Journal:
Cartography and Geographic Information Science

Title: **The development and analysis of quasi-linear map projections**

Author: **Jonathan Charles Iliffe**

Affiliation: **University College London (UCL).**

Department of Civil, Environmental & Geomatic Engineering,
University College London,
Gower Street,
London WC1E 6BT,
UK

E-mail: j.iliffe@ucl.ac.uk
Phone : +44 20 7679 2733

ORCID iD: orcid.org/0000-0001-7536-2952

Biographical note

Jonathan Iliffe is a senior lecturer in geodesy at University College London, with a specialist interest in the datums, map projections and coordinate systems used in land and marine applications.

Abstract

Map projections are an essential component of coordinate systems used in applications such as surveying, topographic mapping and engineering, and care needs to be taken to select ones that minimize distortion for each case. This paper explores the selection process for near-linear features on the surface of the Earth and derives limits for the extent of a project that can be projected within specified distortion tolerances. It is then demonstrated that a multi-faceted set of projections of the Earth may be used to extend this concept to the mapping of features such as highways and railways that are quasi-linear but do not exactly follow a standard geometrical line (a great circle or a small circle) on the surface of the Earth. A continuous, conformal coordinate system may be derived in such situations, extending to indefinite length and applicable over a swath of several kilometers width, but it cannot be extended to cover situations with extensive variations in height. Instead, the Snake Projection is analyzed, and it is shown that this can be used to develop continuous (non-zonal) projected coordinate systems for major engineering projects extending for hundreds of kilometers and having extensive height ranges. Examples are shown of the application of this to railway projects.

Keywords

Map projection. *Scale factor*. *Snake Projection*.

Word count: 8,058 excluding title page information, but including abstract, captions, tables, place indicators, etc.

Introduction

A map projection is a systematic representation of all or part of a body such as the Earth on a plane (Snyder, 1987). It is largely thought of in two-dimensional terms: a curved 2D surface is transformed to a flat 2D surface, with a resulting scale factor, k, that increases with the size of the area being projected. The maximum value of scale factor distortion that can be tolerated will vary with the function or purpose of the map or coordinate system that is being created. General route finding can accommodate distortions of a few percent, while surveying and topographic mapping would generally tolerate a distortion of no more than about 0.1%. Still more exacting, in engineering applications the requirement is often to avoid the need to make any kind of correction to observed data, and so the maximum scale factor distortion that can be tolerated is commensurate with the accuracy of the original data and the control surveys, or around 10 – 20 parts per million (ppm) (Uren and Price, 2010).

It is hence standard practice for cartographers, surveyors, or engineers to select an appropriate projection that minimizes the distortion for their particular project, taking into account location and directional extent (Šavrič et al, 2016; Orihuela, 2016; Strebe, 2016). Several studies have developed expert systems to facilitate this selection (Zhao et al, 2007; Šavrič et al, 2016). Where the areal extent of the project or application is too large to be accommodated on one projection without excessive distortion, it becomes necessary to split it up into smaller zones and project each one separately. Examples include those of the State Plane Coordinate System of the USA (Stem, 1990) or the seven Gauss-Krüger *fajas* of the Argentine coordinate reference system, each a three degree wide band of longitude (Argerich, 2006). Whilst each of the zones in these examples tend to be a few hundred kilometers in extent, for the more exacting engineering applications the maximum dimension of the zone is much smaller. Depending on the projection used, the maximum dimension for a 10 ppm distortion would be up to 25 – 50 km. These tend to be adopted on an ad hoc basis for particular projects, but an example of a published zonal system for engineering applications is that of the Highways Agency of England, which has divided England into 55 zones, with widths between 9 km and 78 km, such that the maximum distortion is less than 50 ppm (Highways Agency, 2007). Clearly, the zonal approach provides a means of minimizing the scale factor distortion, but it does this at the expense of continuity, leading to difficulties in computations that cross zone boundaries.

An alternative approach to controlling the distortion is to avoid the use of projections altogether, and carry out computations in a geocentric three-dimensional coordinate system. This is the approach adopted when determining coordinates (or relative coordinates) with global navigation satellite systems (GNSS) in which the XYZ coordinates no longer have any relationship to the concepts of horizontal and vertical. However, while this is appropriate for individual points of high precision, for most practical applications (and for the software designed to support them) it is essential to maintain the relationship with gravity, and be able to project mapping and orthophotography onto the horizontal plane, to be able to process and compute survey data, design horizontal and vertical curves, and so on. Ideally, an optimum projection needs to be designed, applicable to a particular project and to a band of several kilometers on either side to accommodate associated works and eliminate edge effects.

In most cases, choosing the optimum projection means identifying the principal axis of a project and aligning this with a particular linear feature of a map projection. Examples include the north-south axis of the Transverse Mercator projection, or the east-west axis of the Lambert Conformal Conic, which can be exploited to map regions that are so aligned. However, the question that this paper addresses is whether this approach can be extended to derive projected coordinate systems for engineering applications that are quasi-linear, such as transport links (highways or rail routes) or pipelines. Clearly they can be if the project happens to follow closely a meridian, a parallel, or an

arbitrary great circle, but this paper will be concerned with the more likely general case, in which a railway or a highway meanders to some extent in its route from A to B, and we shall explore systems where the line of constant or true scale may be made to follow a sinuous path on the surface of the Earth. In this way the coordinate system adopted can in principle be almost unlimited in extent, provided that it could be encompassed within a specified corridor.

At the same time, the *three dimensional* nature of the problem in hand needs to be considered. The assumption is usually made that the surface to be projected is the ellipsoidal or spherical figure of the Earth, but in fact the experience of the user is on or near the *actual* surface. This is separated from the ellipsoid by potentially hundreds, even thousands, of meters, and so the additional scaling effect is significant for topographic or engineering applications. Some engineering applications will get round this by introducing an overall re-scaling into the coordinate system to account for the height effect, but for a project of significant extent the problem will be that the height changes – as a railway climbs over mountain passes, for example. Therefore, alongside an exploration of how a projected coordinate system may be derived with respect to a sinuous *horizontal* line, we shall explore how this line of true scale may be made to vary in its *vertical* trajectory.

This paper will proceed by analyzing the linear distortion characteristics of some of the standard projections, and looking at previous studies that have extended the linear approach beyond topographic applications. It will then seek a more general methodology for establishing quasi-linear coordinate systems, by analyzing the characteristics of alternative approaches such as the use of multiply interrupted projections and the Snake Projection. The geometric characteristics and parametric descriptions of these will be explored in the context of their application to real-world examples.

Linear features of projected coordinate systems

The concept of scale factor, k, is generally defined as the ratio of an (infinitesimally small) linear feature on the projection to its corresponding distance on the surface of the ellipsoid or sphere, or:

$$\mathrm{k} = \frac{\mathrm{D_{PROJ}}}{\mathrm{D_{ELLIPSOID}}} \qquad (1)$$

This scale factor may vary according to direction. Tissot's Indicatrix (Snyder 1987, Bildirici, 2015) describes the projection of an infinitesimally small circle on the surface of the sphere into an ellipse with semi-axes *a* and *b* on the projection. The angular distortion, w, is derived from the indicatrix via the expression:

$$\sin(\mathrm{w}/2) = \frac{|\mathrm{a}-\mathrm{b}|}{(\mathrm{a}+\mathrm{b})} \qquad (2)$$

For topographic mapping, surveying, and engineering applications the property of undistorted angles is essential for computational purposes, and so the primary focus here will be on conformal projections in which a=b (as closely as possible) and w is minimized.

Within the conformal category, the Mercator projections all have their linear axis along a great circle (for the spherical model of the Earth). The original projection is aligned along the equator, the transverse form along a meridian, and the oblique form along a generalized great circle (Iliffe and Lott, 2008). Introducing an overall re-scaling factor k_0 the general relationship of the scale factor with respect to the central axis is:

$$\mathrm{k} = \mathrm{k}_0 \sec\theta \qquad (3)$$

Where θ is the angular distance from the central axis, taking the form of the latitude for the original Mercator projection.

To examine the context of this scale factor relationship, let us introduce the symbol Δ_M to represent the maximum scale factor distortion that can be tolerated by a particular application: for example, 10 – 20 ppm for engineering applications, 0.1% for topographic work, and so on. Then the maximum scale factor that can be tolerated is denoted by k_M. Thus:

$$\mathrm{k_M} = 1 + \Delta_\mathrm{M} = \mathrm{k}_0 \sec\theta_\mathrm{M} \qquad (4)$$

The selection of k_0 will generally be made to minimize to some extent the maximum scale factor distortion. For relatively small values of Δ_M a value for k_0 of $(1 - \Delta_M)$ will distribute the distortion evenly around unity. Although not always universally adopted in practice, if we make the assumption that this would be adopted under the strict condition of minimization of scale factor distortion, we can substitute into (4) to derive the maximum angular distance from the line of contact, θ_M, for the given maximum distortion:

$$\theta_\mathrm{M} = \cos^{-1}\left(\frac{1-\Delta_\mathrm{M}}{1+\Delta_\mathrm{M}}\right) \qquad (5)$$

For the topographic condition of 0.1% for Δ_M this results in a value of 3.6° for θ_M, or a band twice this width centred on the line of contact. For engineering applications, 10 ppm corresponds to ±0.36°, or a band of ±40 km on either side of the line of contact.

These figures define the distortion with respect to the central axis of the projection; along the axis the projection may be used for an indefinite length. An important development of this concept was the introduction of the Space Oblique Mercator projection (Snyder, 1978; Ren et al, 2010) in which the central axis is the ground track of an orbiting satellite and thus there is continuity in the projection used for successive satellite images.

If the above set of projections cover projects that are broadly aligned along great circles, then the first order of deviation from this will be projects that are aligned with respect to a slightly curved line, and therefore along an axis based on a small circle of the sphere. In general these may be developed using conic projections: the regular case is known as the Lambert Conformal Conic (Snyder, 1987) in which case the small circle is a parallel of latitude, but the general form also exists and is known as Krovak's Oblique Conic (Skokanová and Havlíček, 2010).

Using the standard Lambert form of the projection, but understanding that the results can be generalized to the oblique Krovak case, the applicability of the projection has a slight dependence on the value of the standard parallel, but for values that range from close to the equator up to 70° (north or south) the topographic condition of 0.1% for Δ_M results in a band between 3.5° and 3.7° on either side of the standard parallel. For engineering applications, 10 ppm again corresponds to ±0.36° (as with the cylindrical family), but this time with respect to a small circle rather than to a great circle.

Again, an extension of this concept beyond the standard applications is provided by Ren et al (2013), who developed a Space Oblique Conic form for use with imagery derived from side-looking synthetic aperture radar.

In conclusion to this section, it can be seen that there are several options that can be drawn from the standard set of projections to derive coordinate systems for projects aligned closely to a great circle or a small circle. For any more complex pattern, we need to explore suitable alternatives.

Interrupted Projections

It has been demonstrated so far that scale factor increases as the area to be projected increases, although this growth may be controlled along one selected axis. It has also been established that an alternative method to limit the increase of scale factor is through the use of projection *zones* of limited size. An analogous technique is the use of the interrupted projection: this is certainly well known in the context of global applications (Maling, 1992), but its use in topographic or engineering applications is less well documented. In essence, the interrupted projection limits the amount of "stretch" by introducing a "cut" and doing this in a location that is generally not relevant to the conceived use of the map.

Most examples of interrupted projections tend to be equal area projections, with few ready examples of conformal ones (Šavrič et al, 2015). However, the concept of a *myriahedral projection* was introduced by van Wijk (2008) in which the development of scale factor distortion is controlled by making multiple interruptions, essentially joining together a sequence of individual projections in such a way that the scale factor is always sufficiently close to unity for the projection to be considered both conformal and equal area. In summary, the method proceeds by dividing the Earth into a polyhedral mesh, and separately projecting each face. Each side of each polygonal face is then designated as either a cut or a fold, with the former being the basis of the interruption in the projection and the latter forming the join between two individual projections. As demonstrated by van Wijk (2008), the Platonic solids form an obvious basis for the projection onto separate faces of a polyhedron but in fact *any* mesh covering the Earth can be used, provided it folds out without overlaps.

In the examples of standard projections covered in the previous section, it was seen that each possesses a particular "linearity" that could be exploited to optimize the projection for a particular purpose. In the case of the myriahedral projection, we potentially have the chance to optimize for the case where a project does not follow a convenient standard linearity (a great or a small circle). This paper will not repeat the derivation and analysis presented by van Wijk (2008), but instead will focus on analyzing whether the distortion of such a projection can be controlled to an extent sufficient to allow such projections to be used in engineering applications.

Consider the case of the sample project shown in Figure 1, representing a quasi-linear feature such as a railway. Exaggerated for clarity, this shows a project that stretches across a portion of the Earth's surface and we require an appropriate projection. Note also that this example contains a branch from the main linear route.

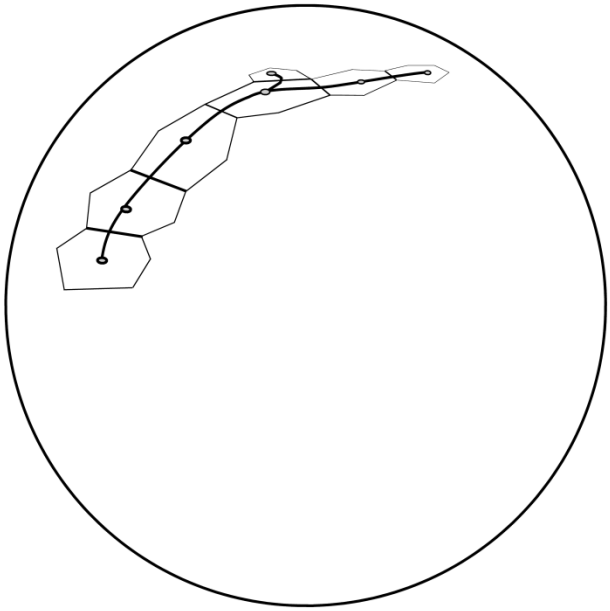


Figure 1: Quasi-linear project (with sub-branch) on the surface of the sphere.

Let this linear feature be defined in space by a set of *seed points*, p_i(defined by their geographic coordinates). Each of these points sits in its own face of the polyhedron; let the folded edge between faces occupied by $p_i$and p_j be designated f_{ij}. The non-fold sides of each polygon have been cut; we are not at this stage interested in where or how, and so the polygons are shown with arbitrary shapes on the non-fold sides.

For the fold sides, it is required to have the same scale factor on either side and the fold must project as a straight line whatever its orientation. This can be satisfied by defining the fold as a great circle of the sphere that forms the perpendicular bisector of p_i to p_j, and projecting each face as a gnomonic projection (Figure 2). There is no specific requirement for the seed points to be equally spaced, but a regular pattern makes it easier to control scale factor distortion.

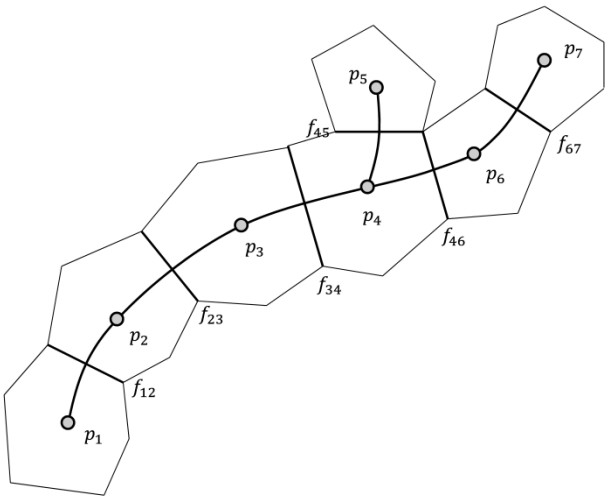


Figure 2: Projected fold-out of the quasi-linear project of Figure 1

The components of scale factor of the gnomonic projection can be labelled k_r and k_c for the radial and circumferential directions respectively (radiating from the seed point). These are given by:

$$\mathrm{k_r} = \mathrm{k_0}\sec^2\psi \tag{6}$$

$$\mathrm{k_c} = \mathrm{k_0}\sec\psi \tag{7}$$

Where ψ is the angular distance from the seed point at the centre of the sub-projection. Substituting these terms into (2) gives the angular distortion:

$$\sin\left(\frac{\mathrm{w}}{2}\right) = \frac{|\sec^2\psi - \sec\psi|}{(\sec^2\psi + \sec\psi)} = \frac{|\sec\psi - 1|}{(\sec\psi + 1)} = \frac{|1 - \cos\psi|}{(1 + \cos\psi)} \tag{8}$$

To use the gnomonic projection computationally and for coordinate systems, the overall scale factor distortion and the angular distortion must both be less than the defined accuracy limit. For engineering applications, adopting a target value of 10 ppm for Δ_M on the line p_i to $p_j$allows for values up to 20 ppm away from this. Again using a value for $\mathrm{k_0}$ of $(1 - \Delta_M)$, the critical scale factor is in the radial direction, from which we derive:

$$1 + \Delta_\mathrm{M} = (1 - \Delta_\mathrm{M})\sec^2\psi_\mathrm{M} \tag{9}$$

$$\psi_\mathrm{M} = \cos^{-1}\left\{\left(\frac{1 - \Delta_\mathrm{M}}{1 + \Delta_\mathrm{M}}\right)^{1/2}\right\} \tag{10}$$

From (10), 10 ppm for Δ_M gives 0.256° for ψ_M, the maximum permitted distance from the seed point to the edge of the polygonal cell. This corresponds to around 28 km on the surface of the Earth, and so the polygonal cells can be a maximum size of twice this, or 56 km.

To maintain an approximation to conformality, a maximum value also needs to be set for the angular distortion, w. The direct equivalent of a scale factor distortion of 10 ppm is $(10/10^6)$ in radians, or 2.1”. Substituting 0.256° into (8) shows that the maximum distortion angle occurs at the same value of ψ as the maximum scale factor distortion.

In principle therefore we can envisage a coordinate system for a quasi-linear feature that comprises a set of polygonal cells, each projected onto a gnomonic projection, and 10 ppm maximum distortion on the linear axis where the cell sizes are no more than 56 km across. Away from the linear axis the distortion would grow; 20 ppm would be reached at 0.314° from the central seed point, which at the edge of the cell would be 0.182°, or approximately 20 km, from the linear axis. Thus, the projected coordinate system would be applicable to a swath of 20 km on either side of the actual route, giving an adequate buffer for the mapping of associated infrastructure. However, there is a further consideration: away from the linear axis, although features may be conformal within the required tolerance, on the cell boundary they will be rotated with respect to the corresponding feature on the other side. So a straight line on the surface of the Earth will project with a kink in it across the cell boundary. This also needs to be controlled.

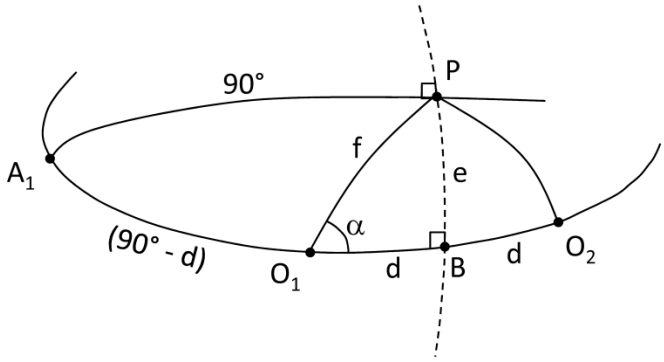


Figure 3: Spherical diagram of two adjacent cells for sub-projections, centered on O1 and O2.

In Figure 3, let O_1 and O_2 be the centers of two adjacent sub-projections, with the great circle perpendicular bisector forming the cell boundary. Both O_1 and O_2 are an angular distance *d* from this boundary. At a general point P, a distance *e* from the linear axis, a representative straight line (a great circle route) cuts the cell boundary at 90°: we shall test the straightness of this line when it is projected on either side of the cell boundary.

The point P is situated at distance f from O_1, at an angle α from the axis O_1O_2. All great circles perpendicular to the cell boundary will converge at a point A_1 at 90° from it (and at a corresponding point A_2 on the other side). The angular distance f is a function of *d* and α, the relationship being derived from the spherical triangle O_1BP and the application of Napier's rules for right angled spherical triangles (Van Brummelen, 2013) to give:

$$\tan(\mathrm{f}) = \frac{\tan(\mathrm{d})}{\cos\alpha} \tag{11}$$

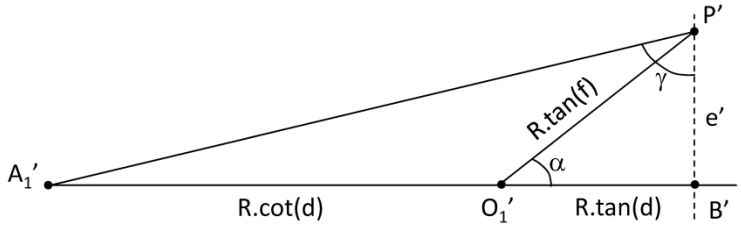


Figure 4: Gnomonic projection of cell centered on O1.

Figure 4 shows these points projected in the gnomonic sub-projection of O_1. A generic radial distance *c* is projected at a distance of $Rtan(c)$ from the central point of any gnomonic projection (Snyder, 1987), where R is the radius of the spherical model of the Earth. Distances radial from the central point O_1 are labelled accordingly (note that O_1A_1 is the complement of *d*). The distance e is projected as e', and is computed from the cosine rule applied to the plane figure O_1'B'P' thus:

$$\mathrm{e'}^2 = \big(\mathrm{R}\tan(\mathrm{f})\big)^2 + \big(\mathrm{R}\tan(\mathrm{d})\big)^2 - 2\big(\mathrm{R}\tan(\mathrm{f})\big)\big(\mathrm{R}\tan(\mathrm{d})\big) \tag{12}$$

$$= \mathrm{R}^2[\tan^2\mathrm{f} + \tan^2\mathrm{d} - 2\tan\mathrm{f}\tan\mathrm{d}\cos\alpha] \tag{13}$$

And with substitution of (11):

$$e'^2 = R^2\left(\frac{\tan^2 d}{\cos^2 \alpha} - \tan^2 d\right) \quad (14)$$

$$= R^2 \tan^2 d(\sec^2 \alpha - 1) \quad (15)$$

$$e' = R\tan(d)(\sec^2 \alpha - 1)^{1/2} \quad (16)$$

The angle γ can then be computed from the right angled triangle A_1'B'P' as:

$$\tan(\gamma) = \frac{R\tan(d)+R\cot(d)}{R\tan(d)(\sec^2 \alpha - 1)^{1/2}} \quad (17)$$

$$= \frac{1+\cot^2 d}{(\sec^2 \alpha - 1)^{1/2}} \quad (18)$$

On the surface of the sphere this angle is a right angle, and so its error in the projected form is 90°-γ and since the situation is symmetrical on the other side of the boundary the total kink in the line is given by $2(90° - \gamma)$. If we again aim for a maximum angular distortion of 2.1", but this time applied to the kink of the line, then at the previously established maximum value for d of 0.256° equation (18) solves as a maximum of 14° for α. Again applying Napier's rules to triangle O_1BP gives the offset of the point P from the central axis, *e*:

$$\tan e = \sin d \tan \alpha \quad (19)$$

This gives a swath of ±7.1 km on either side of the central linear axis, which is less than the 20 km that was defined purely by the maximum values of scale factor and angular distortions.

In summary, the myriahedral approach of a series of gnomonic sub-projections gives a potential methodology for establishing a projected coordinate system for quasi-linear engineering projects that have a "sinuous" route, potentially with branches. It can give engineering standard accuracy over a swath of around ±7km from the central linear axis, and potentially more. However, it does have one drawback: it cannot be adapted for the case where the scale factor varies as the height above the surface of the sphere or ellipsoid changes. Hence, after discussing this effect in the next section, an alternative methodology will be discussed.

The Effect of Elevation

Our objective is to establish a projected coordinate system for a long, quasi-linear project on the surface of the Earth, such that the scale factor distortion is within constrained limits. In practice, this means establishing a unity scale factor at a height that may be several hundred meters above the surface of the ellipsoid. Before deriving a methodology for this, let us examine the basic premise and summarize alternative approaches.

The effect of height on horizontal scale factor is due to the fact that a line at a height h above the surface of a sphere of radius R will be longer than the corresponding distance on the spherical surface by a factor of $(R + h)/R$. Given an Earth radius of approximately 6,400 km this implies a scale factor distortion of 1 ppm for every 6.4 m of elevation above the surface. So for example, for an engineering project situated at a height of 500 m above the surface of the ellipsoid, the distortion

corresponds to 78 ppm. In this case, the scale factor experienced by the user will be a compound of that due to the projection and that due to the height effect, thus:

$$\mathrm{k_{Compound}} = \mathrm{k_{Projection}} \mathrm{k_{Height}} \tag{20}$$

Expressing this in terms of the actual meanings of the scale factor terms gives:

$$\frac{\mathrm{D_{Proj}}}{\mathrm{D_{Ground}}} = \frac{\mathrm{D_{Proj}}}{\mathrm{D_{Ellipsoid}}} \mathrm{k_{Height}} \tag{21}$$

And hence:

$$\mathrm{k_{Height}} = \mathrm{k_h} = \frac{\mathrm{D_{Ellipsoid}}}{\mathrm{D_{Ground}}} = \frac{\mathrm{R}}{(\mathrm{R+h})} \tag{22}$$

In many cases an engineering project may be situated at a height that makes the scale factor distortion significant, but the variation of height *within* the project is minimal. In such cases, a projected coordinate reference system may be devised without distortion by suitable selection of the overall re-scaling parameter k_0. This becomes more difficult for larger engineering projects, and especially for the transport links that we are particularly concerned with in this paper. As an example of what might typically be encountered, consider the main line rail route from London to Glasgow, via the west coast of Great Britain. Figure 5 shows the elevation plotted against distance along the line from London, from which it can be seen that the maximum elevation is 291 m and the minimum is 12 m; hence the spread of scale factor distortions is 44 ppm, which significantly exceeds the accuracy requirement.

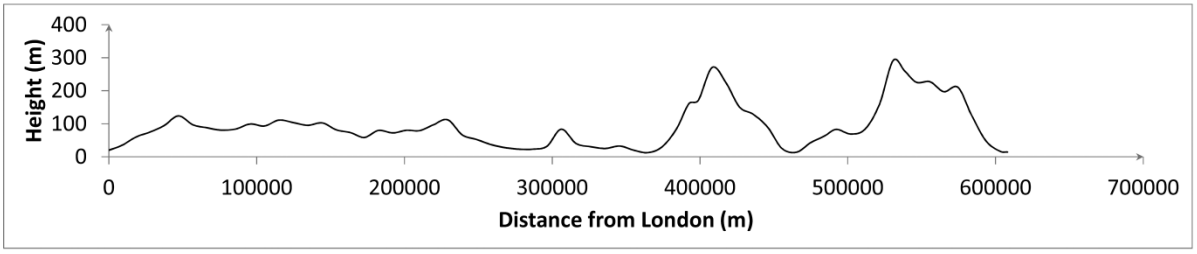


Figure 5: Elevation of the London to Glasgow west coast rail route, with respect to the general trend line of the project.

Even greater distortions are encountered in highway projects, where the height variations can be extreme in mountainous countries. In such cases, in particular in South America, common practice is to introduce separate coordinate systems for each height band. The Plano Topogràfico Local (or PTL) is a coordinate system based on the Transverse Mercator projection but with an overall scaling factor that makes the map scale factor correct at the central height of the band; the height range of the band will depend on the accuracy required, with the designation *primary* corresponding to 25 ppm and applicable to a height range ±150 m on either side of the mean (Gobierno de Chile, 2016). Another example is the use of the same system on the construction of the motorway linking São Paulo and Rio de Janeiro, a length of 405 km and with a height range of 700 m (Idoeta, 2005). This was divided into 12 separate coordinate zones, with the centers of the height bands ranging from 50 m to 700 m above datum.

Thus, the concept of a zonal coordinate system, which was introduced in the context of limiting the growth of scale factor distortion over a large area, is in this context re-introduced for the purpose of controlling the growth of scale factor with respect to height. The effect is the same: the scale factor has been controlled, but continuity has been lost.

Curvilinear Projections in 3D

The preceding sections have established the following requirements for a projected coordinate system to be used with certain types of project: a continuous coordinate system, without separate zones; the scale factor true along the general trend line of a curvilinear route; and the line of true scale to vary in height as well as in plan, such that it follows the route more or less at ground level.

In defining these ideal characteristics, a distinction is made between the *actual* route taken by a road or a railway, and its *general trend line*. A railway, for example, may have some radii of curvature of a few hundred meters or less, and allowing the coordinate system to follow the twists and turns on this scale is likely to have a severe impact on the shapes of features and hence on the conformal properties. Maintaining the scale factor along a general trend line, with a curvature measured in 10s or 100s of kilometers, gives a broad lateral stability away from the trend and gives a coordinate system that is applicable for several kilometers on either side.

This is the approach adopted for the *Snake Projection*, introduced in Iliffe et al (2007). Since its initial publication, extensive further experience has been accumulated of its practical implementation, and some modifications have been introduced. In this paper, the mathematical derivation of the projection in its current form is summarized, in order to highlight its defining parameters: the effect of these is then explored through examples of its practical implementation.

Let an initial base projection be defined: in most instances this will most suitably be the Oblique Mercator projection, its central axis aligned as closely as possible with the general trend line of the linear project.

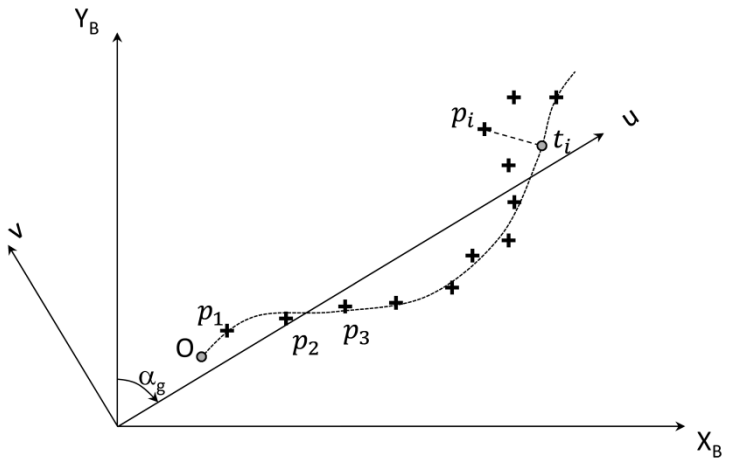


Figure 6: Coordinate systems adopted with respect to seed points of a quasi-linear project, together with fitted horizontal trend line.

With reference to Figure 6, (X_BY_B)is the coordinate system of the base projection, whilst (uv) is a right handed coordinate system with the same origin but aligned along the linear axis of the projection.

This system is defined by its bearing with respect to the Y_B axis, α_g . The coordinates of the actual project route (the railway, for example) are shown as a series of individual crosses (simplified to just a few points): these points are known as the *seed points,* p_i, as they define the project that requires this optimized coordinate system. Within this *uv* coordinate system, let a representative horizontal trend line be defined by a polynomial of degree N, with the form:

$$v = f(u) = a_0 + a_1 u + a_2 u^2 + \cdots + a_N u^N \quad (23)$$

The degree of the polynomial is selected to accomplish two objectives: to maximize the radius of curvature of the trend line and to align it as closely as possible with the original track. Ultimately, the scale factor of the new projection will be unity along the trend line but the distortion will grow as we move away from it, and hence the need to align it as close to the original track as possible. The objective of maximizing its radius of curvature is to keep the conformality of the projection along the track by avoiding any sharp twists to the coordinate space. To some extent these two objectives act against each other, but the practical outcome will depend ultimately on the typical geometry of rail routes. As examples, Table 1 shows the outcome of fitting polynomials to four high profile rail routes, all of which currently make use of the Snake Projection for their engineering coordinate systems, along with the maximum value of scale factor distortion that resulted in each case.

Project/route	Polynomial order	Minimum radius of curvature (km)	Maximum distance from trend line to track (km)	Interpolation interval (km)	Smoothing parameter, d (km)	Maximum scale factor distortion (ppm)
Crossrail	2nd	364	15.5	10	7.5	6.3
High Speed 2 (HS2)	2nd	419	20.2	15	15	15.5
Channel Tunnel Rail Link	2nd	511	7.3	10	10	8.4
East Coast Main Line	4th	103	20.8	10	10	11.9

Table 1: Selected parameters and principal characteristics of four sample uses of the Snake Projection.

Next, each of the seed points p_i is mapped to a corresponding point t_i at the foot of the perpendicular to the trend line, using an iterative search for the shortest distance, the u coordinate of this point being expressed as u_i. Then introducing an arbitrary starting point O for the trend line, with the coordinate u_O, the distance along the trend line from this point to each point t_i is computed as:

$$s_i = \int_{u_O}^{u_i} \sqrt{1 + \left(\frac{df(u)}{du}\right)^2}\, du \quad (24)$$

In practice this may be evaluated by numerical integration. As a result of this, the longitudinal profile of the project may be shown as in Figure 7 by plotting the heights h_i of each of the seed points against the distances s_i. Note that with the exaggeration of the vertical scale, what are shown as apparent “spikes” in Figure 9 actually result from gradients of less than 1%, and so are not unrealistic as representing points along a railway track.

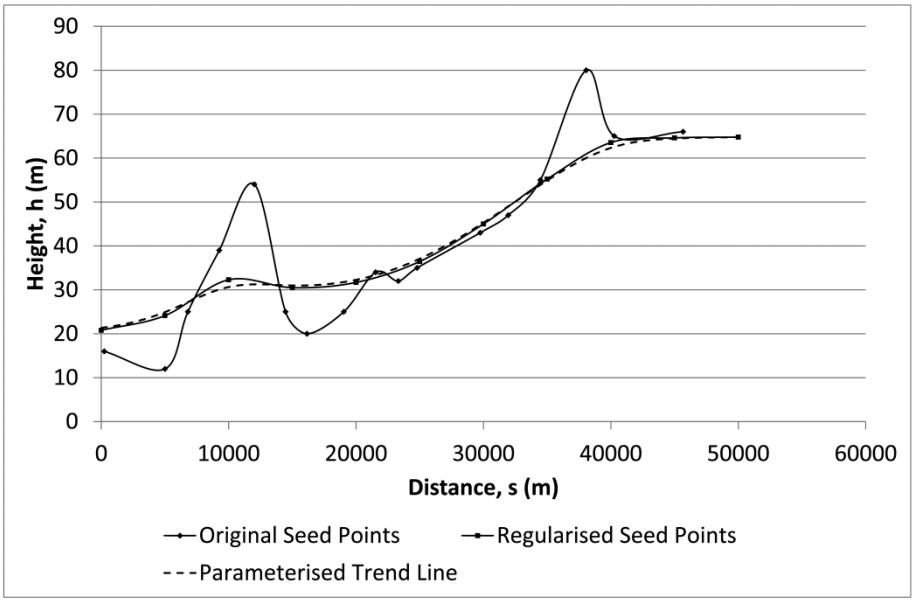


Figure 7: Sample vertical profile, with regularized seed points fitted at a spacing of 5 km and with a 5 km smoothing distance; also the parameterized trend line.

For the general case the spacing of the seed points is irregular; the initial step is therefore to interpolate the data onto a regularly spaced set of points. This was carried out using an indirect distance weighting function, in which the height $h(r_j)$ of the j[th] interpolated point, r_j is given by:

$$\mathrm{h(r_j)} = \left(\sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{h(t_i)} \left[\mathrm{e}^{-|\mathrm{s(t_i)}-\mathrm{s(r_j)}|/\mathrm{d}}\right]\right) \left(\sum_{\mathrm{i=1}}^{\mathrm{n}} \left[\mathrm{e}^{-|\mathrm{s(t_i)}-\mathrm{s(r_j)}|/\mathrm{d}}\right]\right)^{-1} \quad (25)$$

where $h(t_i)$ are the heights at the original irregular seed points and $s(t_i)$ and $s(r_j)$ are the distances along the trend line to the original (irregular) and regularized seed points respectively (Lancaster and Salkauskas, 1986). The parameter *d* is an arbitrary distance parameter that governs the level of smoothing at this stage of the process: a shorter value of *d* means that the heights at the regularized points are more strongly influenced by closer seed points, and hence that limited smoothing takes place. This means that the vertical trend line closely follows the height of the original track, but also means that sudden changes in height can introduce sharp changes in scale factor and hence lead to a non-conformal coordinate system. Table 1 shows the parameters that were adopted for the examples shown in order to minimize scale factor and angular distortion, and the process is illustrated in Figure 7.

The final step in establishing the vertical trend line is to define the elements of the cubic B-splines (Lancaster and Salkauskas, 1986). These are used to determine the height of the trend line,h_T, at any required point from:

$$\mathrm{h_T} = \mathbf{qMp} \quad (26)$$

Where:

$$\mathbf{q} = (\mathrm{q^3} \quad \mathrm{q^2} \quad \mathrm{q} \quad 1) \quad (27)$$

and q is a value between 0 and 1 to indicate how far through a particular bay a required point lies and the four successive heights of the r_j points are collected in the vector **p**:

$$\mathbf{p} = (\mathrm{h}(\mathrm{t}_{\mathrm{i}-1}) \quad \mathrm{h}(\mathrm{t}_{\mathrm{i}}) \quad \mathrm{h}(\mathrm{t}_{\mathrm{i}+1}) \quad \mathrm{h}(\mathrm{t}_{\mathrm{i}+2}))^{\mathrm{T}} \tag{28}$$

The term **M** is known as the blending matrix and is defined as:

$$\mathbf{M} = {}^{1}\!/_{6}\begin{pmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{pmatrix} \tag{29}$$

An example of the parameterized trend line is also shown in Figure 7, in which it can be seen that a further degree of smoothing has been introduced in the process.

The steps summarized in equations (23) to (29) have established a trend line in plan and height that follows the general direction of a project but smooths out any sudden turns in the horizontal plane or spikes in height. We now define a new projected coordinate system with respect to this trend line.

The requirement on the trend line is that a distance measured on the ground between two points must be equal to the distance computed from the coordinates of these two points, or simply that the compound scale factor $k_{Compound}$ is unity. If we denote the scale factor on the trend line at zero height by the symbol k_{T-0} then Equations (20) and (22) give:

$$k_{Compound} = 1 = \mathrm{k}_{\mathrm{T}-0}\mathrm{k}_{\mathrm{h}} \tag{30}$$

And therefore the required scale factor at zero height of the trend line is given by:

$$\mathrm{k}_{\mathrm{T}-0} = \frac{1}{\mathrm{k}_{\mathrm{h}}} = \frac{\mathrm{R}+\mathrm{h}_{\mathrm{T}}}{\mathrm{R}} \tag{31}$$

If we denote the scale factor on this line in the base projection by k_B then to force the compound scale factor at the design height to equal unity each element of the trend line needs to be multiplied by a variable re-scaling factor $k'(s)$ that is a function of the distance along the trend line from the origin:

$$\mathrm{k}'(\mathrm{s}) = \frac{\mathrm{k}_{\mathrm{T}-0}}{\mathrm{k}_{\mathrm{B}}} \tag{32}$$

We note here that both k_{T-0} and k_B are themselves variable with respect to s.

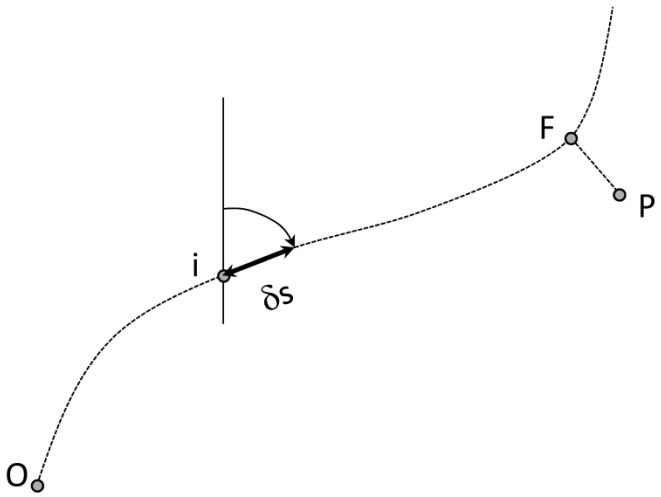


Figure 8: Trend line from the origin O to the foot of the perpendicular, F, of a representative point P.

With reference to Figure 8, the practical implementation of this in terms of the computation of the coordinates of a point P is achieved by considering the changes to the coordinates that result from the cumulative re-scaling of the elements of the trend line from the origin O up to the foot of the perpendicular F; and then to maintain local conformality the perpendicular FP is itself re-scaled by the same factor as was applied along the trend line at F, or $k'(F)$. Thus the coordinates of the point P are given as:

$$x_P = x_{P(Base)} + \sum_O^F (k'(s) - 1) sin\theta_i \delta s + FP(k'(F) - 1)\, sin(\theta_F + 90°) \qquad (33)$$

$$y_P = y_{P(Base)} + \sum_O^F (k'(s) - 1) cos\theta_i \delta s + FP(k'(F) - 1)\, cos(\theta_F + 90°) \qquad (34)$$

θ_i is the bearing of the trend line at the general summation point *i*, and δs is the length of the summed element. The summation is an approximation to pure integration, and in practice can be achieved with step lengths of 500 m.

Equations 33 and 34 are reversible, to determine the base coordinates from the Snake Projection coordinates, provided that an iteration is carried out to account for the approximation of computing the scale factor $k'(s)$ with respect to the coordinates $(x_P\ , y_P)$ rather than the base coordinates. The geographic coordinates may then be determined by reversing the base projection.

Practical Applications

The Snake Projection was initially developed for use in upgrade projects on the west coast railway line from London Euston to Glasgow, achieving a continuous coordinate system with a maximum scale factor distortion of 11 ppm on the track (Iliffe et al, 2007). Subsequently it was adopted as the standard methodology by Network Rail and other organizations (Rigby, 2012; Iliffe et al, 2013). In this section, its use in three example applications is explored: in turn, these demonstrate its use in a major infrastructure project, a project demonstrating a large height range, and a project with multiple integrated branches.

The proposed HS2 (High Speed Route 2) rail route will run from London to Birmingham, and eventually to Manchester and Leeds, and is due to begin construction in 2017 (HS2 Ltd, 2016). The first phase of the project has been given parliamentary approval, and will cover the 225 km section from London to the West Midlands. HS2 is the largest infrastructure project in Europe, and required a continuous low distortion coordinate system to use from the engineering design through to final implementation. In 2014 the project commissioned the creation of the coordinate system *HS2P1+14* [derivation: HS2, Phase 1 of the project, including an extending spur, commissioned in 2014]. Going through some of the choices presented in this paper in turn, a standard choice might have been to tailor the Transverse Mercator projection to the project, as it has a greater north-south extent than east-west in its initial phase. However, even optimizing the choice of the central meridian and the re-scaling factor, substantial sections of the route go far beyond the 20 ppm limit on the distortion, in fact reaching over 50 ppm in parts, as can be seen in Figure 9a. A better choice would be the Oblique Mercator projection, selecting a central line of contact at an approximate azimuth of 320° at the central point; most of the track itself is within 10 km of this axis, and since it was previously shown that engineering standard distortion could be achieved within a corridor of ±40km with respect to the center line, this leaves up to 30 km as an additional buffer for associated works. However, as seen in Figure 9b, the dominant influence on the distortion is now the height effect, with features such as Cannock Chase or the Chilterns (up to approximately 250 m above sea level) clearly visible in terms of the scale distortion that they have caused.

The coordinate system HS2P1+14 itself, based on the Snake Projection, was created with a maximum scale factor distortion on the track of 15.5 ppm, with an overall r.m.s. value of 3.9 ppm. Moreover, as can be seen from Figure 9c, the effects of the hills have been extensively moderated, and the coordinate system is now applicable to the whole swath on either side of the central corridor.

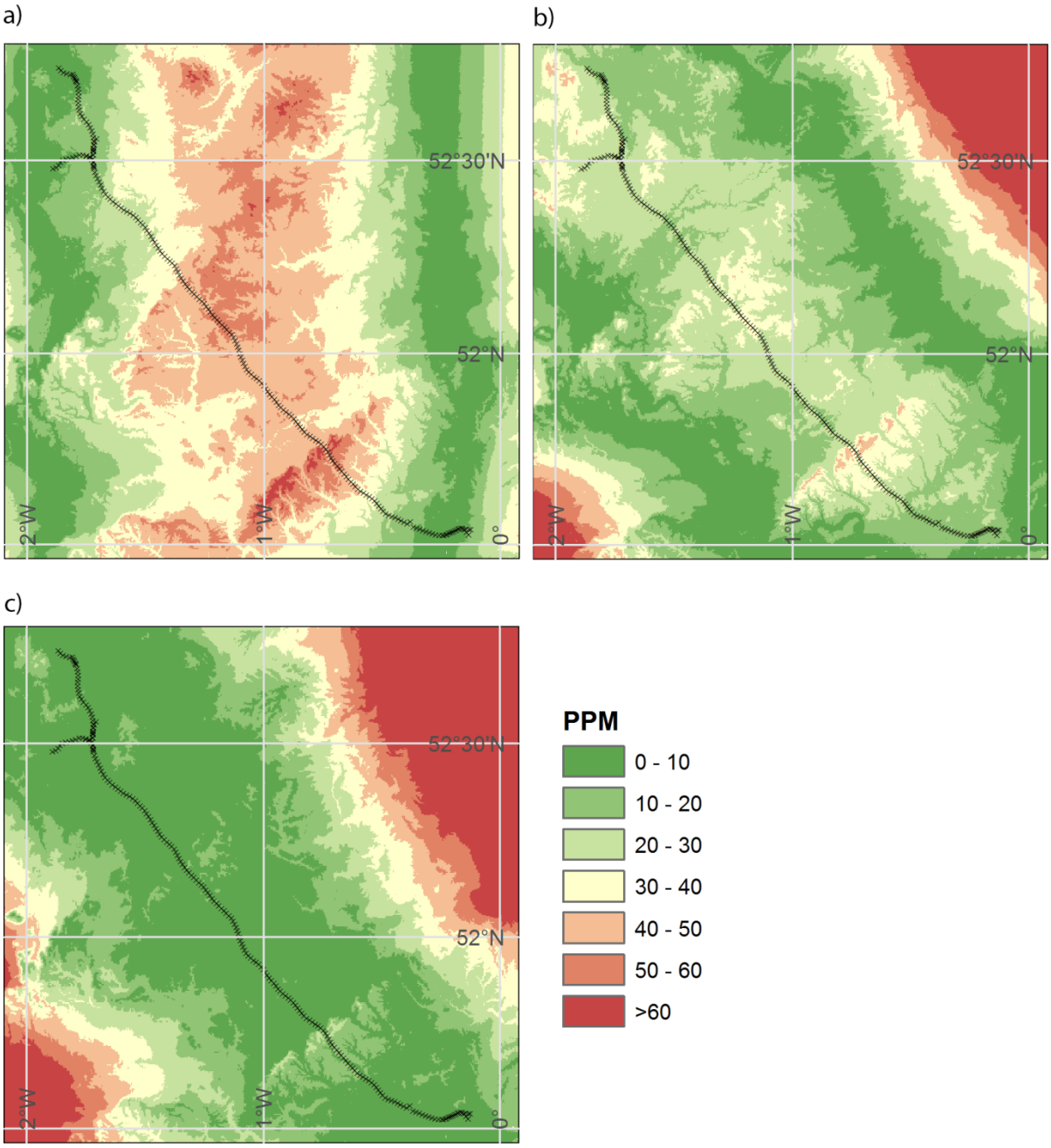
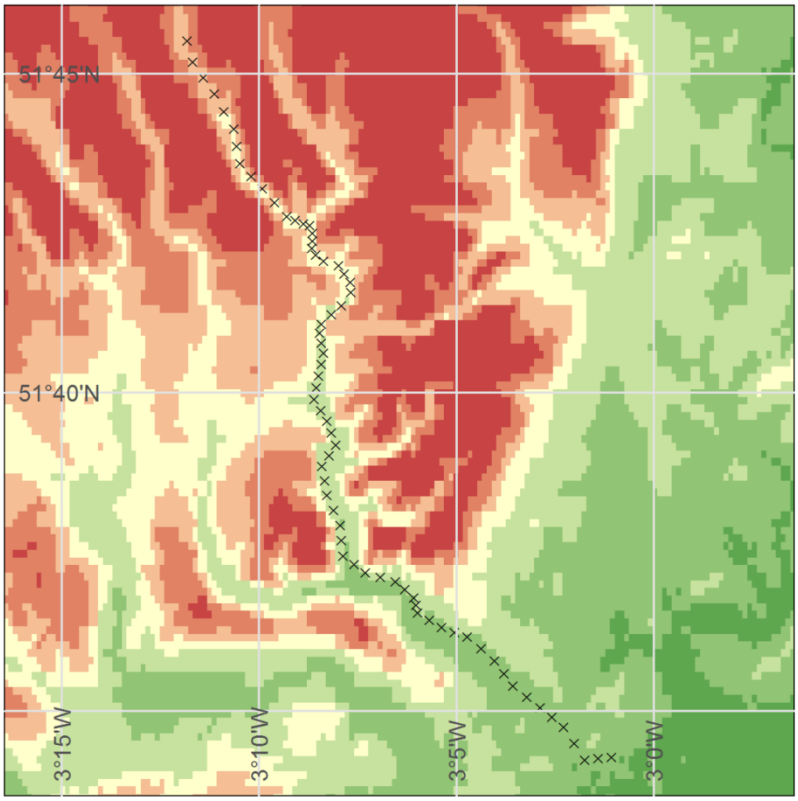


Figure 9. Tailored map projections for HS2, showing levels of scale factor distortion in parts per million: a) Transverse Mercator; b) Oblique Mercator; c) Snake Projection. Note that the distortion for each tailored projection in Figures 9, 10, and 11 is shown on a common background projection of Web Mercator.

An even starker illustration of the effect of elevation is seen in the example of the coordinate system created for the redoubling work on the line from Newport to Ebbw Vale in South Wales. Although the overall extent of the project was not large (less than 30 km direct distance from one end to the other), the elevation changed by almost 250 m on the track itself, which would cause approximately 40 ppm distortion in a conventional projection. Figure 10a shows the optimally selected Oblique Mercator projection, which shows the distortion on the track of up to 40 ppm and the even larger distortions in the hills rising on either side. For this project, the coordinate system EBBWV14 was

created based on the Snake Projection, resulting in a maximum distortion of 8.0 ppm and an r.m.s value of 3.8 ppm. The resulting distortion is illustrated in Figure 10b.

a)



b)

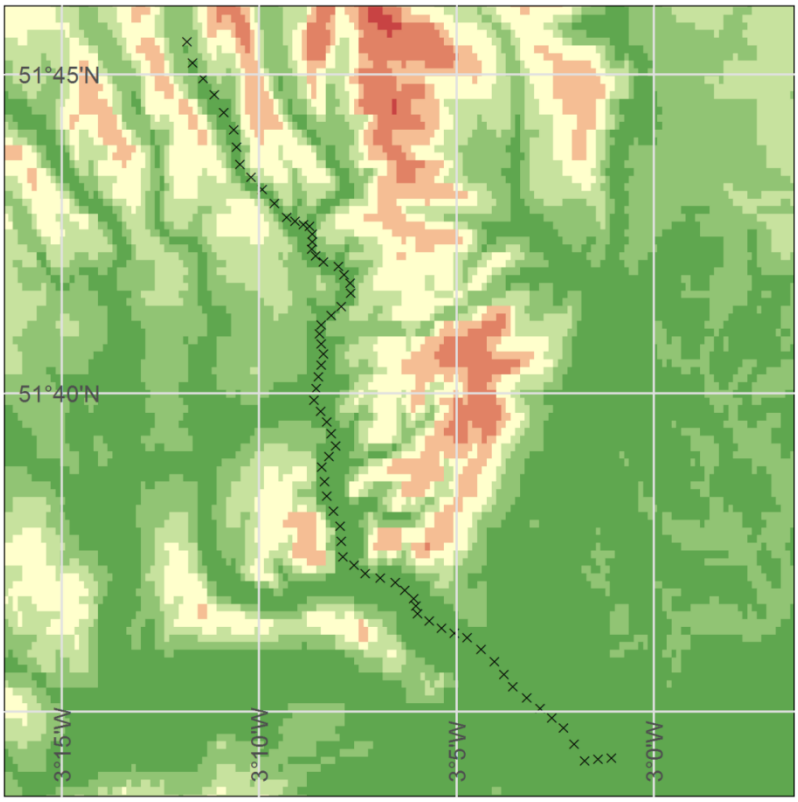


Figure 10. Tailored map projections for Newport to Ebbw Vale: a) Oblique Mercator; b) Snake Projection. [Key as for Figure 9.].

When considering the myriahedral approach in a previous section, it was noted that it was also able to cope with *branches* off the main trend line by having more than one fold side in a selected face of the polyhedron. The same effect can be achieved with the Snake Projection by rectifying the final XY or EN grid; that is, by specifying the grid coordinates that should be assigned to a point of given latitude and longitude, and the angle that the grid should make with true north at this point. However, to achieve continuity of coordinates between the main grid and the branch, a necessary condition is that the two grids have the same scale factor distortion: if the existing grid has a distortion of +10ppm, and the new one has a distortion of -10ppm (both perfectly acceptable for engineering applications) then the scale distortion difference is 20 ppm and within 1 km of the junction the coordinates will differ by 20 mm. On the other hand, if the scale can be made the same within 1 ppm (principally by localized adjustments to the height of the trend line) then the coordinates will agree to within 2 – 3 mm up to 2 – 3 km from the junction, and through computations can be made on either branch on a seamless coordinate system. As with the myriahedral approach, multiple branches can be introduced (in a "dendritic" pattern) but (as noted by van Wijk, 2008) there can never be a "re-join" at another junction.

This approach was adopted as part of the East-West Rail project, a major initiative to establish a strategic railway connecting East Anglia with Central, Southern and Western England (East West Rail, 2016). Initially, the grid OBB12 was commissioned in 2012 to connect Oxford, Bicester and Bletchley. This was subsequently extended as far as Bedford by merging a new grid, OBB12ext at Bletchley. Then in 2015 the grid CAPR15 was merged at Claydon Junction, covering the route south through Aylesbury to Princes Risborough. Figure 11 shows the extent of applicability of the combined grids, showing those regions where the coordinates of each one differ by less than 1 mm from the neighboring grid. In this way, a project that is *quasi-linear* but *multi-branched* has been assigned a coordinate system that is both low distortion and continuous.

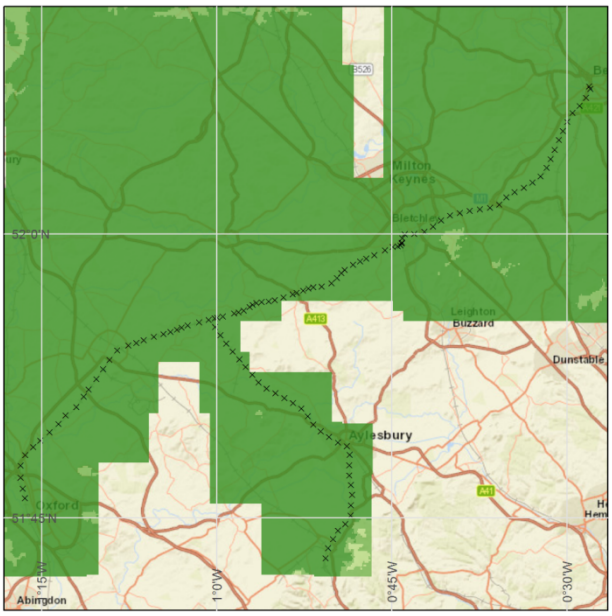


Figure 11. Tailored map projection for the East-West Rail project, showing the overlap area of the three grids (OBB12, OBB12ext and CAPR15) and the level of scale factor distortion. Note that due to

the way the distortion is computed, the overlap area has been clipped by 1 km in each direction, and so is slightly larger than shown here. [Key as for Figure 9. Background mapping © World Street Map].

Conclusions

Exploiting the geometrical properties of standard map projections such as Transverse Mercator or Lambert Conformal Conic is a standard cartographic practice, in which the most promising projection type is selected for a particular task and its defining parameters optimized. The work presented here has analyzed the way in which this concept can be extended to the more general application of a project that broadly follows a linear route over hundreds of kilometers, meandering as it does so, and changing its scale as it rises and falls over hills and valleys. The interrupted approach of the Myriahedral Projection, or the projected trend line of the Snake Projection, are presented here as candidates to provide solutions to this problem. Both derive projected coordinate systems that are effectively distortion-free over swaths several kilometers wide, following the route of a quasi-linear project for hundreds of kilometers (although only the Snake Projection deals with the height issue at the same time).

The significance of such a coordinate system lies in the fact that it is both low-distortion and continuous. In any engineering application this is important, but particularly so in the context of railways. New high speed rail lines require minimum radii of curvature of 7.2 km for operation up to 400 kph (HS2 Ltd, 2016), which underlines the fact that a coordinate system split into separate zones is likely to lead to difficulties in computation and design of the geometric properties of the railway. At the same time, the low-distortion property means that the design may be carried out with a 1:1 relationship between the coordinate system and the ground. By comparison, a design carried out in a national or regional coordinate reference system would face scale factor distortions of as much as 500 ppm: this would need to be allowed for when construction or maintenance work is commenced, and adjustments may need to be made to ensure that the design was still within specification. Working directly in a continuous true-scale coordinate system of the type presented in this paper eliminates this type of problem, and ensures that a single, fit-for-purpose coordinate system may be used across the whole of a project, almost irrespective of its length.

Acknowledgements

The contribution of Dr James Turner to the coding and the production of imagery is acknowledged, as is the collaboration of Network Rail in providing data on railway routes.

References

Argerich, A. I. (2006). Issues about Satellite Images Georeferencing in Argentina. Proceedings of XXIII FIG Congress, Munich, Germany, October 8-13, 2006.

Bildirici, I. O. (2015) Quasi indicatrix approach for distortion visualization and analysis for map projections, International Journal of Geographical Information Science, 29:12, 2295-2309, DOI: 10.1080/13658816.2015.1074236.

East West Rail (2016). http://www.eastwestrail.org.uk/

Gobierno de Chile (2016). Edición 2016 del Manual de Carreteras de la Dirección de Vialidad. Ministerio de Obras Públicas, Gobierno de Chile. Santiago de Chile, Chile.

Highways Agency (2007). Implementation of Local Grid Referencing System for England. Interim Advice Note 99/07.

HS2 Ltd (2016). https://www.gov.uk/government/organisations/high-speed-two-limited

Idoeta, I. V. (2005). Uso del plan topográfico local en obras lineales de ingeniería: caso Autopista Presidente Dutra [The use of a local topographic plan in linear engineering works: case study of the President Dutra Highway]. Proceedings of the 6th Geomatics Week, Barcelona, Spain, 2005.

Iliffe, J. C., Arthur, J. V., and Preston, C. "The Snake Projection: a customised grid for rail projects". Survey Review, Vol. 39, No. 304, pp 90 – 99, 2007.

Iliffe, J. C. and Lott, R., 2008. Datums and Map Projections, for Remote Sensing, GIS and Surveying. Second Edition. Whittles Publishing. 208 + ix pp.

Iliffe, J.C, Preston, C., and Hewitt, C. (2013). "Coordinate Systems for High Speed Track Alignment". Journal of the Permanent Way Institution. Volume 131, January 2013, pp 52 – 55.

Lancaster, P. and Salkauskas, K (1986). Curve and Surface Fitting: An Introduction. London. Academic Press.

Maling, D. (1992). Coordinate Systems and Map Projections. 2nd Edition. ISBN 9780080372334. Pergamon.

Orihuela, S. Generalization of the Lambert–Lagrange projection. The Cartographic Journal Vol. 53 No. 2 pp. 158–165 May 2016.

Ren, L., Clarke . K.C., Chenghu, Z. , Ding, L. and Li, G. (2010) Geometric Rectification of Satellite Imagery with Minimal Ground Control Using Space Oblique Mercator Projection Theory, Cartography and Geographic Information Science, 37:4, 261-272, DOI: 10.1559/152304010793454309.

Ren, L., Clarke . K.C., Chenghu, Z. and Ren, H. (2013) The space oblique conic projection, Cartography and Geographic Information Science, 40:4, 282-288, DOI: 10.1080/15230406.2013.810804

Rigby, R. (2012). Corridor Projections Solved. Professional Surveyor Magazine. November 2012. Vol. 32. No. 11.

Šavrič, B., Jenny, B., White, D. and Strebe, D.R. (2015) User preferences for world map projections, Cartography and Geographic Information Science, 42:5, 398-409, DOI: 10.1080/15230406.2015.1014425.

Šavrič, B., Jenny, B. and Jenny, H. (2016) Projection Wizard – An Online Map Projection Selection Tool, The Cartographic Journal, 53:2, 177-185, DOI: 10.1080/00087041.2015.1131938

Skokanová, H. & Havlíček, M. Acta Geod. Geoph. Hung (2010) 45: 120. doi:10.1556/AGeod.45.2010.1.17

Snyder, J.P. 1978. The space oblique Mercator projection. Photogrammetric Engineering and Remote Sensing 44(5): 585-96.

Snyder, J. P. (1987). Map projections: A working manual, Geological Survey Professional Paper 1395, Washington DC, United States Government Printing Office.

Stem, J. E. "State Plane Coordinate System of 1983". (1990). NOAA Manual NOS NGS 5. Rockville, MD. January 1989. Reprinted with minor corrections March 1990.

Strebe, D.R. (2016) An adaptable equal-area pseudoconic map projection, Cartography and Geographic Information Science, 43:4, 338-345, DOI: 10.1080/15230406.2015.1088800.

Uren, J and Price, B. (2010). *Surveying for Engineers*, Fifth Edition. Basingstoke: Palgrave Macmillan.

Van Brummelen, G. (2013) Heavenly Mathematics: The Forgotten Art of Spherical Trigonometry. Princeton University Press. Xvi + 192 pp. Princeton, New Jersey.

van Wijk, J. J. (2008) Unfolding the Earth: Myriahedral Projections, The Cartographic Journal, 45:1, 32-42, DOI: 10.1179/000870408X276594

Zhao, H., Zhu, H., Li, L. and Xing, Y. (2007). 'COM-based expert system for map projection selection', Geoinformatics 2007: Cartographic Theory and Models, ed. By Li, M. and Wang, J., Bellingham: Society of Photo-Optical Instrumentation Engineers, pp. 675115–65125.