Article

A seasonality and a zoning model for Puerto Rico extracted from ten years of MODIS data

Israel O Dilán-Pantojas [1,‡], Marcelo Francia [2,‡], and Manuel J. La Torre-Poueymirou [3,‡]

1 University of Puerto Rico, Rio Piedras Campus; israe.dilan@upr.edu
2 University of Puerto Rico, Rio Piedras Campus; marcelo.francia@upr.edu
3 University of Puerto Rico, Rio Piedras Campus; manuel.latorre@upr.edu
* Correspondence: israel.dilan@upr.edu; Tel.: +1-787-241-4148
† Current address: 14 Ave. Universidad Ste. 1401, San Juan, PR 00925-2534
‡ These authors contributed equally to this work.

Version May 15, 2019 submitted to Remote Sens.

Abstract: This work attempts to create a zoning and seasonality model of Puerto Rico by utilizing 10 years of MODIS data. The motivation for this work comes from the necessity to mitigate the potential damage that climate change and other atmospheric disturbances can effect on the island. Through spatial analysis of distinct variables using exploratory analysis and unsupervised clustering methods, the markedly distinct grouping of zones across the island was observed. A temporal analysis consisting of an exploratory analysis, statistical pattern analysis and trend analysis of the average value of the same variables over time, yielded evidence of abrupt breaks and some changes in the seasonal patterns of the Puerto Rico region, along with a model that proposed an increasing trend in LST and a slight reduction in the EVI trend. Interactions between these variables may prove useful for the construction of preventative strategies to protect vulnerable regions.

Keywords: MODIS; Enhanced Vegetation Index; Land Surface Temperature; Unsupervised Clustering; Autoregressive Integrated Moving Average model; Seasonality; Zoning; Puerto Rico

1. Introduction

Over the last decade, the world has experienced immense climatological changes with the continuous growth of industrialization, and human population. Humanity has been conscient, or at least warned, of the world's environmental changes along with the divulgence of the idea of what global warming is, through social media and other means communication. Today experiencing longer winters, longer drought seasons as well as discerning the increasing number of major storm formations all around the globe, consciousness on this matter has become relatively common to the average individual. The abnormal climatological events Puerto Rico has experienced in the last decade are an example of our environment's general reaction to the adversities it has faced since the accumulation of human waste as a result of the industrial growth along with human population increase. In the last 10 years, Puerto Rico has undergone persisting drought seasons, forest fires, as well as experience re-occurring catastrophic landfall events, like Hurricane Maria and Irma in 2017. Puerto Rico has also experienced a redistribution of age groups with the younger populations migratory patterns out of the island, which has left Puerto Rico with an increasingly elderly population. (*)

Remote sensing databases are currently being used for characterization of ecosystem structure, diversity, and function [27]. Due to the spatial and temporal dimensions that these databases contain, they have facilitated the study of ecological spatial distributions [11] and seasonal trends [2] of vegetation variables across land regions. Integrating these types of data sets into land models enables the monitoring and prediction of biomass production [17], study vegetation response to growth during climatological disturbing events [20], as well as study climate and land change [5]. The MODIS instrument in the NASA Terra spacecraft expresses features of the land that are commonly used for wildfire monitoring, global temperature changes and ecosystem dynamics [7]. According to Wulder, MODIS enables regular monitoring of global vegetation productivity, but not a spatially detailed characterization of vegetation. Thus, the MODIS data can be used to describe and study general trends of vegetation state and global temperature among other characteristics of our biome. With these tools and data sets, available to the world's general public, through NASA's AppEEARS portal [4,18,25], we will express a 10-year zoning and seasonality model of the archipelago of Puerto Rico. 10 years' worth of data were selected due to data set availability from 2008 to 2018 in NASA's AppEEARS portal database [4,18,25]. No evidence presenting a zoning and seasonality model based on remote sensing data has been found in the existing literature. With this model, we hope to spread awareness of Puerto Rico's previous, current and future ecological and climatological status to the numerous governmental agencies, with the purpose of contributing to the optimization of preparedness of forest preservation and emergency management teams when dealing with natural catastrophes.

Seasonality is a distinctive behavioral characteristic of a time series data summary, where a graph may present predictable change when recurring over time. Zoning, containing numerous forms of definition; in our study zoning, is defined as the classification of land area by characteristics pertaining to the remote sensing variables being used in this study. This zoning and seasonality model will be generated by analyzing the Enhanced vegetation index (EVI) and land surface temperature (LST) data sets of a plotted area containing Puerto Rico's archipelago.

LST provides information regarding land surface temperature in Kelvin, that we then converted to Celsius. Data sets for LST are generated by the averaging 8-day per-pixel 1000m resolution scans of land through NASA's Terra MODIS spacecraft, MOD11A2 [14]. NDVI provides consistent spatial and temporal comparisons of vegetation canopy greenness, a composite property of leaf area, and canopy structure [4,15]. The Terra spacecraft generates vegetation indices 16-day composite eight days apart, and are retrieved from daily, atmosphere-corrected, bidirectional surface reflectance scan by the NASA's Terra MODIS spacecraft, MOD13Q1 [4]. Clouded pixels are removed from the produced NDVI and EVI data packages with the purpose of eliminating low quality and absent data recordings with the purpose of eliminating non-significant memory space [15]. NDVI has previously been used to assess the vegetation response to episodic disturbance events [20]. Considering possible relationships between LST and NDVI, both have been previously used to quantify live fuel moisture content in forests to estimate fire danger rating [3]. Additionally, both are currently used for agricultural drought monitoring and food security status assessment in various countries [8]. Both variables, have the potential to be used for projecting ecological and climate risk factor situations, with the purpose of ameliorating preparedness when facing natural anomalies in the archipelago of Puerto Rico.

2. Materials and Methods

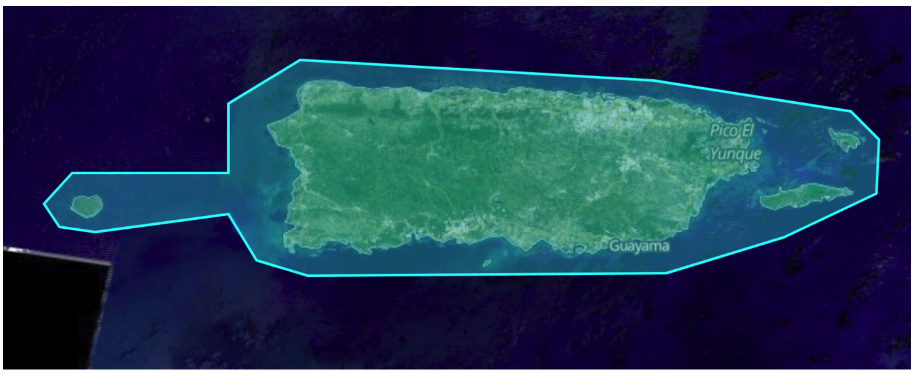

Figure 1. The polygon utilized to extract data for the region of interest, from AppEEARS.

In order to extract patterns and create models which may be used to optimize Puerto Rico's public policy, a significant amount of relevant data is necessary. This research focuses on utilizing some of the vast amounts of freely available, high-quality data obtained through the MODIS instrument cluster from the Terra satellite. These instruments record moderate resolution images of our planet's surface light reflectance captured by a spectroradiometer, from which different information variables are extracted. This study centers on the Enhanced Vegetation Indices (EVI) and Land Surface Temperature (LST) variables as recorded by MODIS, through a ten-year time-period from January 1^{st}, 2018 through December 31^{st}, 2018. This time period was chosen because it contained an appropriate window of observations within the constraints of available data for both variables.

EVI was used instead of NDVI due to enhanced sensitivity for high biomass regions and reduced atmospheric influence [10] on measurements, that EVI presents over traditional NDVI. EVI is a composite of near-infrared, red and blue wavelength measures, used to calculate chlorophyll density contained in existing vegetative cover[9]. EVI is commonly used for monitoring vegetation condition in various parts of the world[21]. We used the MOD13Q1 product, which provides EVI data every 16 days at a 250-meter spatial resolution. Measurements were scaled by the 0.001 factor required for appropriate use of the data, as appointed by NASA's AppEEARS portal(*). The LST data was obtained from the MOD11A2 product. LST is a measure of the average surface temperature over an 8 days time window, expressed as Kelvin units, within a 1-kilometer spatial resolution. This data is specific to the temperature of the land surface, not taking into consideration air temperature when measurements are made. Measurements were scaled by the 0.02 factor required for appropriate use of the data, as appointed by NASA's AppEEARS portal(*). Kelvin where converted to degrees Celsius by subtracting 273.15, for our analysis.

Our Zoning model consisted of an exploratory data analysis and an unsupervised clustering approach in order to extract information about trends in the spatial data.The exploratory analysis consisted of observing the mean pixel value over time and frequency analysis of pixel value for the whole plot. For the clustering of zones across the Puerto Rico archipelago, unsupervised clustering was performed grouping the zones from 2 to 11 clusters, scoring a stratified sample of 2,000 pixels within each zone, and then selecting the cluster with the best overall score. This method was used to obtain the best spatial representation of different possible clusters. This unsupervised clustering method was chosen because it does not require training samples, and it has been previously used for the classification of remotely sensed images[1].Two algorithms were used, K-means (KM) and Clustering LARge Application (CLARA), ******which would output clusters where within-group distances are minimized and between-group separation is maximized.******The K-means algorithm does this through defining the centroid of each cluster with its mean value, assigning the data to the cluster with the nearest mean. The CLARA algorithm uses medoids, which are the most centrally located point in the cluster, from different randomly selected data-sets and comparing them to the original data, looking for the arrangement with less dissimilarity. Each approach was compared using the silhouette index to obtain the best number of clusters. The silhouette index value for an object indicates how well it is matched to its own cluster and how poorly it relates to other neighboring clusters.

The Seasonality analysis was performed on the average value for the whole Puerto Rico region on each time-point and it consisted on three phases; an initial exploratory analysis, an examination of abrupt changes on seasonal trend patterns over time and the creation of a model/forecast with the information obtained from the downloaded data products. The exploratory analysis on time-series data consisted on fitting a line through linear regression and a curve in order to express clearer observable trends and patterns contained in the data

One of these more granular analysis is the Breaks For Additive Season and Trend (BFAST) approach to detect tendencies and seasonal changes in our time series data [24]. This method iteratively estimates the number and the time of drastic changes within a time series, allowing the characterization of the change by its magnitude and direction. Finally, for model/forecast preparation the Auto-Regressive Integrated Moving Average(ARIMA) algorithm was used. This model is utilized in order to detect tendencies and seasonal changes in our averaged time series data. This kind of model is applied to non-stationary data and it uses past values in the regression equation for the time series in order to inform automatic model creation. The size of our 240 data points for EVI and 480 data points for LST of time series data exploits the ARIMA's high reliance on previous data points. This model creation approach has already been used to forecast NDVI in coniferous areas [6] and for drought monitoring [8].

All analysis where performed utilizing R-Statistical Software and the code for this project is found in Appendix D.Code.

3. Results

3.1. Zoning analysis

Results for the zoning analysis are divided into two segments, one is based on an exploratory analysis of the data and the other part is based on the unsupervised clustering approach. The first, exploratory analysis refers to the EVI in the spatial dimension, shown in Figure 2.A. Figure 2.B presents the average amount of vegetation on the island is of substantial density as observed around the 0.51 EVI frequency.

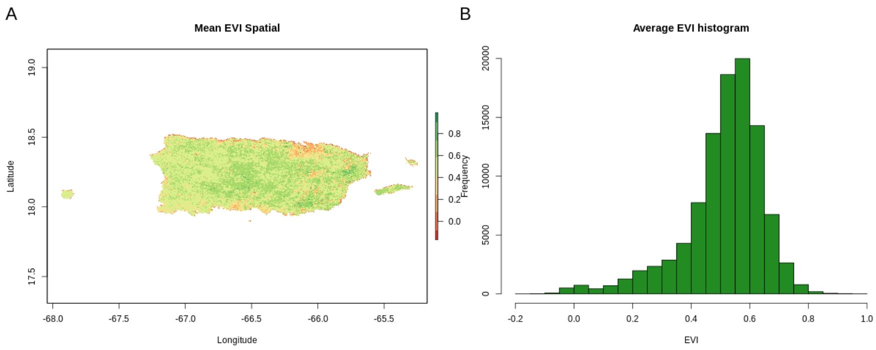

Figure 2. 2.A Spatial representation average EVI over 10 years. 2.B EVI frequency centers around 0.51 value.

Similarly, Figure 3.A shows 2 significantly distinct spatial regions between high-temperature areas and low-temperature areas. Figure 3.B shows the average frequency of LST in the whole archipelago, which over the last decade ranges from 24° to 26° degrees Celsius.

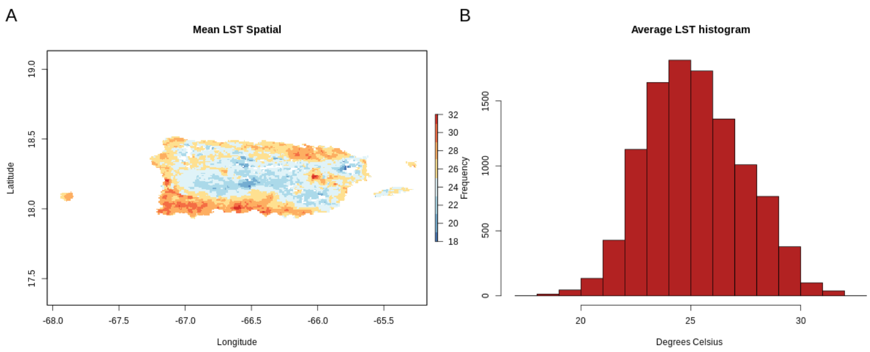

Figure 3. 3.A Spatial representation average LST over 10 years. 3.B LST frequency centers around 25.19° Celsius value.

After performing the clustering analysis figures for all clustering runs presented in supplementary material on Appendix A and evaluating the results corresponding figures and tables in Appendix B, the highest scoring clusters are presented in Figure 5 and Figure 6.

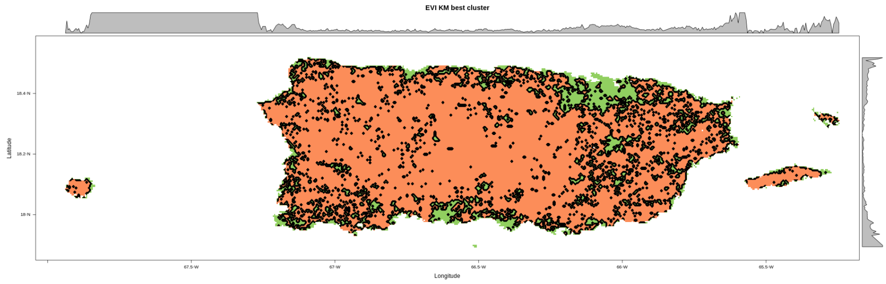

Figure 4. Best cluster number distribution according to the silhouette index for enhanced vegetation index across the Puerto Rico archipelago for the whole time series. Produced by the K-means algorithm

Figure 4 depicts the cluster classification with the highest silhouette index for EVI, achieved by the k-means algorithm. Two main clustered regions where identified, one in orange representing areas with low EVI, and a green region representing areas with high EVI. White represents areas without EVI cluster classification.

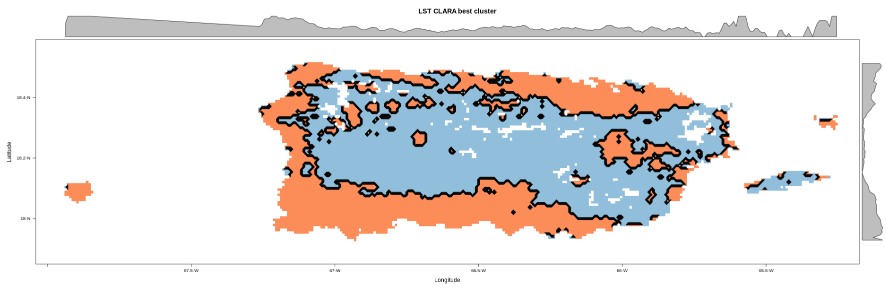

Figure 5. Best cluster number distribution according to the silhouette index for land surface temperature across the Puerto Rico archipelago for the whole time series. Produced by the CLARA algorithm

In figure 5 we present the cluster classification with the highest silhouette index for LST, obtained with the CLARA algorithm. Two main clustered regions where identified, blue zones cover areas with low LST, and orange zones cover areas with high LST. White represents areas without LST cluster classification.

3.2. Seasonality analysis

Seasonality analysis results are divided into three segments, the first segment is based on the exploratory analysis of the data focused on observing the yearly trends, the second segment contains the results for the BFAST analysis of abrupt pattern deviations, and the third segment that is based on the utilization of the ARIMA approach to presenting a model and a forecast for seasonality behavior of EVI and LST in the next two years, namely 2019 and 2020.

The linear models presented, suggest a decrease in average EVI value Figure 6.A and an increase in average LST value Figure 6.B, while the curves suggest a seasonal and periodic repeat of patterns for both variables. It is important to note the variation in the amplitude and frequency of the EVI curve in the periods late 2009, late 2015, and late 2016 when compared to the rest of the curve in Figure 6.A, similarly it is important to observe the variation in the amplitude and frequency the LST curve in the periods 2008-2019 and late 2015 in Figure 6.B. The ARIMA model analysis are presented in Appendix C.

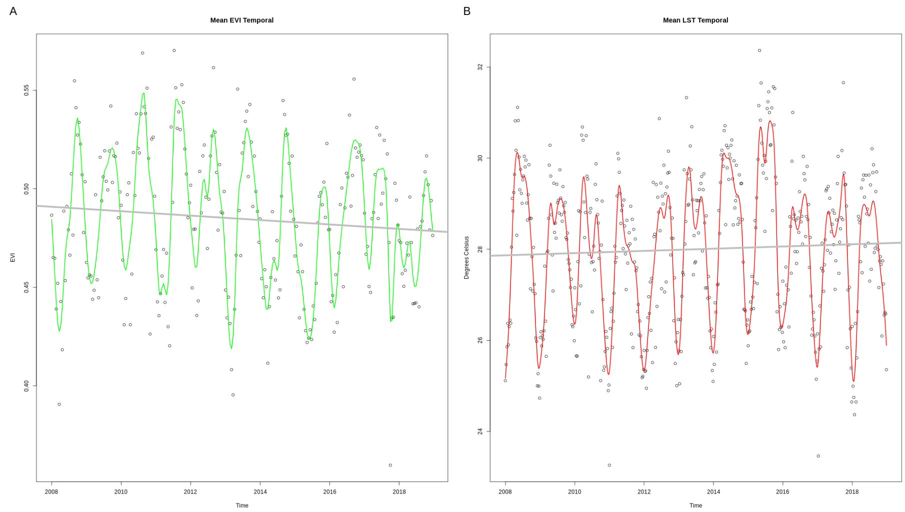

Figure 6. Lineal trends across the whole time series for EVI(A) and LST(B)

Closer examination and comparison of yearly behavior show deviations from yearly trends in years 2010, 2011, 2015, 2017, and 2018, further analysis was required in order to determine if this changes where significant when compared to the rest of the trend Figure 7.A. Similarly, deviations from yearly trends in years 2008, 2010, 2011, 2014, and 2015, further analysis was required in order to determine if this changes where significant when compared to the rest of the trend Figure 7.B.

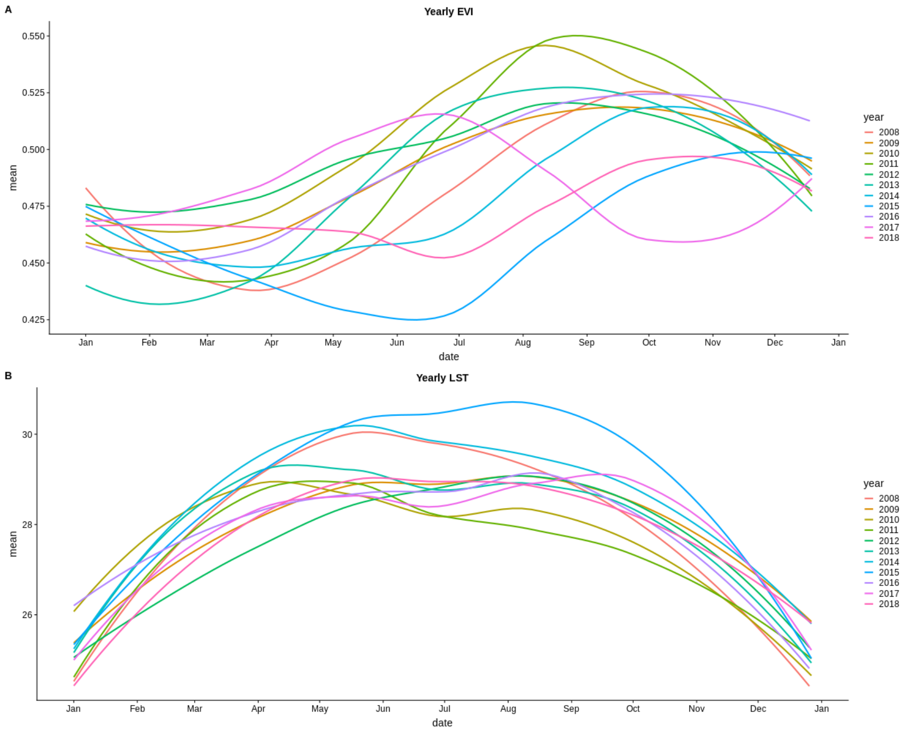

Figure 7. Yearly average change for EVI(A) and LST(B) for each year of the whole time series

In order to observe the significance of the changes, the BFAST analysis was performed. The results are contained in figure 8 for EVI trends and Figure 9 for LST trends. It is shown in the trend analysis, Figure 8.Tt, that significant changes in EVI trend occurred during the years 2015 and late 2017. Similarly, the trend analysis, Figure 9.Tt, shows significant deviations from the trend of LST average values that occurred during the years 2010 and late 2015.

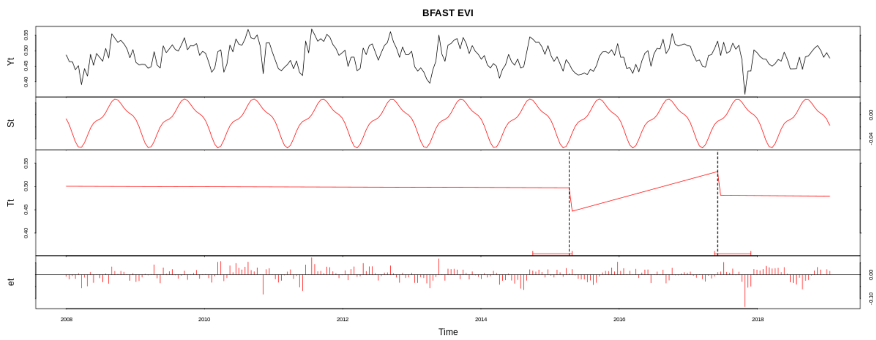

Figure 8. BFAST analysis results on EVI.

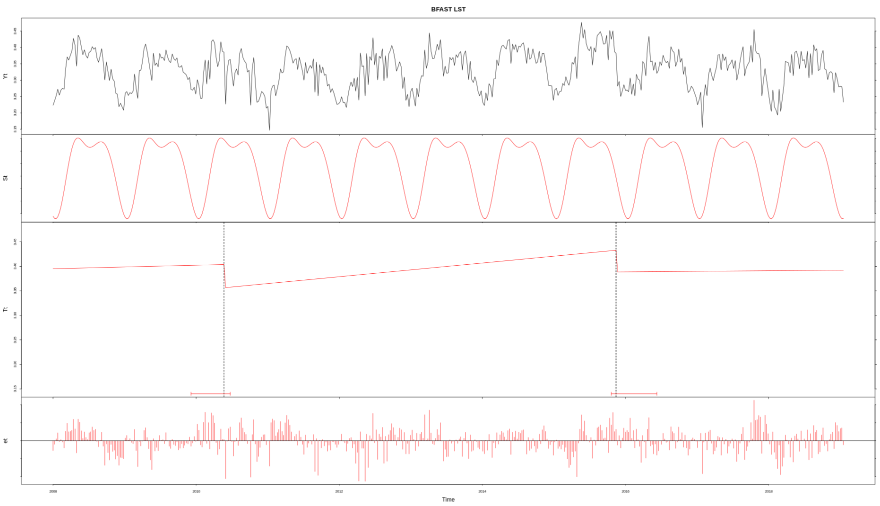

Figure 9. BFAST analysis results on LST.

Model creation was automatized with the utilization of the ARIMA algorithm, this model was then used to create a forecast. The forecast results are presented in Figures 10 for the EVI model and Figure 11 for the LST model, suggesting a possible behavior of the trends for the next year in the colored regions of the graph.

In Figure 10. EVI suggests a trend for the next year with an 80% confidence in darker purple regions. However, a 90% confidence, for the behavior of the trend is suggested with lighter purple regions on the graph. These purple regions show the possible directions the EVI trend might take in the next year. These projections are extracted from our accounted data considering 10 years of previous EVI in the archipelago of Puerto Rico that had been used for our exploratory analysis.

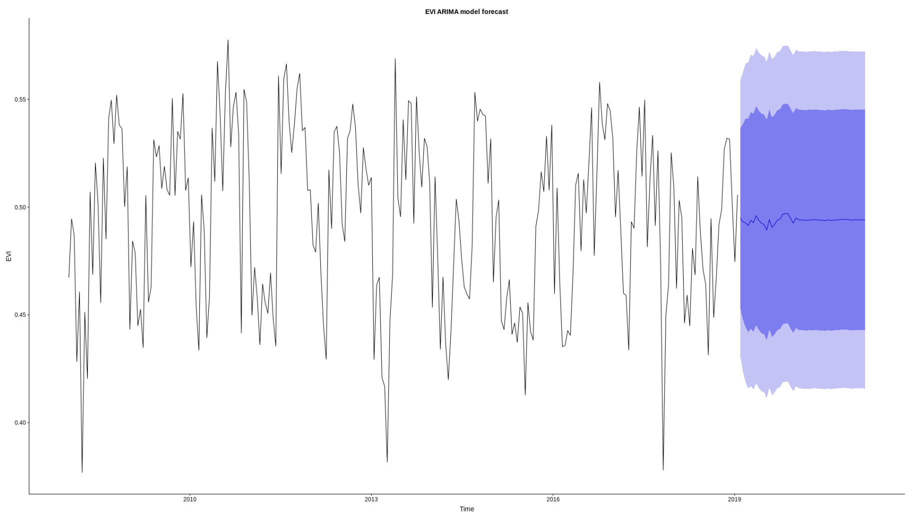

Figure 10. ARIMA EVI seasonal model shows the observed behavior of average EVI (Black) and a forecast (Blue) for the following two years.

LST forecast for years 2019 and 2020 results contained in Figure 11, closely follow the observed pattern of previous years, it is worth noting that the expected average LST peaks and valleys both at the 80% and 90% confidence levels seem to accommodate peaks and valleys close to the highest and lowest observed in previous years.

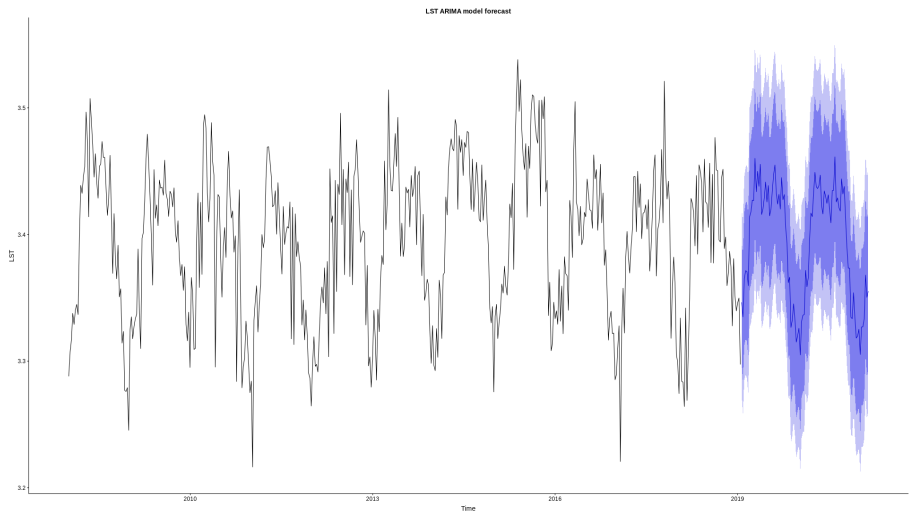

Figure 11. ARIMA LST seasonal model shows the observed behavior of average LST (Black) and a forecast (Blue) for the following two years.

4. Discussion

The zoning model was created to asses the adequate number of zones which should be considered by the Puerto Rican government when executing public policy. The exploratory analysis of the EVI in figure 2.A showed us vegetation rich areas in the interior of the island, which is mainly composed of forest and herbaceous pasture or agriculture[26]. It also suggests a grouping of lower vegetation areas such as the ones surrounding the coasts, and some in densely populated areas such as the San Juan, Caguas, and Ponce areas. Figure 2.B illustrates the frequency of all pixel values of EVI across the Puerto Rico archipelago. This also shows that there is a larger amount of areas with high EVI when compared with areas of low EVI.

The exploratory analysis of LST in Figure 3.A shows two significantly distinct spatial regions between high-temperature areas and low-temperature areas. This difference in zones happens to be very similar to the grouping observed when exploring the EVI data. Coastal LST is on average higher than that for the center of the island, except for densely populated areas such as the San Juan, Caguas, and Ponce areas. However, the smaller the patch of land of the composing archipelago, the more density for higher LST may be present. It is important to observe the overlap between a region around 17.9° to 18.2° North and 67.3° to 66.3° West of moderately high EVI values near 0.4 and high LST nearing the 30° to 32° Celsius, this region seems more prone to spontaneous wildfires.

From this exploratory observations, we can infer the presence of very distinct zones which might share common characteristics that could play a major factor on the general ecological health on the island, when it comes to the effects of climate change as experienced through seemingly abnormal seasonal patterns.

Figure 4 depicts the cluster classification for the EVI. We can mainly identify two regions, a large area with high EVI and areas with low EVI, located either in highly urbanized regions[12] or in the shore. This kind of cluster classification does not capture the gradient of vegetation that other studies have shown[26], but this is because the MODIS does not provide enough definition for vegetation classification[27]. It does, however, illustrate the general trend of vegetation distribution between high urbanized and low urbanized areas.

For the cluster classification with the highest silhouette index for LST in Figure 5, we observe that highly urbanized zones and their surrounding areas are classified within the high LST cluster. Other areas, which coincide with areas with high EVI in figure 4, had lower LST.

In the seasonality analysis for the whole time series(Figure 17) we can observe a decreasing trend for EVI(Figure 17.A), and an increasing trend for LST (Figure 17.B) in the Puerto Rico Archipelago. This constant lowering of EVI is consistent with the increasing urbanization and deforestation in

Puerto Rico. The increase in LST could be attributed to the current rise of temperature that the earth is experiencing.

When comparing the different variables across the 10 years(Figure 18), we observe for EVI(Figure 18A) a tendency to rise during the months of April to August, although this trend was not observed during the year 2017. In 2017 we can see a decrease in EVI that started by July and extended the month of September, by the time that hurricane Irma and Maria passed the Puerto Rico archipelago. A lower EVI implies lower photosynthetic activity, which would also be reflected by a lower exchange of plant water with the atmosphere. Examining the annual trends of LST(Figure 18.B) we can observe a bimodal distribution, where there is a small decrease followed by an increase in temperature between the months of June and July.

These results are in no way conclusive and therefore require further and more granular analysis.

Author Contributions: conceptualization was brought about by all authors; methodology was devised in discussion with our professors as acknowledged; software utilized was R-Statistical Software, managed by Dilán-Pantojas; validation was observed by all authors; formal analysis was performed by all authors; investigation, X.X.; resources, X.X.; data curationmanaged by Dilán-Pantojas and La Torre-Poueymirou; the writing—original draft preparation, was performed by all authors; writing—review and editing, was performed by all authors; visualization, managed by Dilán-Pantojas; supervision, was carried by Dilán-Pantojas; project administration, was performed by all authors;

Funding: This research received no external funding.

Acknowledgments: *We would like to acknowledge the contributions to this study made by our professors Ph.D. Jose E. Garcias Arraras & Ph.D. Carla Restrepo. We would also like to thank our individual mentors Ph.D. Humberto Ortiz-Zuazaga, M.S. Eileen Poueymirou-Yunque, Ph.D. Jose Luis Agosto, for all the mentorship and guidance they've provided us. We would also like to thank Jose A. Reyez-Zayas for voluntarily donating some of his time in proofreading our article and script. Finally, we would like thank and acknowledge our fellow classmates in this course who where very supportive and embarked on a similar journey to ours' in producing this creative work.*

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MDPI	Multidisciplinary Digital Publishing Institute
DOAJ	Directory of open access journals
AppEEARS	Application for Extracting and Exploring Analysis Ready Samples
NASA	National Aeronautics and Space Administration
LP DAAC	Land Processes Distributed Active Archive Center
MODIS	Moderate Resolution Imaging Spectroradiometer
NDVI	Normalized difference vegetation index
EVI	Enhanced vegetation index
LST	Land surface temperature
KM	K-means algorith
CLARA	Cluster LARge Applications
BFAST	Breaks For Additive Season and Trend
ARIMA	Auto-Regressive Integrated Moving Average

Appendix A. Clustering

K-means clustering of EVI was performed generating 2 to 11 clusters. These cluster images would be genarated randomle for each pixel contained in our EVI dataset. Thus, randomizing our EVI image

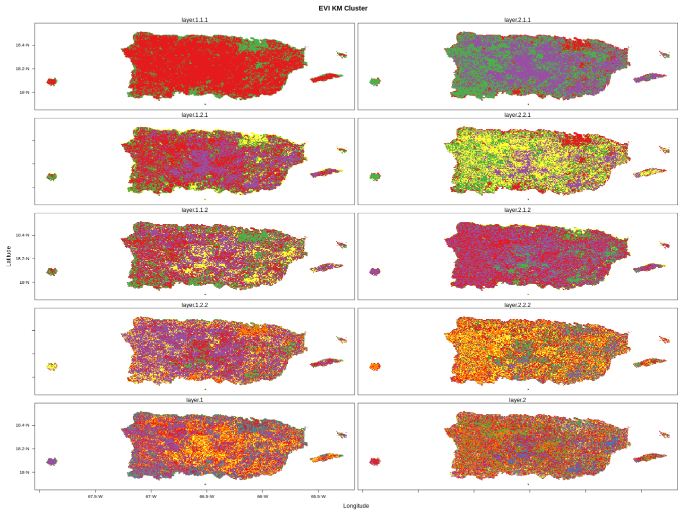

Figure A1. Results for K-means clustering of EVI from 2 to 11 clusters.

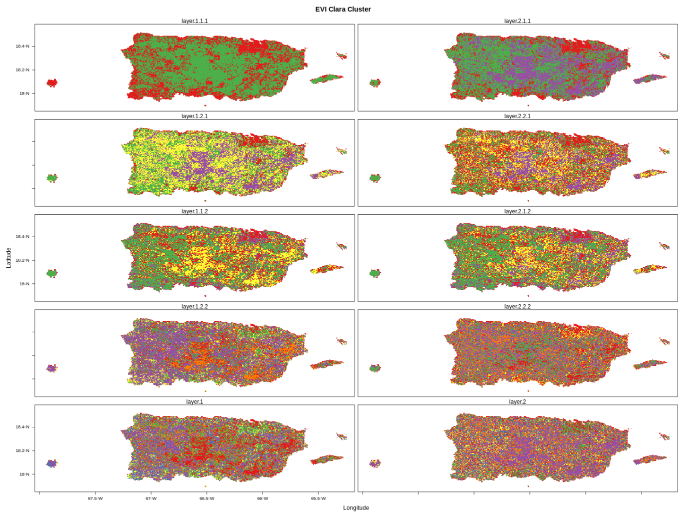

Figure A2. Results for CLARA clustering of EVI from 2 to 11 clusters

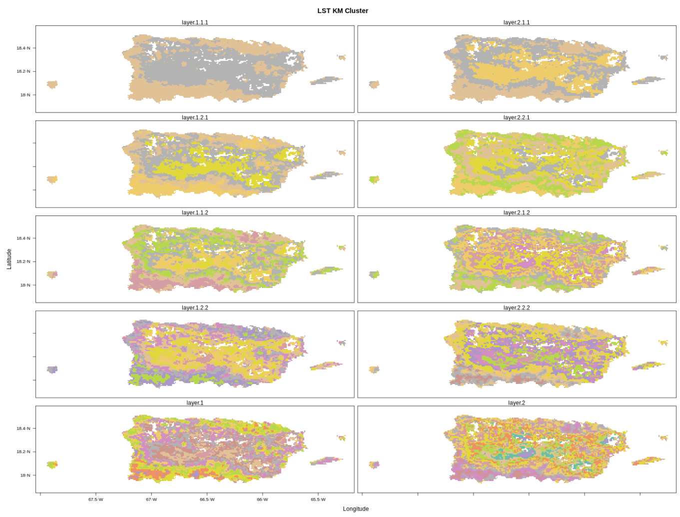

Figure A3. Results for k-means clustering of LST from 2 to 11 clusters

Table A1. Silhouette index for EVI Clusters.

Number of clusters	K-means	CLARA
2	0.5754129	0.4972576
3	0.5325603	0.4666588
4	0.5147639	0.5038890
5	0.5140073	0.4612106
6	0.5279513	0.4745793
7	0.5340362	0.4978624
8	0.5280703	0.4883521
9	0.5327400	0.4888689
10	0.5292343	0.4936729
11	0.5257909	0.4976052

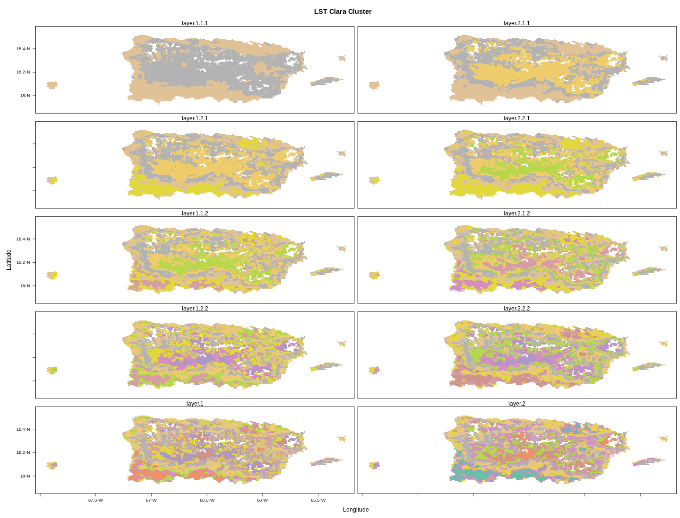

Figure A4. Results for CLARA clustering of LST from 2 to 11 clusters.

249 ## Appendix B. Unsupervised clustering analysis

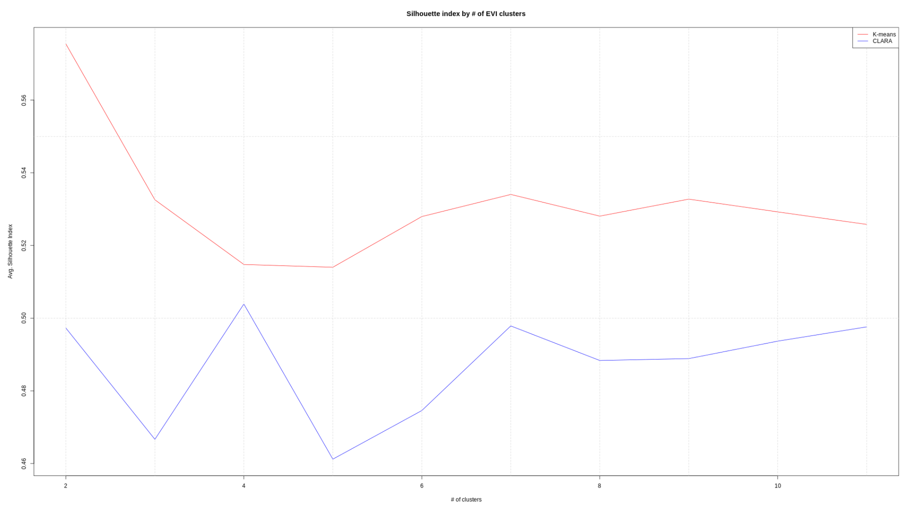

Figure A5. EVI clustering result comparison on Silhouette index.

Table A2. Silhouette index for LST Clusters.

Number of clusters	K-means	CLARA
2	0.5669281	0.5743742
3	0.5440286	0.5277574
4	0.5306771	0.5262508
5	0.5237698	0.5106174
6	0.5191691	0.5127648
7	0.5212166	0.5016520
8	0.5119441	0.5060061
9	0.5089707	0.5003322
10	0.5284022	0.4972170
11	0.5143890	0.5010930

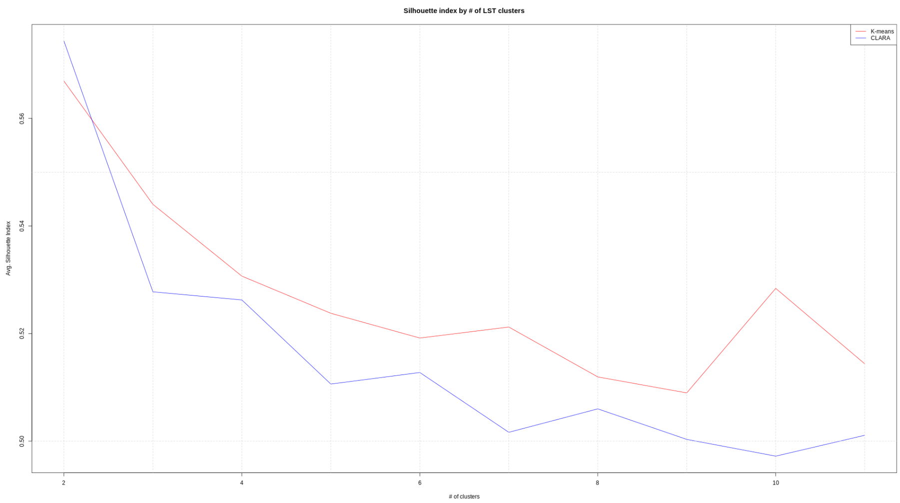

Figure A6. LST clustering result comparison on Silhouette index.

Appendix C. Seasonal model evaluation

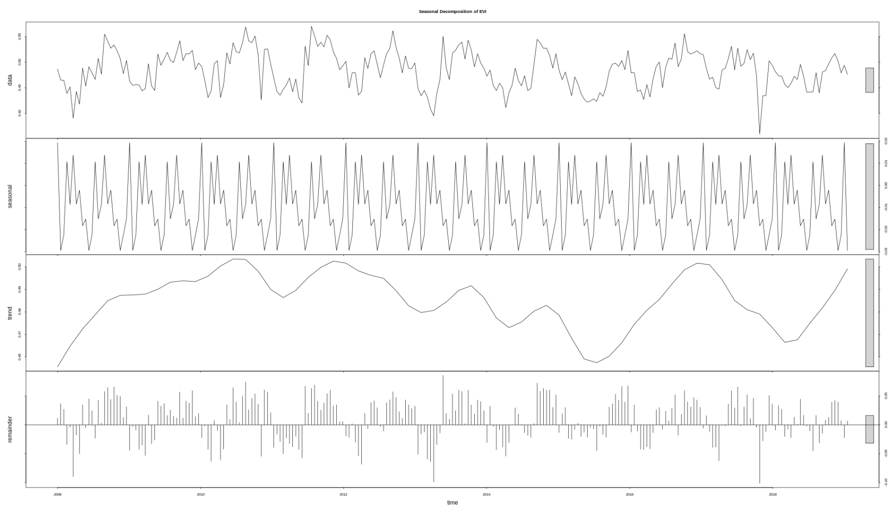

Figure A7. Average EVI decomposition into seasonal component and trend.

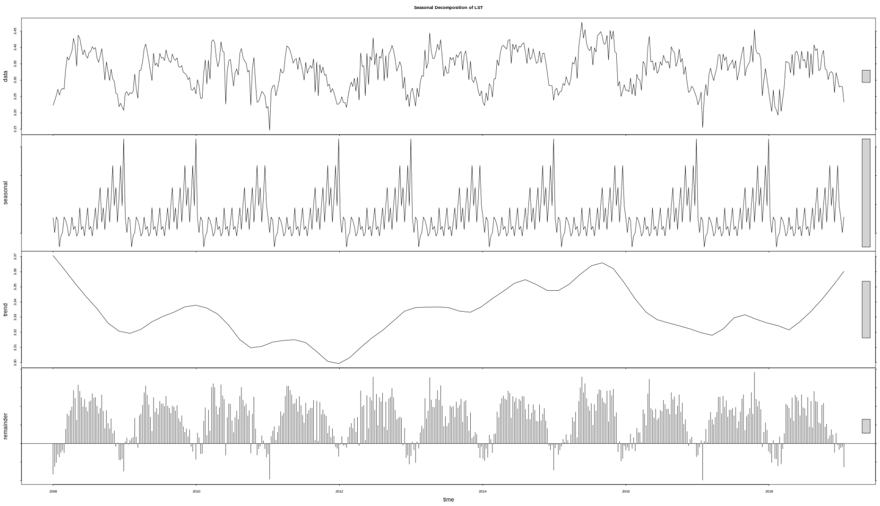

Figure A8. Average EVI decomposition into seasonal component and trend.

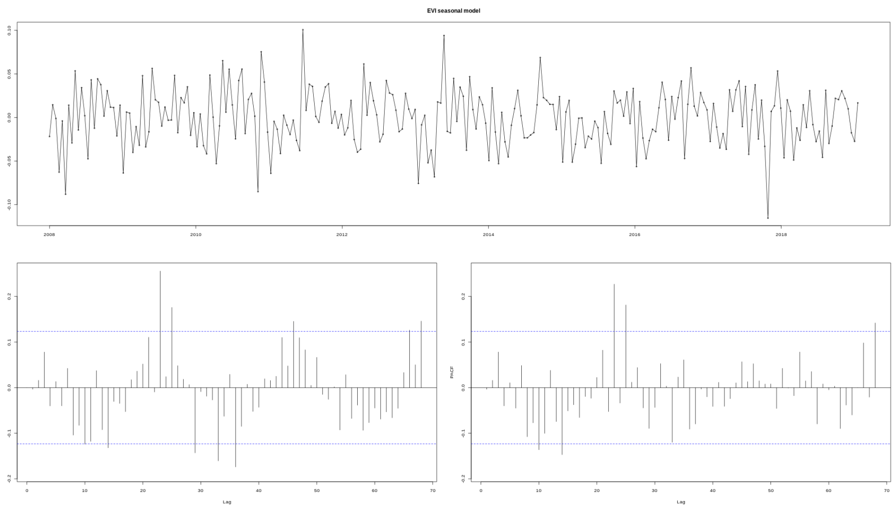

Figure A9. EVI model residuals.

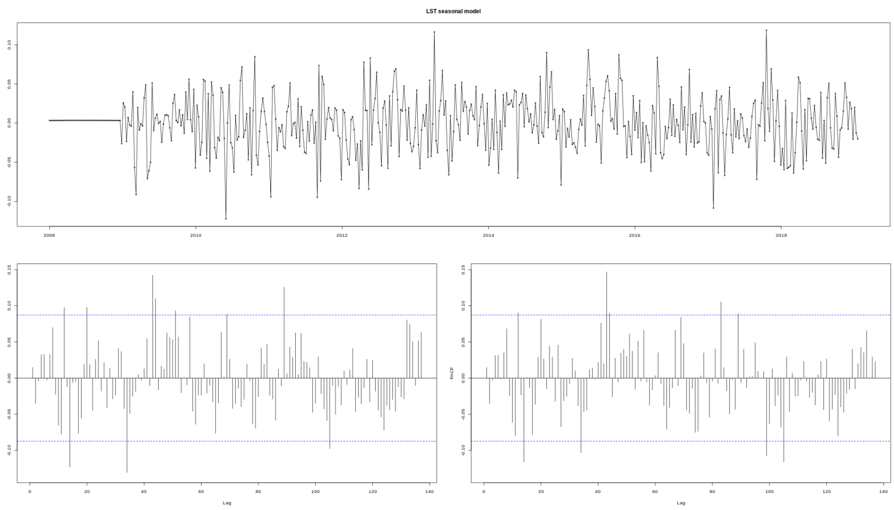

Figure A10. LST model residuals.

Appendix D. Code

```
# [plot solution]
fig_label <-
  function(text,
           region = "figure",
           pos = "topleft",
           cex = NULL,
           ...) {
    region <- match.arg(region, c("figure", "plot", "device"))
    pos <- match.arg(
      pos,
      c(
        "topleft",
        "top",
        "topright",
        "left",
        "center",
        "right",
        "bottomleft",
        "bottom",
        "bottomright"
      )
    )

    if (region %in% c("figure", "device")) {
      ds <- dev.size("in")
      # xy coordinates of device corners in user coordinates
      x <- grconvertX(c(0, ds[1]), from = "in", to = "user")
      y <- grconvertY(c(0, ds[2]), from = "in", to = "user")

      # fragment of the device we use to plot
      if (region == "figure") {
        # account for the fragment of the device that
        # the figure is using
        fig <- par("fig")
        dx <- (x[2] - x[1])
        dy <- (y[2] - y[1])
        x <- x[1] + dx * fig[1:2]
        y <- y[1] + dy * fig[3:4]
      }
    }

    # much simpler if in plotting region
    if (region == "plot") {
      u <- par("usr")
      x <- u[1:2]
      y <- u[3:4]
    }

    sw <- strwidth(text, cex = cex) * 60 / 100
    sh <- strheight(text, cex = cex) * 60 / 100

    x1 <- switch(
      pos,
      topleft     = x[1] + sw,
      left        = x[1] + sw,
      bottomleft  = x[1] + sw,
      top         = (x[1] + x[2]) / 2,
      center      = (x[1] + x[2]) / 2,
      bottom      = (x[1] + x[2]) / 2,
      topright    = x[2] - sw,
      right       = x[2] - sw,
      bottomright = x[2] - sw
    )

    y1 <- switch(
      pos,
      topleft     = y[2] - sh,
      top         = y[2] - sh,
      topright    = y[2] - sh,
      left        = (y[1] + y[2]) / 2,
      center      = (y[1] + y[2]) / 2,
      right       = (y[1] + y[2]) / 2,
      bottomleft  = y[1] + sh,
      bottom      = y[1] + sh,
      bottomright = y[1] + sh
    )

    old.par <- par(xpd = NA)
    on.exit(par(old.par))

    text(x1, y1, text, cex = cex, ...)
    return(invisible(c(x, y)))
  }

##################
# Beging Script
##################

setwd("~/Hot_zones/new")

library(ncdf4) # package for netcdf manipulation
library(raster) # package for raster manipulation
library(rgdal) # package for geospatial analysis

library(lattice)
library(RColorBrewer)
library(rasterVis)
```

```r

library(cluster)
library(clusterCrit)
library(ggplot2) # packages for plotting
library(scales)

library(tidyverse)
library(lubridate)
library(cowplot)

library(forecast)
library(tseries)
library(bfast)
# library(map)

# library(zoo)
# library(xts)

############
# Load Data
############
# Load EVI data
eviStack <- stack(raster(x = "MOD13Q1.006_250m_aid0001.nc"))
# Apparently scale factor is already applied in NCDF4
# eviStack$X250m.16.days.EVI <- (eviStack$X250m.16.days.EVI * 0.0001)

# Load LST data
lstStack <- stack(raster(x = "MOD11A2.006_1km_aid0001.nc"))
# Apparently scale factor is already applied in NCDF4
# lstStack$X8.day.daytime.1km.grid.Land.surface.Temperature <- (lstStack$X8.day.daytime.1km.grid.Land.surface.Temperature * 0.02)
lstStack$X8.day.daytime.1km.grid.Land.surface.Temperature <-
  (lstStack$X8.day.daytime.1km.grid.Land.surface.Temperature - 273.15)

# -----------------------------------------------
# Zonification
# -----------------------------------------------

######################
# Exploratory analysis
######################
# EVI mean and histogram plots
png(filename = "../Figures/Exploratory_EVI.png", width = 1240)
par(mfrow = c(1, 2))
plot(
  mean(eviStack),
  main = "Mean EVI Spatial",
  xlab = "Longitude",
  ylab = "Latitude",
  col = brewer.pal(8, "RdYlGn")
)
fig_label("A", cex = 2)
hist(eviStack,
     col = "forestgreen",
     xlab = "EVI",
     main = "Average EVI histogram")
fig_label("B", cex = 2)
dev.off()

# LST mean and histogram plots
png(filename = "../Figures/Exploratory_LST.png", width = 1240)
par(mfrow = c(1, 2))
plot(
  mean(lstStack),
  main = "Mean LST Spatial",
  xlab = "Longitude",
  ylab = "Latitude",
  col = rev(brewer.pal(8, "RdYlBu"))
)
fig_label("A", cex = 2)
hist(lstStack,
     col = "firebrick",
     xlab = "Degrees Celsius ",
     main = "Average LST histogram")
fig_label("B", cex = 2)
dev.off()

###########################
# Individual Countour Plots
###########################
par(mfrow = c(1, 1))
png(filename = "../Figures/Exploratory_EVI_Mean.png",
    width = 1920,
    height = 1080)
levelplot(
  eviStack,
  main = "Average EVI 10 year composite",
  col.regions = colorRampPalette(brewer.pal(8, "RdYlGn")),
  margin = list(FUN = 'mean'),
  contour = T
)
dev.off()

png(filename = "../Figures/Exploratory_LST_Mean.png",
    width = 1920,
    height = 1080)
levelplot(
  lstStack,
  main = "Average LST 10 year composite",
  col.regions = colorRampPalette(rev(brewer.pal(8, "RdYlBu"))),
  margin = list(FUN = 'mean'),
```

```
  contour = T
)
dev.off()

##############################
# EVI Zonification Clustering
##############################
par(mfrow = c(1, 1))
eviBrick <- brick(eviStack)
eviDF <- values(eviBrick)

# Check NA's in the data
idx <- complete.cases(eviDF)

# Max Clusters
n <- 11
# Initiate the raster datasets that will hold all clustering solutions
# from 2 groups/clusters up to n
rstKM <- raster(eviBrick[[1]])
rstCLARA <- raster(eviBrick[[1]])

for (nClust in 2:n) {
  cat("-> Clustering data for nClust =", nClust, "......")
  km <- kmeans(eviDF[idx,], centers = nClust, iter.max = 50)
  cla <- clara(eviDF[idx,], k = nClust, metric = "manhattan")

  # Create a temporary integer vector for holding cluster numbers
  kmClust <- vector(mode = "integer", length = ncell(eviBrick))
  claClust <- vector(mode = "integer", length = ncell(eviBrick))

  # Generate the temporary clustering vector for K-means (keeps track of NA's)
  kmClust[!idx] <- NA
  kmClust[idx] <- km$cluster

  # Generate the temporary clustering vector for CLARA (keeps track of NA's too ;-)
  claClust[!idx] <- NA
  claClust[idx] <- cla$clustering

  # Create a temporary raster for holding the new clustering solution
  # K-means
  tmpRstKM <- raster(eviBrick[[1]])
  # CLARA
  tmpRstCLARA <- raster(eviBrick[[1]])

  # Set raster values with the cluster vector
  # K-means
  values(tmpRstKM) <- kmClust
  # CLARA
  values(tmpRstCLARA) <- claClust

  # Stack the temporary rasters onto the final ones
  if (nClust == 2) {
    rstKM    <- tmpRstKM
    rstCLARA <- tmpRstCLARA
  } else{
    rstKM    <- stack(rstKM, tmpRstKM)
    rstCLARA <- stack(rstCLARA, tmpRstCLARA)
  }

  cat(" done!\n\n")
}

png(filename = "../Figures/Clusters_EVI_KM.png",
    width = 1920,
    height = 1080)
levelplot(
  rstKM,
  main = "EVI KM Cluster",
  col.regions = brewer.pal(6, "Set1"),
  contour = F,
  colorkey = FALSE
)
dev.off()
png(filename = "../Figures/Clusters_EVI_Clara.png",
    width = 1920,
    height = 1080)
levelplot(
  rstCLARA,
  main = "EVI Clara Cluster",
  col.regions = brewer.pal(6, "Set1"),
  contour = F,
  colorkey = FALSE
)
dev.off()

#######################################
# EVI Zonification Clustering Analysis
#######################################
# Start a data frame that will store all silhouette values
# for k-means and CLARA
clustPerfSI <- data.frame(nClust = 2:n,
                          SI_KM = NA,
                          SI_CLARA = NA)

for (i in 1:nlayers(rstKM)) {
  # Iterate through each layer

  cat("-> Evaluating clustering performance for nClust =",
      (2:n)[i],
      "......")
```

```
  # Extract random cell samples stratified by cluster
  cellIdx_RstKM <- sampleStratified(rstKM[[i]], size = 2000)
  cellIdx_rstCLARA <- sampleStratified(rstCLARA[[i]], size = 2000)

  # Get cell values from the Stratified Random Sample from the raster
  # data frame object (rstDF)
  rstDFStRS_KM <- eviDF[cellIdx_RstKM[, 1], ]
  rstDFStRS_CLARA <- eviDF[cellIdx_rstCLARA[, 1], ]

  # Make sure all columns are numeric (intCriteria function is picky on this)
  rstDFStRS_KM[] <- sapply(rstDFStRS_KM, as.numeric)
  rstDFStRS_CLARA[] <- sapply(rstDFStRS_CLARA, as.numeric)

  # Compute the sample-based Silhouette index for:
  # K-means
  clCritKM <- intCriteria(
    traj = as.matrix(rstDFStRS_KM),
    part = as.integer(cellIdx_RstKM[, 2]),
    crit = "Silhouette"
  )
  # and CLARA
  clCritCLARA <- intCriteria(
    traj = as.matrix(rstDFStRS_CLARA),
    part = as.integer(cellIdx_rstCLARA[, 2]),
    crit = "Silhouette"
  )

  # Write the silhouette index value to clustPerfSI data frame holding
  # all results
  clustPerfSI[i, "SI_KM"]    <- clCritKM[[1]][1]
  clustPerfSI[i, "SI_CLARA"] <- clCritCLARA[[1]][1]

  cat(" done!\n\n")

}

EVI_CLUST <- as.data.frame(
  clustPerfSI,
  digits = 3,
  align = "c",
  col.names = c(
    "#clusters",
    "Avg. Silhouette (k-means)",
    "Avg. Silhouette (CLARA)"
  )
)

View(EVI_CLUST)

# Sublime text 3 regex to format tabular output
# \d+\s(\d+)\n(\d\.\d+)\n(\d\.\d+)
# \1 & \2 & \3 \\\\

png(filename = "../Figures/Cluster_EVI_SIL.png",
    width = 1920,
    height = 1080)
plot(
  clustPerfSI[, 1],
  clustPerfSI[, 2],
  xlim = c(2, n),
  ylim = range(clustPerfSI[, 2:3]),
  type = "n",
  ylab = "Avg. Silhouette Index",
  xlab = "# of clusters",
  main = "Silhouette index by # of EVI clusters"
)

# Plot Avg Silhouette values across # of clusters for K-means
lines(clustPerfSI[, 1], clustPerfSI[, 2], col = "red")
# Plot Avg Silhouette values across # of clusters for CLARA
lines(clustPerfSI[, 1], clustPerfSI[, 3], col = "blue")

# Grid lines
abline(v = 2:n, lty = 2, col = "light grey")
abline(h = seq(0.20, 1.0, 0.05),
       lty = 2,
       col = "light grey")

legend(
  "topright",
  legend = c("K-means", "CLARA"),
  col = c("red", "blue"),
  lty = 1,
  lwd = 1
)
dev.off()

png(filename = "../Figures/Best_Cluster_EVI.png",
    width = 1920,
    height = 1080)
levelplot(
  rstKM,
  layer = 1,
  main = "EVI KM best cluster",
  col.regions = colorRampPalette(brewer.pal(3, "RdYlGn")),
  contour = T,
  colorkey = FALSE
)
dev.off()
```

```

##############################
# LST Zonification Clustering
##############################
lstBrick <- brick(lstStack)
lstDF <- values(lstBrick)

# Check NA's in the data
idx <- complete.cases(lstDF)

# Max Clusters
n <- 11
# Initiate the raster datasets that will hold all clustering solutions
# from 2 groups/clusters up to n
rstKM <- raster(lstBrick[[1]])
rstCLARA <- raster(lstBrick[[1]])

for (nClust in 2:n) {
  cat("-> Clustering data for nClust =", nClust, "......")
  km <- kmeans(lstDF[idx,], centers = nClust, iter.max = 50)
  cla <- clara(lstDF[idx,], k = nClust, metric = "manhattan")

  # Create a temporary integer vector for holding cluster numbers
  kmClust <- vector(mode = "integer", length = ncell(lstBrick))
  claClust <- vector(mode = "integer", length = ncell(lstBrick))

  # Generate the temporary clustering vector for K-means (keeps track of NA's)
  kmClust[!idx] <- NA
  kmClust[idx] <- km$cluster

  # Generate the temporary clustering vector for CLARA (keeps track of NA's too ;-)
  claClust[!idx] <- NA
  claClust[idx] <- cla$clustering

  # Create a temporary raster for holding the new clustering solution
  # K-means
  tmpRstKM <- raster(lstBrick[[1]])
  # CLARA
  tmpRstCLARA <- raster(lstBrick[[1]])

  # Set raster values with the cluster vector
  # K-means
  values(tmpRstKM) <- kmClust
  # CLARA
  values(tmpRstCLARA) <- claClust

  # Stack the temporary rasters onto the final ones
  if (nClust == 2) {
    rstKM    <- tmpRstKM
    rstCLARA <- tmpRstCLARA
  } else{
    rstKM    <- stack(rstKM, tmpRstKM)
    rstCLARA <- stack(rstCLARA, tmpRstCLARA)
  }

  cat(" done!\n\n")
}

png(filename = "../Figures/Clusters_LST_KM.png",
    width = 1920,
    height = 1080)
levelplot(
  rstKM,
  main = "LST KM Cluster",
  col.regions = colorRampPalette(rev(brewer.pal(8, "Set2"))),
  contour = F,
  colorkey = FALSE
)
dev.off()

png(filename = "../Figures/Clusters_LST_Clara.png",
    width = 1920,
    height = 1080)
levelplot(
  rstCLARA,
  main = "LST Clara Cluster",
  col.regions = colorRampPalette(rev(brewer.pal(8, "Set2"))),
  contour = F,
  colorkey = FALSE
)
dev.off()

#######################################
# LST Zonification Clustering Analysis
#######################################
# Start a data frame that will store all silhouette values
# for k-means and CLARA
clustPerfSI <- data.frame(nClust = 2:n,
                          SI_KM = NA,
                          SI_CLARA = NA)

for (i in 1:nlayers(rstKM)) {
  # Iterate through each layer

  cat("-> Evaluating clustering performance for nClust =",
      (2:n)[i],
      "......")

  # Extract random cell samples stratified by cluster
  cellIdx_RstKM <- sampleStratified(rstKM[[i]], size = 2000)
```

```
  cellIdx_rstCLARA <- sampleStratified(rstCLARA[[i]], size = 2000)

  # Get cell values from the Stratified Random Sample from the raster
  # data frame object (rstDF)
  rstDFStRS_KM <- lstDF[cellIdx_RstKM[, 1], ]
  rstDFStRS_CLARA <- lstDF[cellIdx_rstCLARA[, 1], ]

  # Make sure all columns are numeric (intCriteria function is picky on this)
  rstDFStRS_KM[] <- sapply(rstDFStRS_KM, as.numeric)
  rstDFStRS_CLARA[] <- sapply(rstDFStRS_CLARA, as.numeric)

  # Compute the sample-based Silhouette index for:
  #
  # K-means
  clCritKM <- intCriteria(
    traj = as.matrix(rstDFStRS_KM),
    part = as.integer(cellIdx_RstKM[, 2]),
    crit = "Silhouette"
  )
  # and CLARA
  clCritCLARA <- intCriteria(
    traj = as.matrix(rstDFStRS_CLARA),
    part = as.integer(cellIdx_rstCLARA[, 2]),
    crit = "Silhouette"
  )

  # Write the silhouette index value to clustPerfSI data frame holding
  # all results
  clustPerfSI[i, "SI_KM"]    <- clCritKM[[1]][1]
  clustPerfSI[i, "SI_CLARA"] <- clCritCLARA[[1]][1]

  cat(" done!\n\n")

}

LST_CLUST <- as.data.frame(
  clustPerfSI,
  digits = 3,
  align = "c",
  col.names = c(
    "#clusters",
    "Avg. Silhouette (k-means)",
    "Avg. Silhouette (CLARA)"
  )
)

View(LST_CLUST)

png(filename = "../Figures/Clusters_LST_SIL.png",
    width = 1920,
    height = 1080)
plot(
  clustPerfSI[, 1],
  clustPerfSI[, 2],
  xlim = c(2, n),
  ylim = range(clustPerfSI[, 2:3]),
  type = "n",
  ylab = "Avg. Silhouette Index",
  xlab = "# of clusters",
  main = "Silhouette index by # of LST clusters"
)

# Plot Avg Silhouette values across # of clusters for K-means
lines(clustPerfSI[, 1], clustPerfSI[, 2], col = "red")
# Plot Avg Silhouette values across # of clusters for CLARA
lines(clustPerfSI[, 1], clustPerfSI[, 3], col = "blue")

# Grid lines
abline(v = 2:n, lty = 2, col = "light grey")
abline(h = seq(0.20, 1.0, 0.05),
       lty = 2,
       col = "light grey")

legend(
  "topright",
  legend = c("K-means", "CLARA"),
  col = c("red", "blue"),
  lty = 1,
  lwd = 1
)
dev.off()

png(filename = "../Figures/Best_Clusters_LST.png",
    width = 1920,
    height = 1080)
levelplot(
  rstCLARA,
  main = "LST CLARA best cluster",
  layer = 1,
  col.regions = colorRampPalette(rev(brewer.pal(3, "RdYlBu"))),
  contour = T,
  colorkey = FALSE
)
dev.off()

#############
## ZONING END
#############

####################
```

```
## SEASONALITY BEGIN
####################

# ---------------------------------------------
# Seasonality
# ---------------------------------------------

#########################
# Exploratory Seasonality
#########################
# ---------------------------------------------
#  Load EVI seasonality data
# ---------------------------------------------
evi_stat <-
  read.csv("MOD13Q1-006-Statistics.csv", stringsAsFactors = FALSE)[-1,]
evi_stat$Date <-
  as.Date(evi_stat$Date, format = "%Y-%m-%d", deltat = 16 / 365)

# ---------------------------------------------
# Load LST Seasonality data
# ---------------------------------------------
lst_stat <-
  read.csv("MOD11A2-006-Statistics.csv", stringsAsFactors = FALSE)[-1,]

lst_stat$Mean <- lst_stat$Mean - 273.15
lst_stat$Date <-
  as.Date(lst_stat$Date, format = "%Y-%m-%d", deltat = 8 / 365)


png(filename = "../Figures/Exploratory_Seasonality.png",
    width = 1920,
    height = 1080)
par(mfrow = c(1, 2))
plot(
  x = evi_stat$Date,
  y = evi_stat$Mean,
  main = "Mean EVI Temporal" ,
  xlab = "Time" ,
  ylab = "EVI"
)
lines(
  smooth.spline(
    x = evi_stat$Date,
    y = evi_stat$Mean,
    cv = T
  ),
  lwd = 2,
  col = "green"
)
abline(
  reg = lm(evi_stat$Mean ~ evi_stat$Date),
  col = "grey",
  lw = 5
)
fig_label("A", cex = 2)
plot(
  x = lst_stat$Date,
  y = lst_stat$Mean,
  main = "Mean LST Temporal" ,
  xlab = "Time" ,
  ylab = "Degrees Celsius"
)
lines(
  smooth.spline(
    x = lst_stat$Date,
    y = lst_stat$Mean,
    cv = T
  ),
  lwd = 2,
  col = "red"
)
abline(
  reg = lm(lst_stat$Mean ~ lst_stat$Date),
  col = "grey",
  lw = 5
)
fig_label("B", cex = 2)
dev.off()

#########################
# EVI Seasonality Yearly
#########################
# Parse the dates, and use lower case names
df <- as_tibble(evi_stat) %>%
  rename_all(tolower) %>%
  mutate(date = ymd(date))

# Define the plot
p <- df %>%
  mutate(year = factor(year(date)),  # use year to define separate curves
         date = update(date, year = 1)) %>% # use a constant year for the x-axis) %>%
  ggplot(aes(date, mean, color = year)) +
  scale_x_date(date_breaks = "1 month", date_labels = "%b")

# Raw montly data
# p + geom_line()

# Hone in on a single year
# p + geom_line(aes(group = year), color = "black", alpha = 0.1) +
#   geom_line(data = function(x) filter(x, year == 2010), size = 1)
```

```
# Smoothed version
p1 <-
  p + geom_smooth(se = F) + ggtitle("Yearly EVI") # for the main title

#########################
# LST Seasonality Yearly
#########################
# Parse the dates, and use lower case names
df <- as_tibble(lst_stat) %>%
  rename_all(tolower) %>%
  mutate(date = ymd(date))

# Define the plot
p <- df %>%
  mutate(year = factor(year(date)),   # use year to define separate curves
         date = update(date, year = 1)) %>%  # use a constant year for the x-axis
  ggplot(aes(date, mean, color = year)) +
  scale_x_date(date_breaks = "1 month", date_labels = "%b")

# Raw daily data
# p + geom_line()

# Hone in on a single year
# p + geom_line(aes(group = year), color = "black", alpha = 0.1) +
#   geom_line(data = function(x) filter(x, year == 2010), size = 1)

# Smoothed version
p2 <-
  p + geom_smooth(se = F) + ggtitle("Yearly LST") # for the main title

png(filename = "../Figures/Seasonality_All.png",
    width = 1240,
    height = 1000)
plot_grid(p1,
          p2,
          labels = "AUTO",
          nrow = 2,
          ncol = 1)
dev.off()


#########################
# EVI Seasonality Yearly
#########################
b <-
  bfast(
    Yt = ts(
      evi_stat$Mean,
      deltat = 16 / 365,
      start = c(2008, 1)
    ),
    season = "harmonic",
    max.iter = 50
  )

png(filename = "../Figures/Sesonal_BFAST_EVI.png", width = 1240)
plot(b, main = "BFAST EVI")
dev.off()

####################
# EVI ARIMA Forecast
####################
X <-
  tsclean(ts(
    evi_stat$Mean,
    deltat = 16 / 365,
    start = c(2008, 1)
  ))
decomp <- stl(X, s.window = "periodic")
un_seasoned <- seasadj(decomp)
model1 <- auto.arima(un_seasoned, seasonal = T)
f1 <- forecast(model1, 48)
accuracy(f1)

png(filename = "../Figures/Sesonal_Decomp_EVI.png",
    width = 1920,
    height = 1080)
plot(decomp, main = "Seasonal Decomposition of EVI")
dev.off()

png(filename = "../Figures/Sesonal_Residuals_EVI.png",
    width = 1920,
    height = 1080)
tsdisplay(residuals(model1), main = "EVI seasonal model")
dev.off()

png(filename = "../Figures/Sesonal_Model_EVI.png",
    width = 1920,
    height = 1080)
autoplot(f1, xlab = "Time", ylab = "EVI") + ggtitle("EVI ARIMA model forecast")
dev.off()

#########################
# LST Seasonality Yearly
#########################
b <-
  bfast(
    Yt = ts(
      log(lst_stat$Mean),
```

```
      deltat = 8 / 365,
      start = c(2008, 1)
    ),
    season = "harmonic",
    max.iter = 50
  )

png(filename = "../Figures/Sesonal_BFAST_LST.png",
    width = 1920,
    height = 1080)
plot(b, main = "BFAST LST")
dev.off()

#####################
# LST ARIMA Forecast
#####################
X <-
  tsclean(ts(
    log(lst_stat$Mean),
    deltat = 8 / 365,
    start = c(2008, 1)
  ))
decomp <- stl(X, s.window = "periodic")
un_seasoned <- seasadj(decomp)
model1 <- auto.arima(un_seasoned, seasonal = T)
f1 <- forecast(model1, 96)
accuracy(f1)

png(filename = "../Figures/Sesonal_Decomp_LST.png",
    width = 1920,
    height = 1080)
plot(decomp, main = "Seasonal Decomposition of LST")
dev.off()

png(filename = "../Figures/Sesonal_Residuals_LST.png",
    width = 1920,
    height = 1080)
tsdisplay(residuals(model1), main = "LST seasonal model")
dev.off()

png(filename = "../Figures/Sesonal_Model_LST.png",
    width = 1920,
    height = 1080)
autoplot(f1, xlab = "Time", ylab = "LST") + ggtitle("LST ARIMA model forecast")
dev.off()

# RasterVis Documentation
# https://oscarperpinan.github.io/rastervis/

# Unsupervised Clustering on Raster Data
# https://www.r-exercises.com/2018/02/28/advanced-techniques-with-raster-data-part-1-unsupervised-classification/

# Plot Year Stacking
# https://stackoverflow.com/questions/48722758/r-how-to-create-a-seasonal-plot-different-lines-for-years

# Seasonality ETS
# https://www.statmethods.net/advstats/timeseries.html

# Seasonality Auto.Arima
# https://www.datacamp.com/community/tutorials/time-series-r

# TimeSeries Analysis
# https://www.datascience.com/blog/introduction-to-forecasting-with-arima-in-r-learn-data-science-tutorials
```

References

1. Atkinson, P.M.; Lewis, P. Geostatistical classification for remote sensing: an introduction. *Computers & Geosciences* **2000**, *26*, 361-371, DOI:10.1016/S0098-3004(99)00117-X.
2. Bertin, R.I. Plant Phenology And Distribution In Relation To Recent Climate Change. *The Journal of the Torrey Botanical Society* **2008**, *135*, 1:126-1:46, DOI:10.3159/07-rp-035r.1.
3. Chuvieco, E.; Cocero, D.; Riaño, D.; Martin, P.; Martínez-Vega, J.; de la Riva, J.; Pérez, F. Combining NDVI and Surface Temperature for the Estimation of Live Fuel Moisture Content in Forest Fire Danger Rating. *Remote Sensing of Environment* **2004**, *92*, 3:322-3:31, DOI:10.1016/j.rse.2004.01.019.
4. Didan, K. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. [Data set] (Accessed April 28, 2019)
5. Eastman, J.; Sangermano, F.; Machado, E.; Rogan, J.; & Anyamba, A. Global trends in seasonality of normalized difference vegetation index (NDVI), 1982–2011. *Remote Sensing* **2013**, *5*, 10:4799-10:4818, DOI:10.3390/rs5104799.
6. Fernández-Manso, A.; Quintano, C.; Fernández-Manso, O. Forecast of NDVI in coniferous areas using temporal ARIMA analysis and climatic data at a regional scale *International Journal of Remote Sensing* **2011**, *32*, 1595-1617, DOI:10.1080/01431160903586765

7. García-Mora, T.J.; Mas, J.; Hinkley, E.A. Land cover mapping applications with MODIS: a literature review. *International Journal of Digital Earth* **2012**, *5:1*, 63-87, DOI:10.1080/17538947.2011.565080
8. Han, P.; Wang, P.X.; Zhang, S.Y.; Zhu, D.H. Drought forecasting based on the remote sensing data using ARIMA models *Mathematical and Computer Modelling* **2010**, *51*, 1398-1403, DOI:10.1016/j.mcm.2009.10.031
9. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* **2002**, *83*, 195-213, DOI:10.1016/S0034-4257(02)00096-2
10. Huete, A.; Justice, C.; Van Leeuwen, W. MODIS vegetation index (MOD13) algorithm theoretical basis document. Ver. 3 *NASA: Washington, DC, USA* **1999**.
11. Jennings, M.D. Gap analysis: Concepts, methods, and recent results. *Landscape Ecology* **2000**, *15*, 5-20, DOI:10.1023/A:1008184408300.
12. Martinuzzi, S.; Gould, W.A.; Gonzalez, O.M.R. Land development, land use, and urban sprawl in Puerto Rico integrating remote sensing and population census data. *Landscape and Urban Planning* **2017**, *79*, 288-297.
13. MODIS Evapotranspiration. https://modis.gsfc.nasa.gov/data/dataprod/mod16.php (Accessed April 28, 2019)
14. MODIS Land Surface Temperature and Emissivity (MOD11).https://modis.gsfc.nasa.gov/data/dataprod/mod11.php (Accessed April 28, 2019)
15. MODIS Vegetation Index Products (NDVI and EVI). https://modis.gsfc.nasa.gov/data/dataprod/mod13.php (Accessed April 28, 2019)
16. Reiter, M.E.; Elliott, N.K.; Jongsomjit, D.; Golet, G.H.; Reynolds, M.D. Impact of extreme drought and incentive programs on flooded agriculture and wetlands in California's Central Valley. *PeerJ* **2018**, *6:e5147*, DOI:10.7717/peerj.5147.
17. Rigge, M.; Smart, A.; Wylie, B.; Gilmanov, T.; Johnson, P. Linking Phenology and Biomass Productivity in South Dakota Mixed-Grass Prairie. *Rangeland Ecology & Management* **2013**, *66*, 579-87, DOI:10.2111/rem-d-12-00083.1.
18. Running, S., Mu, Q., Zhao, M. MOD16A2 MODIS/Terra Net Evapotranspiration 8-Day L4 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC.[Data set] (Accessed April 28, 2019)
19. Senay, G.; Velpuri, N.;Bohms, S.; Budde, M.; Young, C.; Rowland, J.; Verdin, J. Drought monitoring and assessment: remote sensing and modeling approaches for the famine early warning systems network. *Hydro-Meteorological Hazards, Risks and Disasters* **2015**, 233-262. DOI:10.1016/B978-0-12-394846-5.00009-6.
20. Steyer, G.D.; Couvillion, B.R.; Barras, J.A. Monitoring vegetation response to episodic disturbance events by using multitemporal vegetation indices *Journal of Coastal Research* **2013**, *Spring*, 118-130, DOI:10.2112/SI63-011.1.
21. Sjostrom, M.; Ardo, J.; Arneth, A.; Boulain, N.; Cappelaere, B.; Eklundh, L.; De Grandcourt, A.; Kutsch, W.L.; Merbold, L.; Nouvellon, Y.; Scholes, R.J.; Schubert, P.; Seaquist, J.; Veenendaal, E.M. Exploring the potential of MODIS EVI for modeling gross primary production across African ecosystems *Remote Sensing of Environment* **2011**, *115*, 1081-1089, DOI:10.1016/j.rse.2010.12.013.
22. Torres-Degró, A. Envejecimiento demográfico:Un acercamiento a los métodos cuantitativos. *CIDE digital* **2010**, *1*, 79-102.
23. Psilovikos, A.; Elhag, M. Forecasting of Remotely Sensed Daily Evapotranspiration Data Over Nile Delta Region, Egypt *Water Resources Management* **2013**, *27*, 4115-4130, DOI:10.1007/s11269-013-0368-2
24. Verbesselt, J.; Hyndman, R.; Newnham, G.; Culvenor, D. Detecting trend and seasonal changes in satellite image time series. *Remote Sensing of Environment* **2009**, *114*, 106–115, DOI:10.1016/j.rse.2009.08.014.
25. Wan, Z.; Hook, S.; and Hulley, G. MOD11A2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V006. [Data set] (Accessed April 28, 2019)
26. Wang, C.; Yu, M.; Gao, Q. Continued Reforestation and Urban Expansion in the New Century of a Tropical Island in the Caribbean *Remote Sensingl* **2017**, *9*, 731.
27. Wulder, M.; Michael A.; Hall, R.; Coops, N.; Franklin, S. High Spatial Resolution Remotely Sensed Data for Ecosystem Characterization. *BioScience* **2004**, *54*, 511–521, DOI:10.1641/0006-3568(2004)054[0511:HSRRSD]2.0.CO;2.

Sample Availability: The data used for this works was obtained from and is available through NASA's AppEEARS portal where the request ID for the data is $5643f426 - f351 - 429d - 8607 - d4555eabc307$. Projection and request information for the data utilized in the project can be downloaded from:
https://drive.google.com/a/upr.edu/file/d/1nS5aIuj7DH0MxjrtQLywf1tgBZBsxWJ4/view?usp=sharing