DOI 10.1590/s1982-21702022000100007

ORIGINAL ARTICLE

Partitions of normalised multiple regression equations for datum transformations

Andrew Carey Ruffhead[1] - ORCID: 0000-0002-7836-6702

[1]University of East London, School of Architecture, Computing and Engineering, UK.
E-mail: ruffhead40b@yahoo.co.uk

Received in 24th May 2021.
Accepted in 25th November 2021.

Abstract:

Multiple regression equations (MREs) provide an empirical direct method of transforming coordinates between geodetic datums. Since they offer a means of modelling distortions, they are capable of a more accurate fit to datum-shift datasets than more basic direct methods. MRE models of datum shifts traditionally consist of polynomials based on relative latitude and longitude. However, the limited availability of low-power terms often leads to high-power terms being included, and these are a potential cause of instability. This paper introduces three variations based on simple partitions and 2 or 4 smoothly conjoined polynomials. The new types are North/South, East/West and Four-Quadrant. They increase the availability of low-order terms, enabling distortions to be modelled with fewer side effects. Case studies in Great Britain, Slovenia and Western Australia provide examples of partitioned MREs that are more accurate than conventional MREs with the same number of terms.

Keywords: multiple regression equations; surface polynomials; datum transformations; geodetic datums.

How to cite this article: RUFFHEAD, A.C. Partitions of normalised multiple regression equations for datum transformations. *Bulletin of Geodetic Sciences*. 28(1): e2022007, 2022.

1. Introduction

One method of approximating a variable over a surface is a bivariate polynomial which is a linear combination of terms $x^i y^j$. (For the moment, *x* and *y* are simply horizontal coordinates measured in orthogonal directions.)

There are two commonly-used ways of defining the range of terms in a bivariate polynomial. One is the polynomial *with top power n*. This has $(n+1)^2$ possible terms and is described in (1).

$$p(x,y) = \sum_{i=0}^{n} \sum_{j=0}^{n} c_{i,j} x^i y^j \tag{1}$$

The other – adopting terminology used by Berberich, Emeliyanenko and Sagraloff (2010) – is the polynomial *of total degree n*. This has *(n+1)(n+2)/2* terms and is described in (2).

$$p(x,y) = \sum_{i=0}^{n} \sum_{j=0}^{i} c_{i,j} x^i y^j = \sum_{i,j} \left\{ c_{i,j} x^i y^j : i \geq 0, j \geq 0, i+j \leq n \right\} \tag{2}$$

To complete the polynomial model of the dependent variable, the coefficients of the monomials can be computed by least-squares optimisation, based on the model's quality-of-fit to data values at control points. To some users of this approach, this meets the broad definition of *regression,* in that it is a statistical method that attempts to determine the strength and character of the relationship between the dependent variable and the independent variables.

At best, the derivation of a model by least-squares is a limited interpretation of regression. The contributions of individual terms are quantified by the size of the coefficients, but the model is liable to include terms which are more likely to be the result of noise then any trend in the dependent variable. For a polynomial model to be a regression model in the full sense of the word, there should be some process to ensure that the components are statistically significant.

A widely-respected text on regression methods is Draper and Smith (1998), which also exists in earlier editions (from 1966 onwards). The chapter "Selecting the 'Best' Regression Equation" describes several approaches including stepwise regression.

Appelbaum (1982) introduced polynomials obtained by stepwise regression as a means of modelling datum shifts. He adopted the term *multiple regression equations* (MREs), although it is arguable that multiple regression functions would have been a more accurate description. The paper presented new computations of coordinate shifts from European Datum 1950 (ED50) to World Geodetic System 1972 (WGS72).

MREs were adopted by the US Defense Mapping Agency (DMA) to relate local datums to the World Geodetic System 1984 (WGS84). A selection covering large areas can be found in Appendix D of NIMA (2004). A large collection of MREs for local datums to WGS84 is contained in Sections 19 and 20 of DMA (1987b). It should be noted that DMA was absorbed into the National Imagery and Mapping Agency (NIMA) which is now the National Geospatial-Intelligence Agency (NGA).

The main attraction of polynomial models to compute coordinate shifts between geodetic datums is that they can cater for distortions (which are mainly caused by the limitations of older surveying methods). Similarity transformations cannot do this, neither can near-conformal methods such as Standard Molodensky and Bursa-Wolf. An affine transformation also has its limitations, precisely because it preserves collinearity and the ratio of distances between points on a line. Polynomial models can be more accurate within a defined region provided there is a sufficiently dense distribution of points where the shifts are known. This was borne out in the author's derived transformations for AGD84→GDA94 in Western Australia where even the affine models produced residuals around twice those of polynomials with top power 4.

The nature of polynomials, particularly those with high-degree terms, makes them prone to instability if used outside the region covered by control points. This can be countered by adding surrounding pseudo data to the control points, but this can worsen the fit to actual data.

There are alternatives to MREs for distortion modelling. Perhaps the most notable alternative is grid look-up models (see, for example, Greaves 2004). These offer greater accuracy than MREs but require considerably more storage. They also require substantial observational data to set up, as well as sophisticated interpolation to regularise the data. MREs therefore represent a compromise between direct models and the most precise methods of surface-fitting.

The purpose of the investigation behind this paper is to establish whether accuracy and stability of MREs can be improved by a system of partitioning that allows a wider choice of low-order terms.

2. Datum transformations by multiple regression equations

2.1 Construction of MREs

The DMA MREs were derived by the same stepwise regression process as Appelbaum's. (See Section 7.2.4.3.3 of DMA 1987a.) Each step begins with the addition of a variable, the choice being the one that gives the biggest improvement to accuracy. A statistical-significance test determines whether one or more already-included variables should be removed. The process continues until the desired precision is achieved.

MREs are usually polynomial functions of intermediate coordinates *U* (based on latitude) and *V* (based on longitude). The range of possible ways of defining these intermediate coordinates is summarised in Ruffhead (2018). The form used by DMA is that given by (3) and (4) except that a single scale factor *K* is used. φ and λ are in degrees, and the point $(\varphi_{off}, \lambda_{off})$ is an offset based on a central point. Appelbaum made K a small-integer multiple of $\pi/180$, and this practice has been followed by DMA and others.

$$U = K_1\left(\phi - \phi_{\text{off}}\right) \tag{3}$$

$$V = K_2\left(\lambda - \lambda_{\text{off}}\right) \tag{4}$$

DMA's MREs for the datum shifts $\Delta\phi$ and $\Delta\lambda$ take the form shown in equations (5) and (6), the units being arc-seconds. The top power *n* adopted by DMA is 9, but the statistical-significance tests ensure that less than half the possible terms are used. Obviously, when modelling is done in arc-seconds, residuals need to be converted to linear units via standard formulae based on radii of curvature.

$$\Delta\phi = \sum_{i=0}^{n}\sum_{j=0}^{n} a_{i,j} U^i V^j \tag{5}$$

$$\Delta\lambda = \sum_{i=0}^{n}\sum_{j=0}^{n} b_{i,j} U^i V^j \tag{6}$$

MREs for ellipsoidal height shifts Δh are perfectly possible. Several such MREs can be found in DMA (1987b) (and are ellipsoidal despite being denoted ΔH). In practice, they are used less often than horizontal-shift equations, partly because ellipsoidal heights are not always known for older datums. Some survey organisations prefer to compute datum shifts in terms of Cartesian coordinates (geocentric or nearly so) or plane coordinates (based on compatible map projections). This paper concentrates on datum shifts $\Delta\phi$ and $\Delta\lambda$.

Polynomial functions of a predetermined form, usually with a total degree as in (2), have been proposed to model other datum shifts. 6-term functions (with total degree 2) have been used by Abd-Elmotaal (1994), Kutoglu (2009) and Mitsakaki, Agatza-Balodimou and Papazissi (2006). 10-term functions (with total degree 3) have been used by Ayer, Boakye-Yiadom and Smith-Ephraim (2010). Varga, Grgić and Bašić (2017) used functions with total degree 3, but augmented by a V^4 term. 16-term polynomials with top power 3 have been used by Chen and Hill (2005). In all these examples except Abd-Elmotaal (1994), the equivalence of the polynomials to multiple regression equations is stated or implied, despite the absence of tests for statistical significance.

Kutoglu (2009) does provide a rationale for restricting terms to total degree 3, and that is the danger of "erratic behaviours" from polynomials of higher order. For IFRF94 to ED50 over Istanbul, even the 3rd-degree terms had to be excluded for that reason.

Dawod and Alnaggar (2000) and Gledan and Azzaidani (2014) claim to use stepwise regression, but the set of monomials from which the terms are selected is unusual in both cases. The first omits any products of latitude and longitude terms. The second includes terms to the power 12 and 13.

Although Soycan (2005) also uses the term "multiple regression equation" for a polynomial transformation, polynomial coefficients are tested for significance in an entirely different way from stepwise regression. Polynomials of type (2) were tested, with the total power set to 5, 4, 3, 2 and 1. The 5th-degree terms were discarded *en bloc*, then the 4th-degree terms, then the 3rd-degree terms. The polynomial adopted for each datum-shift model had total degree 2, with all 6 possible terms included.

The inverse (or reverse transformation) of an MRE can be computed by applying the model with a change of sign for each coefficient. This was recommended by Appelbaum (1892) and is described by Ruffhead and Whiting (2020) as a simple same-formula inverse (SSFI). It is not exact relative to the original model, because U and V are calculated from ϕ and λ in the other datum.

2.2 Fully-normalised MREs

In general, the regions over which MREs are defined have a latitude-extent that differs from the longitude-extent. Consequently, the single-scale-factor convention adopted by DMA results in scaling that is too large or too small, thereby causing coefficients to be misleadingly small or misleadingly large. It was noted in Ruffhead (2018) that in one of NIMA's transformations, the scaling of V gave V^9 a maximum value of 1487; the coefficient of that term was not only deceptively small, but should have been computed to 3 extra decimal places.

Ruffhead (2018) recommended separate scale factors to ensure full normalisation, so that both U and V span the range 1 to 1 over the defined region. That convention is adopted here. The constants K_1 and K_2 in (3) and (4), are therefore defined as follows.

$$\phi_{\text{off}} = \left(\phi_{\max} + \phi_{\min}\right)/2 \tag{7}$$

$$\lambda_{\text{off}} = \left(\lambda_{\max} + \lambda_{\min}\right)/2 \tag{8}$$

$$K_1 = 2/\left(\phi_{\max} - \phi_{\min}\right) \tag{9}$$

$$K_2 = 2/\left(\lambda_{\max} - \lambda_{\min}\right) \tag{10}$$

The full normalisation of U and V adopted here ensures that the monomials U^iV^j never numerically exceed 1 in the region of interest. It means that $\left|a_{i,j}\right|$ and $\left|b_{i,j}\right|$ give a good indication of the magnitudes of the terms in (5) and (6). In addition,

$$\left|a_{i,j}U^iV^j\right| \le \left|a_{i,j}\right| \tag{11}$$

$$\left|b_{i,j}U^iV^j\right| \le \left|b_{i,j}\right| \tag{12}$$

Ruffhead (2018) also recommended improvements on the SSFI method for computing reverse transformations to the original source-datum-to-target-datum model. The 4th and most accurate method computes the SSFI, then applies the original model, then subtracts the misclosure from the SSFI. The precision relative to the original transformation is easily sub-millimetre provided it is within the region for which the MRE is defined.

2.3 Ratio-based significance tests

Deriving MREs where the number of terms is not pre-determined differs from the process of deriving other transformation models, in that the number of parameters is in theory unlimited. The objective of obtaining a particular level of accuracy depends largely on the number of terms, and their coefficients are the parameters to be optimised. At the same time, MREs are potentially lengthy, so it is desirable to limit the number of terms to those which make a significant contribution to the quality-of-fit.

There is an undeniable element of trial-and-error about finding the "best" model. Every time a combination of monomial terms U^iV^j is tried, the coefficients are obtained by least-squares optimisation. Statistical significance of terms is deduced from the statistical properties of the solution.

Evaluating standard errors of the computed coefficients is explained in Sections D.14-D.15 of Bomford (1980), and summarised below. The components needed are:

- the number N_{obs} of observations, in this case the number of control points;
- the number N_{par} of parameters, in this case the number of coefficients;
- the normal matrix A^TWA where W is the weight matrix;
- the residual vector v.

The standard error of an observation of unit weight σ_0 requires $N_{obs}>N_{par}$ and is given by $\sigma_0^2 = vW^Tv/\left(N_{\text{obs}} - N_{\text{par}}\right)$. The standard errors of the computed coefficients are the square roots of the diagonal elements of $\sigma_0^2(A^TWA)^{-1}$.

One measure of a coefficient's statistical significance, albeit simplistic, is its absolute value divided by its standard error. This ratio can be denoted AP/SE where AP and SE are "absolute parameter" and "standard error" respectively. If a given polynomial is more accurate than needed, the term with the lowest AP/SE can be removed and a new polynomial tested. This step is repeated until there are just enough terms to ensure the desired level of accuracy. This appears to be the process used in Chapter 4 of Gao (2017), although the ratio with the minimum value has to be based on the *absolute* value of the coefficient (not the numerical value as stated).

If a coefficient with the smallest AP/SE has a value for that ratio smaller than 1, the corresponding term in the polynomial differs from zero by less than the uncertainty. Empirical tests indicate that its removal reduces the value of σ_0. Conversely, if that term has a value for AP/SE greater than 1, its removal increases the value of σ_0, but not necessarily by much. Further investigation may be needed if the ratio AP/SE is small because of a standard error that is unusually large (although so far this has not occurred in the author's computations).

Ideally the accuracy of a derived model should be based on a form of cross-validation. The simplest form ("holdout" cross-validation) is where the least-squares solution is based on the root-mean-square (RMS) of the residuals at test points that are distinct from those used as control points. This is a feasible option for a large dataset, provided a representative subset of data points can be set aside for verification without seriously affecting the coverage of the region of interest by control points. The RMS will be an independent estimate of the accuracy.

For small datasets, the argument for deriving a model from all available data may outweigh the case for setting points aside for verification. The RMS of the residuals at the control points can be used as an error estimate, depending on N_{par} and N_{obs}. If they are close in value, the RMS is going to be misleading. In the extreme case N_{par}=N_{obs}, the RMS is zero, but that is no guarantee that the model is an exact representation of the datum shift over the entire region of interest. Indeed, having N_{obs} polynomial terms could introduce massive instability into the model. However, the RMS is likely to be reasonable if N_{par} is much smaller than N_{obs}, ensuring plenty of redundancies.

Obviously, for models of the form (5) and (6) given in arc-seconds, residuals need to be converted to linear units via standard conversion formulae based on radii of curvature.

3. Partitioned MREs

Multiple regression equations as defined by (5) and (6) present a dilemma. The more terms are used, the better the fit to control points. However, the more the number of terms is increased, the greater the dependence on high-order terms; such terms increase the danger of instability away from the control points. MREs with a higher preponderance of low-order terms could be achieved with monomials that are local to a sector or quadrant.

The proposed partitioned variations on MREs are based on conjoined polynomials with C^1 continuity, that is to say they are continuous with continuous first-order partial derivatives throughout the region of interest. As shown in Figure 1, there are three ways that the U and V axes can be used to partition the region:

- across the V-axis, creating a north/south (N/S) split;
- across the U-axis, creating an east/west (E/W) split;
- across both axes, creating 4 quadrants (4Q).

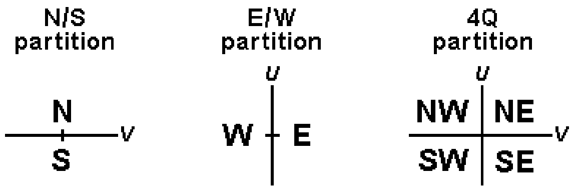

Figure 1: Partitions based on either or both of the U and V axes.

The procedures for the reverse transformation discussed in Subsections 2.1 and 2.2 are equally applicable to the partitioned MREs described in this Section.

3.1 N/S MREs

This variation on normalised MREs allows for different polynomials either side of the V-$axis$. The terms which are *not* divisible by U^2 are common to both polynomials. As the one-sided terms are divisible by U^2, they are zero along the line $U=0$ and so are their partial derivatives with respect to U. This ensures that the datum shift has C^1 continuity.

All N/S MREs can be expressed in the form (13) and (14) below. U and V are fully normalised intermediate coordinates (as before); $\left[U^iV^j\right]_N$ denotes a monomial only defined north of the V-axis; $\left[U^iV^j\right]_S$ denotes a monomial only defined south of the V-axis.

$$\Delta\phi('') = \sum_{i=0}^{1}\sum_{j=0}^{n} a_{i,j}U^iV^j + \sum_{i=2}^{n}\sum_{j=0}^{n} a_{i,j,N}\left[U^iV^j\right]_N + \sum_{i=2}^{n}\sum_{j=0}^{n} a_{i,j,S}\left[U^iV^j\right]_S \tag{13}$$

$$\Delta\lambda('') = \sum_{i=0}^{1}\sum_{j=0}^{n} b_{i,j}U^iV^j + \sum_{i=2}^{n}\sum_{j=0}^{n} b_{i,j,N}\left[U^iV^j\right]_N + \sum_{i=2}^{n}\sum_{j=0}^{n} b_{i,j,S}\left[U^iV^j\right]_S \tag{14}$$

The attraction of N/S MREs is that they allow more low-order terms than conventional MREs. There are $2n(n+1)$ possible terms for which $max(i,j)\leq n$, and this number is shown in Table 1 for different values of n.

3.2 E/W MREs

This variation on normalised MREs allows for different polynomials either side of the *U-axis*. The terms which are *not* divisible by V^2 are common to both polynomials. As the one-sided terms are divisible by V^2, they are zero along the line $V=0$ and so are their partial derivatives with respect to V. This ensures that the datum shift has C^1 continuity.

All E/W MREs can be expressed in the form (15) and (16) below. U and V are fully normalised intermediate coordinates (as before); $\left[U^iV^j\right]_E$ denotes a monomial only defined east of the *U-axis*; $\left[U^iV^j\right]_W$ denotes a monomial only defined west of the U-axis.

$$\Delta\phi('') = \sum_{i=0}^{n}\sum_{j=0}^{1} a_{i,j}U^iV^j + \sum_{i=0}^{n}\sum_{j=2}^{n} a_{i,j,E}\left[U^iV^j\right]_E + \sum_{i=0}^{n}\sum_{j=2}^{n} a_{i,j,W}\left[U^iV^j\right]_W \tag{15}$$

$$\Delta\lambda('') = \sum_{i=0}^{n}\sum_{j=0}^{1} b_{i,j}U^iV^j + \sum_{i=0}^{n}\sum_{j=2}^{n} b_{i,j,E}\left[U^iV^j\right]_E + \sum_{i=0}^{n}\sum_{j=2}^{n} b_{i,j,W}\left[U^iV^j\right]_W \tag{16}$$

The attraction of E/W MREs is that they allow more low-order terms than conventional MREs. As in the case for N/S MREs, there are $2n(n+1)$ possible terms for which $max(i,j)\leq n$, and this number is shown in Table 1 for different values of n.

3.3 4Q MREs

This variation on normalised MREs allows for different polynomials in the four quadrants formed by the *U-axis* and V-axis. The terms which are *not* divisible by U^2V^2 are common to all polynomials. As the quadrant-specific terms are divisible by U^2V^2, they are zero along the lines $U=0$ and $V=0$ and so are their partial derivatives. This ensures that the datum shift has C^1 continuity.

All 4Q MREs can be expressed in the form (17) and (18) below. U and V are fully normalised intermediate coordinates (as before); the quadrant subsipts, such as NE in $\left[U^iV^j\right]_{NE}$, indicate the quadrant in which the monomial is defined.

$$\Delta\phi('') = \sum_{i=0}^{1}\sum_{j=0}^{n} a_{i,j}U^iV^j + \sum_{i=2}^{n}\sum_{j=0}^{1} a_{i,j}U^iV^j + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{i,j,NE}\left[U^iV^j\right]_{NE} \tag{17}$$
$$+ \sum_{i=2}^{n}\sum_{j=2}^{n} a_{i,j,SE}\left[U^iV^j\right]_{SE} + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{i,j,SW}\left[U^iV^j\right]_{SW} + \sum_{i=2}^{n}\sum_{j=2}^{n} a_{i,j,NW}\left[U^iV^j\right]_{NW}$$

$$\Delta\lambda\left(''\right)=\sum_{i=0}^{1}\sum_{j=0}^{n}b_{i,j}U^iV^j+\sum_{i=2}^{n}\sum_{j=0}^{1}b_{i,j}U^iV^j+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{i,j,NE}\left[U^iV^j\right]_{NE}$$
$$+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{i,j,SE}\left[U^iV^j\right]_{SE}+\sum_{j=2}^{n}b_{i,j,SW}\left[U^iV^j\right]_{SW}+\sum_{i=2}^{n}\sum_{j=2}^{n}b_{i,j,NW}\left[U^iV^j\right]_{NW} \quad (18)$$

Again, the 4Q MRE allows more low-order terms than conventional MREs. There are $4(n^2-n+1)$ possible terms for which $max(i,j) \leq n$, and this number is shown in Table 1 for different values of n.

Table 1: Number of possible terms in each type of MRE for a given top power n.

Type of MRE	n=2	n=3	n=4	n=5	n=6
Conventional	9	16	25	36	49
N/W	12	24	40	60	84
E/W	12	24	40	60	84
4Q	12	28	52	84	124

4. Case Studies

The purpose of the following case studies is to find partitioned MREs that are more accurate and more stable than conventional MREs.

The process of applying statistical significance MREs is based on repeated removal of term for which AP/SE is smallest. The criteria for ending the process is less important than the need for consistency, to ensure the comparisons of different types of MRE are performed on a “level playing field”. For this study, the process continues until the minimum AP/SE is greater than 1. The method can be described as *Eliminating Ratios Less Than One* (ERLTO). The final number of terms is denoted by $N(\Delta\phi)$ or $N(\Delta\lambda)$ depending on whether the MRE represents $\Delta\phi$ or $\Delta\lambda$.

There is an obstacle to comparing the accuracy of a partitioned MRE with a conventional MRE with the same number of terms. It is possible that none of the conventional MREs actually has the same number of terms. To overcome this, four conventional MREs are obtained from polynomials with top powers 3, 4, 5 and 6 respectively. A smooth curve labelled “Conv” (for “conventional”) is drawn through the RMS residual values of the conventional MREs, as in Figures 2 to 4. Using that curve, the RMS residual value of each partitioned MRE can be compared with the interpolated conventional equivalent (from the “Conv” curve) corresponding to the number of terms.

4.1 Case study 1 (Great Britain)

The datum transformation considered here is from European Terrestrial Reference System 1989 (ETRS89) to Ordnance Survey Great Britain 1936 (OSGB36). The area of application is Great Britain and there are 4315 data points known in both datums. This dataset was split between 4011 control points and 304 test points. The latter were extracted from all round the region by an algorithm that ensured the squares they came from were only marginally depopulated.

It is noted in Section 6.2 of Ordnance Survey (2018) that the best similarity transformations is not very accurate with 2D errors of up to 3m at the 2σ level. In fact the dataset was created to cater for distortions by means of rubber sheeting leading to grid look-up (ibid, Section 6.3). It is none-the-less suitable for testing MREs and comparing the different types.

The latitude limits of the region of interest are 62.924° and 49.284°. The longitude limits are -10.799° and 4.863°. Applying (7) - (10) to (3) and (4), U and V are given by

$$U = (\phi - 56.104°) / 6.820° \tag{19}$$

$$V = (\lambda + 2.968°) / 7.831° \tag{20}$$

MREs of all 4 types were derived from the 4011 control points and are described in Table 2. Statistical significance was determined by the ERLTO process.

Table 2: Accuracy of conventional and partitioned MREs for ETRS89→OSGB36 datum shift, estimated from test points.

Type of MRE [top power = $max(i,j)$]	Abbreviated type	No of possible terms	N(Δϕ)	RMS in m of Δϕ residuals	N(Δλ)	RMS in m of Δλ residuals
Conventional, top power 6	Conv(6)	49	45	0.1287	47	0.1364
Conventional, top power 5	Conv(5)	36	35	0.1503	36	0.1722
Conventional, top power 4	Conv(4)	25	23	0.3006	24	0.1834
Conventional, top power 3	Conv(3)	16	16	0.3331	15	0.3374
N/S, top power 4	N/S(4)	40	40	0.1485	32	0.1427
N/S, top power 3	NS(3)	24	24	0.2306	22	0.1630
E/W, top power 4	E/W(4)	40	39	0.2482	37	0.1707
E/W, top power 3	E/W(3)	24	24	0.2730	24	0.2970
4Q, top power 4	4Q(4)	52	51	0.1786	49	0.1461
4Q, top power 3	4Q(3)	28	27	0.2620	26	0.2029

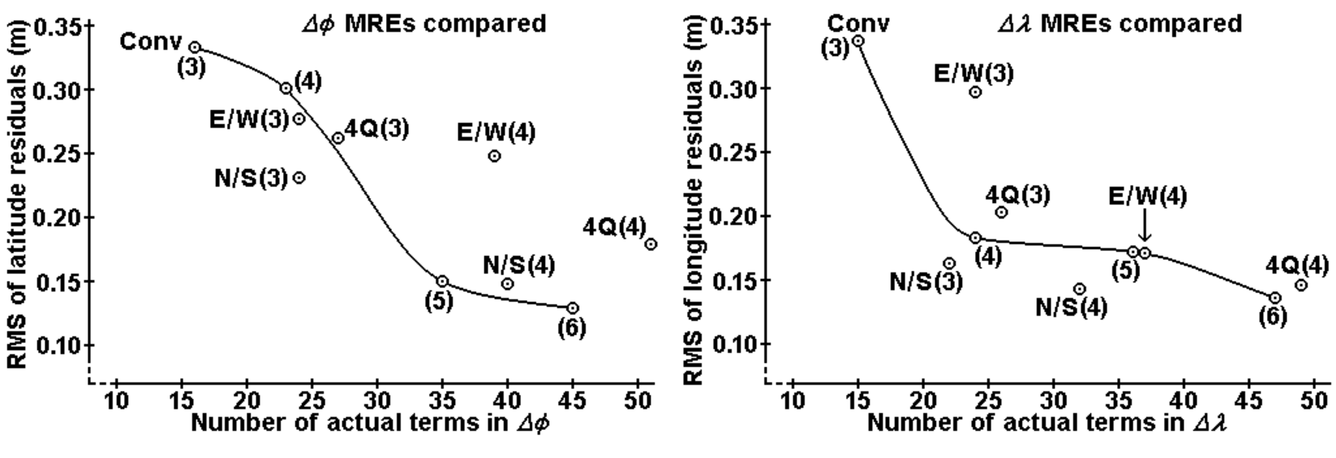

Figure 2: Partitioned-vs-conventional accuracy comparison for ETRS89→OSGB36datum shift.

Figure 2 shows that N/S is superior to E/W and 4Q for both $\Delta\phi$ and $\Delta\lambda$. This may be because the north-south extent of Great Britain is greater than its east-west extent. Three out of four of the N/S MREs are superior to interpolated "Conv". (Table 5 will provide further comparisons, both in terms of accuracy and simplicity for a given level of accuracy.)

4.2 Case study 2 (Slovenia)

The datum transformation considered here is from local datum D48 to D96, the latter being a locally-adopted name for European Terrestrial Reference System 1989. The area of application is Slovenia and there are 3331 data points known in both datums. This dataset was split between 3123 control points and 208 test points. The latter were extracted from all round the region with no two points close together.

The inadequacy of a 7-parameter transformation for the country as a whole is evident from Table 5 of ESRI (2012). Different transformations for D48 to ETRS89 covering different regions of Slovenia are listed, and there are wide variations in the parameters. The above dataset was created to cater for distortions by means of rubber sheeting. It is none-the-less suitable for testing MREs and comparing the different types.

The latitude limits of the region of interest are 46.879° and 45.421°. The longitude limits are 13.372° and 16.596°. Applying (7) - (10) to (3) and (4), *U* and *V* are given by

$$U = \left(\phi - 46.150°\right) / 0.729° \tag{21}$$

$$V = \left(\lambda - 14.984°\right) / 1.612° \tag{22}$$

MREs of all 4 types were derived from the 3123 control points and are described in Table 3. Statistical significance was determined by the ERLTO process.

Table 3: Accuracy of conventional and partitioned MREs for D48→D96 datum shift, estimated from test points.

Type of MRE [top power = *max(i,j)*]	Abbreviated type	No of possible terms	N($\Delta\phi$)	RMS in m of $\Delta\phi$ residuals	N($\Delta\lambda$)	RMS in m of $\Delta\lambda$ residuals
Conventional, top power 6	Conv(6)	49	46	0.0851	45	0.0788
Conventional, top power 5	Conv(5)	36	34	0.0893	34	0.0817
Conventional, top power 4	Conv(4)	25	25	0.0968	22	0.0896
Conventional, top power 3	Conv(3)	16	13	0.1022	14	0.0981
N/S, top power 4	N/S(4)	40	35	0.0901	39	0.0877
N/S, top power 3	NS(3)	24	24	0.0975	24	0.0931
E/W, top power 4	E/W(4)	40	34	0.0897	36	0.0778
E/W, top power 3	E/W(3)	24	22	0.0954	22	0.0882
4Q, top power 4	4Q(4)	52	48	0.0888	50	0.0780
4Q, top power 3	4Q(3)	28	27	0.0928	28	0.0883

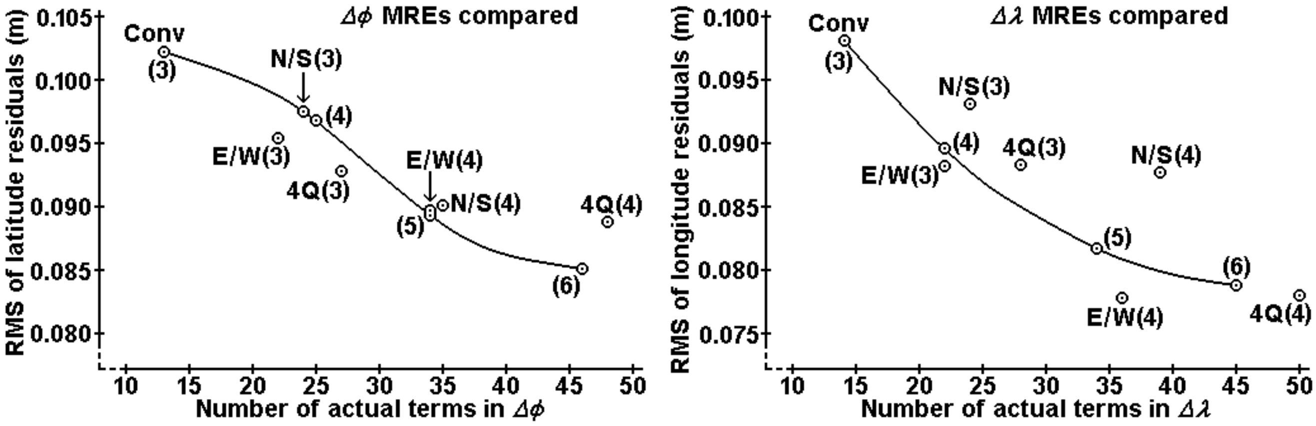

Figure 3: Partitioned-vs-conventional accuracy comparison for D48→D96 datum shift.

Figure 3 shows that E/W is superior to N/W and 4Q for both $\Delta\phi$ and $\Delta\lambda$. This may be because the east-west extent of Slovenia is greater than its north-south extent. Three out of four of the E/W MREs are superior to interpolated "Conv" and the other is a near-tie. (Table 5 will provide further comparisons, both in terms of accuracy and simplicity for a given level of accuracy.)

4.3 Case study 3 (Western Australia)

The datum transformation considered here is from Australian Geodetic Datum 1984 (AGD84) to the Geocentric Datum of Australia 1994 (GDA94). The area of application is Western Australia and the 82 data points of the STATEFIX GPS network are known in both datums. All were adopted as control points.

From the author's own derived transformations, RMS horizontal residuals were as follows:

- Standard Molodensky, 2.8943m (2.2917m for $\Delta\phi$, 1.7678 for $\Delta\lambda$);
- Helmert and Bursa-Wolf, 0.7142m (0.4452m, 0.5585m);
- 12-parameter affine, 0.6825m (0.4123m, 0.5439m).

The latitude limits of the region of interest are -13.595° and -35.105°. The longitude limits are 112.901° and 128.997°. Applying (7) - (10) to (3) and (4), U and V are given by

$$U = (\phi + 24.350°)/10.755° \quad (23)$$

$$V = (\lambda - 120.949°)/8.048° \quad (24)$$

The author's conclusion that there were not enough data points to set some aside as test points was supported by the presence of only 8 data points in the NW quadrant. This created a problem for the 4Q model with top power 4 because in general it has 9 NW-specific terms. It was resolved by removing the 3 highest-order terms $\left[U^4V^4\right]_{NW}$, $\left[U^3V^4\right]_{NW}$ and $\left[U^4V^3\right]_{NW}$ from both (17) and (18); in each instance, that created two redundancies in the quadrant.

MREs of all 4 types were derived from the 82 control points. Statistical significance was determined by the ERLTO process. It can be seen from Table 4 that for each of the "Conv" MREs, the same number of terms were retained for $\Delta\phi$ and $\Delta\lambda$; there is no particular reason for this, and the terms eliminated by ERLTO were not the same.

Because of the absence of test points, contour maps were applied to the Western Australia models to check that they were realistic. They showed that 4Q(4) for $\Delta\phi$ and Conv(6) for $\Delta\lambda$ give rise to volitivity in the north-eastern part of the State. They also showed extreme volitivity in the case of 4Q(4) for $\Delta\lambda$ along the northern coast between 114° and 120°. The RMSs for these models are daggered in Table 4 and blacked-out in Figure 4.

Table 4: Accuracy of conventional and partitioned MREs for AGD84→GDA94 datum shift, estimated from control points.

Type of MRE [top power = $max(i,j)$]	Abbreviated type	No of possible terms	N($\Delta\phi$)	RMS in m of $\Delta\phi$ residuals	N($\Delta\lambda$)	RMS in m of $\Delta\lambda$ residuals
Conventional, top power 6	Conv(6)	49	35	0.1638	35	0.1471†
Conventional, top power 5	Conv(5)	36	28	0.2002	28	0.1994
Conventional, top power 4	Conv(4)	25	18	0.2531	18	0.2539
Conventional, top power 3	Conv(3)	16	13	0.2780	13	0.3029
N/S, top power 4	N/S(4)	40	35	0.1728	30	0.2111
N/S, top power 3	NS(3)	24	18	0.2570	16	0.2585
E/W, top power 4	E/W(4)	40	33	0.1632	36	0.1604
E/W, top power 3	E/W(3)	24	19	0.2359	15	0.2869
4Q, top power 4	4Q(4)	49*	35	0.1465†	41	0.1118†
4Q, top power 3	4Q(3)	28	19	0.2152	22	0.2494

*Reduced from 52 by the exclusion of high-order terms in the NW quadrant.

†Achieved at the cost of volatility which made the model impractical.

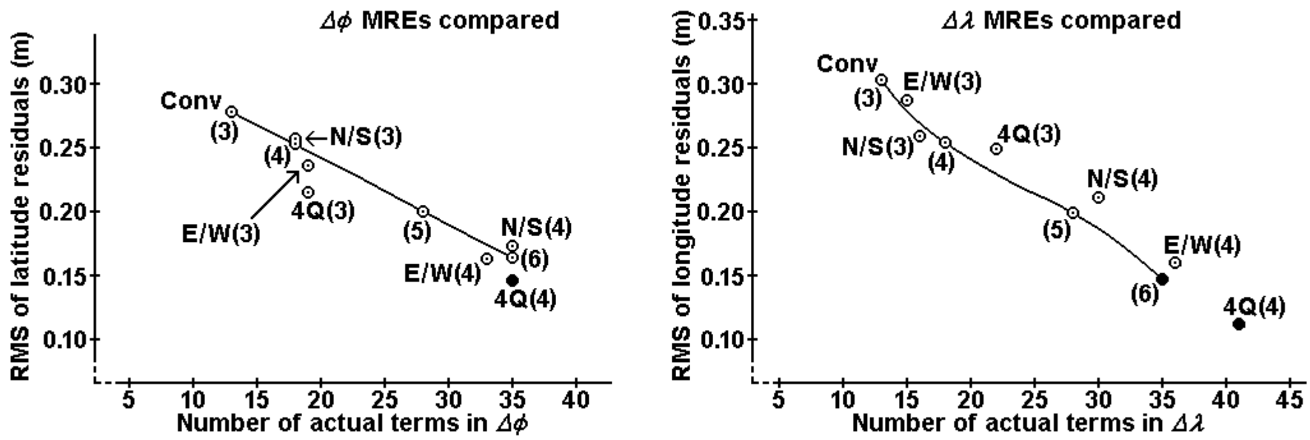

Figure 4: Partitioned-vs-conventional accuracy comparison for AGD84→GDA94 datum shift

Figure 4 shows that no single type stands out overall for Western Australia. For $\Delta\phi$, 4Q is superior to E/W, and both are superior to N/S and interpolated "Conv". (Table 5 will provide further comparisons, both in terms of accuracy and simplicity for a given level of accuracy.)

5. Conclusions

The models derived from the case studies are described in Tables 2 to 4 and illustrated in Figures 2 to 4. Table 5 brings together the accuracy comparisons in Figures 2 to 4, which show that in most of the comparisons the conventional MREs were outperformed by at least one of the partitioned MREs.

Table 5: Summary of MRE comparisons from the case studies.

	Best model for simplicity relative to a given level of accuracy		**Best model for accuracy disregarding impractical models**	
	$\Delta\phi$	$\Delta\lambda$	$\Delta\phi$	$\Delta\lambda$
Great Britain	N/S(3)	N/S(3)	Conv(6)	Conv(6)
Slovenia	E/W(3)	E/W(3)	Conv(6)	E/W(4)
Western Australia	4Q(3)	N/S(4)	E/W(4)	E/W(4)

For Great Britain, N/S proved better than the conventional MREs in terms of economy. Conv(6) was only more accurate than N/S(4) because of more terms, 15 more in the case of $\Delta\lambda$. E/W and 4Q were generally inferior to N/S and Conv.

For Slovenia, E/W proved generally better than the conventional MREs. Conv(6) was only more accurate than N/S(4) for $\Delta\phi$, and that was largely due to 9 more terms. Neither N/S and 4Q had a clear advantage over Conv.

For Western Australia the picture was mixed. For $\Delta\phi$ models with less than 25 terms, 4Q3 gave the best fit at the control points, but the 4Q4 models were too unstable to be practical. E/W4 was the most accurate of the models that were stable.

If multiple regression equations are the chosen tool for modelling a datum transformation, it makes sense to derive forms which allow more low-order terms and to apply them when they provide a better fit. The evidence so far indicates that N/S and E/W partitioning each have the potential to improve accuracy. Therefore they should be considered as alternatives to the conventional approach. The case for 4-quadrant partitioning is limited at present, but further investigation is recommended.

The partitioning proposed in this paper is worth considering in any application of MREs, since it retains C^1-continuity and provides a wider choice of low-order terms. This in turn makes MREs less reliant on high-order terms which introduce the risk of instability.

Future research into the use of regression for datum-shift models should not limit itself to the stepwise approach. Faraway (2004, Chapter 8) describes its limitations and expresses a strong preference for criterion-based procedures. These include the Akaite Information Criterion (AIC), the Bayes Information Criterion (BIC), adjusted R^2 (R_a^2) and Mallow's C_p statistic; see Faraway (2004, Section 8.3).

ACKNOWLEDGEMENT

The author thanks Professor Joseph Awange of Curtin University for providing control data for Western Australia, Sandi Berk of the Surveying and Mapping Authority of the Republic of Slovenia for providing control data for Slovenia, and Mark Greaves of Ordnance Survey for providing control data for Great Britain. The author also thanks the reviewers for their constructive comments.

REFERENCES

Abd-Elmotaal, H. 1994. Comparison of polynomial and similarity transformation based datum-shifts for Egypt. *Journal of Geodesy*, 68(3), pp.168-172.

Appelbaum, L.T. 1982. Geodetic datum transformation by multiple regression equations. *Proceedings of the Third International Geodetic Symposium on Satellite Doppler Positioning*. Las Cruces, New Mexico, pp.207-223.

Ayer, J., Boakye-Yiadom, E. and Smith-Ephraim, A. 2010. A Java Program for Direct Conversion between the Ghana National Grid and the Universal Transverse Mercator Projection Coordinates.

Berberich, E., Emeliyanenko, P. and Sagraloff, M. 2010. An Elimination Method for Solving Bivariate Polynomial Systems: Eliminating the Usual Drawbacks. Paper presented to SIAM Workshop on Algorithm Engineering & Experiments (ALANEX11), 22 January 2011, California USA.

Bomford, G. 1980. *Geodesy* (fourth edition). Oxford, Clarendon Press.

Chen, W. and Hill, C. 2005. Evaluation Procedure for Coordinate Transformation. *Journal of Surveying Engineering*, 131(2), pp.43-49.

Dawod, G. M. and Alnaggar, D. S. 2000. Optimum Geodetic Datum Transformation Techniques for GPS Surveys in Egypt. *Proceedings of Al-Azhar Engineering Sixth International Engineering Conference*, Al-Azhar University, 1-4 September, 4, pp.709-718.

DMA 1987a. *Supplement to Department of Defense World Geodetic System 1984 Technical Report: Part I – Methods, techniques, and data used in WGS 84 development*, DMA TR 8350.2-A, National Imagery and Mapping Agency, Washington, WA, USA.

DMA 1987b. *Supplement to Department of Defense World Geodetic System 1984 Technical Report: Part II – Parameters, formulas, and graphics for the practical application of WGS 84*, DMA TR 8350.2-B, National Imagery and Mapping Agency, Washington, WA, USA.

Draper, N. R. and Smith, H. 1998. *Applied regression analysis*. 3rd Edition. John Wiley and Sons, Inc., New York. 705p.

ESRI 2012. *ArcGIS 10.1 Geographic and Vertical Transformation Tables.* Resources.arcgis.com/en/help/main/10.1/003r/pdf/geographic_transformations.pdf

Faraway, J. J. 2004. *Linear Models With R*. Texts in Statistical Science. CRC Press. 255p.

Gao, Y. 2017. *Analysis of coordinate transformation with different polynomial models*. BSc thesis, Universität Stuttgart, Geodätisches Institut.

Gledan, J. A. and Azzaidani, O. A. 2014. ELD79-LGD2006 Transformation Techniques – Implementation and Accuracy Comparison in Tripoli Area, Libya. *International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering*, 8, pp.255-258.

Greaves, M. 2004. OSTN02: A new definition transformation from GPS derived coordinates to national grid coordinates in Great Britain. *Survey Review*, 37(293), pp.502-519.

Kutoglu, H.S. 2009. Alternative methods for improving Transformation Consistency between Geocentric and Non-Geocentric Coordinate Systems. *Survey Review*, 41(314), pp.408-418.

Mitsakaki, C., Agatza-Balodimou, A.M. and Papazissi, K. 2006. Geodetic reference frames transformations. *Survey Review*, 38(301), pp.608-618.

NIMA 2004. *Department of Defense World Geodetic System 1984: Its definition and relationships with local geodetic systems. 3rd ed.* Technical Report no. 8350.2, National Imagery and Mapping Agency, Washington, WA, USA. (Date on cover page is 2000, but 2004 is date of enclosed amendment.) Available at <https://gis-lab.info/docs/nima-tr8350.2-wgs84fin.pdf>

Ordnance Survey 2018. *A guide to coordinate systems in Great Britain*. Version 3.3.

Ruffhead, A. C. 2018. Introduction to multiple regression equations in datum transformations and their reversibility. *Survey Review*, 50(358) pp.82-90.

Ruffhead, A. C. and Whiting, B. M. 2020. *Introduction to geodetic datum transformations and their reversibility.* Surveying Working Paper No 01/2020, School of Architecture, Computing and Engineering, University of East London. 30p.

Soycan, M. 2005. Polynomial versus similarity transformations between GPS and Turkish reference systems. *Survey Review*, 38(295), pp.58-69.

Varga, M., Grgić, M. and Bašić, T. 2017. Empirical comparison of the geodetic coordinate transformation models: a case study of Croatia. *Survey Review*, 49(352), pp.15-27.