, W. L., & Petitti, D. B. (2013). Neighborhood effects on heat deaths: social and environmental predictors of vulnerability in Maricopa County, Arizona. *Environmental Health Perspectives*, 121(2), 197-204. <https://doi.org/10.1289/ehp.1104625>
- Hondula, D., Davis, R. E., Leisten, M. J., Saha, M. V., Veazey, L. M., & Wegner, C. R. (2012). Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA from 1983-2008: A case-series analysis. *Environmental Health* 11(1), 1-11. <https://doi.org/10.1186/1476-069X-11-16>
- Hondula, D. M., Davis, R. E., Saha, M. V., Wegner, C. R., & Veazey, L. M. (2014). Geographic dimensions of heat-related mortality in seven U.S. cities. *Environmental Research*, 138, 439-452. <https://doi.org/10.1016/j.envres.2015.02.033>

- Hondula, D. M., Davis, R. E., & Georgescu, M. (2018). Clarifying the connections between green space, urban climate, and heat-related mortality. *American Journal of Public Health*, 108(S2), S62-S63. <https://doi.org/10.2105/AJPH.2017.304295>
- Johnson, D. P., & Wilson, J. S. (2009). The socio-spatial dynamics of extreme urban heat events: The case of heat-related deaths in Philadelphia. *Applied Geography*, 29(3), 419-434. <https://doi.org/10.1016/j.apgeog.2008.11.004>
- Jin, M., & Liang, S. (2006). An improved land surface parameter for land surface models using global remote sensing observations. *American Meteorological Society*, 2867-2881. <https://doi.org/10.1175/JCLI3720.1>
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141-151. <https://doi.org/10.1177/001316446002000116>
- Kaplan, G., Avdan, U., & Avdan, Z. Y. (2018). Urban heat island analysis using the Landsat 8 satellite data: a case study in Skopje, Macedonia. *Proceedings*, 2(258), 1-5. <https://doi.org/10.3390/ecrs-2-05171>
- Kumar, K. S., Bhaskar, P. U., & Padmakumari, K. (2012). Estimation of land surface temperature to study urban heat island effect using Landsat ETM+ image. *International Journal of Engineering Science and Technology*, 4(02), 771-778
- Lugo-Amador, N. M., Rothenhaus, T., & Moyer, P. (2004). Heat-related illness. *Emergency Medicine Clinics of North America*, 22(2), 315-327. <https://doi.org/10.1016/j.emc.2004.01.004>
- McGeehin, M. A., & Mirabelli, M. (2001). The potential impacts of climate variability and change on temperature-related morbidity and mortality in the United States. *Environmental Health Perspectives*, 109 Suppl 2, 185-189. <https://doi.org/10.1289/ehp.109-1240665>
- Mushore, T. D., Mutanga, O., Odindi, J., & Dube, T. (2018). Determining extreme heat vulnerability of Harare Metropolitan City using multispectral remote sensing and socio-economic data. *Journal of Spatial Science*, 63(1), 173-191. <https://doi.org/10.1080/14498596.2017.1290558>
- Oke, T.R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, 108, 1-24. <https://doi.org/10.1002/qj.49710845502>
- Olmedo, G. F., Ortega-Farías, S., de la Fuente-Sáiz, D., Fonseca-Luego, D. & Fuentes-Peñailillo, F. (2016) Water: tools and function to estimate actual evapotranspiration using land surface energy balance models in R. *The R Journal*, 8:2, 352-369.

- City of Philadelphia, Office of Sustainability (2019). Beat the Heat - Hunting Park. Retrieved February 20, 2019 from https://www.phila.gov/media/20190719092954/HP_R8print-1.pdf
- Steadman, R (1984). A universal scale of apparent temperature. *Journal of Climate and Applied Meteorology*, 23, 1674-1687. [https://doi.org/10.1175/1520-0450\(1984\)023<1674:AUSOAT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1984)023<1674:AUSOAT>2.0.CO;2)
- Uejio, C. K., Wilhelmi, O. V., Golden, J. S., Mills, D. M., Gulino, S. P., & Samenow, J. P. (2011). Intra-urban societal vulnerability to extreme heat: the role of heat exposure and the built environment, socioeconomics, and neighborhood stability. *Health & place*, 17(2), 498-507. <https://doi.org/10.1016/j.healthplace.2010.12.005>
- U.S. Geological Survey Earth Resources Observation And Science Center. (2014). Landsat 8 OLI Level-2 Surface Reflectance (SR) Science Product [Data set]. U.S. Geological Survey. <https://doi.org/10.5066/f78s4mzj>
- Urban, A., & Kysely, J. (2014). Comparison of UTCI with other thermal indices in the assessment of heat and cold effects on cardiovascular mortality in the Czech Republic. *International Journal of Environmental Research and Public Health*, 11(1), 952-967. <https://dx.doi.org/10.3390%2Fijerph110100952>
- Vermote, E., Wolfe, R. (2015). MYD09GA MODIS/Aqua Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2020-04-03 from <https://doi.org/10.5067/MODIS/MYD09GA.006>
- Wan, Z., Hook, S., Hulley, G. (2015). MYD11A1 MODIS/Aqua Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2020-04-03 from <https://doi.org/10.5067/MODIS/MYD11A1.006>
- Weber, S., Sadoff, N., Zell, E., & de Sherbinin, A. (2015). Policy-relevant indicators for mapping the vulnerability of urban populations to extreme heat events: A case study of Philadelphia. *Applied Geography*, 63, 231-243. <https://doi.org/10.1016/j.apgeog.2015.07.006>
- Zhang, J., Wang, Y., & Li, Y. (2006). A C++ program for retrieving land surface temperature from the data of Landsat TM/ETM+ band6. *Computers & Geosciences*, 32(10), 1796-1805. <https://doi.org/10.1016/j.cageo.2006.05.001>

9. Appendix

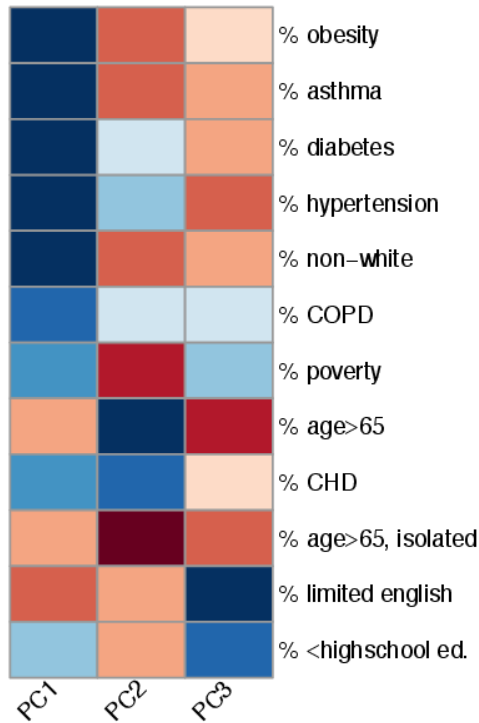
Table A1: Overview of data set name, spatial resolution, temporal coverage, creator, and retrieval date of datasets included in this analysis.

Dataset	Subdivisions/ Resolution	Date/ Time	Creator	Retrieval
B01001: Sex by Age	Census Tracts	2018, 5-year estimate (2014-2018)	American Community Survey (ACS)	Tidycensus package in R
B15003: Educational Attainment for the Population 25 Years and Over	Census Tracts	2018, 5-year estimate (2014-2018)	ACS	Tidycensus package in R

C16002: Household Language by Household Limited English-Speaking Status	Census Tracts	2018, 5-year estimate (2014-2018)	ACS	Tidycensus package in R
B17020: Poverty Status by Age	Census Tracts	2018, 5-year estimate (2014-2018)	ACS	Tidycensus package in R
B02001: Race	Census Tracts	2018, 5-year estimate (2014-2018)	ACS	Tidycensus package in R
B11007: Households by Presence of People 65 Years and Over, Household Size and Household Type	Census Tracts	2018, 5-year estimate (2014-2018)	ACS	Tidycensus package in R
500 Cities: Census Tract-level Data (GIS Friendly Format), 2018 release	Census Tracts	2018	Center for Disease Control and Prevention (CDC)	CDC Website
USGS Landsat 8 Collection 1 Tier 1 Surface Reflectance	30 meter	June 1 to September 1, 2017-2019	NASA	Google Earth Engine
MODIS Aqua Surface Reflectance	500m	June 1 to September 1, 2017-2019	NASA	Google Earth Engine
MODIS Aqua Land Surface Temperature and Emissivity	1km	June 1 to September 1, 2017-2019	NASA	Google Earth Engine
Philadelphia City Limits Shapefile	City of Philadelphia	2014	City of Philadelphia	OpenDataPhilly
TIGER/Line Shapefile, 2018, state, Pennsylvania, Current Census Tract State-based	Pennsylvania	2018	US Census Bureau, Department of Commerce	Data.gov

Citywide_StreetTreeDensity_2017	street level vector file	2017	Lauren Medsker, Philadelphia Horticultural Society modified from Philadelphia Parks and Recreation Street Tree Inventory	ArcGIS Online
TIGER/Line Shapefile, 2018, county, Philadelphia County, PA, All Roads County-based Shapefile	Philadelphia County	2018	US Census Bureau, Department of Commerce	Data.gov

A Heat Sensitivity Index



B Heat Exposure Index

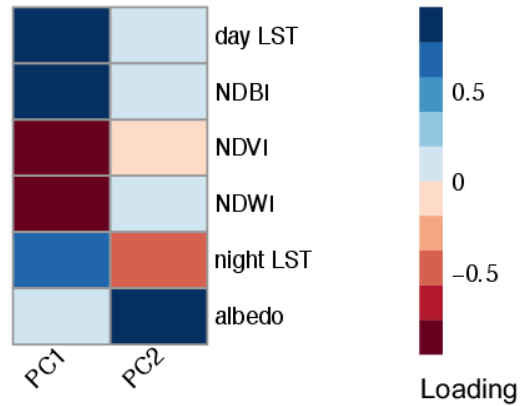


Figure A1. Loadings of Principal Component Analyses for A) Heat Sensitivity Index and B) Heat Exposure Index. For the Heat Sensitivity Index, the principal component 1 (PC1) is driven by socioeconomic factors and chronic health conditions and accounts for 51 percent of data variance. PC2 is representative of age and age-related health conditions. PC3 can be described as the socio-cultural component with strong loadings on percent population in limited English-speaking households, with less than a high school education, and living below the federal poverty level. Together the three components explain 83 percent of data variance. In the Heat Exposure Index, PC1 is driven by daytime LST, NDBI, NDVI, and NDWI. PC2 is representative of heat retention and driven by nighttime LST and albedo. Together the two components account for 81 percent of data variance.

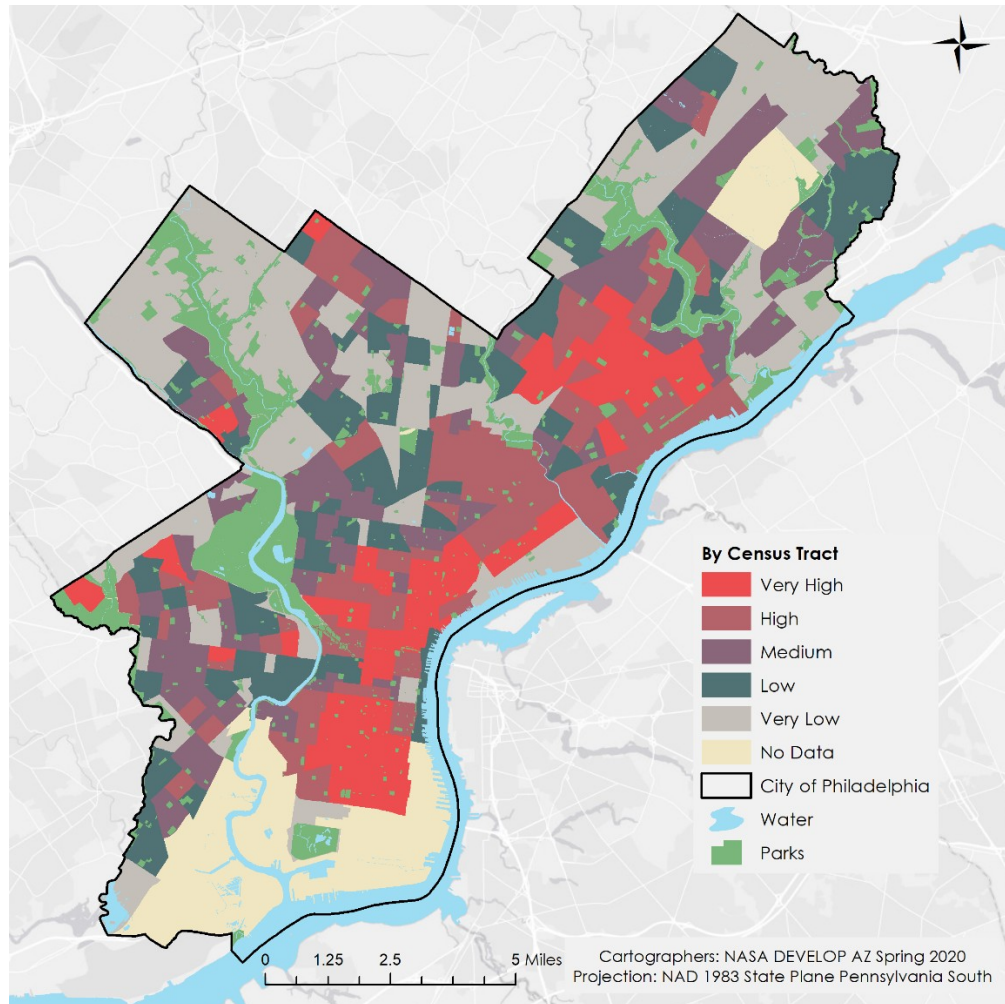


Figure A2. 2018 Heat Sensitivity Index for Philadelphia by census tract. The scores are categorized in quintiles. The Heat Sensitivity Index based on socioeconomic and health indicators only.

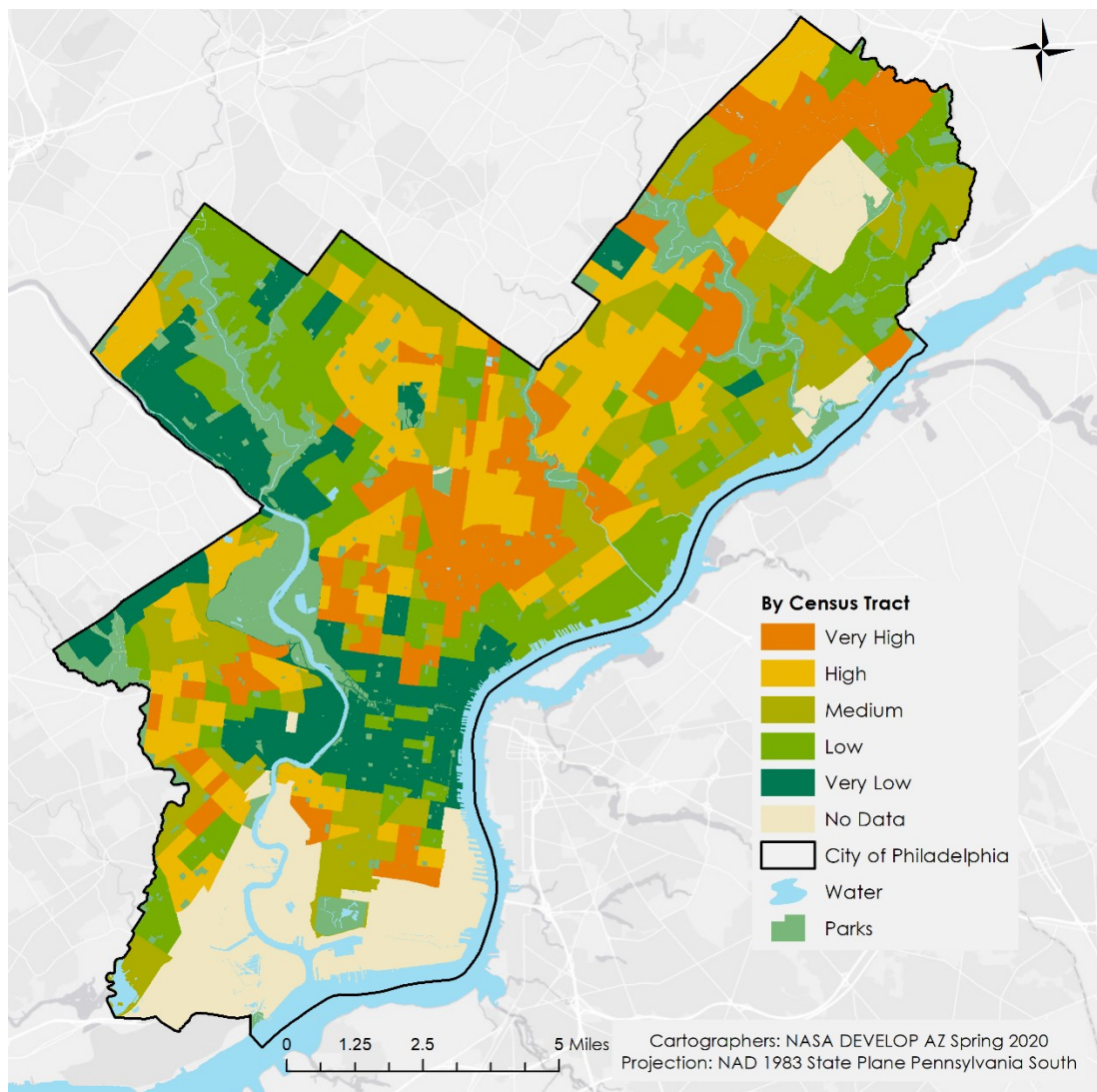


Figure A3. 2018 Heat Exposure Index for Philadelphia by census tract. The scores are categorized in quintiles. The Heat Exposure Index is based on environmental variables only.