INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE, 2017
VOL. 31, NO. 4, 637–657
http://dx.doi.org/10.1080/13658816.2016.1215466

REVIEW ARTICLE

Calculating on a round planet

Nicholas R. Chrisman

Sciences géomatiques, Université Laval, Québec, Canada; School of Science, RMIT University, Melbourne, Australia

ABSTRACT

Despite incontrovertible scientific evidence to support a round Earth, GIS software implementation typically models the Earth with flat map projections. This choice has consequences that vary from mild to severe. This paper explores solutions that range from moderate measures to correct for map projection errors to radical revisions of standard practice that place all calculations on the ellipsoid. On examination, the best practice requires a distinct projection choice for each geometric operation. Multiple reasons, from the technical to ethical, justify revised practices and increased vigilance for the errors induced by inappropriate projections.

ARTICLE HISTORY
Received 5 May 2016
Accepted 17 July 2016

KEYWORDS
Coordinates; projections; geodetic calculations; ellipsoidal distance; ellipsoidal area; buffers on the ellipsoid; line intersection on the ellipsoid

Introduction

The Earth is round, and that fact should surprise nobody. Whether general acceptance occurred 500 or 2000 years ago (Snyder 1987, p. 3) makes little difference to the current practice of geographic information technologies. Despite the overwhelming evidence for a round Earth, the software practices and technology of geographic information systems (GIS) remain firmly attached to flat surfaces, indexed by Cartesian coordinates. Tobler remarked some years ago that none of the textbooks on spatial analysis treated the Earth as 'homeomorphic to a sphere' (Tobler 1993, p. 19). Blindness to the obvious is thus not limited to software designers, but this paper will direct its attention to the software community and GIS practitioners. As one of the guilty parties who designed software in a prior era, I will discuss some of the reasons why the original developments were coded in planar geometry. Then I will describe the disadvantages and limitations of the current approach, and finally discuss the remedies and changes required to repair the damage and to start calculating on a round Earth.

Calculating on a plane

It is accepted practice to calculate in planar Cartesian coordinates in many fields from scientific disciplines such as physics to practical settings in engineering surveying. Analytical geometry, like all technology, has multiple origins. Descartes (1637) made

CONTACT Nicholas R. Chrisman nchrisman@fastmail.fm sciences géomatiques, Université Laval, Québec, QC, Canada

significant contributions that established the notation and procedures that have become common (and therefore do not require citation by all but the most punctilious). For example, the naming of the coordinate axes as x,y,z was introduced in this monumental book (Descartes 1637, p. 10). Descartes offered his mathematical presentation as a demonstration of his scientific method, one based on foundationalism. Descartes chose the plane to demonstrate the application of algebraic (hence numerical) treatment to the realm of geometry. This discipline had firm rules about construction limited to straight edges and compasses, a non-numerical approach. Descartes was not motivated by any urge to model the Earth, although the word geometry derives etymologically from Earth measurement.

Subsequently, analytical geometry became an underpinning of many fields in science and engineering. A sign of this implicit acceptance comes in the school curriculum; this technology is introduced to young students around the world. What was a matter of natural philosophy became a routine matter for primary students [for instance: the controversial Common Core in the United States introduces planar coordinate systems in year 5 (Common Core 2015); the English National Standards (Department of Education 2013, p. 29) introduce 2D coordinates in year 4]. Analytical geometry is taught in the planar case, typically, using just x and y axes; perimeter and area calculations are taught in primary school – in the case of England, year 5 (Department of Education 2013, p. 36). As a mathematical exercise, it makes sense to work with simple models. Planar geometry allows simple calculations, providing a good fit for elementary teaching.

Unlike the Earth, maps are flat, a property inherited from the various support technologies from parchment to printing presses to handheld tablet screens. Centuries of practice have established means to reduce measurements observed in the field (on ship or on land) to a portable planar surface. Mapping technique absorbed a series of technologies all dependent on graphics arts, from etching plates, lithography, and photographic reproduction. Each has been just as flat as its predecessor. Therefore, when early developments occurred in digital cartography, adopting a flat surface in numerical form was no big surprise. Planar geometry was embodied directly in the orthogonal axes of digitizer tables (Boyle 1964) and in the software to reproduce maps on line-printers (Horwood *et al*. 1963). The choice had already been made, since these techniques were all built on digital conversion of the stock of printed maps.

The first generation of automated cartography worked for the most part on single map sheets, not on massive worldwide databases. For the period, the choice was well-justified. Even the Canada Geographic Information System (CGIS) (Tomlinson 1968), with a vision to cover a continental extent, stuck close to the planar model of its raster scanner since each vector was reduced to incremental coordinates (Tomlinson 1998, p. 23). Tomlinson attributes these codes to Galton, though they are more commonly termed 'Freeman codes' in the computer graphics literature (Freeman 1961). Wherever the origin, this technique lies on the plane.

World Data Banks and cartographic projection software

Not all data remained tied to a flat map. Early GIS and automated cartography built archives of data for whole continents and soon the whole world. In Canada, CGIS was conceived to cover a continent with a spatial index (Morton 1966). This index divided into degree-based units (in the illusion of planar geometry), then, inside each map sheet, it used planar coordinates for the linework (in incremental coordinates tied to a pixel on a flat map). Therefore, the actual area measurements were made in planar mode, not on a sphere or ellipsoid.

In the same period, the US Central Intelligence Agency (CIA) developed World Data Bank I (WDB-I) in a very short period to respond to some short term needs (Schmidt 1969). Separate map sheets were digitized, and back-projected out of their specific projections into a homogenous framework of latitude-longitude. My own work was influenced by access to WDB-I, as I set out to convert it to a topological data structure (Chrisman 2006, p. 105). Working with this source, and its follow-on WDB-II (Anderson *et al*. 1977), led me to remark that WDB-II was 'no more structured than spaghetti on a plate' (Chrisman 1974a, p. 167). However, I did not give them credit for representing the whole planet in an ellipsoidal frame of reference (latitude-longitude in radians). [The datum for WDB-I was not clearly specified, and its accuracy levels were probably not high enough to distinguish between the various geodetic datums of the period.] The CAM software (written by IBM for CIA) provided elaborately detailed sixth-order expansions for ellipsoidal projections (personal inspection of the source code, 1975). Still the goal was to produce printed maps using traditional reproduction technologies. Scribing onto peel-coat material was the preferred output, so it was not important to fill polygons. A cartographer would lift the inside of the polygon as required for the particular color scheme. This was an era of computer-assisted cartography, not attuned to thematic mapping driven by attribute tables, let alone a full GIS.

In my own software, POLYVRT (Chrisman 1974b), I implemented the capability to project these data sources into planar coordinates so polygons could be used by the existing cartographic packages of the era. None of this software (such as SYMAP) had any projection capability. I implemented projections on a spherical Earth model, using the formulae from textbooks such as Richardus and Adler (1972). Considering that the databases derived from sources with errors in the tens of kilometres, it made little sense to insist on geodetic accuracy. I used single-precision code for POLYVRT and much of my later work. On the 60-bit floating point hardware that I used for POLYVRT, that made for very little limitation. However, as the industry converged on a 32-bit word, single precision set practical limits on resolution and project extent (Chrisman 1984).

Software for cartographic projections (and their inverses) became one of the elements of the GIS toolkit. Snyder's (1987) magisterial collection of map projections was written to support this development. The General Cartographic Transformation Package (GCTP) code written by Elassal (1987) under the direction of Snyder permeated the industry, in part because it was placed in the public domain. These efforts used the full ellipsoidal derivations of cartographic projections required for topographic mapping. GCTP calculated entirely in double precision (64 bit) floating point, which was considered somewhat extravagant (for the era) though technically valid. Few software environments maintained 64-bit floating point resolution in storage until much later.

Coding early GIS

The software packages called GIS emerged in the 1980s from this background. The prototype ODYSSEY was written on timesharing computers like the Digital Equipment

VAX (Chrisman *et al.* 1992). Storage was limited to single precision (effectively six digits of resolution), but we found that we needed to calculate intersections in double precision to avoid overflow and other computational issues. We used planar algorithms, expecting each project (and database) to be in some local coordinate system and projection. ODYSSEY implemented basic projections intended for visualization rather than precise georeferencing. Following the conversion to NAD83, I converted to much more sophisticated implementations in full 64-bit floating point and code from GCTP.

As far as I can determine from outside the shield of non-disclosure agreements and license conditions, the commercial packages of the 1980s (ARC/INFO, System 9, Intergraph MGE) were just as limited to single precision floating point storage in the early implementations. Planar geometry was the rule, with only limited dissenters. Lukatela (1987) presented a geodetically adept computational framework called HIPARCHUS, to be discussed below.

Current situation

Since the mid-1980s when the first generation of GIS software emerged on workstation platforms, the speed of computing devices has increased by a factor of many thousand. The conceptual underpinnings of the software have barely budged. Computation still relies on projected planar coordinates. Official coordinate systems remain entrenched where they were in the 1930s, based on conformal map projections. Eventually, most software platforms moved to double precision, for reasons of resolution.

One development in the 1990s bears some mention. GIS implementations began to rely on greater integration with database engines. To provide for indexing, floating point representations created troubles best solved by some form of integer coordinates. For example, Oracle Spatial originally used a latitude-longitude octree called the Helical Hyperspatial Code (Varma *et al.* 1990); ArcSDE imposed integer representation of coordinates. Granularity was often established by default, despite some consequences on project extent or resolution. In most cases, the database indexing requirement constrained choice of projections. The index was built one time, based on a particular reference frame. This index would not survive local rotation, often the result of changing projection. This limitation persists in more recent developments, as far as can be determined outside the cloak of product secrecy.

Representing global coordinates

As in many situations, convention plays a huge role in developing standards for geographic coordinates. ISO (International Standards Organization) and OGC (Open Geospatial Consortium) have combined to codify the practice of spatial referencing by coordinates (ISO 19111:2007, OGC (2010)).

The convention of positioning objects on the Earth with two angles dates back to Hipparchus (second century BCE) and Claudius Ptolemy (150 CE). Latitude was easier to establish than longitude. The current system became worldwide with the 1884 international agreement on the Meridian of Greenwich (Howse 1985). This angular notation has the weight of convention, but it does have peculiar numerical properties. Trigonometric functions (such as tangent) return useless (infinite) results at right angles, and therefore programmers must pay attention to avoid extraneous error messages and wild results.

Certain programmers and GIS users over the years have blurred the distinction between a geodetic coordinate system and a Cartesian, projected coordinate system. If degrees are treated as distances, a distorted pseudo-projection results (often termed Plate Carrée). Of course, distances calculated by the Pythagorean Theorem from angular measures bear no relation to the Earth, and areas in square degrees are meaningless. This has not stopped whole scientific communities built upon around dubious databases of these so-called 'square degrees' (references suppressed to protect the guilty).

Lukatela (1987, 2000, p. 118) has described a scheme for representing angular values of a geodetic point by the direction cosines of the ellipsoid normal. With this representation, useful properties can be computed without trigonometrical functions. In addition, fewer checks for valid operations (division by zero; infinite results) are required. While this representation may be technically superior, adoption remains limited.

Much of modern geodetic calculations are not carried out with angular measures, but on a Cartesian triaxial system, centered on the center of mass of the Earth, measured in metres (IERS 2013). Called the International Terrestrial Reference System (ITRS), it is implemented by the International Terrestrial Reference Frame (ITRF). ITRF is regularly redefined as more precise measurements become available. The international agreements behind ITRS and ITRF only date from 1991, and adoption by mapping authorities lagged due to attachment to local datums from prior eras. One tricky part of ITRF is the 'no-net-rotation' condition, which is the attempt to remain Earth-Centered, Earth-Fixed (ECEF) in the face of tectonic movements. ITRF removes the safety of a firmly anchored brass marker as the origin of surveying.

Geopositioning using satellite methods works in a Cartesian ECEF framework, since timing signals provide distances for trilateration. Geodetic XYZ does not recognize the surface of the ellipsoid in any special way. Distances calculated using Cartesian measures will cut through land, water, crust, mantle or atmosphere with no distinction. There is, consequently, no surface on which to measure area. It would take some massive rethinking of cartographic conventions to switch to Cartesian ECEF for GIS operations. Results obtained in geodetic XYZ are usually converted to angular coordinates on an established datum, then usually converted to projected coordinates (confusingly called xy (or more properly, Easting and Northing), though decidedly not on the same geodetic axes).

Through some kind of collective amnesia, after all these manipulations the resulting projected coordinates become an infinite plane. The finite extent of the Earth is forgotten. As Tobler (1993, p. 19) notes, working on the sphere is counter-intuitive:

> Should everyone know that the circumference of a circle increases as two π times the sine of the radius, (which means it eventually goes to zero), and that the area of a circle increases in proportion to the square of the sine of one half of the circular radius? Or that the circumference of a circle on the surface of an ellipsoid depends not only on the circle's radius but also on where one puts the center (Blaschke 1949)?

Many features of the actual round Earth overrule the simplicity of planar assumptions. This essay will review what goes wrong when GIS data gets separated from the Earth model, then present some solutions.

What goes wrong

Planar coordinate systems inevitably distort something (Grafarend *et al*. 2014, p. 1). Projection choice is fundamentally a choice of which property to preserve by distorting everything else. It also determines how difficult it will be to correct these distortions.

Mercator-bashing has become a routine phenomenon in certain sectors. Journalists continue to present true areas as some kind of revelation (Taylor 2015). Sadly, none of this attention reduces the dead weight of cylindric map projections on the GIS sector. Mercator is not the only projection that distorts.

There are a few prominent cases where an error in understanding map projections has led to major mistakes. On 3 May 2003, The Economist (2003a) published an article showing circles around the missile base in North Korea. Due to a fundamental misunderstanding of the Mercator projection, these circles on the map presented a gross underestimate of the range of these weapons systems (still not operational according to the article). The Economist (2003b) issued a correction in 12 days with the subtitle 'Flat-earth thinking' (with the result that the original mistake is no longer visible apart from bandit copies around the network). The Economist stayed with their projection choice, forcing a radically different geometry for missile ranges over the North Pole. A switch to an azimuthal equidistant projection would have permitted them to retain the circles of missile range. In an article dated 12 December 2012, The Economist (2012) had switched to an azimuthal for their repeat coverage of this topic. Figure 1 shows a similar image (from a Creative Commons source). Washington Post (Fisher 2013) also chose an azimuthal projection. Perhaps the outcry of geospatial professionals may have actually reached through to the journalism sector.

There are other major policy issues that are influenced by calculating on the wrong surface. At the global scale, Usery and others (2003) consider the errors that accumulate in global datasets due to raster resolution interacting with projection issues. Many more local policies are implemented by a payment based on the area of the land parcel in question. As Nievergelt (1997, p. 2) illustrates with a quotation from the Rhind papyrus, the pharaohs of Egypt were concerned with measurement of area around 3700 years ago. Most North American jurisdictions assess a portion of property tax based on area of parcel. In the opposite fiscal direction, the Common Agricultural Policy of the European Union provides subsidies to farmers based on eligible land area (European Commission 2013; EU 2013). The 'Single Area Payment Scheme' remains in force through 2020 to be replaced by a 'Basic Payment Scheme' (EU 2013). Curiously, the massive regulation provides no definition for hectares or how they are to be measured. In England, the name change has already occurred, tying payments to hectares of land and other conditions (RPA 2015). Rural Payments Agency (RPA) mapping procedure appears to operate on the UK National Grid (a Transverse Mercator), although RPA will accept area measurements (including slope area) based on independent certified surveys (RPA 2015, Annex 3, p. 49). Thus, England may have a patchwork with some areas on a projection plane and other areas from field survey. It is unlikely that fields on the margins of the Transverse Mercator will opt for field survey if the result is smaller than the inflationary distortion provided by the projection. Those in the center of the zone will have an incentive to pay for a survey. There would be a neat research project to verify if this incentive is strong enough to influence behavior.

Figure 1. North Korean missile range, depicted on wikimedia commons using an orthographic projection. Produced by 'Cmglee' (2013). Creative-commons attribution-share alike 3.0 license.

Harmel (2009, p. 29) reports on the case of France in the period of the single area payment. Here, the subsidies were based on area measurements calculated on the national coordinate system, a Lambert conformal conic. In the conversion to the revised national coordinate system (RGF 93), subsidies to the agriculture sector were reduced by 17 million euros. This reduction in payments was due to reduced scale error in the projection, particularly in the south of France and Corsica. Presumably, many other countries would have similar savings to make, but the magnitude is unreported and unexamined.

Scale factors

The largest impact of map projection comes from scale factor issues (for more treatment, see Chrisman and Girres 2016). Projections distort by manipulating the scale of the Earth in various ways.

The official coordinate systems adopted around the world favor angles and shapes over distance and area (Chrisman and Girres 2016). The historical record links the adoption of these conformal projections to gunnery calculations during the Great War of 1914–18 (O'Keefe 1952). Both the old and new projections for France (cited by Harmel 2009) are Lambert conformal conics. When used as the base for GIS, the properties of perimeter and area are distorted with a known and avoidable error. In addition, tolerances and thresholds (as in buffers) are different from their intended ground distances (as in the North Korean missile cases). All these errors can be modeled by scale factors with a known and computable spatial distribution. These effects may not be very large when projection guidelines are carefully observed. However, they remain. The scale error on the standard projections of France was considered reasonable by the professionals who recommended adoption. Perhaps they did not consider how many public decisions might be based on that extended projection zone. The magnitude of that decision is around 17 million euros for just one subsidy program.

Berk and Ferlan (2016) confirm the magnitude of area distortions in their exhaustive study of all the parcels of Slovenia. The Transverse Mercator projection official coordinate system overestimates the western side of the country by 1.85 m^2/ha and underestimates along the central meridian by 1.99 m^2/ha.

In a similar manner to the more common conformal case, an equivalent projection removes area distortion by manipulating the distortion of other properties. Unlike conformal projections, equal-area projections have different scales in different directions. Hence, it is much more complicated to correct distance and azimuth calculations on an equal-area projection.

To my knowledge, no current GIS software provides a tool to generate a surface of distance scale errors or area scale errors based on projection parameters. Instead, properties calculated on the projection plane are presented as true measures despite solid evidence to the contrary.

Straight lines

Another issue concerns the concept of straight lines, a concept fundamental to geometry for thousands of years. The definition of a straight line is unambiguous in a planar context. In fact, the whole conceptual concentration on straight lines may come from the use of tabletop models of geometry (originally implemented in the sand, according to ancient practice). By unwritten computer-graphics convention, unless otherwise specified, the gap between two points is assumed to be 'straight'. Cartographers have accepted the null hypothesis of straight lines without much debate (notably, Peucker 1976). On the ellipsoid, the simplicity of 'straight' disappears (Deakin *et al*. 2011).

To remind the reader, and to move in stages, consider spherical geometry as an intermediary. On the sphere, the shortest path between two points (P_1, P_2) is along a 'great circle' – the intersection of the sphere with a plane P defined by the two points plus the center of the Earth. All great circles have the same circumference and curvature (it is a sphere). Another neat property is that the perpendiculars at the two end points (a line toward the center of the Earth) are contained on the plane Q, and thus determines the shortest path in the field. Basic surveying uses a plumb bob to 'level' the instrument with respect to this local perpendicular.

The ellipsoid brings some additional complexities. Sections through the ellipsoid are ellipses, though the equator and other parallels of latitude remain circular (Deakin 2009). By definition, the distance to the center of the Earth varies from a maximum on the equator to a minimum at the poles. Polar flattening is around 23 km, big enough to make a difference in surface distances and areas. Meridians remain easy to model, but

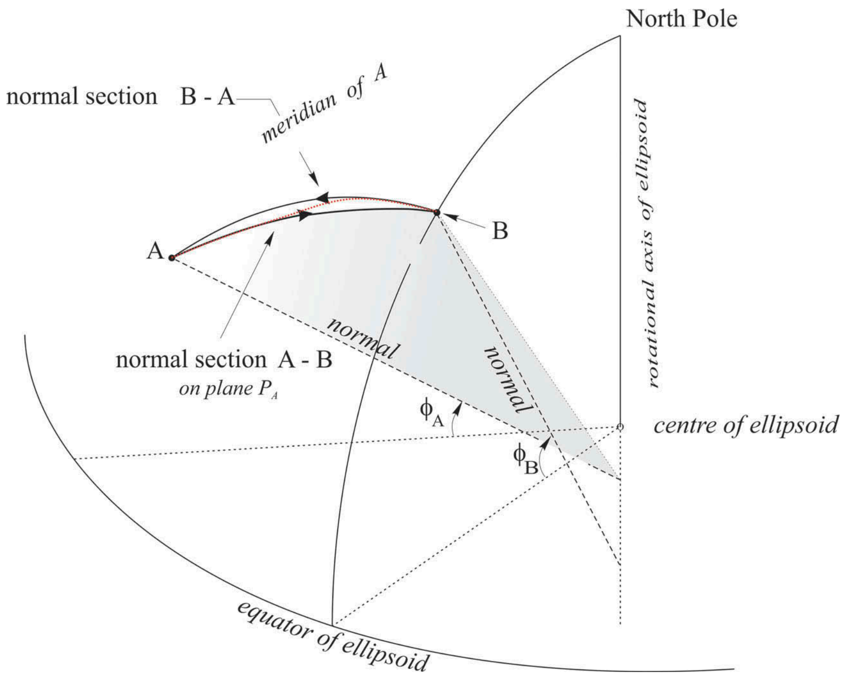

Figure 2. Geodesic from A to B (in red dashed line) between two normal sections. The plane P_A including the normal at A is shown in grey.To render the lines distinguishable, the ellipsoid has been dramatically flattened. Source: Modified from (Deakin 2009); with kind permission of the author.

other orientations of lines get more complex. At the points A and B (not at the same latitude), the two local perpendicular lines will not meet at the center of the Earth or at the same position on the polar axis (Deakin 2009). Consequently, there are two 'normal sections' based on the two local perpendiculars; each is an ellipse on its distinct plane (P_A, P_B) (Figure 2). These two lines do not coincide, except in special cases like the equator or meridians. These two normal sections are, neither one, the shortest path. The distance along the normal section ellipse can be estimated by integration, so it is not a shortcut to distance calculations. A third elliptical section – the 'Great Elliptical Arc' – can be obtained by including the center of the Earth and the two endpoints, in keeping with the procedure for great circles. It is not the shortest path, but it does trace the ground footprint of satellites (constrained to orbit the center of mass of the planet).

The shortest path is termed the 'geodesic'. It traces the locus of normal to the surface, so it coincides with each normal section near their respective endpoints. In between the endpoints, the geodesic on the ellipsoid follows an S-shaped curve often contained between the two normal sections, but not necessarily (see Figure 2 for an exaggerated illustration). The length of the geodesic can be calculated, but it is certainly not 'straight' in the classical meaning of that term. It does not have a neat analytical form to aid in calculating intersections and other properties.

The distance between the two normal sections is not large, though it depends on latitude difference and location. In the case of the Black-Allan Line (on the border between Victoria and New South Wales, some 176 km long) the two normal sections are 0.03 m apart (Deakin *et al*. 2011, p. 13). Following either normal sections or the geodesic, ellipsoidal calculations are required to obtain the distance between two points. As distances increase, geodesics diverge up to a maximum (equatorial dimension) then converge on the antipodal point. In the spherical case, the point on the opposite side of the sphere can be reached on any plane that passes through the axis between the two points. Thus, the whole sphere is on the 'shortest' path (the longest possible distance on the sphere). The ellipsoid is different, due to the flattening. In the equatorial case it is clear to see that an azimuth to either pole will be shorter than around the wider equator (Rapp 1993, p. 41).

Geodesics on the ellipsoid also have a curious property as they continue on past their end point. While great circles close neatly back on the starting point, non-meridional geodesics on the ellipsoid wander off in a spiral (Rapp 1993, p. 46). This infinite behavior is more similar to the infinite path of straight lines on the plane, though ellipsoidal geodesics intersect on each rotation.

In the recent geodetic literature, Karney (2013, p. 44) defines straightness as a property on a projection (local azimuthal at every point). This brings us right back to the plane, even for the most basic definition of geodesics. To my knowledge, no GIS software bothers to consider the path of ellipsoidal geodesics. Even Hipparchus (Lukatela, 1987) uses local spherical models with great circles. By some reports, certain database implementations use spherical great circles as well when data are referenced in angular coordinates. Straight lines on projections are applied in most cases despite clear evidence of distortions. The mistake of the North Korean missile map is not so unusual.

Summary of problems

- Geometric properties calculated by a GIS on planar projections will have errors that are predictable and preventable.
- The current affordances of GIS software packages encourage laziness with respect to projection error.
- Insufficient tools are provided to correct known errors.

Simple remedies

This section will set out some simple fixes that stay within the current structure of systems and current standards. A following section will consider more radical solutions. This treatment will review a variety of existing solutions, and provide citations to other sources for full mathematical treatment. Otherwise the mass of formulae would burden the text and diminish the breadth of the argument.

The first and simplest solution is simply to warn users. Disclaimers and dense 'Terms of Use' contracts pop up in our daily routine, enjoined by lawyers everywhere (Hunter and Goodchild 1993, p. 59). At a minimum, GIS software should say something like 'Estimates of distance and area are approximations calculated on a projection plane, and some adjustment may be required to derive more accurate values.' But such a statement begs the question. It serves lawyers more than readers. If there is a known effect, why is it not treated? If you went to all the effort, don't you want the real answer? These errors are not random or unpredictable; the effect is totally understood and programmable.

A more aggressive stance could be copied from the data quality setting. Hunter and Reinke (2000) suggested a scheme to prohibit certain operations if the data quality is insufficient. It would be simple enough to prohibit the calculation of areas, for example, unless the projection was equivalent. Few software systems have implemented anything quite so complex. The current permissive approach requires other solutions.

Correcting for known error

The errors introduced by projections can be corrected. A local scale factor can be calculated point by point, and therefore corrected estimates can be produced. For distance measures, the length calculated on the projection plane can be adjusted, larger or smaller, depending on distance from the standard parallels or central meridian – depending on the specific projection used (Vincenty 1985, Girres 2011, 2012). Sampled at the points along a line, the variation in scale factor can be approximated if the sampling is relatively dense and uniform. Long segments aligned with a meridian or parallel might require a bit more treatment, but generally the result should be adequate with an estimate at either end.

Adjustment of measurements according to scale error is commonplace in surveying (not just for the local error on the ellipsoid, but also raised up to ground level), yet somehow the practice (grid to ground) never penetrated GIS software. This deficiency may contribute to the distrust of GIS calculations in the surveying community. In the early period, GIS results could be defended as approximate, and fit for a different set of users. There is no reason now to continue this logic. Computing models should return the actual answer, not just what is easy to program. Using an inadequate model limits the tools to approximate data sources. Sadly, the knowledge of each discipline in the geospatial sector is not fully shared with the other participants. This lack may be due to different training pathways, but after 40 years, ignorance is no excuse.

For some operations dependent on distance, the parameters used in the calculation should be adjusted. For example, the width of a buffer should be modified to adjust for the local distortion of distances. This value would change across the data coverage, making some complexity, but it would resolve the projection scale issue. These adjustments will be small, and some will argue that the intent of a buffer is not so precise. However, it is not the programmer's role to make that decision. GIS software should deliver what it advertises. Any treatment of uncertainty should not be compounded by adding uncertainties due to flaws in the numerical model.

Adjusting area estimates

Adjustment for errors in area are also possible, though more complicated. Taken at a point, like the distance factor, the error in area (for conformal projections) is approximately twice the distance error (CERTU 2010, p. 4, Girres 2012, p. 112). The squared term tends to be much smaller in typical applications.

$$e_A = 2 * e_L + {e_L}^2, \tag{1}$$

where e_L = linear error; e_A = area error factor.

Unlike a distance calculation, area extends far away from the boundary points. Accumulating the average projection error around the perimeter leaves the vast inside of some polygons unsampled. Girres (2012) took the conservative approach of estimating the true area of a polygon by tessellating into a series of triangles on the ellipsoid and calculating the area using geodetic trigonometry. This approach moves past the adjustment of area to the topic of a subsequent section (abandoning projections altogether).

Using projections intelligently

Some of the difficulties cited above could be resolved by adopting different projections. The conformal properties of current official coordinate systems permitted simple hand calculations for surveying or artillery, but do not serve all the functions of a GIS. If the principal geometric concern is area (as in a crop estimation service or a forest inventory), then an equivalent (equal-area) projection may be the simplest solution (for example, Balcerzak and Pedzich 2007). For many statistical purposes covering the conterminous United States, the most common projection is Alber's Equal Area imported from nineteenth century work in Germany (Snyder 1987, p. 98). This projection (in its typical configuration for the 48 states) has differential linear scale error up to about 1% between the standard parallels, and a bit more at the northern and southern extremities. To make the area work out correctly, the scale along parallels compensates for the scale on meridians. Therefore, correcting for distance calculations on this projection is complicated. But, if the goal is areas, there is no need for correction. The main requirement is to insert sufficient points along 'straight' lines to ensure that they approximate the right geometry (Gillissen 1993). Most GIS avoid this step.

The European Union has adopted a Lambert Azimuthal Equal Area (LAEA) projection to serve a similar role for coverage of the European member states (though not extending to overseas territories) (INSPIRE 2010, p. 7). Sadly, the INSPIRE working group adopted too many incompatible systems. They also allow the major two conformal mapping systems (Lambert conformal conic and Transverse Mercator), not surprising since these are the common choices for national official mapping. These conformal mapping projections require adjustments such as those described above.

The INSPIRE working group also condone (in some manner) the use of the non-projection called 'Plate-Carée' since it has come into use on (unnamed) internet viewing services. Planar calculations on this surface are wrong in every way imaginable, and bear no useful relationship to the Earth. In other words, even the power of 26 nation states cannot set a standard which opposes the economic power of an international corporation of global reach. This is particularly distressing when said corporation makes a selection without scientific merit. Yet, these sub-optimal choices become enshrined as standards, and we must work around them for decades to come. The First Mover effect is frighteningly powerful.

For other procedures in a GIS beyond area calculation, a change in projection is not a full solution. For buffer calculations, the appropriate projection would be equidistant. The problem with this class of projections is that the equidistance property is only valid from one point. Thus, there is no single projection that will permit correct buffer calculations without considerable re-scaling of distances. As the North Korean missile case shows, buffers of geodetic proportions require careful treatment.

Abandoning projections altogether

While it is possible to counter each deficiency of map projections, the result becomes a patchwork. It is time to consider the more radical solution of abandoning projections for all calculations. I hasten to clarify that projections would remain important for display purposes. But if the need for specific kinds of calculations is removed, the choice of projection can be tailored to the requirements of visualization instead of calculation.

This paper will consider four calculations that together provide adequate primitives for the computational geometry of GIS. In most of these cases, there will be variants to deal with the distinct kind of straight lines on an ellipsoid. The first two are simple distance and area, the classic measures considered in projection design. These two are not enough; fundamental GIS is also built on set theory queries implemented by map overlay procedures. These involve line intersection as the basic geometric step. Other operations calculate the distance from a line to construct buffers and similar objects. As announced previously, in an aim for brevity this article will not include the detailed formulae. They can be found in the sources cited.

Distance on the ellipsoid

Calculating the distance between two points on the ellipsoid is termed the 'inverse' problem (or the 'Gauss inverse'). Perhaps due to the navigational connections or the terrestrial surveying background, the 'direct' problem determines the end-point of a geodesic given an initial azimuth and a distance from a starting point. Both problems have been treated by geodesists for over 200 years (for example, Clairaut 1743, Bessel 1825, Gauss 1828, Helmert 1884, Rapp 1993). This distance typically follows the shortest path, the geodesic. Distances along other lines, such as normal sections, can also be computed.

In recent years, many programmers have used Vincenty's (1975a) iterative solution, derived from Rainsford (1955). This requires at least six trigonometric function calls per iteration; convergence takes two to four iterations (Vincenty 1975a, p. 93) except in antipodal cases. As the author admitted (Vincenty 1975a, p. 92), the solution is not accurate for nearly antipodal points, but it provides a decent solution (less than 1 mm) in more common cases (Vincenty 1975a, p. 90). Vincenty (1975b) worked out some solutions for antipodal points, but did not publish the paper. Rapp (1993) and others have resolved most of the numerical issues to deliver stable solutions.

Karney (2011) has developed numerical procedures that expand the series a few more terms, and reduce the calculations by careful grouping of terms that can be precalculated. Karney increased the accuracy to 15 nm, and eliminated the convergence problems (Karney 2011, p. 52). On a computer of the current era, Karney (2013, p. 53) reports 2.3 μs on average for the inverse problem versus 1.3 μs for Vincenty's algorithm on the same configuration. Just for reference, on the computer used by Vincenty in his tests (an IBM 7094 – already obsolete in 1975), the memory cycle (time to retrieve one element) was 2.2 μs. In other words, one simple memory reference took as long as it now takes to estimate distance to the nearest micrometer. For the purposes of a GIS operating on ellipsoidal geometry, the calculation of distances is no longer a barrier.

Area on the ellipsoid

Area calculation on the ellipsoid is more complicated, as Lukatela (2000) has explained. The well-established planar approach slices a polygon into trapezoids (with parallel rays heading off to one axis or another). On the ellipsoid, latitude and longitude can be treated as pseudo-orthogonal axes creating pseudo-trapezoids to the Equator. However, the geodesic (diagonal) boundary of this trapezoid is not a straight line, so the area between the geodesic and the Equator has to be approximated by integration (Danielsen 1989, Sjoberg, 2006). Kimerling (1984) derived the areas of pseudo-trapezoids through approximations calculated on the UTM projection. This strategy defeats the purpose to calculate directly on the ellipsoid.

In the geodesy literature, a few publications provide a direct solution following two distinct strategies. The first, more traditional, approach works on an 'authalic sphere' with corrections for the ellipsoid. Danielsen (1989) draws an ellipsoidal pseudo-trapezoid formed by the boundary line in question, the equator, and the two meridians connecting to the end points of the line. All four lines are geodesics on the ellipsoid. Danielsen maps them onto a sphere (the reduced latitude) iteratively to preserve the azimuth. The area is accumulated on the appropriate radius. He provides the basic formulae to calculate to 1 m^2 (third order terms in the iterative procedure). Working directly on the ellipsoid, Sjoberg (2006) improves on these results. Karney (2013, p. 50) worked out the terms to higher order and recast them as trigonometric sums. As is often the case, the estimation of these terms can be unstable when distances are short- very much the case of most GIS applications. Karney mobilizes the work of Bessel (1825) to limit this issue. This technique involves integration where each term invokes four trigonometric functions and some multiplications and divisions in a traditional series expansion. The computation will cost more than the planar case, for certain.

Lukatela (2000) follows another strategy, based on a tessellation of the ellipsoid into Voronoi-like cells, each of which is a spherical facet with a locally adjusted radius. This is the Hipparchus model (Lukatela 1987). In this model, each cell is a composite of spherical triangles – bounded by great circles, whose area is easily derived. Unlike the other more mathematically oriented publications, Lukatela (2000, p. 125) provides a full-scale test of his calculation algorithm. The area of his 2432 Voronoi cells falls short of the expected total for the Earth by 45 km^2 or one part in 1.1×10^7. When applied to the Digital Chart of the world data for land and ocean, the algorithm evaluated 1.3×10^6 triangles. The sum of land + ocean fell short by 4160 km^2, about 100 times larger than the discrepancy with the smaller set of objects. Inevitably, there is some loss to roundoff, but the same applies to area estimates performed on the plane. There are few studies with equal rigor published for other techniques. At one part in 1.2×10^5, Hipparchus is clearly superior to the accuracy in area for any areas calculated on conformal projections. In normal zones, linear scale factors range up to one part in 4×10^3, so area scale factors can be of the order of one in 8×10^3.

In the Hipparchus model, the spherical calculations for area used the local radius of the ellipsoid (following the algorithm of Legendre 1806). The Hipparchus system promoted by Lukatela (1987) tessellates the ellipsoid into spherical Voronoi-like cells.

Using either the method of Karney or Lukatela, areas can be estimated on the ellipsoid. All area measurements, either performed on an equal-area projection or on the ellipsoid, can suffer from numerical issues dealing with thin objects. There may be rather limited attention to this kind of roundoff accumulation in GIS. Adding many small items to a big accumulation loses digits (Kahan 1965, Higham 1993). The well-established techniques to counter round-off losses are rarely deployed in practical software. Still, this kind of issue is of much lower concern than the scale factor issues on conformal projections (by a factor of 100, perhaps).

Line intersection and distance to a line on the ellipsoid

Beyond distance and area, GIS depends on additional geometric procedures. One of the most notable in the history of GIS was polygon overlay. This capability provided the basis to merge attributes attached to distinct geometric figures, and to implement topological queries. At the base, this depends on the ability to calculate the intersection of linework. Once established, an overlay engine does a number of additional procedures, from point in polygon to topological cleaning. It is thus fundamental to operational GIS.

The basic geometric operation to calculate intersection on the plane begins with calculating the distance from a point to a straight line (Little and Peucker 1979). Two opposing signs on the oriented distance provide a better test for intersections than many of the other solutions (Dougenik, personal communication, 1978). Even on the plane, the intersection problem is non-trivial. Douglas (1974) had a famous rant titled 'It makes me so CROSS', citing all the special checks required for coincident points, near parallel cases, and numerical effects of roundoff. Saalfeld (1987) countered with 'It doesn't make me nearly as CROSS', citing better choices for representing lines. This background prepares for further complication on the ellipsoid where, as discussed previously, lines are not as well-behaved.

Karney (2013, p. 54) includes the intersection problem in his set of geodesic calculations, linked to what he calls the interception problem – the calculation of the point on a line closest to a point off the line. This is an alternate way of formulating the distance from a point to a line. For the ellipsoid, Karney suggests an iterative procedure using an ellipsoidal gnomonic projection. The spherical gnomonic projection preserves great circles as straight lines. On the ellipsoid, it is not as exact. Karney's projection provides a solution where intersections within 1000 km of the center are within 1.7 m of true position anywhere along a geodesic (Karney 2013, p. 54). Karney is not satisfied with that level of error, so he iterates the intersection procedure with the calculated intersection as the new center of the projection, until convergence is reached. Since the epsilon tolerance associated with many polygon overlay procedures is on the order of a meter, the use of a single gnomonic for a region might be enough. The main point is that Karney does not provide a procedure to calculate the intersection of two geodesics directly on the ellipsoid. The reason is that points along a general geodesic do not hold to a specific equation, as do points on a planar line. It makes the analytical geometry quite difficult.

Once the interception problem is solved on the gnomonic projection, the best practice would be to calculate the distance using the ellipsoidal procedure provided by Karney. Distances on the gnomonic projection are notoriously distorted, even quite close to the origin. Alternatively, distance on the gnomonic can be adjusted by these known scale factors.

Discussion

Initially I expected to reject planar geometry for GIS, and to reconstruct computational geometry on the ellipsoid. The errors of projection use seemed too pervasive and complex.

There are good solutions for precise calculation of distance on the ellipsoid. In some user interfaces, the distance calculation allows *ad hoc* query of ellipsoidal distances, although the database still contains perimeter and distance measures calculated in the projection in which the coordinates are stored. The alternative requires some estimation of scale factors on each projection surface which are a bit messy but not out of the question. For area, the reverse seems to be the case. The calculation on the ellipsoid is possible, but tricky. Calculating area on an equivalent projection gives a precise answer with no need for re-engineering. Berk and Ferlan (2016) confirm this result in an exhaustive study of all parcels in Slovenia. They also demonstrate that the choice of projection parameters does matter.

The final case of line intersection does not seem feasible on the ellipsoid. Even the most die-hard geodesist suggests that we use a particular projection (though not a typical projection).

On balance, the case for direct calculation on the ellipsoid is not as strong as I expected. Rather than rejecting projections, I conclude by recommending the use of different projections tailored for each specific problem.

Technical conclusion

GIS professionals must learn to apply the appropriate projection for each kind of calculation, following the advice of Tobler (1993). Blanket adoption of official (conformal) coordinate systems leads to incorrect results, as documented above. The following guidelines move from the easier cases to the harder ones. Regular users can immediately adopt the easier ones, while Recommendation 3 requires change in the software sector.

Recommendation 1: area

The area column in spatial databases should only be calculated on an equivalent projection. There are many to choose from, if appropriately designed for the study area. As Usery and others (2003) document, choice of projection parameters do matter. Berk and Ferlan (2016) confirm this result by obtaining much better performance from an Alber's Equal Area Conic than the Lambert (LAEA) often suggested by European agencies. A user can simply specify a current display projection, then calculate the areas into a new column.

Recommendation 2: distance

For measures of length (perimeter), the correct result can be obtained rather easily from a proper ellipsoidal calculation (for example, Vincenty 1975a; Karney 2013). This provides a distance on the ellipsoid undistorted by the projection. Alternatively, a projection scale factor can be estimated based on the end points. For a conformal projection, these correction factors are a simple function of distance from the central meridian of a cylindric projection such as UTM, or the standard parallels of a conic. A surface of these values could be constructed using a raster map calculator. Producing this surface should be an automatic function of any advanced software package (though currently unavailable to my knowledge). As Girres (2012) points out, using a grid of correction values would also permit adjustment of the distances to actual ground values, taking elevation and slope distance into proper account with a digital elevation model. The ground distance model may deliver a more useful answer than a distance derived on the projection plane or the ellipsoid.

Recommendation 3: intersections and buffer distance

Given that most polygon overlay is calculated with a fuzzy tolerance, it may seem to be too much effort to use the iterative procedure of Karney. Perhaps just using an ellipsoidal gnomonic projection centered on the study area will be adequate for regions up to 1000 km. This projection will have the side effect of grossly distorting the area measurements often included in the process. Until software developers recognize the issue, it may be necessary to continue with the current solutions based on planar calculations on conformal projections.

Ethical conclusion

Ultimately, the argument is not about the nuance of high order terms in some geodetic equation or about the correct estimation procedure. Professionals practicing in the realm of geographic representation must deal with the Earth as it is – round. Ellipsoidal models provide the correct answer to many spatial queries, and the surrogate answers calculated on some other object must be properly adjusted.

The current solutions produced by the default planar algorithms produce systematically biased results. Since these factors are totally predictable, a professional must adjust for them.

Ignorance is no excuse.

Acknowledgments

While this essay benefits from the experience gained in many research projects funded by agencies in multiple countries, there is no funding agency responsible for this work. This topic has been discussed with many colleagues over decades, and I recognize the discussions and arguments, particularly with Rodney Deakin. Taking over the Map Projections course (GEOM 2117) from Rod at RMIT forced me to recognize how much more I had to learn about calculations on the ellipsoid. The original spark came, as usual, from a PhD student, Jean-François Girres, whose research included a treatment of projection error. This article develops material past the prior collaborations, with the assistance from anonymous reviewers; all conclusions derive solely from the author.

Disclosure statement

No potential conflict of interest was reported by the author.

ORCID

Nicholas R. Chrisman http://orcid.org/0000-0002-6542-7561

References

Anderson, D.E., Angel, J.L., and Gorny, A.J., 1977. World data bank II. Content, structure and application. *In*: G. Dutton, ed. *An advanced study symposium on topological data structures for geographic information systems*. Cambridge, MA: Laboratory for Computer Graphics and Spatial Analysis, Harvard University, 16–21.

Balcerzak, J. and Pedzich, P., 2007. The methods of calculation of ellipsoidal polygon areas based on some map projection properties. *In*: Proceedings ICC 2007. Paris, paper 2.2.3. Available from: http://icaci.org/files/documents/FICC_proceedings/FICC2007/documents/doc/THEME 2/oral /2.2.3 THE METHODS OF CALCULATION OF ELLIPSOIDAL P.doc

Berk, S. and Ferlan, M., 2016. Accurate area determination in the Cadaster: case study of Slovenia. *Cartography and Geographic Information Science*, in press. doi:10.1080/15230406.2016.1217789.

Bessel, F.W., 1825. Uber die Berechnung der geographischen Längen und Breiten aus geodätischen Vermessungen. *Astronomy Nachrichten*, 4 (86), 241–254. translated into English by C. F. F. Karney and R. E. Deakin. 2010. as The calculation of longitude and latitude from geodesic measurements, Astron. Nachr. 331 (8),852–861,E-print arXiv:0908.1824 doi:10.1002/asna.201011352

Blaschke, W., 1949. *Kreis und Kugel*. New York: Chelsea Publishing.

Boyle, R., 1964. *Graph plotting mechanism*. UK Patent GB1030331/US Patent 3258776A granted 28 June 1966. Available from: https://www.google.com/patents/US3258776

CERTU. 2010. Distances et surfaces, comment prendre en compte l'altération linéaire? Géoréférencement et RGF93, Théorie et concepts - Fiche T9. Lyon, 1–6. Available from: http://www.certu.fr/fr/_Information_géographique-n32/Géoréférencement_et_RGF93-n795/IMG/pdf/RGF93_theorie_et_concept_T9.pdf

Chrisman, N., 1974a. The impact of data structure on geographic information processing. *In*: Proceedings AUTO-CARTO I, 165-177. Available from: http://mapcontext.com/autocarto/proceedings/auto-carto-1/pdf/cartographic-data-structures-panel.pdf

Chrisman, N., 1974b. *POLYVRT: a program to convert geographic base files*. Cambridge, MA: Laboratory for Computer Graphics and Spatial Analysis, Harvard University.

Chrisman, N., 1984. On storage of coordinates in geographic information systems. *Geoprocessing*, 2, 259–270.

Chrisman, N., 2006. *Charting the unknown: how computer mapping at Harvard became GIS*. Redlands, CA: ESRI Press.

Chrisman, N., Dougenik, J.A., and White, D., 1992. Lessons for the design of polygon overlay processing from the ODYSSEY WHIRLPOOL algorithm. *In*: P.J. Bresnahan, E. Corwin, and D. Cowen, eds. *Proceedings 5th international symposium on spatial data handling*, Vol. 2, 3–7 August, Charleston, SC. Columbia, SC: University of South Carolina, 401–410.

Chrisman, N. and Girres, J.F. 2016. First do no harm: eliminating systematic error in analytical results of GIS applications. *In*: W. Shi, B. Wu, and A. Stein, eds. *Uncertainty modelling and quality control for spatial data*. Boca Raton, FL: CRC Press, Chapter 3, 27–44.

Clairaut, A.C., 1743. *Théorie de la figure de la terre, tirée des principes de l'hydrostatique*. Paris: David fils, à la Plume d'or.

Cmglee. 2013. *North Korean missile range.svg. Wikimedia Commons*. Available from: https://commons.wikimedia.org/wiki/File:North_Korean_missile_range.svg

Common Core. 2015. Grade 5 » Geometry » Graph points on the coordinate plane to solve real-world and mathematical problems. *Common Core Standard CCSS.MATH.CONTENT.5.G.A.1*. Washington, DC: Council of Chief State School Officers. Available from: http://www.corestandards.org/Math/Content/5/G/A/1/

Danielsen, J.S., 1989. The area under the geodesic. *Survey Review*, 30 (232), 61–66. doi:10.1179/sre.1989.30.232.61

Deakin, R.E., 2009. *The normal section on an ellipsoid*. unpublished lecture notes. RMIT, Melbourne.

Deakin, R.E., Sheppard, S.W., and Ross, R., 2011. The Black-Allen line revisited. Presented at 24th Victorian Regional Survey Conference, 1–3 April, Shepparton VIC.

Department of Education. 2013. Mathematics programmes of study: key stages 1 and 2 national curriculum in England. *DFE-00180-2013*. London: Department of Education. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335158/PRIMARY_national_curriculum_-_Mathematics_220714.pdf

Descartes, R., 1637. Géometrie. *In*: *Discours de la méthode*. Leyden: Jan Maire, Appendix 3, 297–314. Page citations to the 1887 edition. Paris: A. Hermann. Available from: http://www.gutenberg.org/files/26400/26400-pdf.pdf

Douglas, D.H., 1974. It makes me so CROSS. *In*: D. Peuquet and D. Marble, eds. *Introductory readings in GIS*. London: Taylor & Francis, Vol. 1990, 303–307.

Elassal, A.A., 1987. *General cartographic transformation package (GCTP), Version II*. Rockville, MD: US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Charting and Geodetic Services, NOAA Technical Report NOS 124 CGS 9. Available from: http://www.ngs.noaa.gov/PUBS_LIB/Brunswick/NOAATRNOS124CGS9.pdf

EU (European Union). 2013. Regulation EU no 1307/2013. Official Journal of the European Union. Available from: http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:32013R1307

European Commission. 2013. *CAP reform – an explanation of the main elements*. Available from: http://europa.eu/rapid/press-release_MEMO-13-937_en.htm

Fisher, M., 2013. Map: this is how far those North Korean missiles can actually reach. *Washington Post*, 4 April. Available from: https://www.washingtonpost.com/news/worldviews/wp/2013/04/04/map-this-is-how-far-those-north-korean-missiles-can-actually-reach/

Freeman, H., 1961. On the encoding of arbitrary geometric configurations. *IRE Transactions on Electronic Computers*, EC-10 (2), 260–268. doi:10.1109/TEC.1961.5219197

Gauss, C.F., 1828. *Disquisitiones Generales circa Superficies Curvas*. Göttingen: Dieterich. Available from: http://books.google.com/books? id=bX0AAAAAMAAJ. translated into English by Morehead JC and Hiltebeitel AM, 1902. as General investigations of curved surfaces of 1827 and 1825. Princeton NJ: Princeton Univ Lib.

Gillissen, I., 1993. Area computation of a polygon on an ellipsoid. *Survey Review*, 32 (248), 92–98. doi:10.1179/sre.1993.32.248.92

Girres, J.-F., 2011. A model to estimate length measurements uncertainty in vector databases. *In*: *Proceedings of the 7th international symposium on spatial data quality*, 12–14 October, Coimbra, 83–88.

Girres, J.-F., 2012. *Modèle d'estimation de l'imprécision des mesures géométriques de données géographiques: Application aux mesures de longueur et de surface*. PhD dissertation. Université Paris-Est, France.

Grafarend, E., You, R.-J., and Syffus, R., 2014. *Map projections*. Berlin: Springer.

Harmel, A., 2009. Le nouveau système réglementaire Lambert 93. *Géomatique Expert*, 68, 26–30. Available from: http://www.geomag.fr/sites/default/files/68_91.pdf [Accessed 11 December 2015].

Helmert, F.R., 1884. *Die mathematischen und physikalischen theorieen der höheren geodäsie*. Vol. 2. Leipzig: BG Teubner. Available from: https://books.google.com/books?hl=en&id=p9-CAAAAIAAJ&oi=fnd&pg=PA1&dq=Helmert+Geodasie&ots=WkeiL7bUwB&sig=jUI24nO340F1TAdshHxO3ke0iwU

Higham, N., 1993. The accuracy of floating point summation. *SIAM Journal on Scientific Computing*, 14 (4), 783–799. doi:10.1137/0914050

Horwood, E., *et al*., 1963. *Computer methods of graphing, data positioning and symbolic mapping*. Seattle: University of Washington Civil Engineering Department

Howse, D., 1985. 1884 and longitude zero. *Vistas in Astronomy*, 28 (1), 11–19. doi:10.1016/0083-6656(85)90003-0

Hunter, G.J. and Goodchild, M.F., 1993. Managing uncertainty in spatial databases: putting theory into practice. *URISA Journal*, 5 (2), 55–63. Available from: http://www.urisa.org/clientuploads/directory/Documents/Journal/vol5no2.pdf

Hunter, G.J. and Reinke, K.J., 2000. Adapting spatial databases to reduce information misuse through illogical operations. *In* G.P.M. Heuvelink and M.J.P.M. Lemmens, eds. *Proceedings of 4th international symposium on spatial accuracy assessment in natural resources and environmental sciences (accuracy 2000)*, 12–14 July, Amsterdam. Delft: Delft University Press, 313–319.

IERS. 2013. *The international terrestrial reference system*. Paris: International Earth Rotation and Reference Systems Service. Available from: http://www.iers.org/IERS/EN/DataProducts/ITRS/itrs.html

INSPIRE, Infrastructure for Spatial Information in the European Community. 2010. *INSPIRE Specification on coordinate reference systems - Guidelines*. Ispra, Italy: INSPIRE Thematic Working Group on Coordinate Reference Systems and Geographical Grid Systems. Available from: http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_Specification_CRS_v3.1.pdf

ISO. 2007. *Geographic information - Spatial referencing by coordinates (ISO 19111:2007)*. Geneva: International Standards Organization.

Kahan, W., 1965. Pracniques: further remarks on reducing truncation errors. *Communications of the ACM*, 8 (1), 40. doi:10.1145/363707.363723

Karney, C.F.F., 2011. Geodesics on an ellipsoid of revolution. *arXiv: 1102. 1215*. Available from: http://arxiv.org/pdf/1102.1215.pdf

Karney, C.F.F., 2013. Algorithms for geodesics. *Journal of Geodesy*, 87 (1), 43–55. Available from: http://link.springer.com/article/10.1007%2Fs00190-012-0578-z

Kimerling, J.A., 1984. Area computation from geodetic coordinates on the spheroid. *Surveying and Mapping*, 44 (4), 343–351.

Legendre, A.M., 1806. Analyse des triangles tracés sur la surface d'un sphéroide. *In*: *Mémoires de la classe des sciences mathématiques et physiques de l'Institut Nationale de France*, Premier semestre. Paris: Baudouin, 130–161. Available from: http://books.google.com/books?id=-d0EAAAAQAAJ&pg=PA130-IA4

Little, J.J. and Peucker, T.K., 1979. A recursive procedure for finding the intersection of two digital curves. *Computer Graphics and Image Processing*, 10 (2), 159–171. doi:10.1016/0146-664X(79)90048-0

Lukatela, H., 1987. Hipparchus geopositioning model: an overview. *In*: Proceedings AUTO-CARTO 8, 87–96. Available from: http://mapcontext.com/autocarto/proceedings/auto-carto-8/pdf/hipparchus-geopositioning-model-an-overview.pdf

Lukatela, H., 2000. Ellipsoidal area computations of large terrestrial objects. *In*: M.F. Goodchild and J.A. Kimerling, eds. *Discrete global grids*. Santa Barbara: NCGIA, Chapter 6, 114–126. Available from: http://escholarship.org/uc/item/9492q6sm

Morton, G.M., 1966. *A computer oriented geodetic database and a new technique for file sequencing*. Ottawa: IBM.

Nievergelt, J., 1997. Introduction to geometric computing: from algorithms to software. *In*: M. van Kreveld, *et al*., eds. *Algorithmic foundations of geographic information systems*. Berlin: Springer, 1–19. doi:10.1007/3-540-63818-0_1

OGC (Open Geospatial Consortium). 2010. The OGC abstract specification, topic 2: spatial referencing by coordinates. OGC08-015r2. Available from: http://portal.opengeospatial.org/files/?artifact_id=39049

O'Keefe, J.A., 1952. The universal transverse mercator grid and projection. *The Professional Geographer*, 4 (5), 19–24. doi:10.1111/j.0033-0124.1952.45_19.x

Peucker, T.K., 1976. A theory of the cartographic line. *In*: G.M. Kirschbaum and K.H. Meine, eds. *International yearbook of cartography, 16*. Bonn: Kirschbaum Verlag, 134–143. Available from: http://mapcontext.com/autocarto/proceedings/auto-carto-2/pdf/a-theory-of-the-cartographic-line.pdf

Rainsford, H.F., 1955. Long geodesics on the ellipsoid. *Bulletin Geodesique*, 37 (1), 12–22. doi:10.1007/BF02527187

Rapp, R.H., 1993. *Geometric geodesy part II*. Columbus, OH: Department of Geodetic Science and Surveying, The Ohio State University. Available from: http://hdl.handle.net/1811/24409

Richardus, P. and Adler, R.K., 1972. *Map projections for geodesists, cartographers and geographers*. Amsterdam: North-Holland.

RPA (Rural Payments Agency). 2015. *Rural Land and Entitlements (RLE1) guidance*. London: HMSO. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/424004/RLE_Guidance_2015_v1.pdf

Saalfeld, A., 1987. It doesn't make me nearly as CROSS some advantages of the point-vector representation of line segments in automated cartography. *International Journal of Geographical Information Systems*, 1 (4), 379–386. doi:10.1080/02693798708927823

Schmidt, W., 1969. The automap system. *Surveying and Mapping*, 29 (1), 101–106.

Sjöberg, L.E., 2006. Determination of areas on the plane, sphere and ellipsoid. *Survey Review*, 38 (301), 583–593. doi:10.1179/sre.2006.38.301.583

Snyder, J., 1987. Map projections: a working manual. USGS Professional Paper 1395. Reston, VA: US Geological Survey.

Taylor, A. 2015. This interactive map shows how 'wrong' other maps are. *Washington Post*, 18 August. Available from: https://www.washingtonpost.com/news/worldviews/wp/2015/08/18/this-interactive-map-shows-how-wrong-other-maps-are/ [Accessed 20 August 2015].

The Economist. 2003a. *North Korea: when bluff turns deadly*. Available from: http://www.economist.com/node/1748566 [3 May].

The Economist. 2003b. *Correction: North Korea's missiles*. Available from: http://www.economist.com/node/1788311 [13 May].

The Economist. 2012. *Daily chart: missile diplomacy*. Available from: http://www.economist.com/blogs/graphicdetail/2012/12/daily-chart-7

Tobler, W., 1993. Global spatial analysis. Santa Barbara: NCGIA, Technical Report 93-1, 19–24. Available from: http://escholarship.org/uc/item/05r820mz

Tomlinson, R.F., 1968. A geographic information system for regional planning. *In*: G.A. Stewart, ed. *Land evaluation*. Melbourne, VIC: Macmillan, 200–210.

Tomlinson, R.F., 1998. The Canada geographic information system. *In*: T.W. Foresman, ed. *The history of geographic information systems: perspectives from the pioneers*. Upper Saddle River, NJ: Prentice-Hall, 21–32.

Usery, E.L., *et al*., 2003. Projecting global datasets to achieve equal areas. *Cartography and Geographic Information Science*, 30 (1), 69–79. doi:10.1559/152304003100010956

Varma, H.P., Boudreau, H., and Prime, W., 1990. A data structure for spatio-temporal databases. *The International Hydrographic Review*, 67 (1), 71–78.

Vincenty, T., 1975a. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Survey Review*, 23 (176), 88–93. addendum: Survey Review 23(180), 294 (1976). Available from: http://www.ngs.noaa.gov/PUBS LIB/inverse.pdf

Vincenty, T., 1975b. Geodetic inverse solution between antipodal points. *unpublished report dated August 28*. Available from: http://geographiclib.sf.net/geodesic-papers/vincenty75b.pdf. [Accessed 1 December 2015].

Vincenty, T., 1985. Precise determination of the scale factor from Lambert conical projection coordinates. *Surveying and Mapping*, 45 (4), 315–318.