5. Georeferencing

Geographic Information Systems and Science

SECOND EDITION

Paul A. Longley, Michael F. Goodchild, David J. Maguire, David W. Rhind

Outline

- Introduction
- Placenames
- Postal addresses and postal codes
- Linear referencing systems
- Cadasters
- Latitude and longitude
- Projections and coordinate systems
- The Global Positioning System
- Converting georeferences

Georeferencing

- Is essential in GIS, since all information must be linked to the Earth's surface
- The method of georeferencing must be:
  - Unique, linking information to exactly one location
  - Shared, so different users understand the meaning of a georeference
  - Persistent through time, so today's georeferences are still meaningful tomorrow

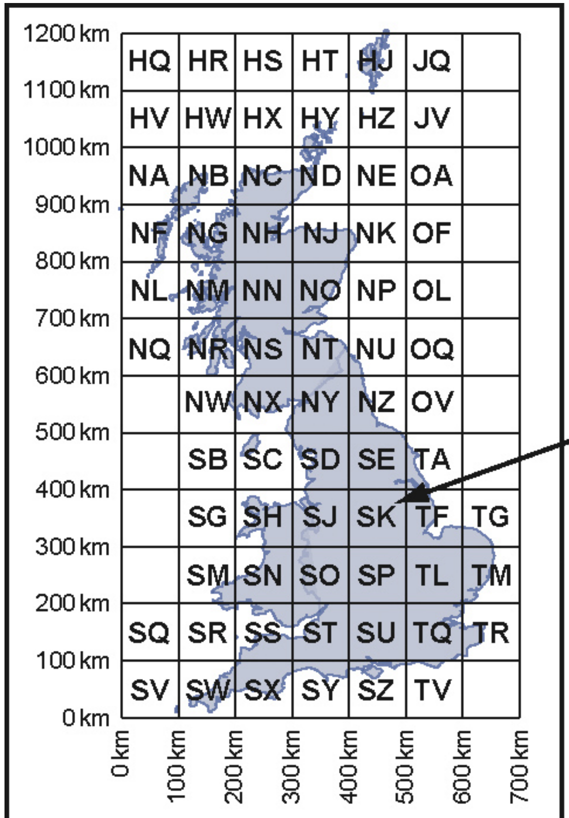
Uniqueness

- A georeference may be unique only within a defined domain, not globally
  - There are many instances of Springfield in the U.S., but only one in any state
  - The meaning of a reference to London may depend on context, since there are smaller Londons in several parts of the world

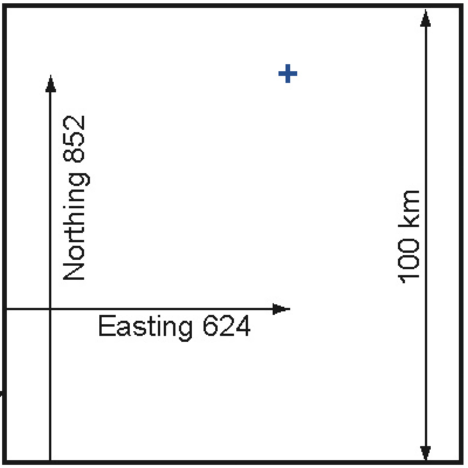
Georeferences as Measurements

- Some georeferences are metric
  - They define location using measures of distance from fixed places
    - E.g., distance from the Equator or from the Greenwich Meridian
- Others are based on ordering
  - E.g. street addresses in most parts of the world order houses along streets
- Others are only nominal
  - Placenames do not involve ordering or measuring

The National Grid is a system of metric georeferencing used in Great Britain. It is administered by the Ordnance Survey of Great Britain, and provides a unique georeference for every point in England, Scotland, and Wales. The first designating letter defines a 500 km square, and the second defines a 100 km square. Within each square, two measurements, called easting and northing, define a location with respect to the lower left corner of the square. The number of digits defines the precision—three digits for easting and three for northing (a total of six) define location to the nearest 100 m.



British National Grid 100 km Squares



SK624852

Placenames

- The earliest form of georeferencing
  - And the most commonly used in everyday activities
- Many names of geographic features are universally recognized
  - Others may be understood only by locals
- Names work at many different scales
  - From continents to small villages and neighborhoods
- Names may pass out of use in time
  - Where was Camelot?

Postal Addresses and Postcodes

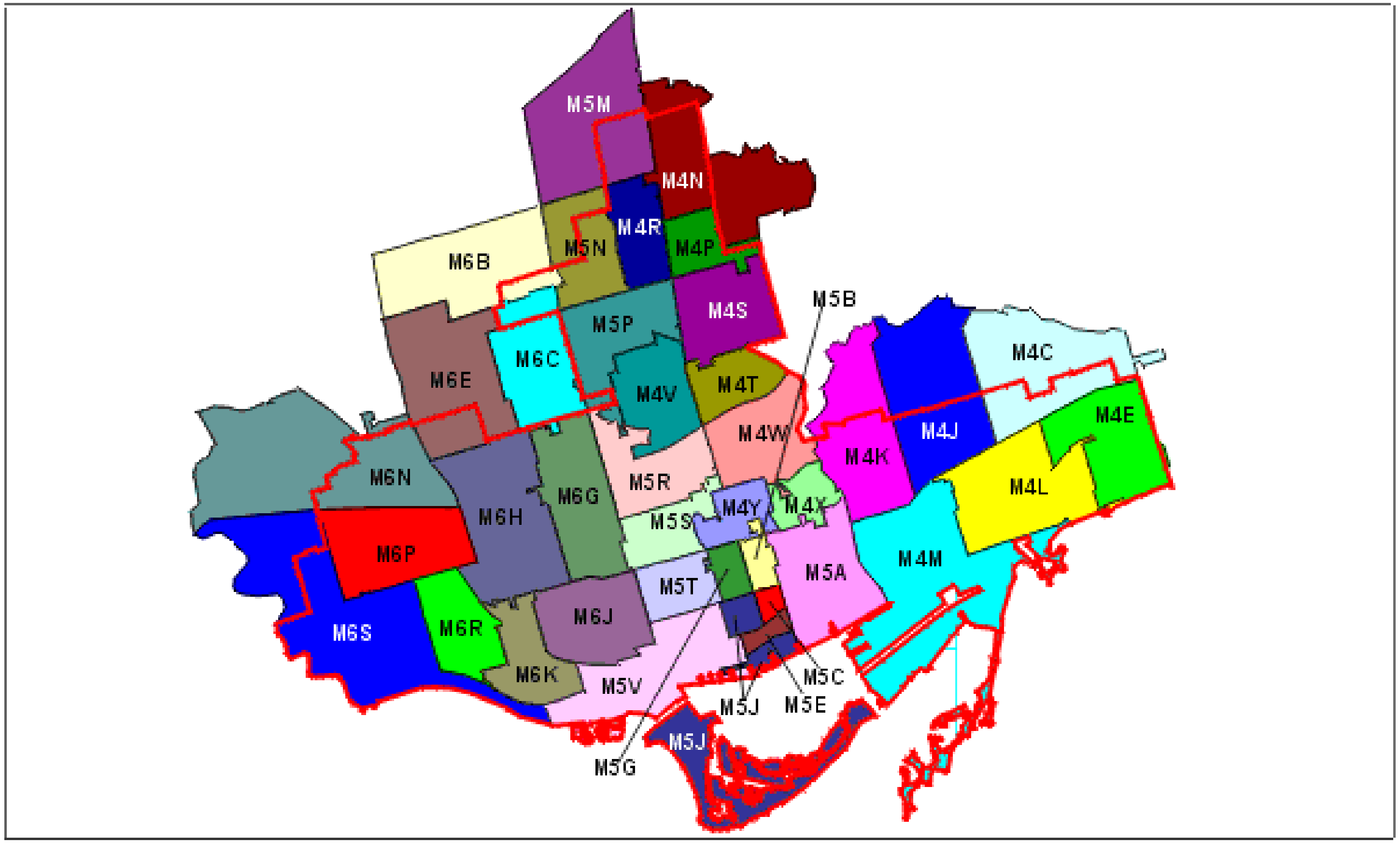
- Every dwelling and office is a potential destination for mail
- Dwellings and offices are arrayed along streets, and numbered accordingly
- Streets have names that are unique within local areas
- Local areas have names that are unique within larger regions
- If these assumptions are true, then a postal address is a useful georeference

Where Do Postal Addresses Fail as Georeferences?

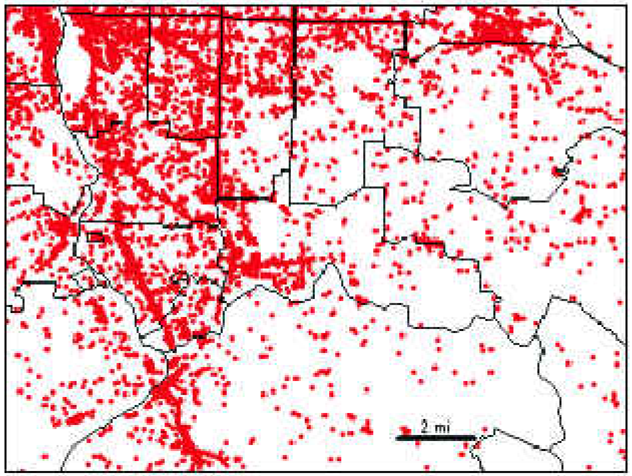
- In rural areas
  - Urban-style addresses have been extended recently to many rural areas
- For natural features
  - Lakes, mountains, and rivers cannot be located using postal addresses
- When numbering on streets is not sequential
  - E.g. in Japan

Postcodes as Georeferences

- Defined in many countries
  - E.g. ZIP codes in the US
- Hierarchically structured
  - The first few characters define large areas
  - Subsequent characters designate smaller areas
  - Coarser spatial resolution than postal address
- Useful for mapping

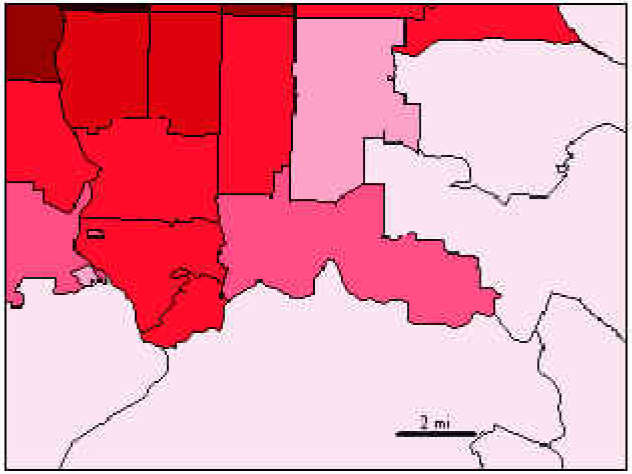


Forward sortation areas (FSAs) of the central part of the Toronto metropolitan region. In Canada the first three characters of the six-character postcode form the FSA



Business
Zip Code Boundary

Locations of businesses, with Zip Code boundaries.



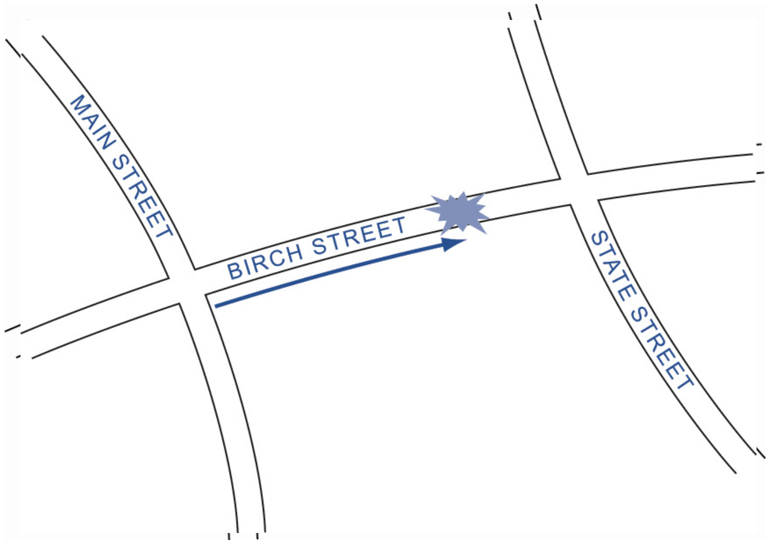
0 - 12
13 - 36
37 - 67
68 - 116
117 - 176
177 - 307
308 - 558

Zip Codes shaded by density of businesses per square mile.

ZIP code boundaries are a convenient way to summarize data in the US. The dots on the left have been summarized as a density per square mile on the right

Linear Referencing

- A system for georeferencing positions on a road, street, rail, or river network
- Combines the name of the link with an offset distance along the link from a fixed point, most often an intersection



Users of Linear Referencing

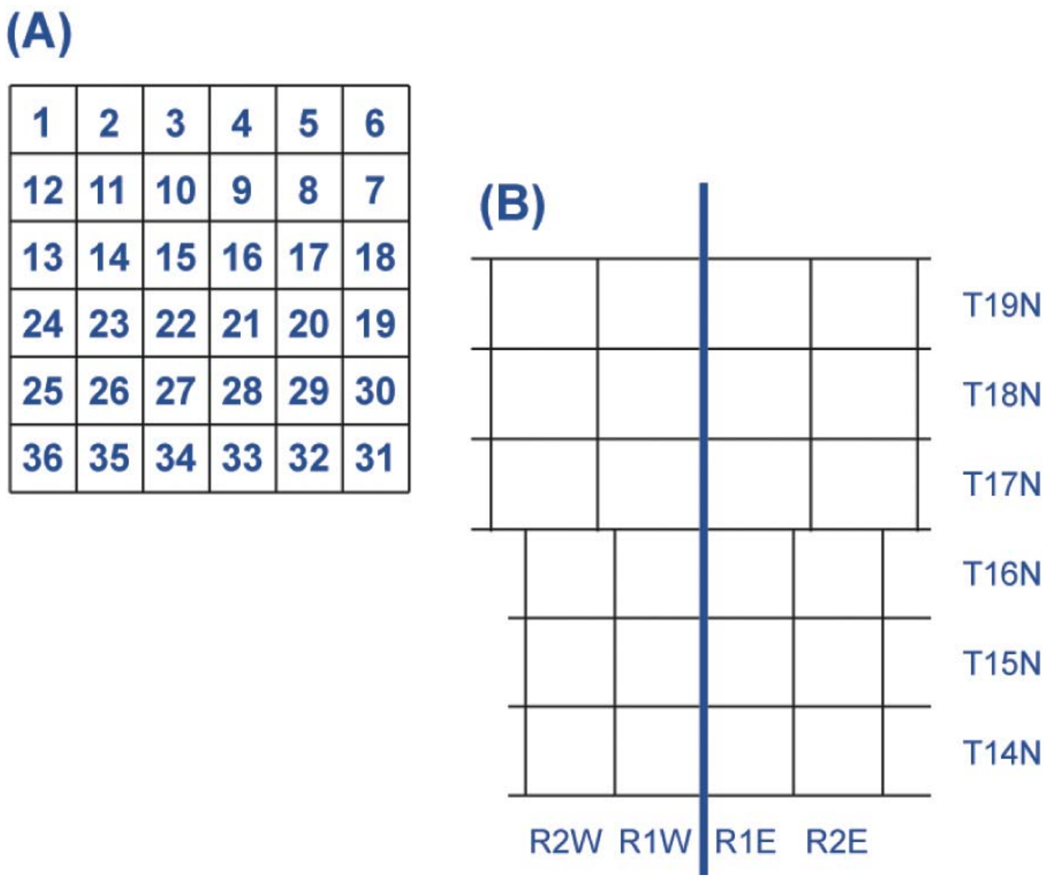
- Transportation authorities
  - To keep track of pavement quality, signs, traffic conditions on roads
- Police
  - To record the locations of accidents

Problem Cases

- Locations in rural areas may be a long way from an intersection or other suitable zero point
- Pairs of streets may intersect more than once
- Measurements of distance along streets may be inaccurate, depending on the measuring device, e.g. a car odometer

Cadasters

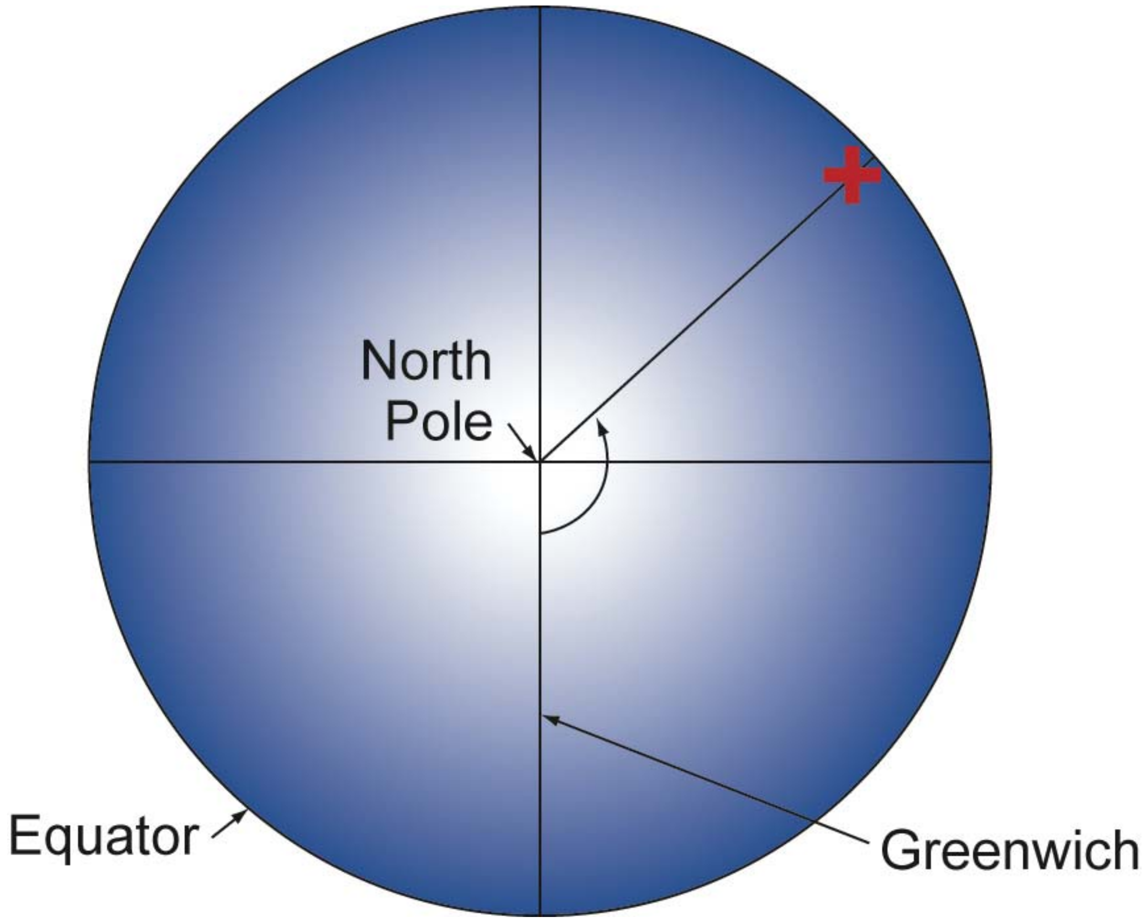
- Maps of land ownership, showing property boundaries
- The Public Land Survey System (PLSS) in the US and similar systems in other countries provide a method of georeferencing linked to the cadaster
- In the Western US the PLSS is often used to record locations of natural resources, e.g. oil and gas wells



Portion of the Township and Range system (Public Lands Survey System) widely used in the western US as the basis of land ownership. Townships are laid out in six-mile squares on either side of an accurately surveyed Principal Meridian. The offset shown between townships 16N and 17N is needed to accommodate the Earth's curvature (shown much exaggerated). The square mile sections within each township are numbered as shown in (A) east of the Principal Meridian, and reversed west of the Principal Meridian.

Latitude and Longitude

- The most comprehensive and powerful method of georeferencing
  - Metric, standard, stable, unique
- Uses a well-defined and fixed reference frame
  - Based on the Earth's rotation and center of mass, and the Greenwich Meridian



Definition of longitude. The Earth is seen here from above the North Pole, looking along the Axis, with the Equator forming the outer circle. The location of Greenwich defines the Prime Meridian. The longitude of the point at the center of the red cross is determined by drawing a plane through it and the axis, and measuring the angle between this plane and the Prime Meridian.

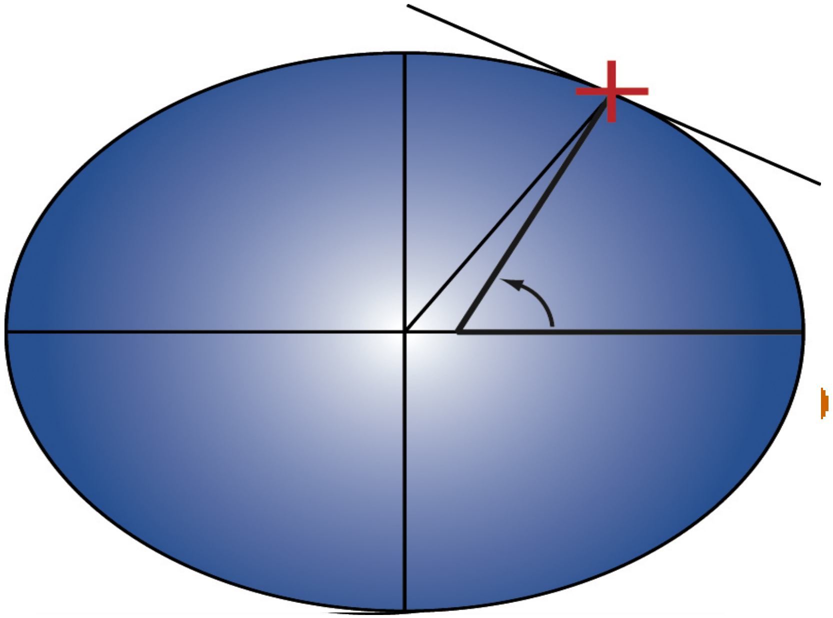
Definition of Latitude

- Requires a model of the Earth's shape
- The Earth is somewhat elliptical
  - The N-S diameter is roughly 1/300 less than the E-W diameter
  - More accurately modeled as an ellipsoid than a sphere
  - An ellipsoid is formed by rotating an ellipse about its shorter axis (the Earth's axis in this case)

The History of Ellipsoids

- Because the Earth is not shaped precisely as an ellipsoid, initially each country felt free to adopt its own as the most accurate approximation to its own part of the Earth
- Today an international standard has been adopted known as WGS 84
  - Its US implementation is the North American Datum of 1983 (NAD 83)
  - Many US maps and data sets still use the North American Datum of 1927 (NAD 27)
  - Differences can be as much as 200 m

Latitude and the Ellipsoid



- Latitude (of the blue point) is the angle between a perpendicular to the surface and the plane of the Equator
- WGS 84
  - Radius of the Earth at the Equator 6378.137 km
  - Flattening 1 part in 298.257

Projections and Coordinates

- There are many reasons for wanting to project the Earth's surface onto a plane, rather than deal with the curved surface
  - The paper used to output GIS maps is flat
  - Flat maps are scanned and digitized to create GIS databases
  - Rasters are flat, it's impossible to create a raster on a curved surface
  - The Earth has to be projected to see all of it at once
  - It's much easier to measure distance on a plane

Distortions

- Any projection must distort the Earth in some way
- Two types of projections are important in GIS
  - *Conformal* property: Shapes of small features are preserved: anywhere on the projection the distortion is the same in all directions
  - *Equal area* property: Shapes are distorted, but features have the correct area
  - Both types of projections will generally distort distances

Cylindrical Projections

- Conceptualized as the result of wrapping a cylinder of paper around the Earth
- The Mercator projection is the best-known cylindrical projection
  - The cylinder is wrapped around the Equator
  - The projection is conformal
    - At any point scale is the same in both directions
    - Shape of small features is preserved
    - Features in high latitudes are significantly enlarged

Conic Projections

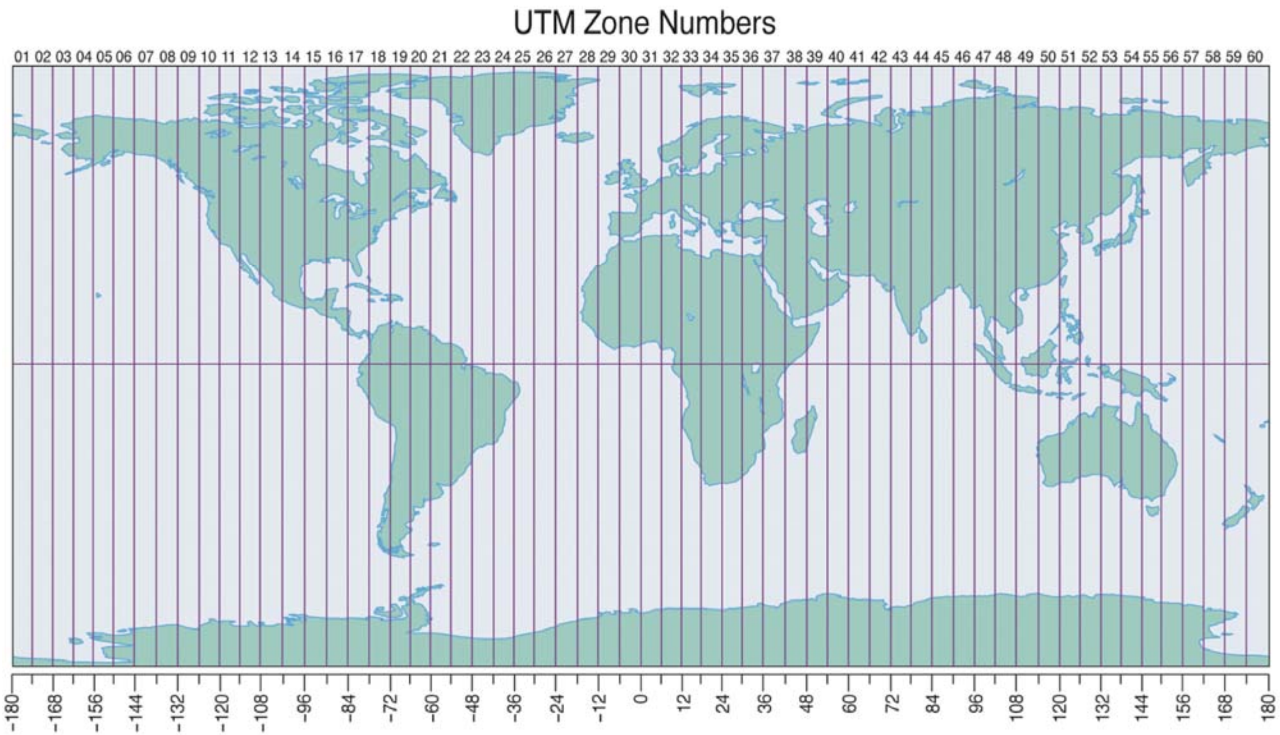
- Conceptualized as the result of wrapping a cone of paper around the Earth
  - Standard Parallels occur where the cone intersects the Earth
  - The Lambert Conformal Conic projection is commonly used to map North America
  - On this projection lines of latitude appear as arcs of circles, and lines of longitude are straight lines radiating from the North Pole

The "Unprojected" Projection

- Assign latitude to the *y* axis and longitude to the *x* axis
  - A type of cylindrical projection
  - Is neither conformal nor equal area
  - As latitude increases, lines of longitude are much closer together on the Earth, but are the same distance apart on the projection
  - Also known as the Plate Carrée or Cylindrical Equidistant Projection

The Universal Transverse Mercator (UTM) Projection

- A type of cylindrical projection
- Implemented as an internationally standard coordinate system
  - Initially devised as a military standard
- Uses a system of 60 zones
  - Maximum distortion is 0.04%
- *Transverse* Mercator because the cylinder is wrapped around the Poles, not the Equator



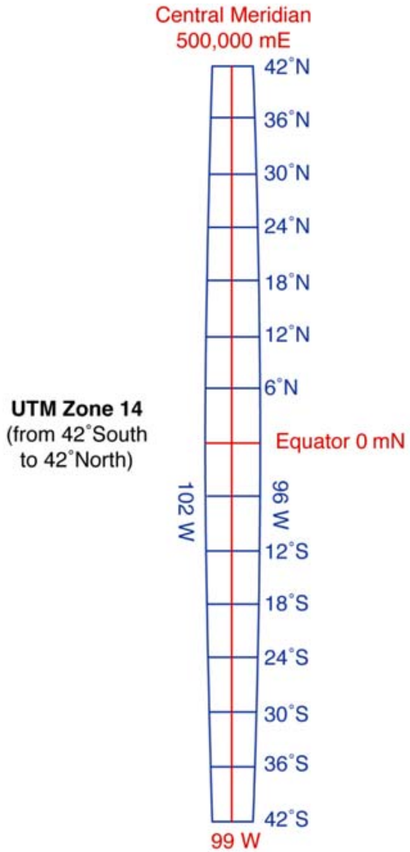
Universal Transverse Mercator (UTM) System

Zones are each six degrees of longitude, numbered as shown at the top, from W to E

Implications of the Zone System

- Each zone defines a different projection
- Two maps of adjacent zones will not fit along their common border
- Jurisdictions that span two zones must make special arrangements
  - Use only one of the two projections, and accept the greater-than-normal distortions in the other zone
  - Use a third projection spanning the jurisdiction
  - E.g. Italy is spans UTM zones 32 and 33

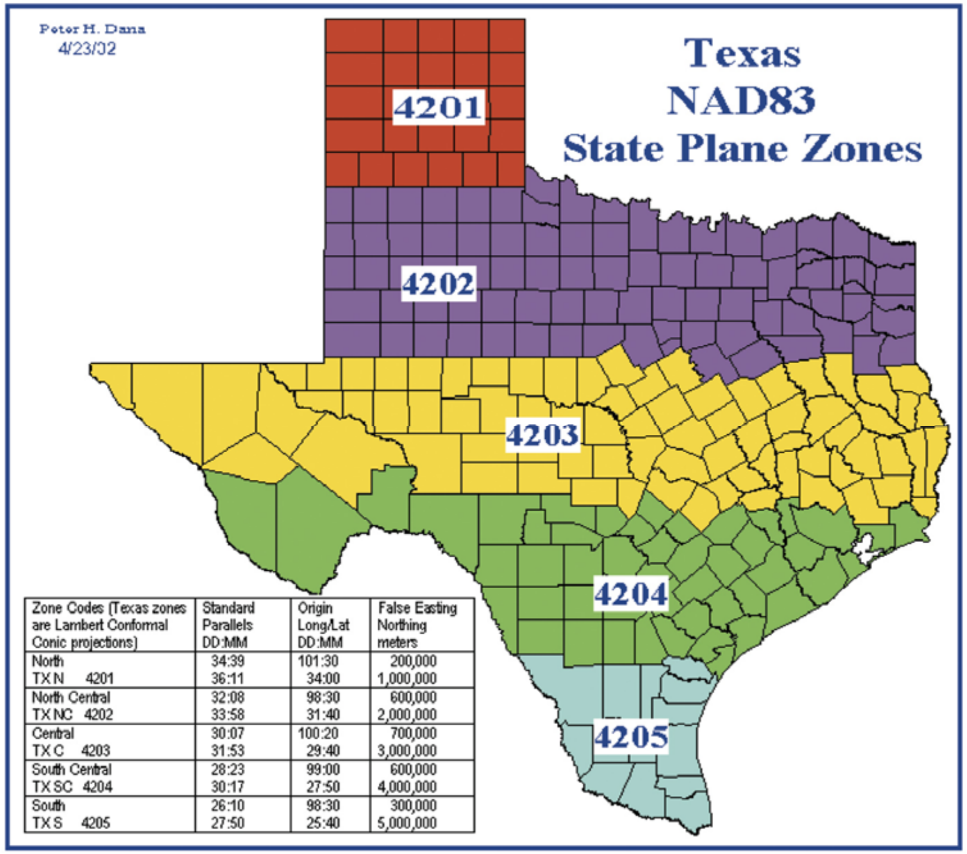
UTM Coordinates



- In the N Hemisphere define the Equator as 0 mN
- The central meridian of the zone is given a false easting of 500,000 mE
- Eastings and northings are both in meters allowing easy estimation of distance *on the projection*
- A UTM georeference consists of a zone number, a hemisphere, a six-digit easting and a seven-digit northing
  - E.g., 14, N, 468324E, 5362789N

State Plane Coordinates

- Defined in the US by each state
  - Some states use multiple zones
  - Several different types of projections are used by the system
- Provides less distortion than UTM
  - Preferred for applications needing very high accuracy, such as surveying



Zone Codes (Texas zones are Lambert Conformal Conic projections)	Standard Parallels DD:MM	Origin Long/Lat DD:MM	False Easting Northing meters
North TX N 4201	34:39 36:11	101:30 34:00	200,000 1,000,000
North Central TX NC 4202	32:08 33:58	98:30 31:40	600,000 2,000,000
Central TX C 4203	30:07 31:53	100:20 29:40	700,000 3,000,000
South Central TX SC 4204	28:23 30:17	99:00 27:50	600,000 4,000,000
South TX S 4205	26:10 27:50	98:30 25:40	300,000 5,000,000

Converting Georeferences

- GIS applications often require conversion of projections and ellipsoids
  - These are standard functions in popular GIS packages
- Street addresses must be converted to coordinates for mapping and analysis
  - Using *geocoding* functions
- Placenames can be converted to coordinates using *gazetteers*

The Global Positioning System

- Allows direct, accurate measurement of latitude and longitude
- Accuracy of 10m from a simple, cheap unit
  - Differential GPS capable of sub-meter accuracy
  - Sub-centimeter accuracy if observations are averaged over long periods