Understanding Perception of Different Urban Thermal Model Visualizations

Gunjan Barua

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
In
Geography

Thomas Pingel, Chair/Advisor
Theodore Lim
Stephanie Zick
Santosh Rijal

February 13, 2023
Blacksburg, VA

Keywords: colormaps, perception study, structure from motion, thermal maps, urban heat island

Understanding Perception of Different Urban Thermal Model Visualizations

Gunjan Barua

Abstract

While satellite-based remote sensing techniques are often used for studying and visualizing the urban heat island effect, they are limited in terms of resolution, view bias, and revisit times. In comparison, modern UAVs equipped with infrared sensors allow very fine-scale (cm) data to be collected over smaller areas and can provide the means for a full 3D thermal reconstruction over limited spatial extents. Irrespective of the data collection method, the thermal properties of cities are typically visually represented using color, although the choice of colormap varies widely. Previous cartographic research has demonstrated that colormap and other cartographic choices affect people's understanding. This research study examines the difference in map reading performance between satellite and drone-sourced thermal pseudo-color images for three map reading tasks, the impact of color map selection on map reading, and the potential benefits of adding shading to thermal maps using high-resolution digital surface models for improved interaction. Participants expressed a preference for the newly designed rainbow-style color map "turbo" and the FLIR "ironbow" colormap. However, user preferences were not strongly related to map reading performance, and differences were partly explained by the extra information afforded by multi-hue and shading-enhanced images.

Understanding Perception of Different Urban Thermal Model Visualizations

Gunjan Barua

General Audience Abstract

While satellite-based remote sensing techniques are often used for studying and visualizing the urban heat island effect, they are limited in terms of resolution, view bias, and revisit times. In comparison, modern drones or Unmanned Aerial Vehicles (UAVs) equipped with infrared sensors allow very fine-scale (cm) data to be collected over smaller areas and can provide the means for a full 3D thermal reconstruction over a small area. Irrespective of the data collection method, the thermal properties of cities are typically visually represented using color, although the choice of colormap varies widely. Previous cartographic research has demonstrated that colormap and other cartographic choices affect people's understanding. This research study examines the difference in map reading performance between satellite and drone-sourced thermal pseudo-color images for three map reading tasks, the impact of color map selection on map reading, and the potential benefits of adding hillshade augmentation to thermal maps using high-resolution digital surface models for improved interaction. Participants expressed a preference for the newly designed rainbow-style color map "turbo" and the FLIR "ironbow" colormap. However, user preferences were not strongly related to map reading performance, and differences were partly explained by the extra information afforded by multi-hue and shading-enhanced images.

Dedication

This thesis is dedicated to my wife, parents, in-laws, and family. Thank you for always believing in me.

Acknowledgments

I would like to take this opportunity to express my heartfelt gratitude to my parents, in-laws, and family members who have supported me throughout my academic journey. Their unwavering encouragement and unwavering faith in me have been invaluable in helping me pursue my higher studies.

I am grateful to my wife, Lumbini Barua, for her unconditional support, love, and encouragement throughout my academic journey. Her constant belief in me, understanding, and patience have been a source of inspiration and motivation in pursuing my research and completing this thesis. I am forever grateful for her presence in my life, and I could not have accomplished this without her support.

I am grateful to Dr. Theodore Lim for his support during the conceptualization of my study and for providing expert guidance and feedback throughout the research process. I am also thankful for his support during the drone data collection process, as well as his role in helping me understand the potential of using technology to address environmental challenges that affect human populations. I am also thankful to Dr. Santosh Rijal and Dr. Stephanie Zick for their valuable feedback from time to time that helped shape my research.

I am grateful to Dr. Tom Crawford and the Department of Geography for supporting me with the Sidman P. Poole grant that helped me to collect data from participants for my study. I am also thankful to the department and the College of Natural Resources & Environment (CNRE) for providing me with a conference grant that helped me to present my paper at the AutoCarto 2022 conference. Thanks is also due to Karen Bland for always being there to provide any support administrative support whenever required.

I am thankful to my NEIL lab members for always being the first ones to listen to all the ideas, failures, and successes related to the research, and for extending their support and collaboration in finishing my research.

I want to express my heartfelt appreciation to my friends within and outside the department, who have been there for me every step of the way. Their constant encouragement, positivity, and willingness to listen have been a source of comfort and inspiration throughout my academic journey. I am grateful for their support and friendship, and I look forward to cherishing these relationships for years to come.

Lastly, I am deeply grateful for the guidance, support, and mentorship of my thesis supervisor, Dr. Thomas Pingel, throughout my academic journey. His expertise, insights, and feedback have been instrumental in shaping my research and helping me navigate the challenges of graduate studies. His unwavering commitment to excellence and dedication to helping me succeed has been a constant source of inspiration and motivation. I am fortunate to have had such a supportive and knowledgeable supervisor, and I will always be grateful for his contributions to my academic and personal growth.

Contents

List of Figures

List of Tables

1. Introduction

The term Urban Heat Island Effect (UHIE) was coined more than 200 years ago by Luke Howard in *The Climate of London* (1818) and describes the phenomenon of warmer temperatures in urban areas compared to the surrounding areas. In the United States, cities are typically 1 to 4 °C hotter compared to surrounding rural areas (US EPA, 2022). As a result, human health (Ebi *et al*., 2018; Habeeb *et al*., 2015) and urban infrastructure (Clark *et al.,* 2018; Sailor *et al.,* 2019) are negatively impacted. Growing evidence suggests that heat exposure also disproportionally impacts low-income communities and people of color (Wilson, 2020; Klinenberg, 2003).

Urban heat can be a significant issue, but it is not always immediately noticeable to residents (Guardaro *et al.,* 2020). There are gaps in data and understanding of the causes and effects of extreme heat (Wilson, 2020), which makes it difficult to prioritize efforts to reduce its impact (Meerow & Keith, 2021). Making temperature distribution maps easily accessible and understandable could lead to greater engagement and more effective decision-making for heat-resilient community planning as visualizations can impact decision-makers and stakeholders (Zhu & Chen, 2008).

Image resolution in visualizations plays an important role in decision-making as higher-resolution visualizations enhance performance due to better visual perception details (Yeshurun & Carrasco, 1998). Lowering the resolution decreases the signal-to-noise ratio, resulting in difficulty in differentiating signals (Rajashekar *et al.,* 2006). On the other hand, the selection of colormaps influences a user's data visualization comprehension (Schloss *et al.,* 2019). A proper colormap can effectively enhance the expressiveness and persuasiveness of data in visual

representations (Nagoor et al., 2017). An accurate selection of colormap can improve the objective performance of tasks, arouse affective resonance, and raise visual immersion (Zhang et al., 2021).

For this research, a user study tested ways to visualize the urban thermal environment using the drone and satellite-sourced imagery. The user performance (accuracy and response time) on three map-reading tasks was tested. In scientific cartography, it is common to evaluate maps using techniques like experimental psychology (Board, 1978). Like any other maps, thermal maps require content and construct validation (Anastasi, 1968). Thus, map reading tasks are efficient techniques for evaluating the effectiveness of the maps. Apart from the colormaps, and resolution, shading augmentation was also applied to develop thermal visualizations in this study. Thermal maps represented using colors and shading can help alleviate the confusion usually developed by map users (Patterson and Jenny, 2014).

The map reading tasks consisted of visualizations where mosaiced orthoimagery from both satellite and drone using several different colormaps and additional augmentation (e.g, shading) were tested. An initial pilot testing round helped narrow down the variables before full user testing. The study looked into three hypotheses.

- It is expected that some colormaps will show significant improvement over others, according to an ANOVA analysis. Turbo and ironbow colormaps are expected to perform best.
- Based on users' perceptions, The high-resolution urban heat map is expected to perform better than the lower-resolution, satellite-derived heat map.

- Maps with cartographic augmentations, such as shading, is expected to perform better than those without the augmentation.

2. Background

2.1.The problem with the UHIE

The UHIE is caused by using concrete and other human-made materials that absorb and re-emit heat to a greater degree than natural materials (Howard, 1818). The problem is expected to increase during the next century, as cities expand in size and population, simultaneously increasing temperatures and exposing more people to them. The UN Environment Programme (2021) estimates that by the year 2100, the average temperature of cities may increase up to 4 ^{o}C over a 2006-2015 baseline and expose 2.3 billion people to severe heat waves which are more common with the UHIE and are associated with many negative health outcomes (Harlan *et al*., 2007). In the United States, death from heat exposure is not particularly common in absolute terms (3.1 of 1 million deaths; Lowe, 2016), but it is the number one weather-related cause of death averaged over the past several decades (Lee, 2014; Sheridan and Allen, 2018). Another study of 93 European cities revealed that in 2015, 4.3% of all summer deaths could be attributed to the impacts of urban heat island (UHI) effects (Iungman et al., 2023). However, apart from a direct impact on deaths, heat exposure also increases the risk of hospitalization (Basu, 2009), respiratory illness (D'Ippoliti et al., 2010), other heat-related mortality risks (Hajat et al., 2014), causes increased energy bills, which are especially burdensome for low-income families (Sailor et al 2019; Thomson et al 2019), as well as missed work days or decreased productivity (Lundgren et al 2013), significant physical discomfort/inconvenience (Guardaro et al 2020), and mental stress (Hansen et al, 2008).

Increased urban heat has other important effects as well. Increased temperatures in urban areas increase air contamination by influencing atmospheric pollutants' concentration and spatial

distribution (Sarrat, *et al*. 2006). The UHEI has been shown to impact precipitation over urban areas (Baik *et al*., 2001; Bornstein and Lin, 2000; Dixon and Mote, 2003). Vegetation in urban areas tends to decay exponentially with the increase in land surface temperature (Zhou, et al. 2016).

2.2.Factors contributing to the Urban Heat Island Effect

Ryu and Baik (2012) identified impervious surfaces, anthropogenic heat, and three-dimensional geometry as the primary contributing factors for UHIE. Impervious surfaces in urban areas such as concrete and asphalt are significant drivers of UHIE. Dark features such as rooftops and asphalt can additionally lower albedo, meaning more of the sun's energy is absorbed rather than reflected. Anthropogenic heat is caused by human activity such as waste heat from automobile traffic and heating systems in homes and businesses. 3D geometry includes things such as heat trapped in vertical walls, radiation entrapment, and wind speed reduction. The so-called "sky view factor", or the proportion of sky visible from the ground is related to 3D geometry and can reduce the UHIE in places such as downtown areas where high-rise buildings can occlude the sun (Bernard et al., 2018).

Zhao et al. (2018) examined the relationship between land surface temperature variation and lidar-derived terrain factors, land cover composition, and landscape pattern metrics and confirmed that higher building density is associated with an increase in the UHIE whereas vegetation cover decreases the effect. In addition, areas near water bodies showed a lower UHIE value compared to areas farther from water. Deilami et al. (2018) conducted a meta-analysis of 75 related research articles on UHIE and found that key spatiotemporal factors included the percentage of vegetation, seasonal variation, area, day/nighttime variation, population, the

proportion of water body in an urban area, percentage of road pavement, biophysical components, and social and economic variables.

Benz & Burney (2021) found that the extreme heat burden is disproportionately borne by non-white populations and lower-income neighborhoods in 70% of the US counties. Another study showed that around 94% of the 108 studied US urban areas displayed persistent patterns of increased land-surface temperature in the formerly redlined areas compared to non-redlined areas by an average of 2.6 ^{o}C (Hoffman et al., 2020). Apart from the physical and meteorological factors, economics plays a role: Cui et al. (2016) found a positive relationship between higher GDP and higher land surface temperatures in major cities in the US and China.

2.3.Measuring Urban Heat Island Effect

Temperatures in urban areas are warmer both with respect to air and surface temperatures. Air temperatures more closely correspond to human comfort (Hiemstra et al., 2017) but are more difficult to directly measure at densities that would result in high-resolution maps. The recently launched National Integrated Heat Health Information System - Climate Adaptation Planning and Analytics (NIHHIS-CAPA) Urban Heat Island Mapping Campaign seeks to remedy this and has been providing funding since 2017 to produce high-resolution air temperature maps of cities, often using a citizen-science-based effort to sample air temperatures with a vehicle “traverses” approach that then form the basis of an interpolation (Shandas *et al*., 2019, Figure 1).

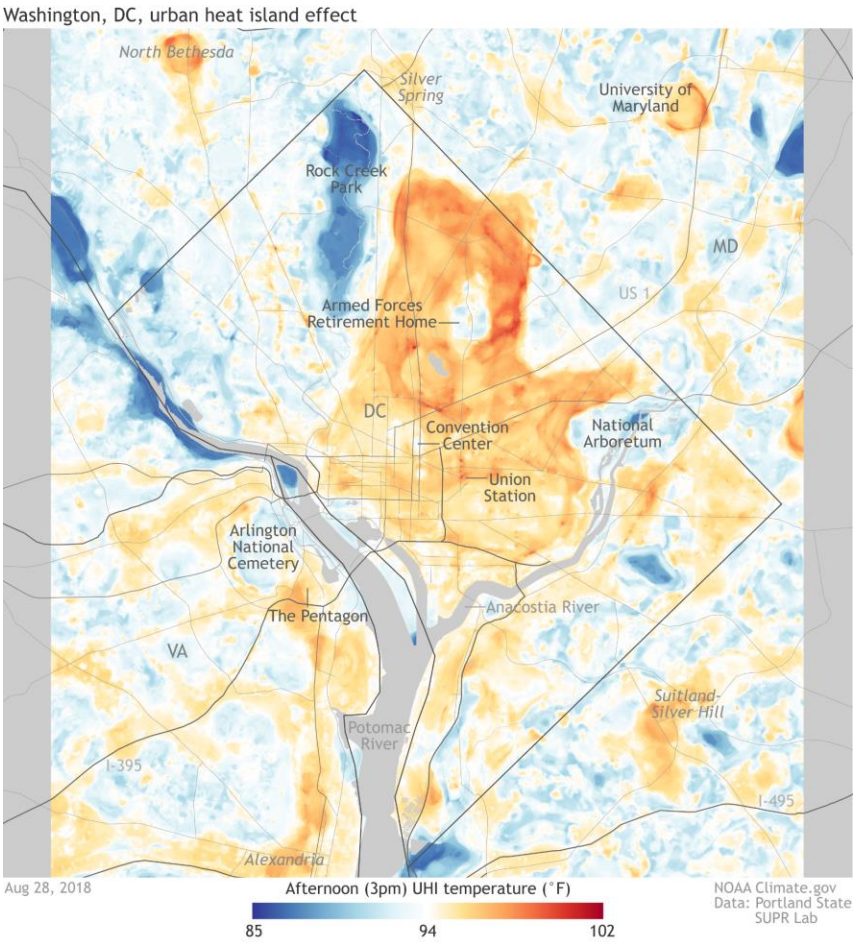


Figure 1 High-resolution map of urban heat island effect generated from vehicle traverses. From Shandas et al. (2019) viewable at climate.gov.

Surface temperatures can be measured over an entire city using satellites and provide a relatively high-resolution product that allows for coarse comparisons of how land cover types or other classifications produce different temperature profiles. At present, there are several satellites in operation that collect thermal infrared data from the earth's surface, including Landsat TM/ETM+, MODIS, ASTER, HCMM, AVHRR, TIMS, and others (Weng 2009). Many researchers have used data from these satellites to measure land surface temperature (LST) in urban areas. For example, Nichol (2005) compared the daytime and nighttime LST using ETM+ thermal image and ASTER sensors respectively which found that densely built high-rise

buildings have lower temperatures in the daytime but appear warmer during the night. Chen *et al*. (2017) used MODIS satellite data to identify the positive relationship between urbanization and urban diurnal land surface temperature variation. Krehbiel and Henebry (2016) used MODIS satellite data to compare air temperature and land surface temperature as a measure of urban heat islands in the upper-midwest US where it was found that the diurnal variation for air temperature is insignificant while the variation for land surface temperature was much higher. Imhoff *et al*. (2010) used Landsat-based NLCD data and thermal satellite images from MODIS to determine that impervious surface area is the main driver for 70% of the total variance in land surface temperature in 38 of the most populous cities in the US. A study by Adaktylou (2020) used Landsat 8 satellite images to develop UHIE visualizations for elementary school children and found it useful for teachers as they found a new tool for engaging students in science topics. The post-survey of the study also indicated an enhanced interest among students in learning about the complexity of the environment.

Recently, Land Surface Temperature data layers have been made available as part of an Analysis Ready Data (ARD) package, distributed via the USGS EarthExplorer platform, thus obviating the need for end-users to apply the sensitive calculations to convert raw data from relevant Landsat 8 bands to surface temperature estimates (Dwyer et al., 2018). An updated version of the Surface Temperature dataset is also available via Google Earth Engine (Ermida *et al*., 2020; Figure 2).

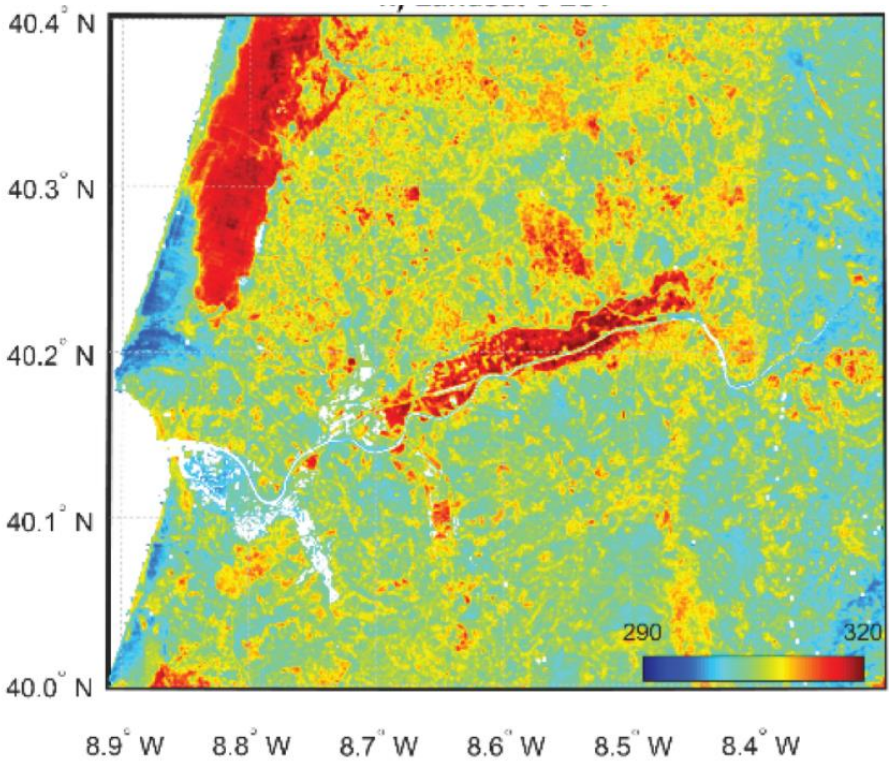


Figure 2 Surface Temperature (Kelvin) algorithmic estimation from Landsat-8 shown for central Portugal, distributed via Google Earth Engine. From Ermida et al. 2020, Figure 4.

2.4.Measuring Urban Heat Island Effect using UAVs

More recently, unmanned aerial vehicles (UAVs) or drones equipped with infrared (IR) cameras have been used to study the UHIE. UAVs allow the users to fly at a lower altitude which allows for the generation of higher resolution (cm level) imagery. In addition to that, it also allows the generation of 3D models with reliable metrics that result in cost-effective qualitative and quantitative inspections (Luis-Ruiz et al., 2021). Naughton and McDonald (2019) used a DJI Matrice 100 quadcopter UAV to collect thermal imagery and found a 3.9 to 15.8°C land surface temperature variability in Milwaukee, WI, and El Paso, TX. Song and Park (2020) used a FLIR Vue Pro R thermal infrared (TIR) camera mounted on a DJI Inspire 1 UAV to compare thermal imagery with temperatures collected on the ground using infrared thermometers. They found a

high correlation of 0.70 and RMSE of 4 to 4.5 ℃, with higher variability reported for surfaces such as concrete and wooden decks with higher temperatures. Thermal cameras are often of substantially lower resolution (5 to 15% of the resolution) than RGB cameras and have less contrast, making conventional Structure from Motion reconstructions difficult. Previtali et al. (2014) suggested an easy approach by generating a geometric 3D model with RGB images and projecting thermal images as textures. Another method by Daffara et al. (2020) suggested projecting thermal images onto a 3D point cloud by manually inserting ground control points (GCP). One study proposed fusing visible and thermal point clouds where the camera is calibrated to acquire its internal orientation parameters (Dahaghin et al., 2019). Lastly, Yang and Lee (2019) introduced a four-band thermal mosaicking method that combines the thermal image with a cropped version of the RGB image, using the RGB channels for scene reconstruction and the thermal image for orthophoto generation. Most thermal mapping projects focus on orthoimage generation, and as mentioned above, some projects constructed full 3D thermal scenes. Another example is from Morrison et al. (2018), who developed a ground-based remote sensing application for three-dimensional (3D) urban thermography using an LWIR camera simulated in Discrete Anisotropic Radiative Transfer (DART) model and Blender software. An example of a thermal image developed from a UAV is shown in Figure 3.

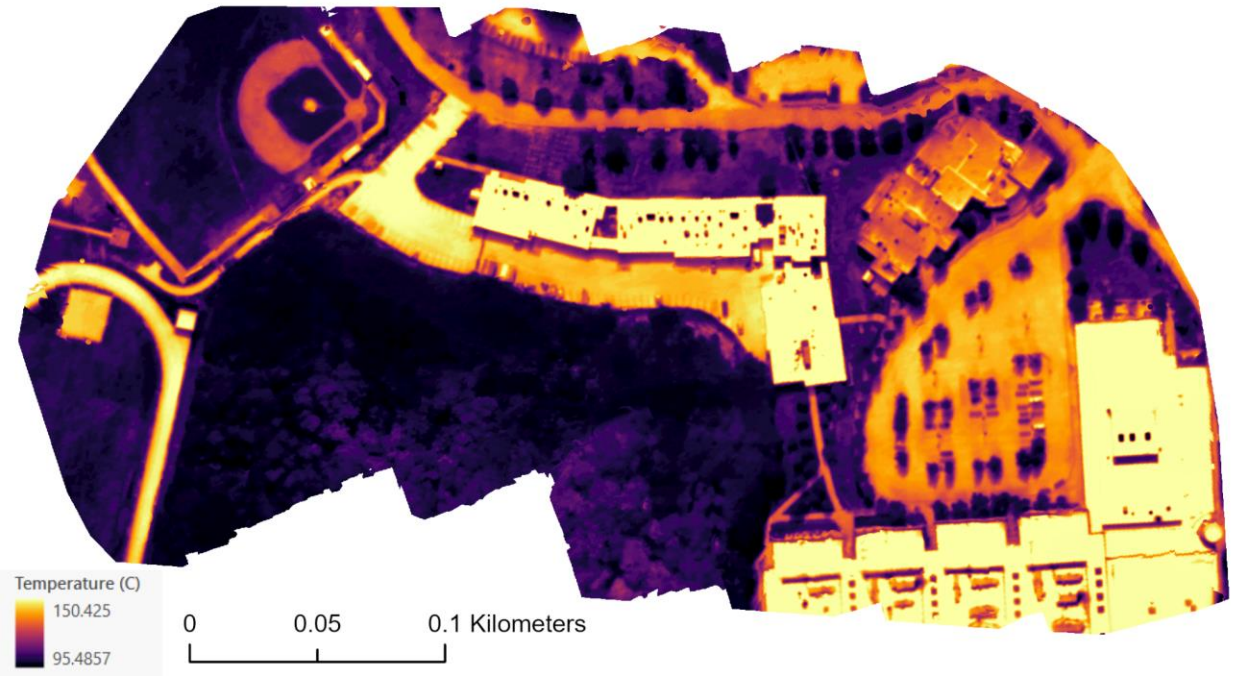


Figure 3 Thermal image orthomosaic constructed from images taken during a flight in Roanoke, VA using Mavic Pro Duo thermal camera.

2.5.The importance of higher-resolution thermal imaging

Researchers have been using satellite images for studying the urban thermal environment, however, they have some limitations when studied for smaller, more detailed extents such as neighborhoods or specific sites. Satellite images have low revisit times, usually between 1-14 days (Morrison *et al*., 2021; Naughton & McDonald, 2019). The images also have a low spatial resolution, where a large area is covered by a single pixel, making the identification of specific features difficult. Landsat's thermal bands, for instance, are collected at 100 m resolution, although they are resampled to 30 m to match other bands such as RGB, which are natively collected at that resolution. Satellite-derived images of surface temperature may work well at the city or region scale, but they are often hard to interpret at finer scales (Naughton & McDonald, 2019). To tackle these challenges, some studies have used multi-satellite combination and

downscaling/interpolation techniques using land cover classes in combination with thermal imagery to produce 10 m resolution interpolated images (Nichol, 2005; Nichol & Wong, 2009), however, they are still unable to capture the true complexity of thermal variation in an urban area. The thermal map appears overly smooth in comparison to natively finer resolution platforms such as those deployed by UAVs (Figure 4). NASA's G-LiHT, similar to other portable airborne imaging devices, combines lidar, imaging spectroscopy, and thermal imaging to examine the makeup, layout, and operation of land-based ecosystems. G-LiHT has been employed in several investigations to specifically design the thermal three-dimensional layout of urban areas with a resolution of 1-2.5 meters (Berger et al., 2014; Chen et al., 2016).

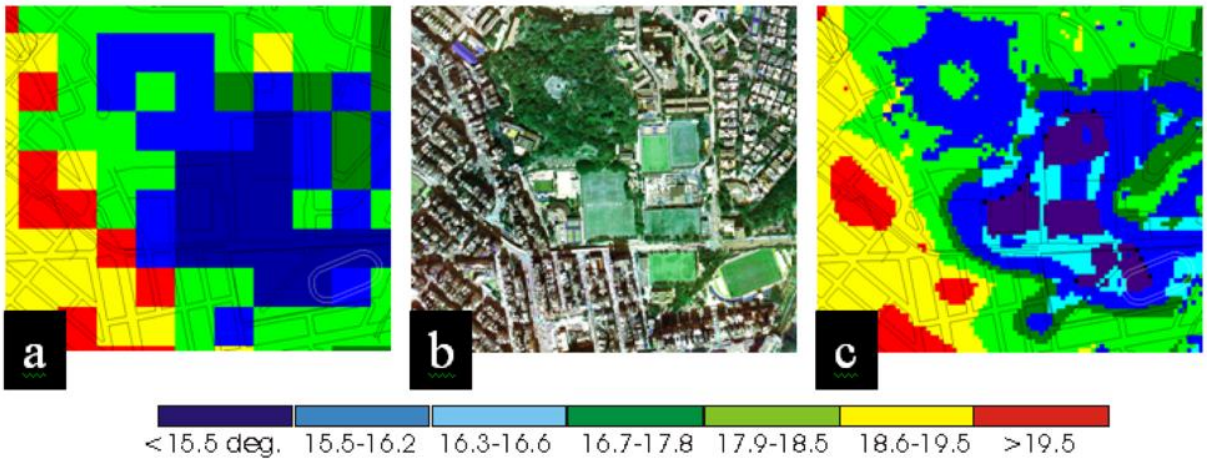


Figure 4 (a) ASTER 90 m thermal data combined with land cover classes derived from (b) RGB imagery to create (c) an interpolated 10 m surface temperature estimate for Hong Kong (from Nichol and Wong, 2009).

UAVs help researchers overcome many of the limitations of satellites by collecting high-resolution data at high temporal resolution (Laliberte *et al*., 2011). This also allows the user to investigate the desired area repeatedly in shorter time intervals (Vasterling & Meyer, 2013) and change speed, direction, and altitude as needed (Villa et al., 2017). Improvements in spatial resolution of thermal imagery make buildings, trees, and sidewalks directly visible in the scene,

removing a layer of generalization that may make novice users directly apprehend their thermal properties. Deploying UAVs several times throughout the day can illustrate how various surfaces heat up differently (Lim *et al*., 2022). To date, most visualizations of fine-scale thermal environments have used symbolized orthoimagery to communicate scenes. Some examples of directly sensed 3D SfM-derived thermal models (Previtali *et al.*, 2014; Daffara *et al.*, 2020; Dahaghin *et al.*, 2019; Yang and Lee, 2019) were reviewed. However, previous research has found that 3D, user-manipulatable models in other contexts improve map reading (Carbonell Carrera *et al*., 2017). Further, 3D maps may allow users to see how structures such as vertical walls, windows, and doors of buildings - features much more common to everyday experience than rooftops - emit thermal energy into the environment.

2.6.Addressing the problem with scientific cartography

Thermal maps are often communicated in ways that may be problematic, but scientific cartography provides a method for rigorously testing such geovisualizations, thereby improving their use in informing the public and improving decision-making (Robinson, 1950; Hayek et al., 2016). Early cartographers mostly developed their understanding of map design based on trial-and-error methods and through an empirical study of design decisions affecting map usage (Griffin, 2017). However, in the twentieth century, the importance of map users was explicitly acknowledged (Montello, 2002). Many studies have benefitted from a connection to Psychology and utilize a strong tradition of empirical research (e.g., Hegarty *et al*., 2009). Eye-tracking has been used in empirical map research since the 1970s to help determine how people interact with maps (Steinke, 1987), and newer, minimally obtrusive eye-tracking technology has created a renewed interest in these methods (Fuhrmann *et al*., 2009). Pingel and Clarke (2014) used both map rotation and profile selection tasks, and measured response time and accuracy rates to

measure the efficacy of a novel lidar-based visualization, and these tasks will be adapted for this study. To determine how easily a visualization can be interpreted based on its elements, the subband entropy measure of visual clutter was discussed by Rosenholtz *et. al.* (2007). Visual clutter is the state where excess items, or their representation in a visual, lead to performance degradation at some tasks.

2.7.How do different colormaps and visualization techniques affect people's understanding of geographic information?

A viewer's comprehension of geographic information is mediated by the symbology and other elements of the cartographic representation. Slocum et al. (2001) posited that the usability of a geovisualization should be evaluated under six categories: 1) geospatial virtual environments; 2) dynamic representations; 3) metaphors and schemata in user interface design; 4) individual and group differences; 5) collaborative geovisualization, and 6) evaluating the effectiveness of geovisualization methods. In terms of the visualization method, it has been claimed by researchers in human-computer interaction that users prefer dynamic displays compared to static displays and 3D displays over 2D displays (Scaife & Rogers, 1996). However, Hegarty et al. (2009) argue that what they call "naive cartography" assumes that realistic and/or dynamic maps perform better because users say that they prefer them when the reverse is more often true.

Thermal maps generally use variations in color to map temperature. Color is generally broken down by hue and value as so-called "visual variables" in cartography, and Bertin outlined their appropriate use in Semiology of Graphics (1967). Hue - corresponding to what often is thought of as color - means differences in kinds of colors: greens, blues, yellows, and so on, and cartographers generally use these to express nominal differences or differences in kind. Value,

sometimes called gray tone, corresponds to the lightness or darkness of a color and is often used to indicate quantitative changes. The darker-is-more convention indicates that the darker the color, the larger the quantity (Cuff, 1973; Palsky, 1999). Color is often used in the form of a colormap, a scale of colors defined by the mapmaker that is used to denote low and high values of the variable being mapped.

Colormaps may be randomly or purposefully sorted when used to map qualitative variables, but more typically they are used to map quantitative variables. The two main types are sequential and diverging colormaps. Sequential maps progress from one color to the next in a way that generally indicates an order. Figure 5a shows a daily forecast high-temperature map of the contiguous United States, where temperatures are indicated by a colormap that starts in light purple, moves through blues, then greens and yellows, and with dark red indicating high temperatures. In contrast, diverging colormaps use neutral values (often white or yellow) to indicate values near a mean or meaningful zero value, with blue often indicating a low value and red indicating a high value for thermal maps specifically and many other kinds of maps as well. As with all colormaps, the specific colors used may vary, and cartographers have user-tested many of these, especially with concern for special populations, such as those with colorblindness (Brewer, 2003). An example of a diverging color map is shown in Figure 5b, where thermal anomalies (deviations from a given value) are displayed.

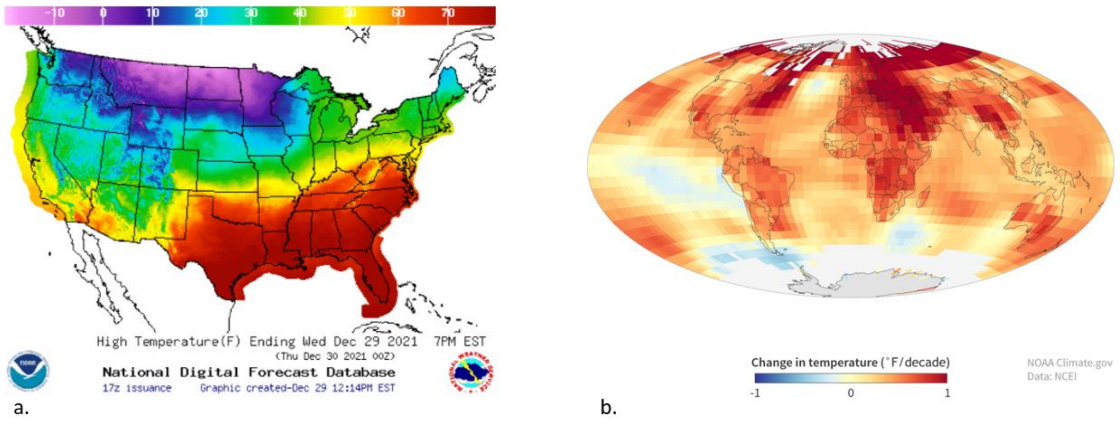


Figure 5 Daily high-temperature values from weather.gov are shown in a conventional sequential colormap (a) and global thermal anomalies from climate.gov are shown using a conventional diverging colormap (b).

Thermal images from FLIR cameras can be programmed to use a variety of color maps, including sequential schemes such as gray tone, rainbow, ironbow, and many more. Examples of several colormaps are shown in Figure 6. Map users are generally cautioned to use "perceptually uniform" colormaps when possible, such that the lightness and darkness of the color don't change haphazardly throughout the ramp (Moreland, 2009).

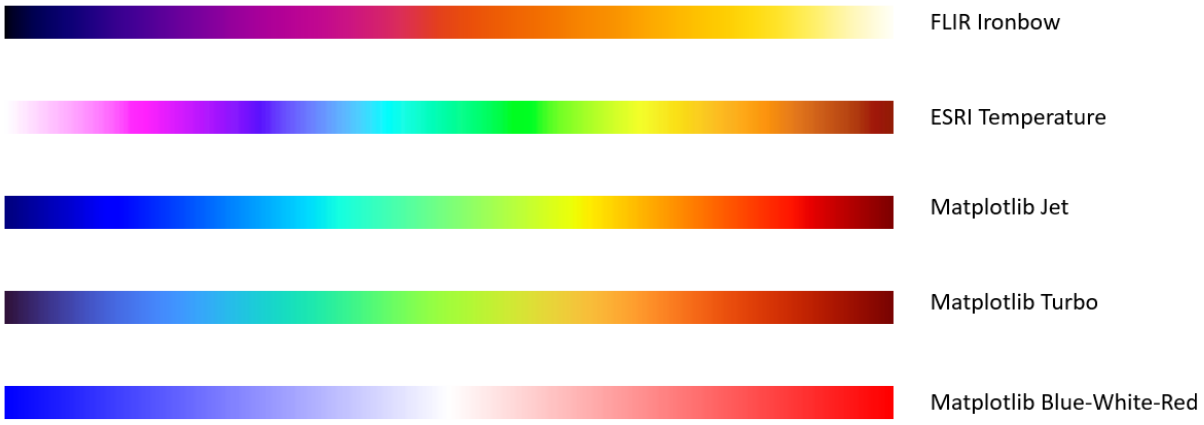


Figure 6 Several colormaps popular for showing thermal data

2.8.Application of terrain visualization lessons on thermal mapping

Since thermal maps are generally portrayed orthographically, they can be imagined as 2.5D "surfaces", much like a terrain surface. Color is often used to map elevation through hypsometric tints which use green for low-lying areas and browns and whites to represent high elevation, but such colors alone can be misinterpreted by a large number of map users (Jenny and Patterson, 2013). Additional commonly used cartographic methods of representing elevation include contours, spot heights, and shaded relief (Imhof, 1982). More advanced techniques can use multi-directional shading, curvature derivatives, and atmospheric shading to improve the reader's sense of the terrain (Kennelly, 2017; Jenny and Patterson, 2021). Figure 7 shows an example of hypsometric tinting, relief shading, and contours used by Imhof (1982) to create a visually powerful image of the terrain. It may be that these ancillary methods could be used to add value to thermal images as well, and many (contours, shading, spot heights) can be readily adapted.

Figure 7 A combination of methods is used to communicate a sense of terrain (Imhof, 1982)

3. Methods

The study was designed to test the performance of three different aspects of thermal model visualizations - color scheme, resolution, and shading augmentation - for their impact on map readability. Three different map reading tasks were used to gauge performance as measured by response time and accuracy. The map reading task was combined with each of the two resolution types to produce six total blocks, the order of which was randomized for each participant. After the completion of the tasks, participants completed a questionnaire to solicit explicit user preferences and collect basic demographic information.

3.1.Participants

A total of 72 participants were tested. All participants were 18 years of age or older. Participants were recruited via wide circulation in list-serve emails and other related platforms within Virginia Tech and received a small gift card in exchange for participation. The study was run following human subjects testing guidelines and received approval from the Institutional Review Board (IRB) (approval number IRB-22-299). Six participants were omitted from the data analysis. In four cases, the system failed to record the complete data record. In two cases, participants paused several times during the study despite instructions to the contrary, resulting in outlier response times. The analysis was based on the data from the resulting 66 participants.

3.2.Materials

The whole survey was administered in PsychoPy® (Peirce, 2007), a free cross-platform package that allows administering a range of experiments in the behavioral sciences. Stimuli were presented on a monitor and participant responses were provided via keyboard. Instructions for

each task were provided on the screen. An administrator was always present to dictate the instructions and assist wherever required. The task was conducted using the Dell Precision 7540 laptop equipped with an Intel core i7 processor with 6 cores, 32 GB RAM, 1 TB storage, and a 6 GB NVIDIA Quadro RTX 3000 graphics card. The external monitor used by participants for the study was a Dell 24-inch monitor with a 16:9 aspect ratio and 1920 x 1080 resolution.

3.2.1. UAV-derived Thermal Imagery

A DJI Mavic 2 Enterprise Dual Drone (Figure 8) was used to capture thermal and RGB images for the study. The drone weighs 899 grams and has a maximum flight time of 31 mins in no wind conditions. The RGB sensor produces a 12-megapixel (MP) image at 4056 x 3040 resolution. The FLIR Lepton thermal sensor captures data at a relatively coarse 160 x 120 resolution before upscaling it for the output image of 640 x 480 pixels. Output images for both sensors were captured in the direction of the automated flight-planning software (DroneDeploy) and stored on the onboard SD card in JPEG format.

Figure 8 DJI Mavic 2 Enterprise Dual Drone used in the study

The drone was flown over Patrick Henry High School in Roanoke, Virginia at an altitude of 120 m Above Ground Height (AGH) to collect imagery. The site was selected because it contained various types of land cover, including large buildings, parking lots, grassy areas, and tree cover. A total of 480 images were captured during the flight, 240 IR images and 240 RGB images. The collected images were processed using the Pix4D Structure from Motion (SfM) software to construct a thermal orthoimage. Reconstruction of thermal imagery is generally done in concert with RGB imagery since thermal imagery is often of low resolution and contrast to work well alone. RGB image processing provides the spatial structure onto which texturing from the thermal imagery is applied. Thus, initial matching (Step 1) for RGB and thermal images is run separately. Step 1 computes key points for the images which are later used to match images together. The software then runs an automatic aerial triangulation and bundle block adjustment from the initial matches to estimate 3D structure. After this, the two image set projects were then merged to build the final high-resolution thermal model in which the point clouds and the mesh geometry were produced from the RGB images, and the mesh textures and orthoimage were generated from the thermal images. In this case, thermal images were collected in grayscale to allow for simple translation via interpolation from grayscale to other color palettes.

To improve and enhance the accuracy of the model, thermal ground control points (TGCPs) were used. TGCPs combine high-precision Real Time Kinematic (RTK) GNSS points with measured ground temperatures at salient points within the scene. An Emlid Reach RS2 RTK GNSS was used to collect the coordinates and a handheld infrared thermometer was used to collect the temperature at those locations. GCP locations were varied throughout the scene as is standard practice when collecting drone imagery, with the additional parameter that a variety of land surfaces were also selected to acquire a range of temperatures. Six TGCPs were used in total.

3.2.2. Satellite Thermal Imagery

Atlanta, Georgia was selected for the city-scale portion of the experiment because it featured a mix of both urban and forested landscapes, leading to a varied temperature profile. The city is one of the largest metropolitan cities in the US and is home to half a million people (U.S. Census Bureau, 2020). The city has a variety of land covers and according to a land cover percentage calculation using 2019 MRLC land cover data, the city contains 76.8% developed area. The developed areas account for 20-100% impervious surface which is a major cause of increasing surface temperature.

Rather than using raw Landsat-8 data to calculate surface temperature, the Analysis Ready Data (ARD) dataset (Landsat Collection 2 Surface Temperature) prepared by the USGS was used. The data combines information from multiple bands and information from other satellites to produce surface temperature estimates (Dwyer et al., 2018). According to the Landsat C2 U.S. ARD Data Format Control Book, the thermal band had to be multiplied with a scale factor and added an additive factor. Following is the equation.

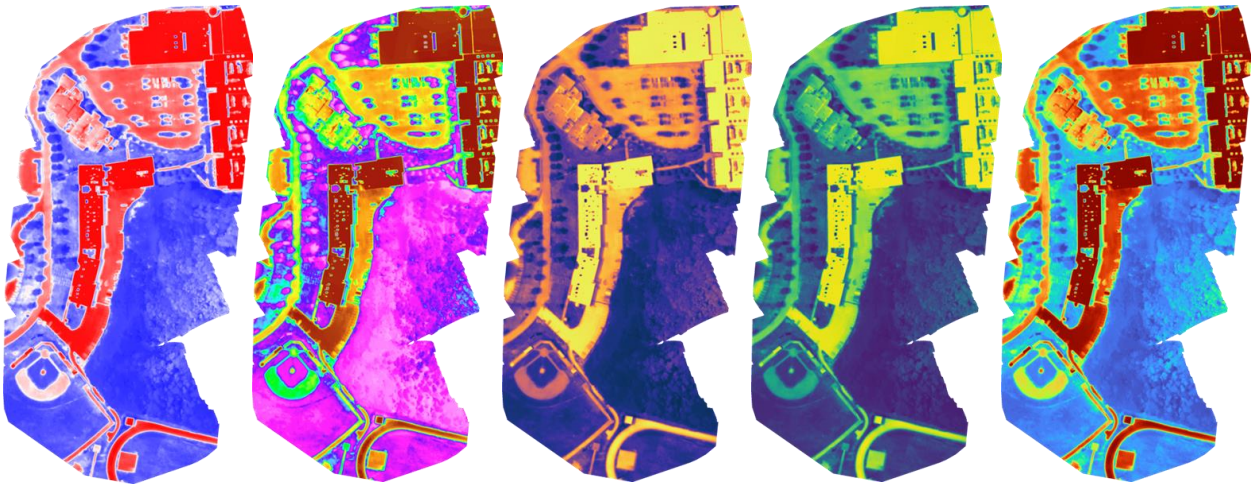
$$\textit{Thermal image} * 0.00341802 + 149.0$$

The temperature values in the resulting image were in Kelvin. To convert it into Celsius, -273.15 was subtracted from the pixel values. Colormaps were then applied to the images for visualization, as described in the next section.

3.3.Independent Variables

3.3.1. Colormap

The study tested 5 color schemes (colormaps) which are commonly used to visualize thermal data. In practice, dozens are colormaps are used in science, industry, and popular media to communicate temperature. Five were selected for this study as representative of the range of those often used: (a) diverging blue-white-red, (b) Esri's standard conventional temperature ramp, (c) FLIR's ironbow designed specifically for thermal mapping, (d) viridis, a colorblind-friendly sequential scheme and the default colormap in Python's matploblib, and (e) turbo, a perceptually uniform version of the both popular and much-maligned "jet" colormap (Figure 9). ArcGIS Pro (version 2.9) was used to apply the colormap to the thermal image data for the Esri temperature map, and Python was used to translate the other four. The application of the ironbow colormap required the installation of the cmocean package ("thermal" colormap, Thyng et al., 2016), and the rest used standard colormaps included in the matplotlib package.



(a) (b) (c) (d) (e)

Figure 9 Colormaps a) blue-white-red, b) Esri's temperature, c) ironbow, d) viridis, e) turbo

3.3.2. Shading

In topographic maps, relief shading helps to visualize variations in the surface and allows their realistic depiction and interpretation (Imhof, 1982; Onorati et al., 1992). Similarly, this element when applied to thermal images may afford more accurate or faster interpretation. The initial experiment was with shading based on the thermal gradient. However, an alternative shading technique based on the underlying digital surface model (DSM) that was created during the 3D reconstruction was ultimately selected since it conveyed a better sense of the scene. For drone imagery, the method was as follows. First, a slope layer is calculated from the DSM and then applied to the thermal image by first projecting it into CIELAB colorspace and then modifying the luminosity channel according to indicated slope shading. Lastly, the output image was globally re-brightened to account for the darkening from the slope-shading and projected back to RGB colorspace to record the final image. For the satellite imagery, the same method was used based on a lidar-derived DSM of the area. The lidar dataset was downloaded separately for the city area from the Virginia Geographic Information Network (VGIN), tiled together and then converted into DSM to later integrate with the thermal satellite image. A comparison of shaded vs. unshaded images is shown in Figure 10.

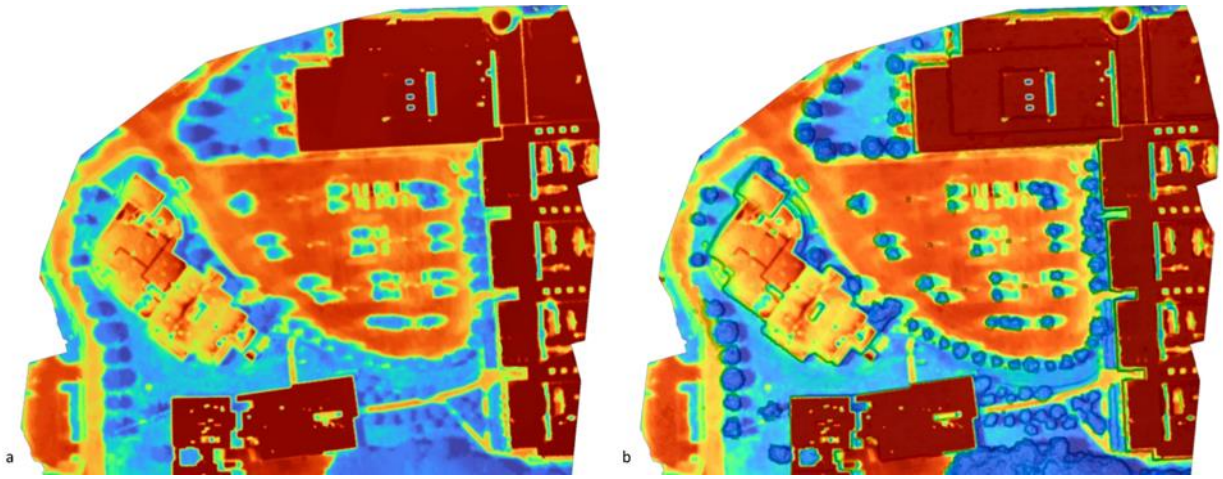


Figure 10 Selected area of (a) unshaded and (b) DSM-shaded thermal images using the turbo colormap.

3.3.3. Resolution

The study also looked into the resolution of the visualization. The lowest unit of measurement that may be observed or recorded for an item is referred to as resolution (Tobler, 1988). imagery produced by drones has a higher compared to that of a satellite image. Therefore, in this study, source and resolution were tied together, although conceptually the resolution of the image is the key factor under consideration. In this study, an infrared camera mounted on a drone generated high-resolution 2D visualization of the urban thermal environment. The resolution of the drone image was 15 cm over a 10.4-hectare extent. The satellite images were produced for 34.31 km^2 extent at 1 m resolution. Though the original resolution of the Landsat-8 ARD thermal data is 30 m, the 1 m resolution was required for the overlay of the higher resolution lidar data for the shading condition. An example of lidar-shaded satellite-derived thermal imagery is shown in Figure 11.

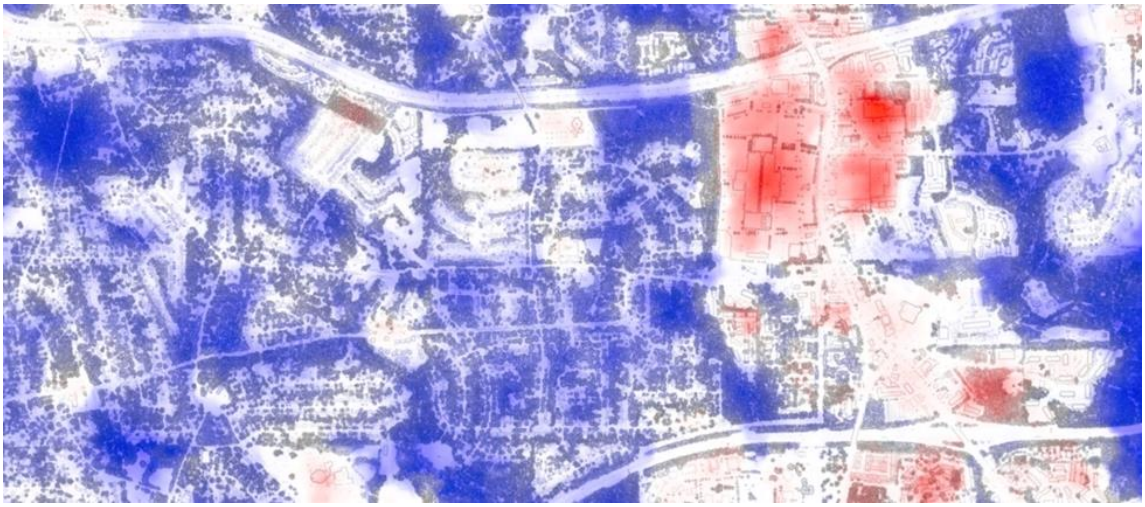


Figure 11 Selection of the satellite and lidar-DSM shaded image using blue-white-red colormap

3.4.Map Reading Tasks

Three map-reading tasks were used in the study to simulate a range of potential thermal-map use cases, and accuracy and response time metrics were used to gauge performance. The tasks were selected based on analogous tasks from both the widely used Topographic Map Assessment (Newcombe *et al.*, 2015) which measures users' ability to interpret topographic maps that represent an elevation surface, and a similar lidar visualization study (Pingel and Clarke, 2014).

The study then encompassed three tasks, five colormaps, and two shading conditions in two resolutions. To accommodate that, each task was divided into two separate blocks based on the image resolution (i.e., drone-sourced high-resolution image and city-scale satellite image). These six (3 x 2) blocks were randomized during the trial. Each combination of colormap and shading was repeated in a trial five times using different locations and/or rotation amounts, making each block consist of 50 questions (5 x 2 x 5). Instructions were provided on the screen. Each block also included an additional training phase of 5 questions that were not scored for the analysis. An administrator was always present to dictate the instructions and assist wherever required.

3.4.1. Task 1: Pairwise Comparison

In the first task, the users were shown a map with two points (A and B) identified. Based on a provided legend, they were asked to identify which point had a higher temperature. This question was modeled after the Topographic Map Assessment Task 7. The task was completed twice, once for the drone-sourced data and again for the satellite-sourced data. Response time and accuracy were recorded for each question.

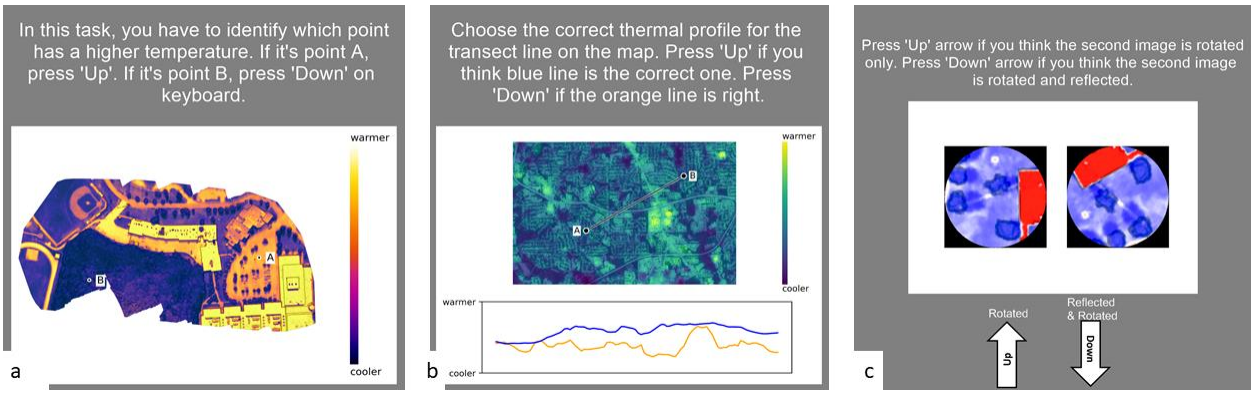


Figure 12 Map reading tasks (a) pairwise comparison, (b) profile estimation, and (c) map rotation.

3.4.2. Task 2: Profile estimation

The next map-reading task asked participants to select the correct thermal profile from two options for a line that was drawn on the map image. As with other tasks, drone and satellite versions were conducted separately, and both accuracy and response time were recorded. This test is analogous to both the Topographic Map Assessment - section 5 (Newcombe et al., 2015) and Pingel and Clarke (2014), and is similar to how UAV-derived thermal data is presented by Song and Park (2020).

3.4.3. Task 3: Map Rotation

The last map-reading task that was tested is the map rotation task based on Shepard and Metzler (1971) and Pingel and Clarke (2014). This task presents two extracts of a thermal map, one of which has been rotated. In some images, however, the second extract was both rotated *and* reflected. Participants were asked to identify whether the second extract was reflected, which requires them to match key points in images and to mentally rotate them to see if they match (Figure 12(c)). The task measures the saliency of visualization. Because rotation amount is correlated to response time (Shepard & Metzler, 1971), the rotation amount was systematically varied and randomized to control for this effect.

3.5.Post-Task Questionnaire

During the post-task questionnaire, participants were asked to report simple demographic information (age, major, level of education), as well as Likert-type questions asking participants to explicitly rate how they regarded the visualizations (concerning colormaps, augmentations, and scales). Such statements were considered in light of Hegarty *et al*. (2009) who demonstrated that stated user preferences do not always match objective measures of performance. They were also asked whether they had a general difficulty in differentiating among colors to assess colorblindness. However, this question may not have been interpreted as intended, as no participants reported difficulty.

4. Results

The results were analyzed using descriptive statistics and repeated measures ANOVA in JMP: Statistical Software and in Python. A standard least squares model (Restricted Maximum Likelihood, REML) was specified in JMP with the main effects of task, shading enhancement, colormaps and source, as well as second-order interactions, and participants as a random effect. F-value was used in the repeated measures ANOVA analysis to determine the statistical significance of the differences. The F-value calculation computes the explained variance to unexplained variance ratio (Feir-Walsh & Toothaker, 1974). Descriptive statistics are reported using Least Squares Mean (M) and Standard Error (SE) as reported by JMP; these describe the centrality and spread of the variables with the effects of other variables factored away. The response times for all three tasks were log-transformed to normalize the data for analysis. Effect sizes were calculated using eta squared value (η^2), used specifically used to report effect sizes in ANOVA models (Lakens, 2013). Eta squared corresponds to the proportion (0 to 1) of variance explained by a given variable. It is a standardized estimate of effect size, which means that it is comparable across outcome variables measured in various units (Adams & Conway, 2014). Lakens's (2013) descriptions of effect sizes are used to interpret magnitudes ($\eta^2 = 0.01$ is small, $\eta^2 = 0.06$ is medium, and $\eta^2 = 0.14$ is large).

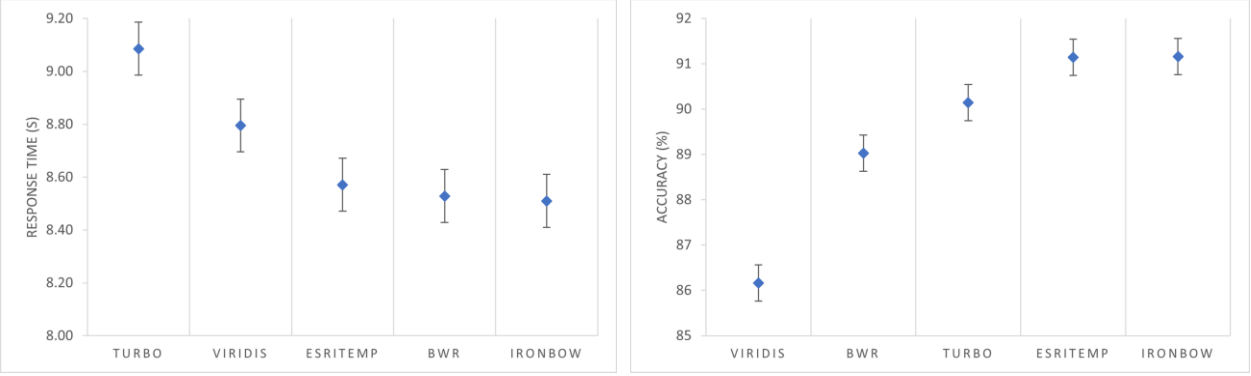
4.1. Task

Participants informally indicated that the profile estimation task was the most difficult and time-consuming among the three map reading tasks. The response time for the profile estimation task (M = 13.50 s, SE = 0.08) was significantly longer than for either the map rotation task (M = 8.22, SE = 0.08) or the A-B comparison task (M = 4.37, SE = 0.0.08), according to repeated measures

ANOVA analysis using log-transformed values ($F_{2,65}$ = 6749.96, $p < 0.001$). The effect size of this difference was large ($\eta^2 = 0.93$). Accuracy was most accurate for the map rotation task (M = 92.04%, SE = 0.36%), followed by the A-B comparison task (M = 89.77%, SE = 0.36%) and profile estimation task (M = 86.77%, SE = 0.36%), a statistically significant result ($F_{2,65}$ = 53.04, $p < 0.001$) with a medium effect size ($\eta^2 = 0.13$).

4.2.Colormaps

It was expected that there would be significant differences between colormaps and that the turbo and ironbow colormaps would perform better than the others, that is, they would have a lower response time indicating more efficiency and/or higher mean accuracy. Aggregated across all tasks and controlling for task differences and other sources of variability, overall response time was indeed lowest for ironbow colormap (M = 8.51, SE = 1.03) but was highest for turbo (M = 9.09, SE = 1.07). Ironbow (M = 91.16%, SE = 3.49%) and Esri's default temperature colormap (M = 91.14%, SE = 3.50%) had the highest accuracy rates. The lowest accuracy was for viridis (M = 86.16%, SE = 4.25%), which also performed poorly with respect to response time. Overall, the results show that ironbow performed the best, and viridis performed the worst (Figure 13). The repeated measures ANOVA showed the mean difference among the colormaps were significant for both the log of response time ($F_{4,65}$ = 12.16, $p < 0.001$) and accuracy ($F_{4,65}$ = 19.62, $p < 0.001$). Although statistically significant, the effect size is small for log-response time ($\eta^2 < 0.01$) but was medium for accuracy ($\eta^2 = 0.09$).

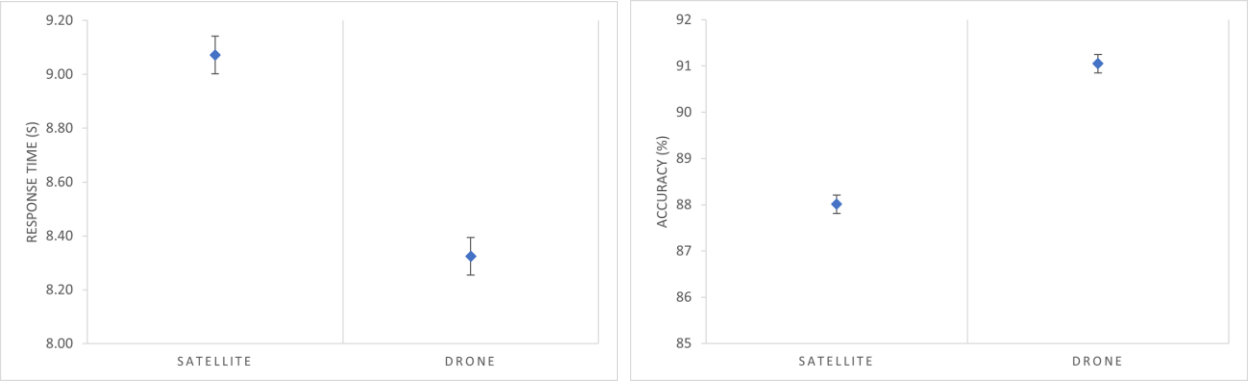


(a) (b)

Figure 13 Least square means and standard error plot of colormaps for (a) response time and (b) accuracy

4.3.Source

It was expected that the drone-sourced higher-resolution imagery would perform better compared to the city-scale satellite imagery. Differences in response time were not statistically significant ($F_{1,65} < 0.01$, $p = 0.96$), but were for accuracy ($F_{1,65} = 52.80$, $p < 0.001$). Judgments made from drone imagery were more accurate ($M = 91.05\%$, $SE = 3.51\%$) than those made from satellite imagery ($M = 88.01\%$, $SE = 4.00\%$). The effect size was medium ($\eta^2 = 0.06$) for accuracy. Figure 14 shows the least square means difference of source for accuracy and response time.

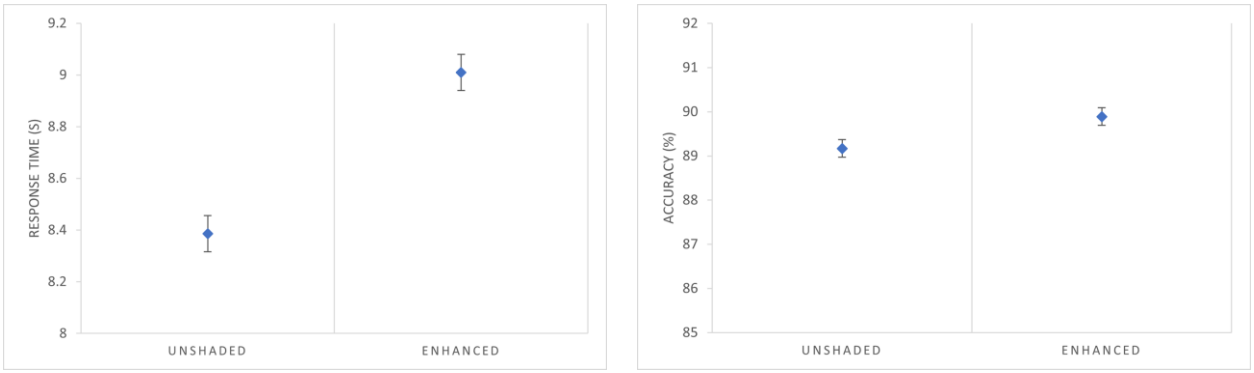


(a) (b)

Figure 14 Least squares means and standard error plots for imagery source

4.4.Shading

It was expected that the addition of lidar-derived surface shading as an augmentation/enhancement would improve performance. Repeated measures ANOVA analysis indicated that differences in accuracy were not significant ($F_{1,65}$ = 2.92, p = 0.09), but that differences in log-response time were ($F_{1,65}$ = 128.77, p < 0.001). Contra to expectations, the unshaded images had a lower (faster) response time (M = 8.39, SE = 1.02) than the enhanced images (M = 9.01, SE = 1.03) (Figure 15), but with a small effect size (η^2 = 0.01)



(a) (b)

Figure 15 Least squares means plots for shading enhancement

Table 1 Least mean square and standard error of response time and accuracy for tasks, colormap, source, and shading. Levels not connected by the same letter are statistically different.

		Response Time (s)			Accuracy (%)		
		M	SE	Level	M	SE	Level
Task	a_or_b	4.37	0.08	A	89.77	0.36	A
	rotation	8.22	0.08	B	92.04	0.36	B
	profile	13.50	0.08	C	86.77	0.36	C
Colormap	Ironbow	8.51	0.10	A	91.16	0.47	A
	Esritemp	8.57	0.10	A	91.14	0.47	A
	Turbo	9.09	0.10	B	90.14	0.47	AB
	BWR	8.53	0.10	A	89.03	0.47	B
	Viridis	8.80	0.10	A	86.16	0.47	C
Source	Satellite	9.07	0.07	A	88.01	0.30	A
	Drone	8.32	0.07	B	91.05	0.30	B
Shading	Unshaded	8.39	0.07	A	89.17	0.30	A
	Enhanced	9.01	0.07	B	89.89	0.30	A

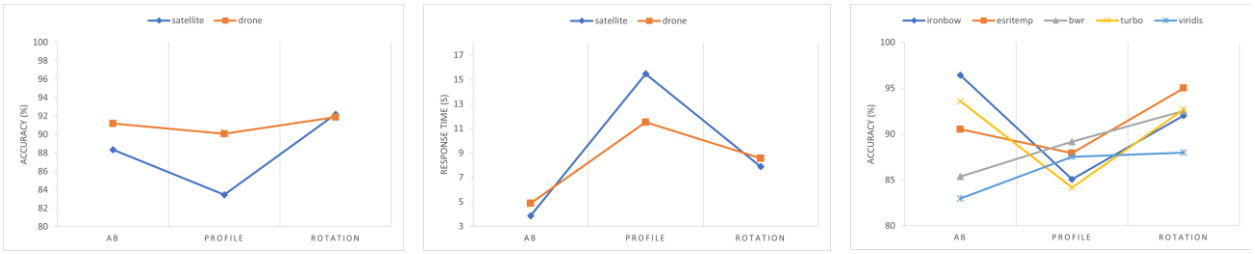
Table 2 ANOVA results table for log-response time and accuracy

		log(Response Time)			Accuracy		
	DF	F	Prob > F	η^2	F	Prob > F	η^2
Task	2	6749.96	< 0.001	0.93	53.04	< 0.001	0.13
Colormap	4	12.16	< 0.001	< 0.01	19.62	< 0.001	0.09
Source	1	< 0.01	0.96	< 0.01	52.80	< 0.001	0.06
Shading	1	128.77	< 0.001	0.01	2.92	0.09	< 0.01
Task * Colormap	8	8.60	< 0.001	< 0.01	22.05	< 0.001	0.21
Task * Source	2	304.42	< 0.001	0.04	23.02	< 0.001	0.06
Task * Shading	2	34.19	< 0.001	< 0.01	5.94	0.003	0.01
Colormap * Source	4	18.06	< 0.001	< 0.01	67.50	< 0.001	0.33
Colormap * Shading	4	18.43	< 0.001	0.01	4.87	< 0.001	0.02
Source * Shading	1	2.98	0.08	< 0.01	64.42	< 0.001	0.08

4.5.Interactions

The ANOVA test results show that the task interacted with both colormap and source. The task-colormap interaction for accuracy ($F_{8,65}$ = 22.05, $p < 0.001$) and log-response time ($F_{8,65}$ = 8.60, $p < 0.001$) were significant. However, the effect size was large ($\eta^2 = 0.21$) for accuracy and small for response time ($\eta^2 < 0.01$). A similar case was found for task-source interaction where both accuracy ($F_{2,65}$ = 23.02, $p < 0.001$, $\eta^2 = 0.06$) and log-response time ($F_{2,65}$ = 304.42, $p < 0.001$, $\eta^2 = 0.04$). The effect size was medium for both. The task-source accuracy plot (Figure 16a) shows that the accuracy was higher for drone images in the a-b comparison task and the profile estimation task, but not for the map rotation task. From the response time plot (Figure 16b) of the

same interaction, it can be seen that the drone images have a higher response time only in the profile estimation task. In effect, the drone images required more time for interaction but afforded higher accuracy. In the task-colormap interaction (Figure 16c) it can be seen that the performance of many colormaps varied significantly by task. Overall, Esri's temperature colormap performed well on all tasks, and viridis performed poorly on most tasks. Ironbow and turbo were particularly effective for the a-b comparison task.

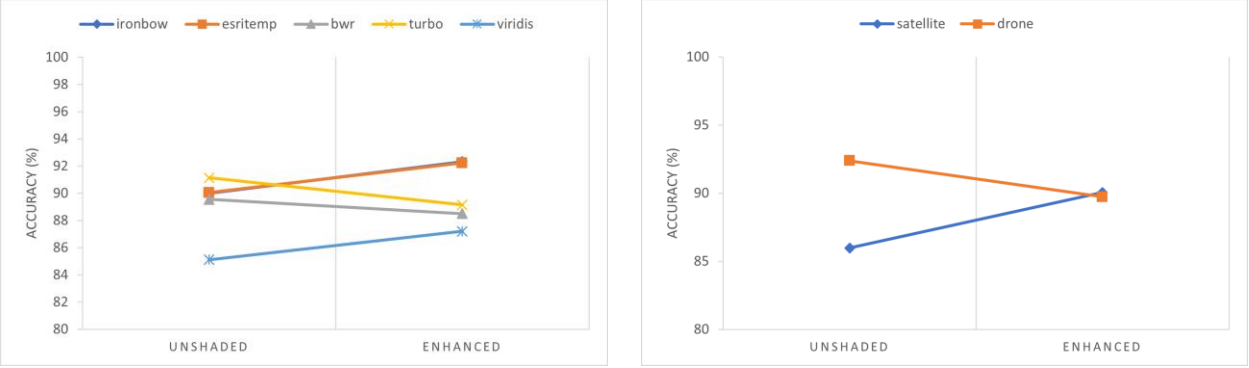


(a) (b) (c)

Figure 16 LSM plots for (a) task-source interaction for accuracy, (b) task-source interaction for response time, and (c) task-colormap interaction for accuracy

Shading enhancement interacted with both colormap and source. The enhancement-colormap interaction for both accuracy ($F_{4,65} = 4.87$, $p < 0.001$) and log-response time ($F_{4,65} = 18.43$, $p < 0.001$) were statistically significant, but the effect size was medium for accuracy ($\eta^2 = 0.02$) and small for log-response time ($\eta^2 = 0.01$). The shading enhancement increased accuracy for the ironbow, Esri temperature, and viridis colormaps, but decreased accuracy for the blue-white-red and turbo colormaps. Ironbow had the highest accuracy for enhanced images, and the turbo had the highest accuracy for unshaded images. The shading-source interaction was not significant for log-response time ($F_{1,65} = 2.98$, $p = 0.08$, $\eta^2 < 0.01$), but was significant for accuracy ($F_{8,65} = 64.42$, $p < 0.001$) with a medium effect size ($\eta^2 = 0.08$). Shading enhancement improved accuracy for the satellite images (from 85.97% to 90.05%), but reduced accuracy for drone-

sourced images (from 92.37% to 89.73%). Figure 17 shows the shading interaction with colormaps and source for accuracy.



(a) (b)

Figure 17 Least square means plot for (a) accuracy of shading-colormap and (b) accuracy of shading-source. Esri temperature and ironbow colormaps were virtually identical for accuracy on unshaded/enhanced conditions.

4.6.Preferences

The post-task questionnaire asked the participants about their experience with geography on a 5-point scale (M = 3.03, SD = 0.98), but it was not correlated with response time ($r(66)$ = -0.05, p = 0.70) or accuracy ($r(66)$ = 0.08, p = 0.54). Among the participants, 74% were at the graduate level and 20% of them were at the undergraduate level. According to the participant's preference, turbo ranks first (M = 2.27, SD = 1.23) and the Esritemp ranks last (M = 4.13, SD = 1.39). 67% of the participants preferred enhanced images and the rest preferred unshaded images. Figure 18 shows a parallel coordinate plot tying together response time, accuracy, and user preferences for the five colormaps tested.

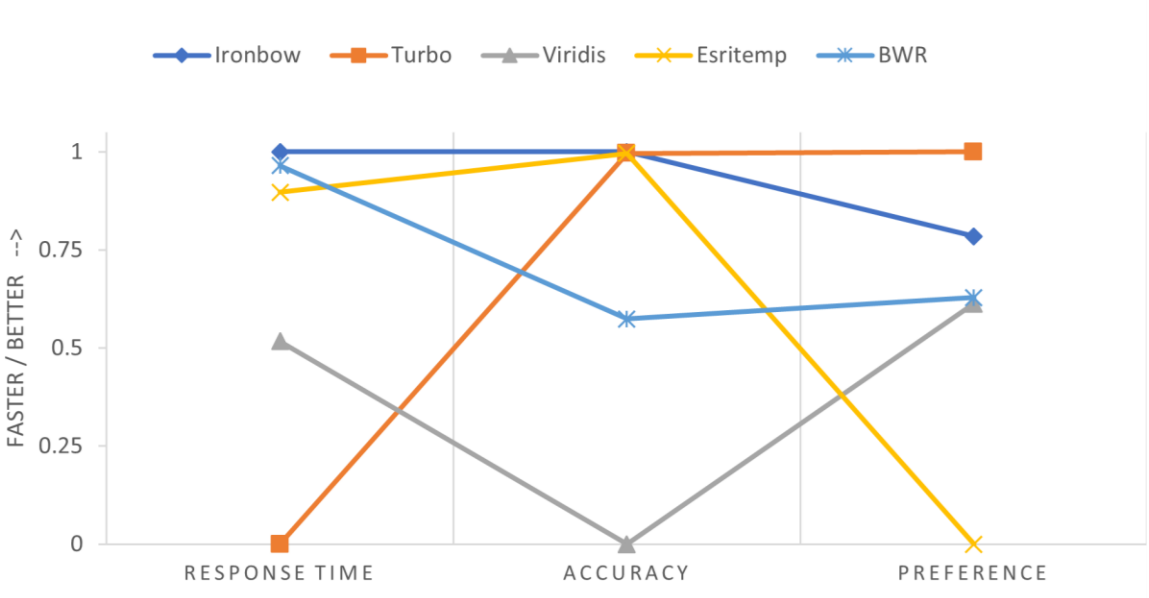


Figure 18 Parallel coordinate plot of normalized response time, accuracy, and user preference.

5. Discussion

Visualizations can significantly influence how information is processed by decision-makers. This study tested how colormap, shading augmentation, and resolution impact the reading of thermal maps, using three different map reading tasks for assessment. The results show a significant difference in the performance of colormaps (both accuracy and response time), accuracy for the resolution, and response time for the shading enhancement. The study also found significance in task interaction with source (accuracy and log-response time) and colormap (accuracy), and shading interaction with colormap (accuracy) and source (accuracy). The post-task questionnaire recorded user demographics and their preference for colormaps and shading. The results provided substantial evidence to support Hegarty et al.'s (2009) findings that user preferences about visualizations do not always correspond to objective measures of performance.

This study found a preference for the rainbow-style turbo colormap, which was accurate (within 0.02% of the ironbow, the best performer) but the slowest in response time. Despite criticism in the cartography and data visualization community (Brewer 1997; Borland and Taylor 2007; Moreland, 2016), rainbow and spectral colormaps are still considered visually appealing (Moreland, 2016; Zeller et al., 2020) as evidenced by their popularity of jet colormap in MATLAB and Python. Turbo was used in this study instead of jet to avoid 'banding' or 'streaking' caused by uneven brightness, a common criticism of jet (Mikhailov, 2019). The study supports the preference for rainbow colormaps (Borkin et al., 2011; Zeller et al. 2020), and provides evidence for more accurate results (Reda and Szafir, 2021) but with a cost of slower response time (Borkin et al., 2011; Liu and Heer, 2018). The post-task question aimed at gathering participants' self-reported difficulties in color differentiation, a major criticism of rainbow colormaps, however, did not yield the expected results as no participants reported any

difficulties. As a result, the impact of colorblindness on the outcome could not be assessed. *Ironbow* had the best performance among the colormaps unambiguously as it had the lowest response time and the highest accuracy. Ironbow utilizes two main colors to create a sequential colormap in a uniform distribution (Sousa *et al*., 2020) with increasing luminance and can be effective for thermal data visualization (Son *et al*., 2019). The land surface temperature data is interval data and an effective colormap for interval data should have uniform hue or luminance distribution (Bergman *et al.*, 1995). In support of an appropriate choice of colormap to visualize thermal data, the study results reflected that the ironbow colormap supports its position as an effective one. Here, user preference did align well with performance as ironbow was the second most preferred colormap. Esri's temperature colormap (*esritemp*) was the least preferred colormap but was a strong objective performer, and viridis was both a poor performer and not well-liked, despite being a favorite of the scientific visualization community for its colorblind-friendliness and perceptual uniformity from dark to light. (Viridis is now the default colormap in both Matlab and Python.) Taken altogether, support was found for Liu and Heer's (2018) hypothesis that using carefully designed multi-hue color schemes can improve the performance of visualizations. Higher-resolution drone images had lower response times and higher accuracy compared to satellite images, although the response time results were not significant.

The DSM-shading enhancement to the thermal images allowed users to see the land patterns underneath more easily, and two-thirds of users preferred these augmented/shaded images. Shading allows for a more direct perception of important considerations such as the orientation of rooftops since solar direction plays a role in surface temperature variation (Najafifar *et al.*, 2019). Performance for this cartographic enhancement was mixed, however, with judgments from shaded images made more accurately, but also more slowly. As with the multi-hued

colormaps, the reason expected for slower interpretation is that the images contain more information, and thus take more time to process. The sub-band entropy metric of "visual clutter" (Rosenholtz *et. al.,* 2007) was used to quantify the information content by colormap and shading condition (Figure 19). Two items are apparent from this plot. First, enhanced images contain more information that is resulting in a higher visual clutter score. Second, multi-hued colormaps such as turbo and esritemp have a higher visual clutter score than the simpler sequential color schemes ironbow and viridis. It is suggested that the difference in information content, as measured by sub-band entropy, explains the difference in response time for the shading condition and is a partial explanation of the accuracy/response time patterns observed for the colormap conditions. Users are taking more time to interpret the more complex images but are also more accurately interpreting them.

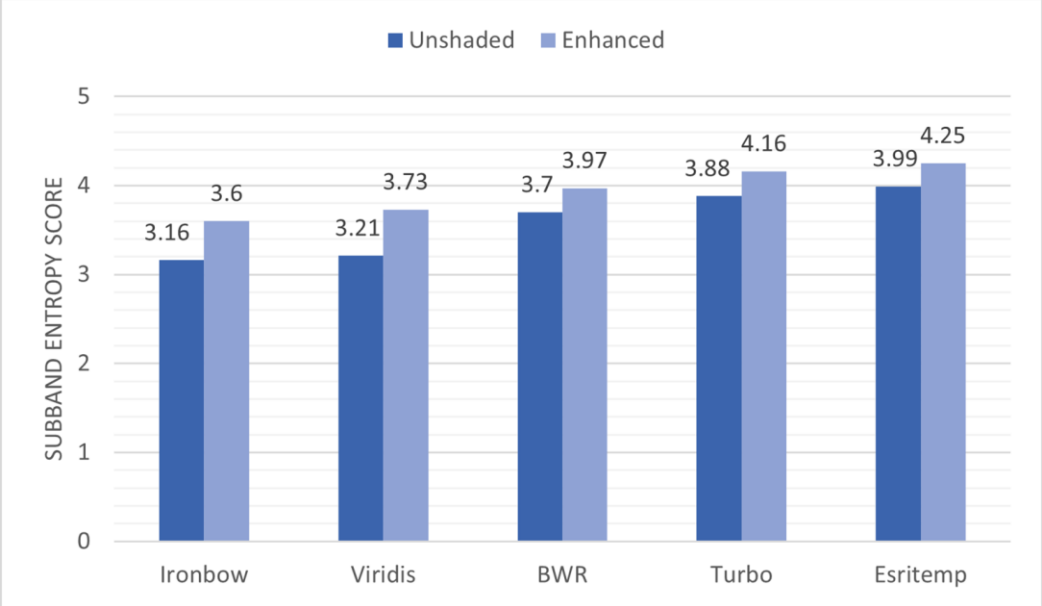


Figure 19 Visual clutter score for images in different colormaps with and without shading enhancement

The study was conducted on the university campus, with 74% of the participants being graduate-level students and 20% being undergraduate-level students, as indicated in the results of the preference. The study's goal was to influence policymakers' ability to interpret urban heat

information, but it should be noted that the results may differ if the participants were from a non-student population. Factors such as sociocultural, economic, and political backgrounds can affect cognition and perception of visual stimuli, as shown in studies comparing different cultures' visual perception, such as American and Japanese (Masuda & Nisbett, 2006; Miyamoto et al., 2006), American and East Asian exchange students (Masuda & Nisbett, 2006), and Caucasian, African, Asian, Latino, and Multiracial American (Choi *et al.*, 2015). The study by Patterson and Jenny (2014) also highlighted differences in how the US and European populations perceive hypsometric tints in maps. The study found out that while U.S. population perceive pale orange-brown representing environmental phenomena, the Europeans describe the color represents topographic phenomena. Moreover, while Ooms et al. (2015) demonstrated that higher education levels result in better map-reading accuracy, Cybulski (2020) found no statistically significant differences in map reading based on age, gender, or frequency of use. Therefore, to gain insight into how diverse sociocultural, demographic, and economic factors may influence thermal map reading, future studies should consider administering the study with a more diverse population.

6. Conclusions and Future Work

It was hypothesized that the appropriate choice of the colormap, shading augmentation, and resolution would improve thermal map reading. Based on the participant's performance in the user study, the ironbow colormap is the most effective one for thermal map reading. Esritemp was a strong performer, but it was not well-liked. The turbo colormap was the most popular but the users took the longest to interpret, although judgments made from it were among the most accurate. Despite being a well-liked colorblind-friendly colormap in the scientific community, viridis was a poor performer overall. Overall, response time and accuracy for colormap were statistically significant, but preference was not strongly related to performance. The DSM-based shading enhancement increases the accuracy of thermal maps; however, the map reading takes longer because the enhancement provides an increased complexity in multi-hue and shaded images. That results in longer cognitive processing time but ensures higher accuracy.

The UAV-derived higher-resolution thermal maps (15 cm) produced by drones allowed users to make more accurate and faster judgments compared to the coarser resolution satellite imagery. The high-resolution drone images also have the advantage of making distinguishable features like buildings, trees, parking lots, and recreation areas possible. The large cool areas provided by urban forests can be directly perceived in these images, as can the ground shade next to specific trees. Contrarily, even the most basic grass-filled spaces next to buildings and parking lots stand out as warmer.

The findings suggest that selecting a colormap with or without shading in a specific resolution can have a significant impact on the user's performance when reading thermal maps. A good combination of these variables can present a more accurate thermal scenario in a spatial context,

potentially adding useful information to thermal maps. The availability of information may result in better decision-making outcomes by reducing uncertainty (Shapiro & Varian, 1998). This statement is true only when a decision-maker has the ability to process the available information (Bazerman et al., 1998; Lachman et al., 1979). As a result of their limited cognitive resources, decision-makers may be easily overburdened when too much information becomes available. The visual representation of information, on the other hand, may improve a decision-ability maker's ability to process information (Coury and Boulette, 1992).

Future research will build on these findings to move beyond concerns about the effectiveness and performance of maps and to concentrate on these larger challenges. The methods used here to produce 2D maps can also be used to produce 3D maps. Future research will examine the effects of 3D and immersive thermal maps on readers' perceptions of the distribution and origins of urban heat. Users using 3D maps might be able to observe how buildings' vertical walls, windows, and doors—features that are much more prevalent in daily life than rooftops—emit heat energy into the surrounding air.

Apart from the visualization perspective, future studies could also look at the high-resolution nighttime heat variation of urban areas. The Federal Aviation Administration (FAA) recently relaxed their flight restriction for operations at night for drones which might open the door to this kind of work. Furthermore, machine-learning-based extrapolation of the thermal data collected using drones for certain areas within the city to get an overview of the whole city could be another aspect that can be explored.

The results provide actionable pieces of evidence for choosing colormaps, resolution, and shading enhancement specifically for thermal map reading. Selecting an appropriate thermal map

visualization method will help urban planners, landscape planners, architects, and policymakers to understand and communicate the distribution of urban heat in an effective manner.

References

Anastasi, A. (1968). *Psychological Testing*, (3rd edition). New York, Macmillan.

Adaktylou, N. (2020). Remote Sensing as a Tool for Phenomenon-Based Teaching and Learning at the Elementary School Level: a Case Study for the Urban Heat Island Effect. *International Journal of Educational Methodology, 6*(3), 517–532. https://doi.org/10.12973/ijem.6.3.517

Adams, M. A., & Conway, T. L. (2014). Eta Squared. *Encyclopedia of Quality of Life and Well-Being Research*, 1965–1966. https://doi.org/10.1007/978-94-007-0753-5_918

Baik, J. J., Kim, Y. H., Kim, J. J., & Han, J. Y. (2006). Effects of boundary-layer stability on urban heat island-induced circulation. *Theoretical and Applied Climatology, 89*(1–2), 73–81. https://doi.org/10.1007/s00704-006-0254-4

Basu, R. (2009). High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008. *Environmental Health, 8*(1), 40. https://doi.org/10.1186/1476-069X-8-40

Benz, S. A., & Burney, J. A. (2021). Widespread Race and Class Disparities in Surface Urban Heat Extremes Across the United States. *Earth's Future, 9*(7). https://doi.org/10.1029/2021EF002016

Berger, C., Riedel, F., Rosentreter, J., Stein, E., Hese, S., & Schmullius, C. (2014). Fusion of Airborne Hyperspectral and LiDAR Remote Sensing Data to Study the Thermal Characteristics of Urban Environments. *Computational Approaches for Urban Environments*, 273–292. https://doi.org/10.1007/978-3-319-11469-9_11

Bergman, L. D., Rogowitz, B. E., & Treinish, L. A. (1995). A rule-based tool for assisting colormap selection. *IEEE Visualization*, 118–125. https://doi.org/10.5555/832271.833862

Bernard, J., Bocher, E., Petit, G., & Palominos, S. (2018). Sky View Factor Calculation in Urban Context: Computational Performance and Accuracy Analysis of Two Open and Free GIS Tools. *Climate, 6*(3), 60. https://doi.org/10.3390/cli6030060

Berry, J. W., Poortinga, Y. H., Breugelmans, S. M., Chasiotis, A., & Sam, D. L. (2012). *Cross-cultural psychology: Research and applications*. Cambridge: Cambridge University Press.

Board, C. (1978). Map Reading Tasks Appropriate in Experimental Studies in Cartographic Communication. *Cartographica: The International Journal for Geographic Information and Geovisualization, 15*(1), 1–12. https://doi.org/10.3138/ag15-v252-3726-w346

Borkin, M., Gajos, K., Peters, A., Mitsouras, D., Melchionna, S., Rybicki, F., Feldman, C., & Pfister, H. (2011). Evaluation of Artery Visualizations for Heart Disease Diagnosis. *IEEE Transactions on Visualization and Computer Graphics, 17*(12), 2479–2488. https://doi.org/10.1109/tvcg.2011.192

Borland, D., & Taylor, R. M. (2007). Rainbow Color Map (Still) Considered Harmful. *IEEE Computer Graphics and Applications, 27*(2), 14–17. https://doi.org/10.1109/mcg.2007.323435

Bornstein, R., & Lin, Q. (2000). Urban heat islands and summertime convective thunderstorms in Atlanta: three case studies. *Atmospheric Environment, 34*(3), 507–516. https://doi.org/10.1016/s1352-2310(99)00374-x

Brewer, C. A. (2003). A transition in improving maps: The ColorBrewer example. *Cartography and Geographic Information Science, 30*(2), 159-162. https://doi.org/10.1559/152304003100011126

Brewer, C.A. (1997). Spectral schemes: controversial color use on maps. *Cartography and Geographic Information Systems, 24*(4), 203-220. https://doi.org/10.1559/152304097782439231

Carbonell Carrera, C., Avarvarei, B. V., Chelariu, E. L., Draghia, L., & Avarvarei, S. C. (2017). Map-reading skill development with 3D technologies. *Journal of Geography, 116*(5), 197-205. https://doi.org/10.1080/00221341.2016.1248857

Chen, Y. C., Chen, C. Y., Matzarakis, A., Liu, J. K., & Lin, T. P. (2016). Modeling of mean radiant temperature based on comparison of airborne remote sensing data with surface measured data. *Atmospheric Research*, 174–175, 151–159. https://doi.org/10.1016/j.atmosres.2016.01.004

Chen, Y.-C., Chiu, H.-W., Su, Y.-F., Wu, Y.-C., & Cheng, K.-S. (2017). Does urbanization increase diurnal land surface temperature variation? Evidence and implications. *Landscape and Urban Planning*, 157. https://doi.org/10.1016/j.landurbplan.2016.06.014

Choi, H., Connor, C. B., Wason, S. E., & Kahan, T. A. (2015). The effects of interdependent and independent priming on Western participants' ability to perceive changes in visual scenes. *Journal of Cross-Cultural Psychology*, 47(1), 97–108.

Clark, S. S., Chester, M. V., Seager, T. P., & Eisenberg, D. A. (2018). The vulnerability of interdependent urban infrastructure systems to climate change: could Phoenix experience a Katrina of extreme heat? *Sustainable and Resilient Infrastructure, 4*(1), 21–35. https://doi.org/10.1080/23789689.2018.1448668

Cook, B., Corp, L., Nelson, R., Middleton, E., Morton, D., McCorkel, J., Masek, J., Ranson, K., Ly, V., & Montesano, P. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing, 5*(8), 4045–4066. https://doi.org/10.3390/rs5084045

Coury, B. G., & Boulette, M. D. (1992). Time Stress and the Processing of Visual Displays. Human Factors: *The Journal of the Human Factors and Ergonomics Society, 34*(6), 707–725. https://doi.org/10.1177/001872089203400605

Crawford, P. v. (1973). The Perception of Graduated Squares as Cartographic Symbols. *The Cartographic Journal, 10*(2), 85–88. https://doi.org/10.1179/caj.1973.10.2.85

Cybulski, P. (2020). Spatial distance and cartographic background complexity in graduated point symbol map-reading task. *Cartography and Geographic Information Science, 47*(3), 244–260. https://doi.org/10.1080/15230406.2019.1702102

Cuff, D. J. (1973). Colour on Temperature Maps. *The Cartographic Journal, 10*(1). https://doi.org/10.1179/caj.1973.10.1.17

Cui, Y., Xu, X., Dong, J., & Qin, Y. (2016). Influence of urbanization factors on surface urban heat island intensity: A comparison of countries at different developmental phases. *Sustainability (Switzerland), 8*(8). https://doi.org/10.3390/su8080706

Daffara, C., Muradore, R., Piccinelli, N., Gaburro, N., de Rubeis, T., & Ambrosini, D. (2020). A Cost-Effective System for Aerial 3D Thermography of Buildings. *Journal of Imaging, 6*(8), 76. https://doi.org/10.3390/jimaging6080076

Dahaghin, M., Samadzadegan, F., & Dadrass Javan, F. (2021). Precise 3D extraction of building roofs by fusion of UAV-based thermal and visible images. *International Journal of Remote Sensing, 42*(18), 7002–7030. https://doi.org/10.1080/01431161.2021.1951875

Dambruch, J., & Krämer, M. (2014). Leveraging public participation in urban planning with 3D web technology. *Proceedings of the Nineteenth International ACM Conference on 3D Web Technologies - Web3D '14*, 117–124. https://doi.org/10.1145/2628588.2628591

Dastani, M. (2002). The Role of Visual Perception in DataVisualization. *Journal of Visual Languages and Computing, 13*, 601–622. https://doi.org/10.1006/S1045-926X(02)00026-5

Deilami, K., Kamruzzaman, M., & Liu, Y. (2018). Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures. In International *Journal of Applied Earth Observation and Geoinformation* (Vol. 67, pp. 30–42). Elsevier B.V. https://doi.org/10.1016/j.jag.2017.12.009

DeValois, R. L., & DeValois, K. K. (1990). *Spatial Vision (Oxford Psychology Series, 14) (Revised)*. Oxford University Press.

Dixon, P. G., & Mote, T. L. (2003). Patterns and causes of Atlanta's urban heat island–initiated precipitation. *Journal of Applied Meteorology, 42*(9), 1273-1284. https://doi.org/10.1175/1520-0450(2003)042%3C1273:PACOAU%3E2.0.CO;2

D'Ippoliti, D., Michelozzi, P., Marino, C., de'Donato, F., Menne, B., Katsouyanni, K., Kirchmayer, U., Analitis, A., Medina-Ramón, M., Paldy, A., Atkinson, R., Kovats, S., Bisanti, L., Schneider, A., Lefranc, A., Iñiguez, C., & Perucci, C. A. (2010). The impact of heat waves on mortality in 9 European cities: results from the EuroHEAT project. *Environmental Health, 9*(1), 37. https://doi.org/10.1186/1476-069X-9-37

Dwyer, J. L., Roy, D. P., Sauer, B., Jenkerson, C. B., Zhang, H. K., & Lymburner, L. (2018). Analysis ready data: enabling analysis of the Landsat archive. *Remote Sensing, 10*(9), 1363. https://doi.org/10.3390/rs10091363

Ebi, K. L., Balbus, J. M., Luber, G., Bole, A., Crimmins, A., Glass, G., Saha, S., Shimamoto, M. M., Trtanj, J., & White-Newsome, J. L. (2018). Human Health. In D. R. Reidmiller, C. W. Avery, D. R. Easterling, K. E. Kunkel, K. L. M. Lewis, T. K. Maycock, & B. C. Stewart (Eds.), Impacts, Risks, and Adaptation in the United States: *Fourth National Climate Assessment* (Vol. 2, pp. 1309–1345).

Ermida, S. L., Soares, P., Mantas, V., Göttsche, F. M., & Trigo, I. F. (2020). Google earth engine open-source code for land surface temperature estimation from the landsat series. *Remote Sensing, 12*(9), 1471. https://doi.org/10.3390/rs12091471

Feir-Walsh, B. J., & Toothaker, L. E. (1974). An Empirical Comparison of the Anova F-Test, Normal Scores Test and Kruskal-Wallis Test Under Violation of Assumptions. *Educational and Psychological Measurement, 34*(4), 789–799. https://doi.org/10.1177/001316447403400406

Fuhrmann, S., Komogortsev, O., & Tamir, D. (2009). Investigating Hologram-Based Route Planning. *Transactions in GIS, 13*, 177–196. https://doi.org/10.1111/j.1467-9671.2009.01158.x

Graham, N. V. S. (1989). *Visual Pattern Analyzers (Oxford Psychology Series, 16)*. Oxford University Press.

Griffin, A. L. (2017). Cartography, visual perception and cognitive psychology. In A. J. Kent & P. Vujakovic (Eds.), *The Routledge Handbook of Mapping and Cartography* (pp. 44–54). Routledge. https://doi.org/10.4324/9781315736822-5

Guardaro, M., Messerschmidt, M., Hondula, D. M., Grimm, N. B., & Redman, C. L. (2020). Building community heat action plans story by story: A three neighborhood case study. *Cities, 107*, 102886. https://doi.org/10.1016/j.cities.2020.102886

Habeeb, D., Vargo, J., & Stone, B. (2015). Rising heat wave trends in large US cities. *Natural Hazards, 76*(3), 1651–1665. https://doi.org/10.1007/s11069-014-1563-z

Haeberling, C. (2005, July). Cartographic design principles for 3D maps–A contribution to cartographic theory. *In Proceedings of ICA Congress Mapping Approaches into a Changing World.*

Hajat, S., Vardoulakis, S., Heaviside, C., & Eggen, B. (2014). Climate change effects on human health: projections of temperature-related mortality for the UK during the 2020s, 2050s and 2080s. *Journal of Epidemiology and Community Health, 68*(7), 641–648. http://dx.doi.org/10.1136/jech-2013-202449

Hansen, A., Bi, P., Nitschke, M., Ryan, P., Pisaniello, D., & Tucker, G. (2008). The Effect of Heat Waves on Mental Health in a Temperate Australian City. *Environmental Health Perspectives, 116*(10), 1369–1375. https://doi.org/10.1289/ehp.11339

Harlan, S. L., Brazel, A. J., Jenerette, G. D., Jones, N. S., Larsen, L., Prashad, L., & Stefanov, W. L. (2007). In the shade of affluence: the inequitable distribution of the urban heat island. https://doi.org/10.1016/S0196-1152(07)15005-5

Hayek, U. W., Neuenschwander, N., & Grêt-Regamey, A. (2016). AM Facilitating well-informed trade-off decision making on land use change : Integrating rules and indicators of ecosystem service provision into procedural 3 D visualization. https://scholarsarchive.byu.edu/iemssconference/2012/Stream-B/247/

Hegarty, M., Smallman, H. S., Stull, A. T., & Canham, M. S. (2009). Naïve cartography: How intuitions about display configuration can hurt performance. In *Cartographica* (Vol. 44, Issue 3, pp. 171–186). https://doi.org/10.3138/carto.44.3.171

Hiemstra, J. A., Saaroni, H., & Amorim, J. H. (2017). The Urban Heat Island: Thermal Comfort and the Role of Urban Greening (pp. 7–19). https://doi.org/10.1007/978-3-319-50280-9_2

Hoffman, J. S., Shandas, V., & Pendleton, N. (2020). The Effects of Historical Housing Policies on Resident Exposure to Intra-Urban Heat: A Study of 108 US Urban Areas. *Climate*, *8*(1), 12. https://doi.org/10.3390/cli8010012

Howard, L. (1833). *The Climate of London* (Vol. 1).

Hu, L., & Wendel, J. (2019). Analysis of urban surface morphologic effects on diurnal thermal directional anisotropy. *ISPRS Journal of Photogrammetry and Remote Sensing*, *148*. https://doi.org/10.1016/j.isprsjprs.2018.12.004

Huang, Q., & Lu, Y. (2015). The Effect of Urban Heat Island on Climate Warming in the Yangtze River Delta Urban Agglomeration in China. *International Journal of Environmental Research and Public Health*, *12*(8). https://doi.org/10.3390/ijerph120808773

Imhof, E. (1982). Cartographic Relief Presentation. ESRI Press.

Imhoff, M. L., Zhang, P., Wolfe, R. E., & Bounoua, L. (2010). Remote sensing of the urban heat island effect across biomes in the continental USA. *Remote Sensing of Environment*, *114(*3). https://doi.org/10.1016/j.rse.2009.10.008

Irvankoski, K., Torniainen, J., & Krause, Ch. M. (2012, May). Visualisation of Elevation Information on Maps: An Eye Movement Study. *Conference Proceedings SWAET 2012, The Scandinavian Workshop on Applied Eye Tracking Karolinska Institute.* http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.716.4831&rep=rep1&type=pdf

Iungman, T., Cirach, M., Marando, F., Pereira Barboza, E., Khomenko, S., Masselot, P., Quijal-Zamorano, M., Mueller, N., Gasparrini, A., Urquiza, J., Heris, M., Thondoo, M., & Nieuwenhuijsen, M. (2023). Cooling cities through urban green infrastructure: a health impact assessment of European cities. *The Lancet*. https://doi.org/10.1016/s0140-6736(22)02585-5

Jenny, B., & Patterson, T. (2021). Aerial perspective for shaded relief. *Cartography and Geographic Information Science*, *48*(1), 21-28. https://doi.org/10.1080/15230406.2020.1813052

Kehrer, L. (1989). Central performance drop on perceptual segregation tasks, *Spatial Vision*, *4*(1), 45-62. doi: https://doi.org/10.1163/156856889X00040

Kennelly, P. (2017). Terrain Representation. *Geographic Information Science & Technology Body of Knowledge, 2017(Q4)*. https://doi.org/10.22224/gistbok/2017.4.9

Kennelly, P. J., & Kimerling, A. J. (2004). Hillshading of Terrain Using Layer Tints with Aspect-Variant Luminosity. *Cartography and Geographic Information Science, 31*(2), 67–77. https://doi.org/10.1559/1523040041649416

Klinenberg. (2003). Heat wave: a social autopsy of disaster in Chicago. *Choice Reviews Online, 40*(07), 40–4319. https://doi.org/10.5860/choice.40-4319

Kniaz, V. v., & Mizginov, V. A. (2018). Thermal texture generation and 3D model reconstruction using SFM and GAN. International Archives of the Photogrammetry, *Remote Sensing and Spatial Information Sciences - ISPRS Archives, 42*(2), 519–524. https://doi.org/10.5194/isprs-archives-XLII-2-519-2018

Krayenhoff, E. S., & Voogt, J. A. (2007). A microscale three-dimensional urban energy balance model for studying surface temperatures. *Boundary-Layer Meteorology, 123*(3). https://doi.org/10.1007/s10546-006-9153-6

Krehbiel, C., & Henebry, G. (2016). A Comparison of Multiple Datasets for Monitoring Thermal Time in Urban Areas over the U.S. Upper Midwest. *Remote Sensing, 8*(4). https://doi.org/10.3390/rs8040297

Kusaka, H., Kondo, H., Kikegawa, Y., & Kimura, F. (2001). A Simple Single-Layer Urban Canopy Model For Atmospheric Models: Comparison With Multi-Layer And Slab Models. *Boundary-Layer Meteorology, 101*(3). https://doi.org/10.1023/A:1019207923078

Lachman, J. L., Lachman, J. L., & Butterfield, C. (1979). *Cognitive Psychology and Information Processing: An Introduction*. Lawrence Erlbaum Associates.

Lagouarde, J.-P., Hénon, A., Kurz, B., Moreau, P., Irvine, M., Voogt, J., & Mestayer, P. (2010). Modelling daytime thermal infrared directional anisotropy over Toulouse city centre. *Remote Sensing of Environment, 114*(1). https://doi.org/10.1016/j.rse.2009.08.012

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00863

Laliberte, A. S., Goforth, M. A., Steele, C. M., & Rango, A. (2011). Multispectral Remote Sensing from Unmanned Aircraft: Image Processing Workflows and Applications for Rangeland Environments. *Remote Sensing, 3*(11). https://doi.org/10.3390/rs3112529

Lazzarini, M., Marpu, P. R., & Ghedira, H. (2013). Temperature-land cover interactions: The inversion of urban heat island phenomenon in desert city areas. *Remote Sensing of Environment, 130*. https://doi.org/10.1016/j.rse.2012.11.007

Lee, W. V. (2014). Historical global analysis of occurrences and human casualty of extreme temperature events (ETEs). *Natural hazards, 70*(2), 1453-1505. https://doi.org/10.1007/s11069-013-0884-7

Leonowicz, A. M., Jenny, B., & Hurni, L. (2009). Automatic generation of hypsometric layers for small-scale maps. *Computers and Geosciences, 35*(10), 2074–2083. https://doi.org/10.1016/j.cageo.2008.12.012

Li, J., Song, C., Cao, L., Zhu, F., Meng, X., & Wu, J. (2011). Impacts of landscape structure on surface urban heat islands: A case study of Shanghai, China. *Remote Sensing of Environment, 115*(12). https://doi.org/10.1016/j.rse.2011.07.008

Lim, T. C., Wilson, B., Grohs, J. R., & Pingel, T. J. (2022). Community-engaged heat resilience planning: Lessons from a youth smart city STEM program. *Landscape and Urban Planning, 226,* 104497. https://doi.org/10.1016/j.landurbplan.2022.104497

Lin, J. R., Cao, J., Zhang, J. P., van Treeck, C., & Frisch, J. (2019). Visualization of indoor thermal environment on mobile devices based on augmented reality and computational fluid dynamics. *Automation in Construction, 103*, 26–40. https://doi.org/10.1016/j.autcon.2019.02.007

Liu, Y. & Heer, J. (2018). Somewhere Over the Rainbow: An Empirical Assessment of Quantitative Colormaps. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3173574.3174172

Lowe, S. A. (2016). An energy and mortality impact assessment of the urban heat island in the US. *Environmental Impact Assessment Review, 56*, 139–144. https://doi.org/10.1016/j.eiar.2015.10.004

de Luis-Ruiz, J. M., Sedano-Cibrián, J., Pérez-Álvarez, R., Pereda-García, R., & Malagón-Picón, B. (2021). Metric contrast of thermal 3D models of large industrial facilities obtained by means of low-cost infrared sensors in UAV platforms. *International Journal of Remote Sensing, 43*(2), 457–483. https://doi.org/10.1080/01431161.2021.2003903

Lundgren, K., Kuklane, K., Gao, C., & Holmér, I. (2013). Effects of Heat Stress on Working Populations when Facing Climate Change. *Industrial Health, 51*(1), 3–15. https://doi.org/10.2486/indhealth.2012-0089

Masuda, T., & Nisbett, R. (2006). Culture and change blindness. *Cognitive Science*, 30(2), 381–399.

McGranaghan, M. (1989). Ordering Choropleth Map Symbols: The Effect of Background. *The American Cartographer, 16*(4). https://doi.org/10.1559/152304089783813918

Meerow, S., & Keith, L. (2021). Planning for Extreme Heat. *Journal of the American Planning Association, 88*(3), 319–334. https://doi.org/10.1080/01944363.2021.1977682

Mersey, J. E. (1990). *Colour and Thematic Map Design : the Role of Colour Scheme and Map Complexity in Choropleth Map Communication*. Toronto: University of Toronto Press.

Miyamoto, Y., Nisbett, R., & Masuda, T. (2006). Culture and the physical environment holistic versus analytic perceptual affordances. *Psychological Science*, *17*(2), 113–119.

Mikhailov, A. (2019, August 20). Turbo, An Improved Rainbow Colormap for Visualization. https://ai.googleblog.com/2019/08/turbo-improved-rainbow-colormap-for.html

Montello, D. R. (2002). Cognitive Map-Design Research in the Twentieth Century: Theoretical and Empirical Approaches. *Cartography and Geographic Information Science, 29(*3), 283–304. https://doi.org/10.1559/152304002782008503

Moreland, K. 2009. Diverging Color Maps for Scientific Visualization. *In Proceedings of the 5th International Symposium on Visual Computing*. https://doi.org.10.1007/978-3-642-10520-3_9

Moreland, K. 2016. Why We Use Bad Color Maps and What You can Do About it. *in Proc. IS&T Int'l. Symp. on Electronic Imaging: Human Vision and Electronic Imaging*, 2016, https://doi.org/10.2352/ISSN.2470-1173.2016.16.HVEI-133

Morrison, W., Kotthaus, S., Grimmond, C. S. B., Inagaki, A., Yin, T., Gastellu-Etchegorry, J. P., Kanda, M., & Merchant, C. J. (2018). A novel method to obtain three-dimensional urban surface temperature from ground-based thermography. *Remote Sensing of Environment, 215*, 268–283. https://doi.org/10.1016/j.rse.2018.05.004

Morrison, W., Kotthaus, S., & Grimmond, S. (2021). Urban surface temperature observations from ground-based thermography: intra- and inter-facet variability. *Urban Climate, 35*. https://doi.org/10.1016/j.uclim.2020.100748

Nagoor, Borgo, & Jones. (2017). Data Painter: A Tool for Colormap Interaction. *CGVC*, 69–76. https://doi.org/10.2312/cgvc.20171280

Najafifar, A., Hosseinzadeh, J., & Karamshahi, A. (2019). The Role of Hillshade, Aspect, and Toposhape in the Woodland Dieback of Arid and Semi-Arid Ecosystems: A Case Study in Zagros Woodlands of Ilam Province, Iran. *Journal of Landscape Ecology, 12*(2), 79–91. https://doi.org/10.2478/jlecol-2019-0011

Naughton, J., & McDonald, W. (2019). Evaluating the variability of urban land surface temperatures using drone observations. *Remote Sensing, 11*(14). https://doi.org/10.3390/rs11141722

Newcombe, N. S., Weisberg, S. M., Atit, K., Jacovina, M. E., Ormand, C. J., & Shipley, T. F. (2015). The lay of the land: Sensing and representing topography. *Baltic International Yearbook of Cognition, Logic and Communication, 10*(1), 6. http://dx.doi.org/10.4148/1944-3676.1099

Nichol, J. (2005). Remote sensing of urban heat islands by day and night. *Photogrammetric Engineering & Remote Sensing, 71*(5), 613-621. https://doi.org/10.14358/PERS.71.5.613

Nichol, J., & Wong, M. S. (2009). High Resolution Remote Sensing of Densely Urbanised Regions: a Case Study of Hong Kong. *Sensors, 9*(6). https://doi.org/10.3390/s90604695

Niethammer, U., James, M. R., Rothmund, S., Travelletti, J., & Joswig, M. (2012). UAV-based remote sensing of the Super-Sauze landslide: Evaluation and results. *Engineering Geology, 128.* https://doi.org/10.1016/j.enggeo.2011.03.012

Ooms, K., De Maeyer, P., Dupont, L., Van Der Veken, N., Van De Weghe, N., & Verplaetse, S. (2015). Education in cartography: what is the status of young people's map-reading skills? *Cartography and Geographic Information Science*, *43*(2), 134–153. https://doi.org/10.1080/15230406.2015.1021713

Onorati, G., Ventura, R., Poscolieri, M., Chiarini, V., & Crucillà, U. (1992). The Digital Elevation Model of Italy for geomorphology and structural geology. *CATENA*, *19*(2), 147–178. https://doi.org/10.1016/0341-8162(92)90022-4

Palsky, G. (1999). The debate on the standardization of statistical maps and diagrams (1857-1901). *Elements of the history of graphical semiotics*. Cybergeo. https://doi.org/10.4000/cybergeo.148

Patterson, T., & Jenny, B. (2014). Evaluating Cross-blended Hypsometric Tints: A User Study in the United States, Switzerland, and Germany. *Cartographic Perspectives, 75*, 5–17. https://doi.org/10.14714/cp75.578

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pingel, T., & Clarke, K. (2014). Perceptually Shaded Slope Maps for the Visualization of Digital Surface Models. *Cartographica: The International Journal for Geographic Information and Geovisualization, 49*(4), 225–240. https://doi.org/10.3138/carto.49.4.2141

Popelka, S., & Brychtova, A. (2013). Eye-tracking Study on Different Perception of 2D and 3D Terrain Visualisation. *The Cartographic Journal, 50*(3), 240–246. https://doi.org/10.1179/1743277413Y.0000000058

Previtali, M., Erba, S., Rosina, E., Redaelli, V., Scaioni, M., & Barazzetti, L. (2012). Generation of a GIS-based environment for infrared thermography analysis of buildings. *SPIE Proceedings*. https://doi.org/10.1117/12.930050

Rajashekar, U., Bovik, A. C., & Cormack, L. K. (2006). Visual search in noise: Revealing the influence of structural cues by gaze-contingent classification image analysis. *Journal of Vision, 6*(4), 7. https://doi.org/10.1167/6.4.7

Ratajski, L. (1974). Commission V of the ICA: The Tasks It Faces. *International Yearbook of Cartography, 14*, 140–144.

Reda, K. & Szafir, D. (2021). Rainbows Revisited: Modeling Effective Colormap Design for Graphical Inference. *IEEE Transactions on Visualization and Computer Graphics, 27*(2), https://doi.org/10.1109/TVCG.2020.3030439

Robinson, A. H. (1952). *The Look of Maps (Vol. 16).* University of Wisconsin Press. https://doi.org/10.2136/sssaj1952.03615995001600040024x

Rogowitz, B. E., Treinish, L. A., & Bryson, S. (1996). How Not to Lie with Visualization. *Computers in Physics, 10*(3), 268. https://doi.org/10.1063/1.4822401

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision, 7*(2), 17. https://doi.org/10.1167/7.2.17

Ryu, Y. H., & Baik, J. J. (2012). Quantitative analysis of factors contributing to urban heat island intensity. *Journal of Applied Meteorology and Climatology, 51*(5), 842–854. https://doi.org/10.1175/JAMC-D-11-098.1

Sailor, D. J., Baniassadi, A., O'Lenick, C. R., & Wilhelmi, O. V. (2019). The growing threat of heat disasters. *Environmental Research Letters, 14*(5), 054006. https://doi.org/10.1088/1748-9326/ab0bb9

Sarrat, C., Lemonsu, A., Masson, V., & Guedalia, D. (2006). Impact of urban heat island on regional atmospheric pollution. *Atmospheric Environment, 40*(10). https://doi.org/10.1016/j.atmosenv.2005.11.037

Scaife, M., & Rogers, Y. (1996). External cognition: how do graphical representations work? International *Journal of Human-Computer Studies, 45*(2). https://doi.org/10.1006/ijhc.1996.0048

Schloss, K. B., Gramazio, C. C., Silverman, A. T., Parker, M. L., & Wang, A. S. (2019). Mapping Color to Meaning in Colormap Data Visualizations. *IEEE Transactions on Visualization and Computer Graphics, 25*(1), 810–819. https://doi.org/10.1109/TVCG.2018.2865147

Shandas, V., Voelkel, J., Williams, J., & Hoffman, J. (2019). Integrating satellite and ground measurements for predicting locations of extreme urban heat. *Climate, 7*(1), 5. https://doi.org/10.3390/cli7010005

Shapiro, C., & Varian, H. R. (1998). *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business Review Press.

Shepard, R. N., & Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects. *Science, 171(*3972), 701–703. https://doi.org/10.1126/science.171.3972.701

Sheridan, S. C., & Allen, M. J. (2018). Temporal trends in human vulnerability to excessive heat. *Environmental research letters, 13*(4), 043001. https://doi.org/10.1088/1748-9326/aab214

Son, H., Kim, C., & Choi, H. (2019). High-quality as-is 3D thermal modeling in MEP systems using a deep convolutional network. *Advanced Engineering Informatics, 42*, 100999. https://doi.org/10.1016/j.aei.2019.100999

Song, B., & Park, K. (2020). Verification of accuracy of unmanned aerial vehicle (UAV) land surface temperature images using in-situ data. *Remote Sensing, 12*(2). https://doi.org/10.3390/rs12020288

Sousa, M. J., Moutinho, A., & Almeida, M. (2020). Thermal Infrared Sensing for Near Real-Time Data-Driven Fire Detection and Monitoring Systems. *Sensors, 20*(23), 6803. https://doi.org/10.3390/s20236803

Soux, A., Voogt, J. A., & Oke, T. R. (2004). A Model to Calculate what a Remote Sensor `Sees' of an Urban Surface. *Boundary-Layer Meteorology, 111*(1). https://doi.org/10.1023/B:BOUN.0000010995.62115.46

Steinke, T. R. (1987). Eye movement studies in cartography and related fields. *Cartographica: The International Journal for Geographic Information and Geovisualization, 24*(2), 40-73. https://doi.org/10.3138/J166-635U-7R56-X2L1

Sugawara, H., & Takamura, T. (2006). Longwave radiation flux from an urban canopy: Evaluation via measurements of directional radiometric temperature. *Remote Sensing of Environment, 104(*2). https://doi.org/10.1016/j.rse.2006.01.024

Taha, H. (1997). Urban climates and heat islands: albedo, evapotranspiration, and anthropogenic heat. *Energy and Buildings, 25*(2). https://doi.org/10.1016/S0378-7788(96)00999-1

Thomson, H., Simcock, N., Bouzarovski, S., & Petrova, S. (2019). Energy poverty and indoor cooling: An overlooked issue in Europe. *Energy and Buildings, 196*, 21–29. https://doi.org/10.1016/j.enbuild.2019.05.014

Thyng, K., Greene, C., Hetland, R., Zimmerle, H., & DiMarco, S. (2016). True Colors of Oceanography: Guidelines for Effective and Accurate Colormap Selection. *Oceanography, 29*(3), 9–13. https://doi.org/10.5670/oceanog.2016.66

Tobler, W. R. (1988). Resolution, resampling, and all that. In H. Mounsey and R. Tomlinson (eds.), *Building Database for Global Science*. London: Taylor & Francis, pp. 129-137.

Turner, M. G., & Gardner, R. H. (2015). *Landscape Ecology in Theory and Practice*. Springer New York. https://doi.org/10.1007/978-1-4939-2794-4

UN Environment Programme. (2021). *Beating the Heat: A Sustainable Cooling Handbook for Cities.*

United States Census Bureau. (2020). *QuickFacts Blacksburg town, Virginia; United States*. https://www.census.gov/quickfacts/fact/dashboard/blacksburgtownvirginia,US/POP010220

United States Environmental Protection Agency. (n.d.). *Measuring Heat Islands*. https://www.epa.gov/heatislands/measuring-heat-islands

U.S. Census Bureau (2020). *Population, Census, April 1, 2020*. Retrieved from https://www.census.gov/quickfacts/fact/table/atlantacitygeorgia,GA/PST045221

US EPA. (2022, March 15). *Heat Island Compendium.* https://www.epa.gov/heatislands/heat-island-compendium

Vasterling, M., & Meyer, U. (2013). Challenges and opportunities for UAV-borne thermal imaging. *In Remote Sensing and Digital Image Processing* (Vol. 17, pp. 69–92). Springer International Publishing. https://doi.org/10.1007/978-94-007-6639-6_4

Villa, T. F., Jayaratne, E. R., Gonzalez, L. F., & Morawska, L. (2017). Determination of the vertical profile of particle number concentration adjacent to a motorway using an unmanned aerial vehicle. *Environmental Pollution, 230.* https://doi.org/10.1016/j.envpol.2017.06.033

Voogt, J. A. (2008). Assessment of an Urban Sensor View Model for thermal anisotropy. *Remote Sensing of Environment, 112*(2). https://doi.org/10.1016/j.rse.2007.05.013

Ware, C. (1988). Color sequences for univariate maps: theory, experiments and principles. *IEEE Computer Graphics and Applications, 8*(5), 41–49. https://doi.org/10.1109/38.7760

WEIBEL, & R. (1991). Digital terrain modelling. *Geographical Information Systems; Principles and Applications*, 269–297. http://ci.nii.ac.jp/naid/10007659984/en/

Weng, Q. (2009). Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. *In ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 64, Issue 4, pp. 335–344). https://doi.org/10.1016/j.isprsjprs.2009.03.007

White, D. (1985). Relief Modulated Thematic Mapping By Computer. *The American Cartographer, 12*(1), 62–68.

Wilson, B. (2020). Urban Heat Management and the Legacy of Redlining. *Journal of the American Planning Association, 86*(4), 443–457. https://doi.org/10.1080/01944363.2020.1759127

World Bank. (2020, April). *Urban Development.* https://www.worldbank.org/en/topic/urbandevelopment

Yang, F., & Chen, L. (2016). Developing a thermal atlas for climate-responsive urban design based on empirical modeling and urban morphological analysis. *Energy and Buildings, 111*, 120–130. https://doi.org/10.1016/j.enbuild.2015.11.047

Yang, Y., & Lee, X. (2019). Four-band thermal mosaicking: A new method to process infrared thermal imagery of urban landscapes from UAV flights. *Remote Sensing, 11*(11). https://doi.org/10.3390/rs11111365

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396*(6706), 72–75. https://doi.org/10.1038/23936

Zeller, S., Samsel, F., Navratil, P. (2020). Environmental visualization: moving beyond the rainbows. *In Practice and Experience in Advanced Research Computing*, pp. 321-326. https://doi.org/10.1145/3311790.3396667

Zhang, Y., Fjeld, M., Fratarcangeli, M., Said, A., & Zhao, S. (2021). Affective Colormap Design for Accurate Visual Comprehension in Industrial Tomography. *Sensors, 21*(14), 4766. https://doi.org/10.3390/s21144766

Zhao, C., Jensen, J., Weng, Q., & Weaver, R. (2018). A geographically weighted regression analysis of the underlying factors related to the surface Urban Heat Island Phenomenon. *Remote Sensing, 10*(9). https://doi.org/10.3390/rs10091428

Zhou, L., & Hansen, C. D. (2016). A Survey of Colormaps in Visualization. *IEEE Transactions on Visualization and Computer Graphics, 22*(8), 2051–2069. https://doi.org/10.1109/tvcg.2015.2489649

Zhou, W., Huang, G., & Cadenasso, M. L. (2011). Does spatial configuration matter? Understanding the effects of land cover pattern on land surface temperature in urban landscapes. *Landscape and Urban Planning, 102*(1). https://doi.org/10.1016/j.landurbplan.2011.03.009

Zhou, D., Zhao, S., Zhang, L., & Liu, S. (2016). Remotely sensed assessment of urbanization effects on vegetation phenology in China's 32 major cities. *Remote Sensing of Environment, 176*. https://doi.org/10.1016/j.rse.2016.02.010

Zhu, B., & Chen, H. (2008). Information Visualization for Decision Support. *Handbook on Decision Support Systems 2*, 699–722. https://doi.org/10.1007/978-3-540-48716-6_32