Functional and stochastic models for geometrical detection of spatial deformation in engineering: a practical approach

Author: Halim Bin Setan MSc, BSc.

Qualification: This thesis is submitted for the degree of Doctor of Philosophy.

Submitted: City University,
Department of Civil Engineering,
January 1995.

ABSTRACT

The objective of this study is to formulate a simple and practical but rigorous two-step analysis procedure for the geometrical detection of spatial deformation using geodetic methods. A thorough and critical study of theory and current practice of deformation monitoring has been undertaken, and a practical scheme has been developed for 3-D least squares estimation (LSE) and one-stage detection procedure (i.e. stability determination and localization of spatial deformation) via two-epoch analysis.

In LSE, a simple datum definition via minimum constraints with fixed coordinates has been adopted; a strategy for rank defect analysis of normal equations by simplified eigenvalue decomposition (EVD) has been developed; an optimised computational procedure for S-transformations has been formulated; a mathematical model for additional parameters and pseudo observables (distance differences and ratios) has been extended and established for 3-D application; a procedure for handling of algebraically correlated pseudo observations via observation de-correlation has been established; a procedure for robustified LSE for multiple gross errors detection has been formulated and its effects has been derived; a simple method of variance component estimation (VCE) has been extended; and the use of global and local tests and reliability analyses in LSE has been presented.

In deformation detection, a strategy for determination of common stations between epochs via S-transformations and partitioning has been developed; a flexible one-stage computational procedure for geometrical detection of spatial deformation by iterative congruency testing and S-transformations has been established; the robust method for deformation detection has been modified to allow one-stage computation; and general S-transformations equations have been applied in all cases.

This developed strategy has been implemented in five computer programs (ESTIMATE, COMPS, COMON, DETECT and ROBUST). The developed programs can be executed either on an IBM based personal computer (PC) or under the UNIX environment. Links between these programs and two of the Engineering Surveying Research Centre's (ESRC) programs (GAP and DCRE) have been established. The programs have been successfully applied and evaluated using simulated and real data. Five real photogrammetric monitoring schemes undertaken by the ESRC, with up to 169 stations, were analysed for detecting the significance of spatial deformation between epochs. The results obtained confirmed the suitability of the strategy in practical applications.

Further refinement to the developed programs are suggested to make them more user friendly. Further possible research activities include a combined or integrated approach for deformation analysis and real time deformation monitoring.

ACKNOWLEDGEMENTS

The author would like to thank his supervisor, Prof. M.A.R. Cooper, who suggested the topic of this research, provided continuous valuable guidance, encouraging ideas and suggestions towards the completion of this thesis.

Special thanks are due to: Prof. W.F. Teskey of the Department of Surveying Engineering of The University of Calgary (Canada) for personally discussing and clarifying the presented methodology during his visit to ESRC; Professor A. Chrzanowski of the Department of Engineering Surveying of The University of New Brunswick (UNB, Canada) for providing useful information on the UNB approach for deformation detection.

The author is indebted to Dr. S. Robson, who provided valuable discussion and proof-read the thesis. Dr. J. Chandler is thanked for providing the suitable format of his graphics program DCRE in the MicroStation environment. Co-operation and help obtained from the staff members of the ESRC are highly appreciated.

Professor J. Fryer and Mr. M. Elfick of the Department of Civil Engineering and Surveying at The University of Newcastle (Australia), and Mr. M. Harris of the Department of Surveying at The Queensland University of Technology (Australia) are acknowledged for helpful discussion during their periods of sabbatical leave at ESRC.

The author would also like to thank the University of Technology Malaysia and the Public Service Department of Malaysia for sponsoring this study. Credits are also due to the author's wife (Salina) and children (Hazwani, Khairul Dani, Liana and Hafetz) for their continuous support and understanding throughtout these hard and challenging years.

Declaration

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOTATION AND ABBREVIATIONS

A	first design matrix or configuration or coefficient matrix
A^+	pseudo inverse of A
b	misclosure vector or observed minus computed vector
B	second design matrix
BLUE	best linear unbiased estimate
BIQUE	best invariant quadratic unbiased estimate
cond ()	condition of a matrix
c	weighting factor in RLSE
C	general constraints matrix (for datum definition)
d	vector of coordinate difference or displacement vector or
d or d_d	scalar, number of datum defects
d_c	configuration defect
det ()	determinant of a square matrix
D	cofactor matrix for computing the determinant
e_i	expected frequency
E	matrix of eigenvalues
E{ }	expectation
EDM	electromagnetic distance measurement
ESRC	Engineering Surveying Research Centre
EVD	eigenvalue decomposition
F	Fisher distribution
FOD	first-order design
G	inner constraints matrix
G^g	generalised inverse of G
G_n	normalised G
G^+	pseudo inverse of G
GPS	global positioning system
H_a	alternative hypothesis
H_o	null hypothesis
I	unit or identity matrix
I_p or I_j	diagonal matrix for datum definition
J	Jacobian matrix

k	maximum number of linearly independent rows or columns in a matrix, or number of non-zero singular values or rank of a matrix or number of observation groups for VCE
k_1	scale error or scale bias
k_2	zero or additive or constant error
l	vector of observations or measurements or observables
l_a	LSE of the observables or adjusted observations or adjusted observables
l_o	computed observations
LSE	least squares estimation
m	number of equations in the combined model or the number of stations in a network or
m	metre
mm	millimetre
mp	number of observation equations containing the bias parameter
M	redundancy matrix
$MDGE_i=\nabla_i$	marginally detectable gross error of observation i
n	number of observations
N	normal equation matrix or
N	standard normal distribution
N^{-1}	ordinary or Cayley inverse of N
N^+	pseudo inverse or Moore-Penrose inverse of N
N^r	symmetrical reflexive generalized inverse of N
o_i	actual frequency
order ()	order of matrix
p	additional parameters
P	matrix used in decomposition of matrix S
PC	personal computer
Q	matrix of eigenvector
Q_d	cofactor matrix of d
Q_l	cofactor matrix of the observations
Q_{la}	cofactor matrix of the adjusted observations
Q_{xa}	cofactor matrix of the estimated parameters
Q_v	cofactor matrix of the estimated residuals
r	distance ratio or

r or df	number of degrees of freedom or total of the redundancy numbers or redundancy in LSE
r_i	redundancy number of observation i
rank ()	rank of matrix
R	Cholesky factor of Σ_l or
R	matrix used in decomposition of matrix S
RLSE	robustified LSE
r_x or R_1	rotation about x-axis (α)
r_y or R_2	rotation about y-axis (β)
r_z or R_3	rotation about z-axis (γ)
RMSE	root mean square error
s or λ	scale
s_i	singular value of i
S	matrix of S-transformations
s_{max}	maximum singular value
s_{min}	minimum singular value
SLR	satellite laser ranging
SOD	second-order design
SVD	singular value decomposition
t	Student's distribution
tr ()	trace of matrix
t_x	translation along x-axis
t_y	translation along y-axis
t_z	translation along z-axis
T	test statistic or
T	decorrelation matrix
T_b	Baarda's (or data snooping) test statistic
TOD	third-order design
T_p	Pope's (or Tau) test statistic
u	vector of right hand side of normal equation or
u	scalar, number of parameters
$u_{\alpha o}$	critical value of Baarda's method
U	left singular vector or triangular matrix from Cholesky factorization
v	vector of residuals

v_i'	normalized residual for observation i
V	right singular vector
VCE	variance component estimation
VLBI	very long baseline interferometry
w_i	weight of observation i
W or P	weight matrix of the observations
x	vector of estimated parameters or correction vector
x_a	vector of updated parameters
x_o	approximate values of parameters
(x_o, y_o, z_o)	coordinates of the centroid
$y^{1/2}$	expansion factor of error ellipsoid
ZOD	zero-order design
α	Type I error probability or significance level or risk level or rotation with respect to x axis
β	Type II error probability or rotation with respect to y axis
γ	rotation angle with respect to z axis
Δx	limit used in numerical modelling
Δs	distance difference
δ_i	influential factor or global distortion parameter of observation i (external reliability)
λ_o	non-centrality parameter
∇_i	MDGE of observation i (internal reliability)
ρ	correlation coefficient
σ_i	standard deviation of observation i
σ_i^2	variance of observation i
σ_{xy}	covariance between x and y
σ_o^2	variance factor or variance of unit weight or reference variance or unit variance
Σ	matrix containing singular values
Σ_l	covariance matrix of the observations
Σ_r	covariance matrix of the distance ratio
$\Sigma_{\Delta s}$	covariance matrix of the distance difference
τ	tau distribution
χ^2	chi-square distribution
Ω	quadratic form

1-D	one dimensional
$(1+k_1)$	scale factor
$1-\alpha$	confidence level or probability
$1-\beta$	power of test
2-D	two dimensional
3-D	three dimensional or spatial

1. INTRODUCTION

This chapter introduces some of the important aspects related to the detection of spatial (3-D) deformation, highlights current research trends, and summarizes the work involved in this thesis.

1.1 Background

Surveying is the determination of the relative positions of points on the earth's surface by means of terrestrial and / or space-based measurements. It is customary to divide surveying into geodetic and plane surveying, depending on the area of coverage. Geodetic surveying generally extends over large areas and takes into account the curvature of the earth. Plane surveying involves relatively small areas (e.g. for distances up to 100 m where the linear effect of curvature is about 1 mm) and the earth's surface is considered as a plane. This project is concerned only with plane surveying, but the computations are in three dimensions (3-D).

Deformation survey or monitoring of deformation is an important area of engineering surveying. Its prime purpose is the detection of spatial deformation to provide information on the stability and extent of any movement or deformation of an object occuring over time. Information obtained from the detection process is useful for the purpose of safety assessment, as well as for predicting and preventing the possibility of failure or disaster in the future.

In general, the earth's crust and man-made features undergo deformations. The deformation could be caused by some of the following factors (Vanicek and Krakiwsky, 1986): tidal phenomena; crustal loading and rebound; tectonic phenomena; ground consolidation; combined effects; as well as short and long term movement of engineering structures brought about by loading and ground settlement.

Deformation surveys are mainly carried out to investigate crustal movement, slope stability, glacier and shelf ice movement, ground subsidence and deformation of man-made engineering structures (Caspary, 1987b). Examples of such structures which are commonly studied are dams, bridges, pipelines and tall buildings.

The determination or detection of deformation consists of design, measurement and analysis stages. The measurement techniques (Richardus, 1984) are generally divided into geodetic and geotechnical / structural methods. Geodetic methods can be based on the following measuring techniques: conventional or terrestrial surveys (including photogrammetry) and space-based methods. Figure 1.1 illustrates the performance of the geodetic measuring techniques. In general, the accuracy decreases with increasing baseline.

The geotechnical and structural methods are the most accurate for monitoring over short distances (Teskey, 1986, 1988), say up to a few tens of metres. Special geotechnical equipment is used to measure directly the changes in height (settlement gauge), length (extensometer), water pressure (piezometer) and tilt (inclinometer). The structural measuring equipment measures changes in displacement (displacement meter), strain (strainmeter) and inclination (inclinometer). However, such methods only provide information about local movement, and are suitable only to determine movements within the structures, not the overall movement of an object under investigation.

Geodetic methods are capable of determining movements within structures (only if access is possible) and overall movements. Terrestrial surveys with electromagnetic distance measurement (EDM) instrument and theodolite are the most commonly used for distances ranging up to 10 km. Their accuracy (affected by refraction) can be improved by several means, such as adopting proper measurement schemes (Ashkenazi et al, 1980; Secord, 1986) or modifying the functional model (Gruendig and Teskey, 1984).

The space system using Global Positioning System (GPS) is suitable for baselines of several km, and with the use of dual frequency receivers, allows corrections for ionospheric effect to give millimetre (mm) level accuracy. Very Long Base Line Interferometry (VLBI) and Satellite Laser Ranging (SLR) used for tectonic studies, are highly accurate but very expensive. Photogrammetry (especially the close range configuration) is a powerful non-contact measuring technique using metric cameras, suitable when multiple targets are involved. Inertial methods are typically of lower accuracy and are also expensive. The selection of the most appropriate technique depends on factors such as cost, accuracy required and scale (coverage). Combination of the measuring techniques is also possible.

In engineering surveying, the first scheme of deformation monitoring using geodetic methods was carried out in Switzerland in the 19th century (Caspary, 1987b). Generally,

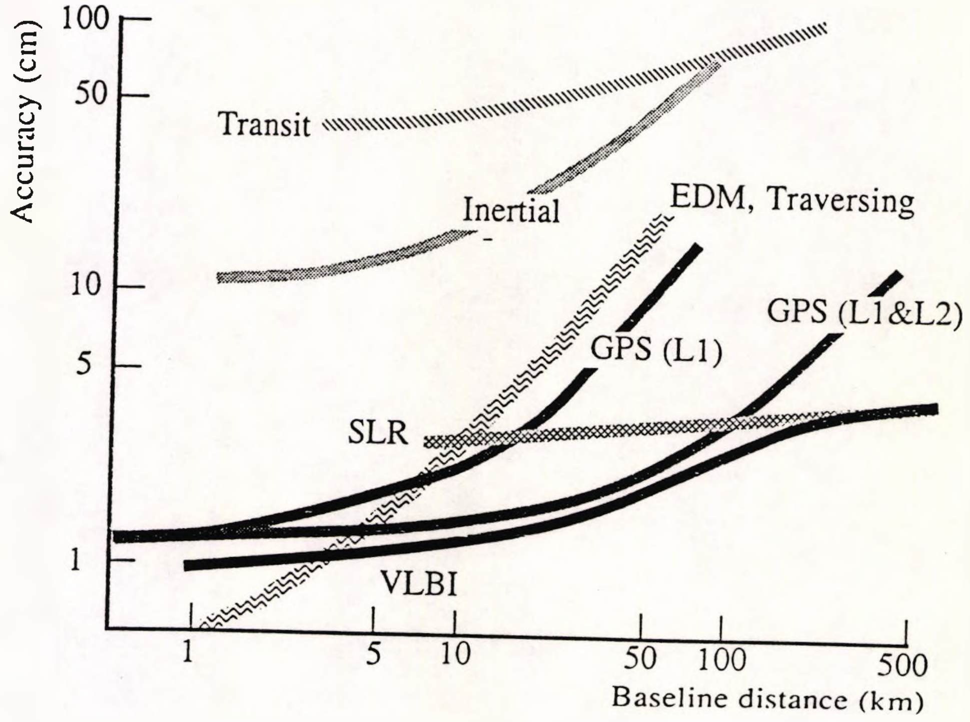

Figure 1.1 The performance of the geodetic measuring techniques (taken from Wells et al, 1986)

although movements are actually three dimensional (3-D), their detection is usually divided into horizontal and vertical components. This research allows 3-D detection of deformation.

Application of the geodetic method is quite simple when the object under investigation is represented by targetted or marked points. A set of observations is used to connect the points into a monitoring network. The observations repeated at different epochs of time provide data for deformation detection.

In theory, all observations at each epoch are carried out simultaneously. However, as this is often impossible from the practical point of view, the effects of deformations during observational periods are usually neglected, provided the rate of deformation is small and the observational period is short compared to the interval between epochs. Alternatively, time can be considered as a parameter, and Papo and Perelmuter (1981) suggest the use of displacement velocities and accelerations to describe the deformation (i.e. kinematic model).

Subsequent analysis to determine significant deformations can be performed in one (simultaneous) step or two-step analysis. One-step analysis involves extensive computation, and is rarely applied in practice. Two-step analysis consists of independent least squares estimation (LSE) of single epoch (or analysis of observation) followed by deformation detection (or analysis of deformation) between epochs. Only two-step analysis will be dealt with from now on.

Determination of any deformation occuring can be based on geometrical, physical or combined geometrical and physical approaches. In the geometrical approach, a high precision survey monitoring network is needed (Biacs, 1989). It describes the estimated deformation in the form of the displacement vectors, without interpretation of the cause of movements. The physical approach (via geotechnical and structural methods) gives only relative measurements (Richardus, 1984), and is beyond the scope of this study. The combined approach is the best, but more research is needed.

Several methods for the geometrical approach to deformation detection via two-step analysis with geodetic methods are available. The method chosen depends on the differencing technique, type of network, coverage and type of model being used.

Deformation detection can be based on either coordinate or observation differencing. Coordinate differencing is used commonly, due to its flexibility, the most important aspect being its ability to handle different observational schemes at different epochs. Observation differencing has the advantage of being datum invariant, but has the major drawback of requiring identical observational schemes at different epochs. Cooper (1987) gives more details on differencing.

There are two types of monitoring networks depending on their purpose: absolute and relative. An absolute monitoring network usually consists of the reference points (expected to be stable) and the object points (under investigation). In a relative monitoring network, all points are considered as object points. The absolute monitoring network approach is more meaningful in engineering applications as the deformation of the object points is determined relative to a set of stable reference datum points. In the relative network approach, only the pattern of the relative displacement between points can be determined. Further details are given by Chen (1983).

The coverage of the deformation survey can be on micro, local, regional, continental or global scales, and is closely related to the measuring techniques. Most engineering applications are within micro, local and regional scales (several m to 10 km), employing terrestrial survey (such as triangulation, trilateration, levelling), photogrammetric and GPS techniques. A common example is dam monitoring. For monitoring at continental and global scales, such as the monitoring of crustal movements, space-based techniques (GPS, VLBI, SLR) combined with high precision gravimetry are usually employed.

Analysis between epochs can be based on two-epoch or multi-epoch analysis. For engineering applications, two-epoch analysis (assuming no correlation between epochs) is generally adequate and provides enough information.

Models chosen for analysis are either static, kinematic or dynamic, depending on the temporal variations (Biacs, 1989). The static model examines only the existence or non-existence of deformations. The kinematic model deals with the motion of network points. The dynamic model takes into account the effects of various underlying forces on the motions of network points.

1.2 Reasons for carrying out research in this area

The commonly adopted methods for monitoring deformation in engineering are based on the repeated observation of a survey monitoring networks at different epochs, followed by two-step analysis. The detection of deformation uses two-epoch analysis, an absolute monitoring approach and a static model to compare the coordinates between the epochs.

The stages of LSE and deformation detection are highly critical and need special attention because the significance of the estimated deformations depends on the observational accuracy and network design. In most engineering cases, the magnitudes of deformation to be detected are small, and at the margin of observational error. In LSE, a realistic mathematical model is needed because an erroneous model will lead to apparent deformation. During deformation detection, it is required to transform the results into a common datum, identify a set of stable points and localize the deformation.

Extensive research work has been carried out on LSE and deformation detection in Europe and North America, involving many sophisticated methods. However, very little effort has been made to arrive at a simple and practical, but rigorous, method suited to the practising surveyor. In most cases, the applications are restricted to two dimensional (2-D) or one dimensional (1-D) only (Dodson, 1990), whilst deformation actually occurs in 3-D. Usually, the deformation detection procedure consists of two-stage computations (Gruendig et al, 1985; Chen et al, 1990a): analysis of the reference points followed by analysis of the whole network.

This thesis sets out to devise practicable means for meeting the following important requirements for geometrical detection of spatial deformation using terrestrial surveying observables: coordinate datum definition; rank analysis and error modelling in LSE; identification of common or stable stations; congruency testing; S-transformations; and localization and testing of deformation.

1.3 Summary of the historical theoretical development

The two-step analysis has been devised over several decades. Originally, the theory of LSE was developed independently by Gauss in 1795 and Legendre in 1806, and later refined by Markoff in 1912 (Cross, 1983; Cooper and Cross, 1988). Since then, the theory has been continually examined, refined and applied using linear equations (Searle, 1971; Lawson and Hanson, 1974).

The theory of generalized inverses, developed by Moore in 1920 and Penrose in 1955, is applied for solving the datum problem in the singular linear model (Cooper and Cross, 1991). The solution uses the inner constraints method (Meissl, 1969; Blaha, 1971). A derivation for the complete inner constraints is presented by Papo (1987) and Dermanis (1994). Network optimization studies (Grafarend, 1974), carried out before any observations are made can be used to estimate network quality.

As an alternative to inner constraints, Baarda (1973) introduced S-transformations for transforming LSE results from one datum to another. A more simple explanation of the theory of S-transformations is given by Strang Van Hees (1982). Teunissen (1985) discusses the concepts of the generalized inverses and S-transformations.

Statistical testing was first used for detailed analysis of LSE results in engineering surveying only in 1960's (Baarda, 1968). This was followed by the development of outlier detection and reliability theories (Baarda, 1977). Aspects of statistical testing are discussed in great detail by Mikhail (1976) and Vanicek and Krakiwsky (1986). Today, the mathematical model for LSE is known as the Gauss-Markov model.

In the past, approximate methods have been employed in analysis for detecting deformation. During the 1970's, statistical testing was used extensively in the analysis. Pelzer (1971) was the first to apply statistical testing to deformation detection (Biacs and Teskey, 1990). Further developments, resulting in several methods for deformation detection were made by Van Mierlo (1975), Heck et al (1977), Neimeier (1979), Koch (1980) and Kok (1982).

In order to compare different approaches of geometrical analysis for deformation monitoring, the International Federation of Surveyors (or Fédération Internationale des Géomètres (FIG)) established a Committee of Working Group on Analysis of Deformation Measurements (Commission 6 on Engineering Surveys) in 1978.

The committee initially identified five main research groups, named after their locations, i.e. Delft, Fredericton, Hannover, Karlsruhe and Munich (Chrzanowski, 1981; Heck et al, 1983).

More groups were established later. FIG also organized several symposia, which were held in Cracow, 1975; Bonn, 1978; Budapest, 1982; Katowice, 1985; and Fredericton, 1988. In 1986, the Committee published its final report (Chrzanowski and Chen, 1986) with details on various methodologies and possible further work.

The real application of deformation monitoring is extensive in Europe (Gruendig et al, 1985) and North America (Chen et al, 1990a; Teskey and Biacs, 1991; Teskey et al, 1992). Several program packages were developed, for example LOKAL (Gruendig et al, 1985), DEFNAN (Chrzanowski et al, 1986) and CANADAS (Biacs, 1989).

Current trends in deformation monitoring include integrated analysis in which the measured displacements are combined with the finite element method (Teskey, 1986; Teskey and Biacs, 1990; Szostak-Chrzanowski and Chrzanowski, 1991), robust estimation of deformation (Caspary and Borutta, 1987a), and real time monitoring using telemetric and automatic data acquisition (Chrzanowski et al, 1991).

The most recent development in monitoring activities is on the micro-scale in real time for close-range industrial applications. Bayly and Teskey (1992) successfully applied an electronic theodolite system to close-range three dimensional high precision machinery alignment surveys in near real time.

1.4 Outline of the thesis

The objective of this research project is to formulate a simple and practical but rigorous procedure for 3-D LSE and the geometrical detection of spatial (3-D) deformation using geodetic methods. Moreover, fully 3-D applications and one-stage detection process (for stability determination and localization of deformation) are anticipated.

To achieve this objective, several important aspects of LSE and deformation detection have been examined, developed, adopted and implemented by writing and testing computer programs. As a consequence, a practical strategy has been developed by the author, and successfully applied to evaluate both simulated and real data.

As summarised in the title of the thesis, 3-D LSE (including error modelling) and geometrical detection of spatial deformation are discussed. Practical aspects and testing have been confined to engineering applications.

1.5 Thesis contributions

The particular contribution of this research includes the following:

(1) A thorough and critical study of theory and current practice of deformation monitoring has been carried out and is described in chapters 2, 3 and 4.

(2) A full 3-D case has been applied and presented throughout this study, as opposed to the commonly used 2-D application.

(3) A strategy for one-stage detection procedure (i.e. stability determination and localization of spatial deformation) has been developed (sections 4.2.3 and 4.4).

(4) A simple method for handling datum defects by means of a minimum constraints solution with fixed coordinates has been adopted (section 2.2.2).

(5) A strategy for rank defect analysis of normal equations via simplified eigen value decomposition (EVD) has been developed (section 2.2.6).

(6) A suitable and efficient computational procedure for the practical application of S-transformations for transforming LSE results from one datum to another has been formulated (section 2.3.5).

(7) The general S-transformations equations have been applied in all cases (sections 2.3.1, 2.3.5, 4.2.1, 4.2.3.2 and 4.2.3.4).

(8) A general functional model for additional parameters, allowing single, combination or multiple errors has been developed (section 3.5.3.1).

(9) Functional and stochastic models for pseudo observables have been extended for the 3-D case (section 3.5.3.2).

(10) A procedure for incorporating the algebraically correlated pseudo observations into ordinary LSE algorithms by de-correlation of observations has been established (section 3.5.3.2.3).

(11) A practical blunder detection strategy via robustified LSE has been formulated (section 3.6.2.2) and its effects on the estimated solution derived (section 3.6.2.3).

(12) A simple method of variance component estimation (VCE) for estimating variances of uncorrelated observations has been extended (section 3.7.3).

(13) A strategy for LSE with global and local tests, together with precision and reliability

analyses has been presented (section 3.9).

(14) A strategy to determine common stations between two epochs and to apply S-transformations to transform LSE results of each epoch into a common datum has been developed (section 4.2.1).

(15) A one-stage computational procedure for geometrical detection of spatial deformation that incorporates flexible initial datum station definitions, relevant global and local statistical testing, and S-transformations has been established (sections 4.2.2 and 4.2.3).

(16) A one-stage computation for robust method has been formulated (section 4.2.3.4).

(17) Five computer programs have been developed based on the practical strategy. Links with relevant programs have been established too (chapter 5).

(18) Simulated and real data have been analysed to evaluate the applicability of the developed strategy (chapter 6).

(19) Recommendations for future work have been presented (chapter 7).

1.6 Thesis structure

The thesis consists of seven main Chapters, seven Appendices and a list of References and Bibliography.

Chapter 2 gives a brief introduction to the principles of LSE. It also highlights the related important aspects of rank defect analysis and datum re-definition. Strategies developed for rank defect analysis and S-transformations of LSE results are discussed.

Chapter 3 summarizes the main sources of model errors and the importance of quality measures and statistical testing in LSE. The remaining part of this chapter describes the strategy developed for modelling of systematic, gross and random errors. Moreover, a strategy for 3-D LSE is presented.

Chapter 4 initially highlights some of the requirements for deformation detection. This is followed by a description of the modules developed for geometrical detection of spatial deformation.

Chapter 5 focuses on the actual implementation of the concepts, described in chapters 2, 3 and 4, into five computer programs developed for LSE and geometrical detection of spatial deformation (programs ESTIMATE, COMPS, COMON, DETECT and ROBUST).

Chapter 6 discusses the results obtained from the application of the developed computer programs using simulated and real data to assess the adopted strategy.

Chapter 7 summarizes the outcome of the research and the developed practical strategy for deformation detection. It also highlights related future work that can be explored.

Appendices contain relevant information not shown in the main sections.

References and the Bibliography contain lists of sources which were consulted in the course of the research. They are listed under the authors' names arranged in alphabetical order, chronologically for each author.

2. LSE, RANK DEFECT ANALYSIS AND S-TRANSFORMATIONS

The process of LSE, as applied in monitoring of deformation, suffers from rank deficiency due to datum and / or configuration defects. Moreover, relevant results from LSE are datum dependent. Consequently, it is important to incorporate checks for rank defects and also a facility for transformation of LSE results into an appropriate datum.

This chapter describes some fundamental aspects of LSE, rank defect analysis and transformation of LSE results from one datum to another. A simple method for handling datum defects has been adopted (section 2.2.2). A strategy for checking rank defects of normal equations has been developed (section 2.2.6), utilizing simplified eigenvalue decomposition (EVD). Practical application of S-transformations has been formulated (section 2.3.5), with emphasis on special computational procedures optimised for stability and speed.

2.1 Estimation process

In engineering surveying, the computational problems are concerned with the determination (or estimation) of parameters or unknowns from the measurements (or observations) by means of the chosen mathematical models. The main tasks involved can be divided into eight inter-related steps (Gracie and Krakiwsky, 1987; Vanicek and Krakiwsky, 1986). These steps are shown in Figure 2.1: identification of the parameters, formulation of the mathematical model, design or pre-analysis, data acquisition (observations), data pre-processing, data processing (estimation), assessment and representation of the results.

This study concentrates on the estimation task, which is actually comprised of estimation, assessment and the representation of results. Whenever necessary, other tasks are considered too, as they are inter-related. For example, formulation of mathematical model and pre-analysis.

Surveying measurements, such as distances, angles and height differences, are used to estimate the parameters (for example 3-D coordinates, orientation unknowns, scale factor, etc.). The relationship between measurements and parameters is known as the mathematical model. While the measurements are made in the physical space or real world, the estimation of parameters is made in abstract space using a particular and suitable mathematical model.

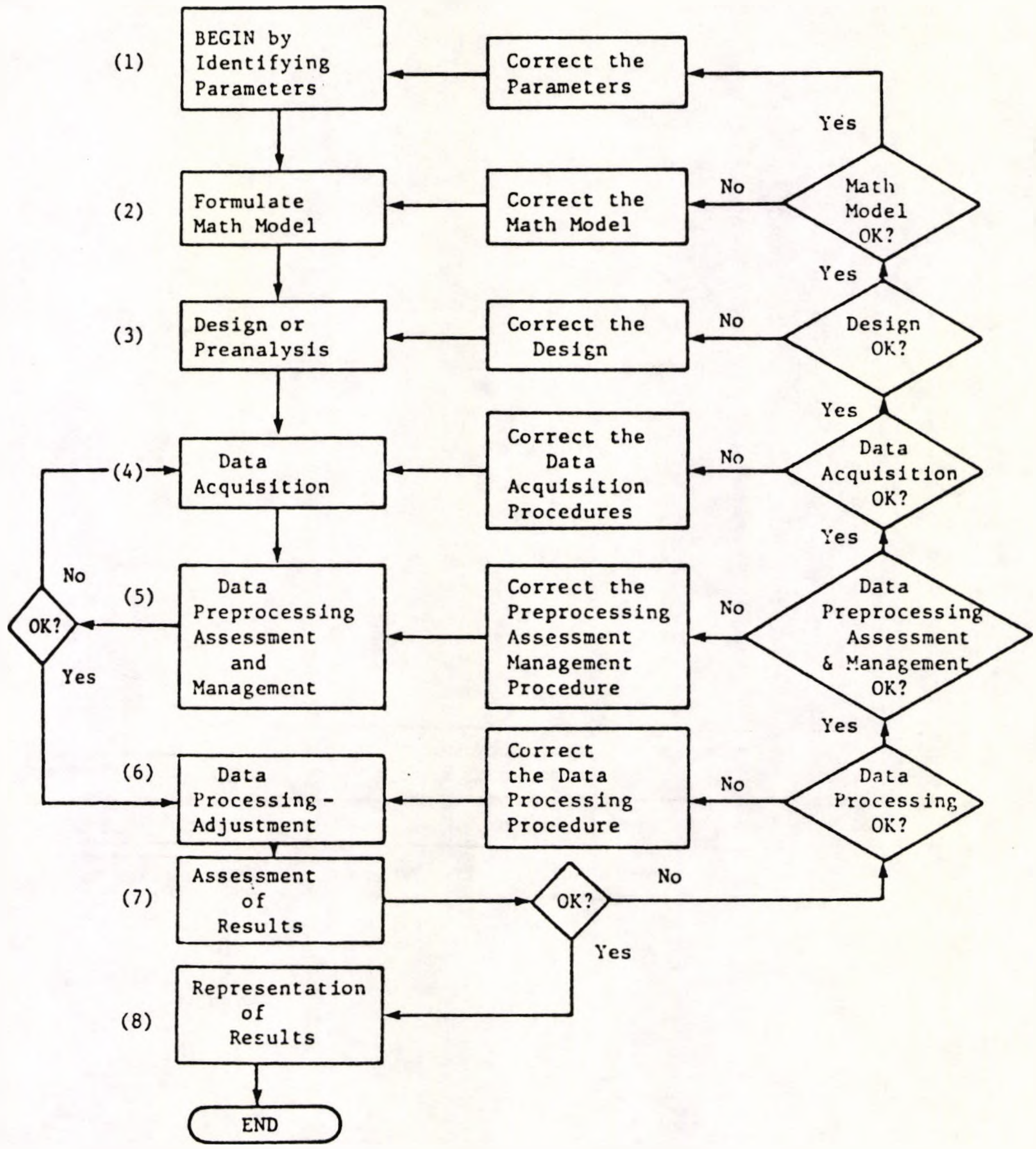

Figure 2.1 The main tasks in the determination of the parameters (taken from Gracie and Krakiwsky, 1987).

The mathematical model is composed of two parts (Mikhail, 1976), functional and stochastic models. The functional model describes the geometrical relationship between the measurements and the parameters to be estimated, and in the general case has an implicit form of

$$f(x,l)=0 \tag{2.1}$$

where x is the vector of parameters to be estimated and l is the vector of observations. Let m equations in the functional model relate n observations and u parameters.

The stochastic model describes the random nature (or statistical properties) of the measurements. It is represented in the form of a covariance matrix (Σ_l) or weight matrix (W) or cofactor matrix (Q_l) of the observations. The relationship between Σ_l, Q_l and W is simply

$$W=\sigma_o^2\Sigma_l^{-1} \tag{2.2}$$

$$\Sigma_l=\sigma_o^2Q_l \tag{2.3}$$

$$W=\sigma_o^2\Sigma_l^{-1}=Q_l^{-1} \tag{2.4}$$

where σ_o^2 is the a priori variance factor (often assumed to be known with value of unity).

The functional model is generally non-linear (as given in Appendix A), and consequently leads to a non-linear computational problem. Also, there are redundancies in the data, i.e. the number of measurements is larger than the number of parameters. In other words, there are more measurements than the minimum needed for a unique solution of the parameters. Redundancy results in an overdetermined system of equations.

For simplicity and convenience, non-linear problems are usually 'linearised' by using Taylor's theorem, and then reduced to the utilization of linear mathematical models, matrix notation and consequently solution of linear algebraic equations. Solution using non-linear models (Press et al, 1988; Rawlings, 1988) is beyond the scope of this study.

2.1.1 Least squares problem

In order to obtain unique estimates of parameters from the overdetermined linear equations, certain criteria are required. The most commonly used is known as the least squares criterion, and the (estimation) problem is termed the linear least squares problem or simply the least squares method. In numerical analysis (Golub and Loan, 1990), the method which uses the least squares criterion is sometimes known as 2-norm or L2 norm minimization.

Essentially, the least squares criterion minimizes the quadratic form of the residuals (corrections to the observations), i.e. the sum of the squares of the weighted residuals

$$v^t Q_l^{-1} v \rightarrow \text{minimum} \tag{2.5}$$

or $$v^t W v \rightarrow \text{minimum}$$

where v is the vector of residuals.

The least squares problem is known by different names in different scientific disciplines. For example, most surveyors use the term least squares adjustment. Statisticians are more comfortable with linear regression. Cooper (1987) prefers the term least squares estimation (LSE) because the term has a more proper statistical meaning than least squares adjustment. The term LSE will be used throughout this study.

2.1.2 Parameter estimation

Linearization of equation (2.1) via the application of Taylor's theorem to the first order (Appendix A) produces a linearised form of the functional model as

$$Ax+Bv=b \tag{2.6}$$

where

$A=\partial f/\partial x$ is the first design matrix, dimensions (m,u)

x=vector of corrections to the parameters, dimensions (u,1)

$B=\partial f/\partial l$ is the second design matrix, dimensions (m,n)

v=vector of residuals, dimensions (n,1)

$b=-f(x_o,l)$ is the misclosure vector, dimensions (m,1)

x_o=approximate values of parameters with dimensions (u,1)

l=vector of the observables, dimensions (n,1)

In equation (2.6), the quantities in A, B and b are computed using x_o and l. The solutions for equation (2.6) with least squares criterion of equation (2.5) can be written as

$$\hat{x}=[A^t(BW^{-1}B^t)^{-1}A]^{-1}A^t(BW^{-1}B^t)^{-1}b \qquad (2.7)$$
$$\hat{v}=-W^{-1}B^t(BW^{-1}B^t)^{-1}(A\hat{x}-b)$$

Further details on the solution are given by Cross (1983) and Cooper and Cross (1988). Usually, it is necessary to iterate the solution due to the approximations (i.e. the initial approximate values of the parameters are not accurate enough because only first order terms of Taylor's series are used in the linearisation procedure). In each iteration, the parameters are updated to give

$$\hat{x}_a=x_o+\hat{x} \qquad (2.8a)$$

where $\hat{x}_a$ represents updated parameters, x_o the approximate values of parameters (updated in each iteration), and $\hat{x}$ is computed from equation (2.7). At the end of LSE, the least squares estimates of the observables ($\hat{l}_a$) are

$$\hat{l}_a=l+v \qquad (2.8b)$$

The general functional model of equations (2.1) and (2.6) is also known as the combined model (Cross, 1983) or general model (Cooper, 1987). In practice, two special cases of the functional model of equations (2.1) and (2.6) are considered. Firstly, if each measurement can be written as an explicit function of the unknowns, the functional model becomes

$$l=f(x) \qquad (2.9)$$

Equation (2.9) is known as LSE of parameters or LSE using observation equations or parametric LSE or LSE of indirect observations or the variation of coordinates.

Secondly, if the model is a function of only the observations and the parameters do not appear in the model the functional model is

$$f(l)=0 \qquad (2.10)$$

Equation (2.10) is called LSE using condition equations or LSE of corrections to measurements or LSE of observations only.

The linearised forms of equations (2.9) and (2.10) are simply obtained by substituting B as unit matrix (with negative sign) and A as null or zero respectively into equation (2.6). The results can be written as

$$Ax=b+v \tag{2.11}$$

$$Bv-b=0 \tag{2.12}$$

In each case, the number of equations in (2.11) and (2.12) will be n and (n-u) respectively.

All of the above techniques of LSE (i.e. general, observation and condition equations) produce identical results when applied to the same problem (Mikhail and Gracie, 1981). However, the most commonly used method is LSE using observation equations (equations 2.9 and 2.11), due to its simplicity. In this method, the formulation of equations is simple and straightforward, as the number of observation equations is exactly the same as the number of observations. Moreover, it is easy to implement the procedure using computer programs. This method of LSE is used throughout this study.

2.1.3 LSE using observation equations

In general (Lawson and Hanson, 1974), there are six cases of the LSE problems using observation equations, depending on the number of observations and parameters, and rank of A, as shown in Figure 2.2. Cases 1, 2 and 3 are called exactly determined (no redundancy), overdetermined (with redundancy) and underdetermined (not enough data) problems respectively. In this study, with data redundancy, only case 2 is applicable, which can be either full rank (case 2a) or rank deficient (case 2b).

Equations for LSE are shown here without further derivation. More details are found extensively in surveying literature, for example Mikhail (1976), Olliver and Clendinning (1978), Mikhail and Gracie (1981), Cross (1983), Cooper (1987), Koch (1987) and Leick (1990). The fundamental equations for LSE of full rank (only if Cayley inverse N^{-1} exist; section 2.1.5) with n observations and u parameters with redundancy r are:

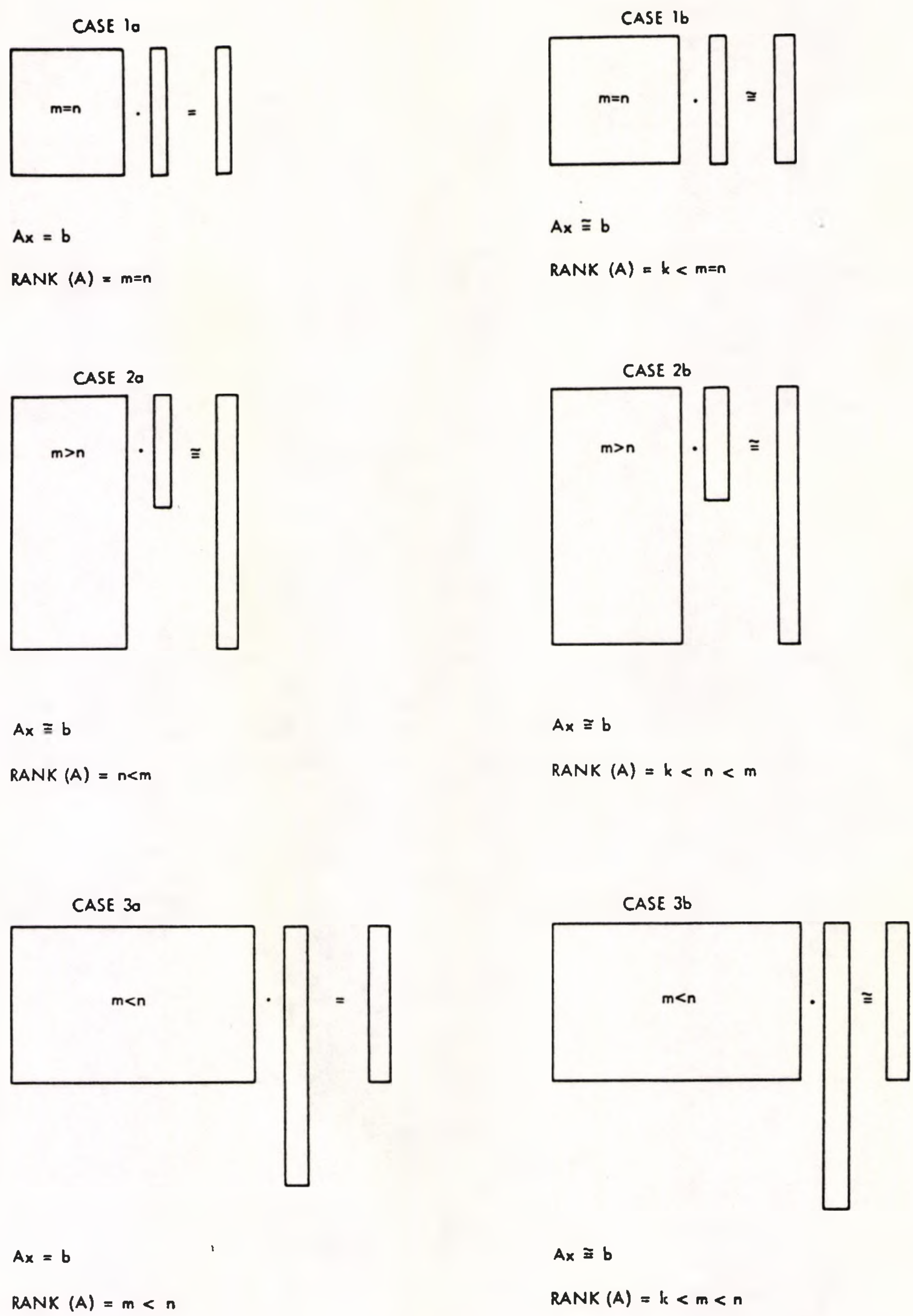

Figure 2.2 The six cases of the least squares problem.
The cases depend on the sizes of observation (m), parameters (n) and rank of (A) (taken from Lawson and Hanson, 1974).

Non-linear model i.e. functional model

$$l=f(x) \tag{2.13}$$

Observation equation (linear model)

$$Ax=b+v \tag{2.14}$$

where

$A=\partial f/\partial x$ is the design matrix, dimensions (n,u), rank u

x=vector of corrections to x_o, dimensions (u,1)

$b=l-l_o$ is the misclosure vector, dimensions (n,1)

l=vector of actual observation (corrected and reduced for systematic errors), dimensions (n,1)

$l_o=f(x_o)$ is the computed observation, dimensions (n,1)

x_o=approximate values of parameters, dimensions (u,1)

v=vector of residuals, dimensions (n,1)

Weight matrix W (equation 2.4) of dimensions (n,n) i.e. stochastic model

$$W=\sigma_o^2\Sigma_l^{-1}=Q_l^{-1} \tag{2.15a}$$

where σ_o^2 is the a priori variance factor.

If observations are uncorrelated, W is diagonal matrix, and weight of observation i is

$$w_i=\sigma_o^2/\sigma_i^2 \tag{2.15b}$$

where σ_i^2 is the variance of observation i.

Normal equations

$$Nx=u \tag{2.16}$$

where

$$N=A^tWA \text{ with dimensions } (u,u) \quad (2.17)$$

$$u=A^tWb \text{ with dimensions } (u,1) \quad (2.18)$$

Estimated parameters ($\hat{x}$) and their cofactor matrix ($Q_{\hat{x}}$)

$$\hat{x}=N^{-1}u=(A^tWA)^{-1}A^tWb \quad (2.19a)$$

$$\hat{x}_a=x_o+\hat{x} \text{ are the updated parameters} \quad (2.19b)$$

$$Q_{\hat{x}a}=Q_{\hat{x}}=N^{-1}=(A^tWA)^{-1} \quad (2.20)$$

Estimated residuals ($\hat{v}$) and their cofactor matrix ($Q_{\hat{v}}$)

$$\hat{v}=A\hat{x}-b \quad (2.21)$$

$$Q_{\hat{v}}=Q_l-AN^{-1}A^t=W^{-1}-AN^{-1}A^t \quad (2.22)$$

The estimated (or a posteriori) variance factor $\hat{\sigma}_o^2$, useful for statistical testing (section 3.3.1) is

$$\hat{\sigma}_o^2=\Omega/r \quad (2.23)$$

where

$r=n-u$ is the number of degrees of freedom or redundancy

$\Omega=\hat{v}^tW\hat{v}$ is the quadratic form of the residuals

Least squares estimates of the observables $\hat{l}_a$ (sometimes called adjusted observations or adjusted observables) and their cofactor matrix ($Q_{\hat{l}a}$)

$$\hat{l}_a=l+\hat{v} \quad (2.24)$$

$$Q_{\hat{l}a}=AN^{-1}A^t=AQ_{\hat{x}}A^t \quad (2.25)$$

Relationship between cofactor matrices is shown by equations (2.22) and (2.25), hence

$$Q_{\hat{v}}=Q_l-Q_{\hat{l}a} \text{ or } Q_{\hat{l}a}=Q_l-Q_{\hat{v}} \quad (2.26)$$

All the cofactor matrices (equations 2.20, 2.22, 2.25 and 2.26) are symmetrical. The process of LSE is generally iterative, as described in section 2.1.2. Equations (2.14) to (2.23) are evaluated repeatedly during iteration. At the end of iteration, Cooper (1987) recommends that the estimated variance factor (equation 2.23) is used for statistical testing only (section 3.3.1), and not for scaling $Q_{\hat{x}}$ (equation 2.20). Such concept is applied in this work.

2.1.4 Properties of least squares estimates

In general, there are an infinite number of solutions to equation (2.14), depending on the chosen criteria. In practice, it is common to apply least squares criterion (equation 2.5), leading to LSE (equations 2.13 to 2.26). The main reasons for this are based on both practical and mathematical considerations (Cross, 1983; Cooper and Cross, 1988), as discussed below.

Practically, LSE is very simple to apply, gives a unique solution of parameters and provides simple quality measures via the error propagation. Equations (2.17), (2.18) and (2.19a) show that the least squares estimate of the parameters is a linear estimate (transformation) of the measurements.

Mathematically, the LSE given by equations (2.19) and (2.20) has the properties (Caspary, 1987b) of the best linear unbiased estimate (BLUE), and consequently the estimated variance factor (equation 2.23) being the best invariant quadratic unbiased estimate (BIQUE). BLUE means that the linear estimate is unbiased and has minimum variance, i.e.

$$E\{\hat{x}\}=x \tag{2.27}$$

$$\text{trace } (Q_{\hat{x}})=\text{minimum} \tag{2.28}$$

Equation (2.27) reflects the unbiased property, where the expectation of $\hat{x}$ is equal to the true value of x. Equation (2.28) indicates that the trace of $Q_{\hat{x}}$ obtained by LSE is smaller than $Q_{\hat{x}}$ of any other linear unbiased estimate. If the observations are normally distributed, LSE has the property of a maximum likelihood estimate (Cross, 1983).

The estimated parameters (usually coordinates) and their cofactor matrix computed from equations (2.19) and (2.20) are datum dependent, depending on the choice of datum constraints. However, there exist functions of x that are invariant or independent of the datum (Caspary, 1987b), known as estimable or datum invariant quantities. Examples of estimable quantities are $\hat{v}$, $Q_{\hat{v}}$, $\hat{l}_a$ and $Q_{\hat{l}a}$, as well as Ω and $\hat{\sigma}_o^2$.

2.1.5 Assumptions in LSE

The process of LSE given in section 2.1.3 is based on the following assumptions:

1. Linearity.

The basic assumption in LSE is that the functional model (equation 2.14) is linear. Generally, the linear model can be used to provide an adequate and satisfactory approximation of the actual model. Non-linear functions are linearised via the application of Taylor's theorem.

2. Computational or full rank system.

The solution from LSE, as shown by equations (2.13) to (2.26), assumes that the Cayley inverse, N^{-1}, exists. Hence, rank of N (or A) is equal to the number of parameters u, i.e. full rank.

3. Model is correct.

It is usually assumed that the mathematical model is a true representation of the physical reality. Both selected functional and stochastic models are considered correct, adequate and complete.

4. Independent and uncorrelated observations.

For simplicity and convenience, observations are assumed to be independent and uncorrelated. This results in a diagonal weight matrix (equation 2.15b).

5. Observational data are free from errors.

The observations are assumed to be free from systematic and gross errors (or blunders or mistakes).

6. Normality.

For the purpose of statistical analysis of LSE results, observations (and consequently residuals) are assumed to be normally distributed, with zero mean. This normality assumption is not necessary for the LSE process, but is required for statistical testing purposes.

It is necessary to check the validity of the above assumptions. If any failures of such assumptions are found, they must be rectified. A strategy for rank defect analysis is developed in section 2.2.6. A procedure for error modelling is formulated in sections 3.5.3, 3.6.2 and 3.7.3. In some cases, the observations are algebraically correlated, the treatment of such correlated observations is established in section 3.5.3.2.3. The detection of gross errors is formulated in section 3.6. Statistical testing is discussed in section 3.3. In engineering surveying, the linear model is adequate, and the linearity assumptions are usually satisfactory.

2.2 Rank defect analysis

In general, LSE suffers from rank deficiency (section 2.2.1). Moreover, critical configurations will introduced ill-conditioned or near singularity situations. Although several means of handling datum defects are available, a simple approach will be useful. In addition, checks for rank defects (due to configuration) during LSE are required.

It is necessary to examine the nature of any defects (section 2.2.1), define the datum (section 2.2.2), and to check the rank defects (section 2.2.3). The powerful numerical technique of singular value decomposition (SVD) and eigenvalue decomposition (EVD) are introduced (sections 2.2.4 and 2.2.5) for determination of the rank of the normal equations. Simplified EVD is formulated and applied as a practical tool for rank defect analysis (section 2.2.6).

2.2.1 The nature of defects

In section 2.1.3, LSE technique (equations 2.13 to 2.26) uses the following linearised equations:

$$Ax=b+v \tag{2.29}$$

$$Nx=u \tag{2.30}$$

where $N=A^tWA$ and $u=A^tWb$

$\hat{x}=N^{-1}u$ and $Q_{\hat{x}}=N^{-1}$ (2.31)

The above equations are only applicable in the case of full rank. Generally, matrices A (and N) are not full rank, due to configuration and / or datum defects (Cooper and Cross, 1991). In such cases, N becomes singular and its ordinary inverse does not exist. Configuration defects are sometimes called internal defects, while datum defects are termed external defects.

The datum defects or datum problem arises when the required parameters are not estimable from the measurements, because the datum or reference system for the parameters is incompletely defined by the measurements. The parameters are therefore not estimable from the measurements. Configuration defects are caused by insufficient measurements for unique determination of size and shape of the network. Normally, datum defects are handled by means of constraints (Koch, 1987), whilst configuration defects can be removed (during the design stage) by introducing additional measurements.

2.2.2 Datum defects and definition

Fortunately, the causes of datum defects are usually known, and any datum defects of N can be removed by defining a proper datum. For a 3-D network, the datum definition requires seven datum elements, three for translation, three for rotation and one for scale. The datum is defined by specifying the minimum number of required datum elements, which in fact is equal to the datum defect of N.

For example, assuming the measurements do not contain any datum information, the datum can be simply defined by fixing six coordinates of two points and one coordinate of another (non-collinear) point. Alternatively, the datum can be defined by means of other combinations as well. The following notation is used in this chapter

t_x=translation along x-axis (2.32)

t_y=translation along y-axis

t_z=translation along z-axis

r_x=rotation about x-axis (rotation matrix R_1)

r_y=rotation about y-axis (rotation matrix R_2)

r_z=rotation about z-axis (rotation matrix R_3)

$s=\lambda$=scale

d=datum defect

Some measurements contain datum information which defines datum elements, and will consequently reduce the number of required datum elements (i.e. datum defect). For example distance measurements provide scale, while zenith angles provide orientation in x and y axes. A typical 3-D network comprising of slope distances, zenith angles and horizontal directions will provide three datum elements (Table 2.1), and hence leaves four datum elements undefined (three translation and one rotation about z-axis).

In general, the number of required datum elements (or datum defects) depends on the dimension of the network and types of measurements, as illustrated in Table 2.1. The relationship between the number of datum defects d, rank and the order of N (u) is

$$d=\text{order }(N) - \text{rank }(N) \tag{2.33}$$

or $\quad \text{rank }(N)=\text{order}(N)-d=u-d$

Mathematically, datum definition is carried out by adding a minimum number of constraint equations to the observation equations for removal of rank deficiency. In such a case, the datum is called minimum constraints datum. It is well known in LSE that the coordinates and their cofactor matrix are datum dependent (section 2.1.4). Therefore, different choices of constraints for datum definition will lead to different solutions of $\hat{x}$ and $Q_{\hat{x}}$.

The most common and useful choices of datum for the monitoring network are minimum constraints (or zero variance computational base), minimum trace (or inner constraints or free network) and partial minimum trace datums. The solutions are called minimum constraints, minimum trace and partial minimum trace solutions respectively. A more detailed explanation can be found in Caspary (1987b) and Biacs (1989).

The general constraint equation (Biacs, 1989) to define the missing datum information of the network can be written as

type of network	type of observations	defined datum elements	defect d	required datum elements
1-D	height differences	s	1	t_z
2-D planar	distances	s	3	t_x, t_y, r_z
	azimuths	r_z	3	t_x, t_y, s
max d=4	coordinates	t_x, t_y, r_z, s	-	-
	coordinate differences	r_z, s	2	t_x, t_y
	horizontal angles or directions or distance ratios	-	4	t_x, t_y, r_z, s
3-D spatial	distances or distance differences	s	6	t_x, t_y, t_z, r_x, r_y, r_z
max d=7	azimuths	r_z	6	t_x, t_y, t_z, r_x, r_y, s
	directions or horizontal angles or zenith or vertical angles	r_x, r_y	5	t_x, t_y, t_z, r_z, s
	height differences	r_x, r_y, s	4	t_x, t_y, t_z, r_z
	coordinates	t_x, t_y, t_z, r_x, r_y, r_z, s	-	-
	coordinate differences	r_x, r_y, r_z, s	3	t_x, t_y, t_z
	distance ratios or distance with scale bias or photo coordinates	-	7	t_x, t_y, t_z, r_x, r_y, r_z, s
	distance, zenith angle and horizontal angle or direction	r_x, r_y, s	4	t_x, t_y, t_z, r_z

Table 2.1. Surveying observables and the datum elements.

$$C^t x = 0 \tag{2.34}$$

where $C = I_p G$ and $C^t = G^t I_p$

C and G are called general and inner constraints matrices respectively. I_p is a diagonal matrix with values of unity for datum stations and zero for non-datum stations. For a 3-D network with m stations, the maximum dimensions of the matrices are C (3m,7), I_p (3m,3m diagonal), G (3m,7) and x (3m,1).

In engineering surveying, matrix G^t (shown in Figure 2.3) is well known. The first three rows of G^t define the translation along x, y and z respectively. The next three rows define the rotations about x, y and z, respectively, while the last row defines the scale of the network. If the observations contain datum information, the rank deficiencies are less than seven (i.e. number of rows of G^t), and the corresponding rows of G^t are omitted.

$$G^t = \begin{bmatrix} +1 & 0 & 0 & +1 & 0 & 0 & \dots & +1 & 0 & 0 \\ 0 & +1 & 0 & 0 & +1 & 0 & \dots & 0 & +1 & 0 \\ 0 & 0 & +1 & 0 & 0 & +1 & \dots & 0 & 0 & +1 \\ 0 & +z_1 & -y_1 & 0 & +z_2 & -y_2 & \dots & 0 & +z_n & -y_n \\ -z_1 & 0 & +x_1 & -z_2 & 0 & +x_2 & \dots & -z_n & 0 & +x_n \\ +y_1 & -x_1 & 0 & +y_2 & -x_2 & 0 & \dots & +y_n & -x_n & 0 \\ +x_1 & +y_1 & +z_1 & +x_2 & +y_2 & +z_2 & \dots & +x_n & +y_n & +z_n \end{bmatrix}$$

Figure 2.3 Full components of matrix G^t for a 3-D network

Let the number of coordinates used for datum definition of m stations in a 3-D network be mdat, and datum defect is d. Also let mm be any number between d and 3m. The three types of the above datums can be easily realised as follows:

mdat	type of datum
3m	minimum trace
d	minimum constraints
mm	partial minimum trace

The application of constraints (equation 2.34) to the observation equations leads to the bordering of the singular normal equations in the form

$$\begin{bmatrix} A^tWA & C^t \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ k \end{bmatrix} = \begin{bmatrix} A^tWb \\ 0 \end{bmatrix} \tag{2.35}$$

The use of equation (2.35) implies that the sum of the corrections to the provisional coordinates will be zero, and the network will have zero rotation and zero scale change with respect to the selected datum points (Biacs, 1989).

The solution for $\hat{x}$ and $Q_{\hat{x}}$ (Caspary, 1987b; Biacs, 1989; Cooper, 1994) becomes:

For a minimum trace datum

$$\hat{x}=(N+GG^t)^{-1}A^tWb \tag{2.36}$$

$$Q_{\hat{x}}=(N+GG^t)^{-1}N(N+GG^t)^{-1} \tag{2.37}$$

$$Q_{\hat{x}}=(N+GG^t)^{-1}-G(G^tGG^tG)^{-1}G^t=N^+ \tag{2.38a}$$

where $N=A^tWA$

If the coordinates for computing G are reduced to the centroid and G is normalized (equations 2.41 to 2.43), expression for $Q_{\hat{x}}$ becomes

$$Q_{\hat{x}}=(N+GG^t)^{-1}-GG^t \tag{2.38b}$$

For a partial minimum trace datum

$$\hat{x}=(N+CC^t)^{-1}A^tWb \tag{2.39}$$

$$Q_{\hat{x}}=(N+CC^t)^{-1}-G(G^tCC^tG)^{-1}G^t=N^r \tag{2.40}$$

N^+ in equation (2.38a) is the pseudo inverse or Moore-Penrose inverse, while N^r in equation (2.40) is called the symmetrical reflexive generalized inverse (Koch, 1987; Biacs, 1989; Appendix D). Both inverses are computed indirectly as shown above. Equations (2.38) and also (2.40) are used instead of equation (2.37) for computing the cofactor matrix, since they are more efficient computationally. This is because the matrix to be inverted in the second term has maximum dimension of 7, hence avoiding the multiplication of large matrices.

To achieve numerical stability (Caspary, 1987b), the approximate or provisional coordinates used for computing matrix G^t are reduced to the centroid (centre of gravity) of the network, followed by the normalization of G^t. Reduction to the centroid for a network of m stations can be written as

$$x_i'=x_i-x_o,\ y_i'=y_i-y_o,\ z_i'=z_i-z_o \tag{2.41}$$
$$[x_o\ y_o\ z_o]=[\Sigma x_i/m\ \ \Sigma y_i/m\ \ \Sigma z_i/m]$$

where (x_o, y_o, z_o) defines the centroid of the network. The normalization is

$$G_n^t=(G^tG)^{-1/2}G^t \text{ and } G_n^tG_n=I \tag{2.42}$$

where $(G^tG)^{-1/2}=[(G^tG)^{1/2}]^{-1}$

$$G^tG=(G^tG)^{1/2}[(G^tG)^{1/2}]^t \tag{2.43}$$

Factorization of equation (2.43) is easily achieved by Cholesky factorization.

Further derivation and computational details are given in Cooper and Cross (1988, 1991). Due to datum dependencies, the estimated $\hat{x}$ and $Q_{\hat{x}}$ are biased, depending on the selected constraints. However, some quantities are datum invariant (section 2.1.4). These invariant quantities play an important role in reliability analysis and statistical evaluation of LSE results.

In practice, the conventional minimum constraints datum is adopted due to its simplicity (Cooper and Cross, 1988). The solution can be obtained either by fixing d coordinates of the network, or using 'fixed' pseudo observations with realistic a priori variances.

By holding fixed a minimum number of coordinates (the number is equal to d) linearly dependent rows and columns of N are deleted. Such approach removes the selected fixed coordinates from the system of equations to form a reduced normal equation set, resulting in a non-singular system of equations as follows (Caspary, 1987b)

$$\begin{bmatrix} A_1 & A_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b+v \end{bmatrix} \tag{2.44}$$

where x_2 are the fixed coordinates used to define the datum (known as zero variance computational base), their number equals the rank defect d. The solution becomes

$$\hat{x}_1=(A_1{}^tWA_1)^{-1}A_1{}^tWb,\ Q_{\hat{x}1}=(A_1{}^tWA_1)^{-1} \tag{2.45a}$$

$$\hat{x}_2=0,\ Q_{\hat{x}2}=0 \tag{2.45b}$$

In this manner, the minimum constraints solution (equation 2.45a) is based on equations (2.13) to (2.26). The choice of the fixed coordinates is arbitrary. Moreover, minimum constraints by fixed distance or azimuth (Cooper, 1987) is possible too.

Suitable pseudo observations with realistic a priori variances will be 'fixed', and may be used to remove the datum defects (Koch, 1987; Biacs, 1989; Chen et al, 1990a). For example, fixed azimuth provides the unknown rotation parameters of the network about z-axis. Let

$$l=[l_1\ \ l_2]^t \tag{2.46}$$

where l_2 are the pseudo observations with weights (W_2). The solution becomes

$$\begin{aligned} &N=(A_1{}^tW_1A_1)+(A_2{}^tW_2A_2)=N_1+\Delta N \\ &u=(A_1{}^tW_1b_1)+(A_2{}^tW_2b_2)=u_1+\Delta u \\ &\hat{x}=N^{-1}u,\ Q_{\hat{x}}=N^{-1} \end{aligned} \tag{2.47}$$

In equation (2.47), the addition of ΔN changes the singular N_1 into non-singular N.

The solution via a minimum trace datum (equations 2.36 to 2.38) has the property of minimum trace and minimum norm, i.e

$$\hat{x}'\hat{x}=\min,\ \mathrm{tr}\ (Q_{\hat{x}})=\min \tag{2.48}$$

According to Cooper and Cross (1991), the datum is defined by the approximate coordinates of all the stations used. After LSE, the centroid, as well as the average direction and average distance (or scale) of all points from the centroid remain constant (equation 2.35). The solution by partial minimum trace datum minimizes both the partial trace and partial norm. Such a solution is useful for deformation detection purposes, because only a subset of stations considered as stable is used to define the datum.

In this study, the minimum constraints solution with fixed coordinates (equations 2.44 and 2.45a) has been adopted. Moreover, a simple procedure has been formulated to expand the reduced $Q_{\hat{x}1}$ to its full size, by appropriate re-ordering and filling with respect to the fixed datum stations. If required, it is possible to transform the minimum constraints solution into the minimum trace, partial minimum trace or other minimum constraints solution via S-transformations (section 2.3).

2.2.3 Qualitative analysis of linear equations

In the method of LSE using observation equations (section 2.1.3), the normal equation matrix N (equations 2.17 and 2.30) is square, symmetric, positive definite, non-singular and of full rank (i.e. its dimension and rank are equal to the dimension of parameters $\hat{x}$). In other words, the linear equation is consistent. The matrix is full rank and N has a unique inverse only if it is square and non-singular (Appendix D)

$$N^{-1}N=NN^{-1}=I \tag{2.49a}$$

It is therefore important to analyse and determine if A (or N) is of full rank. Several quantitative measures can be computed for analysis purposes, such as rank, condition and determinant. Further details on matrix algebra are given by Mikhail (1976), Mikhail and Gracie (1981) and Golub and Loan (1990).

The rank of matrix A (dimensions n x u, where $n \geq u$), denoted by k, is the maximum number of linearly independent rows or columns in the matrix. The matrix is full rank if k = min (n,u), and rank deficient if k < min (n,u). For a square matrix (where n=u), A is non-singular if k=n, and singular if k < n.

Theoretically, the rank of a matrix can be determined (Rawlings, 1988) by Gaussian elimination using elementary row and column operations to reduce it to an equivalent matrix (echelon form). Elements below the diagonal are reduced to zero. The rank of the matrix is the number of non-zero elements remaining on the diagonal.

The condition of a matrix indicates the stability or sensitivity of the solution, and is measured by the condition number. The condition number of a matrix A (of full rank) is the ratio of the largest to the smallest singular values of A (Forsythe et al, 1977; Lawson and Hanson, 1974; equation 2.51).

If the condition number is too large, the system is ill-conditioned or nearly singular, otherwise it is well-conditioned. Methods of computing the condition number are given in Golub and Loan (1990) and Forsythe et al (1977).

The determinant (denoted by det) of a square matrix is a scalar quantity. For matrix N (of dimensions uxu), it is equal to the sum of the products of the elements of the first row of N and their corresponding cofactors (Mikhail and Gracie, 1981)

$$\det N = n_{11}D_{11} + n_{12}D_{12} + \ldots + n_{1u}D_{1u} \qquad (2.49b)$$

where n is element of the first row of N and D is the respective cofactor.

In theory, the determinant of a square matrix is non-zero if it is non-singular, and zero if it is singular. However, Golub and Loan (1990) show that there is little relationship between determinant and condition, as a well-conditioned matrix can have a very small determinant, or vice versa. For this reason, throughout this research, only the rank and condition are used in the qualitative analysis of the linear equations.

2.2.4 Concept of SVD

The solutions of LSE are accomplished routinely on a computer, using double precision floating-point arithmetic. Hence the computations are generally affected by round-off errors (Forsythe et al, 1977; Dyck et al, 1984) because the numbers cannot be represented to their full precision, due to the process of rounding or truncating.

In practice, the determination of the rank of a matrix is not an easy task. In some cases, Gaussian elimination can transform a rank deficient matrix into a full rank matrix (Stewart, 1973). Also, the presence of round-off error may produce the same effect.

In numerical analysis, SVD is a powerful computational tool for analyzing linear equations because it reveals qualitative information about the structure of the matrices, especially rank and condition. The technique is very effective for handling rank deficiency in the presence of round-off errors. The theoretical aspects on SVD can be found in Golub and Loan (1990), Forsythe et al (1977), Lawson and Hanson (1974) and Press et al (1988).

The SVD of a real matrix A (dimensions nxu) is the factorization of

$$A = U\Sigma V^t \qquad (2.50)$$

where U (dimensions nxn) and V (dimensions uxu) are orthogonal matrices, i.e. both U^tU and V^tV are equal to unit matrix. The columns of U and V are called the left and right singular vectors respectively. Matrix Σ (dimensions nxu) has non-negative elements on the diagonal (of uxu sub-matrix) and zeros elsewhere. The diagonal elements of Σ are the singular values of A. The above orthogonal transformation is important because the orthogonal matrix is non-singular (i.e. full rank), and the rank of the diagonal matrix is equal to the number of its non-zero diagonal elements.

Equation (2.50) above involves the actual sizes of the matrices. In practice (Lawson and Hanson, 1974), the maximum number of singular values is u. Hence, the matrices required (i.e. economy size) in computing the SVD are A (n,u), U (n,u), Σ(uxu diagonal matrix) and V (u,u).

Clearly, the rank of A (denoted by k) is the number of non-zero singular values (denoted by s). A is full rank if k is equal to u, and rank deficient if k is less than u. Moreover, the condition number of A is the ratio of the largest to the smallest singular values. Hence

$$\text{rank}\ (A)=k \tag{2.51}$$

$$\text{cond}\ (A)=s_{max}/s_{min} \tag{2.52}$$

where s_{max} and s_{min} are maximum (largest) and minimum (smallest) singular values respectively. Integer k is the number of non-zero singular values.

If A is rank deficient, the condition number can be considered as infinite. Moreover, if the matrix is ill-conditioned, the condition number will be very large, although the matrix is of full rank.

The SVD algorithm is mostly based on Golub and Reinsch (1970). The objective is to determine an orthogonal U and V so that

$$U^tAV = \Sigma \text{ is diagonal} \tag{2.53}$$

The algorithm consists of two stages. First, matrix A is reduced to superdiagonal form using Householder's bidiagonalization. Then, the superdiagonal elements are reduced iteratively using Francis' QR algorithm to a neglible size, leaving the desired diagonal matrix. This algorithm is very fast and effective. Further discussions on these aspects can be found in Forsythe et al (1977) and Golub and Loan (1990).

For proper and effective use of SVD, it is required to set a tolerance or limit, which reflects the accuracy of the data and any round-off error. For example (Forsythe et al, 1977)

$$\text{tolerance} = (\text{accuracy of data})*(\text{largest singular value})$$

In practice, the computed rank (called numerical or effective rank) is the number of singular values greater than the tolerance. Consequently, SVD is a very stable technique for handling rank defects in the presence of round-off error because the tolerance does not enter into the decomposition process.

The numerical rank and condition number computed via SVD (equations 2.51 and 2.52) is very useful for analyzing linear equations. By this means, it is possible to determine whether the system is rank deficient or ill-conditioned, prior to the computation of the inverse.

2.2.5 SVD solution

SVD can also be used to solve the least squares problems. Golub and Loan (1990) show that the SVD solution minimizes the sum of squared residuals, norm and the variance, and hence a solution identical to LSE is obtained. SVD can also handle both full rank and rank defect systems.

The SVD solution can be computed in two ways, by factorization or via a pseudo inverse. The factorization is similar to LU factorization, and involves manipulation of diagonal matrices (Forsythe et al, 1977). After SVD analysis, the following equations are employed

$$A = U\Sigma V^t;\ z = V^t x \text{ and } d = U^t b \tag{2.54}$$

Hence d, z and $\hat{x}$ are solved respectively using

$$d=U^t b,\ \Sigma z=d \text{ and } \hat{x}=Vz \tag{2.55}$$

In the above equations, if any singular value s_i is less than the tolerance, the corresponding z_i must be set to zero.

In engineering applications, it is required to compute the cofactor matrix of the parameters, $Q_{\hat{x}}$. For this reason, a direct solution through the computation of pseudo inverse can be employed. Fortunately, SVD parameters can also be used to evaluate the inverse. The expression for pseudo inverse of A (Forsythe et al, 1977) can be written as

$$A^+= V\Sigma^+U^t \tag{2.56}$$

where $\Sigma^+ = [s_1^+, s_2^+ \ldots s_u^+]$ is the (u,u) diagonal matrix

$s_i^+ = (1/s_i)$ if s_i is greater than the tolerance,

otherwise it is zero

A^+ satisfies all four Moore-Penrose equations (Golub and Loan, 1990; Appendix D), i.e. $AA^+A=A$, $A^+AA^+=A^+$, $(AA^+)^t=AA^+$ and $(A^+A)^t=A^+A$. If A is square and nonsingular, $A^+=A^{-1}$ is the ordinary inverse. The solution for the parameters, their cofactor matrix and residuals then becomes

$$\hat{x}=A^+b \tag{2.57}$$
$$Q_{\hat{x}}=A^+(A^+)^t$$
$$v=A\hat{x}-b$$

An important relationship occurs when A is square and symmetric. According to Lawson and Hanson (1974), A has an eigenvalue-eigenvector decomposition (also known as eigenvalue decomposition (EVD))of the form

$$A=QEQ^t \tag{2.58}$$

where Q is orthogonal and E is diagonal. This is similar to equation (2.50) for SVD. The diagonal elements of E are the eigenvalues of A and the column vectors of Q are the eigenvectors of A. This relationship is useful because the normal equation matrix is usually square, symmetric and positive definite (section 2.2.3). Hence, based on equations (2.50) and (2.58), the eigenvalues of A are equal to its singular values, and the eigenvectors are the columns of U.

Both SVD and EVD give the same qualitative analysis and can be used for analyzing and solving the normal equations. The solution for equation (2.16) via SVD and EVD can be summarized as

$$Nx = A^tWb \tag{2.59}$$
$$N = U\Sigma U^t \text{ or } N = QEQ^t \tag{2.60}$$
$$N^+ = U\Sigma^+U^t \text{ or } N^+ = QE^+Q^t \tag{2.61}$$
$$\hat{x} = N^+(A^tWb) \text{ and } Q_{\hat{x}} = N^+ \tag{2.62}$$

In terms of computer storage, if the dimension of N is (u,u), SVD uses N (u,u), U(u,u), U^t(u,u) and Σ(uxu diagonal). The storage needed for EVD is N(u*(u+1)/2), Q(u,u) and E(uxu diagonal). This shows that the use of eigenvalue decomposition reduces the storage.

Another important application of EVD is related to the establishment of the confidence regions around the estimated coordinates for the assessment of precision (section 3.3.5). The confidence region for the 3-D case is described by an ellipsoid.

The function describing the ellipsoid (Vanicek and Krakiwsky, 1986) for a point with coordinates $\hat{x}$ (actually 3-D) and cofactor matrix $Q_{\hat{x}}$ is

$$y=(\hat{x}-x_c)^t Q_x^{-1}(\hat{x}-x_c) \tag{2.63a}$$

with the probability that the point lies within the ellipsoid as

$$\text{probability } (\chi^2_3 < y) = 1-\alpha \tag{2.63b}$$

where y is the expansion factor, x_c the centre of ellipsoid, α the significance level and $(1-\alpha)$ the confidence level or probability.

The parameters of the ellipsoid (semiaxes and orientation) can be computed by EVD of $Q_{\hat{x}}$ as

$$Q_{\hat{x}}=QEQ^t \tag{2.64}$$

where the diagonal elements of E are eigenvalues of $Q_{\hat{x}}$ and column vectors of Q are eigenvectors of $Q_{\hat{x}}$. The semiaxes of the ellipsoid are the eigenvalues, while the respective orientations are the eigenvectors. Further details are given in section 3.3.5.

2.2.6 Using EVD for rank defect analysis

The least squares criterion mentioned earlier provides a unique solution to an overdetermined system when the system is of full rank (section 2.1.3). However, in general, the system is rank deficient (section 2.2.1). The rank of A (dimension nxu) is therefore

$$\text{rank}(A) = k = u - \text{defect} \tag{2.65}$$

and $\quad u = k + \text{defect}\ ;\ \text{defect} = d_d + d_c$

where d_d (or simply d) is the datum defect and d_c is the configuration defect.

In practice, it is assumed that A is not full rank because of the datum defect only, and consequently free from configuration defect. The handling of datum defects are described in section 2.2.2.

Both SVD and EVD (equations 2.54 to 2.62) may be used for simultaneous analysis and solution of the linear equations. However, ordinary LSE based on equations (2.13) to (2.26) is commonly adopted in practice. For this reason, EVD is used for rank defect analysis of N, prior to the inversion process.

In this research, a strategy has been developed for incorporating EVD (i.e. simplified EVD) into ordinary LSE (Setan, 1993a). The developed strategy can now be summarised. Initially, using minimum constraints, any datum defects (d_d) are removed by means of equations (2.44) and (2.45a). These constraints remove the rows and columns of A that correspond to fixed stations. Qualitative analysis on the reduced normal equations via EVD (equation 2.58) based on rank and condition (equations 2.51 and 2.52) can then be used to determine whether the system is full rank or rank deficient.

If the normal equations are full rank and the condition is not too large, the network is considered as free from the configuration defects, and ordinary LSE is performed. On the other hand, a rank deficient system or a large condition number indicates that the network suffers from configuration defects or ill-conditioning respectively. The number of defects (d_c) reflects the minimum number of additional measurements required to handle the configuration defect. It is then necessary to examine the network and add the relevant measurements where necessary. In the case of ill-conditioning, the observations with very small redundancy numbers (section 3.8) indicate the weak area of the network. The rank analysis is performed at each iteration.

From equations (2.51) and (2.52), only eigenvalues are required for computations of rank and condition. Therefore, only eigenvalues need be computed, as eigenvectors are not needed. Such procedure will speed up computational time and save storage, and it is called a simpified EVD.

2.3 S-transformations and datum re-definition

The three dimensional coordinates and their cofactor matrix obtained from LSE of each epoch for deformation detection are datum dependent (section 2.1.4), and must be referred to a common datum. During the process of deformation detection, it is also required to re-define the datum with respect to a set of stable points. Consequently, a facility to allow for changes of datum or computational base is needed.

The transformation of LSE results and datum re-definition can be carried out via the similarity covariance transformation (S-transformation). In terms of implementation, direct evaluation of S-transformations is time consuming as it involves multiplication of large matrices and hence is not practical. In this study, a special computational procedure has been formulated (section 2.3.5) to handle such transformations efficiently. Moreover, the application of S-transformations use the general transformation equation.

The concepts of datum re-definition and S-transformations are described in sections 2.3.1, 2.3.2 and 2.3.3. The uses of S-transformations in LSE and deformation detection are presented in section 2.3.4.

The formulated computational procedure (section 2.3.5) involves reducing the coordinates to a centroid and the normalization of matrix G to achieve numerical stability. Decomposition of matrix S is also carried out to speed up the computations. Use of the reduced coordinates is also recommended for flexible S-transformations.

2.3.1 Need for datum re-definition

The required variables estimated from LSE of each epoch for the purpose of deformation detection (section 3.9) are estimated three dimensional coordinates ($\hat{x}_a$), their cofactor matrix ($Q_{\hat{x}}$), estimated variance factor ($\hat{\sigma}_o^2$), degrees of freedom (r) and datum defect (d). For simplicity, x and Q_x will be used to represent $\hat{x}_a$ and $Q_{\hat{x}}$ respectively throughout this chapter.

Both x and Q_x are datum dependent. For the purpose of deformation detection, LSE can be based on the minimum trace, minimum constraints or partial minimum trace datum (section

2.2.2). Datum invariant quantities ($\hat{\sigma}_o^2$, r, d) remain the same because of their datum independence property (section 2.1.4).

The concept of datum definition (section 2.2.2) is readily applicable in the monitoring of deformation. The observations at each epoch are processed independently by LSE to estimate x and Q_x. In general, the monitoring network is treated as free network where all stations are assumed to be unstable a priori, and hence a minimum trace datum is used. In some cases, a set of stable points is known in advance, and in this case a partial minimum trace datum can be used. In practice, the conventional minimum constraints datum is favoured due to its simplicity. In this case S-transformations of x and Q_x into either minimum or partial minimum trace datum is needed.

In the initial stage of deformation detection, x and Q_x of any two epochs are differenced to estimate displacement vectors and their cofactor matrix. Theoretically, x and Q_x have to be referred to the same common datum. However, different datum definitions may be necessary for each epoch, possibly because of different defects in the configuration or practical limitations (such as obstruction of the line of sight or destruction of points).

The solution with respect to a common datum can be obtained either from LSE of each epoch where the new datum is defined by a common set of points, or via S-transformations (section 4.2.1) of x and Q_x of each epoch to the new datum. The S-transformations approach is very useful as it replaces the repeated LSE and inversion of the normal equations coefficient matrix. Moreover, during the localization of deformation (section 4.2.3.2.3), S-transformations are used repeatedly for transforming the displacement vectors and their cofactor matrix with respect to new datums defined by different sets of stable points. This approach is analagous to a partial minimum trace solution.

2.3.2 Concept of S-transformations

The S-transformation is based on the work of Baarda carried out in the 1950s which was published later in Baarda (1973).

The following formulation for S-transformations is adopted from Strang Van Hees (1982). In general, it is based on a fundamental linear equation in the form of

$$x=Gp \tag{2.66}$$

or $dx_j=dx_i+Gdp$ in differential form

For a 3-D network, the equation for transforming coordinates x_i to x_j is (Cooper and Cross, 1991)

$$x_j=x_o+\lambda R_1R_2R_3x_i \tag{2.67a}$$

or $\quad x_j=x_o+\lambda Rx_i$

Equation (2.67a) contains a maximum of seven transformation parameters: three translation in x_o, three rotation in R and a scale factor λ. The rotation matrices R_1, R_2 and R_3 are with respect to the x, y and z axes respectively (equation 2.32).

$$R_1=\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\alpha & \sin\alpha\\ 0 & -\sin\alpha & \cos\alpha\end{bmatrix},\ R_2=\begin{bmatrix}\cos\beta & 0 & -\sin\beta\\ 0 & 1 & 0\\ \sin\beta & 0 & \cos\beta\end{bmatrix},\ R_3=\begin{bmatrix}\cos\gamma & \sin\gamma & 0\\ -\sin\gamma & \cos\gamma & 0\\ 0 & 0 & 1\end{bmatrix} \tag{2.67b}$$

Differentiation of the above equation for one point gives

$$\begin{bmatrix}dx\\ dy\\ dz\end{bmatrix}_j=\begin{bmatrix}dx\\ dy\\ dz\end{bmatrix}_o+\begin{bmatrix}d\lambda & d\gamma & -d\beta\\ -d\gamma & d\lambda & d\alpha\\ d\beta & -d\alpha & d\lambda\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}_i+\begin{bmatrix}dx\\ dy\\ dz\end{bmatrix}_i \tag{2.68}$$

or $\quad dx_j=dx_i+Gdp$

where

$$G=\begin{bmatrix}1 & 0 & 0 & 0 & -z_i & y_i & x_i\\ 0 & 1 & 0 & z_i & 0 & -x_i & y_i\\ 0 & 0 & 1 & -y_i & x_i & 0 & z_i\end{bmatrix}$$

$dx_j=[dx\ dy\ dz]_j^t$

$dx_i=[dx\ dy\ dz]_i^t$

$dp=[dx_o\ dy_o\ dz_o\ d\alpha\ d\beta\ d\gamma\ d\lambda]^t$

For a 3-D network with m stations and a maximum datum defect (d) of seven, the dimensions are $dx_j(3m,1)$, $dx_i(3m,1)$, $G(3m,7)$ and $dp(7,1)$. A solution for dp can be obtained using a generalised inverse of G denoted by G^g. Hence

$$dp=G^g(dx_j-dx_i) \tag{2.69}$$

$$dx_j=dx_i+GG^g(dx_j-dx_i)$$

and $$(I-GG^g)dx_j=(I-GG^g)dx_i$$

The relationship between cofactor matrices is obtained by multiplying each side of the above equation with their respective transpose.

$$(I-GG^g)Q_{xj}(I-GG^g)^t=(I-GG^g)Q_{xi}(I-GG^g)^t \tag{2.70}$$

By using the g condition (Appendix D), which is satisfied by all generalised inverses, the fundamental equation of S-transformations may be derived for transforming x_i and Q_{xi} to x_j and Q_{xj} according to a generalised inverse G^g.

$$Q_{xj}=(I-GG^g_j)Q_{xi}(I-GG^g_j)^t$$

or in a simple form

$$x_j=S_jx_i \tag{2.71}$$
$$Q_{xj}=S_jQ_{xi}S_j^t$$
$$S_j=(I-GG^g_j)$$

In order to transform x and Q_x into a minimum trace datum, the pseudo inverse G^+ is employed. Therefore

$$S_p=(I-GG^+) \tag{2.72}$$
$$G^+=(G^tG)^{-1}G^t$$
$$S_p=(I-G(G^tG)^{-1}G^t)$$

In a minimum trace datum, all points are used for datum definition. Therefore, a more general expression for the S-transformations of any arbitrary x_i and Q_{xi} into x_j and Q_{xj} becomes

$$x_j=S_jx_i \tag{2.73}$$
$$Q_{xj}=S_jQ_{xi}S_j^t$$
$$S_j=(I-G(G^tI_jG)^{-1}G^tI_j) \text{ or}$$
$$S_j=(I-G(C^tG)^{-1}C^t) \text{ if } C=I_jG$$

In equation (2.73), I_j (3mx3m) is a diagonal matrix for defining the computational base after S-transformations. If only some of the points are used for datum definition (partial minimum trace datum), the elements of I_j for datum and non-datum points are one and zero respectively. Dimensions of the S-transformations matrices involving m stations and maximum datum defect are x_i (3m,1), x_j (3m,1), Q_{xi} (3m,3m symmetric), Q_{xj} (3m,3m symmetric), G (3m,7), I (3m,3m diagonal), C (3m,7), C^tG (7,7) and S_j (3m,3m generally non-symmetric). Hence, one only need to invert a maximum of a (7 by 7) matrix, which is quite simple.

A closer look at matrix G (equation 2.68) shows that it is actually the same as the inner constraints matrix in equation (2.34). The number of its columns is equal to the number of datum defect. As shown in section 2.2.2, depending on the type of observations, G will have corresponding fewer columns whilst (C^tG) is still of full rank.

The expression for S-transformations can also be obtained by appplying the concept of propagatian of variance and the g condition (Caspary, 1987b). Let the solutions with respect to two different constraints (equation 2.37) be

$$x_1 \text{ and } Q_{x1}=(N+C_1C_1^t)^{-1}N(N+C_1C_1^t)^{-1} \tag{2.74}$$
$$\text{with } C_1^tx=0 \text{ [datum 1]}$$
$$x_2 \text{ and } Q_{x2}=(N+C_2C_2^t)^{-1}N(N+C_2C_2^t)^{-1}$$
$$\text{with } C_2^tx=0 \text{ [datum 2]}$$

The two cofactor matrices can be connected by means of g condition (Appendix D) which is correct for any Q_x

$$\text{g condition: } NQ_xN=N \tag{2.75}$$

Hence the cofactor matrix with respect to datum 2 is

$$Q_{x2}=(N+C_2C_2^t)^{-1}N(N+C_2C_2^t)^{-1} \qquad (2.76)$$

$$=(N+C_2C_2^t)^{-1}NQ_{x1}N(N+C_2C_2^t)^{-1} \text{ because } NQ_{x1}N=N$$

or $\quad Q_{x2}=S_2Q_{x1}S_2^t$ where $S_2=(N+C_2C_2^t)^{-1}N$

The above equation is of the type obtained from the general law of propagation of variance applied to a linear function y, where

$$\text{If } y=Ax,\ Q_y=AQ_xA^t \qquad (2.77)$$

Hence, $x_2=S_2x_1$

The transformation of x_i and Q_{xi} into x_j and Q_{xj} is similar to equation (2.73) above

$$x_j=S_jx_i \qquad (2.78)$$

$$Q_{xj}=S_jQ_{xi}S_j^t$$

$$S_j=(N+CC^t)^{-1}N=(I-G(C^tG)^{-1}C^t) \text{ where } C=I_jG$$

2.3.3 Properties of S-transformations

In general, the square matrix S is non-symmetric, and only symmetric for a minimum trace datum. Two important properties of S-transformations are :

(i) S is idempotent

For successive S-transformations, only the last transformation determines the resulting x and Q_x.

Let $S=S_kS_j$ i.e. first S_j then S_k (equation 2.71)

$$S=(I-GG_k)(I-GG_j)$$

$$=(I-GG_k-GG_j+GG_kGG_j)$$

applying g condition, $GG_kG=G$

$$S=(I-GG_k-GG_j+GG_j)=(I-GG_k) \tag{2.79}$$

hence $S_kS_j=S_k$

Since S is idempotent, rank (S)= trace (S)

<u>(ii) Product of SG is zero.</u>

This property is useful for checking the computation of the S matrix.

$$SG=(I-GG^g)G$$
$$=G-GG^gG$$
$$=G-G \text{ as g condition gives } GG^gG=G$$
$$SG=0 \tag{2.80}$$

Another important aspect is that the datum invariant quantities (section 2.1.4) do not change between either minimum trace, partial minimum trace or minimum constraints datums.

2.3.4 Application in deformation detection

S-transformations can be applied in both LSE and deformation detection. As will be shown below, the general S-transformations equation can be applied directly.

In LSE (section 2.2.2) the types of solutions can be based on either a minimum trace, partial minimum trace or minimum constraint datum. The transformation equations for transforming x_i and Q_{xi} into x_j and Q_{xj} based on the chosen datum are given by equation (2.73) as

$$x_j=S_jx_i,\ Q_{xj}=S_jQ_{xi}S_j^t \tag{2.81}$$
$$S=(I-G(C^tG)^{-1}C^t),\ C=I_jG$$

Let m be the number of stations in the 3-D network, d be the datum defect and mm the number of coordinates chosen for datum definition. The elements of I_j will be unity for datum points and zero for other points. The solutions (section 2.2.2) can be realised as if mm=3m, minimum trace datum (2.82)

mm=d, minimum constraints datum

mm between d and 3m, partial minimum trace datum

In most cases, it may be necessary to divide the points into two groups, datum and non-datum points. Fraser and Gruendig (1985) and Cooper (1987) adopted the partitioning and ordering of x_i and Q_{xi} with respect to datum and non-datum points as the following:

$$x_i=[x_r \quad x_e]^t \quad ; Q_{x_i}=\begin{bmatrix} Q_r & Q_{re} \\ Q_{er} & Q_e \end{bmatrix} \tag{2.83}$$

where r (retain) refers to datum points and e (eliminate) refers to other non-datum points. Hence, equation (2.81) becomes

$$x_j=Sx_i \text{ and } Q_{xj}=SQ_{xi}S^t \tag{2.84a}$$

$$S=I-\begin{bmatrix} G_r(G_r^tG_r)^{-1}G_r^t & 0 \\ G_e(G_r^tG_r)^{-1}G_r^t & 0 \end{bmatrix} \tag{2.84b}$$

This partitioning approach requires re-ordering of x, Q_x, G_r and G_e.

Prior to deformation detection, x and Q_x of any two epochs must be referred to a common datum defined by sets of common points (section 4.2.1). Assume that each epoch has different stations and datum definitions. Let the coordinates and their cofactor matrices for the two epochs be x_1, Q_{x1} (refers to datum A) and x_2, Q_{x2} (refers to datum B), and it is required to be referred to datum C. Reordering of x and Q_x via equation (2.83) gives

$$x_i=[x_r \quad x_e]^t \quad ;Q_{x_i}=\begin{bmatrix} Q_r & Q_{re} \\ Q_{er} & Q_e \end{bmatrix} \tag{2.85}$$

where x_r and x_e are referred to common (datum) and non-common points respectively. The transformation equations become

$$x_{1c}=S_1x_1;\ Q_{x1c}=S_1Q_{x1}S_1^t; \tag{2.86}$$
$$S_1=(I-G_1(C_1^tG_1)^{-1}C_1^t);\ C_1=I_cG_1$$
$$x_{2c}=S_2x_2,\ Q_{x2c}=S_2Q_{x2}S_2^t;$$
$$S_2=(I-G_2(C_2^tG_2)^{-1}C_2^t);\ C_2=I_cG_2$$

This is the equivalent of obtaining a solution with respect to a partial minimum trace datum. Once x and Q_x are referred to a common datum, the displacement vector (d) and the cofactor matrix of displacement (Q_d) for the common stations can be computed as

$$d=x_{2c}-x_{1c};\ Q_d=Q_{x2c}+Q_{x1c} \tag{2.87}$$

If both epochs utilize the same stations, but need to be referred to the same datum (say datum C), d and Q_d can be determined directly

$$x_{1c}=Sx_1;\ x_{2c}=Sx_2;\ d=x_{2c}-x_{1c}=S(x_2-x_1) \tag{2.88}$$
$$Q_d=Q_{x2c}+Q_{x1c}=S(Q_{x2}+Q_{x1})S^t$$

During the localization of deformation, S-transformations of this type (equation 2.88) are very useful for the iterative transformation of displacement vector and its cofactor matrix into a datum defined by a set of stable points. The S-transformations, used together with a partial congruency test and test of the largest quadratic form, is in effect removing each unstable or suspected point interactively, one at a time. Elements of d and Q_d need to be re-ordered each time a suspected point is removed from the computational base. Details on such procedures are given in section 4.2.3.2.

Another application of S-transformations is demonstrated by Chen et al (1990a) for identification of stable points via a robust method. Using the general S-transformation, I is interpreted as a weighting factor (weight matrix) for obtaining an iterative weighted similarity transformation of d and Q_d. Further aspects on robust methods are discussed in section 4.2.3.4.

The deformation detection procedure developed in chapter 4 uses the general S-transformations formulation effectively for simultaneous identification of stable points and estimation of the deformation of unstable points. This is conceptually correct due to the basic property of the S matrix being idempotent. Hence, once the final datum is defined by the stable points, the final transformation determines the resulting d and Q_d.

2.3.5 Computational procedure

The general equation for S-transformations of x_i and Q_{xi} to x_j and Q_{xj} (equation 2.73) is simply

$$
\begin{aligned}
& x_j=S_j x_i \\
& Q_{xj}=S_j Q_{xi} S_j^t \\
& S_j=(I-G(G^t I_j G)^{-1} G^t I_j) \text{ or} \\
& S_j=(I-G(C^t G)^{-1} C^t) \text{ if } C=I_j G
\end{aligned}
\tag{2.89}
$$

Although the above equation looks simple, its direct implementation is not practical because the transformation matrix S is non-symmetric in general and the computation of Q_{xj} is time consuming.

In terms of storage requirements, for a 3-D network of m stations, the major storage areas are occupied by S, Q_{xi} and Q_{xj} (section 2.3.2). Matrix S is full and non-symmetric (3m,3m), while Q_{xi} and Q_{xj} are symmetric, each requires [(3m)(3m+1)/2] spaces. The main task is in the computation of Q_{xj} as it involves multiplication of large matrices. Another problem is the numerical instability that might occur.

In this study, a computational strategy has been formulated (Setan, 1993b) based on the following criteria in order to produce an efficient implementation of the equation (2.89) into a working computer program:

i. Working with single arrays in most cases and only the triangular matrix Q_x is needed.
ii. Reduction of all approximate coordinates for computing G to their centroid to avoid numerical instability.
iii. Further normalization via Cholesky factorization to achieve numerical stability.
iv. A special procedure by decomposition of S matrix (Biacs, 1989; Biacs and Teskey, 1990) to speed up the computation.
v. Using reduced coordinates during S-transformations.

To save storage, single arrays are used and some of the spaces are used again, the main storage being for S (full matrix), and the triangular matrices Q_{xi} and Q_{xj}. To achieve numerical stability, the provisional coordinates used for computing elements of matrix G^t are reduced to the centroid of the network, followed by normalization of G^t. Reduction to the centroid (equation 2.41) is obtained from

$$x_i'=x_i-x_o,\ y_i'=y_i-y_o,\ z_i'=z_i-z_o \tag{2.90}$$
$$x_o=\Sigma x_i/m,\ y_o=\Sigma y_i/m,\ z_o=\Sigma z_i/m$$

where (x_o, y_o, z_o) are coordinates of the centroid of m stations, which are simply the means of each x_i, y_i and z_i coordinates.

The purpose of normalization of a matrix is to make the norm (or length) or the rows equal to unity. The normalization of G^t (equations 2.42 and 2.43) is

$$G_n^t=(G^tG)^{-1/2}G^t \text{ and } G_n^tG_n=I \tag{2.91}$$
$$(G^tG)^{-1/2}=[(G^tG)^{1/2}]^{-1}$$

The factorization of the square matrix G^tG can be carried out by means of the standard Cholesky factorization in the form $A=U^tU$ where U^t and U are lower and upper triangular matrices respectively. Hence

$$G^tG=(G^tG)^{1/2}[(G^tG)^{1/2}]^t \tag{2.92}$$

Evaluation of the cofactor matrix during S-transformations involves multiplication of large matrices (equation 2.89). To speed up the computation, Biacs (1989) decomposed the S matrix as follows:

$$x_j=S_jx_i,\ Q_{xj}=S_jQ_{xi}S_j^t \tag{2.93}$$
$$S_j=I-G[C^tG]^{-1}C^t,\ C=I_jG$$
$$Q_{xj}=[I-G[C^tG]^{-1}C^t]Q_{xi}[I-C[C^tG]^{-1}G^t]$$
$$=Q_{xi}-G[C^tG]^{-1}C^tQ_{xi}-Q_{xi}C[C^tG]^{-1}G^t+G[C^tG]^{-1}C^tQ_{xi}C[C^tG]^{-1}G^t$$

also $x_j=[I-G[C^tG]^{-1}C^t]x_i$

letting $P^t=[C^tG]^{-1}C^tQ_{xi}$ or similarly $P=Q_{xi}C[C^tG]^{-1}$

and $\quad R^t=[C^tG]^{-1}C^t$

The S-transformations equation can be written as

$$x_j=[I-GR^t]x_i \text{ and } Q_{xj}=Q_{xi}-GP^t-PG^t+GR^tPG^t \tag{2.94}$$

In equation (2.94), the computation of Q_{xj} which involves multiplication of large matrices (equation 2.89) is reduced to an addition of matrices, which can be performed more quickly.

The computational scheme developed for practical applications of S-transformations (section 5.1.2) uses equations (2.83), (2.93) and (2.94), together with the options for datum defects, reduction to centroid and normalization. Hence, the equations for transforming x_i and Q_{xi} to x_j and Q_{xj} are:

$$x_i=[x_r \quad x_e]^t \quad ;Q_{x_i}=\begin{bmatrix} Q_r & Q_{re} \\ Q_{er} & Q_e \end{bmatrix} \tag{2.95}$$

where

$x_j=[I-GR^t]x_i$

$Q_{xj}=Q_{xi}-GP^t-PG^t+GR^tPG^t$

$P^t=[C^tG]^{-1}C^tQ_{xi}$

$R^t=[C^tG]^{-1}C^t$

$C=I_jG$

In equation (2.95), elements of I_j are one and zero for datum and non-datum points respectively. The partitioning procedure is adopted here to simplify the uses of S-transformations in localization of deformation (section 4.2.3.2.3). However, if the purpose is only to transform results of LSE, there is no need for partitioning, and all the remaining equations are still applicable.

In this work, further refinement to the computational scheme has been developed by using reduced coordinates during the computations of S-transformations. The general procedure for transforming x_i and Q_{xi} to x_j and Q_{xj} is:

(i) Reduce x_i to x_i' with respect to the approximate coordinates.

$$x_i'=x_i-x_o$$

where x_o is the approximate coordinates

(ii) Apply reduction to centroid in computing matrix G (equation 2.90).

(iii) Normalize matrix G (equation 2.91), if necessary.

(iv) Compute S-transformations

$$x_j'=S_jx_i'$$

$$Q_{xj}=S_jQ_{xi}S_j^t$$

(use equation 2.95 for optimization).

(v) Compute final coordinates

$$x_j=x_o+x_j'$$

This scheme allows transformation into minimum trace, partial minimum trace or minimum constraints solutions.

3. ERROR MODELLING IN LSE

A realistic mathematical model is needed for LSE, with respect to systematic, gross and random errors. Such requirements are essential because any significant errors (especially gross errors) will lead to apparent deformation.

This chapter examines the modelling of significant systematic, random and gross errors in measured uncorrelated surveying data. The applications of the precision and reliability analyses, together with the statistical testing for assessment of LSE results are described. A strategy has been formulated for error modelling and LSE (sections 3.5.3, 3.6.2, 3.7.3 and 3.9).

3.1 Sources of error

The LSE process as described in section 2.1 uses observational data for estimation of parameters via a linearised mathematical model (i.e. Gauss-Markov Model) consisting of both functional and stochastic models (Appendices A and C). In reality, it is important that every possible source of error be considered. The main sources of error (Caspary, 1987b) are in the mathematical model (functional and stochastic), observational data and in the computations.

Theoretically, the chosen functional model should represent the reality. Unfortunately, this is rarely achieved in practice due to unmodelled effects. The difference between the functional model and reality is called systematic error or bias. Typical sources of systematic error are instrumental factors (maladjustment), physical effects (for example atmospheric condition), choice of the mathematical model and observer's limitations such as personal bias (Cooper, 1974, 1987; Davies et al, 1981).

The stochastic model (section 2.1) describes the random errors of the measurements. Random (or accidental) errors are the unavoidable small differences between the measurements and their expectations, and follow statistical distributions. In surveying, observations are considered as random variables, and random errors are assumed to follow the normal distribution. Further details are given in Cooper (1974) and Cross (1983).

Errors in observational data are normally large errors attributed to gross errors, also known as blunders, outliers or mistakes (Hawkins, 1980). This type of error arises due to the malfunction of equipment, techniques and / or observers. Common examples of gross error are incorrect reading and recording of measurements.

In this study only systematic, random and gross errors in the observations are considered, while other error sources are assumed to be handled separately and beyond the scope of this research.

In deformation monitoring, both functional and stochastic models have to be correct and accurate, while errors in the observational data need to be detected and removed. Any significant errors (especially gross errors) will lead to apparent deformations and hence will contaminate the results of deformation detection (section 1.2).

Modelling of these errors requires an assessment of the quality of the LSE results and the application of the concept of hypothesis or statistical testing. Aspects related to quality measures and statistical testing are discussed in sections 3.2 and 3.3 respectively. The remaining sections describe the strategy developed for error modelling and detection.

3.2 Quality criteria

The quality of LSE results is usually assessed using some form of quality criteria. The measures of precision, reliability and accuracy are useful to describe the quality with respect to random, gross and systematic errors respectively (Cooper and Cross, 1988). For monitoring networks, it is also important to assess the capability of the network to detect the expected significant movement, i.e. sensitivity. As monitoring networks generally need to be of high accuracy, it is very important to assess the quality of a network with respect to precision, reliability, accuracy and sensitivity (Niemeier et al, 1982).

Precision is indicated by random errors. In general, a relatively high precision network (small random errors) is required to guarantee detection of instabilities. Criteria for precision are derived from the cofactor matrix of the estimated parameters $Q_{\hat{x}}$. Most measures of precision are therefore datum dependent, and can be either global or local. Extensive discussions are given by Caspary (1987b) and Cross (1983). Examples of global measures are the trace of $Q_{\hat{x}}$, the maximum eigenvalue of $Q_{\hat{x}}$ and criterion matrices. Local measures include standard deviations of the individual parameters (coordinates) and confidence regions in the form of absolute or relative error ellipses and ellipsoids.

Reliability is dependent on the network configuration and the precision of measurements. It indicates the extent to which the network measurements as a whole are self-checking. The reliability of a network (Baarda, 1968) is its capability of detecting gross errors by suitable testing procedures. Reliability measures consist of two parts, internal and external reliability. Criteria for reliability can be computed from the cofactor matrix of the estimated residuals $Q_{\hat{v}}$.

Internal reliability is related to the probability of gross error detection, whilst external reliability measures the effect of undetected errors on the parameter estimation. The measures for internal reliability of a measurement can be based on the sizes of the corresponding (local) redundancy number and marginally detectable gross error (MDGE). Both measures of internal reliability are datum independent.

For external reliability, the basic measure is datum dependent (Caspary, 1987b). However, datum independent measures can be determined by means of influential factors. In deformation monitoring, we seek a high reliability network, since higher internal reliability (larger redundancy number, smaller MDGE) increases the probability of detecting gross errors. In addition, higher external reliability (smaller influential factor) indicates that the model responds insignificantly to undetected errors. It is obvious that high internal reliability leads to high external reliability.

Accuracy indicates the quality with respect to systematic errors. This approach is used by Cooper and Cross (1988) by extending the functional model via bias parameters to include any suspected systematic errors. After LSE, the variances of the bias parameters (obtained from $Q_{\hat{x}}$) are examined to determine if such parameters are significant. Some researchers use the term precision, as the measure is also based on $Q_{\hat{x}}$.

Measures of the sensitivity with respect to certain deformation models provide quality to determine movements. Niemeier et al (1982) and Cooper (1987), among others, show the importance of sensitivity analysis. The measure is in terms of a form vector, which is actually the minimum size of detectable deformation for the assumed model.

The quality criteria mentioned above are very important for proper assessment of LSE results. In practice, all the measures of precision, reliability and sensitivity can be determined at the design stage. In this thesis, it is assumed that the monitoring network is being designed and optimized by taking into consideration all or most of these aspects.

In network optimization study (Grafarend, 1974; Cross, 1983), the LSE problems are classified into four sections:

(i) Zero-order design (ZOD) or datum problem, to search for an optimal datum.
(ii) First-order design (FOD) or configuration problem, to optimize the configuration.
(iii) Second-order design (SOD) or the generalised weight problem, to determine the optimal distribution of the observational work in a fixed configuration.
(iv) Third-order design (TOD) or the densification problem, to optimize the improvement of an existing network.

In monitoring networks, the required precision is generally known, and only ZOD, FOD and SOD are applicable (Neimeier, 1987). Basically, it is required to optimize the configuration (i.e. station positions and observation scheme) that will satisfy the observation weights and the required precision, and to optimize the selection of instruments and observing procedure.

The solution to the network optimization problem can be based on either computer simulation (or pre-analysis) or analytical approach (Cross, 1983; Cooper and Cross, 1988). Pre-analysis is commonly used, while the analytical approach is still under investigation. In pre-analysis, equations (2.20) and (2.22) are utilised for this purpose where

$$
\begin{aligned}
Q_{\hat{x}} &= N^{-1} = (A^t W A)^{-1} \qquad (3.1a)\\
Q_{\hat{v}} &= W^{-1} - A Q_{\hat{x}} A^t
\end{aligned}
$$

The subject of network optimization is beyond the scope of this research, but details can be found in Grafarend (1974), Cross and Thapa (1979), Schaffrin (1981), Grafarend and Sanso (1985) and Kuang et al (1991).

Measures of both precision and reliability are used in this study. In addition, pre-analysis (equation 3.1a) has been adopted for computing precision and reliability of the network prior to LSE of the actual data. The relevant formulae for precision and reliability analyses are given in section 3.8.

3.3 Statistical testing

The linear mathematical model for LSE relates the observations and parameters (section 2.1.3). Statistical testing on the LSE results is performed to ensure that the results obtained using the adopted model are satisfactory and there are no significant errors in the observational data.

The linearity and computational assumptions are necessary for the LSE process, while distributional assumptions are required for valid statistical testing (section 2.1.5). It is usually assumed that the observations are normally distributed, and the mathematical model is correct and complete.

The statistical testing procedure consists of three steps: formulation of a null hypothesis, computation of a suitable test statistic and the selection of risk level to determine critical value of the test statistic (Caspary, 1987b). In LSE and deformation detection, such tests can be either global or local.

Initially, the null hypothesis H_0 is formulated to express the condition to be tested. To get an idea of what is true if H_0 fails, an alternative hypothesis H_a is formulated, although it is not always expressed explicitly. For example, in the global test on the estimated variance factor (section 3.3.1), H_0 and H_a can be expressed as

$$H_0: \hat{\sigma}_0^2=\sigma_0^2 \qquad (3.1b)$$

$$H_a: \hat{\sigma}_0^2>\sigma_0^2 \qquad \text{one tailed test}$$

or $$H_a: \hat{\sigma}_0^2\neq\sigma_0^2 \qquad \text{two tailed test}$$

The test statistic T is usually chosen so that its distributional properties are known if H_0 is true (i.e. normal distribution) and it is sensitive to small departures from H_0. For the above global test, the computed T has a χ^2 distribution (under H_0) with expectation r (number of degrees of freedom in LSE)

$$T=\Omega/\sigma_o^2 \sim \chi^2_r \qquad (3.2)$$

where Ω is the quadratic form of the residuals (equation 2.23).

Finally, the risk level α is selected (normally based on experience) and the critical value of the test statistic for α (as determined from tables or computation) is compared with T to determine the outcome of the testing. A decision is made whether to accept H_o (actually do not reject H_o) or reject it (i.e. accept H_a). Hence, the result of the testing is either rejection of H_o or no rejection. In the above example (equation 3.2), H_o is not rejected at risk level α if

$|T| \leq \chi^2_{r,\alpha}$ (one tailed test)

T is between $\chi^2_{r,1-\alpha/2}$ and $\chi^2_{r,\alpha/2}$ (two tailed test)

Otherwise, H_o is rejected at risk level α.

In statistical testing, two types of errors may be made, namely Type I and Type II errors. A Type I error is the rejection of H_o although it is actually true or correct, e.g. rejection of a good observation. The probability of this error is called risk or significance level α. In practice, typical values for α (Caspary, 1987b) are 0.1, 0.05, 0.01 and 0.001. A Type II error is the acceptance of H_o although H_a is actually true (or H_o is wrong), e.g. accepting a bad observation. The probability of this error is β. The quantity $(1-\beta)$ is called the power of the test. A typical value of β is 0.20 or 20%.

Ideally, it is required to minimize the probability of both types of error. However, β increases as α decreases and vice-versa (Cross, 1983), indicating that the probabilities of Type I and Type II errors cannot be reduced at the same time, and both errors need to be optimised. Testing procedures that handle both types of errors are attributed to Baarda (1968), i.e. Baarda's data snooping and B-method. Such testing procedure is also closely related to reliability analysis. Basically, the methods standardize both the risk levels and commonly used values are α_o 0.1% and β_o 20%. The related aspects of testing in deformation detection is discussed in section 4.3. In practice, only Type I error is considered, due to the difficulty of knowing probability density function (pdf) for test statistic under H_a.

The applications of statistical testing in surveying are extensive and only the relevant tests will be discussed here. A more detailed explanation may be found in Vanicek and Krakwisky (1986), Cooper (1987), Cross (1983), Mikhail (1976) and many others.

In LSE, statistical testing can be applied either before LSE (to assess the quality of observed data) or after LSE (to assess the results of estimation) (Cooper, 1987), the latter will be examined here.

Statistical tests are useful as an aid in assessing the results of LSE. In the following discussion, it is generally assumed that the a priori variance factor is known, and its value is unity. The most commonly used tests following LSE are:

1. Test on the estimated variance factor (global).
2. Goodness of fit test on the estimated residuals (global).
3. Significance test on parameters estimated by LSE (local).
4. Test for outlying estimated residuals (local).
5. Test on confidence region of parameters (local).

Wherever applicable, the tests should be used for statistically checking on LSE results.

3.3.1. Test on the estimated variance factor

The test on the estimated (or a posteriori) variance factor, often termed a global (model) test, is used to check that both functional and stochastic models are acceptable. Such a test checks the validity of the following important assumptions: the model is correct and complete, on average the observations are normally distributed (i.e. contain random errors only), and no systematic and gross errors are present in the measurements. The above assumptions are stated as the null hypothesis H_0.

As the outcome of LSE, an unbiased estimate of the variance factor can be found (equation 2.23). The idea behind the global test is to determine whether the a posteriori variance factor is significantly different from the a priori variance factor or not.

The basic hypotheses for the test are formulated as equation (3.1b)

$$H_o: \hat{\sigma}_o^2=\sigma_o^2 \quad \text{and} \quad H_a: \hat{\sigma}_o^2\neq\sigma_o^2 \tag{3.3}$$

To test H_o, the test statistic T (equations 3.2 and 2.23) is computed

$$\begin{aligned} T&=\hat{v}^tW\hat{v}/\sigma_o^2 \sim \chi^2_r \\ &=\Omega/\sigma_o^2=r\hat{\sigma}_o^2/\sigma_o^2 \end{aligned} \tag{3.4}$$

where quantities $\hat{v}$, W, Ω, σ_o^2 and $\hat{\sigma}_o^2$ and r are as defined in section 2.1.3. H_o is accepted and the test passes at the chosen significance level α if the computed T lies within the specified interval (based on percentage point of a χ^2 distribution) as follows

$$\chi^2_{r,1-\alpha/2}<T<\chi^2_{r,\alpha/2} \tag{3.5}$$

If T falls outside the interval, H_o is rejected and the test fails.

The above test (equation 3.5) is known as two-tailed test. It is likely that significant errors will increase the value of the estimated variance factor. For this reason, especially in the detection of gross errors, it is usual to adopt a one-tailed test (equation 3.1b) where

$$H_o: \hat{\sigma}_o^2=\sigma_o^2 \quad \text{and } H_a: \hat{\sigma}_o^2>\sigma_o^2 \tag{3.6}$$
H_o is accepted if $T\leq\chi^2_{r,\alpha}$

Interpretation of the test is useful as the acceptance of H_o (i.e. the test passes) indicates that there is no objection to H_o, but does not prove it. Also, rejection of H_o (i.e. the test fails) only shows that either the model or the observations or both are wrong. Hence, to determine the causes, further investigations or statistical tests are needed. The most often accepted reasons for the failure of the global test are: an incorrect or incomplete functional model (i.e. systematic errors); unrealistic stochastic model (weighting of observations and correlations) and; gross errors in the measurements. Such reasoning is called a posteriori reasoning by Cooper (1987).

The global test described above can also be performed using an F test based on the Fisher statistic (Gruendig and Bahndorf, 1984), and the same results are obtained. An important relationship between the critical values of χ^2 and F is

$$F_{r,\infty,\alpha}=(\chi^2_{r,\alpha})/r \tag{3.7}$$

3.3.2. Goodness of fit test (on the estimated residuals)

For the purpose of statistical testing, it is assumed that the observations (and also estimated residuals) are normally distributed. Hence, it is important to test the normality assumption using a goodness of fit test. The hypotheses are

H_o: observations are normally distibuted (3.8)
H_a: observations have some other distribution

The test statistic is (Cooper, 1987)

$$T= \Sigma_{i=1,nc}(o_i-e_i)^2/e_i \sim \chi^2_{nc-1} \tag{3.9}$$

where o_i is the actual frequency (or number) of observations (or residuals) in each class i, e_i is the expected frequency in each class (based on normal distribution) and nc is the number of classes. The normality test passes, and H_o is accepted at the selected significance level α if T is less than critical value $\chi^2_{nc-1,\alpha}$

$$0 < T < \chi^2_{nc-1,\alpha} \tag{3.10}$$

Otherwise, H_o is rejected, and the test fails, indicating the failure of the normality assumption.

In reality, surveying observables are of different types, such that although the observations are normally distributed, the estimated residuals across all types of measurements are not. In such a case it is necessary to normalize or standardize the estimated residuals prior to testing. Standardized or normalized residuals can be computed by

$$\hat{v}_i'=\hat{v}_i/(Q_{\hat{v}i})^{1/2}=\hat{v}_i/(\sigma_o\sigma_{\hat{v}i}) \tag{3.11}$$
$$Q_{\hat{v}}=[W^{-1}-AN^{-1}A^t]$$

where $\hat{v}_i'$, $\hat{v}_i$ and $Q_{\hat{v}i}$ ($\sigma_{\hat{v}i}$) are standardized residuals, estimated residuals and cofactor matrix (standard deviation) of the residuals respectively. Values for $\hat{\sigma}_{vi}$ are obtained from the square root of the diagonal elements of $Q_{\hat{v}}$. An approximation (Pope, 1976) for $\hat{\sigma}_{vi}$ can also be computed with the following expression

$$\sigma_{\hat{v}i}=(\hat{\sigma}_o*\sigma_i/\sigma_o)*((n-u)/n)^{1/2} \tag{3.12}$$

where σ_i is the standard deviation of observation i, n the number of observations and u the number of parameters.

If the variance factor is known, the standardized residuals will have a normal distribution with zero mean and unit variance. The goodness of fit test can be applied either to the whole or to a group of observations (residuals).

3.3.3. Significance test on parameters (estimated by LSE)

In LSE, it is also possible to extend the functional model by means of additional parameters, for example, the incorporation of a scale factor to the measured EDM distance (section 3.5.3.1). The significance of any such parameters must be tested statistically.

For each estimated additional parameter ($\hat{p}$), the test is based on (Cooper, 1987)

$$H_o: \hat{p}=0 \qquad H_a: \hat{p}\neq 0 \tag{3.13}$$

The test statistic is

$$T=\hat{p}/\hat{\sigma}_p \sim t_{mp-1} \tag{3.14}$$

where $\hat{\sigma}_p$ is the square root of the diagonal element of the cofactor matrix (i.e. standard deviation) of the estimated parameter and mp is the number of observation equations containing the bias parameter. The test passes, and H_o is accepted, indicating the additional parameter is not significant (i.e. no scale discrepancy) if T lies within the interval

$$t_{mp-1,1-\alpha/2} < T < t_{mp-1,\alpha/2} \tag{3.15a}$$

Otherwise H_o is rejected, and the estimated additional parameter is considered to be significant.

If H_o is accepted, LSE should be repeated without the additional parameter.

If more than one additional parameter are included in the functional model, any correlations between the parameters have to be determined. The correlation coefficient between parameters x and y (Mikhail, 1976) is

$$\rho_{xy}=\sigma_{xy}/(\sigma_x\sigma_y) \tag{3.15b}$$

where ρ_{xy} is the correlation coefficient, σ_{xy} is the covariance between x and y, and σ_x, σ_y are the standard deviations of x and y respectively. If the coefficient is zero, parameters x and y are uncorrelated. On the other hand, a coefficient close to unity indicates highly correlated parameters. If $\rho_{xy}=1$, x and y are functionally correlated.

3.3.4. Tests for outliers

LSE provides a means of checking individual observations through the examination of their estimated residuals. The use of statistical tests in this respect is related to the concept of data snooping, and can be very useful for outlier detection.

The test for outlying estimated residuals (or outlier test) examines the standardized residuals $\hat{v}_i'$ (equation 3.11; Caspary, 1987b)

H_o: $E\{\hat{v}_i'\}=0$ (3.16)

or each $\hat{v}_i'$ is free from gross error

or each $\hat{v}_i'$ belongs to $N(0,\sigma_{\hat{v}i}^2)$

H_a: one residual is an outlier or $E\{\hat{v}_i'\}\neq 0$

or one residual contains gross error

The test statistic is

$$T=\hat{v}_i'=|\hat{v}_i/(\sigma_o\sigma_{\hat{v}i})|\sim N(0,1) \tag{3.17a}$$

or

$$T=\hat{v}_i'=|\hat{v}_i/(\hat{\sigma}_o\sigma_{\hat{v}i})|\sim \tau_r \tag{3.17b}$$

Equation (3.17a) is used when σ_o^2 is known, and it is called un-studentized test, while equation (3.17b) (studentized test) is employed when σ_o^2 is unknown and is estimated by $\hat{\sigma}_o^2$. Evaluation of $Q_{\hat{v}i}$ in order to get $\sigma_{\hat{v}i}$ will require extensive computation, as shown in equations (2.22) and (3.11). For uncorrelated observations, only diagonal elements of $Q_{\hat{v}}$ need to be computed, and expression for $\sigma_{\hat{v}i}$ (section 3.8) can be written as

$$\sigma_{\hat{v}i}=\sigma_i\sqrt{(r_i)} \tag{3.18}$$

$$r_i=1-a_iN^{-1}a_i^t w_i$$

where

σ_i=standard deviation of observation i

r_i=redundancy number of observation i

a_i=elements of design matrix A for observation i

N^{-1}=Cayley inverse of N

w_i=weight of observation i

In this work, equation (3.18) is adopted because it involves less computation than equation (3.11) and computes values of r_i, needed for reliability analysis.

In principle, both tests (equations 3.17a and 3.17b) examine each standardized residual individually, 'out of the context' of the other residuals. In fact, since all residuals are tested simultaneously ('in context'), the probability of the test is higher than (1-α), and some form of standardization is required. Further details on 'in context and out of context' testing are given by Vanicek and Krakiwsky (1986).

In practice, the application of the above tests (i.e. equations 3.17a and 3.17b) with the appropriate standardization procedure is known as data snooping (Schwarz and Kok, 1993), and consists of either Pope's (Tau) or Baarda's (data snooping) methods. Pope's method uses the studentized test and standardizes for Type I error (α) only by invoking Bonferroni's inequality (Vanicek and Krakiwsky, 1986). Baarda's method uses the un-studentized test and standardizes both Type I and Type II errors (α and β).

Pope's or Tau method (Pope, 1976) assumes σ_o^2 as unknown, and applies the estimated $\hat{\sigma}_o^2$ in computing the normalized residuals. The test statistic (equation 3.17b) is one-dimensional

$$T_p=\hat{v}_i/(\hat{\sigma}_o\sigma_{\hat{v}i})\sim\tau_r \tag{3.19}$$

where r is the number of degrees of freedom (redundancy), and $\sigma_{\hat{v}i}$ may be computed rigorously (equation 3.18) or by approximation (equation 3.12). Based on the concept of in context testing, ignoring the correlation amongst the residuals, α is standardized as

$$\alpha_n=1-(1-\alpha)^{1/n}\approx\alpha/n \tag{3.20}$$

where n is the number of the observations. H_o is accepted if

$$|T_p| \leq \tau_{r,\alpha o} \qquad \text{where } \alpha_o=\alpha_n/2 \tag{3.21}$$

Otherwise (if $T>\tau_{r,\alpha o}$), H_o is rejected, and the corresponding observation must be examined.

As tables of τ-distribution are very rare in statistical books, the following relationship between τ and the t-distribution may be used (Pope, 1976)

$$\tau_r=r^{1/2}t_{r-1}/(r-1+t_{r-1}^{\ 2})^{1/2} \tag{3.22}$$

The computation of the critical value of τ is given by Pope (1976), together with the listing of a useful Fortran subroutine.

Baarda's method (Baarda, 1968) assumes that σ_o^2 is known a priori, and employs a multi-dimensional test. The test statistics (equation 3.17a) is

$$T_b=\hat{v}_i/(\sigma_o\sigma_{\hat{v}i})\sim N(o,\sigma_{\hat{v}i}) \tag{3.23}$$

H_o is accepted if $|T_b| < N_{\alpha/2}$. Given α 0.05 (5%), the critical value of N is 1.96.

In the actual implementation of Baarda's method, both Type I and Type II errors are taken into account. Typical values for standardized α_o and β_o are 0.1% and 20% respectively (section 3.3) leading to the critical value $u_{\alpha o}$ of 3.29 (Figure 3.1). H_o is accepted if

$$|T_b| \leq u_{\alpha o} \tag{3.24}$$

Otherwise H_o is rejected if $T>u_{\alpha o}$. Interpretation of the test is similar to Pope's method.

Baarda's method also provides a measure of reliability, both internal and external. Some expressions for reliability (see section 3.8) of each observation i (datum independent quantities) are:

internal reliability

$r_i=1-a_iN^{-1}a_i^tw_i$ (3.25a)

$MDGE_i=\nabla_i=\sigma_i(\lambda_o/r_i)^{1/2}$ (3.25b)

external reliability

$\delta_i=\lambda_o(1-r_i)/r_i$ (3.26)

where r_i=redundancy number of observation i

$MDGE_i=\nabla_i$=size of the MDGE in observation i.

δ_i=influential factor or global distortion parameters of observation i

λ_o=non-centrality parameters, computed from α_o and β_o

A typical value of λ_o (with α_o 0.1% β_o 20%) is 17. Baarda (1968) provides a nomogram for the evaluation of $u_{\alpha o}$ (with respect to type I error) and λ_o (with respect to degrees of freedom and type I and II errors). Figure 3.1 shows the nomogram for β_o=20%.

When σ_o^2 is known (section 3.3.2), the normalized residuals are normally distributed (Steeves and Fraser, 1987) with N(0,1). For simplicity, as an alternative to equation (3.24), standard normal distribution can be used (equation 3.23), and H_o is accepted if

$$|T| < N_{\alpha/2} \tag{3.27}$$

In equations (3.24) and (3.27), the critical values are independent of r.

3.3.5. Testing on confidence region of parameters

The application of statistical analysis to LSE results allow the establishment of confidence regions around the estimated coordinates, via the utilization of cofactor matrix $Q_{\hat{x}}$ (see section 3.8), where

$$Q_{\hat{x}}=\sigma_0^2(A'WA)^{-1} \tag{3.28}$$

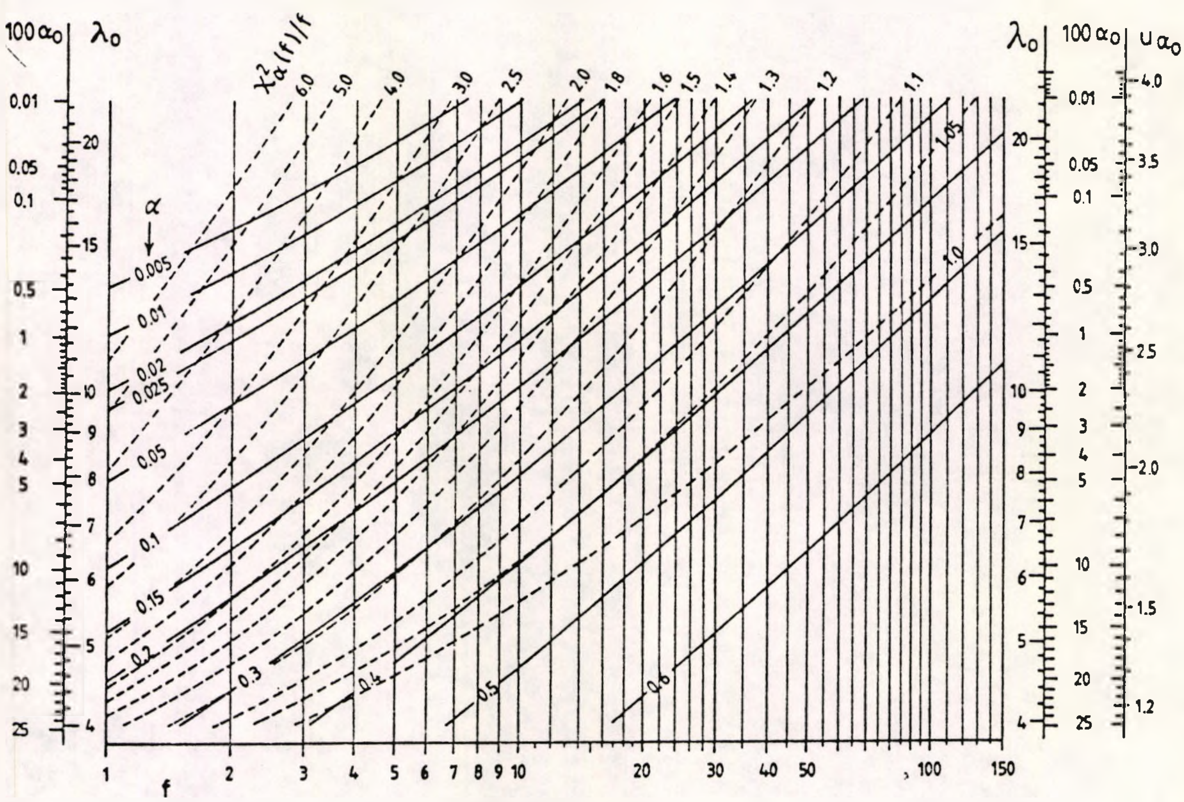

Figure 3.1 Baarda's nomogram for β_0 20%
(taken from Caspary, 1987b).

The confidence region can be either absolute or relative in nature (Dodson, 1990). The absolute (or point) confidence region is datum dependent and gives an overall picture of network precision. The relative confidence region is datum independent and reflects the relative precision between stations.

In multi-dimensional space, the confidence region is in the form of a hyper ellipsoid, and for 3-D is ellipsoid. The parameters of a confidence region are the semiaxes and their orientation, which can be determined via EVD, as outlined in section 2.2.5. For details, see Mikhail (1976).

The probability for standard confidence region (when expansion factor $y^{1/2}$ is equal to one) is 0.683, 0.394 and 0.199 for 1-D, 2-D and 3-D cases respectively. In spatial (3-D) space, the probability that the point lies within the standard point error (confidence) ellipsoid is about 20%. It is then necessary to increase the probability with respect to the appropriate or selected significance level (typically 5%) by multiplying the semiaxes of the ellipsoid by the expansion factor. For example $y^{1/2}$ is 2.796 for 5% significance level or 95% confidence level. The following values are taken from Mikhail (1976):

probability	expansion factor($y^{1/2}$)		
	1-D	2-D	3-D
0.500	0.678	1.177	1.538
0.900	1.646	2.146	2.500
0.950	1.960	2.447	2.796
0.990	2.575	3.035	3.368

Table 3.1 Expansion factor for confidence region

3.4 Strategy for error modeling

In principle, the method of LSE is valid without any assumption with respect to errors. However, in qualitative analysis and the statistical evaluation of LSE results, it is generally assumed that the selected mathematical models are correct, i.e. all the systematic and gross errors have been eliminated prior to LSE, and the measurements contain only realistic random errors and are regarded as random variables. These assumptions are important (Caspary, 1987b) due to the high sensitivity of LSE to both systematic and gross errors. Equally important is the proper handling of random error.

In practice or reality, the model may be affected by errors (section 3.1). Treatment of these errors is very important in the measurement and analysis of survey data, such a procedure is termed error modelling. The purpose of modelling systematic and gross errors is to reduce the effects of such errors and to ensure that their magnitudes are insignificant, either before (preferably) or after LSE. Stochastic modelling is useful for realistic estimation of random errors. In this research, a strategy for error modelling has been developed by examining the nature of each type of error in turn.

Systematic errors have a constant effect on repeated measurements and consequently cannot be recovered or detected via repetitive measurements. Hence, measurements need to be corrected and reduced for systematic effects. Corrections are related to known physical effects, while reductions are related to the geometry involved (Cooper, 1987).

Traditionally, the effects of systematic error are minimized either by calibration of instruments and applying the appropriate corrections and reductions to measurements, and / or adopting a suitable measurement scheme. In practice, systematic errors are usually either neglected or assumed to be insignificant.

It is equally important to detect any gross errors in the measurements. During the measurement process, large gross errors can be handled or avoided by screening, i.e. adopting a proper and suitable measurement scheme (Cooper, 1974, 1982; Secord, 1986) that provides some independent checks. In this way, suspect measurements can be examined, rejected and re-measured as necessary.

If any gross errors remain undetected in the measurements, it is also possible to detect them after LSE (post-LSE) based on the techniques employing the analysis of residuals. Generally (actually not necessarily) large residuals will indicate erroneous measurements. Whenever such residuals are detected, the corresponding measurement is examined to find out if a gross error can be found. That measurement is then deleted or remeasured.

In general, post-LSE techniques for gross error detection are based on either the mean shift model (such as Baarda's and Pope's methods) or the variance inflation model (i.e. de-weighting of observations). Details on this aspect are given by Cross and Price (1985), Chen et al (1987), Kubik and Wang (1991), Gao et al (1992) and Schwarz and Kok (1993). The most popular techniques for gross error detection in surveying are Baarda's (data snooping), Pope's (Tau) and the Danish method (Caspary, 1987b).

Both Pope's and Baarda's methods assume that just one measurement is affected by a gross error, and are based on rigorous statistical theory. In fact, it is likely that the measurements contain multiple gross errors. The general procedure for both methods in this situation is by successive re-application of the relevant tests (section 3.3.4) to identify suspected erroneous observation and then to eliminate that particular observation and repeat the LSE until no gross errors are detected.

The Danish method is not based on rigorous statistical theory and uses a suitable de-weighting strategy to locate and eliminate the gross errors (Kubik et al, 1987). Moreover, the Danish method can provide a simultaneous solution in the presence of multiple gross errors, and the erroneous measurements are not deleted completely. Hence, the Danish method can be considered as robust in nature.

Although measurements may not contain significant systematic and gross errors, they will still be inconsistent (i.e. repeated measurements of same elements will give rise to different measured values) due to random errors. In LSE, the stochastic model describes these random effects by means of a weight coefficient or cofactor or covariance matrix of measurements (section 2.1). It is therefore essential to determine the cofactor matrix prior to LSE. Estimates of the cofactor matrix of the measurements may be obtained either from experiments or previous performance (Cooper and Cross, 1988)

Survey measurements can be correlated algebraically (for example horizontal angles) or uncorrelated (such as photo measurements, distances, directions and height differences). For simplicity and practical purpose, measurements are generally assumed to be uncorrelated and independent. In this simplified case, only the variances are needed, and the cofactor or weight matrix becomes diagonal (equation 2.15b).

As an extension to the existing method of error modelling, a strategy has been developed in this study to cope with uncorrelated surveying measurements. The most significant systematic errors are in distance measurements. Modelling of systematic errors in measured EDM distances can be carried out via the application of additional parameters (scale and / or zero errors) and pseudo observations (distance ratio and distance difference). The modelling of systematic errors is formulated in section 3.5. As pseudo observations are theoretically correlated algebraically, a de-correlation technique is applied (section 3.5.3.2.3).

In gross error detection, successful implementation of the Danish method requires a proper termination criterion and a suitable de-weighting function. Section 3.6 gives a description of the developed strategy of robustified LSE, which consists of a modified Danish method and the incorporation of global and local tests, together with reliability analysis. The effects of de-weighting the observations on the parameters are also derived (section 3.6.2.3).

In stochastic modelling, a more suitable method of estimating the cofactor matrix of observations is based on an iterative numerical technique of VCE (Caspary, 1987b; Chen et al, 1990b). Unfortunately, the computational effort required is extensive for the general case. Assuming uncorrelated and independent observations, the iterative and simplified technique of VCE for estimating the variances of observations is extended by adding tests on the global and group estimated variance factors as termination criterion (section 3.7).

The procedure developed for error modelling and LSE also incorporates global and local tests, together with precision and reliability analyses (section 3.8). Such a procedure is discussed in section 3.9.

3.5 Modelling of systematic errors

Sources contributing to systematic error are described in section 3.1. Some examples of the systematic errors in surveying are given by Anderson and Mikhail (1988). Details on sources of error affecting EDM measurements are discussed in Rueger (1988).

The method for modelling systematic error consists of both pre-LSE and post-LSE techniques, and is composed of combinations of the following:

.instrument calibration and applying corrections

.a proper measurement scheme

.improvements to the functional model

The first two are known as field or pre-LSE methods, while the later (focus of this study) is a computational technique that also includes post-LSE analysis. The above methods are described in sections 3.5.1, 3.5.2 and 3.5.3 respectively.

3.5.1 Instrument calibration and applying corrections

Most surveying observables consist of angular and linear measurements. The instruments normally used for this purpose are theodolite, tacheometers and levels. The instruments, being man made, are subject to mechanical or optical defects due to imperfections in their construction or lack of adequate adjustment. Most of these mechanical or optical defects can be reduced by careful calibration and adjustment of instruments before their use for measurement. Cooper (1982) provides a more detailed description on the calibration and adjustment of such instruments.

Nowadays, with the wide and accepted use of EDM for distance measurement, regular calibration and performance evaluation of EDM instruments becomes very important, if their suitability and precision are to be assured at the time of usage. The method of calibration (Kennie, 1990) to determine the systematic errors of an EDM instrument can be either laboratory or field based. Calibration procedures for EDM are given in detail by Ashkenazi and Dodson (1975), Rueger (1977), Deeth et al (1978), Sprent (1980) and Dracup et al (1982).

It is generally necessary to apply the appropriate corrections and reduction to the raw observations for known systematic effects before computation. Atmospheric refraction is a source of serious systematic errors in the measurements, especially zenith (or vertical) angles and EDM distances. Cooper (1987) describes in detail the necessary systematic corrections for the measured angles, height differences and EDM distances.

In EDM measurement, variations in the atmospheric conditions will change the refractive index along the EDM wavepath, and thereby limit the accuracy. It is then necessary to determine the refractive index of the atmosphere, and apply the computed atmospheric correction to the measured EDM distances. The most common method to determine atmospheric refractive index is by measuring temperature, pressure and humidity at the two ends of the measured line, and applying the computed correction to the measured distance. More details are given in Rueger (1988) and Cooper (1987), amongst others. Several approaches for handling this problem include the use of airborne sensors (Savage and Prescott, 1973), multi-mavelength EDM (Hugget and Slater, 1975), and improving the atmospheric model (Fraser, 1984; Dodson and Zaher, 1985).

3.5.2 A proper measurement scheme

Sometimes it is not always possible or convenient to calibrate the instruments frequently, especially for angular measurements. In this situation, errors due to the instrumental imperfection and / or non-adjustment can contaminate the measurements. Fortunately, most of the systematic errors (including observer's error) can be eliminated or reduced to a negligible amount via a proper observational procedure or measurement scheme.

For example (Cooper 1982), the principle of reversal (i.e. reading angular measurements on both faces of a theodolite) is often used to reduce the effects of horizontal collimation errors in the measured horizontal angles or directions. Also, errors due to natural causes, especially atmospheric conditions or refraction on horizontal angles, can be rendered negligible by choosing appropriate times for observing (Anderson and Mikhail, 1988), for example, at night when temperature and atmospheric conditions are almost constant. Further aspects on this approach can be found in Cooper (1982), Anderson and Mikhail (1988) and Davies et al (1981). High precision horizontal angle measurements can be achieved with the use of forced centring and good targeting (Ashkenazi et al, 1980). Teskey and Biacs (1990) adopt a special procedure of precise trigonometric heighting that enables zenith angle to be corrected for earth curvature and refraction.

3.5.3 Improvements to the functional model

Measured EDM distances are subject to instrumental error (Burnside, 1982; Rueger, 1988): constant error independent of distance such as zero error, reflector constant and centring error; error dependent on distances such as scale and frequency errors ; cyclic error such as electromagnetic coupling and effects of signal strength. In practice, the cyclic error should be calibrated and corrected before LSE because it is difficult to estimate the parameters related to cyclic error at the stage of LSE processing. Therefore, during network LSE, systematic errors in EDM distances can be either dependent (example scale error) or independent (such as zero error) of the measured distances.

As discussed in section 3.5.1, instrumental errors of EDM can be determined via calibration. The most common being scale and zero errors. As an alternative, it is possible to model the systematic error by extending, refining or improving the functional model for LSE. Modelling of this type can be either via the inclusion of additional parameters or using pseudo observables. The development made during this research is discussed in sections 3.5.3.1 and section 3.5.3.2.

3.5.3.1 Inclusion of additional parameters

In this approach, the effects of the particular systematic errors are represented and included in the functional model as additional or bias parameters (Cooper, 1987; Gruendig and Bahndorf, 1984). The most common approach in this case is to improve the functional model with respect to the measured EDM distances by introducing scale or zero errors as additional parameters.

A more general concept has been developed in this study, allowing the modelling of either scale (or bias), zero (or constant) errors or both in combination. It is also possible to model all or only a group of measured distances. Moreover, multiple (and independent) scale or zero error can be modelled as well, and are treated in the same manner as direction (Appendix A). In all cases, the significance of the parameters are determined by means of significance testing (section 3.3.3). For combination of scale and zero errors, correlation analysis (equation 3.15b) will indicate whether the parameters are correlated. Ideally, there should be little correlation between parameters.

The functional model is extended from the basic functional model of a spatial (slope) distance (s_k) between points i and k (Appendix A)

$$f(x)=s_k-[(x_k-x_i)^2+(y_k-y_i)^2+(z_k-z_i)^2]^{1/2}=0 \tag{3.29}$$

where $s_k=[dx^2+dy^2+dz^2]^{1/2}$, $dx=x_k-x_i$, $dy=y_k-y_i$, $dz=(z_k+ht_k)-(z_i+ht_i)$

ht_i=height of instrument at i, ht_k=height of reflector at k

Let k_1 and k_2 be the scale (or scale bias) and zero errors of the measured distance. The improved functional models and parameters to be estimated become:

(i) for combination of scale and zero errors

$$s_{co}=(1+k_1)[dx^2+dy^2+dz^2]^{1/2}+k_2 \tag{3.30}$$

parameters=$[x_i\ y_i\ z_i\ x_k\ y_k\ z_k\ ...\ k_1\ k_2]^t$ where $(1+k_1)$ is known as scale factor

(ii) for scale error only

$$s_{se}=(1+k_1)[dx^2+dy^2+dz^2]^{1/2} \tag{3.31}$$

parameters=$[x_i\ y_i\ z_i\ x_k\ y_k\ z_k\ ...\ k_1]^t$

(iii) for zero error only

$$s_{ze}=[dx^2+dy^2+dz^2]^{1/2}+k_2 \tag{3.32}$$

parameters=$[x_i\ y_i\ z_i\ x_k\ y_k\ z_k\ ...\ k_2]^t$

The above expressions show that equations (3.31) and (3.32) are the special case of equation (3.30) when k_2 is zero and $(1+k_1)$ is unity respectively.

Linearization of the above equations via Taylor's series expansion (Appendix A) results in the following observation equations:

(i) combination of scale and zero errors

$$\begin{aligned} &a_1\delta x_i+a_2\delta y_i+a_3\delta z_i+a_4\delta x_k+a_5\delta y_k+a_6\delta z_k \\ &+a_7\delta k_1+a_8\delta k_2=(s_{obs}-s_{co})+v_k \end{aligned} \tag{3.33}$$

where $a_1=\partial f/\partial x_i=-(1+k_1)dx/s$; $a_2=\partial f/\partial y_i=-(1+k_1)dy/s$

$a_3=\partial f/\partial z_i=-(1+k_1)dz/s$; $a_4=\partial f/dx_k=(1+k_1)dx/s$

$a_5=\partial f/\partial y_k=(1+k_1)dy/s$; $a_6=\partial f/\partial z_k=(1+k_1)dz/s$

$a_7=\partial f/\partial k_1=s$; $a_8=\partial f/\partial k_2=1.0$

(ii) scale error only

$$a_1\delta x_i+a_2\delta y_i+a_3\delta z_i+a_4\delta x_k+a_5\delta y_k+a_6\delta z_k \tag{3.34}$$
$$+a_7\delta k_1=(s_{obs}-s_{se})+v_k$$

where $a_1=\partial f/\partial x_i=-(1+k_1)dx/s$; $a_2=\partial f/\partial y_i=-(1+k_1)dy/s$

$a_3=\partial f/\partial z_i=-(1+k_1)dz/s$; $a_4=\partial f/dx_k=(1+k_1)dx/s$

$a_5=\partial f/\partial y_k=(1+k_1)dy/s$; $a_6=\partial f/\partial z_k=(1+k_1)dz/s$

$a_7=\partial f/\partial k_1=s$

(iii) zero error only

$$a_1\delta x_i+a_2\delta y_i+a_3\delta z_i+a_4\delta x_k+a_5\delta y_k+a_6\delta z_k \tag{3.35}$$
$$+a_7\delta k_2=(s_{obs}-s_{ze})+v_k$$

where $a_1=\partial f/\partial x_i=-dx/s$; $a_2=\partial f/\partial y_i=-dy/s$

$a_3=\partial f/\partial z_i=-dz/s$; $a_4=\partial f/dx_k=dx/s$

$a_5=\partial f/\partial y_k=dy/s$; $a_6=\partial f/\partial z_k=dz/s$

$a_7=\partial f/\partial k_2=1.0$

s_{obs}=observed spatial distance between i and k

s_{co}=computed spatial distance (equation 3.30)

s_{se}=computed spatial distance (equation 3.31)

s_{ze}=computed spatial distance (equation 3.32)

s=computed spatial distance (equation 3.29)

Equations (3.33), (3.34) and (3.35) can be used to model the whole or a group of measured distances. Furthermore, single or multiple (independent) scale or zero errors can be modelled via equations (3.34) and (3.35) respectively. The same formulation can also be written for horizontal distances, i.e. a special case of the above requiring 2-D coordinates only.

Another approach to improve the functional model is via the use of scaled distances (Angus-Leppan, 1972; Vincenty, 1969, 1979). Gruendig and Teskey (1984) demonstrate an appropriate functional model, observational scheme and procedure in using scaled distances.

The functional model can be written as

$$f(x)=s_k(1+k_1)-[dx^2+dy^2+dz^2]^{1/2}=0 \tag{3.36}$$

$$s_k(1+k_1)=[dx^2+dy^2+dz^2]^{1/2}$$

where scale error k_1 obtained from LSE is applied to the observed distance.

In this research, the linearised observation equation for 2-D case (Gruendig and Teskey, 1984) has been modified for 3-D case, and becomes

$$a_1\delta x_i+a_2\delta y_i+a_3\delta z_i+a_4\delta x_k+a_5\delta y_k+a_6\delta z_k=(s_{obs}(1+k_1)-s)+v_k \tag{3.37a}$$

where $a_1=\partial f/\partial x_i=-dx/s$; $a_2=\partial f/\partial y_i=-dy/s$

$a_3=\partial f/\partial z_i=-dz/s$; $a_4=\partial f/dx_k=dx/s$

$a_5=\partial f/\partial y_k=dy/s$; $a_6=\partial f/\partial z_k=dz/s$

Equations (3.36) and (3.37a) indicate that the estimated scale error is applied to the observed distance, instead of the computed distance as in equation (3.34).

The concept of scaled distances is useful in deformation detection. Once the scale and / or zero errors are estimated via equations (3.33), (3.34) or (3.35), they can be applied into the LSE process in the same manner as equation (3.37a). For this purpose, the functional model of equation (3.30) can be re-arranged as

$$[s(1+k_1)+k_2]=[dx^2+dy^2+dz^2]^{1/2} \tag{3.37b}$$

Equation (3.37b) allows the estimated scale and / or zero errors to be applied to the observed distances. The components of observation equations (a_1 to a_6) are as in equation (3.37a).

3.5.3.2 Using pseudo observables

In practice, sometimes not all the significant systematic errors in EDM distances can be reduced, by either calibration, measurement scheme or via additional parameters. To make matters worse, some systematic errors are difficult to model. In this situation, it is recommended that the appropriate pseudo observables (i.e. observables that are derived from the measurements) be used. Two types of pseudo observables can be used to cope with systematic error, either dependent or independent of the distance, i.e. distance ratios and distance differences respectively. With pseudo observables, both functional and stochastic models are taken into account.

The effects of atmospheric refraction will cause an error that is dependent on the distance measured. If two or more such distances are quickly measured from the same point by EDM, it can be assumed that the effects are the same for each measurement and the corresponding ratios (i.e. distance ratios) will be free from such errors. In addition, the effects of other errors that are linear and dependent on the distances will be eliminated too.

Constant errors (such as zero error, reflector constant and centring errors) will cause the measured distances to be either too short or too long. By taking differences (i.e. distance differences) between the distances (measured at the same point), the effects of the constant error which are linear, independent of the distances and common to all measurements will be cancelled.

During this study, the basic formulation of pseudo observables for spatial (slope) distances (distance ratios and differences) together with the expression for their covariance matrix, ignoring physical correlation, has been derived (sections 3.5.3.2.1 and 3.5.3.2.2).

Consider three stations i, j, k, with their corresponding coordinates (x_i, y_i, z_i), (x_j, y_j, z_j) and (x_k, y_k, z_k). Let s_j and s_k represent the spatial distances from i to j and i to k respectively, with the standard deviation of the measurements σ_j and σ_k. Let ht_i be the heights of instrument at i, ht_j and ht_k be heights of targets at j and k respectively.

3.5.3.2.1 Distance difference

In this work, the functional model for the 2-D case given by Cooper (1987) has been extended for 3-D case, such that the functional model for distance difference (Δs) is

$$f(x)=\Delta s-(s_k-s_j)=0 \qquad (3.38)$$

$$\Delta s=s_k-s_j$$

where $s_k=[dx_k^2+dy_k^2+dz_k^2]^{1/2}$, $s_j=[dx_j^2+dy_j^2+dz_j^2]^{1/2}$

$$dx_k=(x_k-x_i),\ dy_k=(y_k-y_i),\ dz_k=(z_k+ht_k)-(z_i+ht_i)$$

$$dx_j=(x_j-x_i),\ dy_j=(y_j-y_i),\ dz_j=(z_j+ht_j)-(z_i+ht_i)$$

The observation equation for distance difference becomes

$$a_1\delta x_i+a_2\delta y_i+a_3\delta z_i+a_4\delta x_j+a_5\delta y_j+a_6\delta z_j \qquad (3.39)$$
$$+a_7\delta x_k+a_8\delta y_k+a_9\delta z_k=(\Delta s_{obs}-\Delta s)+v$$

where $a_1=(dx_j/s_j)-(dx_k/s_k)$; $a_2=(dy_j/s_j)-(dy_k/s_k)$

$a_3=(dz_j/s_j)-(dz_k/s_k)$; $a_4=-dx_j/s_j$; $a_5=-dy_j/s_j$;

$a_6=-dz_j/s_j$; $a_7=dx_k/s_k$; $a_8=dy_k/s_k$; $a_9=dz_k/s_k$

Δs_{obs}= difference of measured distance s_k-s_j

Δs=difference of computed distance s_k-s_j

The variance $\sigma_{\Delta s}^2$ of the above "observable" can be determined by applying the principle of variance propagation.

If $y=Ax$, $\Sigma_y=J\Sigma_x J^t$ where $J=\partial y/\partial x$ (3.40)

$$\Delta_s=s_k-s_j=\begin{bmatrix}-1 & 1\end{bmatrix}\begin{bmatrix}s_j\\ s_k\end{bmatrix}$$

$$\sigma_{\Delta s}^2=\Sigma_{\Delta s}=\begin{bmatrix}-1 & 1\end{bmatrix}\begin{bmatrix}\sigma_j^2 & 0\\ 0 & \sigma_k^2\end{bmatrix}\begin{bmatrix}-1\\ 1\end{bmatrix}$$

$$\sigma_{\Delta s}^2=(\sigma_j^2+\sigma_k^2) \qquad (3.41)$$

In this work, the covariance matrix for multiple distance differences has been derived using equation (3.40), and for three observables can be obtained as follows. Let $\Delta s_1=s_2-s_1$, $\Delta s_2=s_3-s_2$ and $\Delta s_3=s_4-s_3$, and the variances of s_1, s_2, s_3 and s_4 be σ_1^2, σ_2^2, σ_3^2 and σ_4^2 respectively. Then

$$J=\begin{bmatrix}-1 & 1 & 0 & 0\\ 0 & -1 & 1 & 0\\ 0 & 0 & -1 & 1\end{bmatrix} \; ; \; \Sigma_s=\begin{bmatrix}\sigma_1^2 & 0 & 0 & 0\\ 0 & \sigma_2^2 & 0 & 0\\ 0 & 0 & \sigma_3^2 & 0\\ 0 & 0 & 0 & \sigma_4^2\end{bmatrix}$$

$$\Sigma_{\Delta s}=J\Sigma_s J^t$$

$$\Sigma_{\Delta s}=\begin{bmatrix} \sigma_1^2+\sigma_2^2 & -\sigma_2^2 & 0 \\ -\sigma_2^2 & \sigma_2^2+\sigma_3^2 & -\sigma_3^2 \\ 0 & -\sigma_3^2 & \sigma_3^2+\sigma_4^2 \end{bmatrix} \tag{3.42a}$$

Similarly, the covariance matrix of n observables can be written as

$$\Sigma_{\Delta s}=\begin{bmatrix} \sigma_1^2+\sigma_2^2 & -\sigma_2^2 & 0 & . & . & 0 \\ -\sigma_2^2 & \sigma_2^2+\sigma_3^2 & -\sigma_3^2 & . & . & 0 \\ 0 & -\sigma_3^2 & \sigma_3^2+\sigma_4^2 & . & . & . \\ . & . & . & . & . & 0 \\ . & . & . & . & \sigma_{n-2}^2+\sigma_{n-1}^2 & -\sigma_{n-1}^2 \\ 0 & 0 & . & 0 & -\sigma_{n-1}^2 & \sigma_{n-1}^2+\sigma_n^2 \end{bmatrix} \tag{3.42b}$$

3.5.3.2.2 Distance ratio

In this research, a functional model for spatial distance ratio (r) has been extended from the 2-D case (Vincenty, 1979)

$$f(x)=s_k/s_j-r=0 \tag{3.43}$$
$$r=s_k/s_j$$

The observation equation becomes

$$a_1\delta x_i+a_2\delta y_i+a_3\delta z_i+a_4\delta x_j+a_5\delta y_j+a_6\delta z_j \tag{3.44}$$
$$+a_7\delta x_k+a_8\delta y_k+a_9\delta z_k=\Delta r+v$$

where $a_1=r[(dx_j/s_j^2)-(dx_k/s_k^2)]$; $a_2=r[(dy_j/s_j^2)-(dy_k/s_k^2)]$

$a_3=r[(dz_j/s_j^2)-(dz_k/s_k^2)]$; $a_4=-rdx_j/s_j^2$; $a_5=-rdy_j/s_j^2$;

$a_6=-rdz_j/s_j^2$; $a_7=rdx_k/s_k^2$; $a_8=rdy_k/s_k^2$; $a_9=rdz_k/s_k^2$

$\Delta r=(r_{obs}-r)$

The variance (σ_r^2) is derived by propagation of variance via equation (3.40)

$r=s_k/s_j$

$J=[\partial r/\partial s_j \quad \partial r/\partial s_k]=[-r/s_j \quad r/s_k]$

$$\sigma_r^2=\Sigma_r=\begin{bmatrix}-r/s_j & r/s_k\end{bmatrix}\begin{bmatrix}\sigma_j^2 & 0\\ 0 & \sigma_k^2\end{bmatrix}\begin{bmatrix}-r/s_j\\ r/s_k\end{bmatrix}$$

$$\sigma_r^2=r^2[(\sigma_j/s_j)^2+(\sigma_k/s_k)^2] \tag{3.45}$$

The covariance matrix for multiple distance ratios (e.g. three) has been derived in a similar fashion (equation 3.40). Let $r_1=s_2/s_1$, $r_2=s_3/s_2$, $r_3=s_4/s_3$, with variances of σ_1^2, σ_2^2, σ_3^2 and σ_4^2 respectively. Hence, the expression for the covariance matrix becomes

$$\Sigma_r=\begin{bmatrix}c_{11} & c_{12} & c_{13}\\ c_{21} & c_{22} & c_{23}\\ c_{31} & c_{32} & c_{33}\end{bmatrix} \tag{3.46}$$

where $c_{11}=r_1^2[(\sigma_1/s_1)^2+(\sigma_2/s_2)^2]$; $c_{12}=c_{21}=-r_1r_2(\sigma_2/s_2)^2$; $c_{13}=c_{31}=0$

$c_{22}=r_2^2[(\sigma_2/s_2)^2+(\sigma_3/s_3)^2]$; $c_{23}=c_{32}=-r_2r_3(\sigma_3/s_3)^2$

$c_{33}=r_3^2[(\sigma_3/s_3)^2+(\sigma_4/s_4)^2]$

To simplify the computation, use of natural logarithms of distance ratios as observables, instead of the ratios themselves is recommended by Vincenty (1979). The following formulation for 3-D has been modified from Vincenty (1979).

The observation equation given by equation (3.44) can also be expressed as

$$v_r=r[a_1\delta x_i+a_2\delta y_i+a_3\delta z_i+a_4\delta x_j+a_5\delta y_j+a_6\delta z_j +a_7\delta x_k+a_8\delta y_k+a_9\delta z_k]-\Delta r \tag{3.47}$$

where $a_1=[(dx_j/s_j^2)-(dx_k/s_k^2)]$; $a_2=[(dy_j/s_j^2)-(dy_k/s_k^2)]$

$a_3=[(dz_j/s_j^2)-(dz_k/s_k^2)]$; $a_4=-dx_j/s_j^2$; $a_5=-dy_j/s_j^2$;

$a_6=-dz_j/s_j^2$; $a_7=dx_k/s_k2$; $a_8=dy_k/s_k^2$; $a_9=dz_k/s_k^2$

Let $\sigma_{01}=\sigma_1/s_1$ and $\sigma_{02}=\sigma_2/s_2$. Then equation (3.45) becomes

$$\sigma_r^2=r^2(\sigma_{01}^2+\sigma_{02}^2) \tag{3.48}$$

The observation equation for the logarithm of a ratio is obtained simply by dividing equation (3.47) by r. The new observation equation is

$$\begin{aligned} &a_1\delta x_i+a_2\delta y_i+a_3\delta z_i+a_4\delta x_j+a_5\delta y_j+a_6\delta z_j \\ &+a_7\delta x_k+a_8\delta y_k+a_9\delta z_k=(\Delta r/r)+v_{ln} \end{aligned} \tag{3.49}$$

where $a_1=[(dx_j/s_j^2)-(dx_k/s_k^2)]$; $a_2=[(dy_j/s_j^2)-(dy_k/s_k^2)]$

$a_3=[(dz_j/s_j^2)-(dz_k/s_k^2)]$; $a_4=-dx_j/s_j^2$; $a_5=-dy_j/s_j^2$;

$a_6=-dz_j/s_j^2$; $a_7=dx_k/s_k2$; $a_8=dy_k/s_k^2$; $a_9=dz_k/s_k^2$

Similarly, equation (3.48) becomes

$$\sigma_{ln}^2=(\sigma_{01}^2+\sigma_{02}^2) \tag{3.50}$$

The covariance matrix for multiple observables has been derived as follows. Let $r_1=s_2/s_1$, $r_2=s_3/s_2$, $r_3=s_4/s_3$, with variances of σ_1^2, σ_2^2, σ_3^2 and σ_4^2 respectively. Also, $\sigma_{01}=\sigma_1/s_1$; $\sigma_{02}=\sigma_2/s_2$; $\sigma_{03}=\sigma_3/s_3$ and $\sigma_{04}=\sigma_4/s_4$. Therefore, applying the principle of variance propagation

$$\begin{bmatrix} \ln r_1 \\ \ln r_2 \\ \ln r_3 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \ln s_1 \\ \ln s_2 \\ \ln s_3 \\ \ln s_4 \end{bmatrix} \tag{3.51}$$

Hence $\Sigma_{ln}=J\Sigma_s J^t$ where

$$J=\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \; ; \; \Sigma_s=\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

$$\Sigma_{ln} = \begin{bmatrix} \sigma_{01}^2+\sigma_{02}^2 & -\sigma_{02}^2 & 0 \\ -\sigma_{02}^2 & \sigma_{02}^2+\sigma_{03}^2 & -\sigma_{03}^2 \\ 0 & -\sigma_{03}^2 & \sigma_{03}^2+\sigma_{04}^2 \end{bmatrix} \tag{3.52}$$

Equation (3.52) is similar to equation (3.42) and covariance matrix of correlated horizontal angles (Appendix C). This demonstrates that covariance matrices for distance difference, correlated angles and logarithms of distance ratio can be computed in the same manner.

3.5.3.2.3 De-correlation of observations

Most of the surveying observables described in Appendix A are assumed to be uncorrelated, for example distances, directions and height differences. However, as shown by the derived covariance matrix of observations (for example, equations 3.42 and 3.52), the formulated pseudo observables are algebraically correlated. Hence, the covariance matrix of the observations will not be diagonal anymore.

In this study, the correlated pseudo observables have been applied into ordinary LSE using the concept of transformation or de-correlation (Milbert, 1985). The transformation process decorrelates the transformed observations, and the covariance matrix of the transformed observation equations becomes a unit matrix. In this way, the normal equations and variance factors are formed and solved as in ordinary LSE. However, the final residuals need to be transformed back again. Details on the formulation of de-correlation as given by Milbert (1985) are described here.

Let A be the design matrix, b the misclosure vector, v the residuals, Σ_l be the (positive definite) covariance matrix of measurements and R the Cholesky factor of Σ_l.

Normal equations are conventionally formed as

$$N=A^t\Sigma_l^{-1}A \text{ and } u=A^t\Sigma_l^{-1}b \tag{3.53}$$

The Cholesky factor R of Σ_l is (equation 2.43)

$\Sigma_l = R^t R$ or $\Sigma_l^{-1} = R^{-1} R^{-t}$ (3.54)

$T = R^{-t} = (R^{-1})^t$

where T is the de-correlation matrix

The de-correlation of A, b, v, Σ_l and Σ_{la} via matrix T results in A', b', v', $\Sigma_{l'}$ and $\Sigma_{la'}$

$$A' = TA = R^{-t}A \text{ or } R^t A' = A \quad (3.55)$$

$$b' = Tb = R^{-t}b \text{ or } R^t b' = b$$

$$v' = Tv = R^{-t}v \text{ or } R^t v' = v$$

$$\Sigma_{l'} = I$$

$$\Sigma_{la'} = A' N^{-1} A'^t$$

Also the quadratic form becomes

$$v^t \Sigma_l^{-1} v = v'^t v' \quad (3.56)$$

The normal equations are obtained by inserting equation (3.55) into equation (3.53), resulting in

$$A^t \Sigma_l^{-1} A = A'^t A' \text{ and } A^t \Sigma_l^{-1} b = A' b' \quad (3.57)$$

Equations (3.53) and (3.57) show that the normal equations remain invariant after de-correlation or transformation. Factor R for each set of correlated measurements needs to be computed once only at the beginning of LSE and then stored. The computations of A', B' and v' involve lower triangular linear equations, and can be computed quickly columnwise by forward reduction.

At the end of LSE, residuals are needed, especially for the purpose of gross error detection. In this case, the transformed residuals (v') need to be transformed back. From equation (3.55) it can be seen that computation of v is straightforward via a direct solution of the triangular equations.

The redundancy matrix (M), required for reliability analysis, may be obtained easily without additional computations. Usually,

$$M=(I-AN^{-1}A^t\Sigma_l^{-1})$$

For de-correlation, utilizing equation (3.55)

$$M=(I-A'N^{-1}A'^t)=\Sigma_{l'}-\Sigma_{la'} \qquad (3.58)$$

Equation (3.58) shows that the redundancy matrix is the covariance matrix of residuals $\Sigma_{v'}$ after de-correlation (see also equation 2.22 in section 2.1.3).

3.6 Modelling of gross errors

In common with the treatment of systematic errors, it is required to check and detect the presence of gross errors during the measurement process (i.e. prior to LSE) and also during LSE. Consequently, a method for gross error detection incorporating both pre-adjustment (pre-LSE checking on gross errors) and also robustified LSE (the focus of study) is presented. Both methods are described in sections 3.6.1 and 3.6.2 respectively.

3.6.1 Pre-LSE checking on gross errors

It is of prime importance that all observations are checked against systematic (section 3.5.2) and gross errors, prior to LSE. Checks on gross errors are important for initial assessment of data quality. In fact, independent checks during measurement and checks on data consistency during preliminary computation will reduce most of the gross errors in the measurement prior to LSE.

The adopted measurement scheme designed to eliminate systematic errors as mentioned in section 3.5.2, will at the same time provide quick independent checks on gross error. For example, the application of the principle of reversal in measuring horizontal angles or directions on both faces of the theodolite eliminates horizontal collimation error and provides an independent check on any measuring gross error. Further examples of independent checks on gross errors (for measuring and recording errors) are rounds of horizontal angles on different zeros, taking stadia reading in geodetic levelling, and independent sets of measured EDM distances. With the trend towards automation in the measuring process, such checks should be performed automatically.

In the preliminary computations, data can be checked against gross errors by misclosure analysis of horizontal angles, heights and triangles. Moreover, preliminary computation of provisional coordinates can also be used as a check on gross errors. A detailed explanation on this aspect may be found in Cooper (1987). It should be emphasised here that the outlined initial check procedure is capable of detecting some (not all) of the gross errors in the measurements.

3.6.2 Robustified LSE

Assuming a correct stochastic model, and no significant systematic errors in the measurements, only significant and undetected gross errors remain in the uncorrelated measurements. In this research, a procedure for robustified LSE (RLSE) has been developed for simultaneous detection and localization of the remaining multiple gross errors. RLSE is a modification of the Danish method.

The next three sub-sections examine the concept and develop the procedure of RLSE (sections 3.6.2.1 and 3.6.2.2) followed by the derivation of effects of de-weighting during RLSE (section 3.6.2.3).

3.6.2.1 Concept of RLSE

The basic concept used in RLSE is based on the Danish method proposed by Krarup (Caspary, 1987b; Straub, 1983). In the Danish method, large (estimated) residuals are associated with gross errors.

The objective of this method is to find observations which are not consistent with the majority and to exclude them from the LSE by reducing their weights (i.e. de-weighting). In this manner, weights are treated as dynamic quantities such that only the consistent measurements are used effectively in the LSE process. It is expected that the effects of gross error on the final estimation will be insignificant.

Application of the Danish method is very simple (Straub, 1983). After a conventional LSE using a priori weights, the estimation is repeated several times during which the weights of certain measurements are reduced according to their residuals after the preceeding estimation. Weights of the observations with higher residuals are reduced (i.e. low weights), while the weights of observations with lower residuals (i.e. lower than a certain limit) are held stable (equation 3.59a, Appendix B).

With a proper and suitable choice of de-weighting function and constant (factor) c, convergence is achieved. In the final solution, weights of outlying observations will approach zero, effectively removing them from the LSE. Moreover, the observations affected by gross error are found with corrections of the same order of the magnitude of their corresponding residuals but with reversed sign.

In practice, the de-weighting schemes for the Danish method are based on trial and error, see for example Kubik et al (1985), Jorgensen et al (1985), Straub (1983) and Kubik et al (1988). Some of the de-weighting schemes are given in Appendix B. However, simulation studies with known gross errors indicate that most of the de-weighting schemes either flagged additional measurements, or were unable to detect some of the errors. During the course of this research (Setan, 1992), it is found from experience that the following de-weighting scheme (modified from Caspary, 1987b) is acceptable:

$w_i=1/\sigma_i^2$ (3.59a)

if $|\hat{v}_i|\leq$limit; $p_i=1.0$; $w_i'=p_i*w_i$ (weight unchanged)

if $|\hat{v}_i|>$limit; $p_i=e^{-f}$; $w_i'=p_i*w_i$ (weight changed)

limit$=c*\sigma_i*\hat{\sigma}_o$

$f=|\hat{v}_i|/(c*\sigma_i*\hat{\sigma}_o)$

the weighting factor c is usually between 2.0 and 3.0.

If use the normalized residuals (equations 3.17b and 3.18)

$\hat{v}_{i'}=\hat{v}_i/(\hat{\sigma}_o\sigma_i r^{1/2})$:

if $|\hat{v}_i'|\leq c$; $p_i=1.0$; $w_i'=p_i*w_i$ (weight unchanged)

if $|\hat{v}_i'|>c$; $p_i=e^{-f}$; $w_i'=p_i*w_i$ (weight changed)

In principle, the developed technique of RLSE uses the above de-weighting scheme (equation 3.59a) and provides a facility for variation of the weighting factor. In addition, global and local tests are used as stopping criteria, together with reliability analysis (section 3.8) to determine the capability of gross error detection.

A global test on the estimated variance factor, employing the chi-square test (section 3.3.1) is useful as an initial check on the residuals of LSE. For the purpose of gross error detection, it is expected that the estimated variance factor will be bigger, and hence a one-tailed test can be adopted (equations 3.4 and 3.6).

The passing of the global test does not necessarily indicate that the measurements are free from gross errors as such a test is not very sensitive. Therefore, a local test is applied to each normalized residuals thereby examining each observation in turn. Tests such as those based on Baarda (equations 3.23 and 3.24) and Pope (equations 3.19 and 3.21) can be used for this purpose.

In practice, only type I error is considered (section 3.3). In addition, Pope's test is more adaptable than Baarda's test (Teskey, 1994), due to its high sensitivity. Taking all these into consideration, Pope's test is found more suitable, and is recommended for RLSE.

3.6.2.2 Procedure for RLSE

The formulation of the developed procedure for RLSE of uncorrelated measurements can then be summarized as follows:

(a) During and after ordinary LSE (section 2.1.3), using the weights w_i, compute the estimated residuals $\hat{v}_i$ and the estimated variance factor $\hat{\sigma}_o^2$. The one-tailed global test (equations 3.4 and 3.6) is employed. As the global test is not sensitive enough, a local test based on Pope (equations 3.19 and 3.21) is also employed. If both tests pass, the LSE results are acceptable, and step (e) (below) can be executed for reliability analysis. Otherwise, a RLSE must be performed, via steps (b), (c) and (d).

(b) Define the de-weighting scheme (equation 3.59a) and weighting factor c. It is recommended that RLSE is begun with a factor c of 3.0 (i.e. bigger factor) to avoid the possibility of additional flagging of good or acceptable measurements. The same factor c is applied during RLSE until no observation weights are changed. Factor c is then reduced interactively by 0.1 (new c becomes 2.9, 2.8 and so on) until the weights of some observations are changed. If c becomes too low, for example less than 1.5, RLSE should be stopped. The use of normalized residuals (equation 3.59a) is also useful.

(c) Compute new weights for all the measurements (i.e. w_i'). In this manner, weights of measurements with $\hat{v}_i$ greater than the limit (i.e. suspect measurements), will be reduced. Otherwise, the previous weights are maintained (equation 3.59a).

(d) A new LSE is carried out using the new weights. This process of re-estimating and de-weighting (a, b and c) is iterated until convergence is achieved, using the selected c factor in (b). If both global (equations 3.4 and 3.6) and local (equations 3.19 and 3.21) tests are passed, the procedure is stopped. Otherwise, the c factor is reduced again, and the procedure is repeated until the termination criteria are met (i.e. both global and local tests are passed).

(e) At this stage, two options are possible, either direct or indirect use of RLSE. In both options, statistical testing (section 3.3) and reliability analysis (section 3.8) are useful for proper assessment of the final result.

<u>(i) Direct use of RLSE</u>

Once the stopping criteria are met, final computations of RLSE can also be performed, and all the changed weights will be drastically reduced to zero, for example weights of 1e-10. The final weights must not be too small in order to avoid numerical instability. The purpose is to minimize or greatly reduced the effects of measurements with gross errors on the final solutions of coordinates, trace, rmse and residuals. Results of the RLSE will be the same as ordinary LSE without the erroneous or flagged observations, as in (ii) below.

<u>(ii) Indirect use of RLSE/ new LSE</u>

It is advisable to check the results of LSE. In particular, erroneous measurements need to be carefully examined, and deleted or remeasured as necessary. In this case, a new LSE must be carried out with the new data, to arrive at a final result.

To speed up the computations of RLSE, the earlier de-weighting scheme (equation 3.59a) has been modified as:

$$w_i=1/\sigma_i^2 \tag{3.59b}$$

if $|\hat{v}_i|\leq$limit; p_i=1.0; $w_i'=p_i*w_i$

(weight unchanged)

if $|\hat{v}_i|>$limit; p_i=f; $w_i'=p_i$

(weight changed)

limit=$c*\sigma_i*\hat{\sigma}_o$

f is the smallest weight, for example 1e-10

the weighting factor c is between 2.0 and 3.0

In equation (3.59b), observations with residuals greater that the limit are de-weighted drastically closer to zero. Another possibility has been developed by using the normalized residuals (equation 3.17b) as the following:

if $|\hat{v}_{i'}|\leq$limit; p_i=1.0; $w_i'=p_i*w_i$ (3.59c)

(weight unchanged)

if $|\hat{v}_{i'}|>$limit; p_i=f; $w_i'=p_i$

(weight changed)

normalized residual=$\hat{v}_{i'}=\hat{v}_i/(\hat{\sigma}_o\sigma_i r^{1/2})$

limit=$\tau_{\alpha,r}$

f is the smallest weight, for example 1e-10

Equation (3.59c) incorporates Pope's method (section 3.3.4) into the RLSE. In the presence of multiple gross errors, equations (3.59a), (3.59b) and (3.59c) will de-weight more than one observation simultaneously. Residuals can also be normalized by adopting Baarda's method (section 3.3.4).

A simulation study carried out during this research indicated the use of normalized residuals with de-weighting scheme of equation (3.59a) as the most suitable RLSE scheme. If observations are de-weighted drastically to speed up the computation (equations 3.59b and 3.59c), it is possible that some good observations may be flagged as gross errors and some gross errors may be left undetected.

The method of RLSE is capable of detecting and localizing gross error correctly since both global and local tests are used to verify the results. Another important aspect is related to reliability analysis (section 3.8). Redundancy numbers together with MDGE will indicate whether a gross error can be detected or not. If the redundancy number for a particular measurement is very small (i.e. very large MDGE), gross errors in that measurement will be left undetected since there is very little controllability. On the other hand, a redundancy number close to unity indicates that the effects of that measurement (and gross error in it) on the solution are almost insignificant.

3.6.2.3 Effects of de-weighting during RLSE

The method of iterative RLSE for the detection and localization of multiple gross errors (section 3.6.2.2) has the following properties: approximate magnitudes of gross errors (with opposite signs) are directly recovered in the bigger residuals, and the final solutions are relatively insensitive to the presence of the detected and localized gross errors. Moreover, if weights of suspected measurements are greatly reduced (for example to 1e-10 but not exactly zero to avoid numerical instability), the final estimated parameters, precisions and residuals from RLSE are the same as ordinary LSE if the gross errors were removed.

Clearly, RLSE is based on the concept of de-weighting suspected measurements iteratively via a suitable weighting function. In principle, as will be shown later, the de-weighting process will change the estimated parameters, their cofactor matrix and redundancy numbers (i.e. $d\hat{x}$, $dQ_{\hat{x}}$ and r). In this study, the effects of the de-weighting process on the solutions of RLSE has been derived via a step by step procedure.

The relevant equations for linear LSE (section 2.1.3) are:

Observation equations:

$$Ax=b+v;\ W=Q_l^{-1}=\sigma_o^2\Sigma_l^{-1} \tag{3.60}$$

Normal equations:

$$A^tWAx=A^tWb \text{ or} \tag{3.61}$$
$$Nx=u \text{ where } N=A^tWA \text{ and } u=A^tWb$$

Solution:

$$\hat{x}=(A^tWA)^{-1}A^tWb=N^{-1}u=Q_{\hat{x}}u \tag{3.62}$$

where $Q_{\hat{x}}=N^{-1}$

Let W=the a priori weight matrix

W'=new weight matrix=W+dW

$\hat{x}$'=new parameters=$\hat{x}$+d$\hat{x}$

For simplicity, x, Q_x and Q_v will be used to represent $\hat{x}$, $Q_{\hat{x}}$ and $Q_{\hat{v}}$ respectively, in this section.

The new normal equations with new weight become

$$A^t(W+dW)A(x+dx)=A^t(W+dW)b$$

$$(A^tWA+A^tdWA)(x+dx)=A^tWb+A^tdWb \tag{3.63}$$

From equation (3.61), $N=A^tWA$ and $u=A^tWb$. Let

$$dN=A^tdWA \text{ and } du=A^tdWb \tag{3.64}$$

Combining equations (3.63) and (3.64) gives

$$(N+dN)(x+dx)=u+du$$

$$Nx+Ndx+dNx+dNdx=u+du \tag{3.65}$$

But Nx=u and Nx-u=0. Equation (3.65) can then be written as

$$(N+dN)dx+dNx=du$$

$$(N+dN)dx=du-dNx$$

$$dx=(N+dN)^{-1}(du-dNx) \tag{3.66}$$

Substituting equation (3.64) into equation (3.66) produces

$$dx=(N+dN)^{-1}(A^tdWb-A^tdWAx)$$

$$dx=(N+dN)^{-1}(A^tdW[b-Ax]) \tag{3.67}$$

As Ax=b+v, then b-Ax=-v. Equation (3.67) can be simplified as

$$dx=-(N+dN)^{-1}A^t dWv$$
$$dx=-(N+A^t dWA)^{-1}A^t dWv \qquad (3.68)$$

Employing the matrix equality (Mikhail, 1976)

$$\text{if} \quad X=Y+UZV \qquad (3.69)$$
$$\text{then} \quad X^{-1}=Y^{-1}-Y^{-1}U(Z^{-1}+VY^{-1}U)^{-1}VY^{-1}$$

in this case, $(N+A^t dWA)$ is actually $(Y+UZV)$. Therefore

$$N=Y,\ A^t=U,\ dW=Z \text{ and } A=V \qquad (3.70)$$

Applying equation (3.69) into part of equation (3.68) results in

$$(N+A^t dWA)^{-1}=N^{-1}-N^{-1}A^t[dW^{-1}+AN^{-1}A^t]^{-1}AN^{-1}$$
$$=N^{-1}-N^{-1}A^t[dW(I+AN^{-1}A^t dW)^{-1}]AN^{-1}$$
$$(N+dN)^{-1}=N^{-1}-N^{-1}A^t dW(I+AN^{-1}A^t dW)^{-1}AN^{-1} \qquad (3.71)$$

Equation (3.71) can also be written as

$$Q_{x'}=Q_x+dQ_x \qquad (3.72)$$

where

$$Q_{x'}=(N+dN)^{-1},\ Q_x=N^{-1}$$
$$dQ_x=-N^{-1}A^t dW(I+AN^{-1}A^t dW)^{-1}AN^{-1}$$

Substituting equation (3.71) into (3.68) leads to

$$dx=-(N+A^t dWA)^{-1}A^t dWv$$
$$=-[N^{-1}-N^{-1}A^t dW(I+AN^{-1}A^t dW)^{-1}AN^{-1}]A^t dWv$$
$$=-[N^{-1}A^t dWv-N^{-1}A^t dW(I+AN^{-1}A^t dW)^{-1}AN^{-1}A^t dWv]$$
$$=-N^{-1}A^t dW[v-(I+AN^{-1}A^t dW)^{-1}AN^{-1}A^t dWv]$$

$$dx=-N^{-1}A^tdW(I+AN^{-1}dW)^{-1}(I+AN^{-1}A^tdW)[v-$$
$$(I+AN^{-1}A^tdW)^{-1}AN^{-1}A^tdWv]$$
$$=-N^{-1}A^tdW(I+AN^{-1}A^tdW)^{-1}[(I+AN^{-1}A^tdW)v-AN^{-1}A^tdWv]$$
$$=-N^{-1}A^tdW(I+AN^{-1}A^tdW)^{-1}[v+AN^{-1}A^tdWv-AN^{-1}A^tdWv]$$

$$dx=-N^{-1}A^tdW(I+AN^{-1}A^tdW)^{-1}v \quad (3.73)$$

The effects of de-weighting on parameters and cofactor matrix are given by equations (3.72) and (3.73) respectively, i.e.

$$dx=-N^{-1}A^tdW(I+AN^{-1}A^tdW)^{-1}v \quad (3.74)$$
$$dQ_x=-N^{-1}A^tdW(I+AN^{-1}A^tdW)^{-1}AN^{-1}$$

The new parameters and cofactor matrix become

$$x'=x+dx \quad (3.75)$$
$$Q_{x'}=Q_x+dQ_x$$

Further simplification is obtained by incorporating cofactor matrices Q_l and Q_v (Cooper, 1994). An expression for Q_v is (equation 2.22)

$$Q_v=W^{-1}-AN^{-1}A^t \text{ and } AN^{-1}A^t=W^{-1}-Q_v=Q_l-Q_v \quad (3.76)$$
where $W^{-1}=Q_l$ from equation (2.4).

Substituting equation (3.76) into equation (3.74) produces

$$dx=-Q_xA^tdW[I+(Q_l-Q_v)dW]^{-1}v \quad (3.77)$$

Similarly, an expression for the effects on cofactor matrix is

$$dQ_x=-Q_xA^tdW[I+(Q_l-Q_v)dW]^{-1}AQ_x \quad (3.78)$$

The above general expression can be written specifically for de-weighting of some observations only. This is very important because during RLSE, only a group of observations are de-weighted.

Close examination of equation (3.74) for dx and dQ_x shows the same pattern as the expression for sequential LSE with the same parameters, as shown in Mikhail (1976), Cooper (1987) and Cross (1983). Due to the similar pattern, the derivation of an expression for de-weighting of some observations only, will be the same as in sequential LSE, and will be given here without going into detail.

The measurements can be divided into two groups: group 1 contains the original observations (or more precisely, observations of unchanged weights), and group 2 contains the de-weighted observations. Observations equations are

$$A_1x=b_1+v_1 \quad \text{for group 1} \qquad (3.79)$$
$$A_2x=b_2+v_2 \quad \text{for group 2}$$

Let n be the number of original observations, m the number of de-weighted observations and u be the number of parameters. Expression for dx and dQ_x (i.e. equation 3.74) become

$$dx=-N^{-1}A_2^t dW_2(I+A_2N^{-1}A_2^t dW_2)^{-1}v_2 \qquad (3.80)$$
$$dQ_x=-N^{-1}A_2^t dW_2(I+A_2N^{-1}A_2^t dW_2)^{-1}A_2N^{-1}$$

Dimensions of the relevant matrices are dx(u,1), N^{-1}(u,u), A_2(m,u), dW_2(m,m), I(m,m), v_2(m,1) and dQ_x(u,u). The dimensions of the new matrix to be inverted is (m,m).

From the relevant dimensions of matrices involved, it is clear only smaller matrices (m,m) need to be inverted. This is the same as the number of de-weighted observations. As in sequential LSE, the same provisional coordinates are required for computation. In general, the step by step procedure for computing the effect of de-weighting has similar patterns and advantages as sequential LSE.

The above formulae (equation 3.80) also demonstrate that the effect of de-weighting on x and Q_x is a function of the design matrix in addition to the changes in weights.

A more direct approach of computing effects of de-weighting is obtained via simultaneous computation of RLSE. By taking differences of estimated coordinates and cofactor matrices (equation 3.62) before and after de-weighting, the effects can be determined directly. This approach is more practical as it forms part of the RLSE process.

Redundancy numbers, useful for reliability analysis will also be affected by de-weighting. A rule of thumb indicates that, as weights of a group of observations are reduced, their redundancy numbers will be increased, and redundancy numbers for other observations will be slightly reduced. The expression for the redundancy number of observation i (equation 3.18) is given by

$$r_i=1-a_iN^{-1}a_i^tw_i \tag{3.81}$$

De-weighting of observation i will reduce w_i and hence increase r_i. As the sum of r_i is the number of degree of freedom (section 3.8), redundancy numbers for other observations will be reduced or decreased.

If weights of suspected observations are greatly reduced to zero, their redundancy numbers will be close to one, indicating the insignificance of such observations towards the final estimation. As mentioned in section 3.6.2.2, the final estimated parameters and cofactor matrix via RLSE are expected to be close to the results of ordinary LSE (after removal of the suspected observations).

3.7 Stochastic modelling

In LSE, assuming no significant systematic and gross errors, the remaining errors are in the stochastic model. This section deals with the stochastic modelling of uncorrelated surveying observables. The method of simplified variance component estimation (VCE) has been extended for this purpose.

3.7.1 Need for stochastic modelling

In general, although being ignored in most cases, the process of LSE requires a realistic stochastic model (section 3.4), and hence proper stochastic modelling.

Two important aspects in stochastic modelling are the information on a priori Q_l and the correlations amongst the observations. For simplicity, it is generally assumed that survey observations are independent and uncorrelated. This approach leads to diagonal Q_l (sometimes called the weight matrix W) matrices and easy manipulation for LSE (section 2.1.3). The relationship between Σ_l, Q_l and W for uncorrelated observations is given by equation (2.15b)

$$\Sigma_l=\sigma_o^2 Q_l \text{ where } q_{ii}=\sigma_i^2 \tag{3.82}$$
$$W=\sigma_o^2\Sigma_l^{-1} \text{ where } w_i=\sigma_o^2/\sigma_i^2$$

Current advances in instrumentation and observational methods allow very high precision observations (i.e. small variances or small weights) to be made with angular measurements standard deviations of within one second of arc, and of linear measurements (height difference and distance) at mm level (Appendix C).

The methods for estimating the variances (or standard deviations) of the observations can be based on study of repeated measurements (Cooper, 1974), experiments or previous performance (Cross, 1983; Cooper and Cross, 1988; section 3.4). The estimated variances are used as variances of observations during LSE. The method of experiments and previous performance are often used in practice.

Experiments utilise the instruments and procedures similar to the actual field work in order to estimate the variances, for example calibration of EDM instruments. Previous performance of instruments and methods are in the form of previous experience, manufacturer's specifications or reports, and research findings in scientific publications. In practice, variance based on previous performance is found to be acceptable.

Appendix C gives an example of the formulae for computing the variances of common uncorrelated surveying observables. Further details are given by Blachut et al (1979) and Secord (1986).

In some cases, the adopted angular measurements are horizontal angles. However, such angles are actually not measured directly, but are derived indirectly from the measured uncorrelated horizontal directions. In fact, angles are derived from the differences of directions. Hence, horizontal angles, in the same manner as distance differences, are algebraically correlated. The cofactor matrix for angles may be derived using the concept of error propagation (Appendix C), and the expression for Q_l is analagous to equation (3.42) in section 3.5.3.2.1 describing distance difference error propagation.

In section 3.5, pseudo observables (distance ratios and differences) are introduced for handling systematic errors in distances. Although distance measurements are uncorrelated, the derived pseudo observables are algebraically correlated, and the cofactor matrix is non-diagonal. Computations of the elements of the cofactor matrix for pseudo observables have been presented in detail in section 3.5.3.2.

In a more general situation of correlated observations, proper stochastic modelling can be carried out via a numerical method of variance component estimation (VCE), as discussed in detail by Chen et al (1990b) and Caspary (1987b). However, such a procedure requires extensive computations.

In the case of independent and uncorrelated observations, the most important aspect in stochastic modelling is the proper determination of variance (or standard deviation) of the observations. A more simple method for modelling in terms of computational effort, known as simplified VCE is adopted and extended in the following section.

3.7.2 Principle of simplified VCE

LSE (section 2.1.3) is based on the following linearised equations

$$Ax=b+v \tag{3.83}$$
$$\hat{x}=N^{-1}u=(A^tWA)^{-1}A^tWb$$
$$\Sigma=\sum_{i=1}^{k} \sigma_{oi}^2 Q_{li}$$

where k is the number of observation groups.

The basic problem related to the general VCE is to simultaneously estimate parameters $\hat{x}$ and k variance factors $\hat{\sigma}_{oi}^2$. The general stochastic model can be written as

$$\Sigma=\sigma_{o1}^2Q_{l1}+\sigma_{o2}^2Q_{l2}+...+\sigma_{o1}^2Q_{l1} \tag{3.84}$$

In the above equation, each group of observations has its own error characteristics and should be modelled separately.

The rigorous solution for general VCE is iterative and requires considerable computational effort in order to arrive at a unique Best Invariant Quadratic Unbiased Estimation (BIQUE) of $\hat{\sigma}_o^2$. See Caspary (1987b) for details on VCE and the relevant computational aspects.

To reduce the computations and arrive at the same results, Caspary (1987b) uses an approximate method, termed simplified VCE. It is assumed that all observations are independent and of equal variance. The following equations (section 2.1.3) are relevant

$$Ax=b+v \tag{3.85}$$
$$\Sigma=\sigma_o^2I$$
$$\hat{\sigma}_o^2=v^tWv/r=v^tWv/(n-u)$$

where r is number of the degrees of freedom, n the number of observations, u the number of parameters and v the estimated residuals ($\hat{v}$).

The observations are partitioned, according to their types, into j groups, for example group one contains distance measurements, group two direction observations, and so on. The partitioning of observations, residuals, variances and weights can be expressed as

observations: $l^t=[l_1^t,l_2^t,...,l_j^t]$ (3.86a)

residuals: $v^t=[v_1^t,v_2^t,...,v_j^t]$

variances: $\sigma^2=[\sigma_{o1}^2,\sigma_{o2}^2,...\sigma_{oj}^2]$

weights: $w^t=[w_1^t,w_2^t,...,w_j^t]$

$=[1/\sigma_{o1}^2,1/\sigma_{o2}^2,...1/\sigma_{oj}^2]$

Hence, $v^tWv=v_1^tw_1v_1+v_2^tw_2v_2+...+v_j^tw_jv_j$

The variance factor for group j of observations is

$$(\hat{\sigma}_{oj}^2)_{i+1}=(v_j^t W_j v_j/r_j)_i \tag{3.86b}$$

It is then required simultaneously to determine parameters and the variance factor for each group of observations. Some form of iterative procedure is required, and the extended procedure is described in section 3.7.3.

3.7.3 Computational procedure

In this research, the computational procedure for simplified VCE has been extended from Caspary (1987b), with the addition of statistical test for global and local (or group) estimated variance factors, as stopping criterion.

The procedure can be summarized as the following. Let j be the number of groups, and i the number of iterations. Initially, set the global variance factor σ_o^2 to be unity (1.0) and select the initial diagonal weight matrix W as I. In other words, for each group,

$$(\hat{\sigma}_{oj}^2)=1.0,\ W_j=1.0 \tag{3.87}$$

The computations can be summarised below:

<u>a. In i iterations</u>

$$(W_j)_i=1/(\hat{\sigma}_{oj}^2)_i \tag{3.88}$$

Perform ordinary LSE and estimate both the global and local (group) variance factors,

(i) global variance factor

$$\hat{\sigma}_o^2=v^tWv/r \text{ (as equation 2.23)} \tag{3.89}$$

(ii) local variance factor for j groups

$$(\hat{\sigma}_{oj}^2)_{i+1}=(v_j^t W_j v_j/r_j)_i \tag{3.90}$$

where r_j is sum of redundancy numbers for group j.

The new standard deviation becomes

$$(\hat{\sigma}_{oj}^2)_{i+1}=(\hat{\sigma}_{oj}^2)_1(\hat{\sigma}_{oj}^2)_2...(\hat{\sigma}_{oj}^2)_{i+1} \qquad (3.91)$$

b. Test statistically the estimated variance factors

Both global and local variance factors can be tested statistically using the test on the estimated variance factor of section 3.3.1 (via equations 3.4 and 3.5). In this case

For the global variance factor

$$H_o: \hat{\sigma}_o^2=1.0 \text{ and } H_a: \hat{\sigma}_o^2\neq1.0 \qquad (3.92)$$
$$T=v^tWv\sim\chi^2_r$$
$$\text{accept } H_o \text{ if } \chi^2_{1-\alpha/2,r}<T<\chi^2_{\alpha/2,r}$$

For the local variance factor

$$H_o: \hat{\sigma}_{oj}^2=1.0 \text{ and } H_a: \hat{\sigma}_{oj}^2\neq1.0 \qquad (3.93)$$
$$T=v_j^tW_jv_j\sim\chi^2_{rj}$$
$$\text{accept } H_o \text{ if } \chi^2_{1-\alpha/2,rj}<T<\chi^2_{\alpha/2,rj}$$

If all estimated variance factors are not equal to one statistically (i.e. H_o is rejected), the procedure outlined in (a) is iterated using new weights until the tests are passed. At the end of iteration, all variance factors will be statistically equal to one.

The purpose of incorporating the statistical test is to speed up the computations. The relevant test here is the two-tailed chi-square test on the estimated variance factors (equations 3.92 and 3.93). During VCE, it is required that the estimated variance factors be equal to one as a termination limit for iteration. By applying this test, the VCE procedure can be stopped once the estimated global and local variance factors are not significantly different from one and pass the test.

In using VCE, it is important that sufficient redundancies are available for each group of observations. Because VCE assumes the errors as normally distributed, the presence of systematic or gross errors will corrupt the results of VCE. For proper usage of VCE, both systematic and gross errors need to be eliminated first.

3.8 Precision and reliability analyses

In section 3.2, aspects of precision and reliability are highlighted to demonstrate their importance as measures of quality in a monitoring network. Both precision and reliability measures are functions of observable precision and network geometry. In this section, the relevant equations for uncorrelated observations are presented. Further details are given by Cross (1983), Cooper (1987), Caspary (1987b) and Gruendig and Bahndorf (1984).

The most commonly used measures of precision for u parameters are: trace; mean variance; variances of parameters and; error ellipses and ellipsoids.

(i) Trace of the cofactor matrix of the parameters (global)

The trace of $Q_{\hat{x}}$ is the sum of diagonal elements of $Q_{\hat{x}}$ (equation 2.20).

$$tr(Q_{\hat{x}})=[\hat{\sigma}_1^2+\hat{\sigma}_2^2+...+\hat{\sigma}_u^2)$$

(ii) Mean variance (global)

$$\hat{\sigma}_m^2=tr(Q_{\hat{x}})/u$$

Sometimes, the square root of the mean variance is used, known as root mean square error (RMSE). For a 3-D network, both trace and mean variance may be computed separately for x, y and z axes. Usually, the highest precision LSE solution is the minimum trace solution.

(iii) Variances of individual parameters (local)

The precision of each parameter is given by the appropriate diagonal elements of $Q_{\hat{x}}$.

(iv) Error ellipses and ellipsoids (local)

Local precision can be represented graphically by the confidence regions of the parameters (section 3.3.5). These are derived from $Q_{\hat{x}}$, in the form of absolute and relative error ellipses (2-D network) and ellipsoids (3-D network). The computations of the axes and orientations of the error ellipsoid can be carried out using EVD as outlined in section 2.2.5.

Reliability analysis provides two measures, internal and external reliability (Baarda, 1977). Datum independent quantities are useful (equations 3.25 and 3.26).

(i) Internal reliability (redundancy number and MDGE)

Internal reliability shows the controllability of observations. A smaller MDGE indicates a more reliable observation, whilst a larger MDGE indicates a less reliable observations. Gross error in any observations that is less than its MDGE will be left undetected. An expression for MDGE of observation i (equation 3.25b) is (Gruendig and Bahndorf, 1984; Biacs, 1989)

$$MDGE_i=\nabla_i=\sigma_i(\lambda_o/r_i)^{1/2}=(\lambda_o/r_iw_i)^{1/2}$$

where σ_i=standard deviation of observation i

r_i=redundancy number of observation i

λ_o=non-centrality parameters, computed from α_o and β_o

w_i=weight of observation i

Value of λ_o is 17 for α_o 0.1 % and β_o 20 % (Figure 3.1). Alternatively, λ_o may also be computed using Tau factor as outlined by Cooper and Cross (1988).

The redundancy number (or local redundancy) describes the contribution of the ith observation to the overall redundancy of the model. Redundancy number r_i can be computed for each uncorrelated observation i as (equations 3.18 and 3.25a)

$$r=tr\ (Q_{\hat{v}}W)$$

$$Q_{\hat{v}}=W^{-1}-AN^{-1}A^t$$

$$r_i=1-a_iN^{-1}a_i^tw_i$$

$$=1-h_i \text{ if } h_i=a_iN^{-1}a_i^tw_i$$

From the above equations, relationship between redundancy and the number of parameters can be seen. The sum of all r_i is equal to the redundancy, and the sum of all h_i is equal to the number of the parameters.

If r_i is unity, observation i will not influence the solution. In relation to RLSE, the effects of gross errors are greatly reduced or insignificant, as de-weighting of observation i will increases r_i close to unity. If w_i is closer to zero, r_i will be closer to unity. On the other hand, if r_i is zero, gross errors in observation i will never be detected, indicating a weak area of the network, and the MDGE becomes too large. The mean of all redundancy numbers is also useful.

The use of redundancy numbers is recommended as they are very useful in determining weak areas within a network. An observation with a redundancy number close to zero indicates that it is unreliable and a gross error in it could be undetectable. In such a case, additional measurements may be needed to strengthen the network.

<u>(ii) External reliability</u>

External reliability measures the influence of a gross error on the unknown parameters, i.e. the change of x_j caused by the MDGE ∇_i. The expression of external reliability is (Caspary, 1987b)

$$\nabla x_j = -(A^t W A)^{-1} A^t W \nabla_i$$

The above measure is datum dependent. A more useful datum independent measure can be obtained for each observation. Gruendig and Bahndorf (1984) show the expression for the datum independent external reliability. For each observation i, the external reliability is the measure of the maximum effect of undetected MDGE, in the form of influential factor or global distortion parameters (δ_i)

$$\begin{aligned}\delta_i &= (1-r_i) w_i \nabla_i^2 \\ &= \lambda_o (1-r_i)/r_i\end{aligned}$$

A good and high reliability is guaranteed if observations control each other (i.e. smaller MDGE, larger r_i), and if the influence of a blunder on the parameters of the network is small (i.e. smaller influential factor). The minimum value of any redundancy number for general purpose surveying observables (Caspary, 1987b) should be 0.3. For a monitoring network of high accuracy, r_i should be greater than 0.8 (Cooper, 1994), the mean of all r_i about 0.5 (Caspary, 1987b), MDGE ranges between (6-8)σ_i, and δ_i smaller than 100 (Baarda, 1977), preferably between 40-70.

3.9 A strategy for LSE of single epoch

The strategy developed for LSE (using observation equations) of a single epoch (incorporating chapters 2 and 3) can be summarised as Figure 3.2. Starting with a chosen mathematical model, which consists of both the functional (Appendix A, section 3.5) and stochastic (Appendix C, section 3.7) parts, the observation and normal equations are formed (section 2.1.3). The datum definition (section 2.2.2) being simply constructed via minimum constraints of fixed coordinates.

The normal equation is analysed for rank deficiency by means of simplified EVD (section 2.2.6). As the datum defect is known, any deficiency is related to configuration, and additional observations may be required. If the normal equation matrix is non-singular, the LSE solutions are computed (section 2.1.3), followed by the global test (section 3.3.1), local test (section 3.3.4), precision and reliability analyses (section 3.8). Failure of the global test indicates the presence of either systematic or gross errors. If the system is ill-conditioned, low redundancy numbers (section 3.8) indicate weak area of the network, and it may be necessary to check the configuration.

Systematic errors in EDM distances (i.e. scale, zero errors) can be tackled by improving the functional model via additional parameters (sections 3.5.3.1 and 3.3.3) or pseudo observables (3.5.3.2). The correlated pseudo observables can be handled via de-correlation of observations (section 3.5.3.2.3). Gross errors can be detected by means of iterative RLSE procedure (section 3.6.2). Assuming no systematic and gross errors in the observations, the method of VCE (section 3.7.3) is useful for estimating variances of groups of observations.

It is also possible to carry out precision and reliability analyses (section 3.2) via pre-analysis. In this manner, the network quality can be determined prior to the actual LSE.

It is recommended that LSE results are accepted only if both the global and local tests (sections 3.3.1, 3.3.2 and 3.3.4) are passed. This is important because sometimes global test is not sensitive enough. In addition, the precision and reliability analyses (section 3.8) must be acceptable too. The results from LSE for the purpose of deformation detection consist of the estimated variance factor ($\hat{\sigma}_o^2$), degrees of freedom (r or df), number and types of datum defect (d), number of stations (m), provisional 3-D coordinates (x_o), estimated 3-D coordinates ($\hat{x}_a$) and their cofactor matrices ($Q_{\hat{x}}$).

If required, S-transformations (section 2.3.5) can be used to transform the results of LSE (i.e. coordinates and cofactor matrix) from minimum constraints datum into minimum trace, partial minimum trace or other minimum constraints datum. Invariant quantities ($\hat{v}$, $\hat{\sigma}_o^2$, r and d) remain the same during such transformation.

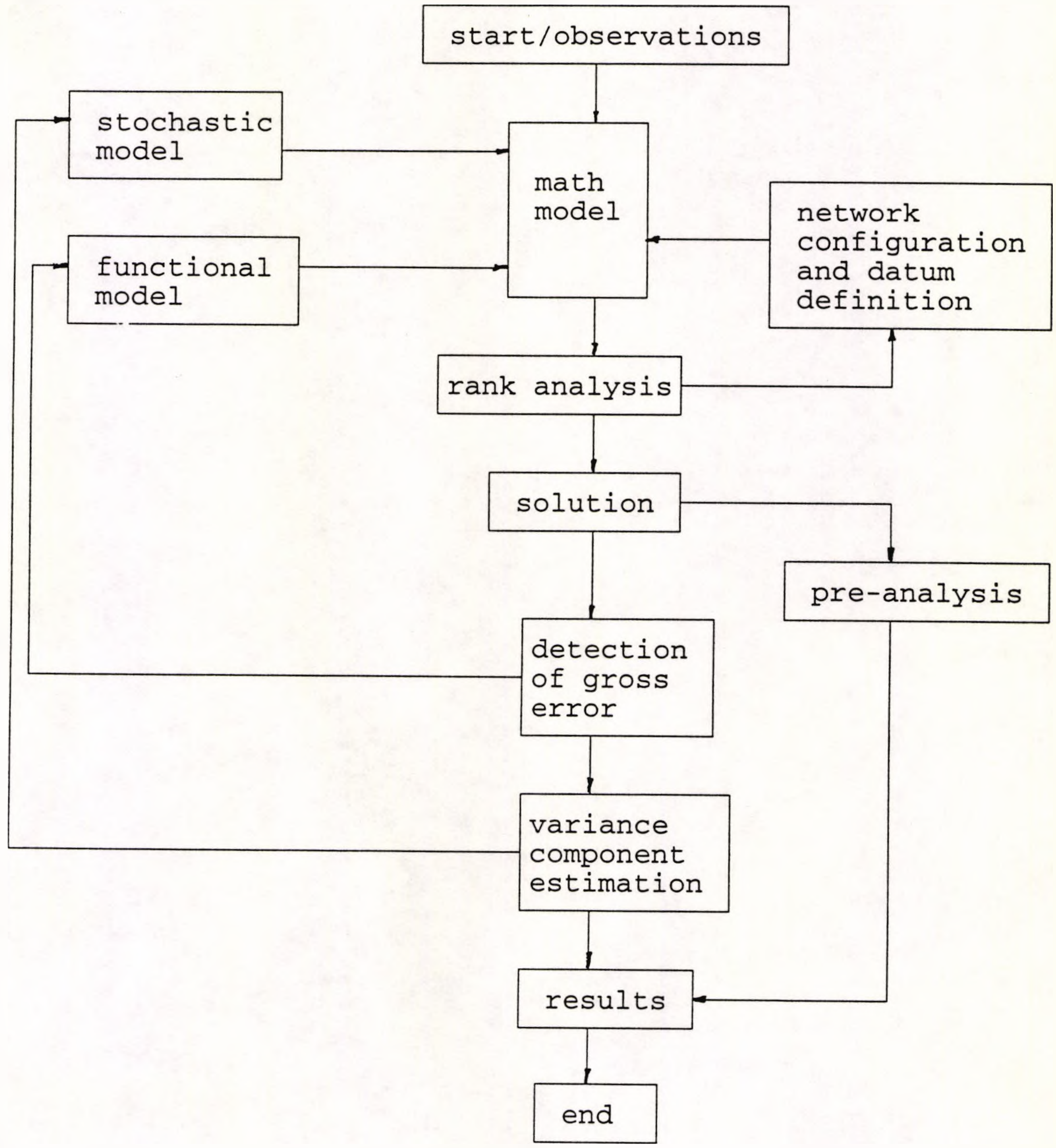

Figure 3.2 Strategy for LSE of a single epoch

4. DETECTION OF SPATIAL DEFORMATION

During the detection of deformation, displacement vectors are referred to a set of common stable datum points between any two epochs. Hence, the analysis procedure requires amongst others, the transformation of LSE results into a common datum, identification of a set of stable common points, and the localization of deformation together with the appropriate statistical testing (globally and locally).

The main aim of this chapter is to highlight the above requirements and present the analysis procedure developed for the geometrical detection of spatial deformations via one-stage computation.

4.1 Requirements for detection of deformation

For most engineering applications, deformation detection is based on two-epoch analysis (section 1.1). The main aims of deformation detection are (Caspary, 1987b): to confirm the stability of datum or reference points; to detect and determine any significant deformation or movement (with respect to the stable datum points) and; to provide a graphical representation of deformation vectors. Statistical tests are used to verify the estimated results.

The general procedure for detection of deformation assumes common stations, and similar datum definition and defects between the two epochs. In other words, the analysis is restricted to common points only. However, in practice, epochs may have differing network configuration, different numbers and types of observations, different numbers of stations, and possibly different datum definition and defects. Therefore, it may be necesary to transform the LSE results of each epoch into common stations and datum prior to detection of deformation (section 2.3.1).

Another important aspect is that no stations are to be assumed stable until tested for stability. Hence, a method for identifying and testing the stable common points to be used as datum (or computational base) is needed, followed by the localization of deformations (i.e. transformations of results with respect to the selected datum points). In terms of statistical verification of the estimated results, both global and local tests are needed. Such statistical testing on the estimated deformations are used to establish whether significant movements have or have not occured between the two epochs.

Hence, important requirements for deformation detection can be summarised as the following:

1. For each epoch, the same number of stations, station names and datum definition are required. If different configurations, numbers of stations or datum definition are used in each epoch, the analysis can be applied to common stations and datum only. Therefore, data must be transformed into the common stations and datum prior to analysis.

2. Assuming common stations and datum definition in both epochs, global tests are needed to decide whether to continue the analysis or not, and also to determine whether significant movements have occured between epochs.

3. The stable points (used as datum) can either be known or unknown in advance. If they are known, a global test is needed to verify their stability. Otherwise, a method of identifying and testing the stable points is required. In both cases, it is necessary to transform the results with respect to this new datum.

4. Once the global test is passed, a local test is needed to verify the significance of the estimated movements that have occured between the two epochs. Final results should consist of both numerical and graphic output.

In this study, a procedure for geometrical detection of spatial deformation based on the above needs has been developed (section 4.2). The developed procedure uses one-stage computation, employs two-epoch analysis, absolute monitoring networks, static model and coordinate approach. Most of the statistical tests used for verification of results are based on Fisher's F-distribution, and aspects of the significance level of testing are briefly described in section 4.3. Consequently, the developed procedure is summarized in section 4.4.

4.2 Main modules of analysis

To meet the requirements imposed in section 4.1, the main modules developed for geometrical detection of spatial deformation can be divided into three main stages:

.Transformation into common datum (section 4.2.1)

.Initial check and preliminary testing (section 4.2.2)

.Stability determination and localization of deformation (section 4.2.3)

Each module is discussed in sections 4.2.1, 4.2.2 and 4.2.3 respectively.

4.2.1 Transformation into common datum

In this study, S-transformations and partitioning scheme (section 2.3.4) have been employed in order to transform LSE results of each epoch into a common datum, defined by points common to each epoch. Let the main data (coordinates and cofactor matrix) for epochs one and two be $\hat{x}_1$, $Q_{\hat{x}1}$ and $\hat{x}_2$, $Q_{\hat{x}2}$ (or simply x_1, Q_{x1}, x_2, Q_{x2} from now on). Assume that different numbers of stations and datum definition are used in each epoch. Let the computational bases for epochs one and two be i and j respectively. It is then required to transform the main data into the new datum defined by common stations (n points) for both epochs.

Initially, it is necessary to determine the common stations, and then re-arrange the data so that the common stations are ordered at the beginning of arrays x and Q_x. By using the partitioning scheme, with common stations ordered at the beginning of arrays x and Q_x, equation (2.85) becomes

for epoch one: $x_1^{(i)}$ and $Q_{x1}^{(i)}$ with computational base i

for epoch two: $x_2^{(j)}$ and $Q_{x2}^{(j)}$ with computational base j

$$x_1^{(i)}=\begin{bmatrix} x_{1r} \\ x_{1e} \end{bmatrix}, \quad Q_{x1}^{(i)}=\begin{bmatrix} Q_{1r1r} & Q_{1r1e} \\ Q_{1e1r} & Q_{1e1e} \end{bmatrix} \tag{4.1a}$$

$$x_2^{(j)}=\begin{bmatrix} x_{2r} \\ x_{2e} \end{bmatrix}, \quad Q_{x2}^{(j)}=\begin{bmatrix} Q_{2r2r} & Q_{2r2e} \\ Q_{2e2r} & Q_{2e2e} \end{bmatrix} \tag{4.1b}$$

where r refers to common stations, and e refers to non-common stations.

The transformation of the LSE results of each epoch into the new datum defined by n common stations is similar to equation (2.86)

$$\text{new } x_1^{(n)} = \begin{bmatrix} x_{1r}^{(n)} \\ x_{1e}^{(n)} \end{bmatrix} = S_1 x_1^{(i)} \text{ , new } Q_{x1}^{(n)} = \begin{bmatrix} Q_{1r1r}^{(n)} & Q_{1r1e}^{(n)} \\ Q_{1e1r}^{(n)} & Q_{1e1e}^{(n)} \end{bmatrix} = S_1 Q_{x1}^{(i)} S_1^t \tag{4.2}$$

$$\text{new } x_2^{(n)} = \begin{bmatrix} x_{2r}^{(n)} \\ x_{2e}^{(n)} \end{bmatrix} = S_2 x_2^{(j)} \text{ , new } Q_{x2}^{(n)} = \begin{bmatrix} Q_{2r2r}^{(n)} & Q_{2r2e}^{(n)} \\ Q_{2e2r}^{(n)} & Q_{2e2e}^{(n)} \end{bmatrix} = S_2 Q_{x2}^{(j)} S_2^t \tag{4.3}$$

where $S_1=I-G_1(G_1^tI_nG_1)^{-1}G_1^tI_n$, $S_2=I-G_2(G_2^tI_nG_2)^{-1}G_2^tI_n$

Values of I_n are 1 and 0 for common (or datum) and non-common (non-datum) stations respectively.

The optimised computational procedure outlined in section (2.3.5) is recommended for the evaluation of equations (4.2) and (4.3). After this transformation, the useful results for deformation detection are the coordinates and cofactor matrix of the common stations in each epoch:

epoch one: $x_{1r}^{(n)}$ and $Q_{1r1r}^{(n)}$

epoch two: $x_{2r}^{(n)}$ and $Q_{2r2r}^{(n)}$

The displacement vector d and its cofactor matrix Q_d for the common stations can be simply computed via equation (2.88)

$$d=x_{2r}^{(n)}-x_{1r}^{(n)} \text{ and } Q_d=Q_{1r1r}^{(n)}+Q_{2r2r}^{(n)} \tag{4.4}$$

As shown earlier in section 2.3.3, the transformation process does not change the datum invariant quantities, and hence variance factors and degrees of freedom for each epoch can be used straight away. However, the deformation detection will be based on common stations only.

Equation (4.4) for computing d and Q_d clearly demonstrate the advantage of ordering the common or datum stations at the beginning of arrays x and Q_x. One only needs to extract the upper parts of the arrays, without the need of additional computations, for the purpose of deformation detection.

In this study, LSE is computed using a simple minimum constraints solution with fixed coordinates (section 2.2.2). Depending on whether stable points are known or not in advance, the S-transformations (section 2.3.4 and equation 2.95) can be used to transform the minimum constraints solution into minimum trace, partial minimum trace or other minimum constraints solutions respectively. In general, any of these LSE solutions can be used for detection of deformation.

4.2.2 Initial check and preliminary testing

The data required for detection of deformation are obtained directly from the results of the LSE of each epoch, i.e. the estimated variance factor $\hat{\sigma}_o^2$, degrees of freedom (df or r), datum defect d, estimated coordinates x and their cofactor matrix Q_x (section 3.9). In this work, a procedure has been established for initial checking and preliminary testing of data.

As an initial check prior to deformation detection, it is important to examine that for both epochs, the same (common) stations (number and names) and datum definition (computational base) are being used in LSE. This is very important because of the requirement for common stations and also x and Q_x (hence d and Q_d) are datum dependent. If needed, x and Q_x can be transformed into the common datum prior to analysis as shown in 4.2.1.

Once the initial check is acceptable, and before commencing with the stability determination, it is required to test the compatibility of the independent variance factors of the two epochs. For this purpose, a preliminary test on variance ratios, as given by Biacs (1989) and Caspary (1987b) should be performed. The test can be either one or two-tailed, and the most commonly used is the one-tailed test.

The null hypothesis examines whether the estimated variance factors of each epoch have same expectation, and for one-tailed test can be written as

H_o: $\hat{\sigma}_{oi}^2=\hat{\sigma}_{oj}^2$ at significance level α (4.5)

H_a: $\hat{\sigma}_{oi}^2>\hat{\sigma}_{oj}^2$ or $\hat{\sigma}_{oj}^2>\hat{\sigma}_{oi}^2$

where $\hat{\sigma}_{oi}^2$ and $\hat{\sigma}_{oj}^2$ are the estimated variance factors of epochs i and j. Let their respective degrees of freedom be df_i and df_j.

The test statistic is in the form of a ratio of the variance factors

$$T=\hat{\sigma}_{oj}^2/\hat{\sigma}_{oi}^2 \sim F_{dfj,dfi} \quad (4.6a)$$

assuming j and i referred to the larger and smaller variance factors respectively. Their relevant degrees of freedom become df_j and df_i. The outcome of the one-tailed test on the variance ratio is

if $T < F_{dfj,dfi,\alpha}$, test passes, accept H_o (4.6b)

if $T \geq F_{dfj,dfi,\alpha}$, test fails, reject H_o

For the two-tailed test, the relevant equations are:

H_o: $\hat{\sigma}_{oi}^2=\hat{\sigma}_{oj}^2$ and H_a: $\hat{\sigma}_{oi}^2\neq\hat{\sigma}_{oj}^2$ (4.7a)

$$T=\hat{\sigma}_{oj}^2/\hat{\sigma}_{oi}^2$$

The test passes, and H_o is accepted at significance level α if

$$F_{dfj,dfi,1-\alpha/2} < T < F_{dfj,dfi,\alpha/2} \quad (4.7b)$$

where $F_{dfj,dfi,1-\alpha/2}=1/(F_{dfj,dfi,\alpha/2})$

If T lies outside the region, the test fails, and H_o is rejected.

If H_o is accepted, indicating the two variance factors are statistically equivalent, the variance ratio test passed, and the pooled (or combined or common) variance factor $\hat{\sigma}_o^2$ may be computed as

$$\hat{\sigma}_o^2=[(\hat{\sigma}_{oi}^2)(df_i)+(\hat{\sigma}_{oj}^2)(df_j)]/df \tag{4.8}$$

where $df=df_i+df_j$.

Further analysis (i.e. stability determination and localization of deformation) are developed in sections 4.2.3. On the other hand, failure of the above preliminary test (i.e. rejection of H_o) indicates (Chen et al, 1990a) improper weighting of observations, and requires the examination of observational data and / or LSE results. The analysis should be stopped at this stage.

4.2.3 Stability determination

Following the initial check, the stability determination and localization of deformation are performed iteratively. In this study, a procedure has been established for geometrical detection of spatial deformation. The developed procedure uses congruency testing together with the sucessive removing of the unstable points from the datum, re-ordering and S-transformations of d and Q_d with respect to the re-defined datum points. This procedure is a combination of the Hannover, Karlsruhe and Stuttgart (Fraser and Gruendig, 1985) methods (Appendix E). It is slightly different from Biacs (1989), which is based on the Bonn method.

The procedure also modifies the robust approach of the Fredericton method (developed at the University of New Brunswick Canada) as an alternative for the automatic identification of stable stations or deformation detection.

Moreover, the procedure allows manual selection of the datum stations, and change of significance level α between global and local tests. Another important aspect is that the procedure enables simultaneous analysis of datum stations and detection of spatial deformation, via one-stage computation. Hence, no additional computations are required.

The procedure can be summarized into three stages: congruency tests of common stations (section 4.2.3.1), localization of deformation (section 4.2.3.2) and final local testing of the estimated deformation (section 4.2.3.3). A brief description of the robust method is included in section 4.2.3.4.

4.2.3.1 Congruency test

A statistical test known as the congruency test is required to determine whether significant movements have occured between any two epochs. Its purpose is to determine whether or not a set of the 'tested' points have moved between any two epochs (Caspary, 1987b; Fraser and Gruendig, 1985). The tested points can be either all points common to both epochs (a global congruency test) or selected points used for datum definition (a partial congruency test).

The application of congruency tests is very simple. Initially, the congruency of common datum points at each epoch is tested by the global congruency test. If the test indicates significant movements, localization is performed followed by a similar test on the remaining (partial) datum points through the partial congruency test.

Let the estimated coordinates and cofactor matrices for both epochs be x_1, Q_{x1} and x_2, Q_{x2}. During deformation detection, it is assumed that these data are referred to a common datum, defined by the same datum points in each epoch.

According to Caspary (1987b), the outcome of the global congruency test is independent of the a priori selected datum. Hence, either minimum constraints, minimum trace or partial minimum trace datums may be used. However, in general, the minimum trace datum is recommended as the initial datum, if no information on the stability is available. Otherwise, a partial minimum trace datum is used. Transformation from one datum to another is easily achieved via S-transformations (section 2.3.5).

At the start of deformation detection process, the computation of d and Q_d follows equation (4.4)

$$d=x_2-x_1$$
$$Q_d=Q_{x2}+Q_{x1}$$

The global congruency test (based on Pelzer, 1971) examines the null hypothesis of:

H_o: $E\{x_2\}-E\{x_1\}=E\{d\}=0$ (4.9)

i.e. no significant deformation

or H_o: $E\{x_2\}=E\{x_1\}$

i.e. coordinates are the same

or H_o: $d=x_2-x_1=0$ and H_a: $d\neq 0$

In other words, the null hypothesis states that the common points in both epochs are stable or have not moved.

The test statistic is datum independent (Fraser and Gruendig, 1985)

$$T=\Omega/(h*\hat{\sigma}_o^2) \quad (4.10)$$
$$=(d^tQ_d^+d)/(h*\hat{\sigma}_o^2)\sim F_{h,r}$$

where

$\Omega=d^tQ_d^+d$ is the quadratic form

$d=x_2-x_1$ is the displacement vector

$Q_d=Q_{x2}+Q_{x1}$ is the cofactor matrix of displacement vector d

$h=\text{rank}(\Omega)=\text{rank}(Q_{x1}+Q_{x2})$ is the rank of Q_d

$=3n-d$ for 3-D network of n common datum stations and datum defect d

$\hat{\sigma}_o^2$=pooled variance factor (equation 4.8)

$r=df=r_1+r_2$ is the sum of degrees of freedom in both epochs

Q_d^+ is the pseudo inverse of Q_d

$Q_d^+=(Q_d+GG^t)^{-1}-G(G^tGG^tG)^{-1}G^t$ (Caspary, 1988)

Matrix G^t is given in Figure 2.3. If G is normalized,

$Q_d^+=(Q_d+G_nG_n^t)^{-1}-G_nG_n^t$ as $G_n^tG_n=I$

α is the chosen significance level, typically $\alpha=0.05$

In the deformation detection process, the displacement vectors are usually computed with respect to the first epoch. In computing matrix G, the provisional coordinates of the first epoch are used, and they are referred to the centroid to avoid numerical instability. Further numerical stability may be obtained by normalizing G. Details on reduction to centroid and normalization are given in section 2.3.5. The above means of computing the pseudo inverse is adopted because of its simplicity, involving inversion of a small matrix. Details may be found in Biacs (1989).

The outcome of the congruency test at significance level α is that if T is less than the critical value (i.e. $T<F_{h,r,\alpha}$), the test passes, and H_o is accepted. This means that there is no significant deformation within the group of reference points and analysis can be stopped at this stage. Otherwise, if ($T\geq F_{h,r,\alpha}$), the test fails and H_o is rejected, indicating the existence of significant deformations or movements. It is then necessary to examine the nature of the movements via localization, followed by the partial congruency test.

In the case of very large degrees of freedom, it is possible to replace the pooled variance factor (equation 4.8) with known variance factor equal to unity. The test statistic of equation (4.10) becomes (Biacs and Teskey, 1990)

$$T=\Omega/h \sim F_{h,\infty} \tag{4.11a}$$

and H_o is accepted if $T < F_{h,\infty,\alpha}$

The partial congruency test examines the stability of the partial network formed by the selected or retained datum points only. This is applicable because the set of the retained datum points is actually part of the initial computational base. Let the vector of deformation d and the cofactor matrix Q_d be partitioned as (Fraser and Gruendig, 1985)

$$d=\begin{bmatrix} d_r \\ d_e \end{bmatrix}, \; Q_d=\begin{bmatrix} Q_r & Q_{re} \\ Q_{er} & Q_e \end{bmatrix} \tag{4.11b}$$

where r refers to (retained) datum points, and e refers to non-datum points (i.e. datum points eliminated from the computational base).

The null hypothesis for the partial congruency test becomes

$$H_o\text{: } E\{d_r\}=0 \tag{4.12}$$

i.e. the partial network has not changed in shape

The test statistic

$$T=(d_r^t Q_r^+ d_r)/((h-3k)\hat{\sigma}_o^2) \sim F_{h-3k,r} \tag{4.13}$$

where k is the number of eliminated points in sub-vector d_e. Interpretation of the test is similar to the global congruency test. If T is less than $F_{h-3k,r,\alpha}$, the test passes, H_0 is accepted and the analysis may be stopped. Otherwise, the test fails, H_0 rejected and further localization is required.

4.2.3.2 Localization of deformation

If the global congruency test fails, and hence indicates occurrence of significant deformation, it is required to locate and isolate any suspect unstable points, and at the same time to re-define a new datum for the computations (with respect to datum points). Several methods of localization are available (Caspary, 1987b; Chrzanowski and Chen, 1986; Appendix E). In this study, the localization procedure has been modified from Fraser and Gruendig (1985).

Starting with a chosen computational base (usually based on a set of known reference datum points), the procedure removes one point at a time from the computational base. Points are removed via the successive application of decomposition (of quadratic form), re-ordering and partitioning (with respect to the datum points), S-transformations (of d and Q_d), and partial congruency test until the congruency test passes.

The computational base can be either chosen manually (based on some priori information, for example) or computed by means of the congruency testing or the robust method. The procedure for localization via decomposition, re-ordering and S-transformations is discussed in sections 4.2.3.2.1, 4.2.3.2.2 and 4.2.3.2.3 respectively.

4.2.3.2.1. Decomposition of quadratic form

If the global congruency test (equation 4.10) fails (i.e. rejection of H_0), the required information on non-congruency between the two epochs is contained in the quadratic form Ω.

In this case, the main task is to investigate the individual contributions of each point (Ω_j) to the total quadratic form Ω. The point with the largest (maximum) Ω_j is usually considered as the most significantly deformed.

Computation of Ω_j is carried out using a decomposition or splitting method. In this study, the decomposition procedure is based on Fraser and Gruendig (1985) using techniques of partitioning adopted by Niemeier (1979) and Van Mierlo (1981). The vectors d and Q_d^+ are partitioned for each point as

$$d=\begin{bmatrix} d_r \\ d_j \end{bmatrix}, \quad Q_d^+=\begin{bmatrix} P_r & P_{rj} \\ P_{jr} & P_j \end{bmatrix} \tag{4.14}$$

$$\Omega=\Omega_r+\Omega_j$$

$$d_j=[dx_j\ dy_j\ dz_j]^t$$

$$d_{jj}=P_j^{-1}P_{rj}d_r+d_j$$

$\Omega_j=d_{jj}{}^tP_jd_{jj}$ for each point j

The above computation is repeated for each point giving rise to Ω_j. This computational scheme removes the effect of other points in the computed Ω_j, i.e. the effect of other points is excluded.

Following this decomposition procedure, the point with the largest Ω_j (considered as the most significantly deformed) is interactively removed from the computational base.

As an alternative to the decomposition procedure, a single point test can also be performed for each point as an aid for identifying the most suspect point, ignoring the correlation amongst points (Biacs, 1989). It is expected that this test is slightly less sensitive than using decomposition (equation 4.14). The single point test is based on

$$H_o\text{: } d_j=[dx_j\ dy_j\ dz_j]^t=0 \tag{4.15}$$

The test statistic for the 3-D case is

$$T_j=\Omega_j/(3\hat{\sigma}_o^2)=(d_j{}^tP_jd_j)/(3\hat{\sigma}_o^2)\sim F_{3,df} \tag{4.16}$$

If $T<F_{3,df,\alpha}$, the point has not significantly moved. Otherwise, it is considered as being unstable. For practical purpose, it is recommended that, only the point with the largest test statistic T_j is considered as significantly moved and hence removed from the computational base.

In order to remove the suspected point j from the computational base, it is required to re-order some of the relevant data followed by S-transformations (sections 4.2.3.2.2 and

4.2.3.2.3).

4.2.3.2.2. Re-ordering with respect to datum points

Re-ordering with respect to datum points is necessary everytime a point is removed from the computational base. Re-ordering is also required initially for easy extraction of common and / or reference stations, and for easy manipulation of pre-defined datum points.

The relevant data to be ordered are displacement d, cofactor matrix Q_d, matrix G (related to datum defect) and diagonal matrix I_p. In this study, the re-ordering strategy has been established as equation (4.17), using the same notation as equation (4.11b)

$$d=\begin{bmatrix} d_r \\ d_e \end{bmatrix} ; \; I_p=\begin{bmatrix} I_r \\ I_e \end{bmatrix} ; \; G=\begin{bmatrix} G_r \\ G_e \end{bmatrix} ; \; Q_d=\begin{bmatrix} Q_r & Q_{re} \\ Q_{er} & Q_e \end{bmatrix} \qquad (4.17)$$

In the above equation, d_e refers to the suspected point with significant deformation, i.e.

$d_e=d_j=[dx_j \; dy_j \; dz_j]^t$

Q_e is a (3x3) symmetric cofactor matrix for d_e. Elements d_r and Q_r refer to the partial or remaining datum points. Matrix I_p and G are re-ordered to facilitate the use of general equations for the S-transformations (section 2.3.5) and partial congruency test (section 4.2.3.1) respectively. Elements of I_e corresponding to the non-datum points are set to zero, while elements of I_r (for datum points) remain one. Matrix G is computed once only.

4.2.3.2.3. S-transformations

In this study, the general S-transformations equation (section 2.3.5) has been applied for transforming d and Q_d into the new computational base defined by the remaining datum points. The S-transformations scheme of d and Q_d into d' and Q_d' (equation 2.85) becomes

$$d=\begin{bmatrix} d_r \\ d_e \end{bmatrix} , \; Q_d=\begin{bmatrix} Q_r & Q_{re} \\ Q_{er} & Q_e \end{bmatrix} \qquad (4.18a)$$

$$d' = \begin{bmatrix} d_r' \\ d_e' \end{bmatrix} = Sd \ , \ Q_d' = SQ_dS^t = \begin{bmatrix} Q_r' & Q_{re}' \\ Q_{er}' & Q_e' \end{bmatrix} \qquad (4.18b)$$

Fraser and Gruendig (1985) use the following strategy

$$G=[G_r\ G_e]^t \qquad (4.19)$$

$$S=I-\begin{bmatrix} G_r(G_r^tG_r)^{-1}G_r^t & 0 \\ G_e(G_r^tG_r)^{-1}G_r^t & I \end{bmatrix}$$

In this research, use of the general S-transformations equation (section 2.3.5) has been established as

$$S=[I-G(G^tI_dG)^{-1}G^tI_d] \qquad (4.20)$$

$$d'=Sd;\ Q'=SQ_dS^t$$

Elements of I_d are one (unity) and zero for datum and non-datum points respectively. In the above equations, the elements G, I_d, d and Q_d have been ordered (section 4.2.3.2.2). The computational strategy is similar to the evaluation of the S-transformations (section 2.3.5). Vectors d_r' and d_e' refer to the datum and non-datum points respectively.

After this transformation, the remaining network formed by the retained points (d_r') must be tested for stability by means of the partial congruency test described in section 4.2.3.1. The test statistic given by equation (4.13) can be written as

$$T=(d_r'^tQ_r'^+d_r')/((h-3k)\hat{\sigma}_o^2)\sim F_{h-3k,df}$$

where k is the number of points removed from the computational base. If the test fails, the process of decomposition of the quadratic form Ω (section 4.2.3.2.1), re-ordering (section 4.2.3.2.2) and S-transformations (section 4.2.3.2.3) are repeated until the partial congruency test passess.

At the end of the localization stage, d_e' (equation 4.18b) represents the vectors of deformation of the non-datum points with respect to the datum defined by d_r'. In other words, the solution is in the form of the partial minimum trace datum. The same applies to their cofactor matrix. In order to confirm the localization finding, final testing of deformation may be performed, as discussed in section 4.2.3.3.

As a final confirmation of the localization procedure, combined LSE of the data from both epochs can be performed, using the stable points (d_r') as datum. The points in epoch two suspected as significantly moved (i.e. d_e') are assigned different numbers in each epoch. Alternatively, S-transformations similar to section 4.2.1 can be used for this purpose. The vectors of deformation of the non-datum points may be computed directly from their differences (equation 4.4) in coordinates.

Fraser and Gruendig (1985) show that the difference betweeen the vector of deformations from combined LSE and the final significant deformation vectors obtained by localization (d_e' in equation 4.18b) will be insignificant. Hence, results of localization can be used directly for demonstrating the deformation trends.

4.2.3.3 Final testing of deformation

Having determined the significant vectors of deformation by means of the localization procedure, the final testing of deformation for verification or confirmation purpose is in the form of a local test known as single point test. A graphical plot to represent the deformation vectors against their point confidence ellipsoid is also useful.

The single point test (equation 4.15) is based on the null hypothesis (Cooper, 1987)

$$H_o: d_j=[dx_j\ dy_j\ dz_j]^t=0 \tag{4.21a}$$
for each point j

The test statistic for point j is based on the multi-dimensional F-test

$$T_j=(d_j^t Q_{dj}^{-1} d_j)/(m\hat{\sigma}_o^2) \sim F_{m,df} \tag{4.21b}$$

where m is the dimension of the network. The test statistic for a 3-D network (m=3) is

$$T_j=\Omega_j/(3\hat{\sigma}_o^2)=(d_j^t Q_{dj}^{-1} d_j)/(3\hat{\sigma}_o^2)\sim F_{3,df} \tag{4.22a}$$

where Q_{dj} is the cofactor matrix of the displacement vector d_j. This local test is performed for each individual point. If T_j is less than $F_{3,df,\alpha}$, point j is considered as stable, i.e. the displacement vector is not significant. Otherwise it is considered as moved or unstable. In the above test, any correlation between points is neglected. A more general approach of computing Ω_j is by applying the concept of decomposition of quadratic form (section 4.2.3.2.1). The single point test can also be computed using a t-test. In this study, the developed procedure permits change of significance level α (section 4.3) between congruency (section 4.2.3.1) and single point tests (equation 4.22a).

It is expected that all datum points will be stable (i.e. the test passes), while non-datum points can either be stable or unstable. Hence, the points with significant movements are expected to be unstable.

The final displacement of each point can be shown graphically, by comparing the displacement vector of each point against its confidence region at a specific significance level. In a 3-D case, such graphical representation is not straight forward. The first and simplest method is by splitting the information into a plot of horizontal and vertical deformations.

The second method, adopted in this study is by considering the confidence region in all three axes, i.e. xy, xz and yz axes (section 5.2). The displacement vector of any point that lies outside the corresponding confidence region (i.e. error ellipse in horizontal and confidence interval in vertical in the first method; and error ellipses in the xy, xz and yz axes for the second method) indicates significant movements. For stable datum points, the plots of displacement vectors will be within the confidence region.

Both the local single point test and plot of deformation vectors (and error ellipses) give similar interpretation. The plot is very useful as it gives an overall picture of any trends in the estimated deformation, both in direction and magnitude. Examination of the plot will also indicate if there are any group movements.

Another useful test is a one-dimensional single point test for testing the significance of the magnitude of a spatial deformation vector d_s in the specified direction (Heck, 1984; Biacs and Teskey, 1990)

$$T=d_s/\sigma_{ds}\sim N(0,1) \tag{4.22b}$$

The movement is insignificant if

$|T| < N_\alpha$ if use variance factor of unity

$|T| < t_{r,\alpha}$ if use estimated variance factor

4.2.3.4 Robust method in the detection of deformation

The congruency testing method described earlier (sections 4.2.3.1 and 4.2.3.2) identifies the stable datum reference stations and localizes the deformation. The method will iteratively remove one station (i.e. suspected as unstable) at a time from the initial chosen computational base until the partial congruency test passes. Alternatively, a type of robust method known as an iterative weighted similarity transformation may be used.

This robust method has been developed at the University Of New Brunswick, Canada (Chen et al, 1990; Chen, 1983; Chrzanowski and Chen, 1986; Chrzanowski et al, 1986). In this method, the strategy is to minimize the first norm of the displacement vectors of the reference points via a weighted transformation. Stations with less movement should have most influence in the definition of a datum (more weights), or vice versa. The method can produce a datum that is robust to the unstable reference points and gives less distorted displacements.

Let d and Q_d be the displacement vector and its cofactor matrix for the common points in both epochs. By partitioning or ordering with respect to the reference datum points (equation 4.18a), elements d_r and Q_r for reference points can be easily extracted.

$$d=\begin{bmatrix} d_r \\ d_e \end{bmatrix},\ Q_d=\begin{bmatrix} Q_r & Q_{re} \\ Q_{er} & Q_e \end{bmatrix} \tag{4.23}$$

where r and e refer to the reference (datum) and object (non- datum) points, respectively. The transformation of d_r and Q_r into another datum can be written as

$$d_r'=S_rd_r=[I-G_r(G_r^tW_rG_r)^{-1}G_r^tW_r]d_r \tag{4.24a}$$

$$Q_r'=S_rQ_rS_r^t \tag{4.24b}$$

The weight matrix W_r in the above equation is selected so that the first norm of the displacement vector d_r' approaches a minimum, hence

$$|d_r'|_1=\min \tag{4.25}$$

This equation indicates that the sum of magnitudes of all displacement components is a minimum. Let the transformation parameters (t) be

$$t=(G_r^tW_rG_r)^{-1}G_r^tW_rd_r \tag{4.26}$$

$$|d_r'|_1=\Sigma|d_r-G_rt|=\min$$

The weight matrix W_r for iterative weighted similarity transformation is taken as identity in the first iteration, and in the (k+1) iteration becomes

$$W_r^{(k+1)}=\text{diagonal}\ \{1/|d_r^{(k)}{}_{(i)}|\} \tag{4.27}$$

where $d_r^{(k)}{}_{(i)}$ is i^{th} component of vector d_r' after the k^{th} iteration. The iterative transformation, utilising equations (4.24a) and (4.27), continues until the absolute difference between successive transformed displacement components is smaller than the tolerance δ (for example, half of the average accuracy of displacement components, such as 0.0001 m)

$$|d_r^{(k+1)}-d_r^{(k)}|<\delta \tag{4.28}$$

To avoid numerical instability in computing W_r (when $d_r^{(k)}{}_{(i)}$ approaching zero), two approaches are possible, either by expanding the expression for W_r or setting a lower bound. In the first approach (Chen et al, 1990a), the expression for W_r can be written as

$$W_r^{(k+1)}=\text{diagonal}\{1/(|d_r^{(k)}{}_{(i)}|+\delta)\} \tag{4.29}$$

With the second approach, if $|d_r^{(k)}{}_{(i)}|$ is smaller than the lower bound (say δ), its weight is set to zero. Hence, approximate solutions are obtained.

In the final (k+1) iteration, the cofactor matrix is computed as

$$Q_r' = S_r^{(k+1)} Q_r [S_r^{(k+1)}]^t \tag{4.30}$$

Identification of unstable reference points at a specified significance level α is by means of a single point test or by comparing the displacement of each point against its confidence region. The interpretation is similar to the testing of deformation outlined in section 4.2.3.3.

Examination of the equations involved indicated that the iterative weighted similarity transformation is equivalent to a weighted S-transformations, and hence the general S-transformations equations (section 2.3.5) are equally applicable. The weight matrix W_r (or I_r in the general S-transformations) is initially taken as identity in order to transform the solutions with respect to all the points.

Compared to congruency testing, this robust method is quicker if there are more unstable reference points, as the main computation modules (equation 4.24a) involve only the displacement vectors, and the cofactor matrix (equation 4.30) is only computed once during the final stage. Also, there is no need for the ordering of relevant elements necessary for the congruency method. However, this method only can be used to analyse reference stations only, and as indicated by Chen et al (1990a), further computations are needed for localization of deformation. Also, the number of iterations may be large. On the other hand, the congruency testing method can be used to analyse both reference and object stations simultaneously in one-stage computation, and the results can be used directly. In this study, the robust method has been modified to allow for one-stage computation. In the final S-transformations, the stable datum points are given weight of unity, while weight for other points are assigned as zero. The computational procedure for S-transformations developed in section 2.3.5 has been applied into equation (4.24).

In terms of application, the robust method may be used as an alternative method to determine a set of stable reference points initially, prior to congruency testing.

4.3 Significance level in testing

During statistical testing of LSE results, the selection of a significance level (α) is to some extent arbitrary. Associated with the test are the Type I and Type II errors. A Type I error is simply rejecting a good observation with probability of α, whilst Type II error is in fact accepting a bad observation at probability level β (section 3.3). In deformation detection (Biacs, 1989), a Type I error occurs when significant movements are detected but did not occur, i.e. false alarm. On the other hand, a Type II error exists when existing movements are not detected, i.e. missed detection. In practice, selection of of α (and also β) to be used for global (i.e. congruency) and local (single point) tests (sections 4.2.3.1 and 4.2.3.3) is quite important. Generally, smaller significance levels are required for local (single point) test.

Standardization of significance levels is discussed in section 3.3.4. To date, only the B-method (developed at Delft Technological University, Netherlands) synchronized the significance levels via α and β. However, this method has been shown to lead to a high probability of false alarm (Biacs, 1989). In most applications, only Type I errors are considered, and standardization of α is carried out via Bonferroni's inequality.

During the detection of deformation, an arbitrary value of α for the global test may be selected. For a local single point test, neglecting correlation between stations, standardization of α via the application of Bonferroni's inequality (Vanicek and Krakiwsky, 1986; equation 3.20) gives

$$\alpha_l=1-(1-\alpha_g)^{1/m}\approx\alpha/m \tag{4.31}$$

where m is the number of stations, α_g and α_l are significance levels for global and local tests, respectively.

However, as the number of stations is increased, α_l becomes too small, and consequently the critical values become too large, leading to missed detection. A more practical expression (Cooper, 1994) is simply

$$\alpha_l=\alpha_g/(m^{1/2}) \tag{4.32}$$

In monitoring works, α is usually chosen as 0.05 and 0.01 for global and local test respectively. The procedure developed for detection of deformation (section 4.4) allows user to select any α for both global and local tests, with standardization via equation (4.32).

4.4 Procedure for deformation detection

From a practical point of view, it is necessary to divide the monitoring network into reference (or datum) and object (non-datum) points. Initially, analysis needs to be carried out on the reference points only (Teskey, 1994), and not on the whole network. Usually, with a properly designed network, survey control stations can be used as an initial datum stations.

In addition, the stability of the initial datum stations can be based on the geotechnical, geological or engineering knowledge. In the case of no previous information, it is also possible to use the points with the smallest local statistics as initial datum stations, and applying the Bonn method (Appendix E).

With this in mind, the application of the 3-D deformation detection modules developed in section 4.2 can be summarized into several stages as shown in Figure 4.1:

i. Preliminary checks and if necessary, transformation into a common datum (section 4.2.1). At this early stage, information related to the common reference and datum stations need to be extracted. By rearranging the data of common stations at the beginning of arrays, the extraction process is straightforward.

ii. Initial check and preliminary testing to determine whether the analysis is to be continued (section 4.2.2). A one-tailed test on the variance ratio (equations 4.5, 4.6 and 4.8) is used.

iii. One-stage computation for stability determination of the datum or reference points, and localization of deformation. The initial datum points can be considered as either unknown or known. In the case of unknown datum points, a minimum trace datum is adopted. Otherwise, a partial minimum trace datum is employed.

Starting with a chosen computational base (datum), and after appropriate ordering, the datum points are analysed via the methods of congruency testing and localization until the congruency test passes and a set of stable datum points are found. Sequences of congruency testing (section 4.2.3.1), decomposition (section 4.2.3.2.1), re-ordering (section 4.2.3.2.2) and S-transformations (section 4.2.3.2.3) are performed iteratively in the analysis. This is then followed by final testing of deformation via the single point test (section 4.2.3.3). During testing, significance level α can be changed between congruency and single point tests. Standardization of α can be determined computationally (equation 4.32) or manually.

All the datum stations must pass both the congruency (global) and single point (local) tests (i.e. stable). Otherwise, only the stable datum stations must be used to define the datum, and the detection procedure are repeated.

Alternatively, the robust method (section 4.2.3.4) may be used for the determination of initial stable datum points or deformation detection. It is expected that both congruency and robust methods will give similar results on the stability of datum points. However, the congruency testing method is more flexible with respect to datum definition and therefore more suitable in practice.

The above approach is purely geometric and produces the vectors of deformations, showing the movement trends. Further analysis may be performed to check any group movements. Theoretically, the method requires that three times the number of datum stations be equal to or greater than the number of datum defect, to avoid the singularity. In the extreme case of photogrammetric data with a maximum of seven datum defects, a minimum of three (hence nine coordinates) stable datum points are needed. If no stable points can be found, other methods such as strain analysis (Brunner, 1979; Cooper, 1987) may be used.

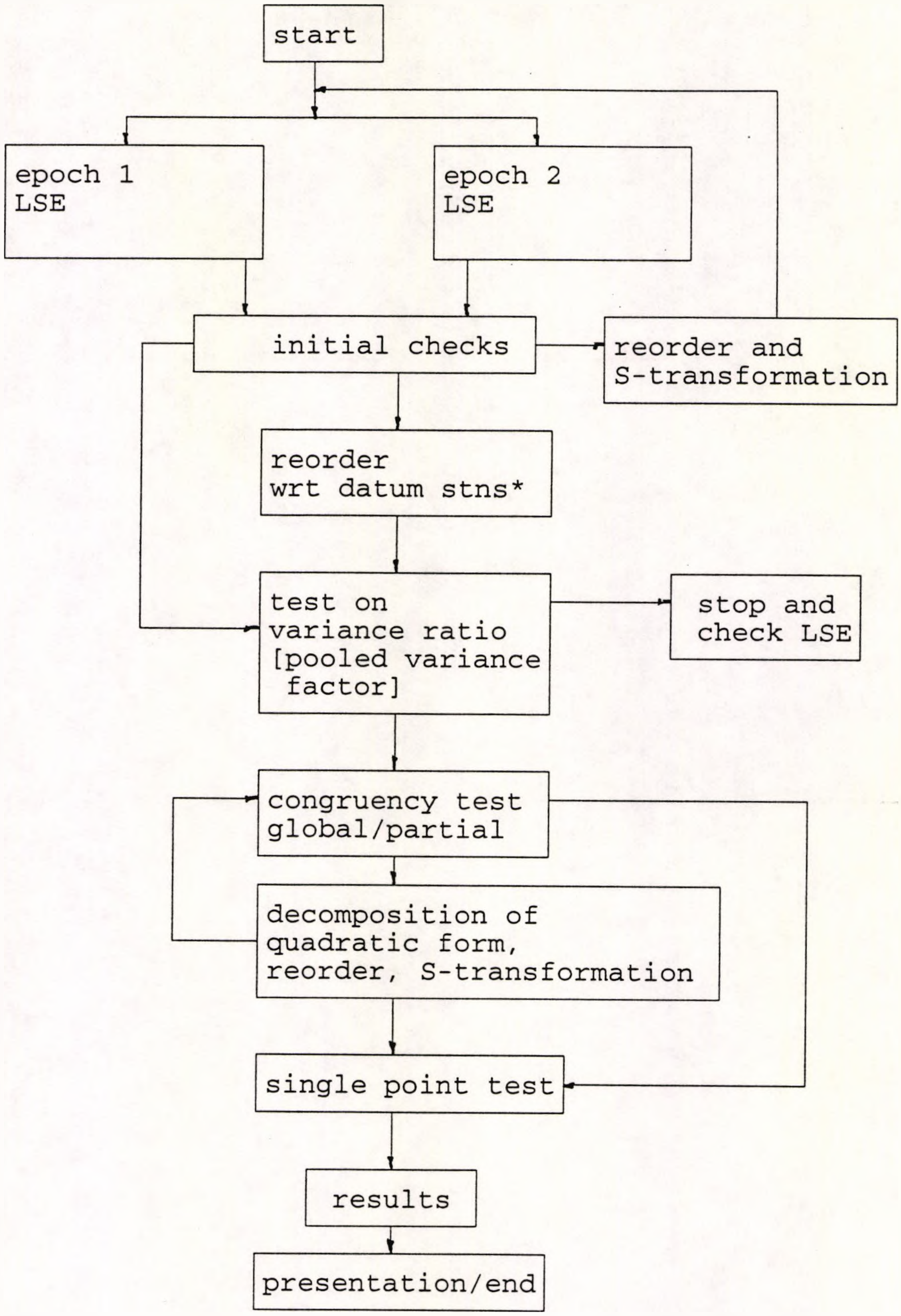

Figure 4.1 Procedure for geometrical detection of spatial deformation

(*datum stations can be determined manually or computationally)

5. IMPLEMENTATION

This chapter describes the implementation of the concepts and methodologies discussed in sections 2, 3 and 4 into five computer programs for 3-D application. The description of program modules developed for LSE, S-transformations, determination of common stations and detection of spatial deformation (two programs) between two epochs are given. The links between these programs and two of ESRC's programs (GAP and DCRE) for deformation detection have been established.

5.1 Description of programs

The procedure adopted for deformation detection between two epochs is summarized in Figure 4.1 (section 4.4). The main tasks are independent LSE at each epoch, S-transformations of LSE results into a common datum, determination of common stations between epochs, and the detection of spatial deformation.

In this research, the tasks have been implemented into five computer programs for 3-D application as the following:

a. ESTIMATE for 3-D LSE of a single epoch (section 5.1.1).
b. COMPS for 3-D S-transformations of LSE results into the selected datum prior to deformation detection (section 5.1.2).
c. COMON for 3-D determination of common stations between two epochs, with the necessary re-ordering and S-transformations with respect to the common stations (section 5.1.3).
d. DETECT for geometrical detection of spatial deformation between two epochs based on congruency testing (section 5.1.4).
e. ROBUST for geometrical detection of spatial deformation between two epochs based on a robust method (section 5.1.5).

All the computer programs were written in FORTRAN77, and were developed for applications on both personal computer (PC) and UNIX environments. FORTRAN77 (Dyck et al, 1984) was used mainly due to its powerful computing ability, and compatibility between PC and UNIX compilers. Another reason is familiarity of the author with FORTRAN77.

In general, all the programs are interactive in nature to allow users some flexibility in decision making, such as computation modes, continuation of computations, selection of significance level for statistical testing, creation of output files and termination of the programs. Whenever necessary, the programs automatically compute approximations of critical values for statistical tests, using the formulations of Cooke et al (1990). For program portability between PC and UNIX environments, all routines are developed independently, and no external routines are used.

In terms of storage, single arrays are used in most cases, using the strategy outlined by Healy (1986). Some of the dummy arrays are used repeatedly. For cofactor matrix, only upper triangles are stored in single arrays. Error checking and trapping methods are included, on both input errors and singularity. The program will be terminated automatically if either data input files are not available, on the existance of input errors or singularity.

Another important task in deformation detection (Figure 4.1) is the presentation of results. In general, results can be presented numerically and / or graphically. Although the above programs provide the results numerically, a more helpful presentation of the results is in the form of graphics plots.

For this purpose, a special graphics program developed by Chandler (1994) at the Engineering Surveying Research Centre (ESRC) called DCRE is used. This program runs under the INTERGRAPH MicroStation environment, and has a flexible on-screen graphics capability of showing plots of results of LSE and deformation detection. For example, plots of points, station names, error ellipses and deformation vectors can be produced. Error ellipses can be also portrayed in three axes, xy, xz and yz.

In order to use DCRE, the outputs of all five implemented programs are produced so that they are compatible with input for DCRE. Each program can produce an additional specialised output file as input for DCRE if requested. Three types of input format for DCRE are used for plots of the networks, plots of stations with ellipses, and plots of deformation vectors with ellipses.

Another useful program available at ESRC is GAP for LSE of single epoch (section 5.1.1). Link between GAP, DCRE and all the five implemented programs, as well as their usage have been established and described in section 5.2.

5.1.1 Program ESTIMATE [LSE]

ESTIMATE is a program designed for the LSE of local monitoring 3-D networks using terrestrial observations. The basic structure is shown in Figure 5.1 (based on Figure 3.2).

The program can currently handle thirteen types of (assumed) uncorrelated observations (Appendix A and section 3.5.3.1): horizontal and slope (or spatial) distances, height differences, horizontal directions (with orientation parameters), horizontal angles, zenith angles (or zenith distances), azimuths, vertical angles and specialised slope distances (with scale, zero or combination errors, and multiple scale or zero errors). The program is also capable of processing three types of algebraically correlated observations: horizontal angles and pseudo observables (spatial distance ratios (equation 3.44) and distance differences) (section 3.5.3.2).

Two data input files are required for ESTIMATE, one contains the provisional coordinates and the other contains the observations and their variances. As engineering monitoring surveys are confined to a specific area requiring a special network, a locally defined 3-D cartesian coordinates system is employed. Datum definition (via minimum constraints) is also included in the provisional coordinate file. The observations are assumed to have been corrected and / or reduced for systematic errors where appropriate.

The adopted procedure for LSE of parameters is based on equations (2.13) to (2.26). Handling of the minimum constraints datum is carried out by holding fixed a minimum number of coordinates (equal to the datum defect d). These points are then removed from the system of equations to form reduced non-singular normal equations (equations 2.44 and 2.45a).

The program starts with the selection of computational modes, reading and checking of input files, followed by checks on the datum deficiency of the network. The datum deficiency is determined automatically from the types of observations, as indicated in Table (2.1). For a minimum constraints solution (adopted for deformation detection), the datum defect must be equal to the number of fixed coordinates. An overconstrained solution is obtained if the number of fixed coordinates is greater than the datum defect. If the number of fixed coordinates is less than the datum defect, the program is terminated automatically.

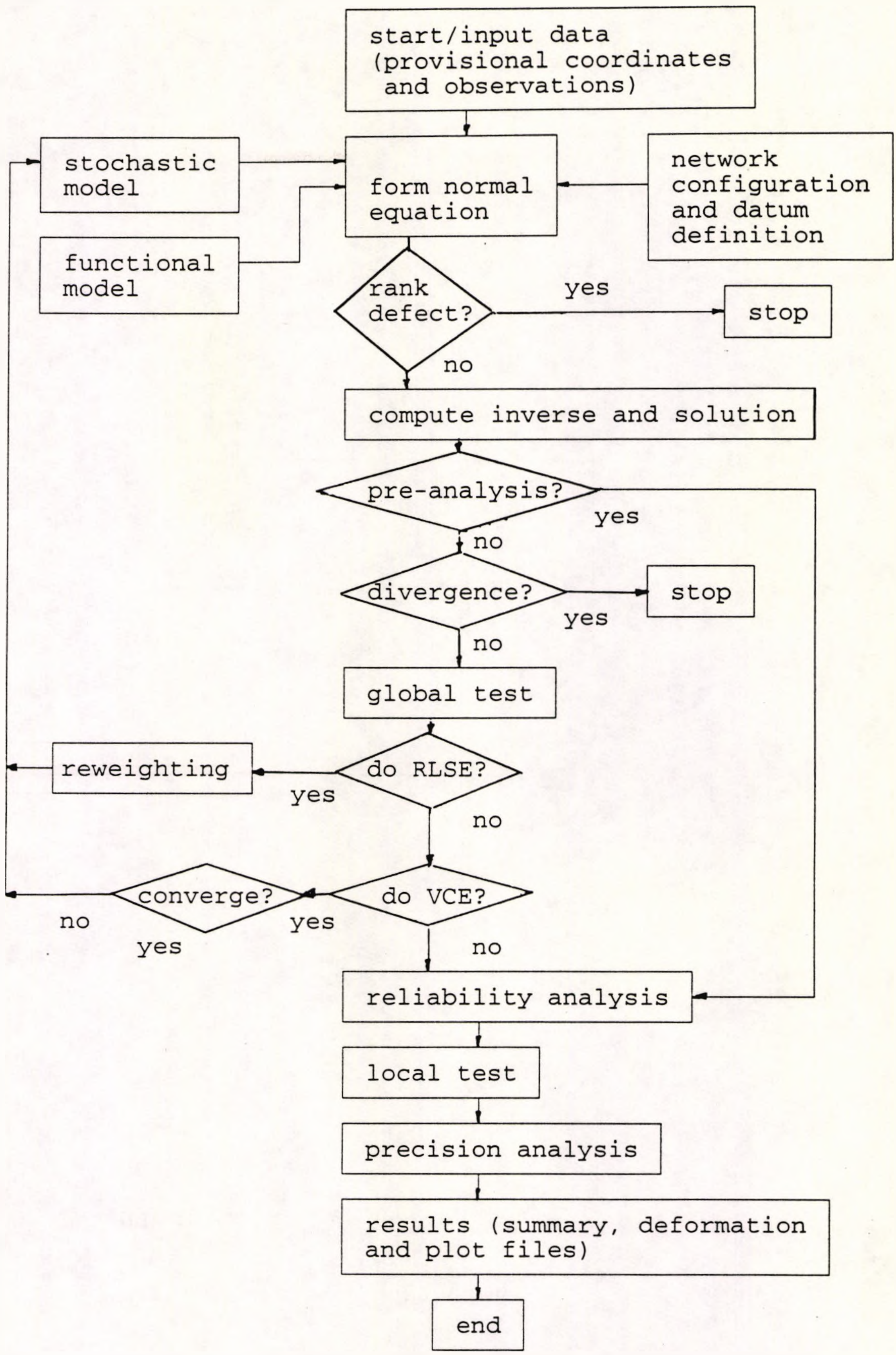

Figure 5.1 Flowchart for program ESTIMATE

In ESTIMATE, the observations are not stored. During the formation of the normal equations, an initial check on gross errors is incorporated by means of misclosure vectors (equation 2.14). Any misclosure vector bigger than a pre-defined limit will be flagged for checking and the program will terminate. Consequently, it is required to check the observational data. The limit must not be too small to allow for inaccuracy in provisional coordinates, for example 10 m for linear and 5° for angular observations. In practice, this check is very useful for detecting any very large gross error prior to further analysis.

For deformation detection purposes, the inverse of the normal equations is required, and hence an ordinary inversion routine is adopted. During LSE, prior to inversion, a rank defect analysis on normal equations via simplified EVD is carried out, utilising equations (2.51), (2.52) and (2.58). The program allows for the manual entry of limits for checking the rank and condition. Depending on the effective rank, non-singular normal equations allow further computation, whilst any singular normal equations will cause automatic termination of the program. In the singular case, it is necessary to check the network configuration and datum definition. In certain cases, although the normal equation is non-singular, large condition number will indicate that it is ill-conditioned.

The LSE process computes the solution, updates the estimated coordinates, and computes the variance factor. The solution is usually iterated due to the linearisation and approximation process. During iteration, a check for convergence and divergence of the variance factor is also carried out. Divergence will automatically stop the program.

Relevant global and local statistical tests (section 3.3) are incorporated in the program. In most cases, users can define the significance level for testing, while critical values are computed automatically. It is also possible to enter the critical value manually.

The global (chi-square) test is computed once the solution has converged. Such a test can be either a one or two-tailed test, following equations (3.3) to (3.6). In practice, a one-tailed test of equation (3.6) is commonly used. Following the outcome of the global test, redundancy numbers (equation 3.18) necessary for reliability analysis are computed, and the user has the option for executing an outlier detection module and creating output files.

The outlier detection module is based on robustified LSE using equation (3.59a) in the ordinary case and equation (3.59b) to speed up the computations. During the process of iterative robustified LSE, a weighting factor is interactively determined by the user. In relation to de-weighting, it is also possible to compute the effect of changing observation weights on the estimated parameters (section 3.6.2.3).

All the adopted statistical tests are based on the normality assumption (section 2.1.5). This assumption is tested by means of a chi-squared goodness of fit test, given by equations (3.9) and (3.10), using all observations. The local test is related to gross error detection, and the adopted test is based on Pope's Tau method (equations 3.19 and 3.21), with automatic computation of the critical values. If an additional parameters approach is adopted for handling systematic errors, a significance test on each additional parameters follows (equations 3.14 and 3.15a).

All the foregoing discussion assumes uncorrelated and independent observations. Horizontal angles are algebraically correlated, and need to be treated differently. In handling systematic errors, introduction of pseudo observables (distance differences and ratios) also creates algebraically correlated observations (sections 3.5.3.2.1 and 3.5.3.2.2). In order to use this ordinary LSE scheme, the concept of observation de-correlation (section 3.5.3.2.3) is applied and implemented in ESTIMATE.

Precision and reliability analyses follow section 3.8. For reliability analysis, values of λ_o (Figure 3.1) need to be entered manually, which are available in Baarda (1968). For internal reliability, both redundancy number and MDGE (equations 3.25a and 3.25b) are computed, whilst the infuential factors (equation 3.26) are computed as a measure of external reliability. Precision measures are based on trace, mean variance and the variances of individual parameters (section 3.8).

In relation to the precision measures, the parameters defining the error ellipsoid can be computed via EVD using equation (2.64). Such an ellipsoid allows graphics representation of the confidence region of the estimated parameters (section 3.3.5). However, in this case the measure is not computed, as graphics are handled separately by program DCRE. A compatible format suitable for using DCRE is produced, where the sub-cofactor matrix for each station is extracted.

The computational mode in effect so far is the ordinary or robustified LSE. Two other computational modes are pre-analysis and VCE. The program is designed so that all three computational modes can be executed, without any modification to the original input data files.

In pre-analysis, equations (2.20) and (2.22) are used to determine the expected precision and reliability of the solution based on the provisional coordinates and precision of the observations. The actual observation vectors are not used in this mode.

Assuming observations as uncorrelated and free from both systematic and gross errors, the computational mode for simplified VCE (section 3.7.3) is executed iteratively for estimating variances of observations in groups. Two-tailed global and local tests (equations 3.92 and 3.93) are used for stopping criteria.

Program ESTIMATE will produce up to three output files, depending on the user's selection. The output consists of summary file (compulsory) containing LSE results, a file for detection of deformation or S-transformations (i.e. deformation file), and a plot file for DCRE. The adopted LSE solution (equation 2.45a) is based on reduced normal equations. In order to produce output files for deformation detection and DCRE, the reduced cofactor matrix Q_x obtained from LSE is expanded to its full size (section 2.2.2).

The summary file contains information on the LSE such as estimated coordinates, global and local tests, reliability and precision analysis. Data in the deformation file (section 3.9) consist of the estimated variance factor, degrees of freedom, the datum defect, provisional 3-D coordinates, estimated 3-D coordinates and their full (i.e upper triangular) cofactor matrix. The plot file for DCRE contains estimated coordinates and sub-cofactor matrix for each station.

ESTIMATE only deals with surveying data. At ESRC, most of the monitoring activities use combinations of photogrammetric and ordinary surveying data, and LSE is carried out by a program called GAP (General Adjustment Program), developed by Clark (1992). GAP is capable of processing the combination of photogrammetric and uncorrelated surveying data (horizontal and slope distances, height differences, horizontal and vertical angles). Program GAP also produces summary, deformation and plot files. During development of ESTIMATE, GAP was used to check and verify some computations.

5.1.2 Program COMPS [S-transformations]

COMPS is a program for the S-transformations of the results of single epoch LSE (i.e coordinates and cofactor matrices) from one datum to another, prior to deformation detection. This facility is useful as the solution obtained from ESTIMATE is with respect to a minimum constraints datum only. With the aid of COMPS, it is possible to transform the minimum constraints solution into a minimum trace, partial minimum trace or other minimum constraints solution, depending on the selected datum.

At the beginning of the deformation detection process, it is required that coordinates and cofactor matrices of common stations for each epoch are based on the same datum. If these requirement is not fulfilled, the LSE results at each epoch must be transformed into the same datum via program COMPS.

The fundamental of program COMPS is demonstrated in Figure 5.2, and the adopted computational procedure is described in section 2.3.5. The required data input file for COMPS (i.e. deformation file) is generated by ESTIMATE (or GAP) for each epoch (section 5.1.1). The same data format is applicable for COMON, DETECT and ROBUST.

Datum definition for S-transformation is included in the estimated coordinates simply by assigning codes with values of one and zero for datum and non-datum points respectively. It may be necessary to edit the input (deformation) file to define the required datum. For convenience, COMPS provides automatic datum definition if minimum trace datum is selected.

The program begins with reading and checking of input file, and consequently checking on the number of datum defects. Depending on the number of coordinates used for datum definition (equation 2.81), transformation results can be based on either a minimum trace, partial minimum trace or minimum constraints datum. If the number of selected coordinates is less than the number of datum defects, the program is terminated automatically.

In COMPS, the general form of S-transformations is used (section 2.3.5 and equation 2.95). Options are included for datum definition, reduction to centroid and normalization. For numerical stability in evaluating matrix G, the computational mode allows for reduction of coordinates to their centroid (equation 2.90) and normalization of G (equations 2.91 and 2.92).

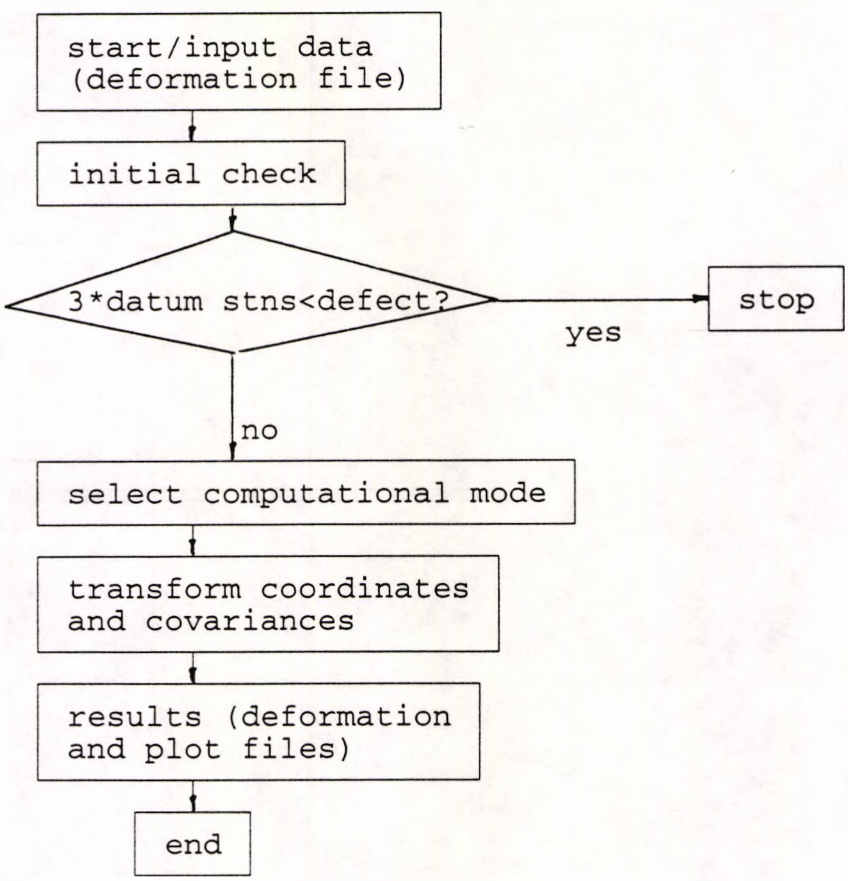

Figure 5.2 Flowchart for program COMPS

To speed up the computation, matrix S is decomposed by equations (2.93) and (2.94) to change multiplication into addition of large matrices. During development of program COMPS, computational check via equations (2.79) and (2.80) was performed.

Output of program COMPS consists of two files, as chosen by the user. The first file (compulsory) is in the same format as deformation file and is suitable for deformation detection, while the second file is a plot file for DCRE with the format similar to program ESTIMATE.

5.1.3 Program COMON [determination of common stations]

COMON is a program which allows the determination of common stations between two epochs, with the necessary re-ordering and S-transformations with respect to the common stations. Figure 5.3 summarizes the steps involved, based on section 4.2.1.

The program requires two data input files (i.e. deformation files), one from each epoch. Both epochs can be based on different numbers of stations and datum definition. For flexibility, common stations for each epoch do not necessarily have the same approximate coordinates. At the moment, for convenience in computation, the number of stations in the first epoch must be greater than those in the second epoch.

The program starts by reading and checking the input files and data. If the datum defect between epochs is different, the program automatically determines the biggest defect as the common defect for checking purposes. This is followed by searching and indexing of the common stations with respect to the first epoch. If three times the number of common stations is equal to or greater than the common defect, the procedure continues with re-ordering and S-transformations. Otherwise (or if no common stations are found), the program will automatically stop.

Re-ordering involves re-arrangement of the data in each epoch (station names, provisional and estimated coordinates together with their cofactor matrix) so that common stations are ordered at the beginning of arrays, based on the partitioning scheme of equations (4.1a) and (4.1b). If provisional coordinates of common stations are different, it is also possible to continue the process using values of first epoch.

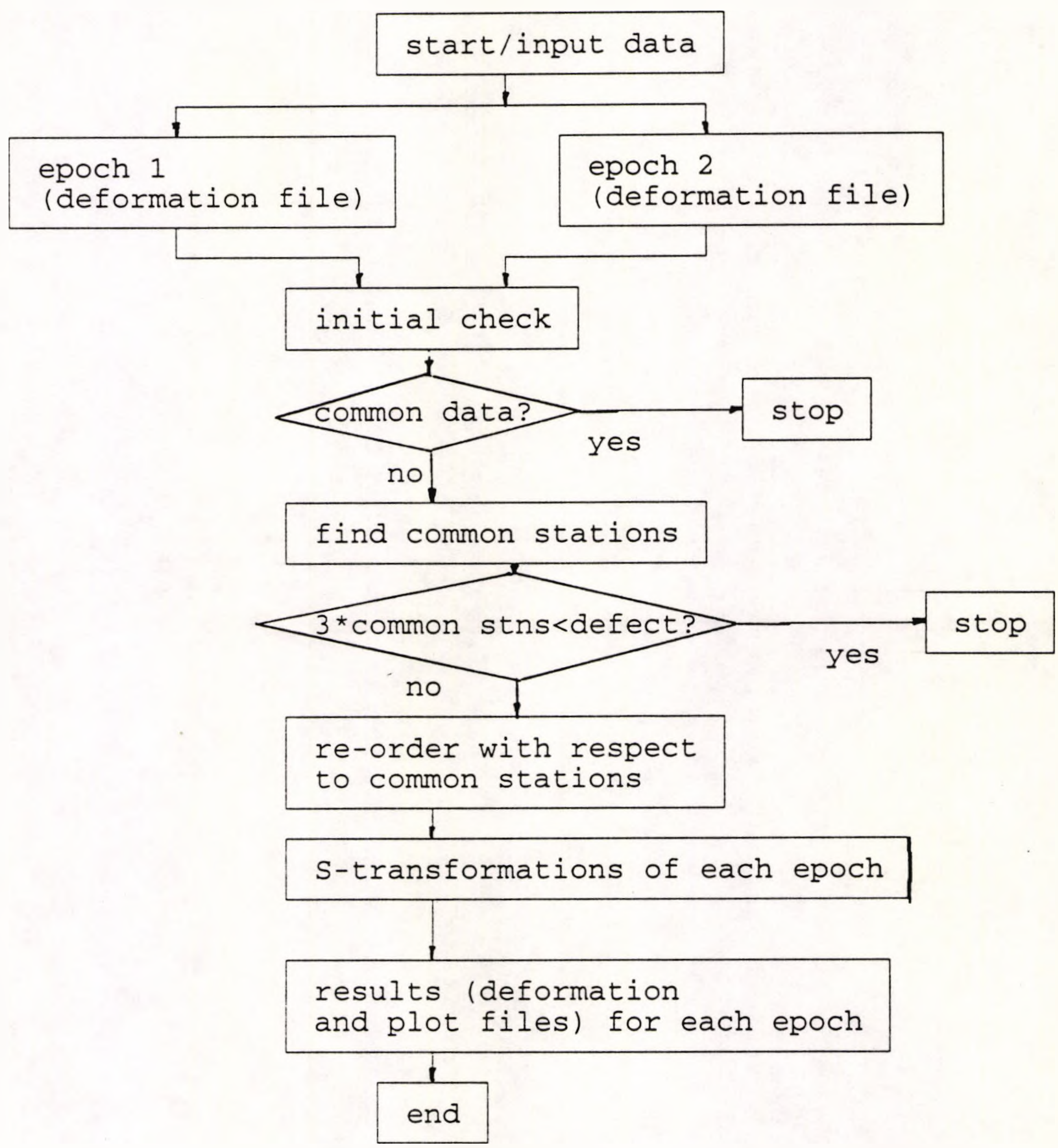

Figure 5.3 Flowchart for program COMON

The next step is the independent S-transformations (section 2.3.5) of the estimated coordinates and cofactor matrix from each epoch into the new datum defined by the common stations (equations 4.2 and 4.3). The transformation process does not change the datum invariant quantities (including degrees of freedom and variance factor) for each epoch.

COMON produces up to four output files, two for each epoch, namely deformation and plot files. The deformation file (compulsory) contains data with respect to the common stations only, suitable for the application of deformation detection via DETECT or ROBUST. The format for the plot file for graphics display via DCRE is the same as that of ESTIMATE.

5.1.4 Program DETECT [deformation detection by congruency testing]

DETECT is a computer program for geometrical detection of spatial deformation with the following properties: geometrical method; based on surveying and / or photogrammetric data; uses two-epoch analysis, an absolute monitoring network, static model, coordinate differencing and; assumes no correlation between epochs. The one-stage computational procedure for DETECT is based on section 4.4 and is summarised in Figure 5.4.

Two data input files are required for DETECT, one from each epoch. The data files (deformation files) can be obtained from the appropriate output of ESTIMATE, GAP, COMPS, or COMON. The data format is described in section 5.1.1.

Program DETECT starts with the reading and checking of input files, selection of datum definition, and initial checks on the data. The initial checks (section 4.2.2) examine that the same common stations, provisional coordinates and datum definition are being used in the LSE of each epoch. Any discrepancies will cause termination of the program, and if applicable, followed by the on screen advice for executing COMPS or COMON.

At the start of deformation detection, the datum used for LSE at each single epoch can be either minimum trace, minimum constraints or partial minimum trace (section 4.2.1). Datum and non-datum points are related to stable (or reference) and unstable (or object) points respectively. Program DETECT allows the user to define the status of datum points, whether known (i.e. partial minimum trace) or unknown (minimum trace or minimum constraints) in advance. This can be carried out by editing the input files.

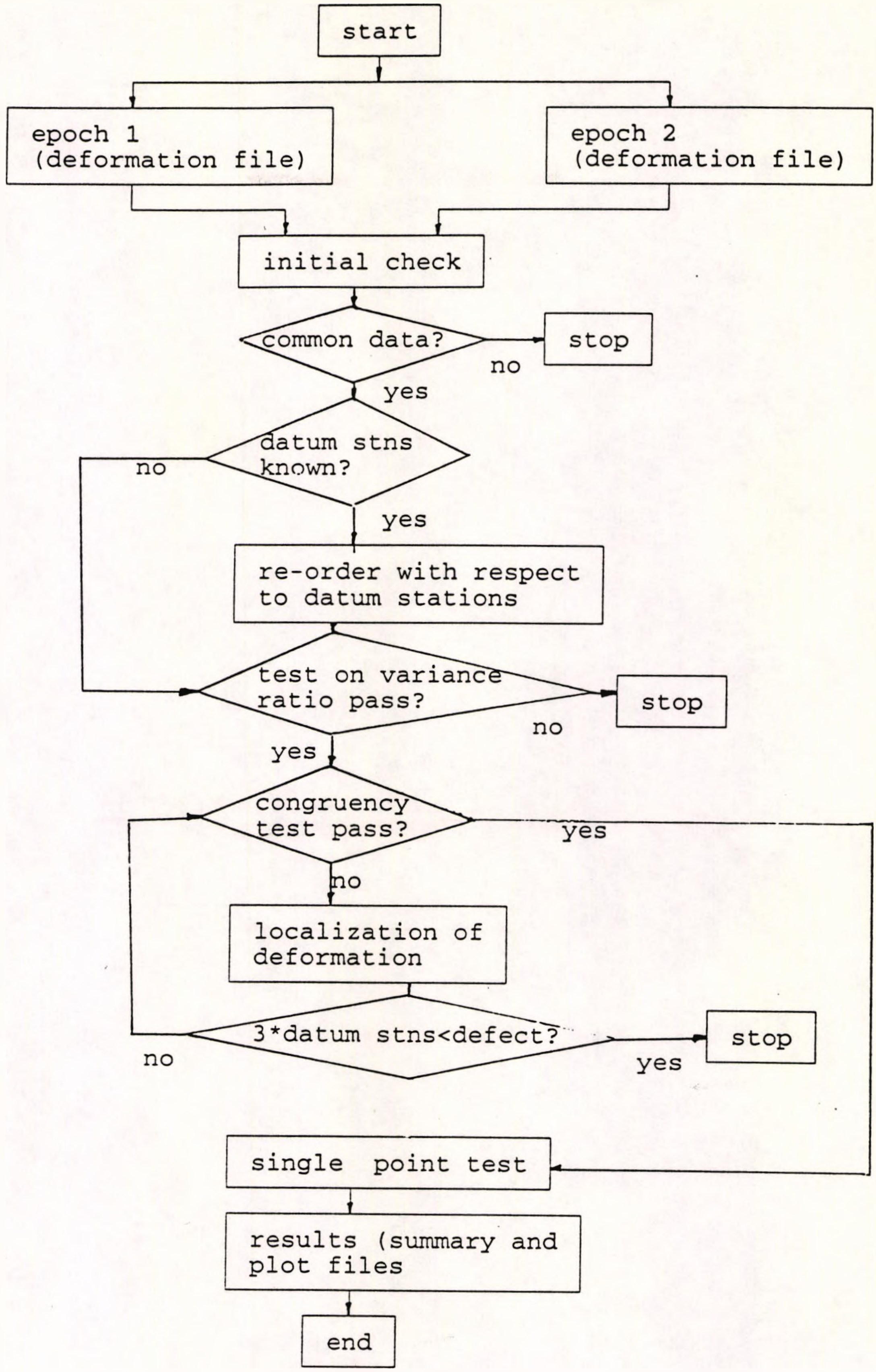

Figure 5.4 Flowchart for program DETECT

If datum points are not known, all stations will be used for datum definition. Otherwise, the relevant data for the datum stations at each epoch (coordinates, cofactor matrix and datum codes) are rearranged at the beginning of their respective arrays for simplicity in further computations (equation 4.17). A check on the number of coordinates used for datum definition against the datum defect is also carried out at this stage, and whenever the datum is redefined. The program is terminated automatically if three times the number of coordinates is less than the number of datum defects.

Three statistical tests are employed in DETECT, one-tailed test on the variance ratio (equation 4.6b), congruency test (section 4.2.3.1) and single point test (section 4.2.3.3). During this testing, the user can select the appropriate significance levels, while approximate critical values are computed automatically. The user also has the option of entering the critical values manually.

Following initial checks, a preliminary test on the variance ratio examines the compatibility of the independent variance factors at each epoch. Acceptance of the test leads to computation of common variance factor (equation 4.8) and stability determination. Failure of the test will terminate the program automatically, and requires the examination of LSE results and observational data.

Stability determination starts with checks on the stability of initial datum points in both epochs via congruency tests (equation 4.10). If the test indicates significant movements of the datum points, localization of deformation (section 4.2.3.2) is performed. Stability determination and localization of deformation consists of an iterative process of congruency testing, decomposition of the quadratic form (equation 4.14), re-ordering with respect to datum points (equation 4.17) and S-transformations (equations 4.18b and 4.20) until the partial congruency test (equation 4.13) passes. During this iterative procedure, the datum point with the largest quadratic contribution is removed from the computational base.

In DETECT, approximate coordinates are used for computing matrix G. Computation of displacement is with respect to the first epoch. The program also provides automatic and manual modes of the above iterative procedure. In the manual mode, user can also select which datum point is to be removed from the computational base. This facility is not possible in automatic mode. During datum re-definition, a check against datum defect is also carried out.

Once the congruency test passes, the localization procedure will compute any significant deformation vectors of non-datum points with respect to the final datum points. Final testing of deformation is in the form of a single point test (equation 4.22a). Program DETECT computes the standardized significance level for this test via equation (4.32), and also allows the user to change the level and its critical value.

Another useful feature of DETECT is the direct comparison between two epochs (equation 2.88). In this mode, each epoch must be based on the same datum.

DETECT will produced summary and plot files, as selected by the user. The summary file (compulsory) contains the full results of deformation detection, whilst the plot file contains useful information for DCRE. In the plot file, data for each station consists of station names, coordinates, deformation vectors, and their respective sub-cofactor matrix. By using DCRE, the deformation vectors and error ellipses (in three axes) can be portrayed graphically.

5.1.5 Program ROBUST [deformation detection by robust method]

ROBUST is also a program for detection of spatial deformation, but on a different principle to DETECT. Whilst DETECT uses congruency testing, ROBUST is based on the robust method (section 4.2.3.4) of iterative weighted S-transformations. Figure 5.5 shows the concept and one-stage computational procedure of ROBUST.

The program has two computational modes, deformation detection by robust method and direct S-transformations assuming known datum. In deformation detection, the user can decide on the maximum number of iterations. The data input handling, initial checks and the test on variance ratio are similar to DETECT.

The procedure for deformation detection via ROBUST is quite simple. Following the test on the variance ratio, the program computes the iteration limit (equation 4.28) based on the RMSE of the variances (section 3.8). It is also possible to enter the limit manually. At the beginning of iteration, the weight matrix is taken as identity, and in successive iterations, the weighting scheme of equation (4.29) is applied in equation (4.24a). In effect, stations with less movements are given more weight, and hence have more influence in the datum definition. At the end of iteration, the cofactor matrix of displacement is updated via equation (4.30). This is

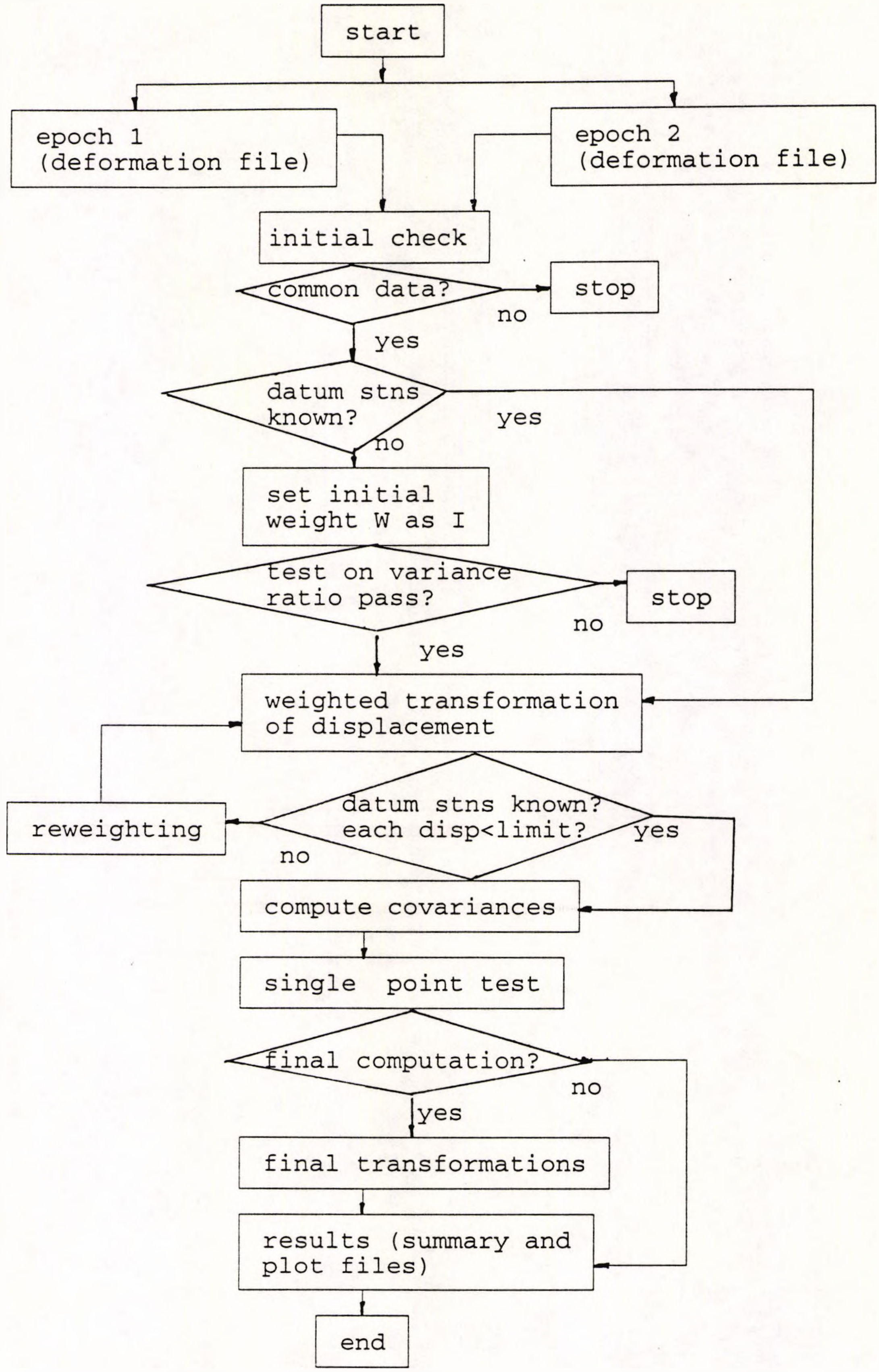

Figure 5.5 Flowchart for program ROBUST

followed by the single point test (section 4.2.3.3) in a similar manner to DETECT.

Program ROBUST computes S-transformations via equation (2.95). It also allows the user to perform final computation with the datum defined by the stable points. In this case, S-transformations and the single point test are applied again. Similarly, if datum points are known in advance, program ROBUST performs direct S-transformations, followed by the single point test. The output files from ROBUST are similar to those from DETECT (section 5.1.4), consisting of summary and plot files.

5.2 Using the programs

All the five implemented programs and GAP are connected for deformation detection purposes, and can produce special output for DCRE. During this research, the links between all seven programs have been established, as shown in Figure 5.6.

If datum (and hence stable) stations are known prior to deformation detection, results from LSE can be transformed via COMPS. Generally, assuming different stations and datums, and before proceeding with deformation detection, COMON can be used to search for common stations between epochs, with the appropriate S-transformations.

Programs ESTIMATE, GAP, COMPS and COMON produced deformation files suitable for deformation detection. For initial datum definition in detection process these files may be edited.

DETECT or ROBUST can be used for geometrical detection of spatial deformation via one-stage computation (i.e. stability determination and localization of deformation), and provide numerical results. For graphics display, a plot file for DCRE can be created by both programs.

For LSE at each epoch, ESTIMATE is suitable for processing terrestrial surveying data as described in section 5.1.1, while GAP is capable of processing both photogrammetric and ordinary surveying data. Programs COMPS, COMON, DETECT and ROBUST, however, are capable of processing LSE results obtained from any combinations of surveying, photogrammetry and / GPS data, because the datum defect forms part of the input data.

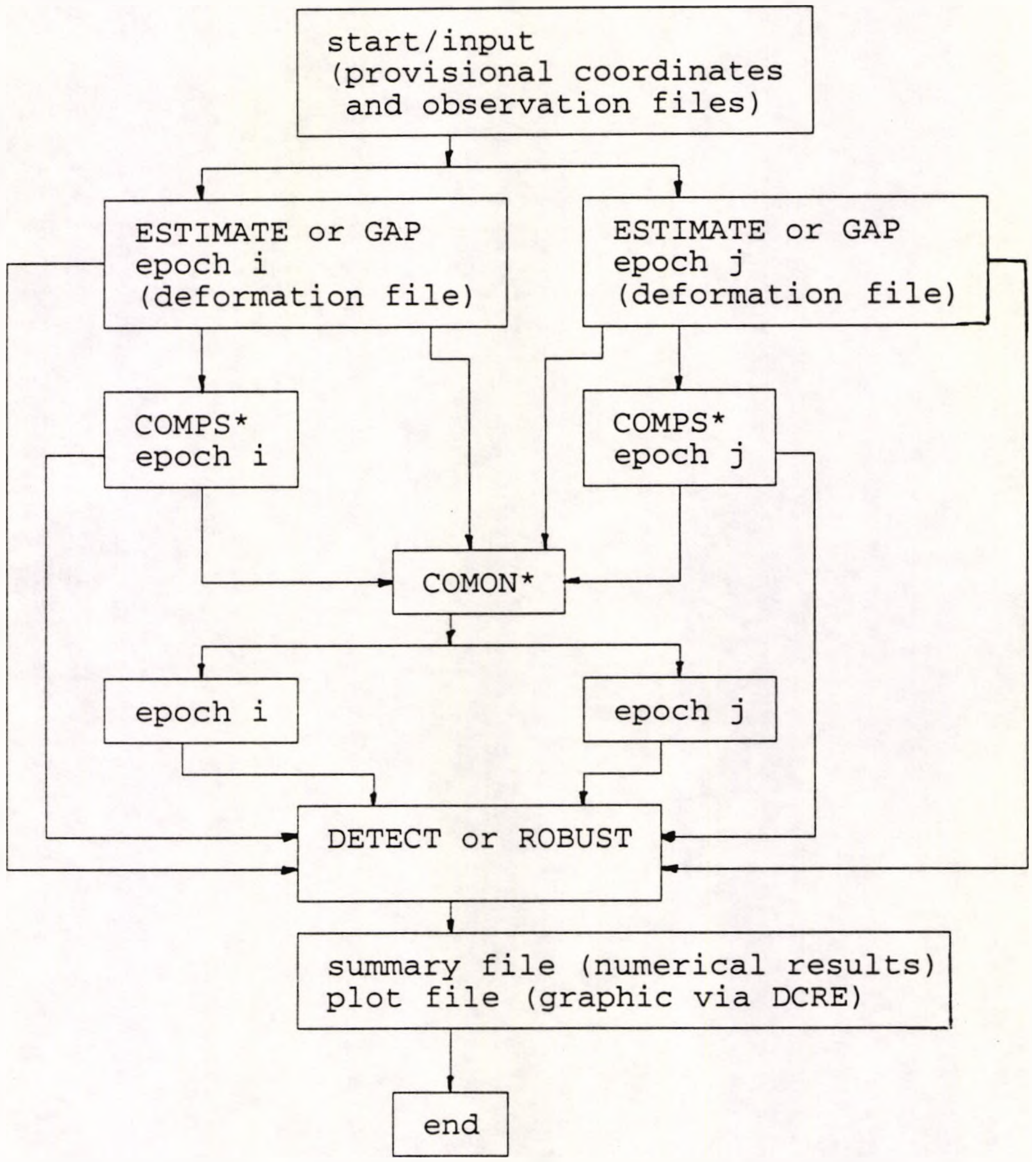

Figure 5.6 Linking the programs for geometrical detection of spatial deformation (* indicates that the program is only executed if necessary)

In practice, interpretation of the results obtained from the usage of the above programs is very helpful. In LSE (via ESTIMATE or GAP), the solutions should pass both the global and local tests. Failure of the above tests requires the application of strategies outlined in section 3.9. Moreover, the precision and reliability analyses must be acceptable too. In geometrical detection of spatial deformation (program DETECT or ROBUST), it is expected that all datum points will be stable, while non-datum points can be either stable or unstable as indicated by the single point test (section 4.4).

In program DCRE, views can be rotated about three axes, and the standard views are: top, bottom, right, left, front, and isometric view. Top (or plan), front and right views can be used to show ellipses and deformation vectors in xy, xz and yz planes respectively. The isometric view is useful for portraying the trend of overall movements. Figure 5.7 is an isometric view of a cube to show the concept of 3-D views.

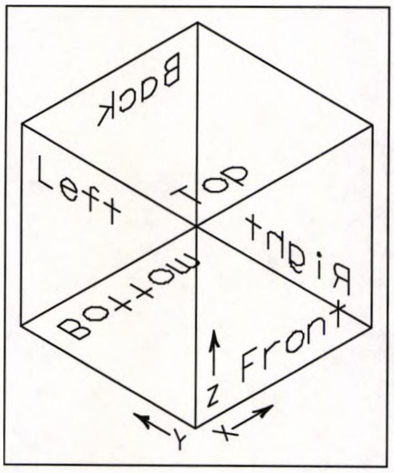

Figure 5.7 Concept of 3-D views

Graphically (program DCRE), plots of displacement vectors for stable datum points will be within the ellipses describing the confidence regions. On the other hand, plots of displacement vectors for unstable non-datum points (i.e. with significant movement) will be outside the ellipse.

A guide for using the developed programs is given in Setan (1995), whilst sample input and output of the programs are listed in Appendix G.

6. APPLICATION

This chapter discusses the actual application and testing of the implemented programs described earlier in chapter 5, for processing simulated and real data. Analysis and interpretation of the results obtained are also discussed in this chapter.

6.1 Simulation tests

During initial testing, simulated data were used extensively to verify the correctness of the adopted procedure. The simulated data used were based on properly generated normally distributed observations and also published data. The simulation tests were divided into four parts: rank analysis (section 6.1.1), datum definition and S-transformations (section 6.1.2), LSE and errors (section 6.1.3), and deformation detection (section 6.1.4). A significance level α=0.05 was chosen for most of the statistical testing.

6.1.1 Rank analysis

Two networks have been used to illustrate the importance of rank analysis: levelling and 3-D networks. The levelling network shown in Figure 6.1 is adopted from Caspary (1987b). The observations (all of equal weights as I) are

$$l^t=[l_1\ l_2\ l_3\ l_4\ l_5\ l_6]^t=[1.2\ 1.6\ 1.7\ 1.2\ 2.1\ 1.3]^t \text{ mm}$$

After forming the observation equations, the normal equation coefficient matrix N(4,4) according to equation (2.17) is

$$N=\begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}$$

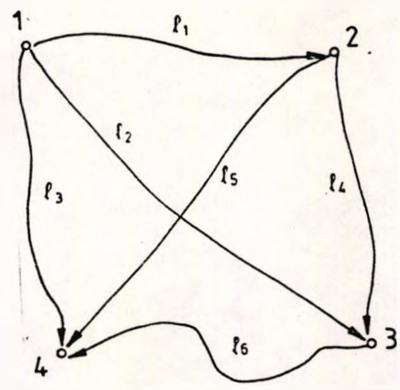

Figure 6.1 Simple levelling network

[number of observations n=6, number of parameters u=4]

The eigenvalues (equation 2.58) of N=[4.0 2.2e-16 4.0 4.0]^t. If the limit or tolerance is 0.00001, rank (N) is 3. Hence, rank deficiency (equation 2.33) of N is

d=u-rank (N)=4-3=1

Matrix N is singular, and it does not possess the ordinary (Cayley) inverse. This is because the levelling network has one datum defect (Table 2.1). Although the limit can be chosen to be very small (for example 1e-20), such that N becomes full rank, the condition cond(N) will be too large (2e16) indicating an ill-conditioned situation. Moreover, any ordinary inversion routine will not be able to compute the inverse of N.

In order to solve the above singular equation, it is required to define a datum. The most simple can be defined by means of a minimum constraints (equations 2.44 and 2.45a), achieved for example by keeping point 1 as fixed hence u reduces to 3. The reduced N(3,3) becomes

$$N=\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$$

With eigenvalues of N=[4 1 4]^t, r(N)=3, and cond(N)=4. Matrix N is now non-singular, its inverse and subsequent solution can be computed as usual.

Figure 6.2 is a simulated six station 3-D network consisting of 54 observations (12 slope distances, 30 horizontal directions and 12 height differences). Details on the simulated data are given in section 6.1.3. To demonstrate the handling of these deficiencies, all observed distances and height differences related to station 2 were removed from the scheme, leaving only angular observations connecting station 2 into the network. The number of observations reduces to 47, and the number of parameters is 24 (18 unknown coordinates and 6 orientation parameters).

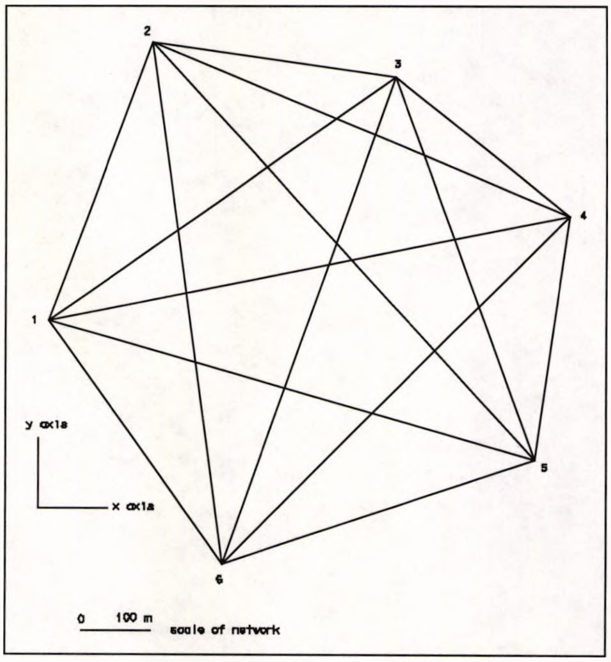

Figure 6.2 Network 1 (plan view)

Initially, no station was considered fixed, and program ESTIMATE in pre-analysis mode revealed a datum defect of four (i.e. 3 translation along all axes and one rotation about the z axis). These datum defects may be removed by fixing 4 coordinates, in this case, x_1, y_1, z_1 and y_3, such that the number of parameters becomes 20. The following LSE with a minimum constraints datum gives a set of normal equations with an effective rank of 19, indicating one configuration defect.

Geometrically, this is due to not enough observations to locate station 2. By adding one observation (and its standard deviation), such as the slope distance from 1 to 2, the normal equation matrix becomes non-singular. However, the normal equation was found to be ill-conditioned via rank analysis (section 2.2.6). In reliability analysis, the redundancy number (section 3.8) of the added observation was found to be close to zero, indicating a weak area in the network, and more observations connected to station 2 are required. By iteratively adding relevant observations and performing pre-analysis, an acceptable scheme can be designed.

6.1.2 Datum definition and S-transformations

Several tests were carried out to demonstrate the important aspects of datum definition and S-transformations. In the first test, the levelling network shown in Figure 6.1 was used. LSE by the method of observation equations (section 2.1.3) utilized equations (2.29), (2.30), (2.31) and (2.34).

Assuming all observations have equal weights (given by I), and using approximate values of parameters as $[10.0\ 11.1\ 11.5\ 11.6]^t$, three types of solutions were computed:

#1. Ordinary minimum constraints with station 1 chosen as the datum point.
#2. Minimum trace where all stations are used for datum definition.
#3. Partial minimum trace with stations 2 and 3 as datum points.

The results obtained are similar to Caspary (1987b) and are depicted in Table 6.1. Solution #2 gives minimum norm and minimum trace whilst solution #3 minimizes the partial norm and partial trace, with respect to the datum points. The residuals remain unchanged due to their independence on the selection of datum constraints. Although not computed here, other invariant quantities (section 2.1.4) are adjusted observations, cofactor matrices of the adjusted observations and the residuals.

S-transformations of the LSE results in Table 6.1 were performed using equation (2.89) where $G=[1\ 1\ 1\ 1]^t$. Three types of computational bases were chosen:

1. Minimum trace to minimum constraints
Datum point = station 1; $I_p = [1\ 0\ 0\ 0]^t$
2. Minimum constraints to minimum trace
Datum points = all stations (1,2,3,4); $I_p = [1\ 1\ 1\ 1]^t$
3. Minimum trace to partial minimum trace
Datum points = stations 2 and 3; $I_p = [0\ 1\ 1\ 0]^t$

	solution #1 minimum constraints datum	solution #2 minimum trace datum	solution #3 partial minimum trace datum
datum defined by stn	1	1,2,3,4 [all]	2,3
parameter			
x_1	10.0	9.925	10.25
x_2	10.6	10.525	10.85
x_3	11.5	11.425	11.75
x_4	12.4	12.325	12.65
norm	0.94	0.93	1.14
partial norm	-	-	0.35
cofactor matrix			
Q_{11}	0.00	0.1875	0.375
Q_{21}	0.00	-0.0625	0.000
Q_{22}	0.50	0.1875	0.125
Q_{31}	0.00	-0.0625	0.000
Q_{32}	0.25	-0.0625	-0.125
Q_{33}	0.50	0.1875	0.125
Q_{41}	0.00	-0.0625	0.125
Q_{42}	0.25	-0.0625	0.000
Q_{43}	0.25	-0.0625	0.000
Q_{44}	0.50	0.1875	0.375
trace	1.500	0.750	1.000
partial trace	-	-	0.250
residuals			
v_1	-0.6	-0.6	-0.6
v_2	-0.1	-0.1	-0.1
v_3	0.7	0.7	0.7
v_4	-0.3	-0.3	-0.3
v_5	-0.3	-0.3	-0.3
v_6	-0.4	-0.4	-0.4

Table 6.1 LSE results for levelling network (unit in mm).

Results of the transformations are summarised in Table 6.2. Comparison of the results of LSE in Table 6.1 with results of S-transformations in Table 6.2 (columnwise) shows that the parameters and cofactor matrices are identical. Invariant quantities are not affected by this transformation.

	sol#2 to #1 minimum trace to minimum constraints	sol#1 to #2 minimum constraints to minimum trace	sol#2 to #3 minimum trace to partial minimum trace
datum defined by stn *	1	1,2,3,4 [all]	2,3
parameter			
x_1	10.0	9.925	10.25
x_2	10.6	10.525	10.85
x_3	11.5	11.425	11.75
x_4	12.4	12.325	12.65
cofactor matrix			
Q_{11}	0.00	0.1875	0.375
Q_{21}	0.00	-0.0625	0.000
Q_{22}	0.50	0.1875	0.125
Q_{31}	0.00	-0.0625	0.000
Q_{32}	0.25	-0.0625	-0.125
Q_{33}	0.50	0.1875	0.125
Q_{41}	0.00	-0.0625	0.125
Q_{42}	0.25	-0.0625	0.000
Q_{43}	0.25	-0.0625	0.000
Q_{44}	0.50	0.1875	0.375

Table 6.2 S-transformations results for the levelling network (unit in mm).

The second test was based on a simulated 3-D network of four stations (Figure 6.3) with 21 survey measurements (slope distance, height difference, vertical angle and uncorrelated horizontal angle), giving rise to a datum defect of four. The above observables can be handled by GAP, and hence allowed an independent check of computations via GAP. LSE were carried out (using ESTIMATE and GAP) with a minimum constraints datum (fixing x_1, y_1, z_1 and x_3) and a minimum trace datum (using GAP). The results are shown in Table 6.3. The minimum constraints solution is transformed into the minimum trace solution using COMPS with the following options for handling matrix G:

(a) Ordinary computation without reduction to centroid and normalization
(b) Reduction to centroid and normalization
(c) Reduction to centroid only
(d) Normalization only

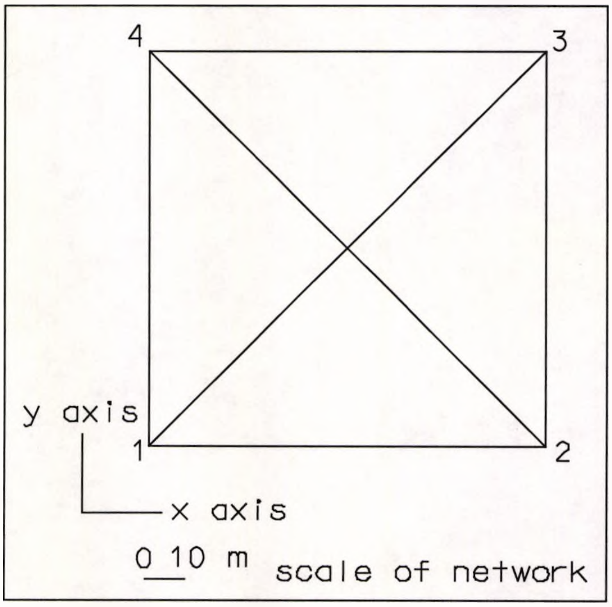

Figure 6.3 Network 2 (plan view)

All the options were found to produce identical results (Table 6.3). Differences between the transformed solution and results obtained from GAP are found to be insignificant. For practical purposes, reduction to the centroid is recommended, since S-transformations results are automatically referred to the centroid defined by the datum points. Normalization provides further computational stability.

Similar tests were performed using the six station 3-D network shown in Figure 6.2. Initially, ESTIMATE was used to obtain the minimum constraints solution (fixing x_1, y_1, z_1 and x_3. Error ellipses in all 3 axes are shown in Figures 6.4, 6.5 and 6.6 respectively. All the graphics in this chapter will be shown in plan, front and right views to portray ellipses and deformation vectors in xy, xz and yz directions respectively. Isometric view is used to show the trend of deformation.

	stn	coordinates			standard deviation		
		x	y	z	σ_x	σ_y	σ_z
1.0							
minimum	1	100.0000	100.0000	10.0000	0.0000	0.0000	0.0000
constraints	2	199.9969	99.9967	19.9999	0.0025	0.0028	0.0018
datum	3	200.0000	199.9954	14.9989	0.0000	0.0042	0.0018
	4	100.0014	199.9968	4.9994	0.0025	0.0025	0.0018
2.0							
minimum	1	100.0016	100.0016	10.0005	0.0013	0.0013	0.0010
trace	2	199.9985	100.0007	20.0004	0.0014	0.0014	0.0013
datum	3	199.9993	199.9993	14.9993	0.0013	0.0013	0.0011
(from GAP)	4	100.0007	199.9984	4.9998	0.0014	0.0014	0.0013

3.0 S-transformations from minimum constraints to minimum trace datums

		x	y	z	σ_x	σ_y	σ_z
3.1							
ordinary	1	100.0016	100.0016	10.0004	0.0013	0.0013	0.0010
computation	2	199.9985	100.0006	20.0004	0.0014	0.0014	0.0013
	3	199.9993	199.9993	14.9994	0.0013	0.0013	0.0011
	4	100.0007	199.9984	4.9998	0.0014	0.0014	0.0013
3.2							
reduction	1	100.0016	100.0016	10.0004	0.0013	0.0013	0.0010
to centroid	2	199.9985	100.0006	20.0004	0.0014	0.0014	0.0013
and	3	199.9993	199.9993	14.9994	0.0013	0.0013	0.0011
normalize G	4	100.0007	199.9984	4.9998	0.0014	0.0014	0.0013
3.3							
reduction	1	100.0016	100.0016	10.0004	0.0013	0.0013	0.0010
to centroid	2	199.9985	100.0006	20.0004	0.0014	0.0014	0.0013
only	3	199.9993	199.9993	14.9994	0.0013	0.0013	0.0011
	4	100.0007	199.9984	4.9998	0.0014	0.0014	0.0013
3.4							
normalize	1	100.0016	100.0016	10.0004	0.0013	0.0013	0.0010
G only	2	199.9985	100.0006	20.0004	0.0014	0.0014	0.0013
	3	199.9993	199.9993	14.9994	0.0013	0.0013	0.0011
	4	100.0007	199.9984	4.9998	0.0014	0.0014	0.0013
differences	1	0.0000	0.0000	-0.0001	0.0000	0.0000	0.0000
(3.2)-(2.0)	2	0.0000	-0.0001	0.0000	0.0000	0.0000	0.0000
	3	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 6.3 Comparison between the results of S-transformations for a small 3-D network (unit m).

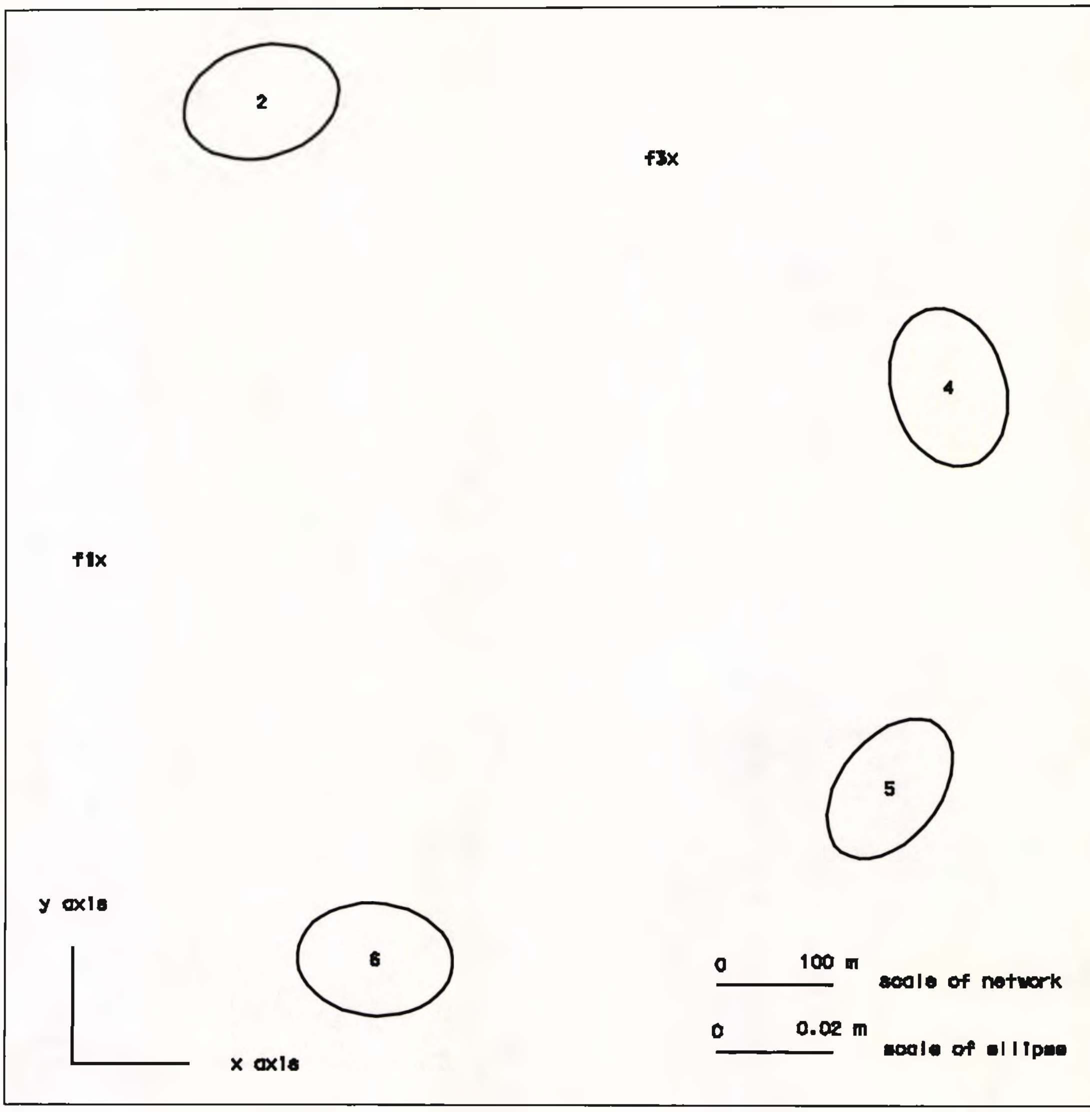

Figure 6.4 Minimum constraints solution for network 1 (plan view and xy ellipse)
[datum defined by fixing x_1, y_1, z_1 and y_3]

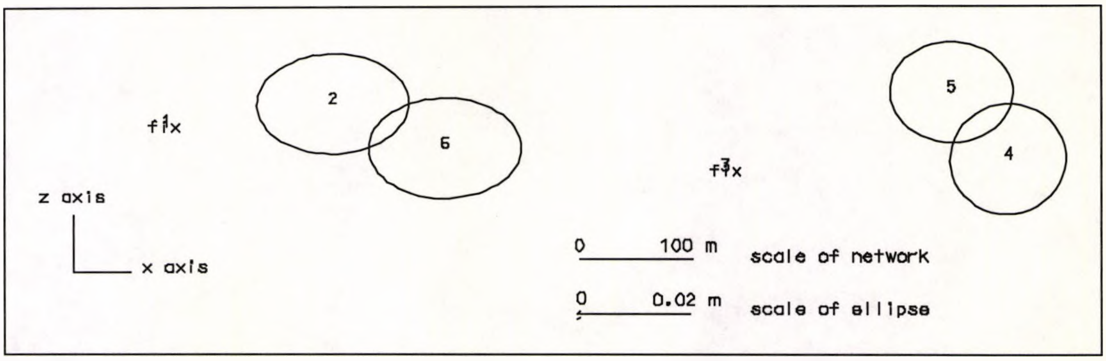

Figure 6.5 Minimum constraints solution for network 1 (front view and xz ellipse)

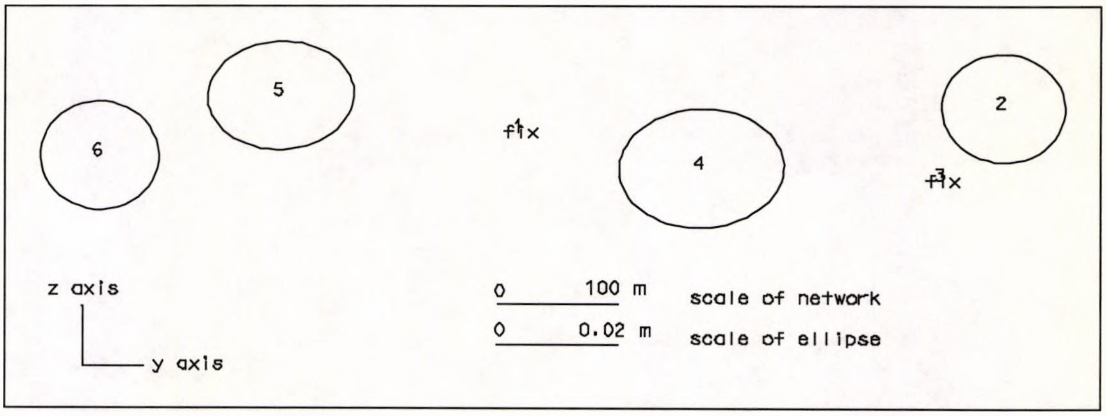

Figure 6.6 Minimum constraints solution for network 1 (right view and yz ellipse)

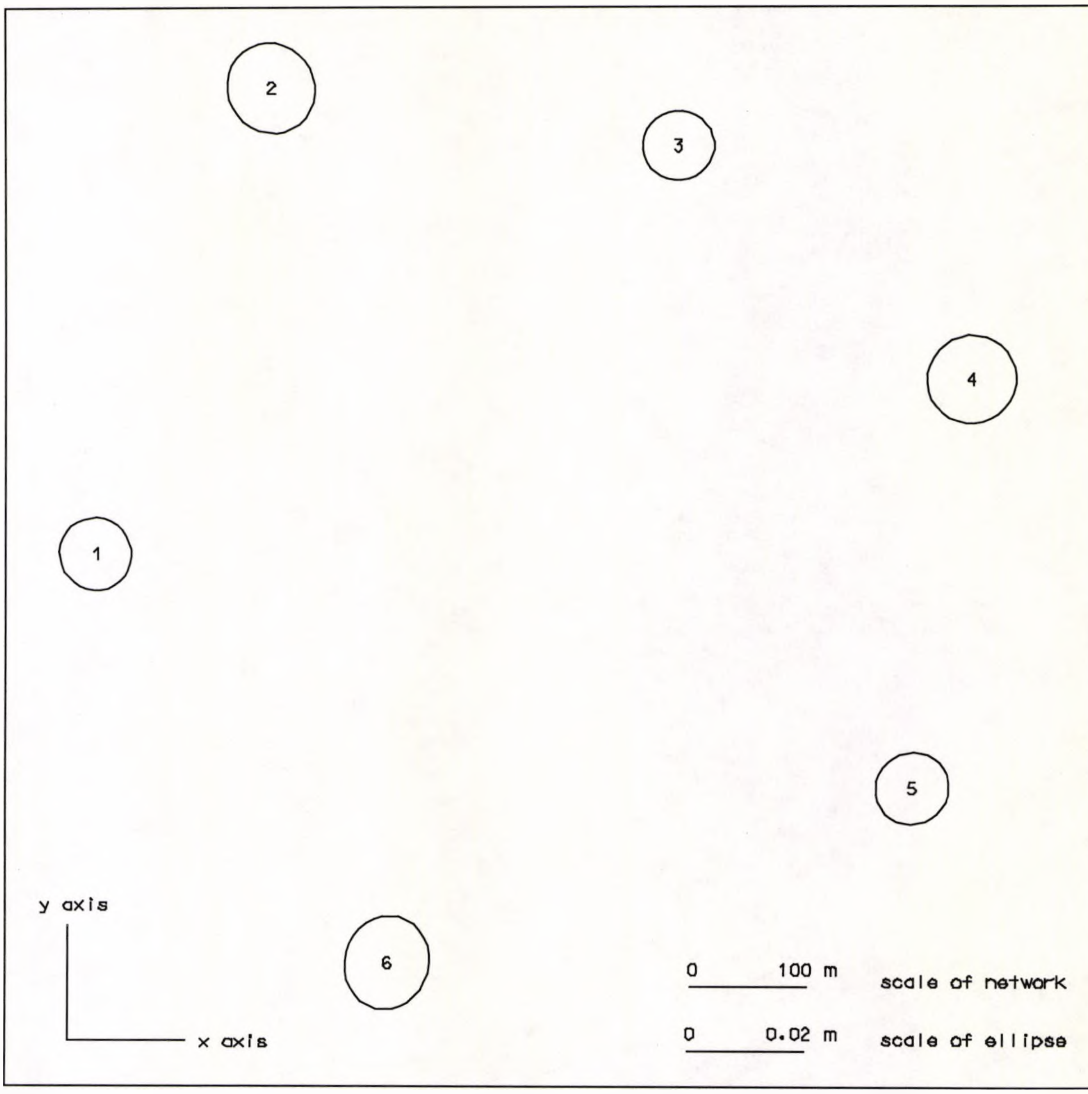

Figure 6.7 Minimum trace solution for network 1 (plan view and xy elipse)
[datum defined by all stations]

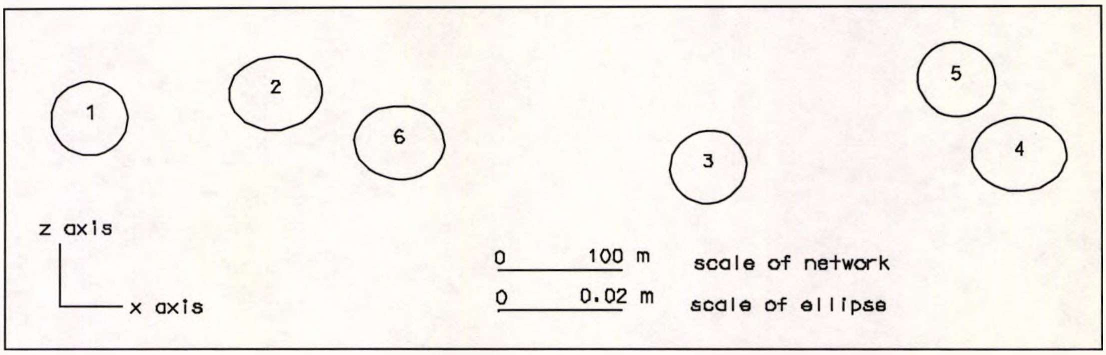

Figure 6.8 Minimum trace solution for network 1 (front view and xz ellipse)

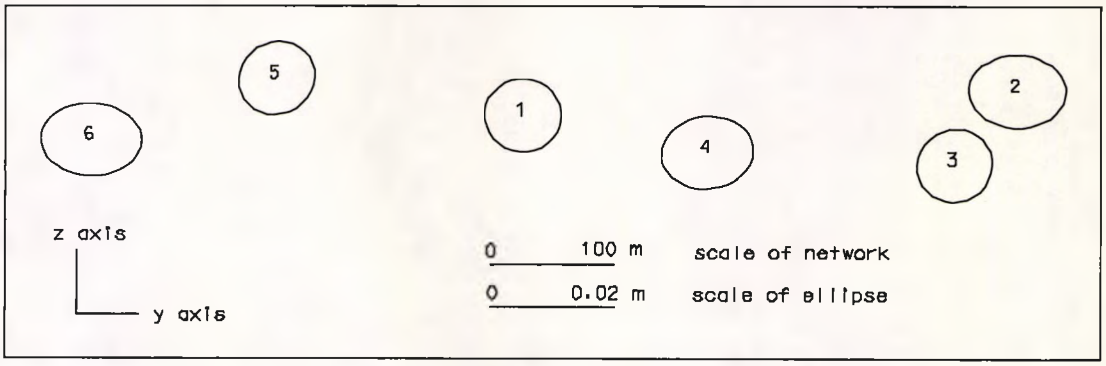

Figure 6.9 Minimum trace solution for network 1 (right view and yz ellipse)

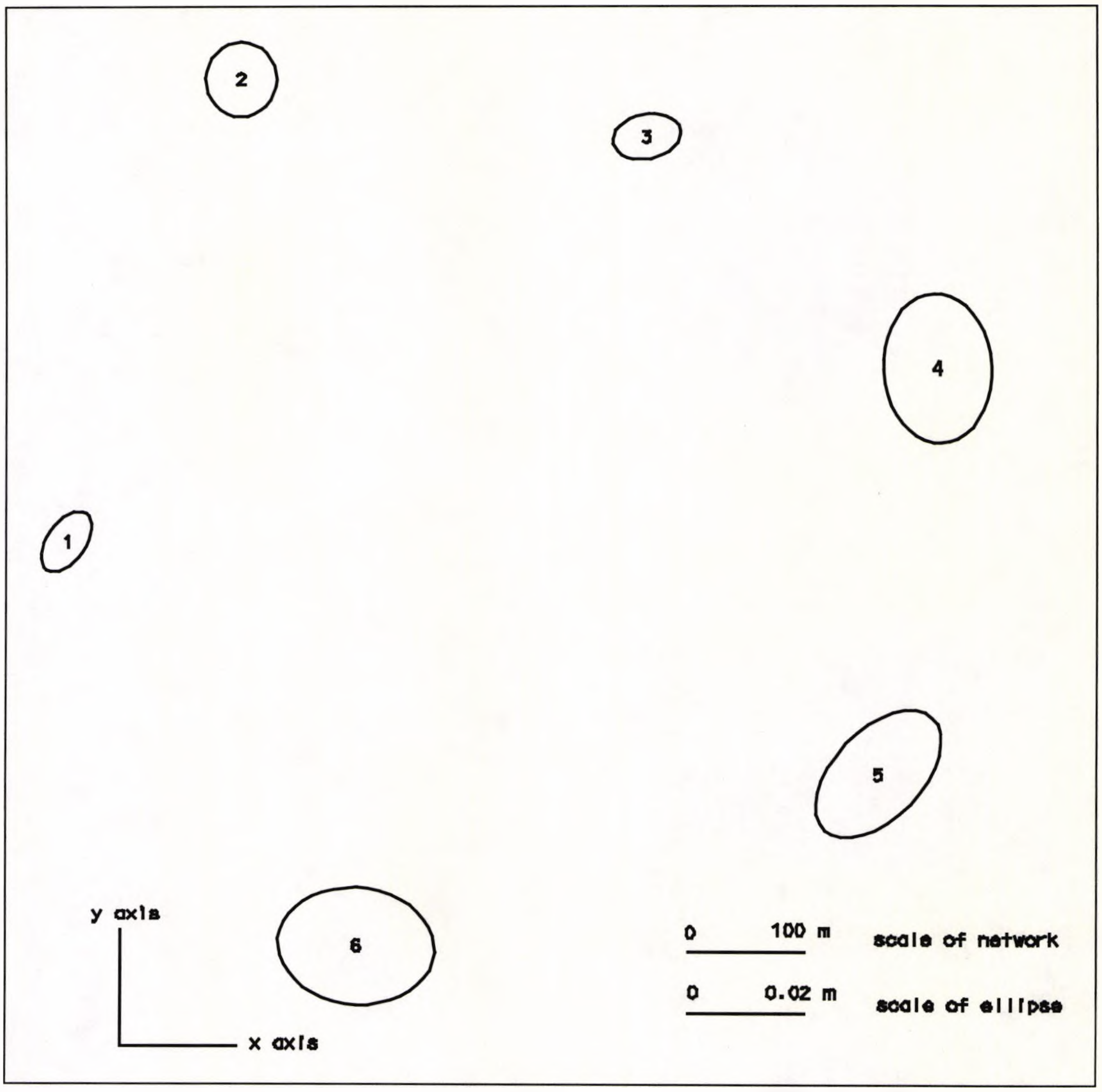

Figure 6.10 Partial minimum trace solution for network 1 (plan view and xy ellipse)
[datum defined by stations 1, 2 and 3]

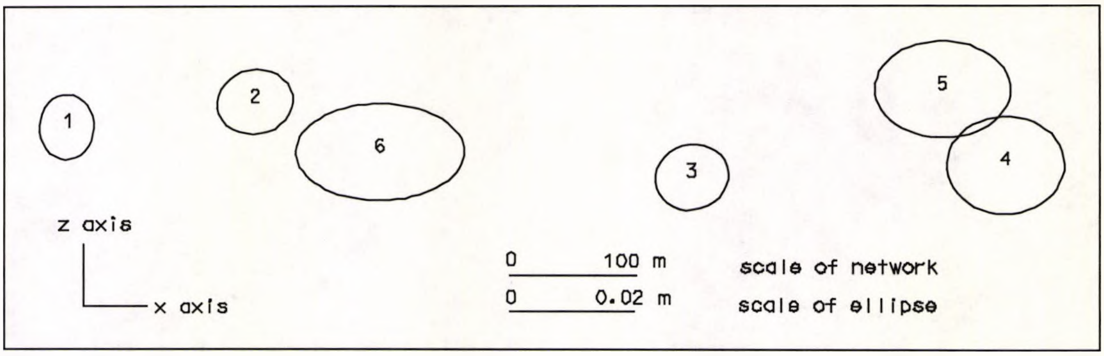

Figure 6.11 Partial minimum trace solution for network 1 (front view and xz ellipse)

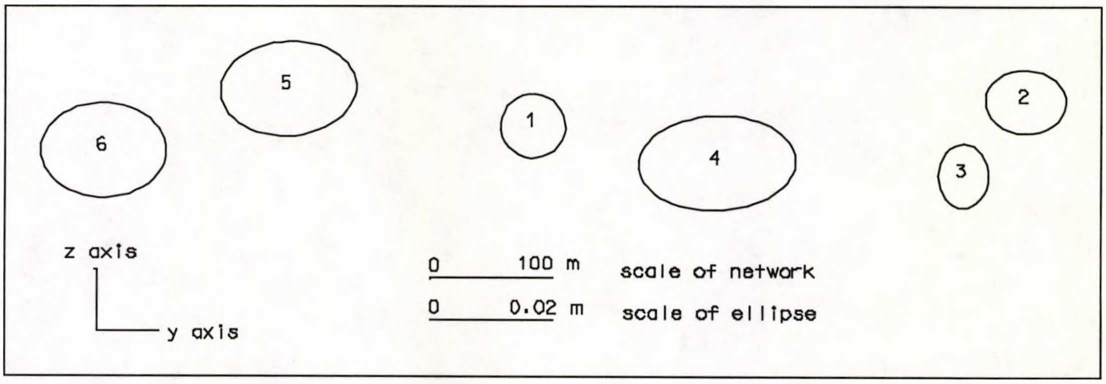

Figure 6.12 Partial minimum trace solution for network 1 (right view and yz ellipse)

The minimum constraints solution was then transformed by COMPS into minimum and partial minimum trace (with respect to stations 1, 2 and 3) solutions. The pattern of error ellipse for the minimum trace solution is shown in Figures 6.7 to 6.9, while that for partial minimum trace solution is given in Figures 6.10 to 6.12. The minimum trace solution (Figures 6.7 to 6.9) gives overall smaller ellipses, whilst the partial minimum trace solution (Figures 6.10 to 6.12) gives smaller ellipses at the datum stations.

The third test examined the time necessary to compute S-transformations by means of equations (2.89) and (2.94). LSE of a nine stations 3-D photogrammetric network was performed using GAP via minimum trace datum. It is required to transform the solution into partial minimum trace datum with respect to the first four stations.

The computational time for evaluation of equations (2.89) and (2.94) with reduction to centroid and normalization are 104 secs and 20 secs respectively, using an IBM compatible 286 (12 Mhz) personal computer. This shows that the decomposition process greatly improves the speed of computations. It is anticipated that the computational time will be quicker on a high performance computer or workstation.

6.1.3 LSE and errors

Simulation tests were conducted to evaluate pre-analysis, handling of random, systematic and gross errors. The 3-D network used is shown in Figure 6.2. The simulated network consists of 6 stations and 54 uncorrelated observations (12 slope distances (sd) with a simulated random error σ of 5 mm, 30 directions (dir) with σ of 5 secs and 12 height differences (dh) with σ of 5mm). The randomized observations are free from both systematic and gross errors because they are derived from known assigned coordinates. The data are listed in Appendix F.

The rank deficiency (d) is 4 (i.e. 3 translation and one rotation about z axis), and is removed by fixing coordinates x_1, y_1, z_1 and y_3. The number of parameters is 20 (14 coordinates and 6 orientation unknowns), giving rise to 34 degrees of freedom. The commonly accepted significance level α of 0.05 is used for statistical test. In subsequent reliability analysis, α_o 0.1% and β_o 20% are used, leading to a value of λ_o as 17.0 (Figure 3.1).

1. Pre-analysis and handling of random error

Pre-analysis is very important as it allows a preliminary evaluation of the network and computation of precision and reliability measures, through the use of provisional coordinates and observation precisions without actually using the observations.

Results from pre-analysis are shown in Table 6.4 (column 2), indicating a reasonable network with high precision (trace of 0.0001 m^2 in each axis) and reliability. In reliability analysis, all the redundancy numbers are greater than 0.3, and their average is 0.6. Also, the MDGE ranged between five to eight times σ. The external reliability is quite small, with the maximum influential factor less than 40.

LSE of the above network is then performed, and the results are indicated in Table 6.4 (column 3). After one iteration, the global, local and goodness of fit tests passed. Precision and reliability measures are the same as those from pre-analysis.

Assuming the precisions of the observations as unknown, VCE mode is executed, and the observations are divided into three groups. The results obtained are portrayed in Table 6.4 (column 4). After three iterations, global and local tests on the estimated variance factors passed. The usual global, local and goodness of fit tests also passed as well. The observation estimated precisions are 4.4 mm, 4.4 secs and 5.0 mm for distances, directions and height differences respectively. Such values are very close to the simulated σ. The precision and reliability measures are similar in magnitude to the pre-analysis results.

2. Handling of systematic error

For testing purposes, two types of systematic error (zero and scale) were introduced into all the distances measured from stations 1 and 3.

Firstly, a 50 mm zero error was introduced into the randomized distance observations. Three types of model were used to examine the effect of this zero error on the solution: using ordinary distances; multiple zero errors as additional or bias parameters and; distance differences as pseudo observables.

	pre-analysis	ordinary LSE	VCE
simulated random errors sd 5 mm dir 5 secs dh 5 mm			estimated 4.4 mm 4.4 secs 5.0 mm
rank analysis	ok	ok	ok
iteration	1	1	3
global test		pass (0.82)	pass (0.94)
goodness of fit test		pass	pass
local test		pass	pass
trace x axis y axis z axis (unit m^2)	0.0001 0.0001 0.0001	0.0001 0.0001 0.0001	0.0001 0.0001 0.0001
redundancy numbers	0.32-0.78	0.32-0.78	0.33-0.78
internal reliability	23.33-36.53	23.33-36.53	20.57-31.33
external reliability	4.77-36.38	4.77-36.78	4.86-33.27

Table 6.4 Pre-analysis, LSE and VCE for network 1
(units for internal reliability are mm for linear and secs for angular observations)

The solutions obtained are summarized in Table 6.5. Solution with ordinary LSE failed both global and local (2 observations) tests, and hence cannot be accepted. With the use of zero error in a bias parameter approach, both global and local tests passed. The simulated 50 mm zero error at stations 1 and 3 was recovered and found to be significant as 48±8 and 47±8 mm respectively. With distance differences as pseudo observables, the solution also passed both the global and local tests.

	ordinary LSE	LSE / add parameter [zero error]	LSE / pseudo observables [distance difference]
no of obs	54	54	52
df	34	32	32
rank analysis	ok	ok	ok
iteration	1	3	1
variance factor	5.67	0.89	0.86
global test	fail	pass	pass
local test	fail (2)	pass	pass
zero error (m)	-	0.048±0.008 0.047±0.008 significant	-
trace x axis y axis z axis	0.0001 0.0001 0.0001	0.0001 0.0001 0.0001	0.0003 0.0001 0.0001
redundancy number	0.32-0.78	0.00-0.78	0.14-0.77
internal reliability	23.33-36.53	23.38-too large	23.44-54.58
external reliability	4.77-36.38	5.06-too large	4.97-102.17

Table 6.5 Test with 50 mm zero error for network 1

Secondly, a scale error of 40 ppm (i.e. scale factor 0.99996) was incorporated into the randomized distance observations from stations 1 and 3. Three types of models were used for testing: using ordinary distances; multiple scale errors as additional parameters and; distance ratios as pseudo observables.

The solutions are shown in Table 6.6. The ordinary solution passed the global test but failed the local (one observation) test. Solutions using additional parameter and pseudo observable approaches passed both global and local tests. With the inclusion of an additional parameter, the scale factor was estimated to be significant as 0.99997±0.00001 at stations 1 and 3.

The results of ordinary LSE in Table 6.6 demonstrate that sometimes global test is not sensitive enough, and it is necessary to incorporate local test in LSE.

As shown in Tables 6.5 and 6.6, solutions based on the additional parameter and pseudo observable approaches are similar, but not exactly the same, due to different configurations. However, both approaches lead to a less reliable network as indicated by reliability analysis. The uses of pseudo observables and additional parameters provided satisfactory results. However, the use of additional parameters caused the redundancy numbers of some observations to be very close to zero.

	ordinary LSE	LSE / add parameter [scale factor]	LSE / pseudo observables [distance ratio]
no of obs	54	54	52
df	34	32	32
rank analysis	ok	ok	ok
iteration	1	2	1
variance factor	1.22	0.82	0.82
global test	pass	pass	pass
local test	fail (1)	pass	pass
scale error	-	0.99997±0.00001 0.99997±0.00001 significant	
trace x axis y axis z axis	0.0001 0.0001 0.0001	0.0003 0.0001 0.0001	0.0003 0.0001 0.0001
redundancy number	0.32-0.78	0.17-0.77	0.17-0.77
internal reliability	23.33-36.47	23.45-49.41	23.45-49.41
external reliability	4.77-36.20	4.98-80.66	4.99-80.66

Table 6.6 Test with 0.99996 scale factor for network 1

For practical purposes, the use of additional parameters is recommended because it is simple to apply, as compared to pseudo observables. However, one should be very careful in using either additional parameters or pseudo observables as the reliability of the network will be decreased.

To maintain the network reliability, the estimated (significant) additional parameters, especially scale, can be used for correcting the observed distances during LSE process, as outlined in section 3.5.3.1 (i.e. scaled distances). By applying the estimated scale factor (Table 6.6) to scale the distances, the solution converged after one iteration, with estimated variance factor of 0.79, and both global and local tests passed. The precision and reliability measures is similar to the ordinary LSE (column 1 of Table 6.6).

3. Handling of gross error

To study the capability of gross error detection, 6 gross errors with the magnitudes of 10σ were introduced into the randomized data (see Table 6.7, column 1). Four cases were examined: ordinary LSE; robustified LSE; Pope's and; Baarda's methods.

The computed solutions are displayed in Table 6.7 (column 3). Ordinary LSE is unacceptable as both global and local tests are failed. With robustified LSE, using equation (3.59b) to speed up the computation, the solutions converged after 13 iterations, with the estimated variance factor of 0.73 and weighting factor c of 2.3. The solution passed both global and local tests. RLSE procedure deweights all 6 observations that contain gross errors. During RLSE, the weights of the 6 observations were drastically reduced to zero. Consequently, degrees of freedom for computing variance factor may be reduced by 6, i.e from 34 to 28, and the variance factor was estimated as 0.89.

The final residuals (with opposite sign) of the deweighted observations indicated that the estimated magnitudes of gross error were close to the simulated gross error (columns 1 and 3 of Table 6.7). Hence, robustified LSE is able to detect and locate all the gross errors in this network correctly. Also, the redundancy numbers of the deweighted observations are close to unity showing the effects of gross error in the solution as being negligible.

In Pope's and Baarda's methods, the sequence of LSE with global and local tests followed by successive elimination of suspected observations were repeated until both global and local tests passed. These methods both resulted in variance factors of 0.89, the same as from robustified LSE. As shown in Table 6.7, all three methods detect the gross errors correctly.

	ordinary LSE	robustified LSE	Baarda's method	Pope's method
simulated ge sd 1-2 +50 mm sd 1-5 -50 mm dir 1-3 -50 sec dir 2-1 +50 sec dh 2-3 +50 mm dh 4-6 -50 mm		residuals **[-43.71] **[+40.52] **[+44.90] **[-47.21] **[-54.50] **[+48.21]	 * * * * * *	 * * * * * *
rank analysis	ok	ok	ok	ok
variance factor	10.68	0.89	0.89	0.89
iteration	1	13	1@run	1@run
global test	fail	pass	pass	pass
local test	fail(1)	pass	pass	pass
trace x axis y axis z axis	0.0001 0.0001 0.0001	0.0001 0.0001 0.0001	0.0001 0.0001 0.0001	0.0001 0.0001 0.0001
redundancy number	0.32-0.78	0.19-0.78	0.19-0.78	0.19-0.78
internal reliability	23.33-36.53	23.33-47.12	23.33-47.12	23.33-47.12
external reliability	4.77-36.38	4.79-71.83	4.79-71.83	4.79-71.83

Table 6.7 Test with 6 gross errors for network 1

(note: * indicates deleted observations (one at a time) by Pope's Tau and Baarda's methods
** indicates deweighted observations during robustified LSE)

Comparison between the solutions via robustified LSE and Pope's method also revealed that the final coordinates, their standard errors and trace were exactly the same in both cases. Moreover, residuals and reliability measures for non-deweighted observations are also identical. This result shows that by adopting robustified LSE, the solution can be used directly without the need to eliminate the observations.

The method of robustified LSE is also used to detect gross error in the data shown in Table 6.8, based on Brownlee (section 13.12) (1965). The equation for these data (Daniel and Wood, 1980) is in the form Ax=b where x represents the unknowns, b the misclosure vector and A the design matrix

$$a_1x_1+a_2x_2+a_3x_3+a_4x_4=b$$

The values of A (21,4) and b (21,1) are given in Table 6.8. These data were examined by many researchers (Daniel and Wood, 1980; Gao et al, 1992; Schwarz and Kok, 1993). Initially, the weight matrix was taken as I. The solution using RLSE converged after 12 iterations, and in the final computations, the weights of the de-weighted observations were reduced close to zero (1e-10). A total of 5 observations were deweighted as shown in Table 6.9 (observations 1, 3, 4, 13 and 21 with *), and the estimated variance factor was 1.05. As 5 observations are deweighted, the number of degrees of freedom was reduced from 17 to 12.

observation number	vector b (21,1)	design matrix A (21,4)
1	42	1 80 27 89
2	37	1 80 27 88
3	37	1 75 25 90
4	28	1 62 24 87
5	18	1 62 22 87
6	18	1 62 23 87
7	19	1 62 24 93
8	20	1 62 24 93
9	15	1 58 23 87
10	14	1 58 18 80
11	14	1 58 18 89
12	13	1 58 17 88
13	11	1 58 18 82
14	12	1 58 19 93
15	8	1 50 18 89
16	7	1 50 18 86
17	8	1 50 19 72
18	8	1 50 19 79
19	9	1 50 20 80
20	15	1 56 20 82
21	15	1 70 20 91

Table 6.8 Misclosure vector and design matrix
(taken from Schwarz and Kok, 1993)

Daniel and Wood (1980) and Gao et al (1992) found that observations 1, 3, 4 and 21 contain gross errors. By applying iterated data snooping (i.e. Tau and Baarda methods), Schwarz and Kok (1993) recovered observations 1, 3, 4, 13 and 21 as containing gross errors, resulting in an estimated variance factor of 1.05. Hence, both RLSE and iterated data snooping find observations 1, 3, 4, 13 and 21 as erroneous. Results in Table 6.9 also show that the final estimated residuals and redundancy numbers are the same for robustified LSE (non-deweighted observations) and the Tau method. Moreover, both solutions passed both global and local tests. Redundancy numbers of erroneous observations are close to one showing their insignificance in relation to the final solution.

observation number	Tau method v	r	robustified LSE v	r
1*	-	-	5.92	1.00
2	0.82	0.29	0.82	0.29
3*	-	-	6.13	1.00
4*	-	-	8.30	1.00
5	-0.81	0.92	-0.81	0.92
6	-1.26	0.89	-1.26	0.89
7	-0.15	0.74	-0.15	0.74
8	0.85	0.74	0.85	0.74
9	-0.88	0.82	-0.88	0.82
10	-0.30	0.71	-0.30	0.71
11	0.54	0.78	0.54	0.78
12	-0.11	0.67	-0.11	0.67
13*	-	-	-3.11	1.00
14	-1.54	0.76	-1.54	0.76
15	1.31	0.78	1.31	0.78
16	0.03	0.85	0.03	0.85
17	-0.71	0.55	-0.71	0.55
18	-0.06	0.82	-0.06	0.82
19	0.58	0.79	0.58	0.79
20	1.69	0.91	1.69	0.91
21*	-	-	-9.32	1.00

Table 6.9 Results of Tau method and robustified LSE

6.1.4 Deformation detection

Two test networks have been used to demonstrate the detection of spatial deformation, surveying and photogrammetric networks. The significance level for testing in LSE and deformation detection was chosen as 0.05, except for the single point test, where a significance level of 0.01 was used.

The 6 station surveying network of Figure 6.2 was again used, with 54 surveying observations and 34 degrees of freedom (simulated random error of 5 mm for slope distances and height differences and 5 secs for directions). The following deformations were simulated at stations 3, 5 and 6 to generate data for a second epoch:

station	simulated deformation (m)		
	dx	dy	dz
1	-	-	-
2	-	-	-
3	-0.050	+0.100	-0.100
4	-	-	-
5	+0.010	+0.050	-
6	-	-	+0.300

Table 6.10 Simulated deformation for network 1

The simulated random error for the second epoch is 5 mm for slope distances and height differences and 7 secs for directions.

To demonstrate the applicability of the adopted method for handling gross and systematic errors as applied to deformation detection, the following errors were introduced into the data of second epochs:

(i) Epoch (2a) contains 6 gross errors of 10σ (see Table 6.7)
(ii) Epoch (2b) contains multiple scale factor of 0.99996 (see Table 6.6) for distances measured from stations 1 and 3.

LSE for each epoch was carried out, using program ESTIMATE, by minimum constraints, i.e. fixing x_1, y_1, z_1 and x_3. This was followed by deformation detection using DETECT and ROBUST.

Seven cases of LSE for the second epoch were considered:

(i) LSE using randomized data without gross error.
(ii) Robustified LSE (equation 3.59a) but suspect observations are de-weighted drastically, i.e.

scheme 1.

(iii) Robustified LSE (equation 3.59a) via ordinary de-weighting, i.e. scheme 2.

(iv) LSE using data after gross errors were eliminated.

(v) LSE using the multiple scale errors as additional parameters.

(vi) LSE using distance ratio.

(vii) LSE using multiple scale errors estimated in (v), i.e. scaled distances.

Cases (ii), (iii) and (iv) use data from epoch (2a), whilst case (v), (vi) and (vii) use data from epoch (2b). The LSE results for each epoch passed both global and local tests. The variance factor for the first epoch was estimated as 0.814 with 34 degrees of freedom. The results of the second epoch are displayed in Table 6.11.

cases	estimated variance factor	degrees of freedom	number of iteration
(i)	0.593	34	1
(ii)	0.837	28	11
(iii)	1.364	34	33
(iv)	0.837	28	1
(v)	0.603	32	2
(vi)	0.603	32	1
(vii)	0.578	34	2

Table 6.11 LSE results for second epoch (network 1)

The number of degrees of freedom for case (iv) were reduced to 28 due to the deletion of 6 observations. Case (ii) converges after 11 iterations, with a variance factor of 0.837 and degrees of freedom of 28, as 6 observations were deweighted drastically. After 33 iterations, case (iii) converges. Results from cases (ii) and (iv) were identical. For case (v), scale factors were estimated as 0.99997±0.00001 and 0.99996±0.00001 for stations 1 and 3 respectively. Results of cases (v) and (vi) were very close. In case (vii), the multiple scale errors estimated in (v) are used for scaling the observed distances during LSE.

As both epochs use the same stations and datum, deformation detection can be proceeded straight away. LSE results for each epoch passed the test on variance ratio, indicating the compatibility of the variance factor of each epoch. An initial run of DETECT lead to the failure of global congruency test, confirming the existence of deformation.

Initially, all stations were used to define the datum. Starting with 6 datum stations, the successive process of removing suspected points from the datum was repeated until a (partial) congruency test passed. This resulted in 3 datum stations (1, 2 and 4) for the final computations. All the datum points passed the single point test and were confirmed as stable. Points 3, 5 and 6 failed the single point test and were suspected as significantly deformed. The estimated deformation obtained via DETECT for each case was very close to the simulated deformation (Table 6.10), as summarized in Table 6.12.

In ROBUST, appropriate weightings were applied followed by weighted S-transformations, until the solution convergenced. Results from ROBUST are shown in Table 6.13. For cases (i) to (iv), 3 stations (1, 2 and 4) were found to be stable and passed the single point test. Significant deformations were detected at stations 3, 5 and 6. However, for cases (v), (vi) and (vii), only two stations (1 and 2) were found as stable, indicating the effect of weighting (section 4.2.3.4) on the solution. Consequently, the results of ROBUST were rejected.

For further verification, COMPS was used to transform the LSE results of each epoch (for case (i) only) with respect to a new datum defined by stations 1, 2 and 4. The coordinate differences between epochs were computed using DETECT, and the results are the same as solution (i) in Table 6.12. Solution (i) is the expected result for acceptable data (Table 6.12). Solutions (ii) and (iv) are identical, showing that the effect of gross error in robustified LSE with scheme 1 for deformation detection is insignificant. Solution (iii) differs slightly from the other solutions, demonstrating the small effect of scheme 2. Solutions (v), (vi) and (vii) were similar and very close to solution (i), demonstrating that the effect of scale error is almost negligible.

The graphical presentation of solution (i) is depicted in Figures 6.13 to 6.16, and clearly shows that stations 3, 5 and 6 lie outside the ellipses. The deformation trend shown in Figure 6.13 indicates the movement of stations 3, 5 and 6. Figure 6.14 demonstrates movement of stations 3 and 5 in the xy directions, while Figures 6.15 and 6.16 show movement of station 6 in the z direction and station 3 in the xz and yz directions respectively.

Figure 6.17 is a plan view of a real 19 station photogrammetric monitoring network (Cooper, 1994), with seven datum defects. Deformations were simulated at stations A1 as (-0.100, 0.050, 0.030) m and B1 as (-0.100, 0.100, 0.100) m to generate data for a second epoch.

station	solution (i) dx	dy	dz	solution (ii) dx	dy	dz
1	-	-	-	-	-	-0.001
2	0.001	-	-	-	-0.001	-
4	-0.001	-	-	-	-	0.001
6	0.001	-0.001	0.300	-0.001	-	0.298
3	-0.050	0.100	-0.100	-0.050	0.100	-0.101
5	0.012	0.051	-	0.008	0.050	-0.001

station	solution (iii) dx	dy	dz	solution (iv) dx	dy	dz
1	-	-0.002	-0.001	-	-	-0.001
2	0.001	0.001	-0.001	-	-0.001	-
4	-0.001	-	-0.001	-	-	0.001
6	-0.002	-	0.297	-0.001	-	0.298
3	0.051	0.101	-0.100	-0.050	0.100	-0.101
5	0.006	0.050	-0.001	0.008	0.050	-0.001

station	solution (v) dx	dy	dz	solution (vi) dx	dy	dz
1	0.003	0.001	-	0.003	0.001	-
2	-0.001	-	-	-0.001	-	-
4	-0.002	-0.001	-	-0.002	-0.001	-
6	-0.002	-0.001	0.300	-0.002	-0.001	0.300
3	-0.051	0.102	-0.100	-0.051	0.101	-0.100
5	0.009	0.050	-	0.009	0.050	-

station	solution (vii) dx	dy	dz
1	0.004	0.002	-
2	-0.001	-0.001	-
4	-0.003	-0.001	-
6	-0.002	-0.001	0.300
3	-0.052	0.100	-0.100
5	0.009	0.050	-

Table 6.12 Estimated deformation for network 1 via DETECT (Datum stations comprised of stations 1, 2 and 4. Unit m)

station	solution (i) dx	dy	dz	solution (ii) dx	dy	dz
1	-	-	-	-	-	-0.001
2	0.001	-	-	-	-0.001	-
4	-0.001	-	-	-	-	0.001
6	0.001	-0.001	0.300	-0.001	-	0.298
3	-0.050	0.100	-0.100	-0.050	0.100	-0.101
5	0.012	0.051	-	0.008	0.050	-0.001

station	solution (iii) dx	dy	dz	solution (iv) dx	dy	dz
1	-	-0.002	-0.001	-	-	-0.001
2	0.001	0.001	-0.001	-	-0.001	-
4	0.001	0.001	-0.001	-	-	0.001
6	-0.002	-	0.297	-0.001	-	0.298
3	-0.051	0.101	-0.100	-0.050	0.100	-0.101
5	0.006	0.050	-0.001	0.008	0.050	-0.001

station	solution (v) dx	dy	dz	solution (vi) dx	dy	dz
1	0.001	0.001	-	0.001	0.001	-
2	0.007	-0.004	-	0.007	-0.004	-
4	-	-0.020	-	-	-0.020	-
6	-0.013	-0.007	0.300	-0.013	-0.007	0.300
3	-0.044	0.089	-0.100	-0.044	0.089	-0.100
5	0.002	0.032	-	0.002	0.032	-

station	solution (vii) dx	dy	dz
1	0.002	0.001	-
2	0.009	-0.005	-
4	-0.001	-0.023	-
6	-0.014	-0.007	0.300
3	-0.043	0.085	-0.100
5	0.001	0.030	-

Table 6.13 Estimated deformation for network 1 via ROBUST (unit m)

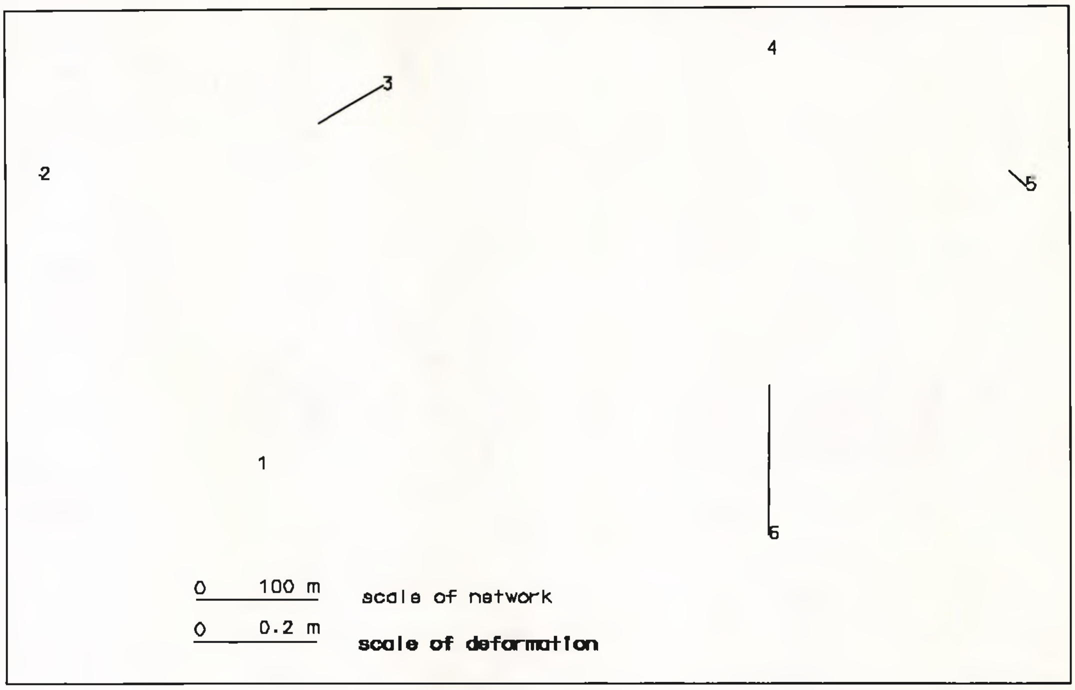

Figure 6.13 Estimated deformation for network 1 (isometric view)

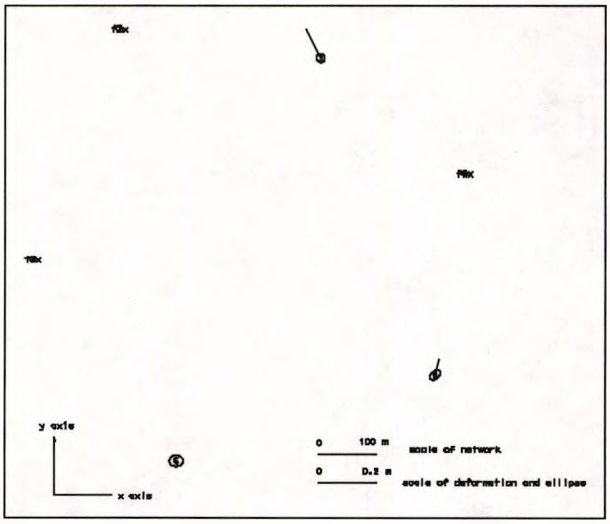

Figure 6.14 Estimated deformation for network 1 (plan view)

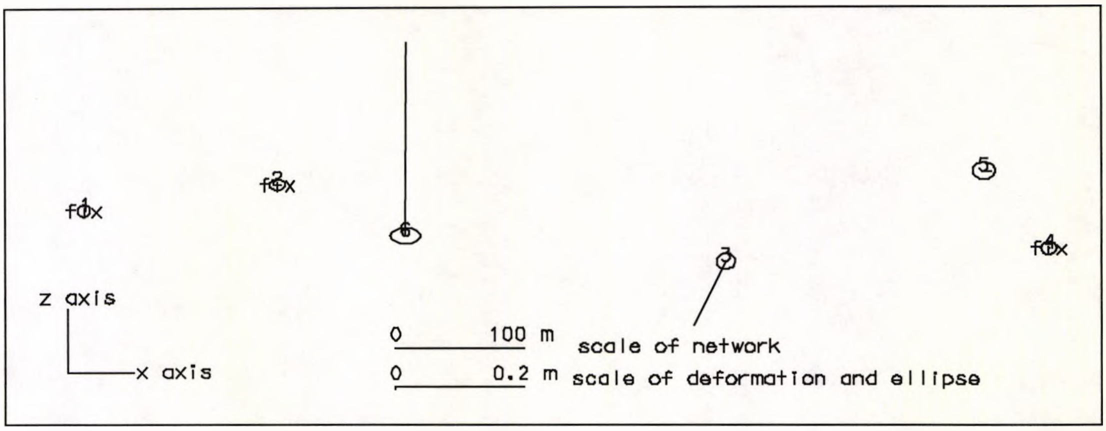

Figure 6.15 Estimated deformation for network 1 (front view)

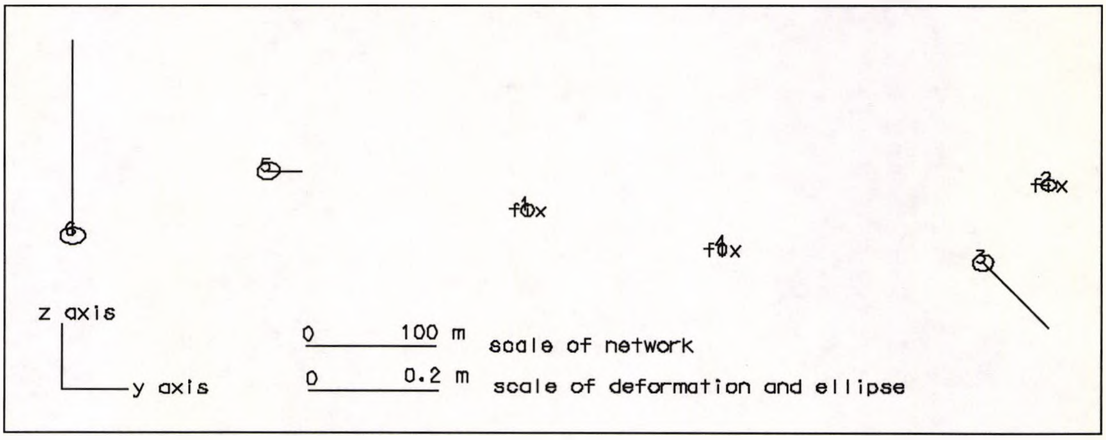

Figure 6.16 Estimated deformation for network 1 (right view)

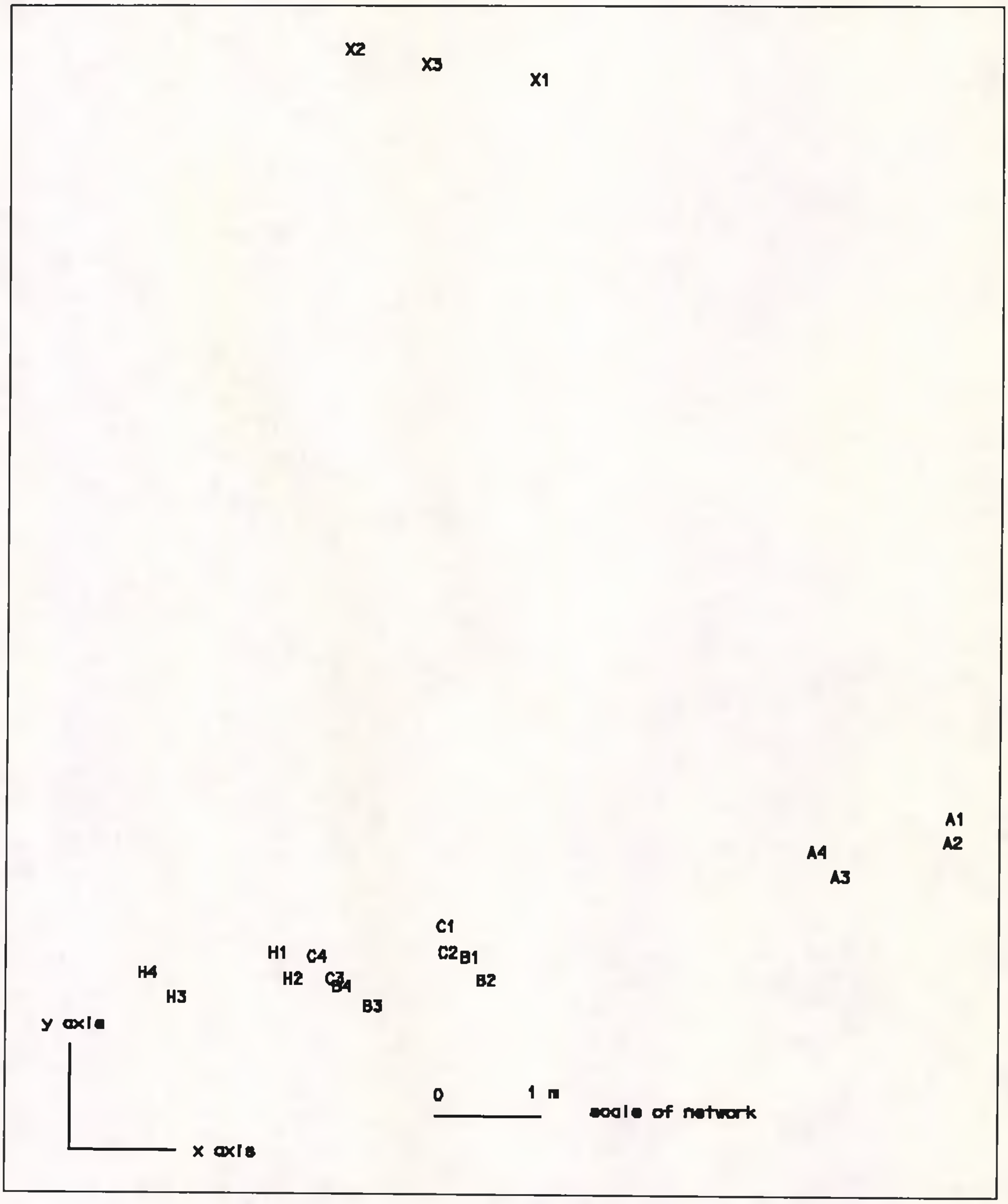

Figure 6.17 Network 3 (plan view)

In each epoch, a LSE with minimum trace solution (Cooper, 1994) was obtained using GAP. Each solution was found to pass the global test. The estimated variance factors and degrees of freedom for the first epoch were 1.1963 and 128, whilst those for the second epoch were 1.1779 and 94. Special files for deformation detection were created by GAP.

The results of deformation detection from DETECT and ROBUST are quite close as summarised in Table 6.14. Starting with 19 datum stations, DETECT resulted in 15 stable datum stations and 4 non-datum (1 stable and 3 unstable) stations. The unstable stations were found to be A1, B1 and A2. Station C1 is stable, although it is actually a non-datum point. Results of ROBUST indicated 16 stable datum stations, and detected stations A1, B1 and A2 as unstable. All the datum stations were confirmed as stable, passing the single point test. Station A2 is flagged as unstable because its computed statistic for single point test was bigger than the critical value at α=0.01. A2 has small deformation vector (0.003 m) compared to stations A1 and B1, and is significant but small. Significant deformations for stations A1 and B1 are shown below:

	station A1 (m)	station B1 (m)
simulated	[-0.100 +0.050 +0.030]	[-0.100 +0.100 +0.100]
DETECT	[-0.101 +0.048 +0.031]	[-0.099 +0.100 +0.101]
ROBUST	[-0.101 +0.048 +0.031]	[-0.099 +0.099 +0.101]

Differences between the simulated and estimated deformation are very small, and almost negligible. Graphically, Figure 6.18 shows the displacement of stations A1 and B1. The results from DETECT (congruency testing) are shown in Figures 6.19 to 6.21, where deformation vectors of stations A1, B1 and A2 are outside the ellipse. Similarly, results from ROBUST (robust method) as given by Figures 6.22 to 6.24 also indicated stations A1, B1 and A2 as significantly deformed.

6.2 Processing real data for deformation detection

In addition to the simulation tests, real data are also processed, in order to determine the capability of the adopted procedure for deformation detection. The real data consist of up to 169 stations.

	DETECT [congruency testing]			ROBUST [robust method]		
	dx	dy	dz	dx	dy	dz
A1	-0.101	0.048	0.031 *	-0.101	+0.048	+0.031 *
A4	0.000	0.000	-0.001	0.000	-0.001	0.000
C2	0.001	0.000	0.000	0.001	0.000	-0.001
C1	0.000	0.001	0.000	0.001	0.001	0.000
H3	0.000	0.000	0.000	-0.001	0.000	0.000
A2	0.001	-0.003	-0.001 *	0.001	-0.003	-0.001 *
A3	0.000	-0.001	0.000	0.000	-0.001	0.000
B1	-0.099	0.100	0.101 *	-0.099	0.099	0.101 *
B2	0.000	0.001	0.000	0.000	0.001	0.000
B3	-0.001	0.001	0.001	-0.001	0.000	0.001
B4	0.000	0.000	0.000	0.000	0.000	0.000
C3	-0.001	0.000	0.000	-0.001	0.000	0.000
C4	0.001	0.000	0.000	0.001	0.000	0.000
H1	0.001	0.000	-0.001	0.001	0.000	-0.001
H2	0.000	0.000	0.000	0.000	0.000	0.000
H4	0.000	0.000	0.000	0.000	0.000	0.000
X1	0.000	0.002	0.001	0.000	0.002	0.001
X2	0.000	-0.003	0.000	0.000	-0.003	0.000
X3	0.000	0.000	0.001	0.000	0.000	0.001
global test						
#1	pass			pass		
#2	fail					
#3	pass					
local test (single point test) (α 0.01)	pass			pass		

global test:
#1 test on variance ratio (significance level α 0.05)
#2 global congruency test (α 0.05)
#3 partial congruency test (α 0.05)

Table 6.14 Estimated deformation for network 3
(unit in m, * indicates unstable stns)

Data obtained from five real photogrammetric monitoring schemes with seven datum defects were processed in order to detect their spatial deformation: Mary Rose network, wood panel testing, A55 network A, A55 network B and deformation study of wood panel. The first three and the final scheme (Mary Rose, wood panel testing, A55 network A (September 1992-

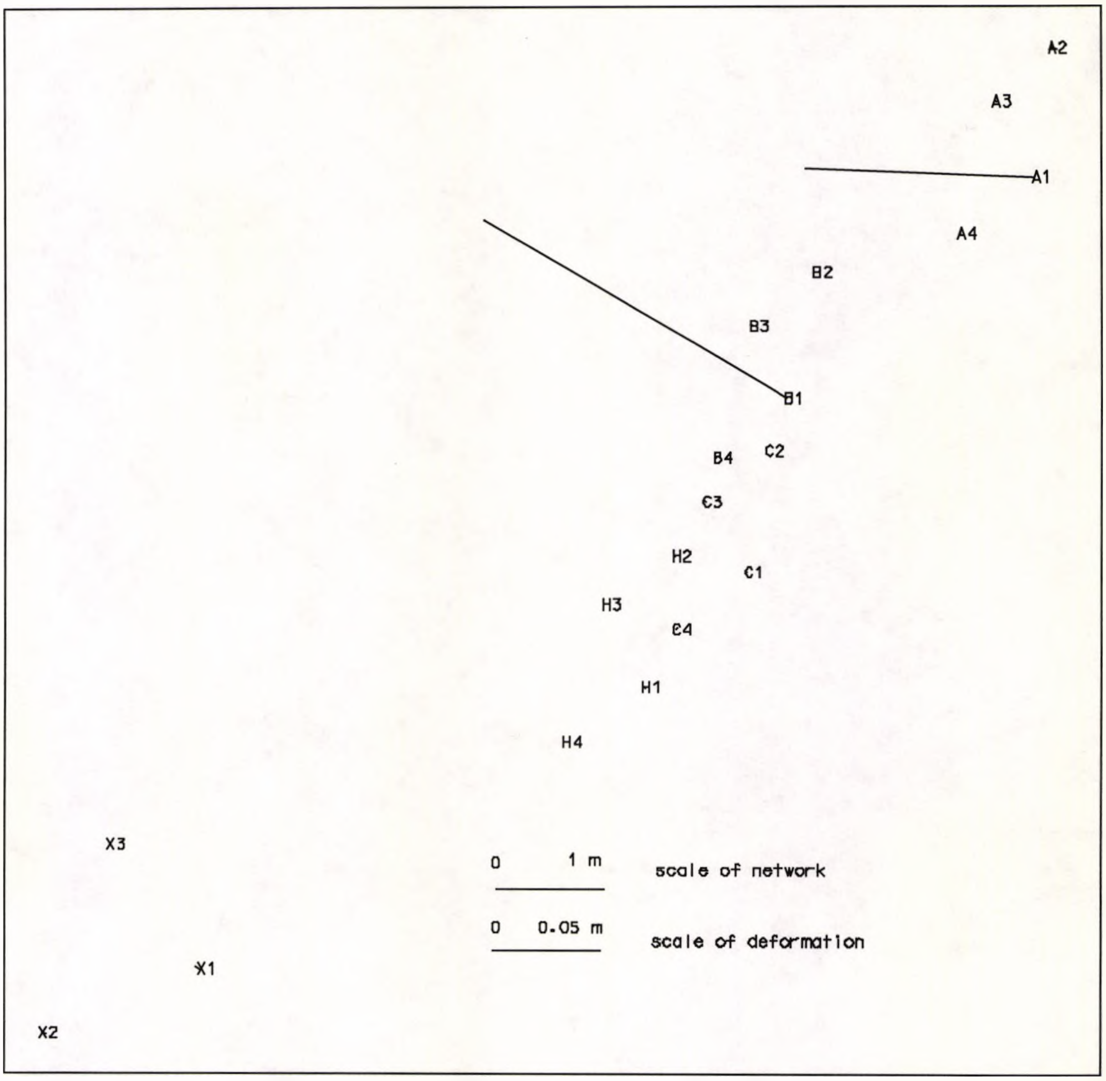

Figure 6.18 Estimated deformation for network 3 by congruency testing (isometric view)

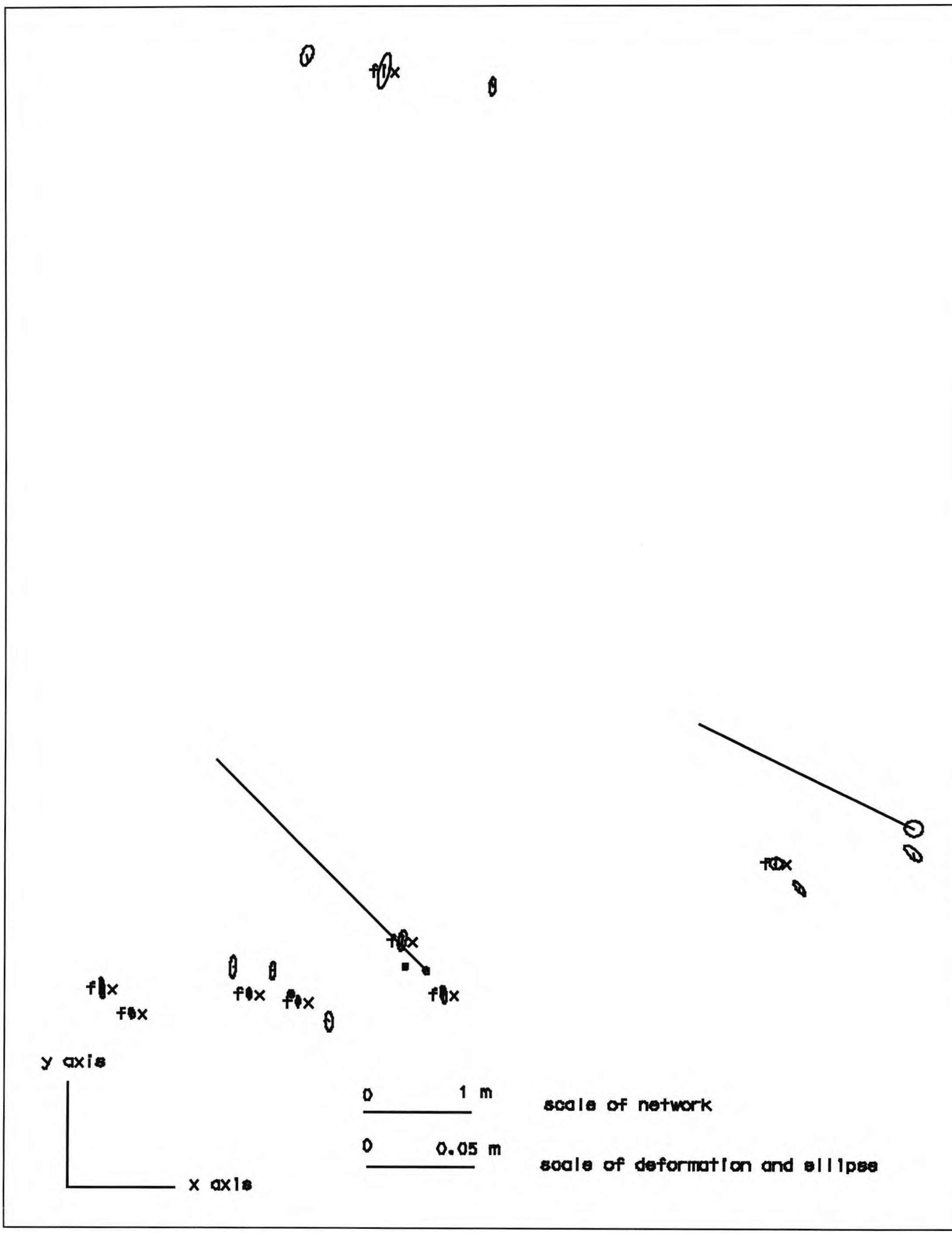

Figure 6.19 Estimated deformation of network 3 by congruency testing (plan view)

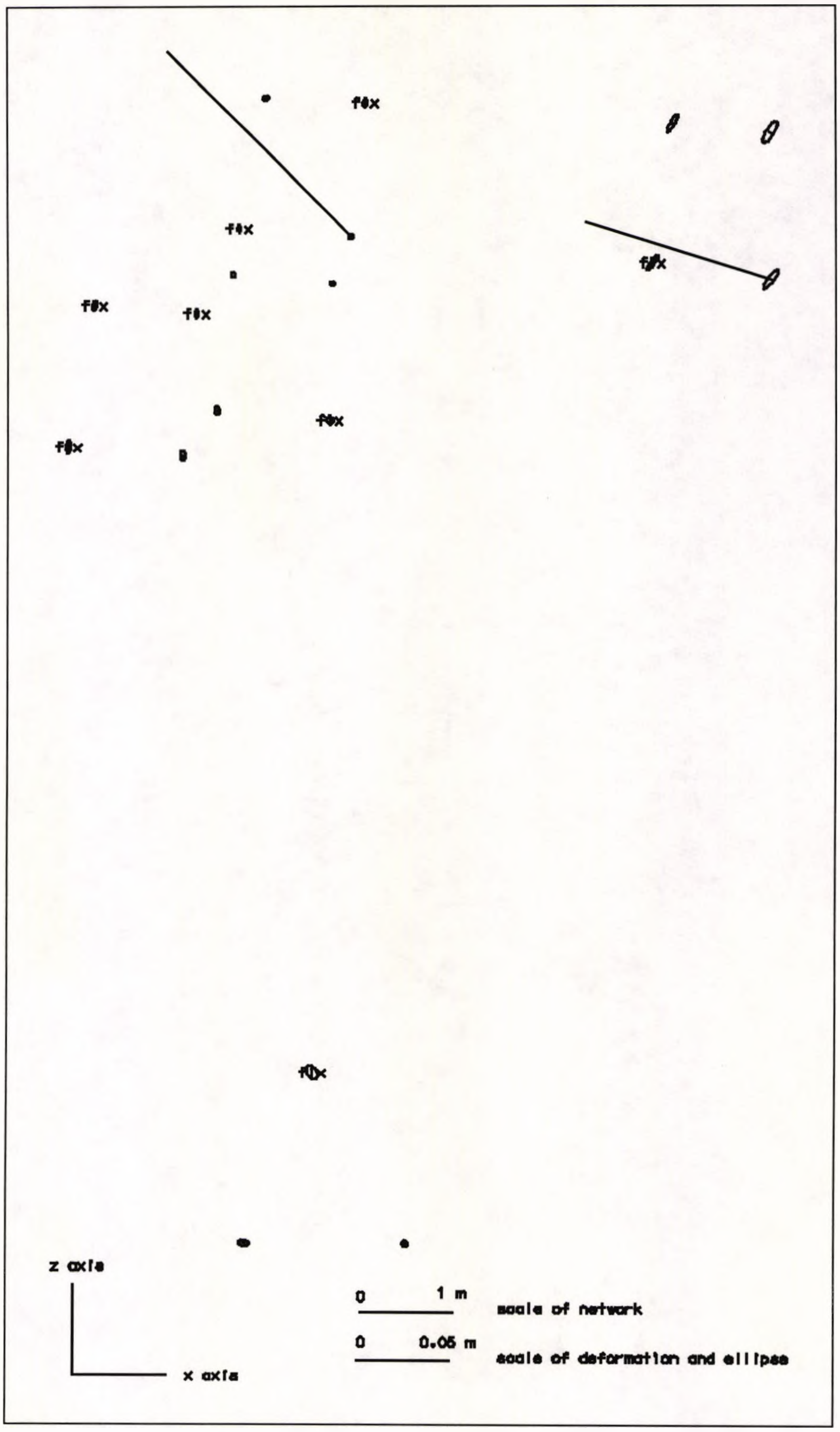

Figure 6.20 Estimated deformation for network 3 by congruency testing (front view)

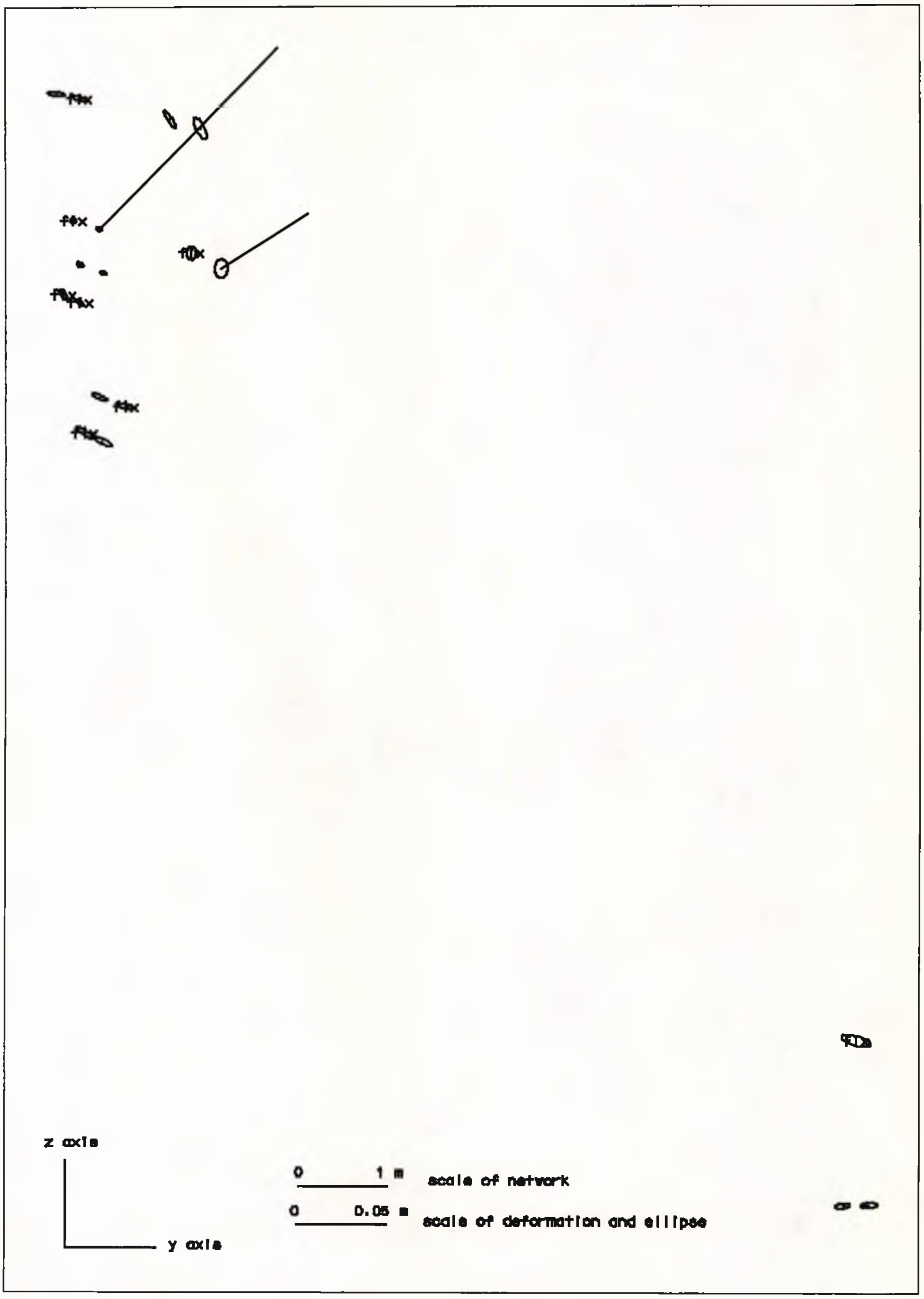

Figure 6.21 Estimated deformation for network 3 by congruency testing (right view)

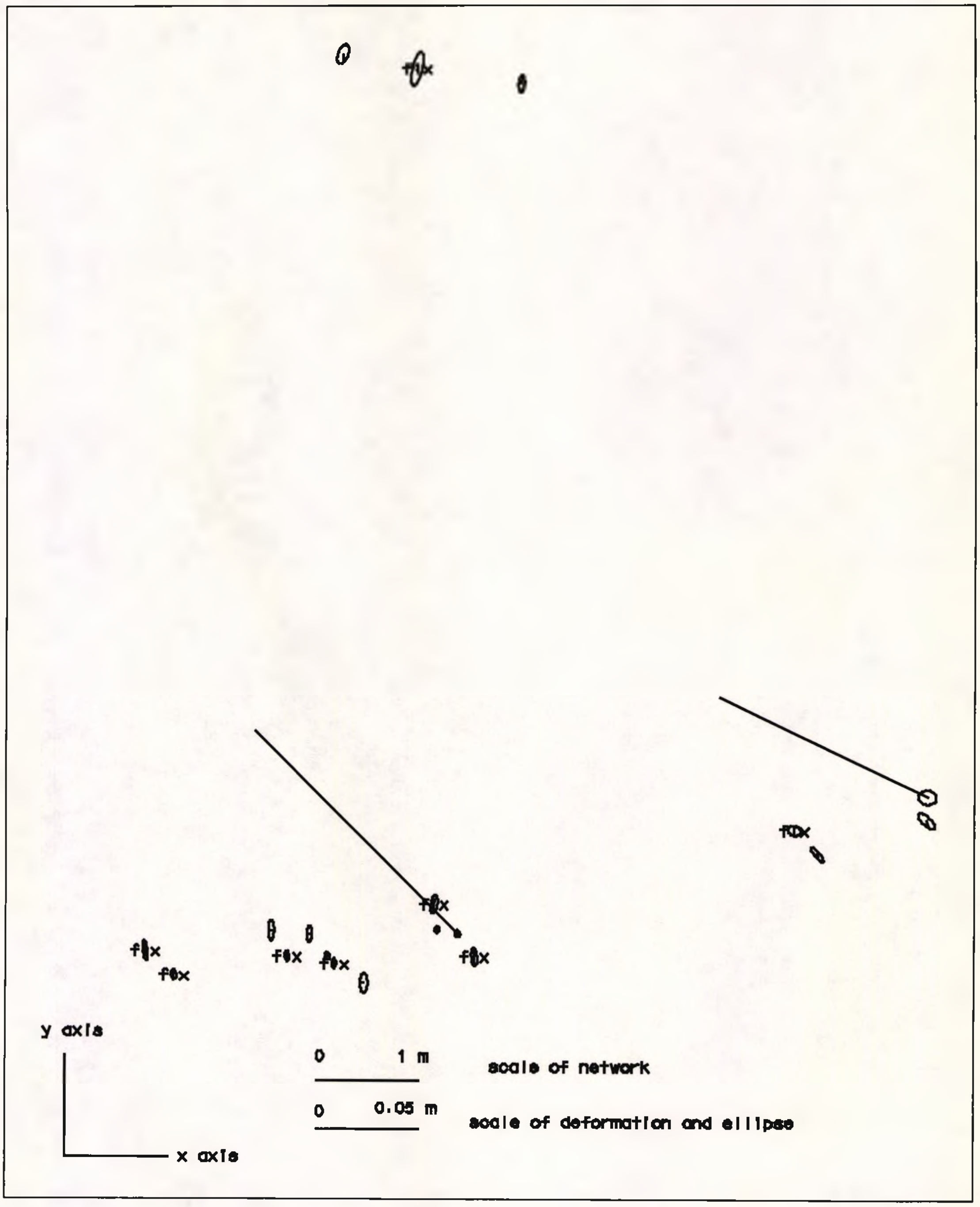

Figure 6.22 Estimated deformation for network 3 by robust method (plan view)

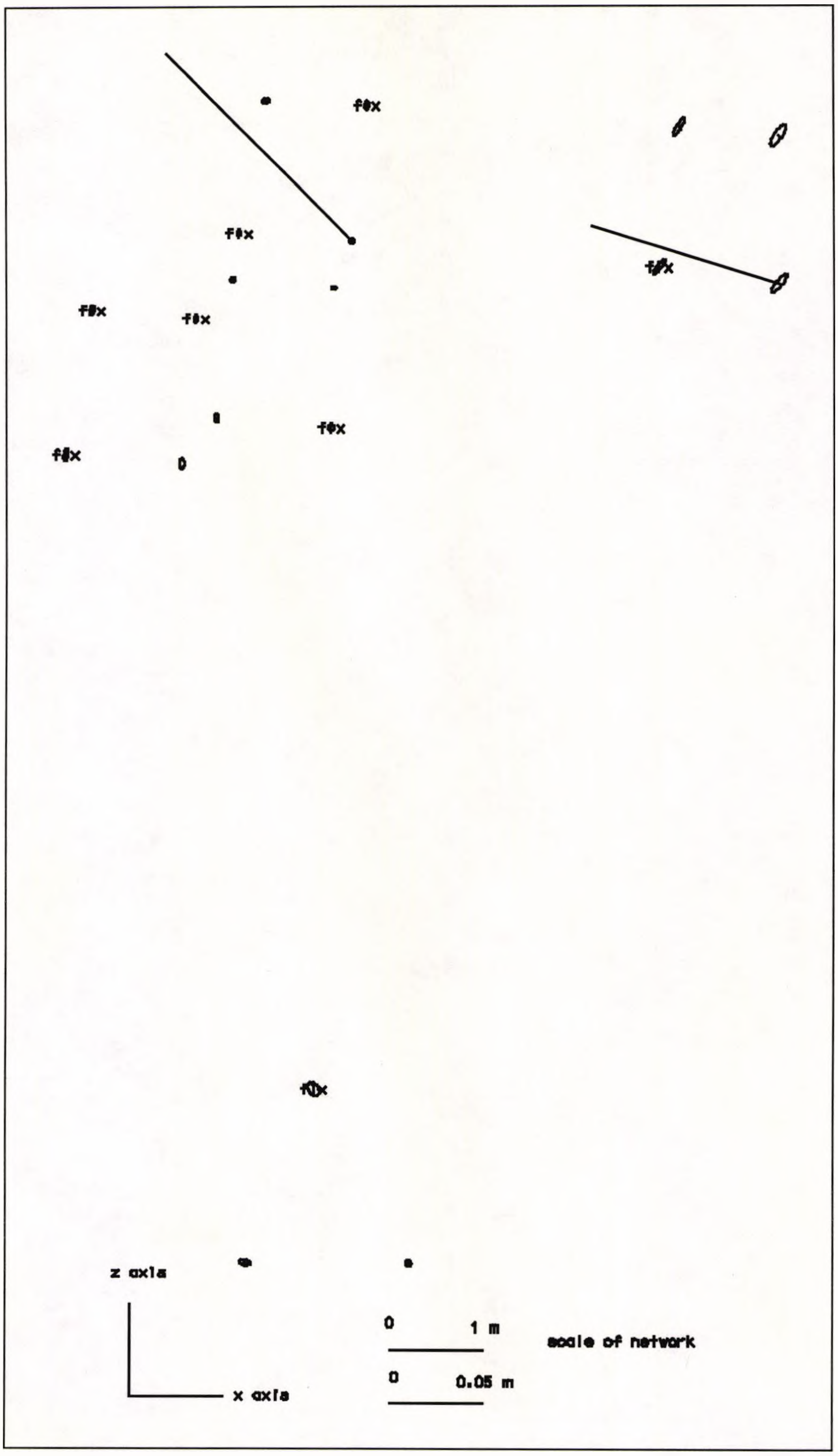

Figure 6.23 Estimated deformation for network 3 by robust method (front view)

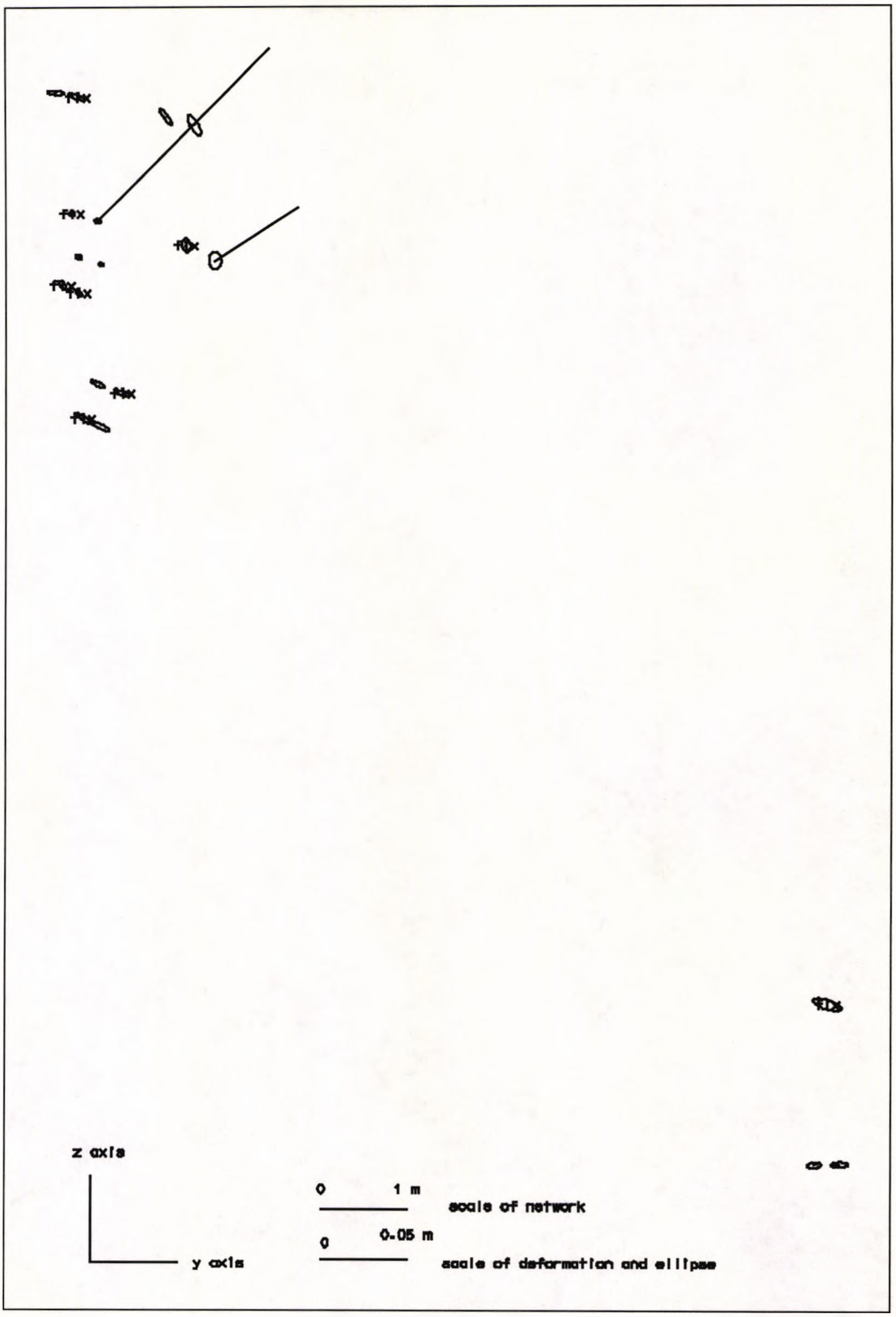

Figure 6.24 Estimated deformation for network 3 by robust method (right view)

December 1993) and deformation study of wood panel), each consisted of two epochs of data with the same number of stations in each epoch. The fourth scheme, (A55 network B (July 1991-December 1993)) had a different number of stations in each epoch.

Robustified LSE for each epoch was not carried out using ESTIMATE since it is currently limited to processing surveying data, not photogrammetric data. LSE was carried out using GAP with a minimum trace solution, because this is the only mode in which GAP will produce deformation files. However, the solution could be transformed to a minimum constraints datum using COMPS if required.

All the LSE results passed the global test and were found to be acceptable. Deformation detection for all the schemes was performed using DETECT and graphics output was obtained via DCRE. For processing the fourth scheme, program COMON was used to extract information on common stations, prior to deformation detection.

In a minimum trace solution, all stations are used to define the datum for LSE. The strategy adopted for deformation detection is composed of two steps, initial and final computations. Initially, DETECT was used to provide a set of datum points. Only datum points that passed the single point tests were used to define the datum for the final computation carried out using DETECT. In the final computation, all datum points should pass the single point test. Otherwise they must be removed from the computational base and the procedure repeated until stable datum points are found. A significance level of 0.05 was used for global testing. However, for the single point test, a significance level of 0.01 was adopted, as the computed standardized level was very close to 0.01. In the sections 6.2.1 to 6.2.5, only aspects of deformation detection will be discussed.

6.2.1 Mary Rose

The Mary Rose monitoring scheme consists of 41 stations as shown in Figure 6.25. The first and second epochs were measured in August 1991 and February 1993 respectively. In LSE, the number of degrees of freedom were found to be 110 and 174, whilst the estimated variance factors were 0.9997 and 0.9996, for the first and second epochs respectively.

Deformation detection resulted in 15 stable datum stations and 26 non-datum (13 stable

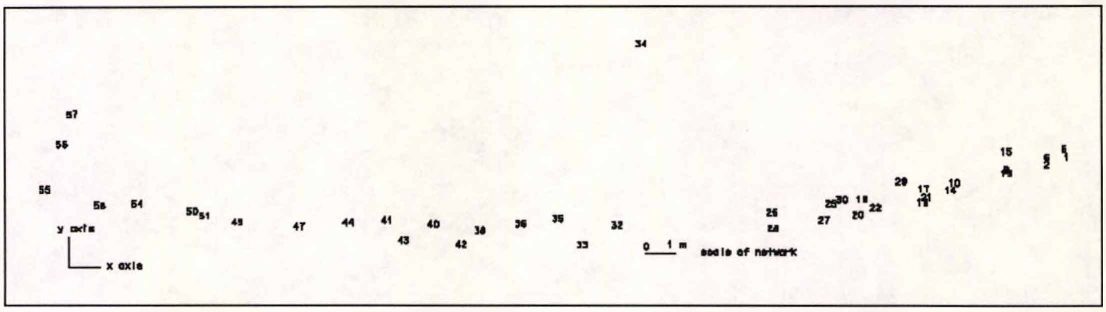

Figure 6.25 Mary Rose network (plan view)

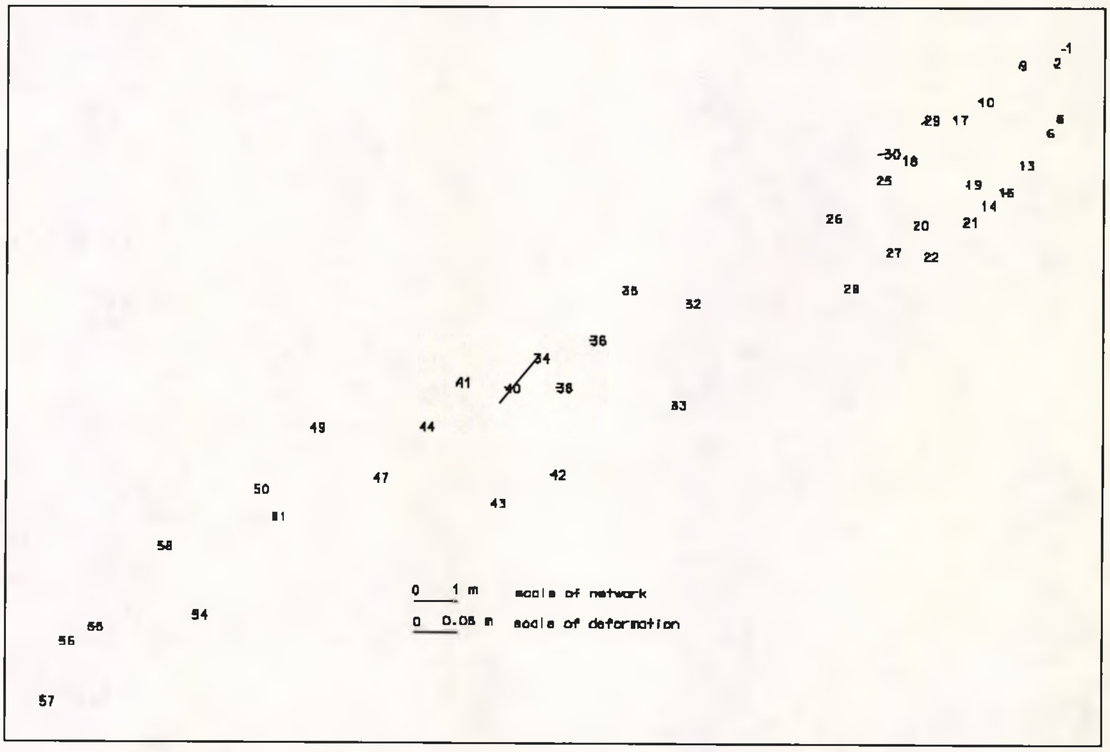

Figure 6.26 Estimated deformation for Mary Rose network (isometric view)

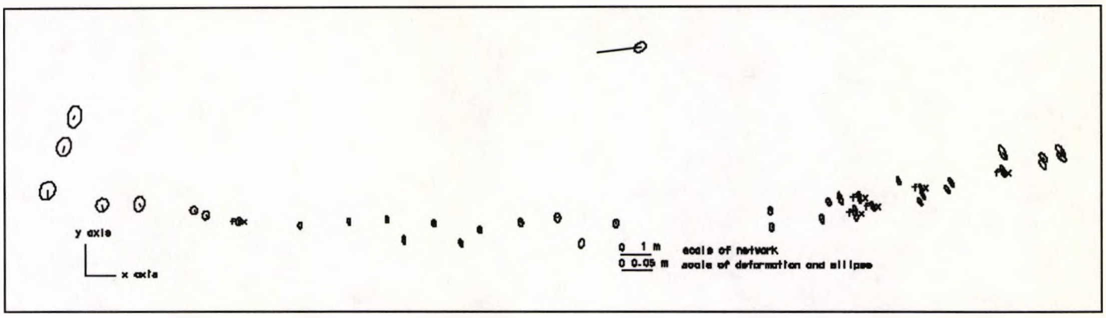

Figure 6.27 Estimated deformation for Mary Rose network (plan view)

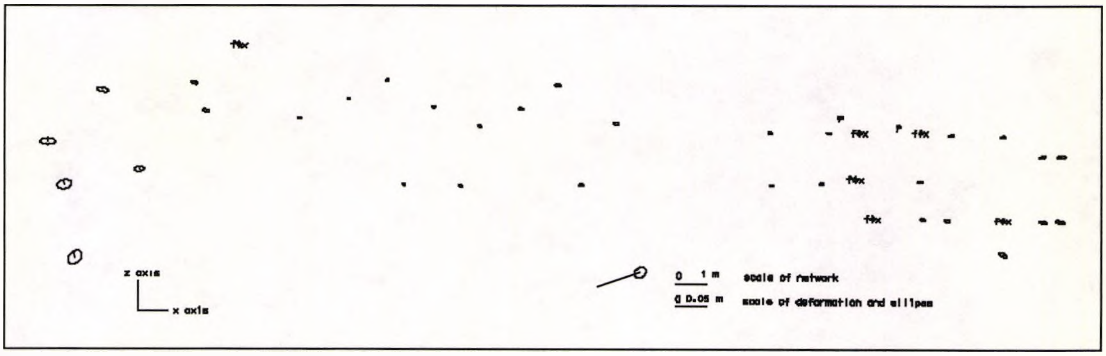

Figure 6.28 Estimated deformation for Mary Rose network (front view)

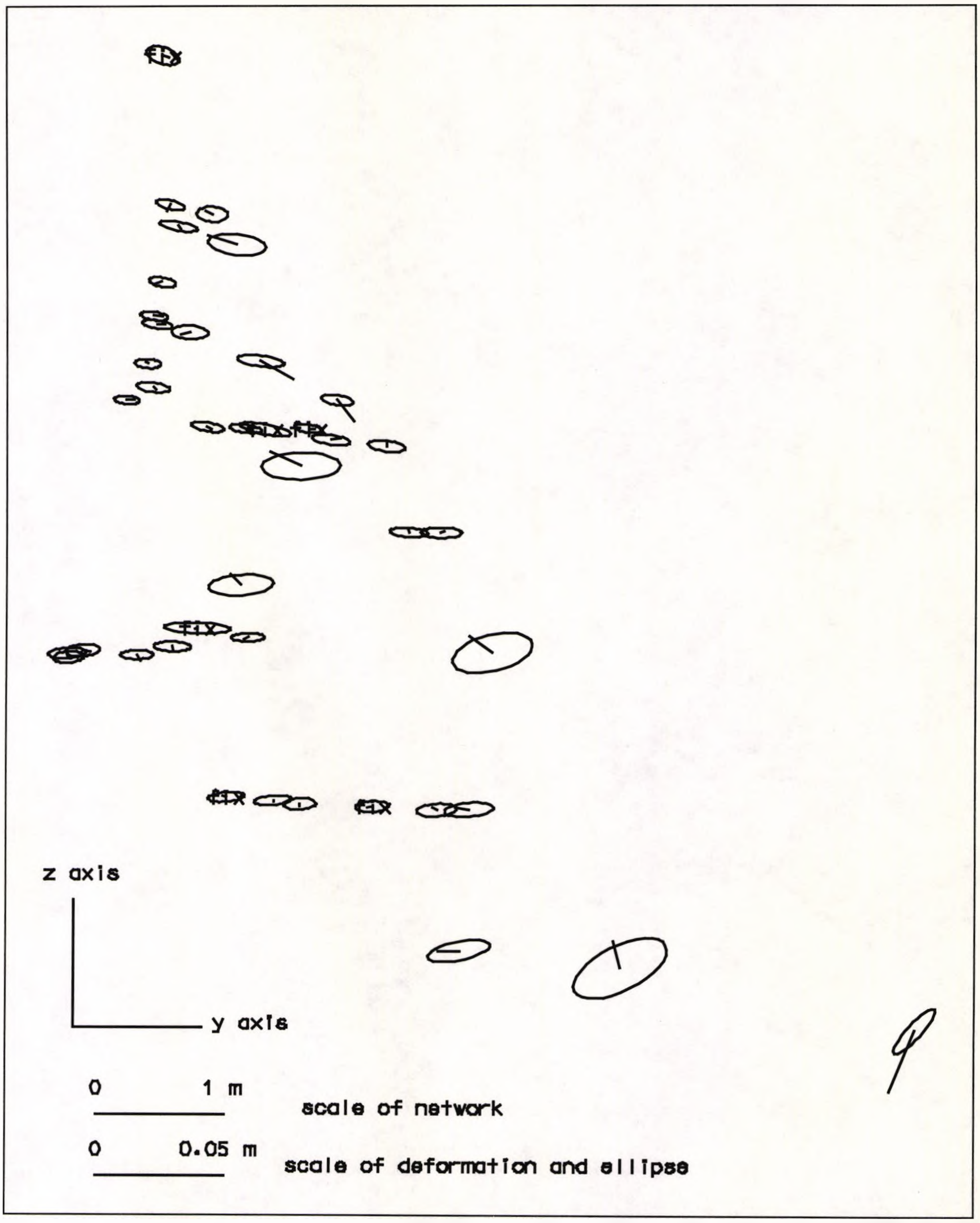

Figure 6.29 Estimated deformation for Mary Rose network (right view)

and 13 unstable) stations. The stable datum stations were: 1, 13, 17, 21, 32, 2, 6, 18, 29, 22, 26, 33, 42, 47 and 50. The results are shown graphically in Figures 6.26 to 6.29, and indicate that station 34 moved significantly. The components of the movement at station 34 were (-0.072, -0.010,-0.024) m, the displacement vector being 0.077 m. Station 34 also has maximum movement in the x and z directions. Station 30 has maximum movement in the y axis of 0.013 m. It was later confirmed by the Mary Rose Trust that point 34 was deliberately moved during construction works.

6.2.2 Wood panel testing

The wood panel testing is an experimental scheme to investigate the deformation of a wood panel brought about by changes in the relative humidity. A test wood panel with 74 pre-marked points was imaged by five digital cameras in a square-based pyramid configuration, to obtain data for the first epoch (Robson, 1994). The layout of the targetted points is shown in Figure 6.30. The test panel was about 1 m square (i.e. micro-scale application). Before the second epoch, a constraining baton had been glued to the right-hand section of the panel, and the panel was deformed by wetting it with water.

During LSE (Robson, 1994), the number of degrees of freedom for each epoch was 495, whist the estimated variance factors were 0.9989 and 1.0915 for epochs one and two respectively.

The outcome of the deformation detection resulted in 13 stable datum stations and 61 non-datum (3 stable and 58 unstable) stations. The stable datum station includes stations 100, 104, 105, 109, 110, 111, 112, 115, 116, 118, 119, 121 and 143. Most of the stable stations were in the right section of the panel, where the constraining baton was fixed (Figure 6.31). Three stations were found to have the largest overall movements especially in the z axis of up to 13 mm, i.e stations 173 (0.3, -0.5, 13.3) mm, 162 (0.3, -0.6, 9.8) mm and 167 (0.3, -0.6, 9.5) mm. Maximum movements in the x and y directions were detected at stations 124 (1.3 mm) and 168 (4.2 mm). Figures 6.31 to 6.34 show the pattern of the estimated displacements. Figure 6.31 indicates that the left section of the panel was bent upward. Overall movement in xy was detected in the upper region of the panel (Figures 6.31 and 6.32), while the most significant movement is in the z direction (figures 6.33 and 6.34). Such deformation trends were expected, since the constraining baton had been glued to the right section of the panel.

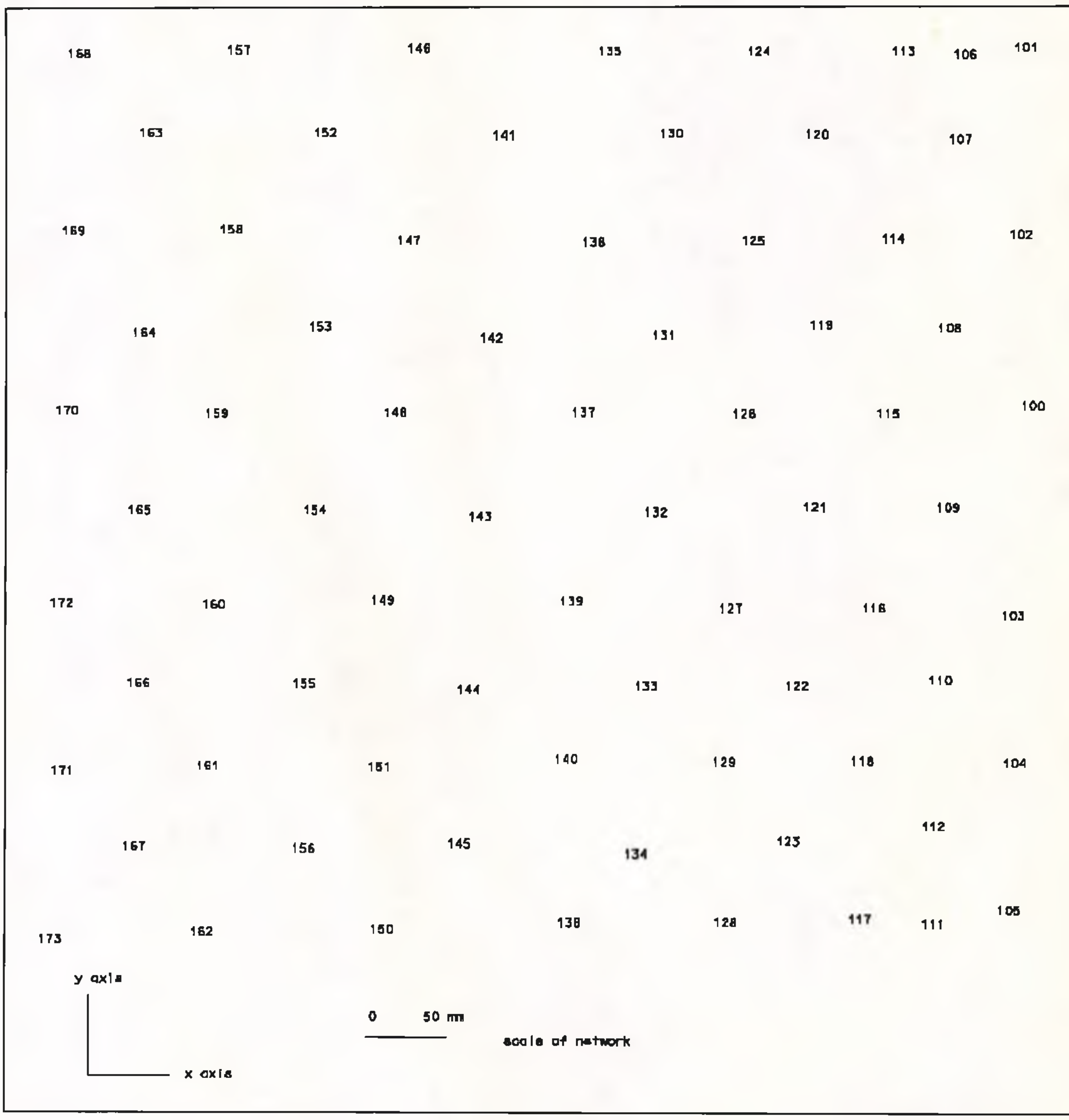

Figure 6.30 Wood panel network (plan view)

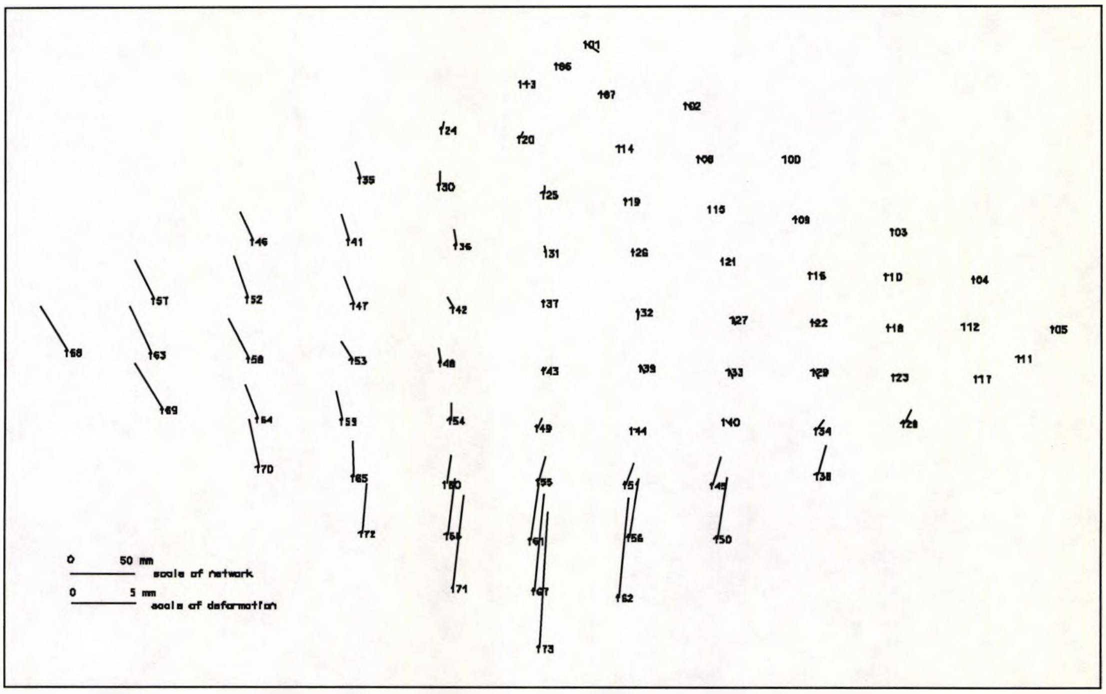

Figure 6.31 Estimated deformation for wood panel (isometric view)

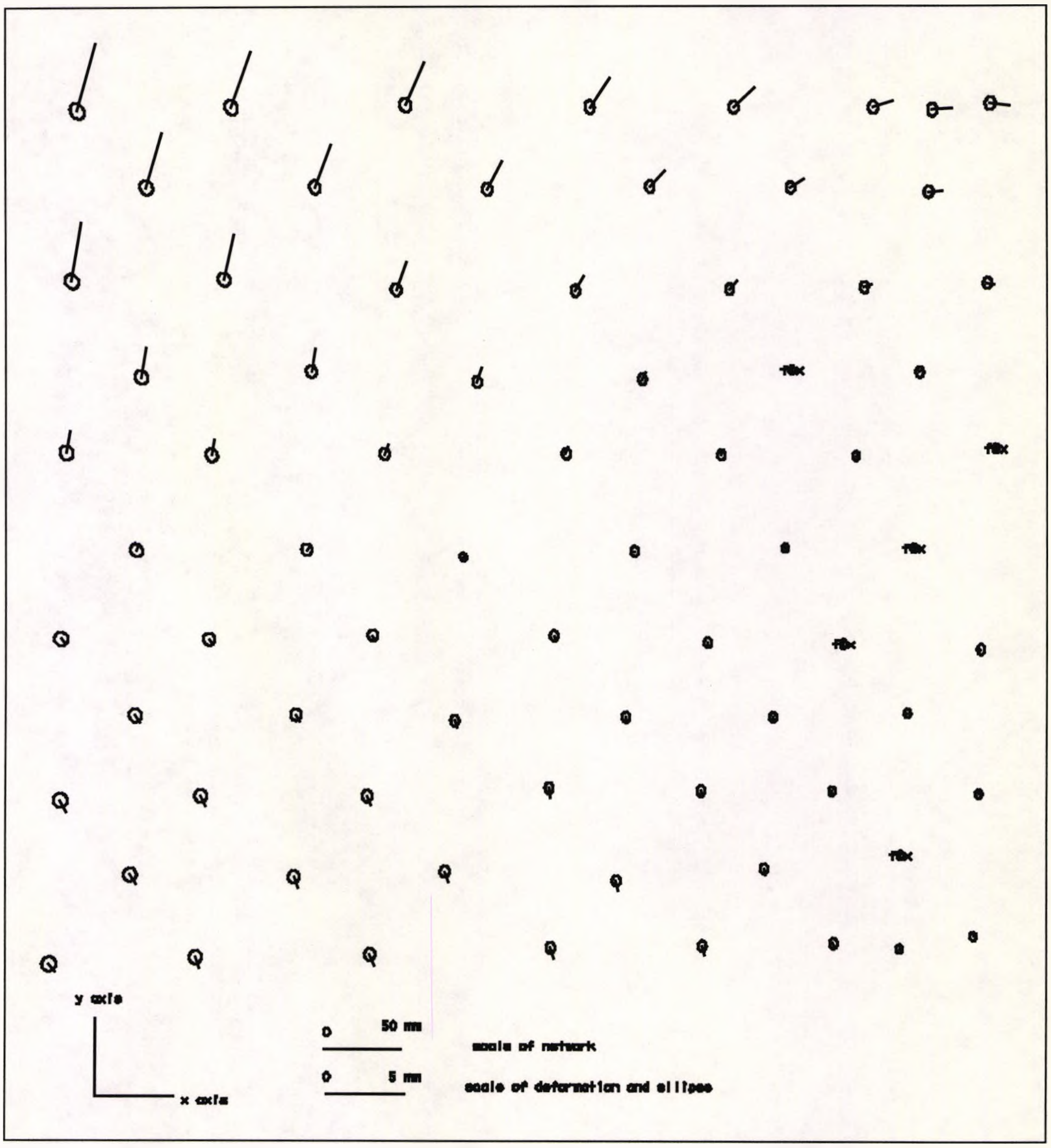

Figure 6.32 Estimated deformation for wood panel (plan view)

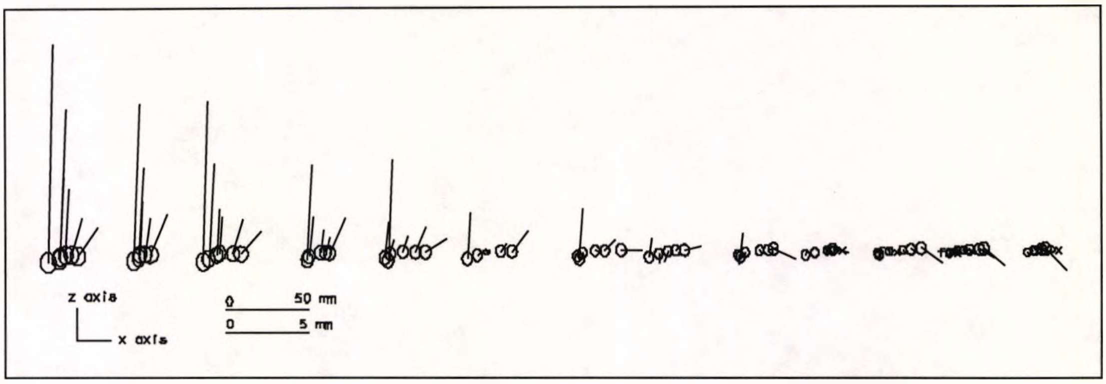

Figure 6.33 Estimated deformation for wood panel (front view)

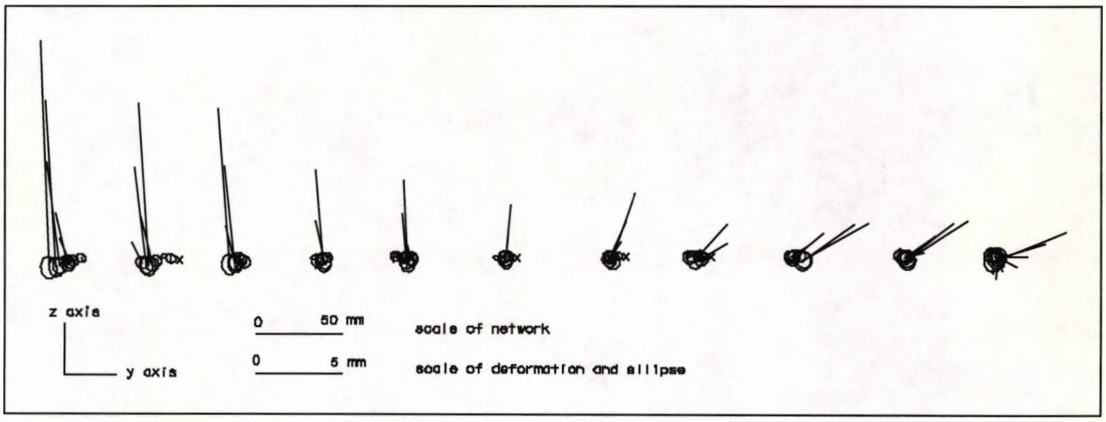

Figure 6.34 Estimated deformation for wood panel (right view)

6.2.3 A55 network A (September 1992-December 1993)

Monitoring engineering works along a part of A55 highway in North Wales is one of the activities being undertaken by ESRC. The data used here are taken from the east site, which consisted of 72 stations, as shown in Figure 6.35. The first epoch was measured in September 1992 whilst the second epoch was in December 1993. In robustified LSE (Cooper, 1994) the number of degrees of freedom for first and second epochs were 629 and 539 respectively, whilst their estimated variance factors were 1.0854 and 1.0591.

Results of deformation detection found 15 stable datum stations and 57 non-datum (16 stable, 41 unstable) stations. The stable datum stations comprise of station D32, D33, D35, E91, E92, E94, E95, F91, F94, F95, G94, G95, H94, F111 and G101. The most significant overall deformation and movement in the y direction were detected at station b2 (-0.004, -0.103, 0.003) m, giving rise to a 0.103 m displacement vector. Maximum movements in the x and z directions were found at stations k17 (0.028 m) and 68 (-0.018 m) respectively. Figure 6.36 shows that most of the points on the west region were unstable. Plots of deformation vectors and ellipses are shown in Figures 6.37 to 6.39.

6.2.4 A55 network B (July 1991-December 1993)

In this scheme, there were 70 stations in epoch one (July 1991) and 79 in epoch two (December 1993), with corresponding degrees of freedoms of 603 and 616. During LSE (Cooper, 1994) the estimated variance factors were computed as 1.0383 and 1.0501 respectively.

Program COMON was used to determine the common stations between these two epochs. A total of 60 common stations was found, as shown in Figure 6.40. After re-ordering the coordinates and cofactor matrices with respect to the common stations, S-transformations of each epoch were carried out, using common stations for datum definition. Hence, a partial minimum trace solution was computed for each epoch. The variance factors and degrees of freedom remain unchanged since they are independent of the S-transformations.

It was found that some of the provisional coordinates of the common stations were slightly different in each epoch. In DETECT, the deformation detection must be based on the same provisional coordinates for common stations. To arrive at the same provisional

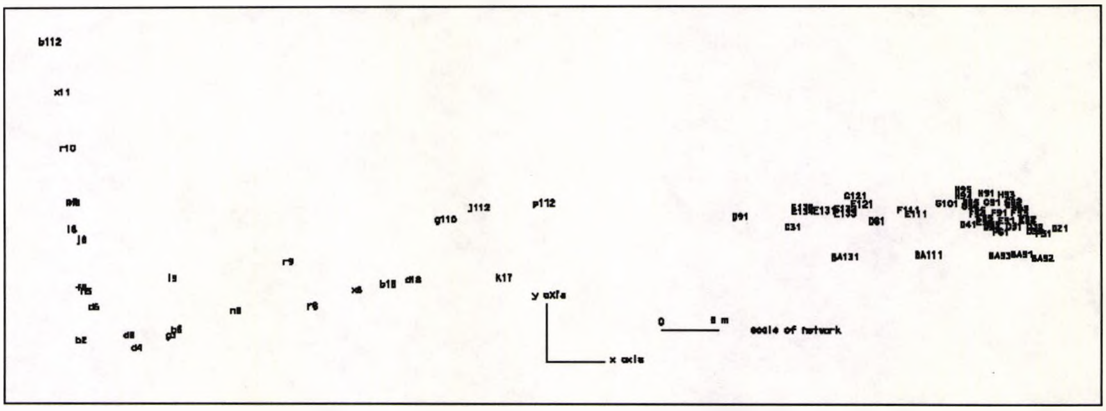

Figure 6.35 A55 network A (plan view)

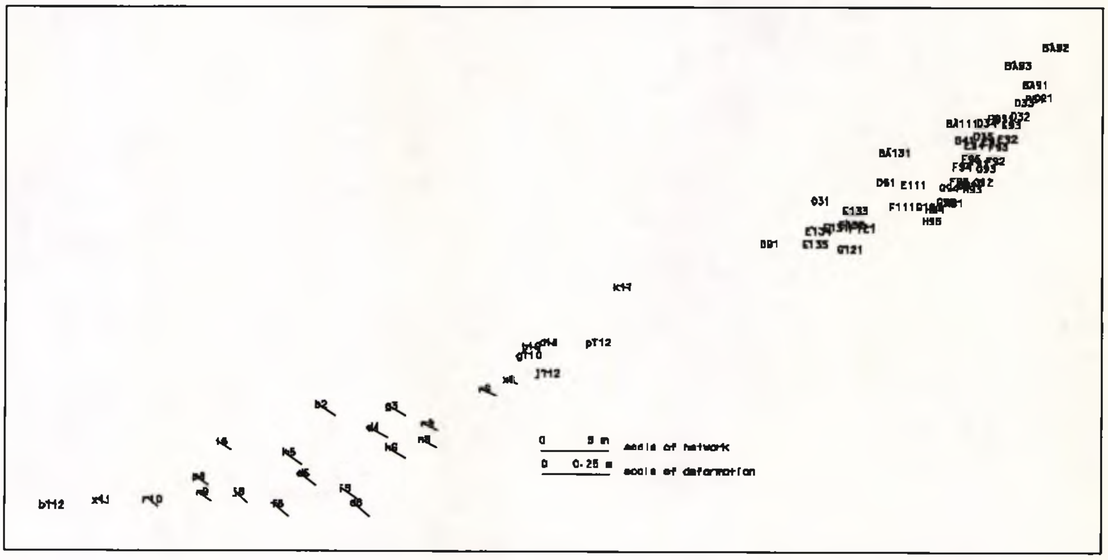

Figure 6.36 Estimated deformation for A55 network A (isometric view)

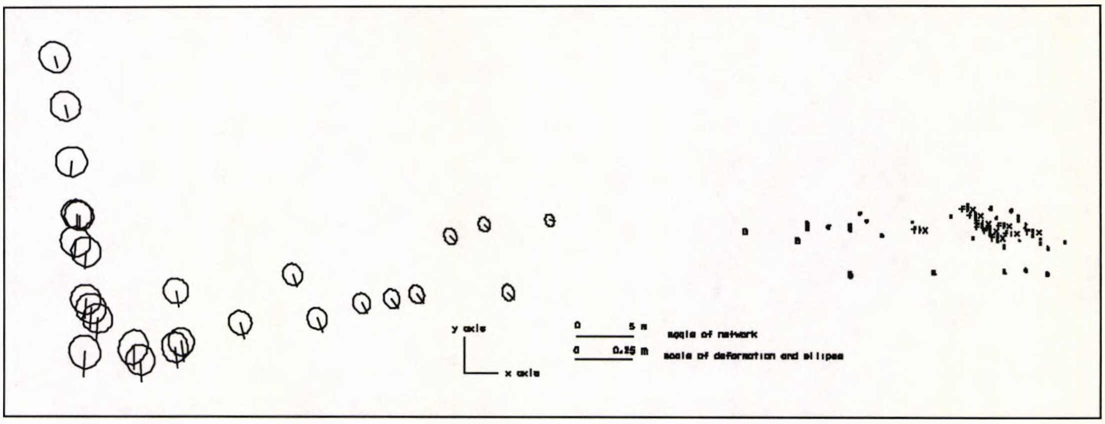

Figure 6.37 Estimated deformation for A55 network A (plan view)

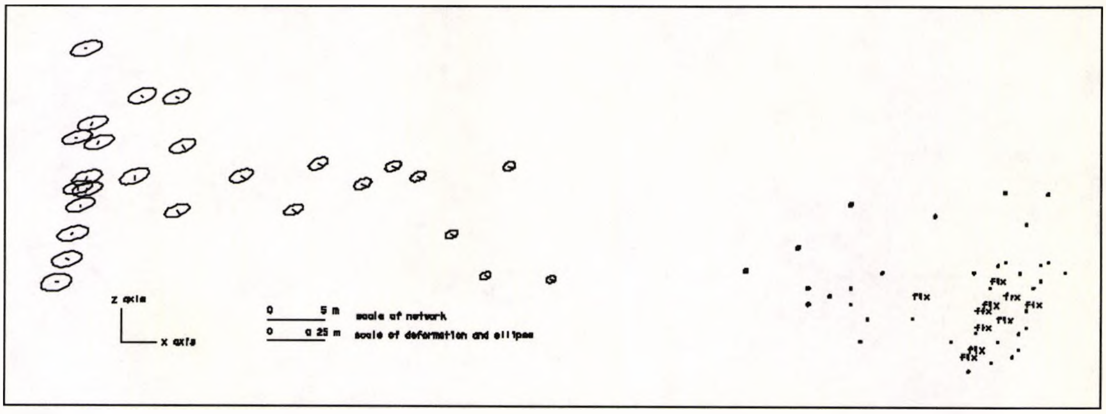

Figure 6.38 Estimated deformation for A55 network A (front view)

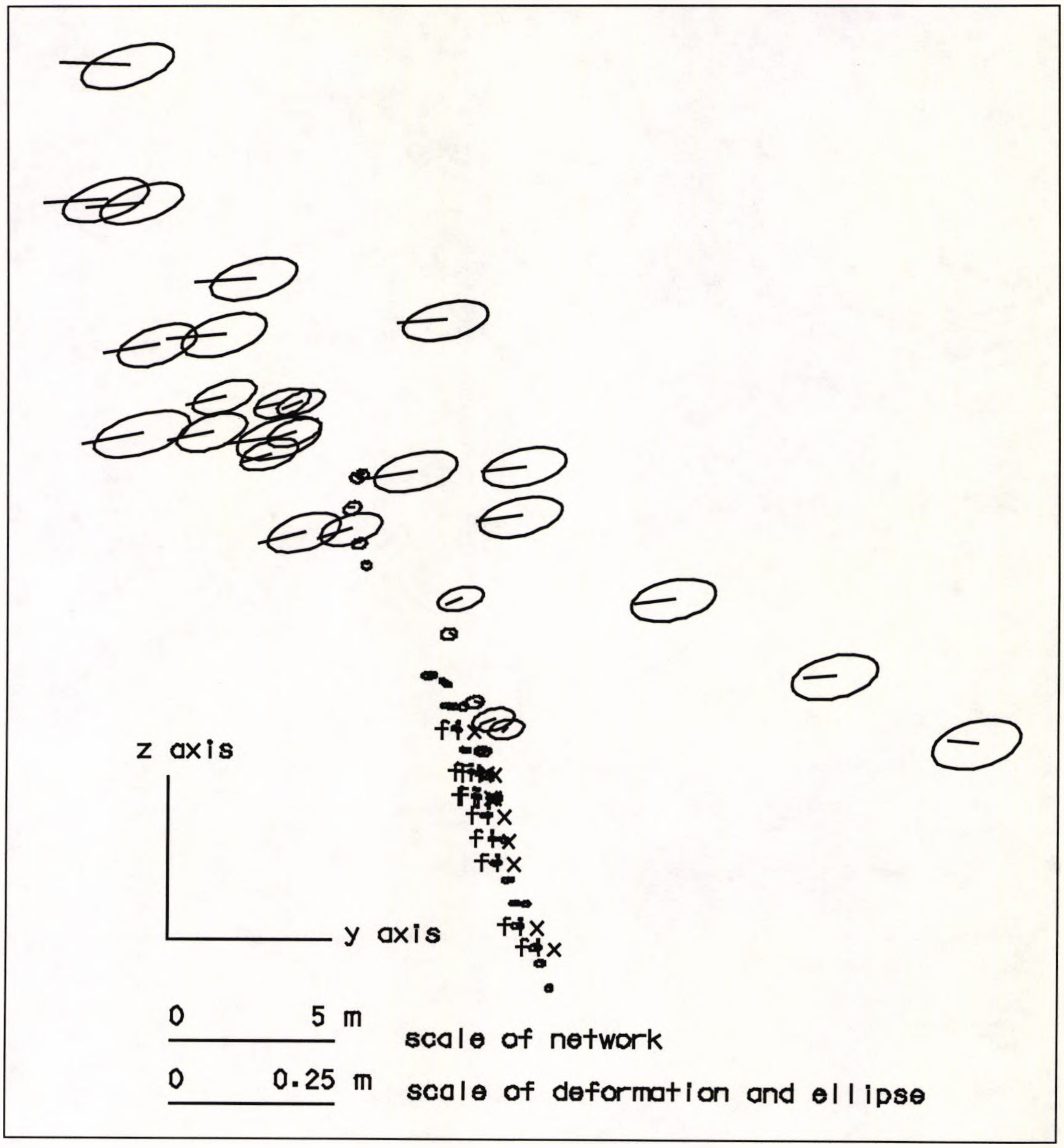

Figure 6.39 Estimated deformation for A55 network A (right view)

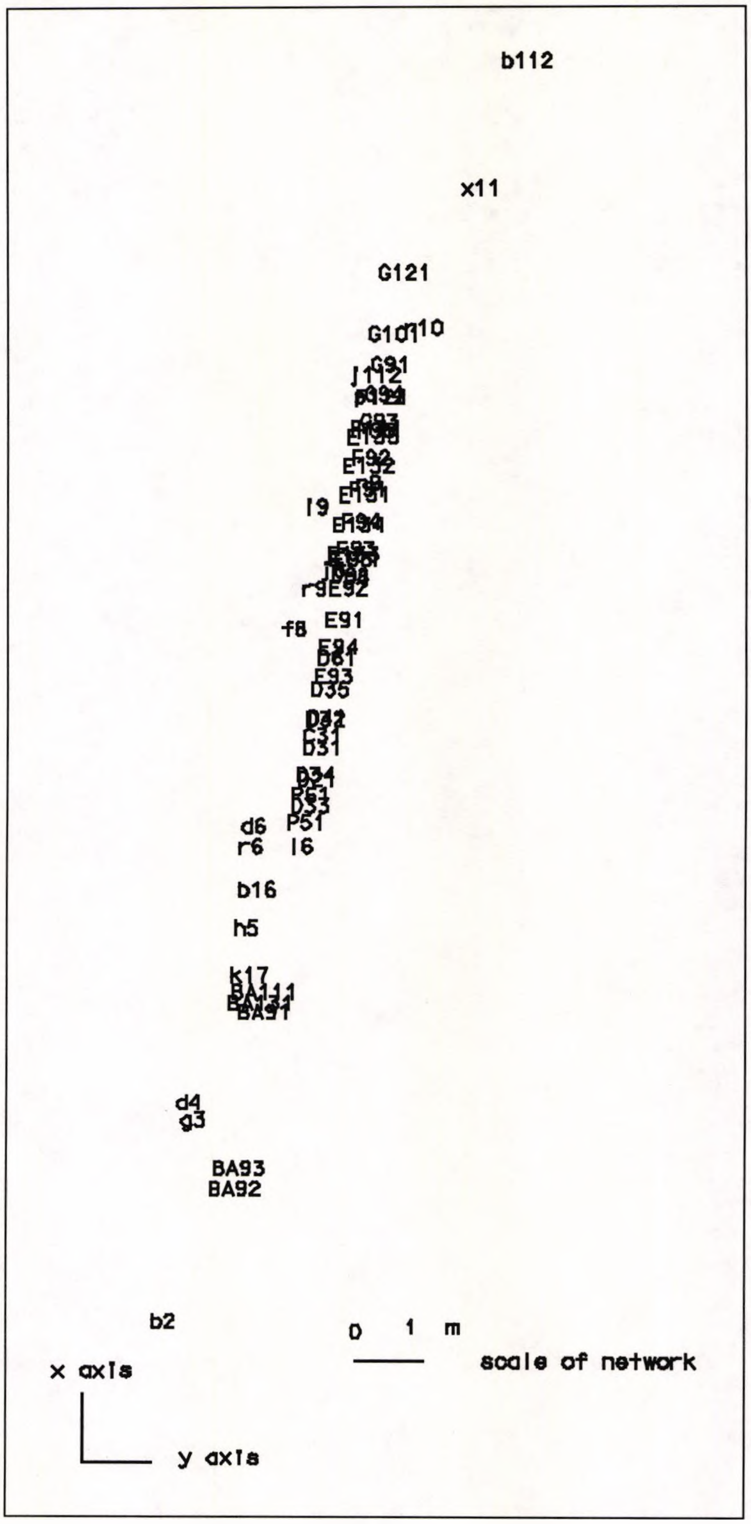

Figure 6.40 A55 network B (plan view)

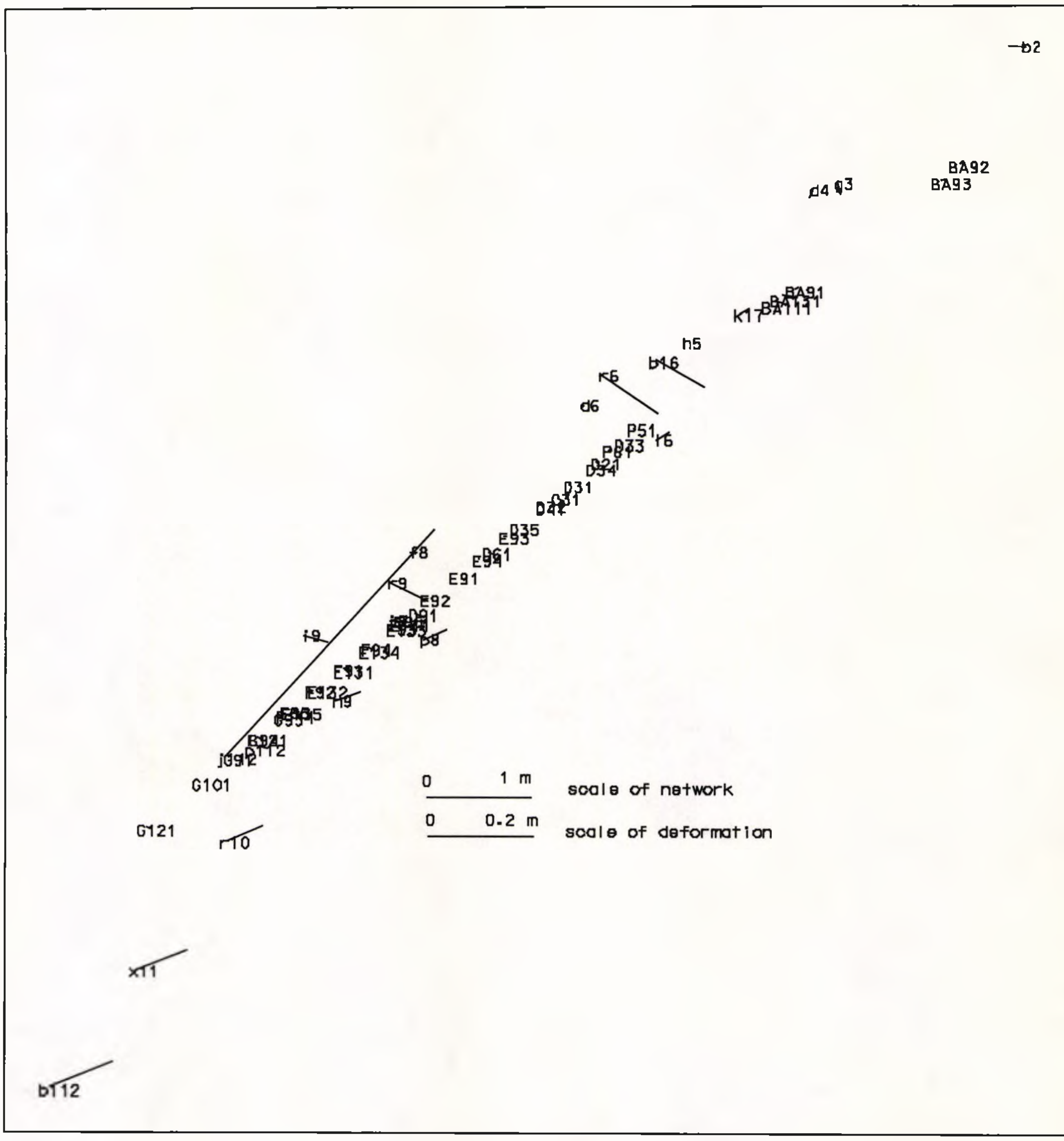

Figure 6.41 Estimated deformation for A55 network B (isometric view)

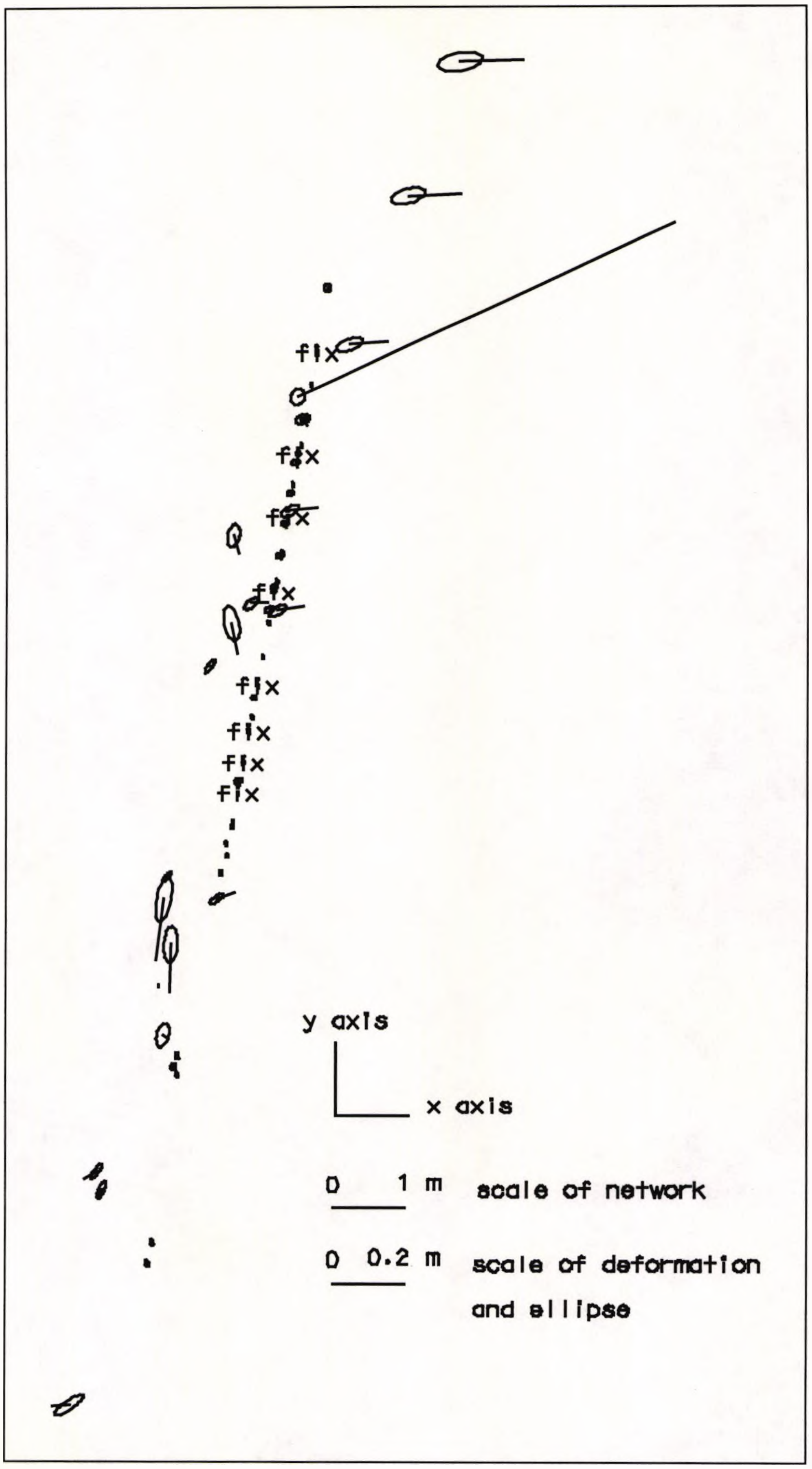

Figure 6.42 Estimated deformation for A55 network B (plan view)

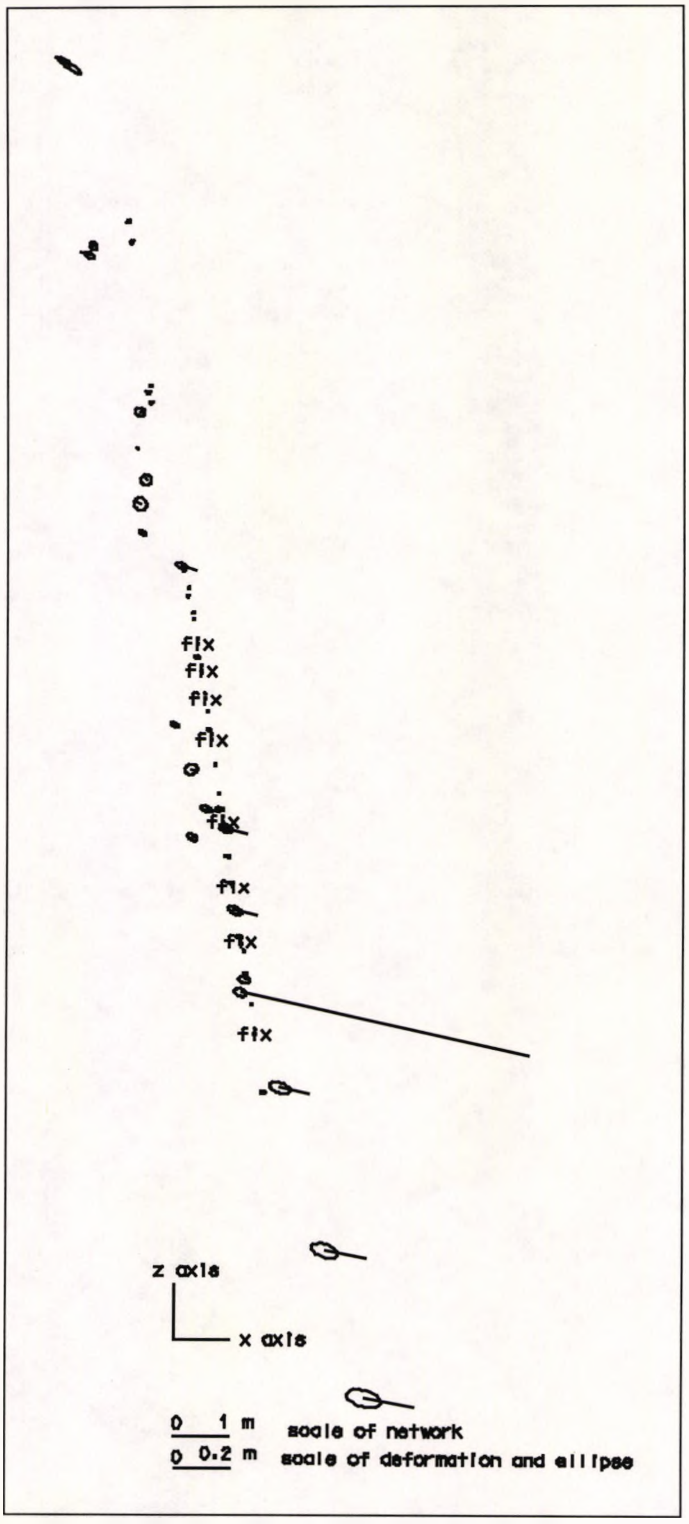

Figure 6.43 Estimated deformation for A55 network B (front view)

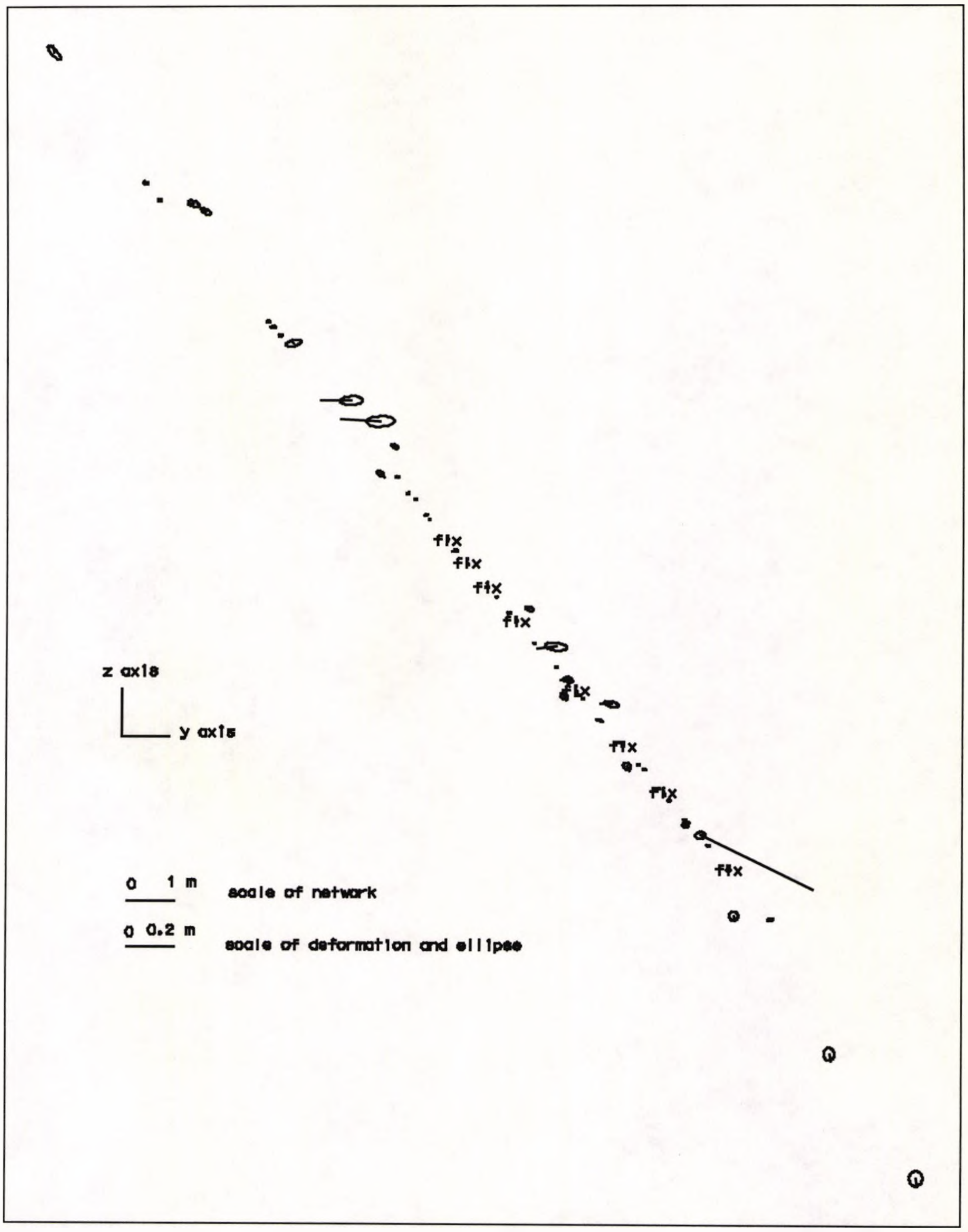

Figure 6.44 Estimated deformation for A55 network B (right view)

coordinates, values from the first epoch were chosen, since deformation detection was to be based on the first epoch. To facilitate the process of deformation detection, only common data (numbers and names of stations, provisional and estimated coordinates, and also the cofactor matrix) for each epoch were written to the deformation files. Such facility is a part of module for program COMON.

Deformation detection on the 60 common stations found 17 stable datum and 43 non-datum (7 stable, 36 unstable) stations. The stable datum stations were stations h5, D21, D31, D32, D33, D34, D35, D41, E91, E94, E95, F91, F94, F95, G94, P51, E111. The most significant deformation (Figure 6.41) was detected at station j112 with 1.163 m displacement vector, i.e (1.036, 0.478, -0.228) m. Other unstable stations were found to have small movements. Later, site investigation revealed that station j112 was moved deliberately during site works. Figures 6.41 to 6.44 show the trend of the deformation.

6.2.5 Deformation study of wood panel

This scheme is the actual application of the wood panel testing (section 6.2.2), to investigate the effect of moisture (temperature and humidity) on the behaviour of the wood panels. Such panels are used widely as a frame for the preservation of painting. This study is being carried out by ESRC at Hamilton Kerr Institute, Cambridge.

This scheme is still under investigation, and involves several epochs and various types of wood panels. A special room that allows temperature and relative humidity to be controlled are used in this study. In general, each panel with pre-marked points, is slotted into a fixed frame, and then imaged by 5 digital cameras. Data are automatically transferred into a PC, using the concept of 3-D measuring system. Automated methods of target location, identification and matching are followed by LSE of camera parameters and object coordinates. Details of the data collection procedure and LSE are given in Robson et al, (1995).

Data from two epochs that comprised 169 points in each epoch, were analysed (Figure 6.45). Twelve points are situated on the fixed frame (A1, A2, A3, B1, B2, B3, C1, C2, C3, D1, D2 and D3), whilst the remaining 157 points are on the panel. In the second epoch, the room's temperature and relative humidity were increased drastically. During LSE, the number of degrees of freedom were 1154 and 1156, whilst the estimated variance factors were 0.6835 and

0.5497, for the first and second epochs respectively.

Initially, the twelve points on the fixed frame were used to define the starting datum for deformation detection. The results of the detection indicated station B2 as unstable, and 11 stable datum stations and 158 non-datum (16 stable, 142 unstable) were found. The pattern of the deformation is displayed in Figures 6.46 to 6.49, indicating the significant upward movement of both left and right sides of the panel. The maximum displacement vector (18.1 mm) and movement in z direction (18.1 mm) were detected at station 233. Maximum movements in the x and y directions were found at points 258 (-0.7 mm) and B2 (-0.8 mm) respectively.

6.3 Chapter summary

Simulated numerical tests were carried out using a PC, since only small data sets were involved. All the real data consisted of between 40 to 169 stations, and were processed using workstation under UNIX environments. The graphics were produced using Intergraph MicroStation. In these applications, the relevant files were transferred between the PC, workstation and MicroStation using the UNIX based file transfer utility (program FTP).

All the developed programs were originally written and checked using the PC, and only transferred to the workstation for processing large data. Since FORTRAN 77 is compatible between the PC and UNIX based environment, only the dimension parameters need to be changed.

The results obtained in sections 6.1 and 6.2 show that the developed strategy and programs are applicable for the geometrical detection of spatial deformation. The strategy adopted for handling systematic and gross errors is directly applicable for deformation detection. Agreement of the results with known data, published results and actual situations demonstrated that the adopted procedure has fulfilled its expectations.

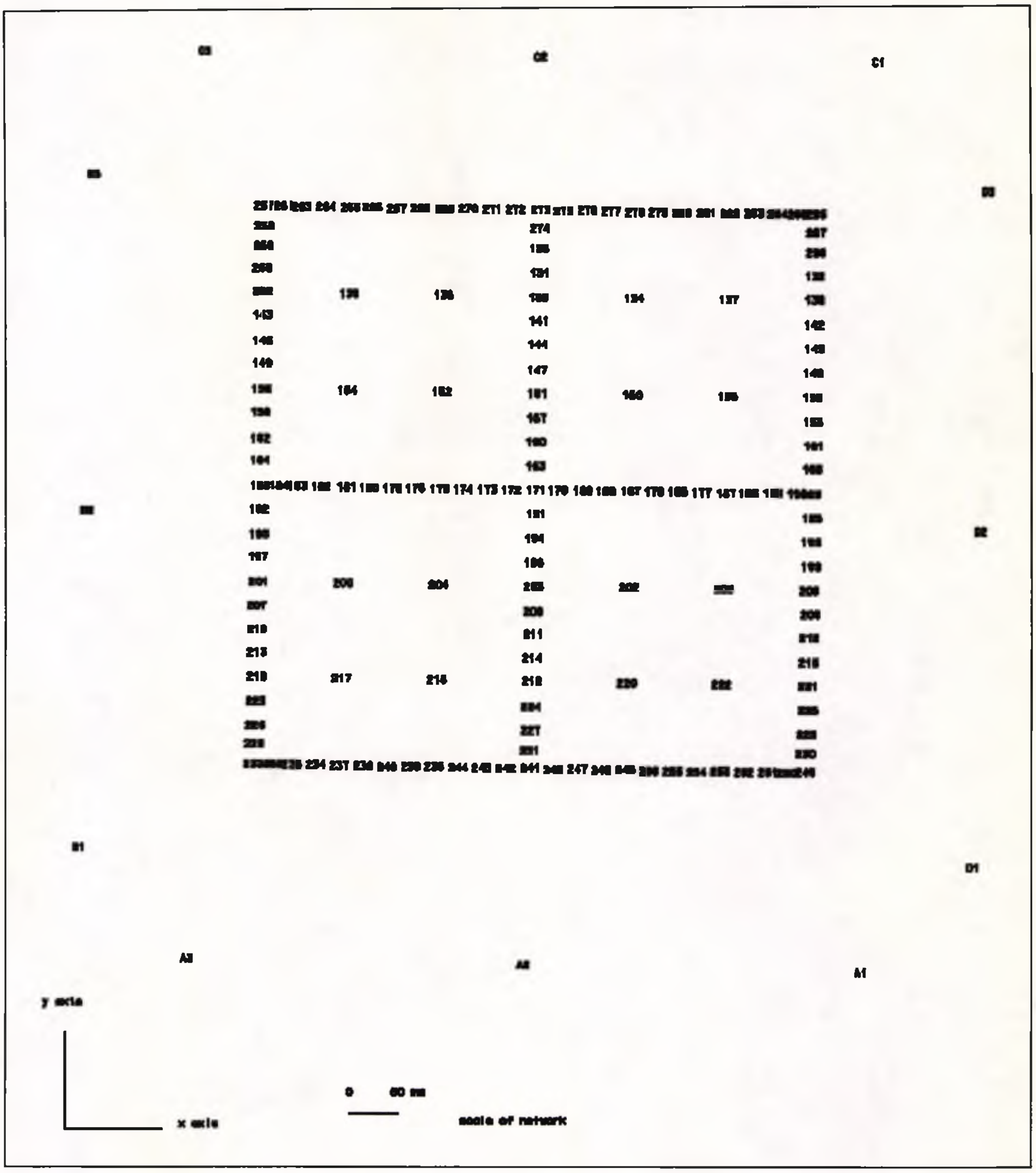

Figure 6.45 Wood painting network (plan view)

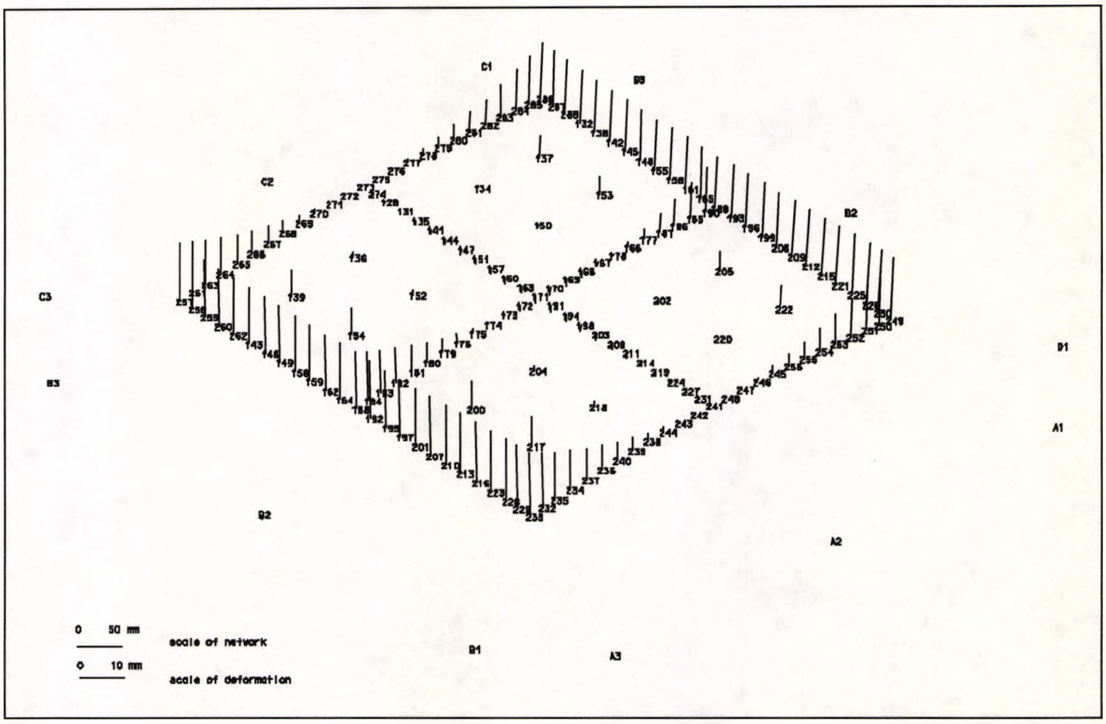

Figure 6.46 Estimated deformation for wood painting network (isometric view)

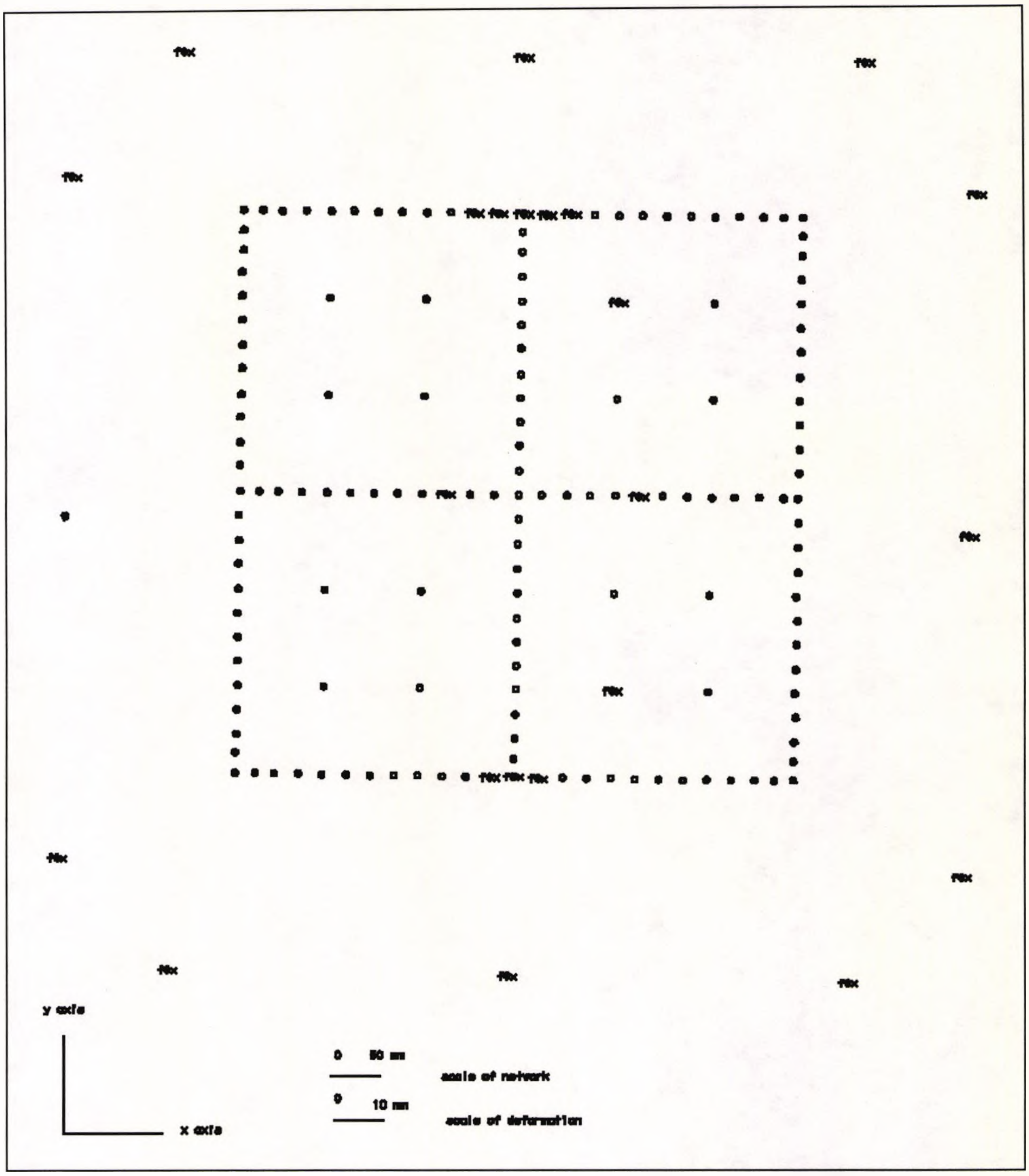

Figure 6.47 Estimated deformation for wood painting network (plan view)

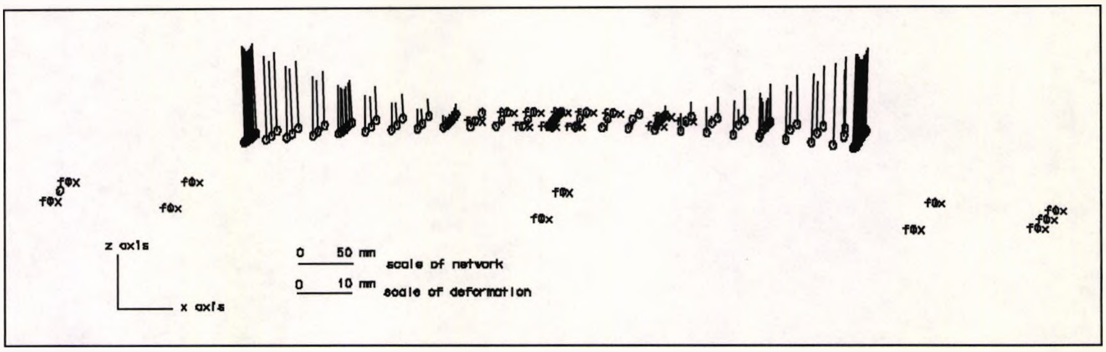

Figure 6.48 Estimated deformation for wood painting network (front view)

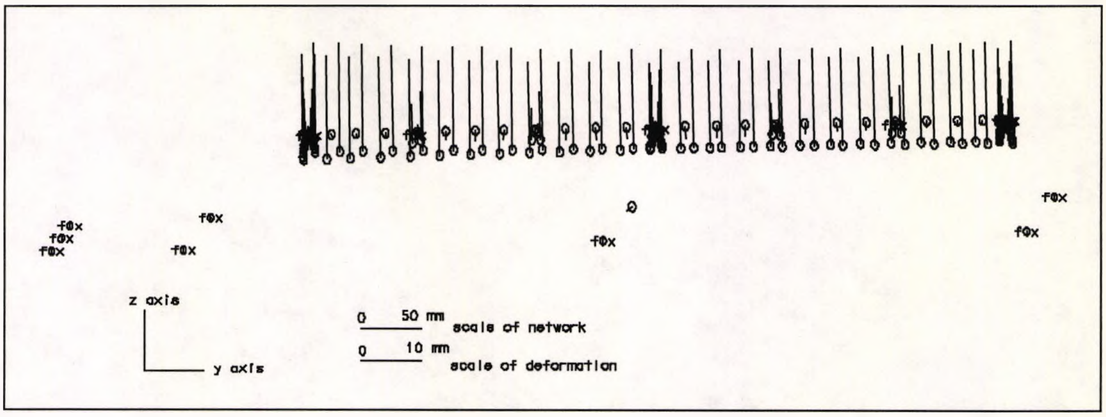

Figure 6.49 Estimated deformation for wood painting network (right view)

7. CONCLUSIONS AND RECOMMENDATIONS

This chapter summarizes the outcome of this program of research, and highlights the developed practical strategy for the detection of spatial deformation. In addition, some areas for possible further work are suggested.

7.1 Outcome of the research

The study is focused on practical, but rigorous, two-step geometrical analysis for the detection of spatial (3-D) deformation using geodetic methods. The two-steps consist of independent least squares estimation (LSE) of each epoch followed by deformation detection using two-epoch analysis. Since LSE and deformation detection are very important and critical, an extensive critical review of theory and current practice has been undertaken (chapters 1, 2, 3 and 4). Relevant aspects were examined and strategies were formulated in order to arrive at a practical approach for LSE and deformation detection in 3-D.

Several important LSE aspects that have been analysed and implemented (section 3.9) include: pre-analysis (section 3.2); datum definition via minimum constraints with fixed coordinates (section 2.2.2); rank defect analysis via simplified eigenvalue decomposition (EVD) (section 2.2.6); treatment of scale and zero errors in EDM distances by means of additional parameters (section 3.5.3.1) and pseudo observables (section 3.5.3.2); handling of the algebraically correlated pseudo observations via de-correlation of observations (section 3.5.3.2.3); detection of multiple gross errors by robustified LSE (section 3.6.2); stochastic modelling (section 3.7.3) by simplified variance component estimation (VCE); assessment of LSE results (section 3.9) using statistical testing (section 3.3) together with precision and reliability analyses (section 3.8); and datum re-definition via general S-transformations equation (section 2.3.5) with optimised computational procedure.

In deformation detection, the following aspects were implemented (section 4.4): initial testing (section 4.2.2); transformation into common stations using partitioning and S-transformations (section 4.2.1); determination and verification of stable datum stations by congruency testing (sections 4.2.3.1) and robust method applying S-transformations (4.2.3.4); localization of deformation through decomposition, re-ordering and S-transformations (section 4.2.3.2); and the testing of spatial deformations via single point tests (section 4.2.3.3).

The developed strategy for 3-D LSE and geometrical detection of spatial deformation (via one-stage computational procedure) is summarized in sections 3.9 and 4.4. The strategy is applied using five computer programs (section 5.1) developed during this research for independent 3-D LSE of surveying data for each epoch (program ESTIMATE (section 5.1.1)) and the geometrical detection of spatial deformation between any two epochs. The deformation detection modules include the determination of common stations between two epochs (program COMON (section 5.1.3)), S-transformations of single epoch LSE results (program COMPS (section 5.1.2)) and the detection of spatial deformation (programs DETECT (section 5.1.4) and ROBUST (section 5.1.5)) via one-stage computation. Programs COMON, COMPS, DETECT and ROBUST use the general S-transformations equations and the computational procedure outlined in section 2.3.5.

All five programs are independent of external routines, and can be executed under the personal computer (PC) (IBM compatible) environment for small data sets or under the UNIX environment for processing large data sets. Links between these programs and two of the ESRC's programs (GAP and DCRE) have also been established (section 5.2).

Tests carried out with simulated and real data show that the developed procedure and programs are applicable for the detection of spatial deformation (chapter 6). Rank analysis (section 6.1.1) was found useful for checking the effective rank and condition of normal equations. Optimised computational procedures for S-transformations (section 6.1.2) were tested and verified (Tables 6.2 and 6.3). In simulation tests, the simulated errors (section 6.1.3) and deformation (section 6.1.4) were correctly recovered. The use of published data (Tables 6.8 and 6.9) also indicate that the robustified LSE strategy is applicable. Moreover, the developed strategy for handling systematic and gross errors was shown applicable for deformation detection (Tables 6.12 and 6.13).

Real data from five monitoring schemes with up to 169 stations were processed under the UNIX environment. In processing these data (section 6.2), the links established between programs (section 5.2) were found advantageous, especially when different stations were used in each epoch (section 6.2.4). Flexibility in datum definition was proven useful (section 6.2.5). Agreement between the estimated results and real situations (section 6.2) confirmed the suitability of the procedure for deformation detection. Moreover, the procedure was found to be suitable for wide range of uses, from local (sections 6.2.1, 6.2.3 and 6.2.4) to micro-scale (sections 6.2.2 and 6.2.5) applications.

7.2 Practical strategy for deformation detection

A practical strategy for the geometrical detection of spatial deformation is summarized below. It is derived from the outcome of this research. The procedures for 3-D LSE, S-transformations and deformation detection (via one-stage computation) adopted during this research are found to be flexible and useful for practical engineering work.

In the LSE of each epoch, for simplicity, the observations need to be uncorrelated. Prior to the actual LSE, pre-analysis (section 3.2) should be carried out first to evaluate the expected precision and reliability. An initial datum definition via minimum constraints (section 2.2.2) is recommended for reasons of simplicity. Rank defect analysis (section 2.2.6) is important to check the deficiency of the normal equations. A suitable computational scheme allows practical S-transformations (section 2.3.5) of LSE results from one datum to another prior to and during the detection process.

It is necessary to handle systematic and gross errors during measurement prior to LSE. This can be carried out by regular calibration (section 3.5.1), proper measurement scheme (section 3.5.2) and independent checks (section 3.6.1) or combination. Observations also need to be reduced and corrected properly for systematic effects, before they are used in LSE.

During LSE, any significant scale and zero errors in the measured EDM distances can be handled by improving the functional model based on either the additional parameters (section 3.5.3.1) or pseudo observables (section 3.5.3.2) approach. Use of the algebraically correlated pseudo observables requires application of de-correlation technique (section 3.5.3.2.3). Although additional parameters seem to be the most obvious choice, precision and reliability analyses (section 3.8) should be used to determine the most suitable approach. For the purpose of deformation detection, use of additional parameters is recommended, due to its simplicity. Consequently, significant parameters must be included with the measured distances in the same manner as scaled distances (section 3.5.3.1) in order to maintain network precision and reliability.

Multiple gross errors in the observations can be handled conveniently by adopting robustified LSE (section 3.6.2), without the need to eliminate the suspected observations. This is possible because the effects of gross errors on the final estimation are almost negligible.

Of equal importance is the proper handling of random error, which can be determined in practice with reference to experience and some prior knowledge. The use of a simplified VCE (section 3.7.3) is recommended for this purpose. However, both systematic and gross errors have to be tackled first in order to obtain a reliable estimation of variances.

During LSE, the effects of any significant systematic and gross errors need to be controlled, and results are only acceptable if the solution passes both global and local tests. This is important because the global test is sometimes not sensitive. In addition, precision and reliability analyses must be acceptable as well (section 3.9).

Prior to deformation detection, it is necessary to verify that LSE results of each epoch are referred to common stations and datum (section 4.2.2). In the case of differing configurations or datum definition between epochs, it is necessary to apply S-transformations to transform the LSE results into a common datum (section 4.2.1).

In most monitoring activities, the reference stations are known in advance, and such stations can be used as an initial datum for deformation detection. During deformation detection, the stability determination and testing of datum points, localization and testing of deformation can carried out effectively in one-stage computation (section 4.2.3.1, 4.2.3.2 and 4.2.3.3). The results of this computation can be used directly without additional computations, due to the idempotent property of S-transformations matrix. To obtain uniformity in testing, it is necessary to vary or standardize the significance level α between global and local test (section 4.4). In practice, regular values of α are 0.05 and 0.01 for global and local tests respectively.

Interpretation of the results of deformation detection is very important. All the datum stations must be confirmed as stable, otherwise, it is necessary to repeat the computations until a set of stable datum stations is found. Unstable non-datum stations represent the stations that have been deformed significantly between epochs (section 4.4).

With the available knowledge on these aspects and developed strategy, engineering surveyors will be able to routinely carry out deformation monitoring. Correct interpretation of the estimated solutions is necessary in order to fully understand the underlying concepts.

7.3 Suggestion for further work

This study has concentrated on the 3-D LSE processing of surveying data and the detection of spatial deformation by geometrical means. Some aspects on the developed programs need to be refined and open the possibility of further work. In addition, related subject matter can also be pursued in the future.

The developed programs can be refined to make them more attractive and user friendly. For example by the inclusion of on-screen graphics display so that plots of deformation vectors can be viewed directly during computation. Other possible refinements include menu templates and on-line help.

Related subject matter includes the application of integrated method and a near real time scheme. For engineering purpose, it is not sufficient to simply compute deformation vectors, and the estimated deformation must be linked to the physical process occuring during the deformation study. The most complete approach is via the integrated method which combines the geometric model with the finite element method. Such an integrated method combines all data (geodetic, geotechnical and structural) into one solution. By these means both geometrical and physical information are taken into account. This research has provided a sound geometric basis for the integrated method.

Technological development in instrumentation allows automatic and remotely controlled telemetric data acquisition. This advancement has led to the development of real time 3-D positioning and analysis, and opens the possibility of near real time industrial applications and continuous monitoring with telemetric data acquisition.

APPENDIX A. THE FUNCTIONAL MODEL FOR SURVEYING OBSERVABLES

In the method of LSE using observation equations (Chapters 2 and 3), the functional model relating n observations with u unknown parameters is generally non-linear

$$l=f(x)$$

In order to apply linear algebra for LSE, the model needs to be linearised. The linearised model is in the form

$$Ax=b+v$$

where $A=\partial f/\partial x$

A is the design (or linearisation or coefficient or configuration) matrix, x vector of parameters, b the misclosure (i.e. observed minus computed) vector, and v the vector of residuals.

Matrix A can be determined either by applying differential calculus or numerical modeling. In both cases, Taylor series expansion is used. Differential calculus is commonly adopted and is applied in this study. Numerical modelling is suitable for evaluation of complex functions. However, a proper selection of the limit is required, depending on the types of functions. Aspects on numerical modelling are discussed by Seager and Shortis (1993).

The use of Taylor series expansion in differential calculus (Cooper, 1974; Mikhail and Gracie, 1981) is simple. Let

$$l=f(x)$$

then $l=l_o+(\partial f/\partial x)\delta x+\text{higher order terms}$

$l_o=f(x_o)$

$\delta x=x-x_o$

$\partial f/\partial x$=first derivatives of f with respect to x evaluated at $x=x_o$

In LSE, the higher order terms are ignored, and the result is

$l = l_o + (\partial f/\partial x)\delta x$ or $Ax = b$

where $A = \partial f/\partial x$, $b = l - l_o$, $x = \delta x$

Similarly, if

$l = f(x_1, x_2, \ldots, x_u)$

then $l = l_o + (\partial f/\partial x_1)\delta x_1 + (\partial f/\partial x_2)\delta x_2 + \ldots + (\partial f/\partial x_u)\delta x_u$

or $Ax = b$

where $A = [\partial f/\partial x_1 \; \partial f/\partial x_2 \; \ldots \; \partial f/\partial x_u]$, $b = l - l_o$

$x = [\delta x_1 \; \delta x_2 \; \ldots \; \delta x_u]^t$

The relevant functional models and linearised observation equations for ordinary surveying observables can be found in many textbooks, for example Cooper (1987), Mikhail and Gracie (1981), and Blachut et al (1979). In most cases, only the 2-D case is considered.

Here, the functional models and observation equations are formulated to enable 3-D LSE. Whenever applicable, both 1-D and 2-D cases are also shown. Consider three stations i, j and k, with their coordinates (x_i, y_i, z_i), (x_j, y_j, z_j) and (x_k, y_k, z_k) respectively. Let

s=spatial or slope distance from i to k

d=horizontal distance from i to k

h=height difference from i to k

α=azimuth from i to k, measured clockwise

ϕ=direction from i to k, measured clockwise

β=horizontal angle at i, measured clockwise from j to k

ζ=zenith angle or distance from i to k

θ=vertical angle from i to k

ht_{tar}=height of reflector at k (if applicable)

ht_{ins}=height of instrument at i if applicable)

1. Spatial (slope) distance (s) i to k

Functional model:

$$s=[(x_k-x_i)^2+(y_k-y_i)^2+(z_k+ht_{tar}-z_i-ht_{ins})^2]^{1/2}$$
$$=[dx^2+dy^2+dz^2]^{1/2}$$

where $dx=(x_k-x_i)$, $dy=(y_k-y_i)$, $dz=(z_k+ht_{tar})-(z_i+ht_{ins})$

Observation equation:

$$-(dx/s)\delta x_i-(dy/s)\delta y_i-(dz/s)\delta z_i$$
$$+(dx/s)\delta x_k+(dy/s)\delta y_k+(dz/s)\delta z_k=(s_{obs}-s_o)+v$$

where s_{obs}=observed value (after applying appropriate corrections)

s_o=computed value (using functional model)

2. Horizontal distance (d) i to k

Functional model:

$$d=[dx^2+dy^2]^{1/2}$$

Observation equation:

$$-(dx/d)\delta x_i-(dy/d)\delta y_i+(dx/d)\delta x_k+(dy/d)\delta y_k=(d_{obs}-d_o)+v$$

3. Height difference (h) i to k

Functional model:

$$h=(z_k+ht_{tar})-(z_i+ht_{ins})=dz$$

Observation equation:

$$-\delta z_i+\delta z_k=(h_{obs}-h_o)+v$$

4. Azimuth (α) i to k (clockwise)

Functional model:

$$\alpha=\tan^{-1}(dx/dy)$$

Observation equation:

$$-(dy/d^2)\delta x_i+(dx/d^2)\delta y_i+(dy/d^2)\delta x_k-(dx/d^2)\delta y_k=(\alpha_{obs}-\alpha_o)+v$$

<u>5. Direction (ϕ) i to k (clockwise)</u>

Functional model:

$$\phi=\alpha-w=\tan^{-1}(dx/dy)-w$$

where α is azimuth, and w is orientation parameter to be estimated for each group of the directions measured.

Observation equation:

$$-(dy/d^2)\delta x_i+(dx/d^2)\delta y_i+(dy/d^2)\delta x_k-(dx/d^2)\delta y_k-\delta w=(\phi_{obs}-\phi_o)+v$$

<u>6. Horizontal angle (β) at i, from j, to k (clockwise)</u>

Functional model:

$$\beta=\tan^{-1}(dx_k/dy_k)-\tan^{-1}(dx_j/dy_j)$$

where $dx_k=(x_k-x_i)$, $dy_k=(y_k-y_i)$

$dx_j=(x_j-x_i)$, $dy_j=(y_j-y_i)$

Observation equation:

$$[(dy_j/d_j^2)-(dy_k/d_k^2)]\delta x_i+[(dx_k/d_k^2)-(dx_j/d_j^2)]\delta yi$$
$$-(dy_j/d_j^2)\delta x_j+(dx_j/d_j^2)\delta y_j$$
$$+(dy_k/d_k^2)\delta x_k-(dx_k/d_k^2)\delta y_k=(\beta_{obs}-\beta_o)+v$$

where $d_j=[dx_j^2+dy_j^2]^{1/2}$ is the horizontal distance between i and j

$d_k=[dx_j^2+dy_j^2]^{1/2}$ is the horizontal distance between i and k

7. Zenith angle (ξ) i to k

Functional model:

$$\xi=\cos^{-1}(dz/s)=\cos^{-1}[dz/(dx^2+dy^2+dz^2)^{1/2}]$$

Observation equation:

$$-[dx/(s^2\tan\xi)]\delta x_i-[dy/(s^2\tan\xi)]\delta y_i+(d/s^2)\delta z_i$$
$$+[dx/(s^2\tan\xi)]\delta x_k+[dy/(s^2\tan\xi)]\delta y_k-(d/s^2)\delta z_k=(\xi_{obs}-\xi_o)+v$$

8. Vertical angle (θ) i to k

Functional model:

$$\theta=\tan^{-1}(dz/d)$$

Observation equation:

$$(dx)(\tan\theta/s^2)\delta x_i+(dy)(\tan\theta/s^2)\delta y_i-(d/s^2)\delta z_i$$
$$-(dx)(\tan\theta/s^2)\delta x_k-(dy)(\tan\theta/s^2)\delta y_k+(d/s^2)\delta z_k=(\theta_{obs}-\theta_o)+v$$

APPENDIX B. VARIOUS WEIGHTING FUNCTIONS FOR RLSE

RLSE is actually a modification of the Danish method. The procedure of RLSE is outlined in section 3.6.2., where iterative process of LSE and de-weighting of observations are repeated until convergence is achieved. In this section, various weighting functions used by others are highlighted.

The original Danish method is an iterative algorithm to minimize the weighted square sum of the residuals when the solution converge as follows

$$\Sigma(v^tWv) \rightarrow \min$$

The original de-weighting scheme uses the following exponential weights

if $|\hat{v}_i| <$ limit; p_i=1.0; weight unchanged as $w_i'=p_i*w_i$

if $|\hat{v}_i| \geq$ limit; $p_i=\exp^{-f}$; new weight becomes $w_i'=p_i*w_i$

$\text{limit}=c*\sigma_{input}*\hat{\sigma}_o$

$f=|\hat{v}_i|/\text{limit}$

where $\hat{v}_i$=estimated residual for observation i

w_i'=new weight for observation i

σ_{input}=original standard deviation of observation i

$\hat{\sigma}_o$=square root of the estimated variance factor

c=weighting constant

Constant c is usually set to 3. This value is chosen (Straub, 1983) because the probability that a true error exceed 3σ is 0.0027.

The statistic of the Danish method is simply

$$t_{danish}=\hat{v}_i/(\sigma_{input}\hat{\sigma}_o) \leq c$$

In practice, variations of the above deweighting functions are used for different problems. Selection of constant c is usually based on experience.

For photogrammetric problem, Kubik et al (1988) use the following

if $|\hat{v}_i|$ < limit; p_i=1.0; $w_i'=p_i*w_i$

if $|\hat{v}_i|$ ≥ limit; $p_i=\exp^{-f}$; $w_i'=p_i*w_i$

limit=$2\sigma_{input}$

$f=|\hat{v}_i|^2/(\text{limit})^2$

Jorgensen et al (1985) suggested a slightly different function for photogrammetric bundle adjustment

$p_i=\exp^{-f1}$; $w_i'=p_i*w_i$ for first 3 iterations

$f1=0.05(|\hat{v}_i|/\text{limit})^{4.4}$

$p_i=\exp^{-f2}$; $w_i'=p_i*w_i$ for further iterations

$f2=0.005(|\hat{v}_i|/\text{limit})^{3.0}$

limit=$\sigma_{input}*\hat{\sigma}_o$

Kubik et al (1987) give examples of weight functions for levelling and resection problems. For levelling network:

$p_i=\exp^{-f1}$; $w_i'=p_i*w_i$ for first 5 iterations

$f1=0.01(|\hat{v}_i|/\text{limit1})^{4.4}$

limit1=$\sigma_{input}*\hat{\sigma}_o$

$p_i=\exp^{-f2}$; $w_i'=p_i*w_i$ for further iterations

$f2=0.05(|\hat{v}_i|/\text{limit2})^{20}$

limit2=$2*\sigma_{input}*\hat{\sigma}_o$

For resection problem

$p_i=\exp^{-f1}$; $w_i'=p_i*w_i$ for first 3 iterations

$f1=0.03(|\hat{v}_i|/\text{limit1})^{25}$

limit1=$0.4*\sigma_{\hat{v}i}*\hat{\sigma}_o$

$p_i=\exp^{-f2}$; $w_i'=p_i*w_i$ for further iterations

$f2=0.05(|\hat{v}_i|/\text{limit2})^{20}$

limit2=$15*\sigma_{\hat{v}i}*\hat{\sigma}_o$

where $\hat{\sigma}_{vi}$=estimated standard deviation of residuals

Chong (1987) uses normalized residuals and adopted

if $|v_{i'}|$ < limit; $w_i'=1$

if $|v_i'| \geq$ limit; $w_i'=0$

$v_i'=\hat{\sigma}_i/(\hat{\sigma}_o\sigma_{\hat{v}i})$ is normalized residual

limit=critical value of Tau statistic

Jianjun (1991) uses the following function

if $|\hat{v}_i|$ < limit; $w_i'=p_1$

if $|\hat{v}_i| \geq$ limit; $w_i'=p_1/fac$

limit=$c*\sigma_{\hat{v}i}$

$p_1=\hat{\sigma}_o^2/\hat{\sigma}_i^2$

$fac=1+(v_i/(\sigma_i r_i))^2$

c=critical value of Baarda's statistic

APPENDIX C. THE STOCHASTIC MODEL FOR SURVEYING OBSERVABLES

The surveying observables normally use in monitoring activities consist of angular (directions, azimuths, horizontal, vertical and zenith angles) and linear measurements (distances and height differences). Such observations are assumed to be uncorrelated but with different precisions. The angular, distance and height difference measurements are usually obtained using theodolites, EDMs and levels respectively.

The expression for variances of these observables are given below, based on Blachut et al (1979), Secord (1986) and Reuger (1988). The units of variances are $secs^2$ and mm^2 for angular and linear measurements respectively. Further details may also be found in Cooper (1982).

The precisions of the observations depend on the instruments used, the method of observations and to some extent, the environmental circumstances. Errors in the angular measurements are due to the instrument and / or observer (typically pointing, reading, levelling and centring) and due to the refraction or environment (proportional to distance). Errors in distance measurements are due to phase determination and refraction.

1. Direction

The variance (σ_{di}^2) of the measured direction is the combination of pointing, reading, levelling and centering errors. It can be expressed as (Blachut et al, 1979; Secord, 1986)

$$\sigma_{di}^2=\sigma_p^2+\sigma_r^2+\sigma_l^2+\sigma_c^2$$

where

σ_{di}^2=variance of the direction

σ_p=pointing error, ranges from (30 secs/m) to (60 secs/m)

m=telescope magnification

σ_r=reading error, ranges from 0.3d to 2.5d

d=least division of the micrometer (horizontal) in secs

σ_l=mislevelment error=$\sigma_v\cot(z)$

σ_v=levelling error, ranges from 0.02v to 0.2v

v=sensitivity of the vertical spirit level in secs

z=zenith angle to target

σ_c=centering error of theodolite=$\rho(\sigma_{c1}^2+\sigma_{c3}^2)^{1/2}/d$

σ_{c1}=centering error of the target

σ_{c3}=centering error of the theodolite

d=distance to target

ρ=206265 secs

The final right-hand term of the above expression is known as external error, dependent on distance. Levelling error is negligible for small value of vertical angle.

2. Horizontal angle

A horizontal angle is the difference of two directions, and the variance (σ_{ha}^2) is (Secord, 1986; Blachut et al, 1979)

$\sigma_{ha}^2=[2(\sigma_p^2+\sigma_r^2)+\sigma_{l1}^2+\sigma_{l2}^2+\sigma_{c'}^2]$

where

$\sigma_{l1}=(\sigma_v \cot(z_1))$

$\sigma_{l2}=(\sigma_v \cot(z_2))$

z_1 and z_2=zenith angles to the targets

$\sigma_{c'}^2=\rho^2[(\sigma_{c1}^2/d_1^2)+(\sigma_{c2}^2/d_2^2)+(\sigma_{c3}^2/d_1^2d_2^2)(d_1^2+d_2^2-2d_1d_2\cos\alpha)]$

σ_{c1} and σ_{c2}=centring errors of the targets

σ_{c3}=centring error of the theodolite

d_1 and d_2=distances to the targets

α=measured horizontal angle

The horizontal angles are algebraically correlated, and the covariance matrix Σ_l is not diagonal. Elements of Σ_l may be determined by the propogation of the variance as follows

measured directions s_1,s_2,s_3,s_4

standard deviations $\sigma_1,\sigma_2,\sigma_3,\sigma_4$

horizontal angles $\alpha_1=s_2-s_1$, $\alpha_2=s_3-s_2$, $\alpha_3=s_4-s_3$

$$\Sigma_l = \begin{matrix} \sigma_1^2+\sigma_2^2 & -\sigma_2^2 & 0 \\ -\sigma_2^2 & \sigma_2^2+\sigma_3^2 & -\sigma_3^2 \\ 0 & -\sigma_3^2 & \sigma_3^2+\sigma_4^2 \end{matrix}$$

3. Zenith or vertical angle

The variance of a zenith angle (σ_{za}^2) is (Secord, 1986)

$$\sigma_{za}^2=\sigma_p^2+\sigma_r^2+\sigma_v^2+\sigma_{c'}^2+\sigma_{ref}^2$$

where

σ_v=levelling error, ranges from 0.02v to 0.2v as before

$\sigma_{c'}$=centering error=$\sigma_t(\sin(z)/d)$

σ_t=precision of the target height

σ_{ref}=effect of refraction=$\sigma_k(d/2R)$

k=coefficient of refraction

d=distance between stations

R=mean radius of curvature of earth

Typical formulae for computing k is (Bomford, 1980)

$$k=502(p/t^2)(0.0341+\partial t/\partial h)$$

where p is pressure in mbar, t is temperature in ° Kelvin (K) and ($\partial t/\partial h$) is vertical temperature gradient in km^{-1}.

4. EDM distance

Variance of EDM distance (σ_s^2) can be written as (Blachut et al, 1979; Secord, 1986)

$$\sigma_s^2=a^2+b^2s^2+\sigma_{\Delta s}^2$$

where a and b represent constant (i.e.zero) and distance dependent (i.e.scale) errors of the EDM respectively, as given by the manufacturers. Parameter b is based on the knowledge of the refractive index of the air and the stability of the modulation frequency. $\sigma_{\Delta s}$ is the error due to the reduction applied to the instrument output.

In practice, the variance of EDM distances is usually adopted as

$\sigma_s^2=a^2 + b^2s^2$ where a is in mm, b in ppm (parts per million), and s in m.

5. Height difference

Typical precision of orthometric height difference determination via spirit levelling is in the form (Blachut et al, 1979)

$\sigma=\sigma_h(L)^{1/2}$ mm

where σ_h=standard deviation per unit distance (typically few mm)

L=length of the levelling line (km)

The observed height differences are referred to the actual surface of geoid, whilst the computations are referred to the model surface of ellipsoid. In general, the observed height differences need to be reduced to the ellipsoid. However, for small areas, geoid-ellipsoid separation is very small, and height differences can be used directly without reduction to the ellipsoidal surface.

Nowadays, modern instrumentation together with proper observing scheme allow high precision measurements. The direction, horizontal angles, vertical and zenith angles can be measured using a high precision theodolite with precision of equal to or less than 1 sec. The vertical and zenith angles are less accurate due to the uncertainty of vertical refraction. In distance measurements, precise EDM instruments such as Geomensor and Mekometer allow measurement to 0.2 mm $\pm$ 0.2 ppm. Height differences may be obtained either by spirit levelling or trigonometric heighting. Spirit levelling allow height differences to be determined to mm level, whilst trigonometric heighting is less accurate due to the refraction effects.

In order to utilise the potential of the present precise surveying instruments for mm level monitoring, special attachment devices are required (Ashkenazi et al, 1980), such as precision (forced) centring and targetting devices.

APPENDIX D. CONDITIONS SATISFIED BY THE GENERALISED INVERSE

The generalised inverse (or g-inverse) of an (mxn) matrix A (denoted by A^g) satisfies the following equation

$AA^gA=A$ [g condition]

There is a variety of generalised inverses, and all the inverses satisfy the g condition. The types of generalised inverses applicable for LSE are reflexive generalised, pseudo (or Moore-Penrose) and normal (or Cayley) inverses.

The reflexive generalised inverse (A^r) of A (Caspary, 1987b) satisfies

$AA^rA=A$ [g condition]

$A^rAA^r=A^r$ [r condition]

According to Cooper and Cross (1991) and Forsythe et al (1977), the pseudo-inverse or Moore-Penrose inverse of an nxu matrix A (denoted by A^+) satisfy the following four important conditions (Moore-Penrose conditions)

$AA^+A=A$ [g condition]

$A^+AA^+=A^+$ [r condition for reflexive property]

$(AA^+)^t=AA^+$ [l condition for least squares property]

$(A^+A)^t=A^+A$ [m condition for minimum norm property]

When A is square and full rank, normal inverse (A^{-1}) exists. The inverse satisfies all the above four conditions

$AA^{-1}A=A$

$A^{-1}AA^{-1}=A^{-1}$

$(AA^{-1})^t=AA^{-1}$

$(A^{-1}A)^t=A^{-1}A$

Reflexive generalised inverse is used in the partial minimum trace solution, while pseudo inverse is used in the minimum trace or free network solution. Normal inverse is computed during minimum constraints solution (section 2.2.2).

APPENDIX E. THE MAIN APPROACHES TO THE ANALYSIS OF DEFORMATION SURVEYS

The FIG ad hoc committee on the analysis of deformation surveys (Chrzanowski and Chen, 1986) identifies ten main groups or approaches to the analysis of deformation surveys. The groups, named after their locations, are Bonn, Delft, Fredericton, Haifa-Tel Aviv, Hannover, Karlsruhe, Munich I, Munich II, Stuttgart and Warsaw I.

The differences between these groups are in the strategies of deformation modelling, identification of deformation models, determination of deformation parameters, selection of the significance level for testing and quality control. Details on the different approaches are given by reports of the ad hoc committee (Chrzanowski, 1981; Heck et al, 1983; Chrzanowski and Secord, 1983; Chrzanowski and Chen, 1986).

In deformation modelling, three groups (Fredericton, Haifa-Tel Aviv and Munich I) concentrate on the deformation of the whole body. The other groups focus on the determination of the movements of the single points or a group of points, and (except Warsaw I) are based on the congruency testing on a group of reference points. The Warsaw I group applies congruency testing on the differences of the individual observations.

In the identification of deformation model, the applications of the congruency testing differ in the strategy of localization of the unstable reference points. The Hannover, Karlsruhe and Stuttgart groups remove the suspected unstable reference points from the computational base, one by one until the congruency test passes. The Bonn group uses a minimum number of selected points as reference points, and adds stable point one by one until the the congruency test fails. The Munich II group uses the Cholesky decomposition, whilst the Delft group adopts a trial and error method. The Warsaw I group applies trend analysis to identify the stable reference points.

The Haifa-Tel Aviv, Munich I and Fredericton groups adopt different approaches. The Haifa-Tel Aviv group uses the velocity model for the simultaneous estimation of the velocity and coordinates. The Munich I group applies the finite element method to determine the deformation pattern of the individual triangular elements of the network. The Fredericton group uses a robust method to identify the displacement pattern and the deformation model.

All the groups apply least squares criterion to estimate the deformation parameters. However, the determination of the single point movements is different. Most of the groups (Delft, Fredericton, Hannover, Karlsruhe, Munich II, Warsaw) consider each reference points as the same point between two epochs. The Bonn and Stuttgart groups consider such points as two separate points, and hence the single point movements are referred to the datum defined by the reference points.

The selection of the significance level in testing also varies. The Hannover, Stuttgart and Karlsruhe groups fix the significance level as 0.05 for the congruency test. The Delft group applies Baarda's test. The Fredericton group uses an arbirtarily selected significance level in all tests. The Bonn group uses the Bayesian inference to obtain a less sensitive test.

The groups use different tests for the detection of gross errors. The Bonn and Hannover groups fix the significance level for the simultaneous tests on all the observations, whilst the Delft, Fredericton and Karlsruhe groups fix the standardized significance level for tests on the individual observations.

APPENDIX F. SIMULATED DATA FOR NETWORK 1

Data for network 1 (see Figure 6.2) consist of 54 uncorrelated observations (12 slope distances, 12 height differences and 30 directions) derived from 6 stations. The coordinates for the stations are as follows:

station	Provisional coordinates (m)		
	x	y	z
1	1200.000	2600.000	120.000
2	1350.000	3000.000	140.000
3	1700.000	2950.000	80.000
4	1950.000	2750.000	90.000
5	1900.000	2400.000	150.000
6	1450.000	2250.000	100.000

Observations obtained from the above coordinates are perfect and contain no random, systematic or gross errors. To generate normally distributed observations, random errors σ of 5 mm and 5 secs were simulated into the linear (slope distances and height differences) and angular (directions) observations respectively. The randomized observations are listed below.

(i) 12 slope distances

from	to	slope distance (m)
1	2	427.6666
1	3	611.6399
1	4	765.4421
1	5	728.6376
1	6	430.5789
3	4	320.3127
3	5	589.4070
3	6	743.5670
3	2	358.6066
5	6	476.9706
5	2	814.0091
5	4	358.6068

(ii) 12 height differences

from	to	height difference (m)
1	2	19.9952
2	3	-59.9933
3	4	9.9963
4	5	59.9983
5	6	-50.0033
6	1	20.0009
1	3	-40.0076
3	5	70.0027
5	1	-29.9849
6	2	40.0011
2	4	-49.9887
4	6	10.0026

(iii) 30 directions

from	to	direction (decimal degree)
1	2	20.5546
1	3	55.0070
1	4	78.6900
1	5	105.9452
1	6	144.4608
2	3	98.1299
2	4	112.6194
2	5	137.4913
2	6	172.4069
2	1	200.5545
3	4	128.6593
3	5	160.0191
3	6	199.6534
3	1	235.0065
3	2	278.1310
4	5	188.1318
4	6	225.0016
4	1	258.6916
4	2	292.6198
4	3	308.6606
5	6	251.5679
5	1	285.9467
5	2	317.4883
5	3	340.0176
5	4	8.1317
6	1	324.4617
6	2	352.4068
6	3	19.6539
6	4	45.0018
6	5	71.5674

APPENDIX G. SAMPLE PROGRAM INPUT AND OUTPUT

This section gives the sample program input and output. The relevant files contain the data necessary for executing programs ESTIMATE, COMPS and DETECT. The same data format are applicable for programs COMON and ROBUST.

Input files for ESTIMATE

epoch 1:
sep.tar (provisional coordinates)
sepa.obs (observational data of epoch 1)

epoch 2:
sep.tar (provisional coordinates)
sepb.obs (observational data of epoch 2)

Output files of ESTIMATE

epoch 1:
1.res (summary of LSE)
1.def (deformation file)
1.plo (plotting file)

epoch 2:
2.res (summary of LSE)
2.def (deformation file)
2.plo (plotting file)

Deformation file can be used directly as input for COMPS, DETECT, COMON and ROBUST. Plotting file is used as input file for DCRE.

Example of using COMPS

epoch 1:
1.def (input file)
1a.def (deformation file)
1a.plo (plotting file)

epoch 2:
2.def (input file)
2a.def (deformation file)
2a.plo (plotting file)

Example of using DETECT

1a.def 2a.def (input files of two epochs)
def.sum (summary of detection)
def.plo (plotting file)

The following nine files are listed for demonstration purpose: sep.tar, sepa.obs, 1.res, 1.def, 1.plo, 1a.def, 1a.plo, def.sum and def.plo. The note is intended to explain the relevant information related to the files.

(i) sep.tar

1	1200.000	2600.000	120.000	1	1	1
2	1350.000	3000.000	140.000	0	0	0
3	1700.000	2950.000	80.000	0	1	0
4	1950.000	2750.000	90.000	0	0	0
5	1900.000	2400.000	150.000	0	0	0
6	1450.000	2250.000	100.000	0	0	0

Note:
The data is in free format.
column 1 is an integer representing station name.
columns 2, 3 and 4 contains 3-D coordinates (x, y, z) in m.
columns 5, 6 and 7 are integer code for datum definition (1 for datum, 0 for non-datum).

(ii) sepa.obs

0	0	1	2	427.6666	5.0000	1.000	0.000	0.000	1
0	0	1	3	611.6399	5.0000	1.000	0.000	0.000	0
0	0	1	4	765.4421	5.0000	1.000	0.000	0.000	0
0	0	1	5	728.6376	5.0000	1.000	0.000	0.000	0
0	0	1	6	430.5789	5.0000	1.000	0.000	0.000	0
0	0	3	4	320.3127	5.0000	1.000	0.000	0.000	0
0	0	3	5	589.4070	5.0000	1.000	0.000	0.000	0
0	0	3	6	743.5670	5.0000	1.000	0.000	0.000	0
0	0	3	2	358.6066	5.0000	1.000	0.000	0.000	0
0	0	5	6	476.9706	5.0000	1.000	0.000	0.000	0
0	0	5	2	814.0091	5.0000	1.000	0.000	0.000	0
0	0	5	4	358.6068	5.0000	1.000	0.000	0.000	0
1	0	1	2	19.9952	5.0000	0.000	0.000	0.000	1
1	0	2	3	-59.9933	5.0000	0.000	0.000	0.000	0
1	0	3	4	9.9963	5.0000	0.000	0.000	0.000	0
1	0	4	5	59.9983	5.0000	0.000	0.000	0.000	0
1	0	5	6	-50.0033	5.0000	0.000	0.000	0.000	0
1	0	6	1	20.0009	5.0000	0.000	0.000	0.000	0
1	0	1	3	-40.0076	5.0000	0.000	0.000	0.000	0
1	0	3	5	70.0027	5.0000	0.000	0.000	0.000	0
1	0	5	1	-29.9849	5.0000	0.000	0.000	0.000	0
1	0	6	2	40.0011	5.0000	0.000	0.000	0.000	0
1	0	2	4	-49.9887	5.0000	0.000	0.000	0.000	0
1	0	4	6	10.0026	5.0000	0.000	0.000	0.000	0
3	1	1	2	20.5546	5.0000	0.000	0.000	0.000	1
3	0	1	3	55.0070	5.0000	0.000	0.000	0.000	0
3	0	1	4	78.6900	5.0000	0.000	0.000	0.000	0
3	0	1	5	105.9452	5.0000	0.000	0.000	0.000	0
3	0	1	6	144.4608	5.0000	0.000	0.000	0.000	0
3	1	2	3	98.1299	5.0000	0.000	0.000	0.000	0

3	0	2	4	112.6194	5.0000	0.000	0.000	0.000	0
3	0	2	5	137.4913	5.0000	0.000	0.000	0.000	0
3	0	2	6	172.4069	5.0000	0.000	0.000	0.000	0
3	0	2	1	200.5545	5.0000	0.000	0.000	0.000	0
3	1	3	4	128.6593	5.0000	0.000	0.000	0.000	0
3	0	3	5	160.0191	5.0000	0.000	0.000	0.000	0
3	0	3	6	199.6534	5.0000	0.000	0.000	0.000	0
3	0	3	1	235.0065	5.0000	0.000	0.000	0.000	0
3	0	3	2	278.1310	5.0000	0.000	0.000	0.000	0
3	1	4	5	188.1318	5.0000	0.000	0.000	0.000	0
3	0	4	6	225.0016	5.0000	0.000	0.000	0.000	0
3	0	4	1	258.6916	5.0000	0.000	0.000	0.000	0
3	0	4	2	292.6198	5.0000	0.000	0.000	0.000	0
3	0	4	3	308.6606	5.0000	0.000	0.000	0.000	0
3	1	5	6	251.5679	5.0000	0.000	0.000	0.000	0
3	0	5	1	285.9467	5.0000	0.000	0.000	0.000	0
3	0	5	2	317.4883	5.0000	0.000	0.000	0.000	0
3	0	5	3	340.0176	5.0000	0.000	0.000	0.000	0
3	0	5	4	8.1317	5.0000	0.000	0.000	0.000	0
3	1	6	1	324.4617	5.0000	0.000	0.000	0.000	0
3	0	6	2	352.4068	5.0000	0.000	0.000	0.000	0
3	0	6	3	19.6539	5.0000	0.000	0.000	0.000	0
3	0	6	4	45.0018	5.0000	0.000	0.000	0.000	0
3	0	6	5	71.5674	5.0000	0.000	0.000	0.000	0

Note:
The data is in free format.
column 1 is an integer of observation code.
(0=spatial or slope distance, 1=height differences, 2=uncorrelated horizontal angle, 3=direction, 4=zenith angle, 5=correlated horizontal angle, 6=azimuth, 7=vertical angle, -1=horizontal distances, -2=spatial distance ratio, -3=spatial distance difference, -4=spatial distance plus scale error, -5= spatial distance plus zero error, -6=spatial distance plus scale plus zero errors, -8=spatial distance with multiple scale errors, -9=spatial distance with multiple zero errors).
column 2 is an integer representing reference station used as reference in measurement of horizontal angle. Its value is zero if not applicable. For direction measurement, value of 1 indicates start of set at a station.
column 3 and 4 are integers representing instrument and target stations respectively (in m).
column 5 is observation in unit m for linear and unit decimal degree for angular. The observation is assumed as been corrected or reduced properly.
column 6 is standard deviation of measurements in mm for linear and secs for angular.
column 7 is mainly for manipulating correlated observations. It contains standard error for start of horizontal angle, distance ratio and difference. For uncorrelated observation, its value is 0. For slope distance, its value is scale for the distance (default value is 1).
column 8 and 9 are heights of instrument and target respectively in m.
column 10 is mostly 0. Value of 1 indicates start of the observational group for VCE.

(iii) 1.res

LSE MODE: full analysis

coord file:sep.tar
obsns file:sepa.obs

global variance factor: 0.81415
no of iteration: 1 degree of freedom: 34

GLOBAL TEST
significant level for test: 0.050
chi-square test [one-tailed]:pass
0.570 is less or equal to 1.000

OBSERVATIONS/RESIDUALS

no	code	ro	at	to	obs	se	v	r	vn
1	0	0	1	2	427.6666	0.5000E+01	0.4650	0.33	0.18
2	0	0	1	3	611.6399	0.5000E+01	-1.6775	0.54	-0.50
3	0	0	1	4	765.4421	0.5000E+01	2.8161	0.50	0.88
4	0	0	1	5	728.6376	0.5000E+01	-4.3053	0.53	-1.31
5	0	0	1	6	430.5789	0.5000E+01	-1.9678	0.32	-0.77
6	0	0	3	4	320.3127	0.5000E+01	2.6472	0.40	0.93
7	0	0	3	5	589.4070	0.5000E+01	2.4858	0.54	0.75
8	0	0	3	6	743.5670	0.5000E+01	-0.5356	0.44	-0.18
9	0	0	3	2	358.6066	0.5000E+01	3.3965	0.38	1.22
10	0	0	5	6	476.9706	0.5000E+01	-0.8132	0.34	-0.31
11	0	0	5	2	814.0091	0.5000E+01	-1.1071	0.47	-0.36
12	0	0	5	4	358.6068	0.5000E+01	2.4538	0.37	0.90
13	1	0	1	2	19.9952	0.5000E+01	-4.9844	0.58	-1.44
14	1	0	2	3	-59.9933	0.5000E+01	-2.2100	0.59	-0.64
15	1	0	3	4	9.9963	0.5000E+01	3.6952	0.58	1.07
16	1	0	4	5	59.9983	0.5000E+01	-0.0167	0.58	0.00
17	1	0	5	6	-50.0033	0.5000E+01	4.0755	0.58	1.18
18	1	0	6	1	20.0009	0.5000E+01	5.3405	0.58	1.55
19	1	0	1	3	-40.0076	0.5000E+01	2.3056	0.58	0.67
20	1	0	3	5	70.0027	0.5000E+01	-4.4215	0.59	-1.28
21	1	0	5	1	-29.9849	0.5000E+01	-8.0840	0.58	-2.34
22	1	0	6	2	40.0011	0.5000E+01	-4.6440	0.58	-1.35
23	1	0	2	4	-49.9887	0.5000E+01	-6.8148	0.58	-1.98
24	1	0	4	6	10.0026	0.5000E+01	-3.5412	0.58	-1.03
25	3	1	1	2	20.5546	0.5000E+01	1.8952	0.69	0.51
26	3	0	1	3	55.0070	0.5000E+01	1.0615	0.77	0.27
27	3	0	1	4	78.6900	0.5000E+01	-1.9227	0.77	-0.49
28	3	0	1	5	105.9452	0.5000E+01	-1.6687	0.77	-0.42
29	3	0	1	6	144.4608	0.5000E+01	0.6347	0.68	0.17
30	3	1	2	3	98.1299	0.5000E+01	1.8687	0.68	0.50
31	3	0	2	4	112.6194	0.5000E+01	2.6556	0.76	0.68
32	3	0	2	5	137.4913	0.5000E+01	-5.6379	0.78	-1.42
33	3	0	2	6	172.4069	0.5000E+01	-5.1468	0.76	-1.31

34	3	0	2	1	200.5545	0.5000E+01	6.2604	0.69	1.67
35	3	1	3	4	128.6593	0.5000E+01	1.8348	0.62	0.52
36	3	0	3	5	160.0191	0.5000E+01	-8.0810	0.75	-2.07
37	3	0	3	6	199.6534	0.5000E+01	2.7516	0.76	0.70
38	3	0	3	1	235.0065	0.5000E+01	6.2264	0.75	1.59
39	3	0	3	2	278.1310	0.5000E+01	-2.7318	0.62	-0.77
40	3	1	4	5	188.1318	0.5000E+01	-2.6422	0.64	-0.73
41	3	0	4	6	225.0016	0.5000E+01	0.0723	0.75	0.02
42	3	0	4	1	258.6916	0.5000E+01	-1.2653	0.78	-0.32
43	3	0	4	2	292.6198	0.5000E+01	3.6279	0.75	0.93
44	3	0	4	3	308.6606	0.5000E+01	0.2074	0.65	0.06
45	3	1	5	6	251.5679	0.5000E+01	-3.0629	0.70	-0.81
46	3	0	5	1	285.9467	0.5000E+01	-1.1386	0.77	-0.29
47	3	0	5	2	317.4883	0.5000E+01	7.0869	0.77	1.79
48	3	0	5	3	340.0176	0.5000E+01	-0.1157	0.77	-0.03
49	3	0	5	4	8.1317	0.5000E+01	-2.7697	0.67	-0.75
50	3	1	6	1	324.4617	0.5000E+01	3.0084	0.70	0.80
51	3	0	6	2	352.4068	0.5000E+01	-3.1783	0.76	0.81
52	3	0	6	3	19.6539	0.5000E+01	3.2005	0.78	0.80
53	3	0	6	4	45.0018	0.5000E+01	-1.4515	0.76	-0.37
54	3	0	6	5	71.5674	0.5000E+01	-1.5791	0.71	-0.41

sum of r: 34.00

CHI-SQUARE GOODNESS OF FIT TEST ON
normalized/standardized estimated residuals
[number of group = 2]
[degrees of freedom= 1]

obs f	exp f	[o-e]	[o-e]**2/e
29.000	27.000	2.000	0.148
25.000	27.000	-2.000	0.148
		sum=	0.296

0.296 < 3.840
test passes...res are normally distributed

LOCAL TEST
[critical value of tau statistic= 3.11]
no of obsns fail the test[ie v norm > tau]= 0

RELIABILITY ANALYSIS

obs no	rnum	irel	erel
1	0.3340	35.6692	33.8918
2	0.5429	27.9789	14.3128
3	0.4989	29.1874	17.0762
4	0.5305	28.3050	15.0468
5	0.3184	36.5322	36.3841
6	0.3964	32.7454	25.8905
7	0.5417	28.0113	14.3854
8	0.4436	30.9537	21.3253
9	0.3832	33.3011	27.3585
10	0.3393	35.3917	33.1029
11	0.4748	29.9175	18.8024

12	0.3676	34.0000	29.2400
13	0.5847	26.9607	12.0752
14	0.5859	26.9335	12.0165
15	0.5850	26.9541	12.0610
16	0.5849	26.9560	12.0651
17	0.5848	26.9580	12.0693
18	0.5843	26.9694	12.0939
19	0.5844	26.9669	12.0886
20	0.5851	26.9513	12.0549
21	0.5841	26.9737	12.1032
22	0.5840	26.9760	12.1082
23	0.5841	26.9734	12.1025
24	0.5838	26.9809	12.1188
25	0.6914	24.7937	7.5891
26	0.7688	23.5116	5.1118
27	0.7664	23.5486	5.1815
28	0.7726	23.4543	5.0042
29	0.6839	24.9278	7.8557
30	0.6833	24.9395	7.8792
31	0.7562	23.7075	5.4819
32	0.7780	23.3732	4.8522
33	0.7595	23.6552	5.3827
34	0.6917	24.7875	7.5769
35	0.6208	26.1646	10.3836
36	0.7510	23.7893	5.6372
37	0.7560	23.7095	5.4857
38	0.7545	23.7338	5.5317
39	0.6224	26.1318	10.3148
40	0.6394	25.7807	9.5859
41	0.7540	23.7418	5.5469
42	0.7751	23.4156	4.9316
43	0.7458	23.8713	5.7936
44	0.6491	25.5887	9.1913
45	0.6967	24.6987	7.4010
46	0.7705	23.4854	5.0625
47	0.7660	23.5555	5.1944
48	0.7675	23.5311	5.1485
49	0.6712	25.1627	8.3264
50	0.6996	24.6477	7.3004
51	0.7644	23.5796	5.2398
52	0.7809	23.3284	4.7686
53	0.7628	23.6039	5.2857
54	0.7139	24.4000	6.8144

STN COORDS/SE

stn	x	y	z
1	1200.0000	2600.0000	120.0000
	0.0000	0.0000	0.0000
2	1349.9985	2999.9999	139.9902
	0.0049	0.0036	0.0032
3	1700.0009	2950.0000	79.9947

	0.0041	0.0000	0.0032
4	1950.0039	2749.9990	89.9947
	0.0037	0.0049	0.0035
5	1900.0043	2399.9978	149.9930
	0.0039	0.0044	0.0032
6	1450.0012	2250.0062	99.9938
	0.0049	0.0035	0.0032

trace and rmse:

0.0001	0.0001	0.0001
0.0043	0.0041	0.0033

PARAMETERS FOR DIRECTIONS

stn	orientation (deg)	se(sec)
1	0.00073	2.46813
2	-0.00038	2.74444
3	-0.00020	2.69119
4	-0.00105	2.90129
5	-0.00091	2.82163
6	-0.00083	2.78661

Note:
This file is summary of LSE. It contains results of global and local tests, estimated coordinates, together with precision and reliability analyses.

(iv) 1.def

```
#coord file:sep.tar          ;obsn file:sepa.obs
   34 #degrees of freedom
      0.814147 #posteriori variance factor
    4 #rank defect
 1 1 1 0 0 1 0 #datum code
    6 #no of stns
#prov coords
      1  1200.0000  2600.0000  120.0000
      2  1350.0000  3000.0000  140.0000
      3  1700.0000  2950.0000   80.0000
      4  1950.0000  2750.0000   90.0000
      5  1900.0000  2400.0000  150.0000
      6  1450.0000  2250.0000  100.0000
#estimated coords
      1  1200.0000  2600.0000  120.0000 1 1 1
      2  1349.9985  2999.9999  139.9902 0 0 0
      3  1700.0009  2950.0000   79.9947 0 1 0
      4  1950.0039  2749.9990   89.9947 0 0 0
      5  1900.0043  2399.9978  149.9930 0 0 0
      6  1450.0012  2250.0062   99.9938 0 0 0
#covariances
    1          0.000000000000000E+00
```

2	0.000000000000000E+00
3	0.000000000000000E+00
4	0.000000000000000E+00
5	0.000000000000000E+00
6	0.000000000000000E+00
7	0.000000000000000E+00
8	0.000000000000000E+00
9	0.000000000000000E+00
10	0.236183885669249E-04
11	0.000000000000000E+00
12	0.000000000000000E+00
13	0.000000000000000E+00
14	0.318689947526138E-05
15	0.133093456840842E-04
16	0.000000000000000E+00
17	0.000000000000000E+00
18	0.000000000000000E+00
19	0.292294604460585E-06
20	-0.259234592607668E-06
21	0.103827347930409E-04
22	0.000000000000000E+00
23	0.000000000000000E+00
24	0.000000000000000E+00
25	0.121601761515521E-04
26	0.159591717351124E-05
27	-0.841677250603603E-07
28	0.171197227643384E-04
29	0.000000000000000E+00
30	0.000000000000000E+00
31	0.000000000000000E+00
32	0.000000000000000E+00
33	0.000000000000000E+00
34	0.000000000000000E+00
35	0.000000000000000E+00
36	0.000000000000000E+00
37	0.000000000000000E+00
38	0.000000000000000E+00
39	0.000000000000000E+00
40	-0.923877759816369E-07
41	-0.128679687536849E-06
42	0.520950994117585E-05
43	0.540198853457363E-06
44	0.000000000000000E+00
45	0.103894613411439E-04
46	0.000000000000000E+00
47	0.000000000000000E+00
48	0.000000000000000E+00
49	0.570871105645641E-05
50	0.306472478974394E-06
51	0.848559801470249E-07
52	0.721931610389486E-05

53	0.000000000000000E+00
54	0.340750523513292E-06
55	0.134185425902084E-04
56	0.000000000000000E+00
57	0.000000000000000E+00
58	0.000000000000000E+00
59	-0.787433113386810E-05
60	-0.628624098083962E-06
61	0.119879569006635E-06
62	-0.755341818786170E-05
63	0.000000000000000E+00
64	-0.340779443787870E-06
65	-0.332350422512976E-05
66	0.241566006473863E-04
67	0.000000000000000E+00
68	0.000000000000000E+00
69	0.000000000000000E+00
70	0.859061576317848E-07
71	-0.158258321427431E-06
72	0.623050254771579E-05
73	0.232603667390580E-06
74	0.000000000000000E+00
75	0.624430265377008E-05
76	0.235799129219890E-06
77	0.460753735614827E-06
78	0.124747084939587E-04
79	0.000000000000000E+00
80	0.000000000000000E+00
81	0.000000000000000E+00
82	0.105428849813906E-05
83	-0.237005119803232E-05
84	0.612291441257390E-07
85	0.100433824021053E-05
86	0.000000000000000E+00
87	-0.120576193252030E-06
88	0.324307801765336E-05
89	0.606382040241949E-05
90	-0.403335653117243E-07
91	0.152279439882472E-04
92	0.000000000000000E+00
93	0.000000000000000E+00
94	0.000000000000000E+00
95	-0.794000573068074E-05
96	0.620714577369019E-06
97	-0.196625252110583E-06
98	-0.639978966271142E-05
99	0.000000000000000E+00
100	-0.498974248051832E-06
101	-0.211491840728978E-05
102	0.134157856381365E-04
103	-0.259629678017574E-06

104	0.710564767845698E-05
105	0.189987669570203E-04
106	0.000000000000000E+00
107	0.000000000000000E+00
108	0.000000000000000E+00
109	-0.135186282106819E-06
110	0.380160366358185E-08
111	0.415599808289926E-05
112	0.196392400602223E-06
113	0.000000000000000E+00
114	0.520684978593921E-05
115	0.925879047424945E-07
116	-0.143695175151357E-06
117	0.624692007313709E-05
118	-0.114445636293589E-06
119	0.401696861938354E-06
120	0.103967707657788E-04
121	0.000000000000000E+00
122	0.000000000000000E+00
123	0.000000000000000E+00
124	-0.146594562034047E-05
125	-0.196027725515107E-05
126	0.486968380053489E-07
127	-0.661306121131798E-06
128	0.000000000000000E+00
129	-0.359757678289836E-07
130	0.168583260148196E-05
131	0.728585177587548E-05
132	0.881143661380316E-07
133	0.108633884837057E-04
134	0.842496554760888E-05
135	0.412757504476116E-06
136	0.238748650035446E-04
137	0.000000000000000E+00
138	0.000000000000000E+00
139	0.000000000000000E+00
140	0.113815827084233E-05
141	0.795570860755044E-06
142	-0.156527405267326E-06
143	0.188447802152499E-05
144	0.000000000000000E+00
145	-0.135045722980187E-06
146	0.698390356251991E-06
147	0.485314944362405E-06
148	-0.126555546135624E-06
149	0.709962443442664E-06
150	0.145324906016725E-05
151	0.322396517696441E-07
152	-0.911559808208819E-06
153	0.122980971019547E-04
154	0.000000000000000E+00

155	0.000000000000000E+00
156	0.000000000000000E+00
157	0.361730762146149E-08
158	-0.138265287616770E-06
159	0.518777370420365E-05
160	0.920073639328803E-08
161	0.000000000000000E+00
162	0.416070448802979E-05
163	0.115232045268282E-06
164	0.265381530471828E-06
165	0.623110446042786E-05
166	0.229449080497125E-06
167	0.184523144866072E-06
168	0.520455976580990E-05
169	0.112692481109321E-06
170	-0.162107224002597E-06
171	0.103921145563762E-04

Note:
This deformation file contain data necessary for detection purpose, i.e degrees of freedom, estimated variance factor, datum defect information, number of stations, provisional coordinates together with estimated coordinates and cofactor matrix (upper triangle).

(v) 1.plo

```
1 1200.0000 2600.0000 120.0000 0.0000 0.0000 0.0000
# se:    0.0000    0.0000    0.0000
# var: 0.0000E+00 0.0000E+00 0.0000E+00 0.0000E+00 0.0000E+00 0.0000E+00
2 1349.9985 2999.9999 139.9902 0.0000 0.0000 0.0000
# se:    0.0049    0.0036    0.0032
# var: 0.2362E-04 0.1331E-04 0.1038E-04 0.3187E-05 0.2923E-06 -0.2592E-06
3 1700.0009 2950.0000 79.9947 0.0000 0.0000 0.0000
# se:    0.0041    0.0000    0.0032
# var: 0.1712E-04 0.0000E+00 0.1039E-04 0.0000E+00 0.5402E-06 0.0000E+00
4 1950.0039 2749.9990 89.9947 0.0000 0.0000 0.0000
# se:    0.0037    0.0049    0.0035
# var: 0.1342E-04 0.2416E-04 0.1247E-04 -0.3324E-05 0.2358E-06 0.4608E-06
5 1900.0043 2399.9978 149.9930 0.0000 0.0000 0.0000
# se:    0.0039    0.0044    0.0032
# var: 0.1523E-04 0.1900E-04 0.1040E-04 0.7106E-05 -0.1144E-06 0.4017E-06
6 1450.0012 2250.0062 99.9938 0.0000 0.0000 0.0000
# se:    0.0049    0.0035    0.0032
# var: 0.2387E-04 0.1230E-04 0.1039E-04 -0.9116E-06 0.1127E-06 -0.1621E-06
```

Note:
This file provides suitable data for plotting of coordinates and error ellipses via DCRE.

(vi) 1a.def

```
#       RESULTS OF S-TRANSFORMATIONS
#    input file:1.def
   34 #degrees of freedom
      0.814147 #posteriori variance factor
    4 #rank defect
 1 1 1 0 0 1 0 #datum code
    6 #no of stns
#prov coords
        1   1200.0000   2600.0000    120.0000
        2   1350.0000   3000.0000    140.0000
        3   1700.0000   2950.0000     80.0000
        4   1950.0000   2750.0000     90.0000
        5   1900.0000   2400.0000    150.0000
        6   1450.0000   2250.0000    100.0000
#estimated coords
        1   1199.9986   2599.9993    120.0056 1 1 1
        2   1349.9969   2999.9993    139.9958 1 1 1
        3   1699.9993   2949.9996     80.0003 1 1 1
        4   1950.0024   2749.9987     90.0003 1 1 1
        5   1900.0029   2399.9975    149.9986 1 1 1
        6   1449.9999   2250.0057     99.9994 1 1 1

#covariances
    1          0.483515970784850E-05
    2         -0.143926387321563E-06
    3          0.499959880687926E-05
    4          0.724147237042138E-07
    5         -0.432184575703288E-07
    6          0.450534002656987E-05
    7         -0.127903494625284E-05
    8         -0.149973694366921E-05
    9          0.376360980739105E-07
   10          0.725755665439951E-05
   11          0.307574851420544E-06
   12         -0.184663620105944E-05
   13          0.736226572353050E-07
   14         -0.343499225530364E-06
   15          0.773916579713363E-05
   16          0.400242025980504E-08
   17          0.473317276509568E-07
   18         -0.689079818269369E-06
   19          0.270164089362218E-06
   20         -0.983038840031267E-07
   21          0.449923512993229E-05
   22         -0.740288545345684E-06
   23         -0.754405098970337E-06
   24         -0.845606441574500E-07
   25         -0.202428847971961E-05
   26         -0.107263077984547E-05
```

27	-0.229575693674104E-06
28	0.484233898036141E-05
29	0.430338450171833E-06
30	-0.909220668258191E-06
31	-0.318793446210695E-07
32	0.741422867635749E-07
33	-0.824043744888095E-06
34	0.478637277600519E-07
35	0.486840931886210E-07
36	0.469480248702341E-05
37	-0.555828485969781E-08
38	-0.430210990978524E-07
39	-0.696464675106598E-06
40	-0.320349359071107E-06
41	0.154995803957332E-07
42	-0.681374578769990E-06
43	0.213493612211466E-06
44	0.202848854600743E-06
45	0.449119196436083E-05
46	-0.181131528591609E-06
47	-0.107026625045615E-06
48	-0.758579406623945E-07
49	-0.341382476816685E-05
50	-0.480729490203568E-06
51	-0.561721301077291E-07
52	-0.817320642833034E-06
53	0.172542078382677E-06
54	0.116560322062185E-06
55	0.736178166887803E-05
56	0.174631270827266E-06
57	-0.412402999596825E-06
58	-0.864664908387838E-07
59	0.233947569243452E-05
60	-0.120098090456151E-05
61	0.107752594128891E-06
62	0.101258986042732E-05
63	-0.142141418212310E-05
64	-0.752523150301732E-07
65	0.156819302781020E-06
66	0.712772480070406E-05
67	-0.309950001552210E-07
68	-0.871483114416700E-08
69	-0.173258301159838E-05
70	0.411252039468850E-07
71	-0.580042898036984E-07
72	-0.696500308721831E-06
73	0.630018874617417E-07
74	-0.236165583206883E-07
75	-0.690085059504770E-06
76	0.644037167357949E-07
77	0.369429611013534E-06

78	0.450420244419207E-05
79	-0.949322799162536E-06
80	0.824991909490967E-06
81	0.743965639229974E-07
82	-0.246990713269599E-06
83	-0.331890459030529E-06
84	0.628905594682620E-07
85	-0.736809646185436E-06
86	-0.698596806697350E-06
87	-0.303403015932098E-07
88	-0.829517289820531E-06
89	-0.887192621670245E-06
90	-0.798430612984739E-07
91	0.497380101826121E-05
92	-0.485024083332771E-06
93	-0.298405089368933E-06
94	0.341578291212280E-07
95	0.526632791900110E-06
96	-0.160453695661097E-06
97	-0.870471957442992E-07
98	0.619573024466683E-06
99	-0.122426114710237E-05
100	-0.136275883409057E-06
101	0.364887563517050E-06
102	-0.213741404582686E-05
103	-0.227705398646277E-06
104	0.198540410709275E-06
105	0.478948398451307E-05
106	-0.148227040005214E-07
107	-0.124166213695252E-07
108	-0.696509719024170E-06
109	-0.103280554814232E-06
110	0.776609507534110E-07
111	-0.173493148096415E-05
112	0.959130035480232E-07
113	-0.102961329574531E-06
114	-0.691464634761431E-06
115	-0.399423172783557E-07
116	-0.352185561520734E-06
117	-0.687512684055332E-06
118	-0.168040028537033E-06
119	0.324019177644415E-06
120	0.449841130116059E-05
121	-0.168538188849583E-05
122	0.168010314551575E-05
123	-0.240288008812771E-07
124	-0.293417746990606E-06
125	0.192117510318939E-05
126	-0.513092453084518E-07
127	-0.523631666277637E-06
128	-0.271101018093563E-07

129	0.261940112503630E-07
130	-0.211998743946601E-05
131	-0.279632350479988E-05
132	-0.576927466907266E-07
133	-0.221116056982311E-05
134	-0.122460970726035E-05
135	0.230172601082118E-06
136	0.683357931105319E-05
137	-0.283594101765309E-06
138	-0.153293384859587E-05
139	0.537838066736491E-07
140	-0.109701460189864E-05
141	-0.370705125096349E-05
142	-0.175969697924734E-07
143	0.146188900733195E-06
144	-0.315862744651661E-06
145	0.362008625406066E-07
146	-0.106492829431562E-06
147	-0.195551266859576E-05
148	-0.513885330987034E-07
149	0.894147567197883E-06
150	-0.968950006553818E-06
151	0.658833840669645E-07
152	0.446765065164430E-06
153	0.848031051936060E-05
154	-0.250411549485787E-07
155	0.600392815309167E-07
156	-0.690702802571360E-06
157	0.747045225023251E-07
158	-0.104750145776242E-07
159	-0.697348943206953E-06
160	-0.582721653896770E-07
161	-0.922553498445062E-07
162	-0.173180301621804E-05
163	-0.899165074950041E-08
164	0.367221622472666E-07
165	-0.697521380311752E-06
166	0.140936268037457E-06
167	0.928514710339906E-07
168	-0.687992782355503E-06
169	-0.123335819452026E-06
170	-0.868825503900435E-07
171	0.450536892466361E-05

stn	sx	sy	sz
1	0.0022	0.0022	0.0021
2	0.0027	0.0028	0.0021
3	0.0022	0.0022	0.0021
4	0.0027	0.0027	0.0021
5	0.0022	0.0022	0.0021
6	0.0026	0.0029	0.0021

Note:
This file (same format as file 1.def) contains results of S-transformations, and can be used directly for detection purpose.

(vii) 1a.plo

```
      1  1199.9986  2599.9993  120.0056  0.0000  0.0000  0.0000
# se:     0.0022     0.0022     0.0021
# var:  0.4835E-05  0.5000E-05  0.4505E-05 -0.1439E-06  0.7241E-07 -0.4322E-07
      2  1349.9969  2999.9993  139.9958  0.0000  0.0000  0.0000
# se:     0.0027     0.0028     0.0021
# var:  0.7258E-05  0.7739E-05  0.4499E-05 -0.3435E-06  0.2702E-06 -0.9830E-07
      3  1699.9993  2949.9996   80.0003  0.0000  0.0000  0.0000
# se:     0.0022     0.0022     0.0021
# var:  0.4842E-05  0.4695E-05  0.4491E-05  0.4868E-07  0.2135E-06  0.2028E-06
      4  1950.0024  2749.9987   90.0003  0.0000  0.0000  0.0000
# se:     0.0027     0.0027     0.0021
# var:  0.7362E-05  0.7128E-05  0.4504E-05  0.1568E-06  0.6440E-07  0.3694E-06
      5  1900.0029  2399.9975  149.9986  0.0000  0.0000  0.0000
# se:     0.0022     0.0022     0.0021
# var:  0.4974E-05  0.4789E-05  0.4498E-05  0.1985E-06 -0.1680E-06  0.3240E-06
      6  1449.9999  2250.0057   99.9994  0.0000  0.0000  0.0000
# se:     0.0026     0.0029     0.0021
# var:  0.6834E-05  0.8480E-05  0.4505E-05  0.4468E-06 -0.1233E-06 -0.8688E-07
```

Note:
This file is similar to file 1.plo (plotting of coordinates and error ellipse).

(viii) def.sum

```
DETECTION OF SPATIAL DEFORMATION/summary

epoch1 file:1a.def          epoch2 file:2a.def

...test on variance ratio...[significance level: 0.050]
df1=    34.0;df2=    34.0
[fcom.le.ftab?]      1.373     1.772
test passes
pooled var fac     0.704

...congruency testing...
[significance level: 0.050]

...global congruency test...
df1=    14.0;df2=    68.0
test fails [w.ge.f]   888.150     1.840
existence of deformation
station removed from datum automatically:       6
```

```
...partial congruency test...
df1=      11.0;df2=      68.0
test fails [w.ge.f]  182.727      1.932
existence of deformation
station removed from datum automatically:       3

...partial congruency test...
df1=       8.0;df2=      68.0
test fails [w.ge.f]    18.151     2.076
existence of deformation
station removed from datum automatically:       5

...partial congruency test...
df1=       5.0;df2=      68.0
test passes [w.lt.f]     0.017      2.345
no significant deformation

single point test
[significance level: 0.010]
df1= 3.0 ;df2=      68.0

...single point test...
stn  dx       dy       dz       disp vect     fcom  ftab  info
1  0.0000 -0.0001  0.0000  0.0001      0.00 4.09:stable[1.0]
2  0.0008 -0.0002  0.0000  0.0009      0.02 4.09:stable[1.0]
4 -0.0008  0.0003  0.0000  0.0009      0.03 4.09:stable[1.0]
6  0.0015 -0.0006  0.3000  0.3000  3179.88 4.09:moved [0.0]
3 -0.0499  0.0998 -0.1001  0.1499    640.44 4.09:moved [0.0]
5  0.0118  0.0510  0.0000  0.0523     47.75 4.09:moved [0.0]
no of stns:    6
no of datum stns/stable     :   3
no of datum stns/moved      :    0
no of non-datum stns/stable:    0
no of non-datum stns/moved :    3
[codes:1.0 datum pts,0.0 non-datum pts]

Note:
This file is summary of detection procedure.
```

(ix) def.plo

```
#       DETECTION OF DEFORMATION/plotting input
#      epoch1 file=1a.def           ;epoch2 file=2a.def
#      significance level of single point test= 0.010
#      critical value of chi-square=     11.3717
       1  1199.9986  2599.9993   120.0056
    To:  1199.9986  2599.9992   120.0056
   Def:      0.0000     -0.0001      0.0000 (      0.0001)
    Se:      0.0028      0.0028      0.0028
    Va:  0.7818E-05  0.7919E-05  0.7850E-05  0.1298E-05  0.2561E-06 -0.1028E-06
```

```
  Test:    0.0012
#      datum point/stable
#
     2  1349.9969  2999.9993   139.9958
   To:  1349.9977  2999.9991   139.9958
  Def:     0.0008    -0.0002     0.0000 (     0.0009)
   Se:     0.0036     0.0034     0.0025
   Va:  0.1327E-04  0.1160E-04  0.6464E-05 -0.2225E-05  0.1977E-06 -0.1631E-06
  Test:    0.0525
#      datum point/stable
#
     4  1950.0024  2749.9987    90.0003
   To:  1950.0016  2749.9990    90.0003
  Def:    -0.0008     0.0003     0.0000 (     0.0009)
   Se:     0.0036     0.0018     0.0028
   Va:  0.1277E-04  0.3267E-05  0.7859E-05 -0.2337E-05  0.2013E-06  0.6358E-07
  Test:    0.0615
#      datum point/stable
#
     6  1449.9999  2250.0057    99.9994
   To:  1450.0014  2250.0051   100.2994
  Def:     0.0015    -0.0006     0.3000 (     0.3000)
   Se:     0.0071     0.0057     0.0037
   Va:  0.4974E-04  0.3251E-04  0.1341E-04  0.8657E-06  0.5020E-07 -0.1917E-06
  Test: 6712.5346
#      non-datum point/moved
#
     3  1699.9993  2949.9996    80.0003
   To:  1699.9494  2950.0994    79.9002
  Def:    -0.0499     0.0998    -0.1001 (     0.1499)
   Se:     0.0044     0.0046     0.0037
   Va:  0.1895E-04  0.2118E-04  0.1336E-04 -0.6094E-06  0.8644E-06  0.6330E-06
  Test: 1351.9228
#      non-datum point/moved
#
     5  1900.0029  2399.9975   149.9986
   To:  1900.0147  2400.0485   149.9986
  Def:     0.0118     0.0510     0.0000 (     0.0523)
   Se:     0.0052     0.0052     0.0038
   Va:  0.2684E-04  0.2697E-04  0.1478E-04  0.1110E-04  0.1782E-06  0.1536E-05
  Test:  100.7903
#      non-datum point/moved
#
#  no of stns:   6
#  no of datum pts/stable     :   3
#  no of datum pts/moved      :   0
#  no of non-datum pts/stable:   0
#  no of non-datum pts/moved :   3
```

Note:
This file contains data for plotting of deformation vectors and the error ellipsoids.

REFERENCES AND BIBLIOGRAPHY

ANDERSON, J.M. and MIKHAIL, E.M. (1988). Introduction to surveying. (London: McGraw-Hill), 703pp.

ANGUS-LEPPAN, P.V. (1972). Adjustment of trilateration using length ratios, Survey Review XXI(166): 355-368.

ASHKENAZI, V. and DODSON, A.H. (1975). The Nottingham multipillar baseline, Proceedings of the XVI General Assembly of IAG, Grenoble.

ASHKENAZI, V., DODSON, A.H. and CRANE, S.A. (1980). Monitoring deformations to millimetre accuracy, Paper No. 8.3, Proceedings of the Industrial and Engineering Survey Conference, London.

BAARDA, W. (1968). A testing procedure for use in geodetic networks, Netherlands Geodetic Commission Publications on Geodesy, New Series, 2 (5), Delft.

BAARDA, W. (1973). S-transformation and criterion matrices, Netherlands Geodetic Commission Publications on Geodesy, New Series, 5(1), Delft, 168pp.

BAARDA, W. (1977). Measures for the Accuracy of Geodetic Networks. In: Optimization of Design and Computation of Control Network, ed. by F. Halmos and J. Somogyi. (Budapest: Akadémiai Kiadó). 733pp.

BAYLY, D.A. and TESKEY, W.F. (1992). Close-range high-precision surveys for machinery alignment, CISM Journal ACSGC 46(4): 409-421.

BIACS, Z. F. (1989). Estimation and hypothesis testing for deformation analysis in special purpose networks, Department of Surveying Engineering, The University of Calgary, UCSE report 20032, Calgary, 171pp.

BIACS, Z.F. and TESKEY, W.F. (1990). Deformation analysis of survey networks with interactive hypothesis testing and computer graphics, CISM Journal ACSGC 44(4): 403-416.

BIACS, Z.F., KRAKIWSKY, E.J. and LAPUCHA, D. (1990). Reliability analysis of phase observations in GPS baseline estimation, Journal of Surveying Engineering 116(4): 204-224.

BLACHUT, T.J., CHRZANOWSKI, A. and SAASTAMOINEN, J.H. (1979). Urban surveying and mapping. (New York: Springer-Verlag), 372pp.

BLAHA, G. (1971). Inner Adjustment Constraints with Emphasis on Range Observations, Department of Geodetic Science, The Ohio State University, Report No. 148, Columbus.

BOMFORD, G. (1980). Geodesy, 4th edition. (Oxford: Clarendon Press), 855pp.

BROWNLEE, K.A. (1965). Statistical Theory and Methodology in Science and Engineering, second edition. (New York: Wiley), 590pp.

BRUNNER, F.K. (1979). On the analysis of geodetic networks for the determination of the incremental strain tensor, Survey Review XXV 192: 56-67.

BURNSIDE, C.D. (1982). Electromagnetic Distance Measurement, second edition. (London: Granada), 278pp.

CASPARY, W.F. and BORUTTA, H. (1987a). Robust estimation in deformation models, Survey Review 29(223): 29-45.

CASPARY, W.F. (1987b). Concepts of network and deformation analysis, School of Surveying, The University of New South Wales, Monograph 11, Kensington, N.S.W., 183pp.

CHANDLER, J.H. (1994). Personal communication.

CHEN, Y.Q. (1983). Analysis of deformation surveys - a generalized method, Department of Surveying Engineering, University of New Brunswick, Technical Report no. 94, New Brunswick, 262pp.

CHEN, Y.Q., KAVOURAS, M. and CHRZANOWSKI, A. (1987). A strategy for detection of outlying observations in measurements of high precision, The Canadian Surveyor 41(4): 529-540.

CHEN, Y.Q., CHRZANOWSKI, A. and SECORD, J.M. (1990a). A strategy for the analysis of the stability of reference points in deformation surveys, CISM Journal ACSGC 44(2): 141-149.

CHEN, Y.Q., CHRZANOWSKI, A. and KAVOURAS, M. (1990b). Assessment of observations using minimum norm quadratic unbiased estimation (MINQUE), CISM Journal ACSGC 44(4): 39-46.

CHONG, A.K. (1987). A Robust Method for Multiple Outliers Detection in Multi-Parametric Models, Photogrammetric Engineering and Remote Sensing 53(6): 617-620.

CHRZANOWSKI, A. (with contributions by members of the "ad hoc" committee on the analysis of deformation measurements). (1981). A Comparison of Different Approaches into the Analysis of Deformation Measurements, Paper No. 602.3, Proceedings of the FIG XVI International Congress, Montreaux.

CHRZANOWSKI, A. and SECORD, J. (editors) (1983). Report of the Ad Hoc Committee on the Analysis of Deformation Surveys, Proceedings of the FIG XVII Congress, Paper No. 605.2.

CHRZANOWSKI, A. and Y.Q. CHEN (1986). Report of the ad-hoc committee on the analysis of deformation surveys, Paper No. 608.1, Proceedings of the FIG XVIII International Congress, Toronto.

CHRZANOWSKI, A., CHEN YONG-YI and J. SECORD (1986). Geometrical analysis of deformation surveys, Proceedings of the Deformations Measurements Workshop, Oct.31-Nov.1, Massachusetts Institute of Technology, Cambridge, MA. , 170-206.

CHRZANOWSKI, A., CHEN, Y.Q., SECORD, J.M. and CHRZANOWSKI, A.S. (1991). Problems and solutions in the integrated monitoring and analysis of dam deformations, CISM Journal ACSGC 45(4): 547-560.

CLARK, J.S. (1992). Personal communication.

COOKE, D., CRAVEN, A.H. and CLARKE, G.M. (1990). Basic statistical computing, second edition. (London: Edward Arnold), 178pp.

COOPER, M.A.R. (1974). Fundamentals of Survey Measurement and Analysis. (London: Granada), 107pp.

COOPER, M.A.R. (1982). Modern theodolites and levels. (London: Granada), 156pp.

COOPER, M.A.R. (1987). Control Surveys in Civil Engineering. (London: Collins), 381pp.

COOPER, M.A.R. and CROSS, P.A. (1988). Statistical concepts and their application in photogrammetry and surveying, Photogrammetric Record 12(71): 637-663.

COOPER, M.A.R. and CROSS, P.A. (1991). Statistical concepts and their application in photogrammetry and surveying (continued), Photogrammetric Record 13(77): 645-678.

COOPER, M.A.R. (1994). Personal communication.

CROSS, P.A. and THAPA, K. (1979). The optimal design of levelling networks, Survey Review 25(192): 68-79.

CROSS, P.A. (1983). Advanced least squares applied to position fixing. Department of Land Surveying, North East London Polytechnic, Working paper 6, 205pp.

CROSS, P.A. and PRICE, D.R. (1985). A strategy for the distinction between single and multiple gross errors in geodetic networks, Manuscripta Geodaetica 10(3): 172-178.

DANIEL, C. and WOOD, F.S. (1980). Fitting Equations to Data. second edition. (New York: Wiley), 458pp.

DAVIES, R.E., FOOTE, F.S., ANDERSON, J.M. and MIKHAIL, E.M. (1981). Surveying: theory and practice. sixth edition. (London: McGraw Hill), 992pp.

DEETH, C.P., DODSON, A.H. and ASHKENAZI, V. (1978). EDM accuracy and calibration, Proceedings of Symposium of EDM. London.

DERMANIS, A. (1994). The photogrammetric inner constraints, ISPRS Journal of Photogrammetry and Remote Sensing 49(1): 25-39.

DODSON, A.H. and ZAHER, H. (1985). Refraction effects on vertical angle measurements, Survey Review 28(217): 169-183.

DODSON, A. H. (1990). Analysis of control networks and their application to deformation monitoring. In: Engineering surveying technology. ed. by T. J. M. Kennie and G. Petrie. (London: Blackie). 485pp.

DRACUP, J.F., FRONCZEK, C.J. and TOMLINSON, R.W. (1982). Establishment of calibration baseline, NOAA Technical Memorandum NOS NGS 8. US Department of Commerce, National Oceanic and Atmospheric Administration, Rockville, MD, Washington, 25pp.

DYCK, V. A., LAWSON, J.D. and SMITH, J.A. (1984). FORTRAN 77: an introduction to structured problem solving. (Reston, Virginia: Reston), 696pp.

FORSYTHE, G.E., MALCOLM, M.A. and MOLER, C.B. (1977). Computer methods for mathematical computations. (London: Prentice- Hall), 259pp.

FRASER, C.S. (1984). The turbulent transfer model applied to Geodolite measurements, The Canadian Surveyor 38(2): 79-90.

FRASER, C.S. and GRUENDIG, L. (1985). The analysis of photogrammetric deformation measurements on Turtle Mountain, Photogrammetric Engineering and Remote Sensing 51(2): 207-216.

GAO, Y., KRAKIWSKY, E.J. and CZOMPO, J. (1992). Robust testing procedure for detection of multiple blunders, Journal of Surveying Engineering 118(1): 11-23.

GOLUB, G.H. and REINSCH, C. (1970). Singular Value Decomposition and Least Squares Solutions, Numer. Math. 14: 403-420.

GOLUB, G.H. and LOAN, C.F.V. (1990). Matrix computations, second edition. (London: John Hopkins), 642pp.

GRACIE, G. and KRAKIWSKY, E.J. (1987). The need for adjustment and analysis. In: Papers for CIS adjustment and analysis seminars, ed. by E. J. Krakiwsky, Calgary.

GRAFAREND, E.W. (1974). Optimization of geodetic networks, Bolletino di Geodesia a Science Affini 33(4): 352-406.

GRAFAREND, E.W. and SANSO, F. (editors) (1985). Optimization and Design of Geodetic Networks. (Berlin: Springer), 606pp.

GRUENDIG, L. and BAHNDORF, J. (1984). Accuracy and reliability in geodetic networks - program system OPTUN, Journal of Surveying Engineering 110(2): 133-145.

GRUENDIG, L. and TESKEY, W.F. (1984). Improvement of network precision and reliability by the use of scaled distances, Proceedings of the Commission 6 Engineering surveys, Washington, D.C., pp 104-113.

GRUENDIG, L., NEUREITHER, M. and BAHNDORF, J. (1985). Detection and localization of geometrical movements, Journal of Surveying Engineering 111(2): 118-132.

HAWKINS, D. M. (1980). Identification of outliers. (London: Chapman and Hall), 188pp.

HEALY, M.J.R. (1986). Matrices for statistics. (Oxford: Oxford University Press), 89pp.

HECK, B., E. KUNTZ and B. MEIER-HIRMER (1977). Deformationanalyse mittels relativer Fehlerellipsen, Allgemeine Vermessungs- Nachrichten vol. 84: 78-87.

HECK, B., J.J. KOK, W. WELSCH, R. BAUMER, A. CHRZANOWSKI, Y.Q. CHEN and J.M. SECORD (1983). Report of the FIG Working Group on the analysis of deformation measurements. In: Deformation Measurements, Proceedings of the 3rd (FIG) International Symposium on Deformation Measurements, ed. by I. Joó and A'. Detrekői. (Budapest:

Akadémiai Kiadó)

HUGGET, G.R. and SLATER, L.E. (1975). Precision electromagnetic distance measuring instrument for determining secular strain and fault movement, Technophysics 29: 19-27.

JIANJUN, Z. (1991). Robust estimate with minimum squared error, The Australian Surveyor 36(2): 111-115.

JORGENSEN, P.C., KUBIK, K., FREDERIKSEN, P and WENG, W. (1985). Ah, robust estimation, Aust. J. Geod. Photo. Surv. 42: 19-32.

KENNIE, T.J.M. (1990). Electronic angle and distance measurement. In: Engineering surveying technology, ed. by T. J. M. Kennie and G. Petrie. (London: Blackie). 485pp.

KOCH, K.R. (1980). Parameterschätzung und Hypothesentests in linearen Modellen, Dümmler, Bonn.

KOCH, K.R. (1987). Parameter estimation and hypothesis testing in linear models. (London: Springer-Verlag), 378 pp.

KOK, J.J. (1982). Statistical Analysis of Deformation Problems using Baarda's Testing Procedures, 40 Years of Thought, Anniversary Volume on the Ocassions of Prof. Baarda's 65th Birthday Vol II: 470-488, Delft TU Geodetic Computing Centre, Delft.

KUANG, S.L., CHRZANOWSKI, A. and CHEN, Y.Q. (1991). A unified mathematical modelling for the optimal design of monitoring networks, Manuscripta Geodaetica 16: 376-383.

KUBIK, K., WENG, W. and FREDERIKSEN, P. (1985). Oh, gross errors!, Aust. J. Geod. Photo. Surv. 42: 1-18.

KUBIK, K., MERCHANT, D. and SCHENK, T. (1987). Robust estimation in photogrammetry, Photogrammetry Engineering and Remote Sensing 53(2): 167-169.

KUBIK, K., LYONS, L. and MERCHANT, D. (1988). Photogrammetric work without blunders, Photogrammetric engineering and remote sensing 54(1): 51-54.

KUBIK, K. and WANG, Y. (1991). Comparison of different principles for outlier detection, Aust. J. Geod. Photogram. Surv. 54: 67-80.

LAWSON, C.L. and HANSON, R.J.(1974). Solving least squares problems. (London: Prentice-Hall), 340pp.

LEICK, A. (1990). GPS satellite surveying. (New York: John Wiley & sons), 352pp.

MEISSL, P. (1969). Zusammenfassung und Ausbau der inneren Fehlertheorie eines Punkthaufens, Deutsche Geodätische Kommission, Reihe A, Nr. 61, pp. 8-21.

MIKHAIL, E. M. (1976). Observations and Least Squares. (New York: University Press of America), 497pp.

MIKHAIL, E.M. and GRACIE, G. (1981). Analysis and Adjustment of Survey Measurements.

(London: Van Nostrand Reinhold), 340pp.

MILBERT, D.G. (1985). A note on observation decorrelation, variances of residuals, and redundancy numbers, Bull. Geod. 59: 71-80.

NIEMEIER, W. (1979). Zur Kongruenz mehrfach beobachteter geodätischer Netze, Wissenschaftliche Arbeiten der fachrichtung Vermessungswesen Nr. 88, Universität Hannover, Hannover.

NIEMEIER, W., TESKEY, W.F. and LYALL, R.G. (1982). Precision, reliability and sensitivity aspects of an open pit monitoring network, Aust. J. Geod. Photo. Surv. 37: 1-27.

NIEMEIER, W. (1987). Workshop on engineering networks and deformation analysis, School of Surveying, The University of New South Wales, Kensington, N.S.W.

OLLIVER, J.G. and CLENDINNING, J. (1978). Principles of surveying, volume 2, 4th edition. (Wokingham: Van Nostrand Reinhold), 205pp.

PAPO, H. and PERELMUTER, A. (1981). Kinematic analysis of deformation, Personal paper, 16th International Congress, FIG, Montreux.

PAPO, H. (1987). Bases of null-space in analytical photogrammetry, Photogrammetria 41: 233-244.

PELZER, H. (1971). Zur Analyse geodätischer Deformationsmessungen, Deutsche Geodätische Kommission C.164, Munich.

POPE, A.J. (1976). The statistics of residuals and the detection of outliers, NOAA Technical Report NOS 65 NGS 1, National Ocean Service, National Geodetic Survey, US Department of Commerce, Rockville, MD, Washington, 133pp.

PRESS, W.H., FLANNERY, B.P., TEUKOLSKY, S.A. and VETTERLING, W.T. (1988). Numerical recipes: the art of scientific computing. (Cambridge: Cambridge University Press), 818pp.

RAWLINGS, J.O. (1988). Applied regression analysis: a research tool. (California: Wadsworth & Brooks), 553pp.

REUGER, J.M. (1977). Design and use of baselines for the calibration of EDMI, Proceedings of 20th Australian Survey Congress, Darwin.

REUGER, J.M. (1988). Introduction to EDM, second edition, School of Surveying, The University of New South Wales, Monograph, Kensington, N.S.W., 132pp.

RICHARDUS, P. (1984). Project Surveying. (Rotterdam: A.A. Balkema), 628pp.

ROBSON, S. (1994). Personal communication.

ROBSON, S., BREWER, A., COOPER, M.A.R., CLARKE, T.A., CHEN, J., SETAN, H.B. and SHORT, T. (1995). Seeing the Wood from the Trees-An example of optimised digital photogrammetric deformation detection, Paper to be presented at ISPRS Intercommission

Workshop "From Pixels to Sequences", March 1995, Zurich, 6pp.

SAVAGE, J.C. and PRESCOTT, W.H. (1973). Precision of Geodolite distance measurements for determining fault movements, Journal of Geophysical Research 78(26): 6001-6008.

SCHAFFRIN, B. (1981). Some proposals concerning the diagonal Second Order Design of geodetic networks, Manuscripta Geodetica 6(3): 303-326.

SCHWARZ, C.R. and KOK, J.J. (1993). Blunder detection and data snooping in LS and robust adjustments, Journal of Surveying Engineering 119(4): 127-136.

SEAGER, J.W. and SHORTIS, M.R. (1993). Linearisation without calculus - a network adjustment example, The Australian Surveyor 38(3): 197-203.

SEARLE, S.R. (1971). Linear Models. (New York: John Wiley & Sons). 532pp.

SECORD, J.M. (1986). Terrestrial survey methods for precision deformation measurements, Proceedings of the Deformations Measurements Workshop, Oct.31-Nov. 1, Massachusetts Institute of Technology, Cambridge, MA.

SETAN, H.B. (1992). Multiple gross errors detection by Danish method. ESRC Internal report. 35 pp.

SETAN, H.B. (1993a). Applications of Singular Value Decomposition in rank defect Least Squares Estimation. ESRC Internal report. 22 pp.

SETAN, H.B. (1993b). S-transformations and datum re-definition. ESRC Internal report. 34 pp.

SETAN, H.B. (1995). User guide to the deformation detection programs. ESRC Internal report. 21 pp.

SPRENT, A. (1980). EDM calibration in Tasmania, The Australian Surveyor 30(4): 213-227.

STEEVES, R.R. and FRASER, C.S. (1987). Statistical Post- Analysis of Least Squares Adjustment Results, In: Papers For the CISM Adjustment and Analysis Seminars, ed. by E.J. Krakiwsky, Calgary, Alberta.

STEWART, G.W (1973). Introduction to matrix computations. (London: Academic Press), 441pp.

STRANG VAN HEES, G. L. (1982). Variance-covariance transformations of geodetic networks, Manuscripta Geodaetica 7(1): 1-20.

STRAUB, R. (1983). A strategy to trace gross errors. In : Precise levelling: Contributions to the Workshop on precise levelling, ed. by H. Pelzer and W. Niemeier. (Bonn: Dümmler). 490pp.

SZOSTAK-CHRZANOWSKI, A. and A. CHRZANOWSKI (1991). Use of software FEMMA in 2-D and 3-D modelling of ground subsidence, Proceedings of the Second Canadian Conference on Computer Applications in the Mineral Industry, Vancouver, B.C., 15-19 September, pp.689-700.

TESKEY, W.F. (1986). Integrated analysis of deformations, Proceedings of the Deformations Measurements Workshop, Oct.31-Nov. 1, Massachusetts Institute of Technology, Cambridge, MA.

TESKEY, W.F. (1988). Special survey instrumentation for deformation measurements, Journal of Surveying Engineering 114(1): 2-12.

TESKEY, W.F. and BIACS, Z.F. (1990). A PC-based program for adjustment and deformation analysis of precise engineering and monitoring networks, Aust. J. Geod. Photo. Surv. 52: 37-55.

TESKEY, W.F. and BIACS, Z.F. (1991). Geometrical deformation analysis of large concrete roof structure, Journal of Surveying Engineering 117(1): 36-49.

TESKEY, W.F., BAYLY, D.A. and COLQUHOUN, I.R. (1992). Measurement of deformations in buried pipeline, Journal of Surveying Engineering 118(1): 1-10.

TESKEY, W.F. (1994). Personal communication.

TEUNISSEN, P. (1985). Zero order design: generalized inverses, adjustment, the datum problem and S-transformation. In: Optimization and design of geodetic networks, ed. by E. W. Grafarend and F. Sanso. (Berlin: Springer). 606pp.

VANICEK, P. and KRAKIWSKY, E. (1986). Geodesy: the concept, second edition. (Amsterdam: Elsevier), 697pp.

VAN MIERLO, J. (1975). Testing and adjusting 2 and 3 dimensional networks for detecting deformations, Proceedings of I (FIG) International Symposium on Deformation Measurements by Geodetic Methods, Cracow.

VAN MIERLO, J. (1981). A Testing Procedure for Analysing Geodetic Measurements, In: Hallermann, L. (editor) Proceeding of II International Symposium on Deformation Measurements by Geodetic Methods, Bonn.

VINCENTY, T. (1969). The scale problem in adjustments of linear measurements, Survey Review XX(154): 164-170.

VINCENTY, T. (1979). Methods of adjusting relative lateration networks, Survey Review XXV(193): 103-117.

WELLS, D.E., BECK, N., DELIKARAOGLOU, D., KLEUSBERG, A., KRAKIWSKY, E.J., LACHAPELLE, G., LANGLEY, R.B., NAKIBOGLU, M., SCHWARZ, K.P., TRANQUILLA, J.M. and VANICEK, P. (1986). Guide to GPS positioning. Canadian GPS Associates, Fredericton, N.B., Canada.