

# Geospatial Analysis in R

## Working with sf and geospatial data

### Setup

This document is for showing the different things you can do with geospatial data in R.

```
sf::sf_use_s2(FALSE) # We'll need to do this since the data is planar

## Spherical geometry (s2) switched off
library(sf) # This is a major existing R library for managing geospatial data

## Linking to GEOS 3.10.2, GDAL 3.4.3, PROJ 9.0.0; sf_use_s2() is FALSE

p <- paste( # This is the path to the geospatial data
  "/media/thellinger/cs3pace1/data/choctawatchee_bay/",
  "choctawatchee_bay_lssm.geojson",
  sep=""
) # To keep this under 80 characters to show up in the PDF
gsobj <- read_sf(p)
print(gsobj)

## Simple feature collection with 2982 features and 31 fields
## Geometry type: MULTILINESTRING
## Dimension: XY
## Bounding box: xmin: -86.63939 ymin: 30.36933 xmax: -86.09757 ymax: 30.5213
## Geodetic CRS: WGS 84
## # A tibble: 2,982 x 32
##   OBJECTID ID bnk_height Beach WideBeach MxQExpCode bathymetry roads
##   <int> <int> <chr> <chr> <chr> <chr> <chr> <chr>
## 1 1 1 0 0-5 No No Low Shallow No
## 2 2 2 0 0-5 Yes Yes Moderate Shallow No
## 3 3 3 0 0-5 Yes No Moderate Shallow No
## 4 4 4 0 0-5 No No Moderate Shallow No
## 5 5 5 0 0-5 Yes No Low Shallow No
## 6 6 6 0 0-5 No No Low Shallow No
## 7 7 7 0 0-5 No No Low Shallow Roads
## 8 8 8 0 0-5 No No Low Shallow No
## 9 9 9 0 0-5 No No Low Shallow No
## 10 10 10 0 0-5 Yes Yes Moderate Shallow No
## # i 2,972 more rows
## # i 24 more variables: PermStruc <chr>, PublicRamp <chr>, RiparianLU <chr>,
## # canal <chr>, SandSpit <chr>, Structure <chr>, offshorest <chr>, SAV <chr>,
## # marsh_all <chr>, tribs <chr>, Needs_QC <chr>, defended <chr>,
## # rd_pstruc <chr>, lowBnkStrc <chr>, ShlType <chr>, Fetch_ <chr>,
## # selectThis <chr>, StrucList <chr>, SMMv5Class <chr>, bmpCountv5 <int>,
## # SMMv5Def <chr>, DefDate <chr>, Shape__Length <dbl>, ...
```

## Coordinate Reference Systems

So now we have the geospatial data loaded into R. In this case, we are looking at the data from Choctawatchee Bay. As is common for LSSM models, this dataset uses a MULTILINE geometry as we are classifying stretches of coast. Now let's see some of the functions we can use to work with this data.

```
# This will give us the BBox or Bounding Box, the lon/lat rectangle which covers  
# the extent of all the observations in the dataset  
st_bbox(gsobj)
```

```
##      xmin      ymin      xmax      ymax  
## -86.63939  30.36933 -86.09757  30.52130
```

Next we'll look at the CRS - Coordinate Reference Systems for the data. This is a pretty important topic that Leah Wasser has a great lesson for on the website [Earthdatascience.org](http://Earthdatascience.org)

```
st_crs(gsobj)
```

```
## Coordinate Reference System:  
##   User input: WGS 84  
##   wkt:  
##   GEOGCRS["WGS 84",  
##     DATUM["World Geodetic System 1984",  
##       ELLIPSOID["WGS 84",6378137,298.257223563,  
##         LENGTHUNIT["metre",1]],  
##     PRIMEM["Greenwich",0,  
##       ANGLEUNIT["degree",0.0174532925199433]],  
##     CS[ellipsoidal,2],  
##       AXIS["geodetic latitude (Lat)",north,  
##         ORDER[1],  
##         ANGLEUNIT["degree",0.0174532925199433]],  
##       AXIS["geodetic longitude (Lon)",east,  
##         ORDER[2],  
##         ANGLEUNIT["degree",0.0174532925199433]],  
##     ID["EPSG",4326]]
```

So this dataset is using EPSG 4326. This is the same CRS used by Google Earth and OpenStreet. It's important that all the objects that you're working with are of the same CRS, or else the projections will not be aligned and any calculations made between those objects will be wrong.

Let's look at how to do things by looking at Joe's Bayou. Here are the coordinates I pulled from QGIS to form a bounding box:

Longitude	Latitude
-86.50168	30.42225
-86.47472	30.42225
-86.47472	30.42225
-86.47472	30.39530
-86.47472	30.39530
-86.50168	30.39530
-86.50168	30.39530
-86.50168	30.42225

```
bbox_matrix <- matrix(  
  c(-86.50168, 30.42225, -86.47472, 30.42225, -86.47472, 30.42225,  
    -86.47472, 30.39530, -86.47472, 30.39530, -86.50168, 30.39530,
```

```

    -86.50168, 30.39530, -86.50168, 30.42225),
    ncol=2,
    byrow=TRUE
)

joes_bayou_bbox <- st_polygon(list(bbox_matrix))
joes_bayou <- subset(gsobj, st_within(gsobj, joes_bayou_bbox, sparse = FALSE))

## although coordinates are longitude/latitude, st_within assumes that they are
## planar

print(joes_bayou)

## Simple feature collection with 74 features and 31 fields
## Geometry type: MULTILINESTRING
## Dimension: XY
## Bounding box: xmin: -86.50145 ymin: 30.40251 xmax: -86.47477 ymax: 30.41767
## Geodetic CRS: WGS 84
## # A tibble: 74 x 32
##   OBJECTID ID bnk_height Beach WideBeach MxQExpCode bathymetry roads
##   <int> <int> <chr> <chr> <chr> <chr> <chr> <chr>
## 1 641 0 0-5 No No Low Shallow Roads
## 2 642 0 0-5 No No Moderate Shallow No
## 3 643 0 0-5 Yes Yes Low Shallow No
## 4 644 0 0-5 No No High Shallow No
## 5 645 0 0-5 No No Low Shallow No
## 6 936 0 0-5 No No Low Shallow No
## 7 937 0 0-5 No No Low Shallow No
## 8 938 0 0-5 No No Moderate Shallow No
## 9 939 0 0-5 No No High Shallow No
## 10 940 0 0-5 No No Moderate Shallow No
## # i 64 more rows
## # i 24 more variables: PermStruc <chr>, PublicRamp <chr>, RiparianLU <chr>,
## # canal <chr>, SandSpit <chr>, Structure <chr>, offshorest <chr>, SAV <chr>,
## # marsh_all <chr>, tribs <chr>, Needs_QC <chr>, defended <chr>,
## # rd_pstruc <chr>, lowBnkStrc <chr>, ShlType <chr>, Fetch_ <chr>,
## # selectThis <chr>, StrucList <chr>, SMMv5Class <chr>, bmpCountv5 <int>,
## # SMMv5Def <chr>, DefDate <chr>, Shape__Length <dbl>, ...

plot(joes_bayou['SMMv5Class'])

```

## SMMv5Class

