Canopy Cover and Surface Temperature in Mystic Vale

Magdalen Thot

2023-09-28

1 Introduction

The objective of this report is to discover if there is a relationship between canopy cover and surface temperature, more specifically, the goal is to learn if this relationship exists in Mystic Vale on the University of Victoria campus. To answer this research question, groups of students developed sampling plans, then went into the field and collected temperature data around different sections of the study area. Learning if there is a relationship between canopy cover and surface temperature is significant because canopy cover can help mitigate the growing effects of climate change. Tree cover usually cools down the surrounding environment, so with temperatures rising, canopy cover can play an important role in mitigating the adverse effects of climate change on biodiversity (Aalto et al., 2022). The scope of the study looks at 30 locations around the study area where three temperature samples each were collected and subsequently analyzed.

2 Methods

2.1 Study Site

The study site for the report is Mystic Vale, a forested ravine located southeast of the University of Victoria campus outside of the ring road. Its canopy contains trees such as Douglas-fir, grand fir, western red cedar, conifers, bigleaf maple, black cottonwood, willow trees, and more. The study site covers 11.6 acres and is regularly used by students and members of surrounding communities for recreational activities. The study site contains a portion of a parking lot, a large forested area, and a portion of a grassy field; these three areas served as the stratified areas in our sampling plan.

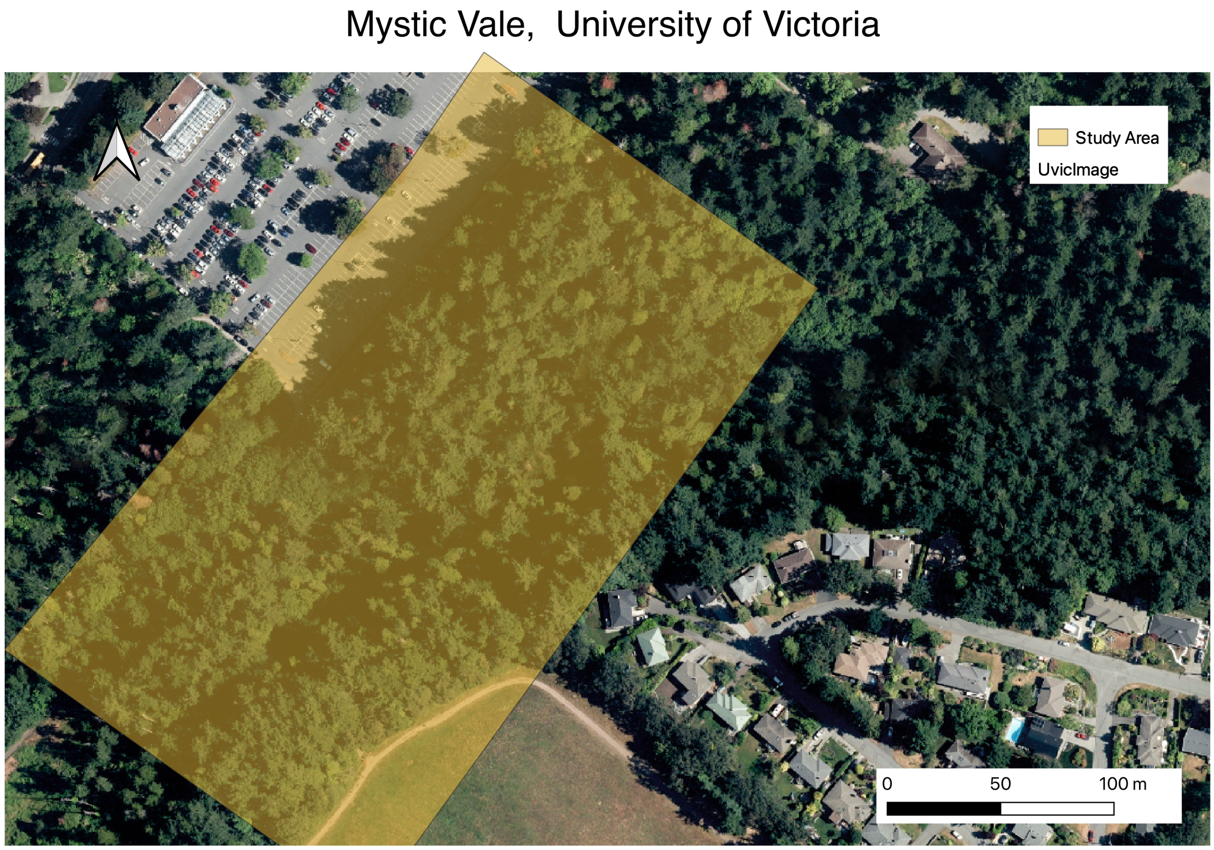

Figure 1: A map of the study site.

2.2 Data Collection

The data was collected by a team of four students on September 14th, 2023, from 2:30 p.m. to 4:20 p.m. PDT. The weather was about 22 degrees Celsius, and the sky was clear with little cloud cover.

For the study, we decided to use a stratified-proportional-random sampling plan. To develop the sampling plan, we did the following. First, for the stratified portion of the plan, we split the study area into a 15-by-15-meter grid and then divided the grid into three stratified areas: the parking lot, a forested area, and the flat grassy field. Next, the sample points were proportionally placed in the stratified regions. More specifically, we planned to take 25 percent of the samples from the parking lot, 25 percent from the field, and 50 percent from the forested area. The stratification helped ensure that each region was properly represented and helped reduce sampling errors, such as inconsistencies when collecting the data (McGrew and Monroe, 2000, Chapter 6).

Additionally, the proportional method was selected instead of the non-proportional due to the nature of the study area. We decided that because the grass and parking lot is relatively homogeneous, taking many samples in those regions was not essential; however, the forested area is heterogeneous, so sampling more in the area was necessary to understand the landscape.

Lastly, in the stratified sections points were randomly placed, this made the sampling probabilistic and increased precision. To conclude, our sampling plan met both statistical requirements and practical restraints. It provides proportional representation to all the regions in the study area, and the randomness helped avoid periodicities in the spatial pattern (McGrew and Monroe, 2000, Chapter 6).

Figure 2: This map displays how to study site was stratified. The yellow is the parking lot, the green is the forested area, and the pink is the field

2.3 Analysis

For this report, the surface temperature data was analyzed using the R programming language. Specifically, the **e1071**, **dplyr**, **sf**, **terra**, **tmap**, **ggplot2**, **knitr**, **terra** and **spatstat** packages were utilized. The first step in the analysis was to import the CSV temperature data. Next, utilizing the dplyr package, the max, mean, and standard deviation of the temperature samples were calculated using pipes in R.

```r
# An example of calculating the mean using pipes
csv <- csv %>%
  group_by(ID) %>%
  mutate(Mean = mean (c(Temperature1,
                       Temperature2,
                       Temperature3)))
```

Next, the data collected was made spatial using the Simple Features library, a collection of functions that allows you to create and manipulate spatial data objects in R (r-spatial, Simple features for R). Additionally, after the data is spatialized, a tif of the study area was imported using the terra library. Then, using the tif, maps showing the mean, max, and median temperatures at each point were made using the tmap library.

```r
# An exampke of how a map is made
MeanTempMap <- map +
  tm_shape(shp) +
  tm_dots(col = "Mean",
          shape = 21,
          border.col = "black",
          title = "Mean Temperature",
          size = 0.9,
          style = "jenks",
          palette = "Reds", n = 6) +
  tm_scale_bar(width = .2, text.size = 0.60) +
  tm_layout(frame = TRUE,
            bg.color = "#EFEFEF",
            fontfamily = "Helvetica",
            legend.title.size = 1.2,
            legend.text.size = 0.9,
            legend.width = 0.4,
            legend.title.fontface = "bold",
            inner.margins = c(0.02, 0.02, 0.02, 0.3),
            legend.position = c(0.72, "top"))
```

After creating the maps, it was important to describe the main features of the data by calculating descriptive statistics. In this section, the Mean, Median, Standard Deviation, Mode, Skewness, and Kurtosis were calculated using the median of all temperatures as an input. The median was chosen because the data was not normally distributed, so it was more of an appropriate choice.

```r
# Example of using the median to calculate standard deviation
sdTemp <-sd(shp$Median)
```

Subsequently, to display the data in a spatial format, multiple spatial descriptive statistics were calculated. In this section, it was not as easy as using an R function like mean( ) to derive an answer; it was necessary to use formulas and build the equations in R. For example, the weighted mean center was calculated using the following code:

```r
weightsFi <- shp$Temperature1 + shp$Temperature2 + shp$Temperature3
Xi <- shp$X
Yi <- shp$Y

XwcTop <- sum(weightsFi * Xi)
XwcFinal <- XwcTop/ sum(weightsFi)

YwcTop <- sum(weightsFi * Yi)
YwcFinal <- YwcTop/ sum(weightsFi)

WeightedMeanCentre <- data.frame(X = XwcFinal, Y = YwcFinal)
```

The same process was followed for the other spatial descriptive statistics: mean center, weighted standard distance, and relative distance. Lastly, the spatial descriptive statistics were placed on maps using the tmap library.

2.4 Results

This table displays the descriptive statistics mean, median, mode, skewness, kurtosis, standard deviation and coefficient of variation calculated for the report. Descriptive statistics help summarize the data for easier interpretation.(Table 1).

Table 1: Descriptive Statistics

Mean	Median	Mode	Standard_Deviation	CoV	Skewness	Kurtosis
22.17	19.35	18.4	6.65	1.098	1.59	0.95

This table displays the spatial descriptive statistics of standard distance, weighted standard distance, and relative distance. They represent the spatial equivalents of descriptive statistics mentioned above.(Table 2).

Table 2: Spatial Descriptive Statistics

Standard_Distance	Weighted_Standard_Distance	Relative_Distance
217.46	204.11	0.00058

This table displays the mean center and weighted mean center. The mean center and weighted mean center represent the "center of gravity" of the points sampled for the study. (Table 3).

Table 3: Mean and Weighted Mean Center

Mean_Center.X	Mean_Center.Y	Weighted_Center.X	Weighted_Center.Y
477344.2	5367551	477343.6	5367560

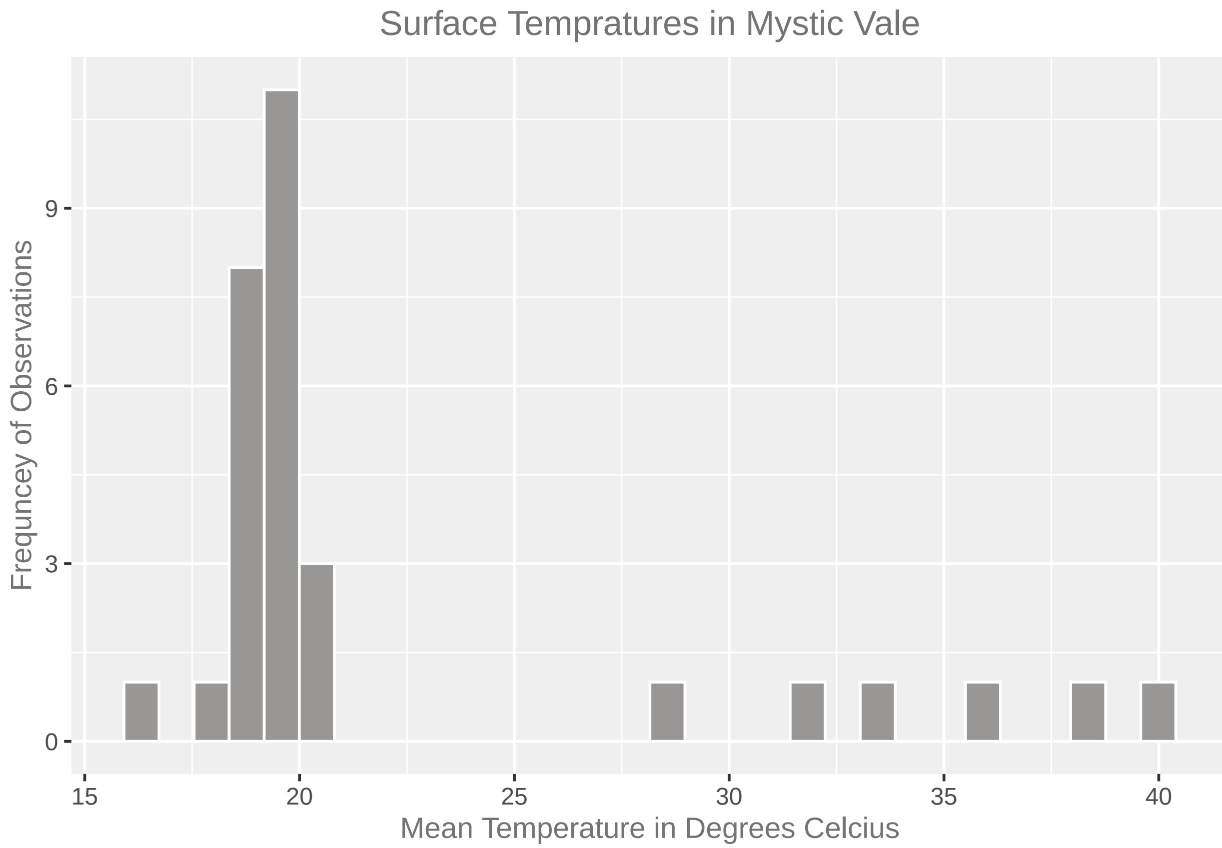

Figure 3: This histogram displays the mean temperatures around the study area, with mean temperature in degrees celsius on the x-axis and the frequency of observations on the y-axis.

Figure 4: This map displays the maximum temperatures at each of the 30 sampled locations in the study area. It uses a jenks classification method with six classes

Figure 5: This map shows the mean temperatures at each of the 30 sampled locations in the study area. It utilizes a Jenks classification method with six classes.

Figure 6: This map shows the median temperatures at each of the 30 sampled locations in the study area.It utilizes a Jenks classification method with six classes.

Figure 7: This map shows the mean center and mean weighted center at each of the 30 sampled locations in the study area. The weights used for the weighted mean center are the sum of temperatures at each sampled location.

Figure 8: This map displays the 30 sampled locations, mean center, and standard distance, which measures the amount of absolute dispersion in a point pattern.

3 Summary

The results suggest that there is a relationship between surface temperatures and tree canopy. Looking at the maps of mean, median, and max surface temperatures at each location, it is clear that the areas with plenty of tree canopy have the lowest temperatures, and the areas with no tree cover, for example, the parking lot have the highest temperatures. Furthermore, when considering the descriptive statistics calculated in the study, it is evident that the mean is greater than the median; this indicates that the data is positively skewed, with outliers on the right side of the distribution (McGrew and Monroe, 2000, Chapter 3). In addition, the mode value, which is less than the mean and the median, and the positive skewness value imply that the majority of data points are concentrated on the left side of the distribution, with lower temperatures (Figure 9). Consequently, the kurtosis value suggests that the data distribution is platykurtic, indicating that it has fewer extreme values (McGrew and Monroe, 2000, Chapter 3). Lastly, the coefficient of variation value and standard deviation indicate that the data has high relative variability (McGrew and Monroe, 2000, Chapter 3).

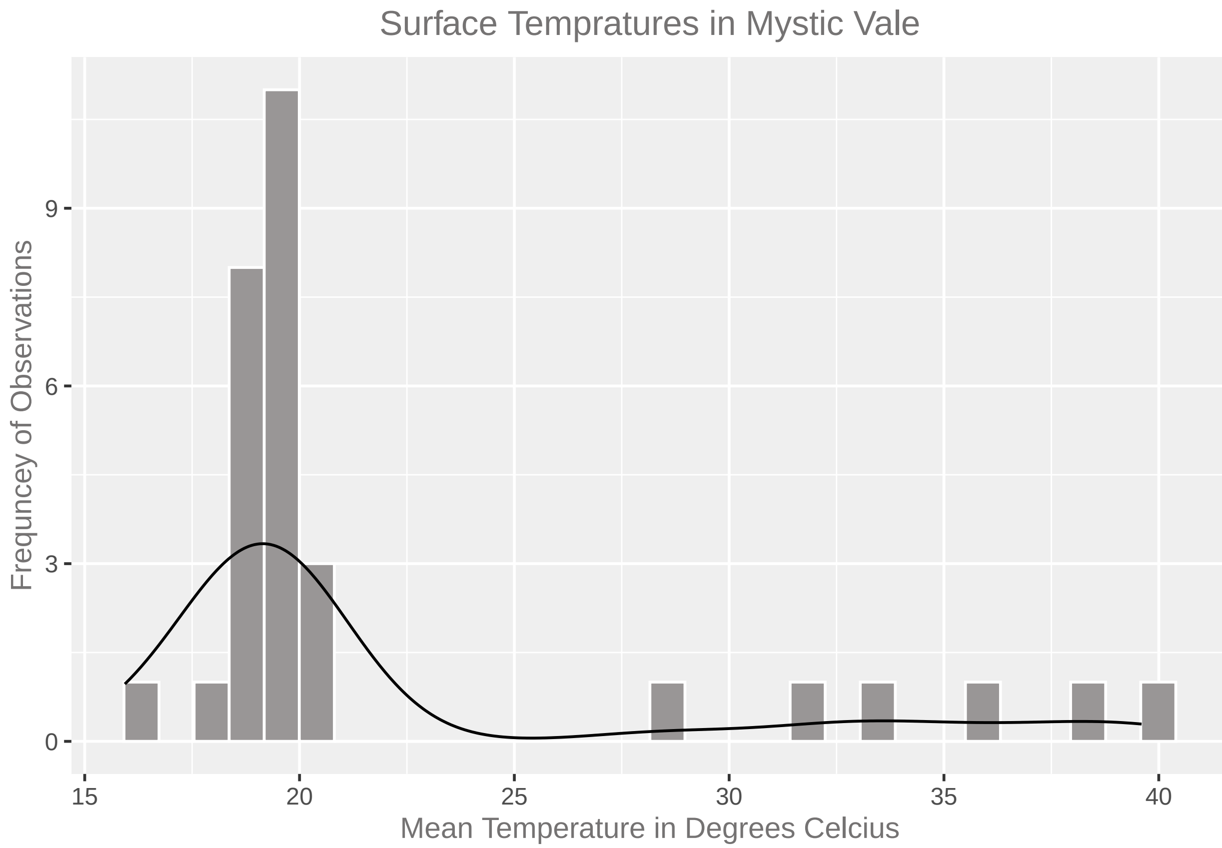

Figure 9: This histogram displays the mean temperatures around the study area, with mean temperature in degrees celsius on the x-axis and the frequency of observations on the y-axis, it also has a curve to highlight the shape of the distribution

Moreover, the standard distance further supports the descriptive statistics because it is a higher value, and implies that the points are more dispersed (McGrew and Monroe, 2000, Chapter 4). In conclusion, with values skewed toward the lower end of temperatures, the results of the study suggests that a relationship between surface temperatures and tree canopy cover exists in Mystic Vale.

4 References

Aalto, I. J., Maeda, E. E., Heiskanen, J., Aalto, E. K., & Pellikka, P. K. (2022). Strong influence of trees outside forest in regulating microclimate of intensively modified Afromontane landscapes. Biogeosciences, 19(17), 4227–4247. https://doi.org/10.5194/bg-19-4227-2022

Chapter 3 - Descriptive Statistics from McGrew and Monroe (2000). An Introduction to Statistical Problem Solving in Geography, McGraw Hill; Boston.

Chapter 4 - Descriptive Spatial Statistics from McGrew and Monroe (2000). An Introduction to Statistical Problem Solving in Geography, McGraw Hill; Boston.

Chapter 6 - Basic Elements of Spatial Sampling from McGrew and Monroe (2000). An Introduction to Statistical Problem Solving in Geography, McGraw Hill; Boston.

Simple features for R. • sf. (n.d.). https://r-spatial.github.io/sf/