FIRST, DO NO HARM: ELIMINATING SYSTEMATIC ERROR IN ANALYTICAL RESULTS OF GIS APPLICATIONS

N. Chrisman[a, *], J-F Girres[b]

[a] Geospatial Sciences, RMIT University, Melbourne, Australia – nicholas.chrisman@rmit.edu.au
[b] Université Paris-Est, IGN/SR, COGIT, 73 avenue de Paris, 94160 Saint-Mandé, France – jean-francois.girres@ign.fr

KEY WORDS: Systematic error, map projections, scale error, area, perimeter, correction of error

ABSTRACT:

GIS applications compute analytical results comprised of geometric measures such as perimeter, distance between objects, and area. Usual practice operates in Cartesian coordinates on map projections, inducing a systematic variation due to scale error. The magnitude of these errors is easy to foresee, though rarely corrected in analytical reports. Solutions to this error can be implemented though a number of alternative procedures.

1. INTRODUCTION

1.1 Motivation

Research on spatial data quality considers many sources of variation in the data used for spatial analysis, usually with a stochastic approach. This paper returns to a subject often neglected, though totally obvious: map projections. As we describe in all entry courses into the discipline, all map projections distort geometric properties in predictable, systematic ways (Snyder, 1987, p.3). The current set of projections most used for GIS applications were established as official coordinates. The purposes of these official adoptions privileged angles, therefore chose conformal projections. Angles were central to the data collection procedures of an earlier era. However, most of the analytical uses of GIS are more likely to report distances and areas as measures, though these are not preserved by conformal projections. Though the map scale errors are well-known and perfectly predictable, they remain untreated in reporting results. While the focus of DQ research has been on the statistical variability of features, we should consider all sources that make our results deviate from ground truth.

1.2 Aims and Overview

This paper adopts the precautionary principle that any profession should consider: 'First, do no harm'. The current practice of GIS typically moves directly into a map projection with only the most limited consideration of alternatives. This choice of projection may do harm to the application, and we intend to demonstrate the magnitude of this effect in some typical use cases.

This paper will consider the intended variability in common projection systems and official coordinate systems, then continue to the actual practice of these systems that extend frequently into zones beyond the design. The current practice of GIS applications involves study areas of much greater extent than the designers of official coordinate systems envisioned. Hence, study areas frequently extend over zone boundaries, and convenience of treatment leads to bending the rules.

This paper builds on a recent PhD from France (Girres, 2012) on the error consequences for measures of perimeter and area. One section of this work dealt with projections, the subject of this paper. These results will demonstrate that the deviations are potentially significant. In addition, the error is strongly spatial so that certain regions will have predictable under- and over-representation. Certain practices will be advanced to mitigate these effects, but a radical solution must be contemplated.

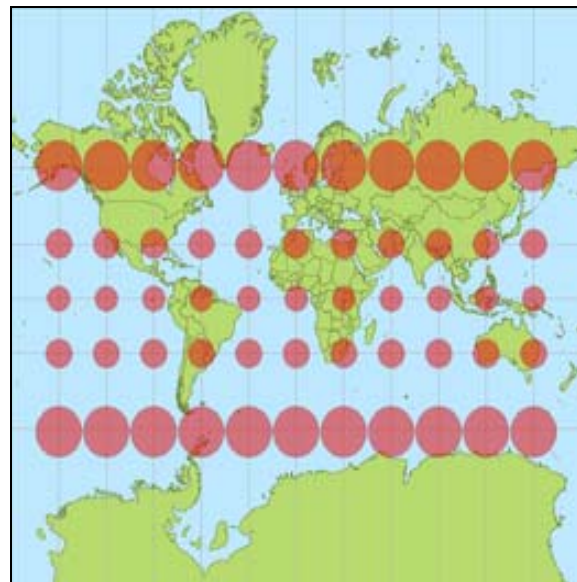## 2. ERROR IN PROJECTIONS

2.1 Overview

Figure 1: A conformal projection (Mercator, equatorial aspect) with exaggerated Tissot circles. (Source: Stefan Kuhn, http://commons.wikimedia.org/wiki/File:Tissot_mercator.png, Creative Common attribution license)

The distortion introduced by a map projection is an analytical result known since Tissot (1881) or before (for full coverage, see Snyder, 1987, p. 20). For the purposes of this paper, we will focus on two functions typically calculated with GIS data:

---

* Corresponding author.

perimeter and area. Both of these are related to the scale factor of the projection, a local parameter that measures the instantaneous variation in distances. It is possible to have no scale distortion at one point, or along one or two lines. These are the locus of tangency, where the projection surface is in contact with the reference ellipsoid. In certain cases, the scale factors in latitude and longitude balance out so that the area result is 'equivalent' (or equal-area). Equivalent projections distort distances, but in a manner that areas are preserved. However, many of the projections used routinely are designed to preserve angles (conformal) by having the same scale factor in all directions (isotropic).
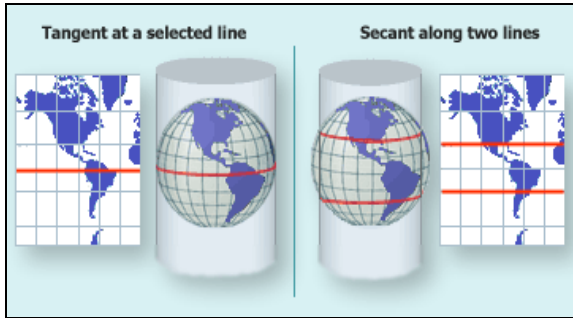

Figure 2: Cylindrical projections with tangent or secant cylinders (Source: US Geological Survey, National Atlas – public domain)
http://www.nationalatlas.gov/articles/mapping/a_projections.html

2.2 Specific Characteristics of Conformal Projections used in Reference Systems

One heritage of topographic mapping practices is the institution of official projections (called National Grids or State Plane Coordinate Systems or some other title). The most usual choice for official coordinates is the transverse Mercator (TM). In the version termed the Universal Transverse Mercator (UTM), developed for military mapping, the coordinate system applies worldwide with strips 6-degrees of longitude wide. Certain jurisdictions have adopted the Lambert Conformal Conic. Together with some oblique versions, they cover the vast majority of official coordinates worldwide.
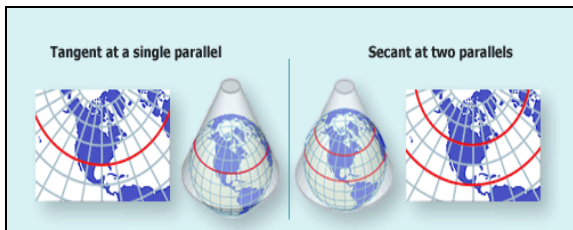

Figure 3: Conic projections with tangent or secant cones. (Source USGS, National Atlas, as Figure 2, public domain)

These official projections were instituted to reduce the effort for surveyors to convert field measurements onto topographic maps. By choosing conformal projections, the angles measured in the field need no conversion (though in some cases adjustments for Grid North). Distances are corrected using a single scale factor that varies smoothly in a known manner.

Official projections typically use a secant geometry, where the reference surface (cylinder or cone) is deliberately smaller than the Earth. For example, the UTM sets the scale factor at 0.9996 in the central meridian (Figure 2). Therefore distances are -

.04% (short of true) in the neighbourhood of the central meridian. For simplicity, we will present the error factor (1-scale factor). This error reduces outwards across the UTM zone (symmetrically) until it reaches 0.0000 at a distance of 180 km approximately. Beyond this distance, the scale factor exceeds 1 so the error is positive, and distances are exaggerated. If applied according to the specification (in 6-degree bands), this error in distance is kept within bounds. However, the balance is maintained mostly at low latitudes. At 50N, the three degree width of the zone is 215.1 km wide, so much more of the zone is underestimated than overestimated.
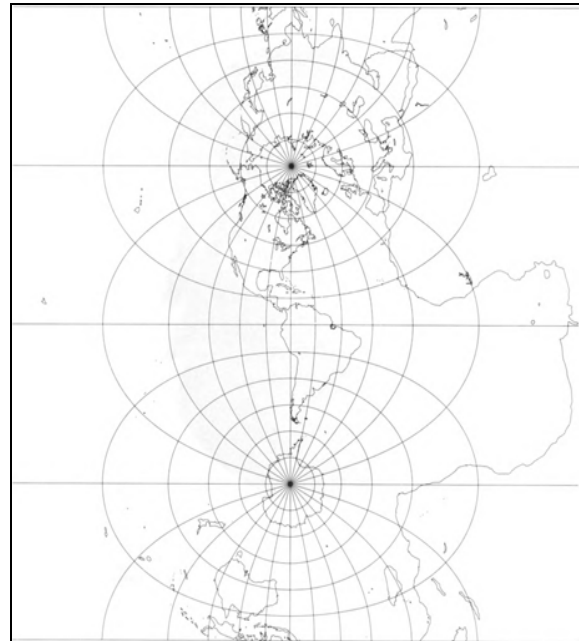

Figure 4: Transverse Mercator projection, central meridian at 75W – 105E. (From Snyder, 1989, p. 50, public domain). *Note that distortions make Africa unrecognizable, and scale error increases away from the central meridian.*

For more precise work, official coordinate systems typically apply scale factors closer to 1.000, and therefore are limited to smaller regions. The Gauss-Kruger coordinates for Germany and a number of other Central and Eastern European countries adopt 3-degree wide zones. The scale factor of 1.0000 applies on the central meridian so all distances are overestimated proportionately away from it. Canada adopted a Modified TM, based on 3-degree zones and a scale factor of 0.9999. The design criteria for the State Plane Coordinates of the United States specify a maximum absolute error factor of 0.0001 for each of 127 zones (Stem, 1989, p 3). For some small states, such as Delaware, the scale factor on the central meridian is set at 0.9999995. At this level, projection error is not significant for GIS applications.

Some zone systems are designed to cover larger spans than UTM. For example, the TM projection for the UK National Grid is used for the whole of Great Britain and adjacent islands, which extend more than 8 degrees of longitude. Though at the latitude of Northern Scotland the distance from the central meridian is smaller and, thus, the error effect is reduced.

2.3 GIS Practices that Increase Error

Zone systems are designed to apply over a limited area, but GIS applications tend to span these zones to encompass bigger landscapes and whole political jurisdictions. GIS databases are intended to be enterprise-wide, not just project based. So, for example, the Wisconsin Department of Natural Resources did not operate in the three zones of Lambert Conformal Conic projections that covered the state in three bands from North to Central to South. Instead, they designed their own Transverse Mercator with central meridian of 90W (a zone boundary of UTM). This DNR-TM is the de facto statewide projection system for the state, with a local origin that keeps coordinates in reasonable bounds. DNR was lucky in that Wisconsin is mostly confined between 87W and 93W, therefore complying with the 6 degree width of UTM zones. (Only Washington Island off the Door Peninsula is located outside the band, and the overestimate of distances on this island will not harm decision-making).
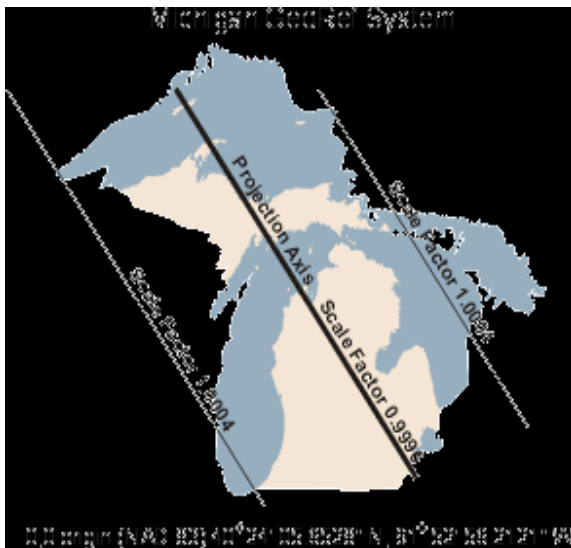

Figure 5: Michigan GeoRef oblique Mercator Source: http://www.michigan.gov/documents/DNR_Map_Proj_and_MI _Georef_Info_20889_7.pdf (Public domain)

Similarly, the State of Michigan sought a solution to having three zones of Lambert Conformal Conic for an oddly shaped entity. They devised an Oblique Transverse Mercator that fits the UTM design criteria. With judicious choice of azimuth, they have a single coordinate system for statewide purposes that balances the error factor between +0.0004 and -0.0004 (Figure 5). At least ten jurisdictions in USA dispensed with multiple zones in the recent overhaul related to the North American Datum changes in the 1980s and 90s (WAGIC, 2010, p. 2)

Other GIS designers are not so lucky. Their jurisdictions are larger, and the drive to have a unitary coordinate system is too strong. They have sometimes taken the strategy of extending coordinate systems beyond their design specifications. For example, the State of Washington uses the South Zone of their Lambert Conformal Conic to extend over the whole north zone for statewide use (Washington Information Services Board, 2011, p.3). Figure 6 demonstrates the error factor in area (not the error in distances) which exceeds +0.0017 on the northern border of Washington.
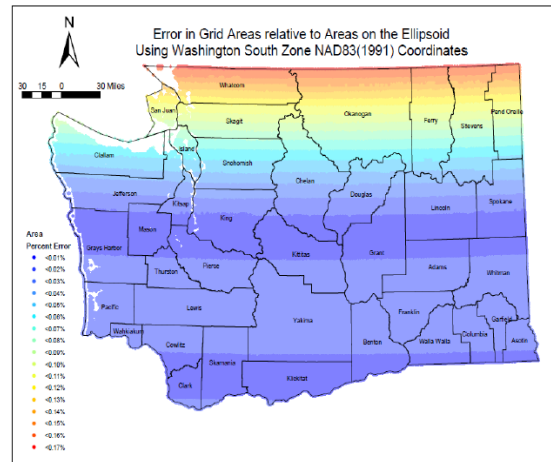

Figure 6: Error factor in area (maximum value +0.17%) for statewide use of Washington South Zone. Source: WAGIC, 2010)

Similarly, France has established official coordinates called the "grille Lambert", based on the same Lambert Conformal Conic. The old zone system had four zones (in east-west bands as conical coordinate systems operate). It was accepted practice to use zone II (which is applied in the center of the French continental area) for the whole of France. As a consequence, the error factor reached +0.005 at 41N (in Corsica). Since 2010, the replacement national system for France is the RGF Lambert 93 projection system. This system has error factors of -0.001 in the center of France, +0.002 at Dunkerque in the north, and +0.003 in Corsica (Figure 7).
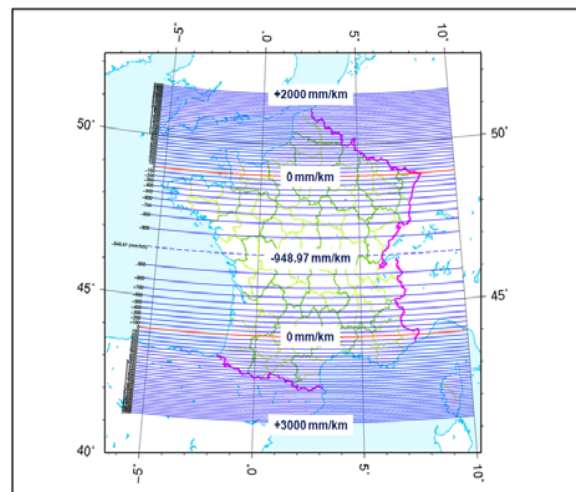

Figure 7: Scale factor of the RGF Lambert93 projection. Source: IGN-SGN/SICA, 2009

For specific usages of geospatial information which requires minimising the error factor, the addition of local systems was decided. One proposal suggested 39 zones for France to bring the maximum scale error down to 1cm/km (0.00001). However the implementation of so many zones was deemed impractical. Finally, 9 zones were selected, in order to bring the maximum local scale error to 8 cm/km (Figure 8).
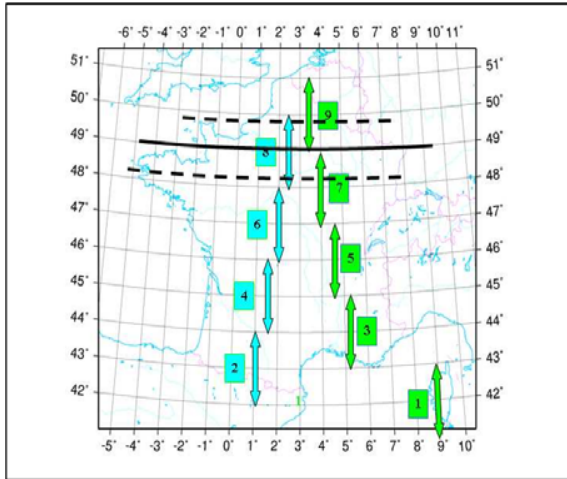
Figure 8: Layout of the "Lambert 9 zones" conformal conic projections for France used to minimise scale error. Source: http://geodesie.ign.fr/index.php?page=documentation

3. CONSEQUENCES OF PROJECTION ERROR ON ANALYTICAL MEASURES

3.1 Estimates of error in perimeter and area

Most treatment of projection error stops with the scale factor, the deviation of distances at a point. This factor can be directly applied to measurements of perimeter for polygons, or lengths of polylines. Girres (2012, p. 119) performed a test on the Lambert93 projection for one department in the Southwest of France (Pyrénées-Atlantiques, Figure 9). The projection error for the primary road network generates an overestimation of about 0.05% of the total length.
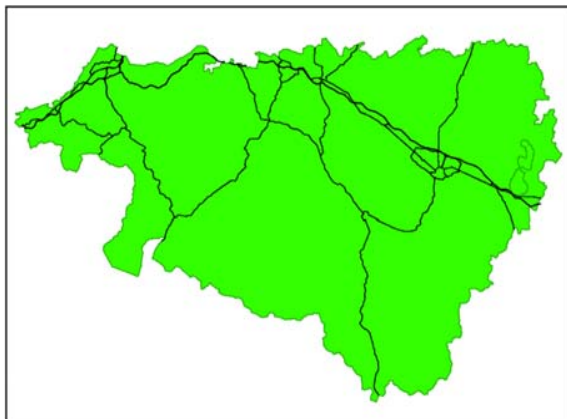

Figure 9: Primary road network of Pyrénées-Atlantiques (From Girres, 2012, p. 119). Source: IGN-BDTOPO.

For conformal projections, since they maintain isotropic scale factors, the area error factor e_A can be obtained from the linear error e_L according to formula (1) (Girres, 2012, p. 112; CERTU, 2010, p.4).

$$e_A = 2 * e_L + e_L^2 \quad (1)$$

Therefore, it is typically double the linear scale error. For the area of Pyrénées-Atlantiques, the overestimation on the Lambert93 projection was 0.12% of the total area. As shown in Figure 10, the computation of the ellipsoidal area for the department of Pyrénées-Atlantiques requires the oversampling of the polygon (using triangles), since the scale error varies

non-linearly on the entire zone. The area of each triangle was calculated using ellipsoidal trigonometry. This figure is consistent with the estimates for the State of Washington case.
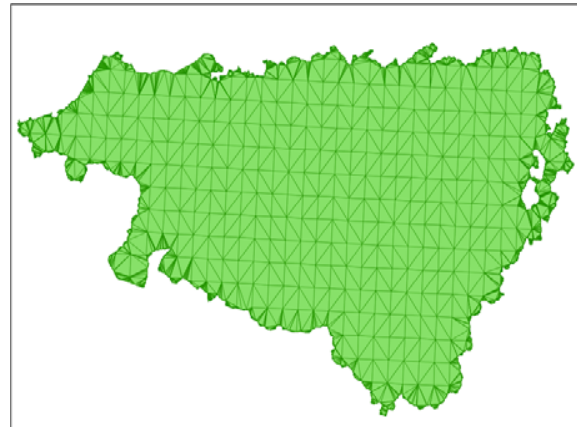

Figure 10: Spherical triangles used to estimate the true area of Pyrénées-Atlantiques using geodetic (ellipsoidal) calculations. (From Girres, 2012, p. 117)

Overall, these figures may seem small, but since they are systematic, they should not occur at all.

3.2 Policy results

The scale errors reported above can add up to differences in public policy. Harmel (2009, p. 29) reports that France was able to save 17 million euro annually (in the subsidies of the Common Agricultural Policy) after the areas of farm properties were recalculated on the new official RGF Lambert93 projection, by comparison with the old reference system (Zone II extended, shown in black). As shown in Figure 11, the error of Lambert93 projection is systematically inferior to Lambert II (extended) in any latitude, and reduces the error factors overall.
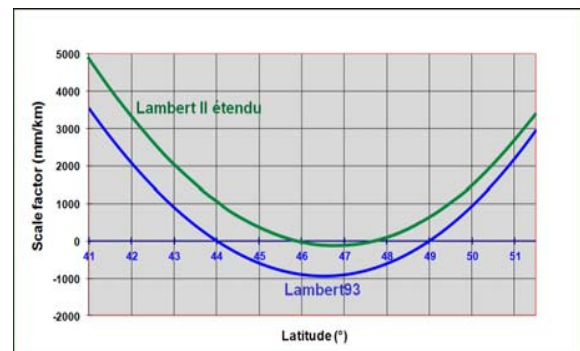

Figure 11: Scale factors differences between RGF Lambert93 and NTF Lambert 2 projections according to the latitude (From Harmel, 2009, p. 30.)

Due to repositioning of the standard parallels, the scale factor difference between the two projection systems increases southwards. Through these differences, the overestimates by the prior projection system in Southern France were significant and cumulative. The area estimates of farm properties revised on the new official projection system (while still subject to some error) were substantially lower. Other situations in other countries may be at least as large, though undocumented as yet.

4. MITIGATING ERROR FROM PROJECTIONS

As described above, projection errors are systematic, not the result of some random and unknown influence. Thus they can be predicted with nearly perfect accuracy- in principle. We argue that from an ethical perspective, a professional has the duty to remove any errors which are known. This section will consider solutions to manage this process.

4.1 Use of local scale error estimate

The calculations to project a point from geographic (angular) coordinates to the projection can also be solved to provide the scale error at each point (Snyder, 1987; Stem, 1989, p.18; IGN-SGN, 2010). These values vary systematically and symmetrically from the central meridian or standard parallels. For objects like parcels or urban street segments, the scale error will not change substantially across the object. Therefore a measure such as perimeter could simply be scaled by the average scale error (or simply a representative value). If the scale error is considered to be an attribute of the feature, then the user can see how the values computed on the projection are scaled in a quite transparent way.

Some objects are large enough that they will have substantial variation internally. The scalar would need to be a weighted average. The exact requirements for this may not be clear in the case of areas when the internal portions of a large polygon may be at a different scale error from the perimeter. Further geometric consideration would have to be undertaken, as shown in Figure 10.

4.2 Changing projections

Another solution would be to change from conformal projections. An equal-area (equivalent) projection might provide a solution for area calculations that would not require rescaling. For instance, the INSPIRE directive recommends the use of the ETRS89 Lambert Azimuthal Equal-Area coordinate reference system for statistical analysis and display reporting in the European Union, where true area representation is required (INSPIRE, 2010; Figure 12).
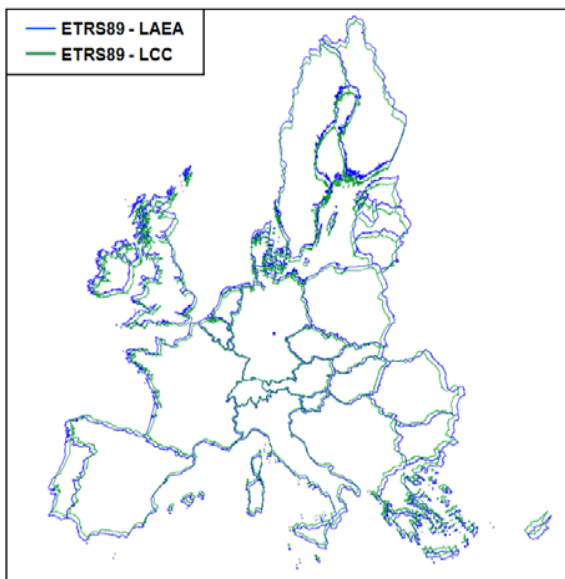

Figure 12: A representation of European Union countries using Lambert Azimuthal Equal-Area (in blue) and Lambert

Conformal Conic (in green) projections. (inspired from Dana, 1997). Data Source: http://epp.eurostat.ec.europa.eu/

Sadly, equivalent projections do not preserve distances (or angles). Therefore, the perimeter would still need to be rescaled using the scale error estimates given above. Note that these are more complicated for non-conformal projections in that they are non-isotropic. The distance would have to be recalculated with weighting for each segment, depending on its orientation.

Use of non-standard projections will require the typical vigilance in integrating sources, but that is a part of normal practice for different datums and projection parameters.

4.3 A radical alternative

An alternative must be considered. If all GIS measurements are maintained on the ellipsoid and all analytical results calculated with ellipsoidal trigonometry, there would be no projection error. The formulae for distances on the ellipsoid are well-known and covered in basic geodesy courses (Vincenty, 1975; Deakin and Hunter, 2007). But they are then often ignored by GIS software designers.

Indeed, the OpenSource GIS software Quantum GIS (Quantum Gis Development Team, 2013) allows the manually computation of ellipsoidal measures, based on Vincenty's formula (Vincenty, 1975). Unfortunately, the "on the fly" computation of geometric length or area on vector objects is still realised using map projections, without considering the scale error.
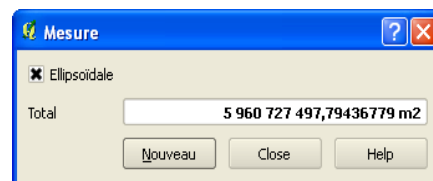

Figure 13: Measurement tool of Quantum GIS 1.7.2 software, allowing the computation of ellipsoidal length and area. Source: http://www.qgis.org/

With modern computational capacity the increased burden would not be unacceptable. Our preliminary estimate is that a calculation that takes 5 floating point operations (probably in double precision), and one square root will be replaced by maybe 8 multiplications and 4 trigonometric function calls.

Areas are not so easy. The planar calculation usually divides the polygon into trapezoids (extending to the axis or some reference line). On the ellipsoid the strategy would have to be reconsidered, perhaps with ellipsoidal triangles radiating from the centroid, or a Delaunay triangulation as demonstrated in Figure 10. The authors do not doubt however that a solution can be found.

5. CONCLUSIONS

In order to reduce the harm done by map projections, GIS professionals have a duty to report measures like perimeter and area after correction for known systematic distortions. In the same way, other well-known sources of error in length and area measures, such as the systematic underestimation involved by the non-consideration of the relief, should also be estimated and

reported to the final user (Girres, 2011; 2012). Some will argue that these are small effects, but even so, they should not persist.

REFERENCES

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

Dana, P.H., 1997. The Geographer's Craft Project. Department of Geography, The University of Colorado, Boulder, USA.
http://www.colorado.edu/geography/gcraft/notes/mapproj/gif/threepro.gif

Deakin, R. E. and Hunter, M. N., 2007, Geodesics on an ellipsoid – Pittman's method, Proceedings of the Spatial Sciences Institute Biennial International Conference (SSC2007), Hobart, Tasmania, Australia, 14-18 May 2007, pp. 223-242.

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, pp. 83-88, Coimbra, Portugal.

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 Université Paris-Est, France.

Harmel, A., 2009. Le nouveau système réglementaire Lambert 93. *Géomatique Expert,* 68, pp. 26-30.
http://www.geomag.fr/sites/default/files/68_91.pdf

IGN-SGN/SICA, 2009. *Réseau géodésique français, concepts et applications*. Institut Géographique National – Service de Géodésie et Nivellement / Service IGN Conseil et Applications, St. Mandé, France.

IGN-SGN, 2010. *Circé france version 4*. Institut Géographique National - Service de Géodésie et Nivellement, St. Mandé, France.
http://geodesie.ign.fr/contenu/fichiers/Plaquette_CirceFrance_v4.pdf

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

Ordnance Survey, 2010. *A guide to coordinate systems in Great Britain,* OS, Southampton, UK.
http://www.ordnancesurvey.co.uk/oswebsite/gps/docs/A_Guide_to_Coordinate_Systems_in_Great_Britain.pdf

Quantum GIS Development Team, 2013. *Quantum GIS Geographic Information System*. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

Sillard, P, 2000. *Les projections et référentiels cartographiques*. Ecole Nationale des Sciences Géographiques, Marne-la-Vallée, France.
http://fad.ensg.eu/moodle/file.php/11/Geodesie/projections.pdf

Snyder, J.P., 1987, *Map projections – A working manual*, Professional Paper 1395, US Geological Survey, Reston VA, USA, third revised printing 1994.
http://pubs.usgs.gov/pp/1395/report.pdf

Stem, J. 1989. *State Plane Coordinate System of 1983*. NOAA Manual NOS NGS 5, National Geodetic Survey, Rockville MD, USA.
http://www.ngs.noaa.gov/PUBS_LIB/ManualNOSNGS5.pdf

Vincenty, T., 1975. *Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations*. Survey Review 22, pp. 88-93.
http://www.ngs.noaa.gov/PUBS_LIB/inverse.pdf

WAGIC (Washington Geographic Information Council), 2010. *Adopting a single State Plane Coordinate System: Call to Action*. Draft document (proposal not implemented).
http://geography.wa.gov/wagic/zonetostate.pdf

Washington Information Services Board, 2011, *Geographic Information Systems (GIS) Geodetic Control Standards*. Standard No. 601-S2.
http://ofm.wa.gov/ocio/policies/documents/161.10.pdf