MAP PROJECTION DESIGN

Alan Vonderohe (January 2020)

Background

For millennia cartographers, geodesists, and surveyors have sought and developed means for representing Earth's surface on two-dimensional planes. These means are referred to as "map projections". With modern BIM, CAD, and GIS technologies headed toward three- and four-dimensional representations, the need for two-dimensional depictions might seem to be diminishing. However, map projections are now used as horizontal rectangular coordinate reference systems for innumerable global, continental, national, regional, and local applications of human endeavor. With upcoming (2022) reference frame and coordinate system changes, interest in design of map projections (especially, low-distortion projections (LDPs)) has seen a reemergence.

Earth's surface being irregular and far from mathematically continuous, the first challenge has been to find an appropriate smooth surface to represent it. For hundreds of years, the best smooth surface was assumed to be a sphere. As the science of geodesy began to emerge and measurement technology advanced, it became clear that Earth is flattened at the poles and was, therefore, better represented by an oblate spheroid or ellipsoid of revolution about its minor axis. Since development of NAD 83, and into the foreseeable future, the ellipsoid used in the United States is referred to as "GRS 80", with semi-major axis a = 6378137 m (exactly) and semi-minor axis b = 6356752.314140347 m (derived).

Specifying a reference ellipsoid does not solve the problem of mathematically representing things on a two-dimensional plane. This is because an ellipsoid is not a "developable" surface. That is, no part of an ellipsoid can be laid flat without tearing or warping it. There are a number of developable surfaces. The most frequently used are cones and cylinders. A cone can be cut from its base to its apex and laid flat. Similarly, a cylinder can be cut parallel with its axis, unrolled, and laid flat. Two-dimensional rectangular coordinate axes (e.g., northing and easting, X and Y) can then be established on these surfaces. See Figure 1.

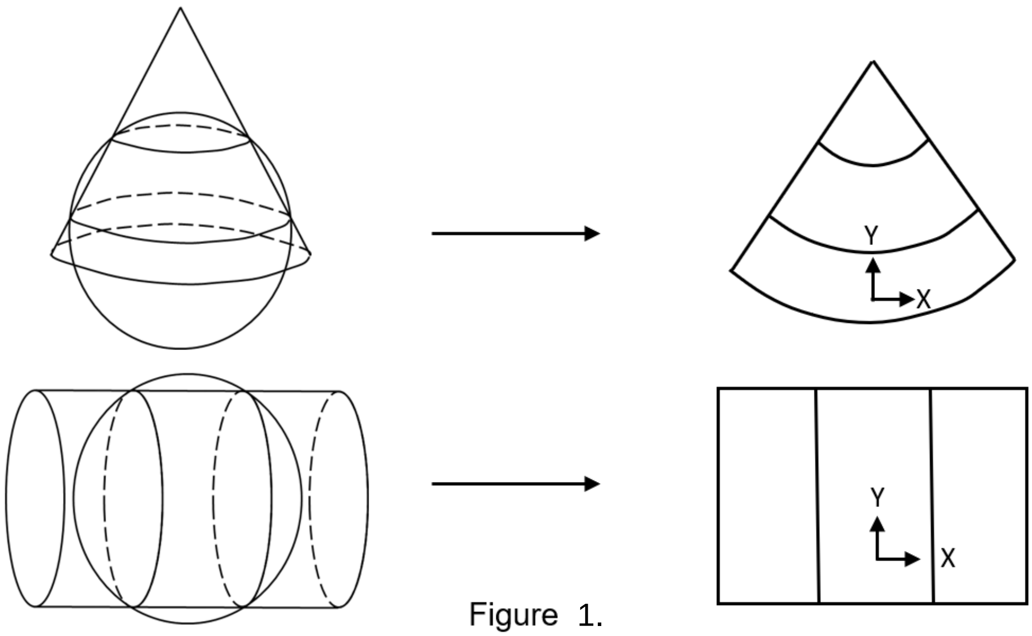


Figure 1.

Conical and Cylindrical Map Projections

Map projection surfaces can be secant to the reference ellipsoid, intersecting it along two lines, as shown in Figure 1. They can be tangent to it, intersecting it along a single line, or they can be non-intersecting with it.

There are functional relationships between points having geodetic coordinates (latitude and longitude) on the reference ellipsoid and corresponding points with two-dimensional rectangular coordinates on a map projection surface. A map projection can be described by mathematical transformations between these two types of coordinates. A “direct” transformation computes northing and easting (N,E) from latitude and longitude (ϕ,λ):

$$N = f_1(\phi,\lambda, \text{ellipsoid parameters, map projection parameters})$$
$$E = f_2(\phi,\lambda, \text{ellipsoid parameters, map projection parameters}) \qquad 1$$

An “inverse” transformation computes ϕ,λ from N,E:

$$\phi = g_1(N,E, \text{ellipsoid parameters, map projection parameters})$$
$$\lambda = g_2(N,E, \text{ellipsoid parameters, map projection parameters}) \qquad 2$$

In the equations above, ellipsoid parameters are two descriptors that define the size and shape of the reference ellipsoid (e.g., a and b). Map projection parameters are descriptors that define the size and shape of the map projection surface and its location and orientation with respect to the reference ellipsoid. Map projection parameters also define the location of the rectangular coordinate origin and its false northing and false easting values.

Because map projection surfaces do not coincide with the reference ellipsoid, except at lines of intersection or tangency, features projected from the ellipsoid to the map projection will be distorted. The first step in map projection design is selection of the type of map projection based upon what spatial aspects are least distorted. The appropriate choice depends upon applications to be supported and user desires.

Some projections tend to preserve areas and are referred to as “equal-area”. On such projections, areas of equal size on Earth’s surface appear as areas of equal size on the map projection.

“Equidistant” projections show no scale variation between one or two points and every other point on the map, or along every meridian. “Azimuthal” projections are such that the azimuths to all points on the map are shown correctly with respect to its center (Snyder (1987)).

Some projections tend to preserve shape, that is, local angles are the same on Earth’s surface as they are on the map projection surface. At any particular point on such a projection, scale is constant and independent of direction, although scale will vary from point to point. These projections are referred to as “conformal” and are the type of projection suitable for a host of applications in geodesy, surveying, and mapping.

The conic and cylindrical projections shown in Figure 1 are referred to as “Lambert conformal conic” and “transverse Mercator”, respectively. On these projections, each point has a scale factor (k) such that

$$k = h_1(\phi,\lambda, \text{ellipsoid parameters, map projection parameters}) \qquad 3$$

where k is a positive real number expressing the ratio of an infinitesimal distance on the map projection surface to the corresponding infinitesimal distance on the reference ellipsoid. If $k < 1$, then the map projection surface is interior to the reference ellipsoid at ϕ,λ. If $k > 1$, then the map projection surface is exterior to the reference ellipsoid at ϕ,λ. If $k = 1$, then the map projection surface and the reference ellipsoid intersect at ϕ,λ. See Figure 2.

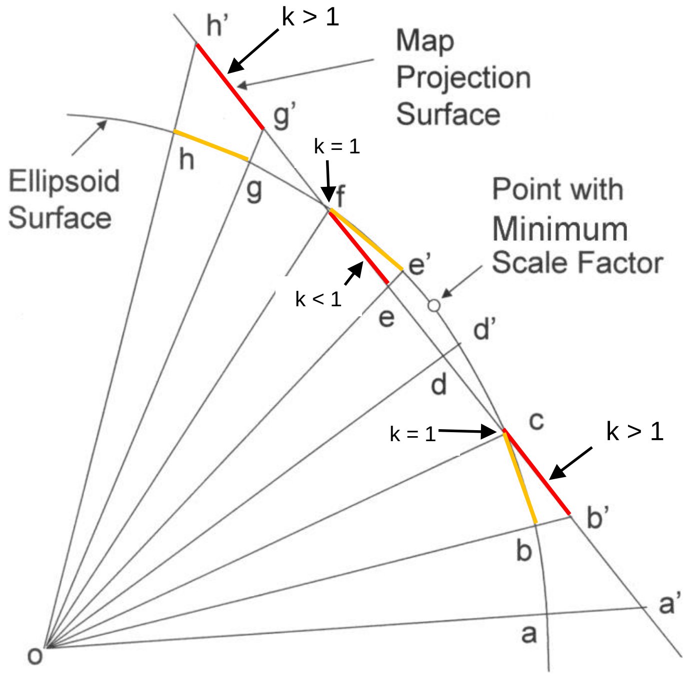


Figure 2.
A Secant Map Projection Surface Indicating Scale Factors

Non-zero scale distortion, arising from k ≠ 1, is expressed as either a ratio of 1 : XXXXXX where

$$XXXXXX = | \text{int} [1 / (k - 1)] | \quad (4)$$

or in integer parts per million (ppm), calculated as int [round { (k – 1) * 10^6, 0 }]. In the former case, XXXXXX is always positive. In the latter case, the parts per million expression can be positive or negative. In any case, if k = 1, scale distortion equals zero.

The primary criterion for map projection design is often a specified maximum absolute value for tolerable distortion. For conformal projections, scale distortion is the controlling variable, with the tolerance specified as in equation 4 or as a range such as ±100 ppm.

On Lambert conformal conic projections, scale factors vary most dramatically in the north-south direction. Therefore, they are well-suited for areas whose east-west extents are greater than their north-south extents. Conversely, on transverse Mercator projections, scale factors vary most dramatically in the east-west direction, making them well-suited for areas whose north-south extents are greater than their east-west extents.

After choosing the appropriate type of map projection for the applications to be supported, the next step in design involves either:

1. Specifying map projection parameters that meet the design criteria, then finding the geographic extents of the area to be covered; or
2. Specifying geographic extents of the area to be covered, then finding map projection parameters for optimal distribution of distortion.

The first method was used by the US Coast and Geodetic Survey (USC&GS) for design of the State Plane Coordinate System of 1927 (SPCS 27). The second method was used for design of the Wisconsin County Coordinate System (WCCS).

Map Projection Parameters

Parameters for Lambert conformal conic map projections depend upon whether the projection surface is secant to or non-intersecting with the reference ellipsoid. If secant to, the parameters are:

1. ϕ_N = Latitude of the northern standard parallel.
2. ϕ_S = Latitude of the southern standard parallel.
3. ϕ_b = Latitude of the coordinate origin.
4. λ_o = Longitude of the central meridian and the coordinate origin.
5. N_b = False northing of the coordinate origin.
6. E_o = False easting of the coordinate origin.

where the standard parallels are lines of constant latitude at the intersections of the cone and the reference ellipsoid, as in Figure 1; the latitude of the coordinate origin and the longitude of the central meridian are often chosen central to the geographic extents of the projection, and the false northing and false easting are chosen to make positive all northings and eastings within the geographic extents of the projection.

If non-intersecting, the parameters for a Lambert conformal conic projection are:

1. ϕ_o = Latitude of the central parallel and the coordinate origin.
2. λ_o = Longitude of the central meridian and the coordinate origin.

3. k_o = Scale factor along the central parallel.
4. N_o = False northing of the coordinate origin.
5. E_o = False easting of the coordinate origin.

where the central parallel and the central meridian are often chosen central to the geographic extents of the projection, the scale factor on the central parallel is chosen to optimize scale factors across the geographic extents of the projection, and the false northing and false easting are chosen to make positive all northings and eastings within the geographic extents of the projection. If a Lambert conformal conic projection is tangent to the reference ellipsoid, the line of tangency is the central parallel and its scale factor is 1. Also, the central parallel and its scale factor can be computed from the parameters of a secant Lambert conformal conic projection.

The parameters for a transverse Mercator projection, whether it is secant to, tangent to, or non-intersecting with the reference ellipsoid are:

1. ϕ_o = Latitude of the coordinate origin.
2. λ_o = Longitude of the central meridian and the coordinate origin.
3. k_o = Scale factor along the central meridian.
4. N_o = False northing of the coordinate origin.
5. E_o = False easting of the coordinate origin.

where the latitude of the coordinate origin is often chosen as the southernmost geographic extent of the projection, the longitude of the central meridian is chosen central to the geographic extents of the projection, the scale factor along the central meridian is chosen to optimize scale factors across the geographic extents of the of the projection, and the false northing and false easting are chosen to make positive all northings and eastings within the geographic extents of the projection.

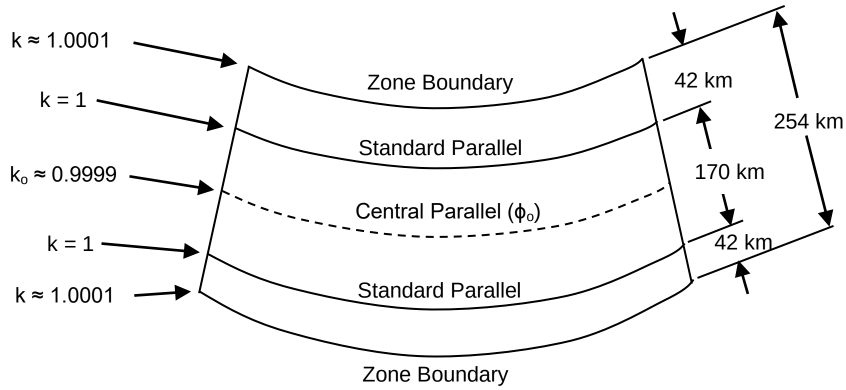
For secant projections, $k_o < 1$. For tangent projections $k_o = 1$. For non-intersecting projections, $k_o > 1$. For all, k_o is the minimum scale factor across the geographic extents of the projection. k_o is also one of two critical parameters for both Lambert conformal conic projections and transverse Mercator projections. In the former case ϕ_o is the second critical parameter. In the latter case λ_o is the second critical parameter. These pairs of critical parameters fix the size of their respective projection surface and its orientation with respect to the reference ellipsoid. Therefore, they control the size and distribution of scale distortion across their projections' geographic extents and their selection or computation is based upon the primary design criterion.

State Plane Coordinate System 1927, 1983, 2022

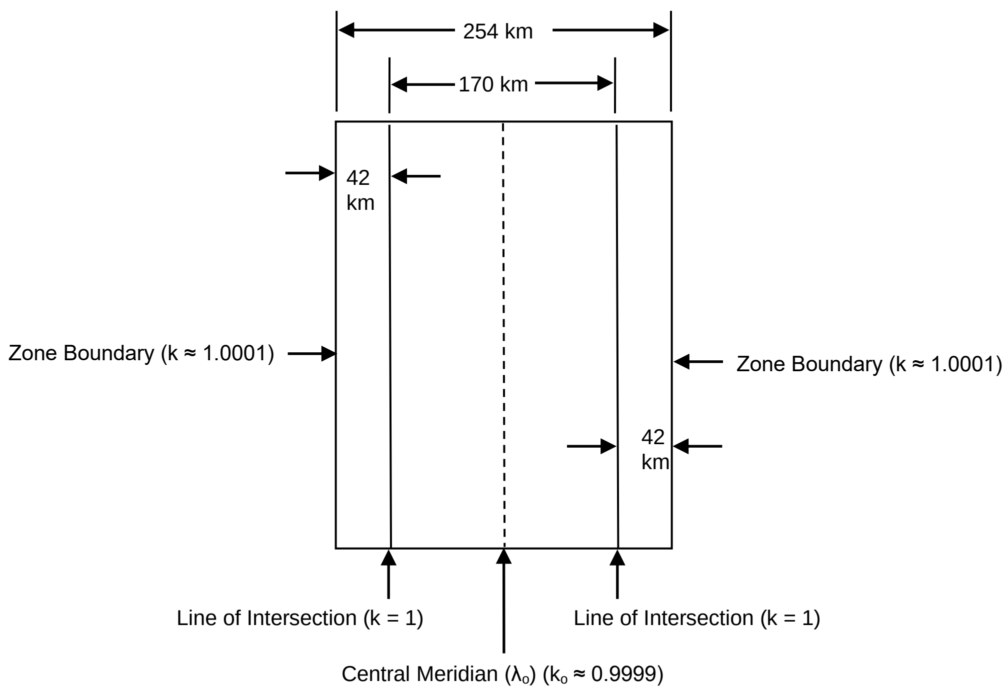
During the mid-1930s, USC&GS designed and published the nationwide SPCS 27, based upon NAD 27 and the Clarke 1866 reference ellipsoid. The intent was to enable surveyors, mappers, and engineers to tie their land and engineering surveys to NAD 27 (Stem (1989)). Three types of conformal map projections were used: 1) Lambert conformal conic with two standard parallels, 2) transverse Mercator, and 3) oblique Mercator in which the central axis does not coincide with a meridian. Design criteria included defining map projection zones as aggregations of counties, covering entire states with as few zones as possible, and restricting the maximum scale distortion to less than 1:10000. This final primary criterion was based upon typical survey accuracies in the 1930s. The most prominent distance measuring device was a 100-foot steel tape with a least graduation of 0.01 foot. That least graduation is 1 part in 10000 of the full length of a tape.

For both Lambert conformal conic with two standard parallels and transverse Mercator projections, the 1:10000 scale distortion criterion translates into maximum zone widths of approximately 254 km (158 miles) in the United States. If 2/3^{rds} of a zone width is inside the lines of intersection and 1/6th of its width lies outside each line of intersection, scale distortion varies from +100 ppm near its long sides to -100 ppm near its central meridian or parallel (Moffitt and Bossler (1998)). See Figure 3.

With these concepts and the other design criteria, the nationwide SPCS 27 came into being with Wisconsin having its familiar three Lambert conformal conic projections (North Zone, Central Zone, South Zone). As specified, zone boundaries coincide with county boundaries. It is important to note that, as with any map projection, the zone boundaries are administrative, not mathematical. That is, it is possible to compute South Zone coordinates and scale factors for points that are actually in the Central Zone, or the North Zone for that matter. It is just that the primary design criterion of 1:10000 maximum scale distortion might be violated for points outside any given zone.



a. Lambert Conformal Conic



b. Transverse Mercator

Figure 3.
Zone Configurations for Maximum Scale Distortions of 1:10000

With the advent of a new horizontal datum (NAD 83) in 1986, the National Geodetic Survey (NGS), successor to USC&GS, designed and published SPCS 83 which is tied to NAD 83 and the GRS 80 reference ellipsoid. At the time, many states had adopted SPCS 27 zone definitions and parameters into legislation, the surveying and mapping community was used to working with SPCS 27, and there were no compelling reasons to make significant changes to map projection types and zone boundaries. With some exceptions requested by particular states, SPCS 83 has the same map projection types, zones, and zone boundaries. Nearly all map projection parameters also remain the same with exceptions for special cases and for false northings and false eastings. The latter were changed for two reasons: 1) SPCS 27 false northings and false eastings had been published in US survey feet and SPCS 83 false northings and false eastings were to be published in multiples of 100000 meters with some exceptions and 2) there were to be noticeable differences between coordinates on SPCS 27 and SPCS 83 to avoid confusing them (Stem (1989)).

NGS is taking a very different approach to SPCS2022 (NGS (2019a), NGS (2019b)). The North American Terrestrial Reference Frame of 2022 (NATRF2022) will provide a highly accurate, time-dependent, three-dimensional geometric reference frame tied to the GRS 80 reference ellipsoid (NGS (2017a)). The North American Pacific Geopotential Datum of 2022 (NAPGD2022) will also be time dependent and will provide, among other things, a geoid height model referenced to NATRF2022 and accurate to 1-2 cm (NGS (2017b)). NATRF2022 and NAPGD2022 will enable highly accurate Global Navigation Satellite System (GNSS) surveys in three dimensions, including orthometric heights or elevations. Given these developments, NGS will also develop a modernized State Plane Coordinate System. Under SPCS2022, each state can have as many as three “layers” of map projections:

1. A single-zone statewide layer designed by NGS.
2. A multiple-zone layer, similar to SPCS 83, designed by NGS.
3. A multiple-zone layer of “low-distortion projections” (LDPs), designed by the individual states with approval from NGS.

Only one of layers 2 and 3 can have statewide coverage and the LDPs in layer 3 are subject to a number of NGS specifications. The collection of layers is referred to as “SPCS2022” and will be published and supported by NGS in their publicly-available software.

Perhaps the most innovative aspect of SPCS2022 is that map projection design will be performed at Earth’s surface, not the reference ellipsoid surface. That is, scale distortion will not be a standalone primary design criterion. Rather, it will be coupled with an “ellipsoid height” factor to produce “linear” distortion as the primary design criterion.

Designing at Earth's Surface

We understand scale factor to be the ratio of an infinitesimal distance on a map projection surface to the corresponding infinitesimal distance on the reference ellipsoid. However, we live upon, perform surveys upon, and want maps of Earth's surface. Therefore, measurements and features must be transformed from Earth's surface to the reference ellipsoid before they can be projected onto a map. The multiplier for transforming an infinitesimal distance on Earth's surface into its corresponding infinitesimal distance on the reference ellipsoid was formerly referred to as the "sea level" factor, then later as the "elevation" factor. This multiplier might now best be referred to as the "ellipsoid height" factor, because that is what it actually is.

Figure 4 illustrates the relationship between an Earth's surface distance and its corresponding reference ellipsoid distance. h is ellipsoid height, H is orthometric height or elevation, and N is geoid height (above the reference ellipsoid). Across Wisconsin, N is negative, not positive as shown in Figure 4. R_G is the Gaussian or geometric mean radius of curvature of the reference ellipsoid at the latitude (ϕ) of the point in question:

$$R_G = \frac{a\sqrt{1-e^2}}{1-e^2 sin^2\phi} \qquad 5$$

where

$$e^2 = \frac{a^2-b^2}{a^2}$$

and e is the eccentricity of the reference ellipsoid. From Figure 4:

$$D_{ellip} = D_{Earth} * E \qquad 6$$

where E is the ellipsoid height factor, given by:

$$E = R_G / (R_G + h) = R_G / (R_G + H + N) \qquad 7$$

NOTE: Until 2022, values of N obtained from NGS geoid models are not really geoid heights. Rather, they are heights of the vertical datum surface (NAVD 88) above or below GRS 80 on NAD 83. NAPGD2022 will provide true geoid heights.

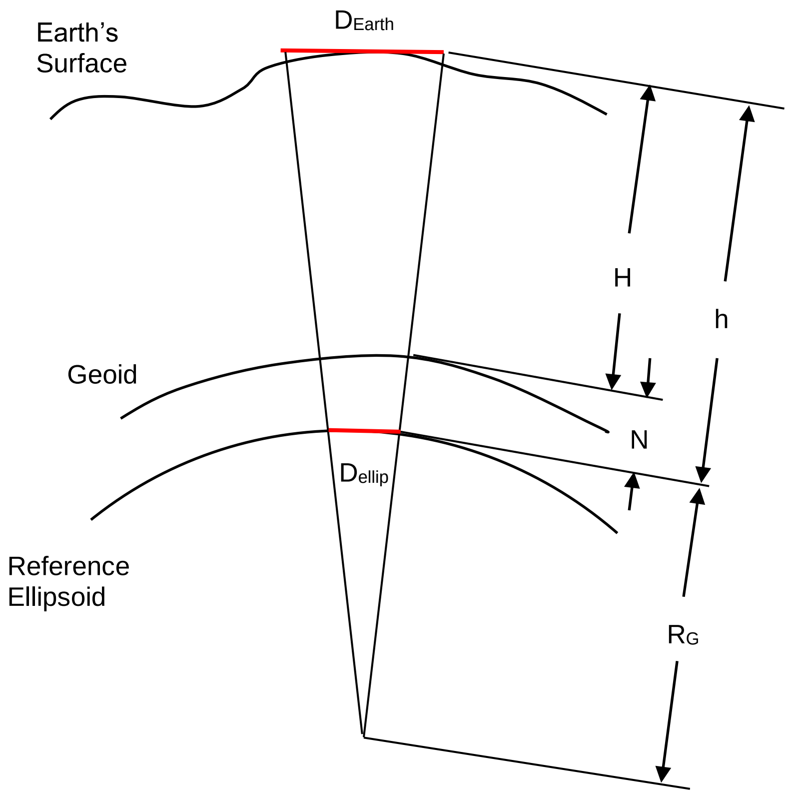


Figure 4.
Ellipsoid Height Factor (Greatly Exaggerated)

To relate Earth surface distances to map projection distances, we use both scale factor and ellipsoid height factor:

$$D_{map} = D_{Earth} * C \qquad 8$$

Where C is the "combined" factor (also referred to as the "combination" factor), given by:

$$C = (k)(E) \qquad 9$$

The value of C minus 1 is the "linear" distortion at the point in question. Linear distortion is the primary criterion when designing at Earth's surface. k_o, ϕ_o (for Lambert conformal conic projections), and k_o, λ_o (for transverse Mercator projections) remain the critical parameters, but ellipsoid heights as well as geographic extents must be taken into account when selecting or computing their values.

If there is a general slope across the geographic extents, it might influence the appropriate choice of map projection type. Lambert conformal conic projections are amenable to general north-south slopes and transverse Mercator projections are amenable to general east-west slopes (Dennis (2016)). Of course, these characteristics must be balanced against those of scale distortion. The effects of variation in scale factor and variation in ellipsoid height factor can tend to offset one another. Table 1, taken from NGS (2019b), indicates relationships among linear distortion, zone width, and range in elevation.

Table 1.
Relationships among Linear Distortion, Zone Width, and Range in Elevation

Linear distortion range Expressed in parts per million (ppm) and as a ratio	**Corresponding zone dimension and height limits**	
	Zone width perpendicular to projection axis (for no variation in topographic height)	Topographic height range (independent of zone width)
±1 ppm (1:1,000,000)	25 km (16 miles)	13 m (42 ft)
±5 ppm (1:200,000)	57 km (35 miles)	64 m (209 ft)
±10 ppm (1:100,000)	81 km (50 miles)	127 m (418 ft)
±20 ppm (1:50,000)	114 km (71 miles)	255 m (836 ft)
±30 ppm (1:33,333)	140 km (87 miles)	382 m (1,254 ft)
±40 ppm (1:25,000)	161 km (100 miles)	510 m (1,672 ft)
±50 ppm (1:20,000)	180 km (112 miles)	637 m (2,090 ft)
±75 ppm (1:13,333)	221 km (137 miles)	956 m (3,135 ft)
±100 ppm (1:10,000)	255 km (158 miles)	1,274 m (4,180 ft)
±150 ppm (1:6,667)	312 km (194 miles)	1,911 m (6,271 ft)
±200 ppm (1:5,000)	360 km (224 miles)	2,548 m (8,361 ft)
±300 ppm (1:3,333)	441 km (274 miles)	3,823 m (12,541 ft)
±400 ppm (1:2,500)	510 km (317 miles)	5,097 m (16,722 ft)

Low-Distortion Projections (LDPs)

The author finds no consensus definition of “low-distortion projection” in the literature. However, for SPCS2022, NGS will not design map projections when linear distortions are required to be less than 1:20000 or ±50 ppm (NGS (2019b)). Many existing LDPs, including Wisconsin’s, were designed such that differences between ground and grid distances could be ignored for moderate-accuracy applications such as engineering design and construction and cadastral surveying and mapping. For such projections, linear distortion is near enough to zero that it can be considered negligible across their geographic extents. Such requirements are met by limiting the size of a projection zone (e.g., to counties) and either

1. Enlarging the reference ellipsoid so it is in proximity to Earth's surface. The map projection is secant to the enlarged reference ellipsoid and is carried right along with it (see Figure 5a). This technique causes ellipsoid height factors (E) to be near unity and, thus, scale distortion is the major component of linear distortion. This method was employed in Minnesota (Johnson (1985)) and in Wisconsin for its first set of LDPs (WisDOT (1993)).

or

2. Retaining the original reference ellipsoid and rescaling k_o with a multiplier of $1/E_p$ where E_p is a pseudo-ellipsoid height factor with a single value across the geographic extents of the map projection. This causes the map projection surface to be in proximity to Earth's surface and also be non-intersecting with the reference ellipsoid (see Figure 5b). k and E are nearly reciprocals of one another at individual points and the combined factor is near unity across the geographic extents of the map projection. This method was employed in Oregon (Dennis (2016)) and in Wisconsin for its second set of LDPs (Vonderohe (2006, 2019)).

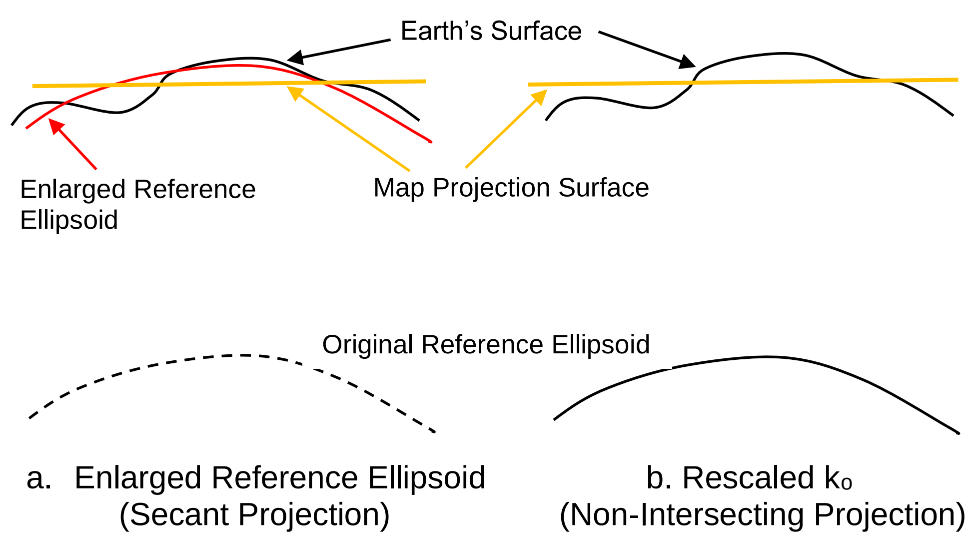


Figure 5.
Two Approaches to Designing Low-Distortion Projections

Wisconsin County Coordinate System (WCCS) Design

For many years, the Wisconsin Department of Transportation (WisDOT), county and municipal governments, land surveyors, and others used SPCS 27 and SPCS 83. Inherent scale distortions as large as 1:10000 and the need to first reduce ground distances by both the ellipsoid height factor and scale factor often posed nuisances and confusion among users. Some counties developed single combined factors from averages across their extents. WisDOT sometimes employed project-based combined factors. Lack of metadata and misunderstandings by users led to confusion about whether to multiply or divide and whether or not factors had already been applied to distances by previous users.

In 1993, WisDOT contracted for development of the Wisconsin County Coordinate System (WisDOT (1993)), a set of LDPs. The expectation was that confusion and misunderstanding concerning "ground-to-grid" reduction of distances would be greatly reduced, if not eliminated, because effects of combined factors could be ignored.

Under WCCS, each county was to be covered by a single projection and maximum linear distortions were to be 1:50000 (±20 ppm) in urban areas and 1:30000 (±33 ppm) in rural areas. In the final design, adjacent counties were placed under a single projection whenever doing so did not violate the maximum linear distortion criteria. WCCS used secant and tangent transverse Mercator projections and secant Lambert conformal conic projections. These projections intersect with their respective reference ellipsoids because the central strategy for WCCS design was using enlarged reference ellipsoids for each zone (see Figure 5a).

The enlarged reference ellipsoids were developed by adding a "pseudo" ellipsoid height (h_p) to both semi-axes of GRS 80:

$$a_{WCCS} = a_{GSR80} + h_p = a_{GSR80} + (\text{median elevation} + \text{mean geoid height}) \quad (10)$$

$$b_{WCCS} = b_{GSR80} + h_p = b_{GSR80} + (\text{median elevation} + \text{mean geoid height})$$

The median elevation for a county was computed from spot elevations and profiles drawn from the county's hardcopy USGS 1:100000-scale (10-foot contour) map. At least fifteen data points were used for each county. The county seat and populated areas were included in the county profile. The mean geoid height for a county was computed from the geoid heights at all NGS first and second order geodetic control points within the county. For each county, the enlarged ellipsoid, scale factors computed from Stem (1989), and the county's geographic extents were used to select a projection type and to place the projection with respect to the county boundaries. In a number of cases, more than one projection type and/or placement were tried and tested before the design was finalized. The final design often had standard parallels (for Lambert conformal conic projections) or central meridians and k_o (for transverse Mercator projections) shifted or modified to account for general slopes of the terrain. WCCS adopted existing projections when applicable (e.g., Brown County, Jackson County)

The linear distortion design criteria were met in all but four minor places across the state where the requirement that each county be covered entirely by a single projection forced exceeding the target maximum linear distortion. Ultimately, WCCS includes 59 zones for the 72 counties in Wisconsin (See Figure 6).

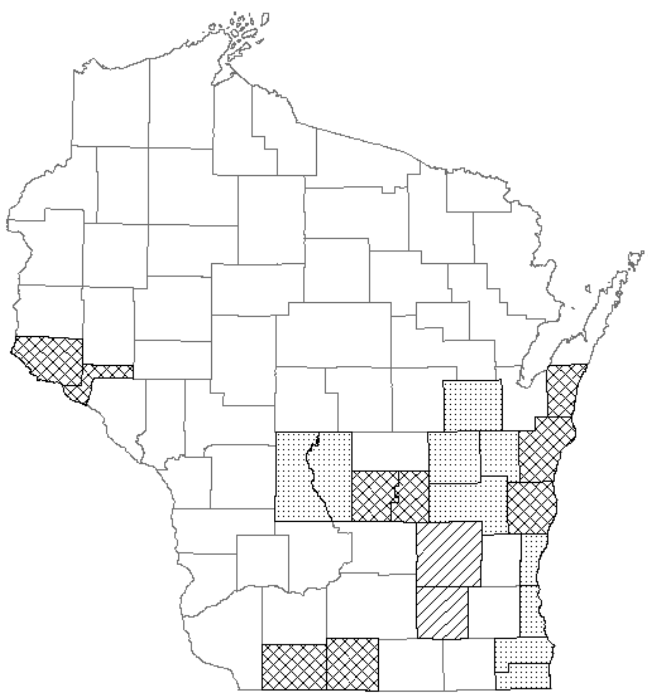


Figure 6.
Wisconsin County Coordinate System (Adjacent Counties with Shared Symbols Share a Single Map Projection – There are 72 Counties and 59 Map Projections (24 Lambert Conformal Conic and 35 Transverse Mercator))

In 1995, the Wisconsin State Cartographer's Office published WCCS (SCO (1995)) and its popularity began to grow. Among other activities, county governments were being funded by the Wisconsin Land Information Program to modernize their land records. These efforts often involved development of countywide digital cadastral maps, using survey plats and coordinate geometry (COGO) computations. WCCS enabled such work without first having to multiply ground-measured and platted distances by combined factors.

Wisconsin Coordinate Reference Systems (WISCRS) Design

NOTE: The design methodology for WISCRS, including full mathematical development, is thoroughly described in Vonderohe (2006, 2019). What follows is a brief summary.

Over time, difficulties with the enlarged reference ellipsoid approach began to emerge in the user and vendor communities. The vast majority of geospatial technology users were unaware of the fundamentals of geodesy and map projections. Even for those who were technically well-grounded, this innovative design approach was not well-understood. Changing ellipsoid parameters implies changing the underlying geodetic datum. Some commercial software did not support customized reference ellipsoids at all. Other software, that allowed user-specified ellipsoids, usually led down datum transformation paths that were inappropriate for WCCS.

In 2004, the Wisconsin Land Information Association (WLIA) formed a Task Force on Wisconsin Coordinate Systems. After considerable deliberation and development of supporting documentation, the Task Force recommended re-design of WCCS in such a way that all projections used the GRS 80 ellipsoid only. Keenly aware of legacy databases and desiring to diminish confusion and minimize costs in the user and vendor communities, the Task Force further recommended that the re-designed WCCS be such that coordinate differences between the original and re-designed systems be negligible. That is, there should be no need to modify or transform any existing data before adopting the re-designed WCCS. It was expected that data sets referenced to either the original or re-designed WCCS could be used together without concern for positional discrepancies. The Task Force chose a target maximum tolerance of five millimeters in any coordinate shift (northing or easting) for all projections.

Each of the 59 WCCS zones was to retain its map projection type, with the parameters modified to meet the above-described design constraints. Modifications included causing all map projection surfaces to be non-intersecting with the GRS 80 reference ellipsoid on NAD 83. Re-design of a given projection was accomplished in three general steps:

1. A map projection, secant or tangent to the GRS 80 reference ellipsoid, was constructed from the parameters of the WCCS projection on its enlarged ellipsoid. This involved use of functions published in Stem (1989) and Bomford (1985).
2. The scale factor parameter (k_o), computed or used in step 1, was multiplied by $1/E_p$, where E_p was computed using h_p in equation 10 and R_G at the mean latitude of the map projection. This caused the re-designed map projection surface to come into proximity to Earth's surface and be non-intersecting with the GRS 80 reference ellipsoid (see Figure 5b).
3. A least squares method was used to adjust the WCCS false northing and WCCS false easting of the coordinate origin, and the new k_o computed in step 2, so coordinates of points on the re-designed projection best fit coordinates of the same points on the WCCS projection. This was facilitated by generating a 0.5-mile grid of points across the county buffered by an additional two miles. At each point on the grid, coordinate differences between WCCS and the preliminary re-designed system formed the basis for linear equations with unknowns S (a multiplier for the new k_o), ΔN_o (to be added to the WCCS false northing), and ΔE_o (to be added to the WCCS false easting). The least squares solution to the collection of these equations produced a final design that met its required tolerances everywhere.

The re-designed system of LDPs was dubbed the "Wisconsin Coordinate Reference Systems (WISCRS)" by the WLIA Task Force. WISCRS was published by the State Cartographer's Office in 2009 and now supports a host of applications at many levels and by many users throughout Wisconsin.

Computer-Assisted Design

Map projection design is not common practice and is not done by a lot of people, so there is not a large commercial market for software to support it. However, with NGS now accepting LDPs in SPCS2022 (if the LDPs meet requirements) and with all SPCS2022 designs being done at Earth's surface, interest in and demand for computer-assisted map projection design has increased such that, as of this writing, at least one company (Geedop, LLC) is offering a web-based

GIS application for map projection design. Their product is named “LDP Design” and can be used to design Lambert conformal conic, transverse Mercator, and oblique Mercator map projections anywhere in the United States (see https://ldp.geedop.com/).

LDP Design is supported by a nationwide dense digital elevation model (DEM), supplemented with GEOID12B, so that linear distortions can be computed anywhere on any of the three mentioned map projection types. Linear distortions are color-coded and presented as “heat maps” for user visualization (see Figure 7). A user specifies or selects geographic extents and then interacts with LDP design, using trials, to find the map projection that suits their needs in terms of linear distortion and its statistical and spatial distributions. Users have control over map projection type and critical map projection parameters and can manipulate them as they please, then visualize the results.

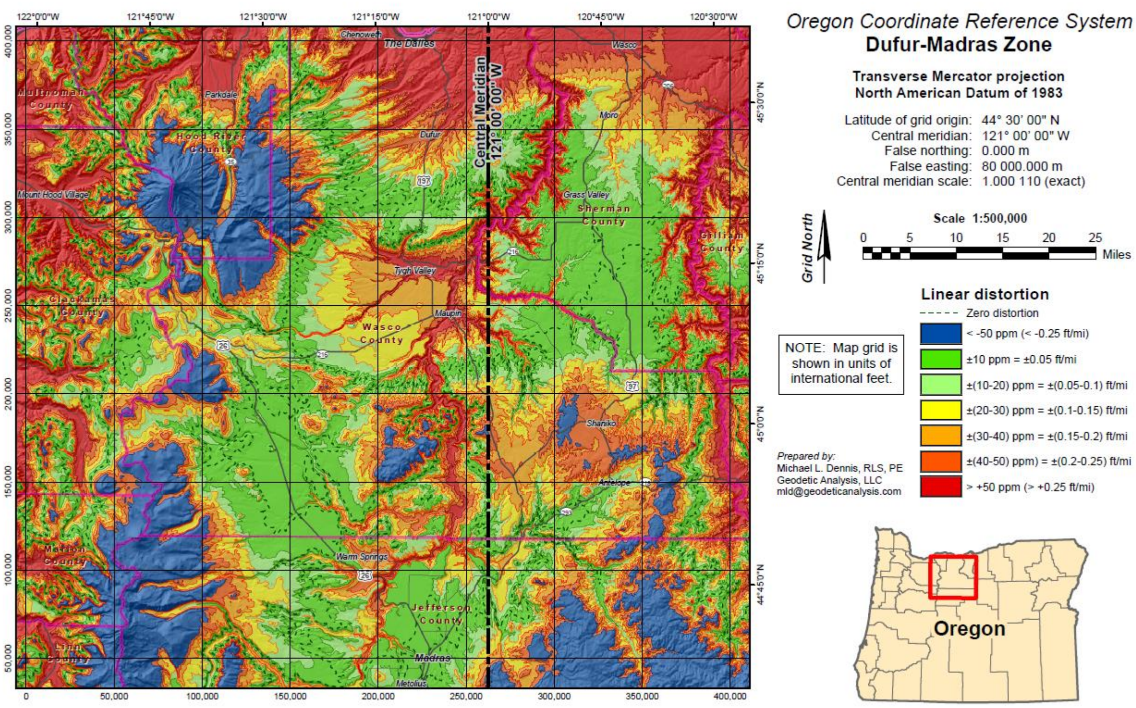


Figure 7.
Heat Map of Linear Distortions for a Transverse Mercator Projection Developed for Oregon DOT
(ftp://ftp.odot.state.or.us/ORGN/39OCRS_Zone_maps_tabloid/)

Fully Automated Design

It is possible to compute optimal values for critical map projection parameters without user interaction. All the user needs to supply is a set of spot elevations, perhaps in the form of a DEM. If the horizontal coordinates for the spot elevations are already on a map projection, the user must also supply the parameters of that projection. Optimization software can be made to include extraction of geoid heights from GEOID18. These are then added to the spot elevations to obtain ellipsoid heights.

The approach can be described by assuming that a perfect, but impossible, LDP would have no linear distortion anywhere. This condition can be expressed with an equation:

$$1 = k_i E_i \qquad 11$$

where the subscript, "i", refers to point i. In equation 11, E_i is the ellipsoid height factor at point i and is a function of a, b, ϕ_i, H_i, and N_i. However, k_i (the scale factor at point i) is a function of a number of variables including the critical map projection parameters (i.e., k_o and ϕ_o for Lambert conformal conic projections and k_o and λ_o for transverse Mercator projections). We can write a different equation 11 for each data point. Therefore, if we have enough data points, we can solve a collection of equations 11 for the critical parameters. k_i is non-linear in terms of the critical parameters, so equation 11 must be linearized for ease of solution:

$$1 - k_i^a E_i = E_i \left(\frac{\partial k_i}{\partial k_o}\right)^a \Delta k_o + E_i \left(\frac{\partial k_i}{\partial \phi_o}\right)^a \Delta \phi_o \qquad 12$$

for Lambert conformal conic projections and

$$1 - k_i^a E_i = E_i \left(\frac{\partial k_i}{\partial k_o}\right)^a \Delta k_o + E_i \left(\frac{\partial k_i}{\partial \lambda_o}\right)^a \Delta \lambda_o \qquad 13$$

for transverse Mercator projections. In equations 12 and 13, the superscript "a" means "evaluated at approximations for the critical parameters". Δk_o, $\Delta\phi_o$, and $\Delta\lambda_o$ are corrections to the approximations. See the appendix for expressions of the partial derivatives.

If there are more than two data points, the conditions expressed in equations 12 and 13 cannot be fully enforced because the unknowns will be overdetermined and there will be more than one solution for them. This means there will be residual linear distortions at the data points and the equations must be modified to accommodate them. For a Lambert conformal conic projection, equation 12 is modified to:

$$v_i + 1 - k_i^a E_i = E_i \left(\frac{\partial k_i}{\partial k_o}\right)^a \Delta k_o + E_i \left(\frac{\partial k_i}{\partial \phi_o}\right)^a \Delta \phi_o \qquad 14$$

where v_i is the residual linear distortion at point i. There is one solution for a collection of more than two equations 14 that is considered optimal for many applications. The “least squares” solution finds final values for the critical map projection parameters that minimize the sum of the squares of the residuals:

$$\sum_{i=1}^{n} v_i^2 \rightarrow minimum \qquad 15$$

where n is the number of available data points. We can think of this as analogous to minimizing the separation between Earth’s surface and the map projection surface (see Figure 8).

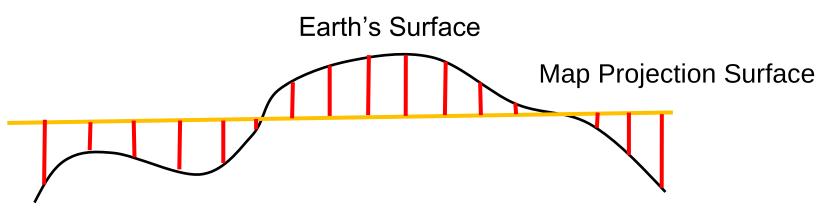


Figure 8.
Least Squares Analogy: Position the Map Projection Surface Such that the Sum of the Squares of Red Distances at Data Points is Minimized

The least squares solution manipulates the full set of n equations 14 to produce a set of two “normal” equations:

$$\left[\sum_{i=1}^{n} \left(E_i \left(\frac{\partial k_i}{\partial k_o}\right)^a\right)^2\right] \Delta k_o + \left[\sum_{i=1}^{n} \left(E_i \left(\frac{\partial k_i}{\partial \phi_o}\right)^a E_i \left(\frac{\partial k_i}{\partial k_o}\right)^a\right)\right] \Delta \phi_o = \sum_{i=1}^{n} \left[E_i \left(\frac{\partial k_i}{\partial k_o}\right)^a (1 - k_i^a E_i)\right]$$

$$\left[\sum_{i=1}^{n} \left(E_i \left(\frac{\partial k_i}{\partial \phi_o}\right)^a E_i \left(\frac{\partial k_i}{\partial k_o}\right)^a\right)\right] \Delta k_o + \left[\sum_{i=1}^{n} \left(E_i \left(\frac{\partial k_i}{\partial \phi_o}\right)^a\right)^2\right] \Delta \phi_o = \sum_{i=1}^{n} \left[E_i \left(\frac{\partial k_i}{\partial \phi_o}\right)^a (1 - k_i^a E_i)\right]$$

16

Similar sets of equations 14 and normal equations 16 can be developed for transverse Mercator map projections by replacing ϕ_o with λ_o everywhere in equations 14 and 16.

The normal equations 16 are solved for the corrections which are then added to the approximations for the unknown critical map projection parameters and the process is repeated until the corrections become negligible. The author uses cut-off thresholds of Δk_o < |0.0000000001|, $\Delta \phi_o$ < | 0.000001| sec, and $\Delta \lambda_o$ < | 0.000001| sec. For initial approximations, k_o^a = 1 and mean values from the data points for ϕ_o^a and λ_o^a have been sufficient for tested cases. In the final solution, the critical map projection parameters will be optimal, the sum of the squares of linear distortions at the data points will be minimized, and the sum of linear distortions at the data points will equal zero.

Subsets of data points can be weighted differently if, for example, a user desires elevations in major urban areas to have twice the effect as those in villages and those in villages to have twice the effect of those in rural areas. This can be done if the optimization software is capable of overlaying the collection of spot elevations with polygon files that outline the areas to be weighted differently.

LiDAR-derived spot elevations from Columbia and Waukesha Counties were available for testing the optimization algorithms. Each of the two data sets were on uniform grids at 100-foot spacing. The Columbia County data set included 2716086 spot elevations and the Waukesha County data set included 1616021 of them. In each case, the data set covered the entire county (see Figures 9 and 10). For all spot elevations, geoid heights were extracted from GEOID18 so ellipsoid heights and linear distortions could be computed.

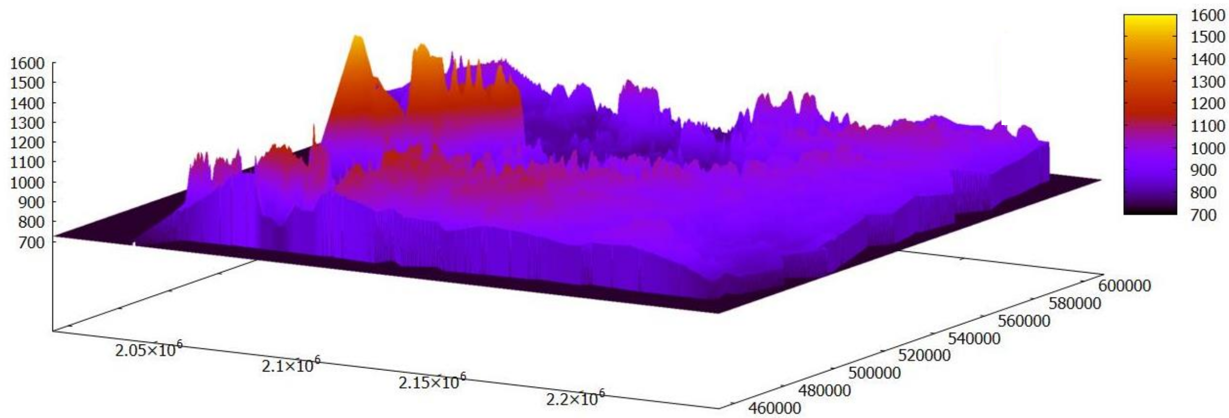


Figure 9.
Columbia County Elevations (U.S. Survey Feet)

In WISCRS, Columbia County's map projection is Lambert conformal conic, so an optimal projection of that type was computed and used for comparison to its WISCRS companion. The non-linear solution converged in three iterations

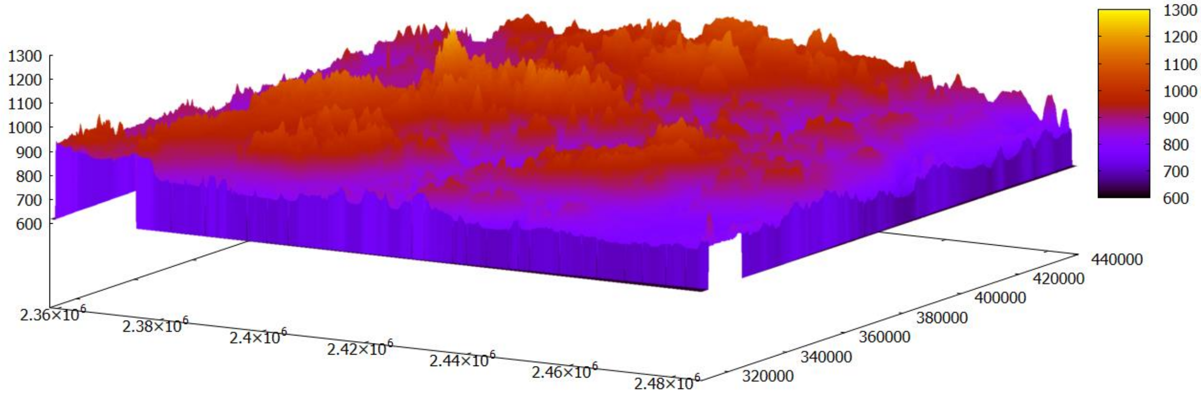


Figure 10.
Waukesha County Elevations (U.S. Survey Feet)

starting with initial approximations of k_o = 1 and ϕ_o = mean latitude of the spot elevations. Geo-statistics for the county appear in Table 2, critical map projection parameters appear in Table 3, and linear distortion comparisons appear in Table 4.

Table 2.
Some Statistics for Columbia County

Extents and Averages	
Maximum Latitude	43 39 50.56984
Minimum Latitude	43 14 40.43847
Maximum Longitude	89 49 34.02891
Minimum Longitude	88 56 40.09426
Maximum Geoid Height	-34.575 m
Minimum Geoid Height	-35.844 m
Maximum Elevation	588.200 m
Minimum Elevation	220.900 m
Maximum Ellipsoid Height	552.516 m
Minimum Ellipsoid Height	185.207 m
Mean Geoid Height	-35.288 m
Mean Elevation	274.628 m
Mean Ellipsoid Height	239.339 m

k_o on the central parallel differs by about 0.4 parts per million with the optimal being smaller than that of WISCRS. The latitude of the central parallel differs by 2' 25.64590" (about 2.7 miles) with the optimal being north of that of WISCRS.

For WISCRS, the extremes of linear distortion are actually smaller in absolute value than those for the optimization algorithm. However, the mean linear distortion for the optimization algorithm is zero as it must be while that of

Table 3.
Critical Map Projection Parameters for Columbia County

	WISCRS	Optimal
k_o on Central Parallel	1.0000349800	1.0000353980
Latitude of Central Parallel	43 27 45.16792	43 30 10.81382

Table 4.
Linear Distortion Statistics for Comparison

	WISCRS	Optimal
Max Combined Factor	1.000011916	1.000015181
Min Combined Factor	0.999954097	0.999952371
Max Distortion	1:83918 or 12 ppm	1:65871 or 15 ppm
Min Distortion	-1:21785 or -46 ppm	-1:20995 or -48 ppm
Mean Combined Factor	0.999999338	1.000000000
Mean Distortion	-1 ppm	0 ppm
Root-Mean-Square (RMS) Linear Distortion	0.00000469 or 1:213078 or 5 ppm	0.00000444 or 1:225244 or 4 ppm

WISCRS is not. Also, the root-mean-square (RMS) linear distortion is smaller for the optimization algorithm. It must be, because the value of RMS is minimized during optimization.

Figure 11 is a bar chart of linear distortions for both WISCRS and the optimization. The least squares solution draws the collection of linear distortions statistically closer to their mean. Figure 12 is a view of the optimal linear distortions in Columbia County.

For Columbia County, it must be noted that during design of WCCS, and reflected in WISCRS, map projection parameters were selected to emphasize urban areas (SCO (1995)). The optimization algorithm did not do this. Rather, all spot elevations were treated equally because no polygon files of urban areas were available for weighting.

In WISCRS, Waukesha County's map projection is transverse Mercator, so an optimal projection of that type was computed and used for comparison to its WISCRS companion. The non-linear solution converged in three iterations starting with initial approximations of $k_o = 1$ and λ_o = mean longitude of the spot elevations. Geo-statistics for Waukesha County appear in Table 5, critical map projection parameters appear in Table 6, and linear distortion comparisons appear in Table 7.

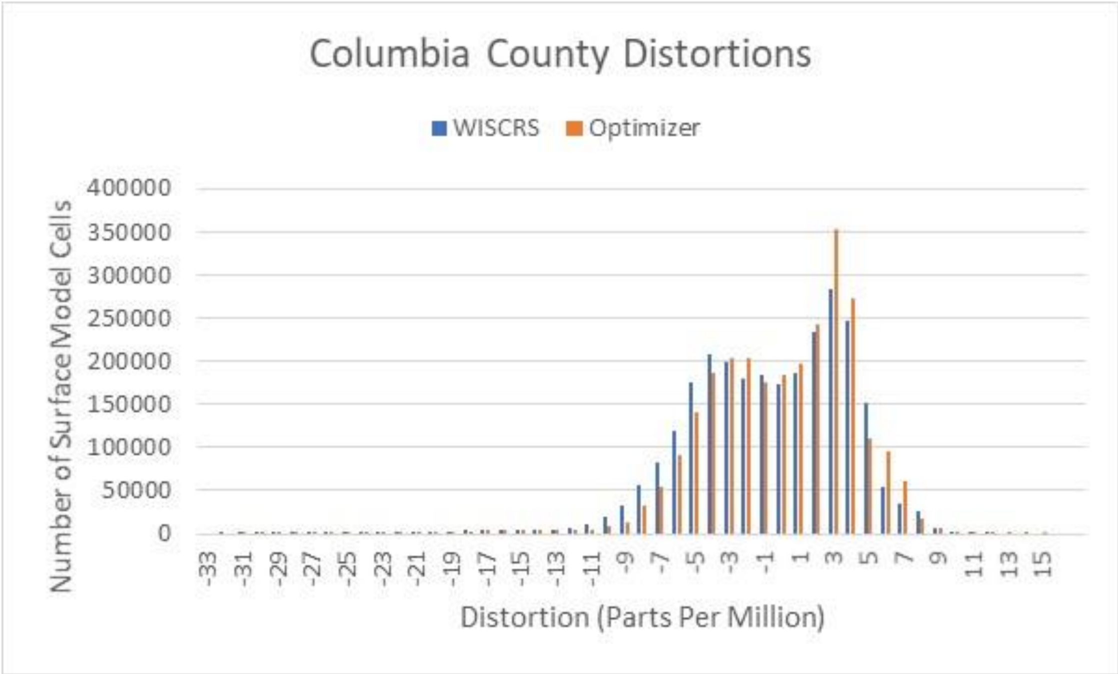


Figure 11.
Bar Chart of WISCRS and Optimal Linear Distortions in Columbia County

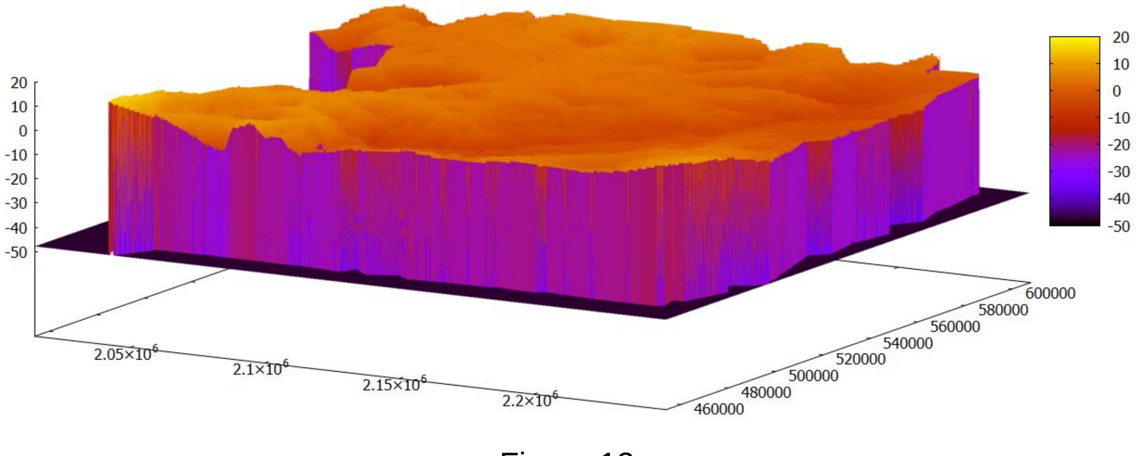


Figure 12.
Optimal Linear Distortions in Columbia County (Parts per Million)

Table 5.
Some Statistics for Waukesha County

Extents and Averages	
Maximum Latitude	43 11 45.65026
Minimum Latitude	42 50 31.70082
Maximum Longitude	88 32 31.86307
Minimum Longitude	88 03 49.53348
Maximum Geoid Height	-34.211 m
Minimum Geoid Height	-35.001 m
Maximum Elevation	373.655 m
Minimum Elevation	188.065 m
Maximum Ellipsoid Height	339.121 m
Minimum Ellipsoid Height	153.286 m
Mean Geoid Height	-34.649 m
Mean Elevation	270.737 m
Mean Ellipsoid Height	236.087 m

Table 6.
Critical Map Projection Parameters for Waukesha County

	WISCRS	Optimizer
k_o on Central Meridian	1.0000346179	1.0000352922
Longitude of Central Meridian	88 13 30.00000	88 15 14.73218

Table 7.
Linear Distortion Statistics for Comparison

	WISCRS	Optimizer
Max Combined Factor	1.000010598	1.000011414
Min Combined Factor	0.999983745	0.999983689
Max Distortion	1:94358 or 11 ppm	1:87613 or 11 ppm
Min Distortion	-1:61519 or -16 ppm	-1:61310 or -16 ppm
Mean Combined Factor	0.999999636	1.000000000
Mean Distortion	0 ppm	0 ppm
Root-Mean-Square (RMS) Linear Distortion	0.00000394 or 1:253518 or 4 ppm	0.00000387 or 1:258159 or 4 ppm

k_o on the central meridian differs by about 0.7 parts per million with the optimal being larger than that of WISCRS. The longitude of the central meridian differs by

1’ 44.73218” (about 1.3 miles) with the optimal being west of that of WISCRS. During design of WCCS, and reflected in WISCRS, the central meridian was shifted east to compensate for a general rise in elevation from east to west (SCO (1995)). Optimization pulled the central meridian slightly back west. Also, during WCCS design, urban areas and transportation corridors were given more emphasis than other areas. The optimization algorithm did not do this because no polygon files were available for weighting.

As in Columbia County, the extremes of linear distortion for WISCRS are smaller in absolute value than those for the optimization algorithm. However, the mean linear distortion for the optimization algorithm is once again zero and its RMS linear distortion is smaller than that of WISCRS. Figure 13 is a bar chart of linear distortions for both WISCRS and the optimization algorithm in Waukesha County. Figure 14 is a view of the optimal linear distortions.

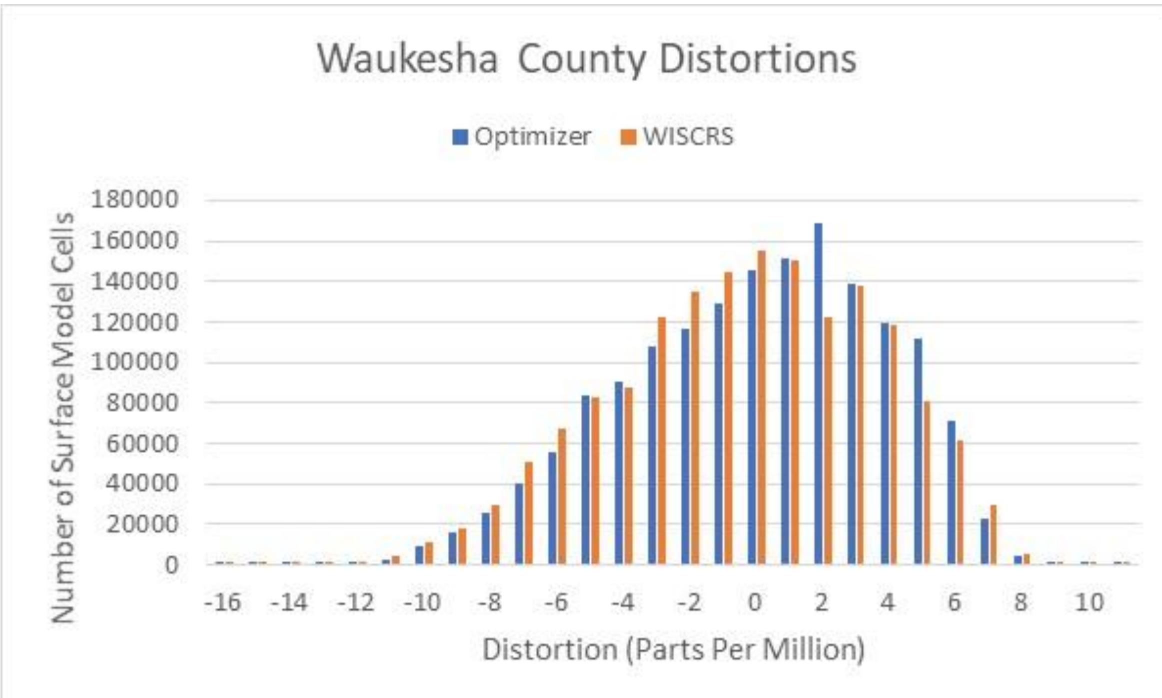


Figure 13.
Bar Chart of WISCRS and Optimal Linear Distortions in Waukesha County

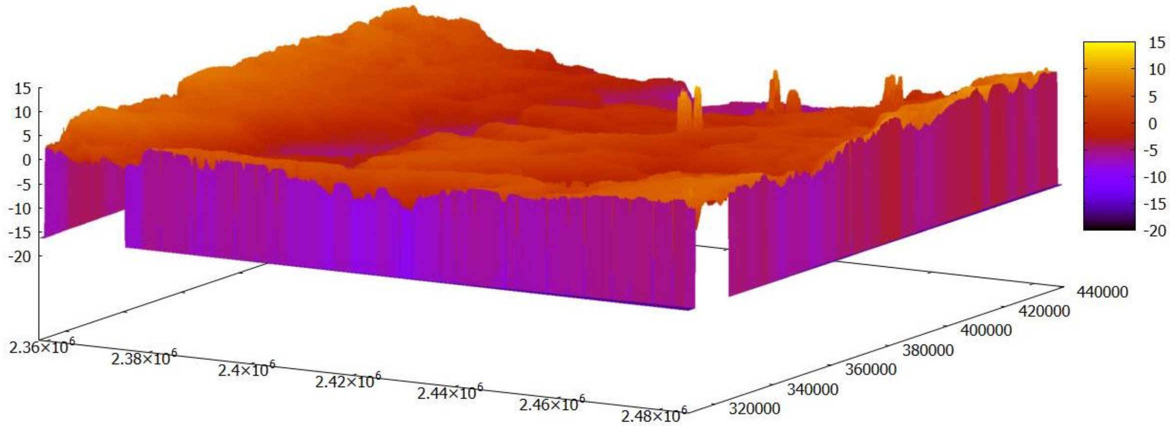


Figure 14.
Optimal Linear Distortions in Waukesha County (Parts per Million)

Summary

Because neither Earth's surface nor that of any reference ellipsoid is developable, all map projections contain distortions. The kinds of distortions and their magnitudes depend upon the type of map projection and its parameters. Selection or determination of map projection type and parameters, under user-determined constraints, constitutes a primary component of map projection design.

Conformal map projections tend to preserve shapes and produce scale factors that are independent of direction at any specific point. The geospatial community in the United States is most familiar with Lambert conic and transverse Mercator conformal map projections. These, along with oblique Mercator map projections have been institutionalized in the State Plane Coordinate System, the Universal Transverse Mercator system, and a few states' LDPs.

SPCS2022 is bringing about dramatic changes and spurring renewed interest in map projection design, particularly with regard for designing at Earth's surface. When doing so, design tolerances are based upon the combined factor, which is the product of the ellipsoid height factor and the scale factor. The combined factor can be easily converted into an expression of linear distortion which is, at any given point, the ratio of an infinitesimal distance on the map projection surface to the corresponding infinitesimal distance on Earth's surface. Low-distortion projections are such that combined factors are nearly equal to 1 and, thus, ground distances can be used for grid distances in nearly all practical applications.

Wisconsin has had two sets of LDPs: 1) the Wisconsin County Coordinate System which is based upon reference ellipsoids enlarged to mean terrain over each of its zones and 2) the Wisconsin Coordinate Reference Systems, which has the same zone boundaries, and nearly the same coordinates as WCCS, but uses controlling scale factors on the central parallels or central meridians that bring the map projection surfaces into proximity to mean terrain without changing the reference ellipsoid.

When designing conformal map projections, there are two critical parameters that define the size of the map projection surface and its orientation with respect to the reference ellipsoid. Determination of values for these parameters has typically involved combinations of hardcopy and digital techniques. A recent innovation is interactive web-based GIS applications that use underlying DEMs and geoid models, combined with programmed map projection functions, to present visualizations, or heat maps, of linear distortions to users. Trials, with various combinations of projection types and critical parameters, are then used to find design solutions that meet user needs.

An alternative, fully-automatic approach to LDP design is presented in this paper. A least squares, or optimal, solution for the critical map projection parameters (Lambert conformal conic or transverse Mercator) can be computed from nothing more than a collection of spot elevations. The solution minimizes the sum of the squares of linear distortions at the data points. Examples were presented for two counties in Wisconsin and compared with results from WISCRS.

Disclaimer

In this document, mention of any company, its products, or its services does not constitute endorsement by the author.

List of References

Bomford, G., (1985), Geodesy, 4th Edition, Oxford University Press, New York, 855 pp.

Bugayevskiy, L.M. and J.P. Snyder, (1995), Map Projections – A Reference Manual, Taylor & Francis, Bristol, PA.

Dennis, M.L., (2016), "Ground Truth: Low Distortion Map Projections for Engineering, Surveying, and GIS", Proceedings, ASCE Pipelines Conference. Available at https://www.researchgate.net/publication/305335178_Ground_Truth_Low_Distortion_Map_Projections_for_Engineering_Surveying_and_GIS

Johnson, G.W., (1985), "A Minnesota County Coordinate System," presented at the 33rd Annual Meeting of the Minnesota Land Surveyors Association, February.

Moffitt, F.H. and J.D. Bossler, (1998), Surveying, 10th Edition, Addison Wesley Longman, New York, NY.

National Geodetic Survey, (2017a), Blueprint for 2022, Part 1: Geometric Coordinates, NOAA Technical Report NOS NGS 62.

National Geodetic Survey, (2017b), Blueprint for 2022, Part 2: Geopotential Coordinates, NOAA Technical Report NOS NGS 64.

National Geodetic Survey, (2019a), State Plane Coordinate System of 2022 Policy, Policy Document, Tracking Number NGS 2019-1214-01.

National Geodetic Survey, (2019b), Procedures for Design and Modification of the State Plane Coordinate System of 2022, Procedures Document, Tracking Number 2019-1214-01-A1.

Snyder, J.P., (1987), Map Projections – A Working Manual, U.S. Geological Survey Professional Paper 1395, United States Printing Office, Washington, D.C.

Stem, J., (1989), State Plane Coordinate Systems of 1983, NOAA Manual NOS NGS 5, Rockville, MD.

Vonderohe, A.P., (2006, 2019), A Methodology for Re-Design of the Wisconsin County Coordinate Systems, Report Submitted to the Jackson County, Wisconsin Land Information Office.

Vonderohe, A.P., (2006), WISCRS (Wisconsin Coordinate Reference Systems): Redesign of the WCCS (Wisconsin County Coordinate System), Final Report.

Wisconsin State Cartographer's Office, (1995), Wisconsin Coordinate Systems, Madison, WI.

Wisconsin State Cartographer's Office, (2009), Wisconsin Coordinate Reference Systems, 2nd Edition, Madison, WI.

WisDOT, (1993), Wisconsin County Coordinate Systems, Report to the Wisconsin Department of Transportation, Technical Services, Phase I Project 0656-24-00, DOT FOS Object Code 5501.

APPENDIX
Partial Derivatives in Equations 12 and 13

In this appendix:

a = semi-major axis of the reference ellipsoid
b = semi-minor axis of the reference ellipsoid
e = first eccentricity of the reference ellipsoid
e' = second eccentricity of the reference ellipsoid
ν_o = radius in the prime vertical at ϕ_o
ρ_o = radius in the meridian at ϕ_o

$$e^2 = \frac{a^2 - b^2}{a^2}$$

$$e'^2 = \frac{e^2}{1 - e^2}$$

$$\nu_o = \frac{a}{\sqrt{1 - e^2 sin^2 \phi_o}}$$

$$\rho_o = \frac{a(1 - e^2)}{(1 - e^2 sin^2 \phi_o)^{\frac{3}{2}}}$$

For Lambert conformal conic map projections, the partial derivatives in equation 12 are:

$$\frac{\partial k_i}{\partial k_o} = 1 + J_{1i} + J_{2i} + J_{3i}$$

where

$$J_{1i} = \frac{M_i^2}{2\rho_o \nu_o}$$

$$J_{2i} = \frac{M_i^3 tan\phi_o}{6\rho_o \nu_o^2}$$

$$J_{3i} = \frac{M_i^4 (5 + 3tan^2 \phi_o)}{24\rho_o^2 \nu_o^2}$$

$$M_i = a(A_o \phi_i - A_2 sin2\phi_i + A_4 sin4\phi_i - A_6 sin6\phi_i) - M_o$$

$$M_o = a(A_o \phi_o - A_2 sin2\phi_o + A_4 sin4\phi_o - A_6 sin6\phi_o)$$

$$A_o = 1 - \frac{e^2}{4} - \frac{3e^4}{64} - \frac{5e^6}{256}$$

$$A_2 = \frac{3}{8}\left(e^2 + \frac{e^4}{4} - \frac{15e^6}{128}\right)$$

$$A_4 = \frac{15}{256}\left(e^4 + \frac{3e^6}{4}\right)$$

$$A_6 = \frac{35e^6}{3072}$$

and

$$\frac{\partial k_i}{\partial \phi_o} = k_o\left(\frac{\partial J_{1i}}{\partial \phi_o} + \frac{\partial J_{2i}}{\partial \phi_o} + \frac{\partial J_{3i}}{\partial \phi_o}\right)$$

where

$$\frac{\partial J_{1i}}{\partial \phi_o} = \frac{M_i\left(\frac{\partial M_i}{\partial \phi_o}\right)}{\rho_o \nu_o} - \frac{M_i^2\left[\rho_o\left(\frac{\partial \nu_o}{\partial \phi_o}\right) + \nu_o\left(\frac{\partial \rho_o}{\partial \phi_o}\right)\right]}{4\rho_o^2 \nu_o^2}$$

$$\frac{\partial J_{2i}}{\partial \phi_o} = \frac{tan\phi_o\left[M_i\left(\frac{\partial J_{1i}}{\partial \phi_o}\right) + J_{1i}\left(\frac{\partial M_i}{\partial \phi_o}\right)\right] + J_{1i}M_i\left(\frac{1}{cos^2\phi_o} - \frac{tan\phi_o}{\nu_o}\right)}{3\nu_o}$$

$$\frac{\partial J_{3i}}{\partial \phi_o} = \frac{J_{1i}\left(\frac{\partial J_{1i}}{\partial \phi_o}\right)[5 + 3tan^2\phi_o]}{3} + \frac{J_{1i}^2 sin\phi_o}{cos^3\phi_o}$$

$$\frac{\partial M_i}{\partial \phi_o} = -a(A_o - 2A_2cos2\phi_o + 4A_4cos4\phi_o - 6A_6cos6\phi_o)$$

$$\frac{\partial \rho_o}{\partial \phi_o} = \frac{3\left(\frac{\partial \nu_o}{\partial \phi_o}\right)}{1 - e^2sin^2\phi_o}$$

$$\frac{\partial \nu_o}{\partial \phi_o} = \frac{ae^2sin\phi_o cos\phi_o}{\left(1 - e^2sin^2\phi_o\right)^{\frac{3}{2}}}$$

For transverse Mercator map projections, the partial derivatives in equation 13 are:

$$\frac{\partial k_i}{\partial k_o} = 1 + F_{2i}L_i^2(1 + F_{4i}L_i^2)$$

and

$$\frac{\partial k_i}{\partial \lambda_o} = -2k_oF_{2i}L_icos\phi_i(1 + 2F_{4i}L_i^2)$$

where

$$L_i = (\lambda_i - \lambda_o)cos\phi_i$$

$$F_{2i} = \frac{1 + \eta_i^2}{2}$$

$$F_{4i} = \frac{5 - 4t_i^2 + \eta_i^2(9 - 24t_i^2)}{12}$$

$$\eta_i^2 = e'^2cos^2\phi_i$$

$$t_i^2 = tan^2\phi_i$$

and longitudes are positive west.