An introduction to acquiring, downloading, and processing Remote Sensing Data with RGISTools, a new R package.

Pérez-Goya U. [1] Montesino-SanMartin M. [1] Militino A.F. [1,2,3]
Ugarte M.D. [1,2,3]

[1]Department of Statistics, Computer Science and Mathematics
Public University of Navarre, Spain

[2]Institute for Advanced Material (*InaMat*), Public University of Navarre, Spain

Presentada por **Ana F Militino**
militino@unavarra.es

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Overview

Background

- Undoubtedly, free satellite imagery is a great source of information for many researchers
- Remote sensing data are the data derived from satellites (sensors). They are very useful in meteorology, agriculture, forestry, geology, hydrology, and natural environment sciences since several decades ago [1].
- Platforms as Landsat, Modis Terra, Modis Aqua or Sentinel provide for free and routinely, high quality images with different temporal (periodicity) and spatial resolutions (pixel size).
- "RGISTools" package is a new R package very useful for statisticians and practitioners of remote sensing data coming from multi-spectral images.

Satellites: Landsat 7-8, and Sentinel-2

- LANDSAT, meaning Land+Satellite, (https://www.usgs.gov/land-resources/nli/landsat). It is available since 1972 from six satellites in the Landsat series. 250 and 500 $meters^2$ spatial resolution are common, weekly and by-weekly temporal resolution.
- Landsat tiles
- SENTINEL-2 (from Copernicus project) are European satellites launched from 2013 onwards [2]. See https://sentinel.esa.int/web/sentinel/home. 10 m^2 spatial resolution and 4 day temporal resolution.
- Sentinel tiles
- Sentinel Orbits

Satellites: Modis

- MODIS (Moderate Resolution Imaging Spectroradiometer) with TERRA and AQUA satellites. https://modis.gsfc.nasa.gov/about/. Daily images of 500 m^2 of spatial resolution
- TERRA's orbit around the Earth is timed so that it passes from north to south across the equator in the morning, while AQUA passes south to north over the equator in the afternoon, providing a high temporal resolution of images over the world.
- TERRA MODIS and AQUA MODIS are viewing the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands, or groups of wavelengths.
- Tiles Modis tiles

Satellites: Landsat 7-8, Sentinel-2 and Modis

Program	Landsat		MODIS		Sentinel	
Mission	Landsat-7	Landsat-8	-		Sentinel-2	
Satellite	Landsat-7	Landsat-8	Terra	Aqua	A	B
Sensor	ET+	TIRS/OLI	MODIS	MODIS	MSI	MSI
Number of Bands	8	8	36	36	12	12
Time Revisit (days)	16	16	1	1	10	10
Resolution (m)	30-60	15-30	250	250	10-60	10-60
Format	GeoTIFF	GeoTIFF	HDF-EOS	HDF-EOS	JP2	JP2

Table 1: Major satellite missions devoted to multi-spectral images and details about their datasets.

Tiles of Modis, Landsat 7-8 and Sentinel-2

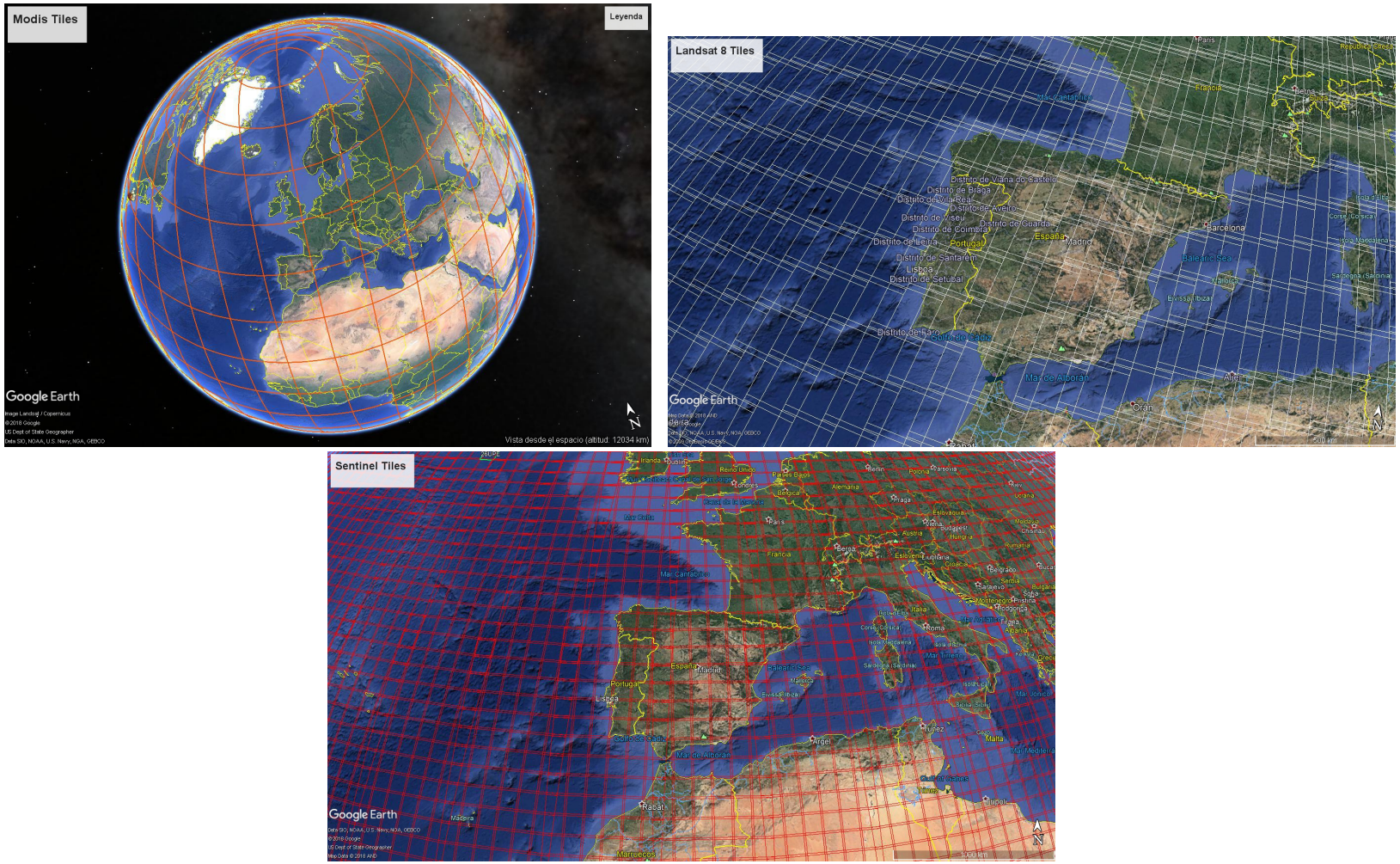
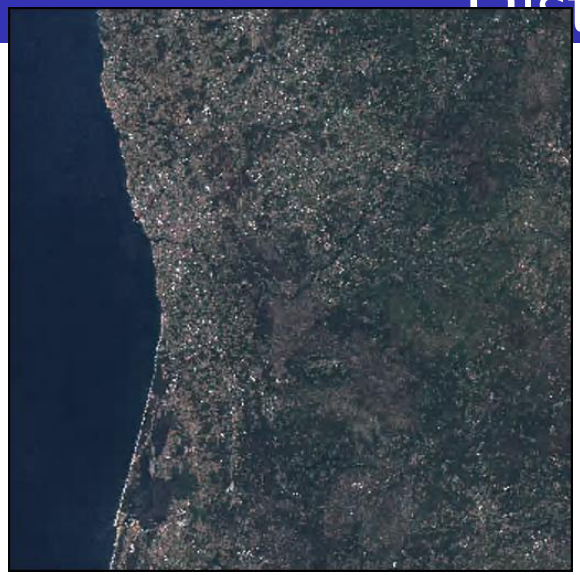
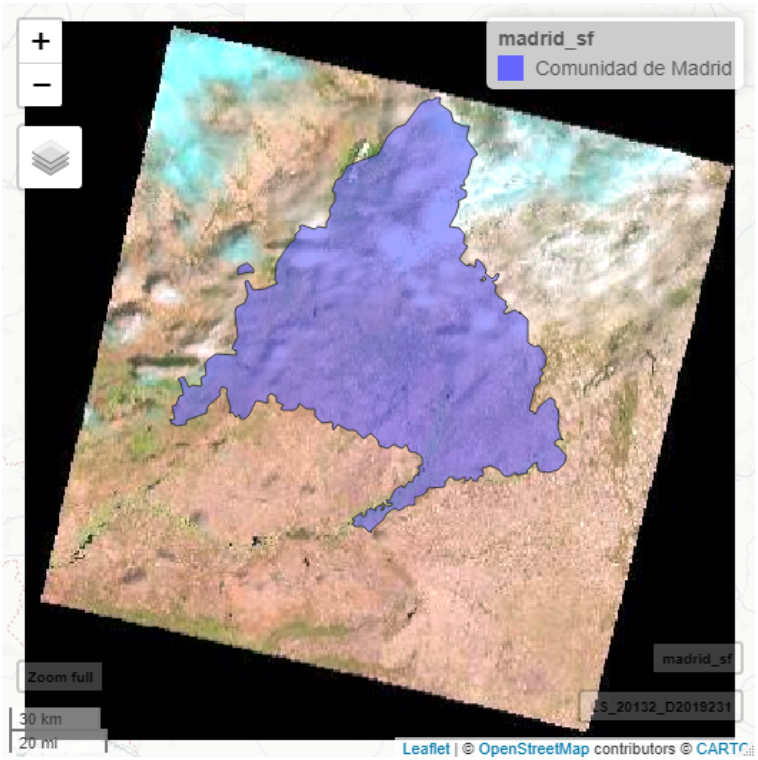


Figure 1: Modis, Landsat 7-8 and Sentinel-2 tiles

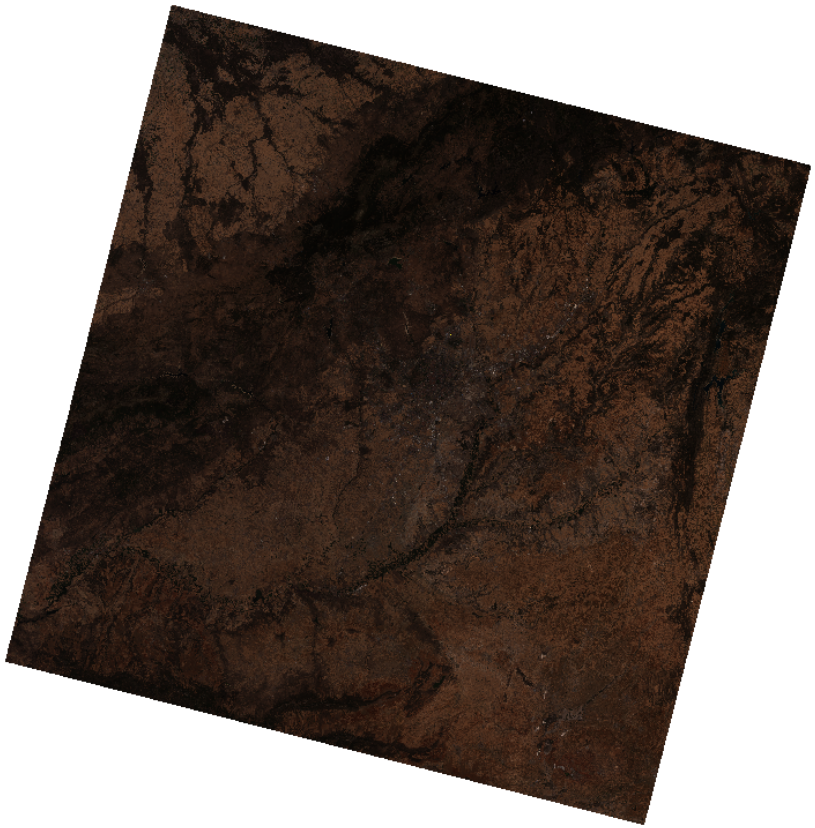
Sentinel-2, Landsat 7-8 and Modis RGB images of Porto District



Landsat 8 RGB image of Madrid



Landsat 8 RGB images of Madrid



Mosaicing Sentinel-2, RGB tiles images of Madrid

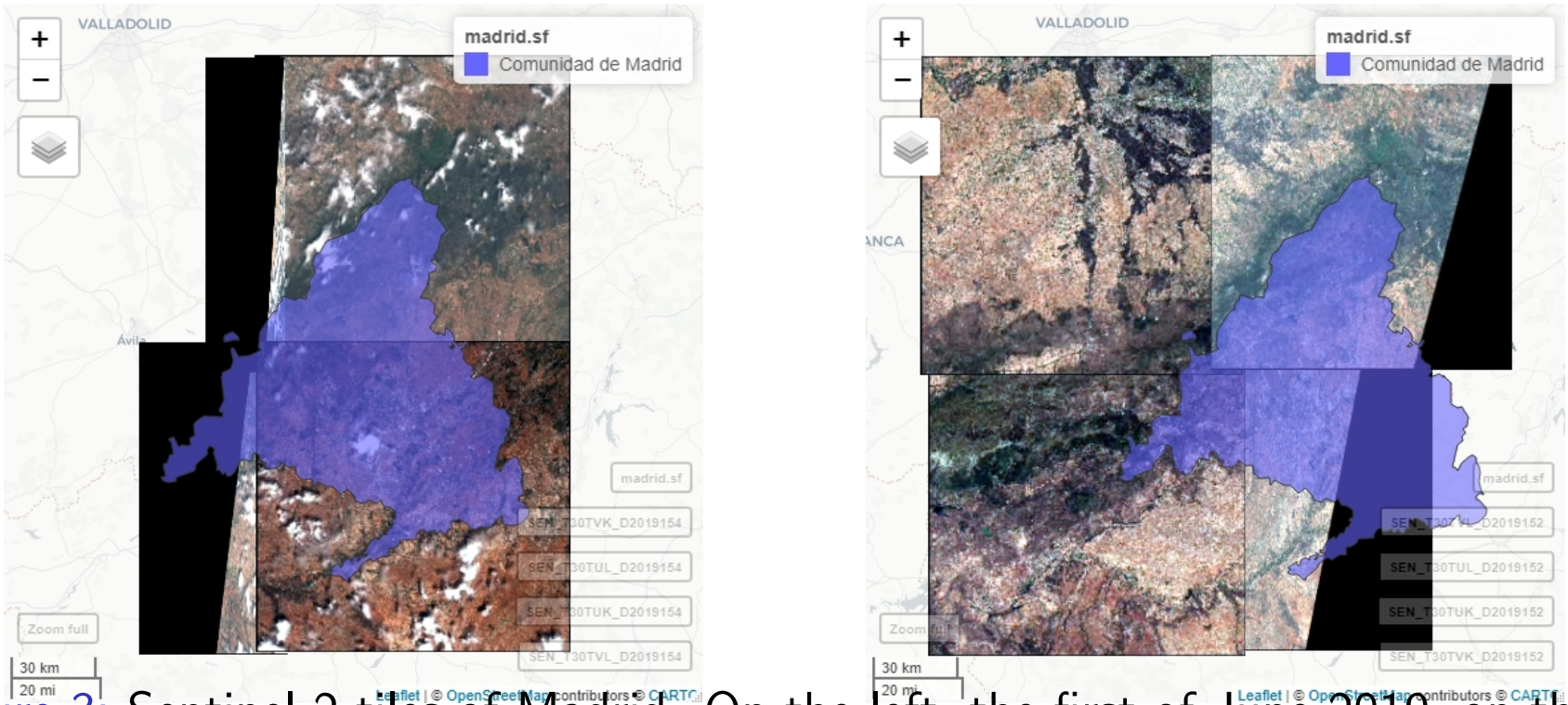


Figure 3: Sentinel-2 tiles of Madrid. On the left, the first of June 2019, on the right, the 6th of June 2019

Compositing, cropping and masking Sentinel-2, RGB tiles images of Madrid

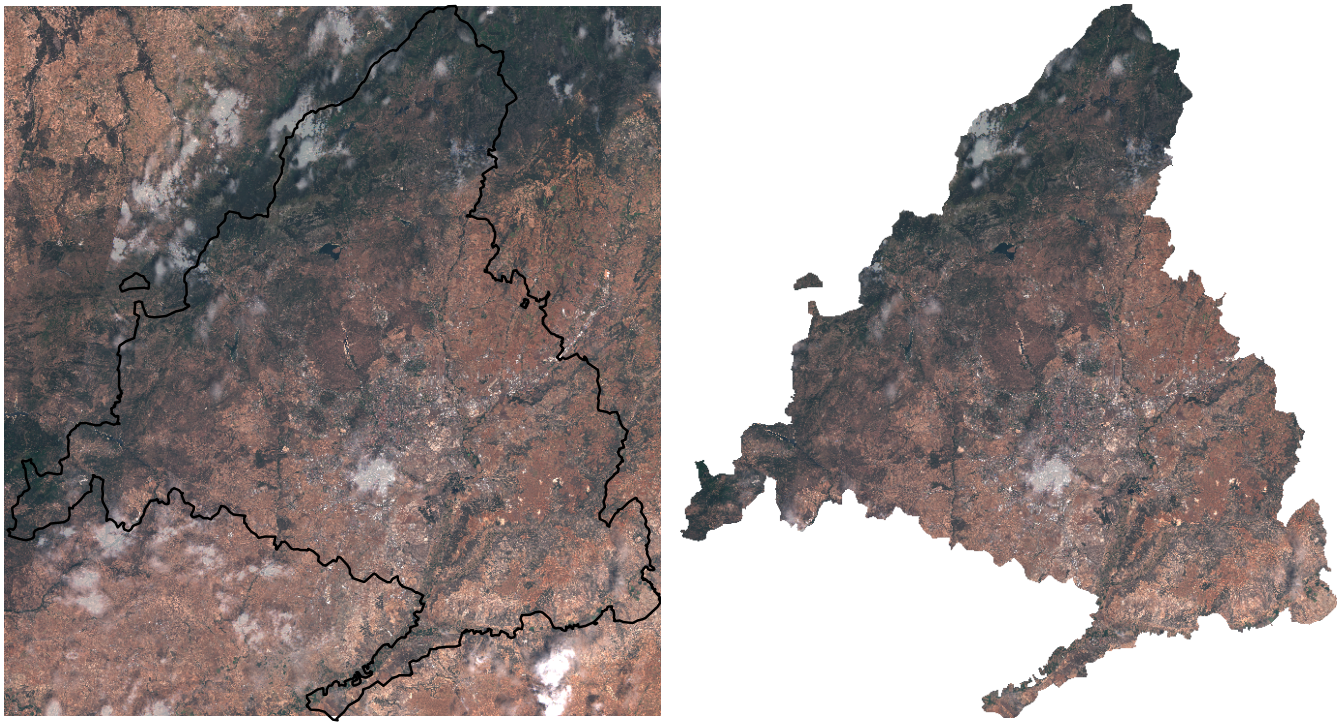


Figure 4: Composite image of Madrid from the first of June 2019, and the 6th of June 2019

Multiespectral Satellite Images

- Satellite images are data archives (zip, hdf files) with an assorted number of bands.
- Download the complete zip (hdf) files is required, yet when uncompresssing we can choose the bands.
- Bands are different according to the satellites. There are also bands for clouds, quality indicators, etc.
- The files also include information about projections, coordinates, extension, etc.
- When these bands are transformed in tif files, they are ready to load in R as "rasters", or multi-layer rasters
- The rasters are optimized matrices, with rows and columns. The rasters save the coordinates of the border of the images, instead all the coordinates for the whole image.

Remote Sensing Data

- The bands and the combination of these bands give different types of remote sensing data (derived variables).
- For example, the Normalized Difference Vegetation Index (NDVI) is given by a simple combination of bands through the radiometric information obtained with the red (R) and near-infrared (NIR) wavelengths of the electromagnetic spectrum ([3])

$$NDVI = ((NIR) - R)/((NIR) + R)$$

.

- NDVI is a vegetation index and takes values from -1 to 1 with no units.
- Land surface temperature (LST) is more difficult to define as simple combination of bands [4].

Processing of images

- Frequently, images present mechanical or atmospheric errors
- For reducing the number of missing or altered data, processing of daily images is usual.
- Compositing methods are used to select the best observation available for each pixel over the compositing period [5].
- The most common method with NDVI is maximum value compositing (MVC) assigning the maximum value across all observations for that pixel during the compositing period.
- Composite LST images are obtained averaging daily images
- Some composite images can be directly downloaded with different spatial and temporal resolutions. For example, with weekly or bi-weekly temporal resolution.

RGISTools: Accessing and analysing satellite imagery(I)

- This package downloads, customizes, and processes time series of multispectral satellite images from Landsat-7, Landsat-8, MODIS and Sentinel-2 in a standardized way.
- The usual process consists of:
  - Searching the available target images in time and space
  - Downloading and decompressing the archives
  - Mosaicking (merging) the tiles to contain the region of interest
  - Defining Remote Sensing Data. For example indices (NDVI, EVI, etc)
  - Uploading the images in RasterStack files of R
  - Masking and cropping images
  - Compositing, or smoothing the images for removing clouds of abnormal data

RGISTools: Accessing and analysing satellite imagery(II)

- 'RGISTools' converts automatically the platform-specific file formats into GTiff, so they can be uploaded in 'R'.
- The package includes a set of functions for filling and smoothing multispectral satellite images with [6] or without covariates [7].
- 'RGISTools', contains 60 functions divided into 7 categories identified by the first 3 characters of the function names.
- It can be downloaded from CRAN [8] https://CRAN.R-project.org/package=RGISTools and from https://spatialstatisticsupna.github.io/index.html

Classification of RGISTools commands

RGISTools uses common prefixes for identifying the same functions used for different satelites

- **mod** identifies MODIS Terra and Aqua satellite functions.
- **sen** identifies Sentinel functions.
- **ls7** identifies Landsat-7 functions.
- **ls8** identifies Landsat-8 functions.
- **ls** identifies both Landsat-7 and Landsat-8 functions.
- **gen** identifies functions for being used in any of the three platforms.
- **var** identifies functions for deriving variables in any of the three platforms.

Porto application with MODIS images

1. Searching for Porto Modis images in August 2016, 2017 and 2018
2. Downloading 93 images (with bands) and plot them in RGB version
3. Removing clouds of RGB images
4. Creating remote sensing NDVI data
5. Creating Composite NDVI images
6. Downloading, Compositing and Smoothing LST images

RGISTools commands for Porto application I

1. **getData** Upload "Porto" shape file (from raster package)
2. **modSearch** Search for images intersecting "Porto" District
3. **modDownload** Download Modis satellite images, collection 6 of MOD09GA
4. **modExtractHDF** Extract Gtiff files from downloaded hdf files
5. **modMosaic** Merge images containing Porto (if necessary)
6. **modFolderToVar(fun = varRGB)** Define RGB images
7. **spTransform** Transforms shp Porto long-lat file to sinusoidal projection
8. **mask** Mask RGB images with porto shape file in sinu
9. **plotRGB** Plot raster stack of RGB images

1.- **getData** Upload "Porto" shape file

Figure 5: Porto district shape file

```
> PortugalRegions<- raster::getData('GADM', country='Portugal', level=1)
> Porto<-subset(PortugalRegions["NAME_1"],NAME_1=="Porto")
> Porto
class       : SpatialPolygonsDataFrame
features    : 1
extent      : -8.786806, -7.875585, 41.00138, 41.47124  (xmin, xmax, ymin, ymax)
crs         : +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0
variables   : 1
names       : NAME_1
value       :  Porto
```

stack provides a Raster Stack of RGB images

```
> src3 <- file.path(src, "Modis","MOD09GA","Variables")
> src3.rgb<- file.path(src3, "RGB")
> tiles.path <- list.files(src3.rgb,
                           full.names = TRUE,
                           recursive = TRUE,
                           pattern = "\\.tif$") ## Locates the tif fil
> rgb.tiles<-stack(tiles.path) ## Provides a raster stack
> rgb.tiles[[93]]
class       : RasterStack
dimensions  : 113, 175, 19775, 3  (nrow, ncol, ncell, nlayers)
resolution  : 464.0102, 462.3513  (x, y)
extent      : -737373.1, -656171.3, 4559151, 4611396  (xmin, xmax, ym
crs         : +proj=sinu +lon_0=0 +x_0=0 +y_0=0 +a=6371007.181 +b=637
names       : RGB_2018243.1, RGB_2018243.2, RGB_2018243.3
min values  :               0,               0,               0
max values  :             255,             255,             255
```

spTransform, lapply and mask

```
########### 7.- Transform longlat in sinusoidal projection
> porto.sinu<-spTransform(porto,proj4string(rgb.tiles))
########### 8.- Mask images with shp Porto file in sinu proj
> fun1<-function(x) mask(x,porto.sinu)
> rgb.tiles.rec<-lapply(rgb.tiles,fun1)
########### 9.- Plot RGB images
> win.graph()
> par(mfrow=c(6,6))
> for (i in 1:31) {plotRGB(rgb.tiles.rec[[i]])}
> win.graph()
> for (i in 32:62) {plotRGB(rgb.tiles.rec[[i]])}
> win.graph()
> for (i in 63:93) {plotRGB(rgb.tiles.rec[[i]])}
```

9.- **plotRGB:** Plot RGB images of 2016 before masking

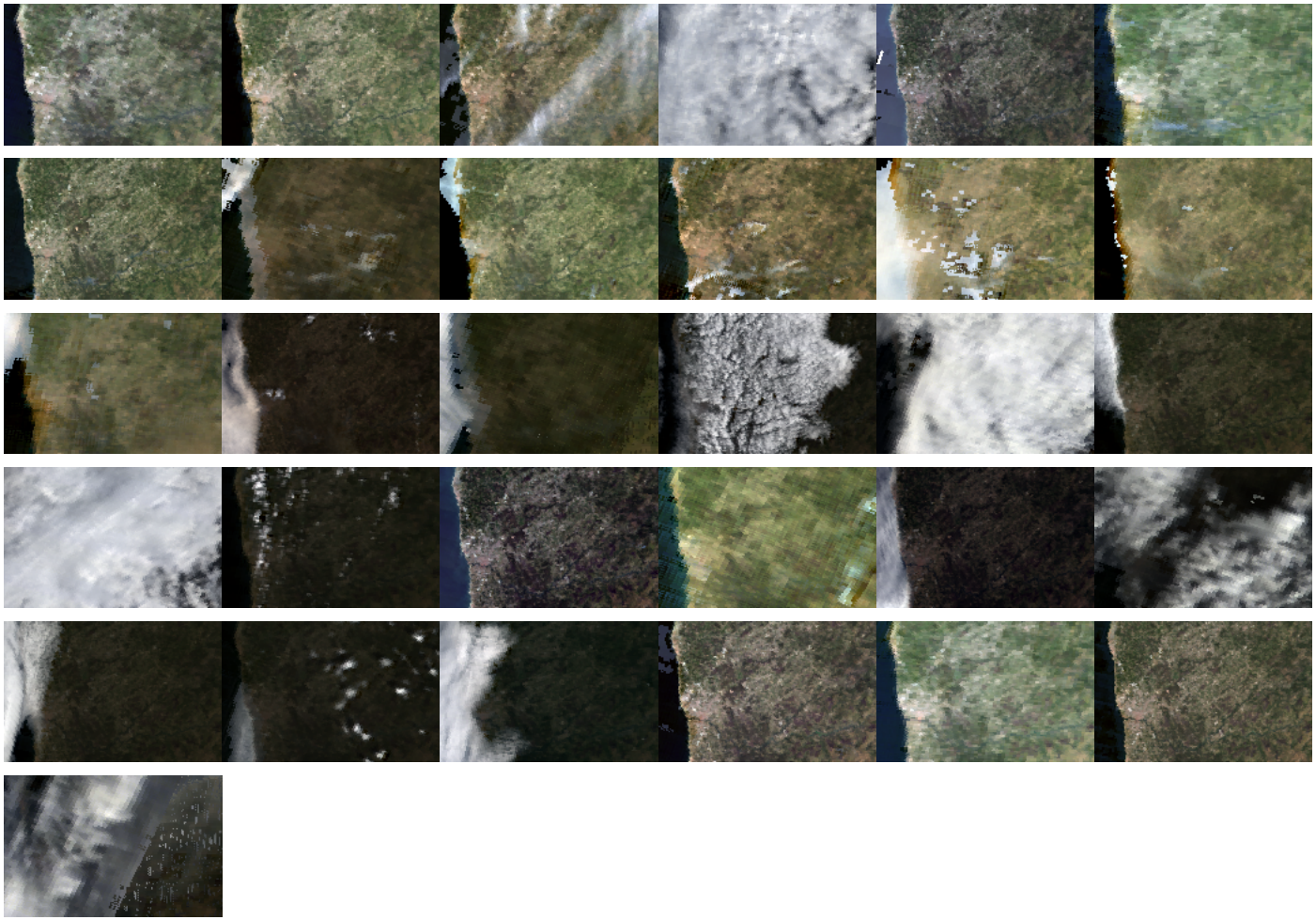


Figure 6: Porto Modis Terra daily images in August of 2016

9.- **plotRGB:** Plot RGB images of 2017 before masking

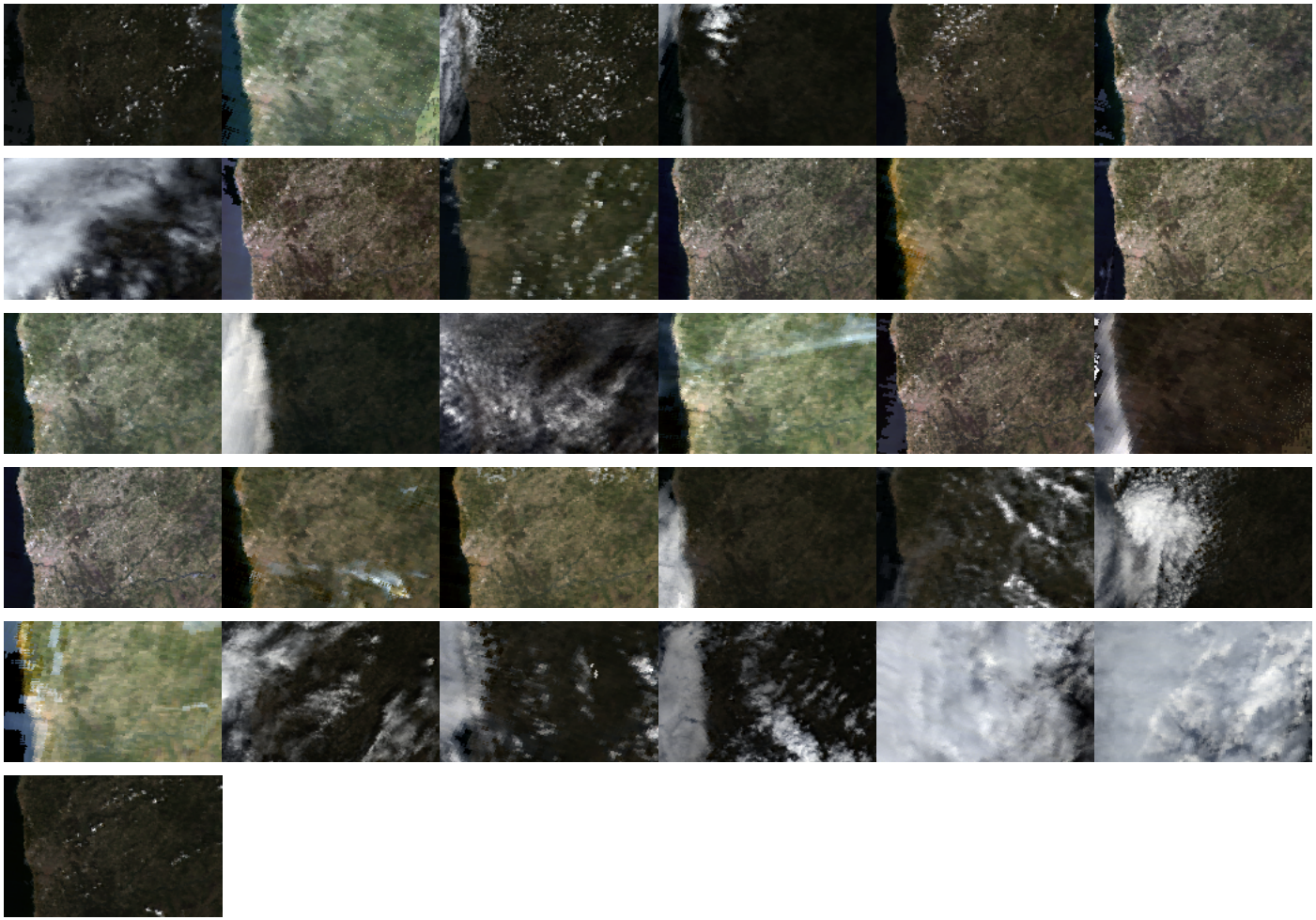


Figure 7: Porto Modis Terra daily images in August of 2017

9.- **plotRGB:** Plot RGB images of 2018 before masking

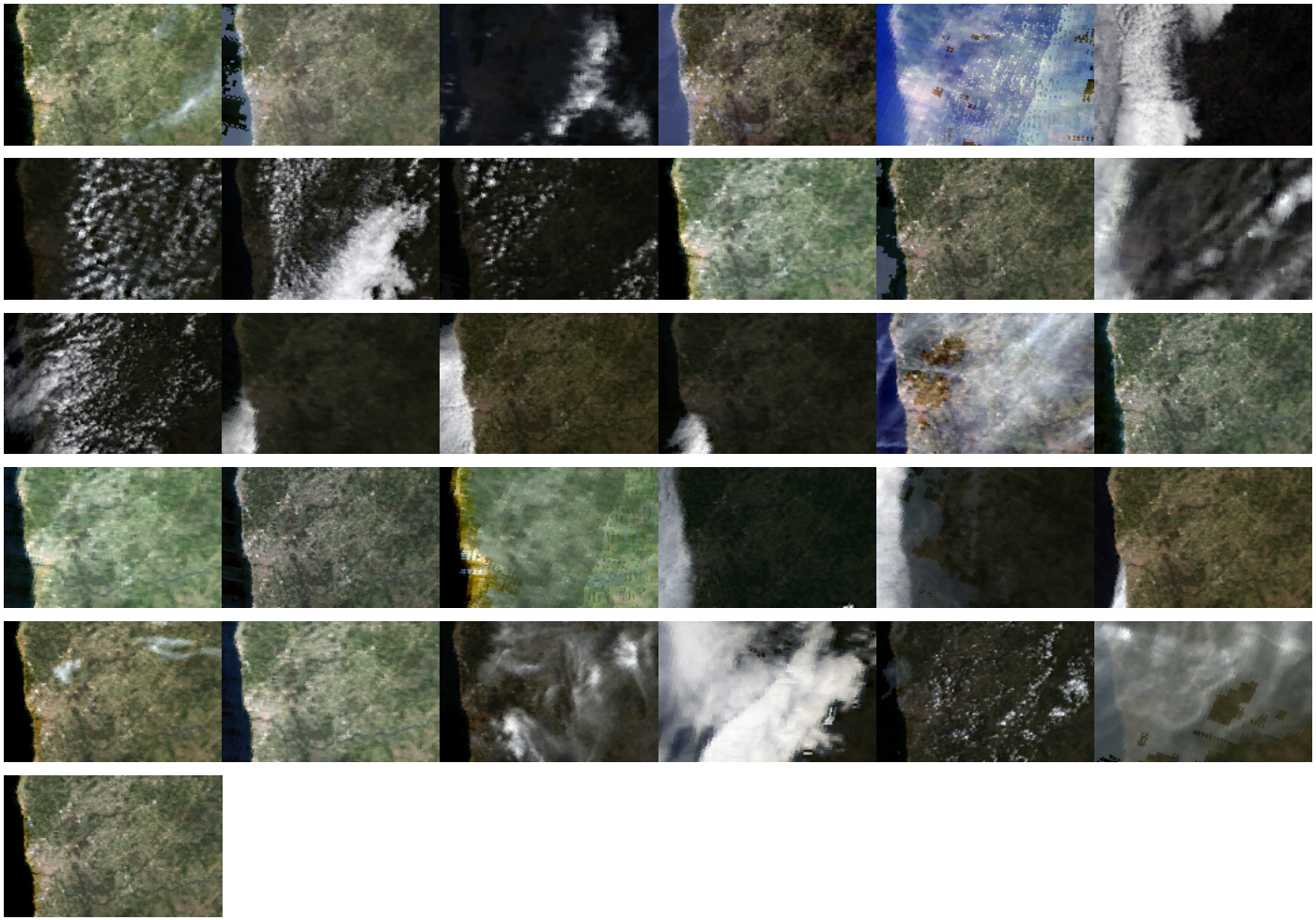


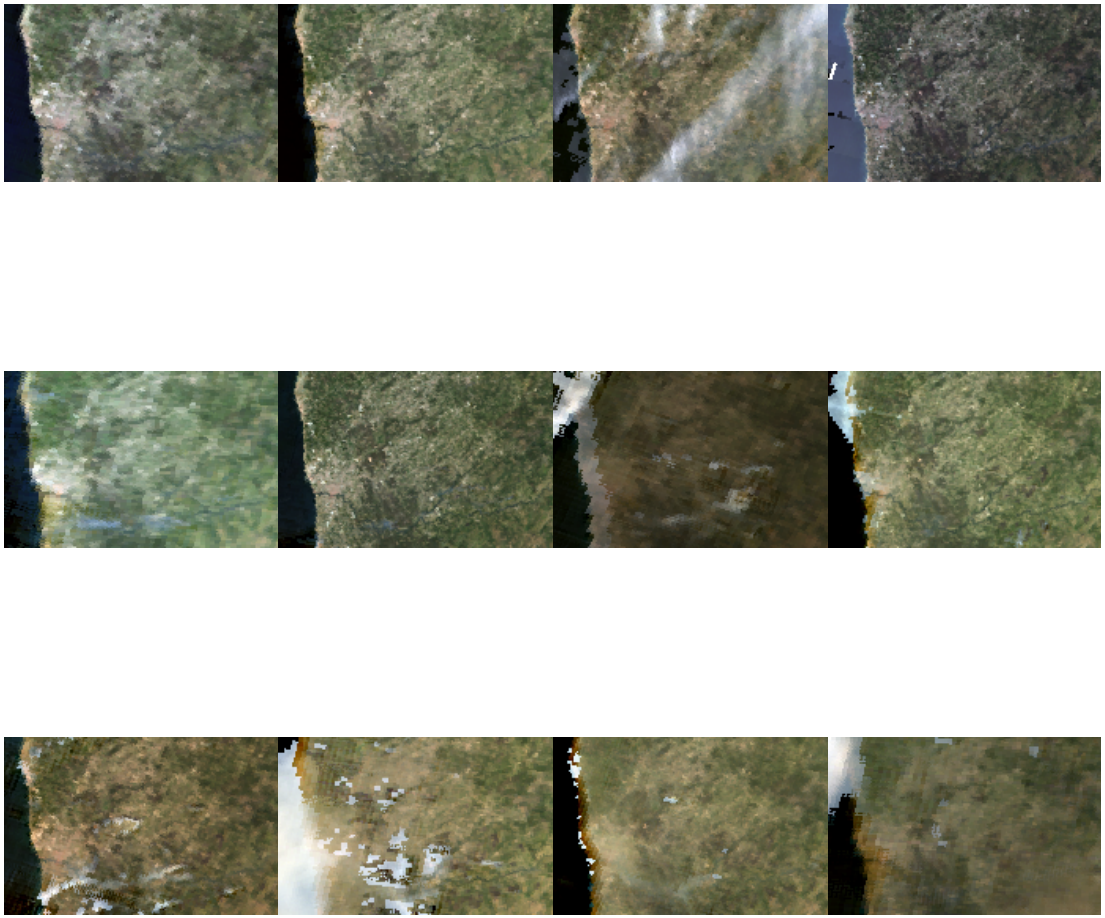
Figure 8: Porto Modis Terra images in August of 2018

IMA (Interpolating mean anomalies) method

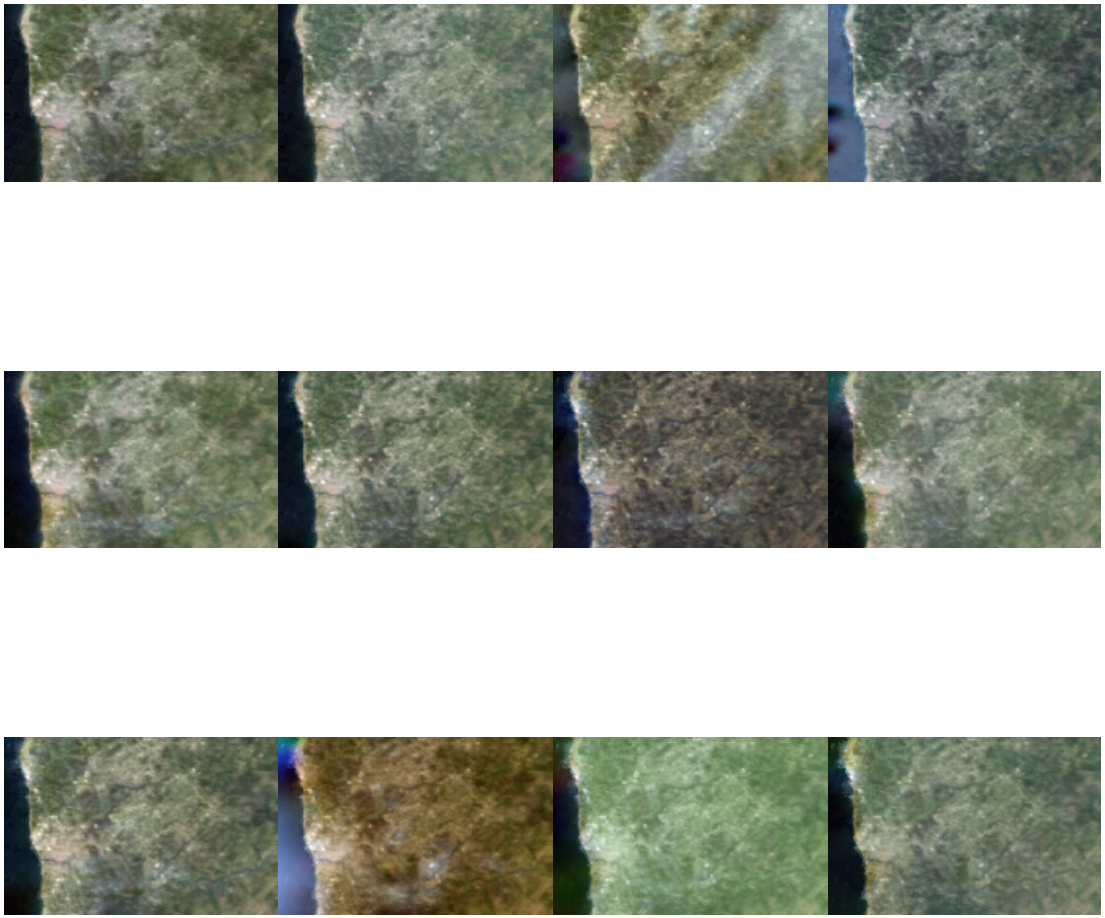
The method consists of:

- 1) defining a neighbourhood for the target image from previous and subsequent images across previous and subsequent years
- 2) computing the mean target image of the neighbourhood
- 3) estimating the anomalies in the target image by subtracting the mean image from the target image
- 4) filtering the anomalies
- 5) averaging the anomalies over a predefined window for reducing the resolution
- 6) interpolating with Tps the averaged anomalies and
- 7) adding the interpolated anomalies to the mean image

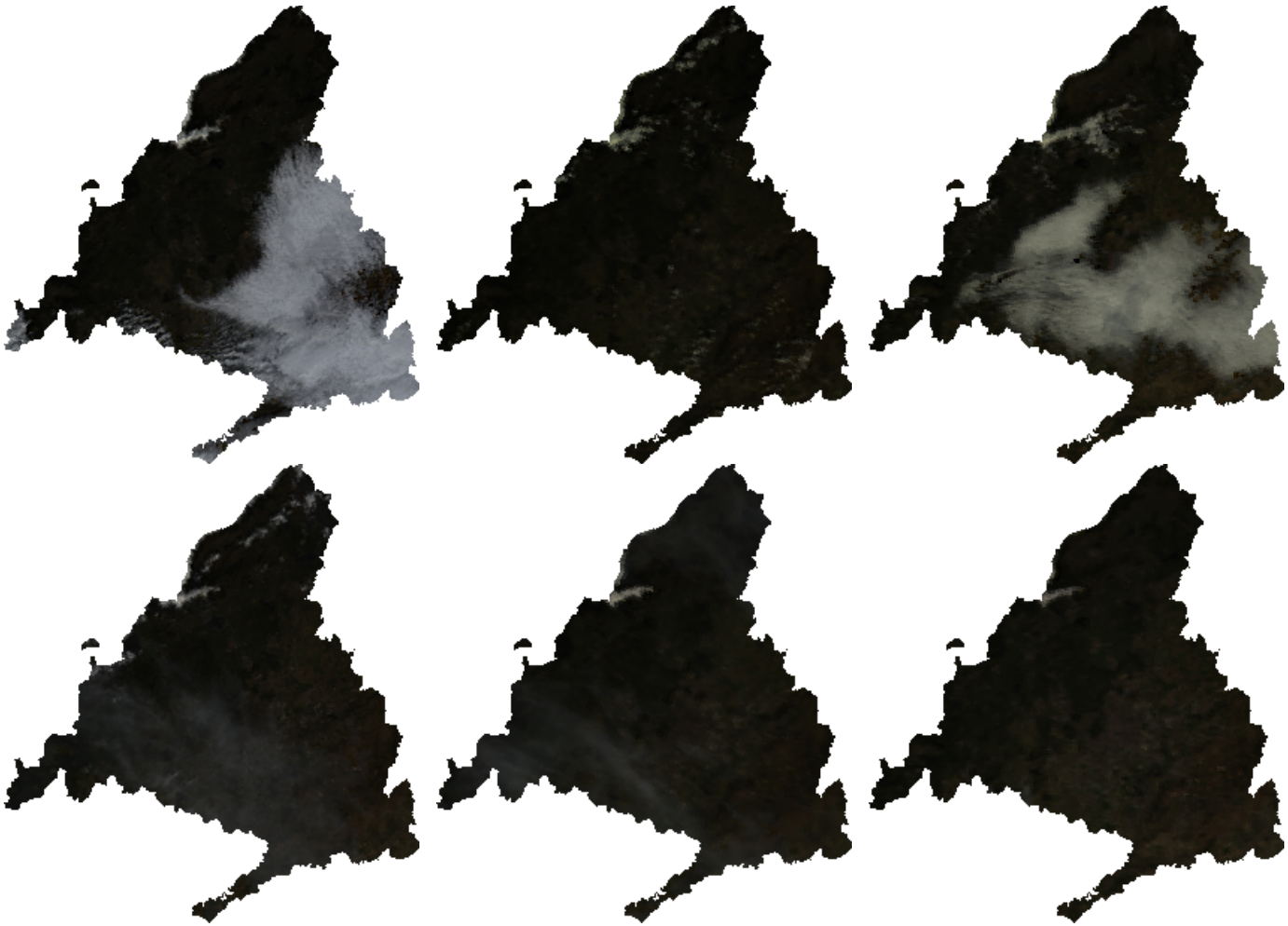
Raw Porto RGB images of August 2016



Smoothed Porto RGB images of August 2016

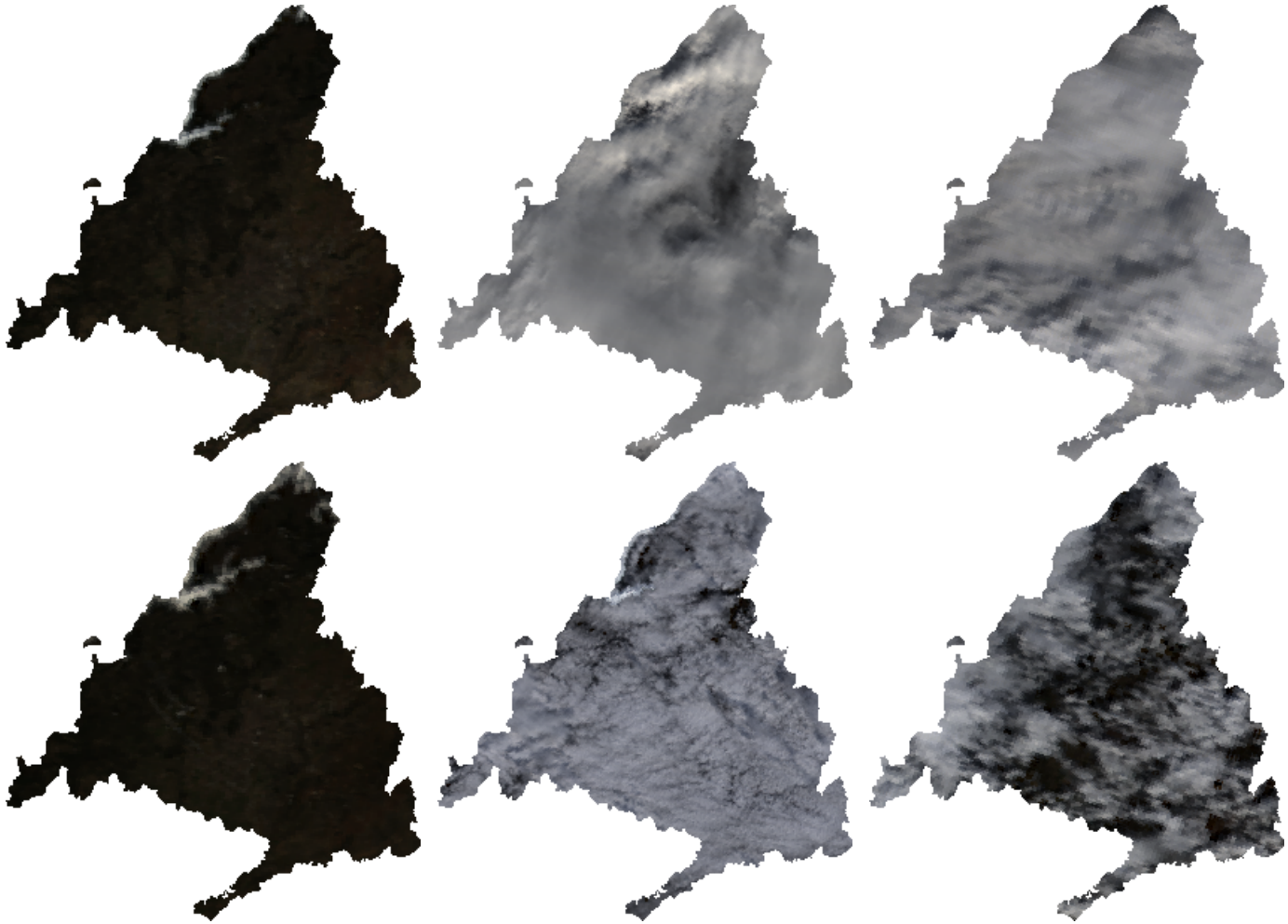


Modis Terra daily images of Madrid in February 2017



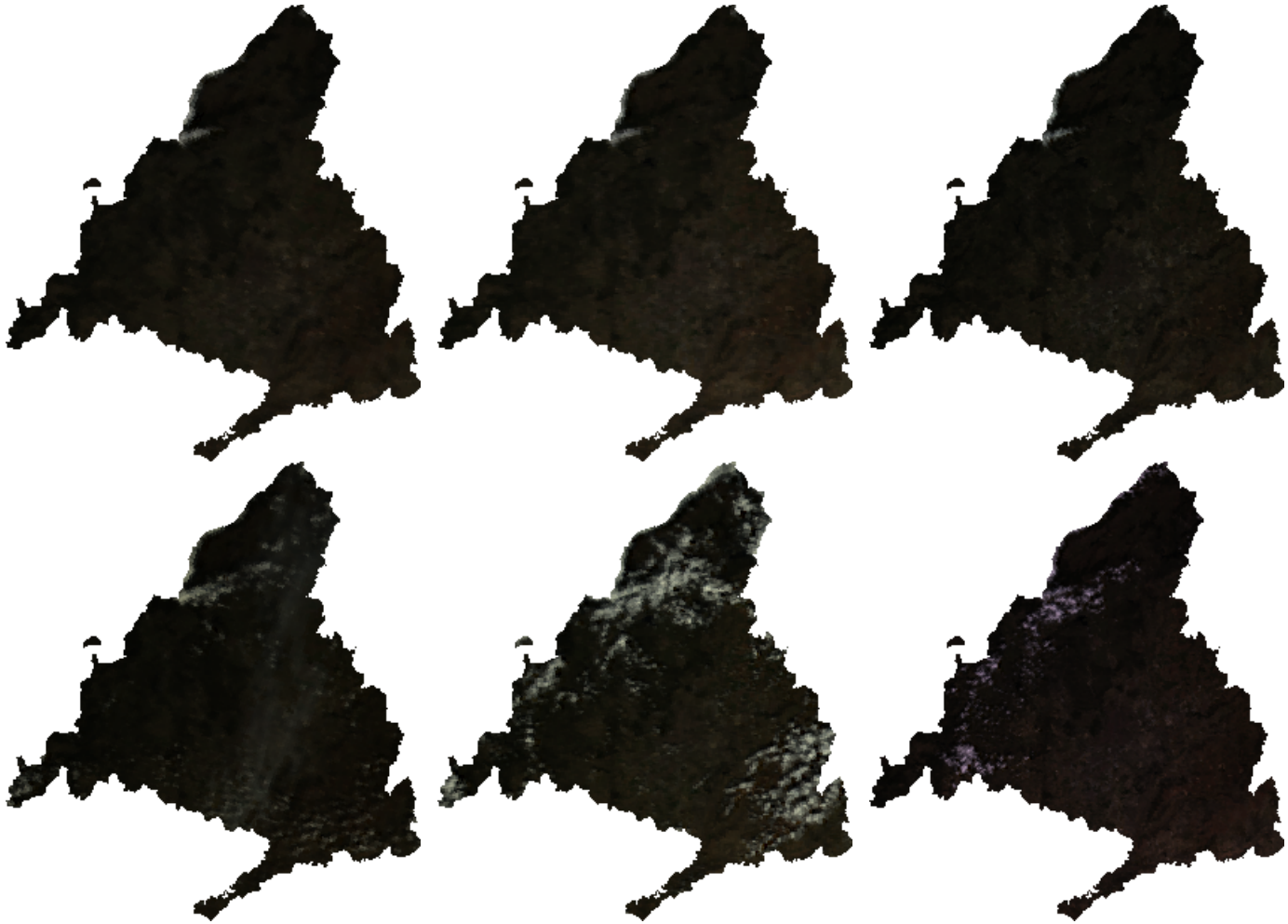
Cloud Mask images of Madrid in February 2017

Modis Terra daily images of Madrid in February 2018



Cloud Mask images of Madrid in February 2018

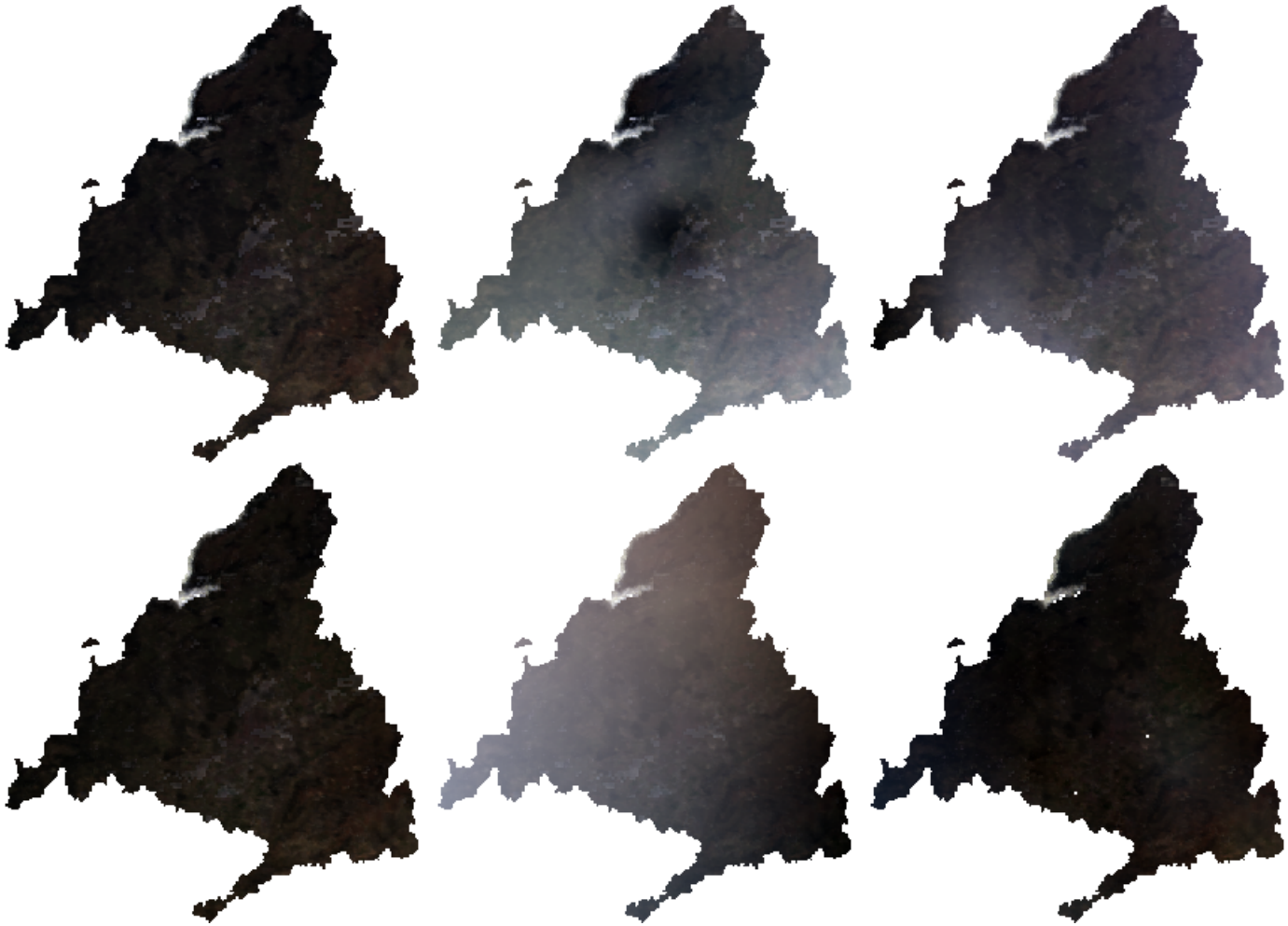
Modis Terra daily images of Madrid in February 2019



Cloud Mask images of Madrid of Madrid in February 2019

Smoothed daily images of Madrid in February 2017

Smoothed daily images of Madrid in February 2018



Smoothed daily images of Madrid in February 2019

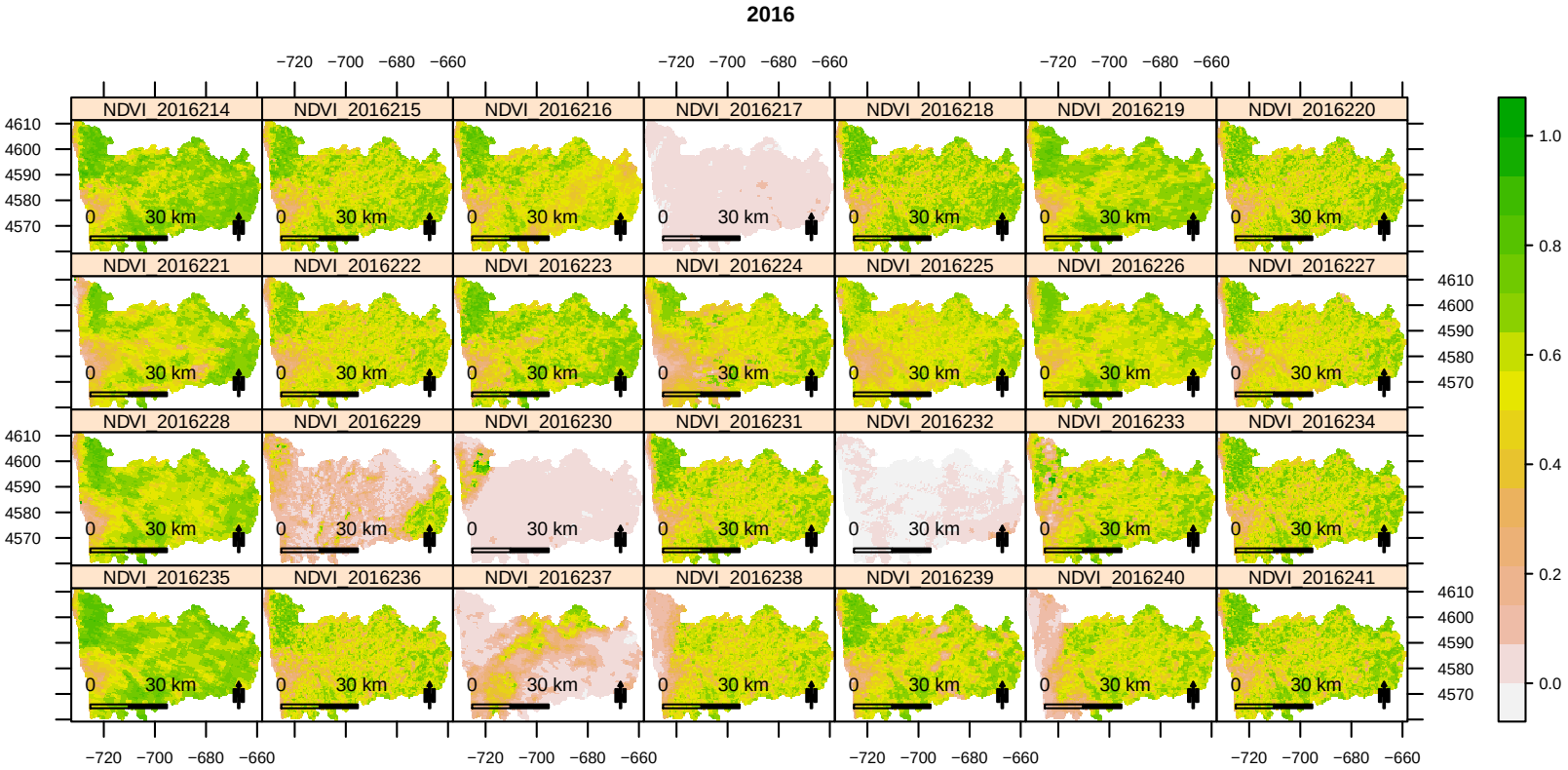
RGISTools commands for Porto application II

1. **modFolderToVar(fun = varNDVI)** Define NDVI images from **varNDVI** function
2. **list.files** and **stack** Define a Raster stack with NDVI images
3. **raster::clamp** Limit the range of NDVI images between -1 and 1
4. **genCompositions** Calculate Maximum Value Composite NDVI images

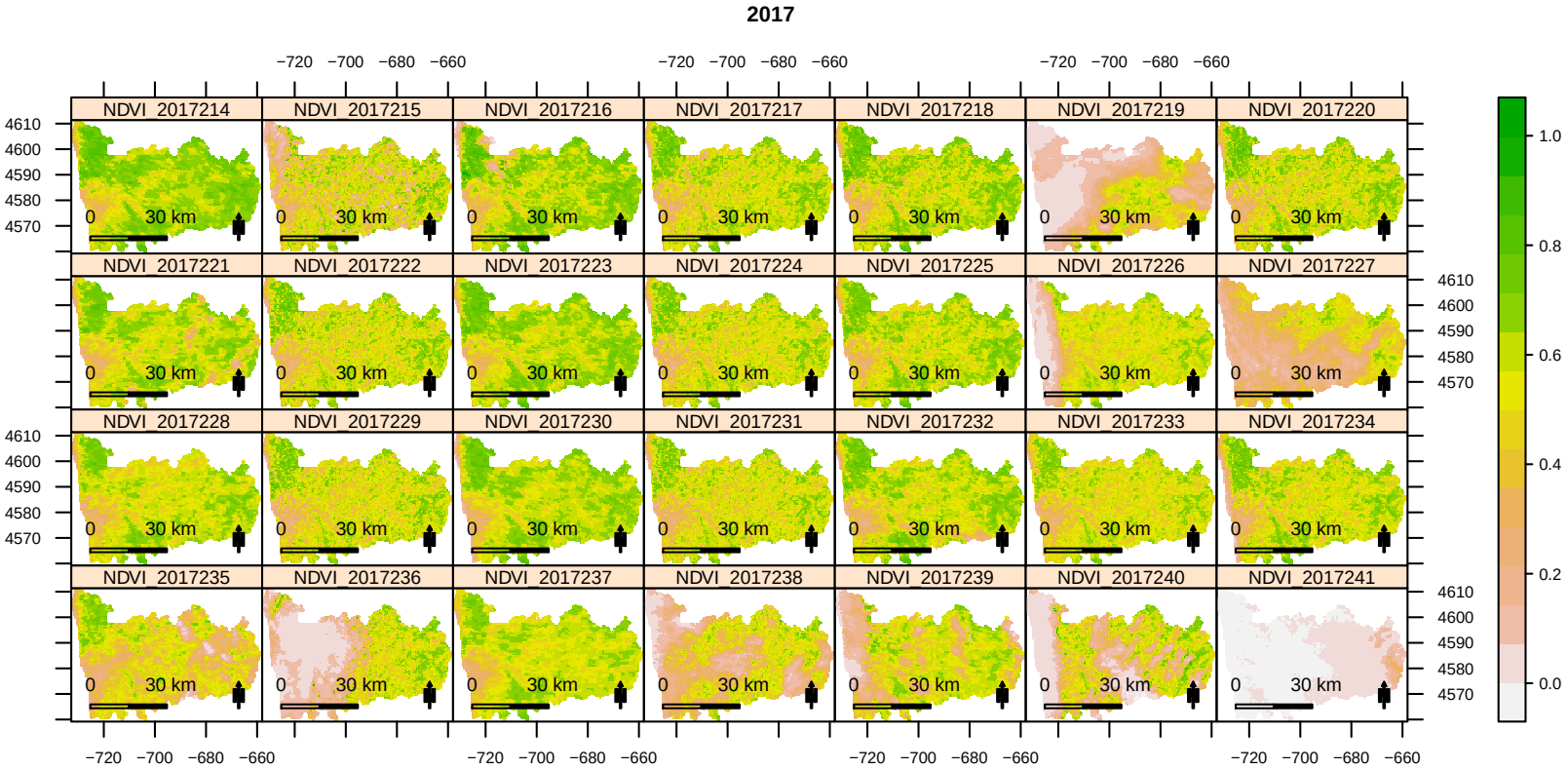
10.-12 **ModFolderToVar** Define NDVI images and creates a Raster Stack

```
> src3 <- file.path(src, "Modis","MOD09GA","Variables")
> src4 <- file.path(src, "Modis","MOD09GA","Oporto")
> modFolderToVar(src4,
                 fun = varNDVI,
                 AppRoot = src3,
                 overwrite = F)
> src3.ndvi<-file.path(src3,"ndvi")
> tif.ndvi <- list.files(src3.ndvi,
                         full.names = TRUE,
                         recursive=TRUE,
                         pattern ="\\.tif")
> porto.ndvi<-stack(tif.ndvi)
> porto.ndvi.limit<-raster::clamp(porto.ndvi1618,mn=-1,mx=1)
> porto.ndvi.mask<-mask(porto.ndvi.limit,porto.sinu)
```

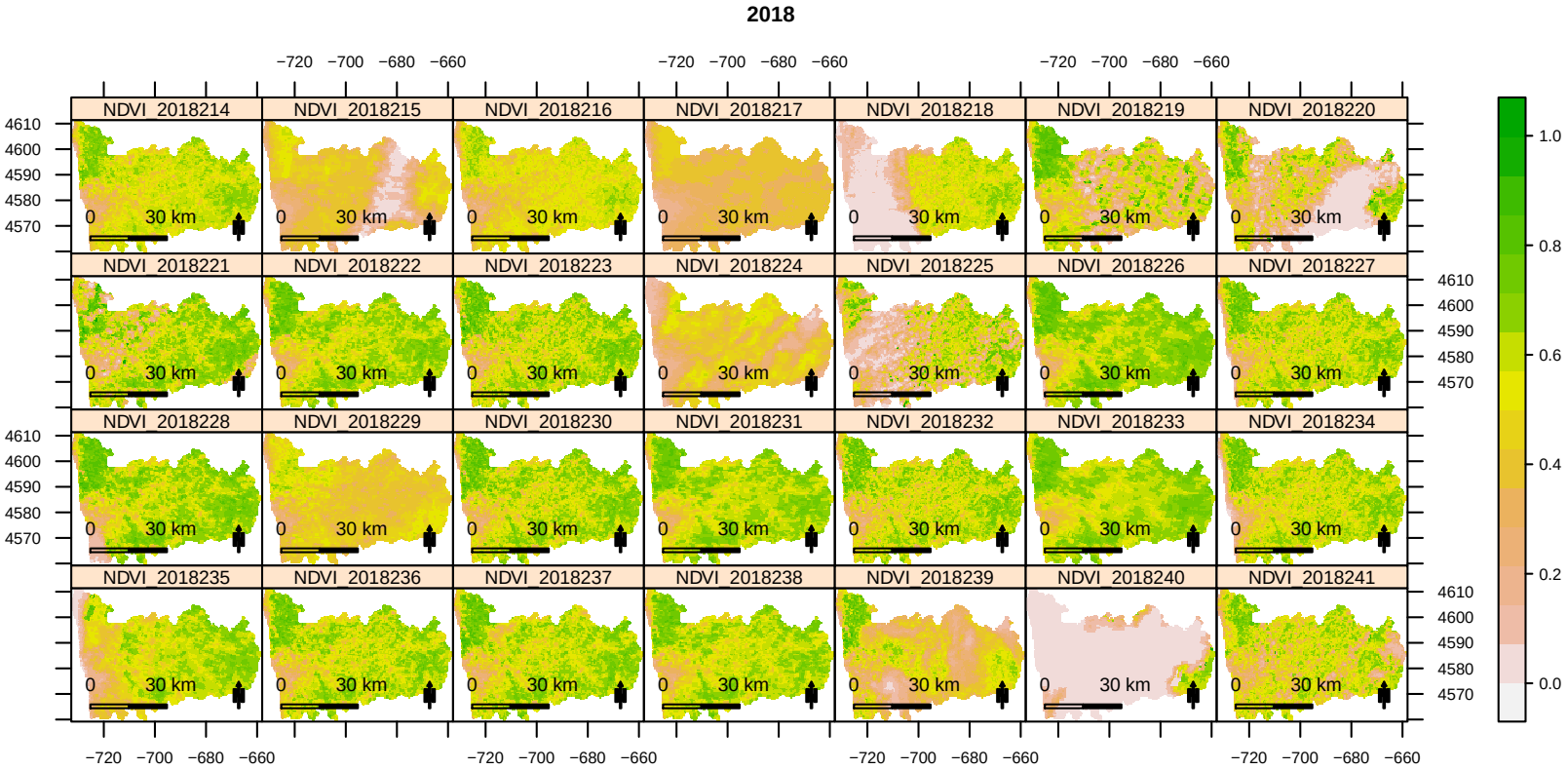
Modis NDVI daily images of Porto District in August of 2016



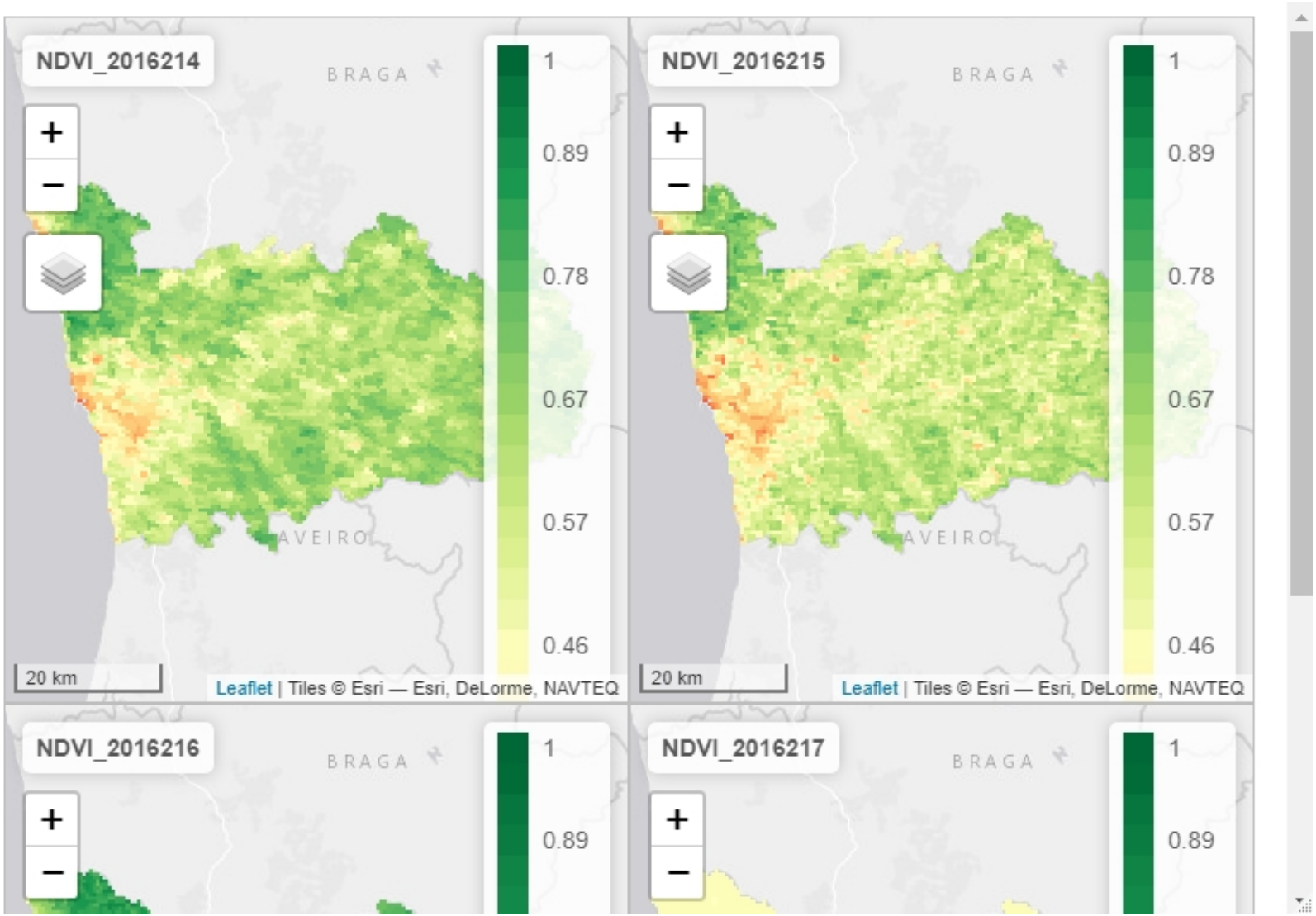
Modis NDVI daily images of Porto District in August of 2017



Modis NDVI daily images of Porto District in August of 2018



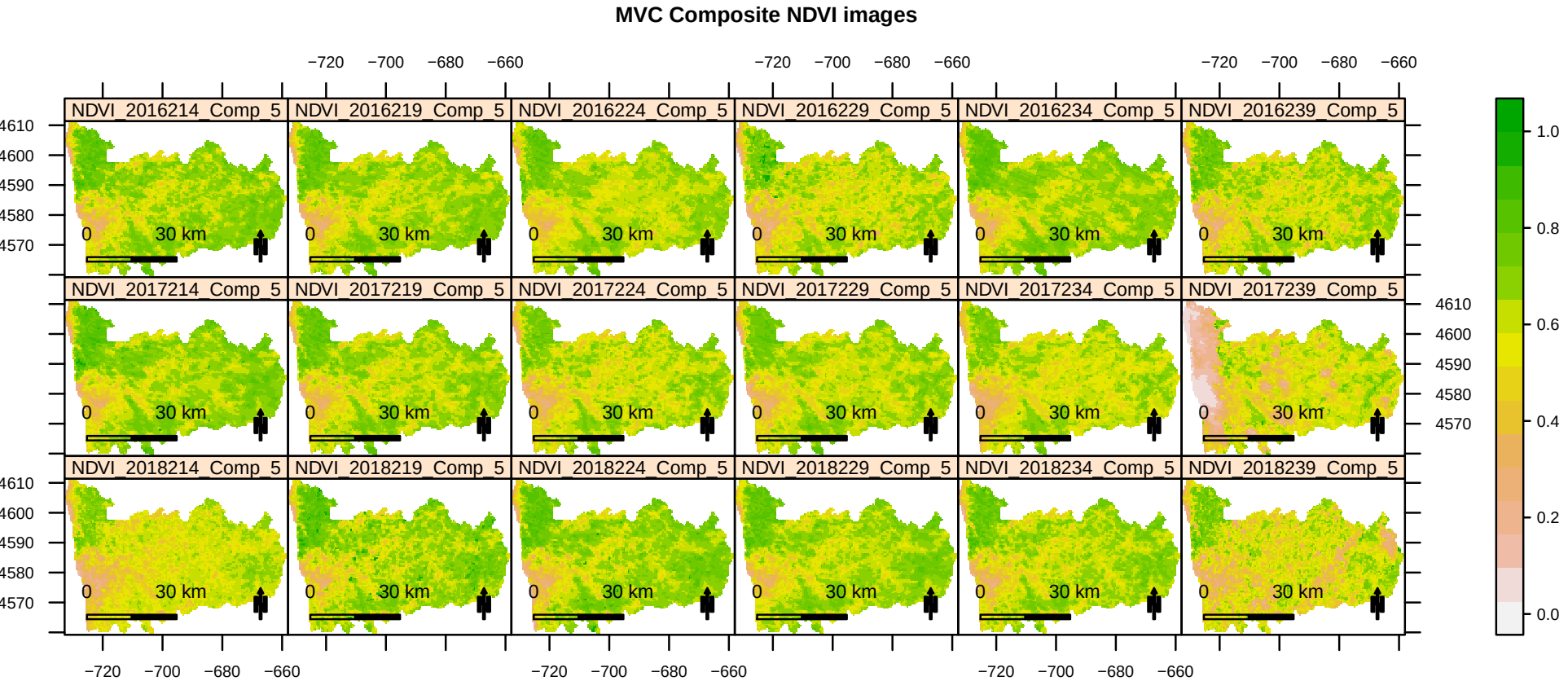
Modis NDVI daily images of Porto District



13. **genCompositions** for MVC Composite Images

```
> porto.ndvi.comp1<-genCompositions(porto.ndvi.mask[[1:31]],
                      n=5,fun=max)
> porto.ndvi.comp2<-genCompositions(porto.ndvi.mask[[32:62]],
                      n=5,fun=max)
> porto.ndvi.comp3<-genCompositions(porto.ndvi.mask[[63:93]],
                     n=5,fun=max)
> porto.ndvi.comp<-stack(porto.ndvi.comp1,porto.ndvi.comp2,
                      porto.ndvi.comp3)
> porto.ndvi
class       : RasterStack
dimensions  : 113, 175, 19775, 93  (nrow, ncol, ncell, nlayers)
resolution  : 464.0102, 462.3513  (x, y)
extent      : -737373.1, -656171.3, 4559151, 4611396  (xmin, xmax, yr
crs         : +proj=sinu +lon_0=0 +x_0=0 +y_0=0 +a=6371007.181 +b=637
names       :  NDVI_2016214,  NDVI_2016215,  NDVI_2016216,  NDVI_2016
min values  : -1.900000e+01, -3.500000e+01, -2.700000e+01, -2.341305e
max values  :    23.0000000,    57.0000000,    57.0000000,     0.3926
```

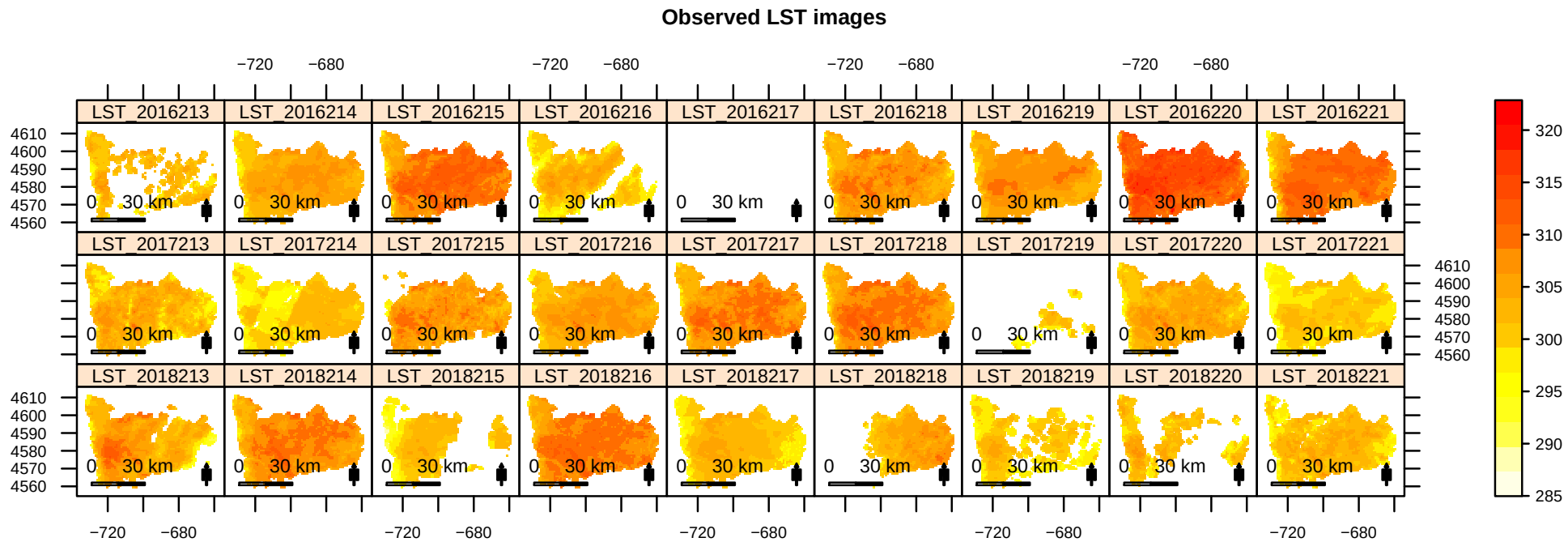
Composite NDVI Modis images of Porto District in August 2016, 2017 and 2018



RGISTools commands for Porto Application III: LST images

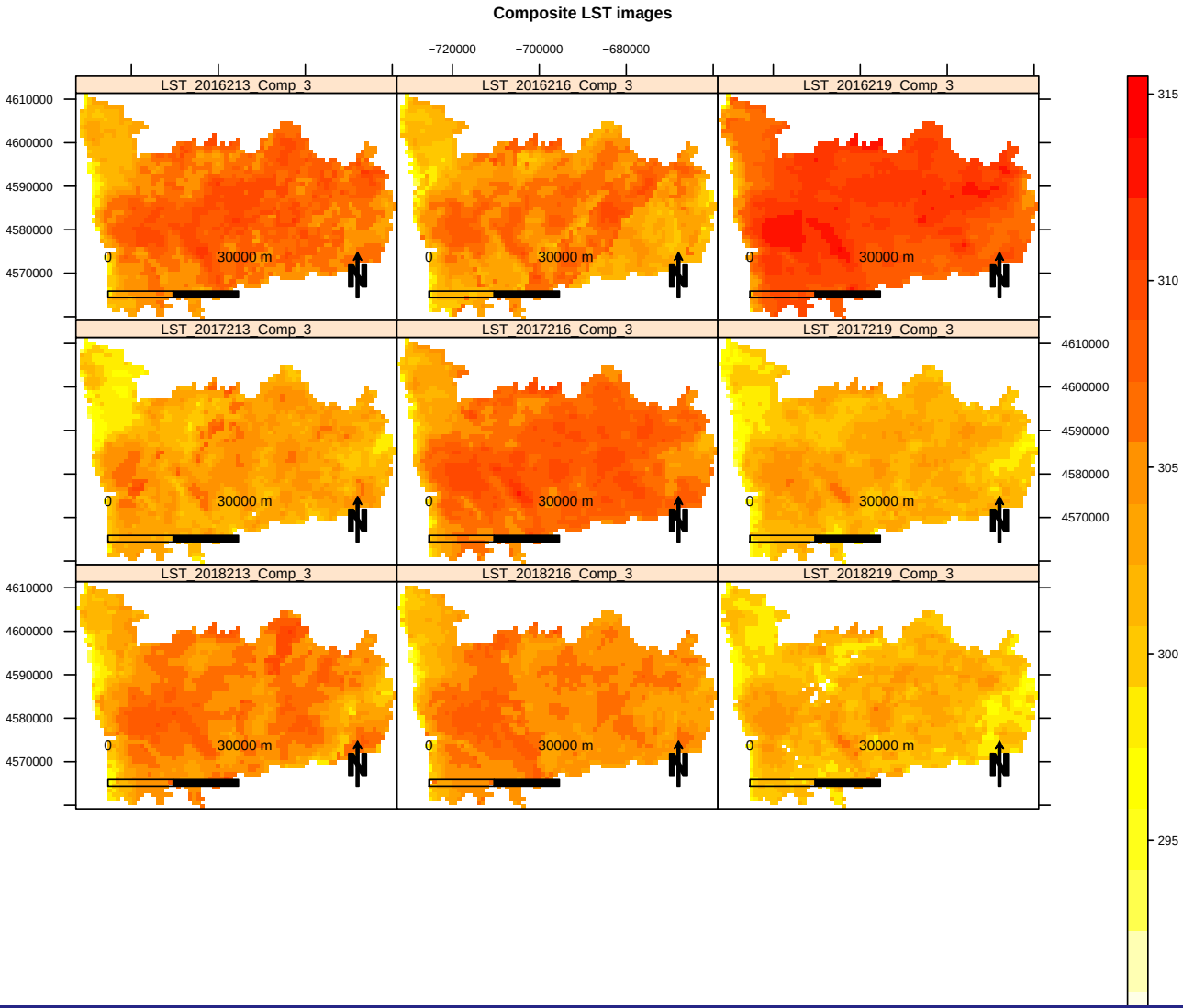
1. **getData** Upload "Porto" shape file (from raster package)
2. **modSearch** Search for August images intersecting "Porto" District
3. **modDownSearch** Download Modis satellite images, collection 6 of MOD09GA
4. **modExtractHDF** Extract Gtiff files from downloaded hdf files
5. **modMosaic** Merge images containing Porto (if necessary)
6. **spTransform** Transforms shp long-lat porto file to sinusoidal projection
7. **mask** Mask LST images with porto shape file in sinusoidal projection
8. **genCompositing** Composite LST images
9. **genSmoothingIMA** Smoothing raw LST images with IMA

Porto - Observed LST from the 1st to the 9th of August 2016, 2017 and 2018



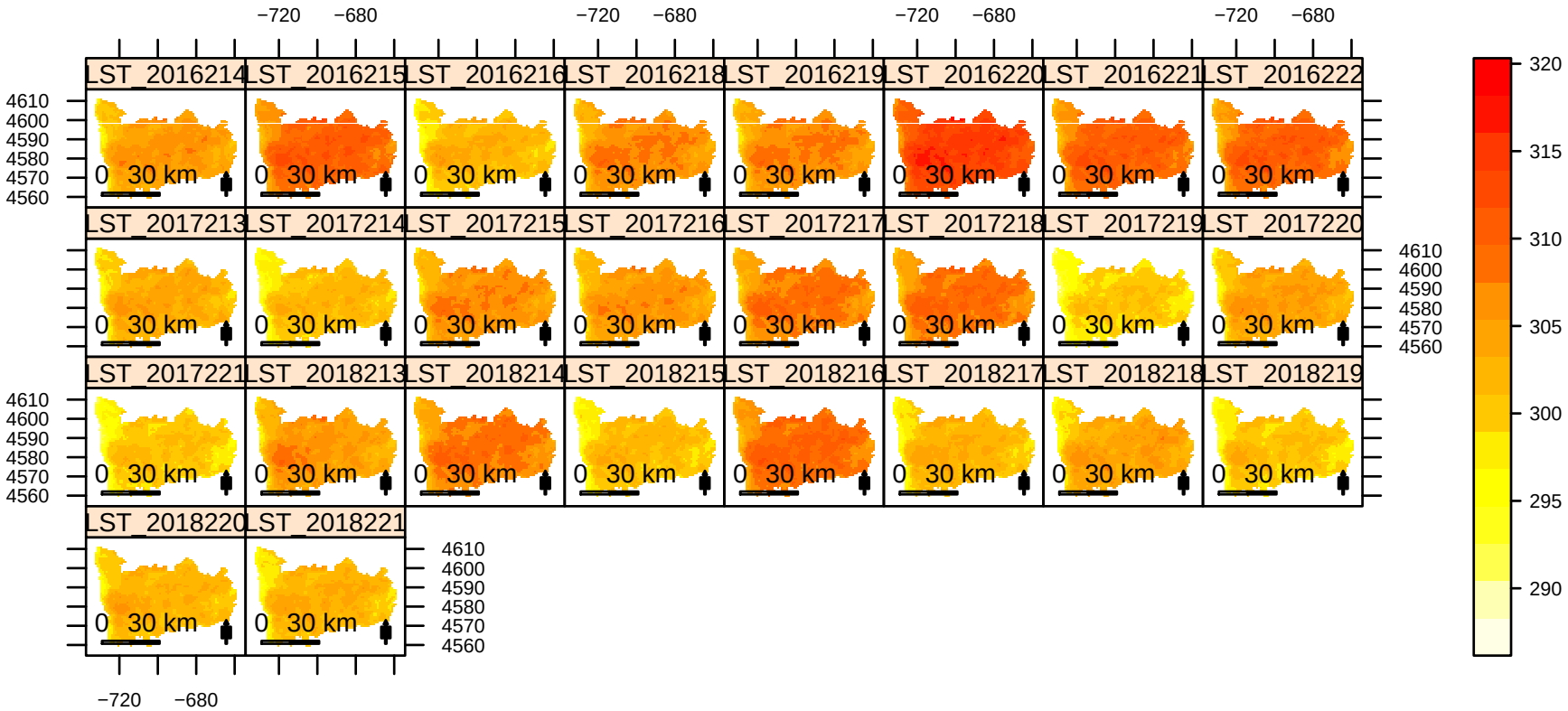
Porto Composite LST in August 2016, 2017 and 2018

Compositing reduces the number of images

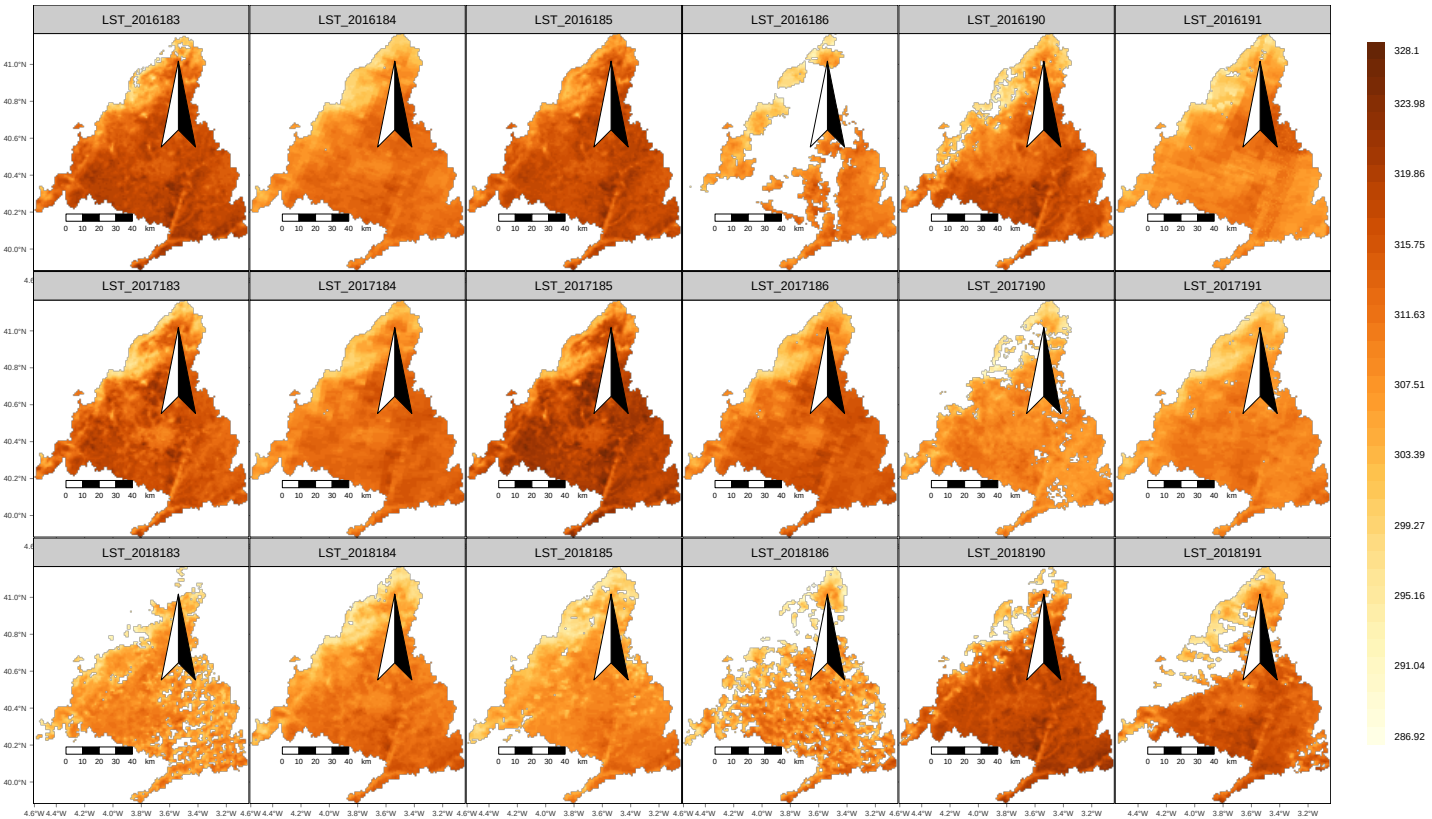


Smoothed Land Surface Temperature (LST) from the 1st to the 9th of August 2016, 2017 and 2018

IMA method does not reduce the number of images when smoothing

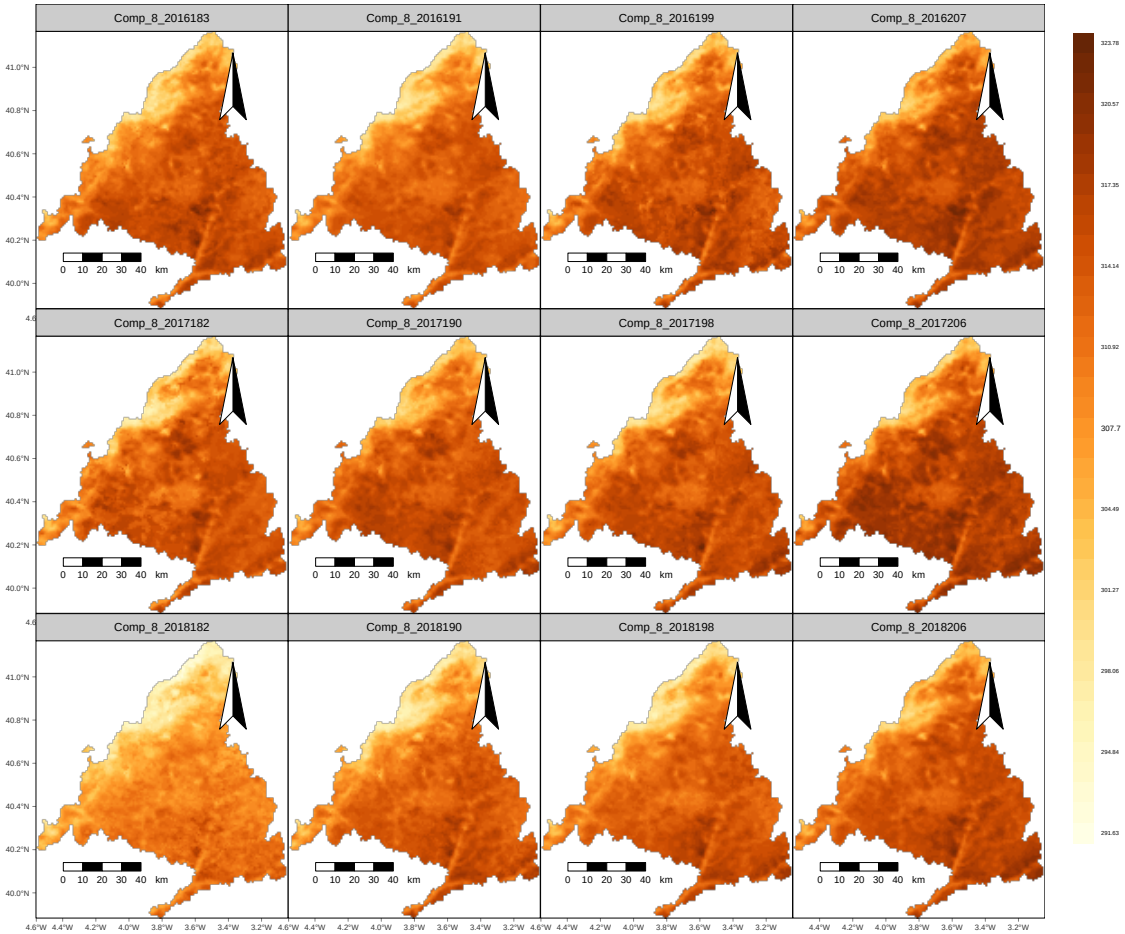


Madrid Observed LST. August 2016, 2017 and 2018



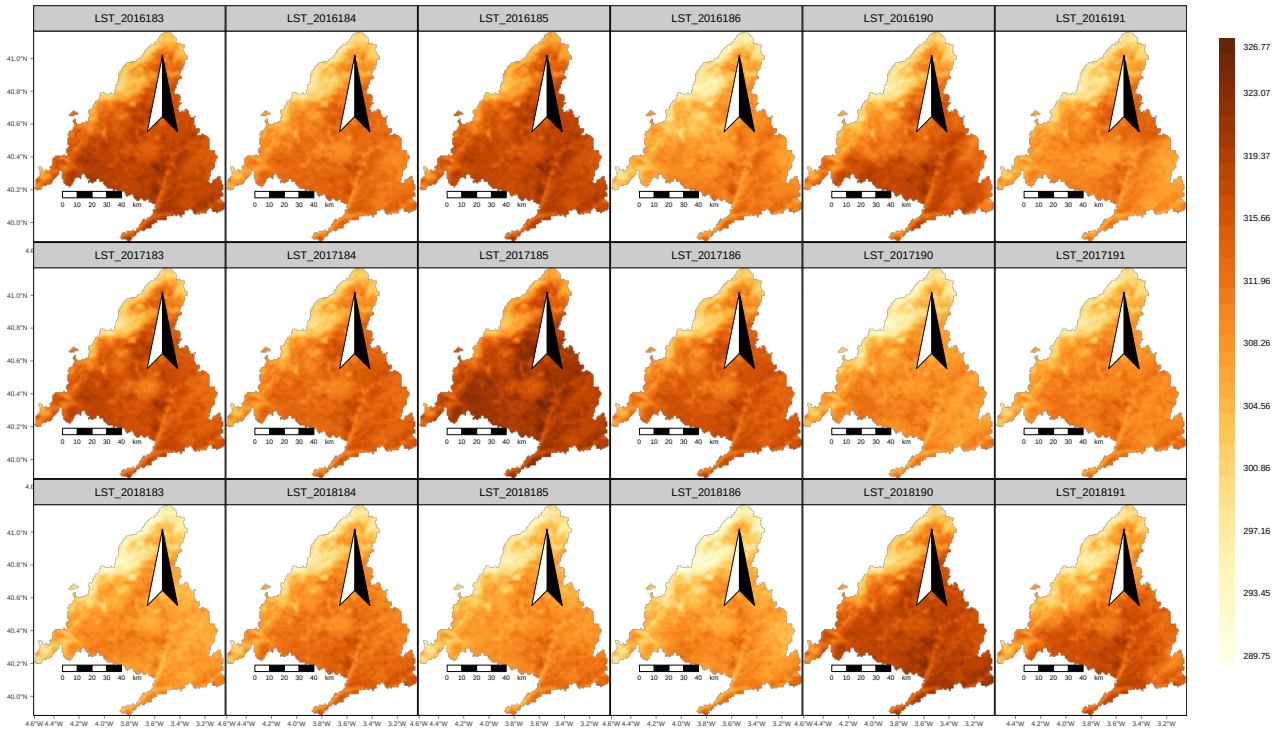
Madrid Composite LST in August 2016, 2017 and 2018

Compositing reduces the number of images

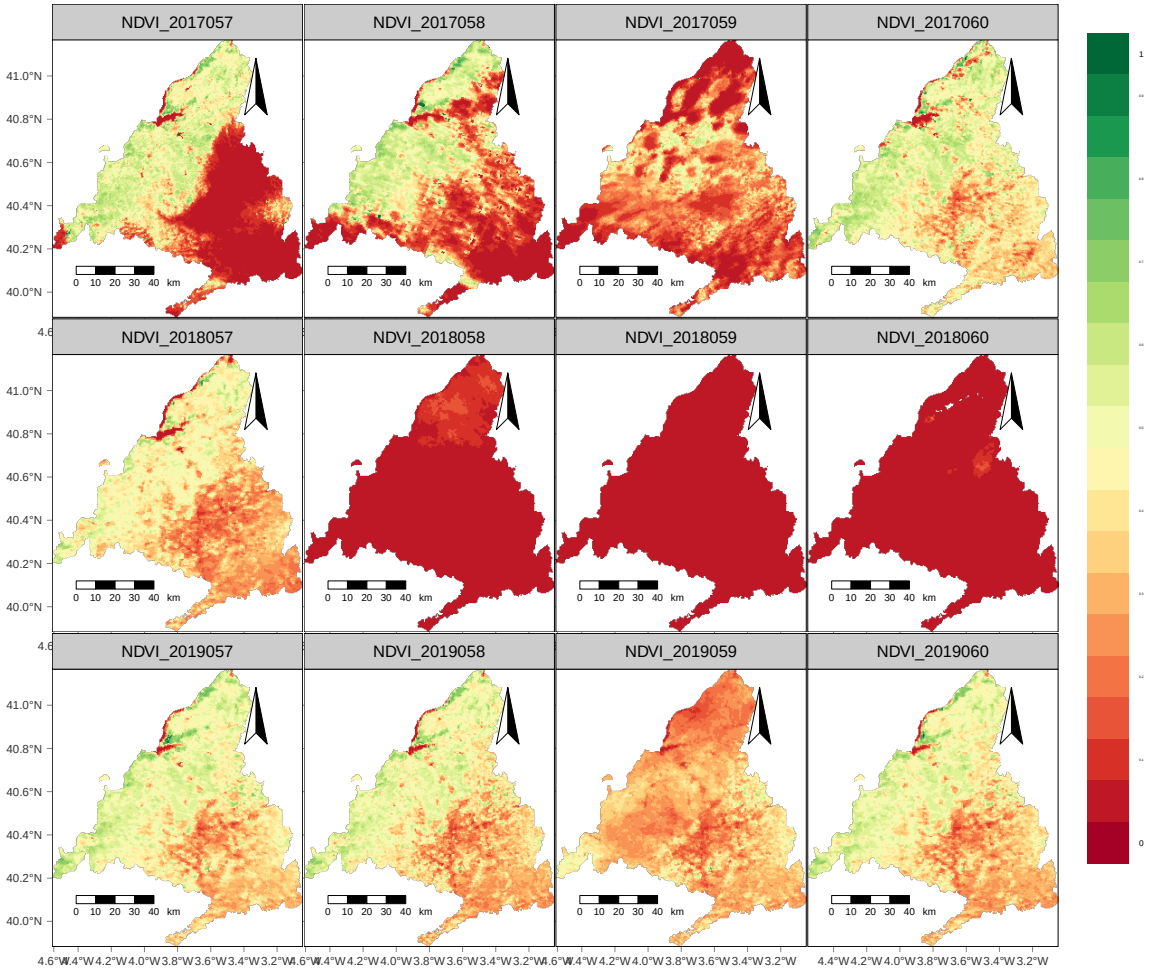


Madrid Smoothed LST in August 2016, 2017 and 2018

IMA method does not reduce the number of images when smoothing

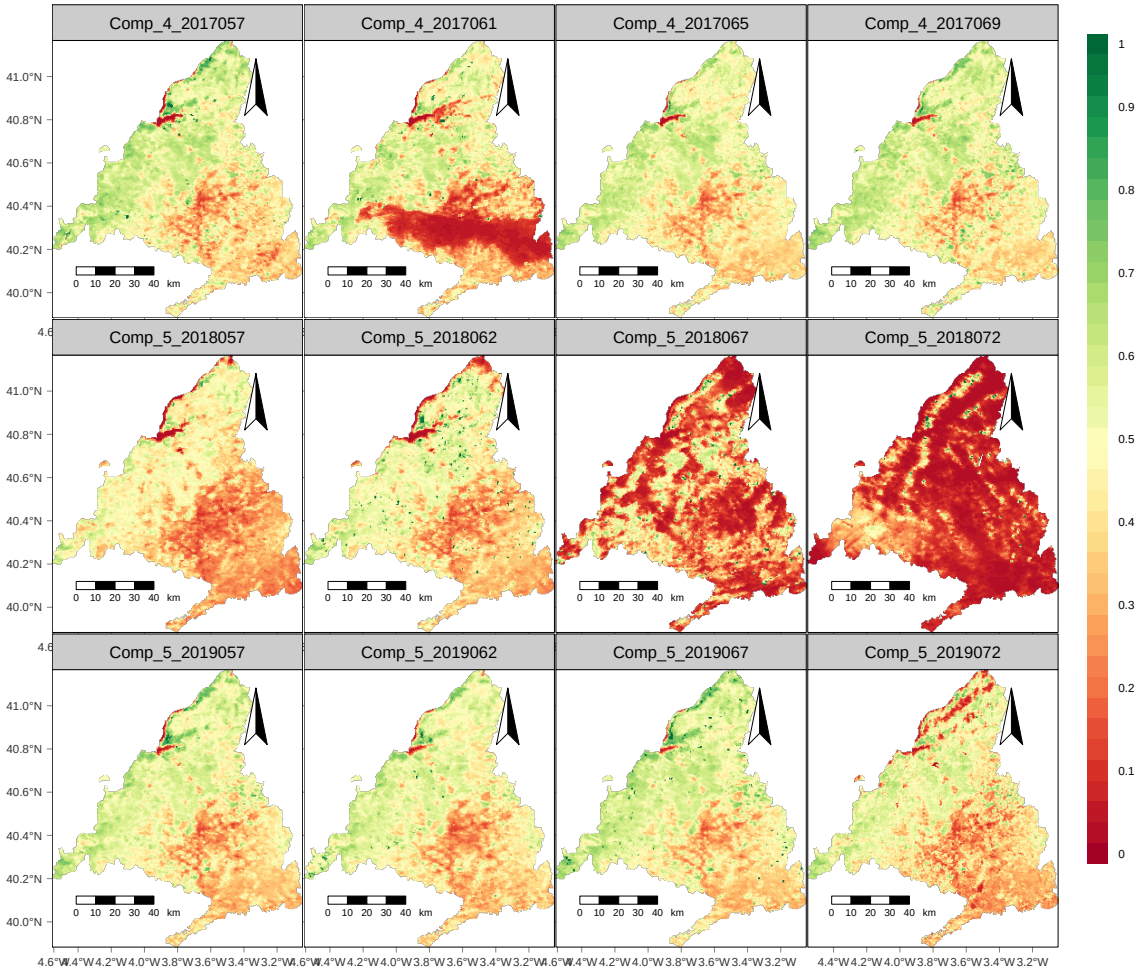


Madrid Observed NDVI. August 2016, 2017 and 2018



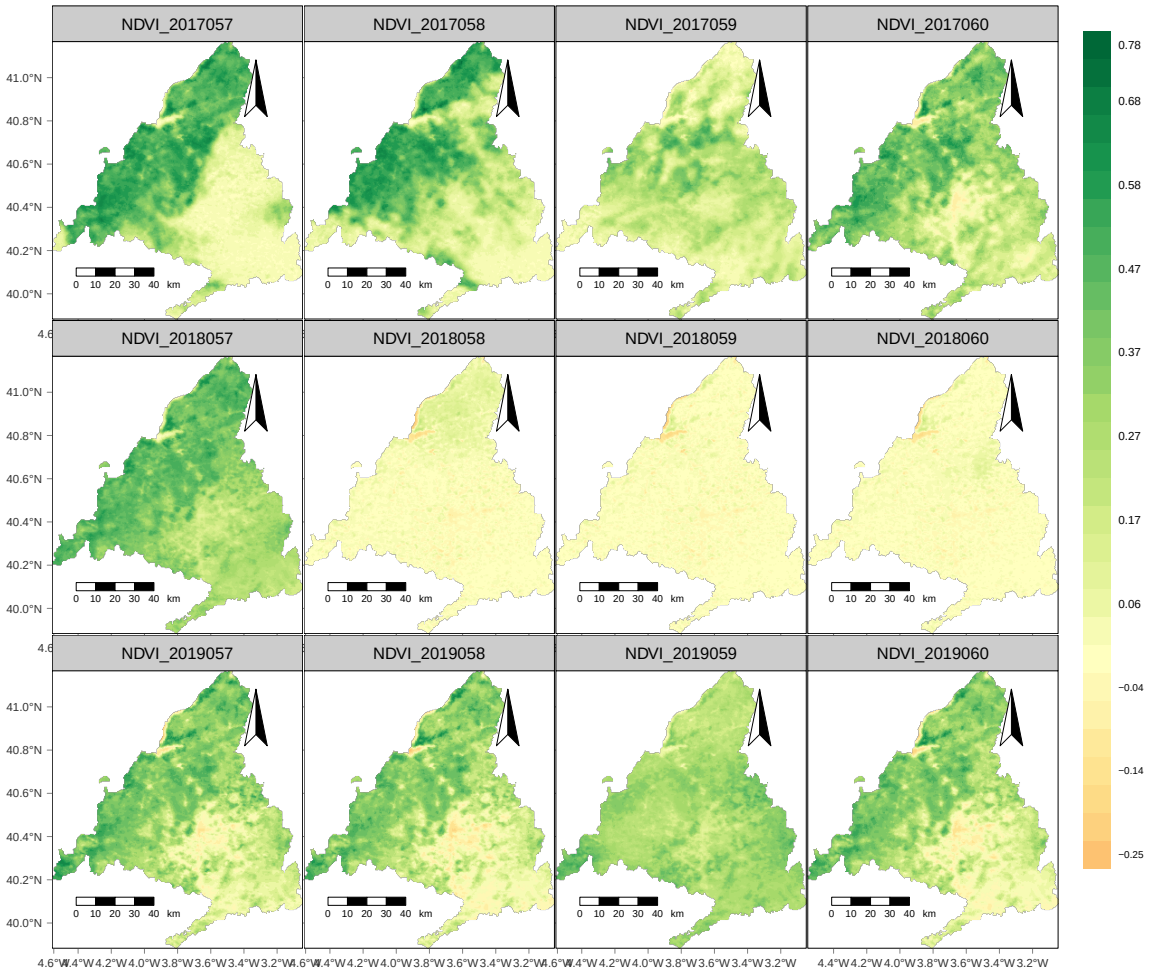
Madrid Composite NDVI in August 2016, 2017 and 2018

Compositing reduces the number of images



Madrid Smoothed NDVI in August 2016, 2017 and 2018

IMA method does not reduce the number of images when smoothing



Conclusions

- Statisticians should know the capabilities of the new resources for accessing and managing data coming from satellites and sensors
- "RGISTools" provides an easy way of acquiring satellite imagery for being used with many statistical procedures
- It can be downloaded from CRAN https://cloud.r-project.org/ or from github, running in Rstudio

```
> library(devtools)
> install_github("spatialstatisticsupna/RGISTools")
```

References I

[1] A. F. Militino, M. D. Ugarte, and U. Pérez-Goya, *An Introduction to the Spatio-Temporal Analysis of Satellite Remote Sensing Data for Geostatisticians*, pp. 239–253.
Cham: Springer International Publishing, 2018.

[2] J. Aschbacher and M. P. Milagro-Pérez, "The european earth monitoring (gmes) programme: Status and perspectives," *Remote Sensing of Environment*, vol. 120, pp. 3–8, 2012.

[3] J. Rouse Jr, R. Haas, J. Schell, and D. Deering, "Monitoring vegetation systems in the great plains with erts," *NASA special publication*, vol. 351, p. 309, 1974.

[4] J. A. Sobrino, J. C. Jiménez-Muñoz, and L. Paolini, "Land surface temperature retrieval from landsat tm 5," *Remote Sensing of environment*, vol. 90, no. 4, pp. 434–440, 2004.

References II

[5] I. L. Hudson and M. R. Keatley, *Phenological research: methods for environmental and climate change analysis*.
Springer Science & Business Media, 2009.

[6] A. F. Militino, M. D. Ugarte, U. Pérez-Goya, and M. Genton, "Interpolation of the Mean Anomalies for Cloud Filling in Land Surface Temperature and Normalized Difference Vegetation Index," *IEEE Transactions on Geoscience and Remote Sensing (Open Access)*, vol. 57, no. 8, pp. 6068–6078, 2019.

[7] A. F. Militino, M. D. Ugarte, and U. Pérez-Goya, "Improving the Quality of Satellite Imagery Based on Ground-Truth Data from Rain Gauge Stations," *Remote Sensing*, vol. 10, no. 3, p. 398, 2018.

References III

[8] U. Pérez-Goya, M. Montesino-SanMartin, A. F. Militino, and M. D. Ugarte, *RGISTools: Handling Multiplatform Satellite Images*, 2020. R package version 1.0.0.