



5 Georeferencing

LEARNING OBJECTIVES

After Studying this chapter you will:

- Know the requirements for an effective system of georeferencing.
- Be familiar with the problems associated with place-names, street addresses, and other systems used every day by humans.
- Know how the Earth is measured and modeled for the purposes of positioning.
- Know the basic principles of map projections and the details of some commonly used projections.
- Know about conversion between different systems of georeferencing.
- Understand the principles behind GPS and learn some of its applications.

Geographic location is the element that distinguishes geographic information from all other types of information, so methods for specifying location on the Earth's surface are essential to the creation of useful geographic information. Many such techniques have been developed over the centuries, and this chapter provides a basic guide for GIS students—what you need to know about georeferencing to succeed in GIS. The first section lays out the principles of georeferencing, including the requirements that any effective system must satisfy. Subsequent sections discuss commonly used systems, starting with the ones closest to everyday human experience, including place-names and street addresses, and moving to the more accurate scientific methods that form the basis of geodesy and surveying, and the most recent methods developed for the Web. The final sections deal with issues that arise over conversions between georeferencing systems, with the Global Positioning System (GPS), with georeferencing of computers and cell phones, and with the concept of a gazetteer.

5.1 Introduction

Chapter 3 introduced the idea of an atomic element of geographic information: an atom made up of location, time (optionally), and attribute. To make GIS work there must be techniques for assigning values to all three aspects, in ways that are commonly understood by people who wish to communicate. Almost all the world agrees on a common calendar and time system, so there are only minor problems associated with communicating that element of the atom when it is needed (although different time zones, different names of the months in different languages, the annual switch to summer or Daylight Saving Time, and systems such as the classical Japanese convention of dating by the year of the emperor's reign all sometimes manage to confuse us).

Time is optional in a GIS, but location is not, so this chapter focuses on techniques for specifying location, and the problems and issues that arise. Locations are the basis for many of the benefits of GIS: the ability to map, to tie different kinds of information together because they refer to the same place, or to measure distances and areas. Without locations, data are said to be *nonspatial* or *aspatial* and would likely have no value at all within a geographic information system.

Time is an optional element in geographic information, but location is essential.

Several terms are commonly used to describe the act of assigning locations to atoms of information. We use the verbs georeference, geolocate, and geocode, and say that facts have been georeferenced or geocoded. We talk about tagging records with geographic locations, or about locating them. The term georeference will be used throughout this chapter.

The primary requirements of a georeference are (1) that it be *unique*, so that there is only one location associated with a given georeference, and therefore no confusion about the location that is referenced; and (2) that its meaning be *shared* among all of the people who wish to work with the information, including their geographic information systems. For example, the georeference 909 West Campus Lane, Goleta, California, USA, points to a single house—there is no other house anywhere on Earth with that address—and its meaning is shared sufficiently widely to allow mail to be delivered to the address from virtually anywhere on the planet. The address may not be meaningful to everyone living in China, but it will be meaningful to a sufficient number of people within China's postal service, so that a letter mailed from

China to that address will likely be delivered successfully. Uniqueness and shared meaning are sufficient also to allow people to link different kinds of information based on common location: for example, a driving record that is georeferenced by street address can be linked to a record of purchasing.

To be as useful as possible, a georeference must be persistent through time because it would be very confusing if georeferences changed frequently, and very expensive to update all of the records that depend on them. This can be problematic when a georeferencing system serves more than one purpose or is used by more than one agency with different priorities. For example, a municipality may change its boundaries by incorporating more land, creating problems for mapping agencies and for researchers who wish to study the municipality through time. Street names sometimes change, and postal agencies sometimes revise postal codes. Changes even occur in the names of cities (Saigon to Ho Chi Minh City) or in their conventional transcriptions into the Roman alphabet (Peking to Beijing).

To be most useful, georeferences should stay constant through time.

Every georeference has an associated spatial resolution, equal to the size of the area that is assigned that georeference. A mailing address could be said to have a spatial resolution equal to the size of the mailbox, or perhaps to the area of the parcel of land or structure assigned that address. A U.S. state has a spatial resolution that varies from the size of Rhode Island to that of Alaska, and many other systems of georeferencing have similarly wide-ranging spatial resolutions.

Many systems of georeferencing are unique only within an area or *domain* of the Earth's surface. For example, many cities in the United States have the name Springfield (18 according to a recent edition of the Rand McNally Road Atlas; similarly, there are 9 places called Whitchurch in the 2003 AA Road Atlas of the United Kingdom). However, city name is unique within the domain of a U.S. state—a property that was engineered with the advent of the postal system in the Nineteenth Century—whereas in the UK city names are unique within counties. Today there is no danger that there are two Springfields in Massachusetts; a driver can therefore confidently ask for directions to "Springfield, Massachusetts" in the knowledge that there is no danger of being sent to the wrong Springfield. But people living in London, Ontario, Canada are well aware of the dangers of talking about "London" without specifying



Figure 5.1 Place-names are not necessarily unique at the global level. There are many Londons, for example, besides the largest and most prominent one in the UK. People living in other Londons must often add more information (e.g., London, Ontario, Canada) to resolve ambiguity. (Photolink/Photolibrary Group Limited)

the appropriate domain. Even in Toronto, Ontario a reference to "London" may be misinterpreted as a reference to the older (UK) London on a different continent rather than to the one 200 km away in the same province (Figure 5.1). Street name is unique in the United States within municipal domains, but not within larger domains such as county or state. There are 120 places on the Earth's surface with the same Universal Transverse Mercator coordinates (see Section 5.8.2), and a zone number and hemisphere must be added to make a reference unique in the global domain.

Although some georeferences are based on simple names, others are based on various kinds of measurements and are called *metric* georeferences. They include latitude and longitude and various kinds of coordinate systems, many of which are discussed in more detail later in this chapter, and are essential to the making of maps, the display of mapped information in GIS, and any kind of numerical analysis. One enormous advantage of such systems is that they provide the potential for infinitely fine spatial resolution: provided we have sufficiently accurate measuring devices and use enough decimal places, such systems will allow us to locate information to any level of accuracy. Another advantage is that from measurements of two or more locations it is possible to compute distances, a very important requirement of georeferencing in GIS.

Metric georeferences are much more useful because they allow maps to be made and distances to be calculated.

Other systems simply order locations. In most countries mailing addresses are ordered along streets, often using the odd integers for addresses on one side and the even integers for addresses on the other. This means that it is possible to say that 3000 State Street and 100 State Street are farther apart than 200 State Street and 100 State Street, allowing postal services to sort mail for easy delivery. In the western United States, it is often possible to infer estimates of the distance between two addresses on the same street by knowing that 100 addresses are assigned to each city block and that blocks are typically between 120 m and 160 m long.

This section has reviewed some of the general properties of georeferencing systems, and Table 5.1 shows some commonly used systems. The following sections discuss the specific properties of the systems that are most important in GIS applications.

5.2 Place-names

Giving names to places is the simplest form of georeferencing and was most likely the one first developed by early hunter-gatherer societies. Any distinctive feature on the landscape, such as a particularly old tree, can serve as a point of reference for two people who wish to share information, such as the existence of good game in the tree's vicinity. Human landscapes rapidly became littered with names, as people sought distinguishing labels to use in describing aspects of their surroundings and as other people adopted them. Today, of course, we have a complex system of naming oceans, continents, cities, mountains, rivers, and other prominent features. Each country maintains a system of authorized naming, often through national or state committees assigned with the task of standardizing geographic names. Nevertheless, multiple names are often attached to the same feature, for example when cultures try to preserve the names given to features by the original or local inhabitants (Mount Everest to many, but Chomolungma to many Tibetans), or when city names are different in different languages (Florence in English, Firenze in Italian).

Many commonly used place-names have meanings that vary between people and with the context in which they are used.

Language extends the power of place-names through words such as "between," which serves to refine

Table 5.1 Some commonly used systems of georeferencing.

System	Domain of uniqueness	Metric?	Example	Spatial resolution
Place-name	varies	no	London, Ontario, Canada	varies by feature type
Postal address	global	no, but ordered along streets in most countries	909 West Campus Lane, Goleta, California	size of one mailbox
Postal code	country	no	93117 (U.S. zip code); WC1E 6BT (UK Unit Postcode)	area occupied by a defined number of mailboxes
Telephone calling area	country	no	805	varies
Cadastral system	local authority	no	Parcel 01452954, City of Springfield, MA	area occupied by a single parcel of land
Public Land Survey System	Western Canada and United States only, unique to Prime Meridian	yes	Sec 5, Township 4 N, Range 6E	defined by level of subdivision
Latitude/longitude	global	yes	119 degrees 45 minutes West, 34 degrees 40 minutes North	infinitely fine
Universal Transverse Mercator	zones six degrees of longitude wide, and N or S Hemisphere	yes	563146E, 4356732N	infinitely fine
State Plane Coordinates	United States only, unique to state and to zone within state	yes	55086.34E, 75210.76N	infinitely fine

references to location, or “near,” which serves to broaden them. “Where State Street crosses Mission Creek” is an instance of combining two place-names to achieve greater refinement of location than either name could achieve individually. Even more powerful extensions come from combining place-names with directions and distances, as in “200 m north of the old tree” or “50 km west of Springfield.”

But place-names are of limited use as georeferences. First, they often have very coarse spatial resolution. “Asia” covers over 43 million sq km, so the information that something is located “in Asia” is not very helpful in pinning down its location. Even Rhode Island, the smallest state of the United States, has a land area of over 2700 sq km. Second, only certain place-names are officially authorized by national or subnational agencies. Many more are recognized only locally, so their use is limited to communication between people in the local community. Place-names

may even be lost through time: although there are many contenders, we do not know with certainty where the “Camelot” described in the English legends of King Arthur was located, if indeed it ever existed.

The meaning of certain place-names can become lost through time.

The growing Web phenomenon of user-generated content is changing this situation, however, and creating an alternative to the traditional top-down system of naming. Individuals are now able to assign names to features quite independently of officialdom by using sites such as Wikimapia (www.wikimapia.org; see Figure 5.2). At time of writing, individuals worldwide had added descriptions of close to 12,000,000 features, from the largest cities to the smallest buildings. Descriptions can be in any language and a wide range



Figure 5.2 Wikimapia coverage of the University of California, Santa Barbara, and surrounding neighborhood. Each rectangle represents one entry, where a volunteer has provided a description and possible images and links to other information, all of which can be exposed by clicking on the rectangle. More rectangles will be displayed as the user zooms.

of scripts, and can include photographs and links to other sources of information.

Wikimapia’s mantra is “Let’s describe the whole world,” and it is one example of a growing phenomenon that has been termed *volunteered geographic information* (VGI), and that draws individual citizens into the process of creating geographic knowledge. In essence, it echoes an earlier era before the establishment of national mapping agencies and authoritative committees, when place-names were created by explorers and local citizens. Many names fell by the wayside, but some were adopted and shown on maps. The process by which America was named, in 1507, was essentially of this nature: an act by an individual cartographer, Martin Waldseemüller, in St-Dié-des-Vosges in France. Waldseemüller had recently read letters from the Florentine explorer Amerigo Vespucci claiming credit for recognizing that the lands discovered to the west of Europe formed a previously unknown continent. Accordingly, he invented a name by feminizing Vespucci’s first name, believing that all continents should have feminine names (Figure 5.3).

Copies of the map were distributed across Europe, and the name stuck.

5.3 Postal Addresses and Postal Codes

Postal addresses were introduced after the development of mail delivery in the Nineteenth Century. They rely on several assumptions:

- Every dwelling and office is a potential destination for mail.
- Dwellings and offices are arrayed along paths, roads, or streets, and are numbered sequentially.
- Paths, roads, and streets have names that are unique within local areas.
- Local areas have names that are unique within larger regions.
- Regions have names that are unique within countries.



Figure 5.3 Detail of the Waldseemüller map of 1507, which for the first time showed the name the cartographer selected for the new continent. (Courtesy Library of Congress Map Collection)

If the assumptions are true, then mail address provides a unique identification for every dwelling and office on Earth.

Today, postal addresses are an almost universal means of locating many kinds of human activity: delivery of mail, place of residence, or place of business. They fail, of course, in locating anything that is not a potential destination for mail, including almost all kinds of natural features (Mount Everest does not have a postal address, and neither does Manzana Creek in Los Padres National Forest in California). They are not as useful when dwellings are not numbered consecutively along streets, as happens in some cultures (notably in Japan, where street numbering can reflect date of construction, not sequence along the street; it is temporal, rather than spatial) and in large building complexes like condominiums. Many GIS applications rely on the ability to locate activities by postal address and to convert addresses to some more universal system of georeferencing, such as latitude and longitude, for mapping and analysis.

Postal addresses work well to georeference dwellings and offices, but not natural features.

Postal codes were introduced in many countries in the late Twentieth Century in order to simplify the

sorting of mail. In the UK system, for example, the first characters of the code identify an Outward Code, and mail is initially sorted so that all mail directed to a single Local Delivery Office (LDO) is together. Incoming mail is accumulated in a local sorting station and sorted a second time by the last characters of the code so that it can be delivered by an LDO. Figure 5.4 shows a map of the Outward Codes for Southend-on-Sea (SS). The full six or seven characters of the postal code (e.g., SS5 4PJ) are unique to roughly 13 houses, a single large business, or a single building. Different countries are more or less insistent on completeness of postal codes; in the UK, for example, sorters will deliver mail with only an Outward Code since LDOs are required to know their local communities.

Postal codes have proven very useful for many purposes besides the sorting and delivery of mail. Although the area covered by a Canadian FSA, a U.S. ZIP code, or a UK postcode varies, and can be changed whenever the postal authorities want, it is sufficiently constant to be useful for mapping purposes, and many businesses routinely make maps of their customers by counting the numbers present in each postal code area, as well as by dividing by total population to get a picture of market penetration. Figure 5.5 shows an example of summarizing data by postal code. In Figure 5.5A it has been necessary to suppress the locations of postcodes where there are

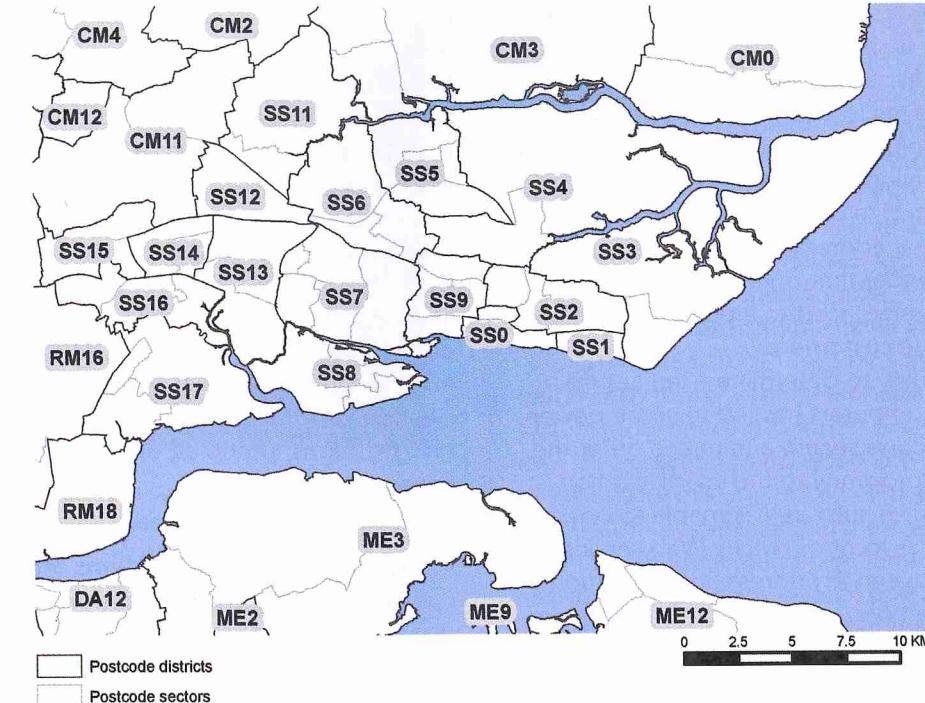


Figure 5.4 Outward Codes for the Southend-on-Sea, UK, Local Delivery Offices (LDOs). Outward codes form the first two, three, or four characters of the UK postal code and are separated from the three characters that identify the Inward Codes. (Map courtesy of Maurizio Gibin)

fewer than five resident patients; it is not necessary to hide any data in the choropleth map shown in Figure 5.5B, since privacy is not violated at this coarser level of granularity. Most people know the postal code of their home, and in some instances postal codes have developed popular images (the ZIP code for Beverly Hills, California, 90210, became the title of a successful television series).

5.4 IP Addresses

Every device (computer, printer, etc.) connected to the Internet has a unique IP (Internet Protocol) address, such as 128.111.106.183, the address of the computer on which this text is being written. IP addresses are allocated to organizations, and the street address included in the registration record can

Figure 5.5 The use of Outward Code boundaries as a convenient basis for summarizing data. In this instance (A) points identifying the residential unit (Inward) postcode have been used to plot the locations of a doctor's patients, and (B) the Outward Code areas have been shaded according to the density of patients per square kilometer.



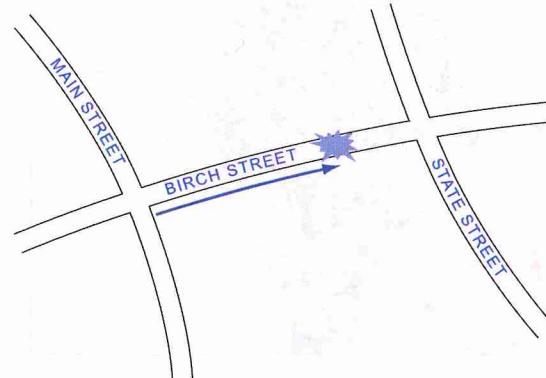
provide an approximate location. The IP address of the user's computer is provided whenever the computer is used to access a Web site, allowing the operators of major sites to determine the user's location. Spatial resolution will vary, however, since all of the IP addresses allocated to a university campus may be georeferenced to a single point on the campus, and all of the addresses allocated to an Internet Service Provider that operates a broadband service may be georeferenced to a point that is many kilometers from some customers.

Recently, the ability to determine even approximate locations for computers has led to some powerful applications. It allows search engines to order the results of search by proximity to the user's location, and it also permits sites such as Wikimapia to open centered on the user's location. Many Web services offer conversion between IP address and geographic coordinates. For example, www.networldmap.com resolves the address 128.111.106.183 to 34.4119 north, 119.7280 west, approximately 10 km from the office at the University of California, Santa Barbara, where the computer with this address is located. Wikimapia does rather better, opening centered on a location only 1 km away.

5.5 Linear Referencing Systems

A linear referencing system identifies location on a network by measuring distance from a defined point of reference along a defined path in the network. Figure 5.6 shows an example, an accident whose

Figure 5.6 Linear referencing—an incident's position is determined by measuring its distance (87 m) along one road (Birch Street) from a well-defined point (its intersection with Main Street).



location is reported as being a measured distance from a street intersection, along a named street. Linear referencing is closely related to street address, but uses an explicit measurement of distance rather than the much less reliable surrogate of street address number.

Linear referencing is widely used in applications that depend on a linear network. This includes highways (e.g., Mile 1240 of the Alaska Highway), railroads (e.g., 25.9 miles from Paddington Station in London on the main line to Bristol, England), electrical transmission lines, pipelines, and canals. Highway agencies use linear references to define the locations of bridges, signs, potholes, and accidents, and to record pavement condition.

Linear referencing systems are widely used in managing transportation infrastructure and in dealing with emergencies.

Linear referencing provides a sufficient basis for georeferencing for some applications. Highway departments often base their records of accident locations on linear references, as well as their inventories of signs and bridges. (GIS has many applications in transportation that are known collectively as GIS-T, and in the developing field of intelligent transportation systems or ITS; see Section 2.3.4.) But for other applications it is important to be able to convert between linear references and other forms, such as latitude and longitude. For example, the OnStar system that is installed in many Cadillacs sold in the United States is designed to radio the position of a vehicle automatically as soon as it is involved in an accident. When the airbags deploy, a GPS receiver determines position, which is then relayed to a central dispatch office. Emergency response centers often use street addresses and linear referencing to define the locations of accidents, so the latitude and longitude received from the vehicle must be converted before an emergency team can be sent to the accident.

Linear referencing systems are often difficult to implement in practice in ways that are robust in all situations. In an urban area with frequent intersections, it is relatively easy to measure distance from the nearest one (e.g., on Birch St 87 m east of the intersection with Main Street). But in rural areas an incident may be a long way from the nearest intersection. Even in urban areas it is not uncommon for two streets to intersect more than once (e.g., Birch may have two intersections with Columbia Crescent), or for a street to intersect with itself (Figure 5.7).



Figure 5.7 The intersection of a street with itself (Calgary, Alberta, Canada), causing problems for linear referencing.

5.6 Cadasters and the U.S. Public Land Survey System

The cadaster is defined as the map of land ownership in an area, maintained for the purposes of taxing land or of creating a public record of ownership. The process of subdivision creates new parcels by legally subdividing existing ones.

Parcels of land in a cadaster are often uniquely identified, by number or by code, and are also reasonably persistent through time, thereby satisfying the requirements of a georeferencing system. Indeed, it has often been argued that the cadaster could form the universal basis of mapping (a *multipurpose cadaster*), with all other geographic information tied to it. But very few people know the identification code of their home parcel; thus use of the cadaster as a georeferencing system is limited largely to local officials, with one major exception.

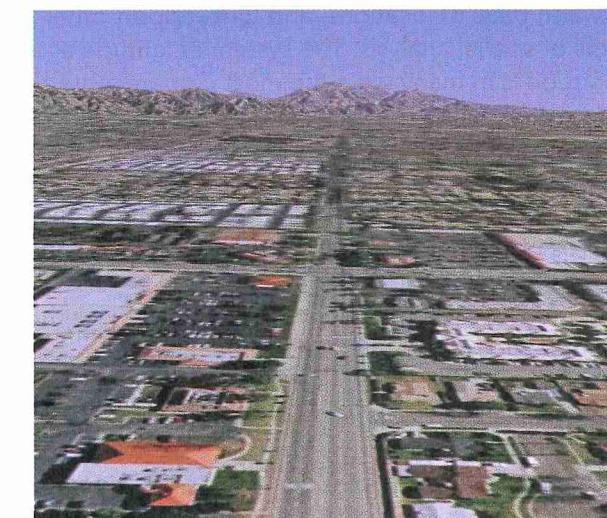
The U.S. Public Land Survey System (PLSS) evolved out of the need to survey and distribute the vast land resources of the western United States, starting in the early Nineteenth Century, and expanded to become the dominant system of cadaster for all of the United States west of Ohio, and all of western Canada. Its essential simplicity and regularity make it useful for many purposes and understandable by the general public. Its geometric regularity also allows it to satisfy the requirement of a metric system of georeferencing because each georeference is defined by measured distances.

The Public Land Survey System defines land ownership over much of western North America and is a useful system of georeferencing.

To implement the PLSS in an area, a surveyor first laid out an accurate north-south line or *principal meridian*. An east-west baseline was then laid out perpendicular to this (Figure 5.8). Rows were then laid out 6 miles apart and parallel to the baseline, to become the townships of the system. Then blocks or ranges were laid out in 6 mile by 6 mile squares on either side of the principal meridian (see Figure 5.9). Each square is referenced by township number, range number, whether it is to the east or to the west, and the name of the principal meridian. Thirty-six sections 1 mile by 1 mile were laid out inside each township and numbered using a standard system (note how the numbers reverse in every other row). Each section was divided into four quarter-sections of 1/4 square mile, or 160 acres, the size of the nominal family farm or homestead in the original conception of the PLSS. The process can be continued by subdividing into four to obtain any level of spatial resolution.

The PLSS would be a wonderful system if the Earth were flat and if survey measurements were always exact. To account for the Earth's curvature, the squares are not perfectly 6 miles by 6 miles, and the rows must be offset frequently; errors in the original surveying complicate matters still further, particularly in rugged landscapes. Figure 5.9 shows the offsetting exaggerated for a small area. Nevertheless, the PLSS remains an efficient system and one with which many people in the Western USA and Western Canada are

Figure 5.8 Google Earth simulation of the view looking east along Baseline Road in Ontario, California—the road that follows the original survey baseline for Southern California laid out by Colonel Henry Washington in 1852. The monument marking the intersection between the baseline and the principal meridian is atop Mount San Bernardino, which appears on the horizon. (Image © 2009 131 DigitalGlobe, Image USDA Farm Service Agency, Image County of San Bernardino)



1	2	3	4	5	6
12	11	10	9	8	7
13	14	15	16	17	18
24	23	22	21	20	19
25	26	27	28	29	30
36	35	34	33	32	31

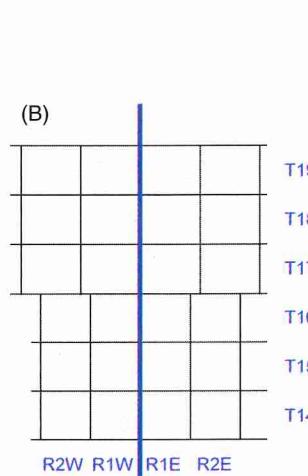


Figure 5.9 Portion of the Township and Range system (Public Lands Survey System) widely used in the western United States as the basis of land ownership (B). Townships are laid out in 6 mile squares on either side of an accurately surveyed Principal Meridian. The offset shown between T16 N and T17 N is needed to accommodate the Earth's curvature (shown much exaggerated). The square mile sections within each township are numbered as shown in (A).

familiar. It is often used to specify location, particularly in managing natural resources in the oil and gas industry and in mining, and in agriculture. Services have been built to convert PLSS locations automatically to and from latitude and longitude (Section 5.7).

5.7 Measuring the Earth: Latitude and Longitude

The most powerful systems of georeferencing are those that provide the potential for very fine spatial resolution, that allow distance to be computed between pairs of locations, and that support other forms of spatial analysis. The system of latitude and longitude is in many ways the most comprehensive and is often called the geographic system of coordinates, based on the Earth's rotation about its center of mass.

To define latitude and longitude, we first identify the axis of the Earth's rotation. The Earth's center of mass lies on the axis, and the plane through the center of mass perpendicular to the axis defines the equator. Slices through the Earth parallel to the axis, and perpendicular to the plane of the equator, define lines of constant longitude (Figure 5.10), rather like the segments of an orange. In 1884 a conference attended by delegates from 25 nations agreed that zero

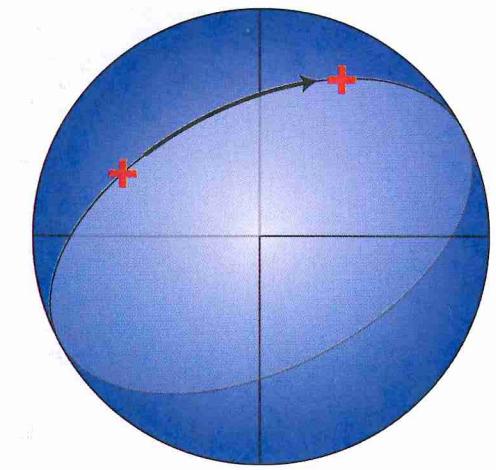


Figure 5.10 Definition of longitude. The Earth is seen here from above the North Pole, looking along the axis, with the equator forming the outer circle. The location of Greenwich defines the Prime Meridian. The longitude of the point at the center of the red cross is determined by drawing a plane through it and the axis, and measuring the angle between this plane and the Prime Meridian.

longitude, the *prime meridian*, should be defined by a line marked on the ground at the Royal Observatory in Greenwich, England; the angle between this slice and any other slice defines the latter's measure of longitude. Each of the 360 degrees of longitude is divided into 60 minutes and each minute into 60 seconds. But it is more conventional to refer to longitude by degrees East or West, so longitude ranges from 180 degrees West to 180 degrees East of the prime meridian. Finally, because computers are designed to handle numbers ranging from very large and negative to very large and positive, we normally store longitude in computers as if West was negative and East was positive; and we store parts of degrees using decimals rather than minutes and seconds. A line of constant longitude is termed a *meridian*.

Longitude can be defined in this way for any rotating solid, no matter what its shape, because the axis of rotation and the center of mass are always defined. But the definition of latitude requires that we know something about the shape. The geoid is defined as a surface of equal gravity formed by the oceans at rest, and by an imaginary extension of this surface under the continents; it has a complex shape that is only approximately spherical. A much better approximation or *figure of the Earth* is the *ellipsoid of rotation*, the figure formed by taking a mathematical ellipse and rotating it about its shorter axis (Figure 5.11). The term *spheroid* is also commonly used.

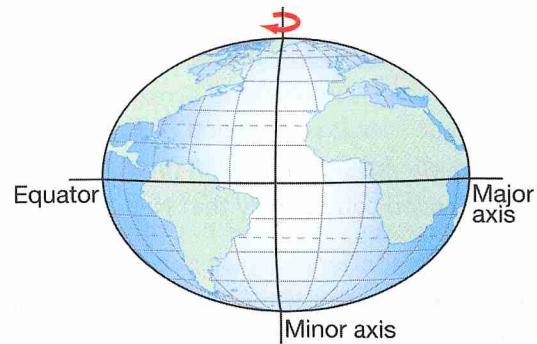


Figure 5.11 Definition of the ellipsoid, formed by rotating an ellipse about its minor axis (corresponding to the axis of the Earth's rotation).

The difference between the ellipsoid and the sphere is measured by its *flattening*, or the reduction in the minor axis relative to the major axis. Flattening is defined as

$$f = (a - b)/a$$

where a and b are the lengths of the major and minor axes, respectively (we usually refer to the *semi-axes*, or half the length of the axes, because these are comparable to radii). The actual flattening is about 1 part in 300.

The Earth is slightly flattened, such that the distance between the poles is about 1 part in 300 less than the diameter at the equator.

Much effort was expended over the past 200 years in finding ellipsoids that best approximated the shape of the Earth in particular countries, so that national mapping agencies could measure position and produce accurate maps. Early ellipsoids varied significantly in their basic parameters and were generally not centered on the Earth's center of mass. But the development of intercontinental ballistic missiles in the 1950s and the need to target them accurately, as well as new data available from satellites, drove the push to a single international standard. Without a single standard, the maps produced by different countries using different ellipsoids could never be made to fit together along their edges, and artificial steps and offsets were often necessary in moving from one country to another (navigation systems in aircraft would have to be corrected, for example).

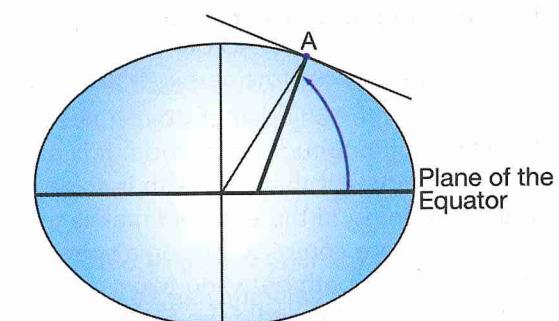
The ellipsoid known as WGS84 (the World Geodetic System of 1984) is now widely accepted, and North American mapping is being brought into

conformity with it through the adoption of the virtually identical North American Datum of 1983 (NAD83). It specifies a semi-major axis (distance from the center to the equator) of 6378137 m, and a flattening of 1 part in 298.257; and its line of zero longitude passes about 100 m to the east of the Greenwich Observatory. But many other ellipsoids remain in use in certain parts of the world, and much older data still adhere to earlier standards, such as the North American Datum of 1927 (NAD27). For points in the United States the difference between two points with identical latitude and longitude, but determined according to the NAD27 and NAD83 datums, can be as much as 100 m. Thus GIS users sometimes need to convert between datums, and functions to do that are commonly available.

We can now define latitude. Figure 5.12 shows a line drawn through a point of interest perpendicular to the ellipsoid at that location. The angle made by this line with the plane of the equator is defined as the point's latitude and varies from 90 South to 90 North. Again, south latitudes are usually stored as negative numbers and north latitudes as positive. Latitude is often symbolized by the Greek letter phi (ϕ) and longitude by the Greek letter lambda (λ), so the respective ranges can be expressed in mathematical shorthand as: $-180 \leq \lambda \leq 180$; $-90 \leq \phi \leq 90$. A line of constant latitude is termed a *parallel*.

It is important to have a sense of what latitude and longitude mean in terms of distances on the surface. Ignoring the flattening, two points on the same north-south line of longitude and separated by one degree of latitude are 1/360 of the circumference of the Earth apart, or about 111 km apart. One minute of latitude corresponds to 1.86 km, and also defines one nautical mile, a unit of distance that is still commonly used in navigation. One second of latitude corresponds to about 30 m. But things are

Figure 5.12 Definition of the latitude of Point A, as the angle between the equator and a line drawn perpendicular to the ellipsoid.



Applications Box 5.1

Newton, Descartes, and the Shape of the Earth

Newton's understanding of centrifugal forces led him to conclude that a planet such as the Earth rotating about its axis should bulge at the equator, where the diameter should therefore be greater than the distance between the poles. Descartes, on the other hand, argued that the reverse should be true—a greater distance between the poles than the diameter at the equator. Given the definition illustrated in Figure 5.12, lines of latitude should grow further apart as one moves away from the equator if Newton is right, and closer together if Descartes is right. In 1735 the French Academy of Sciences dispatched expeditions to Finland and Peru to make accurate deter-

minations of the distance between lines of latitude. The results proved conclusively that Newton was right.

One of the members of the expedition to Peru, Jean Godin, married a Peruvian and promised to take her home to France via a difficult route over the Andes and down the Amazon, at a time when the Spanish authorities in Peru were at odds with the Portuguese authorities in Brazil. He left first, and the 20-year story of Isabel's epic and ultimately successful adventure to rejoin him is admirably told in Robert Whitaker's *The Mapmaker's Wife*.

more complicated in the east-west direction, and these figures only apply to east-west distances along the equator, where lines of longitude are furthest apart. Away from the equator the length of a line of latitude gets shorter and shorter, until it vanishes altogether at the poles. The degree of shortening is approximately equal to the cosine of latitude, or $\cos \phi$, which is 0.866 at 30 degrees North or South, 0.707 at 45 degrees, and 0.500 at 60 degrees. So a degree of longitude is only 55 km along the northern boundary of the Canadian province of Alberta (exactly 60 degrees North).

In GIS, latitude and longitude are often expressed as decimals of degrees, rather than degrees, minutes, and seconds. It is helpful to know that the 5th decimal place of degrees of latitude is about 1 m on the Earth's surface. In GIS it is very uncommon to know positions to greater accuracy, so any additional decimal places that may be displayed or recorded are probably beyond the limits of accuracy and therefore meaningless.

Lines of latitude and longitude are equally far apart only at the equator; toward the poles lines of longitude converge.

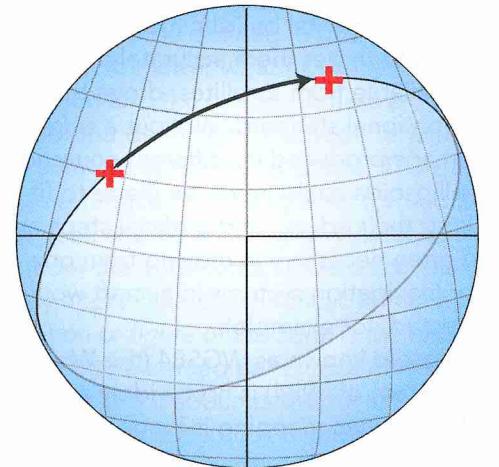
Given latitude and longitude, it is possible to determine distance between any pair of points, not just pairs along lines of longitude or latitude. It is easiest to pretend for a moment that the Earth is spherical because the flattening of the ellipsoid makes the equations much more complex. But on a spherical Earth the shortest path between two points is a great circle, or the arc formed if the Earth is sliced through the two points and through its center

(Figure 5.13; an off-center slice creates a small circle). The length of this arc on a spherical Earth of radius R is given by

$$R \arccos[\sin\phi_1 \sin\phi_2 + \cos\phi_1 \cos\phi_2 \cos(\lambda_1 - \lambda_2)]$$

where the subscripts denote the two points. For example, the distance from a point on the equator at longitude 90° East (in the Indian Ocean between Sri Lanka and the Indonesian island of Sumatra) and the

Figure 5.13 The shortest distance between two points on the sphere is an arc of a great circle, defined by slicing the sphere through the two points and the center (all lines of longitude, and the equator, are great circles). The circle formed by a slice that does not pass through the center is a small circle (all lines of latitude except the equator are small circles).



North Pole is found by evaluating the equation for $\phi_1 = 0, \lambda_1 = 90, \phi_2 = 90, \lambda_2 = 90$. It is best to work in radians (1 radian is 57.30 degrees, and 90 degrees is $\pi/2$ radians). The equation evaluates to $R \arccos 0$, or $R \pi/2$, or one quarter of the circumference of the Earth. Using a radius of 6378 km, this comes to 10,018 km, or close to 10,000 km (not surprisingly, since the French originally defined the meter in the late Eighteenth Century as one ten-millionth of the distance from the equator to the pole).

5.8 Projections and Coordinates

Latitude and longitude define location on the Earth's surface in terms of angles with respect to well-defined references: the prime meridian, the center of mass, and the axis of rotation. As such, they constitute the most comprehensive system of georeferencing and support a range of forms of analysis, including the calculation of distance between points, on the curved surface of the Earth. But many technologies for working with geographic data are inherently flat, including paper and printing, which evolved over many centuries long before the advent of digital geographic data and GIS. For various reasons, therefore, much work in GIS deals with a flattened or projected Earth, despite the price we pay in the distortions that are an inevitable consequence of flattening. Specifically, the Earth is often flattened because:

- Paper is flat, and paper is still used as a medium for inputting data to GIS by scanning or digitizing (see Section 9.3) and for outputting data in map or image form.
- Rasters (Section 8.2.2) are inherently flat, since it is impossible to cover a curved surface with equal squares without gaps or overlaps.
- Photographic film is flat, and film cameras are still used widely to take images of the Earth from aircraft to use in GIS.
- When the Earth is seen from space, the part in the center of the image has the most detail, and detail drops off rapidly, the back of the Earth being invisible; in order to see the whole Earth at once with approximately equal detail, it must be distorted in some way, and it is most convenient to make it flat.

The Cartesian coordinate system (Figure 5.14) assigns two coordinates to every point on a flat surface by measuring distances from an origin parallel to two axes drawn at right angles. We often talk of the two

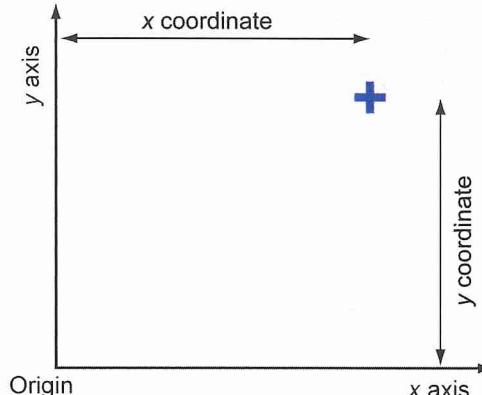


Figure 5.14 A Cartesian coordinate system, defining the location of the blue cross in terms of two measured distances from the origin, parallel to the two axes.

axes as x and y, and of the associated coordinates as the x- and y-coordinate, respectively. Because it is common to align the y-axis with North in geographic applications, the coordinates of a projection on a flat sheet are often termed easting and northing.

Although projections are not absolutely required, there are several good reasons for using them in GIS to flatten the Earth.

One way to think of a map projection, therefore, is that it transforms a position on the Earth's surface identified by latitude and longitude (ϕ, λ) into a position in Cartesian coordinates, (x, y). Every recognized map projection, of which there are many, can be represented as a pair of mathematical functions:

$$x = f(\phi, \lambda)$$

$$y = g(\phi, \lambda)$$

For example, the famous Mercator projection uses the functions:

$$x = \lambda$$

$$y = \ln \tan[\phi/2 + \pi/4]$$

where ln is the natural log function. The inverse transformations that map Cartesian coordinates back to latitude and longitude are also expressible as mathematical functions: in the Mercator case they are:

$$\lambda = x$$

$$\phi = 2 \arctan e^y - \pi/2$$

where e denotes the constant 2.71828. Many of these functions have been implemented in GIS, allowing users to work with virtually any recognized projection and datum, and to convert easily between them.

Two datasets can differ in both the projection and the datum, so it is important to know both for every dataset.

Projections necessarily distort the Earth, so it is impossible in principle for the scale (distance on the map compared to distance on the Earth; for a discussion of scale see Box 4.2) of any flat map to be perfectly uniform or for the pixel size of any raster to be perfectly constant. But projections can preserve certain properties. Two such properties are particularly important, although any projection can achieve at most one of them, not both:

- The *conformal* property, which ensures that the shapes of small features on the Earth's surface are preserved on the projection: in other words, that the scales of the projection in the x- and y-directions are always equal.
- The *equal area* property, which ensures that areas measured on the map are always in the same proportion to areas measured on the Earth's surface.

The conformal property is useful for navigation because a straight line drawn on the map has a constant bearing (the technical term for such a line is a *rhumb line* or *loxodrome*). The equal area property is useful for various kinds of analysis involving areas, such as the computation of the area of someone's property.

Besides their distortion properties, another common way to classify map projections is by analogy to a physical model of how positions on the map's flat surface are related to positions on the curved Earth. There are three major classes (Figure 5.15), but note that they do not cover all known projections:

- *Cylindrical* projections, which are analogous to wrapping a cylinder of paper around the Earth, projecting the Earth's features onto it, and then unwrapping the cylinder.
- *Azimuthal* or *planar* projections, which are analogous to touching the Earth with a sheet of flat paper.
- *Conic* projections, which are analogous to wrapping a sheet of paper around the Earth in a cone.

In each case, the projection's aspect defines the specific relationship, for example, whether the paper is wrapped around the equator or touches at a pole. Where the paper coincides with the surface the scale of the projection is 1, and where the paper is some distance outside the surface the projected feature will be larger than it is on the Earth. Secant projections attempt to minimize distortion by imagining the paper cutting through the surface, so that scale can be both greater and less than 1 (Figure 5.15; projections for which the paper touches the Earth and in which scale is always 1 or greater are called *tangent*).

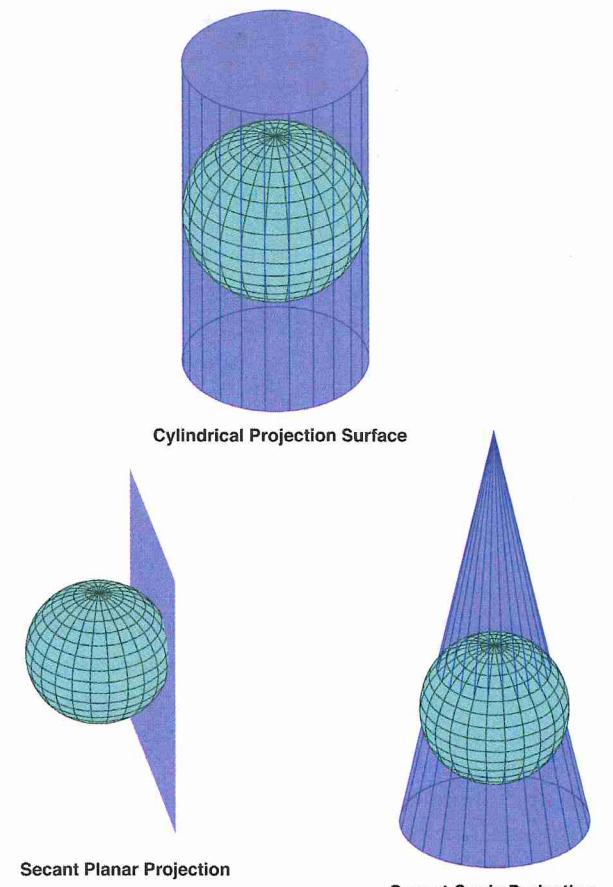


Figure 5.15 The basis for three types of map projections—cylindrical, planar, and conic. In each case a sheet of paper is wrapped around the Earth, and positions of objects on the Earth's surface are projected onto the paper. The cylindrical projection is shown in the *tangent* case, with the paper touching the surface, but the planar and conic projections are shown in the *secant* case, where the paper cuts into the surface. (Reproduced by permission of Peter H. Dana)

All three types can have either conformal or equal area properties, but of course not both. Figure 5.16 presents examples of several common projections and shows how the lines of latitude and longitude map onto the projection, in a (distorted) grid known as a *graticule*.

The next sections describe several particularly important projections in detail, together with the coordinate systems that they produce. Each is important to GIS, and users are likely to come across them frequently. The map projection (and datum) used to make a dataset is sometimes not known to the user of the dataset, so it is helpful to know enough about map projections and coordinate systems to make intelligent guesses when trying to combine such a dataset with other data. Several excellent books on map projections are listed in Further Reading.

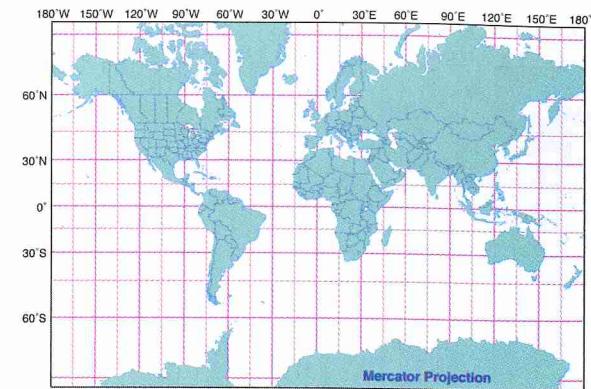
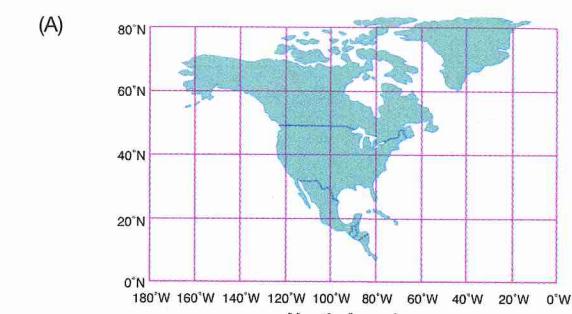


Figure 5.16 Examples of some common map projections. The Mercator projection is a tangent cylindrical type, shown here in its familiar equatorial aspect (cylinder wrapped around the equator). The Lambert Conformal Conic projection is a secant conic type. In this instance, the cone onto which the surface was projected intersected the Earth along two lines of latitude: 20 North and 60 North. (Reproduced by permission of Peter H. Dana)



(A) North America
Unprojected Latitude and Longitude

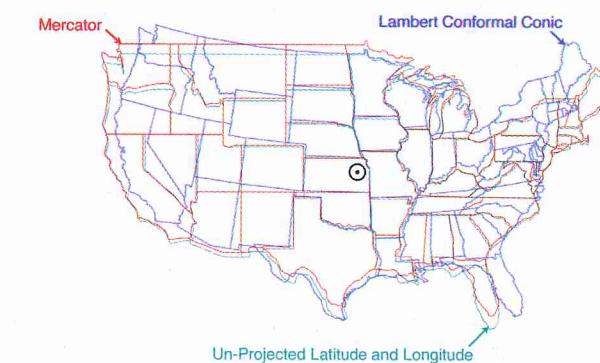
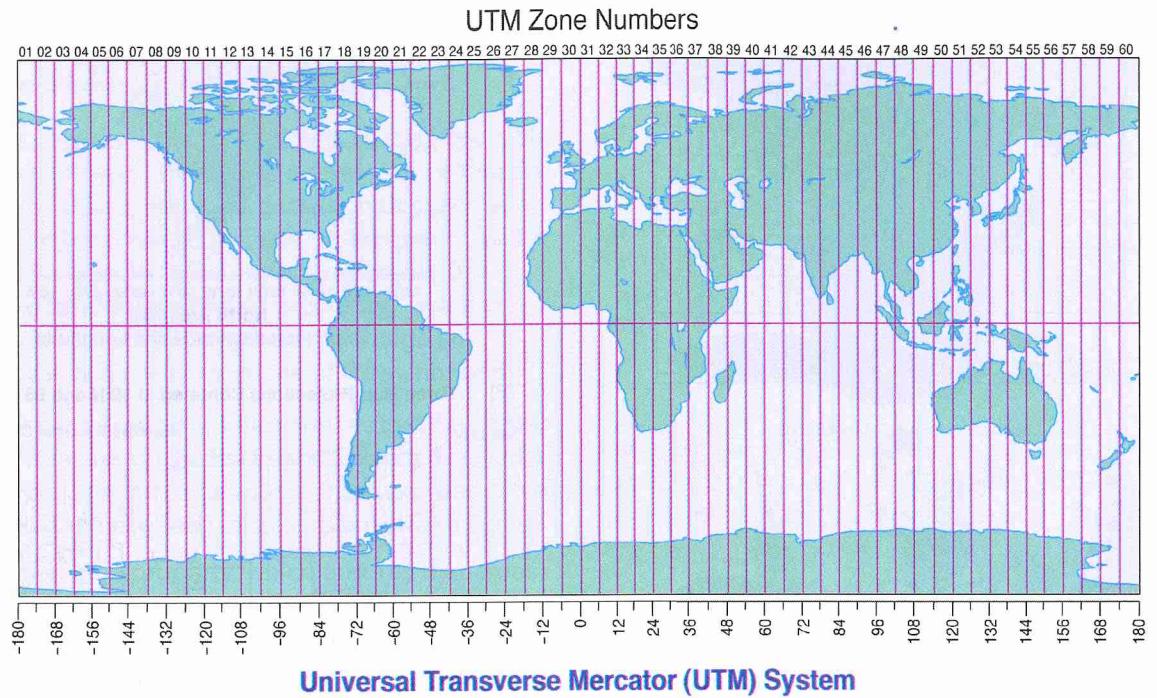


Figure 5.17 (A) The so-called unprojected or Plate Carrée projection, a tangent cylindrical projection formed by using longitude as x and latitude as y. (B) A comparison of three familiar projections of the United States. The Lambert Conformal Conic is the one most often encountered when the United States is projected alone and is the only one of the three to curve the parallels of latitude, including the northern border on the 49th Parallel. (Reproduced by permission of Peter H. Dana)

familiar projections of the United States: the Plate Carrée, Mercator, and Lambert Conformal Conic.

When longitude is assigned to x and latitude to y, a very odd-looking Earth results.

Serious problems can occur when doing analysis using this projection. Moreover, since most methods of analysis in GIS are designed to work with Cartesian coordinates rather than latitude and longitude, the same problems can arise in analysis when a dataset uses latitude and longitude, or so-called geographic coordinates. For example, a command to generate a circle of radius one unit in this projection will create a figure that is two degrees of latitude across in the north-south direction and two degrees of longitude across in the east-west direction. On the Earth's surface this figure is not a circle at all, and at high latitudes it is a very squashed ellipse.



Universal Transverse Mercator (UTM) System

Figure 5.18 The system of zones of the Universal Transverse Mercator system. The zones are identified at the top. Each zone is six degrees of longitude in width. (Reproduced by permission of Peter H. Dana)

It is wise to be careful when using a GIS to analyze data in latitude and longitude rather than in projected coordinates because serious distortions of distance, area, and other properties may result.

5.8.2 The Universal Transverse Mercator Projection

The UTM system is often found in military applications and in datasets with global or national coverage. It is based on the Mercator projection, but in the *transverse* rather than equatorial aspect, meaning that the projection is analogous to wrapping a cylinder around the poles rather than around the equator. There are 60 zones in the system, and each zone corresponds to a half cylinder wrapped along a particular line of longitude, each zone being 6 degrees wide. Thus Zone 1 applies to longitudes from 180 W to 174 W, with the half cylinder wrapped along 177 W; Zone 10 applies to longitudes from 126 W to 120 W centered on 123 W, and so on (Figure 5.18).

The UTM system is secant, with lines of scale 1 located some distance out on both sides of the central meridian. Because the projection is conformal, small features appear with the correct shape, and scale at each point is the same in all directions. Scale

is 0.9996 at the central meridian and at most 1.0004 at the edges of the zone, so the maximum distortion of distances using this projection is about 4/100 of 1%. Both parallels and meridians are curved on the projection, with the exception of the zone's central meridian and the equator. Figure 5.19 shows the major features of one zone.

The coordinates of a UTM zone are defined in meters and are set up such that the central meridian's easting is always 500,000 m (a false easting), so easting varies from near zero to near 1 million m. In the Northern Hemisphere the equator is the origin of northing, so a point at northing 5 million m is approximately 5000 km from the equator. In the Southern Hemisphere the equator is given a false northing of 10 million m and all other northings are less than this.

UTM coordinates are in meters, making it easy to make accurate calculations of short distances between points.

Because there are effectively 60 different projections in the UTM system, maps will not fit together across a zone boundary. Zones become so much of a problem at high latitudes that the UTM system is normally replaced with azimuthal projections centered on each pole (known as the UPS or Universal Polar Stereographic system) above 80 degrees latitude. The problem is

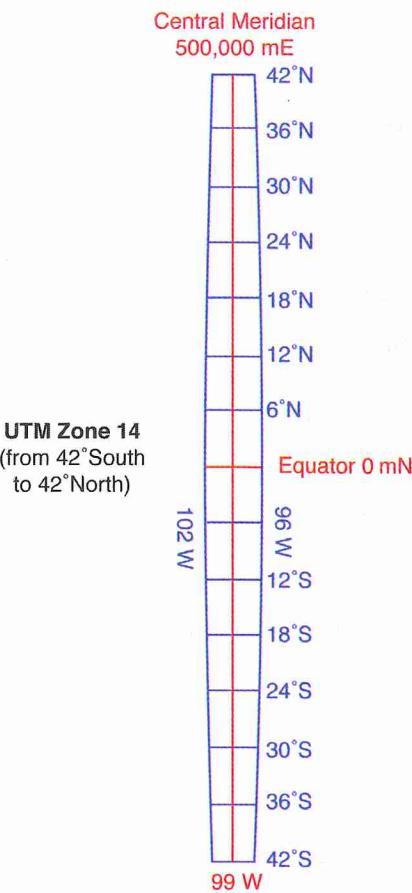


Figure 5.19 Major features of UTM Zone 14 (from 102 W to 96 W). The central meridian is at 99 W. Scale factors vary from 0.9996 at the central meridian to 1.0004 at the zone boundaries. See text for details of the coordinate system. (Reproduced by permission of Peter H. Dana)

especially critical for cities that cross zone boundaries, such as Calgary, Alberta, Canada (which crosses the boundary at 114 W between Zone 11 and Zone 12). In such situations one zone can be extended to cover the entire city, but this results in distortions that are larger than normal. Another option is to define a special zone, with its own central meridian selected to pass directly through the city's center. Italy is split between Zones 32 and 33, and many Italian maps carry both sets of eastings and northings.

UTM coordinates are easy to recognize because they commonly consist of a six-digit integer followed by a seven-digit integer (and decimal places if precision is better than a meter), and sometimes include zone numbers and hemisphere codes. They are an excellent basis for analysis because distances can be calculated from them for points within the same zone with no more than 0.04% error. But they are complicated enough that their use is effectively limited to

professionals (the so-called spatially aware professionals or SAPs) except in applications where they can be hidden from the user. UTM grids are marked on many topographic maps, and many countries project their topographic maps using UTM. It is therefore easy to obtain UTM coordinates from maps for input to digital datasets, either by hand or automatically using scanning or digitizing (Section 9.3.2). In the United States the UTM system is the basis for the National Grid, a system designed to produce a single, unique code for any location in the country (Box 5.2).

5.8.3 State Plane Coordinates and Other Local Systems

Although the distortions of the UTM system are small, they are nevertheless too great for some purposes, particularly in accurate surveying. Zone boundaries also are a problem in many applications because they follow arbitrary lines of longitude rather than boundaries between jurisdictions. In the 1930s each U.S. state agreed to adopt its own projection and coordinate system, generally known as State Plane Coordinates (SPC), in order to support these high-accuracy applications. Projections were chosen to minimize distortion over the area of the state, so choices were often based on the state's shape. Some large states decided that distortions were still too great, and so they designed their SPCs with internal zones (for example, Texas has five zones based on the Lambert Conformal Conic projection (Figure 5.20), while Hawaii has five zones based on the Transverse Mercator projection). Many GIS have details of SPCs already stored, so it is easy to transform between them and UTM, or latitude and longitude. The system was revised in 1983 to accommodate the shift to the new North American Datum (NAD83).

All U.S. states have adopted their own specialized coordinate systems for applications such as surveying that require very high accuracy.

Many other countries have adopted coordinate systems of their own. For example, the UK uses a single projection and coordinate system known as the National Grid that is based on the Oblique Mercator projection and is marked on all topographic maps. Canada uses a uniform coordinate system based on the Lambert Conformal Conic projection, which has properties that are useful at mid- to high latitudes, for applications where the multiple zones of the UTM system would be problematic.

Technical Box 5.2

A National System of Georeferencing: The U.S. National Grid

Recent disasters such as Hurricane Katrina have drawn attention to the benefits of having a simple method of assigning a unique code to every location in the nation. Many other countries have such systems. For example, in Great Britain the National Grid is administered by the Ordnance Survey of Great Britain and is popular among hikers and other users of the outdoors. The U.S. National Grid was first proposed in 2000, adopted by the Federal Geographic Data Committee as a standard in 2001, and is now widely used in the emergency services.

Each National Grid code is composed as follows:

the two-digit UTM zone number

a letter denoting an 8-degree latitude range

a letter denoting a 100-km-wide swath oriented east-west

a letter denoting a 100-km-wide swath oriented north-south

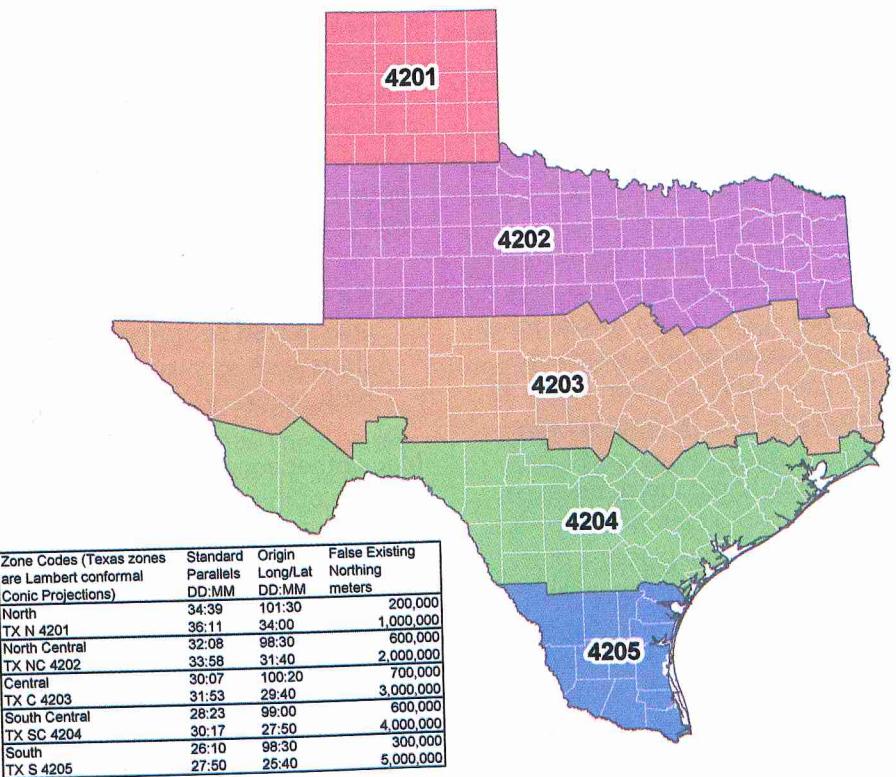
up to five decimal digits of UTM easting, omitting the first digit

the same number of decimal digits of UTM northing, omitting the first two digits

For example, the address 909 West Campus Lane, Goleta, California is at latitude 34 degrees 24 minutes 42.7 seconds North, 119 degrees 52 minutes 14.4 seconds West using the NAD83 datum (the use of one decimal place implies an accuracy of about 3 m). It lies in UTM Zone 11 at 236150 m easting, 3811560 m northing (rounding to the nearest 10 m, which is roughly the accuracy of measurement in this case). The National Grid system uses a variable number of northing and easting digits to reflect accuracy, so that shorter codes correspond to less accurately determined positions. Because accuracy is about 10 m in this case we should use only digits 2 through 5 of the six-digit easting and 3 through 6 of the seven-digit northing, so the resulting code is:

11SKU36151156

Figure 5.20 The five State Plane Coordinate zones of Texas. Note that the zone boundaries are defined by counties, rather than parallels, for administrative simplicity.



5.9 Measuring Latitude, Longitude, and Elevation: GPS

The Global Positioning System and its analogs (GLONASS in Russia and the Galileo system in Europe) have revolutionized the measurement of position, for the first time making it possible for people to know almost exactly where they are anywhere on the surface of the Earth. Previously, positions had to be established by a complex system of relative and absolute measurements. If one was near a point whose position was accurately known (a survey monument, for example), then position could be established through a series of accurate measurements of distances and directions starting from the monument. But if no monuments existed, then position had to be established through absolute measurements. Latitude is comparatively easy to measure, based on the elevation of the sun at its highest point (local noon), or on the locations of the sun, moon, or fixed stars at precisely known times. But longitude requires an accurate method of measuring time, and the lack of accurate clocks led to massively incorrect beliefs about positions during early navigation. For example, Columbus and his contemporary explorers had no means of measuring longitude, and believed that the Earth was much smaller than it is and that Asia was roughly as far west of Europe as the width of the Atlantic. The strength of this conviction is still reflected in the term we use for the islands of the Caribbean (the West Indies) and the first rapids on the St. Lawrence in Canada (Lachine, or China). The fascinating story of the accurate measurement of longitude is recounted by Sobel.

The GPS consists of a system of 24 satellites (plus some spares), each orbiting the Earth every 12 hours on distinct orbits at a height of 20,200 km and transmitting radio pulses at precisely timed intervals. To determine position, a receiver must make exact calculations from the signals, the known positions of the satellites, and the velocity of light. Positioning in three dimensions (latitude, longitude, and elevation) requires that at least four satellites are above the horizon, and accuracy depends on the number of such satellites and their positions (if elevation is not required, then only three satellites need be above the horizon). Several different versions of GPS exist, with distinct accuracies.

A simple GPS, such as one might buy in an electronics store for \$100, or install as an optional addition to a laptop, cell phone, PDA (personal digital assistant), or vehicle, has an accuracy within 10 m. This accuracy will degrade in cities with tall buildings, or under trees, and GPS signals will be lost entirely under bridges or indoors. Differential GPS (DGPS) combines GPS signals from satellites with correction signals received

via radio or telephone from base stations. Networks of such stations now exist, at precisely known locations, constantly broadcasting corrections; corrections are computed by comparing each known location to its apparent location determined from GPS. With DGPS correction, accuracies improve to 1 m or better. Even better accuracies are possible by using various sophisticated techniques or by remaining fixed and averaging measured locations over several hours.

GPS is very useful for recording ground control points when building GIS databases, for locating objects that move (for example, combine harvesters, tanks, cars, and shipping containers), and for direct capture of the locations of many types of fixed objects, such as utility assets, buildings, geological deposits, and sample points. Other applications of GPS are discussed in Chapter 11.

Some care is needed in using GPS to measure elevation. First, accuracies are typically poorer, and a position determined to 10 m in the horizontal may be no better than plus or minus 50 m in the vertical. Second, a variety of reference elevations or vertical datums are in common use in different parts of the world and by different agencies; for example, in the United States the topographic and hydrographic definitions of the vertical datum are significantly different.

5.10 Converting Georeferences

GIS are particularly powerful tools for converting between projections and coordinate systems because these transformations can be expressed as numerical operations. In fact, this ability was one of the most attractive features of early systems for handling digital geographic data and drove many early applications. But other conversions, for example, between place-names and geographic coordinates, are much more problematic. Yet they are essential operations. Almost everyone knows their mailing address and can identify travel destinations by name, but few are able to specify these locations in coordinates, or to interact with geographic information systems on that basis. GPS technology is attractive precisely because it allows its user to determine his or her latitude and longitude, or UTM coordinates, directly at the touch of a button.

Methods of converting between georeferences are important for:

- Converting lists of customer addresses to coordinates for mapping or analysis (the task known as geocoding; see Box 5.3).
- Combining datasets that use different systems of georeferencing.

Geocoding: Conversion of Street Addresses to Coordinates

Geocoding is the name commonly given to the process of converting street addresses to latitude and longitude, or some similarly universal coordinate system. It is widely used, as it allows any database containing addresses, such as a company mailing list or a set of medical records, to be input to a GIS and mapped. Geocoding requires a database containing records representing the geometry of street segments between consecutive intersections, and the address ranges on each side of each segment (a *street centerline database*; see Section 8.2.3.3). Addresses are geocoded by finding the appropriate street segment record and estimating a location based on

linear interpolation within the address range. For example, 950 West Broadway in Columbia, Missouri, lies on the side of the segment whose address range runs from 900 to 998, or $50/98 = 51.02\%$ of the distance from the start of the segment to the end. The segment starts at 92.3503 West longitude, 38.9519 North latitude, and ends at 92.3527 West, 38.9522 North. Simple arithmetic gives the address location as 92.3515 West, 38.9521 North. Four decimal places suggests an accuracy of about 10 m, but the estimate also depends on the accuracy of the assumption that addresses are uniformly spaced, as well as on the accuracy of the street centerline database.

- Converting to projections that have desirable properties for analysis, for example, no distortion of area.
- Searching the Internet or other distributed data resources for data about specific locations.
- Positioning GIS map displays by recentering them on places of interest that are known by name (these last two are sometimes called *locator services*).

The oldest method of converting georeferences is the *gazetteer*, the name commonly given to the index in an atlas that relates place-names to latitude and longitude, and to relevant pages in the atlas where information about that place can be found. In this form the gazetteer is a useful locator service, but it works only in one direction as a conversion between georeferences (from place-name to latitude and longitude). Gazetteers have evolved substantially in the digital era, and it is now possible to obtain large databases of place-names and associated coordinates and to access services that allow such databases to be queried over the Internet (e.g., the Alexandria Digital Library gazetteer, www.alexandria.ucsb.edu; the U.S. Geographic Names Information System, www.geonames.usgs.gov). It is also possible to employ sophisticated software to detect place-names in text and to convert them to georeferences (e.g., www.metacarta.com).

5.11 Geotagging and Mashups

Online gazetteers, geocoding sites, and other services for converting georeferences have made it extremely easy to determine the geographic locations associated with many types of online data. It is now easy, for example, to take a mailing list containing street addresses, geocode them, and create maps. Services such as Google Earth, Google Maps, and their Yahoo! and Microsoft cousins provide the mapping capabilities and are readily invoked through their application programming interfaces (APIs). The term *mashup*, which derives originally from the popular-music industry, has been adopted as a way of describing the joining of two or more online services to create something that neither was able to do on its own. One of the best known mashups is www.housingmaps.com, which combines real-estate (property) listings from Craig's List (www.craigslist.org) with cartographic data from Google Maps (maps.google.com). A similar service for the UK is provided by www.nestoria.co.uk, and the www.londonprofiler.org site allows georeferenced property listings to be viewed against a range of thematic data. Today hundreds of thousands of such mashups have been created, many of them dealing with georeferenced data.

In many online services, such as Wikipedia (www.wikipedia.org), georeferences are embedded

Vincent Tao, Helping to Direct Microsoft's Virtual Earth Team

One of many graduates of Wuhan University who have gone on to successful careers in GIScience, Vincent Tao (Figure 5.21) obtained his Ph.D. from the University of Calgary in 1997. After several years there on the faculty, he moved to a Canada Research Chair at York University in 2001. His company GeoTango, which he launched in 2002, developed GlobeView, one of the earliest virtual globes, and SilverEye, the first software to build 3-D models from single images. GeoTango was acquired by Microsoft in 2005. Since then Tao has been director with the Microsoft Virtual Earth team, responsible for global product strategy and software development in local search and mapping, location-based services, and related technologies. He sees Virtual Earth developing as a powerful tool in many areas of human activity, from shopping and marketing to tourism and entertainment. Well known in the GIScience community, he has received many honors and awards and is a frequent speaker at conferences worldwide. In May 2009 he announced his departure from Microsoft to pursue opportunities in consulting.



Figure 5.21 Vincent Tao.
(Photo Courtesy of Vincent Tao)

(x and y) and in UTM. Note that the UTM coordinates in meters are rounded to the nearest hundred, given the limited spatial resolution of street intersections.

To convert the alarm map to UTM, we look for the simplest possible equations and often adopt an *affine transformation* of the form:

$$\text{UTM north} = a + bx + cy$$

$$\text{UTM east} = d + ex + fy$$

where a through f are values to be determined from the tie points. Standard methods exist in GIS for doing this, yielding the following values:

$$\text{UTM north} = 4744135 + 501.0x + 1222.2y$$

$$\text{UTM east} = 473668 + 1207.9x - 464.6y$$

Finally, these two equations are used to convert all of the fire alarm data points to UTM so that the analysis can proceed.

Figure 5.22 shows another example of georegistration. The highlighted area was the focus of a proposed project on the campus of the University of California, Santa Barbara, which was intended to restore the area to its natural wetland state. A number of wells were drilled to monitor groundwater, and accurately located using GPS. It was then necessary to merge these data with the campus database of detailed high-resolution imagery and ground elevations, which used a different coordinate system. Registration points were established on features, such as the corners of buildings and street lights, which could be easily located in both datasets. Note the difficulty in this case of registering

Table 5.2 Registration points, with coordinates in both systems.

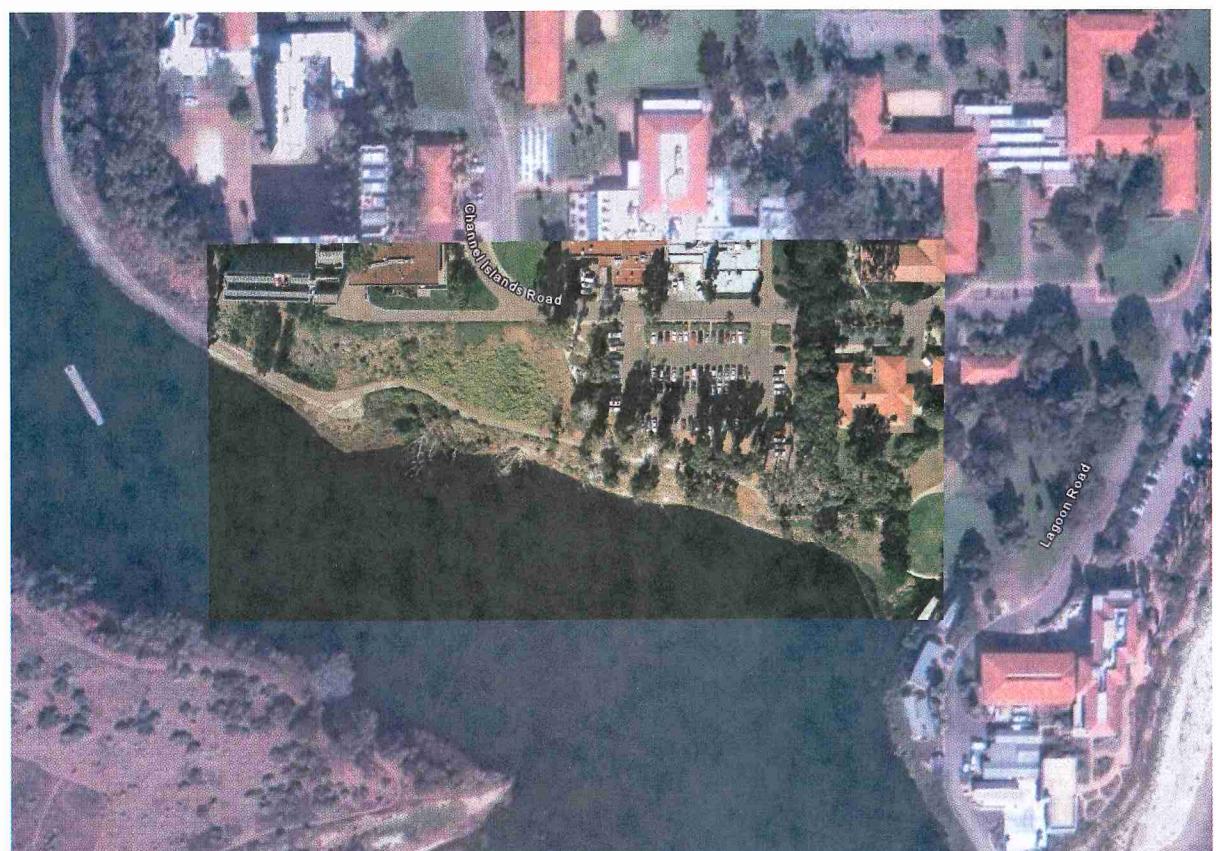
Intersection	x	y	UTM north	UTM east
Oxford and Sanatorium	2.90	9.80	4757500	472700
Wonderland and Southdale	4.86	5.96	4753800	476500
Wharncliffe and Stanley	7.32	8.56	4758200	478600
Oxford and Wharncliffe	7.58	9.67	4759800	478400
Wellington and Southdale	8.24	4.90	4754300	481500
Highbury and Hamilton	11.32	6.67	4758200	484100
Trafalgar and Clarke	13.17	7.56	4759700	486200
Adelaide and Dundas	9.49	8.59	4759400	481100
Highbury and Fanshawe	11.50	12.68	4765400	481500
Richmond and Huron	8.28	10.73	4761500	478800

the lower part of the highlighted area, where there are no readily identifiable features to use for registration.

The figure shows the result, displayed in Google Earth. Note the obvious displacement between the highlighted area and the Google Earth imagery,

but the apparent agreement between the highlighted area and the Google Earth roads. It seems clear that the roads and the highlighted area are registered much more accurately to the Earth than the Google Earth imagery, which is misplaced by more than 10 m.

Figure 5.22 Registration of a high-resolution image (highlighted) of part of the campus of the University of California, Santa Barbara, shown as a Google Earth mashup. (© 2009 Google, Image U.S. Geological Survey)



5.13 Summary

This chapter has looked in detail at the complex ways in which humans refer to specific locations on the planet and how they measure locations. Any form of geographic information must involve some kind of georeference, and so it is important to understand the common methods, together with their advantages and disadvantages. Many of the benefits of GIS rely on accurate georeferencing—the ability to link different items of information together through common geographic location; the ability to measure distances and areas on the Earth's surface, and to perform more complex forms of analysis; and the ability to communicate geographic information in forms that others can understand.

Georeferencing was introduced in early societies to deal with the need to describe locations. As humanity has progressed, we have found it more

and more necessary to describe locations accurately, and over wider and wider domains, so that today our methods of georeferencing are able to locate phenomena unambiguously and accurately anywhere on the Earth's surface. Today, with modern methods of measurement, it is possible to direct another person to a point on the other side of the Earth to an accuracy of a few centimeters. This level of accuracy and referencing is regularly achieved in such areas as geophysics and civil engineering.

But georeferences can never be perfectly accurate, and it is always important to know something about spatial resolution. Questions of measurement accuracy are discussed at length in Chapter 6, and Section 6.2.3 deals with techniques for representation of phenomena that are inherently fuzzy, such that it is impossible to say with certainty whether a given point is inside or outside the georeference.

Questions for Further Study

1. Visit your local map library, and determine: (1) the projections and datums used by selected maps; (2) the coordinates of your house in several common georeferencing systems.
2. Summarize the arguments for and against a single global figure of the Earth, such as WGS84.
3. Access several online mapping services such as Google Maps, Google Earth, or Yahoo! Maps. What projection does each one use, how does it display map scale, and how does map scale vary over the map?
4. Chapter 14 discusses various forms of measurement in GIS. Review each of those methods and the issues involved in performing analysis on databases that use different map projections. Identify the map projections that would be best for measurement of (1) area, (2) length, and (3) shape.

Further Reading

- Bugayevskiy, L.M. and Snyder, J.P. 1995. *Map Projections: A Reference Manual*. London: Taylor and Francis.
- Kennedy, M. 1996. *The Global Positioning System and GIS: An introduction*. Chelsea, MI: Ann Arbor Press.
- Maling, D.H. 1992. *Coordinate Systems and Map Projections*. 2nd ed. Oxford: Pergamon.
- Snyder, J.P. 1997. *Flattening the Earth: Two Thousand Years of Map Projections*. Chicago: University of Chicago Press.
- Sobel, D. 1995. *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of His Time*. New York: Walker.
- Steede-Terry, K. 2000. *Integrating GIS and the Global Positioning System*. Redlands, CA: ESRI Press.
- Whitaker, R.W. 2004. *The Mapmaker's Wife: A True Tale of Love, Murder, and Survival in the Amazon*. New York: Basic Books.