Derivation of rigorously-conformal 7-parameter 3D geodetic datum transformations

A. C. Ruffhead

School of Architecture, Computing and Engineering, University of East London, UK.
*Corresponding author, email ruffhead40b@yahoo.co.uk

This article was published in January 2021 in *Survey Review* 53 (376) 8-15.

http://dx.doi.org/10.1080/00396265.2019.1665614
DOI: 10.1080/00396265.2019.1665614

This is the accepted manuscript, posted on academic social networks on 23 September 2020, 12 months after Taylor & Francis had published the paper online.

ABSTRACT

This paper proposes a new method of deriving rigorously-conformal 7-parameter 3D coordinate transformations between geodetic datums. The problem of linearisation is reduced by distance analysis which provides an estimate of scale-change. The resulting 6-parameter transformation is linearised to enable an initial least-squares estimate of the rotation parameters. The 6-parameter transformation is then re-linearised to obtain a least-squares estimate of the corrections to the rotations. The validity of the scale-change estimate can be tested and is verified in almost all cases. The exception is transformations covering very small areas where short distances maximise the impact of measurement errors in the control data. Even there, the method can be adapted to optimise the transformation. The method can also be used to obtain pseudo-optimal conformal transformations that provide a closest fit to published Bursa-Wolf transformations.
Keywords: datum transformations, geodetic datums, conformal transformations, Helmert.

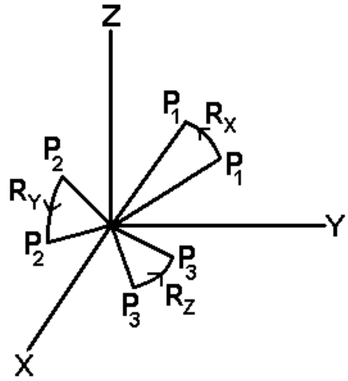
Introduction

Coordinate transformations between geodetic datums fall into 3 categories: conformal, near-conformal and non-conformal. Non-conformal transformations are useful when the intention is to model distortion. Conformal transformations are the ideal choice when the intention is to preserve the shape of the data being transformed. The most general form of conformal three-dimensional transformations are difficult to derive, however. One of the reasons for the popularity of near-conformal transformations is that they are easier to fit to control data using least-squares optimisation.

Conformal three-dimensional transformations involve smaller changes in size and orientation in the context of geodetic datums than in some other applications such as photogrammetry. The 7-parameter datum transformations in ESRI (2012) have no scale charge larger than 269 parts per million and no overall rotation greater than 90″. Even allowing for the fact that the tabulated parameters were designed for use in the near-conformal Bursa-Wolf transformation, limits of that magnitude can be assumed to apply to conformal transformations between geodetic datums.

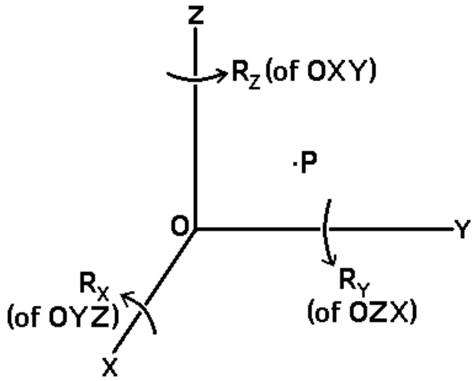
The 7-parameter conformal 3D transformation is generally applied to geocentric Cartesian coordinates and is sometimes known as a similarity translation. The parameters used are:

- Translation parameters ΔX, ΔY, ΔZ.
- Scaling parameter which may either be the scale factor S or the scale change ΔS. They are interchangeable, since S=1+ΔS. ΔS is often expressed in parts per million (ppm).
- Rotation parameters R_X, R_Y, R_Z. In this paper they describe rotation of position vectors about Cartesian axes, as illustrated in Fig. 1. In Europe it is the more commonly used convention. It is used by the International Association of Geodesy and is recommended by ISO (2007). One characteristic is that a positive rotation about the Z-axis has the effect of increasing longitude.



1 Position vector (PV) rotations

Some authors prefer the coordinate-frame rotation convention shown in Fig. 2 which has the opposite effect to those in Fig. 1. Coordinate-frame rotation parameters are therefore opposite in sign to the position-vector rotations used here.



2 Coordinate frame (CF) rotations

In this paper, the rigorous 7-parameter conformal transformation (7PC) is regarded as synonymous with the Helmert transformation. Sources that accept Helmert in that sense include Fan (2005), Hofmann-Wellenhof and Moritz (2006), Sjöberg (2013) and Watson (2005). The author disagrees with NATO (2001) which uses 'Helmert' to describe the simplified form associated with Bursa (1962) and Wolf (1963).

The combined effect of the rotation parameters is a rotation matrix **R**. The transformation equation which transforms the source coordinates (subscript s) to target coordinates (subscript t) is

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + (1+\Delta S)\mathbf{R} \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix}. \tag{1}$$

Helmert versions 1 and 2

The precise form of the rotation matrix **R** in equation (1) depends on the order in which the rotations are applied. There are six possible permutations but only two are used in practice.

One permutation is where R_X is applied to the position vector first and R_Z last. This is the permutation described by Harvey (1986, page 107) as 'the most commonly applied'. It is the permutation used in Dewitt (1996), Fan (2005), Harvey (1986), Hofmann-Wellenhof and Moritz (2006), Reit (1998), Varga *et al* (2017), Watson (2005), Wolf and Ghilani (1997). For the purposes of this paper, the 7PC transformation which uses this permutation of rotations is Helmert Version 1.

The other permutation is where R_Z is applied to the position vector first and R_X last, as in Awange and Grafarend (2002), Sjoberg (2013) and Wang *et al* (2018). For the purposes of this paper, the 7PC transformation which uses this permutation of rotations is Helmert Version 2.

For both possible forms of the rotation matrix, it is convenient to denote $\cos R_X$ by c_X, $\sin R_X$ by s_X, etc.

For Helmert Version 1,

$$\mathbf{R} = \begin{bmatrix} c_Z & -s_Z & 0 \\ s_Z & c_Z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_Y & 0 & s_Y \\ 0 & 1 & 0 \\ -s_Y & 0 & c_Y \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_X & -s_X \\ 0 & s_X & c_X \end{bmatrix}. \tag{2}$$

The Version-1 rotation matrix is therefore

$$\mathbf{R} = \begin{bmatrix} c_Y c_Z & s_X s_Y c_Z - c_X s_Z & s_X s_Z + c_X s_Y c_Z \\ c_Y s_Z & c_X c_Z + s_X s_Y s_Z & c_X s_Y s_Z - s_X c_Z \\ -s_Y & s_X c_Y & c_X c_Y \end{bmatrix}. \tag{3}$$

For Helmert Version 2,

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_X & -s_X \\ 0 & s_X & c_X \end{bmatrix} \begin{bmatrix} c_Y & 0 & s_Y \\ 0 & 1 & 0 \\ -s_Y & 0 & c_Y \end{bmatrix} \begin{bmatrix} c_Z & -s_Z & 0 \\ s_Z & c_Z & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

The Version-2 rotation matrix is therefore

$$\mathbf{R} = \begin{bmatrix} c_Y c_Z & -c_Y s_Z & s_Y \\ s_X s_Y c_Z + c_X s_Z & c_X c_Z - s_X s_Y s_Z & -s_X c_Y \\ s_X s_Z - c_X s_Y c_Z & c_X s_Y s_Z + s_X c_Z & c_X c_Y \end{bmatrix}. \tag{5}$$

The sine terms in equations (2) to (5) change sign if the rotation convention is based on coordinate frame rather than position vector.

The optimisation problem

Whether **R** satisfies (3) or (5), the resulting form of (1) is non-linear with respect to the 7 parameters. This makes it difficult to obtain the parameters that give a least-squares fit to a set of control points.

The following methods have been considered by the author.

- An iterative least-squares optimisation process. Xing Fang (2014) classifies one method as 'quasi indirect errors adjustment' (QIEA) using a quasi-Newton (Broyden-Fletcher-Goldfarb-Shanno) solution.
- The more direct method given by Awange and Grafarend (2002). An algebraic technique called 'Groebner basis' is used to solve the problem. Hashemi *et al* (2013) also describes a method 'using Gröbner bases techniques and solving polynomial systems'.
- A novel RANSAC robust estimation technique described in Paláncz *et al* (2017), which allows for possible outliers among the data points.
- Computation of approximate parameters from a small subset of the common points followed by linearization of the observations equations based on those approximate values. Least-squares optimisation is then applied to the linearised Helmert transformation model. The method is described and demonstrated in Fan (2005). The iterative phase is also described in Wolf & Ghilani (1997) which omits the issue of initial approximation.

None of these sources covers both of the commonly-used versions of Helmert. It is also unclear whether some of these methods actually derive the optimal Helmert parameters. Awange and Grafarend (2002) includes a worked example, but the derived transformation gives a worse fit to the control data than if Helmert had been applied using the optimal Bursa-Wolf parameters.

New optimisation method

The method devised by the author is believed to be original, although stage HO2 has some features in common with the iteration stage used by Fan (2005). In its general form, it consists of the following processes, each involving linear least-squares optimisation. 'HO' denotes Helmert optimisation.

- HO1: Distance analysis to obtain the original estimate of the scale change (ΔS_{DA}). In this case, only the possible impact of measurement errors might stop ΔS_{DA} being optimal because the 7PC transformation applies rotations exactly.
- HO2: Derivation of initial approximate translation and rotation parameters. This is done from what would be the partially-linear version of Bursa-Wolf except that the scale change has been already determined.
- HO3: Iteration using a relinearisation of Helmert (with scale change ΔS fixed) based on corrections to the approximate rotation parameters; the minimised root-mean-square distance residual corresponding to the estimated scale change is denoted MinRMS(ΔS,7PC).
- HO4: Verification of optimality by repeating HO2 and HO3 for small deviations from ΔS_{DA} .

Stage HO1

This is based on distances of the control points to the central point in both datums. The process is given by (6) to (10) below.

$$\left.\begin{aligned} X_{s,m} &= \frac{1}{n}\sum_{i=1}^{n} X_{s,i}, \quad Y_{s,m} = \frac{1}{n}\sum_{i=1}^{n} Y_{s,i}, \quad Z_{s,m} = \frac{1}{n}\sum_{i=1}^{n} Z_{s,i}, \\ X_{t,m} &= \frac{1}{n}\sum_{i=1}^{n} X_{t,i}, \quad Y_{t,m} = \frac{1}{n}\sum_{i=1}^{n} Y_{t,i}, \quad Z_{t,m} = \frac{1}{n}\sum_{i=1}^{n} Z_{t,i}. \end{aligned}\right\} \tag{6}$$

$$\left.\begin{aligned} d_{s,i} &= \sqrt{(X_{s,i} - X_{s,m})^2 + (Y_{s,i} - Y_{s,m})^2 + (Z_{s,i} - Z_{s,m})^2} \\ d_{t,i} &= \sqrt{(X_{t,i} - X_{t,m})^2 + (Y_{t,i} - Y_{t,m})^2 + (Z_{t,i} - Z_{t,m})^2} \end{aligned}\right\} \tag{7}$$

Treating S=1+ΔS as the parameter to be determined, the observation equations take the form

$$d_{t,i} = Sd_{s,i} + v_i \quad \text{or} \quad d_{s,i}\ S + v_i = d_{t,i}. \tag{8}$$

The normal equation takes the form

$$\left(\sum_{i=1}^{n} d_{s,i}^2\right) S = \sum_{i=1}^{n} d_{s,i} d_{t,i} \cdot \tag{9}$$

The least-squares solution, which determines ΔS as well as S, is therefore

$$S_{DA} = \sum_{i=1}^{n} d_{s,i} d_{t,i} \Big/ \sum_{i=1}^{n} d_{s,i}^2 \ . \tag{10}$$

Stage HO2

This stage optimises the translation and rotation parameters for the approximate value of S. Regarding S as fixed, the Helmert equation can be written in the form

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + \mathbf{R} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix} \cdot \tag{11}$$

An alternative form is

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} \\ r_{2,1} & r_{2,2} & r_{2,3} \\ r_{3,1} & r_{3,2} & r_{3,3} \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix} \cdot \tag{12}$$

The matrix elements $r_{i,j}$ in (12) are non-linear functions of the rotations R_X , R_Y and R_Z . This is true whether the rotation matrix takes the form (3) or (5).

In the context of geodetic datum transformations, rotations are generally smaller than 1° (considerably smaller, in fact). That makes their sines approximately their value in radians, their cosines approximately 1, and the products of their sines negligible. Whether the rotation matrix is given by (3) or (5), it can be approximated as follows.

$$\begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} \\ r_{2,1} & r_{2,2} & r_{2,3} \\ r_{3,1} & r_{3,2} & r_{3,3} \end{bmatrix} \cong \begin{bmatrix} 1 & -R_Z & R_Y \\ R_Z & 1 & -R_X \\ -R_Y & R_X & 1 \end{bmatrix} \cdot \tag{13}$$

Substituting (13) into (12),

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} \cong \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + \begin{bmatrix} 1 & -R_Z & R_Y \\ R_Z & 1 & -R_X \\ -R_Y & R_X & 1 \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix} \cdot \tag{14}$$

An equivalent formulation of (14) is

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} \cong \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & SZ_s & -SY_s \\ 0 & 1 & 0 & -SZ_s & 0 & SX_s \\ 0 & 0 & 1 & SY_s & -SX_s & 0 \end{bmatrix} \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \\ R_X \\ R_Y \\ R_Z \end{bmatrix} \cdot \tag{15}$$

The observation equations take the form

$$\begin{bmatrix} X_{t,i} - SX_{s,i} \\ Y_{t,i} - SY_{s,i} \\ Z_{t,i} - SZ_{s,i} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & SZ_{s,i} & -SY_{s,i} \\ 0 & 1 & 0 & -SZ_{s,i} & 0 & SX_{s,i} \\ 0 & 0 & 1 & SY_{s,i} & -SX_{s,i} & 0 \end{bmatrix} \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \\ R_X \\ R_Y \\ R_Z \end{bmatrix} + \begin{bmatrix} v_{X,i} \\ v_{Y,i} \\ v_{Z,i} \end{bmatrix} \cdot \tag{16}$$

Estimates of the six parameters are obtained by ordinary least-squares. $\bar{R}_X, \bar{R}_Y, \bar{R}_Z$ are the estimates that matter because they provide starting approximations to the rotation parameters.

Stage HO3

This stage uses iteration to optimise the translation and rotation parameters for the value of S used in Stage HO2. It starts with the estimates $\bar{R}_X, \bar{R}_Y, \bar{R}_Z$ obtained from stage HO2. The following substitution is made:

$$\begin{bmatrix} R_X \\ R_Y \\ R_Z \end{bmatrix} = \begin{bmatrix} \bar{R}_X \\ \bar{R}_Y \\ \bar{R}_Z \end{bmatrix} + \begin{bmatrix} \delta R_X \\ \delta R_Y \\ \delta R_Z \end{bmatrix}. \tag{17}$$

This contrasts with equation 52 in the iteration phase by Fan (2005) which had four 'delta' quantities to be optimised.

Applying (17) to the quantities in the rotation matrix **R**,

$$r_{i,j} = \bar{r}_{i,j} + \delta R_X(\partial r_{i,j} / \partial R_X) + \delta R_Y(\partial r_{i,j} / \partial R_Y) + \delta R_Z(\partial r_{i,j} / \partial R_Z). \tag{18}$$

The partial derivatives are evaluated for the rotations $\bar{R}_X, \bar{R}_Y, \bar{R}_Z$. Similarly, the overbar in $\bar{r}_{i,j}$ denotes $r_{i,j}$ evaluated for the rotations $\bar{R}_X, \bar{R}_Y, \bar{R}_Z$.

Substitution of (18) into (12) produces the following equation:

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + \begin{bmatrix} \bar{r}_{1,1} & \bar{r}_{1,2} & \bar{r}_{1,3} \\ \bar{r}_{2,1} & \bar{r}_{2,2} & \bar{r}_{2,3} \\ \bar{r}_{3,1} & \bar{r}_{3,2} & \bar{r}_{3,3} \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix} + \begin{bmatrix} m_{1,1} & m_{1,2} & m_{1,3} \\ m_{2,1} & m_{2,2} & m_{2,3} \\ m_{3,1} & m_{3,2} & m_{3,3} \end{bmatrix} \begin{bmatrix} \delta R_X \\ \delta R_Y \\ \delta R_Z \end{bmatrix}. \tag{19}$$

The terms $m_{i,j}$ in (19) can be evaluated from the following equations.

$$m_{i,1} = (\partial r_{i,1} / \partial R_X) SX_s + (\partial r_{i,2} / \partial R_X) SY_s + (\partial r_{i,3} / \partial R_X) SZ_s; \tag{20}$$

$$m_{i,2} = (\partial r_{i,1} / \partial R_Y) SX_s + (\partial r_{i,2} / \partial R_Y) SY_s + (\partial r_{i,3} / \partial R_Y) SZ_s; \tag{21}$$

$$m_{i,3} = (\partial r_{i,1} / \partial R_Z) SX_s + (\partial r_{i,2} / \partial R_Z) SY_s + (\partial r_{i,3} / \partial R_Z) SZ_s. \tag{22}$$

Equations (20) to (22) can be expressed in vector form:

$$\begin{bmatrix} m_{1,1} \\ m_{2,1} \\ m_{3,1} \end{bmatrix} = \begin{bmatrix} \partial r_{1,1} / \partial R_X & \partial r_{1,2} / \partial R_X & \partial r_{1,3} / \partial R_X \\ \partial r_{2,1} / \partial R_X & \partial r_{2,2} / \partial R_X & \partial r_{2,3} / \partial R_x \\ \partial r_{3,1} / \partial R_X & \partial r_{3,2} / \partial R_X & \partial r_{3,3} / \partial R_X \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix}. \tag{23}$$

$$\begin{bmatrix} m_{1,2} \\ m_{2,2} \\ m_{3,2} \end{bmatrix} = \begin{bmatrix} \partial r_{1,1} / \partial R_Y & \partial r_{1,2} / \partial R_Y & \partial r_{1,3} / \partial R_Y \\ \partial r_{2,1} / \partial R_Y & \partial r_{2,2} / \partial R_Y & \partial r_{2,3} / \partial R_Y \\ \partial r_{3,1} / \partial R_Y & \partial r_{3,2} / \partial R_Y & \partial r_{3,3} / \partial R_Y \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix}. \tag{24}$$

$$\begin{bmatrix} m_{1,3} \\ m_{2,3} \\ m_{3,3} \end{bmatrix} = \begin{bmatrix} \partial r_{1,1} / \partial R_Z & \partial r_{1,2} / \partial R_Z & \partial r_{1,3} / \partial R_Z \\ \partial r_{2,1} / \partial R_Z & \partial r_{2,2} / \partial R_Z & \partial r_{2,3} / \partial R_Z \\ \partial r_{3,1} / \partial R_Z & \partial r_{3,2} / \partial R_Z & \partial r_{3,3} / \partial R_Z \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix}. \tag{25}$$

Equation (19) can be expressed in terms of linear combinations of ΔX, ΔY, ΔZ, δR_X, δR_Y and δR_Z.

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} - \begin{bmatrix} \bar{r}_{1,1} & \bar{r}_{1,2} & \bar{r}_{1,3} \\ \bar{r}_{2,1} & \bar{r}_{2,2} & \bar{r}_{2,3} \\ \bar{r}_{3,1} & \bar{r}_{3,2} & \bar{r}_{3,3} \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & m_{1,1} & m_{1,2} & m_{1,3} \\ 0 & 1 & 0 & m_{2,1} & m_{2,2} & m_{2,3} \\ 0 & 0 & 1 & m_{3,1} & m_{3,2} & m_{3,3} \end{bmatrix} \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \\ \delta R_X \\ \delta R_Y \\ \delta R_Z \end{bmatrix}. \tag{26}$$

For the ith point, the observation equations take the form shown below.

$$\begin{bmatrix} X_{t,i} \\ Y_{t,i} \\ Z_{t,i} \end{bmatrix} - \begin{bmatrix} \bar{r}_{1,1} & \bar{r}_{1,2} & \bar{r}_{1,3} \\ \bar{r}_{2,1} & \bar{r}_{2,2} & \bar{r}_{2,3} \\ \bar{r}_{3,1} & \bar{r}_{3,2} & \bar{r}_{3,3} \end{bmatrix} \begin{bmatrix} SX_{s,i} \\ SY_{s,i} \\ SZ_{s,i} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & m_{1,1,i} & m_{1,2,i} & m_{1,3,i} \\ 0 & 1 & 0 & m_{2,1,i} & m_{2,2,i} & m_{2,3,i} \\ 0 & 0 & 1 & m_{3,1,i} & m_{3,2,i} & m_{3,3,i} \end{bmatrix} \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \\ \delta R_X \\ \delta R_Y \\ \delta R_Z \end{bmatrix} + \begin{bmatrix} v_{X,i} \\ v_{Y,i} \\ v_{Z,i} \end{bmatrix}. \tag{27}$$

Since this complies with the linear form $\mathbf{b} = \mathbf{Ax} + \mathbf{v}$, the least-squares estimate of the 6 unknowns can be obtained by solving $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$. The values of $\delta R_X, \delta R_Y, \delta R_Z$ are added to $\bar{R}_X, \bar{R}_Y, \bar{R}_Z$ to obtain new approximations to the rotation parameters. The iteration step is repeated until convergence.

All that remains is to provide explicit formulae for the matrix elements in equation (26).

Consider the case of Helmert Version 1, where the rotation matrix satisfies equation (3). Using $\bar{c}_X$ to denote $\cos\bar{R}_X$, $\bar{s}_X$ to denote $\sin\bar{R}_X$, etc, the definition of $\bar{r}_{i,j}$ ensures that

$$\begin{bmatrix} \bar{r}_{1,1} & \bar{r}_{1,2} & \bar{r}_{1,3} \\ \bar{r}_{2,1} & \bar{r}_{2,2} & \bar{r}_{2,3} \\ \bar{r}_{3,1} & \bar{r}_{3,2} & \bar{r}_{3,3} \end{bmatrix} = \begin{bmatrix} \bar{c}_Y\bar{c}_Z & \bar{s}_X\bar{s}_Y\bar{c}_Z - \bar{c}_X\bar{s}_Z & \bar{s}_X\bar{s}_Z + \bar{c}_X\bar{s}_Y\bar{c}_Z \\ \bar{c}_Y\bar{s}_Z & \bar{c}_X\bar{c}_Z + \bar{s}_X\bar{s}_Y\bar{s}_Z & \bar{c}_X\bar{s}_Y\bar{s}_Z - \bar{s}_X\bar{c}_Z \\ -\bar{s}_Y & \bar{s}_X\bar{c}_Y & \bar{c}_X\bar{c}_Y \end{bmatrix}. \tag{28}$$

From (23), (24) and (25),

$$\begin{bmatrix} m_{1,1} \\ m_{2,1} \\ m_{3,1} \end{bmatrix} = \begin{bmatrix} 0 & \bar{c}_X\bar{s}_Y\bar{c}_Z + \bar{s}_X\bar{s}_Z & \bar{c}_X\bar{s}_Z - \bar{s}_X\bar{s}_Y\bar{c}_Z \\ 0 & \bar{c}_X\bar{s}_Y\bar{s}_Z - \bar{s}_X\bar{c}_Z & -(\bar{s}_X\bar{s}_Y\bar{s}_Z + \bar{c}_X\bar{c}_Z) \\ 0 & \bar{c}_X\bar{c}_Y & -\bar{s}_X\bar{c}_Y \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix}; \tag{29}$$

$$\begin{bmatrix} m_{1,2} \\ m_{2,2} \\ m_{3,2} \end{bmatrix} = \begin{bmatrix} -\bar{s}_Y\bar{c}_Z & \bar{s}_X\bar{c}_Y\bar{c}_Z & \bar{c}_X\bar{c}_Y\bar{c}_Z \\ -\bar{s}_Y\bar{s}_Z & \bar{s}_X\bar{c}_Y\bar{s}_Z & \bar{c}_X\bar{c}_Y\bar{s}_Z \\ -\bar{c}_Y & -\bar{s}_X\bar{s}_Y & -\bar{c}_X\bar{s}_Y \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix}; \tag{30}$$

$$\begin{bmatrix} m_{1,3} \\ m_{2,3} \\ m_{3,3} \end{bmatrix} = \begin{bmatrix} -\bar{c}_Y\bar{s}_Z & -(\bar{s}_X\bar{s}_Y\bar{s}_Z + \bar{c}_X\bar{c}_Z) & \bar{s}_X\bar{c}_Z - \bar{c}_X\bar{s}_Y\bar{s}_Z \\ \bar{c}_Y\bar{c}_Z & \bar{s}_X\bar{s}_Y\bar{c}_Z - \bar{c}_X\bar{s}_Z & \bar{c}_X\bar{s}_Y\bar{c}_Z + \bar{s}_X\bar{s}_Z \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix}. \tag{31}$$

Consider the case of Helmert Version 2, where the rotation matrix satisfies equation (5). Using $\bar{c}_X$ to denote $\cos\bar{R}_X$, $\bar{s}_X$ to denote $\sin\bar{R}_X$, etc, the definition of $\bar{r}_{i,j}$ ensures that

$$\begin{bmatrix} \bar{r}_{1,1} & \bar{r}_{1,2} & \bar{r}_{1,3} \\ \bar{r}_{2,1} & \bar{r}_{2,2} & \bar{r}_{2,3} \\ \bar{r}_{3,1} & \bar{r}_{3,2} & \bar{r}_{3,3} \end{bmatrix} = \begin{bmatrix} \bar{c}_Y\bar{c}_Z & -\bar{c}_Y\bar{s}_Z & \bar{s}_Y \\ \bar{s}_X\bar{s}_Y\bar{c}_Z + \bar{c}_X\bar{s}_Z & \bar{c}_X\bar{c}_Z - \bar{s}_X\bar{s}_Y\bar{s}_Z & -\bar{s}_X\bar{c}_Y \\ \bar{s}_X\bar{s}_Z - \bar{c}_X\bar{s}_Y\bar{c}_Z & \bar{c}_X\bar{s}_Y\bar{s}_Z + \bar{s}_X\bar{c}_Z & \bar{c}_X\bar{c}_Y \end{bmatrix} \tag{32}$$

From (23), (24) and (25),

$$\begin{bmatrix} m_{1,1} \\ m_{2,1} \\ m_{3,1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ \bar{c}_X\bar{s}_Y\bar{c}_Z - \bar{s}_X\bar{s}_Z & -(\bar{s}_X\bar{c}_Z + \bar{c}_X\bar{s}_Y\bar{s}_Z) & -\bar{c}_X\bar{c}_Y \\ \bar{c}_X\bar{s}_Z + \bar{s}_X\bar{s}_Y\bar{c}_Z & \bar{c}_X\bar{c}_Z - \bar{s}_X\bar{s}_Y\bar{s}_Z & -\bar{s}_X\bar{c}_Y \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix}; \tag{33}$$

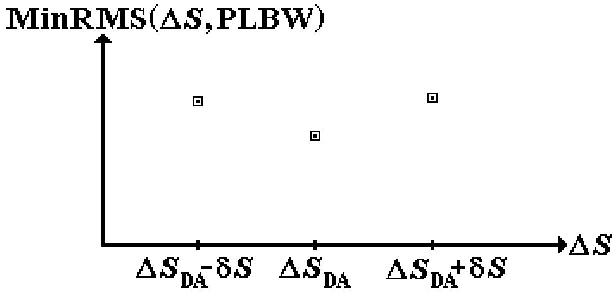
$$\begin{bmatrix} m_{1,2} \\ m_{2,2} \\ m_{3,2} \end{bmatrix} = \begin{bmatrix} -\bar{s}_Y\bar{c}_Z & \bar{s}_Y\bar{s}_Z & \bar{c}_Y \\ \bar{s}_X\bar{c}_Y\bar{c}_Z & -\bar{s}_X\bar{c}_Y\bar{s}_Z & \bar{s}_X\bar{s}_Y \\ -\bar{c}_X\bar{c}_Y\bar{c}_Z & \bar{c}_X\bar{c}_Y\bar{s}_Z & -\bar{c}_X\bar{s}_Y \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix}; \tag{34}$$

$$\begin{bmatrix} m_{1,3} \\ m_{2,3} \\ m_{3,3} \end{bmatrix} = \begin{bmatrix} -\bar{c}_Y\bar{s}_Z & -\bar{c}_Y\bar{c}_Z & 0 \\ \bar{c}_X\bar{c}_Z - \bar{s}_X\bar{s}_Y\bar{s}_Z & -(\bar{c}_X\bar{s}_Z + \bar{s}_X\bar{s}_Y\bar{c}_Z) & 0 \\ \bar{s}_X\bar{c}_Z + \bar{c}_X\bar{s}_Y\bar{s}_Z & \bar{c}_X\bar{s}_Y\bar{c}_Z - \bar{s}_X\bar{s}_Z & 0 \end{bmatrix} \begin{bmatrix} SX_s \\ SY_s \\ SZ_s \end{bmatrix}. \tag{35}$$

Stage HO4

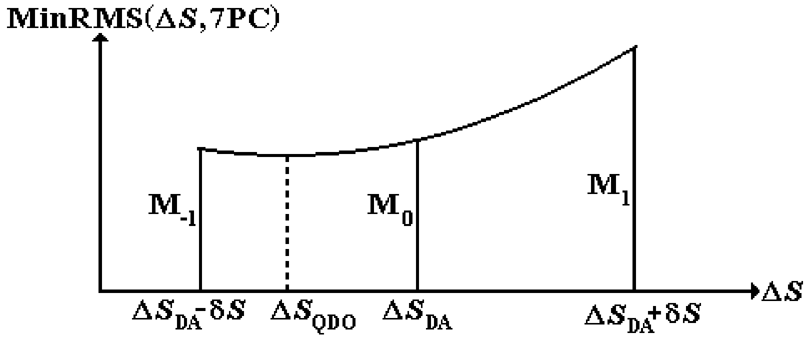
Stage HO4 involves repetition of the stages HO2 and HO3 for two slightly different scale changes, $\Delta S_{DA} - \delta S$ and $\Delta S_{DA} + \delta S$. The residuals enable the computation of three values of MinRMS(ΔS, 7PC). If the latter is evaluated to 8 or 9 decimal places, 0.01 ppm would be a suitable value for δS.

If MinRMS is less for the distance-analysis estimate of ΔS than for the two neighbouring estimates, and the outer values of MinRMS are closer to each other than to the central value, then the distance-analysis estimate of ΔS can be regarded as optimal along with the corresponding translation and rotation parameters obtained by stages HO2 and HO3. The condition is illustrated in Fig. 3, in which MinRMS is near-symmetric about its central value.



3 The pattern of minimum RMS distance residuals from the 7PC transformation that indicates that the ΔS obtained from distance analysis is the optimum value

If the near-symmetry condition is not satisfied, than a quadratic curve can be fitted to the three values of (ΔS, MinRMS), as in Fig. 4. The values of MinRMS are denoted by M_{-1}, M_0 and M_1.



4 The quadratically-derived optimal ΔS based on three minimum RMS distance residuals from the 7PC transformation

The value of ΔS that matches the minimum point on the curve is a quadratically-derived optimal value and can be denoted ΔS_{QDO} . The curve does not need to be plotted, since ΔS_{QDO} can be computed from

$$\Delta S_{QDO} = \Delta S_{DA} + \frac{(M_{-1} - M_1)\delta S}{2(M_{-1} + M_1 - 2M_0)} . \tag{36}$$

Stages HO2 and HO3 are performed a 4th time to obtain the remaining parameters.

The use of formula (36) assumes that MinRMS behaves like a quadratic function around the optimal value of ΔS. All the tests carried out by the author confirm this. To give one example, for the dataset of 44 points in Great Britain common to OSGB 36 and WGS 84, MinRMS had a minimum value (2.519810834 m) when ΔS was −20.6863 ppm. For i in the range −8 to 8, MinRMS(ΔS, 7PC) was almost indistinguishable from $(2.519642890 + 0.0000026244\ i^2)$ m when $\Delta S = (-20.6863 \pm i/100)$ ppm.

Practical implementation

The new optimisation method was applied to three actual datasets to obtain the optimal Helmert Version 1 transformations.

- OSGB 36 to WGS 84 from 44 common points in Great Britain. The optimal parameters were 445.18103 m, −161.83410 m, 542.61595 m, −20.68629 ppm, −0.73244″, 0.27901″, 1.60776″. The RMS distance residual was 2.519643 m.
- AGD 84 to WGS 84 from 82 common points in Western Australia. The optimal parameters were −115.83771 m, −48.37321 m, 144.75955 m, 3.68981 ppm, 0.11971″, 0.38399″, 0.37040″. The RMS distance residual was 0.758323 m.
- Accra datum to WGS 84 from 19 points in Ghana's 'Golden Triangle'. The optimal parameters were −151.19021 m, 31.59316 m, 327.17659 m, −7.16773 ppm, −0.44518″, 0.00582″, −0.02200″. The RMS distance residual was 0.961925 m.

The method was also applied to 18 simulated datasets generated by relatively large rotations and scaling, with pseudorandom numbers used to represent measurement errors.

The first significant finding from these tests (actual and simulated) was that stage HO3 only needed a single iteration. The second iteration merely confirmed convergence. On this evidence, the re-linearisation only needs to be applied once.

The second significant finding was that the 'distance-analysis' scale-change proved to be optimal, except in the case of three simulated datasets for which the area of coverage was significantly smaller than a degree square.

In each of the exceptional cases, stage HO4 showed that the optimal ΔS differed from ΔS_{DA} by 0.04 ppm. Halving the simulated measurement errors reduced the difference to 0.01 ppm. This suggested that small areas increase the size of measurement errors relative to the inter-point distances, undermining the validity of ΔS_{DA} as a scale-change estimate (albeit slightly). This explanation was confirmed when doubling of the measurement errors and doubling of all inter-point distances had a totally neutral effect on the difference between ΔS_{DA} and the optimal ΔS.

One distinguishing feature of the three exceptional cases was that the standard errors of ΔS_{DA} were between 63.663 ppm and 63.669 ppm. This was far higher than in any of the other cases (the next highest being 12.619 ppm). The likely cause is the smallness of the area of coverage, accentuating the relative effect of measurement errors. It is apparent that a small standard error is a good indication that ΔS_{DA} is optimal. In the case of the three actual datasets, the standard errors of ΔS_{DA} were 0.700 ppm, 0.064 ppm and 1.947 ppm respectively.

In practice, therefore, the Helmert optimisation process consists of:

- HO1: Distance analysis to obtain the original estimate of the scale change (ΔS_{DA}).
- HO2: Derivation of initial approximate translation and rotation parameters. This is done by applying least-squares to the observation equations (16).
- HO3: Deriving improved translation and rotation parameters using a relinearisation of Helmert (with scale change ΔS fixed) based on corrections to the approximate rotation parameters; further iteration will be unnecessary (except possibly when testing the software). Least-squares optimisation is applied to the observation equations (27).
- HO4: Verification of optimality by repeating HO2 and HO3 for small deviations from ΔS_{DA} . This is only necessary when testing the software for stages HO1 to HO3, or in cases where all control points are within 150 kilometres of each other. The use of (36) will complete the optimisation of ΔS if necessary.

Conversion of optimal Bursa-Wolf to pseudo-optimal Helmert

The new method of deriving Helmert transformations can be used to obtain the closest-fitting fully-conformal transformation to the near-conformal Bursa-Wolf transformation. Provided the area of application is known, this can be done without knowing the control points from which the Bursa-Wolf parameters were derived.

There are two commonly-quoted versions of the Bursa-Wolf transformation, even if the position-vector rotation convention is the only one considered. The partially linear form is

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + (1+\Delta S)\begin{bmatrix} 1 & -R_Z & R_Y \\ R_Z & 1 & -R_X \\ -R_Y & R_X & 1 \end{bmatrix}\begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix}. \tag{37}$$

The fully-linear version of Bursa-Wolf is

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} + \begin{bmatrix} 1+\Delta S & -R_Z & R_Y \\ R_Z & 1+\Delta S & -R_X \\ -R_Y & R_X & 1+\Delta S \end{bmatrix}\begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix}. \tag{38}$$

The latter form leads directly to observation equations which are linear with respect to ΔX, ΔY, ΔZ, ΔS, R_X , R_Y and R_Z . It is therefore highly probable that published parameters for Bursa-Wolf transformations have been obtained from fully-linear Bursa-Wolf. This is assumed to be the case for the ArcGIS parameters listed in Table 5 of ESRI (2012), even though ESRI (2008) recommends the use of the partially-linear form. The difference between the two forms depends on whether the products ΔSR_X , ΔSR_Y and ΔSR_Z are small enough to be negligible.

It is not possible to derive the optimal Helmert (Version 1) transformation purely from the optimised Bursa-Wolf transformation. However, it is possible to derive a close-fitting 'pseudo-optimal' Helmert transformation by

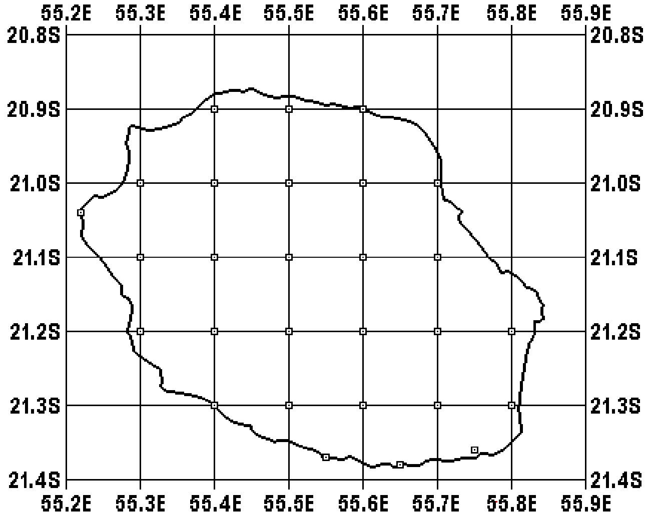
- applying the totally-linear Bursa-Wolf transformation to 25-50 points that give good coverage of the geographical region in question, and
- deriving the Helmert transformation which gives the best fit at these virtual data points. This can be done by the stages HO1, HO2 and HO3 (without further iteration) described earlier.

The reason why stage HO4 should not be necessary is that the virtual data points do not have measurement errors. Residuals from the optimisation will represent the difference between Helmert optimisation and Bursa-Wolf optimisation.

If the scale change and rotations of a Bursa-Wolf transformation are small, then the difference between that transformation and the closest-fitting Helmert transformation is likely to be negligible. Conversely the difference between Helmert and Bursa-Wolf is most likely to be significant when the scale change and at least one rotation are large. For this reason, the author selected the following Bursa-Wolf transformations from Table 5 of ESRI (2012):

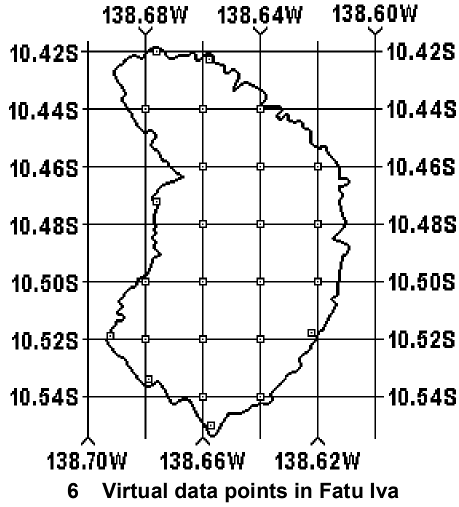
- Réunion 1947 to RGR 1992. The respective ellipsoids are the International Ellipsoid and GRS 80. The Bursa Wolf parameters in ESRI (2012) are 789.524 m, −626.486 m, −89.904 m, −32.3241 ppm, 0.6006″, 76.7946″, −10.5788″.
- Fatu Iva 1972 to WGS 84. The local datum uses the International Ellipsoid. The Bursa Wolf parameters in ESRI (2012), with the rotations converted to position-vector rotations, are 347.103 m, 1078.125 m, 2623.922 m, 186.074 ppm, −33.8875″, 70.6773″, −9.3943″.

28 virtual data points were selected for Réunion Island as shown in Fig. 5. Apart from 4 points from the west and south coast, all are taken from 6′ intersections. Height values between 10 m and 20 m were assigned to the points. RGR 1992 coordinates (Cartesian, then geodetic) were estimated by applying (38) with the ArcGIS parameters.



5 Virtual data points in Réunion Island

25 virtual data points were selected for Fatu Iva as shown in Fig. 6. Apart from 7 points in coastal regions, all are taken from 0.02° intersections. Height values between 20 m and 740 m were assigned to the points. WGS 84 coordinates (Cartesian, then geodetic) were estimated by applying (38) with the ArcGIS parameters.



6 Virtual data points in Fatu Iva

In each case, the method of deriving the Helmert Version 1 transformation was by stages HO1 to HO3. Stage HO4 was only used for checking purposes and it confirmed the optimality of ΔS_{DA} .

The Helmert Version 1 parameters which gave the closest fit were as follows.

- 789.70880 m, −626.93585 m, −89.93390 m, −32.26312 ppm, 0.60127″, 76.79736″, −10.57263″ for Réunion 1947 to RGR 1992. The RMS distance residual was 0.000255 m.
- 346.90967 m, 1078.23235 m, 2623.87087 m, 186.13000 ppm, −33.88457″, 70.66260″, −9.39541″ for Fatu Iva 1972 to WGS 84. The RMS distance residual was 0.000140 m.

The submillimetre RMS distance residuals are an indication of proximity to the optimal Bursa-Wolf transformation. They suggest that whatever the quality of fit of Bursa-Wolf is to the (unavailable) control points, the quality of fit of Helmert Version 1 is only slightly inferior. Basically, an optimal near-conformal transformation has been converted into a near-optimal conformal transformation.

This process is only appropriate if some of the parameters ΔS, R_X, R_Y and R_Z are relatively large (the two examples being extreme cases). If they are all small then Bursa-Wolf is as close to being conformal as to make no difference.

Conclusions

This proposed new method of deriving rigorously-conformal 7-parameter 3D coordinate transformations is suitable for any pair of geodetic datums. The use of distance analysis (stage HO1) provides an estimate of scale-change which only needs correction in very small areas where measurement errors in the data have proportionally the greatest impact. The initial approximation of the rotation parameters (stage HO2) is sufficiently close to the optimal values to ensure that the re-linearisation phase (stage HO3) optimises the 6 size-preserving parameters in a single step.

The verification process (stage HO4) is needed in very rare cases. The scale change from the distance analysis and two variations on that scale change are used to compute a quadratically-derived optimum.

The proposed method is applicable to either version of the Helmert transformation and its computational implementation is straightforward.

The proposed method can also be used to obtain pseudo-optimal conformal transformations that provide a closest fit to published Bursa-Wolf transformations.

Acknowledgements

The author thanks his Director of Studies Dr Brian Whiting for his comments and suggestions. The author also thanks Professor Joseph Awange of Curtin University for providing control data for Western Australia, and Ordnance Survey for providing control data for Great Britain. The author is grateful to the two anonymous reviewers for their comments and suggestions.

References

Awange, L. J. and Grafarend, W. E., 2002. Linearized Least Squares and nonlinear Gauss-Jacobi combinatorial algorithm applied to the 7-parameter datum transformation $\mathbb{C}_7(3)$ problem. *ZFV* 127 (2), 109-116.

Bursa, M., 1962. The Theory of the Determination of the Non-parallelism of the Minor Axis of the Reference Ellipsoid and the Inertial Polar Axis of the Earth, and the planes of the Initial Astronomic and Geodetic Meridians from Observation of Artificial Earth Satellites. *Studia Geophysica et Geodetica*, 6 (3), 209-214.

Dewitt, B., 1996. Initial approximation for the three-dimensional conformal coordinate transformation. *Photogrammetric Engineering and Remote Sensing*, 62, 79–83.

ESRI, 2008. ArcGIS Desktop 9.3 Help: Equation-based methods. Available from: http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Equation-based%20methods

ESRI, 2012. ArcGIS 10.1 Geographic and Vertical Transformation Tables. Available from: resources.arcgis.com/en/help/main/10.1/003r/pdf/geographic_transformations.pdf

Fan, H., 2005. Three-Dimensional Coordinate Transformations with large rotations and scale change. International Workshop on Education in Geospatial Information Technology, 27-28 October, Technical University of Moldova, Chisinau, Moldova.

Fang, X. 2014. A total least squares solution for geodetic datum transformations. *Acta Geodaetica et Geophysica*, 49, 189-207.

Harvey, B. R., 1986. Transformation of 3D Coordinates. *The Australian Surveyor*, 33 (2), 105-125.

Hashemi, A., Kalantari, M. and Kasser, M., 2013. Direct solution of the 7 parameters transformation problem. *Applied Mathematics & Information Sciences*, 7 (4), 1375-1382.

Hofmann-Wellenhof, B, & Moritz, M., 2006. *Physical Geodesy (Second Edition)*. SpringerWienNewYork.

ISO, 2007. *ISO 19111, Geographical information – Spatial referencing by coordinates*, Second Edition, International Organization for Standardization, Geneva, Switzerland.

NATO, 2001. Standardization Agreement (STANAG) 2211 IGEO, *Geodetic Datums, Projections, Grids and Grid References*, Edition 6, NATO Military Agency for Standardization, Brussels.

Paláncz, B., Awange, J. and Völgyesi, L., 2017. A novel RANSAC approach to robustly solve the 3D similarity problem. *Australian Journal of Earth Sciences*. Published online: 08 May 2017. 12 pages.

Reit, B.-G., 1998. The 7-parameter transformation to a horizontal geodetic datum. *Survey Review*, 34 (268), 400-404.

Sjöberg, L. E., 2013. Closed-form and iterative weighted least squares solutions of Helmert transformation parameters. *Journal of Geodetic Science* 3 (1), 7-11.

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), 15-27.

Wang, Q., Chang, G., Xu, T. and Zou, Y., 2018. Representation of the rotation parameter estimation errors in the Helmert transformation model. *Survey Review*, 50 (358), 69-81.

Watson, G. A., 2005. Computing Helmert transformations. *Journal of Computational and Applied Mathematics*, 197 (2), 387-394.

Wolf, H., 1963. Geometric connection and re-orientation of three-dimensional triangulation nets. Bulletin Géodésique, 68, 165-169.

Wolf, P. R. and Ghalani, C. D., 1997. *Adjustment Computations: Statistics and Least-Squares in Surveying and GIS*. 3rd Edition, John Wiley & Sons, Hoboken, New Jersey, 347-353.