Introduction to Coordinate Reference Systems

Bayesian modelling for spatial and spatio-temporal data

MSc in Epidemiology

Week 6

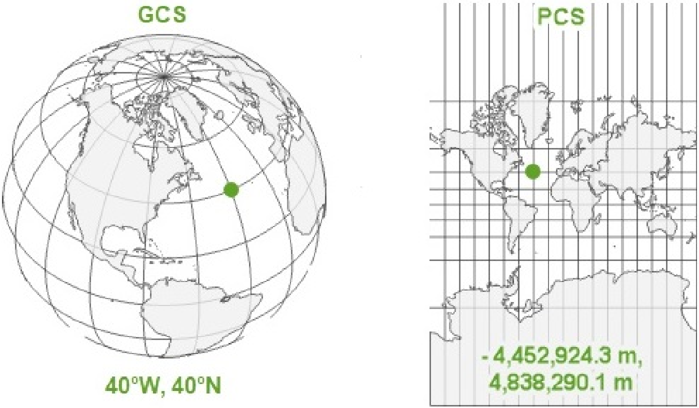
- A coordinate system defines how the spatial elements of the data relate to the surface of the Earth.
- The same CRS can often be referred to in many ways. In fact, each CRS is optimised to best represent the: (i) shape, and/or (ii) the distance and/or (iii) the area of features in a data set.
- There is not a single CRS that does a great job at optimizing all three elements: shape, distance and area.

There are two main types of CRS:

- Geographic coordinate systems (GCS): coordinate systems that identify any location on the Earth's surface using longitude and latitude, with units in decimal degrees or degrees.

- Projected coordinate systems (PCS): coordinate systems that provide various mechanisms to project maps of the Earth's ellipsoid shape onto a two-dimensional Cartesian coordinate plan. They are localised to minimize visual distortion in a particular region. Examples are: Universal Transverse Mercator (UTM), Albers Equal Area, or Robinson, typically with units of feet or meters. Projected coordinate systems are referred to as *map projections*.

GCS *vs* PCS

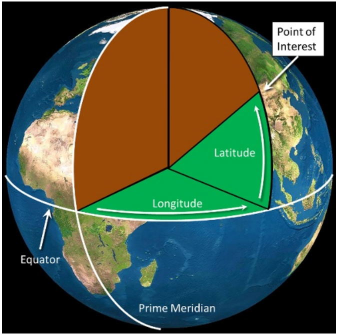


Source: `https://www.esri.com/arcgis-blog/products/arcgis-pro/mapping/gcs_vs_pcs/`

Geographic coordinate systems [1]

A point is referenced by its longitude and latitude, which are angles measured from the Earth's center to a point on the Earth's surface.

- *Longitude* is location in the East-West direction in angular distance from the Prime Meridian plane,
- *Latitude* is angular distance North or South of the equatorial plan.



Source:
`http://gsp.humboldt.edu/OLM/Lessons/GIS/03%20Projections/IntroductionToCoordinateSystems1.html`

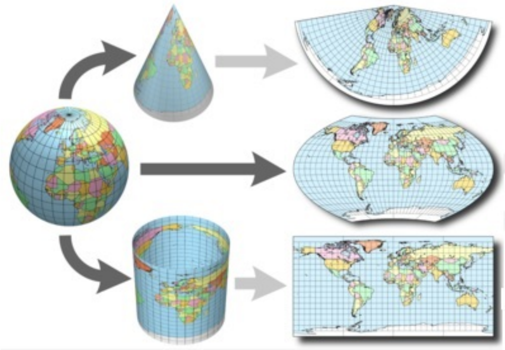
- Obviously we cannot actually measure these angles, but we can estimate them. To do so, we need a model of the shape of the Earth. Such a model is called a datum:
  - it contains information on what *ellipsoid* is used to approximate the Earth's shape and the relationship between the coordinate system and location on the Earth's surface; it also describes the origin $(0, 0)$ of a coordinate system.
  - There are two types of datum:
    - local such as NAD83 (North American Datum 1983) - the ellipsoidal surface is shifted to align with the surface at a particular location,
    - geocentric such as WGS84 (World Geodetic System 1984) - the center is the Earth's center of gravity and the accuracy of projections is not optimized for a specific location. It is used by the Global Positioning System (GPS).

Projected coordinate systems [1]

- A major question in spatial analysis is how to transform this three-dimensional angular system to a two dimensional planar, sometimes called Cartesian system.
- The different types of planar coordinate reference systems are referred to as projections.
- All projected CRSs are based on a geographic CRS and rely on map projections to convert the three-dimensional surface of the Earth into Easting and Northing (x and y) values in a projected CRS.
- The transition leads to some distortions: there is not one best projection. Some projections can be used for a map of the whole world; other projections are appropriate for small areas only.

Projected coordinate systems [2]

There are three main projection types: conic, cylindrical, and planar:



Source: https://www.earthdatascience.org/courses/earth-analytics/spatial-data-r/geographic-vs-projected-coordinate-reference-systems-UTM/

A list of available projections can be found at https://proj.org/operations/projections/

Differences in shape of United States boundaries associated with different projections

Source: Data Carpentry,
https://datacarpentry.org/r-raster-vector-geospatial/09-vector-when-data-dont-line-up-crs/.

Re-projecting

- Re-projecting is the process of changing the representation of locations from one CRS to another.
- But when should data be transformed?
  - The projections of locations on the Earth into a two-dimensional plane are *distortions*, the projection that is best for an application may be different from the projection associated with the data we import. In these cases, data can be re-projected.
  - Another case is when two objects with different CRSs must be compared or combined. Thus, in the case of dealing with multiple data, re-projection permits to transform all data to a common CRS.
- In R, to transform data to a different projection, we can use:
  - `spTransform()` function of the `rgdal` package
  - `st_transform()` function of the `sf` package

CRS in R (before 2020)

- Spatial R packages support a wide range of CRSs and they use the `PROJ` library to perform conversions between cartographic projections. Before 2020, there are two ways of defining a CRS:
  - via the `proj4string` definition, which specifies attributes such as the projection, the ellipsoid and the datum, for example the WGS84 longitude and latitude projection is specified as:
    $+proj = longlat + ellps = WGS84 + datum = WGS84 + no\_defs$
  - the `EPSG` numeric code (EPSG stands for European Petroleum Survey Group). The code refers to only one, well-defined coordinate reference system, for example the EPSG code of the WGS84 projection is 4326.
- Note that the EPSG code could not be available for a particular coordinate system. In R we can create a list of EPSG codes using the `make_epsg()` function in `rgdal` package, or check online at `https://spatialreference.org/ref/epsg/`.

There are new developments at level of `PROJ` (`https://proj.org/`), and shift from `proj4string` to the `OGC WKT2` representation (WKT stands for Well-known Text formats).
Currently, in R:

- The CRS is represented by a list with two components, `input` and `wkt`
- Direct transformations from CRS to CRS (previously, the path was to transform first to WSG84 and from there to the target projection, increasing the risk of error between conversions)
- Capability of time-dependent transformations (previously, static reference frames with no support for time-dependent datums, but on Earth, everything is moving, up to 80 mm/year)

```
library(spData)
library(sf)

# Extract the CRS information from the sf object nz (New Zealand)
st_crs(nz) # here EPSG:2193

# Re-project
nz_wgs <- st_transform(nz, 4326)
st_crs(nz_wgs) # now EPSG:4326

# Extract the CRS information from the sf object world
st_crs(world) #EPSG:4326

# Re-project using the CRS of the sf object world
nz_wgs = st_transform(nz, st_crs(world))
st_crs(nz_wgs)

# Remove the CRS from nz_wgs and plot the result
nz_wgs_NULL_crs = st_set_crs(nz_wgs, NA)

# Plots
par(mfrow = c(1, 3))
plot(st_geometry(nz))
plot(st_geometry(nz_wgs))
plot(st_geometry(nz_wgs_NULL_crs))
```

- For more details on CRS, see section 2.4 and chapter 6 of Lovelace, R., Nowosad, J., and Muenchow, J. (2019), Geocomputation with R, CRC Press; the online version of the book is at `http://geocompr.robinlovelace.net/`

- For more details on new developments linked to PROJ, you can visit the links: `https://proj.org/news.html`, and `https://info.crunchydata.com/blog/waiting-for-postgis-3-st_transform-and-proj6` and `http://rgdal.r-forge.r-project.org/articles/PROJ6_GDAL3.html`