Geodesy, Map Projections and Coordinate Systems

Synopsis of Class 6, GIS in Water Resources, Fall 2011

The fundamentals of the material to be covered in this class are sufficiently extensive that this synopsis extends to three pages, rather than the two pages used previously. Besides this synopsis, please read the overview of Georeferencing and coordinate systems in the *ArcGIS Resource Center* (http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//00v20000000q000000.htm) and the section in the *ArcGIS Resource Center, Help for Desktop 10, Professional Library, Guide books, Map Projections* (http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/What_are_map_projections/003r00000001000000/) for additional information on this subject.

Geodesy is "the branch of the earth sciences that deals with the measurement and representation of the earth, including its gravitational field" (definition from Wikipedia). Geodetic or geographic locations on the earth are defined by latitude (ϕ), longitude (λ) and elevation (z), where the *horizontal coordinates* (ϕ),λ) are measured using angular units, and the *vertical coordinate* (z) is measured using distance units . A *global positioning system* (GPS) uses distance measurements from a collection of satellites to determine geodetic location.

Although we think of the earth as a sphere, it is in fact slightly flattened and best represented as an *oblate spheroid*, with an equatorial radius that is about 1/300 larger than its polar radius. This shape is a form of *ellipsoid* obtained by rotating an ellipse whose major and minor axes are sized to represent the earth's shape, around an axis to form a three-dimensional shape. A *horizontal earth datum* is a combination of a particular ellipsoid shape and an axis of rotation. The two datums used most often in the United States are the *North American Datum of 1927* (NAD27) and the *North American Datum of 1983* (NAD83). NAD27 was designed so that the ellipsoid surface closely matches the land surface in North America; NAD83 was an adjustment with a different ellipsoid shape and a *geocentric* axis of rotation through the poles, which is a consequence of earth measurement from satellites.

The formal definition of the latitude of a point on the earth's surface is that the *latitude* is the angle subtended in the equatorial plane of a line drawn normal to a tangent plane touching the earth's surface at the point of interest. The formal definition of longitude is that *longitude* is the angle subtended at the polar axis between a plane drawn through the polar axis and the meridian passing through the point of interest, and the corresponding plane drawn through the prime meridian.

Distance measurement on the earth's surface along meridians and parallels can be approximated for a spherical earth by using circular geometry, in other words arc length = radius of the earth multiplied by the angle subtended at the center in radians. The *great circle* distance, or shortest distance between two points anywhere on the earth's surface, is calculated using a more complicated formula, and distance measurement on a spheroidal earth surface is even more complex.

Horizontal coordinates are defined geographically or geometrically, but elevation is defined *gravitationally*. Newton's law of gravitation defines the attractive force between two masses separated by a distance, and *gravitational potential* is the potential energy that a unit mass possesses at a particular location in a gravitational field. The *geoid* is a surface of constant gravitational potential at or near the earth's surface that is the reference location for defining a *vertical earth datum*. Two vertical datums are used in the United States: the *National Geodetic Vertical Datum of 1929* (NGVD29), and the *North American Vertical Datum of 1988* (NAVD88). Elevation is defined as the vertical height above the geoid, or geodetic datum. While the ellipsoid forms a nice smooth mathematical surface, spatial variability in the earth's gravitational field causes the geoid to have a lumpy surface, in some cases above the ellipsoid, in other cases below. The difference in heights between the geoid and the ellipsoid is called the *gravity anomaly*.

Elevation along the coast is defined using *tidal heights,* which are elevations measured using water surface elevation gages in coastal waters relative to a statistical average tidal level calculated through time. The best known reference tidal height is *mean sea level*, and land surface elevations are often quoted as being "feet above mean sea level" as if the sea were to extend under the land surface. The concept of mean sea level is outdated for two reasons – first, tidal measurements over long periods of time show consistent and pronounced trends in tidal heights, rising on the order of one foot per century along the US coastline; second, the ocean itself is constantly moving, and has spatial and temporal trends in its temperature and salinity that affect its surface elevation. Elevation measurement is a complicated matter that actually has three different measurement systems – height above ellipsoid (used by GPS), height above geoid (the legal standard used by land surveyors), and tidal heights (used for coastal management). Further complicating all this is that the gravitational field of the earth changes in time because of the season accumulation and release of water storage in snow, reservoirs and aquifers. We live in a complicated world!

Map projection is a method of producing a flat map representation of a portion of the earth's surface. The resulting horizontal coordinates are termed the *Easting* (x) and *Northing* (y), and map projection is actually carried out by a mathematical transformation between (ϕ, λ) and (x,y). There are three types of map projections, each of which can be visualized by placing a sheet of paper around or on a globe, in the form of a cone for a *conic projection*, as a cylinder for a *cylindrical projection*, and as a plane for an *azimuthal projection*. Conic projections are used for regions oriented in the East-West direction, such as the continental United States; cylindrical projections used for regions oriented in the North-South direction, such as for California; azimuthal projections are used for particular regions such as the poles where all the meridians come together.

All map projections have an associated set of *projection parameters*, which specify the particular way the projection will be applied, typically including a reference latitude and meridian (ϕ_o, λ_o) that locate a point of origin on the land surface near the center of the map, and a *false easting* and *false northing* (x_o,y_o) that represent the same origin point on the map. The false easting and northing are sometimes specified as large numbers in meters or feet so that all points on the map end up having positive (x,y) coordinates. Since latitude and longitude vary slightly depending on which horizontal earth datum is used, a complete description of a map projection includes the earth datum, the projection type, and the projection parameters.

There are four properties that map projections seek to preserve: area, shape, orientation, and distance. The *Albers equal area* projection preserves true earth surface area by distorting distance in the North-South direction at just the inverse rate that it distorts distance in the East-West direction. We use this projection extensively in hydrology because preserving true earth area is important when computing mass balances such as from rainfall to runoff. The *transverse Mercator* projection places a cylinder around the earth with its axis through the poles. Meridians and parallels form rectangles in this projection. A variant of this, called the *web Mercator* projection is used for cached basemaps.

A *coordinate system* is a method of establishing standard values of (x,y) for particular locations on the land surface. Within the United States, legal locations for land boundaries are defined using the *state plane coordinate system*, in which each state is described by one more zones, with a particular set of projection parameters associated with each zone. For example, Texas has five state plane zones (North, North-Central, Central, South-Central and South) to minimize distortion of true earth properties within each zone, and in each zone, map units in feet are used to measure distance.

Globally, the best known coordinate system is the Universal Transverse Mercator (UTM) system, which was devised during the US military operations during World War II. In UTM coordinates the map units are in meters. The world is divided into thin vertical strips bounded by meridians, each such zone being six degrees of longitude in width, and stretching from pole to pole in latitude. The zones are numbered beginning with zone 1 at 180°W and increasing eastward. Central Texas lies in zone 14. A *central meridian* (λ_o) runs through the middle of each zone, and for all zones the reference latitude (ϕ_o) is the equator. In the Northern Hemisphere, the corresponding false easting and northing (x_o,y_o) are (500,000, 0). Hence, any point on earth can be located given its UTM zone and (x,y) coordinates. You can imagine how important this was in conducting military operations during World War II using paper maps! Many of the advances in earth mapping have arisen first from military needs, including the global positioning system we now use so much in the civilian world.

In ArcGIS, the *spatial reference* of a feature dataset or feature class contains its datum, projection and coordinate system. It also defines its *map extent*, or the maximum coverage over the earth of a box within which the data reside, and some specialized information called the *tolerance* and *resolution* that have to do with how precisely earth coordinates can be specified given the limits on the number of digits used to store location in the computer system being used.