Transformation of 3D Co-ordinates

BRUCE R. HARVEY

Abstract

As satellite positioning techniques are more frequently used, transformations of three-dimensional co-ordinates will become common practice. Methods of determining and applying transformation parameters for transforming 3D co-ordinates are discussed. The models, appropriate statistical tests and interpretation of the results will also be considered. The methods presented are applicable to many transformations required by surveyors in Australia. Some applications are the conversion of Doppler or GPS co-ordinates into the terrestrial networks and the comparison of Doppler, GPS or terrestrial data with SLR or VLBI data. In this paper particular emphasis will be placed on transformations between terrestrial and GPS data.

Introduction

With the advent of GPS, the transformation of 3D co-ordinates will become a common and frequent computation. Thus a clear understanding of the procedures is essential. Much has been published on this topic in the past. However some incorrect and misleading statements have been made. This paper, which is based on research in Harvey (1985), aims to clarify some misconceptions and present further comments for consideration.

The theory presented in this paper can be applied to many different techniques. However a common application which is addressed here is to transform GPS co-ordinates into local survey co-ordinates e.g. AGD 84. (Points whose geodetic co-ordinates in the survey network are known are referred to as ground points in this paper.) In this case GPS instruments occupy a number of known ground points and a number of other points of interest with unknown ground co-ordinates. The transformation parameters are then determined from the data of those points which have known co-ordinates in both systems i.e. common points. These transformation parameters are then applied to convert the GPS co-ordinates of the other points into corresponding ground co-ordinates.

If correct procedures are followed, the combined adjustment of two independent data sets will improve — (i) the accuracy of a network by controlling the systematic errors, and (ii) the precision of the network because additional data is included. If systematic errors in scale and orientation exist within a geodetic network then not only will the co-ordinates be incorrect, but their estimated accuracies will be optimistic. Therefore the combination of independent data also produces more realistic accuracy estimates.

Two main problems arise when combining data from different networks. They are: 1) Finding the most accurate estimates of the external bias parameters (i.e. scale, rotation, translation and systematic error terms) between the systems, and 2) achieving the best internal combination of the different systems, that is to minimise the corrections to the observations.

GPS and ground networks are at nearly the same scale and are usually oriented

BRUCE R. HARVEY, B.Surv., Ph.D. (NSW), Registered Surveyor (NSW), Rothmans Fellow, School of Surveying, University of New South Wales.

to within a few arc seconds of the same reference co-ordinate system (the Conventional Terrestrial System, CTS), i.e. the CIO, geocentre, Greenwich frame (see e.g. Vanicek and Krakiwsky, 1982). Since both the ground and GPS systems are aligned with the instantaneous rotation pole and corrections are applied to align them with the CTS frame, the rotation angles should be near zero. If the GPS and ground co-ordinates were truly related to the same reference system, they could be more easily combined. However, even though both GPS and ground co-ordinates are supposedly referred to the same CTS frame, there usually are significant differences. The differences are due to systematic errors and errors in the relationship between non-geocentric ground datums and the CTS. In theory, a series of rotation, scale and translation matrices could be applied to each net to represent systematic errors and the differences between reference frames. The elements of these matrices could then be estimated. However, some parameters will be indistinguishable from others. So it may not be possible to solve for them separately. Consequently only single rotation, scale, and translation matrices are estimated.

Transformations

There are a number of ways of transforming one data set to another, An affine transformation transforms straight lines to straight lines and parallel lines remain parallel. Generally the size, shape, position, and orientation of lines in a network are changed. The scale factor depends on the orientation of the line but not on its position within the net. Hence the lengths of all lines in a certain direction are multiplied by the same scalar. For a transformation between two large networks with many common points it may be possible to use a projection transformation where the scale factor is also a function of position. However this leads to a solution with few (if any) degrees of freedom when the data sets contain only a few points. Transformations in which the scale factor does not depend on position within the net, involve fewer parameters but assume that there are no systematic distortions of the networks.

An affine transformation in which the scale factor is the same in all directions is called a similarity transformation. A similarity transformation preserves shape, so angles are not changed, but the lengths of lines and the position of points may be changed. An orthogonal transformation is a similarity transformation in which the scale factor is unity. The angles and distances within the network are preserved and only the positions of points change on transformation.

The general similarity transformation is given by —

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = s\,R \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} Tx \\ Ty \\ Tz \end{bmatrix} \qquad (1)$$

where s is the scale factor and R is a 3 x 3 orthogonal rotation matrix. There are seven parameters which are usually associated with three rotation angles, three translation components and one scale factor. The translation terms (Tx, Ty, Tz) are the co-ordinates of the origin of the xyz net in the frame of the XYZ net. Scale factors and rotation matrices are described below.

If the rotations are small, as expected, then the above equation is approximately linear. A single iteration of a least squares adjustment is then generally sufficient. However, if additional iterations are required, because of poor *a priori* estimates of

the parameters, the convergence is usually very rapid.

It is presumptuous to assume that similarity transformations, rather than affine or projection transformations, correctly describe the differences between any two data sets. However in practice little more than the smoothing of these differences can be accomplished. Moreover, using fewer parameters would seem to be safer when there are few common points. Similarity transformations guard against undue deformations in small regions, but using a similarity transformation on a large network may distort local scale and orientation. Hence the important question is whether local distortions in scale and orientation are significant or not. Similarity transformations will tend to smooth them out. Also, because the similarity transformation is linear in the co-ordinates, it cannot be applied to estimate, or to allow for, any distortions or deformations that are common to the two nets.

In a transformation adjustment the mathematical aspects of the model, statistical analysis, computational errors, and the reliability of the adjustment results, that is, the stability of the solution, all have to be considered. In a stable solution small changes in the observations give small changes in the results. Error analysis associated with least squares adjustment requires *a priori* assumptions or hypotheses which require *a posteriori* testing. The results may also be affected by mathematical approximations and computational errors (e.g. round-off errors).

Rotation Matrices

The rotation matrices about the x, y, and z axes are given by —

$$Rz(k) = \begin{bmatrix} \cos k & \sin k & 0 \\ -\sin k & \cos k & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad Ry(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \quad Rx(\omega) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{bmatrix} \qquad (2)$$

In general, successive rotations of a body about fixed axes are not commutative. Thus the order of the rotations is important unless the rotations are small.

There are at least three ways of rotating a network, Harvey (1985). However, the most commonly applied is the Cardanian rotation matrix (Fig. 1) — Rz(k) Ry(ϕ) Rx(ω).

For small rotations the Cardanian rotation matrix can be approximated by —

$$Rs = \begin{bmatrix} 1 & k & -\theta \\ -k & 1 & \omega \\ \theta & -\omega & 1 \end{bmatrix} \qquad (3)$$

where ω, ϕ and k are the rotation angles about the x, y and z axes respectively.

The assumption of small angles is valid for rotation angles of up to about 3″. However, considerably larger angles can be tolerated if the vectors being rotated are shorter than an earth radius. For example, rotating a 500 km baseline vector with each rotation angle equal to 10″ will cause only 1 mm error in co-ordinates if Rs is used instead of Rz(k) Ry(θ) Rx(ω).

The rotation angles depend on the baseline vectors (i.e. relative positions) and not on the absolute co-ordinates. Thus it does not matter where the origin of co-ordinates is because the estimated rotation angles will be the same, provided care is taken to avoid round-off errors.

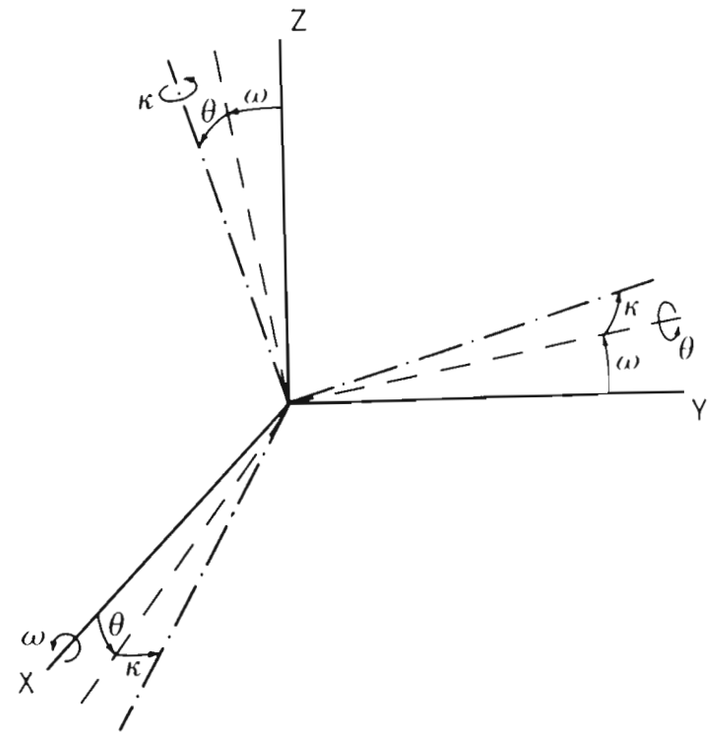

Figure 1
Cardanian Rotations. (Rz(k) Ry(θ Rx(ω))

Scale Factor

A scale factor can be visualised as follows. Imagine a network is drawn on the surface of an inflatable sphere. As the sphere is inflated, the points of the network will be spread apart from each other and from the centre of the sphere. The inflation of the sphere is equivalent to the application of a scale factor greater than unity.

Multiplication of a set of rectangular Cartesian co-ordinates by a scale factor is identical to multiplying the corresponding baseline lengths by the same scale factor. So the scale factor can be determined from either the 3D site co-ordinates or from the baseline lengths. Thus, as in the case of rotation angles, the origin of the co-ordinates has no effect on the results.

In ellipsoidal co-ordinates the longitude is not affected by a scale change but the geodetic latitude is changed slightly. For example, a 1 ppm scale change will change gedodetic latitude by less than about 0.0007″ (2 cm). However the effect on ellipsoidal height is significant. For example, a 1 ppm scale change will produce a change in height of about 6.4 m.

Variance Covariance Matrices of Co-ordinates

In a least squares adjustment the stochastic model must be correct as well as the mathematical model. As always, there is the problem of "internal" and "external" accuracy. The elements of the variance-covariance (VCV) matrices of the data may represent precision estimates. It is difficult to change these values to accuracies, especially the off-diagonal terms. The VCV matrices could be multiplied by factors to account for the differences between internal and external accuracies. But how can these factors be determined? If only the diagonal terms are increased, for example by adding the rms of all error sources, or by scaling them, then the correlations will probably be reduced unless the covariance terms are also increased so as to maintain the original correlations. There is no guarantee that this is a true representation of the actual VCV matrix but it may be a good approximation. Application of the covariance law (Jacobians) to add the effect of systematic errors to the VCV, is the best method. Unfortunately, it is based on functional relationships between the parameters and the systematic errors. These relationships may not be known.

The VCV matrix can be diagonal, block diagonal, or a full matrix. Considerable savings in computer resources can be made if the matrix is sparse. However, in practice the VCV is not diagonal, because the X, Y and Z co-ordinates of a point are correlated. Furthermore, the VCV is generally not block diagonal, because a co-ordinate of one point is usually correlated with co-ordinates of other points. Therefore full VCV matrices should be used in the computation of transformation parameters. Assuming that the VCV is block diagonal or diagonal when in fact it is not would cause errors in the results and their VCV estimates.

Horizontal ground co-ordinates, from a conventional 2D ellipsoidal adjustment, are usually combined with ellipsoidal heights which are the sum of orthometric height and geoid-ellipsoid separation at each point. It is usually assumed that these heights are not correlated with the horizontal co-ordinates.

In both GPS and ground network adjustments, the co-ordinates of one point are conventionally assigned zero variance. If only a portion of the ground net is to be used and it is a considerable distance from the fundamental point (e.g. Johnston origin) then the variances of co-ordinates of the points within this portion will usually be large. Strang van Hees (1982) gives formulae to convert VCV matrices so that they have minimum trace, or zero variances for some point(s) within the selected portion of the network. While the elements of the VCV of the co-ordinates change, the variance of the lengths and angles within the net do not. Thus the translation parameters and absolute co-ordinates estimated from a combined ground-GPS transformation adjustment are physically meaningless.

Co-ordinate Conversion

Both the Bursa-Wolf and Molodensky-Badekas transformation models given below require Cartesian (XYZ) co-ordinates and their VCV matrix. However frequently

ellipsoidal co-ordinates and VCV are available and it is necessary to convert them to co-ordinates and VCV in a Cartesian frame.

The Cartesian co-ordinates can be calculated from —

$$\begin{aligned} X &= (N+h)\cos\phi\cos\lambda \\ Y &= (N+h)\cos\phi\sin\lambda \\ Z &= \{(1-e^2)N+h\}\sin\phi \end{aligned} \qquad (4)$$

where $N = a/\sqrt{1-e^2\sin^2\phi}$ and $e^2 = 2f - f^2$

(For the Australian National Ellipsoid, a = 6378160. and f = 1/298.25.)

The theory of conversion of VCV matrices and the propagation of variances is clearly presented by Mikhail (1976) and others. The converted VCV matrix can be calculated from —

$$VCV_{XYZ} = J\ VCV_{\phi\lambda h} J^T \qquad (5)$$

where J is the Jacobian matrix.

$$J = \begin{bmatrix} \dfrac{Ne^2\sin\phi\cos^2\phi\cos\lambda}{(1-e^2\sin^2\phi)} - (N+h)\cos\lambda\sin\phi & -(N+h)\cos\phi\sin\lambda & \cos\phi\cos\lambda \\ \dfrac{Ne^2\sin\phi\cos^2\phi\sin\lambda}{(1-e^2\sin^2\phi)} - (N+h)\sin\lambda\sin\phi & (N+h)\cos\phi\cos\lambda & \cos\phi\sin\lambda \\ \dfrac{\{Ne^2\sin^2\phi\cos\phi}{(1-e^2\sin^2\phi)} + N\cos\phi\}(1-e^2) + h\cos\phi & 0 & \sin\phi \end{bmatrix}$$

The Jacobian matrix given above is used to transform the VCV of a single point. For conversion of the VCV of more than one point the full Jacobian matrix is —

$$J = \begin{bmatrix} J_1 & 0 & 0 & 0 & . & . \\ 0 & J_2 & 0 & 0 & . & . \\ 0 & 0 & J_3 & 0 & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & J_n \end{bmatrix} \qquad (6)$$

Similarly, the output adjusted co-ordinates and VCV of transformation models is often in a Cartesian frame. Corresponding ellipsoidal co-ordinates can be calculated by several methods, one of which is —

$$\begin{aligned} \phi &= \arctan\{(Z + e^2N\sin\phi) / R\}\ \text{[iterate]} \\ \lambda &= \arctan(Y/X) \\ h &= R/\cos\phi - N \end{aligned} \qquad (7)$$

where $R = \sqrt{(x^2 + y^2)}$

The VCV matrix of these ellipsoidal co-ordinates can be calculated from —

$$VCV_{\phi\lambda h} = J\ VCV_{XYZ}\ J^T \qquad (8)$$

where the Jacobian matrix, J, is —

$$\begin{bmatrix} A & YA/X & B \\ -Y/R^2 & X/R^2 & 0 \\ (X/R\cos\phi) + AC & (Y/R\cos\phi) + YAC/X & BC \end{bmatrix}$$

where —

$$A \simeq X\tan\phi / R^2(e^2 - \sec^2\phi) \qquad B \simeq 1/(R\sec^2\phi - e^2N\cos\phi)$$

$$C = R\sin\phi/\cos^2\phi - Ne^2\sin\phi\cos\phi/(1-e^2\sin^2\phi)$$

The conversion of the VCV matrix of more than one point can be accomplished in a manner similar to that used above.

Transformation Models

Several transformation models for combining different data types have been proposed and are reviewed by Krakiwsky and Thomson (1974) and Adam *et al* (1982). Geodetic networks normally cover only small portions of the globe. Thus some of the transformation parameters will be poorly determined. A number of techniques have been developed to improve the determination of these parameters. For many applications, the internal adjustment of the network, i.e. relative co-ordinates, is also of interest. Each of the models given below is examined to determine their effect on the internal adjustment and on the estimated transformation parameters. The least squares design matrices ('partials') for each of the models below are given by Harvey (1985).

The co-ordinates of the points of both nets are treated as observations and are adjusted. It is not necessary to hold the co-ordinates of any point fixed. However this can be done if desired by, for example, assigning very small variances to the co-ordinates.

Should a model which uses Cartesian or ellipsoidal co-ordinates of sites be applied? Whilst a data set and its VCV can be converted from one type (e.g. ellipsoidal or rectangular Cartesian) to another using the formulae given above, the conversion requires computer time and space and may be undesirable especially for large data sets. In such a case, the model used might be that which requires the least amount of preconversion of data.

Bursa-Wolf Model

This model solves for a 7 parameter transformation — a scale factor, 3 rotation angles, and 3 translation components.

$$\begin{bmatrix} x_B \\ y_B \\ z_B \end{bmatrix} = s\,R \begin{bmatrix} x_A \\ y_A \\ z_A \end{bmatrix} + \begin{bmatrix} Tx \\ Ty \\ Tz \end{bmatrix} \qquad (9)$$

where R (= Rz(k) Ry(θ) Rx(ω), is the rotation matrix about the axes of the xyz_A system, x_A, y_A and z_A are the co-ordinates of points in the A net, x_B, y_B, and z_B are the co-ordinates of points in the B net, s is the scale factor, and Tx, Ty and Tz are the translations, that is, the co-ordinates of the origin of the xyz_A system in the xyz_B system (see figure 2). Allowance for rotation angles of any size can be made (Harvey, 1985).

One well known problem is that generally the adjusted parameters are very highly correlated because the data covers only a small portion of the Earth's surface.

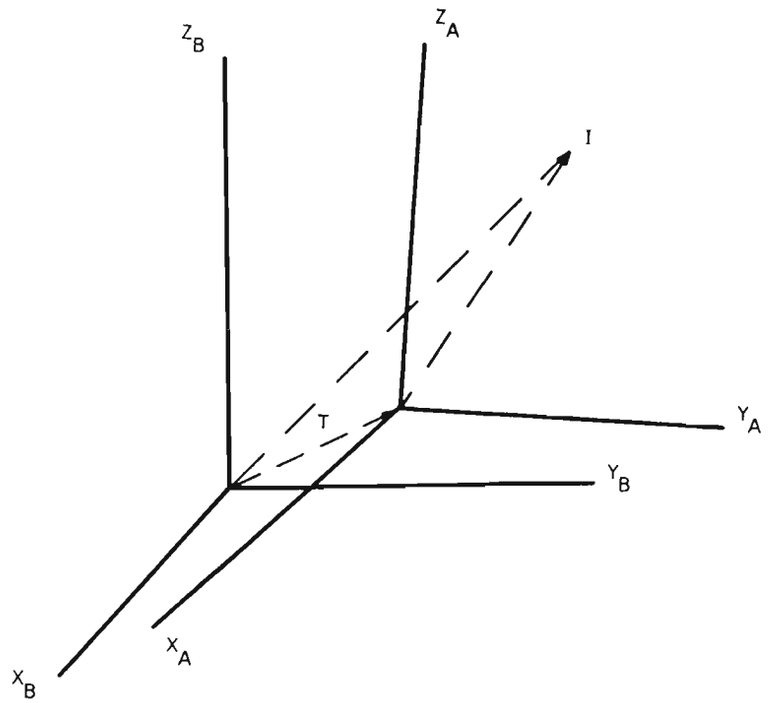

Figure 2. Bursa-Wolf Model

Molodensky-Badekas Model

This model (Badekas, 1969), which removes the high correlation between parameters by relating them to the centroid (m) of the network, is represented by —

$$\begin{bmatrix} x_B \\ y_B \\ z_B \end{bmatrix} = \begin{bmatrix} xm \\ ym \\ zm \end{bmatrix} + \begin{bmatrix} Tx' \\ Ty' \\ Tz' \end{bmatrix} + s\,R \begin{bmatrix} x_A - xm \\ y_A - ym \\ z_A - zm \end{bmatrix} \qquad (10)$$

where $xm = \Sigma x_{Ai/n}$, $ym = \Sigma y_{Ai/n}$, $zm = \Sigma z_{Ai/n}$. Alternatively m can be equal to the co-ordinates of one of the points of the A network.

This model gives the same answers as the Bursa-Wolf model for the internal adjustment, and the scale and rotation parameters. However the translation parameters are different and have smaller *a posteriori* precisions. This can be readily shown by expanding equation 10 as follows —

$$\begin{bmatrix} x_B \\ y_B \\ z_B \end{bmatrix} = P + \begin{bmatrix} Tx' \\ Ty' \\ Tz' \end{bmatrix} + s\,R \begin{bmatrix} x_A \\ y_A \\ z_A \end{bmatrix}$$

where $$P = \begin{bmatrix} xm \\ ym \\ zm \end{bmatrix} - s\,R \begin{bmatrix} xm \\ ym \\ zm \end{bmatrix}$$

P is a constant term, that is, it is the same for all points, and obviously affects the translation terms. The derived values of s and R equal those from the Bursa-Wolf model. If s and R are approximately equal to unity and the identity matrix respectively, then the model is very similar to the Bursa-Wolf model. Clearly the difference in the translation terms obtained from the two models is due to the scaling and rotating of the point m.

The adjusted co-ordinates, baseline lengths, scale factor and rotation angles, their VCV matrices and the *a posteriori* variance factor computed from this model are the same as those from a corresponding Bursa-Wolf solution. However the translations are different and their precisions are generally an order of magnitude smaller.

It is clear that the T′ terms of the Molodensky-Badekas model equal the T terms of the Bursa-Wolf model only when xm = ym = zm = 0, or s=1 and R=I. Although the variances of the T′ terms are smaller than those of the T terms, and are less correlated with the other parameters, the T′ terms are not better estimates of the T terms. They are fundamentally different — this is not commonly understood.

When the transformation parameters from this model are to be applied to transform co-ordinates of points it is essential to know what values of xm, ym and zm were used in the derivation of the parameters. However, in the past they have not always been published with the transformation parameters.

Ellipsoidal Model

This model is given by Harvey (1985) —

$$\begin{bmatrix} \phi_B \\ \lambda_B \\ h_B \end{bmatrix} = \begin{bmatrix} \phi \\ \lambda \\ h \end{bmatrix} + Cs + D \begin{bmatrix} \omega \\ \theta \\ k \end{bmatrix} + E \begin{bmatrix} Tx \\ Ty \\ Tz \end{bmatrix} \qquad (11)$$

where
$$C = \begin{bmatrix} -Ne^2\sin\phi\cos\phi/(M+h) \\ 0 \\ N(1-e^2\sin^2\phi) + h \end{bmatrix}$$

$$D = \begin{bmatrix} -\sin\lambda\{N(1-e^2\sin^2\phi)+h\}/(M+h) & \cos\lambda\{N(1-e^2\sin^2\phi)+h\}/(M+h) & 0 \\ \{(1-e^2)N+h\}\tan\phi\cos\lambda/(N+h) & \{(1-e^2)N+h\}\tan\phi\sin\lambda/(N+h) & -1 \\ -Ne^2\sin\phi\cos\phi\sin\lambda & Ne^2\sin\phi\cos\phi\cos\lambda & 0 \end{bmatrix}$$

$$E = \begin{bmatrix} \frac{-\sin\phi\cos\lambda}{(M+h)} & \frac{-\sin\phi\sin\lambda}{(M+h)} & \frac{\cos\phi}{(M+h)} \\ \frac{-\sin\lambda}{(N+h)\cos\phi} & \frac{\cos\lambda}{(N+h)\cos\phi} & 0 \\ \cos\phi\cos\lambda & \cos\phi\sin\lambda & \sin\phi \end{bmatrix}$$

where $N = \dfrac{a}{\sqrt{1-e^2\sin^2\phi}}$ and $M = \dfrac{a(1-e^2)}{(1-e^2\sin^2\phi)^{3/2}}$

ϕ_B is the geodetic latitude, λ_B is the longitude and h_B is the ellipsoidal height in the B net. ϕ, λ and h are the corresponding co-ordinates in the A net.

This model assumes that the co-ordinates of points in both networks are based on ellipsoids with the same semi-major axis and flattening. If they are not, one network should have its co-ordinates and VCV matrix converted so as to refer to the ellipsoid of the other net, or the model should be extended to cater for different shaped ellipsoids as in Vanicek and Krakiwsky (1982).

The ellipsoidal method should produce the same results as a Bursa-Wolf solution because the same parameters are being estimated. So the two models can be used to check the computer program and any pre-conversion of VCV matrices and co-ordinates.

However, a question of whether the ellipsoidal model or one of the Cartesian models (e.g. Bursa-Wolf or Molodensky-Badekas) is preferable may arise. The least squares adjustment procedure linearises the models by using the first partial derivatives. If higher order partial derivatives of the Cartesian or ellipsoidal models, or both, are significant then the least squares adjustment will not give accurate answers for the adjusted observations and parameters. The differences between the results and the

observation residuals obtained from each method can be statistically tested and the better method chosen. In solutions obtained with real and simulated data, the differences between the results from each of the above models have been less than 1 mm for adjusted lengths and co-ordinates and may therefore be ignored. The differences in the parameters were also found to be insignificant.

Network Geometry

The geometrical formulation of a similarity transformation represents a continuum. Replacing this continuum by a discrete set of common points can cause errors. Moreover, in the presence of observational and computational errors, the accuracy of the estimated parameters may vary considerably. This depends on the spatial distribution of the points used. Hence the implications of the spatial distribution of the points of the network must be studied.

Points should not be collinear because components of rotations about axes parallel to the line of points cannot be determined. For a stable solution it is important that the points are spatially well distributed. For example, a network with an uneven geographical spread of points will cause the solution to be biased towards the areas of high density. Points in areas of low density will be disadvantaged and probably have large corrections to their co-ordinates. Thus an even geographical distribution of points is desirable.

Parameters

The minimum number of observations required to solve for the seven parameters of a similarity transformation is seven. This condition could be satisfied by two common points, with their three co-ordinates known in each net, and a third point with three co-ordinates in one net known and one co-ordinate in the other net known. Three common points, each with their three co-ordinates known in each net, will yield a redundancy of two. However it is desirable to use at least four common sites because there should be at least four degrees of freedom in a solution for seven parameters.

Initially a surveyor should solve for all seven parameters and then test their estimates using the statistical tests given below. If a subset of parameters is found to be insignificantly different to zero then a second solution should be calculated with those parameters held fixed at zero.

It is quite likely that different parts of a continental geodetic network will have distortions in scale and orientation. These distortions may be due to systematic errors in the data. If distortions exist then it is better to solve for a number of sets of local parameters which better represent the local and regional areas, rather than determining one set of parameters for the whole of the continent. Transformation parameters over a large area will represent mean values of scale, rotation and translation for the area and will tend to smooth out distortions. Therefore the total network should be divided into subnets to check consistency throughout the whole area. However it is necessary to be sure that any differences in the parameters from one region to another are statistically significant. An appropriate test is given below.

Parameters representing systematic errors can also be added to the basic similarity transformation models given above (Harvey, 1985). If the model includes a large number of parameters, the adjustment may lead to a poorly conditioned system of equations. Many parameters will usually fit the data better, that is produce smaller residuals, than a few parameters. But the estimates of the parameters may not be accurate. Moreover,

the degrees of freedom of the adjustment will be reduced and therefore the statistical tests will be less effective. Thus, if the data does not contain significant contributions from the systematic errors for which parameters have been added, then using these additional parameters will give smaller residuals but not necessarily a more accurate solution. It should be noted that a scale difference between two nets may be due to systematic errors in the ground heights as well as an error in either GPS or ground distance scale.

The effects of systematic errors can be assessed by procedures such as the use of "check points". One check for systematic model deformations is to compare the estimated transformation parameters from an adjustment using all common sites with those from another solution where one site has been eliminated (the check point). If the changes in the adjustment results are acceptable, then the adjustment is stable and reliable. The co-ordinates of the check point in net A should be transformed by the parameters determined by the second solution. The transformed co-ordinates should then be compared with the corresponding *a priori* co-ordinates in net B.

Check points are a valuable tool but, if they are poorly chosen, can give misleading information about the stability of a solution. For instance, a check point in a cluster of points checks only the consistency within the cluster itself.

The whole process can only be partly automated. The surveyor must decide which parameters to hold fixed, whether to accept the results of statistical tests, and which data to delete. He must also make judgements on the stability of the solution and reliability of the results.

Scale Factor Residuals

As well as combining the two data sets using one of the above models, the surveyor can and should investigate the scale factor as determined from each baseline. A scale factor can be determined for each baseline using —

$$s = L_B/L_A \qquad (12)$$

where L_A is a length in net A and L_B the corresponding length in net B. These scale factors, and possibly also the estimate of the overall scale factor from a transformation model, can then be plotted. Generally, the scale difference, which is the scale factor minus one, would be plotted as in the hypothetical example in Figure 3 below.

Figure 3. Individual Baseline Scale Factors (hypothetical example)

In the plot the error bars of each point represent the standard deviations of the scale factor. These standard deviations are calculated as follows:

$$\sigma^2_S = \begin{bmatrix} \frac{\partial s}{\partial LB} & \frac{\partial s}{\partial LA} \end{bmatrix} \begin{bmatrix} \sigma^2_{LB} & 0 \\ 0 & \sigma^2_{LA} \end{bmatrix} \begin{bmatrix} \frac{\partial s}{LB} \\ \frac{\partial s}{\partial LA} \end{bmatrix}$$

$$= \sigma^2_{LB} / LA^2 + LB^2 \sigma^2_{LA} / LA^4$$

$$\simeq (\sigma^2_{LB} + \sigma^2_{LA}) / LA^2 \qquad (13)$$

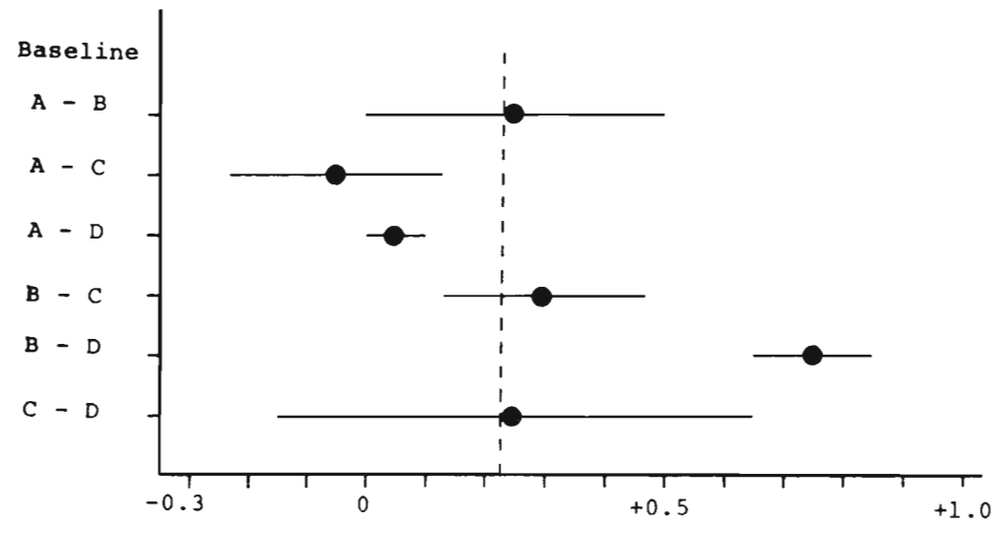

Figure 3. Individual Baseline Scale Factors (hypothetical example)

Note that scale factor, if determined from a single baseline, is not as precise as the baseline lengths. For example, if the baseline is measured to 0.10 ppm in both nets, the derived scale factor will have a precision of 0.14 ppm. It should also be noted that the longest lines do not necessarily give the best determinations of scale factor. The lines with the best fractional (ppm) accuracies give the most precise estimates of scale factor. That is, a very accurate short baseline may have more effect on an estimated scale factor than a poor quality long line. In this way the surveyor can determine which baselines are having the most effect on the overall scale factor as determined by a least squares solution of all data.

This procedure does not consider the correlations between baseline lengths (which are usually small) but it is useful for finding inconsistencies or outliers in the data and the relative power of each baseline to the overall scale factor. The surveyor should look for the following:

1. Single baselines being inconsistent with the overall network, such as line B-D in Figure 3. Standard statistical tests for detecting outliers should be applied.
2. All baselines from one site being inconsistent with the rest.
3. All baselines in one direction (e.g. East-West) being inconsistent.
4. If there are many baselines so that the data can reasonably be separated into regions, then all baselines in a region being inconsistent with the overall network.

Interpretation of the Results of a Transformation Adjustment

The results of a transformation adjustment are two sets of adjusted co-ordinates which differ by the adjusted parameters. In any least squares adjustment the observations are adjusted as well as the parameters being estimated. In this case in particular, the observations (co-ordinates) will be adjusted because, for example, there are two measurements of each length — once in each net. Even though the two input co-ordinate sets are generally not correlated with each other, the adjusted co-ordinate sets will be slightly correlated (the magnitudes of the correlations depend on the data).

The adjusted co-ordinates of network A have the same weighted scale, orientation and location as their *a priori* co-ordinates. Similarly, the adjusted co-ordinates of network B have the same weighted scale, orientation and location as their *a priori* co-ordinates. Thus, the overall scale, location and orientation of each network remains unchanged. Only the individual co-ordinates, angles and lengths are changed (of the order of the standard deviations of the observations). In other words, the movements of points within a network are small but the overall nature of the net does not change. If a transformation adjustment is computed between the *a priori* and adjusted versions of a net, it will be found that the estimated rotation angles, scale difference and translations will all equal zero.

The results are convenient for surveyors because the co-ordinates of points of a network do not change by large transformation parameters. Any large change in co-ordinates of a survey control network necessitates recomputation and design of surveys and construction work based on the control network, and is therefore undesirable. The co-ordinates of the fundamental point of the ground net (e.g. Johnston in Australia) will not change if its *a priori* variances equal zero, even if it is not included in the transformation calculations. This is also important for surveying applications.

The precision of the adjusted length of a line is a function of the *a priori* VCV matrix of the observations (which is usually different for each net), the lengths of the baselines, and the precision of the adjusted scale factor (which is not zero unless the scale factor is held fixed). Therefore the precision of an adjusted baseline length will generally be different in the two networks. They will be the same only when the *a priori* VCV of the observables of each net is the same or when the scale factor is held fixed (not necessarily at unity). Since an adjusted length in the B network, LB, equals the corresponding adjusted length in the A net, LA, times the scale factor, s, it can be shown that —

$$\sigma^2_{LB} = s^2\,\sigma^2_{LA} + 2\,s\,LA\,\sigma_{sLA} + LA^2\,\sigma^2_s \qquad (14)$$

Obviously σ^2_{LB} equals σ^2_{LA} if σ^2_s and σ_{sLA} equal zero, and s equals one, that is when the scale factor is fixed at unity. Note that if s is not held fixed then generally σ_{sLA} will not equal zero because of correlations introduced in the least squares adjustment. Also note that σ_{sLA} may be negative and thus allows σ_{LB} to be less than σ_{LA}. If s is held fixed at some value other than unity then σ_{LB} does not exactly equal

σ_{LA}. However since s rarely differs from unity by more than a few ppm the difference will be insignificant.

Statistical Tests

The application of statistical tests can sometimes be a demanding task, especially when the data sets are large. Moreover, most tests require the full VCV matrices, which may be very large. However, because there are not many parameters in a transformation adjustment and there are usually only a few common points, none of the tasks described below are too demanding and should all be carried out.

The statistical tests discussed below assume that the observations and their residuals are normally distributed. So before the tests can be applied it is necessary to check that the residuals are normally distributed. The tests for normal distribution involve calculating the statistical moments and are discussed by Hoar (1975). If the residuals are not normally distributed then the results of the adjustment may be biased by systematic errors. The standard F test of the *a posteriori* variance factor should also be applied.

After the adjustment the co-ordinates of each point should be tested for outliers. This is done by comparing the correction to a co-ordinate with its standard deviation. Even though the co-ordinates are obtained from previously tested adjustments they may still contain unknown systematic errors such as, the ground and GPS surveys not measuring to exactly the same points. As mentioned above, the corrections to baseline lengths should also be investigated. When the residuals are correlated, it is inappropriate to use a simple rejection criterion based on 2.5 or 3 times the standard deviation of the residuals. Nevertheless this criterion is often applied because of its simplicity. More elaborate methods are given by Pope (1976) and Vanicek and Krakiwsky (1982).

Tests of the Estimated Parameters

The following multivariate tests should be applied to determine whether the estimate of a parameter or group of parameters is significant. Calculate t from (Vanicek and Krakiwsky, 1982) —

$$t = (U-X)^T\,Cx^{-1}\,(U-X) \qquad (15)$$

where X is the vector of parameter estimates being tested, U is the vector of *a priori* values against which each Xi is being compared (often the null vector), and Cx is the estimated VCV of the parameters being tested. The vectors X and U may contain either all the parameters or some subset of them. If a subset is being tested than the corresponding portion of the VCV matrix of the parameters must be used. The hypothesis that $X_i \neq U_i$ should be rejected if —

$$t < \chi^2_{k,\alpha} \qquad (16)$$

where k is the number of parameters being tested and α is the significance level. If X is not significantly different to U_i at the α level then the surveyor can be $(1-\alpha)\times 100\%$ sure that the estimate(s) can be ignored and a readjustment safely computed holding the tested parameters fixed at the U_i values. If only a single parameter is being tested, then it is obvious that

$$t = (u-x)^2 / \sigma_x^2 \qquad (17)$$

For the test of a single parameter at the 95% confidence level, the parameter estimate is insignificant if its difference from the test value (u) is less than 1.96 times its standard deviation.

The above procedure tests the parameters simultaneously and is rigorous. All variances and covariances are taken into account. However, in the test given above (equation 16) it is assumed that the *a priori* variance factor is known. If the *a priori* variance factor is not known, then the following test should be used (Mikhail, 1976) —

$$t' = (U-X)^T\, Cx^{-1}\, (U-X) \,/\, k \qquad (18)$$

where Cx is the estimated VCV matrix of the parameters, or the relevant portion of it, multiplied by the *a posteriori* estimate of the variance factor.

The hypothesis that $X_i \neq U_i$ should be rejected if —

$$t' < F_{k,r,\alpha} \qquad (19)$$

where r is the number of degrees of freedom in the adjustment. If a reliable VCV was used with the input data, then the *a priori* variance factor is known (=1). The estimated *a posteriori* variance factor is then tested using the standard procedures. Otherwise the test in equation 19 should be used. The value of t′ required to reject the hypothesis will be much larger than the value of t required in equation 16, unless r is very large.

If some parameter estimates do not differ significantly from zero, it is advisable to recompute the adjustment with these parameters fixed at zero. The effect, of holding these parameters fixed, on the estimates of the other parameters, the adjusted co-ordinates, and the adjusted lengths should also be studied (see Harvey [1985] for a method).

Tests of Regional Distortions

Two sets of transformation parameters should be statistically tested to determine whether they are significantly different. This can be done by using the same tests as those applied in deformation analysis to determine if two sets of co-ordinates (at different epochs) are significantly different. This test was originally given by Pelzer (1971) and quoted by Chen (1983); rewritten with different symbols it is —

$$T = \frac{(X2 - X1)^T\, (C_{X1} + C_{X2})^{-1}\, (X2 - X1)}{(X2C_{X2}X2 + X1C_{X1}X1)\, k \,/\, (r1 + r2)} < F_{\alpha,k,r1+r2} \qquad (20)$$

where X1 and X2 are the two sets of transformation parameters,
C_{X1} and C_{X2} are the estimated VCV of the parameters.
r1 and r2 are the degrees of freedom of the solutions for X_1 and X_2 (the degrees of freedom usually equal 3 × number of points −7).
k is the number of parameters (usually 7).

If T is less than F, then X2 is not significantly different to X1 at the 1-α (e.g. 95%) confidence level. This tests the parameters simultaneously (i.e. multivariate) and is rigorous in that it uses the variances and *covariances* of the parameters. However this test assumes that X1 and X2 are not correlated. If the parameters are derived from subsets which do not contain any points in common then X1 and X2 will not be correlated.

Some Applications of Transformations

Transformation adjustments are used for two purposes:

1. To relate two datums empirically. That is, obtain measurements of a number of points with co-ordinates known in both systems and solve for the transformation parameters.

2. Use known transformation parameters to convert co-ordinates, or baseline component vectors, from one datum to another. However it must be remembered that a transformation is essentially an interpolation procedure. So the results may be unreliable if the parameters are used to transform the co-ordinates of a point outside the region spanned by the common points which were used to derive the parameters.

It is beyond the scope of this paper to consider all aspects of the combination of GPS with ground data. However a number of questions/points which may need to be considered are listed below. Moreover, the answers to these questions will usually depend on the actual data and its quality in a particular region or for a particular application. Thus generally applicable answers may not exist — further research is needed.

One common application will be the densification of the ground net by GPS observations. In this case GPS observations can be made on several points with known ground co-ordinates — thus giving common points to solve for the transformation parameters. GPS observations will also be made on numerous points of unknown ground co-ordinates. Then a few options exist to determine the corresponding ground co-ordinates of these new GPS points. Should the GPS-derived values be distorted to fit the local control? Should one GPS point be held fixed in the GPS adjustment? Or should all GPS points common to the ground net be held fixed in the GPS adjustment? Problems will arise if the relative positioning observations by GPS are more accurate than that of the existing ground control. If GPS observations are more accurate the ground net may need to be strengthened and perhaps even readjusted using the GPS observations.

Which adjusted net should be adopted?

Should the adjusted co-ordinates of net A or the adjusted co-ordinates of net B be adopted? If it is desired to adopt the GPS scale, orientation and location, then the adjusted GPS co-ordinates should be adopted. Similarly, if it is desired to adopt the ground scale, orientation and location, then the adjusted ground co-ordinates should be adopted. If the GPS data is more accurate then the adjusted GPS net should be adopted if the net with the best internal adjustment is desired.

If the two adjusted nets are of similar quality and if the best internal adjustment is required, then that net with the smallest determinant of the VCV of the adjusted lengths should be adopted. If there are many points in the adjustment and the internal adjustment of data in a certain region is of prime interest, then the determinants of the VCV of the adjusted lengths of only the baselines in that region should be compared. The determinants of the VCV of the adjusted co-ordinates should not be used for this test because the VCV of co-ordinates may contain large components due to a network's datum defects and its distance from the fixed point (e.g. Johnston origin).

However in practice it may not be desirable to change the co-ordinates of the ground points. One way to overcome this is to solve for the transformation parameters with

the ground co-ordinates held fixed. In such a case it would be wise to compare the results with those from a proper adjustment where the ground co-ordinates are not held fixed. Then the errors caused by holding the ground points fixed would be known.

VCV of Transformed Co-ordinates

When known transformation parameters are used to transform co-ordinates it is also desirable to appropriately convert the VCV of those co-ordinates. The VCV of transformed co-ordinates can be determined by applying the law of propagation of variances, as follows. Since —

$$\begin{bmatrix} X_B \\ Y_B \\ Z_B \end{bmatrix} = s \begin{bmatrix} 1 & k & -\theta \\ -k & 1 & \omega \\ \theta & -\omega & 1 \end{bmatrix} \begin{bmatrix} X_A \\ Y_A \\ Z_A \end{bmatrix} + \begin{bmatrix} Tx \\ Ty \\ Tz \end{bmatrix}$$

the

$$VCV_{XYZ_B} = J \begin{bmatrix} VCV_{XYZ_A} & 0 \\ 0 & VCVp \end{bmatrix} J^T \qquad (21)$$

where $J = \begin{bmatrix} s & sk & -s\theta & X_A + kY_A - \theta Z_A & 0 - sZ_A & sY_A & 1 & 0 & 0 \\ -sk & s & s\omega & Y_A - kX_A + \omega Z_A & sZ_A & 0 - sX_A & 0 & 1 & 0 \\ s\theta - s\omega & & s & \theta X_A - \omega Y_A + Z_A - sY_A & sX_A & 0 & 0 & 0 & 1 \end{bmatrix}$

In VCVp (the VCV matrix of the parameters) the parameters are in the order: s, ω, θ, k, Tx, Ty, Tz, VCV_{XYZ_B} and VCV_{XYZ_A} are 3×3 matrices, VCVp is a 7×7 matrix.

Similarly, the VCV matrix of a number of transformed co-ordinates (X_{B1}, Y_{B1}, Z_{B1}) to (X_{Bn}, Y_{Bn}, Z_{Bn}) can be determined as follows —

$$VCV_{XYZ_B} = J \begin{bmatrix} VCV_{XYZ_A} & 0 \\ 0 & VCVp \end{bmatrix} J^T \qquad (22)$$

where

$$J = \begin{bmatrix} F & 0 & 0 & \dots & G1 \\ 0 & F & 0 & \dots & G2 \\ 0 & 0 & F & \dots & G3 \\ & & & \dots & \\ \dots & \dots & 0 & F & Gn \end{bmatrix}$$

$$F = \begin{bmatrix} s & sk - s\theta \\ -sk & s & s\omega \\ s\theta - s\omega & & s \end{bmatrix}$$

and

$$Gi = \begin{bmatrix} X_{Ai} + kY_{Ai} - \theta Z_{Ai} & 0 & -sZ_{Ai} & sY_{Ai} & 1 & 0 & 0 \\ Y_{Ai} - kX_{Ai} + \omega Z_{Ai} & sZ_{Ai} & 0 & -sX_{Ai} & 0 & 1 & 0 \\ \theta X_{Ai} - \omega Y_{Ai} + Z_{Ai} & -sY_{Ai} & sX_{Ai} & 0 & 0 & 0 & 1 \end{bmatrix}$$

VCV_{XYZ_B} and VCV_{XYZ_A} are 3n × 3n matrices.

Thus the VCV matrix of the transformed co-ordinates of a point is a combination of the VCV matrices of the original co-ordinates and the transformation parameters used. Therefore if the estimates of the transformation parameters are to be published, or otherwise distributed for use, it is desirable that the VCV of the parameters also be published. Alternatively the standard deviations and correlations of the parameters could be published because — 1) they provide a clearer indication of the quality of the parameter estimates, and 2) the VCV can be easily formed from the standard deviations and correlations. In either case only 28 terms need be given because the VCV and correlation matrices are symmetric and there are at most seven parameters. An example of desirable output for publication of the parameters is given in Table 1.

Table 1: Example of Output

(a) Adjusted transformation parameters. (Bursa-Wolf model)

Parameter	*Adjusted Value*
Scale factor (s−1)	0.303 ± 0.448 ppm
Rotation about X axis (ω)	−0.790 ± 0.315 secs
Rotation about Y axis (θ)	−1.078 ± 0.659 secs
Rotation about Z axis (k)	−0.142 ± 0.613 secs
Translation along X axis	−109.111 ± 18.428 m
Translation along Y axis	−64.439 ± 16.627 m
Translation along Z axis	118.734 ± 15.767 m

(b) Correlations between adjusted parameters.

	s	ω	θ	k	Tx	Ty	Tz
s	1.000						
ω	−0.109	1.000					
θ	0.320	0.139	1.000				
k	0.340	−0.447	0.795	1.000			
Tx	−0.293	0.118	−0.958	−0.929	1.000		
Ty	−0.346	0.685	−0.603	−0.957	0.790	1.000	
Tz	0.314	0.398	0.961	0.612	−0.847	−0.375	1.000

Concluding Remarks

The following comments may help to clarify some common misconceptions.

- The Molodensky-Badekas model and the Bursa-Wolf model define the translation terms differently. They represent different methods of describing the same transformation relationship between two sets of co-ordinates.

- The ellipsoidal method could also be applied, especially if computer resources are limited. It may save the step which converts the co-ordinates and their VCV to a Cartesian system.
- The Bursa-Wolf model, the Molodensky-Badekas model and the ellipsoidal model all produce the same transformed co-ordinates and VCV when correctly applied, i.e. the transformation parameters and their full VCV are used.
- The software that solves for the transformation parameters should use the full VCV of the input co-ordinates.
- The transformation models and procedures are applicable to many different types of data e.g. Doppler/ground, GPS/ground, GPS/VLBI.
- The abbreviated version of the rotation matrix should not be used if the rotation angles are large.
- When determining the transformation parameters all seven parameters should initially be estimated. If the estimates of some of the parameters are statistically insignificant (using the multivariate test) then, and only then, the solution should be recomputed with those parameters fixed at zero.
- Multivariate statistical tests should be applied to determine whether differences between transformation parameters for different regions are significant. If the differences between two regions are not significant then one set of parameters should be determined for the combined region. Then there will be a substantial increase in the number of redundancies in the solution.
- When publishing or distributing derived transformation parameters for others to use the following should be given —
  the parameters,
  the model used,
  the VCV of the parameters or their standard deviations and correlations (28 terms),
  the degrees of freedom in the adjustment (for statistical tests),
  the co-ordinates of the barycentre (xm, ym, zm), if the Molodensky-Badekas model was used.
- The VCV of transformed co-ordinates is a function of the VCV of the transformation parameters as well as the VCV of the original co-ordinates.
- If co-ordinates and their VCV are transformed by parameters with standard deviations but correlations ignored, then the Molodensky-Badekas model should be used (because the correlation terms will be smaller). However this is not a wise procedure. Note that the co-ordinates of the barycentre will be needed.

Acknowledgements

Bruce Harvey is currently supported by a Rothmans Fellowship. Some of this research was carried out while he was a Commonwealth Postgraduate Award holder.

A/Profs Art Stolz and John Allman, and Don Grant provided useful comments towards this paper.

References

Adam, J., Halmos, F. and Varga, M. (1982) On the concepts of combination of Doppler Satellite and Terrestrial Geodetic Networks, *Acta Geodaet., Geophys. et Montanist. Acad. Sci. Hung* Vol 17(2) pp.147-170.

Badekas, J. (1969) Investigations related to the establishment of a World Geodetic System, Ohio State University Department of Geodetic Science and Surveying. Report 124.

Chen, Y. (1983) Analysis of Deformation Surveys — A Generalised Method,, Dept. Surv. Eng. UNB Canada, Tech Rpt No. 94, 262pp.

Harvey, B. R. (1985) The Combination of VLBI and ground data for Geodesy and Geophysics, UNISURV S-27, School of Surveying, University of NSW, 244pp.

Hoar, G. J. (1975) The Analysis, Precision and Optimisation of Control Surveys, UNISURV S-13, School of Surveying, University of NSW, 200pp..

Krakiwsky, E. J. and Thomson, D. B. (1974) Mathematical Models for the Combination of Terrestrial and Satellite Networks, *Canadian Surveyor* Vol. 28 No. 5.

Krakiwsky, E. J. (1981) A Synthesis of recent advances in the method of Least Squares, Dept. Surv. Eng. Uni. of Calgary, Canada, Publ. 10003.

Mikhail, E. M. (1976) Observations and Least Squares, IEP, Harper & Row, p.497.

Pelzer, H. (1971) Zur Analyse Geodatischer Deformationsmessungen, Deutsche Geodatische Kommission, C 164, Munchen.

Pope, A. J. (1976) The statistics of residuals and the detection of outliers, NOAA Tech. Rpt. NOS 65 NGS 1.

Strang van Hees, G. L. (1982) Variance-Covariance transformations of geodetic networks, *Manus. Geod.* Vol 7 pp.1-20.

Vanicek, P. and Krakiwsky, E. J. (1982) Geodesy: The concepts, North-Holland, 691pp.

(Received 29 November 1985; accepted for publication 24 March 1986.)