A B S T R A C T

Aspects of the standard least squares method of locating earthquakes and its extensions are discussed. It is shown that there is a need to carefully separate and distinguish between the statistical and deterministic properties of the least squares solution and the algorithm used to obtain it. Standard linear statistical analysis gives reasonable confidence regions for the hypocentre provided that the errors in the model travel time to pairs of stations are not correlated. The travel time residuals which result from the overdetermined system are unreliable estimates of the model errors, as are the pooled residuals from groups of events whether or not the data are homogeneous.

The concepts of Absolute and Relative hypocentre determination are clarified and the Homogeneous Station method is developed and demonstrated to be a good relative location method. The application of the method to a group of North Island, New Zealand subcrustal earthquakes chosen for homogeneity revealed that the earthquakes occurred in a thin, fairly flat dipping zone that could be as thin as 9 km and is not thicker than 18 km. The result is a significant refinement of previous estimates for New Zealand.

The method of Joint Hypocentre Determination first described by Douglas (1967) is examined. The advantage of the method is that the error in the travel time model is estimated as well as allowing for and estimating the effect of an interaction of this error with the hypocentre parameters of the earthquakes.

The application of this method to groups of North Island, New Zealand earthquakes allows very significant improvements to the travel time model to be made and confirms the result that there is a velocity contrast for both P and S of between six and ten percent between paths in and entirely out of the downgoing Pacific plate. Estimates of the velocities in the plate are $8.6 \pm .1$ km/sec. for P and $4.73 \pm .05$ km/sec. for S. In addition, station terms are calculated which describe the average departure from the new model of travel times to the stations contributing data to the study. These terms may be interpreted as arising from crustal structure local to the station which is different from that of the average crustal model used.

The conclusion is reached that apart from providing better absolute hypocentre estimates, the method of Joint Hypocentre Determination can be made to yield worthwhile information about structure on the scale considered here.

Euan George Campbell SMITH

THE THEORY OF MULTI-EARTHQUAKE LOCATION BY LEAST SQUARES AND APPLICATIONS TO GROUPS OF NORTH ISLAND, NEW ZEALAND MANTLE EARTHQUAKES

Submitted for the degree of Doctor of Philosophy

Mathematics Department
at the Victoria University of Wellington
Wellington
New Zealand

21st September, 1977

ACKNOWLEDGEMENTS

I wish to thank very much my supervisor, Dr Jim Ansell, for his encouragement and help; Dr Mike Saunders of Applied Mathematics Division, D.S.I.R. who pointed out the advantages of the Householder Method for partitioned matrices; the staff of Computer Services Centre, Victoria University of Wellington for their help with the programming; Martin Reyners and John Haines for the use of their unpublished results and for discussions about their results and mine; the staffs of the Mathematics Department, Victoria University of Wellington and Geophysics Division, D.S.I.R. for their interest and support; and by no means least, my wife who typed the manuscript.

C O N T E N T S

	Page No.
INTRODUCTION	1
CHAPTER I: The Least Squares Method of Locating a Single Earthquake	9
CHAPTER II: Properties of the Least Squares Hypocentre Solution	24
CHAPTER III: The Theory of Multiple Earthquake Analyses	38
CHAPTER IV: The Homogeneous Station Method	58
CHAPTER V: Joint Hypocentre Determinations of North Island, New Zealand Mantle Earthquakes	85
CHAPTER VI: Interpretation of the Model Error Estimates from JHD	127
CHAPTER VII: Further Extensions and Other Applications of JHD	175
APPENDIX I: Algorithm for Solving the Joint Equations of Condition by Least Squares	187
APPENDIX II: Calculation of Least Squares Parameter Estimate Variance/Covariance Matrices with Structured Design Matrices and Using Householder Decomposition	192
APPENDIX III: Residuals from Homogeneous Station Method Locations	198
APPENDIX IV: Maximum Likelihood Estimate of a Plane (Orthogonal Least Squares)	202
APPENDIX V: Preliminary Results: Group 24. (Uniform Jeffreys-Bullen Mantle Model)	206
APPENDIX VI: Final Results: Group 24	211
APPENDIX VII: Statistical Significance of Change in Station Term at a Given Station	219
BIBLIOGRAPHY	222

INTRODUCTION

Where are the earthquakes? The problems of answering this question and using the knowledge derived from the answer to draw inferences about the structure of the earth are old and much discussed ones. It was with two principal objects in mind that the work described herein was undertaken. First, could an answer be found to the question: is there a best way of locating earthquakes? and second: how can one make maximum use of the information contained in the arrival times of seismic waves from earthquakes at seismograph stations?

The first question is easiest. The answer must be that the best method is the one for which the predicted uncertainties in the hypocentre estimate are smallest out of that class of methods for which the errors are fairly predicted. By way of elaboration on this we can say that a method is bad if the predicted uncertainties are themselves badly in error.

The most widely used method for locating earthquakes, the minimum sum of squares of residuals method (Geiger (1910) and described in Chapters I and II) is often a bad method not because the travel time or velocity model used is certain to be in error, but because extra information, often only qualitative, on the nature of this error is frequently available which nullifies the statistical assumptions required to predict the likely errors in the hypocentre.

At this point a remark must be made about the use of the "model". Models of several different classes -- statistical, mathematical (other than statistical), physical -- are referred to in this thesis and I have tried to qualify the term each time it is used to make the sense clear. The most important class of model here is the linear statistical model. From the theory of this model, Flinn (1965) gave hypocentre error estimates (as confidence ellipsoids) for the standard least squares method. Fundamental to the application of the theory are the assumptions that the observations (arrival time readings) are identically and normally distributed and independent. None of these assumptions hold although the failure of some are more important than others. First, the observations almost never have errors that are identically distributed. The assumption of equal mean residuals is widely known to be frequently wrong. Freedman (1968) mentions the inequality of the variance of residuals (observed arrival time - predicted arrival time) at different seismograph stations. Physically, the plausibility of assuming equal variance will depend largely on the particular circumstances - on the microearthquake survey level the assumption is usually quite good. Second, distributions of residuals are almost never normal. Moreover they frequently exhibit the form of the teleseismic residuals from nuclear explosions given by Lomnitz (1970), namely a sharp central peak and long tails. Jeffreys (1970) modelled such distributions as the sum of a normal plus uniform distribution, the latter part representing "blunders" -- reading errors with a large variance. Such an explanation can hardly apply to Lomnitz's distribution.

Finally, when sufficient is known about the earth's structure to suggest that the arrival times at neighbouring seismograph stations are likely to be in error by similar amounts, the assumption of independence is invalid.

This last is the most serious. Lindquist (1971) reports that the F distribution on which Flinn's confidence regions depend is robust to mild departures from normality and homogeneity of variance but can be badly upset by lack of independence. Evernden (1969a) shows that an F statistic can be replaced by a X-squared statistic if the number of degrees of freedom is large enough, but the X-squared distribution suffers from the same problems as F. It is with caution then that one determines accuracy of a hypocentre estimate using standard linear theory.

Let us temporarily leave this problem and turn to the second: maximum information extraction. N arrival times for an earthquake are N pieces of information (degrees of freedom) which are divided by the location method into four location parameters (assuming origin time, latitude, longitude and depth are all to be determined) and N residuals which have of course only N-4 d.f. How can one make best use of these N-4 remaining pieces of information? In particular, how can the information be used to improve the travel time or velocity model? In Chapters I and II it is demonstrated theoretically that the residuals are a very poor estimate of the model errors and that combined residuals from groups of earthquakes have properties which make their interpretation difficult.

Notwithstanding past success of methods which process pooled residuals -- the Jeffreys-Bullen (1948) travel time model obtained by this method has been little improved on as a world average travel-time model by the work of Herrin *et al.* (1968) and others -- it was decided not to persist with this method with its uncertainties of convergence (Douglas & Lilwall, 1972) and its inability to predict the correlation between errors in the hypocentre estimates and the error in the derived model. It is possible to make progress with the method as Veith (1975) has shown in his careful study of Kuriles to Okhotsk earthquakes which produced regional travel time corrections.

The most attractive method to date for accomplishing our task is Douglas's (1967) Joint Hypocentre (or Epicentre) Determination method. The method is to estimate together the hypocentres and the average model error between the group of earthquakes and each station contributing information. The method does not require, as is sometimes stated, the assumption of a constant model error over the group of events. The statistical theory of the General Linear Model (see, for example, Zelen, 1962) is all that is required to furnish estimates of uncertainty in and correlation between the parameters.

The greatest difficulty with the method, reported by Douglas in his introductory paper, is one of ill-conditioning of the linear system produced which manifests itself as a near linear dependence between the station terms and hypocentre estimates.

There are different ways to view this phenomenon. First, the purely numerical problem of solving an ill-conditioned system can be reduced a great deal by using higher precision techniques. The systems that arose in this thesis were all sufficiently well conditioned to obtain stable solutions by the methods described herein. Second, the appearance of large values for the correlations between station terms and hypocentre components in the calculated variance matrix of the parameters warns us at least that we are seeking a result at the limit of the resolving power of the data - a circumstance often obscured by other methods.

The ability of JHD and its offshoots to produce better hypocentre estimates has been demonstrated in a wide variety of contexts. Apart from Douglas's (1967) relocation of a group of earthquakes and the nuclear explosion LONGSHOT, work includes Blaney and Gibbs (1968) who used JHD (more correctly JED - epicentres only) to relocate groups of explosions with demonstrably improved relative position estimates. They used a master event, a given event whose position was fixed to its known position to yield absolute locations, or in the case of explosions at Novaya Zenlya, an event given an estimated position to produce locations of the rest of the group relative to the master. The use of a master event vastly improves the conditioning of the system. The price paid is that in the absence of an absolute master, that is, an explosion of known origin (or equivalent), the error in the assumed solution for the master enters into the picture as a systematic error in all the other parameter estimates. Numerous examples are given in this thesis of the rest of the group moving in parallel with a master as the position of the master is varied.

On a different scale, Dewey (1971) employed JHD on a regional scale in a seismicity study of Venezuela and environs. In this study and in preliminary development involving Nevada explosion data, Dewey invariably found that a master event was necessary for stability. Much of Dewey's data for the seismicity study was teleseismic and his depth control depended on *pP*.

Similarly, Billington and Isacks (1975) apparently had too few observations (a total of nine stations) to operate without a master event. Their work shows the power of JHD to provide excellent relative locations as demonstrated by the simple geometry found for the distribution of 600 km Fiji earthquakes.

Even with a master event, considerable information about model errors is forthcoming from the mean model error or station term estimates. In this the method is superior to Evernden's (1969b) method of using the residuals from a master event as station terms to locate a group relative to the master. (JHD also allows for the effect of reading errors in the arrivals from the master.) However, the modelling aspect of JHD has been somewhat overlooked except for the global study of Lilwall and Douglas (1969). Part of the problem involves suitably pooling the mean model error estimates from different groups of earthquakes. Methods such as Bolt and Freedman (1968) suffer from the problem of ill-conditioning and the imposition of a particular form of azimuthal dependence, such as a sine function, runs the risk of obscuring the true nature of the variation if the applied functional form is totally inappropriate.

Some recent approaches to this problem of modelling are presented in Chapter VII. Our efforts with JHD have been directed towards its evaluation in a regional situation (on a similar scale to Dewey's) where a large amount of data is available from stations within 500 km of the earthquakes. The position of the New Zealand seismograph network to record sub-crustal New Zealand earthquakes has no superior in the world. Subcrustal events were chosen because possibilities of ambiguous interpretation of crustal pulses (common from shallow events in New Zealand) are avoided and because S arrivals from these deeper events are largely non-emergent and thus S data can be used to swell the input information. The work of Chapter V demonstrates the effect, by no means disastrous, of not using a master event, in groups of such events, and Chapter VI is an attempt at a regional travel time model improvement using JHD.

In many cases, absolute hypocentre determination is found to be difficult while relative determination is simpler and all that the particular circumstances demand. The elementary technique (from classical analysis of variance) of ensuring no missing observations - the Homogeneous Station method (Ansell and Smith (1975) and described in full in Chapter IV) is one way of achieving this. The idea is not new but its importance seems to have been overlooked. The excellent relocations obtained by Engdahl (1972) in the Aleutians are largely due to his homogeneity of data.

The occupation of this thesis with least squares methods in their simplest forms has meant that the fairly large class of alternative methods of location proposed in recent years has been pretty-well overlooked. Such methods include James *et al.* (1969) method of computing the origin time from S-P values and advancing to a solution by means of a two-stage iterative process; Keilis-Borok *et al.* (1972) who apply empirically determined functions to the incremental improvement to the hypocentre to improve stability; and Lomnitz's (1977) very novel use of distance rather than time residuals. Most methods involve least squares somewhere. Their objects can be summarised as superior stability. In my experience, limited to regional work but involving depth determinations without special phases, adequate stability can be attained by suitable damping of the incremental improvements (Hartley, 1961). The use of damping also almost always decreases the number of iterations required to achieve a satisfactory hypocentre.

The achievements of this work can be summarised as follows. The Joint Hypocentre Determination method is demonstrated on the regional scale to be a good absolute and relative location method and also to provide worthwhile information for model improvement provided the mild restrictions of Chapter III are observed. The Homogeneous Station method is shown to be an excellent relative location method which can be applied to small networks of stations. The application of these tools to other regions and their extension to crustal earthquakes on the regional level and to micro-earthquakes is work for the future.

CHAPTER I

THE LEAST SQUARES METHOD OF LOCATING A SINGLE EARTHQUAKE

This chapter is in part a summary of known results about the standard least squares method of determining the longitude, latitude, depth and origin time of a single earthquake from body wave arrivals but also includes points about the method which have not before been clearly explained. In this section the notation of following chapters is established.

THE EQUATIONS OF CONDITION

Suppose that an earthquake occurs at an unknown point $\underline{x} = (h, x, y, z)$ in space-time, where h is the origin time, x and y are the longitude and latitude and z is the depth of the source. A set of n body wave arrival times α_j are obtained from seismograph stations. We have models for the travel times t_j of the seismic waves, typically tabulated values of the time given the distance d_j and depth of the source z from the receiving station. We then seek the values h, x, y, z which minimise:

$$s(\underline{x}) = \sum_{j=1}^{n} (\alpha_j - t_j(d_j, z) - h)^2 \qquad \ldots(1.1)$$

and we use these quantities as estimates of the hypocentral parameters. This is the standard least squares hypocentre estimate first suggested by Geiger (1910) but only popularised after the advent of high-speed digital computers made feasible, algorithmically, the finding of the least squares solution.

The numerous references in the literature include Bolt (1960) who describes a computer algorithm for finding the solution and Flinn (1965) who analyses the solution statistically.

It is important to distinguish between the solution which has the least squares property and the method of arriving at the solution.

First, let us deal briefly with the method. From a trial solution ${}_0\underline{x}$ a set of improvements to the trial ${}_0\hat{\delta}\underline{x}$ are calculated by solving the set of equations of condition which result when the travel time model is linearised. From Taylor's Theorem:

$$t_j(d_j,\ z) = t_j({}_0d_j,\ {}_0z) + \left(\frac{\partial t_j}{\partial d}\frac{\partial d}{\partial x}\right){}_0\delta x + \left(\frac{\partial t_j}{\partial d}\frac{\partial d}{\partial y}\right){}_0\delta y + \frac{\partial t_j}{\partial z}{}_0\delta z + {}_0q_j \qquad \dots(1.2)$$

where $\underline{x} = {}_0\underline{x} + {}_0\delta\underline{x}$ and we assume that the travel time model is continuously differentiable, so that q_j, the remainder term or linearisation error, has the property that:

$$|q_j|/\|{}_0\delta\underline{x}\| \rightarrow 0 \quad \text{as} \quad \|{}_0\delta\underline{x}\| \rightarrow 0.$$

Let ε_j be the reading error associated with the observation α_j. Denote by s_j the difference between the true travel time $(\alpha_j + \varepsilon_j - h)$ and the model travel time $t_j(d_j,\ z)$. We have then the equations of condition:

$$\begin{aligned} &{}_0\delta h + \left(\frac{\partial t_j}{\partial d}\frac{\partial d}{\partial x}\right){}_0\delta x \\ &+ \left(\frac{\partial t_j}{\partial d}\frac{\partial d}{\partial y}\right){}_0\delta y + \left(\frac{\partial t_j}{\partial s}\right){}_0\delta y = \alpha_j - t_j({}_0d_j,\ z) - {}_0h + \varepsilon_j - s_j - q_j \qquad \dots(1.3)\end{aligned}$$

which may be solved by premultiplying both sides by the transpose of the matrix A, the j^{th} row of which is:

$$\left(1 \quad \frac{\partial t_j}{\partial d} \frac{\partial d}{\partial x} \quad \frac{\partial t_j}{\partial d} \frac{\partial d}{\partial y} \quad \frac{\partial t_j}{\partial z}\right)$$

and solving the system:

$$A^T A\, {}_0\hat{\delta}\underline{x} = A^T\, {}_0\underline{y} \qquad \ldots(1.4)$$

where ${}_0y_j = \alpha_j - t_j({}_0d_j,\ z) - {}_0h$. An equivalent method such as orthogonal transformation of the equations of condition (Householder Method, Householder (1953)) may be used. The resulting estimates ${}_0\hat{\delta}\underline{x}$ have the property that:

$$\sum_{j=1}^{n} \left(\alpha_j - t_j({}_0d_j,\ z) - {}_0h - {}_0\hat{\delta}h - \frac{\partial t_j}{\partial d}\frac{\partial d}{\partial x}{}_0\hat{\delta}x - \frac{\partial t_j}{\partial d}\frac{\partial d}{\partial z}{}_0\hat{\delta}y - \frac{\partial t_j}{\partial z}{}_0\hat{\delta}z\right)^2$$

is an absolute minimum over all possible values of $\delta\underline{x}$. A new hypocentre estimate is now ${}_1\underline{x} = {}_0\underline{x} + {}_0\hat{\delta}\underline{x}$, and we proceed iteratively until the increments ${}_i\hat{\delta}\underline{x}$ are considered negligible. The general conditions for the convergence of such a scheme are given by Hartley (1961) and depend on the goodness of the linear approximation to the model. Hartley concludes by showing that the only way to ensure convergence to the absolute minimum, which is necessary for statistical purposes, when there may be secondary local minima, is to start in the "well" of the absolute minimum; that is, in some region R about the absolute minimum in which we can find ${}_0\underline{x}$ such that:

$$S({}_0\underline{x}) < \liminf_{\underline{x}\varepsilon R} S(\underline{x})$$

In practice, it is usual not to check Hartley's criteria for convergence but to wait and see if the iterative scheme converges. It has been observed that convergence generally occurs when there is a good distribution of stations about the hypocentre. Flinn (1965) points out that this condition of good distribution is equivalent to a well conditioned system of equations of condition.

CONVERGENCE TO THE LEAST SQUARES SOLUTION

We shall discuss this convergence question a little more before proceeding to discuss the properties of the solution itself.

First, let us note that from Taylor's Theorem, similarly to *(1.2)*:

$$\begin{aligned} {}_{i+1}t_j({}_{i+1}d_j,\ {}_{i+1}z) &= {}_it_j({}_id_j,\ {}_iz) + \left(\frac{\partial\, {}_it_j}{\partial_d}\,\frac{\partial_d}{\partial_x}\right)_i \hat{\delta}x + \left(\frac{\partial\, {}_it_j}{\partial_d}\,\frac{\partial_d}{\partial_y}\right)_i \hat{\delta}y \\ &\quad + \left(\frac{\partial\, {}_it_j}{\partial_z}\right)_i \hat{\delta}z + q_{ij} \qquad \ldots(1.5) \end{aligned}$$

Thus the j^{th} term of the right hand side of *(1.3)* for the $(i+1)^{\text{th}}$ iteration,

$$ {}_{i+1}y_j = (\alpha_j - {}_{i+1}t_j - {}_{i+1}h) $$

differs from the j^{th} residual of the i^{th} iteration:

$$\begin{aligned} {}_ir_j &= {}_iy_j - {}_i\hat{\delta}_h - \left(\frac{\partial\, {}_it_j}{\partial_d}\,\frac{\partial_d}{\partial_x}\right)_i \hat{\delta}x - \left(\frac{\partial\, {}_it_j}{\partial_d}\,\frac{\partial_d}{\partial_y}\right)_i \hat{\delta}y \\ &\quad - \left(\frac{\partial\, {}_it_j}{\partial_z}\right)_i \hat{\delta}z \end{aligned}$$

only by the amount q_{ij}; explicitly we have:

$$_{i+1}y_j = {}_ir_j - q_{ij} \qquad \ldots(1.6)$$

The residuals $_i\underline{r}$ have the property that:

$$_iA^T {}_i\underline{r} = 0 \qquad \ldots(1.7)$$

$_iA^T$ denoting the transpose of $_iA$, and the $(i+1)^{\text{th}}$ increment $_{i+1}\hat{\delta\underline{x}}$ is the solution of:

$$_{i+1}A^T {}_{i+1}A {}_{i+1}\hat{\delta\underline{x}} = {}_{i+1}A^T {}_{i+1}\underline{y} \qquad \ldots(1.8)$$

If we write $_{i+1}A = {}_iA + {}_i\delta A$, the elements of $_i\delta A$ will be differences between first derivatives of the travel time model at points $_i\hat{\delta\underline{x}}$ apart.

The right hand side of *(1.8)* is thus:

$$({}_iA^T + {}_i\delta A^T)({}_i\underline{r} - \underline{q}_i)$$

which, by *(1.7)*, equals:

$$- {}_iA^T\underline{q}_i + {}_i\delta A^T {}_i\underline{r} - {}_i\delta A^T\underline{q}_i$$

For a perfectly linear model, this term is zero implying that one iteration would suffice, except for the possible limitations of finite arithmetic, to find the absolute minimum. Otherwise this quantity depends on the linearisation error and the rate of change of the first derivatives of the model. Thus the rate of convergence will depend on the smoothness and flatness of the model.

Further, the definition of the residual vector ${}_i\underline{r}$

$$ {}_i\underline{r} = {}_i\underline{y} - {}_iA\,{}_i\hat{\delta}\underline{x} \qquad \ldots(1.9) $$

and *(1.4)* show that the least squares process decomposes ${}_i\underline{y}$ into two orthogonal vectors ${}_i\underline{r}$ and ${}_iA\,{}_i\hat{\delta}\underline{x}$ so that:

$$ \|{}_i\underline{y}\|^2 = \|{}_i\underline{r}\|^2 + \|{}_iA\,{}_i\hat{\delta}\underline{x}\|^2 \qquad \ldots(1.10) $$

so that for $\|{}_i\hat{\delta}\underline{x}\| > 0$

$$ \|{}_i\underline{y}\|^2 > \|{}_i\underline{r}\|^2 $$

From *(1.6)*:

$$ \begin{aligned} \|{}_{i+1}\underline{y}\| &\leq \|{}_i\underline{r}\| + \|\underline{e}_i\| \\ &< \|{}_i\underline{y}\| + \|\underline{e}_i\| \end{aligned} $$

so that in the presence of a small linearisation error we might hope to get a sequence of right hand sides getting smaller and smaller but with large linearisation errors, causable in practice by starting a long way from the solution, there is no guaranteed convergence.

Flinn (1965) has examined the linearisation error with the particular view of observing its effects on the confidence regions he calculates for the hypocentre parameters (which will be discussed in a later chapter).

I have used as examples and tests for the results developed in this thesis, locations of New Zealand mantle earthquakes using arrival time data from New Zealand stations only. These stations are always within about 800 km of the epicentres calculated herein and the hypocentres are never deeper than 400 km. The Jeffreys-Bullen Seismological Tables are used as our (initial) model for two reasons: It is the model used by the New Zealand Seismological Observatory in its routine hypocentre determinations; and no better model describing the travel times in the mantle under the North Island hitherto existed. Examination of linearisation errors caused by the quadratic interpolation routine used in this thesis to interpolate in the tables in the distance/depth range of interest revealed that the errors were largest near $z = 33$ km, $d = 0$ where the errors can be several seconds if the distance and depth errors in the final solution are greater than 50 km. The linearisation error decreased with increasing distance and depth, but for hypocentre errors of 50 km the linearisation error is always of the order of one second. When the distance and depth errors are 20 km, the linearisation errors are never greater than 0.8 seconds but always of order 0.2 seconds and when the distance and depth errors are 10 km, the linearisation error is never greater than 0.2 seconds. The magnitude of the linearisation error (in seconds) for errors of 50 km and 10 km at short epicentral distance is given in Table 1.1.

TABLE 1.1

MAXIMUM LINEARISATION ERROR (SEC.)

For an error in epicentral distance and depth of $\leq$ 50 km.

d (degrees) \ Z (km)	96	159	222	285	348	411
.5	4.7	2.75	1.6	1.1	.85	.55
1.0	3.6	2.65	1.9	1.25	1.0	.75
1.5	2.65	2.05	1.6	1.35	.9	.65
2.0	2.35	1.9	1.55	1.3	1.1	.65
2.5	1.6	1.75	1.45	1.15	1.05	.75
3.0	1.25	1.35	1.30	1.05	1.0	.8

For an error in epicentral distance and depth $\leq$ 10 km.

d \ Z	96	159	222	285	348	411
.5	.19	.11	.06	.04	.03	.02
1.0	.14	.11	.08	.05	.04	.03
1.5	.11	.08	.06	.05	.04	.03
2.0	.09	.08	.06	.05	.04	.03
2.5	.06	.07	.06	.05	.04	.03
3.0	.05	.05	.05	.04	.04	.03

IMPROVING THE CONDITIONING IN THE STANDARD METHOD

We have not yet discussed the implications of finite arithmetic in the problem of convergence. A poorly conditioned system of equations has the property that small changes in the right hand side ${}_i\underline{y}$ of *(1.4)* may give rise to disproportionately large increments ${}_i\delta\underline{\hat{x}}$. A measure of the conditioning is the condition number λ of the matrix A^TA (equation *(1.4)*) defined by:

$$\lambda = \sup \|A^TA\underline{x}\| \cdot \sup \|{A^TA}^{-1}\underline{x}\| \quad , \quad \|x\| = 1$$

$$\equiv \sup \|A^TA\underline{x}\| \,/\, \inf \|A^TA\underline{x}\| \quad , \quad \|x\| = 1 \qquad \ldots(1.12)$$

Thus λ is numerically equal to the quotient of the greatest and least eigenvalues of A^TA or ${A^TA}^{-1}$ (see Dahlquist (1972) as a general reference).

λ is a measure of the greatest magnification that can take place when solving:

$$A^TA\underline{x} = A^T\underline{y} \qquad \ldots(1.4a)$$

by which we mean the following. Solving the above and taking moduli we have:

$$\|\underline{x}\| = \|{A^TA}^{-1}A^T\underline{y}\|$$

$$= \|A^T\underline{y}\| \cdot \|{A^TA}^{-1}\underline{u}\|$$

where $\underline{u}$ is a unit vector parallel to $A^T\underline{y}$.

So
$$\|\underline{x}\| \leq \|A^T\underline{y}\| \sup \|A^TA^{-1}\underline{w}\| \,, \quad \|\underline{w}\| = 1$$
$$= \|A^T\underline{y}\| \Lambda_{\max} \qquad \ldots(1.13)$$

where $\Lambda_{\max}$ is the greatest eigenvalue of $(A^TA)^{-1}$. If $\Lambda_{\min}$ is the least eigenvalue ($\neq 0$ if $(A^TA)^{-1}$ exists) then $1/\Lambda_{\min}$ is the largest eigenvalue of A^TA. Using this as a unit for comparison, that is, expressing $\|A^T\underline{y}\|$ as a multiple of $(1/\Lambda_{\min})$:

$$\|A^T\underline{y}\| = M(1/\Lambda_{\min})$$

we have:

$$\|x\| \leq M\lambda \qquad \ldots(1.14)$$

Alternatively we can think of scaling *(1.4a)* so that $\Lambda_{\min} = 1$. Thus λ gives a bound for the magnification which can be achieved when $A^T\underline{y}$ is parallel to the eigenvector corresponding to greatest eigenvalue of A^TA^{-1}.

The second and third columns of the matrix A contain derivatives of the travel time model with respect to distance multiplied by quantities which are well approximated in practice by the cosine and sine of the azimuth of the station from the earthquake (Bullen, 1965). In view of the comparatively small changes in $\partial t/\partial d$ which occur over a wide range of values of (d, z) in models like the Jeffreys-Bullen model (Jeffreys and Bullen, 1948), a wide range of azimuths is required to provide good conditioning.

The last column contains the entries $\partial t_j/\partial z$. This quantity is positive when the ray leaves the source upward -- typically when the epicentral distance is less than the depth -- and negative when the ray leaves downward. For models in which there is an increase in velocity with depth (as distinct from layered models with constant velocities within the layers) there is a tendency for the ray to quickly become flatter at the source. For a wide range of distances $\partial t/\partial z$ will be of small magnitude. Ideally one requires several stations sufficiently close to the source that the rays are steeply upward to these stations, and other stations at a distance where the ray leaves downward.

For example, using the J-B model and considering a 200 km deep event, $\partial t/\partial z$ *(P)* is 1.0 sec./10 km at 100 km from the source and decreases rapidly to .02 sec./10 km at 800 km and continues decreasing more slowly to -0.5 sec./10 km 1900 km (where there is a discontinuity) and does not reach -1.0 sec./10 km until the distance is 5000 km. (The choice of units here will be explained later.) More improvement is obtained if a range of phases which give widely differing values of $\partial t/\partial d$ are available such as *S*, *pP* and core reflections, presuming that *P* is the principal source of information.

It is important to note that the inclusion of extra phases does not automatically provide a better hypocentre estimate. If for some reason the model travel time for such an auxiliary phase has larger errors than allowed for (for example, if *S* is used with a model very much poorer than the *P* model used but with no downweighting of the *S* data) then the solution is likely to be poorer. What these phases do ensure is that, by improving conditioning, an error of given size is less likely to provoke a magnified error in the parameter estimates.

A recent discussion of conditioning of the equations of condition by Buland (1976) demonstrates the value of the Householder method. Buland fails to note the advantage of scaling the hypocentre parameters, that is, choosing units of d and z so that $\partial t/\partial d$ and $\partial t/\partial z$ are the same order of magnitude and the same magnitude as the first column of $A^T A$, the origin time column, which is a column of 1's. As already indicated, units of 10 km for depth (or larger) are commonly required to produce numerical values of $\partial t/\partial z$ close to 1 in magnitude and 10 km is an appropriate unit for d as well. The value of this scaling is fully explained in Smith (1976).

A selection of condition numbers for hypocentres discussed in Chapter IV are plotted in Figure 1.1. The events were located using a P arrival from end of the seven stations shown in the figure. A cross section from SE to NW across the middle of the North Island showing the depth distribution is given in Figure 4.4. The deepest assigned depths are about 270 km along the TNZ-KRP line with the events along the WEL-GNZ line being assigned depths of about 80 km. It is seen immediately that the condition number rapidly increases as one moves outside the convex hull of the seismograph stations, and as the assigned depth increases.

While on the subject of conditioning and stability, we mention a simple device which greatly improves the chances of convergence which is to damp the increments, so that the i+1$^{\text{th}}$

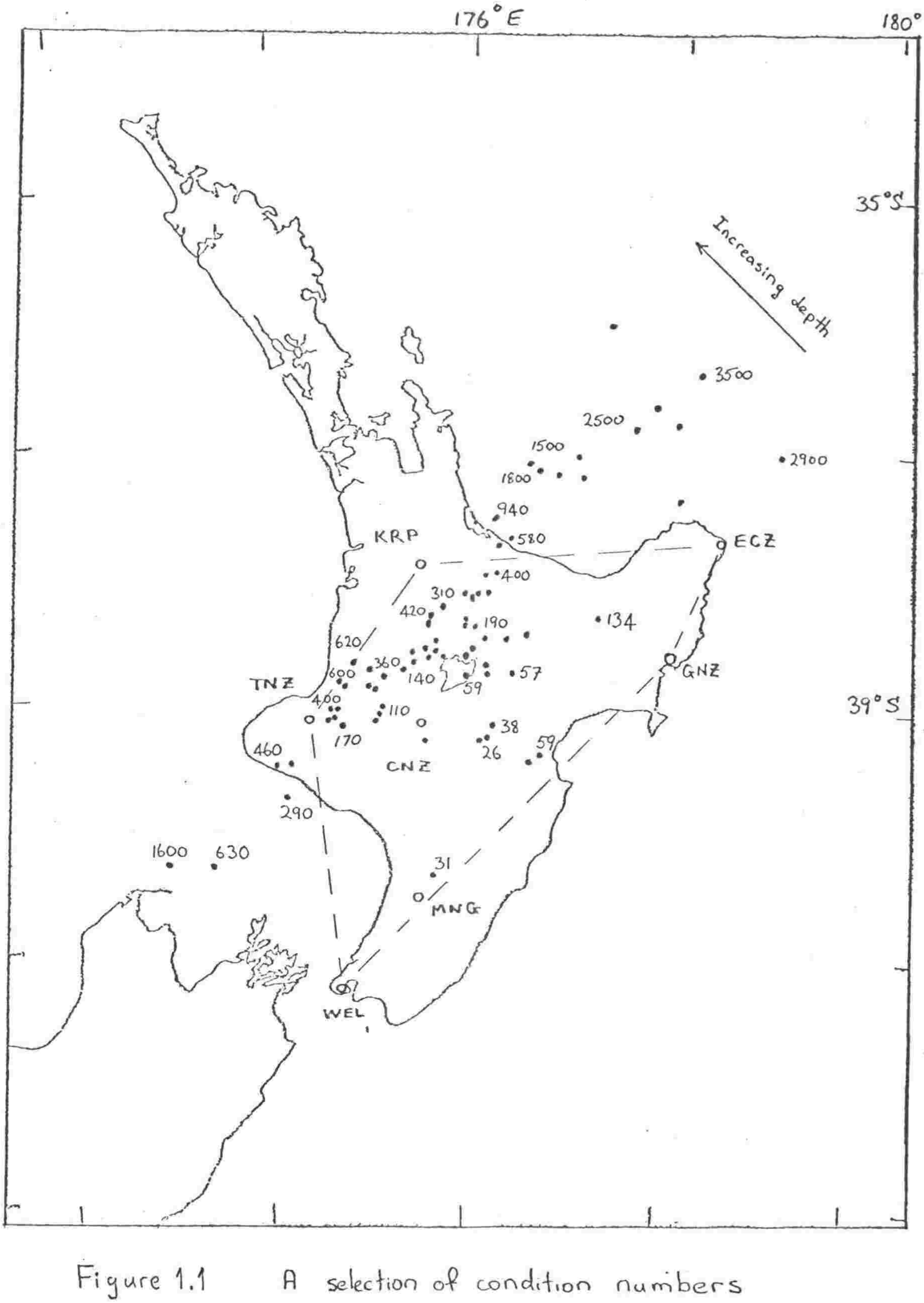


Figure 1.1 A selection of condition numbers

estimate is ${}_i\underline{x} + \phi({}_i\hat{\delta}\underline{x})$, $0 < \phi \leq 1$ (Hartley 1961). In principle, ϕ is chosen to maximise the improvement to the residual sum of squares but, in practice, may be predicted by the apparent rate of convergence. For example, a scheme used herein is: if the magnitude of the second increment is very much less than the first, set $\phi = 1$. If the increments are of the same order, set $\phi = \frac{1}{2}$. If the second increment is larger than the first, set $\phi = \frac{1}{2}\|{}_i\hat{\delta}\underline{x}\| / \|{}_{i+1}\hat{\delta}\underline{x}\|$. More elaborate schemes, such as described by Marquardt (1963), are computationally expensive and not justifiable in most cases.

To conclude this chapter, we mention other means of obtaining hypocentre estimates which are generally employed when the system of equations of condition is ill-conditioned.

Keilis-Borok (1971) uses damping to determine teleseismic hypocentres from a small network of stations in a geographical region which is small compared to the mean epicentral distance. James *et al.* (1969) and others have used the method whereby the origin time is restrained by some means, often from examination of P and S pairs of arrivals, and then the best fit of the latitude, longitude and depth is found. This method works essentially because the near linear dependence of depth and origin time in the equations of condition, caused by the numerically small values of $\frac{\partial t}{\partial z}$ for many values of (d, z), is removed by fixing the origin time. This principle can be extended to the standard method by fixing the origin time, finding a new epicentre and depth and then adjusting the origin time by the mean value of the residuals, since in the absence of an origin time term in the equations of condition, the mean residual need not be zero. (See Chapter II)

Also common is the restriction of the depth of the event when there is data such as a core or surface reflection which indicates a particular depth. In the case of a local earthquake, if the existence of crustal phases indicates a shallow depth and, in the absence of a station near enough to give depth control by providing a sufficiently large numerical value of $\frac{\partial t}{\partial z}$, the depth may be constrained to some nominal value such as 33 km.

The stepwise multiple regression approach to solving the equations of condition (Lee and Lahr (1975)) is another approach to stabilising the iterative means by updating those parameters which make the statistically most significant (by linear theory) improvement to the sum of squares.

The value of solutions obtained by these methods will be discussed in the next chapter. Methods which employ information from more than one event will be discussed in subsequent chapters.

CHAPTER II

PROPERTIES OF THE LEAST SQUARES HYPOCENTRE SOLUTION

ALGEBRAIC PROPERTIES OF THE SOLUTION

We have obtained the hypocentre solution ${}_*\underline{x}$ with the property that:

$${}_*\hat{\delta}\underline{x} = (A^TA)^{-1}A^T{}_*\underline{y} \cong 0 \qquad \ldots(2.1)$$

The vector of residuals for this solution is ${}_*\underline{y}$ which has the property (from *(2.1)*) that:

$$A^T{}_*\underline{y} \cong 0 \qquad \ldots(2.2)$$

Thus the number of degrees of freedom of the residuals are reduced by the number of parameters estimated, generally four. Equations similar to *(1.3)* give the relationship between the residuals, model errors and the error in the final hypocentre:

$$A{}_*\delta\underline{x} = {}_*\underline{y} - \underline{s} + \underline{\varepsilon} - {}_*\underline{q} \qquad \ldots(2.3)$$

Denote the total error $-\underline{s} + \underline{\varepsilon} - {}_*\underline{q}$ by ${}_*\underline{e}$. Then ${}_*\underline{e} = A{}_*\delta\underline{x} - {}_*\underline{y}$ and the relation *(2.2)* gives:

$$||{}_*\underline{e}||^2 = ||A{}_*\delta\underline{x}||^2 + ||{}_*\underline{y}||^2 \qquad \ldots(2.4)$$

So that $||\underline{e}|| \geq ||{}_*\underline{y}||$. This is important when the residuals are used to obtain extra information. For example, if corrections to the travel time model are obtained from an analysis of residuals, there will be a tendency to underestimate the model error.

The exact error ${}_*\delta\underline{x}$ is given by:

$$ {}_*\delta\underline{x} \;=\; (A^TA)^{-1}A^T(-\,\underline{s} + \underline{\varepsilon} - {}_*\underline{q}) \qquad \ldots(2.5) $$

Let us assume we have N seismograph stations observing M earthquakes in a geographical region sufficiently small that the model error s_j can be assumed constant for each station $j = 1\ldots N$ at the expense of introducing an error negligible compared to $\|\,\underline{\varepsilon}\,\|$. Each station may not record each event. The equations of condition for the i^{th} solution are:

$$ P_iA_i\,{}_*\delta\underline{x}_i \;=\; P_i\,{}_*\underline{y}_i - \sum_{j=1}^{N} s_j(P_i\underline{u}_j) + P_i(\underline{\varepsilon}_i - {}_*\underline{q}_i) \quad ..(2.6) $$

where P_i is an $N{\times}N$ diagonal matrix with 1 in the j^{th} position if station j records event i and zero otherwise, and $\underline{u}_j$ is the j^{th} unit vector with 1 in the j^{th} position and zero elsewhere.

From *(2.5)* we have (since $P_i^TP_i = P_i$):

$$ {}_*\delta\underline{x}_i \;=\; \sum_{j=1}^{N} s_j(A_i^TP_iA_i)^{-1}A_i^TP_i\underline{u}_j + (A_i^TP_iA_i)^{-1}A_i^TP_i(\underline{\varepsilon}_i - {}_*\underline{q}_i) \qquad \ldots(2.7) $$

This shows that the contribution to the final error ${}_*\delta\underline{x}_i$ by an error s_j at station j will depend on which other stations were present. It is clear that there is likely to be a systematic error in the ${}_*\delta\underline{x}_i$ if there is some systematic variation with geographic position in the terms s_j, a phenomenon referred to as source biasing. The likelihood of this increases as the number of stations which record all the earthquakes increases. A discussion of whether the least squares estimated is biased in the strict sense will be discussed later in this chapter.

Before turning to the statistical properties of the solution, let us return to equation *(2.5)* and write $\bar{s} = \sum_{j=1}^{n} s_j/n$. Let $\underline{1}$ be a vector of N components, each of which is 1. Let $\delta s_j = s_j - \bar{s}$. Then:

$$ {}_*\delta\underline{x} = (A^TA)^{-1}A^T\{-(\bar{s}\underline{1} + \delta\underline{s}) + \underline{\varepsilon} - {}_*\underline{q}\} \quad ..(2.8) $$

Since the first column of A is $\underline{1}$, $(A^TA)^{-1}A^T\underline{1}$ is a vector with first component, corresponding to the origin time correction, equal to 1 and other components zero. Thus the average model error $\bar{s}$ will produce an error in the origin time of $-\bar{s}$, but will not contribute to the error in the other hypocentre parameters. Because the average model error is always transferred to the origin time, it is impossible from earthquake location data alone to determine the average error of a model. A similar conclusion is reached by Lomnitz (1970) who discusses the problem of improving the accuracy of P travel-time models using earthquake data. Note that there may also be a contribution to the origin time error from the remainder term $\delta\underline{s}$ since the columns of A are not orthogonal.

STATISTICAL PROPERTIES OF THE SOLUTION

We now consider the statistical properties of the least squares solution. In particular, it would be desirable to say how accurate the solution is in terms of confidence intervals for the parameters or joint confidence regions for combinations of the parameters. Flinn (1965) extensively discusses this problem using the assumption that the errors, except for the linearisation error, are normally distributed and independent. He shows that standard linear confidence ellipses based on the use of

$$\hat{\sigma}^2 (A^T A)^{-1}$$

for the estimate of the variance of the parameters, where $\hat{\sigma}^2 = (\sum_{j=1}^{n} {}_{*}y_j^{\,2})/(n-p)$, p being the number of parameters estimated, are satisfactory if the dimensions of the confidence region are small enough to preclude the possibility of the linearisation error being significant compared to the other errors.

The actual distribution of the errors has been a subject of much discussion, controversy and confusion. It is important to realise that the errors come from three sources, as shown by equation *(2.3)*, and that these three errors are not independent, since the linearisation error depends on the location error induced by the model and reading errors.

The reading error is the simplest and most easily dealt with. Freedman (1968) gives a very complete discussion of the reading error and concludes that, if certain precautions are taken, the distribution of reading errors for a particular phase at a particular station will be normal. The precautions must exclude the possibilities of blunders by the reader, failure to identify the correct phase and that the arrivals should not be of a different character as might result from earthquakes of greatly differing magnitude. Models which allow for the existence of such effects can be constructed. Jeffreys (1970) gives the well-known model of normal plus low level uniform distribution to account for blunders and derives a method of weighting observations to remove the effects of blunders. I have derived simple models to account for the possibility of identifying the wrong crustal phase in a situation where two arrivals are expected but only one is observed. These distributions all tend to have the characteristic of a sharp central peak and long persisting tails which are frequently observed in residuals. (See Lomnitz (1970), Freedman (1966b), (1967) and Figure 5.3.

We now come to the problem of modelling the travel time model errors. With the increasing quality of instruments and concomitant improvement in the precision of observations, it is clear that the distribution of residuals obtained from explosion data as in Lomnitz (1970) or Dewey (1971) must reflect the distribution of model errors for these events and that these errors with standard deviations of several seconds tend to be very much greater than the reading errors with standard deviations of a few tenths of a second.

The author was not in the fortunate position of having explosion data available to supplement the earthquake data used in this study. Mention will be made later of the use of such data. Many of the people investigating different location techniques have used explosion data to test their methods. See, for example, Douglas (1967) and Dewey (1971).

The basic difference between explosion residuals and earthquake residuals is that since the location of the explosion is known, the residuals are numerically equal to the total errors with full number of degrees of freedom (equation *(2.3)* when $_*\delta\underline{x} = 0$) whereas in the case of earthquakes the residuals are a linear function of the total error, viz.:

$$_*\underline{y} = \{A(A^TA)^{-1}A^T - I\}_*\underline{e} \qquad \ldots(2.9)$$

with the number of degrees of freedom fewer by p, the number of hypocentre parameters. From *(2.9)* certain deductions can be made about the distribution of the errors. For example, if the distribution of the residuals is not normal then the errors cannot be normally distributed. The operator $\{A(A^TA)^{-1}A^T - I\}$ is a projection from the n dimensional space of the errors to the n-p dimensional subspace which is the null space of the operator A^T. For any postulated error distribution, the distribution of the resultant residual distribution may be readily compared with reality, but, of course, there will be an infinite number of possible error distributions which give the same residual distribution.

MODELLING THE MODEL ERROR AS A RANDOM VARIABLE

We have hitherto begged the question somewhat in talking about the distribution of the travel time model errors, since we have not clearly stated how $\underline{s}$ is a random variable. Consider the following experiment in which, to observe a particular earthquake, a random sample of n points is made from some region of the earth's surface Ω. Define Ω_z to be the subset of Ω such that if $x \in \Omega_z$

$$s(x) \leq z$$

Define:

$$F_s(z) = \int_{\Omega_z} d\tau \Big/ \int_{\Omega} d\tau \qquad \ldots(2.10)$$

F_s is clearly a distribution function and we equate the probability of picking a point from Ω for which the numerical value of the travel time model error is less than z to $F_s(z)$. If the seismograph stations used in the location of the earthquake were chosen in the manner suggested, then the components of $\underline{s}$ would be independent with marginal distribution function $F_s(z)$. Ignoring for the moment the obvious objection that this is not how observations of an earthquake are made, let us continue and calculate the expected value of $\underline{s}$, $E(\underline{s})$. If we define $\mu_s = \int_{-\infty}^{\infty} s\,dF_s$, then:

$$E(\underline{s}) = \mu_s \underline{1} \qquad \ldots(2.11)$$

In order that we may apply the statistical theory of linear models to *(2.3)*, it is mandatory that:

$$E({}_*\underline{y}) = A_* \delta \underline{x} \qquad \ldots(2.12)$$

or equivalently:

$$E({}_*\underline{e}) = 0 \qquad \text{...(2.13)}$$

In our case:

$$E({}_*\underline{e}) = E(-\underline{s}) + E(\underline{\varepsilon}) + E(-{}_*\underline{q}) \qquad \text{...(2.14)}$$

It is fair to take $E(\underline{\varepsilon}) = \underline{0}$. Freedman (1966a) indicates that the mean reading error will be small compared to its standard deviation and we may neglect this quantity as making a negligible contribution to the location error. The linearisation error is a more difficult case and we shall merely assume that the location error is sufficiently small to make these errors negligible compared to the model errors. Unless we are in the position of having several stations within one or two degrees, the linearisation error can be neglected for location errors up to 30-40 km. (See Table 1.1)

Thus we are left with:

$$E({}_*\underline{e}) \cong E(-\underline{s})$$

$$= -\mu_s \underline{1} \qquad \text{...(2.15)}$$

so that the least squares hypocentre estimate will be a biased estimate, but only the origin time will be biased, by the amount $-\mu_s$, because, as we have seen, such an error affects only the origin time.

Under these circumstances, if the distribution of model errors were normal, then the linear confidence regions of Flinn (1965) would be valid for the parameters other than the origin time. The studies of D.W. Norton, reported in Lindquist (1971), compared the *F* distribution with the analogous distributions resulting when the distribution of the errors were neither normal nor homogeneous in form or variance. Norton concluded that the effect of distributions which were sharply peaked with long tails (leptokurtic) would produce confidence regions based on the *F* distribution which were optimistic by a few percent. Skew distributions had little effect on the validity of the use of *F*. If the observations were drawn from populations with markedly unequal form or variance, the resulting confidence region will again be optimistic by a few percent. Independence of the observations is however a very important requirement. Thus with the idealised conditions described above for making the arrival time observations, *F* based confidence regions would be reasonable, if slightly optimistic, estimates of the true confidence region, since under our scheme there is homogeneity of distribution for observations of a particular phase and this distribution may be like Lomnitz's (1970) leptokurtic distribution.

More realistically, seismograph stations are sited to provide what is hoped to be a maximum of information about the earthquakes they record and the regions through which seismic waves pass, and the stations used to locate an event are a subset of those which have information which is at the disposal of the seismologist making the locations.

Consider then a single earthquake and let Ω be the region in which the event might be recorded at a seismograph station. Within this region there are n seismographs which present arrival time information to the seismologist. The difference between the prior situation and this one is the difference between the prior and posterior sampling of Ω. As soon as the seismologist knows or suspects that there may be a model error of a particular amount in the travel time to some station, or that the model errors for a pair of stations may be related to each other, then the preceding theory becomes invalid. If the seismologist has no more information about the region than the arrival times from the one earthquake, then, in view of the robustness of the F statistic, an F based confidence region may be calculated.

WEIGHTED LEAST SQUARES

Homogeneity of the form of the distribution and hence homogeneity of variance of the observations has been mentioned as a desirable property to have. When there is a difference between the variances, then the simple least squares method should be replaced by a weighted least squares method, where the j^{th} observation is given a weight which is inversely proportional to the standard deviation of the j^{th} observation. If, as is usually the case, the ratio of variances of observations is not known exactly, some effort should be made to estimate this quantity. The F confidence regions will sustain mild heterogeneity of variance. Chapter V contains a description of the method used to weight P and S observations which had prior measures of quality assigned. After one has applied the weights, the resulting system of equations is treated as described in Chapter I. The use of correct weights in no way alters the properties of the best hypocentre solution (Zelen (1962)).

MORE THAN ONE EVENT

As soon as information from a second nearby earthquake is available, the situation alters. It is unlikely that the information used to locate the two events is independent — the presence of a single common station will suffice to make the location errors non-independent, since by *(2.7)* there will be a common mislocation component.

This means, for instance, that if by chance the first event lay outside its (100-α) percent confidence region, the probability that the second event also lay outside the corresponding (100-α) percent region is greater than α percent. This phenomenon, commonly referred to as "source biasing", is caused by a lack of independence of errors.

It is possible to construct a statistical model for the model travel time errors in the experiment of observing and locating M earthquakes, known to have occurred in some region V, recorded by various subsets of N stations in a region Ω, where the model travel time error at station j is s_{ij}.

We suppose that all points in V are equally likely to be earthquake sources or that the probability of an event occurring within a small volume $\delta\tau$ centred upon $\underline{x}$ is known for each $\underline{x} \in V$. Let the probability density function be denoted $f_0(\underline{x})$. For any

N vector $\underline{z}$ let $V_{\underline{z}}$ be defined to be the subset of V such that for each $\underline{x} \in V_{\underline{z}}$, the model errors associated with $\underline{x}$, $s_j(\underline{x})$, satisfy:

$$s_j(\underline{x}) \leq z_j \qquad j=1\ldots N$$

Define $F_{\underline{s}}(\underline{z})$ by:

$$F_{\underline{s}}(\underline{z}) = \int_{V_{\underline{z}}} f_0(\underline{x})d\tau \qquad \ldots(2.16)$$

or, in the case of a uniform seismicity:

$$F_{\underline{s}}(\underline{z}) = \left(\int_{V_{\underline{z}}} d\tau\right) / \left(\int_{V} d\tau\right) \qquad \ldots(2.17)$$

$F_{\underline{s}}(\underline{z})$ is a distribution function and we equate the probability that $s_j(\underline{x}) \leq z_j$ for all stations j with $F_{\underline{s}}(\underline{z})$. If this function were available to us we could compute the expected value of $\underline{s}$ and make our linear estimator unbiased. We could compound the distributions of $\underline{s}$ and the distribution of the other errors (which we might fairly approximate as normal) and compute, within a multiplicative factor, the variance-covariance matrix of the errors and transform the equations of condition so that standard linear statistical theory is again applicable (Zelen (1962)).

However, since a knowledge of $\underline{s}$ is one of the objects of our investigations without which the spaces $V_{\underline{z}}$ cannot be ascertained and $F_{\underline{s}}(\underline{z})$ remains unknown, this ideal situation cannot be realised.

At this stage one might adopt a Bayesian approach and from any rudimentary information about $\underline{s}$ construct an approximation to $F_{\underline{s}}$ and hence refine the knowledge of $\underline{s}$. One of the objects of this thesis is to examine methods of obtaining this preliminary knowledge.

THE STANDARD DEVIATION OF THE RESIDUALS

Before we finish with the single earthquake, let us discuss the root mean square residual:

$$\hat{\sigma} = \sqrt{\{(\sum_{j=1}^{n} {}_*r_j^{\,2})/(n - p)\}} \qquad \ldots(2.18)$$

which we have already referred to in passing when discussing confidence regions and which is popularly used as a measure of the quality of a particular location. The justification for this is that if the total error $\underline{e}$ is small, the location error will be correspondingly small (as shown by equation *(2.4)*) as will the residual vector, ${}_*\underline{r} = {}_*\underline{y}$. In the case where the hypocentre estimate is unbiased, $\hat{\sigma}$ is an unbiased estimate of the standard deviation of the total error by the Gauss-Markov Theorem (Zelen (1962)), a property which does not require the assumption of normal errors. Intuitively, we have projected the n vector $\underline{e}$ onto an n-p dimensional subspace in such a way that the directions of $\underline{e}$ and the direction of the projection are completely independent and we wish to compensate for the inevitable loss of magnitude which results from such a projection by dividing $\|{}_*r\|^2$ by the dimension of the subspace; that is, the number of degrees of freedom. However, when there is bias, we cannot say that the directions of $\underline{e}$ and the projection are independent and the result does not hold. We again have the situation that the more we know about the errors $\underline{e}$ the less reliable are results which are based on the assumption of knowing nothing about them.

Even in the ideal situation, however, $\hat{\sigma}$ has a failing when used as a rough guide to the quality of the solution. From the discussion of condition number in Chapter I, it is clear that when two earthquakes have about the same r.m.s. residual, the event with the smaller condition number has a smaller likelihood of being mislocated by errors of a given magnitude than the event with the larger condition number. The condition number is of course the condition number of the parameter variance matrix, being in fact the quotient of the variances of the principal and least components of this matrix. Thus the use of the r.m.s. residual by itself without reference to the variance matrix or condition number is not a satisfactory means of indicating quality.

CHAPTER III

THE THEORY OF MULTIPLE EARTHQUAKE ANALYSES

In this chapter we analyse the different ways of treating the data from a group of earthquakes. Commonly one wishes to obtain more information from the arrival times than just the hypocentre estimates. We might roughly classify such analyses into three general groups: attempts to produce or improve a world average travel time model, attempts to produce a regional travel time model, and other studies which might produce a sub-regional model. The work in Chapter VI falls into this last category.

Until the last decade, seismological efforts in this field tend to have been concentrated on world average models. The author does not propose to discuss fully the methods of Jeffreys and Bullen (Jeffreys (1939)) or Herrin *et al.* (1968), but sufficient of their methods will be presented to compare their work with the different methods of Lilwall and Douglas (1969).

In the past few years, more effort has been directed towards regional modelling or obtaining regional corrections to a world average model since it has been realised that considerable regional time discrepancies exist and that the hypocentre location errors that result are too large if these quantities are ignored. In many ways explosions have been a spur to this work since explosion data provide test cases for the different methods.

The different methods fall into two classes: those which Douglas and Lilwall (1972) refer to as Methods of Successive Approximations, where the information in the residuals from a set of single event locations is pooled in some way to produce travel time corrections, and Joint Methods of different types, first suggested by Douglas (1967).

Before we begin discussing methods in detail, we wish to make a further distinction between different methods which seems not to be commonly recognised. We will refer to a method as a "Relative Hypocentre Determination Method" if the results of the method are a set of locations whose relative errors are satisfactorily small and estimable in magnitude by some statistical means while the average error of the group remains unknown. We will refer to a method as an "Absolute Hypocentre Determination Method" if the results of the method are a set of locations whose absolute errors are statistically described.

Let us begin by establishing our data. Suppose N seismograph stations each record at least some of M earthquakes in a region V (which might be the whole of the seismically active part of the earth). Let n_i be the number which record the i^{th} event and, as in Chapter II, let P_i be the $N{\times}N$ diagonal matrix with the j^{th} diagonal entry 1 if the j^{th} station records the i^{th} event and zero otherwise. Suppose, for convenience, that at the starting point of our analysis each earthquake is assigned its least squares hypocentre as in Chapter I. The equations of condition for the i^{th} event are:

$$P_i A_i \delta \underline{x}_i = P_i(\underline{y}_i - \underline{s}_i + \underline{\varepsilon}_i + \underline{q}_i) \qquad \ldots(3.1)$$

The use of P_i is convenient to keep track of missing observations. The residuals from the i^{th} event are the non-zero entries in $P_i \underline{y}_i$.

VARIATION IN THE MODEL ERROR

We now seek to describe the variation of the model error s_{ij} for the i^{th} event and the j^{th} station. In principle, it is a function of the position of the earthquake and the position of the station and will be indeterminate through insufficient data unless assumptions are made about its behaviour. With certain reasonable assumptions we can calculate a bound for the difference between model errors for a pair of events X_1 and X_2 recorded at station S. Let T_1 and T_2 be the true travel times to S and without loss of generality, let $T_1 > T_2$. Let t_1 and t_2 be the model travel times for the phases to S. Then the difference between the model error is:

$$s_1 - s_2 = T_1 - T_2 - (t_1 - t_2) \qquad \ldots(3.2)$$

Let X_2' be the point on the ray from X_1 to S such that the time from X_2' to S is T_2. (By hypothesis, X_2' lies between X_1 and S.) Let X_2^* be point on the ray from X_1 to S, possibly extended, such that the model travel time from X_2^* to S is t_2. (See Figure 3.1)

Let $\underline{\Phi}$ be the path from X_1 to S. Let $\sigma(\underline{\Phi}, \underline{\Phi}')$ be the slowness function for the phase in question and let $\sigma_m(\underline{\Phi}, \underline{\Phi}')$ be the model slowness. The travel time from X_1 to X_2' is then:

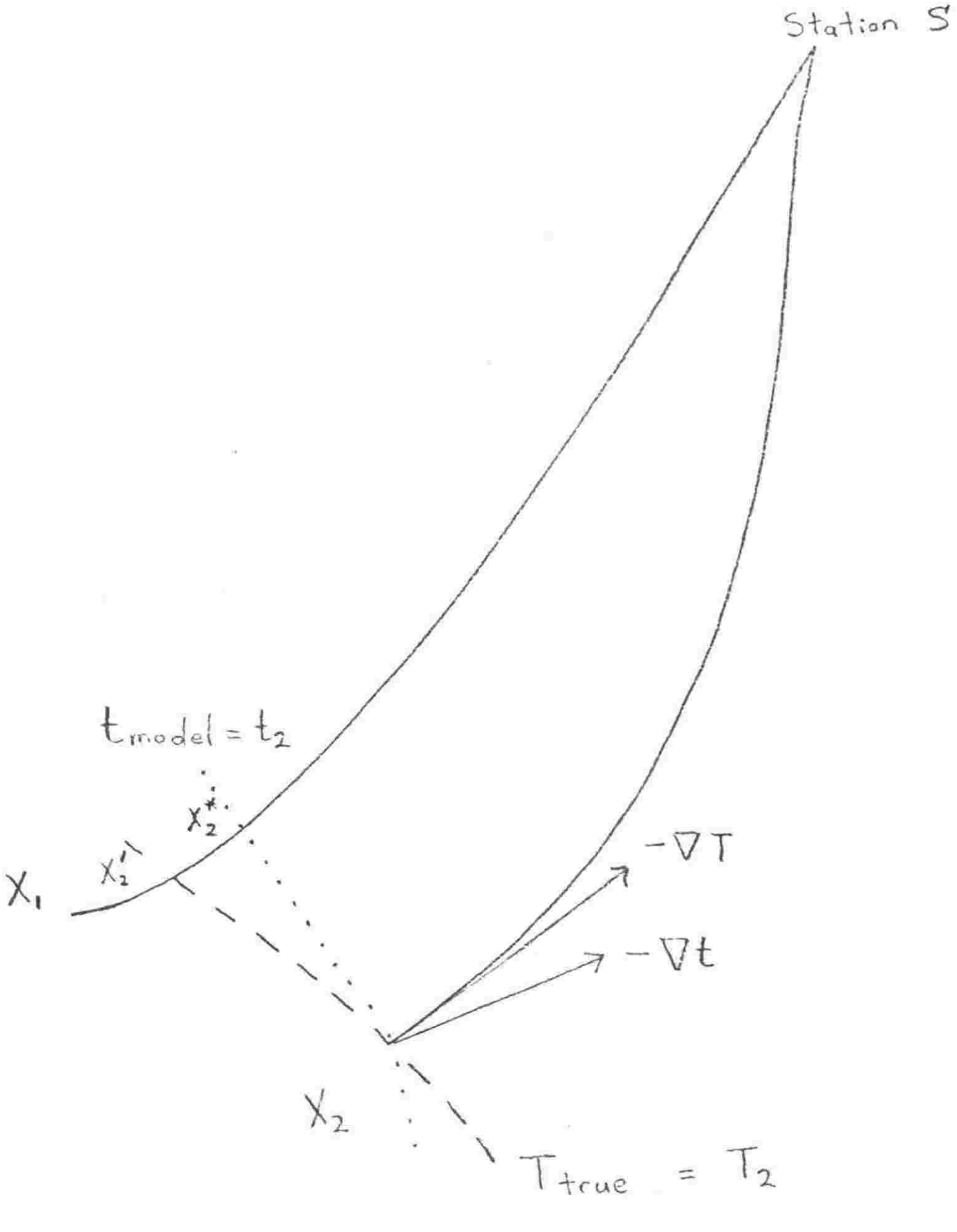


Figure 3.1 Ray paths from X_1 and X_2 to S

$$\int_{X_1}^{X_2'} \sigma(\underline{\Phi}, \underline{\Phi}')dl, \qquad \underline{\Phi}' = \frac{d}{dl}\Phi$$

Since, by Fermet's principle, the path Φ is an extremum time path and, in fact, a minimum for direct body waves (Bath (1969)), this quantity is not greater than:

$$\int_{X_1}^{X_2} \sigma(\underline{y}, \underline{y}')dl$$

where $\underline{y}$ is the straight line path joining X_1 to X_2.

Thus
$$\begin{aligned} T_1 - T_2 &= \int_{X_1}^{X_2'} \sigma(\underline{\Phi}, \underline{\Phi}')dl \\ &\leq \int_{X_1}^{X_2} \sigma(\underline{y}, \underline{y}')dl \\ &= \bar{\sigma}\|X_2 - X_1\| \end{aligned}$$

where $\bar{\sigma}$ is the average slowness along $\underline{y}$.

Similarly:

$$\begin{aligned} \int_{X_1}^{X_2'} \sigma(\underline{\Phi}, \underline{\Phi}')dl &= \bar{\sigma}' l(X_2', X_1) \\ &\geq \bar{\sigma}'\|X_2' - X_1\| \end{aligned}$$

where $\bar{\sigma}'$ is the average slowness along $\underline{\Phi}$ between X_1 and X_2' and l is the arc length of $\underline{\Phi}$ between X_1 and X_2'.

So:
$$\|X_2' - X_1\| \leq \frac{\bar{\sigma}}{\bar{\sigma}'}\|X_2 - X_1\| \qquad \ldots(3.3)$$

The model travel time difference $t_1 - t_2$ will be less than $\int_{X_1}^{X_2^*} \sigma_m(\underline{\Phi}, \underline{\Phi}')dl$ in magnitude by the minimum time path argument.

$$\text{Write:} \qquad t_2 - t_1 = \int_{X_2^*}^{X_1} \sigma_m(\underline{\Phi}, \underline{\Phi})dl + \delta$$

where $\delta < 0$ if $t_2 - t_1 > 0$ and $\delta > 0$ if $t_2 - t_1 < 0$. (δ is the error caused by integrating along the true ray from X_1 to X_2^* instead of the model ray.)

Then from *(3.2)*, writing $\sigma = \sigma_m + \delta\sigma$,

$$\begin{aligned} s_1 - s_2 &= \int_{X_1}^{X_2'} (\sigma_m + \delta\sigma)(\underline{\Phi}, \underline{\Phi}')dl + \int_{X_2^*}^{X_1} \sigma_m(\underline{\Phi}, \underline{\Phi}')dl + \delta \\ &= \int_{X_1}^{X_2'} \delta\sigma dl + \int_{X_2^*}^{X_2'} \sigma_m dl + \delta \end{aligned}$$

$$\text{Therefore:} \qquad |s_1 - s_2| \leq \overline{\delta\sigma}\, l(X_2', X_1) + \bar{\sigma}_m l(X_2', X_2^*) + |\delta|$$

$$\leq \overline{\delta\sigma} \frac{\bar{\sigma}}{\bar{\sigma}'} \|X_2 - X_1\| + \bar{\sigma}_m l(X_2', X_2^*) + |\delta| \qquad \ldots(3.4)$$

Assuming T and t are continuously differentiable where necessary, we have to the first order in $\|X_2 - X_2'\|$ and $\|X_2 - X_2^*\|$, if:

$$\nabla t = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)t$$

$$\text{then} \qquad \nabla T(X_2) \cdot (X_2' - X_2) = 0$$

$$\nabla t(X_2) \cdot (X_2^* - X_2) = 0$$

$$\rightarrow \qquad \nabla t(X_2) \cdot (X_2^* - X_2') = \{(\nabla t - \nabla T)(X_2)\} \cdot (X_2 - X_2')$$

So that:

$$|\nabla t(X_2)\cdot(X_2^*-X_2')| \quad \leq \quad \|(\nabla t-\nabla T)(X_2)\| \cdot \|X_2-X_2'\|$$

$$\leq \quad \frac{\bar{\sigma}}{\bar{\sigma}'}\|(\nabla t-\nabla T)(X_2)\| \cdot \|X_2-X_1\| \qquad \ldots(3.5)$$

We now make the following assumptions: Since the rays are parallel to $-\nabla T$, $X_2^*-X_2'$ will be parallel to $-\nabla T(X_2')$ in the limit as $\|X_2^* - X_2'\| \to 0$. If we assume that ∇T and ∇t are not too different we can take:

$$\frac{\bar{\sigma}}{\bar{\sigma}'}\|(\nabla t-\nabla T)(X_2)\| \cdot \|X_2-X_1\| \,/\, \|\nabla t(X_2)\|$$

as a crude bound for $\|X_2^* - X_2'\|$. If we further assume that $l(X_2^*-X_2')$ $\simeq \|X_2^*-X_2'\|$, that $|\delta|$ is negligible compared to $\overline{\delta\sigma}\|X_2-X_1\|$, and that $\bar{\sigma} \simeq \bar{\sigma}'$, we have that:

$$|s_1-s_2| \quad \lesssim \quad \{\overline{\delta\sigma} + \bar{\sigma}_m \; \|(\nabla t-\nabla T)(X_2)\| \,/\, \|\nabla t(X_2)\| \;\} \; \|X_2-X_1\|$$

$$\ldots(3.6)$$

We shall make use of this bound later in assessing the size of region for which the difference between model errors is of the same order as other errors contributing to hypocentre error. At this stage we will limit ourselves to a few remarks about *(3.6)*. First, the term $\overline{\delta\sigma}\|X_2-X_1\|$ is close to the true error difference when X_2 is on the ray through X_1. The other term contains the fractional error in the gradient which is approximately the angle between the two gradient vectors: this can be guessed very crudely at best.

One can construct situations such as in Figure 3.2 where there is an appreciable angle between true ray and the model ray, which equals the angle between the gradients. Consequently, the second term may sometimes be quite large.

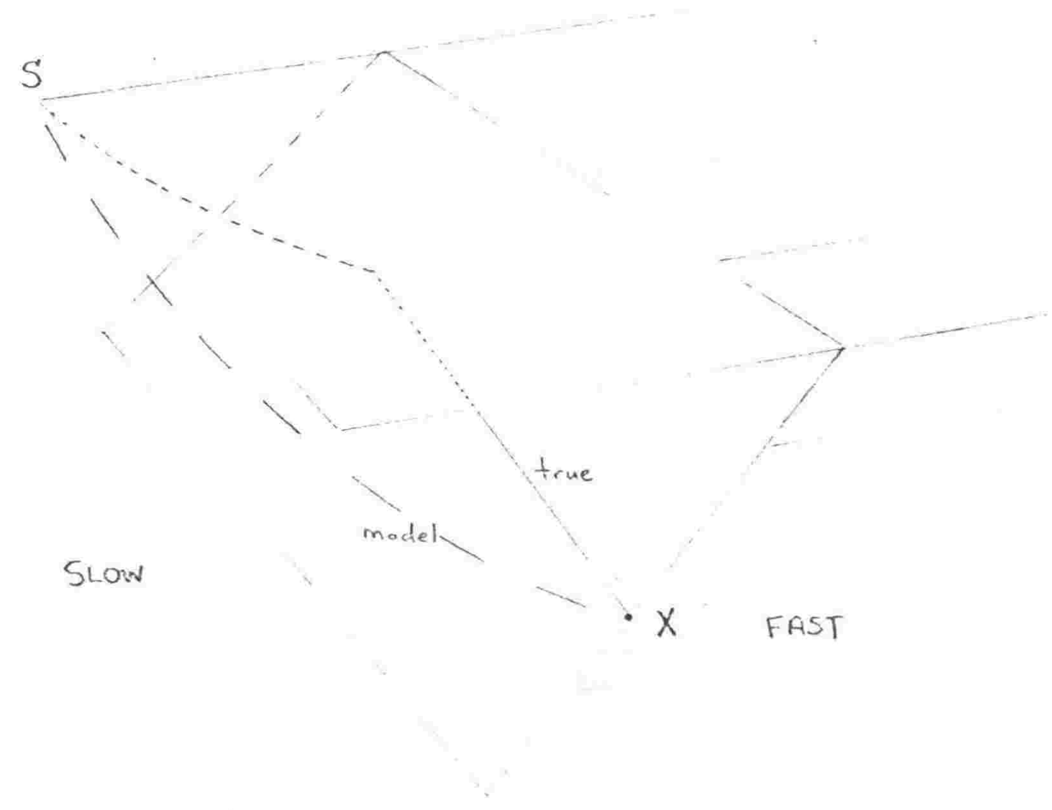


Figure 3.2 True and model rays with lateral inhomogeneity

A representative value for intermediate depth events 30 km apart, which is as we have seen about the maximum displacement for which the linear approximations are worthwhile, with a local average slowness of 0.1 sec/km, a model slowness error of ten percent, and five degree = 0.1 radian ray angle error, is:

$$|s_1 - s_2| \leq 0.6 \text{ sec}$$

It must be reiterated that the expression *(3.6)* applies to model errors at a particular station and does not give a bound for the variation in what has been termed the source error meaning model errors specific to the region containing the source.

We will utilise this expression to give us some idea of the dimensions of a seismic region in which relative location error is likely to be describable using standard statistical methods. We now set out the different methods for dealing with the information in *(3.1)*, recalling that we have, in principle, a distribution function $F_{\underline{s}}(\underline{z})$ describing the model errors $\underline{s}_i$ for our group of events.

THE METHODS OF SUCCESSIVE APPROXIMATIONS

With these methods, the first step is to pool the residuals $\underline{y}_i$ in some way to estimate the model errors. If the earthquakes are distributed over a region such that the differences between model errors for pairs of events is likely to be unsatisfactorily large, then this must be allowed for in some way, either by assuming a functional relationship between the model errors, as for example in the sinusoidal source terms of Herrin *et al.* (1968) or by dividing the regions into subregions and utilising the theory of the analysis of variance as in Bolt and Freedman (1968) or combine both approaches as in Veith (1975). The analysis of variance technique has the advantage of not forcing a possibly incorrect relationship on the errors.

Let us assume then that the model error is estimated by some linear function of the combined residuals, a process which might be no more complicated than taking the mean residual at each station. The model is then updated and the whole process of locating the events and reducing the residuals is repeated until it is considered that convergence has occurred.

We may be critical of such schemes on three counts: The process may not converge to the combined solution for hypocentres and model which gives an absolute minimum to the sum of squares combined residuals, the model update estimates are not unbiased and calculating confidence regions for the resultant solution is difficult. These points are discussed to a certain extent in Douglas and Lilwall (1972) who point out that in poorly conditioned two-step linear iterative procedures (as described above) convergence may be so slow that the occurrence of small increments would be a misleading test of convergence, and also that the method makes no allowance for the correlation between the errors in the model improvement estimates and the errors in the resulting hypocentre. If this correlation is not allowed for, one gets an incorrect confidence region for the hypocentres by using the standard linear statistical theory. We add two more comments. First, even if convergence occurs, it may not be to the absolute minimum. Take the simple example of M events recorded at each of N stations and sufficiently close together to make $\sup|s_j - s_i| \simeq \sigma_\varepsilon$ = reading error standard deviation.

Under this circumstance only the mean model error is worth estimating. If we set:

$$\hat{\underline{s}} = \frac{1}{M}\sum_{i=1}^{M} \underline{y}_i \qquad \ldots(3.7)$$

then compute:

$$\delta\hat{\underline{x}}_i = (A_i^T A_i)^{-1} A_i^T (\underline{y}_i - \hat{\underline{s}}) \qquad \ldots(3.8)$$

which is, by hypothesis, the same as:

$$\delta\hat{\underline{x}}_i = -(A_i^T A_i)^{-1} A_i^T \hat{\underline{s}} \qquad \ldots(3.9)$$

Let $\underline{y}_e \in \{\underline{y}_i : i=1,M\}$ satisfy:

$$\|\underline{y}_e - \hat{\underline{s}}\| \leq \|\underline{y}_i - \hat{\underline{s}}\| \qquad i=1,..M$$

With an increasing number of earthquakes, this $\underline{y}_e$ can be made as close to $\hat{\underline{s}}$ as we please. With the high degree of consistency which is possible among residuals (see Tables in Appendix III), not too many earthquakes might be required to find $\underline{y}_e$ such that:

$$\|\underline{y}_e - \hat{\underline{s}}\| \ll \|\hat{\underline{s}}\|$$

Then by *(3.8)*:

$$\delta\hat{\underline{x}}_e = (A_e^T A_e)^{-1} A_e^T (\underline{y}_e - \hat{\underline{s}}) \qquad \ldots(3.10)$$

so we would not expect a large correction $\delta\hat{\underline{x}}_e$. But the only difference between $\delta\hat{\underline{x}}_e$ and $\delta\hat{\underline{x}}_i$ is due to the change in the matrix A_i across the region. We have seen that over regions with dimensions of several tens of kilometres, the linear approximation to the model, which is equivalent to assuming A_i is constant, is a good one. Thus with the method of successive approximations under these circumstances we would expect only small changes in the hypocentre estimates with consequent small further changes in the model. "Convergence" would have thus occurred, but the evidence of Chapter V, where a similar scheme is compared directly with a joint method, shows that the absolute minimum has not been reached.

The situation is more complicated when it is not the case that every station records every earthquake (which is the common situation for any group of earthquakes one might wish to study) and here particularly we strike the second problem with this method: the model improvement estimates are biased.

Suppose, as before, that the data only warrants an estimate of the mean model error, which we will attempt again by taking the average residual at each station. Then:

$$\hat{\underline{s}} = \left(\sum_{j=1}^{M} P_j\right)^{-1} \sum_{j=1}^{M} P_j \underline{y}_j \qquad \ldots(3.11)$$

Taking expected values:

$$E(\hat{\underline{s}}) = \left(\sum_{j=1}^{M} P_j\right)^{-1} \sum_{j=1}^{M} \{P_j(A_j E(\delta \underline{x}_j) + E(\underline{s}_j))\}$$

which is not the expected value of $\underline{s}_j$ even if $E(\delta \underline{x}_j) = 0$, which is in fact not the case as we have seen.

Having now raised several objections to the successive approximations method, let us consider the family of alternatives which go under the heading of Joint Epicentre (Hypocentre) Determinations, first suggested by Douglas (1967) and since used with considerable success globally (by Lilwall and Douglas (1969)) and locally (for example, by Blaney and Gibbs (1968), Dewey (1971) and Billington and Isacks (1975).

These methods have certain practical difficulties as we shall see, but our theoretical objections to the successive approximation method are overcome to a greater or lesser extent by the different versions of this technique.

Let $\underline{s}_\mu$ be the expected value of the model error for the group of events. Write $\underline{s}_i - \underline{s}_\mu = \delta\underline{s}_i$, so that $E(\delta\underline{s}_i) = \underline{0}$. The equations of condition for the i^{th} event are:

$$P_i(A_i\delta\underline{x}_i + \underline{s}_\mu) \quad = \quad P_i(\underline{y}_i - \delta\underline{s}_i + \underline{\varepsilon}_i - \underline{q}_i) \qquad i=1,..M \qquad ...(3.12)$$

and we solve the overdetermined system:

$$\begin{pmatrix} P_1A_1 & & & P_1 \\ & P_2A_2 & & P_2 \\ & & \ddots & \\ & & P_MA_M & P_M \end{pmatrix} \begin{pmatrix} \delta\underline{x}_1 \\ \vdots \\ \delta\underline{x}_M \\ \underline{s}_\mu \end{pmatrix} = \begin{pmatrix} P_1\underline{y}_1 \\ \vdots \\ \vdots \\ P_M\underline{y}_M \end{pmatrix} \qquad ...(3.13)$$

by a suitable method to obtain least squares estimates for $\delta\hat{\underline{x}}_i$, $\hat{\underline{s}}_\mu$. Since $E(\delta\underline{s}_i) = 0$ these estimates are now unbiased, ignoring the small biasing effect of $E(\underline{\varepsilon}_i)$ and $\underline{q}_i$. To qualify as minimum variance unbiased estimates, however, it is necessary that the variance of the error of each equation in *(3.12)* be the same. The assumption of equal station variances is known to be suspect (Freedman (1968)). Assessing the ratio of the variance of the model errors at one station to variance at another is difficult without some knowledge of the real physical situation.

The problem is eased in that the variance of any error is the sum of the two variances of the model errors and reading errors. If $\sigma^2_{s_1}$ and $\sigma^2_{s_2}$ are the model variances at stations 1 and 2 and the observations at each have a common reading error variance σ^2_{ε}, then the observations should be weighted in the ratio:

$$1 : \sqrt{\frac{\sigma^2_{s_1} + \sigma^2_{\varepsilon}}{\sigma^2_{s_2} + \sigma^2_{\varepsilon}}} = 1 : \sqrt{\frac{\sigma^2_{s_1}/\sigma^2_{s_2} + \sigma^2_{\varepsilon}/\sigma^2_{s_2}}{1 + \sigma^2_{\varepsilon}/\sigma^2_{s_2}}}$$

If $\sigma^2_{\varepsilon}/\sigma^2_{s_2} >> \sigma^2_{s_1}/\sigma^2_{s_2}$, the difference between model error variances does not matter. If, as ideally would be the case, $\sigma^2_{\varepsilon_2}/\sigma^2_{s_2} \simeq 1$, then the ratio becomes:

$$\sqrt{\frac{1 + \sigma^2_{s_1}/\sigma^2_{s_2}}{2}}$$

Unless the variances differ by an order of magnitude no great error is likely to be introduced by assuming equal model error variances. In Chapter V some of the practical considerations of calculating appropriate weights are discussed.

An immediate theoretical difficulty in obtaining a solution to *(3.13)* is that the matrix on the left-hand-side is not of full rank. The sum of the first column of each sub-matrix $P_i A_i$ is the same as the sum of the columns of the part containing the P_i matrices. In theory the system has a rank deficiency of one and needs at least one constraint to ensure the least squares solution to *(3.13)* is unique. It is the choice of constraint that determines the particular Joint Method. Broadly, we might either constrain the model corrections or a subset of the hypocentre improvements.

Appendix I contains the algorithm for reducing *(3.13)* with different constraints and obtaining the least squares solution. For the moment we will confine ourselves to classifying the different constraints.

First, let $s_0 = (\sum_{j=1}^{N} s_{\mu_j})/N$ be the average model correction. Let $\underline{s}'_\mu = \underline{s}_\mu - s_0\underline{1}$. Then equation *(3.12)* can be rewritten:

$$P_i(A_i\delta\underline{x}_i + \underline{s}'_\mu) = P_i(\underline{y}_i - s_0\underline{1} - \delta\underline{s}_i + \underline{\varepsilon}_i - \underline{q}_i) \qquad i=1,..M \qquad ...(3.14)$$

and in equation *(3.13)*, $\underline{s}_\mu$ is replaced by s'_μ which satisfies:

$$\sum_{j=1}^{N} s'_{\mu_j} = 0 \qquad ...(3.15)$$

The augmented system *(3.13)* plus the constraint *(3.15)* now has full rank. The model correction estimates are now the estimates of the deviations about the average model error. The presence of the term $\underline{s_0 1}$ in *(3.14)* means that the new system does not give rise to unbiased parameter estimates. However, we have seen in Chapter II that the presence of an error $k\underline{1}$ only produces a bias of $-k$ in the origin time estimates. Thus, apart from a bias of $-s_0$ in all the origin time estimates, the system gives unbiased parameter estimates.

Appendix II contains a method for obtaining the variance-covariance matrix for an arbitrary subset of the parameters. To construct (joint) confidence regions for a subset of parameters using standard theory requires, as we have seen, the distribution of errors in *(3.12)* to be sufficiently close to normal to validate the use of an F statistic. The assumption that the distribution of the $\delta \underline{s}_i$ is close to normal is much more reasonable than the assumption that the model errors themselves are normal with zero mean. The use of standard linear confidence regions for subsets of the parameters which exclude the origin times is thus probably quite fair. Apart from the unknown bias of the origin time, this method is an absolute location method according to our definition.

There is a practical problem associated with this method reported by Lilwall and Douglas (1972), Dewey (1971) and others, that the system *(3.13)* plus constraint may be severely ill-conditioned; in particular, it may be possible to get almost the same residual sum of squares from widely differing combinations of parameter estimates. In one sense there is nothing one can do about this situation if one insists upon using this method other than using the best numerical techniques available for minimising the effect of roundoff error in the processing of the system. The uncertainties of the parameter estimates should be fairly reflected in the variances of these quantities and parameters which are interacting will tend to have large standard deviations with high correlations between them. It should be noted, however, that the conditioning of the system will improve as the geographical spread of the group of earthquakes increases, and although with the larger region the estimates of the mean model error for the region will have less value, the method still has the important property of producing unbiased parameter estimates.

A simple extension of this method allows the joint procedure to be iterative. Indeed, we seek the combined set of hypocentre estimates plus mean model corrections which yields an absolute minimum sum of squares. There is no reason to suppose that this minimum will be reached in one iteration since the system is still effectively a linearised non-linear system. Douglas (1967) notes that four or five iterations suffice to produce increments δs_i not greater than 0.01 sec. in the case of teleseismic location of Aleutian earthquakes.

However, the worse the conditioning the slower the convergence. My experience (Chapter V) suggests that damping the increments is almost always necessary and that after the first iteration, improvements to the residual sum of squares tend to be insignificantly small while the increments to the solution need not be small.

One method available for improving the conditioning of a given system is to impose further constraints upon the solutions. In particular, if there exists a near linear dependence between the model correction terms and the hypocentre improvements, one can reduce this by fixing one of the hypocentres. This fixed event becomes the Master Event of Douglas (1967), Dewey (1971) and others. This is a singularly appropriate name since the final estimates of the other hypocentres are slaves to the solution which is chosen for the Master. In Chapter V it is demonstrated that the fixing of a Master Event in different positions resulted in the rest of the solutions moving in unison with the Master while retaining their relative spatial distribution for a wide range of positions of the Master.

Let event X_0 be the Master Event. The equations of condition for this event are:

$$P_0\underline{s}_\mu \;=\; P_0(\underline{y}_0 - A_0\delta\underline{x}_0 - \delta\underline{s}_0 + \underline{\varepsilon}_0 - \underline{q}_0)$$

where the hypocentre error term $A_0\delta\underline{x}_0$ now appears among the errors. The system to be solved consists of *(3.13)* with P_0 appended to the left-hand side matrix and $\underline{y}_0$ appended to the right-hand side vector. The system obtained yields quite a different solution from the method where the station terms are taken to be the residuals from the Master Event, even though these two methods tend to produce similar sets of locations, for reasons which we shall examine. An important advantage of the Master Event method described above is that the Master Event need not be recorded at every seismograph station in the network, whereas this is obviously mandatory in the "Master Residuals" method.

The increased stability must be paid for elsewhere, and unless there is additional information about the location error of the Master Event, such as would be the case if this event were an explosion with known hypocentre, the presence of the $A_0\delta\underline{x}_0$ term means that the estimates in this case will be biased. Deterministically, the error in $\hat{\underline{s}}_\mu$ will depend linearly on $\delta\underline{x}_0$, and hence the errors in the $\delta\hat{\underline{x}}_i$, which are coupled with the errors in $\hat{\underline{s}}_\mu$, will depend (linearly) on $\delta\underline{x}_0$. Thus the Master Event Method is a relative location method. Confidence regions constructed for the parameters in this method describe the effect of errors after the bias $A_0\delta\underline{x}_0$ has been removed, or in other words, the effect of errors other than $A_0\delta\underline{x}_0$.

The Master Residuals method is also a relative location method which in an ideal situation produces results similar to the Master Event method. In Chapter II we saw that the residual vector was orthogonal to the part of the total error which contributed to the mislocation. We obtained the equation $_{*}\underline{y} = \{I - A(A^T A)^{-1} A^T\}\underline{e}$. In the situation where all errors are negligible compared to the model errors, the residual vector is a reasonable estimate of that part of the model error which does not contribute to the mislocation. If over some region the model errors change slowly, the change in $\{I - A(A^T A)^{-1} A^T\}\underline{s}$ will depend on how quickly the design matrix A alters across the region. In any case, the Master Residual vector can be interpreted as an estimate of the average value of $\{I - A_i(A_i^T A_i)^{-1} A_i^T\}\underline{s}_i$ for the group of events. The equations of condition for the i^{th} event after correcting the residuals will be:

$$A_i \delta\underline{x}_i = (\underline{y}_i - \underline{y}_0) - (\underline{s}_i - \underline{s}_0) + A_0 \delta\underline{x}_0 + (\underline{\varepsilon}_i - \underline{\varepsilon}_0) - (\underline{q}_i - \underline{q}_0)$$

If we have reached the solution for which $\delta\hat{\underline{x}}_i = 0$, then:

$$\begin{aligned} \delta\underline{x}_i &= (A_i^T A_i)^{-1} A_i^T \{(\underline{s}_0 - \underline{s}_i) + A_0 \delta\underline{x}_0 + \text{other errors}\} \\ &\approx (A_i^T A_i)^{-1} A_i^T (\underline{s}_0 - \underline{s}_i) + \delta\underline{x}_0 \end{aligned}$$

The problems with this method are that one must be certain of the quality of the observations of a particular event - contamination of $\underline{y}_0$ with large reading errors will invalidate the method - and the model correction estimates one obtains are nearly useless. As demonstrated in Chapter II, the model errors are certain to be larger than the residuals in magnitude, and the component that is estimated is the least interesting part of the model error.

It is important to grasp the distinction between absolute and relative location methods. Very often one's purpose is achieved by a relative location method, and then the question is which method produces a set of solutions with the smallest relative errors, the average error of the group being unimportant. A simple technique which the author has found to produce very small relative errors under certain favourable circumstances is the Homogeneous Station Method. Chapter IV comprises the theory of this method and an example of its use.

CHAPTER IV

THE HOMOGENEOUS STATION METHOD

RELATIVE LOCATION ERRORS

In Chapter III we saw that there were disadvantages in the situation which is normal in analysing arrival time data from a group of earthquakes of not having an observation from each earthquake at every station. In particular, the relative location errors of a pair of events will depend on the sets of stations observing each one. Modifying *(2.6)* after *(3.1)* gives the absolute errors of location for events i and j as:

$$\delta \underline{x}_i = (A_i^T P_i A_i)^{-1} A_i^T P_i (-\underline{s}_i + \underline{\varepsilon}_i - \underline{q}_i) \quad \ldots (4.1)$$

$$\delta \underline{x}_j = (A_j^T P_j A_j)^{-1} A_j^T P_j (-\underline{s}_j + \underline{\varepsilon}_j - \underline{q}_j) \quad \ldots (4.2)$$

so that if the events have identical hypocentres, giving $A_i = A_j$, $\underline{s}_i = \underline{s}_j$, $\underline{q}_i = \underline{q}_j$, the difference $\delta \underline{x}_i - \delta \underline{x}_j$ will not be simply a linear function of $\underline{\varepsilon}_i - \underline{\varepsilon}_j$ unless $P_i = P_j$; i.e., if the same stations record both events.

If we insist that this condition apply to a group of earthquakes under study, so that we deal only with those which have an observation of arrival time from each of a given set of stations, we make the arrival time data homogeneous. The P_i, P_j may now be dropped from *(4.1)* and *(4.2)* and the relative location error is:

$$\delta\underline{x}_i - \delta\underline{x}_j = (A_i^T A_i)^{-1} A_i^T(-\underline{s}_i + \underline{\varepsilon}_i - \underline{q}_i) - (A_j^T A_j)^{-1} A_j^T(-\underline{s}_j + \underline{\varepsilon}_j - \underline{q}_j) \quad \ldots(4.3)$$

Take j as a reference event and write:

$$L \equiv (A^T A)^{-1} A^T \quad \ldots(4.4)$$

L is a function of the spatial position of the earthquake and the travel time model which varies according to the departure of the travel time model from linearity: a linear travel time model implies L constant for all the events. If we write:

$$L_i = L_j + \nabla L\cdot(\underline{\delta x}_i - \underline{\delta x}_j) + \text{higher order terms}$$

$$\underline{s}_i = \underline{s}_j + \underline{\delta s}_{ij}$$

then, neglecting terms of order, $\| \delta\underline{x}_i - \delta\underline{x}_j \|^2$ (which include $\underline{q}_i$, $\underline{q}_j$):

$$\delta\underline{x}_i - \delta\underline{x}_j = -L_j \delta\underline{s}_{ij} - \nabla L\cdot(\delta\underline{x}_i - \delta\underline{x}_j)\delta\underline{s}_{ij} - L_j(\underline{\varepsilon}_j - \underline{\varepsilon}_i) + \nabla L\cdot(\delta\underline{x}_i - \delta\underline{x}_j)(\underline{\varepsilon}_i) \quad \ldots(4.5)$$

For a given travel time model we can find a sufficiently small geographical region such that the terms on the RHS of *(4.5)* involving ∇L are negligible compared to the other two terms. The dimensions of such a region will depend on the magnitude of ∇L, that is, on the magnitude of the departure from linearity of the travel time model. Within such a region then:

$$\delta\underline{x}_i - \delta\underline{x}_j \cong -L_j \delta\underline{s}_{ij} + L_j(\underline{\varepsilon}_i - \underline{\varepsilon}_j) \quad \ldots(4.6)$$

If we are assuming that the reading errors are normally distributed with zero mean and variance σ_ε^2, then the difference between two independent errors will have a normal, mean zero, variance $2\sigma_\varepsilon^2$ distribution. This source of error defines a lower limit to the relative location error of two events since this error is present even when the true locations of the events coincide. If we are prepared to treat $\delta \underline{s}_{ij}$ as a random variable and after the fashion of Chapter II define a distribution for $\delta \underline{s}$ and consequently a variance $\sigma_{\delta s}^2$, then if distribution of the total error $(-\delta \underline{s}_{ij} + \underline{\varepsilon}_i - \underline{\varepsilon}_j)$ is such that normal theory applies, we can construct relative location confidence ellipsoids based, as in Chapter II, on:

$$\mathrm{var}\,(\delta \underline{x}_i - \delta \underline{x}_j) = (2\sigma_\varepsilon^2 + \sigma_{\delta s}^2) L_j L_j^T \qquad \ldots(4.7)$$

$$\equiv (2\sigma_\varepsilon^2 + \sigma_{\delta s}^2)(A_j^T A_j)^{-1} \qquad \ldots(4.7a)$$

so that:

$$(\delta \underline{x}_i - \delta \underline{x}_j)^T \frac{1}{(2\sigma_\varepsilon^2 + \sigma_{\delta s}^2)} A_j^T A_j (\delta \underline{x}_i - \delta \underline{x}_j)$$

is approximately a χ^2 random variable with four degrees of freedom. Analogous expressions exist for subsets of the parameters of $(\delta \underline{x}_i - \delta \underline{x}_j)$. If we can estimate $(2\sigma_\varepsilon^2 + \sigma_{\delta s}^2)$ (from the residuals) we can form F confidence regions as in the single earthquake case which now represents the confidence of relative location of an event compared to a reference event.

The most important deviation from the assumption is the possible lack of independence of the components of $(-\delta \underline{s}_{ij} + \underline{\varepsilon}_i - \underline{\varepsilon}_j)$, caused by the likely correlation between entries in $\delta \underline{s}_{ij}$ for stations with similar ray paths to the source region. The effect of this can be minimised by choosing the region small enough so that:

$$\sigma^2_{\delta s} \lesssim 2\sigma^2_{\varepsilon} \qquad \ldots(4.8)$$

If we calculate the correlation between the k^{th} and the l^{th} components of the total error for events i and j, then:

$$\rho_{kl} = \frac{\text{cov}\,\{\delta s_k + \varepsilon_{ik} - \varepsilon_{jk} - \delta s_l + \varepsilon_{il} - \varepsilon_{jl}\}}{\sigma^2_{\delta s} + 2\sigma^2_{\varepsilon}}$$

$$= \frac{E(\delta s_k \delta s_l)}{\sigma^2_{\delta s} + 2\sigma^2_{\varepsilon}}$$

so that:

$$\rho_{kl} = \frac{\rho_{\delta s_k \delta s_q}}{1+\frac{2\sigma^2_{\varepsilon}}{\sigma^2_{\delta s^2}}}$$

Hence if the variation of $\underline{\delta s}$ across the region is such that *(4.8)* is satisfied, ρ_{kl} should be sufficiently small to satisfy the assumptions except when $\rho_{\delta s_k \delta s_l}$ is large, implying high correlations between the model error differences for the k^{th} and l^{th} stations. One can certainly imagine situations where correlations close to 1 would exist, but in the earth there are likely to be present heterogeneities at the source and at the stations such that $\rho_{\delta s_k \delta s_l} << 1$ except when the stations are very close together.

We will establish, shortly, a test that may be applied to determine whether a region is sufficiently small. In Chapter III we obtained the bound of 0.6 sec. for $|s_{ik} - s_{jk}|$ when events i and j are 30 km apart, implying a typical dimension for the region of 60 km. If the differences were distributed uniformly in the range -0.6 to 0.6, then $\sigma_{\delta s} \simeq 0.3$ sec.

Alternatively, we could assume a roughly normal distribution of differences and take 0.6 sec. as being two standard deviations (exceeded by only five percent of the differences). Again, we would obtain:

$$\sigma_{\delta s} \simeq 0.3 \text{ sec.}$$

In the case of arrival times read with a precision of 0.1 sec., σ_ε is in the range 0.1 to 0.2 sec. and so $\sqrt{2}\,\sigma_\varepsilon$ will not be greater than $\sigma_{\delta s}$. This suggests that the regions in which the analysis is valid may be a little smaller than 60 km in diameter.

ESTIMATING THE TOTAL ERROR VARIANCE FROM THE RESIDUALS

We now turn to the residuals which will provide an estimate of $(\sigma^2_{\delta s} + 2\sigma^2_\varepsilon)$ and hence a check that $\sigma^2_{\delta s}$ is not too large compared with $2\sigma^2_\varepsilon$. Following the notation of Chapter III, let $\underline{y}_i$ be the residuals from the i^{th} event for the usual least squares solution. By *(2.9)*:

$$\underline{y}_j = \{A_jL_j - I\}\{-\underline{s}_j + \underline{\varepsilon}_j - \underline{q}_j\} \qquad \ldots(4.10)$$

with the same assumptions as before:

$$\underline{y}_i - \underline{y}_j \simeq \{A_jL_j - I\}\{-\delta\underline{s}_{ij} + \underline{\varepsilon}_i - \underline{\varepsilon}_j\} \qquad \ldots(4.11)$$

Rather than calculating the deviations of the residuals about the residuals of an arbitrary event, we replace $\underline{y}_j$ by $\bar{\underline{y}} = \frac{1}{M}\sum_{i=1}^{M} \underline{y}_i$ and $\delta\underline{s}_{ij}$ by $\delta\underline{s}_i = \underline{s}_i - \frac{1}{M}\sum_{i=1}^{M} \underline{s}_i$ whence:

$$\underline{y}_i - \bar{\underline{y}} \simeq \{A_i L_i - I\}\{-\delta\underline{s}_i + \underline{\varepsilon}_i - \varepsilon_j\} \qquad \ldots(4.12)$$

Thus:

$$(\underline{y}_i-\bar{\underline{y}})^T(\underline{y}_i-\bar{y}) = (-\delta\underline{s}_i+\varepsilon_i-\varepsilon_j)^T\{L_i^T A_i^T-I\} \times \{A_i L_i-I\}(-\delta\underline{s}_i +\underline{\varepsilon}_i-\underline{\varepsilon}_j)$$

$$= (-\delta\underline{s}_i+\underline{\varepsilon}_i-\underline{\varepsilon}_j)^T\{I-A_i L_i\}(-\delta\underline{s}_i+\underline{\varepsilon}_i-\underline{\varepsilon}_j) \qquad \ldots(4.13)$$

The matrix in the quadratic form on the right-hand side of *(4.13)* is idempotent and positive definite and so, by Cochran's Theorem (Scheffé, 1959), if the total error $(-\delta\underline{s}_i + \underline{\varepsilon}_i + \underline{\varepsilon}_j)$ is normally distributed with zero mean and variance σ^2, the right hand side divided by σ^2 is a χ^2 random variable with ρ degrees of freedom, ρ being the rank of the matrix $\{I - A_i L_i\}$ which must be N-4 from elementary considerations. Thus:

$$E\left(\frac{(\underline{y}_i - \bar{\underline{y}})^T(\underline{y}_i - \bar{y})}{\sigma^2}\right) = N - 4$$

or:

$$E\left(\frac{(\underline{y}_i - \bar{\underline{y}})^T(\underline{y}_i - \bar{\underline{y}})}{N - 4}\right) = \sigma^2$$

Thus the mean square deviation of the residuals from the i^{th} event is an unbiased estimate of σ^2. If we pool the mean square residuals we get:

$$\hat{\sigma}^2 = \sum_{i=1}^{M} \frac{(\underline{y}_i - \bar{\underline{y}})^T(\underline{y}_i - \bar{\underline{y}})}{M(N-4)} \qquad \ldots(4.14)$$

Since we have used the residuals to obtain $\bar{\underline{y}}$, this expression ought to be replaced by:

$$\hat{\sigma}^2 = \sum_{i=1}^{M} \frac{(\underline{y}_i - \bar{\underline{y}})^T(\underline{y}_i - \bar{\underline{y}})}{M(N-4)-(N-1)} \qquad \ldots(4.15)$$

with the denominator being reduced by the number of independent parameters in $\bar{\underline{y}}$, namely N-1, since the average residual will satisfy:

$$\sum_{j=1}^{N} \bar{y}_j = 0 \qquad \ldots(4.16)$$

With moderate numbers of stations and earthquakes, the number of degrees of freedom is such that this alteration will produce no significant change in $\hat{\sigma}$.

With our estimate $\hat{\sigma}$ of the standard deviation of the total error we can construct linear relative location confidence regions and also assess the relative contributions to $\hat{\sigma}^2$ by $2\sigma_\varepsilon^2$ and $\sigma_{\delta s}^2$. If we equate $2\sigma_\varepsilon^2 + \sigma_{\delta s}^2$ with $\hat{\sigma}^2$ then with an *à priori* estimate of σ_ε we can estimate $\sigma_{\delta s}$. If $\sigma_{\delta s}$ is too large according to the rule we earlier formulated, then a smaller region should be chosen over which to pool the residuals to give the estimate of $\hat{\sigma}^2$.

AN APPLICATION - THE NORTH ISLAND, NEW ZEALAND, MANTLE SEISMIC ZONE

A summary of the results of this study have been published in Ansell and Smith (1975).

The purpose of the work was to examine the geometry of the North Island Mantle Seismic Zone. It was hoped that the Homogeneous Station Method would yield sufficiently good relative locations that the fine structure of the zone would be revealed, and the study was successful to the extent that it was determined that the earthquakes studied, with assigned depth greater than 120 km, could have originated in a zone only nine kilometres thick.

Briefly, the zone strikes north-east, dipping under the central North Island with depths assigned by Hamilton and Gale (1968) of down to 400 km at the north-western edge of the zone.

Prior to this study, Hamilton and Gale (1969) described the zone having located earthquakes using standard least squares which produced an estimate of 20 to 40 km for the thickness of the zone in different sections. They also attributed appreciable curvature to the zone in the plane perpendicular to the strike.

DATA SELECTION

For this study, we selected all the earthquakes located by the Seismological Observatory, Wellington in the period November 1965 - December 1972 (excepting 1970, for which year readings were unavailable) which had a non-emergent *P* arrival read to 0.1 sec. at each of the North Island stations WEL, MNG, TNZ, CNZ, KRP, GNZ, ECZ and were assigned sub-crustal depths. There were only 73 such earthquakes. The problem in deciding on selection criteria was to have a reasonable number of data per event but also a reasonable total of earthquakes. To ensure reasonable

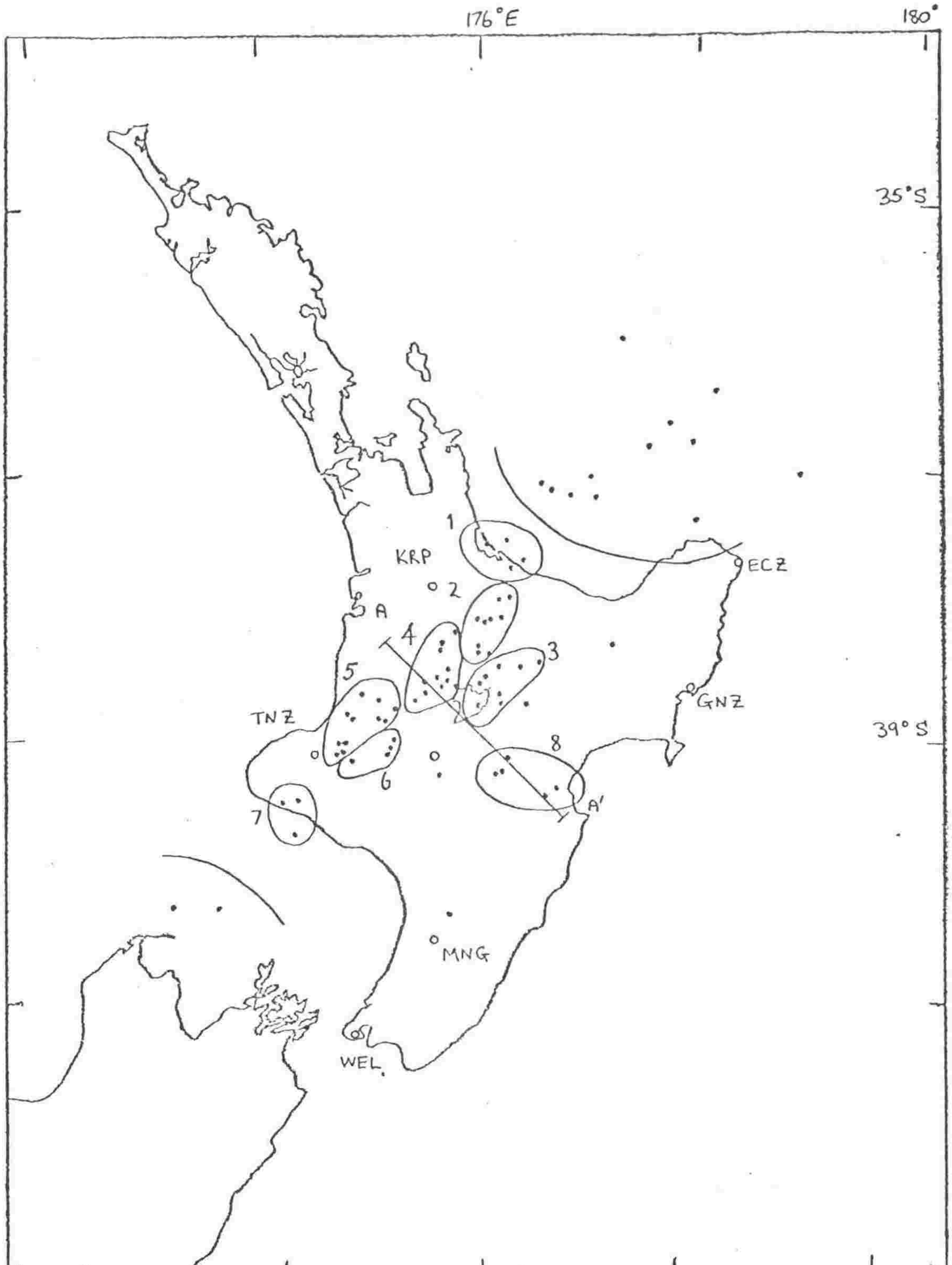


Figure 4.1 Grouped and excluded events

homogeneity of variance, phases designated *eP* by the seismologist who made the readings were disregarded. This caused an appreciable reduction in the number of satisfactory earthquakes but it was hoped that 73 high-quality earthquakes would be a sufficient number. Of these 73, 26 had a single reading given only to the nearest second, the remainder being read to 0.1 sec. The readings were not weighted in any way. At the stage of this study we felt we knew too little about the errors to apply weights to the observations, since the weights should be inversely proportional to the standard deviation of the total error and not just the reading error.

After relocation by the Homogeneous Station Method, 15 earthquakes were excluded from all further analysis because of their position. It was felt that these excluded events lay too far outside the station net and were too sparse to yield reliable relative hypocentre determinations. Those excluded generally had condition numbers $> 10^3$ (see Chapter I, Figure I), while none of the rest had condition numbers as large as this.

The remaining 58 earthquakes were divided by inspection into the eight groups shown in Figure 4.1 for the purpose of analysing residuals. The borders of groups were drawn to encompass hypocentres with similar patterns of residuals and to keep the hypocentres within a volume of about 40 km radius. The four isolated earthquakes in Figure 4.1 could not be assigned to groups, which left 54 earthquakes for the residual analysis.

ANALYSIS OF THE RESIDUALS

The residuals pooled by groups are given in Appendix III. By equation *(4.15)*, with N = 7, M = 54:

$$\hat{\sigma}^2 = (\sum_{i=1}^{8} s_i^2)/(54(7-4)-6\times 8) \qquad \ldots(4.8)$$

where the number of degrees of freedom is reduced by 8×6 = 48 because eight group means are calculated. Thus the total error estimate is obtained:

$$\hat{\sigma} = 0.39 \text{ sec. (114 d.f.)} \qquad \ldots(4.19)$$

If we estimate a standard deviation of 0.2 sec. for the reading error, which is probably a minimum in the view of the Seismological Observatory staff who made the readings, this gives, by *(4.17)*:

$$\hat{\sigma}^2_{\delta s} = 0.39^2 - 2\times 0.2^2$$

$$= 0.27 \text{ sec.} \qquad \ldots(4.20)$$

This agrees well with the theoretical value obtained earlier for $\sigma_{\delta s}$ for a region in which the assumptions hold and in which relative location errors can be described by linear theory.

The residuals from the earthquakes in the eight groups, the means for each group and the standard error

$$\sqrt{(\sum_{i=1}^{7} r_i^2)/3}$$

for each earthquake are listed in Appendix III. It is apparent that the sum of squares of residuals is very significantly reduced by the extraction of the group means. However, as pointed out earlier, no particular physical significance can be attached to these means. In a sense they are estimates of the travel time model errors, but as was shown in Chapter III the true model errors must exceed (in a vector magnitude sense) these group mean values. For this reason, the group means were discarded and not used as station corrections.

DISTRIBUTION OF THE HYPOCENTRES

Since many subcrustal earthquake zones exhibit a planar or nearly planar character, particularly on the scale of the structure being studied here, the first step in describing the structure was to fit a plane by the orthogonal least squares method described in Appendix IV.

It rapidly became apparent that the earthquakes located shallower than about 120 km, mostly in group 8 (and one deeper earthquake mentioned below), did not fit a plane model fitted to the rest and were omitted from this part of the analysis. The isolated earthquake located at 38.3°S, 177.2°E 160 km deep (see Figure 4.1) was the only earthquake of this depth within the net not assigned to a group because of its remoteness from neighbours. Other estimates of depth (which conceivably is grossly in error compared to the other events) are 189 km by the Observatory, 114 km by USCGS, 100 km by ISC.

A plane was then fitted to 49 earthquakes hypocentres. The standard deviation of the displacements of the hypocentres perpendicular to this plane was 5.6 km. By way of immediate comparison, a plane fitted to the Observatory's original locations for the same events produced a standard deviation of perpendicular distance of 10.0 km. This alone suggests a marked improvement in the quality of the relative locations of the homogeneous method.

At this point it should be remarked that there was no great systematic difference between the Observatory's locations, based on many more observations which invariably include some *S* readings, and ours. Generally speaking, there were differences in latitude of 10 km, which increased rapidly to 30 km or more for the earthquakes outside the net to the North. There was a very much smaller difference in longitude, and our depths were often 10 to 20 km deeper than the Observatory's, but not invariably so. The evidence of the best fitting plane clearly indicates that the homogeneous method produces better relative locations than the inhomogeneous standard method. It is not argued that the absolute locations are any better, or even necessarily as good.

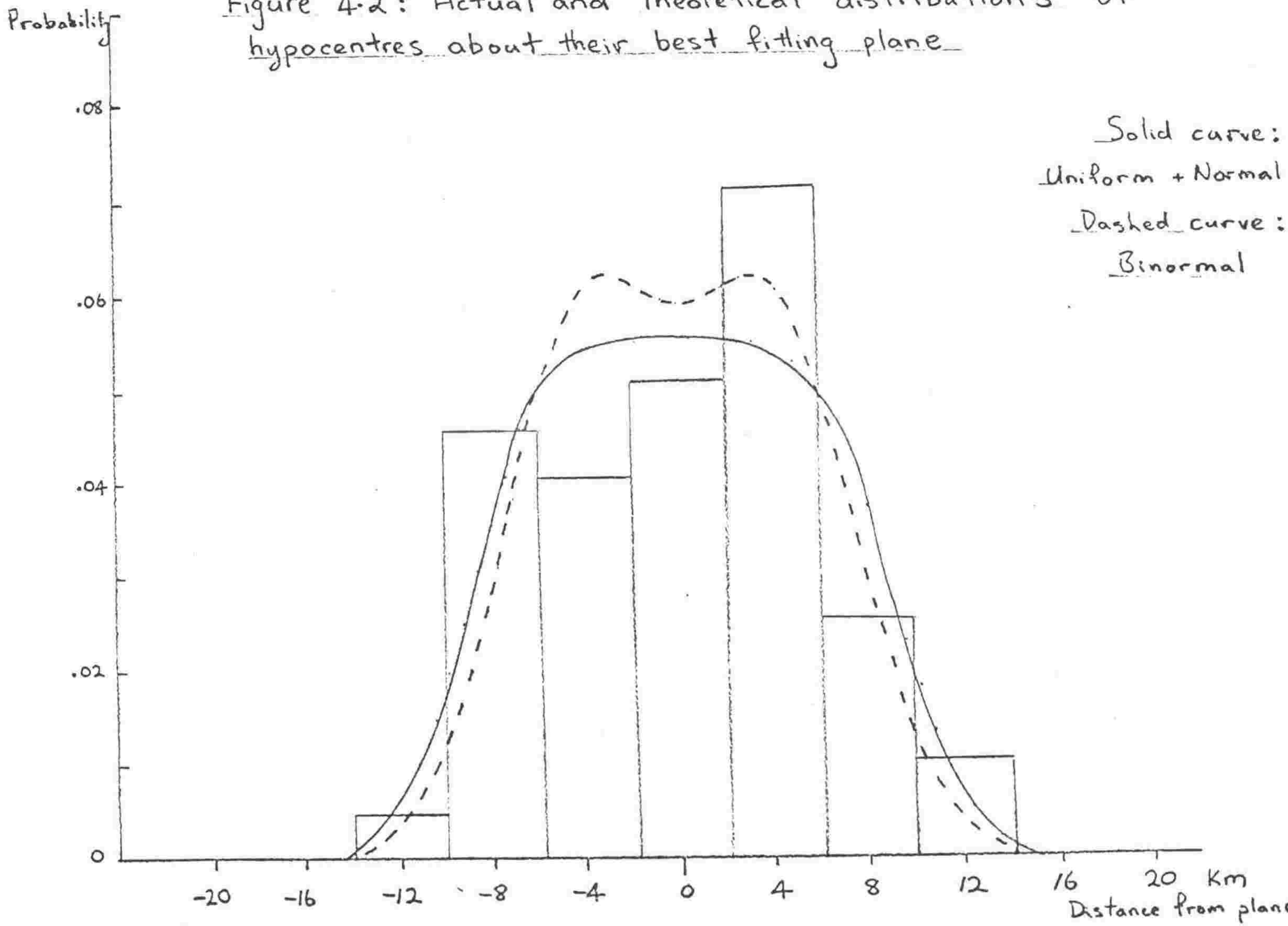
Also by way of comparison, Hamilton and Gale obtained standard deviations about their fitted median curves of 9.7 to 17.5 km. They interpreted this, in view of their estimated location errors, as indicating a standard deviation of the parent hypocentre distribution of 5.0 to 14.3 km. We will now analyse our results with a view to ascertaining the parent distribution.

STATISTICAL PROPERTIES OF THE BEST FITTED PLANE

We can interpret the distribution of hypocentres as being precisely planar if 5.6 km is the standard deviation of the relative location error in the direction normal to the plane. In order to estimate this quantity, and also to check the assumption that this quantity is constant for all the earthquakes, the standard deviations in the direction of the normal were calculated for the 49 earthquakes from their relative location variance matrices using $\hat{\sigma} = 0.39$ sec. as the scale factor, as previously described. The assumption that the standard deviations are equal is certainly not true, which means that the hypocentres $\underline{\hat{x}}_i$ in the likelihood function of Appendix IV, should be replaced by $\underline{\hat{x}}_i/\sigma_i$, where $\sigma_i = \sqrt{\underline{\alpha}^T V_i \underline{\alpha}}$. To solve the resulting maximum likelihood equations directly would be impossible because of the appearance of the new term involving $\underline{\alpha}$. One could proceed in two steps by calculating $\underline{\alpha}$ without weighting and then using this estimate to calculate the σ_i and thence recalculate $\underline{\alpha}$. However, upon examination, the values of σ_i did not differ by more than a factor of 2, with an average value of 3.5 km. In view of the significant difference between 5.6 km and 3.5 km, the hypothesis that the hypocentres could originate in a plane must be dropped. However, the best fitting plane still serves as a convenient reference frame for the description of the distribution.

The next hypotheses to be tested were that the parent distribution could be modelled as a uniform slab of width 2a, or as two planes distance 2a apart. In the first case, the distribution of the distances perpendicular to the best fitting plane, d_i, would be a convolution of a uniform distribution and an

Figure 4.2: Actual and Theoretical distributions of hypocentres about their best fitting plane



error distribution. The error distribution assumed was normal with zero mean and standard deviation 3.5 km (as predicted from the parameter variance analysis). The observations are then described by:

$$d_i = u_i + \varepsilon_i \qquad \ldots(4.21)$$

where $f(u_i) = \frac{1}{2a}$ when $-a \leq u_i \leq a$; $= 0$ otherwise.

$$f(\varepsilon_i) = \frac{1}{\sqrt{2\pi}} \times \frac{1}{3.5} \exp \{-\frac{1}{2} \frac{\varepsilon_i^2}{(3.5)^2}\} \qquad \ldots(4.22)$$

On the assumption that u_i and ε_i are independent, which is strictly valid only where the only source of relative location error is reading error, the distribution of d_i is given by the convolution of $f(u_i)$ and $f(\varepsilon_i)$:

$$f(d_i) = \frac{1}{2a} \int_{\left(\frac{d_i-a}{3.5}\right)}^{\left(\frac{d_i+a}{3.5}\right)} \frac{1}{\sqrt{2\pi}} e^{-w/2} dw \qquad \ldots(4.23)$$

From the log-likelihood function:

$$L = \log \prod_{i=1}^{M} f(d_i)$$

a can be estimated by the method of maximum likelihood. In fact, the equation arising by setting $\frac{\partial L}{\partial a} = 0$ is quite intractable and a was estimated by trial.

Similarly, in the second model with two planes, the distribution of d_i is given by:

$$f(d_i) = p\frac{1}{\sqrt{2\pi}}\frac{1}{3.5}\exp\{-\frac{1}{2}(d_i-a)^2/3.5^2\}$$
$$+ (1-p)\frac{1}{\sqrt{2\pi}}\frac{1}{3.5}\exp\{-\frac{1}{2}(d_i+a)^2/3.5^2\} \quad ..(4.24)$$

From the log-likelihood function, p and a may be estimated by trial since the equations $\frac{\partial L}{\partial a} = 0$ and $\frac{\partial L}{\partial p}$ are analytically insoluble. The distribution seemed so symmetric that only $p = ½$ was tried. The values obtained were: model (1), $a = 9$ km, model (2), $a = 4$ km, implying structure thicknesses of 18 km and 8 km for the two models. The actual values of the maximum likelihood for the two models were almost the same, and so these two models of very different character give quite different estimates of the thickness of the structure. The models are compared with the actual distribution of perpendicular distances in Figure 4.2. Since neither model is particularly realistic, but that each might be considered a limiting case of the family of feasible parent distributions, the two values obtained for the structure thickness, 18 km and 8 km, can be interpreted as approximate bounds for the thickness of a plane slab in which it is hypothesized the earthquakes originate.

FINE STRUCTURE OF THE DISTRIBUTION OF PERPENDICULAR DISTANCES

In Figure 4.3 the positions of the hypocentres with respect to the best fitting plane are plotted with their perpendicular distance in kilometres from the plane.

The horizontal axis is a vector striking 45° E of N, and the vertical axis dips at about 67°. The lateral dimensions of the plane are 300 km × 200 km, but there are large gaps in the

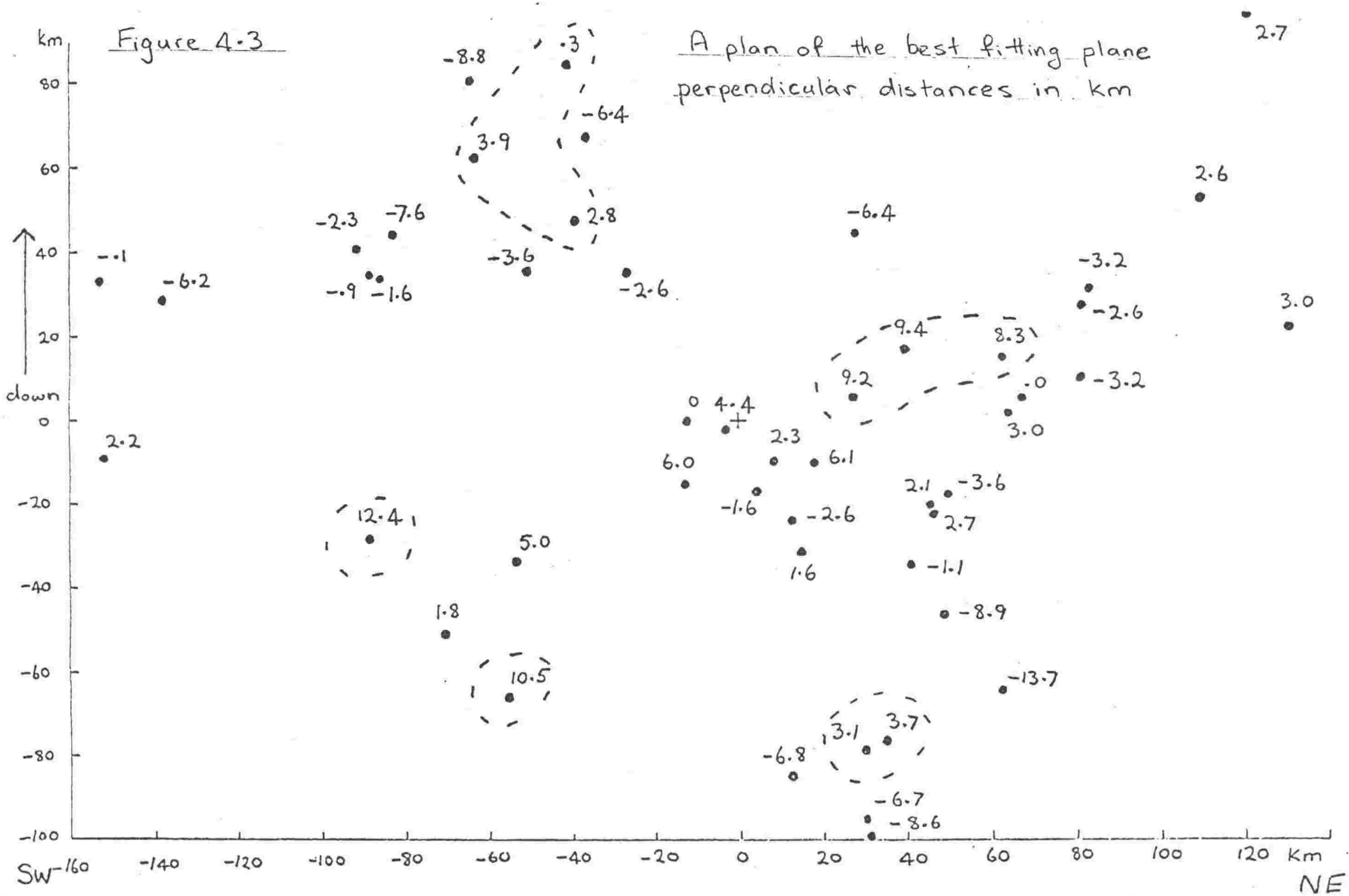

km
down
SW
km
NE
A plan of the best fitting plane
perpendicular distances in km

Figure 4·3

seismicity, particularly at the SW end. This absence of seismicity is readily explained by the increasing failure of station ECZ to obtain satisfactory (non-emergent) readings as one progresses further and further from it. Indeed it would be dangerous to draw any conclusions about the seismicity within the structure from the sample here because of the large dependence on station magnification and signal reception.

If one examines the pattern of perpendicular displacements given in Figure 4.3, it rapidly becomes apparent that there is a high correlation between displacements on a local scale. Because of this, a second order model $d_i = ax_i^2 + by_i^2 + cx_iy_i + dx_i + ey_i + f + \varepsilon_i$ was fitted to the perpendicular displacements. The assumption is that the error in the perpendicular direction is independent of the errors in x and y, and that these latter are negligible. Examination of the orientation of the 90 percent confidence ellipses in Figure 4.4 shows that the first assumption is quite good but that the second is quite wrong. Notwithstanding, a quadratic surface was obtained and the standard deviation of the displacements from the quadratic surface was 4.2 km. Using an F test of significance, this improvement from 5.6 to 4.2 km is significant at the 1 percent level, but the difference between 4.2 km and the predicted 3.5 km is also significant at the 1 percent level. This indicates that a simple second order surface is an inadequate description of the structure also.

If one re-examines Figure 4.3, one notices that there are several small groups (enclosed by dashed lines) of hypocentres which have remarkably similar displacements, but all standing about 10 km above the neighbouring hypocentres. With no more

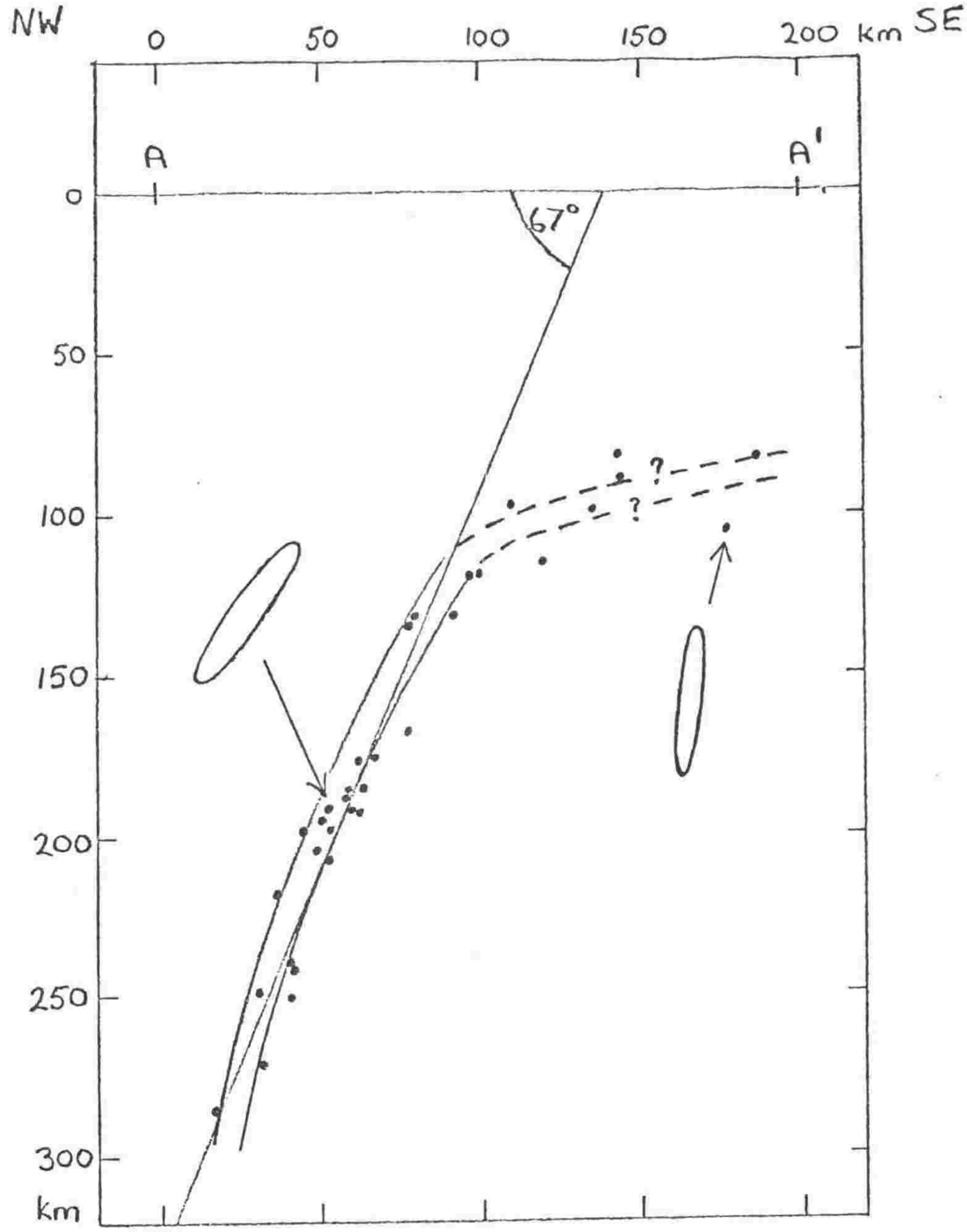


Figure A·4 : Section through North Island Seismic Zone (AA' of Fig. 4·1)
Ellipses are 90% relative location confidence regions.

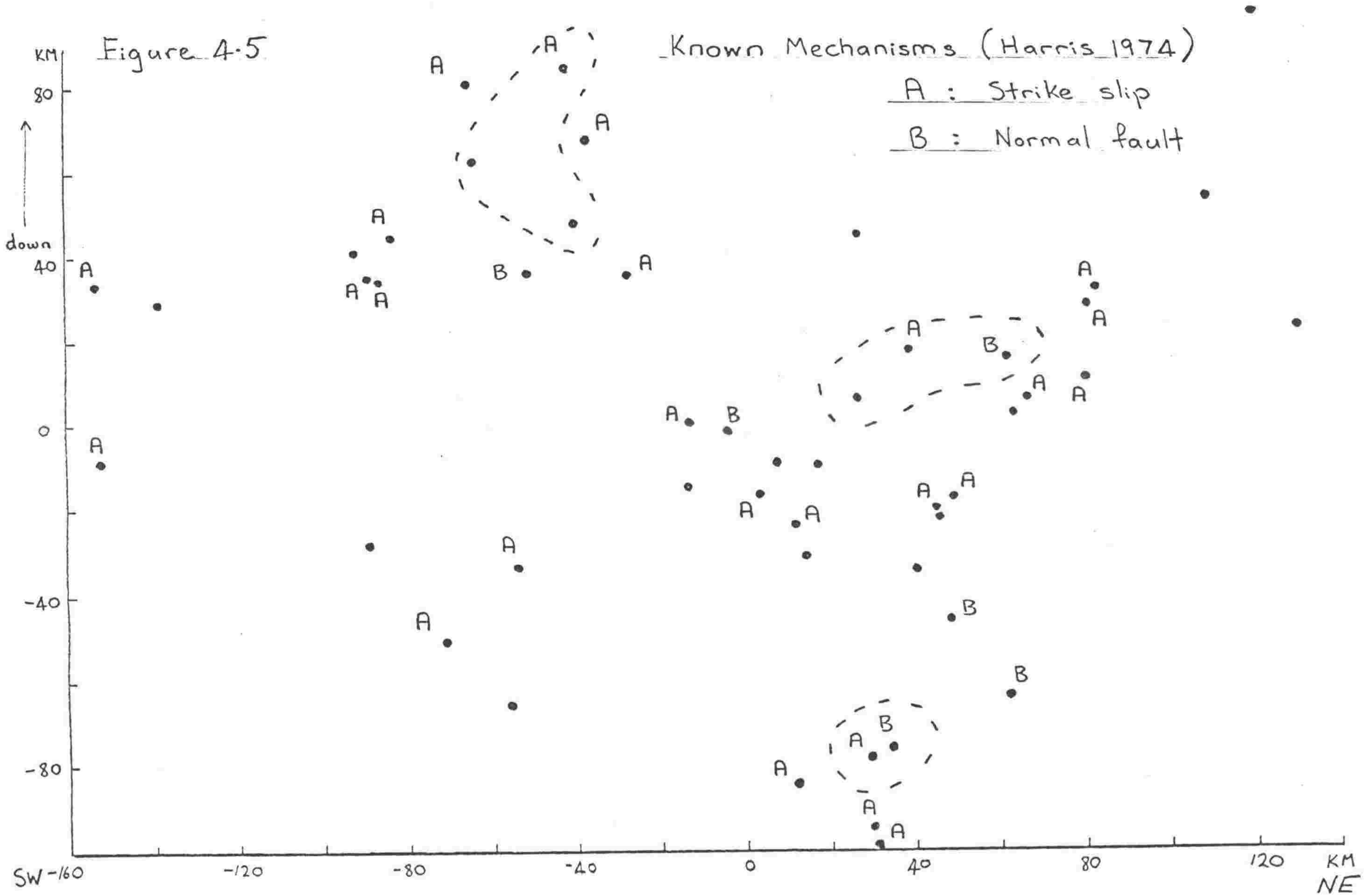

Known Mechanisms (Harris 1974)
A : Strike slip
B : Normal fault
KM
80
40
0
-40
-80
down
SW -160
-120
-80
-40
0
40
80
120
KM
NE

Figure 4.5

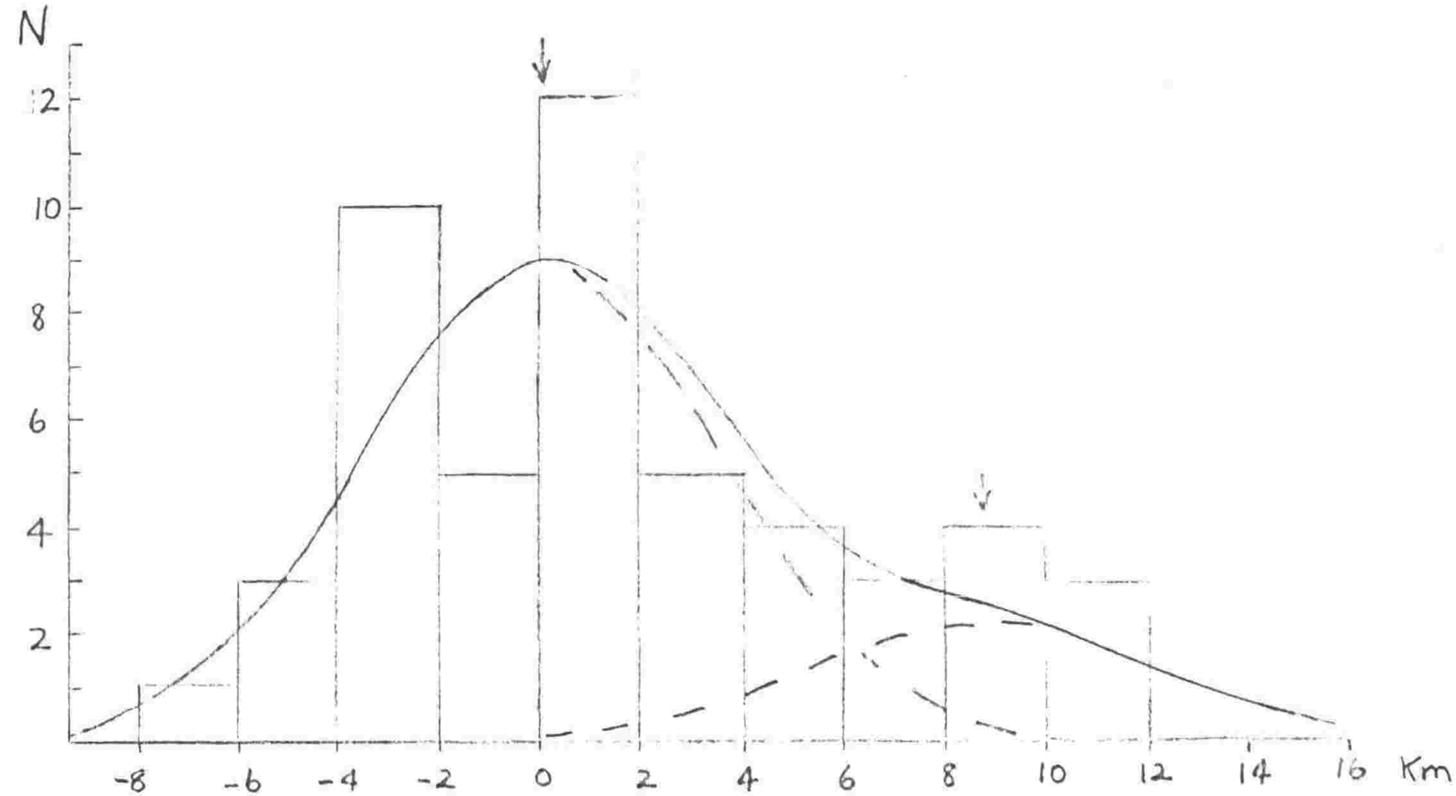


Figure 4.6 Distribution about parallel quadratic surfaces (arrows). Solid curve is the sum of two normal distributions (shown dashed), $\sigma = 3{\cdot}5$ Km.

justification than this, these ten hypocentres were removed and a quadratic fitted to the rest. The standard deviation of the 39 about this new quadratic was 3.4 km. The 10 stood at a mean distance of 8.8 km above this surface, the standard deviation of this group about its mean being 1.5 km.

Figure 4.4 shows a vertical cross section normal to the strike showing the projection of this quadratic surface, a parallel surface 9 km away (and the original best plane) and all these hypocentres lying within 50 km of *AA'* (Figure 4.1). Also shown are the 90 percent confidence ellipses in the plane of the section for two of the events and a hypothetical extension of the zone (dashed lines). In view of the remarkably good fit of this model, one can say that the distribution of hypocentres is described very well by two slightly curved surfaces 9 km apart with rates of seismicity in the two surfaces in the ratio of 1 : 4.

The curvature of surface in the direction of the Strike was an order of magnitude less than in the vertical plane shown. The surface was in fact saddle-shaped being slightly concave toward the NW.

CONCLUSION

It is important not to read too much into these results. First and foremost, the structure in which the earthquakes originate is very thin: two of the models gave thicknesses of 8 and 9 km, the other gave a thickness of 18 km. The physical implications of this will be discussed below, but for the moment it suffices to note that other subcrustal seismic zones have been found to be of similar thickness — see, for example, Engdahl (1973) and Wyss (1973).

It must be remembered that the assumptions lying behind the theory of the relative location errors cannot be expected to hold over the whole extent of the region considered. The thickness estimate depends on the assumptions holding within a region of larger dimensions than the estimated thickness, and so is the most reliable piece of information about the fine structure of the zone. Examination of Figure 4.3 leads one to prefer the twin curved surface model as being the one which tries to account, at least in part, for the pattern of perpendicular displacements. If indeed there are two surfaces or ultra thin sheets in which all the hypocentres originate, the sorting of the hypocentres into the two groups is a nearly impossible job and no pretence is made that the selection and removal of 10 from the 49 picked all or only those from one group. One could of course try different partitions of the set of hypocentres. It is unlikely however that by this means any more would be learned about the structure.

The physical significance of the curvature of the surface is another question. The twin planes model and the quadratic pair model both fit the data well. It is entirely possible that over the downward extent of the structure, about 200 km, the systematic location error alters in such a way as to make the depth errors of the deeper earthquakes larger than the depth errors of the shallower events. This would explain an apparent curve. Thus little physical significance should be attached to the apparent curvature of the structure shown in Figure 4.4.

For the same reason the dip of the best plane may be appreciably in error. Subsequent work (Chapter V, and Adams and Ware (1976)) indicates a dip of 50°-60° rather than 67°.

The question of whether the thinness of the structure continues into the region shallower than 120 km is made difficult to answer because of the orientation of the confidence regions. The long axis of the confidence ellipse of the earthquake at 190 km in the section of Figure 4.4 is almost parallel to the down dip of the structure. It is fortuitous that the direction of minimum error is nearly equal to the direction of the normal to the best fitting plane. For the shallower earthquake, however, the long axis of the ellipse is almost normal to the hypothesised structure thus making quite plausible the idea that the shown earthquakes could have originated between the dashed lines. Again, subsequent work has considerably changed the picture of the seismicity in the range 33-100 km, and the problems of these earthquakes will be discussed in Chapter V.

To summarise, it is beyond doubt that the homogeneous station method in this application gives higher quality relative locations than the standard method, and the high degree of agreement between the theoretical prediction of errors and the scatter of the hypocentres with respect to the derived models of the regions in which they originate suggest that the limit of relative location accuracy has been reached for this quality of data. The quantity of 0.39 sec. for the standard deviation of the total (relative) time error is unlikely to be made much smaller by more sophisticated location techniques without the introduction of more complicated travel time models.

PHYSICAL IMPLICATIONS OF THE RESULTS

Surveys of several seismic zones by different methods have revealed that the earthquakes originate in very thin structures. Engdahl (1973) finds a 10 km thick zone in the Aleutians after examination of cross sections of seismicity, the earthquakes being located using a raytracing method to calculate travel times and a homogeneous station set, so that his work is another example of the homogeneous station method. Engdahl used a velocity model constructed by considering teleseismic observations of arrivals from the nuclear explosion Longshot which featured a 7 percent faster *P* wave velocity within a dipping slab than in the surrounding mantle. The earthquakes were located towards but not exactly in the top of the slab, in a structure parallel to the slab which Engdahl hypothesises is the coldest and brittlest part of the slab.

Wyss (1973) deduces a thickness of 11 km for the thickness of the intermediate seismic zone in the Tonga area from source dimension considerations. Determinations of the thickness of the intermediate and deep seismic zones in other regions generally give values of 10-30 km, but error estimation problems when standard location methods are used tend to make these maximum thicknesses for the zones.

Although there is no way of identifying the physical region within the slab in which the earthquakes studied have originated from the data discussed here alone, it is usually assumed that this zone lies towards the top surface of the down-going slab. There is no reason to suppose that Engdahl's Aleutian model is not a reasonable one for the North Island New Zealand mantle zone.

Some recent work by Harris (1975) classified more than 200 earthquakes in the North Island mantle zone into two groups according to mechanism. The two mechanisms are the common mechanism of tension down the dip of the zone with compression normal to the zone and a strike slip mechanism in the plane of the zone with the tension axis to about 30° to the down dip. About 70 percent of the earthquakes had this latter mechanism; of those located below 200 km, more than 90 percent had this mechanism. Of the 49 earthquakes in Figure 4.3, mechanisms of 31 were given. These are shown in Figure 4.5. There appears to be no correlation between location of an event in one of the surfaces of the two surface model and a particular mechanism. No further discussion of mechanism and its consequences will be attempted in this chapter but will continue in the following chapters in the sections on the model of the North Island mantle.

CHAPTER V

JOINT HYPOCENTRE DETERMINATIONS OF NORTH ISLAND NEW ZEALAND MANTLE EARTHQUAKES

AN EVALUATION OF THE JOINT HYPOCENTRE METHOD

In this chapter we set out to test the Joint Hypocentre method on groups of North Island mantle earthquakes. The theory of the method is given in Chapter III. Joint Hypocentre Determination (JHD) has not before been applied in a situation quite like the one discussed here for, although Dewey's (1971) study in Venezuela involved analysis of some intermediate depth events, Dewey always used a master event and had available substantial teleseismic data including pP observations. For reasons to be discussed, only readings from the net of stations run by the Seismological Observatory at Wellington, New Zealand, were used in this study. Also, I was very keen to discover the problems encountered when one does not use a master event. Various authors using JHD have been somewhat vague about this, saying that they had no success or that convergence was too slow. Because of an anticipated difficulty in being confident about the location of any master event in the region to be studied, I felt that JHD without master event was worth trying. In this section, the theory of JHD is expanded to cover the problems arising in a particular situation and its performance is considered under three headings: relative locations, absolute locations, contribution to model improvement. Chapter VI contains a description of the use of JHD to model the mantle structure of part of the North Island of New Zealand.

DATA SELECTION

It was anticipated that the bulk of the data for the study would come from seismic arrivals at New Zealand seismograph stations of earthquakes subsequently located at depths greater than 33 km by the Seismological Observatory, Wellington. I expected to be able to include in my groups of earthquakes some with sufficient teleseismic readings to make them reliable master events. The discrepancy between locations assigned by the Observatory, particularly the depth, using only New Zealand readings and the Jeffreys-Bullen travel time model, compared with the locations assigned to the same earthquakes by ISC after teleseismic data has been added is well known and Adams and Ware (1977) show that this discrepancy can be reduced to a large extent by assuming 9 percent faster mantle travel times to some stations than others in a way which is consistent with the presence of dense material under the North Island (Hatherton (1970b)). A study of teleseismic residuals by Robinson (1976) independently reinforces the model of a structure dipping at about 50 degrees in which the seismic P velocity is 10-11 percent higher than in the surrounding mantle. Similar values have been found for other Benioff zones - in the Aleutians, for example, Jacob (1972) deduces a 7-10 percent velocity contrast from residual analysis of the nuclear explosion Longshot.

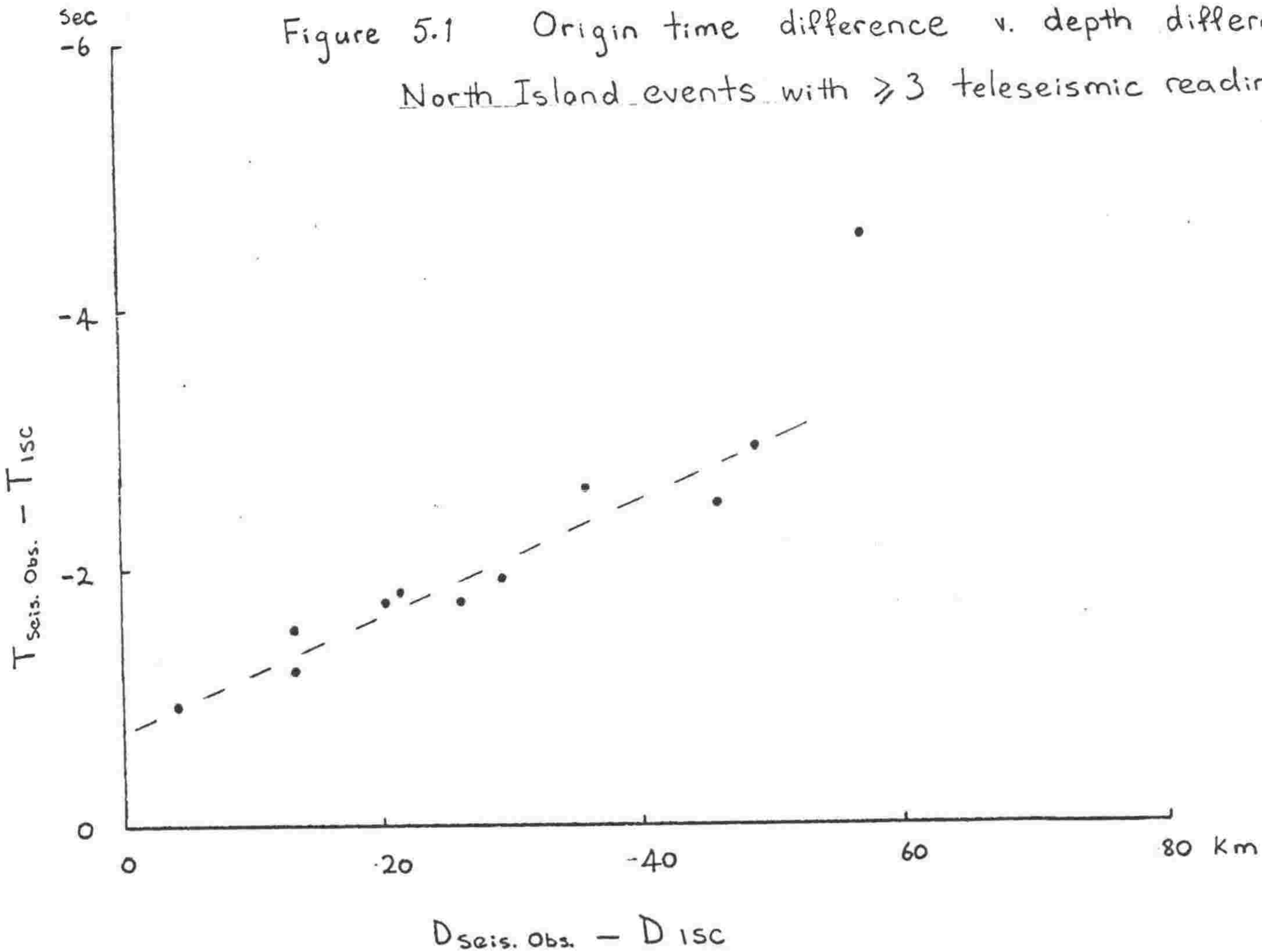
Because of this severe departure from a radially symmetric travel time model, hypocentre determinations based on the latter must be suspect even when substantial teleseismic information is included, for rays to teleseismic stations from an earthquake at, say, 150 km presumably will penetrate down through a significant amount of the fast structure.

Figure 5.1 shows a plot of origin time difference against depth difference, i.e. Observatory origin time - ISC origin time versus Observatory depth - ISC depth for the sample of all the earthquakes from January 1971 to June 1973 inclusive which were assigned subcrustal depths by ISC south of 38°S under New Zealand for which there were three or more teleseismic P readings. The figure illustrates the known linear relationship between the depth and origin time errors. The spread of values shows the difficulty in deciding, assuming that earthquakes in this class are chosen as master events, what an appropriate master event solution would be. Also, in this sample there were no pP observations, and the consistency with which any given teleseismic station provided observations was not high. Thus it was decided to try at first to use JHD without teleseismic data and without a master event.

The next problem was to sort the earthquakes into groups suitable for the application of JHD. The groups should have dimensions not greater than about 60 km diameter (from Chapter III) if the required assumptions for the statistical analysis of the results are to be valid. Moreover, the groups should contain as many earthquakes as possible with as high quality data as possible - in particular, as consistent a set of stations as possible. These last two requirements are bound to conflict since any applied quality criteria are bound to reduce the size of the available population.

First, earthquakes assigned subcrustal depths by the Seismological Observatory for the years 1964-1969, and 1971-1974

Figure 5.1 Origin time difference v. depth difference
North Island events with ≥ 3 teleseismic readings



inclusive, between latitudes 37°S and 42°S were classified into arbitrary groups according to latitude and depth. The following table gives the classifications:

Latitude ↓ \ Depth →	34-49	50-99	100-149	150-199	200-249	250-299	300-
37.0-37.9S	11	12 (13)	13 (53)	14 (85)	15 (104)	16 (68)	17
37.9-38.7	21	22 (23)	23 (62)	24 (317)	25 (117)	26 (26)	27 (10)
38.7-39.2	31	32 (12)	33 (67)	34 (101)	35 (68)	36 (11)	37 (0)
39.2-40.0	41	42 (29)	43 (39)	44 (41)	45 (70)	46	47
40.0-40.7	51	52	53	54	55	56	57
40.7-41.3	61	62	63	64	65	66	67
41.3-42.0	71	72	73	74	75	76	77

The figure given in parentheses is the total population of the group - e.g. the most populous group, 24, had 317 events during the ten years. The groups without population figures had very few earthquakes and it was felt that these would not be processable. Subsequent work revealed that at least some of the groups south of 40°S could be analysed but this has not been attempted in this thesis. One of the features of the group classification is the almost total absence of seismicity in the range 34-50 km (33 km is the depth assigned to "normal" earthquakes by the Observatory). Indeed, most of the seismicity in the 50-100 km range is deeper than 70 km. This sparseness of seismicity has been noted by several authors (see, for example, Eiby (1971)), but the work of Adams and Wave (1977) and the results here suggest this gap is caused by using a too slow model which has the effect here of increasing the depths of subcrustal earthquakes by about 40 km. The earthquake groups north of 37.9°S were not analysed either, because of poor conditioning of the system of equations resulting from these events (lying outside the net of stations).

The question of conditioning will be discussed later, but at this point it suffices to say that this ill-conditioning was deduced from the appearance of very large standard deviations for the parameter estimates.

The next question was the choice of stations. To determine the quality and consistency of the data, the most populous group, 24, was examined to try to find data quality criteria for the selection of subgroups on which to use JHD.

EARTHQUAKE SELECTION WITHIN GROUP 24

Because it was not initially known how stable, well conditioned and sensitive to data heterogeneity the method would be, a fairly severe rule was found which produced what was hoped would be a satisfactory number of earthquakes. Twenty-one "stations" were decided upon, and for the purpose of JHD it is convenient to think of P and S readings at a given seismograph station as separate stations. The stations were selected according to the criterion that each station should record at least half the earthquakes in the group. The stations were CNZ, P, GNZ, P, S, KRP, P, S, TUA, P, S, ECZ, P, S, MNG, P, S, TRZ, P, S,

WTZ, P, S, WEL, P, S, COB, P, S, TNZ, P, GBZ, P. (See Figure 5.5 for locations of stations.)

All those earthquakes from group 24 were then selected which had: non-emergent P arrivals at stations CNZ, KRP and either GNZ or TUA; a total of at least 15 arrivals at the 21 stations; not fewer than 11 non-emergent arrivals (P or S) in total. This provided a sizeable population of 65 events out of the 317.

A note here is necessary about the use of S arrival data. It is common for intermediate depth earthquakes in New Zealand, particularly those located by the Observatory below about 100 km to give clear, non-emergent S arrivals on New Zealand seismographs. The information contained in these S data is too valuable to waste, and so it was intended to use the S arrivals, treating them independently of the P data, but with a suitable weight relative to P.

In the same way, it was decided to use rather than omit emergent arrivals by giving them a suitably low weight relative to a non-emergent arrival. The method used for obtaining these weights and the method of dealing with excessive residuals is discussed in the next section.

WEIGHTING THE EQUATIONS OF CONDITION

The theory of the distribution of the total error in an observation has been discussed in Chapters II and III. Although this distribution cannot be assumed to be normal, even after the removal or special treatment of extreme values, the theory of normal errors will give satisfactory confidence intervals for parameter estimates provided the observations are independent and of equal variance. The problem of non-independence of observations because of coupling through the model error at a station was extensively discussed in Chapter III with the conclusion that the correlation between errors should be satisfactorily decreased for small enough regions. Absence of correlation does not imply independence for non-normal distributions, but this minimising of the correlation is all one can do short of estimating by some means the full variance matrix of the observations and using this to transform the equations of condition into an uncorrelated, common variance system.

With the approach adopted here one must be prepared to examine the confidence intervals obtained for the parameter estimates to see whether they appear to be too small. Highly correlated errors at a given station may produce an effect on the parameters similar to a bias, that is, a systematic error, which might not be encompassed by the calculated confidence regions.

It remains therefore to ensure that the observations have equal variance, or, if as is almost always the case of the equations of condition and certainly in the case now being discussed, that the equations are weighted inversely as the standard deviation of the observations. The resulting transformed system (in the absence of correlations) satisfies the requirements of the Gauss-Markov Theorem (Zelen (1962)) and so the resulting parameter estimates are minimum variance unbiased estimates, irrespective of the error distribution.

To determine the weights, it was assumed that each observation had an error variance of the form:

$$\sigma_{\varepsilon}^2 + \sigma_{s_j}^2$$

where σ_{ε}^2 is the variance of the reading error and $\sigma_{s_j}^2$ is the variance of the deviation about the mean of the travel time model error at the j^{th} station. Initial assumptions were made that the readings classified as emergent and non-emergent come from two populations; the reading errors for each having common variance irrespective of whether the arrivals are P or S, and that the model error variance for P or S was common for all stations. Also it was assumed that $\sigma_s^2 = 3\sigma_p^2$ from the normal relationship that the P wave velocity is about $\sqrt{3}$ times the S wave velocity.

This leads to the assumption that the weight for non-emergent S relative to non-emergent P is given by:

$$w_s = \sqrt{\frac{\sigma_{\varepsilon}^2 + \sigma_p^2}{\sigma_{\varepsilon}^2 + \sigma_s^2}} = \sqrt{\frac{\sigma_{\varepsilon}^2 + \sigma_p^2}{\sigma_{\varepsilon}^2 + 3\sigma_p^2}} \qquad \ldots(5.1)$$

Taking $\sigma_{\varepsilon} \doteqdot \sigma_p$ gives:

$$w_s \doteqdot 0.7 \qquad \ldots(5.2)$$

There was some immediate suggestion that this was too high a relative weight (the assumption of equal reading errors being a very poor one) and the weight for S was taken to be 0.6.

Emergent arrivals were given a weight of 0.5 relative to non-emergent arrivals. This guess was based on personal communications from some of the Seismologists who made the observations but subsequent tests of the residuals revealed that this is quite a reasonable value.

Using these values, the set of 65 earthquakes from group 24 were relocated using JHD. There were 1125 equations in $4\times65 + 21 = 281$ unknowns (no master event was used) to be solved for the least squares solution by the algorithm given in Appendix I. Of the 1125 observations, 579 were non-emergent P arrivals, 421 were non-emergent S, 57 were eP and 58 were eS; 10 observations exceeded 3 standard deviations ($\hat{\sigma} = 0.45$) and were not used in the following analysis.

Under the assumption that the input weights were correct, the distribution of residuals in the three classes, non-emergent P, non-emergent S, and emergent (P and S), would have equal variances. Significantly differing variances would indicate that the wrong weights had been used. Because of the large numbers of residuals in each class, quite small differences in variance would be significant (using an F test). The results were that:

$$\left.\begin{aligned} \hat{\sigma}_p^2 &= (\Sigma r_p^2)/579 = 0.1124 \\ \hat{\sigma}_s^2 &= (\Sigma r_s^2)/421 = 0.1620 \\ \hat{\sigma}_e^2 &= (\Sigma r_e^2)/115 = 0.1119 \end{aligned}\right\} \qquad \ldots(5.3)$$

so that:

$$\frac{\hat{\sigma}_s^2}{\hat{\sigma}_p^2} = 1.44, \qquad \frac{\hat{\sigma}_e^2}{\hat{\sigma}_p^2} = 1.00 \qquad \ldots(5.4)$$

which implies that the weight for emergent arrivals relative to non-emergent ones is satisfactory, but that the weight for non-emergent S should be reduced to:

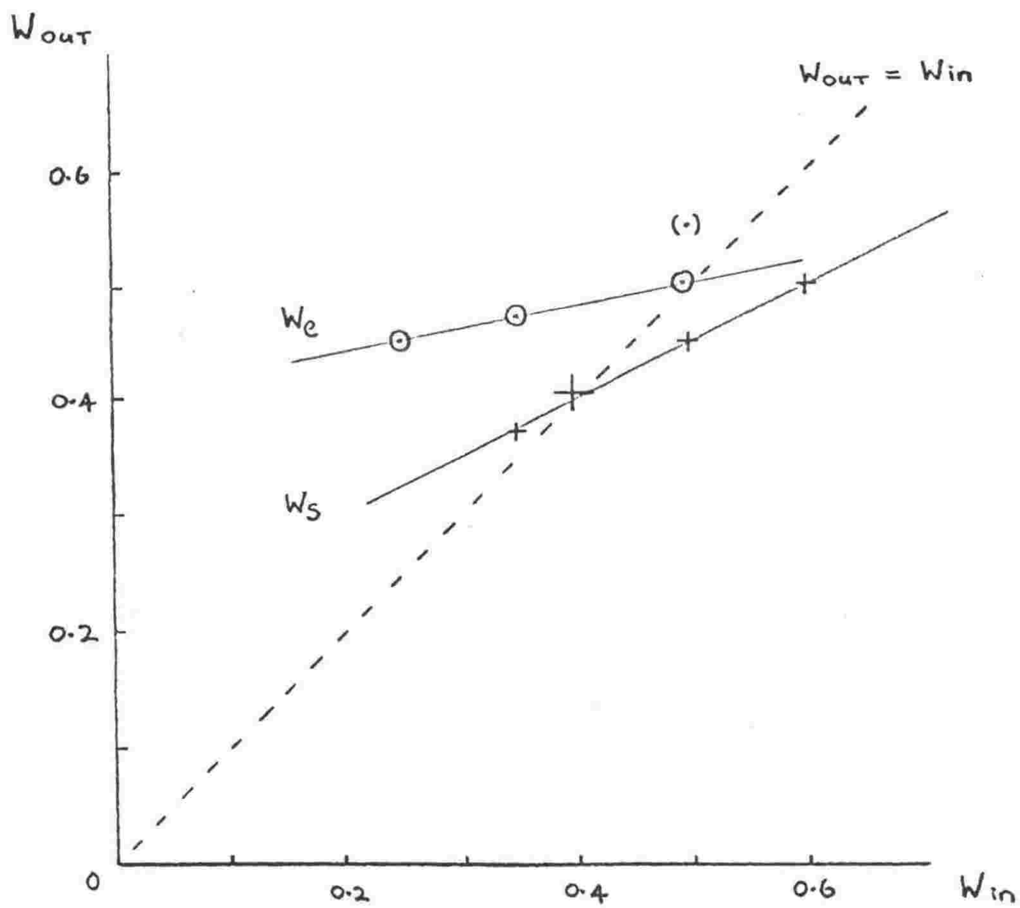
$$\frac{w_s}{\sqrt{\dfrac{\hat{\sigma}_s^2}{\hat{\sigma}_p^2}}} = \frac{0.6}{1.2} = 0.5 \qquad \ldots(5.5)$$

This experiment was twice repeated for different input weights. It was found that the predicted weight plotted against the input weight gave a straight line. The results are shown in Figure 5.2. The intersection of the straight lines through the data points with the line $w_{in} = w_{out}$ gives the values w_s = 0.4, w_e = 0.5. When these weights were used in fact the output weights were w_s = 0.41 and w_e = 0.55 (shown as a large + and (·) on Figure 5.2). If a single large residual were removed from the population of emergent arrivals (and the rule used for the removal of outliers was quite arbitrary), then w_e would be 0.50.

The conclusion of this analysis was that the adoption of the weights w_s = 0.4 and w_e = 0.5 produced internally consistent results. These values are not too far from the initial estimates although the S weight is lower than expected for a reason that I can only assume is due very much to large reading error - doubling the estimate to 0.4 sec. for the S reading error only gives w_s = 0.63 - coupled with greater variation in the S model errors than predicted by simple theory.

No subsequent tests on other groups of earthquakes gave any reason to depart from this weighting scheme which was used

Figure 5.2 Determination of $W_s : W_p$ and $W_e : W_p$



throughout this study. In the section which follows (discussing the stability of the hypocentre and model correction estimates to perturbations of the input control variables), the change in the solution for the various data sets is given. No very great change in the parameter estimates (nothing significant compared to the relative error standard deviation estimates) is observed for small changes in the input weights.

The next aspect of the weighting problem to be dealt with was the method of dealing with outliers - residuals too large to be accounted for by any reasonable theory of errors other than blunders in reading. The number of mistakes detected was in fact very small for the group of 65 events, but rather than omit these, it was felt that since good techniques exist for dealing with such situations, one could be used to deal with the possibility of greater numbers of extreme outliers in other groups.

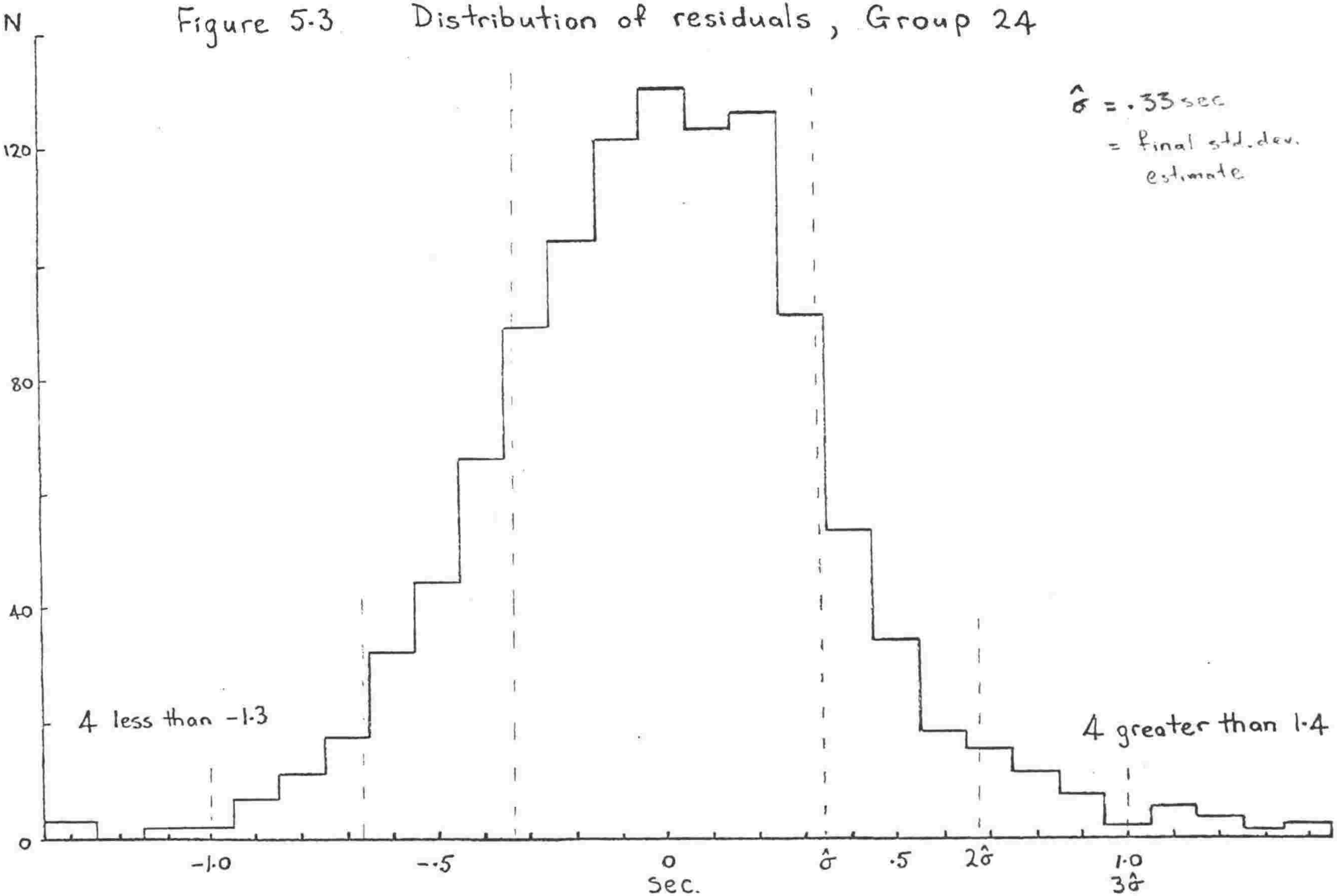
Dewey (1971) used Jeffrey's method (Jeffreys (1939)) which employs a weight:

$$w(r) = 1/[1 + \mu \exp (r^2/2\sigma_0^2)] \qquad \ldots(5.6)$$

where σ_0 is the standard deviation of the population without extreme values, and μ is related to the rate of occurrence of extreme values.

Dewey used Bolt's (1960) value for μ of 0.02 and estimated σ_0 from the distribution of residuals, typical values being of the order of 1 sec. The very low rate of outliers found in the

Figure 5.3 Distribution of residuals, Group 24



group 24 data indicated that 0.02 was much too large. Figure 5.3 shows the distribution of residuals without outlier removal. Because this analysis was being conducted concurrently with the determinations of w_s and w_e, the data shown is from the case $w_s = 0.6$, $w_e = 0.5$. The distribution of residuals is markedly non-normal in that a very much greater than expected number of residuals are within 0.45 sec., the unbiased estimate of the standard deviation, but also that too many exceed three standard deviations - as mentioned before, 10 residuals were greater than 1.3 sec. in magnitude, and most of these were more than four standard deviations. It was not the case that these extreme values tended to occur for emergent arrivals.

This gives a "blunder rate" of 10 per 1125. Calculation of μ as (height of tail)/(excess height of peak) gives a value of 2×10^{-3} - an order of magnitude less than the Bolt-Dewey figure. By trial, a satisfactory value of σ_0 was found to be 0.35 sec. The unbiased estimate of the total error standard deviation which resulted from the use of these weights on the group 24 data was 0.33 sec., which is to be interpreted as the standard deviation of the total error in a non-emergent P observation. This is not inconsistent with the estimate of 0.39 sec. obtained in Chapter IV for the same quantity, since no weighting for outliers was attempted there, and the detection of blunders with only seven observations per event is made very difficult by the least squares method which tends to share the error among all the residuals.

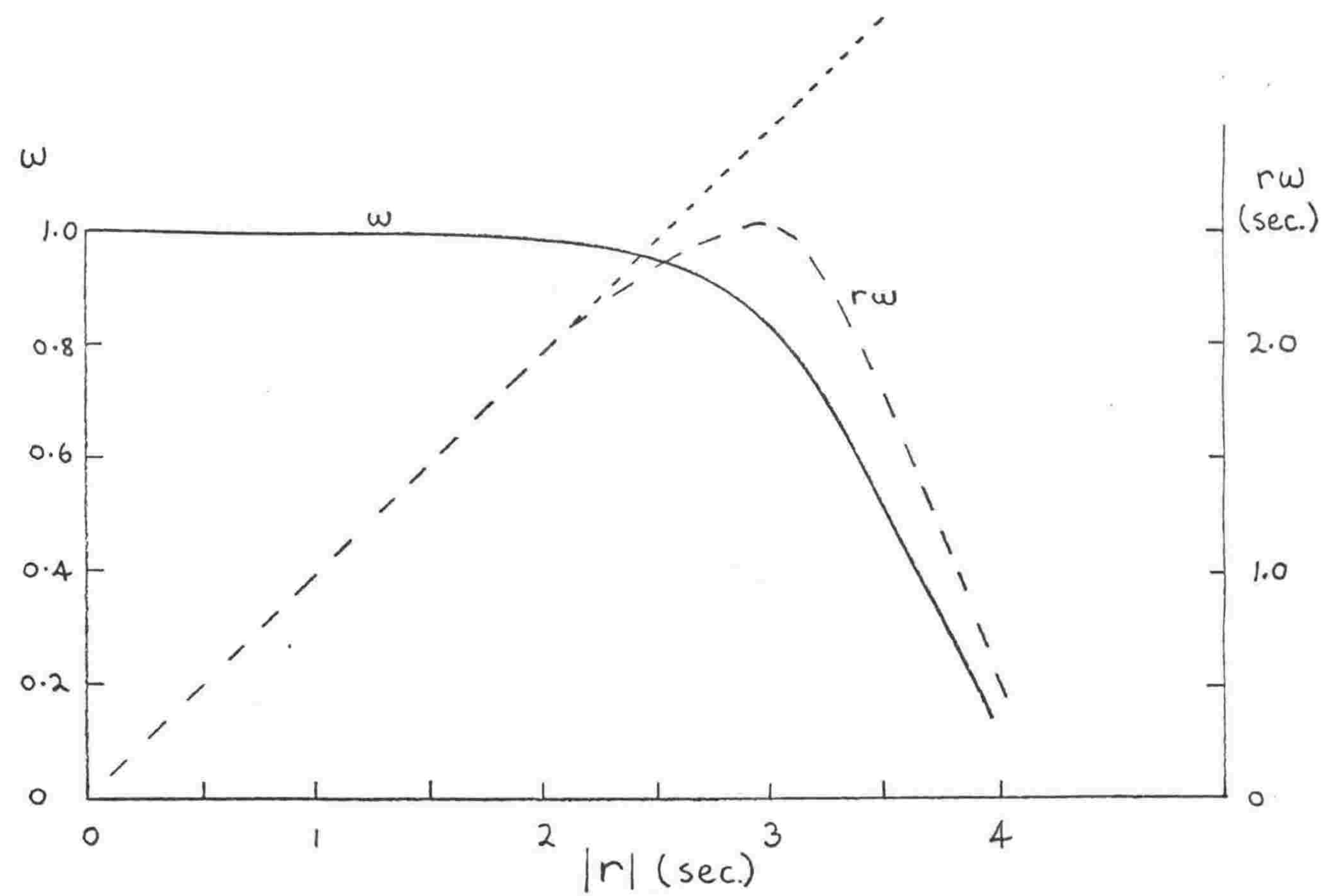
In this regard it cannot be claimed that residuals identified as extreme outliers are certainly blunders and that these are the only such blunders, although this is extremely likely. There is, however, the possibility that a "good" reading will appear to be a blunder and that a blunder will be hidden. The actual method used to apply the Jeffreys weighting scheme was that some number of iterations, generally 3-5, would be performed without Jeffreys-type weighting, then $w(r)$ would be calculated and applied unchanged for a number of iterations until a sufficiently stable solution (as indicated by the magnitudes of the increments) was reached.

Figure 5.4 shows $w(r)$ (normalised to $\sigma_0 = 1$) and $rw(r)$ for $\mu = 0.002$. The weighting produces little effect on residuals less than three standard deviations, but then its effect is rapid. Experiments were carried out with different values of μ and σ_0 and it was found that the solutions were stable to small perturbations of these parameters, but that if too large a value of μ were used, say 0.02, a marked change in the solution occurred after the application of Jeffreys' weighting from which the system took much longer to recover and find a new minimum.

The final distribution of residuals after application of the weighting scheme described was little different from the prior distribution except for the absence of the tails. Before concluding this section on weighting, we return to the assumption of equal station variances made at the beginning of this section. It is generally accepted that this assumption is false (Freedman (1968)), but until one removes the mean model error, detection and removal of the effect of different station variances is difficult.

For each of the 21 stations used, a standard deviation:

Figure 5.4 Jeffreys Weighting : $\omega = 1/(1 + 0.002e^{r^2/2})$



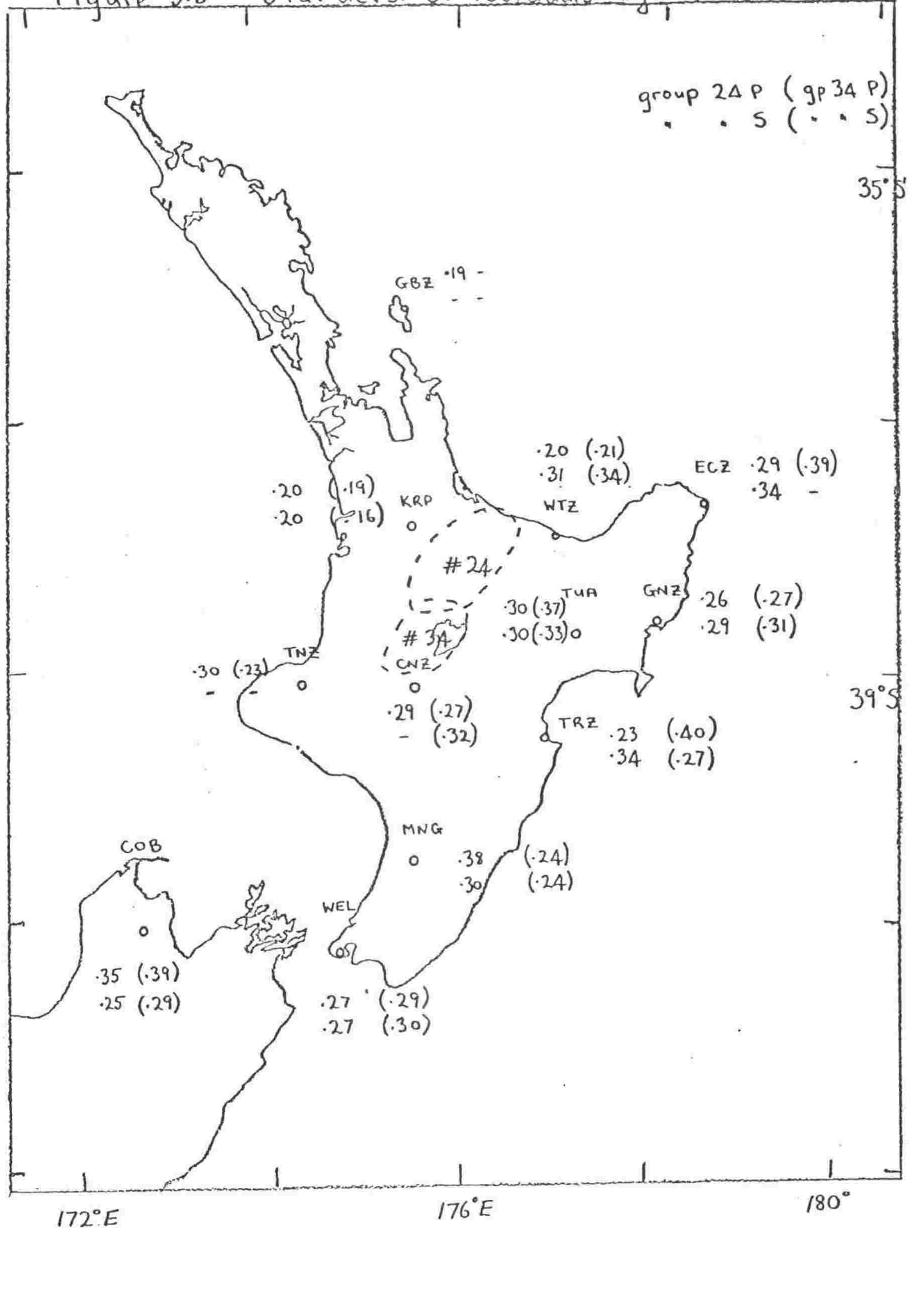
$$\hat{\sigma}_j = \sqrt{\left(\frac{\sum_{i=1}^{N_j} r_{ij}^2}{N_j - 1}\right)}$$

was calculated, where N_j is the number of observations at the j^{th} station and r_{ij} the weighted residual from event i. These quantities ranged from 0.19 (GBZ P) to 0.38 (MNG P). The values obtained are shown in Figure 5.5. There is no readily apparent pattern in the values. Although at most seismographs the P standard deviation was not greater than that for S, the two largest values were MNG P and COB P. A Bartlett's Test of Homogeneity of variance (Snedecor and Cochran (1967)) performed on the station data indicated that the differences between the standard deviations was highly significant (the test statistic is approximately X^2 on 20 d.f. and the value obtained was 73.13).

Ad hoc explanations of these differences might be attempted. For example, the COB P data had by far the highest incidence of emergent arrivals of any station, and so too high a weight for w_e would explain the high value there. However, MNG is one of the better stations in the set and none of its arrivals were described as emergent.

It is perhaps convenient here to give the corresponding results for an adjacent group, group 34. The station standard deviations are also given in Figure 5.5 and it can be seen that for some stations the values are consistent and for others highly inconsistent. In particular, MNG P and TRZ P show large variations from one group to the other. A Bartlett's Test conducted on the group 34 data also showed a significant difference between station variances.

Figure 5.5 Std. devs. of residuals by Stations



In light of these results, I feel it is likely to be difficult to formulate a theory which explains the differences in station variance. A plausible explanation is in fact that the observed differences are attributable to incorrect weighting parameters for S, for emergent arrivals and in the Jeffreys weighting. If by chance it happened that a rather large number of large but not extreme residuals were present among the residuals of MNG P for group 24 and the value of μ used were too small, this would give a larger variance than expected. The Bartlett Test is sensitive to departures from normality in much the same way as the F test (see Chapter II).

Certainly, the great variation that occurs for some stations between groups 24 and 34 shows the dangers of simply weighting the observations at a station inversely as the calculated standard deviation. It was felt that the results so far justified not weighting the observations by stations but to bear in mind the very strong possibility of unequal station variances. This question is mentioned further when the pooled results of all the groups are discussed, when it turned out that there was a large degree of homogeneity of variance between stations.

STABILITY OF PARAMETER ESTIMATES IN GROUP 24

Appendix V contains a complete set of results for group 24 (with no master event), including original Seismological Observatory Solutions. In this section we discuss these results with respect to stability of the solutions. In particular, we discuss the variation in the solution with quantity of data, achieved by altering both the number of stations and events, and also the effect of fixing part or all of the hypocentre of one of the events.

In general terms, the hypocentres were fairly consistently displaced about 30 km to the NW and 40 km shallower than their Seismological Observatory solutions. Average model corrections (the sum over the 21 stations being constrained to be -21.0 sec.) ranged from +5 sec. for KRP S to -5 sec. for GNZ S. The unbiased estimate of the total error:

$$\hat{\sigma} = \sqrt{\frac{\sum_i \sum_j r_{ij}^2}{\text{Number of Observations} - \text{Number of Parameters}}}$$

was 0.326 sec. This is to be interpreted as the estimate of the standard deviation of the total (model plus reading) error of a non-emergent P arrival. Total error estimates for other observations are obtained by dividing through by the weight given to that observation - thus the estimate for non-emergent S is 0.81 sec.

Even with 845 degrees of freedom, one is not entitled to quote 3 significant figures for $\hat{\sigma}$. This is done here for comparison with the values of $\hat{\sigma}$ given by the other solutions with different quantities of data, etc. The change in $\hat{\sigma}$ is so small as to be completely insignificant even where part or all of the hypocentre of one event is fixed. In particular, fixing the depth of a master event to a wide range of values produced insignificant changes in the value of $\hat{\sigma}$ showing that the "valley" of the least squares solution is broad and flat, which must make one cautious about acceptance of the solution. The details of the various solutions are given below. (Solution F is the one given in Appendix V.)

Solution	Number of Station	Number of Events	Master Event	Depth of Master	$\hat{\sigma}$
F	21	65	N	-	0.326
G	21	38	N	-	0.319
H	19	65	N	-	0.335
I	21	19	N	-	0.328
J	17	65	N	-	0.324
K	14	65	N	-	0.322
F	21	65	N	-	0.325
R	21	65	Y	157	0.333
S	21	65	Y	125	0.334
T	21	65	Y	140	0.330
0	(Seismological Observatory Solution, 157 km deep)				

Figure 5.6 shows the changes for F through K in epicentre, depth with latitude, and model error for five arbitrary earthquakes which approximately define the physical extent for group 24, and five stations spread about the group. 90 percent confidence ellipses are calculated from the variance-covariance matrix of solution F. Error bars for the model errors are one standard deviation. It is clear that the changes in the solution produced by altering the number of stations and/or earthquakes are insignificant in terms of the predicted errors. One concludes that the solution is not sensitive to simple changes in the number of data provided there is some minimum number of observations, suggested by solution I which produces the greatest deviation from solution F, particularly in the model errors.

Figure 5.7 shows corresponding changes for F through 0. The model terms here are given for F and R only, plotted against model travel time for the group to the station. The error bars

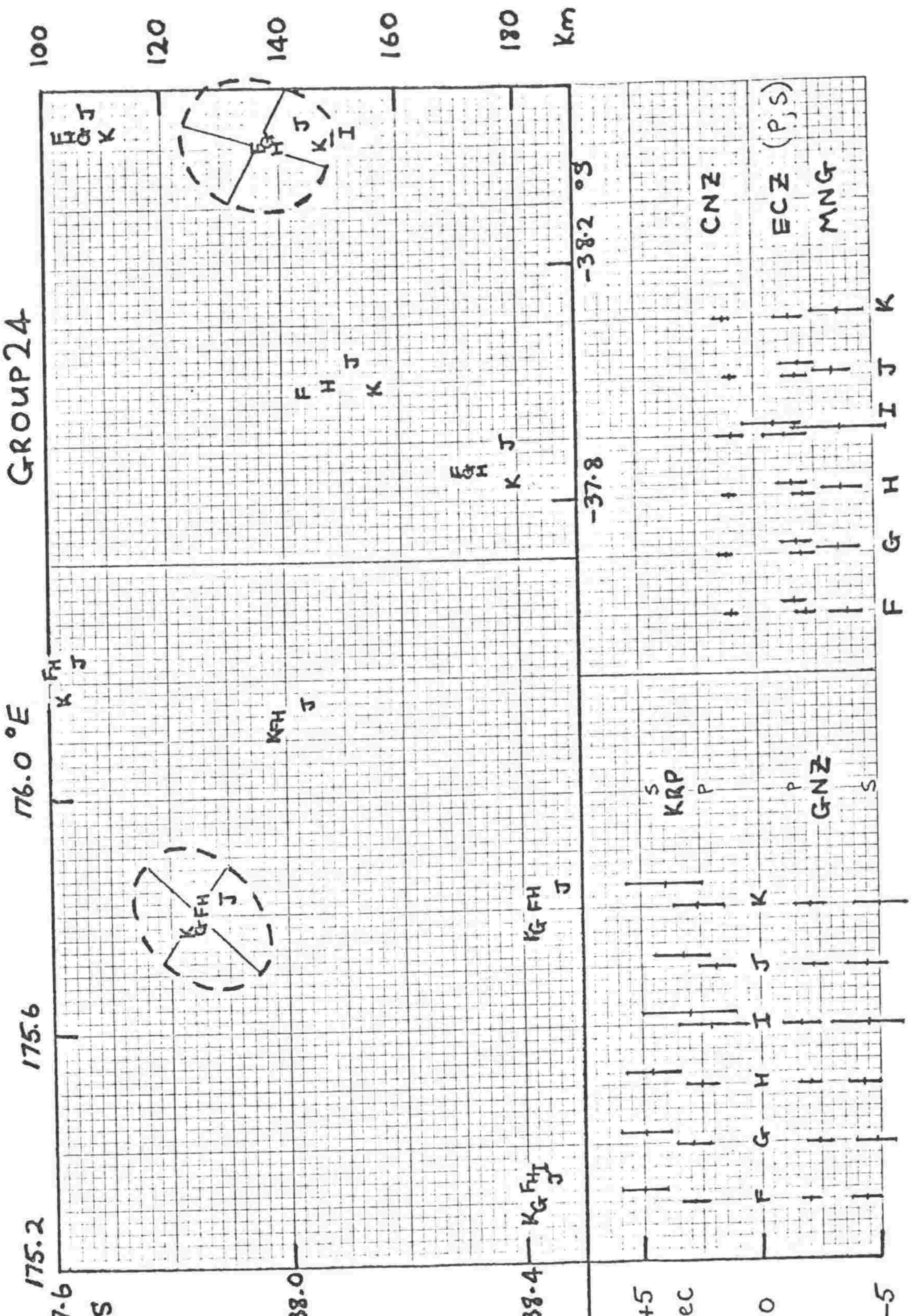

GROUP 24
176.0 °E
175.6
175.2
-37.6 °S
-38.0
-38.4
100
120
140
160
180
Km
-37.8
-38.2 °S
CNZ
ECZ (P,S)
MNG
KRP
GNZ
+5
SeC
0
-5

FIGURE 5.6

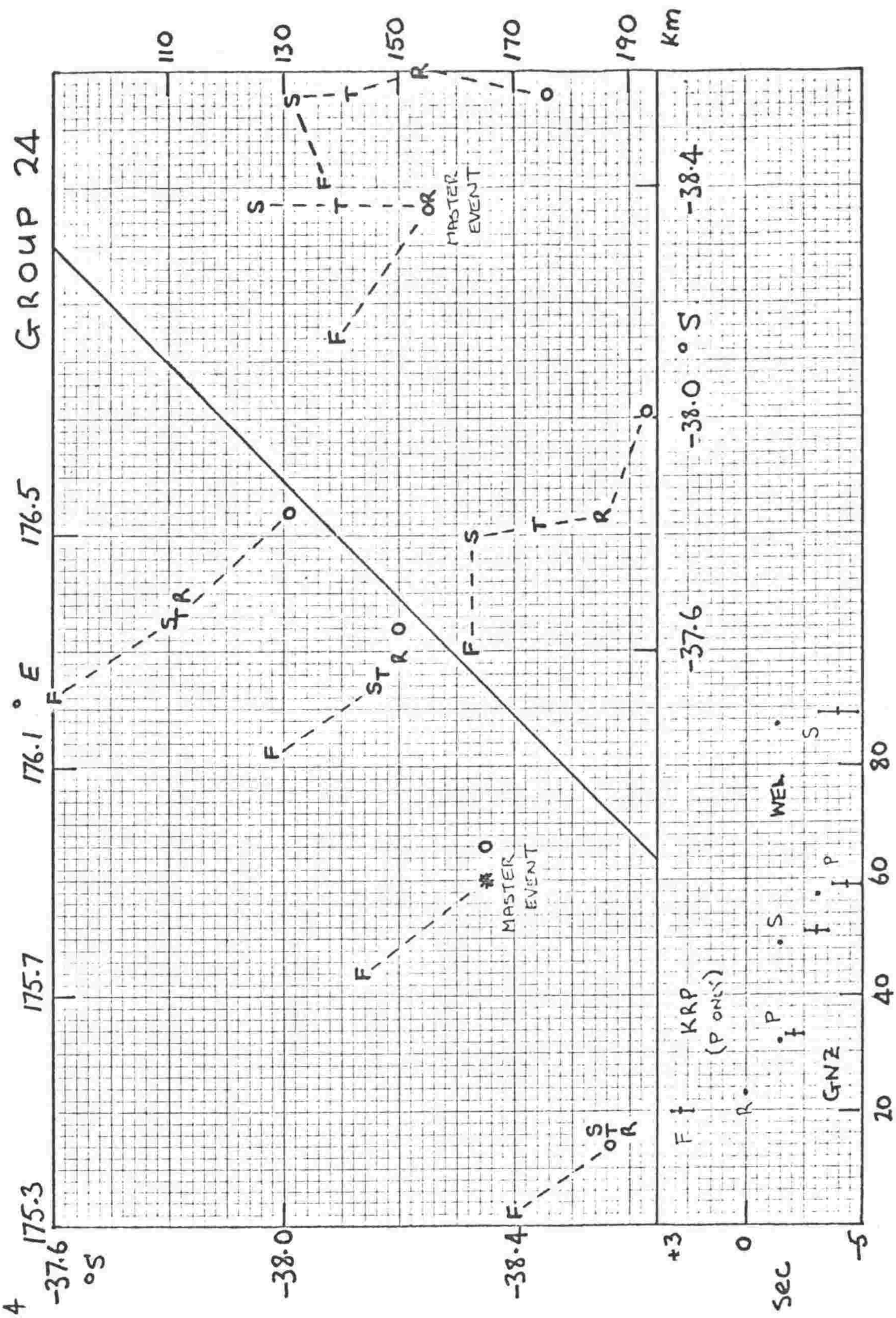


FIGURE 5.7 DIFFERENT DEPTHS OF MASTER EVENT

(on the F values) are one standard deviation. This figure demonstrates the high degree of reliability in the relative location of the group using JHD with or without a master - the solutions move in unison as the master is constrained to different solutions.

The conclusion from the lack of significant variation in $\hat{\sigma}$ between these solutions is that even though the conditioning of the system is good enough for the system of equations to converge quickly and stably (within 5 iterations, often within 3) to a solution, the number of solutions producing almost as small a residual sum of squares cover a vast volume of the solution space. Without some additional information about the physical appropriateness of a solution it would seem that the method is less than satisfactory as far as producing improved absolute hypocentres is concerned. Similarly, the change in the model error estimates between F and R shows that while qualitative information is provided as to whether the average path to a station is fast or slow relative to the original travel time model, the magnitude of the time correction at a given station will depend quite a lot on the values of the hypocentre parameters chosen for the master event.

THE DISTRIBUTION OF HYPOCENTRES IN GROUP 24

Figure 5.8 is a section perpendicular to the strike of the North Island Seismic Zone which serves as a location map for group 24 and again demonstrates the thinness of zone. The distribution of hypocentres in group 24 confirms the result of the Homogeneous Method - there the standard deviation perpendicular

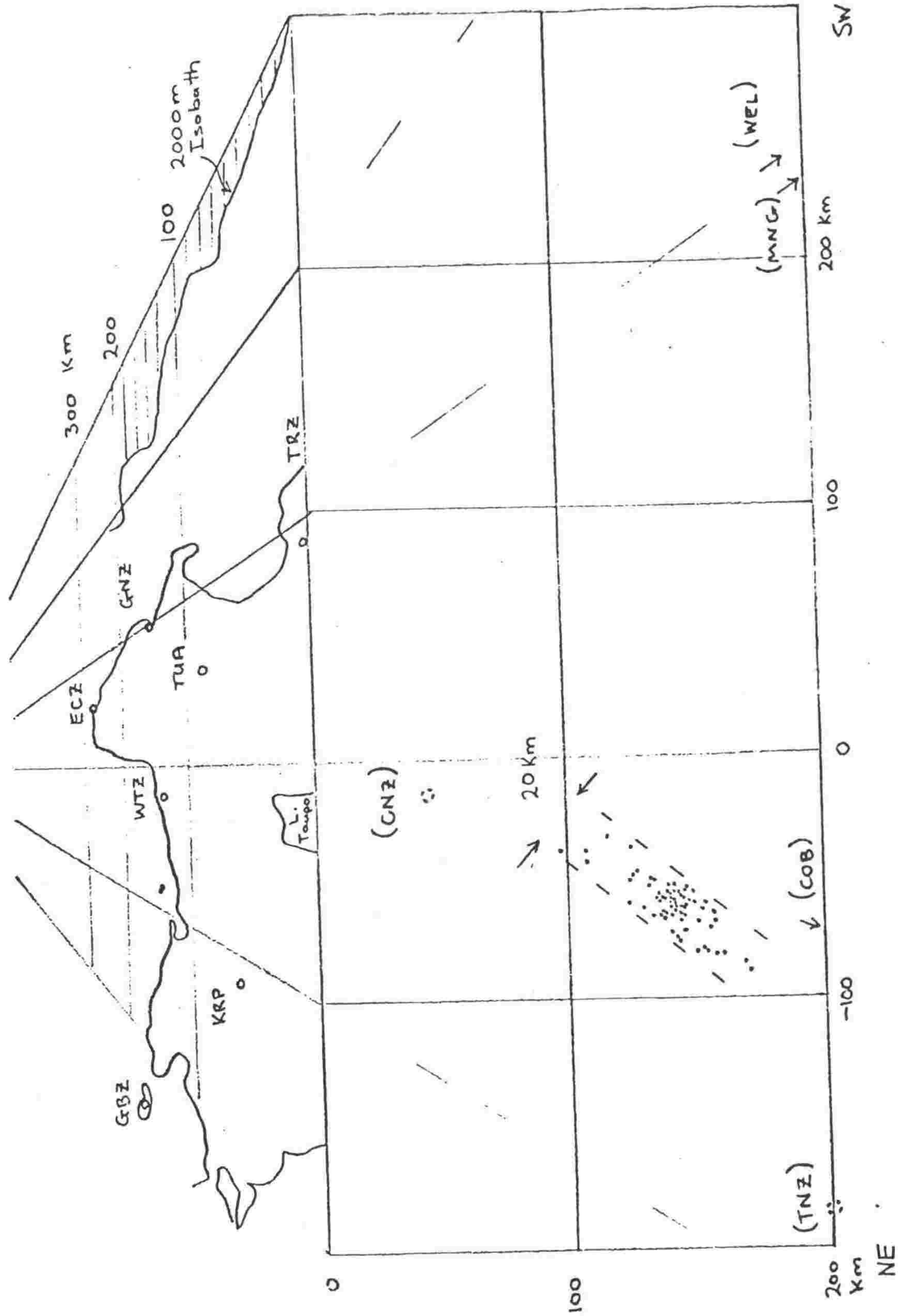


FIGURE 5.8 GROUP 24
SECTION THROUGH 39°S 176°E

to a best fitting plane was 5.5 km; for group 24 the standard deviation was 5.7 km. The geographic extent of group 24 is of course much smaller than the group considered for the Homogeneous Solution Method, with the extent perpendicular to the section in Figure 5.8 being only about ±50 km.

This agreement between the two methods suggests that both achieve the same standard of relative hypocentre determination. Moreover, a plane fitted to the hypocentres of solution R had a standard deviation perpendicular to the plane of 5.6 km, showing that the relative location accuracy is very largely independent of the location errors in the master event.

The other physical characteristics of the best fitting plane for group 24 were quite a bit different from those obtained in the previous study. Since there was about a 30 km displacement of the hypocentres to the North-West relative to their Observatory locations, the best fitting plane was similarly displaced. The strike was 41° East of North compared to 45° - a difference possibly not significant in view of the smaller extent of group 24, but the dip was only 51° compared to 67°. While it is again depth would mean that the model errors for earthquakes at the top and bottom of the group would differ by too much to make the assumption required for JHD valid.

However, the good relative locations and the shallower dip of the zone suggest that the group size chosen is not too large and it was with some confidence that the next stage of considering other groups was embarked on.

MODEL ERROR ESTIMATES FROM GROUP 24

Before discussing the results from the other groups, the model error estimates will be discussed. The great problem with interpreting these results was that the magnitude of the values obtained was often very large compared to the model travel time, the extreme cases being KRP and GNZ where the predicted errors were +15 percent and -6 percent of the average travel times, making a velocity contrast across the dipping zone of 20 percent. Of course the fact that the error terms are calculated only to within an additive constant means that by a suitable choice of constant, this figure might be somewhat reduced. However, one is hardly entitled to choose the additive constant to minimise the contrast. Also, the uncertainties in the model errors make quite dubious the acceptance of the results from a group in isolation.

To begin with, the standard deviation of the model error estimates as calculated on the basis of normal linear theory were of reasonable magnitude ranging from 0.22 sec. for TRZ P and TUA P to 0.95 for KRP S. The S standard deviations were larger because of the lesser weight given to S observations.

The geographic position of the station relative to the hypocentres and the total number of observations at that station are the other factors involved in determining the magnitude of the standard deviation.

The magnitudes of the error estimates and standard deviations were such that it would have been quite impossible to find an additive constant which would have made any appreciable number of the station terms insignificantly different from zero in a statistical sense.

The applicability of the standard errors as a measure of accuracy however depends on the absence of systematic errors and the series O-F reveal, as discussed, that $\hat{\sigma}$ changed very little over a wide range of solutions some of which (R in particular) involved a large reduction in the magnitude of the error estimate.

On the other hand, from the linear theory variance covariance matrix the correlation coefficients for the hypocentre parameters for one of the events and the model error estimates were calculated. The results are given in Appendix V. The results show only a mild correlation between depth or origin time and the station terms but a rather larger correlation between latitude and longitude and the station terms. These correlation coefficients then reflect the dependence between the model error estimates and, particularly, the epicentre estimate. The accuracy of the estimates of the correlations is subject to the same requirement of absence of systematic error, yet the values are not unrealistically small and should be quite fair estimates of the correlation. This indicates that the linear theory is coping with the systematic model errors to a large extent in which case one can accept the values for the standard deviations of the model errors.

Also, as before briefly mentioned, there is a great deal of qualitative consistency in the model error estimates. Those stations which have ray paths through the postulated down-going lithospheric slab have negative terms which increase in magnitude with increasing average travel time implying, as expected, faster than average velocities in the slab.

Thus it was felt that while in isolation the model error information from group 24 was of unknown quantitative value, the pooling of such data from several groups might yield an improvement to the model.

THE OTHER GROUPS

As an initial experiment, earthquakes from four other groups were relocated using JHD; these being the adjacent groups 23, 25, 33, 34 which were generally the next most populous groups. The northern groups 11-17 were rejected on the grounds that they were likely to be so unfavourably placed with respect to the network as to give rise to even more poorly conditioned systems of equations. Finally a total of nine groups were processed covering all but the regions of sparsest seismicity between

37.9S and 39.2S. The set of stations typically used to locate an event south of 39.2S by the Observatory is quite a little different from the set for events north of this latitude. It was hoped that some 250 events in all the groups would be a satisfactory number from which to obtain an improved velocity model.

Table 5.1 gives the details of the data available for the four groups mentioned above together with the number of iterations used and the largest increment to any of the model estimates in the last iteration. The maximum number of iterations allowed was 12 - if the increments became satisfactorily small before the twelfth, the process ceased. The values of the largest increment after 6 iterations, when Jeffrey's weighting was applied, is given for comparison in some cases.

The corresponding data for group 24 is included. In no case was a master event used.

TABLE 5.1

Group/Run	Events	Stations	Iterations	Last Increment	$\hat{\sigma}$
24F	65	21	6	0.04	0.326
G	38	21	8	0.13	0.319
H	65	19	8	0.04	0.355
I	19	21	10	0.05	0.328
J	65	17	9	0.01	0.324
K	65	14	8	0.09	0.322
34A	28	20	8	0.07	0.335
33A	20	17	12	0.46	0.317
			(6	0.59	0.318)
25A	31	20	12	1.68	0.363
			(6	.95	0.362)
B	31	19	12	.25	0.319
			(6	1.59	0.317)
23A	22	16	12	0.00	0.384
			(6	0.06	0.384)

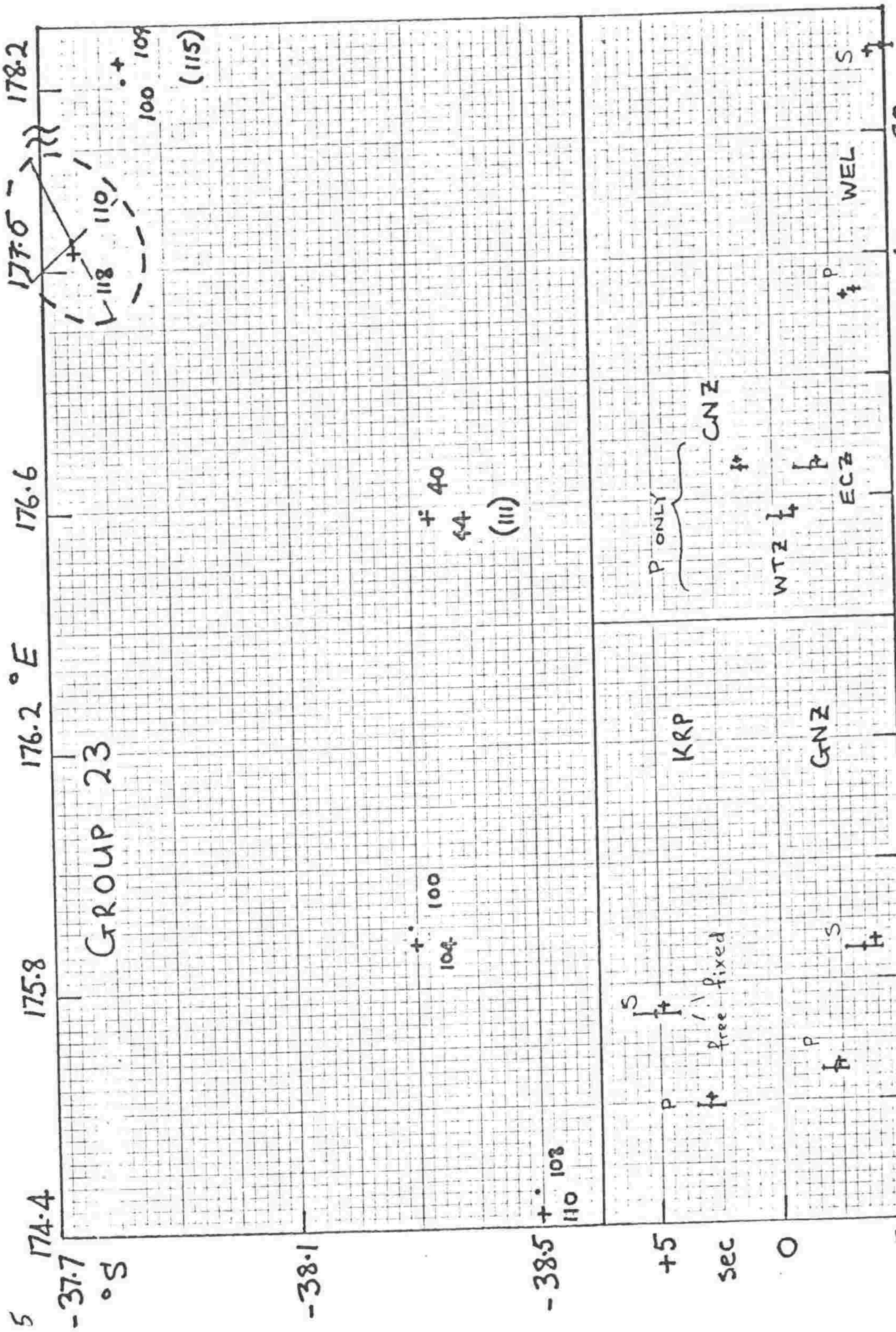
As in the case of group 24, the selection of data was guided by the need to monitor quality (few emergent arrivals) and consistency of stations, bearing in mind the available population of the group. Thus with the more sparsely populated groups, the criteria for selecting data had to be somewhat relaxed. The consistency of the results obtained from group 24 with varying data sets and the use of weighting for quality indicated that the relaxed criteria should be quite satisfactory. From Table 5.1 it will be noted that (i) except for group 23 the total error estimate $\hat{\sigma}$ was fairly constant and that except for 33 and 25 the last increment values are very small. In the case of group 25 it became apparent that P arrivals at station COB had a very large standard deviation. When this station was removed (run B) the stability increased quite appreciably. The extent of the influence of the COB residuals is further indicated by the drop in $\hat{\sigma}$ between runs A and B.

To be able to combine the model error estimates from the different groups to produce an improved model, it is necessary to ensure consistency of solution from group to group. To check this, an appropriate event from group 24 was seeded into each group (an event on the boundary between group 24 and the group in question). The hypocentres of these could be compared with their group 24 positions: if the difference between the positions is small compared with the predicted errors for the hypocentre, the results of the two groups are mutually consistent and the model estimates can be pooled. Further, the seeded events can be used as master events in their new groups fixed at their group 24 hypocentres. The set of model error estimates thus produced can be compared with the set produced when the group is free.

Figures 5.9 to 5.12 summarise the results. The ellipses shown are 90 percent confidence regions. The error bars on the model error estimates represent one standard deviation. The + symbols are the solutions with master event, ⊕ being that event. Confidence regions for all earthquakes in a group are similar and an event central in the group was chosen for which to calculate the confidence region.

Except for group 33, the master event solution fell well within the region and the model error estimates agreed within one standard deviation.

A detailed summary of results from all the groups will not be given at this stage because these results are superseded by the results discussed in the next chapter. At this stage a few immediate points can be noted. First, the average hypocentre displacement relative to the Observatory Bulletin solutions is much the same for all groups as for group 24. The shallower dip of the Benioff zone is seen in all groups, and the overall reduction in depth of about 40 km results in a vanishing of the apparent sparseness of the North Island seismicity between the base of the crust and about 80 km. Adams and Ware (1977) also note this point. The universality of this upward displacement also suggests that the bottom of the Benioff Zone is at a shallower depth than previous estimates based on seismicity profiles using Standard Bulletin Hypocentres. This point, which has the obvious consequence of indicating a shorter descending slab, will be further explored in the next chapter.



GROUP 23
174.4
175.8
176.2 °E
176.6
177.6
178.2
-37.7 °S
-38.1
-38.5
+5
sec
0
-5
KRP
GNZ
P ONLY
CNZ
WTZ
ECZ
WEL
free
fixed
(111)
(115)

FIGURE 5.9

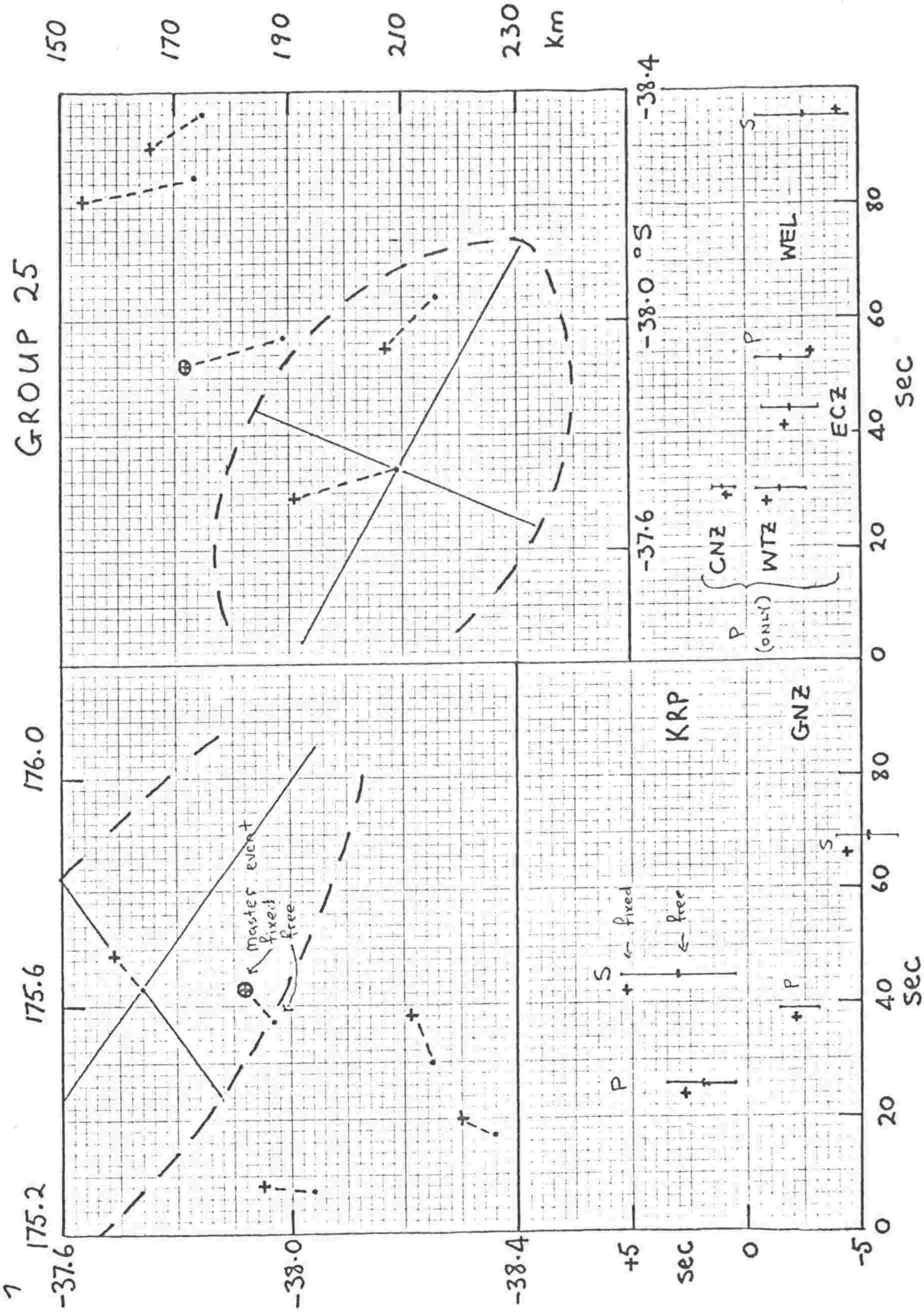

GROUP 25
150
170
190
210
230
Km
175.2
175.6
176.0
-37.6
-38.0
-38.4
-37.6
-38.0 °S
-38.4
master event
fixed
free
P
S
fixed
free
KRP
GNZ
P (ONLY)
CNZ
WTZ
ECZ
WEL
+5
sec
0
-5
0
20
40
60
80
sec
0
20
40
60
80
sec

FIGURE 5.10

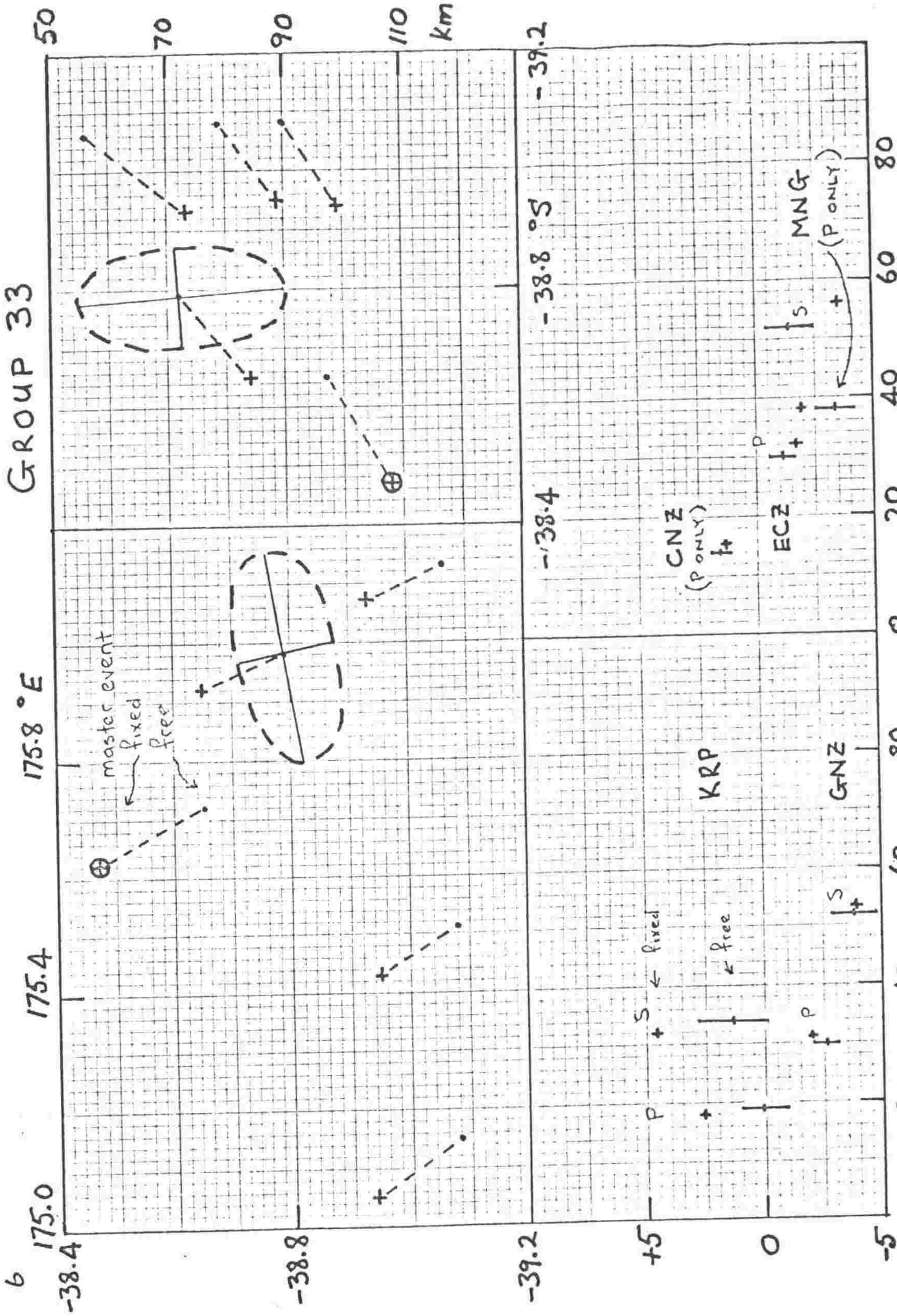

GROUP 33
175.0
175.4
175.8 °E
50
70
90
110 km
-38.4
-38.8
-39.2
-38.4
-38.8 °S
-39.2
+5
0
-5
master event
P fixed
P free
KRP
GNZ
CNZ (P ONLY)
ECZ
MNG (P ONLY)
P
S
20
40
60
80
0

FIGURE 5.11

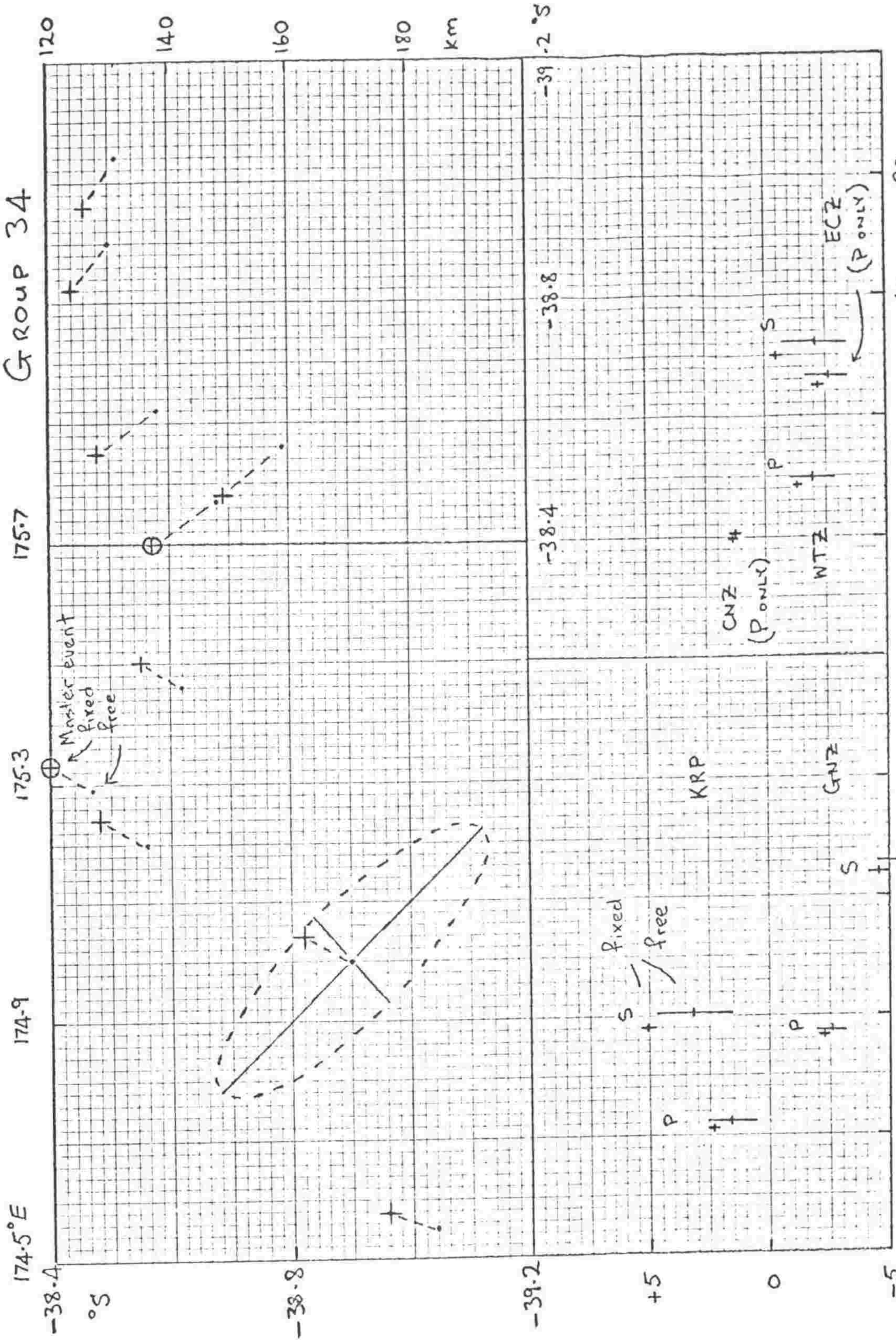

GROUP 34
Master event
fixed
free
KRP
GNZ
CNZ
(P ONLY)
WTZ
ECZ
(P ONLY)
km

FIGURE 5.12

INTERPRETATION OF THE MODEL ERROR ESTIMATES FROM THE GROUPS

This heading is the subject of the next chapter. It became apparent that a simple interpretation of the model error estimates would be unsatisfactory for the following reasons. First, the non-estimable mean model error for each group must be estimated or eliminated before intergroup comparison of model errors is possible. Second, the naive velocity contrast estimate is so large that if it were correct, the linearisation approximation on which JHD hangs would not be valid. (Chapters I and III). Third, the results of Haines (1977) show a large variation in the travel times in the crust locally to N.Z. seismograph stations. Ideally, one should improve the mantle model to the point where the new model error estimates are small enough to validate the linear approximation, and then interpret these estimates in terms of an average group term, an average station term (allowing for a crustal effect local to each station) and a mantle velocity contrast, with an assumed common contrast for stations with similar ray paths to any group.

CONCLUSION: THE ACHIEVEMENTS OF JHD

Briefly, the relocation of hypocentres of subcrustal earthquakes under the North Island, New Zealand using JHD has produced: better absolute locations, good relative locations, more information about the model than standard methods. The absolute location improvement is the hardest to judge in view of the absence of an absolutely determined event. The circumstantial evidence comprises the filling of the apparent gap in seismicity in the range 33-80 km and the generally better agreement in depth with teleseismic estimates. The gain in absolute location improvement is due to the avoidance of a master event. Although in many ways the distribution of stations used about the hypocentres is far from ideal (all looking down, with a large majority looking down through the slab), a master event was not required for stability. On the other hand, the appearance of large model error estimates and correlation between uncertainties in these estimates and location errors indicates that a substantial average location error may still be present.

The relative location question has been little mentioned so far - the evidence for this is the reproduction in all groups of the thinness of the Benioff zone deduced in Chapter IV. On the whole, little emphasis was placed on this feature of JHD, the major effort being directed towards model improvement. However, as an illustration of the capabilities of JHD in this direction we have the following example. Figure 5.13 shows a northern section of the east coast of the North Island, part of the region from which group 22/32 was drawn. Shown are all the epicentres of earthquakes assigned subcrustal (> 33 km) depths by the Seismological Observatory between 1964 and 1975. The arrow-heads show the JHD epicentres of all those which were selected for relocation (in group 22/32). The fact that only a very few are so treated makes the linear grouping of the central five and possibly the easternmost one more remarkable. The central five lie within 2 km of the line drawn. The strike of this lineation is 53°W of N (with appreciable uncertainty) and so is approximately perpendicular to the strike of the Benioff zone. The relocated depths are in the range 33-57 km (compared

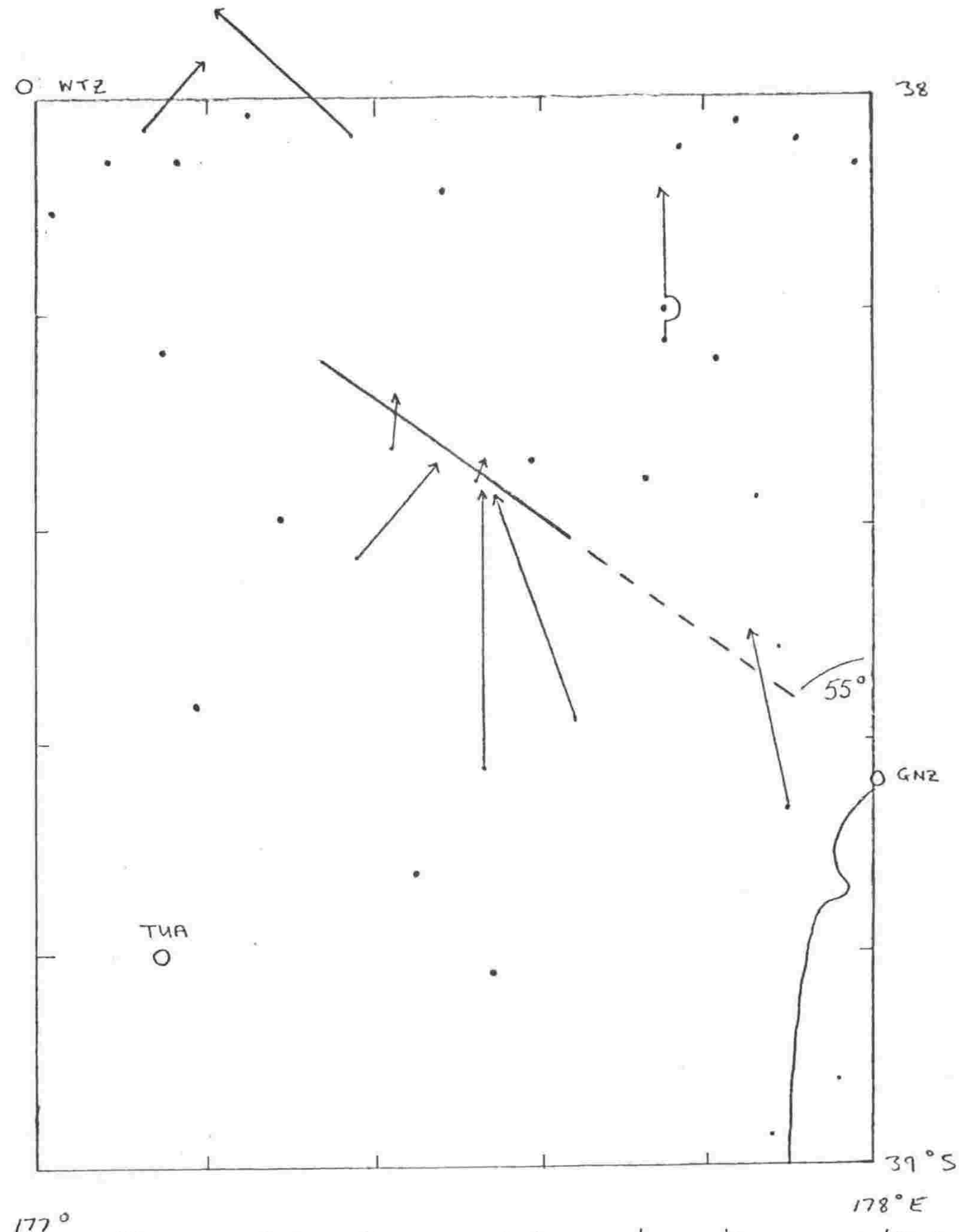


Figure 5.13 Subcrustal earthquake epicentres (Seismological Observatory) and JHD relocations (arrowheads).

with original depths 66-90 km) which places these events in the underlying plate. One might tentatively interpret these events as lying on a fault in the top of the underlying plate which strikes roughly parallel to the down-dip of the plate. Unfortunately the most recent study of mechanisms of North Island subcrustal events (Harris (1975)) lists only two earthquakes in this region and classifies them as uncertain - in neither case is there enough data to determine a focal mechanism, nor do they fit Harris's Categories A and B.

Against this interpretation is the failure to detect any similar lineations anywhere else among the relocated hypocentres. On the other hand, the faulting may not persist with descent of slab, or myriad cracking may make identification of faults impossible. There were not very many hypocentres relocated in the depth range 33-60 km, largely because of the comparative poorness of the data from the shallower events - there tend to be many more emergent arrivals and many more missing stations for such events. A search explicitly for this phenomenon might thus reveal other examples of it.

The final claim made for JHD is the obtaining of better information about model errors. One can see the improvement if one compares the model error estimates for stations with the average residual at the station over a group using the Observatory solutions, and compares both with what one would expect on the basis of a model with high velocities in the down-going slab and low or normal velocities in the surrounding asthenosphere. The model error estimate for KRP is large and positive for group 24. For these events the ray paths must lie almost wholly in the assumed low velocity asthenosphere and normal velocity lithosphere of the Indian Plate. A positive term relative to the stations looking down the slab is expected. The average S residual for KRP before JHD relocation is -1.3 sec, which is in complete disagreement with the model and would suggest in fact a totally different model. The lesson of Chapter II is of course that the mean residual depends too much on the station set to be useful as a source of model improvement. In the next chapter we show that with JHD we have loosened the bond of station set dependence somewhat.

CHAPTER VI

INTERPRETATION OF THE MODEL ERROR ESTIMATES FROM JHD

In this chapter, a travel-time model for part of the North Island, New Zealand is constructed which gives the best fit for the arrival time data from 256 earthquakes in nine groups, subject to certain constraints described below. The positions of the groups are given in Figure 6.1. Group 2367 is a good quality sample from the events placed by the Observatory below 250 km (excluding 600 km events which are to the south west of the area considered (Adams, 1963 and Adams and Ferris, 1976). The groups 23/22 and 22/32 are composites formed to provide adequate data for JHD. Other data on the groups is given in Tables 6.1(a) and (b).

From these groups, a picture of the Benioff Zone in this area is constructed (Figures 6.2 - 6.4) which is somewhat different from that of Adams and Ware (1977) and Hamilton and Gale (1968). Also, mantle and crustal travel-time models for part of the North Island are constructed.

METHOD OF INTERPRETATION

Up until this point, the number of assumptions about the physical situation that we have made has been a minimum: we have assumed that the travel-time model used (Jeffreys-Bullen) was accurate to within about ten percent (Chapter III), which implicitly denies discontinuities or very rapid changes in velocity of more than this amount.

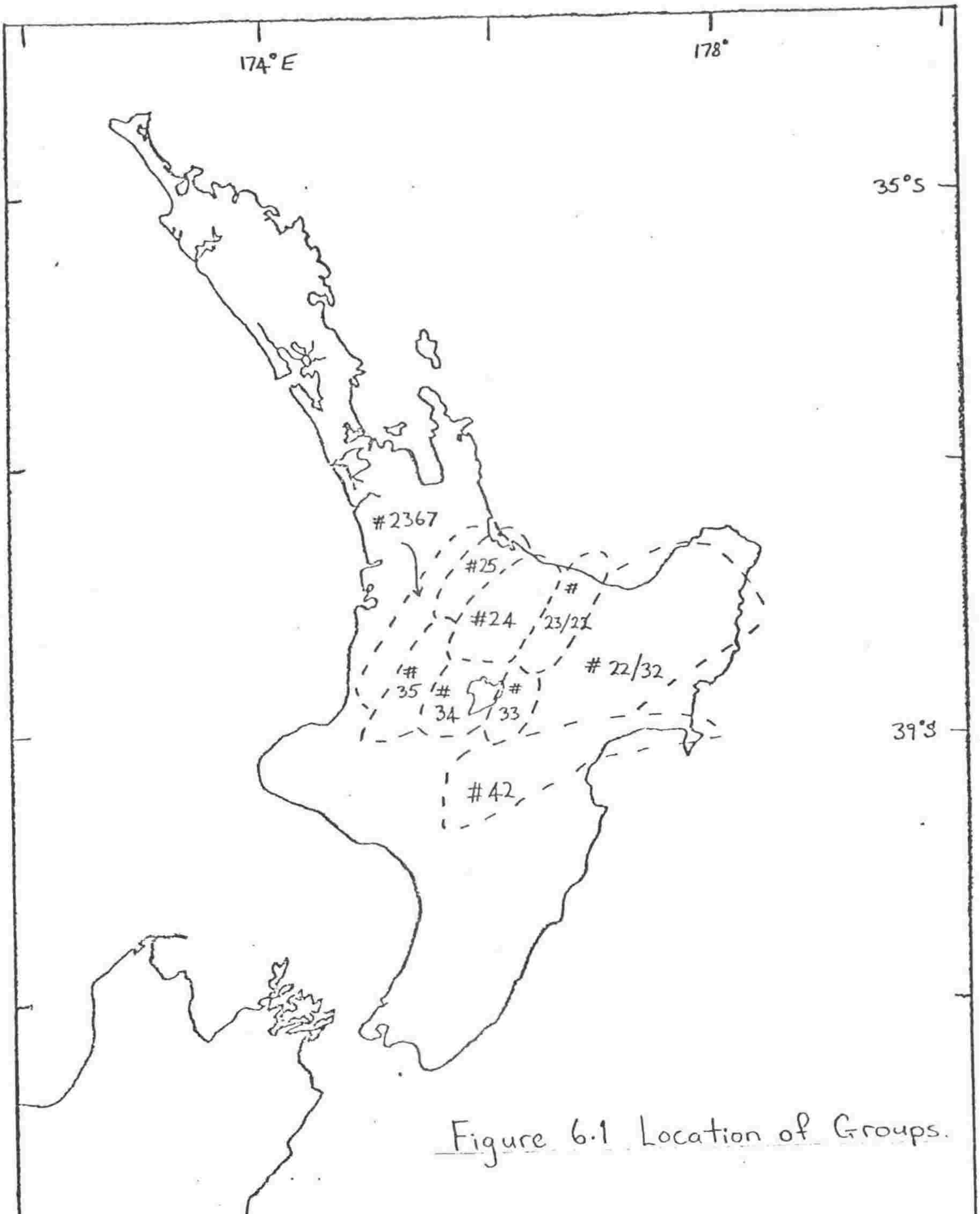


Figure 6.1 Location of Groups.

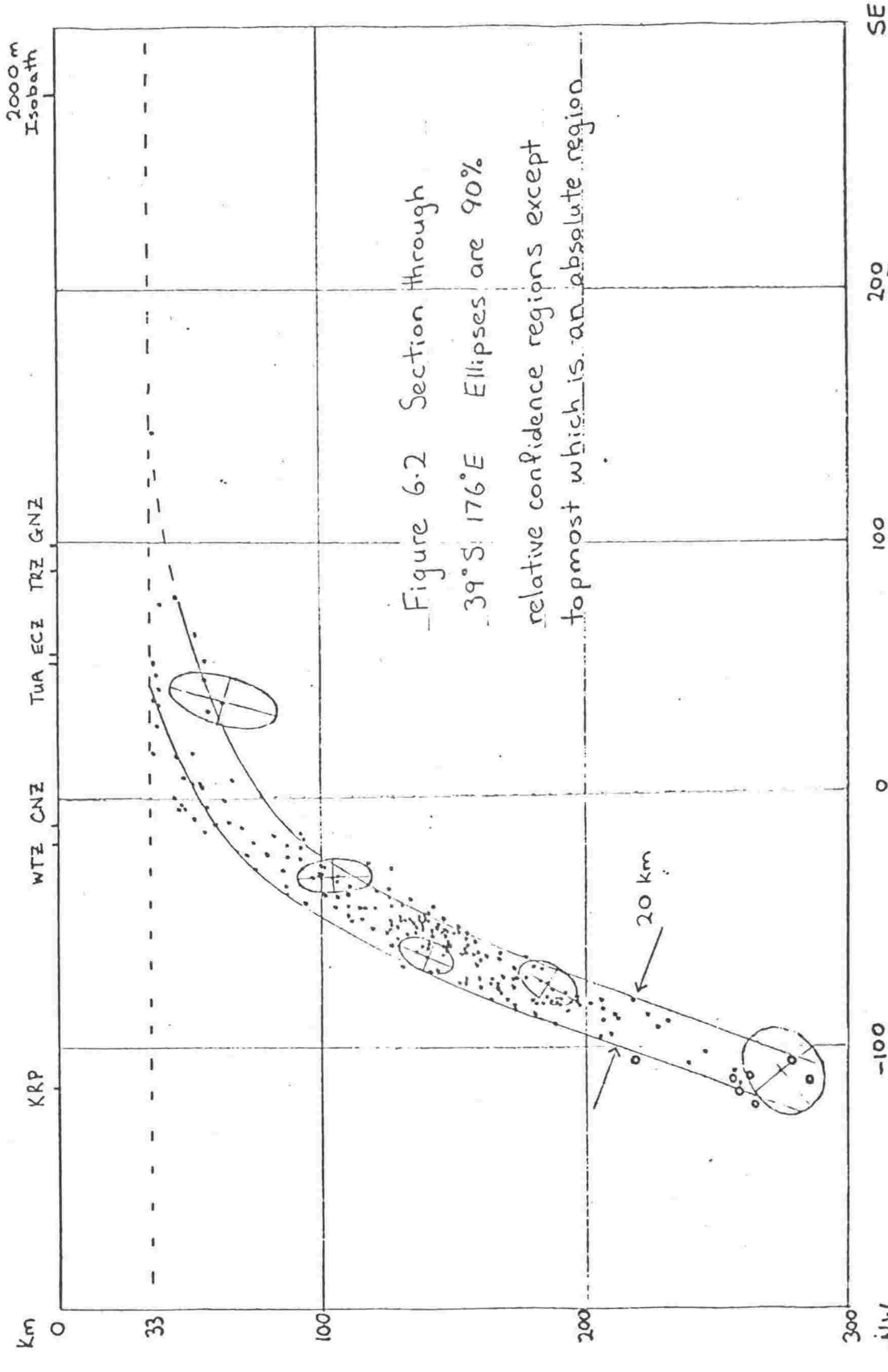


Figure 6.2 Section through 39°S 176°E Ellipses are 90% relative confidence regions except topmost which is an absolute region

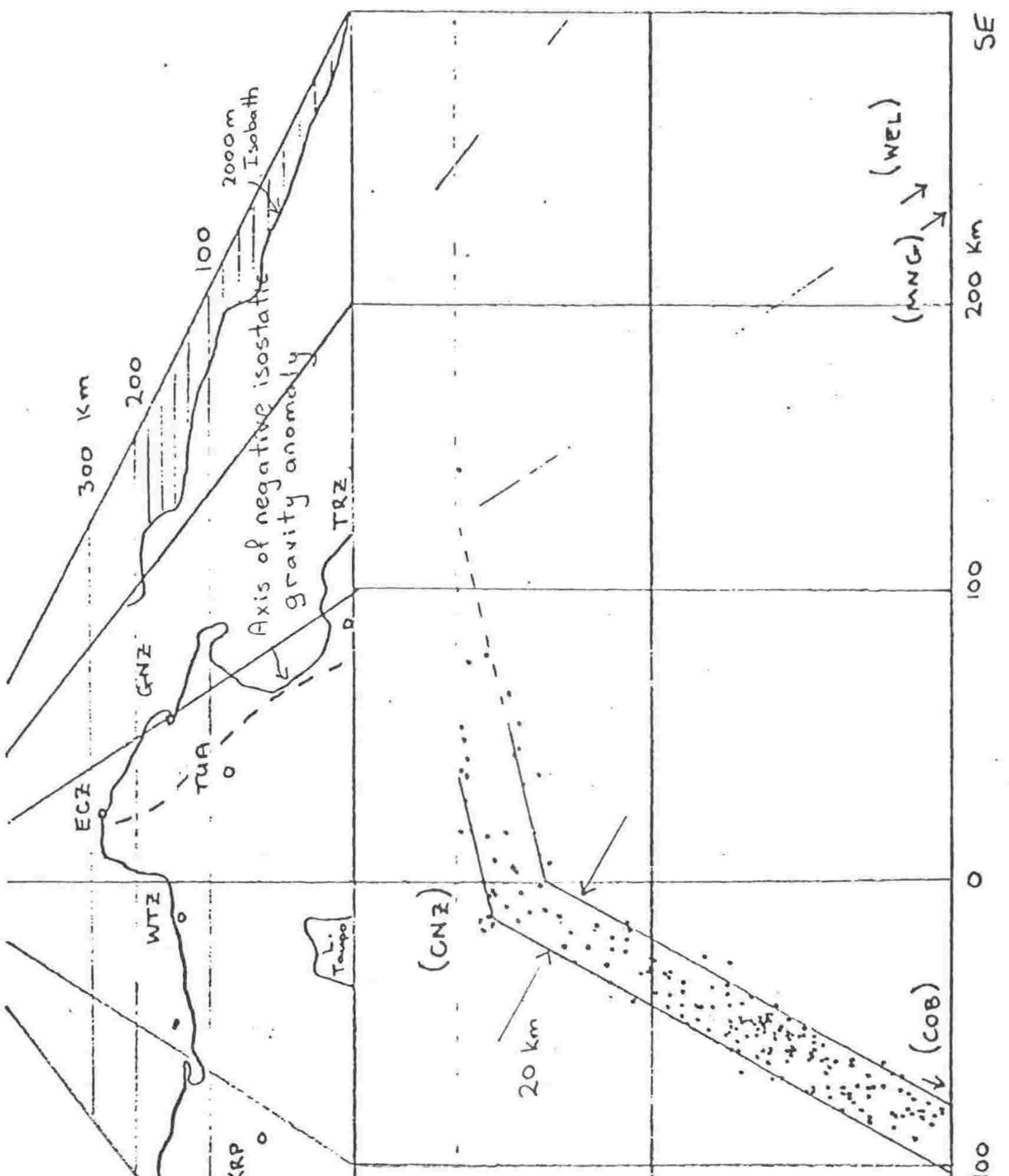


FIGURE 6.3. SECTION THROUGH 39°S 176°E

Figure 6.4

An alternative to Figure 6.3

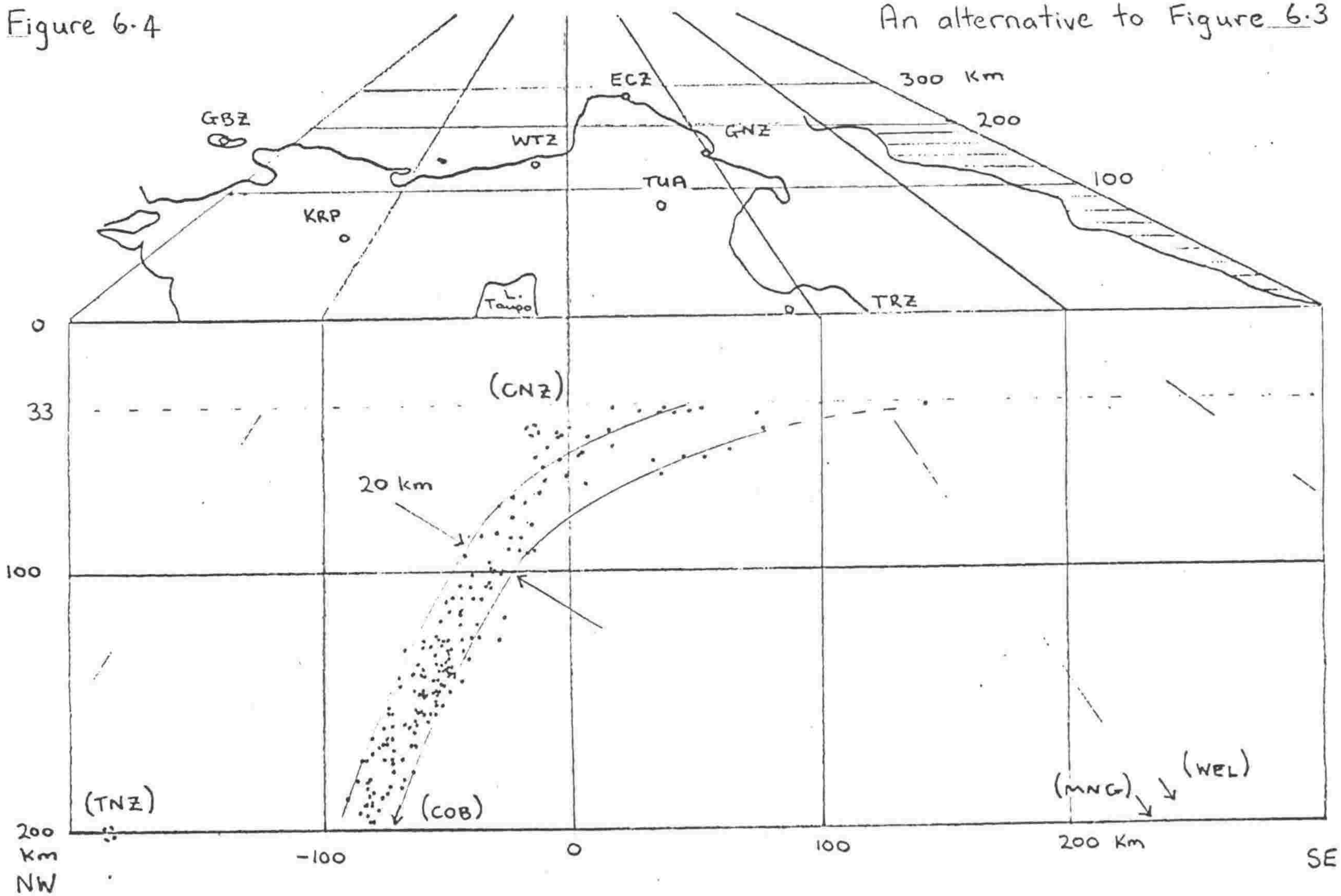


TABLE 6.1(a)

Group	No. of Events	No. of Stations	Master Used	$\hat{\sigma}$ (Total *P* error estimate)	No of Iterations
22/32	31*	15	No	0.371	10
23/22	36*	16	Yes	0.422	10
42	23	17	No	0.426	10
33	21	17	Yes	0.310	10
24	65	21	Yes	0.326	7
34	28	20	Yes	0.344	11
25	31	19	Yes	0.313	11
35	23	20	Yes	0.311	10
2367	19	16	Yes	0.389	10

* 15 earthquakes in common. Group 23 was too sparse and poorly conditioned to process without the additional shocks.

TABLE 6.1(b)

Group	Group(s) contributing to Master Event	MASTER EVENT (FREE - FIXED)					
		Lat. (km)	Std. Dev.	Long. (km)	Std. Dev.	Depth (km)	Std. Dev.
22/32	-	NA		NA		NA	
22/23	22/32	-	-	-	-	-	-
24	22/23	4	2	4	2	3	4
25	24	-5	3	0	3	14	5
2367	25	-	-	-	-	-	-
33	34,42	-	-	-	-	-	-
34	24	-2	3	5	2	19	5
35	25	-	-	-	-	-	-
42	-	-4	3	4	2	5	6*

* This row is the difference between the solutions of the event common to 42 and 33.

The preliminary indication from Chapter V is that the velocity contrast between ray paths entirely within the down-going slab and entirely outside (as the paths to KRP must be for most groups) is greater than ten percent. In order to validate the assumption deemed, in Chapter III, necessary for application of JHD, a simple modification of the travel-time model was made, whereby the mantle J-B travel-time would be increased or decreased according to the hypothesized percentage of the path in the slab. With the limitations of computer speed and instability of the linear system resulting if the average velocity to each station were made a determinable parameter, the decision was made that the velocity contrast, relative to J-B, for each station would be chosen subjectively and arbitrarily except that certain assumptions and constraints should be satisfied. It will be shown that within the limitations of these assumptions and constraints little latitude is allowed in the velocity contrasts for the majority of the stations. However, alternative assumptions are certainly possible and an attempt was made to examine the effect of one of the more fundamental assumptions: that the Jeffreys-Bullen Travel Time model provides satisfactory ray shapes for the region of interest, in particular, that rays lying entirely inside or entirely outside a slab in which velocities are higher than in the surrounding mantle are approximately the same shape as J-B rays. In the absence of a fast ray tracing algorithm (and processing group 24 with 1125 equations, iterating six times, would require 6750 ray calculations) there is no present alternative except the assumption of some standard ray structure. The Jeffreys-Bullen Model may not have been the most appropriate in this case but it is used by the Seismological Observatory for its routine locations and has been used in subcrustal seismicity studies in the past (Hamilton and Gale (1968), Mooney (1970a, b), Adams and Ware (1977) who apply a scale factor of 0.91 for mantle travel times for rays adjudged to lie entirely within the slab). An attempt to test the importance of the ray structure was made by substituting a homogeneous (constant velocity) mantle for the J-B upper mantle and the results of this are discussed in a later section.

With the assumption of the appropriateness of the Jeffreys-Bullen structure, the next step was to decide upon a scaling factor for the sub-crustal part of the travel time for each station and each group. To limit the multitude of possibilities and to keep the consequent model as simple as possible, the following constraints were adopted. First, stations with similar paths to the same group should have the same scale factor and a station having similar paths to different groups should have the same scale factor for those groups. The decision as to what constituted similar paths was arbitrary and subjective, but basically all stations with paths entirely in the slab had a common scale factor assigned. *P* and *S* were treated independently however, and the final scale factors for *P* and *S* paths in the slab were not the same.

Table 6.2 shows the final scale factors for all stations and groups expressed as a percentage contrast to the Jeffreys-Bullen model. To apply a scale factor α one subtracts from the J-B travel time T the crustal portion, T_c. This was carefully done by calculating the incidence angle at the base

TABLE 6.2

STATION	GROUP								
	22/32	42	22/23	33	24	34	25	35	2367
ECZ P	-9		-9	-9	-9	-9	-9	-9	-9
GNZ P	-9	-9	-9	-9	-9	-9	-9	-9	-9
TUA P	-9	-9	-9	-9	-9	-9	-9	-9	-9
TRZ P	-9	-9	-9	-9	-9	-9	-9	-9	
MNG P	-9	-9	-9	-9	-9	-9	-9	-9	-9
WEL P		-9	-9	-9	-9	-9	-9	-9	-9
WTZ P	-9	-9	-9	-9	-9	-9	-9	-9	
CNZ P	-9	-9	-9	-9	-9	-9	-9	-9	-9
COB P			-4	-8	-8	-8		-8	-8
TNZ P	-4	-4	-4	-4	-4	-4	-4	-4	-4
KRP P	-4	-4	-4	-4	0	0	-4	-4	-4
GBZ P	0				0				
ECZ S					-6		-6		
GNZ S	-6	-6	-6	-6	-6	-6	-6	-6	-6
TUA S	-6	-6	-6	-6	-6	-6	-6	-6	-6
TRZ S		-6		-6	-6	-6	-6	-6	
MNG S		-6	-6	-6	-6	-6	-6	-6	-6
WEL S	-6	-6	-6	-6	-6	-6	-6	-6	-6
WTZ S	-6	-6			-6	-6		-6	
CNZ S						-6	-6	-6	-6
COB S					-4	-4	-4	-4	-4
TNZ S		0		-4					
KRP S	0	0	0	0	6	6	0	0	0

of the crust (33 km) using:

$$\frac{\partial T}{\partial \Delta} = \frac{r \sin i}{v} \qquad \ldots(6.1)$$

and assuming a one-layer crust of P velocity 6.1 km/sec., S velocity 3.5 km/sec., thickness 33 km. The resulting travel time is thus:

$$T_c + (1+\alpha/100)(T-T_c) \qquad \ldots(6.2)$$

The station COB and possibly TNZ had mixed paths - paths being partly inside and partly outside the slab. The scale factors for these stations were determined by trial for each group with some attempt to keep as much uniformity as possible, and also to satisfy the other constraints.

The second constraint was that the scale factors should be chosen so that the internal consistency of the hypocentre estimates was as great as possible. This was easily tested for by comparing the positions of the seeded event which a pair of groups had in common. If the hypocentre estimates for these events from each group were close enough together, the model could be judged to be giving internally consistent results for the pair of groups. In fact, with the finally adopted model, differences of only 2 to 3 km in the hypocentre estimates of common events were normal. (See Table 6.1b) The third constraint was that the model error estimates obtained from JHD for this new model should be sufficiently small that the linearisation approximation could be made. In fact the aim was to reduce these times to values which could be accounted for by assuming crustal model variation.

Fourth was the underlying principle that all changes should be as small as possible, that is, that consistent with constraints two and three, the scale factors should be as small as possible.

Within the limited set of satisfactory models left, the programming required to find the best solution, although feasible in principle, would surely be impractical, first because of the limitations of computer speed available, and second because the resulting (linear) system would be very likely to be too poorly conditioned to be meaningfully solved; at best one would expect to have to start the system from an approximate solution quite close to the final solution.

A discussion of some of the recent attempts to extend the model improving powers of JHD by different methods is included in the next chapter. To mention one such method, Aki and Lee (1976) construct a least squares system in which the model error estimates as well as the residuals are minimised in the least squares sense. In view of the third constraint of requiring small model correction, such a scheme could be applied to our problem provided good prior crustal models could be used to calculate initial station terms. Haines (1977) provides a set of station terms for stations of the New Zealand Seismograph Network. From arrival time data, Haines calculates regional P_n and S_n velocities and crustal delay times for P_n and S_n relative to a standard crustal model (5.5, 3.3 km/sec. for P, S above 12 km; 6.5, 3.7 km/sec. between 12 and 33 km). An attempt was made to incorporate Haines' station terms in the model and minimise the resulting error estimates by inspection, but the existence of one or two small discrepancies between Haines' terms and our terms led to the abandonment of this idea. It was decided not to apply prior station terms. The final set of station terms (approximately crustal model corrections) reduced to critically incident ray delay times for the same standard crustal model (to make them directly comparable with Haines' results) are shown in Table 6.3 and will be discussed hereunder.

The method of processing the data was thus reduced to a two-step process with considerable subjective interaction in each step. A subjective decision based on the constraints was made for each group on the quality of the fit of the data from the group using a travel time model selected by trial. Having decided on the best model for each group (Table 6.2 shows that there was minimal variation in the model between groups) the model error estimates from the groups, together with other pertinent information (average model travel time, incident angle, etc.) were pooled and a linear model constructed in which the station terms, average model error for each group, and velocity adjustments for similar ray paths were parameters to be determined. The problem here was in deciding how many different velocity adjustments should be made and what stations could be assumed to have a similar ray path to a given group. However, one can make a quantitative comparison between the sum of squares of residuals for each variation of velocity parameters and decide which combinations give significantly smaller values for the sum of squares.

DERIVATION OF THE LINEAR MODEL FOR DETERMINING STATION TERMS, GROUP MEANS AND VELOCITY ADJUSTMENTS

Let C_j be the time to be added to the model travel time T_0 to give the true average travel time in the crust for rays from station j critically incident at the base of the crust. Let $\hat{C}_{ij}$ be the calculated model error estimate for station j from group i. Let H_i be the average model error estimate for group i: thus the true model error estimate is:

$$C_{ij} = H_i + \hat{C}_{ij} \qquad \ldots(6.3)$$

Let T_{ij} be the average model travel time from the events of group i to station j and let the part of T_{ij} that is the crustal travel time be $f_{ij}T_0$. (T_0 is the critically refracted model crustal travel time.) Let the mantle part of the travel time be in error by α_{ij} percent. Thus the true mantle travel time is:

$$(1+\alpha_{ij}/100)(T_{ij}-f_{ij}T_0)$$

The true crustal travel time is given approximately by:

$$f_{ij}(T_0 + C_j)$$

Thus the true travel time is:

$$(1+\alpha_{ij}/100)(T_{ij}-f_{ij}T_0) + f_{ij}(T_0+C_j)$$

which we assume differs from the calculated travel time:

$$T_{ij} + C_{ij} = T_{ij} + H_i + \hat{C}_{ij} \qquad \ldots(6.4)$$

by a random amount Σ_{ij} which is normally distributed with zero mean and variance ρ^2.

Thus we have the linear model:

$$T_{ij} + H_i + \hat{C}_{ij} = (1+\alpha_{ij}/100)(T_{ij}-f_{ij}T_0) + f_{ij}(T_0+C_j) + \Sigma_{ij} \quad \ldots(6.5)$$

That is:

$$\hat{C}_{ij} = -H_i + f_{ij}C_j + \alpha_{ij}(T_{ij}-f_{ij}T_0)/100 + \Sigma_{ij} \quad \ldots(6.6)$$

With N groups and M_i stations being used for group i, this would furnish $\sum_{i=1}^{N} M_i$ equations in $N + M + \sum_{i=1}^{N} M_i$ unknowns if each station/group combination were allowed a separate velocity adjustment. Instead (by inspection) we group the station/group combinations into classes with a common velocity adjustment. If there are K classes then we replace *(6.6)* with:

$$\hat{C}_{ij} = -H_i + f_{ij}C_j + \alpha_{ij}(k)(T_{ij}-f_{ij}T_0)/100 + \Sigma_{ij} \quad \ldots(6.6a)$$

and solve the equation by least squares for the N terms H_i, the M terms C_j and the K terms $\alpha_{ij}(k)$.

DISCUSSION OF THE RESULTS: FINDING A PRELIMINARY MODEL

The process of reducing the data in the groups in accordance with the constraints can be summarised as follows: find the simplest modification of the J-B mantle model (by scaling mantle travel times) which reduces the model error estimates to a satisfactory small level and produces consistent locations for earthquake(s) common to pairs of groups. It would be wrong to pretend that the solution arrived at in this way is not the result of subjective decisions and that other models might not satisfy the constraints to the same extent. We hereunder present some of the justification in terms of hypocentre consistency for adopting the model summarised in Tables 6.2, 6.3.

A further justification is presented in the next section where the model error estimates (now interpreted as crustal model error estimates) are compared with Haines' values.

The starting point was the model in which *P* and *S* mantle travel times for stations with rays entirely in the slab were decreased by ten percent as suggested by the results of Chapter V, and the scale factors for other stations (in particular KRP, TNZ, GBZ) were determined by trial. From the results from a few groups it became apparent that ten percent was too high and that the scale factor for *S* should be somewhat less than that for *P*. It was intended that once a satisfactory model had been determined, one group would be a master group and processed without a master event, and that the other groups would use as a master event: the event in common with the master group restricted to its master group position (or, if the group were not adjacent to the master group, a master event from an already processed group). The original intent was to use group 24 as the master group. However, it was finally decided to use one of the groups with the most shallow events, group 22/32, the decision being made because the travel times for the shallow events will be less affected by possible deficiencies in the structure of the travel time model and also because the results appeared more consistent that way. Table 6.1(b) shows the source of the master event seeded into each group, and the difference between the hypocentre estimates for this event when the group was run without a fixed event and when the master event was fixed at the position obtained in the seeding group. Not all groups were run free for comparison in this way. The groups omitted from the comparison tended to converge more slowly, usually because of sparseness of data. Where the comparison is made it is apparent that the consistency is quite good - in no case is the latitude or longitude more than 5 km different. The variation in depth is greater, the free depths tending to be greater than the master event controlled depths. However, for group 24, the most stable of the groups, the difference is only 3 km. Moreover, the standard deviations given in the table are the standard deviations for relative error (being obtained from the fixed event solution). Typical absolute error standard deviation estimates for depth for groups 25 and 34 are 12 km and 8 km. Thus 90 percent absolute error confidence regions for the pair of solutions in the cases of groups 25 and 34 would substantially overlap. However, because of this greater discrepancy in the case of group 34, a different test was made in the case of group 42. No master event was seeded into group 42. Instead, group 42 contributed an event to group 33 and the difference between the hypocentre estimates for this event is given in the table (group 33 being processed with a master event from group 34). The hypocentre estimates are within 7 km which we consider to be quite satisfactory.

At this point one should note that the differences between the origin time estimates, hitherto unmentioned, are not easily interpreted because of the unknown average model error for each group. When using a master event, only the latitude, longitude, and depth and not the origin time were fixed. It was hoped that this would allow the average model errors for each group to be estimated later. As it turned out, the linear system *(6.6a)*, although not singular, is too badly conditioned to be sensibly solved without one constraint.

The conclusion from Table 6.1(b) is that the chosen model produces hypocentre estimates with a sufficiently high degree of internal consistency that one can confidently combine the model error estimates from the groups for further analysis.

At this stage, there was one reservation about the results so far obtained. The stations CNZ and WTZ lie on a line trending roughly south-west north-east, that is, parallel to the strike of the Benioff zone, and rather to the north-west of a simple projection of the steeper part of the zone to the surface. (Figure 6.3)

One might thus expect that, first, the two stations would have similar ray paths to groups below about 100 km and second, that the velocity contrast might be somewhat less than that for stations to the south-east. Table 6.2 shows that the stations received identical treatment for all groups. However, a comparison of the (crustal) model error estimates for a sample of stations from the group 24 analysis and Haines' results for those stations shows a particularly large disagreement in the case of WTZ.

Station	Group 24	estimate std. dev.	Haines estimate	Haines std. dev.
KRP P	-0.12	0.28	-0.4	0.3
S	0.00	0.56	0.5	0.6
TUA P	0.47	0.11	0.9	0.25
S	1.84	0.35	2.3	0.4
GNZ P	0.20	0.17	0.5	0.25
S	0.25	0.29	1.2	0.3
WTZ P	-0.59	0.18	0.5	0.25
S	1.06	0.37	2.2	0.5
CNZ P	1.89	0.10	1.1	0.3

(The group 24 results were calculated, as are Haines', relative to MNG P. Variation in angle of incidence at the crust/mantle boundary has been allowed for.) The group 24 estimates for TUA, GNZ and WTZ are all smaller than Haines' estimates; in the cases of TUA and GNZ about half a second smaller but one second for both P and S for WTZ.

At the same time, the CNZ P estimate is greater than Haines' figure. This suggests that the mantle travel time to TUA, GNZ and WTZ, in particular, should be decreased to force an increase in the crustal term. However, to depart from the otherwise very uniform scaling of mantle travel times without more substantial justification was questionable. Thus no change in the model was made. Some attempts to explain this phenomenon in terms of model deficiencies other than the scale factor are discussed below. It should be pointed out that the disagreement is somewhat worse when Haines' figures and the results obtained from group 24 without a fixed event are compared. This might be explained in terms of a systematic error in the hypocentre estimates arising from model errors which is reduced but not removed by the fixing of the master event. In the balance, the total of evidence would seem to indicate that this error is not very large and results will be presented to show that the effect might be attributed to the inappropriateness of using a J-B mantle model. At the same time, that as good a fit as is here obtained using scale factors substantially less in magnitude than -9 percent for *P* and -6 percent for *S* must be regarded as being highly unlikely.

DISCUSSION OF THE RESULTS: LINEAR MODELLING OF THE MODEL ERRORS

As explained above, the model error estimates from all the groups ($\hat{C}_{ij}$ of *(6.3)*) were used to estimate average crustal and mantle model errors. Altogether 160 estimates $\hat{C}_{ij}$ from 9 groups and 22 stations were combined. The only problem in the analysis was to decide which station/group combinations should have a common travel-time scale factor α. In accordance with the assumptions, the number of different classes should be as small as possible. An absolute minimum number would be two: one each for the *P* and *S* travel-times for the bulk of the groups to those stations for which hypothesized ray paths would be entirely in the slab. These estimates and their standard deviation estimates would be one test of the internal consistency of the process used to derive this final model.

By trial it was found that if scale factors for three other classes were determined, the fit of the whole model was dramatically improved. These three classes were: *P* and *S* from the shallowest groups to the stations WTZ, TUA, GNZ and *S* from the deepest groups to KRP. This latter was the only exception to a self-imposed rule to avoid dealing with station/group combinations where only a single station was involved. The exception was made in this case because of the obvious and highly significant improvement obtained.

When the number of scale factor classes was increased from two to five in this manner, the standard deviation estimate of the residuals of the $\hat{C}_{ij}$ was reduced from 0.31 sec. (128 d.f.) to 0.21 sec. (125 d.f.). This value is the estimate of the average uncertainty of a *P* travel-time calculated from the proposed model. In so far as this agrees well with the value predicted in Chapter V for this quantity, and the standard deviation estimates of $\hat{C}_{ij}$ for *P*'s from the JHD processing are typically about 0.2 sec., one can argue that one has reached a practical limit in extracting information from the data by the process described above. The distribution of stations in the classes is given in Table 6.5.

We now give a few more details about the calculation of the final model and summarize the results. The first problem concerned the statistical aspects of *(6.6a)*. It is not the case that the errors Σ_{ij} are independent with common variance. Indeed, the output data from each group provides, from standard linear theory, an estimate of the variance matrix of the station term estimates $\hat{C}_{ij}$. (The variance matrix for group 24 can be determined immediately from correlation matrix for the station terms given in full in Appendix VI and the standard deviation estimates.) However, it would be implausible to assume that there is no correlation in the errors between groups for the same station. The correlations within group 24 are seldom large; the principal systematic features are a small positive correlation between the stations, both *P* and *S*, with all slab paths and a rather larger negative correlation between these stations and KRP. These relationships are far from surprising. They suggest similar correlations between the station term estimates for adjacent groups. In the absence of such complete variance information and since there were few large correlations within groups and the standard deviation estimates for the $\hat{C}_{ij}$ were, within a factor of about two, 0.2 sec. for *P* terms and 0.4 sec. for *S*, the approximation was made that the $\hat{C}_{ij}$ were uncorrelated and that the standard deviation of the *S* terms was twice that of the *P*, implying half-weighting of the *S* equations. As a test, in one model the reciprocals of the estimated standard deviations of the $\hat{C}_{ij}$ were used as weights and the solution in this case was very little different from that obtained by the simpler model.

The second problem was one of conditioning. In practice, it was found to be impossible to separate the group means from the station terms without fixing the value of one of the station terms. This is in fact of comparatively small significance but illustrates once again the known difficulty of determining the average error of any model from arrival-time data alone.

Here, without the constraint of a fixed station, the lack of sufficient variation in the fraction f_{ij} meant that the columns associated with H_i and C_j of the design matrix for *(6.6a)* were close to being linearly dependent in the usual numerical sense. The lack of conditioning reflected itself in very large

TABLE 6.3

STATION TERMS NORMALIZED AS CRITICAL REFRACTION DELAY TIMES;* MNG P = 0

Station	Station Term (sec.)	Std. Dev.	Haines (1977) Station Term	Smith-Haines
KRP P	0.2	0.1	-0.4	-0.6
KRP S	0.1	0.2	0.5	-0.4
WTZ P	-0.1	0.2	0.5	-0.6
WTZ S	1.7	0.5	2.2	-0.5
CNZ P	2.1	0.1	1.1	1.0
CNZ S	6.4	0.5	2.4	4.0
TUA P	0.6	0.1	0.9	-0.3
TUA S	2.3	0.3	2.3	0.0
GNZ P	0.3	0.2	0.5	-0.2
GNZ S	1.0	0.3	1.2	-0.2
ECZ P	0.9	0.2	0.4	0.5
ECZ S	3.6	0.9	1.6	2.0
TRZ P	1.4	0.1	1.0	0.4
TRZ S	3.9	0.4	3.0	0.9
COB P	-0.5	0.3	-0.4	-0.1
COB S	-0.3	0.2	0.9	-0.4
MNG S	0.5	0.3	0.0	0.5
WEL P	-0.1	0.3	0.0	-0.1
WEL S	0.8	0.4	-0.5	1.7
TNZ P	1.3	0.1	0.9	0.4
GBZ P	-1.7	0.2	-1.3	-0.4

* to obtain vertical path station terms, take .66 × (listed value).

TABLE 6.4

GROUP MEANS

Group	Mean travel time error (sec.)	Standard Deviation
24	0.4	0.03
34	0.6	0.04
22/32	0.7	0.05
22/23	0.5	0.05
25	0.7	0.04
42	1.1	0.06
33	0.5	0.04
35	0.5	0.04
2367	0.5	0.04

TABLE 6.5

CLASSES OF STATION/GROUPS WITH COMMON SCALE FACTOR

Station	Groups								
	24	34	22/32	22/23	25	42	33	35	2367
KRP P	0	0	0	0	0	0	0	0	0
KRP S	0	0	0	0	5	0	0	5	5
WTZ P	1	1	3	1	1	3	1	1	1
WTZ S	2	2	4	2	2	4	2	2	2
CNZ P	1	1	1	1	1	1	1	1	1
CNZ S	0	0	0	0	0	0	0	0	0
TUA P	1	1	3	1	1	3	1	1	1
TUA S	2	2	4	2	2	4	2	2	2
GNZ P	1	1	3	1	1	3	1	1	1
GNZ S	2	2	4	2	2	4	2	2	2
ECZ P	1	1	1	1	1	1	1	1	1
ECZ S	2	2	2	2	2	2	2	2	2
TRZ P	1	1	1	1	1	1	1	1	1
TRZ S	2	2	2	2	2	2	2	2	2
COB P	1	1	1	1	1	1	1	1	1
COB S	0	0	0	0	2	0	0	2	2
MNG S	2	2	2	2	2	2	2	2	2
WEL P	1	1	1	1	1	1	1	1	1
WEL S	2	2	2	2	2	2	2	2	2
TNZ P	0	0	0	0	0	0	0	0	0
GBZ P	0	0	0	0	0	0	0	0	0

CLASS	VELOCITY CONTRAST (%)	STANDARD DEVIATION
1 (P)	-0.8	1.1
2 (S)	-1.2	0.4
3 (P)	4.6	7.2
4 (S)	3.4	7.7
5 (S)	4.9	2.3

TABLE 6.6

MODEL RESIDUALS (STD. DEVS. IN PARENTHESES)

Station	Group								
	42	32/22	23/22	33	24	34	25	35	2367
KRP P	0.12	0.27	0.00	-0.13	0.16	0.40	-0.41	-0.32	-0.26
	(0.53)	(0.22)	(0.33)	(0.22)	(0.19)	(0.23)	(0.24)	(0.24)	(0.30)
KRP S	0.22	0.76	-0.42	-0.96	0.04	0.22	-0.24	0.42	-0.20
	(1.12)	(0.46)	(0.54)	(0.39)	(0.38)	(0.41)	(0.41)	(0.39)	(0.54)
WTZ P	-0.16	-0.39	-0.04	-	0.24	0.26	0.14	-0.05	-
	(0.52)	(0.28)	(0.29)	-	(0.13)	(0.19)	(0.18)	(0.22)	-
WTZ S	0.68	-0.02	-	-	0.12	-0.30	-	-0.64	-
	(1.12)	(0.66)	-	-	(0.26)	(0.31)	-	(0.42)	-
CNZ P	0.00	0.26	0.09	-0.19	-0.03	-0.27	0.09	-0.16	0.10
	(0.33)	(0.21)	(0.13)	(0.15)	(0.07)	(0.10)	(0.11)	(0.10)	(0.14)
CNZ S	-	-	-	-	-	-0.16	0.14	0.12	-0.10
	-	-	-	-	-	(0.38)	(0.37)	(0.41)	(0.46)
TUA P	-0.13	-0.06	-0.02	0.15	-0.07	0.00	0.11	-0.05	0.10
	(0.44)	(0.17)	(0.15)	(0.13)	(0.08)	(0.15)	(0.12)	(0.15)	(0.19)
TUA S	0.12	-0.50	-0.42	0.26	0.02	0.08	0.50	0.14	-0.14
	(1.00)	(0.41)	(0.43)	(0.32)	(0.26)	(0.31)	(0.33)	(0.33)	(0.48)
GNZ P	-0.04	-0.18	0.08	0.17	-0.07	-0.12	-0.02	-0.24	-0.08
	(0.84)	(0.27)	(0.21)	(0.16)	(0.13)	(0.19)	(0.17)	(0.20)	(0.25)
GNZ S	-0.08	-1.04	-0.04	-0.72	0.18	0.52	0.12	0.02	-0.34
	(0.84)	(0.47)	(0.42)	(0.27)	(0.23)	(0.28)	(0.29)	(0.30)	(0.44)
ECZ P	-	0.10	-0.04	0.32	-0.15	-0.09	-0.06	-0.09	-0.01
	-	(0.40)	(0.32)	(0.21)	(0.20)	(0.27)	(0.26)	(0.27)	(0.33)
ECZ S	-	-	-	-	0.02	-	-0.02	-	-
	-	-	-	-	(0.27)	-	(0.33)	-	-
TRZ P	-0.10	-0.05	0.24	-0.13	-0.08	0.07	0.05	0.06	-
	(0.52)	(0.29)	(0.18)	(0.13)	(0.09)	(0.12)	(0.12)	(0.14)	-
TRZ S	-0.20	-	-	-0.06	-0.22	0.22	0.36	-0.06	-
	(0.95)	-	-	(0.33)	(0.28)	(0.31)	(0.36)	(0.40)	-
MNG P	0.04	0.37	0.15	-0.14	0.01	-0.27	0.10	-0.24	0.11
	(0.76)	(0.27)	(0.15)	(0.12)	(0.12)	(0.14)	(0.17)	(0.15)	(0.19)
MNG S	-0.12	-	-0.52	0.10	-0.06	0.30	0.08	-0.02	0.20
	(1.25)	-	(0.38)	(0.28)	(0.26)	(0.30)	(0.27)	(0.34)	(0.42)
WEL P	-0.22	-	-0.21	0.32	-0.12	0.09	-0.09	0.21	-0.08
	(0.81)	-	(0.20)	(0.15)	(0.24)	(0.19)	(0.23)	(0.21)	(0.25)
WEL S	0.40	0.26	0.02	0.02	-0.16	-0.22	-0.38	0.52	-0.46
	(1.32)	(0.50)	(0.35)	(0.27)	(0.09)	(0.30)	(0.26)	(0.31)	(0.39)
COB P	-	-	-	-0.13	-0.21	0.03	-	0.36	-0.02
	-	-	-	(0.19)	(0.19)	(0.25)	-	(0.32)	(0.34)
COB S	-	-	-0.20	-0.28	0.26	0.30	-0.76	0.72	0.02
	-	-	0.42	0.31	0.24	0.33	0.31	0.36	0.42
TNZ P	-0.19	0.00	0.06	-0.21	-0.06	-0.27	0.14	0.22	0.41
	(0.50)	(0.24)	(0.22)	(0.20)	(0.15)	0.20	0.21	0.22	0.26
GBZ P	-	-0.19	-	-	0.22	-	-	-	-
	-	(0.35)	-	-	(0.35)	-	-	-	-

standard deviation estimates for all parameters. Thus the station term for MNG P was arbitrarily set to zero. (Haines required that the terms for both MNG P and MNG S be zero in his analysis. Thus his results and those presented here are directly comparable.) Last, the additional classes of scale factors (described above) were determined by an inspection of the residuals of the simple model with only two scale factors. In the case of KRP S, large residuals, many times standard deviation came from the deepest groups 25, 35, 2367. Similarly, consistent residuals were obtained from WTZ, GNZ, TUA *P* and *S* for the shallow groups 32/22, 42. The improvement in the fit at the expense of three additional parameters is very highly significant. A complete set of results is given in Tables 6.3 - 6.5. We shall discuss each set of estimates.

The group means are of little interest in themselves since the overall mean is not estimated. However, the assumption of internal consistency would require that there be no very large difference between means for different groups - too large a variation would imply too large an average difference between the models for the groups. With the exception of group 42 (which was not tied by a master event to the rest), the consistency of the means is satisfactory compared to the overall standard deviation of 0.21 sec.

The travel time factor estimates show that on the whole the original model for the slab paths was correct to about one percent. The most important values are the -0.4 percent (standard deviation 1.2) for *P* times in the slab - an insignificant adjustment - and -1.2 percent (standard deviation 0.4) for *S* in the slab. This value is just significant and indicates that the *S* contrast estimate is -7 percent rather than -6 percent. Nonetheless, these two values are sufficiently small that we may adjudge that our modelling process has entirely satisfied our requirements of internal consistency.

The additional factors: +4.9 percent (2.2) for KRP S for deep groups indicates a significantly slower travel time than that assumed. It can be compared to the original factor of +6 percent for groups 24 and 34 (Table 6.2). However, the absolute isolation of the KRP ray paths from all others, making comparison with other stations impossible, means that any KRP results must be interpreted with caution.

The values +5.0 percent (7.4) for *P* and +3.5 percent (7.6) for *S* for shallower paths to WTZ, TUA, GNZ seem to indicate travel times slower than those assumed with the difficulty that the standard deviations of the calculated values are too large to give these values significant worth. Were it not for the large improvement in the overall fit obtained by including these classes, one would be tempted to say that the factors are not significantly different from zero. However, because of the uncertainty in choosing the right elements for these classes - it may be that some station is inappropriately placed in the wrong class - it is probably correct to deduce that the travel times for this class might be a few percent slower than the model but that the data cannot resolve this difference. This would be in agreement with Haines' velocities for this region which are a few percent slower than in the region south of TRZ.

In order to test the stability of this final model and also to see whether the overall standard deviation could be reduced below 0.2 sec., the combined group data was reprocessed with two additional velocity classes: *P* and *S* for stations WTZ, GNZ, TUA, ECZ to the groups 24, 25, 2367. These stations and groups were chosen because the ray paths from these groups to the stations would be practically straight up the slab with little lateral displacement across the slab. Significant travel time factors might be evidence for lateral velocity variations within the slab; the problem with such an interpretation being the unknown dependence on the assumed ray structure.

However, the result of the experiment was negative in the sense that neither factor was significant, the values being -0.2 percent (1.5) for *P* and -0.8 percent (1.2) for *S*, while the total fit was not significantly altered. At the same time the other parameter estimates hardly varied. It can be concluded that with a station term fixed, the linear system is stable to small perturbations in the station/group classes.

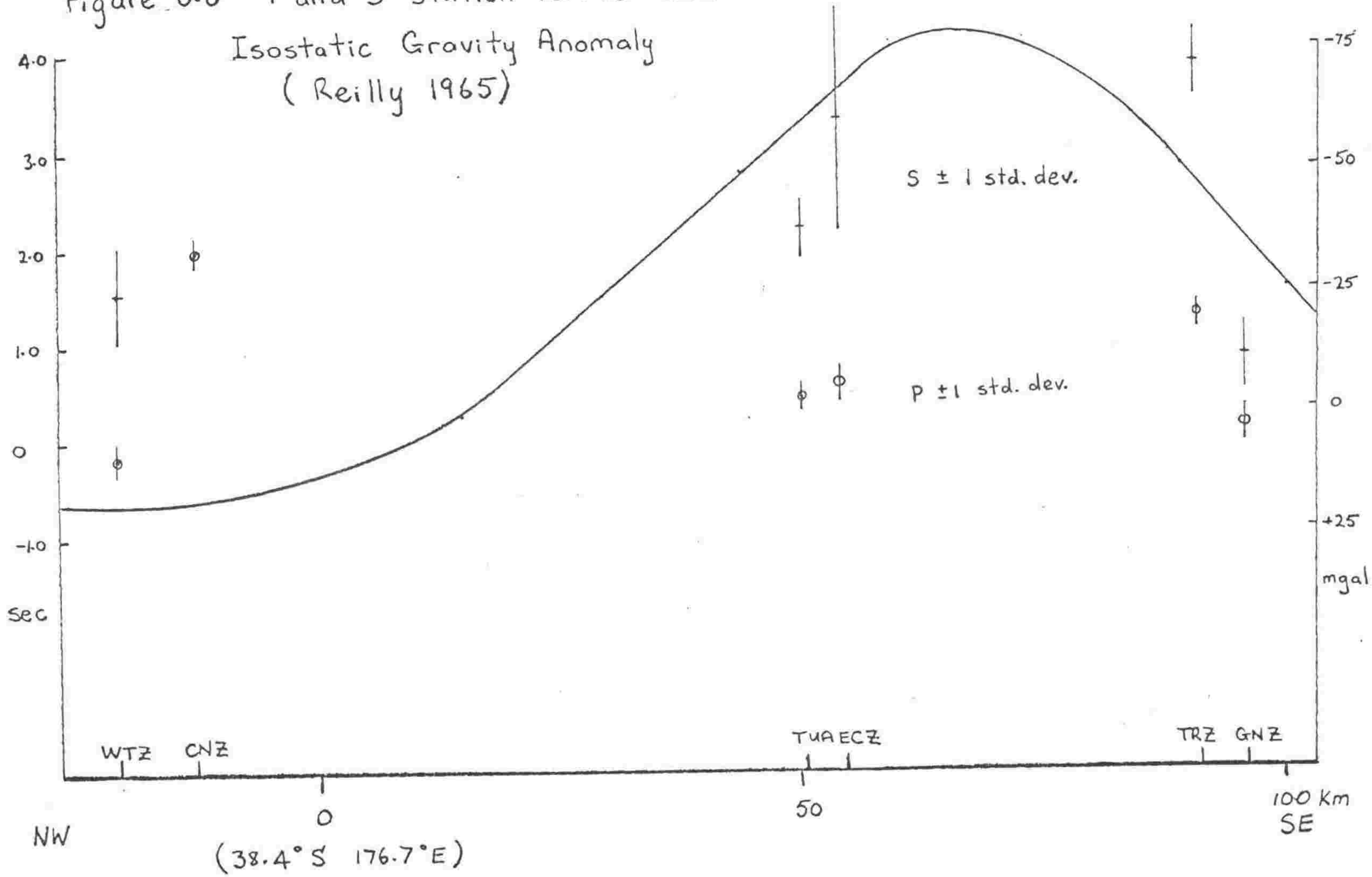
Table 6.3 lists the station terms and Haines' estimates for the same. The agreement between the two sets is good. The most striking difference is in the values for CNZ *P* and *S* which may in part be attributable to a highly anomalous structure at CNZ (located on an active volcano). As an example of the difficulties of interpretation of the station terms here, the CNZ *S* figure, +6.4 sec. (0.5) was based on data from only four groups. Since both the travel time factor and the station term had to be determined from these four data, the station term is far from reliable since the minimum classes policy required grouping CNZ with other stations. An attempt to independently estimate both a station term and a factor would have given ambiguous results. One can conclude, however, that because of the significant difference between the CNZ *S* term estimate and Haines' value, the average mantle travel time for *S* to CNZ is appreciably slower than that of our assumed model.

Haines has shown that his values correlate well with the isostatic gravity anomalies and that negative anomaly and positive station term can be explained in terms of a thicker or lighter (slower) crust or both. Figure 6.6 compares the isostatic gravity anomaly (Reilly (1965)) with some of my East Coast station terms. Because of an unavoidable interaction between mantle model and crustal model in this study and the imposition of uniformity on the former, the station terms of this study are not to be preferred as estimates of crustal model error to Haines' results (which also may suffer somewhat from this latter problem, but probably not as much). That such similar results should have been produced by two such dissimilar studies gives a measure of confidence in both studies and to the station term estimates.

INTERPRETATION OF THE RESULTS

No further comment will be made about the station terms except to note that they are small enough to satisfy the requirement of the third constraint. We now turn our attention

Figure 6.6 P and S station terms and Isostatic Gravity Anomaly (Reilly 1965)



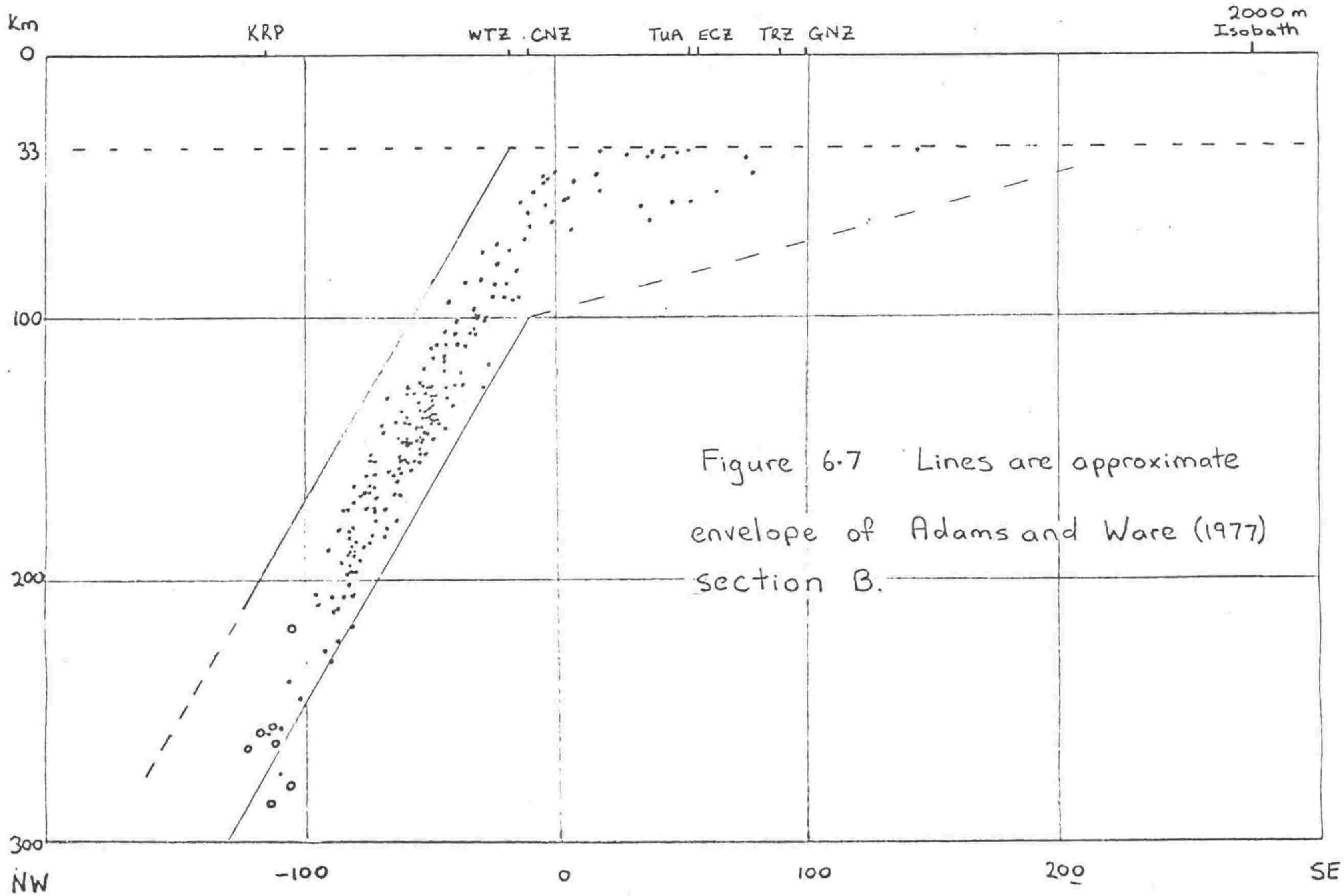
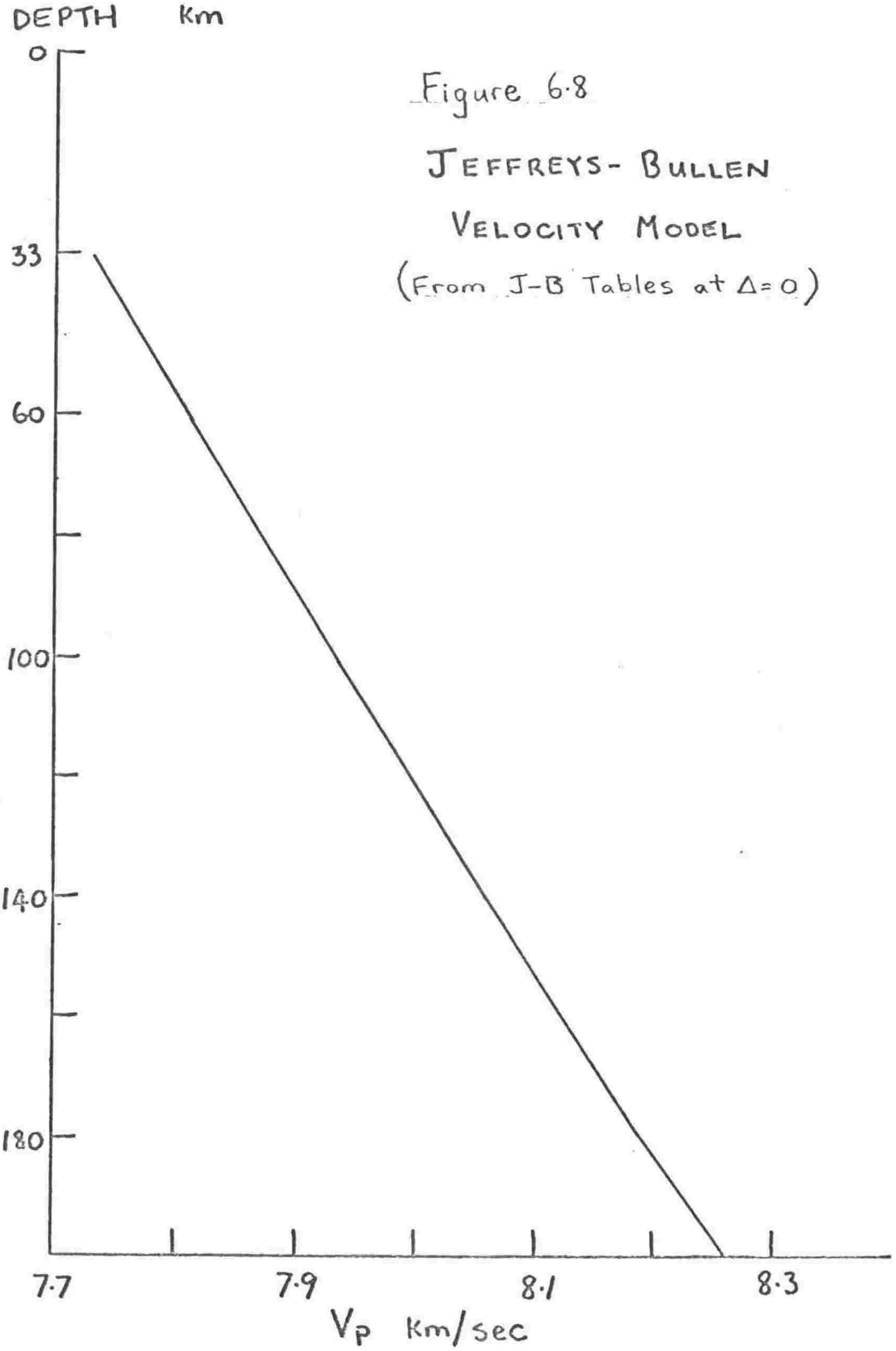


Figure 6.7 Lines are approximate envelope of Adams and Ware (1977) section B.

to the mantle model. The results of the linear analysis indicate that the model presented in Table 6.2 should be modified by altering the -6 percent *S* values to 7 percent and that the new average values, 9 percent fast for *P* and 7 percent fast for *S*, are correct to within one percent. In the case of any particular station however (for example GNZ *S*) the factor may be quite different. We propose only an average mantle model for the stations with paths entirely in the slab. This model is quite close to Adams and Ware's empirically determined model for the hypothesized slab path rays, and it gives rise to a picture of the Benioff Zone (Figure 6.7) which is quite similar to theirs. There is a problem however in interpreting these travel time results in terms of velocities. If the Jeffreys-Bullen ray structure (also used by Adams and Ware) is quite inappropriate for the interior of a downgoing lithospheric slab then it may be dangerous to even make the interpretation that the average velocity is $1/(1 + \alpha/100)$ times the average J-B velocity. It does not follow that a model which satisfactorily gives travel times can be satisfactorily and unambiguously inverted to give velocities. In particular, the nearly linear increase with depth of the J-B velocities (Figure 6.8) means that J-B rays will be flattest near the source and steepest at the station. The velocity structure within subducting lithosphere is difficult enough to theorise about in the absence of enough information about rock properties, temperatures, pressure, etc. However, it might well be that there is less velocity variation with depth than in the J-B model. The Herrin (1968) *P* velocity model, a global model, has this lesser rate of change with depth feature. In such a case, rays would be straighter and the same average velocity would give a faster


DEPTH Km
0
33
60
100
140
180
Figure 6·8
JEFFREYS- BULLEN
VELOCITY MODEL
(From J-B Tables at Δ=0)
7·7
7·9
8·1
8·3
Vp Km/sec

time, or conversely, the same time would imply a slower average velocity. In an attempt to test the effect of such a velocity model, group 24 was reprocessed using constant mantle *P* and *S* velocities for the paths to WTZ, GNZ, TUA (straight line mantle rays), the actual values for which were adjusted to produce, without a fixed event, a solution for the master event of the previous solution as close to its adopted position as possible. There were several interesting features of these experiments. Solutions obtained with greater or lesser velocities than the one arrived at were less stable and yielded a higher overall total error estimate. The final solution adopted with this modification was the best one, in the least squares sense, implying an uncertainty in the velocities adopted of less than one percent. The hypocentre estimates in this solution were within about 3 km of the J-B estimates. The velocities, 8.51 km/sec. for *P* and 4.74 km/sec. for *S* compare with 8.69 km/sec. for *P* and 4.72 km/sec. for *S* from the J-B tables scaled by factors of 0.91 and 0.93, calculated for vertical rays from 145 km - the average depth of group 24. This result does not resolve the problem of inverting the travel times to produce a velocity model, but reinforces one's confidence in the robustness of the method insofar as the results now seem to be somewhat independent of the assumed ray structure. Of course, the homogeneous mantle model was applied to only six out of 21 stations but if there had been severe dependence on structure, it might well have been revealed in such an experiment.

We conclude, tentatively, from the proximity of the solutions obtained using the J-B model and a homogeneous mantle that the average velocities within the slab are 8.6(±0.1) km/sec.

for P and 4.73 (±0.05) km/sec. for S. These results are in good general agreement with other determinations: Robinson (1976) finds an eleven percent average P velocity increase over the J-B average velocity from teleseismic P residual differences. Essentially, Robinson's result depends only on correctly estimating the length of the ray paths in the slab - a longer fast path would imply a smaller velocity contrast. The upward displacement of hypocentres, shortening the apparent length of slab, tends to oppose a longer slab. Haines defines two regions for the East Coast of the North Island. For the northern one, including within it stations GNZ, ECZ, TUA, WTZ, he finds P_n and S_n velocities 8.1 km/sec. (0.1) and 4.65 (0.05). For the southern region, including stations CNZ, TRZ, MNG, WEL he finds 8.5 km/sec. (0.05) and 4.75 (0.02). These velocities were inferred from arrival time differences of P_n and S_n for pairs of stations approximately alligned with shallow earthquake epicentres. The southern region figures agree well with our result. The northern region values are not inconsistent if they pertain to a layer which is thicker in the northern region than in the southern. This would not be inconsistent with the gravity anomalies. In the southern region, P_n and S_n waves would refract through the thinner layer and it would not be seen.

Earlier determinations include Mooney (1970b) who examined the change in residuals with distance from subcrustal events but was unable to produce a quantitative model and Smith, W.D. (1973) who produced a laterally homogeneous model for the North Island mantle a few percent faster than the J-B model above 160 km.

As far as the mantle velocities outside the slab are concerned, little can be concluded from this study. The problem is twofold. Stations not looking down the slab may tend to have ray paths which are a mixture of fast and normal or slow segments. For example, the ray paths to COB for a group such as 24 must consist of a segment in the slab and a segment outside the slab. Without a ray-tracing capability and a good model of structure within a mantle, the exact path cannot be determined.

The second problem is that these stations are relatively isolated. KRP must have ray paths which are entirely outside the slab except for the shallowest events. However, its physical isolation from another station means that inferences about the mantle beneath KRP must be tentative because they cannot be checked. If there were a systematic error in the location estimates all travel times could be substantially in error, but presumably all the pure slab paths would be in error in a similar way or else the observed consistency would not be found. The relative weight of these stations to the rest would indicate that more of the error in such a case would be distributed amongst the non-slab stations. With this reservation, the information from KRP in Table 6.2 is that the average mantle P velocity is approximately four percent faster than the J-B average P velocity and that the S velocity is approximately the J-B average S velocity down to about 100-150 km, deeper than which the S velocity may be about five percent slower. Also, for the two groups 24, 34 (average depth about 140 km) the P velocity is about four percent slower than elsewhere (that is, it is equal to average J-B). The position from the shallowest groups 22/32 and 42 is complicated because the path will be partly in the fast region and partly under the volcanic region where Haines has shown velocities to be quite low. The data from KRP is then quite consistent with the model in which low Q, low velocity material is located above the down-going slab. Mooney (1970a) has shown that the material through which pass rays to KRP certainly has a low Q value. Drawing more inferences about velocities from this study is unwarranted. There is, however, some extra information from an unexpected source which contributes to the picture of the mantle.

VARIATION OF TRAVEL-TIME RESIDUALS BY STATION/GROUPS

In Chapter V while discussing appropriate weighting, the temporary assumption was made that there was no variation in model error between stations which would call for weighting by stations. We can now confirm this to a remarkable extent by considering the distribution of residuals at each station for all the groups. Table 6.7 gives the standard deviation of travel-time residuals at each station by groups. Figure 6.9 shows the average P and S values for each station. With regard to the latter, the values (estimating the total error standard deviation = model error plus reading error) are remarkably constant in the region south-east of the drawn boundary.

One can immediately conclude that there is no significant variation in model error from station to station within this region. Beyond this boundary, in a region roughly corresponding to Mooney's low Q region, the values are lower, but not significantly lower for TNZ or CNZ. The KRP values are very significantly lower which would suggest the interpretation that the mantle under KRP is much more uniform than in the slab. The problem with this interpretation is the result for WTZ. From all other indications, WTZ should show variation of model error like TUA and GNZ. These two stations do have the lowest

TABLE 6.7

STATION	GROUP 22/32	42	22/23	33	24	34	25	35	2367
CNZ P	.30	.15	.34	.29	.30	.27	.32	.31	.30
S						.34	.29	.34	.49
WTZ P	.29	.26	.33		.20	.21	.30	.17	
S	.18	.31			.30	.35		.21	
GNZ P	.37	.37	.24	.24	.27	.30	.30	.25	.29
S	.38	.50	.38	.27	.29	.31	.28	.18	.25
TUA P	.25	.27	.35	.24	.29	.39	.25	.34	.30
S	.33	.25	.39	.26	.30	.32	.21	.37	.37
ECZ P	.28		.19	.25	.29	.40	.31	.48	.38
S					.33		.35		
TRZ P	.42	.34	.42	.29	.25	.40	.22	.24	
S		.31		.43	.32	.30	.24	.21	
MNG P	.33	.39	.43	.16	.39	.25	.28	.29	.28
S		.52	.41	.28	.29	.24	.28	.18	.33
WEL P		.45	.44	.26	.28	.29	.23	.29	.29
S	.24	.30	.28	.32	.27	.30	.26	.21	.31
COB P				.25	.36	.41		.23	.32
S			.38	.20	.25	.29	.17	.30	.29
KRP P	.23	.38	.18	.20	.19	.18	.24	.16	.27
S	.31	.18	.26	.18	.19	.18	.16	.17	.17
TNZ P	.36	.19	.38	.24	.31	.19	.32	.18	.32
S		.27							
GBZ P	.29				.20				

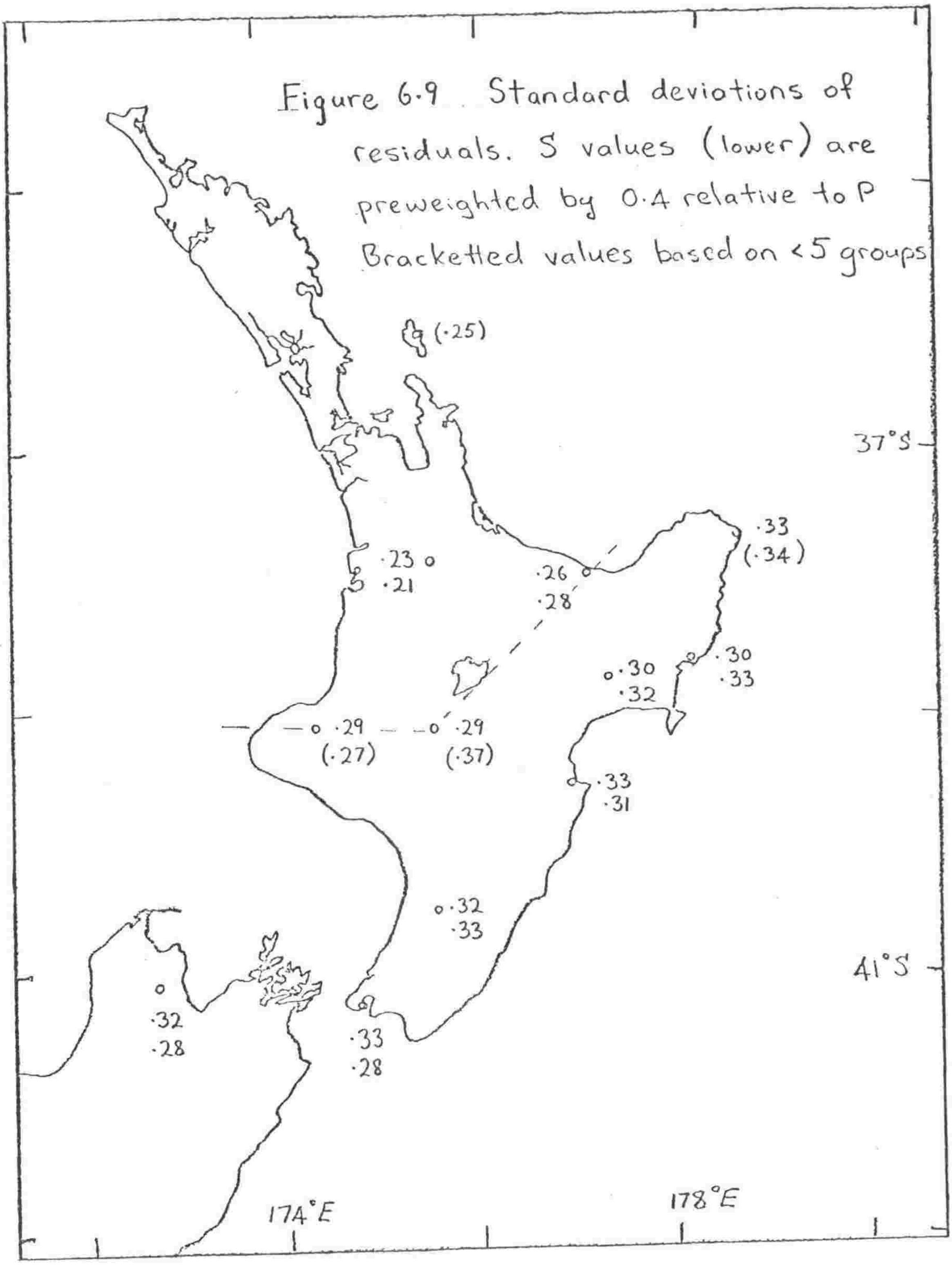


Figure 6·9 Standard deviations of residuals. S values (lower) are preweighted by 0·4 relative to P. Bracketted values based on < 5 groups

P values of any of the above slab stations, and when one considers the variation between groups, the difference between WTZ and GNZ or TUA is barely significant. However, it is a possibility that this lack of variation for WTZ and KRP is due to some artifact of the least squares method whereby the variation at these two relatively isolated and important stations is minimised. On the whole this seems unlikely in view of the overall consistency of the slab station values for a fairly wide geographical spread of groups.

The alternative physical explanation of these low values for WTZ, which would also account for the slightly faster travel times, indicated by lower station terms than Haines', to these three stations (WTZ, GNZ, TUA) is that there is some very small degree of anisotropy in the slab velocity structure. If the slab consists of comparatively narrow fingers moving somewhat independently, one might expect more uniformity of path within a single finger than for paths which cross boundaries, and these latter might have different average velocity if the fingers, because of different rates of subduction, had slightly different velocity structures. This is highly speculative and the only mildly corroborating evidence is the possible existence of faults which strike parallel to the down dip of the zone. One such is suggested by the hypocentres in Chapter V (Figure 5.13). A micro-earthquake survey of the East Coast region of the North Island somewhat south of TRZ, conducted by M. Reyners, revealed a similar fault, almost vertical with similar strike at subcrustal depth. I have endeavoured to find evidence of other faults by looking for south-east/north-west trending seismicity, but since the Seismological Observatory do not attempt depth determinations for shallow earthquakes, and since the error in epicentre estimates appears to be not negligible compared to the dimensions of these faults, while some such sequences of events have been found (compared with the "normal" south-west/north-east trend of seismicity in this region), the evidence for other faults is not particularly convincing. Also, no evidence for any such faulting at depths below 70 km has been found (where the relative hypocentre determination is known to produce superior estimates). It may be that at greater depth the plate becomes more broken up and the faulting so pervasive that no one fault can be isolated. Returning to Table 6.7, which shows the between-groups variation, we note that the majority of the largest values occur amongst the shallowest groups. This may indicate greater homogeneity with depth, or may be a consequence of the greater geographical extent of the two shallowest groups (see Figure 6.1).

One would expect that the wider the aperture of incoming rays to a station, the larger the variation in residuals. This is illustrated by the GNZ values for the shallow groups 22/32 and 42 which are the largest for GNZ of all the groups. On the other hand, WEL *P* has a much higher value for the shallow groups than the deep ones and the apertures of all the groups at WEL must be somewhat similar. The other factor which enters this discussion is simply the number of data at each station in a group. No station has its largest value in group 24, the most populous group, whereas some have largest values in group 34 at a similar average depth. The variation from group to group may be entirely random, although the range of the values would deem the variations to be significant on the basis of assuming a normal distribution of residuals. We conclude this section by summarizing the findings: there is uniformity in the model error variation for all above slab stations with mild indication that there is more variation for shallow ray paths. The mantle outside this region seems much more homogeneous.

A PROFILE OF THE BENIOFF ZONE

The emphasis of this chapter has been on travel-time modelling, but the mosaic of hypocentre estimates from the nine groups gives a compact picture of the Benioff Zone (Figures 6.2 - 6.4). The lateral extent of the groups included in this section is shown in Figure 6.1. All earthquakes located are plotted except that earthquakes further north than 37.9°S are omitted except below 200 km where they are shown as open circles. Several representative 90 percent relative location confidence ellipses are given. The top-most ellipse is an absolute error confidence region. The features of this section, which is similar to that of Adams and Ware in many ways (Figure 6.7), are that the earthquakes below about 70 km are confined to a thin plane sheet dipping at about 60 degrees. The hypocentre estimates here are entirely consistent with the Chapter IV estimate of 10 to 18 km for the zone's thickness below 70 km. The shallower events apparently occupy a thicker zone, although the top-most confidence ellipse indicates the large uncertainty for these events. However, down to 70 km the bottom of the zone, dipping at about 25 degrees seems to be well defined. M. Reyners (personal communication) has found that micro-earthquake hypocentres somewhat south of this section also indicate the bottom of the zone dipping at about 30 degrees.

The seismicity between 33 and 70 km disposes of the apparent gap in the seismicity between these depths as indicated by the standard Observatory locations. Adams and Ware note this also and point out that the apparent gap was almost certainly an artifact of the model used routinely by the Observatory. A number of relocations were assigned depths of 33 km by default. Because of problems of stability, crustal (< 33 km) depths could not be assigned and these events were omitted from the section. It should be pointed out also that this section is not a fair sample of the subcrustal seismicity in the region covered by the groups because the areas of greatest seismicity are <u>under</u>-represented.

One point of disagreement between this section and the corresponding one of Adams and Ware is that they find the dip of the zone to be 50 degrees. Their hypocentre estimates are much more scattered than the solutions given here, but most of this difference appears to be due to differences in the model. The value of the dip for group 24 obtained in Chapter V was 50° also. For the unmodified J-B hypocentres of Chapter IV, the value was 67°. It would seem that the relative vertical error of the hypocentres increases with depth in such a way as to make the apparent dip of the zone quite model dependent. While preferring the result obtained here to that of Adams and Ware, that there is some appreciable uncertainty in the estimate must be allowed.

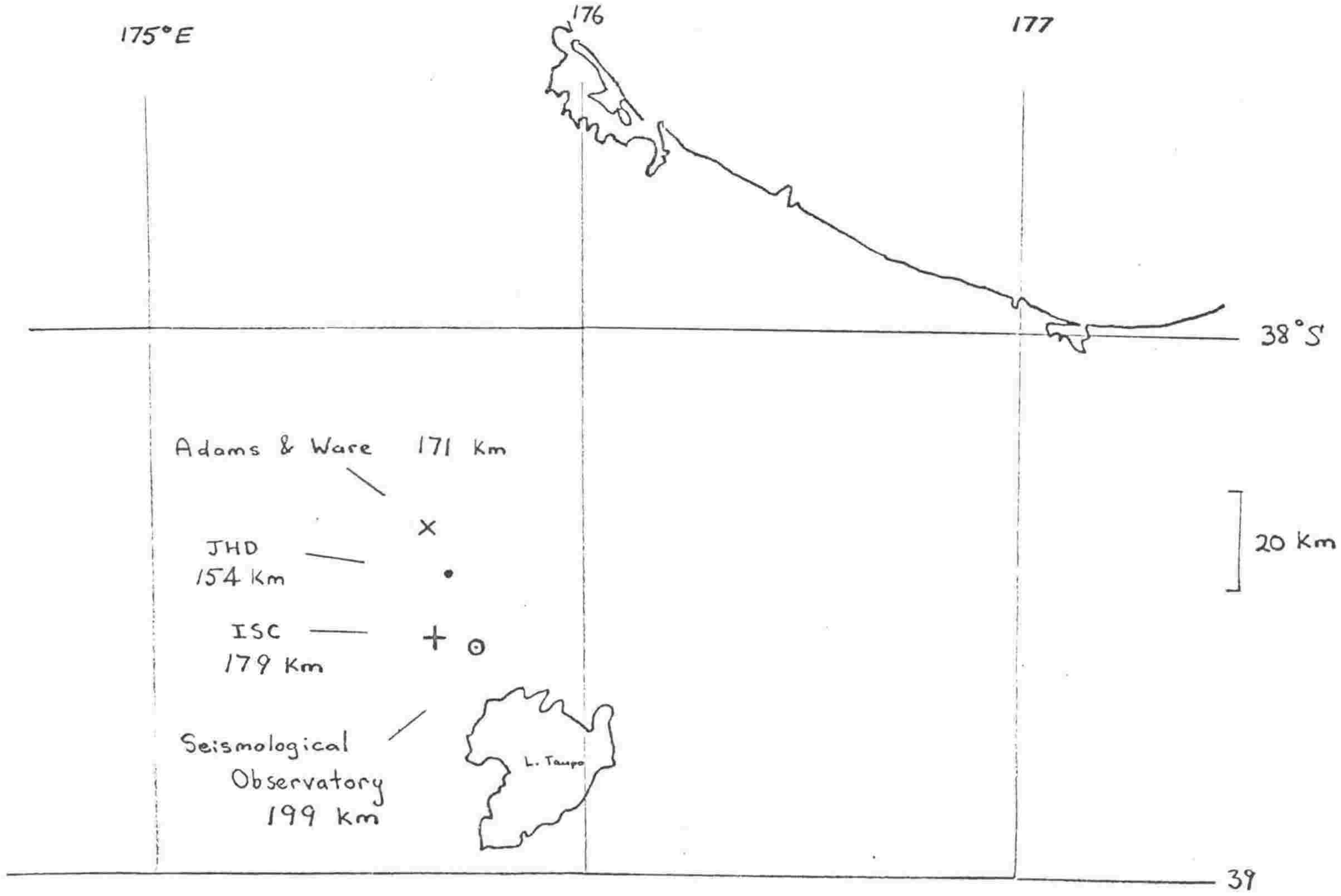


Figure 6.10 Hypocentre estimates for event of 16 Jan 1974

The results of Chapter IV allow us to extrapolate the picture of the Benioff Zone over the North Island. One of the most difficult problems raised by this picture is the sharp transition between moderate dip (25 degrees) and steep dip (50 or 60 degrees) which takes place over about 50 km. It is not feasible to suggest that some 80 km thick oceanic lithosphere (Leeds *et al.* (1975)) can be elastically deformed around such tight curve. If the plate motion is continuous then plastic deformation must be invoked to account for such a transition. An episodic subduction process in which substantial sections of plate partially crack through and hinge downward could explain the observed linearity of the hypocentres.

CONCLUSION

It has been shown that the additional model information provided by Joint Hypocentre determinations can be made to yield a new travel time model which is much superior in terms of goodness of fit. The problem remains of not having some absolute standard by which to judge the results.

First, there is the question of how good the final hypocentre estimates are. Figure 6.10 gives a comparison of solutions for a well observed event from group 24. The differences are due to differences in data and model: the Seismological Observatory solution used P and S data from New Zealand only and a uniform Jeffreys-Bullen model. The International Seismological Centre solution used 22 New Zealand and 3 Teleseismic P observations and a uniform J-B model. Adams and Ware used *P* and *S* New Zealand data and an asymmetric J-B model. The following conclusions can be drawn. First, the hypocentre estimate is highly model dependent - the Adams and Ware epicentre is more than 20 km from the Seismological Observatory Solution. Second, the effect of adding teleseismic data has an effect on the solution, particularly the depth, out of proportion to the number of observations added. This is because the number of New Zealand rays leaving the source at a steep angle and thus strongly contributing to the depth control is not large, whereas the teleseismic rays all leave the source comparatively steeply.

We tentatively draw the following inferences. The Adams and Ware and JHD epicentres are to be preferred to the others because of the more realistic model employed. Of these two, the JHD solution gives a very much better fit than Adams and Ware, whose solutions, in terms of standard deviation of travel time residuals, fit no better than the Seismological Observatory solutions.

The question arises though as to the effect of adding teleseismic data when a better model is employed. For teleseismic ray paths to the Northwest, where most teleseismic *P* observations of New Zealand earthquakes are made, a significant portion of the ray must travel through the slab where the velocities are presumed to be higher than in the straight J-B model. The effect on depth of adding teleseismic observations to the data is thus difficult to estimate. The average difference between Adams and Ware and JHD depths for group 24 was almost 20 km. One should be wary of judging that because the

Adams and Ware depth and the ISC depth are similar, they should be preferred to the JHD depth. On the other hand, the depth uncertainty of the JHD solution (standard deviation 6 km*) is increased by the possibility of systematic errors. It will be remembered from Chapter V that a large range of depths for a master event gave a nearly identical fit.

This problem is not resolvable with the information at present available. Very well located shallow events which could be used as masters to bootstrap down the Benioff Zone using JHD might provide a solution.

We conclude with a brief discussion of the consequences of the velocities deduced for the slab - 8.6 km/sec. for *P* and 4.73 km/sec. for *S*. First, it is physically possible to obtain such velocities. Ringwood and Green (1966) show that the basalt to eclogite transition provides material where velocities can be this high. Marsh and Carmichael (1974) show the transition being gradual and starting at very shallow depths soon after subduction. As subduction proceeds, the transition continues towards eclogite producing increasing velocities. Following such a model implies that it is the old oceanic crust, converted to eclogite, which provides the fast paths. This would explain why WTZ, essentially above the projection of the Benioff Zone, was found to have fast paths. The existence of fast paths to GNZ suggests that the transition must be highly advanced in the shallow dipping part of the zone if this model is correct.

Qualitative support for this model is provided by the pattern of energy transmission from mantle earthquakes. Mooney

---

* See Appendix VI, Table 5.

(1970a) shows that stations looking down the slab receive signals at higher frequency than the out-of-slab stations and that there is a correlation between the high Q region and the pattern of intensities felt from mantle earthquakes - these are felt strongly by observers up and down the east coast and very weakly by observers to the northwest. If the down-going plate consists of normal oceanic lithosphere capped with a higher velocity layer, the slab will act as a wave guide, while the fast layer will transmit high frequencies to above slab stations.

Very little information was provided from this study about the above plate region. If, as was tentatively concluded, the average velocity is a few percent higher than given by the Jeffreys-Bullen model, then the contrast between the in-slab and out-of-slab velocities is reduced to a figure which is more consistent with the contrasts calculated at other Benioff zones.

CHAPTER VII

FURTHER EXTENSIONS AND OTHER APPLICATIONS OF JHD

In this chapter we do not attempt a complete survey of all possible applications of the joint hypocentre methods, but we will discuss some ideas which seem promising and review some extensions which have been recently made to JHD.

JHD AND MODELLING

Crosson (1976 a and b) discusses an extension of JHD in which, for an assumed plane layered geometry, velocity improvements for each layer as well as hypocentres are obtained. The work bears some resemblance to the work of Chapter VI in that subjective decisions about the number and thickness of layers, based on trial solutions, are required. In Crossman (1976 b), quite notable success is achieved in modelling the upper crustal structure of Puget Sound.

This method is inherently that of determining velocities for a given geometry, but can be easily extended to more complicated geometries provided a fast method of calculating travel times (essentially ray tracing) is available. In the case of group 24, 1125 travel time calculations are required at each iteration. The importance of rapid ray tracing is obvious.

An entirely different extension can be found in the work of Fitch & Muirhead (1974), Fitch (1975), Fitch & Rynn (1976) in which the velocity (or velocities) and ray incident angle at the source of a group of earthquakes, assumed constant over a volume containing the earthquakes, are estimated together with the relative positions of the events to a master event. Fitch and Muirhead (1974) and Fitch and Rynn (1976) use this method on crustal earthquakes and Fitch (1975) uses it on deep (600 km) events. We will briefly develop the required equations here since there are some difficulties with the method which seem to have escaped notice.

Let $\underline{r}_j$ be the position of the j^{th} slave with respect to the master event. In polar coordinates we have:

$$\underline{r}_j = r_j(\sin\Theta_j \cos\phi_j, \sin\Theta_j \sin\phi_j, \cos\Theta_j) \quad \ldots(7.1)$$

Let $\hat{\underline{i}}$ be the (unit) vector tangent to the ray to a given station at the master. The assumption is made that rays to all the events of the groups are parallel. Let $\hat{\underline{i}}$ be given by:

$$\hat{\underline{i}} = (\sin\Theta \cos\phi, \sin\Theta \sin\phi, \cos\Theta) \quad \ldots(7.2)$$

where Θ is the incident angle and ϕ the longitude of the ray at the source.

If v is the velocity of the phase being considered, the extra travel time to the slave is given by:

$$\delta T = (\underline{r}_j \cdot \hat{\underline{i}})/v \quad \ldots(7.3)$$

which is equated to the difference in arrival times for slave and master and the station minus the estimated difference in origin times.

If r_{j0}, Θ_{j0}, ϕ_{j0}, v_0, Θ_0, ϕ_0 are initial estimates of the unknown parameters, by a Taylor's expansion of δT we have:

$$\delta T = (\underline{r}_{j0} \cdot \hat{\underline{i}}_0)/v_0 + \frac{\partial}{\partial r_j}\delta T_0 \delta r_j + \frac{\partial\delta}{\partial\Theta_j}T_0\delta\Theta_j$$
$$+ \frac{\partial\delta}{\partial\phi_j}T_0\delta\phi_j + \frac{\partial\delta}{\partial\Theta}T_0\delta\Theta + \frac{\partial\delta}{\partial\phi}T_0\delta\phi + \frac{\partial\delta}{\partial v}T_0\delta v$$
$$+ \text{ higher order terms} \qquad \ldots(7.4)$$

Expanding the scalar product of *(7.3)* gives:

$$\delta T = \frac{r_j}{v}(\sin\Theta_j \sin\Theta \cos(\phi_j - \phi) + \cos\Theta_j \cos\Theta) \quad \ldots(7.5)$$

whence, by inspection, ϕ_j and ϕ are not independently obtainable, as Fitch and Rynn assert.

Writing down the terms $\frac{\partial}{\partial v}\delta T_0$ and $\frac{\partial}{\partial\Theta}\delta T_0$ we have:

$$\frac{\partial}{\partial v}\delta T_0 = -\frac{r_{j0}}{v_0^2}(\sin\Theta_{j0} \sin\Theta_0 \cos(\phi_{j0} - \phi_0)$$
$$+ \cos\Theta_{j0} \cos\Theta_0) \qquad \ldots(7.6)$$

$$\frac{\partial}{\partial\Theta}\delta T_0 = \frac{r_{j0}}{v_0}(\sin\Theta_{j0} \cos\Theta_0 \cos(\phi_{j0} - \phi_0)$$
$$- \cos\Theta_{j0} \sin\Theta_0) \qquad \ldots(7.7)$$

Now consider what happens if the hypocentres happen to be distributed in a plane. Without loss of generality, take this plane to be the equatorial plane of the coordinate system, so that $\Theta_j = \Pi/2$ for each j. Then:

$$\frac{\partial}{\partial v}\delta T_0 = -\frac{r_{j0}}{v_0^2} \sin \Theta_0 \cos(\phi_{j0} - \phi) \qquad \ldots(7.6a)$$

$$\frac{\partial}{\partial\Theta}\delta T_0 \quad = \quad \frac{r_{j0}}{v_0}\cos\Theta_0\,\cos(\phi_{j0} - \phi) \qquad \ldots(7.7a)$$

whence the columns associated with δv, $\delta\Theta$ of the design matrix in the linear model are linearly dependent and so the system will be singular. Moreover, if the events are not coplanar, but are closely distributed about a plane, as the earthquakes studied in this thesis were found to be, one can expect that the system of equations will be poorly conditioned and that errors in δv and $\delta\Theta$ will be highly correlated.

In view of this difficulty, some care must be used in the application of the method. Fitch (1975) tests a simpler method in which only the velocity and not the ray geometry are determined jointly with the hypocentres. In his study, Fitch concluded that inclusion of the ray angles were necessary to obtain a satisfactory solution, but the values obtained for the *P* velocity: 10.7(0.1) km/sec. for the velocity only solution and 11.2(0.4) for velocity plus angles (for Fiji earthquakes in the depth range 600-660 km) are not significantly different, in view of the standard deviations.

Experiments were therefore conducted in which the *P* and *S* source velocities were computed for group 24. The model adopted was that the true travel time difference between slave and master is proportional to the model travel time difference, the constant of proportionality being the fractional error in the velocity which is jointly solved for with the hypocentres and station terms. Briefly, the equations of condition become:

$$r_{ij} = \delta H_i + (\nabla T)_{ij}\delta \underline{x}_i + R_{0j} + \lambda(T_{ij} - T_{0j}) + e_{ij} \qquad \ldots(7.8)$$

for the j^{th} residual of the i^{th} event; δH_i, $\delta \underline{x}_i$, R_{0j} and λ being solved for by minimising the errors e_{ij} in the least squares sense.

Two sets of solutions were obtained. First, it was considered unwise to assume that a uniform velocity applied over the source region in view of the proximity of the source region to the edge of the slab. Thus two sets of velocities were solved for, one for the above slab stations, and one for KRP alone. In the second solution, WTZ, TUA and GNZ were separated from the above slab stations, and so three sets of velocities were obtained. The results are given in Table 7.1.

The first thing to note is that in both sets the ten percent standard deviation values for KRP make these estimates insignificantly different from zero. More than 60 observations each of P and S contributed to this solution. One can conclude that in such circumstances the velocity to a single station cannot be determined by this method.

The other standard deviation estimates were of the order of two to three percent. In no case was the determined source velocity significantly different from the initially assumed value. Neither solution significantly reduced the residual sum of squares.

TABLE 7.1

VELOCITY CONTRASTS

		% CONTRAST	STANDARD DEVIATION
I			
KRP	P	+3.7	11.4
	S	+4.4	10.6
REST	P	-0.2	2.3
	S	+2.1	2.3
II			
KRP	P	-2.3	12.2
	S	+3.7	11.6
WTZ TUA GNZ	P	+0.4	2.4
	S	+3.4	2.5
REST	P	+3.8	2.9
	S	+1.3	4.3

Within the limitation again imposed by assuming a J-B ray structure, we conclude that the velocities initially assumed at the source for the path in the slab are correct to within about two percent. The values are, for 145 km (average for group 24) 8.9(±.2) km/sec. for *P* and 4.86(±.1) km/sec. for *S*.

OTHER USES OF JHD

When this study was begun, the possibility that seismic velocities could vary with time was not considered. There is an implicit assumption throughout this work of invariance of the model with time. With the numerous reports of apparent *P* velocity decrease in a region prior to an earthquake (a New Zealand example being Sutton (1974) discussed below), the possibility arises that, having established a good regional model, one might use JHD to monitor velocity changes by looking for changes in the station terms with time. There are many obvious difficulties in establishing such a scheme, not the least of which is determining the "normal" model sufficiently accurately.

The fortuitous station placement of GNZ with respect to the 1966 Gisborne earthquake (Hamilton, 1969) gave the opportunity for testing the ability of JHD to detect velocity changes. Sutton (1974) reported that for a period of some 400 days prior to the Gisborne earthquake, teleseismic *P* arrivals were about 0.5 seconds later than an average value established during the time before and after the quake. The angles of incidence at GNZ of the rays were in the range 15°-30° and most of the arrivals were from earthquakes in the NW quadrant from

GNZ. The rays would thus pass through almost the same crustal region as rays from group 24 to GNZ. A group of 20 events which occurred in 1965 and were located by the Seismological Observatory in approximately the same geographical position as group 24 were relocated using JHD and the mantle model developed for group 24. Figure 7.1 shows the aperture of these events and the position of group 24. Amongst the 65 events of group 24 was one 1965 event. (The reason for the low numbers of events from the years 1964-1968 is the introduction of TRZ in 1969 and WTZ in 1972 which provided relatively more data for events from later years.)

This event was used as a master event for the 1965 group. The object of the experiment was to compare GNZ *P* station terms for group 24 and the 1965 group, which under the hypothesis, should be more positive than the "normal" value, established by group 24.

A comparison of station terms for the stations common to the two solutions is given in Table 7.2. There is no significant increase in GNZ *P* station term - the difference between GNZ *P* and GNZ *S* for the two groups is almost identical. In order to verify that the method is capable of detecting an average change of 0.5 sec, 0.5 second was added to the GNZ *P* arrivals for the 1965 events. In theory, the GNZ *P* term should be increased by $0.5 - (0.5)/N$, where N is the number of stations, since the sum of the station terms is constrained to be zero. At the same time, the other station terms will increase by $-(0.5)/N$. In this case $N = 16$, so:

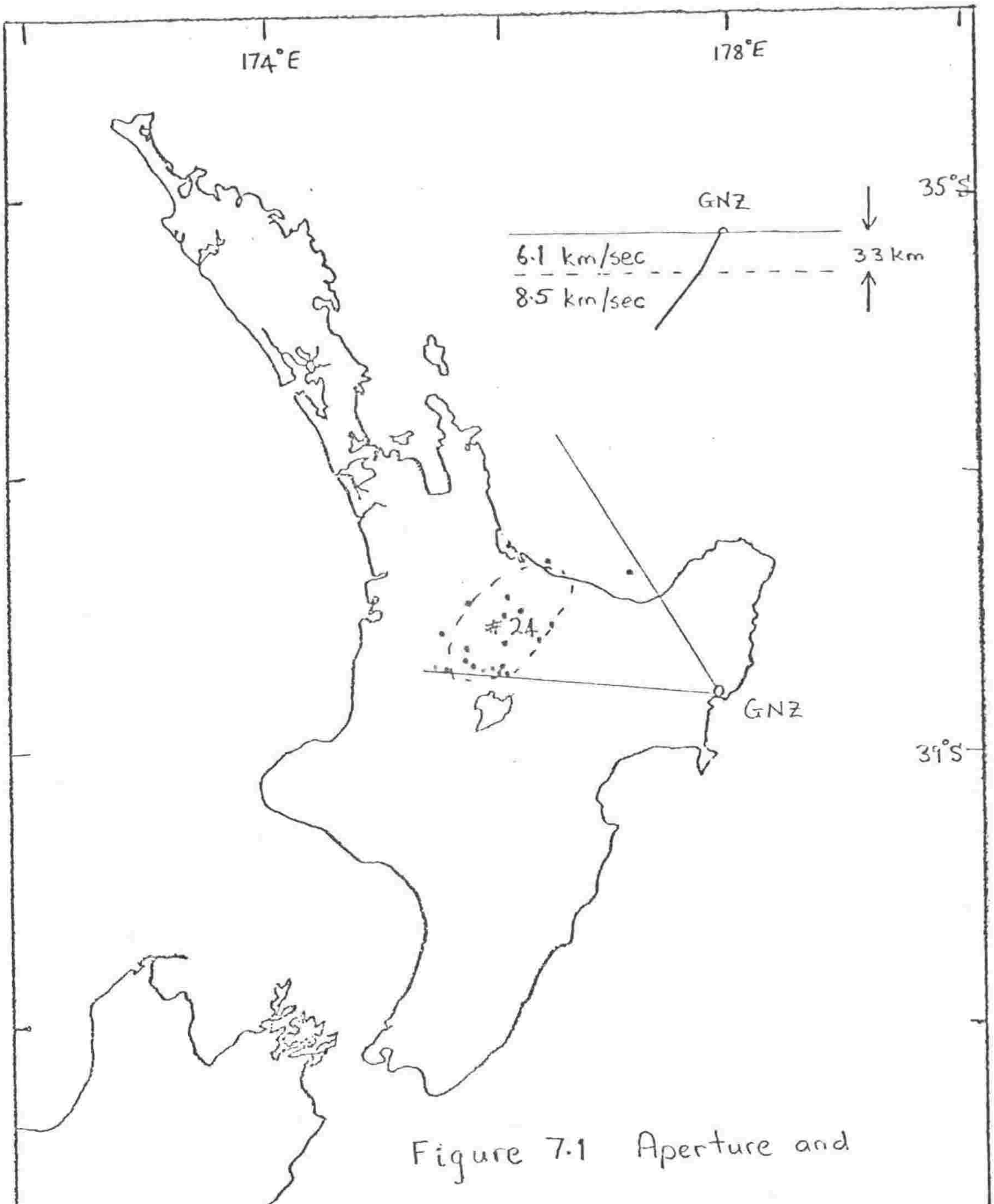


Figure 7·1 Aperture and mean incident angle at GNZ

TABLE 7.2

STATION TERMS 1965 AND GROUP 24

Station	1965	Group 24	1965 - #24
CNZ P	0.95	0.99	-0.04
ECZ P	-0.07	0.20	-0.27
S	1.73	1.95	-0.22
GNZ P	-0.36	-0.24	-0.12
S	-0.30	-0.20	-0.10
KRP P	-0.58	-0.43	-0.15
S	-0.74	-0.35	-0.39
MNG P	-0.80	-0.42	-0.38
S	-0.77	-0.56	-0.21
TNZ P	-0.64	0.70	-0.06
WEL P	-0.76	-0.76	.00
S	-1.06	-0.32	-0.74
COB S	-0.68	-0.91	0.23
TUA P	-0.17	-0.03	-0.14
S	0.10	0.97	-0.87

$$0.5 - (0.5)/N = 0.47$$

and the observed change in the GNZ *P* station term was exactly 0.47 sec.

This demonstrates that the average difference in travel times between 1965 and the normal value established by group 24 (containing only one 1965 event) is estimable by JHD within a factor of $(N-1)/N$. It is possible of course that because of the random errors in the station term estimates, the difference will be undetected. Assuming that the parameter estimates are normally distributed, one can deduce from their standard deviation estimates the probability of obtaining the actual result assuming a true average change of T_c. The results of this analysis are:

X (sec)	Probability that $T_c \geq X$
0.52	0.05
0.43	0.1
0.32	0.2

(The details of this analysis are contained in Appendix VII.)

This strongly suggested that no travel time change as large as 0.5 sec. occurred in arrivals from intermediate depth earthquakes to GNZ in the same apperture as the teleseismic arrivals which showed a +0.5 sec. change. Work on the cause of this discrepancy is continuing. While the non-confirmation of the result is disappointing, the experiment certainly suggests that JHD may have a role in this field.

Finally we mention the use of JHD in locating shallow earthquakes at a regional level where crustal phases are present. The Bulletins of the Seismological Observatory record numerous crustal phases, the identifications being based on arrivals with times close to the model arrival time for the same crustal phase. The crustal model used (layers 12 km and 21 km thick with *P*, *S* velocities 5.5, 3.3 and 6.5, 3.7) may be very poor in some areas, in which case identifications based on an erroneous model are less certain. Herein lies the difficulty of applying JHD. In our terms, a set of arrival times of a given crustal phase from a group of crustal earthquakes at a seismograph station would represent a "station". The presence of wrongly identified arrivals in the set would enormously add to the errors. In brief experiments, the aftershocks of the 1968 Inangahua earthquake (Adams and Lowry 1971) and the aftershocks of a magnitude 7.0 earthquake which occurred close to seismograph station MSZ (Milford Sound) in 1976 were relocated using JHD without any phase reidentification. It was extremely difficult to obtain satisfactory solutions in either case (and quite impossible without a master event). In the case of the Inangahua earthquakes, the problem seemed to be due to the very large residuals of star phases (rays penetrating the second layer but not the mantle) at almost all seismograph stations. In the case of the Milford Sound earthquake, such solutions as could be obtained suggested a model very different from the standard one by which the phase identifications were made. It would appear that this problem will require considerable work before joint locations of crustal earthquakes using crustal phases can be attempted in New Zealand.

APPENDIX I

ALGORITHM FOR SOLVING THE JOINT EQUATIONS OF CONDITION BY LEAST SQUARES

In place of the equations *(3.13)*, we have the system:

$$\begin{pmatrix} {}_{(n_1)}A_1 & & & & \vdots & {}_{(n_1)}I'_1 \\ & {}_{(n_2)}A_2 & & & \vdots & {}_{(n_2)}I'_2 \\ & & \ddots & & \vdots & \vdots \\ & & & {}_{(n_M)}A_M & \vdots & {}_{(n_M)}I'_M \end{pmatrix} \begin{pmatrix} \delta\underline{x}_1 \\ \vdots \\ \delta\underline{x}_M \\ \underline{s}_\mu \end{pmatrix} = \begin{pmatrix} {}_{(n)}\underline{y}_1 \\ \vdots \\ {}_{(n_M)}y_M \end{pmatrix} \qquad \ldots (A1.1)$$

where there are $n_i \leq N$ observations for event i, ${}_{(n_i)}A_i$ is an $n_i \times 4$ coefficient matrix of the condition equation of event i, and ${}_{n_j}I'_j$ is a deleted identity matrix: the k^{th} row of an $N \times N$ identity matrix is deleted if the k^{th} station does not provide an observation.

The least squares solution of this system may be obtained using Householder's QR method (Householder (1953)) which is space conserving and numerically efficient.

The object is to find an orthogonal matrix Q which reduces the coefficient matrix of *(A1.1)* to upper triangular form. If ${}_{n_i}Q_i$ is such that:

$${}_{(n_i)}Q_iA_i = \begin{pmatrix} U_i \\ \underset{\approx}{0} \end{pmatrix} \qquad \ldots (A1.2)$$

where U_i is a 4×4 upper triangular matrix, then:

$$Q_1 = \begin{pmatrix} (n_1)^{Q_1} & & & \\ & (n_2)^{Q_2} & & \\ & & \ddots & \\ & & & (n_M)^{Q_M} \end{pmatrix}$$

is an orthogonal matrix which reduces the left hand side of *(A1.1)* to:

$$Q_1 A = \left[\begin{array}{cccc|c} (n_1)\begin{pmatrix} U_1 \\ \underset{\approx}{0} \end{pmatrix} & & & & (n_1)^{Q_1 I_1'} \\ & (n_2)\begin{pmatrix} U_2 \\ \underset{\approx}{0} \end{pmatrix} & & & (n_2)^{Q_2 I_2'} \\ & & \ddots & & \vdots \\ & & & (n_M)\begin{pmatrix} U_M \\ \underset{\approx}{0} \end{pmatrix} & (n_M)^{Q_M I_M'} \end{array}\right] \qquad \ldots (A1.3)$$

By row interchange, this matrix is replaced by:

$$\Pi Q_1 A = \left(\begin{array}{c|c} U & 4M\times N \\ \hline \underset{\approx}{0} & Z \end{array}\right) \qquad \ldots (A1.4)$$

which is reduced by the orthogonal matrix

$$Q_2 = \left(\begin{array}{cc} (4\times M)^{I} & \underset{\approx}{0} \\ \hline \underset{\approx}{0} & Q_Z \end{array}\right)$$

where $$Q_Z Z = \begin{pmatrix} U_2 \\ \underset{\approx}{0} \end{pmatrix} \qquad \ldots (A1.5)$$

If Π is the matrix of row permentation required to produce the desired row interchanges, then:

$$\mathbf{Q} = \mathbf{Q}_2 \prod \mathbf{Q}_1 \qquad \text{...(A1.6)}$$

is the required orthogonal matrix.

The algorithm then consists of sequentially reducing the individual condition matrices A_i to upper triangular form and separating the first four rows of the matrix Q_i from the remaining n_i - 4, the first being stored in X of *(A1.4)*, the latter being stored in Z. The matrix Z, which is $(\sum_{i=1}^{M} n_i - 4M)$ $\times$ N, is then reduced to D by Q_Z *(A1.5)*.

The conditioning of the A_i is presumed to be sufficiently good that the rank of the R_i *(A1.3)* is 4. The rank of Z, and hence U_Z, is theoretically N - 1 because the original equations have one undetermined parameter (see Chapter III). In practice, it is wise to allow for the possibility that ill-conditioning of Z makes the effective rank of Z less than N - 1 -- a circumstance which in fact was never encountered in this work.

The parameter estimates are then obtained by back substitution against the transformed right hand side of *(A1.1)*. By the nature of $\mathbf{Q}$, this transformed vector can be obtained by sequentially calculating $Q_i \underline{y}_i$ and separately storing the first four and remaining (n_i-4) components and then applying $\mathbf{Q}_2$ to the $\sum_{i=1}^{M} (n_i-4)$ vector that results. Call this vector $\underline{y}'$. If the rank of Z is N - 1, then:

$$N \begin{pmatrix} \diagdown \boxed{U_Z} \\ \underset{\approx}{0} \end{pmatrix} = \begin{pmatrix} \diagdown \boxed{U_L} & \underline{d} \\ \underset{\approx}{0} & 0 \end{pmatrix}$$

where U_L is $(N-1) \times (N-1)$ of rank $N-1$, and $\underline{d}$ is an $(N-1)$ vector. It is easy to solve:

$$\begin{pmatrix} U_L & \underline{d} \\ \underset{\sim}{0} & 0 \end{pmatrix} \begin{pmatrix} \underline{s}_\mu \end{pmatrix} = {}_{(N-1)}\underline{y}' \qquad \ldots(A1.7)$$

so that $\underline{s}_\mu$ satisfies an arbitrary linear constraint. (See Appendix II.)

If the required constraint is $\sum_{j=1}^{N} s_{\mu_j} = 0$, then the required solution is:

$$\underline{s}_\mu = \frac{1}{\delta} \begin{array}{c} {\scriptstyle (N-1)} \\ {\scriptstyle (1)} \end{array} \begin{pmatrix} (\{\delta I - (U_L^{-1}\underline{d})\underline{1}^T\}U_L^{-1}) \\ \hline \underline{1}^T U_L^{-1} \end{pmatrix} \underline{y}' \qquad \ldots(A1.8)$$

where $\delta = \sum_{j=1}^{N-1} (U_L^{-1}\underline{d})_j - 1$. To impose the constraint that one of the events i be fixed, corresponding to the use of a master event, one merely deletes the columns of the matrix A_i corresponding to the parameters fixed. If the first column (origin time of i) is retained, the system still has a rank deficiency of one. If the origin time is fixed, *(A1.8)* is replaced by:

$$\underline{s}_\mu = U_Z^{-1}\,{}_N(\underline{y}') \qquad \ldots(A1.9)$$

The theory is exactly the same if weights are applied. If a weighting matrix $\mathbb{W}$ is used,

$$\mathbb{W} = \begin{pmatrix} {}_{(n_1)}W_1 & & & \\ & {}_{(n_2)}W_2 & & \\ & & \ddots & \\ & & & {}_{(n_M)}W_M \end{pmatrix} \qquad \ldots(A1.10)$$

where W_i is diagonal, then *(A1.1)* is replaced by:

$$\begin{pmatrix} W_1A_1 & & & & \vert & W_1I_1' \\ & W_2A_2 & & & \vert & W_2I_2' \\ & & \ddots & & \vert & \\ & & & W_MA_M & \vert & W_MA_M' \end{pmatrix} \begin{pmatrix} \delta\underline{x}_1 \\ \delta\underline{x}_2 \\ \vdots \\ \delta\underline{x}_M \\ \underline{s}_\mu \end{pmatrix} = \begin{pmatrix} W_1\underline{y}_1 \\ W_2\underline{y}_2 \\ \vdots \\ \vdots \\ W_M\underline{y}_M \end{pmatrix} \qquad \ldots(A1.11)$$

and the procedure is identical.

APPENDIX II

CALCULATION OF LEAST SQUARES PARAMETER ESTIMATE VARIANCE/COVARIANCE MATRICES WITH STRUCTURED DESIGN MATRICES AND USING HOUSEHOLDER DECOMPOSITION

The equations we are solving by least squares can be written:

$$\begin{pmatrix} A_1 & & & & & & \vline & I_1' \\ & A_2 & & & & & \vline & I_2' \\ & & A_3 & & & & \vline & \vdots \\ & & & \ddots & & & \vline & \\ & & & & A_i & & \vline & I_i' \\ & & & & & \ddots & \vline & \\ & & & & & A_M & \vline & I_M' \end{pmatrix} \begin{pmatrix} \delta\underline{x}_1 \\ \delta\underline{x}_2 \\ \cdot \\ \cdot \\ \delta\underline{x}_i \\ \cdot \\ \cdot \\ \delta\underline{s} \end{pmatrix} = \begin{pmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \cdot \\ \cdot \\ \underline{Y}_i \\ \cdot \\ \cdot \\ \underline{Y}_M \end{pmatrix} \qquad \ldots(A2.1)$$

(with $4M + N$ parameters to estimate. The dashes on the matrices I_i denote deleted Identities: if the j^{th} station does not record the i^{th} event, the j^{th} row of I_i is absent.) We make the standard assumption that the LHS is the expected value of the RHS. The least squares parameter estimates are a linear function of the RHS:

$$\hat{\beta} = L\underline{Y} \qquad \ldots(A2.2)$$

such that: $E(\hat{\underline{\beta}}) = LE(\underline{Y}) = LX\underline{\beta} = \underline{\beta}$; i.e. L is a left inverse of the LHS matrix.

Let $\hat{\underline{\beta}}^1$ be any subset of $\hat{\underline{\beta}}$, so:

$$\hat{\underline{\beta}}^1 = L^1\underline{Y} \qquad \ldots(A2.3)$$

(taking the appropriate rows of L). Then:

$$\begin{aligned}
\text{var }(\underline{\hat{\beta}}^1) &= E\{(\underline{\hat{\beta}}^1 - \underline{\beta}^1)(\underline{\hat{\beta}}^1 - \underline{\beta}^1)^T\} \\
&= E\{(L^1\underline{Y} - \underline{\beta}^1)(L^1\underline{Y} - \underline{\beta}^1)^T\} \\
&= L^1E\{\underline{YY}^T\}L^{1T} - \underline{\beta}^1(L^1X\underline{\beta})^T - L^1X\underline{\beta\beta}^{1T} + \underline{\beta}^1\underline{\beta}^{1T} \\
&= L^1\{\text{var }(\underline{Y}) + X\underline{\beta\beta}^TX^T\}L^{1T} - \underline{\beta}^1\underline{\beta}^{1T} \\
&= \sigma^2L^1L^{1T} + L^1X\underline{\beta\beta}^TX^TL^{1T} - \underline{\beta}^1\underline{\beta}^{1T}
\end{aligned}$$

The last two terms cancel so:

$$\text{var }(\underline{\hat{\beta}}^1) = \sigma^2L^1L^{1T} \qquad \ldots(A2.4)$$

Also:

$$\begin{aligned}
\text{cov }(\underline{\hat{\beta}}^1,\underline{\hat{\beta}}^2) &= E\{(\underline{\hat{\beta}}^1 - \underline{\beta}^1)(\underline{\hat{\beta}}^2 - \underline{\beta}^2)^T\} \\
&= E\{(L^1\underline{Y} - \underline{\beta}^1)(L^2\underline{Y} - \underline{\beta}^2)\} \\
&= L^1\{\sigma^2I + X\underline{\beta\beta}^TX^T\}L^{2T} - \underline{\beta}^1(L^2X\underline{\beta})^T \\
&\quad - L^1X\underline{\beta\beta}^{2T} + \underline{\beta}^1\underline{\beta}^{2T} \\
&= \sigma^2L^1L^{2T} + \underline{\beta}^1\underline{\beta}^{2T} - \underline{\beta}^1\underline{\beta}^{2T} - \underline{\beta}^1\underline{\beta}^{2T} + \underline{\beta}^1\underline{\beta}^{2T}
\end{aligned}$$

i.e.

$$\text{cov }(\underline{\hat{\beta}}^1,\underline{\hat{\beta}}^2) = \sigma^2L^1L^{2T} \qquad \ldots(A2.4a)$$

After applying Householder transformations Q sequentially to $(A2.1)$ we obtain:

$$\left(\begin{array}{ccccc|c}
U_1 & & & & & \overbrace{Q_1 \,|\, 4}^{N} \\
& \ddots & & & & \vdots \\
& & {}^{4}_{4}U_i & & & Q_i \,|\, 4 \\
& & & \ddots & & \vdots \\
& & & & U_M & Q_N \,|\, 4 \\
& & & & N\Big\{ & U_L \,|\, \underline{d} \\
& & & & & \underbrace{\quad \,|\, 0}_{N-1}
\end{array}\right)
\begin{pmatrix} \delta\underline{\hat{x}}_1 \\ \vdots \\ \delta\underline{\hat{x}}_M \\ \vdots \\ \delta\underline{\hat{s}} \end{pmatrix}
=
\begin{pmatrix} Q_1\underline{Y}_1 \,|\, 4 \\ \vdots \\ Q_i\underline{Y}_i \,|\, 4 \\ \vdots \\ Q_N\underline{Y}_M \,|\, 4 \\ Q\underline{y} \end{pmatrix} \qquad \ldots(A2.5)$$

where the U's are upper triangular matrices and $\underline{y}$ is the vector formed from the remainders of the vectors $Q_i\underline{y}_i$.

Hence $\delta\hat{\underline{s}}$ is estimated by solving:

$$\begin{array}{c c} & \begin{array}{cc} N-1 & 1 \end{array} \\ \begin{array}{c} N-1 \\ 1 \end{array} & \left(\begin{array}{c|c} U_L & \underline{d} \\ & 0 \end{array}\right) \end{array} \delta\hat{\underline{s}} \quad = \quad Q\underline{y}|_{N-1} \quad = \quad \underline{y}' \qquad \text{...(A2.6)}$$

which is an undetermined system.

Write $\delta\hat{\underline{s}}$ as $\begin{pmatrix} \underline{r}_0 \\ r_e \end{pmatrix} \begin{array}{c} (N-1) \\ (1) \end{array}$

So we solve: $\quad U_L\underline{r}_0 + r_e\underline{d} \quad = \quad \underline{y}'$

And we impose the constraint:

$$\sum_{j=1}^{N-1} r_{0j} + r_e \quad = \quad S \qquad \text{...(A2.7)}$$

where S is arbitrary. This necessary introduction of a constraint will require a reworking of the variance computation, but the result will be similar. We first find the operator L of (A2.3).

From the constraint:

$$\sum_{j=1}^{N-1} \{U_L^{-1}(\underline{y}' - r_e\underline{d})\}_j + r_e \quad = \quad S$$

$$\rightarrow \quad \sum_{j=1}^{N-1} (U_L^{-1}\underline{y}')_j - r_e(\sum_{j=1}^{N-1} (U_L^{-1}\underline{d})_j - 1) \quad = \quad S$$

$$\rightarrow \qquad r_e \quad = \quad \frac{\sum_{j=1}^{N-1} (U_L^{-1}\underline{y}')_j - S}{\sum_{j=1}^{N-1} (U_L^{-1}\underline{d})_j - 1} \qquad \text{...(A2.8)}$$

Let D be the denominator in *(A2.8)* and let $\underline{1}$ be a column of ones so that $\Sigma x_j \equiv \underline{1}^T\underline{x}$, whence:

$$\underline{r}_0 = U_L^{-1}\{\underline{y}' - \underline{d}(\frac{\underline{1}^T U_L^{-1}\underline{y}' - S}{D})\}$$

$$= U_L^{-1}\{I - \frac{\underline{d}}{D}(\underline{1}^T U_L^{-1})\}\underline{y}' + (U_L^{-1}\underline{d})\frac{S}{D}$$

$$= \{I - (\frac{U_L^{-1}\underline{d}}{D})\underline{1}^T\}U_L^{-1}\underline{y}' + (U_L^{-1}\underline{d})\frac{S}{D} \quad \ldots(A2.9)$$

Rewriting *(A2.8)*:

$$r_e = \frac{\underline{1}^T}{D}U_L^{-1}\underline{y}' - \frac{S}{D}$$

allows us to write:

$$\delta\underline{\hat{s}} = \begin{pmatrix} \underline{r}_0 \\ r_e \end{pmatrix} = \begin{pmatrix} I - (U_L^{-1}\underline{d})\frac{\underline{1}^T}{D} \\ \hdashline \frac{\underline{1}^T}{D} \end{pmatrix} U_L^{-1}\underline{y}' + \frac{S}{D}\begin{pmatrix} U_L^{-1}\underline{d} \\ \hdashline -1 \end{pmatrix} \ldots(A2.10)$$

i.e. $$\delta\underline{\hat{s}} = L\underline{y}' + \underline{k} \quad \ldots(A2.10a)$$

Therefore: $$\text{var}\ (\delta\underline{\hat{s}}) = E\{(\delta\underline{\hat{s}} - \delta\underline{s})(\delta\underline{\hat{s}} - \delta\underline{s})^T\}$$

$$= E\{(L\underline{y}' - (\delta\underline{s}-\underline{k}))(L\underline{y}' - (\delta\underline{s}-\underline{k}))^T\}$$

$$= L\{\sigma^2 I + (U_L \,|\, \underline{d})\delta\underline{s}\delta\underline{s}^T(U_L \,|\, \underline{d})^T\}L^T$$

$$- (\delta\underline{s} - \underline{k})(L(U_L \,|\, \underline{d})\delta\underline{s}) - L(U_L \,|\, \underline{d})\delta\underline{s}(\delta\underline{s} - \underline{k})^T$$

$$+ (\delta\underline{s} - \underline{k})(\delta\underline{s} - \underline{k})^T \quad \ldots(A2.11)$$

It remains to show that:

$$L(U_L \,|\, \underline{d})\delta\underline{s} = (\delta\underline{s} - \underline{k}) \quad \ldots(A2.12)$$

Expanding the LHS, we get:

$$\left(\begin{array}{c} I - (U_L^{-1}\underline{d})\frac{\underline{1}^T}{D} \\ \hline \frac{\underline{1}^T}{D} \end{array}\right) U_L^{-1}(U_L \,\vert\, \underline{d})\delta\underline{s}$$

$$= \left(\begin{array}{c|c} I - (U_L^{-1}\underline{d})\frac{\underline{1}^T}{D} & U_L^{-1}\underline{d} - (U_L^{-1}\underline{d})\frac{\underline{1}^T}{D} U_L^{-1}\underline{d} \\ \hline \frac{\underline{1}^T}{D} & \frac{\underline{1}^T}{D}(U_L^{-1}\underline{d}) \end{array}\right)\delta\underline{s}$$

(remembering that $\sum_{j=1}^{N} \delta s_j = S$):

$$= \left(\begin{array}{c} \delta\underline{s}_{N-1} - \frac{U_L^{-1}\underline{d}(S - \delta s_N)}{D} + U_L^{-1}\underline{d}\delta s_N - \frac{U_L^{-1}\underline{d}\underline{1}^T U_L^{-1}\underline{d}\delta s_N}{D} \\ \hline \frac{S - \delta s_N}{D} + \frac{\underline{1}^T}{D}(U_L^{-1}\underline{d})\delta s_N \end{array}\right)$$

$$= \left(\begin{array}{c} \delta\underline{s}_{N-1} + \delta s_N U_L^{-1}\underline{d}\,\{\frac{1}{D} + 1 - \frac{(D+1)}{D}\} - \frac{S}{D}U_L^{-1}\underline{d} \\ \hline \delta s_N\{-\frac{1}{D} + \frac{D+1}{D}\} + \frac{S}{D} \end{array}\right)$$

$$= \delta\underline{s} - \underline{k}$$

Hence:

$$\text{var}\,(\delta\hat{\underline{s}}) = \sigma^2 LL^T \qquad \ldots(A2.13)$$

with L as given in $(A2.10)$.

COMPUTATION OF VAR $(\delta\hat{x}_i)$:

From *(A2.5)*, $\delta\underline{\hat{x}}_i$ is obtained from:

$$U_i\delta\underline{\hat{x}}_i + Q_i|_4\delta\underline{\hat{s}} = Q_iY_i|_4 \qquad \ldots(A2.14)$$

Hence:

$$\begin{aligned}\delta\underline{\hat{x}}_i &= U_i^{-1}\{Q_iY_i|_4 - Q_i|_4(L\underline{y}' + \underline{k})\}\\ &= U_i^{-1}(Q_i|_4)\underline{Y}_i - U_i^{-1}Q_i|_4L\underline{y}' + \underline{k}'\\ &= (U_i^{-1}Q_i|_4 \;\vdots\; -U_i^{-1}Q_i|_4L)\begin{pmatrix}\underline{Y}_i\\ \underline{y}'\end{pmatrix} + \underline{k}' \qquad \ldots(A2.15)\end{aligned}$$

Therefore:

$$\begin{aligned}\text{var }(\delta\underline{\hat{x}}_i) &= \sigma^2(U_i^{-1}Q_i|_4 \;\vdots\; -U_i^{-1}Q_i|_4L)\begin{pmatrix}(U_i^{-1}Q_i|_4)^T\\ \hline (-U_i^{-1}Q_i|_4L)^T\end{pmatrix}\\ &= \sigma^2(U_i^{-1}U_i^{-1T} + U_i^{-1}Q_i|_4LL^TQ_i|_4{}^TU_i{}^{-1T}) \qquad \ldots(A2.16)\end{aligned}$$

where L is the operator of *(A2.10)*, *(A2.10a)*.

COV $(\delta\hat{x}_i, \delta\hat{s})$:

From *(A2.15)* and *(A2.10)* and *(A2.14)*:

$$\begin{aligned}\text{cov }(\delta\hat{x}_i, \delta\hat{s}) &= \sigma^2(U_i^{-1}Q_i|_4 \;\vdots\; -U_i^{-1}Q_i|_4L)(\underset{\sim}{0} \;\vdots\; L)^T\\ &= \sigma^2(-U_i^{-1}Q_i|_4LL^T) \qquad \ldots(A2.17)\end{aligned}$$

COV $(\delta\hat{x}_i, \delta\hat{x}_j)$:

From *(A2.15)* and *(A2.4)*:

$$\begin{aligned}\text{cov }(\delta\hat{x}_i, \delta\hat{x}_j) &= \sigma^2(\underset{\sim}{0} \;\vdots\; U_i^{-1}Q_i|_4 \;\vdots\; -U_i^{-1}Q_i|_4L)\begin{pmatrix}(U_j^{-1}Q_j|_4)^T\\ (\quad 0 \quad)^T\\ (-U_j^{-1}Q_j|_4L)^T\end{pmatrix}\\ &= \sigma^2(U_i{}^{-1}Q_i|_4LL^TQ_j|_4{}^TU_j^{-1T}) \qquad \ldots(A2.18)\end{aligned}$$

APPENDIX III

RESIDUALS FROM HOMOGENEOUS STATION METHOD LOCATIONS

GROUP 1:

	CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL	STD. ERROR
	-0.2	0.1	0.2	-0.8	0.0	1.2	-0.8	1.2
	0.3	0.2	0.1	-1.1	-0.1	1.5	-0.9	1.5
	0.0	0.0	0.5	-0.9	-0.3	1.3	-0.5	1.3
Mean	0.03	0.10	0.27	-0.93	-0.13	1.33	-0.73	
Standard Deviation	0.25	0.10	0.21	0.15	0.15	0.15	0.21	

GROUP 2:

	CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL	STD. ERROR
	0.1	0.4	0.1	-1.1	-0.1	1.4	-0.8	1.4
	-0.1	0.2	0.1	-0.7	0.4	1.0	-0.8	1.1
	0.0	0.0	0.4	-0.8	-0.3	1.1	-0.4	1.0
	0.0	0.3	0.0	-0.7	0.1	0.9	-0.7	1.0
	0.0	0.2	0.3	-1.0	-0.2	1.3	-0.7	1.3
	0.0	0.1	0.0	-0.3	0.1	0.4	-0.4	0.5
	0.0	0.3	-0.2	-0.4	0.4	0.5	-0.6	0.7
	0.9	0.8	-0.5	-1.2	-0.3	1.0	-0.6	1.5
	0.1	0.3	0.0	-0.7	0.1	0.9	-0.7	1.0
Mean	0.11	0.29	0.02	-0.77	0.02	0.94	-0.63	
Standard Deviation	0.30	0.23	0.26	0.30	0.29	0.33	0.15	

GROUP 3:

	CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL	STD. ERROR
	0.3	0.5	-0.1	-1.1	-0.2	1.2	-0.7	1.3
	0.3	0.5	-0.4	-0.7	0.2	0.7	-0.6	1.0
	0.1	0.5	-0.1	-0.9	0.1	1.2	-0.9	1.3
	-0.1	0.2	0.1	-0.7	0.2	1.1	-0.8	1.2
	-0.1	0.5	0.0	-1.0	0.0	1.3	-0.9	1.4
	-0.4	-0.2	0.4	-0.3	0.3	0.7	-0.6	0.8
	-0.3	-0.1	0.4	-0.6	0.2	1.1	-0.8	1.1
	0.0	0.2	0.3	-1.0	-0.3	1.4	-0.7	1.4
Mean	0.00	0.26	0.08	-0.79	0.06	1.09	-0.75	
Standard Deviation	0.27	0.29	0.28	0.26	0.21	0.26	0.12	

GROUP 4:

	CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL	STD. ERROR
	0.5	0.7	-0.3	-1.2	-0.2	1.1	-0.6	1.4
	-0.2	0.0	0.2	-0.4	0.1	0.6	-0.3	0.6
	0.4	0.7	-0.5	-0.8	-0.1	0.6	-0.4	1.0
	-0.1	0.2	0.1	-0.6	0.2	0.8	-0.6	0.9
	0.1	0.3	0.1	-0.8	-0.2	0.9	-0.4	0.9
	-0.1	0.2	0.3	-0.7	0.0	1.0	-0.6	1.0
	-0.3	0.0	0.3	-0.3	0.1	0.5	-0.3	0.6
	-0.3	0.1	0.3	-0.5	0.1	0.8	-0.5	0.8
	-0.1	-0.1	0.3	-0.3	-0.3	0.5	-0.0	0.5
	0.3	0.3	0.0	-0.8	-0.3	0.8	-0.3	0.9
	0.0	0.4	-0.1	-0.8	0.1	0.9	-0.6	1.0
Mean	0.02	0.22	0.06	-0.65	-0.05	0.77	-0.42	
Standard Deviation	0.28	0.27	0.27	0.27	0.18	0.20	0.19	

GROUP 5:

	CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL	STD. ERROR
	0.1	0.1	0.0	-0.3	-0.3	0.2	0.0	0.4
	0.0	0.1	0.0	-0.2	0.0	0.2	-0.1	0.2
	-0.2	0.0	0.3	-0.4	-0.2	0.5	-0.1	0.4
	-0.3	-0.3	0.3	0.2	0.2	0.0	-0.1	0.3
	-0.1	0.1	0.1	-0.2	0.0	0.3	-0.1	0.3
	-0.4	0.0	0.0	0.1	0.6	0.1	-0.4	0.6
	0.2	0.3	-0.3	-0.2	0.1	0.1	-0.1	0.4
	-0.1	0.0	0.0	0.0	0.0	0.1	-0.1	0.1
	-0.3	-0.2	0.1	0.2	0.3	0.0	0.2	0.4
	0.0	0.0	-0.1	0.2	0.1	-0.2	0.0	0.2
	0.2	0.3	-0.3	-0.2	0.0	0.0	0.0	0.3
Mean	-0.08	0.04	0.01	-0.07	0.07	0.12	-0.11	
Standard Deviation	0.20	0.18	0.20	0.22	0.24	0.18	0.11	

GROUP 6:

	CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL	STD. ERROR
	0.1	0.3	-0.4	0.0	0.2	-0.1	-0.1	0.4
	-0.1	-0.3	0.0	0.4	0.3	-0.4	0.0	0.5
	0.0	0.2	-0.2	0.0	0.2	0.0	-0.1	0.2
	-0.2	-0.7	0.5	0.5	0.1	-0.2	0.1	0.7
Mean	-0.05	-0.13	-0.03	0.23	0.20	-0.18	-0.03	
Standard Deviation	0.13	0.46	0.39	0.26	0.08	0.17	0.10	

GROUP 7:

	CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL	STD. ERROR
	0.3	-0.1	0.1	0.0	-0.3	-0.1	0.2	0.4
	-0.3	0.0	0.2	-0.1	0.1	0.3	-0.1	0.3
	-0.1	0.2	-0.1	-0.2	0.1	0.2	-0.1	0.3
Mean	-0.03	0.03	0.07	-0.10	-0.03	0.13	0.00	
Standard Deviation	0.31	0.15	0.15	0.10	0.23	0.21	0.17	

GROUP 8:

	CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL	STD. ERROR
	-0.3	0.1	0.2	-0.6	0.4	0.9	-0.8	1.0
	-0.3	0.2	0.1	-0.7	0.3	1.0	-0.6	1.0
	-0.3	-0.1	0.3	-0.4	0.4	0.8	-0.7	0.9
	-0.3	0.1	0.2	-0.6	0.6	0.9	-0.9	1.1
	-0.2	0.4	0.0	-1.2	0.3	1.5	-0.9	1.5
Mean	-0.28	0.18	0.16	-0.70	0.40	1.02	-0.78	
Standard Deviation	0.04	0.26	0.11	0.30	0.12	0.28	0.13	

MEAN RESIDUALS (STANDARD DEVIATIONS IN PARENTHESES) BY GROUPS

GROUP	CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL
1	0.03 (0.25)	0.10 (0.10)	0.27 (0.21)	-0.93 (0.15)	-0.13 (0.15)	1.33 (0.15)	-0.73 (0.21)
2	0.11 (0.30)	0.29 (0.23)	0.02 (0.26)	-0.77 (0.30)	0.02 (0.29)	0.94 (0.33)	-0.63 (0.15)
3	0.00 (0.27)	0.26 (0.29)	0.08 (0.28)	-0.79 (0.26)	0.06 (0.21)	1.09 (0.26)	-0.75 (0.12)
4	0.02 (0.28)	0.22 (0.27)	0.06 (0.27)	-0.65 (0.27)	0.05 (0.18)	0.77 (0.20)	-0.42 (0.19)
5	-0.08 (0.20)	0.04 (0.18)	0.01 (0.20)	-0.07 (0.22)	0.07 (0.24)	0.12 (0.18)	-0.11 (0.11)
6	-0.05 (0.13)	-0.13 (0.46)	-0.03 (0.39)	0.23 (0.26)	0.20 (0.08)	-0.18 (0.17)	-0.03 (0.10)
7	-0.03 (0.31)	0.03 (0.15)	0.07 (0.15)	-0.10 (0.10)	-0.03 (0.23)	0.13 (0.21)	0.00 (0.17)
8	-0.28 (0.04)	0.18 (0.26)	0.16 (0.11)	-0.70 (0.30)	0.40 (0.12)	1.02 (0.28)	-0.78 (0.13)

COMPARISON OF RESIDUALS FOR GROUP 3: HOMOGENEOUS METHOD AND SEISMOLOGICAL BULLETIN SOLUTION

CNZ	ECZ	GNZ	KRP	MNG	TNZ	WEL
Homogeneous Method						
0.3	0.5	-0.1	-1.1	-0.2	1.2	-0.7
0.3	0.5	-0.4	-0.7	0.2	0.7	-0.6
0.1	0.5	-0.1	-0.9	0.1	1.2	-0.9
-0.1	0.2	0.1	-0.7	0.2	1.1	-0.8
0.1	0.5	0.0	-1.0	0.0	1.3	-0.9
-0.4	-0.2	0.4	-0.3	0.3	0.7	-0.6
-0.3	-0.1	0.4	-0.6	0.2	1.1	-0.8
0.0	0.2	0.3	-1.0	-0.3	1.4	-0.7
Bulletin Solution						
0.6	0.0	-0.1	-1.4	0.1	1.7	-0.4
0.4	-0.5	-1.5	-0.4	0.4	1.7	-0.1
0.1	0.4	-0.1	-1.3	0.5	1.0	-0.6
0.7	-0.5	-0.3	0.0	0.0	1.0	-1.2
-1.0	0.7	-0.1	-2.7	1.0	1.0	0.8
-0.3	-0.4	0.1	0.1	0.4	0.9	-0.5
-0.1	-1.1	-0.9	-0.3	-0.4	0.3	2.3
-0.7	0.8	0.6	-1.8	0.2	0.7	0.3

APPENDIX IV

MAXIMUM LIKELIHOOD ESTIMATE OF A PLANE (ORTHOGONAL LEAST SQUARES)

Given a set of N points $\underline{x}_i^T = (x_{i1}, x_{i2}, x_{i3})$ which lie in a plane:

$$\underline{\alpha}^T(\underline{x}_i - \underline{x}_0) = 0 \qquad \text{...(A4.1)}$$

and a set of estimates $\hat{\underline{x}}_i$ of the points $\underline{x}_i$, the problem is to estimate $\underline{\alpha}$, $\underline{x}_0$. $\underline{x}_0$ is arbitrary to the extent that any point in the plane will serve as a reference point or origin. $\underline{\alpha}^T = (\alpha_1, \alpha_2, \alpha_3)$ is the (unit) vector perpendicular to the plane.

Define $\underline{\varepsilon}_i$ by:

$$\underline{x}_i = \hat{\underline{x}}_i + \underline{\varepsilon}_i$$

where $\underline{\varepsilon}_i$ is a normal random variable with variance V and mean $\underline{\mu}$. Let $\underline{\varepsilon}_i' = \underline{\varepsilon}_i - \underline{\mu}_i$. Hence:

$$\underline{x}_i = \underline{x}_i + \underline{\mu} + \underline{\varepsilon}_i'$$

and

$$\begin{aligned} \underline{\alpha}^T(\underline{x}_i - \underline{x}_0) &= \underline{\alpha}^T(\hat{\underline{x}}_i - (\underline{x}_0 - \mu)) + \underline{\alpha}^T\underline{\varepsilon}_i' \\ &= 0 \qquad \text{...(A4.2)} \end{aligned}$$

Thus the difference between $\underline{x}_0$ and $\underline{\mu}$ rather than $\underline{x}_0$ will have to be estimated, as is intuitively obvious, since the mean location error $\underline{\mu}$ cannot be estimated from the $\hat{x}_i$.

We have reduced the problem to a one-dimensional problem. $\underline{\alpha}^T \underline{\varepsilon}_i'$ is a normal random variable with zero mean and variance $\underline{\alpha}^T V \underline{\alpha} = \sigma^2$. We construct the likelihood function:

$$l = \prod_{i=1}^{N} \frac{1}{\sqrt{2\Pi}} \frac{1}{\sigma} \exp \{-\tfrac{1}{2}\underline{\alpha}^T (\hat{\underline{x}}_i - \underline{x}_0')^2 / \sigma^2\} \quad ..(A4.3)$$

where $\underline{x}_0' = \underline{x}_0 - \mu$. We note that $\underline{\alpha}^T(\hat{\underline{x}}_i - \underline{x}_0')^2$ can be written:

$$\alpha^T (\hat{\underline{x}}_i - \underline{x}_0')(\hat{\underline{x}}_i - \underline{x}_0')^T \underline{\alpha}$$

Let $L = \log l$.

Define X^T to be the $3\times N$ matrix, the i^{th} column of which is $(\hat{\underline{x}}_i - \underline{x}_0')$. Define $\underline{X}_j$ to be the j^{th} column of X.

To estimate $\underline{\alpha}$, $\underline{x}_0'$, σ we wish to solve the equation obtained by setting $\frac{\partial}{\partial \alpha_j}, \frac{\partial}{\partial x_{0j}}, \frac{\partial}{\partial \sigma} L = 0$.

Rewriting:

$$L = -N \log \sqrt{2\Pi\sigma} + \sum_{i=1}^{N} - \tfrac{1}{2}\underline{\alpha}^T (\hat{\underline{x}}_i - \underline{x}_0')(\underline{x}_i - \underline{x}_0')^T \underline{\alpha} / \sigma^2$$

$$= -N \log \sqrt{2\Pi} - N \log \sigma - \frac{1}{2\sigma^2}\underline{\alpha}^T X^T X \underline{\alpha} \quad ...(A4.4)$$

Since $\underline{\alpha}$ is a unit vector, the solution obtained must satisfy the constraint $\underline{\alpha}^T \underline{\alpha} = 1$. By the method of Lagrange multipliers, we replace L by the Lagrangian function:

$$\zeta = L + \frac{\lambda}{2}(\underline{\alpha}^T \underline{\alpha} - 1)$$

$$\frac{\partial}{\partial \alpha_j}\zeta = \frac{\partial L}{\partial \alpha_j} + \lambda \alpha_j$$

$$= -\frac{1}{2\sigma^2}\{\underline{X}_{.j}^T X\underline{\alpha} + \underline{\alpha}^T X^T \underline{X}_{.j}\} + \lambda\alpha_j$$

$$= -\frac{1}{\sigma^2}\{\underline{X}_{.j}^T X\underline{\alpha}\} + \lambda\alpha_j \qquad \ldots(A4.5)$$

(since each term in the brackets is a scalar and hence equal to its transpose).

A. $\frac{\partial\zeta}{\partial\alpha_j} = 0$ for $j = 1,2,3$ implies:

$$-\frac{1}{\sigma^2}\underline{X}_{.j}^T X\underline{\alpha} + \lambda\alpha_j = 0 \qquad j=1,2,3 \qquad \ldots(A4.6)$$

that is:

$$-\frac{1}{\sigma^2}X^T X\underline{\alpha} + \lambda\underline{\alpha} = 0$$

or:

$$X^T X\underline{\alpha} = \sigma^2\lambda\underline{\alpha} \qquad \ldots(A4.7)$$

showing that $\underline{\alpha}$ is an eigenvector of $X^T X$.

B. $\frac{\partial\zeta}{\partial x'_{0j}} = -\frac{1}{\sigma^2}\{\alpha_j(-1)X^T\underline{\alpha}\}$

$$= \frac{1}{\sigma^2}\alpha_j \begin{bmatrix} \sum_i (x_{i1} - x'_{01}) \\ \sum_i (x_{i2} - x'_{02}) \\ \sum_i (x_{i3} - x'_{03}) \end{bmatrix} \underline{\alpha} \qquad j=1,2,3 \quad \ldots(A4.8)$$

So, by inspection, $\frac{\partial\zeta}{\partial x'_{0j}} = 0 \quad j=1,2,3$ if:

$$x'_{0j} = \left(\sum_{i=1}^{N} x_{ij}\right)/N \qquad \ldots(A4.9)$$

C. $\frac{\partial\zeta}{\partial\sigma} = -\frac{N}{\sigma} + \frac{1}{\sigma^3}\underline{\alpha}^T X^T X\underline{\alpha}$. $\frac{\partial\zeta}{\partial\sigma} = 0$ implies:

$$\sigma^2 = \frac{1}{N}\underline{\alpha}^T X^T X\underline{\alpha} \qquad \text{...(A4.10)}$$

which, multiplying *(A4.7)* through by $\frac{1}{N}\underline{\alpha}^T$, and noting that $\underline{\alpha}^T\underline{\alpha}=1$, gives:

$$\sigma^2 = \frac{\sigma^2\lambda}{N} \qquad \text{...(A4.11)}$$

or $\lambda = N$. $L_{\max}$ is now given by:

$$L_{\max} = -N \log \sqrt{2\Pi} - N \log \sigma - \frac{N}{2} \qquad \text{...(A4.12)}$$

Since $\lambda\sigma^2 = N\sigma^2$ is an eigenvalue of $X^T X$, and σ should be a minimum for L to be a maximum, $\underline{\alpha}$ must be the eigenvector associated with the minimum eigenvalue.

Thus the problem is reduced to the mechanical one of finding the eigenvalues and eigenvectors of the 3×3 matrix $X^T X$. The plane is described by the eigenvector corresponding to the minimum eigenvalue. Note that the sum of squares of distances from the plane is given by $\sum_{i=1}^{N} \underline{\alpha}^T(\underline{x}_i-\underline{x}'_0)^2$

$$= \underline{\alpha}^T X^T X\underline{\alpha}$$

$$= N\sigma^2$$

Hence this plane has the property that the sum of squares of deviations perpendicular to the plane is a minimum.

APPENDIX V

PRELIMINARY RESULTS: GROUP 24

(Uniform Jeffreys-Bullen Mantle Model)

TABLE 1

Hypocentre Estimates (and standard deviations)

SERIAL NO.*	SEISMOLOGICAL OBSERVATORY SOLUTION					JHD SOLUTION				
	Origin M	Time S	Lat. (°S)	Long. (°E)	Depth (km)	Origin M	Time S	Lat. (°S)	Long. (°E)	Depth (km)
33/1974	18	33.8	-38.57	175.44	176	18	36.9	-38.40	175.33	137
							0.5	0.04	0.06	6
34/1974	56	38.4	-38.51	175.72	163	56	41.4	-38.31	175.60	128
							0.5	0.04	0.06	6
40/1974	19	34.2	-38.60	175.75	199	19	37.6	-38.33	175.49	154
							0.4	0.05	0.07	6
188/1974	4	44.1	-38.52	175.84	169	4	45.9	-38.33	175.60	147
							0.5	0.05	0.07	6
190/1974	39	1.1	-38.10	176.08	198	39	2.7	-37.85	175.81	170
							0.6	0.05	0.06	7
281/1974	43	7.1	-38.35	175.96	157	43	8.0	-38.13	175.74	139
							0.5	0.05	0.06	6
485/1974	17	7.5	-38.16	176.35	175	17	9.6	-38.03	176.01	147
							0.5	0.05	0.07	6
498/1974	18	24.6	-38.12	176.46	176	18	27.4	-37.91	176.18	138
							0.5	0.04	0.07	6
499/1974	39	25.5	-38.66	175.64	181	39	28.2	-38.40	175.40	145
							0.4	0.04	0.06	6
743/1974	40	9.8	-38.38	176.01	193	40	12.8	-38.15	175.81	156
							0.6	0.05	0.06	6
18/1973	51	31.2	-38.51	175.71	190	51	33.9	-38.22	175.48	143
							0.5	0.06	0.08	7
104/1973	36	46.1	-38.07	176.23	183	36	47.5	-37.83	176.05	161
							0.5	0.05	0.07	6
202/1973	28	56.7	-38.64	175.85	187	28	59.6	-38.37	175.57	153
							0.6	0.05	0.06	6
456/1973	9	23.9	-38.65	175.69	182	9	27.1	-38.40	175.49	147
							0.6	0.05	0.06	6
529/1973	40	57.2	-38.66	175.72	190	41	0.4	-38.40	175.40	148
							0.5	0.04	0.06	6
534/1973	10	32.3	-38.62	175.79	169	10	34.9	-38.42	175.54	134
							0.5	0.04	0.06	6
537/1973	3	9.1	-38.49	175.90	167	3	10.5	-38.30	175.67	145
							0.5	0.05	0.06	6
548/1973	44	59.4	-38.58	175.79	187	45	2.7	-38.31	175.58	151
							0.5	0.05	0.07	6
602/1973	41	19.0	-38.57	175.63	191	41	21.1	-38.34	175.35	157
							0.5	0.05	0.07	6
605/1973	33	0.8	-38.24	176.09	174	33	3.4	-38.04	175.88	141
							0.5	0.05	0.07	6
680/1973	4	6.4	-38.54	175.79	188	4	10.2	-38.31	175.61	145
							0.4	0.05	0.07	6

(For * see page 208)

(Table 1 is continued on the next page)

TABLE 1: Hypocentre Estimates (and standard deviations) *(continued)*

SERIAL NO.*	SEISMOLOGICAL OBSERVATORY SOLUTION					JHD SOLUTION				
	Origin M	Time S	Lat. (°S)	Long. (°E)	Depth (km)	Origin M	Time S	Lat. (°S)	Long. (°E)	Depth (km)
694/1973	18	28.9	-38.56	175.81	182	18	31.6	-38.31	175.58	150
							0.5	0.05	0.07	6
706/1973	45	1.1	-38.43	175.82	186	45	3.1	-38.20	175.60	160
							0.6	0.05	0.07	6
16/1972	53	6.1	-38.31	176.13	172	53	8.0	-38.10	175.87	143
							0.6	0.05	0.06	6
116/1972	17	20.5	-38.20	175.77	198	17	22.2	-37.92	175.63	172
							0.5	0.05	0.06	6
215/1972	2	17.8	-38.50	176.23	154	2	21.5	-38.29	175.96	110
							0.4	0.04	0.06	6
236/1972	12	49.5	-38.54	175.79	191	12	52.1	-38.31	175.53	159
							0.6	0.05	0.07	6
241/1972	43	4.8	-37.97	176.47	168	43	6.1	-37.82	176.17	147
							0.5	0.05	0.07	6
378/1972	41	40.7	-38.34	175.89	196	41	44.1	-38.11	175.72	153
							0.6	0.05	0.06	6
580/1972	19	21.0	-37.99	176.78	152	19	21.8	-37.78	176.55	139
							0.5	0.05	0.07	6
582/1972	51	58.8	-38.39	176.20	153	52	0.7	-38.22	175.88	128
							0.4	0.04	0.07	6
593/1972	28	43.0	-38.61	176.09	161	28	47.9	-38.41	175.81	103
							0.5	0.04	0.05	6
612/1972	28	48.5	-38.39	176.04	183	28	51.4	-38.12	175.80	143
							0.4	0.05	0.08	6
613/1972	53	18.1	-38.28	176.13	176	53	21.1	-38.06	175.96	140
							0.6	0.05	0.06	6
615/1972	34	36.2	-38.25	176.11	168	34	37.8	-38.09	175.88	143
							0.5	0.05	0.07	6
671/1972	56	4.9	-38.11	176.38	167	56	7.0	-37.92	176.22	144
							0.5	0.05	0.07	6
62/1971	3	6.0	-38.27	175.71	193	3	8.2	-38.03	175.61	160
							0.6	0.05	0.07	6
119/1971	1	38.0	-38.17	176.44	176	1	41.0	-37.97	176.26	139
							0.5	0.04	0.07	6
124/1971	36	54.2	-38.69	175.59	178	36	58.0	-38.43	175.38	129
							0.6	0.04	0.06	6
199/1971	17	46.6	-38.36	175.68	182	17	48.1	-38.12	175.51	151
							0.6	0.05	0.07	6
203/1971	14	55.5	-38.21	176.31	179	14	58.6	-38.02	176.06	144
							0.4	0.05	0.09	6
208/1971	21	35.3	-38.35	175.82	167	21	36.5	-38.15	175.57	140
							0.6	0.05	0.06	6
459/1971	55	38.1	-38.43	176.20	152	55	41.5	-38.24	175.97	114
							0.5	0.04	0.06	6
461/1971	33	46.2	-38.31	176.02	179	33	49.3	-38.10	175.79	140
							0.5	0.05	0.06	6
471/1971	21	10.5	-38.42	175.91	180	21	12.7	-38.18	175.68	148
							0.5	0.05	0.06	6
600/1971	40	7.2	-38.03	176.72	175	40	10.6	-37.82	176.49	134
							0.5	0.05	0.07	7

(Table 1 is continued on the next page)

TABLE 1: Hypocentre Estimates (and standard deviations) *(continued)*

SERIAL NO.*	SEISMOLOGICAL OBSERVATORY SOLUTION					JHD SOLUTION				
	Origin Time M	S	Lat. (°S)	Long. (°E)	Depth (km)	Origin Time M	S	Lat. (°S)	Long. (°E)	Depth (km)
612/1971	53	3.3	-38.45	175.89	180	53	6.4	-38.22	175.74	139
							0.5	0.05	0.06	6
631/1971	20	3.8	-38.58	175.74	168	20	6.2	-38.35	175.56	135
							0.5	0.05	0.07	6
641/1971	29	29.0	-38.48	176.36	159	29	31.3	-38.24	176.10	132
							0.5	0.04	0.06	6
47/1969	55	58.0	-38.20	176.34	173	56	0.1	-37.98	176.13	145
							0.4	0.05	0.08	6
48/1969	45	41.1	-38.67	175.76	174	45	43.5	-38.46	175.51	140
							0.5	0.04	0.06	6
202/1969	52	11.8	-38.30	176.05	190	52	14.7	-38.06	175.88	155
							0.4	0.05	0.08	6
288/1969	25	16.8	-38.54	175.89	185	25	19.5	-38.31	175.68	148
							0.5	0.05	0.06	6
364/1969	57	29.8	-38.47	175.62	199	57	33.0	-38.17	175.45	152
							0.5	0.05	0.07	6
519/1969	4	13.9	-38.28	176.13	180	4	15.9	-38.02	175.89	148
							0.5	0.05	0.07	6
550/1969	13	47.5	-38.16	176.27	174	13	48.8	-37.94	176.05	149
							0.6	0.05	0.07	6
655/1969	41	6.0	-38.22	176.40	154	41	7.5	-38.05	176.16	132
							0.6	0.05	0.06	6
720/1969	0	19.4	-37.96	176.36	169	0	20.5	-37.73	176.11	147
							0.5	0.05	0.08	7
778/1969	57	37.3	-38.56	176.31	158	57	40.3	-38.30	176.09	122
							0.5	0.04	0.06	6
156/1968	32	57.6	-38.57	175.69	182	33	0.6	-38.36	175.44	145
							0.5	0.04	0.06	6
171/1968	35	32.0	-38.48	176.02	179	35	34.5	-38.24	175.83	144
							0.5	0.04	0.07	6
548/1968	25	33.8	-38.01	176.53	193	25	34.5	-37.60	176.22	163
							0.7	0.05	0.07	7
625/1968	45	46.3	-38.60	175.70	162	45	47.8	-38.36	175.42	132
							0.6	0.04	0.06	6
454/1967	34	6.9	-38.56	176.02	176	34	9.9	-38.28	175.72	137
							0.6	0.04	0.06	6
601/1965	19	24.1	-38.38	175.89	196	19	26.6	-38.08	175.58	155
							0.7	0.05	0.07	6

* Serial Numbers are not always exactly the same as Seismological Observatory Bulletin numbers due to a different sequencing process, but are always within ±3.

TABLE 2

STATION TERM/HYPOCENTRE CORRELATIONS FOR 190/1974

STATIONS	ORIGIN TIME	LATITUDE	LONGITUDE	DEPTH
CNZ P	-0.01	-0.69	-0.32	-0.00
ECZ P	-0.33	0.40	0.61	0.08
ECZ S	0.01	0.62	0.42	-0.11
GNZ P	-0.13	0.70	0.24	0.25
GNZ S	0.34	0.89	-0.09	-0.01
KRP P	-0.29	-0.66	0.74	-0.53
KRP S	0.01	-0.53	0.55	-0.70
WTZ P	-0.23	0.16	0.78	-0.31
WTZ S	0.19	0.37	0.47	-0.51
MNG P	-0.09	-0.17	-0.62	0.59
MNG S	0.34	0.04	-0.87	0.31
TNZ P	-0.28	-0.89	0.05	-0.03
WEL P	-0.20	-0.29	-0.51	0.59
WEL S	0.19	-0.11	-0.79	0.39
TRZ P	0.14	0.58	-0.65	0.61
TRZ S	0.63	0.67	-0.82	0.14
GBZ P	-0.55	-0.54	0.91	-0.27
TUA P	0.06	0.80	0.09	0.18
TUA S	0.59	0.86	-0.33	-0.15
COB P	-0.38	-0.66	-0.18	0.40
COB S	-0.07	-0.56	-0.46	0.26

TABLE 3

FINAL ITERATION: ITERATION NO. 6

STATIONS	CNZ P	ECZ P	ECZ S	GNZ P	GNZ S	KRP P	KRP S
No. Readings	65	44	37	63	62	65	64
Last Increment	-0.01	0.04	0.03	0.02	0.01	0.00	-0.02
Station Term	1.10	-1.63	-1.74	-2.15	-4.27	2.77	5.00
Standard Deviation	0.23	0.53	0.85	0.37	0.67	0.56	0.96
Average Travel Time	25.08	37.91	66.53	33.25	58.66	20.42	36.10
Sum of Squared Weighted Residuals	5.46	3.66	4.24	4.34	5.01	2.49	2.60
Standard Deviation of Weighted Residuals	0.29	0.29	0.34	0.26	0.29	0.20	0.20

STATIONS	WTZ P	WTZ S	MNG P	MNG S	TNZ P	WEL P	WEL S
No. Readings	38	26	62	61	45	63	65
Last Increment	0.02	0.01	-0.00	-0.04	-0.01	-0.00	-0.04
Station Term	-1.01	-1.04	-2.15	-3.83	2.76	-3.05	-4.05
Standard Deviation	0.40	0.70	0.39	0.73	0.50	0.46	0.80
Average Travel Time	24.44	43.29	41.12	72.52	28.70	50.73	89.57
Sum of Squared Weighted Residuals	1.51	2.40	8.66	5.38	4.04	4.51	4.73
Standard Deviation of Weighted Residuals	0.20	0.31	0.38	0.30	0.30	0.27	0.27

STATIONS	TRZ P	TRZ S	GBZ P	TUA P	TUA S	COB P	COB S
No. Readings	53	51	35	57	58	52	59
Last Increment	-0.00	-0.02	0.04	0.01	-0.01	-0.00	-0.03
Station Term	-0.97	-1.19	2.08	-1.33	-2.21	-2.38	-1.72
Standard Deviation	0.22	0.57	0.88	0.22	0.55	0.55	0.89
Average Travel Time	30.77	54.31	34.71	27.02	47.78	57.19	101.70
Sum of Squared Weighted Residuals	2.70	5.86	1.26	4.93	5.23	6.17	3.49
Standard Deviation of Weighted Residuals	0.23	0.34	0.19	0.30	0.30	0.35	0.25

TOTAL ERROR STANDARD DEVIATION ESTIMATE = 0.326 sec.

APPENDIX VI

FINAL RESULTS: GROUP 24

TABLE 1

STATION TERM/HYPOCENTRE CORRELATIONS FOR 33/1974*

STATIONS	ORIGIN TIME	LATITUDE	LONGITUDE	DEPTH
CNZ P	0.06	-0.36	-0.07	-0.16
ECZ P	-0.15	-0.42	0.53	-0.07
ECZ S	0.21	-0.15	0.29	-0.38
GNZ P	-0.11	0.35	-0.05	0.22
GNZ S	0.44	0.54	-0.33	-0.31
KRP P	-0.33	0.20	0.10	0.42
KRP S	0.23	0.58	-0.24	-0.07
WTZ P	-0.44	-0.03	0.34	0.44
WTZ S	-0.10	0.28	0.17	0.28
MNG P	-0.21	-0.20	-0.27	0.28
MNG S	0.36	0.18	-0.55	-0.23
TNZ P	-0.19	-0.60	0.22	0.02
WEL P	-0.31	-0.31	-0.15	0.34
WEL S	0.21	0.03	-0.51	-0.10
TRZ P	-0.03	0.26	-0.39	0.20
TRZ S	0.44	0.47	-0.51	-0.28
GBZ P	-0.43	-0.44	0.57	0.26
TUA P	-0.41	-0.53	0.11	0.35
TUA S	-0.01	-0.38	-0.20	-0.02
COB P	-0.32	-0.06	0.51	0.22
COB S	0.21	0.27	0.14	-0.23

* (Compare preliminary results, Appendix V, Table 2).

TABLE 2B

CORRELATION MATRIX FOR STATION TERMS

	CNZ P	ECZ P	ECZ S	GNZ P	GNZ S	KRP P	KRP S
CNZ P	1.00	0.26	0.27	-0.56	-0.29	-0.61	-0.57
ECZ P	0.26	1.00	0.74	-0.42	-0.46	-0.32	-0.56
ECZ S	0.27	0.74	1.00	-0.49	0.03	-0.63	-0.27
GNZ P	-0.56	-0.42	-0.49	1.00	0.24	0.70	0.54
GNZ S	-0.29	-0.46	0.03	0.24	1.00	-0.06	0.67
KRP P	-0.61	-0.32	-0.63	0.70	-0.06	1.00	0.38
KRP S	-0.57	-0.56	-0.27	0.54	0.67	0.38	1.00
WTZ P	-0.51	0.09	-0.38	0.50	-0.36	0.80	0.07
WTZ S	-0.68	-0.10	-0.18	0.50	0.17	0.58	0.45
MNG P	0.29	-0.44	-0.67	-0.00	-0.36	0.13	-0.28
MNG S	0.29	-0.60	-0.22	-0.16	0.39	-0.39	0.15
TNZ P	0.63	0.57	0.31	-0.60	-0.67	-0.46	-0.83
WEL P	0.30	-0.28	-0.63	-0.04	-0.53	0.16	-0.41
WEL S	0.38	-0.56	-0.34	-0.20	0.14	-0.34	-0.05
TRZ P	-0.18	-0.77	-0.76	0.43	0.19	0.43	0.36
TRZ S	-0.09	-0.66	-0.17	0.13	0.71	-0.16	0.55
GBZ P	0.00	0.76	0.26	-0.12	-0.71	0.18	-0.53
TUA P	0.40	0.15	-0.31	-0.25	-0.77	0.02	-0.68
TUA S	0.64	-0.03	-0.08	-0.52	-0.37	-0.50	-0.58
COB P	-0.39	0.48	0.16	0.26	-0.27	0.46	-0.00
COB S	-0.36	0.17	0.40	0.05	0.36	-0.05	0.31

	WTZ P	WTZ S	MNG P	MNG S	TNZ P	WEL P	WEL S
CNZ P	-0.51	-0.68	0.29	0.29	0.63	0.30	0.38
ECZ P	0.09	-0.10	-0.44	-0.60	0.57	-0.28	-0.56
ECZ S	-0.38	-0.18	-0.67	-0.22	0.31	-0.63	-0.34
GNZ P	0.50	0.50	-0.00	-0.16	-0.60	-0.04	-0.20
GNZ S	-0.36	0.17	-0.36	0.39	-0.67	-0.53	0.14
KRP P	0.80	0.58	0.13	-0.39	-0.46	0.16	-0.34
KRP S	0.07	0.45	-0.28	0.15	-0.83	-0.41	-0.05
WTZ P	1.00	0.56	0.01	-0.68	-0.17	0.13	-0.58
WTZ S	0.56	1.00	-0.38	-0.45	-0.56	-0.36	-0.54
MNG P	0.01	-0.38	1.00	0.36	0.27	0.87	0.57
MNG S	-0.68	-0.45	0.46	1.00	-0.11	0.20	0.77
TNZ P	-0.17	-0.56	0.27	-0.11	1.00	0.39	0.08
WEL P	0.13	-0.36	0.87	0.20	0.39	1.00	0.45
WEL S	-0.58	-0.54	0.57	0.77	0.08	0.45	1.00
TRZ P	0.10	0.04	0.54	0.38	-0.39	0.44	0.42
TRZ S	-0.53	-0.08	-0.02	0.66	-0.54	-0.21	0.49
GBZ P	0.58	0.14	-0.14	-0.80	0.49	0.06	-0.65
TUA P	0.20	-0.39	0.68	-0.09	0.67	0.81	0.19
TUA S	-0.46	-0.70	0.59	0.43	0.60	0.60	0.59
COB P	0.68	0.49	-0.42	-0.79	-0.06	-0.28	-0.79
COB S	-0.03	0.29	-0.70	-0.24	-0.38	-0.72	-0.43

(Table 4 is continued on the next page)

TABLE 2B: CORRELATION MATRIX FOR STATION TERMS *(Continued)*

	TRZ P	TRZ S	GBZ P	TUA P	TUA S	COB P	COB S
CNZ P	-0.18	-0.09	0.00	0.40	0.64	-0.39	-0.36
ECZ P	-0.77	-0.66	0.76	0.15	-0.03	0.48	0.17
ECZ S	-0.76	-0.17	0.26	-0.31	-0.08	0.16	0.40
GNZ P	0.43	0.13	-0.12	-0.25	-0.52	0.26	0.05
GNZ S	0.19	0.71	-0.71	-0.77	-0.37	-0.27	0.36
KRP P	0.43	-0.16	0.18	0.02	-0.50	0.46	-0.05
KRP S	0.36	0.55	-0.53	-0.68	-0.58	-0.00	0.31
WTZ P	0.10	-0.53	0.58	0.20	-0.46	0.68	-0.03
WTZ S	0.04	-0.08	0.14	-0.39	-0.70	0.49	0.29
MNG P	0.54	-0.02	-0.14	0.68	0.59	-0.42	-0.70
MNG S	0.38	0.66	-0.80	-0.09	0.43	-0.79	-0.24
TNZ P	-0.39	-0.54	0.49	0.67	0.60	-0.06	-0.38
WEL P	0.44	-0.21	0.06	0.81	0.60	-0.28	-0.72
WEL S	0.42	0.49	-0.65	0.19	0.59	-0.79	-0.43
TRZ P	1.00	0.38	-0.49	0.08	0.05	-0.31	-0.35
TRZ S	0.38	1.00	-0.87	-0.54	-0.04	-0.56	0.10
GBZ P	-0.49	-0.87	1.00	0.44	-0.10	0.68	-0.03
TUA P	0.08	-0.54	0.44	1.00	0.60	-0.06	-0.64
TUA S	0.05	-0.04	-0.10	0.60	1.00	-0.56	-0.55
COB P	-0.31	-0.56	0.68	-0.06	-0.56	1.00	0.26
COB S	-0.35	0.10	-0.03	-0.64	-0.55	0.26	1.00

TABLE 3

FINAL ITERATION: ITERATION NO. 7

STATIONS	CNZ P	KRP P	KRP S	TUA P	TUA S	GNZ P	GNZ S
Mantle Contrast (%)	-9	0	6	-9	-6	-9	-6
No. Readings	65	65	64	57	58	63	62
Last Increment	-0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00
Station Term	0.99	-0.43	-0.35	-0.03	0.97	-0.24	-0.20
Standard Deviation	0.07	0.19	0.38	0.08	0.26	0.13	0.23
Average Travel Time	22.67	21.09	38.93	23.19	42.04	28.79	52.06

STATIONS	ECZ P	ECZ S	MNG P	MNG S	TNZ P	WEL P	WEL S
Mantle Contrast (%)	-9	-6	-9	-6	-4	-9	-6
No. Readings	44	37	62	61	45	63	65
Last Increment	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.00
Station Term	0.20	1.95	-0.42	-0.56	0.70	0.76	-0.32
Standard Deviation	0.20	0.27	0.12	0.26	0.15	0.15	0.24
Average Travel Time	33.50	60.27	36.87	66.76	28.26	45.93	83.35

STATIONS	TRZ P	TRZ S	GBZ P	COB P	COB S	WTZ P	WTZ S
Mantle Contrast (%)	-9	-6	0	-8	-4	-9	-6
No. Readings	53	51	35	52	59	38	26
Last Increment	0.00	-0.00	0.00	-0.00	-0.00	0.00	0.00
Station Term	0.59	2.37	-2.15	-1.01	-0.91	-0.78	0.39
Standard Deviation	0.09	0.28	0.35	0.19	0.24	0.13	0.26
Average Travel Time	26.69	48.26	36.39	52.32	98.40	21.69	39.34

TOTAL ERROR STANDARD DEVIATION ESTIMATE = 0.326 (sec.)

TABLE 4

Hypocentre Estimates (and standard deviations)

SERIAL NO.	SEISMOLOGICAL OBSERVATORY SOLUTION					JHD SOLUTION				
	Origin Time M	S	Lat. (°S)	Long. (°E)	Depth (km)	Origin Time M	S	Lat. (°S)	Long. (°E)	Depth (km)
33/1974	18	33.8	-38.57	175.44	176	18	39.4	-38.53	175.52	139
							0.4	0.02	0.03	4
34/1974	56	38.4	-38.51	175.72	163	56	43.7	-38.43	175.79	128
							0.4	0.02	0.03	4
40/1974	19	34.2	-38.60	175.75	199	19	40.1	-38.46	175.69	154
							0.3	0.02	0.03	4
188/1974	4	44.1	-38.52	175.84	169	4	48.4	-38.46	175.80	146
							0.4	0.02	0.03	4
190/1974	39	1.1	-38.10	176.08	198	39	5.4	-37.96	176.04	169
							0.4	0.03	0.03	5
281/1974	43	7.1	-38.35	175.96	157	43	10.5	-38.25	175.95	137
							0.4	0.02	0.03	4
485/1974	17	7.5	-38.16	176.35	175	17	12.2	-38.15	176.23	144
							0.4	0.02	0.03	4
498/1974	18	24.6	-38.12	176.46	176	18	30.0	-38,02	176.40	133
							0.4	0.02	0.03	4
499/1974	39	25.5	-38.66	175.64	181	39	30.6	-38.53	175.60	146
							0.3	0.02	0.03	4
743/1974	40	9.8	-38.38	176.01	193	40	15.3	-38.27	176.02	155
							0.4	0.03	0.03	4
18/1973	51	31.2	-38.51	175.71	190	51	36.5	-38.35	176.68	143
							0.5	0.03	0.04	5
104/1973	36	46.1	-38.07	176.23	183	36	50.3	-37.95	176.28	158
							0.3	0.02	0.04	4
202/1973	28	56.7	-38.64	175.85	187	29	2.1	-38.50	175.77	151
							0.4	0.02	0.03	5
456/1973	9	23.9	-38.65	175.69	182	9	29.6	-38.53	175.69	147
							0.4	0.02	0.03	5
529/1973	40	57.2	-38.66	175.72	190	41	3.0	-38.53	175.60	148
							0.4	0.02	0.03	4
534/1973	10	32.3	-38.62	175.79	169	10	37.3	-38.55	175.73	132
							0.4	0.02	0.03	4
537/1973	3	9.1	-38.49	175.90	167	3	13.0	-38.42	175.87	143
							0.4	0.02	0.03	4
548/1973	44	59.4	-38.58	175.79	187	45	5.3	-38.44	175.78	151
							0.3	0.02	0.03	4
602/1973	41	19.0	-38.57	175.63	191	41	23.7	-38.47	175.54	159
							0.4	0.02	0.03	4
605/1973	33	0.8	-38.24	176.09	174	33	6.0	-38.15	176.10	140
							0.4	0.02	0.03	4
680/1973	4	6.4	-38.54	175.79	188	4	12.7	-38.44	175.81	145
							0.4	0.02	0.03	4
694/1973	18	28.9	-38.56	175.81	182	18	34.2	-38.43	175.78	149
							0.4	0.02	0.03	4
706/1973	45	1.1	-38.43	175.82	186	45	5.9	-38.32	175.81	159
							0.5	0.03	0.04	4

(Table 4 is continued on the next page)

TABLE 4: Hypocentre Estimates (and standard deviations) *(continued)*

SERIAL NO.	SEISMOLOGICAL OBSERVATORY SOLUTION					JHD SOLUTION				
	Origin M	Time S	Lat. (°S)	Long. (°E)	Depth (km)	Origin M	Time S	Lat. (°S)	Long. (°E)	Depth (km)
16/1972	53	6.1	-38.31	176.13	172	53	10.6	-38.22	176.09	141
							0.4	0.03	0.03	5
116/1972	17	20.5	-38.20	175.77	198	17	25.0	-38.04	175.85	173
							0.4	0.02	0.03	4
215*/1972	2	17.8	-38.50	176.23	154	2	24.0	-38.42	176.17	103
							0.1			
236/1972	12	49.5	-38.54	175.79	191	12	54.7	-38.44	175.74	158
							0.4	0.02	0.03	5
241/1972	43	4.8	-37.97	176.47	168	43	8.7	-37.93	176.40	143
							0.4	0.02	0.04	4
378/1972	41	40.7	-38.34	175.89	196	41	46.7	-38.23	175.93	153
							0.4	0.03	0.03	4
580/1972	19	21.0	-37.99	176.78	152	19	24.4	-37.88	176.77	131
							0.4	0.02	0.04	4
582/1972	51	58.8	-38.39	176.20	153	52	3.2	-38.34	176.09	126
							0.4	0.02	0.03	4
593/1972	28	43.0	-38.61	176.09	161	28	50.1	-38.54	176.01	97
							0.4	0.02	0.03	5
612/1972	28	48.5	-38.39	176.04	183	28	54.0	-38.24	176.02	142
							0.4	0.02	0.03	4
613/1972	53	18.1	-38.28	176.13	176	53	23.7	-38.18	176.18	136
							0.4	0.02	0.03	4
615/1972	34	36.2	-38.25	176.11	168	34	40.4	-38.21	176.09	141
							0.4	0.02	0.03	4
671/1972	56	4.9	-38.11	176.38	167	56	9.6	-38.03	176.45	139
							0.4	0.02	0.03	4
62/1971	3	6.0	-38.27	175.71	193	3	10.9	-38.15	175.82	161
							0.4	0.03	0.04	5
119/1971	1	38.0	-38.17	176.44	176	1	43.6	-38.08	176.49	132
							0.4	0.02	0.03	4
124/1971	36	54.2	-38.69	175.59	178	37	0.4	-38.56	175.57	129
							0.5	0.02	0.03	5
199/1971	17	46.6	-38.36	175.68	182	17	50.7	-38.24	175.72	152
							0.4	0.02	0.04	4
203/1971	14	55.5	-38.21	176.31	179	15	1.2	-38.14	176.29	139
							0.4	0.02	0.04	5
208/1971	21	35.3	-38.35	175.82	167	21	39.1	-38.27	175.78	141
							0.4	0.03	0.03	4
459/1971	55	38.1	-38.43	176.20	152	55	43.9	-38.37	176.17	109
							0.4	0.02	0.03	4
461/1971	33	46.2	-38.31	176.02	179	33	51.9	-38.22	176.01	139
							0.3	0.02	0.03	4
471/1971	21	10.5	-38.42	175.91	180	21	15.3	-38.30	175.89	147
							0.4	0.02	0.03	4
600/1971	40	7.2	-38.03	176.72	175	40	13.3	-37.92	176.71	126
							0.4	0.03	0.03	5
612/1971	53	3.3	-38.45	175.89	180	53	8.8	-38.35	175.94	137
							0.4	0.02	0.03	4

* MASTER EVENT

(Table 4 is continued on the next page)

TABLE 4: Hypocentre Estimates (and standard deviations) *(continued)*

SERIAL NO.	SEISMOLOGICAL OBSERVATORY SOLUTION					JHD SOLUTION				
	Origin Time M	S	Lat. (°S)	Long. (°E)	Depth (km)	Origin Time M	S	Lat. (°S)	Long. (°E)	Depth (km)
631/1971	20	3.8	-38.58	175.74	168	20	8.6	-38.48	175.75	134
							0.4	0.02	0.03	4
641/1971	29	29.0	-38.48	176.36	159	29	33.7	-38.36	176.31	126
							0.4	0.02	0.03	5
47/1969	55	58.0	-38.20	176.34	173	56	2.7	-38.10	176.36	140
							0.4	0.02	0.04	4
48/1969	45	41.1	-38.67	175.76	174	45	45.9	-38.59	175.70	139
							0.4	0.02	0.03	4
202/1969	52	11.8	-38.30	176.05	190	52	17.3	-38.18	176.09	153
							0.4	0.02	0.04	4
288/1969	25	16.8	-38.54	175.89	185	25	22.0	-38.44	175.89	146
							0.4	0.02	0.03	4
364/1969	57	29.8	-38.47	175.62	199	57	35.7	-38.29	175.65	154
							0.4	0.02	0.03	4
519/1969	4	13.9	-38.28	176.13	180	4	18.4	-38.13	176.11	146
							0.4	0.02	0.03	4
550/1969	13	47.5	-38.16	176.27	174	13	51.6	-38.06	176.27	145
							0.4	0.03	0.04	4
655/1969	41	6.0	-38.22	176.40	154	41	10.0	-38.17	176.38	129
							0.4	0.02	0.03	4
720/1969	0	19.4	-37.96	176.36	169	0	23.2	-37.84	176.34	143
							0.4	0.03	0.03	4
778/1969	57	37.3	-38.56	176.31	158	57	42.5	-38.41	176.29	117
							0.4	0.02	0.03	4
156/1968	32	57.6	-38.57	175.69	182	33	3.1	-38.49	175.64	145
							0.4	0.02	0.03	4
171/1968	35	32.0	-38.48	176.02	179	35	37.0	-38.37	176.05	141
							0.4	0.02	0.03	4
548/1968	25	33.8	-38.01	176.53	193	25	37.4	-37.70	176.46	159
							0.5	0.03	0.05	5
625/1968	45	46.3	-38.60	175.70	162	45	50.2	-38.48	175.62	133
							0.4	0.02	0.03	4
454/1967	34	6.9	-38.56	176.02	176	34	12.4	-38.40	175.92	135
							0.4	0.02	0.03	4
601/1965	19	24.1	-38.38	175.89	196	19	29.4	-38.20	175.79	154
							0.5	0.03	0.04	5

TABLE 5

Comparison of Solutions
JHD and Adams and Ware (1977)
for a Sample from Group 24

Serial No.	Month	Day	H	M	S	Lat.	Long.	Depth
33/1974	1	12	14	18	39.4 (0.4)†	-38.53 (0.02)†	175.22 (0.03)†	139 (4)†
A & W					34.4	-38.33	175.27	180
34/1974	1	12	16	56	43.7 (0.4)	-38.43 (0.02)	175.79 (0.03)	128 (4)
A & W					40.0	-38.27	175.58	159
40/1974*	1	16	13	19	40.1 (0.3)	-38.46 (0.02)	175.69 (0.03)	154 (4)
A & W					39.4	-38.38	175.64	171
ISC					36.0	-38.58	175.66	179
188/1974	3	6	17	4	48.4 (0.4)	-38.46 (0.02)	175.80 (0.03)	146 (4)
A & W					47.5	-38.33	175.72	159
190/1974	3	7	2	39	5.4 (0.4)	-37.96 (0.03)	176.04 (0.03)	169 (5)
A & W					4.8	-37.77	176.01	173
281/1974	4	9	6	43	10.5 (0.4)	-38.25 (0.02)	175.95 (0.03)	137 (4)
A & W					9.4	-38.09	175.85	152
485/1974	8	2	0	17	12.2 (0.4)	-38.15 (0.02)	176.23 (0.03)	144 (4)
A & W					10.4	-37.96	176.18	163
498/1974	8	10	16	18	30.0 (0.4)	-38.02 (0.02)	176.40 (0.03)	133 (4)
A & W					29.1	-37.84	176.38	145
499/1974	8	11	2	39	30.6 (0.3)	-38.53 (0.02)	175.60 (0.03)	146 (4)
A & W					29.9	-38.42	175.52	161
743/1974	11	12	1	40	15.3 (0.4)	-38.27 (0.03)	176.02 (0.03)	155 (4)
A & W					13.1	-38.07	175.94	177

* See Figure 6.10.

† Relative error standard deviations. Typical absolute error standard deviations are 0.5, 0.05, 0.06, 6.

APPENDIX VII

STATISTICAL SIGNIFICANCE OF A CHANGE IN STATION TERM AT A GIVEN STATION

We wish to compare the estimated model error for GNZ *P* from the 1965 earthquakes with a predicted model for GNZ *P* which is T_c seconds slower (more positive) than normal.

Let the normal model error be T_0. This is estimated by the GNZ *P* value from the years 1969, 1971, 1972, 1973, 1974; the amount is T_0'. To estimate T_0 we must add the average error $\overline{H}_0$ from all the stations (since the sum of the station terms is constrained to be zero).

So: $$T_0\ (\text{est}) = T_0' + \overline{H}_0 \qquad \ldots(A.1)$$

$$T_0' = -0.24^* \qquad \hat{\sigma}_{T_0'} = 0.13^\dagger \qquad \ldots(A.2)$$

In 1965, the GNZ *P* model error estimate is:

$$T_s' + \overline{H}_s \qquad \ldots(A.3)$$

$$T_s' = -0.36^* \qquad \hat{\sigma}_{T_s'} = 0.18^\dagger \qquad \ldots(A.4)$$

And we wish to consider the difference:

$$\delta T = T_s' + \overline{H}_s - (T_0' + \overline{H}_0 + T_c) \qquad \ldots(A.5)$$

---

* See Table 7.2.

† From JHD analysis.

which has expectation zero under the Null Hypothesis that the model travel time is $T_0 + T_c$.

$$\delta T = -0.36 - (-0.24) + (\overline{H}_s - \overline{H}_0) - T_c \quad \ldots (A\ .6)$$

We estimate $\overline{H}_s - \overline{H}_0$ with the average difference between the station terms for stations common to the two analyses (except GNZ P).

Since the hypothesis is that the model for stations other than GNZ is unchanged, the mean difference has an expected value equal to $-(\overline{H}_s - \overline{H}_0)$.

We get:

$$\overline{H}_s - \overline{H}_0 = 0.24$$

$$\hat{\sigma}_{\delta H} = 0.08$$

So:

$$\delta T = -0.12 + 0.24 - T_c$$

$$= 0.12 - T_c$$

and

$$\hat{\sigma}_{\delta T}^{\ 2} = \hat{\sigma}_{T_s'}^{\ 2} + \hat{\sigma}_{T_0'}^{\ 2} + \hat{\sigma}_{\delta H}^{\ 2}$$

so:

$$\hat{\sigma}_{\delta T} = 0.24$$

Then we want:

$$\text{Prob}\left(\left|\frac{0.12 - T_c}{0.24}\right| \geq z_\alpha\right) = \alpha$$

for $\alpha = 0.05,\ 0.1,\ 0.2$.

The corresponding z_α values are 1.65, 1.28, 0.84 (one-sided test) and corresponding T_c values are 0.52, 0.43, 0.32 sec. Thus the probability that T_c could be as great as 0.52 sec. is 0.05; that T_c could be as great as 0.43 is 0.1; that T_c could be as great as 0.32 is 0.2.

BIBLIOGRAPHY

Adams, R.D. (1963), "Source Characteristics of Some Deep New Zealand Earthquakes", *N.Z. Journal of Geology and Geophysics* 6, pp. 209-220.

Adams, R.D. and Ferris, B.G. (1976), "A Further Earthquake at Exceptional Depth Beneath New Zealand", *N.Z. Journal of Geology and Geophysics* 19, pp. 269-273.

Adams, R.D. and Lowry, M.A. (1971), "The Inangahua Earthquake Sequence, 1968", Recent Crustal Movements, Royal Society of New Zealand, *Bulletin 9*, pp. 129-135.

Adams, R.D. and Ware, D.E. (1977), "Subcrustal Earthquakes Beneath New Zealand; Locations Determined with a Laterally Inhomogeneous Velocity Model", *N.Z. Journal of Geology and Geophysics* 20, pp. 59-85.

Aki, K. and Lee, W.H.K. (1976), "Determination of Three Dimensional Velocity Anomalies Under a Seismic Array Using First P Arrival Times from Local Earthquakes. 1: Homogeneous Initial Model", *Journal Geophys. Res.* 81, pp. 4381-4399.

Ansell, J.H. and Smith, E.G.C. (1975), "Detailed Structure of A Mantle Seismic Zone Using the Homogeneous Station Method", *Nature* 253, pp. 518-520.

Bath, M. (1968), *Mathematical Aspects of Seismology*, Elsevier Publishing Co.

Billington, S. and Isacks, B.L. (1975), "Identification of Fault Planes Associated with Deep Earthquakes", *Geophys. Res. Letters* 2, pp. 63-65.

Blamey, C. and Gibbs, P.G. (1968), "The Epicentres and Origin Times of Some Large Explosions", *Geophys. Jnl. Roy. Astr. Soc.* 16, pp. 1-7.

Bolt, B.A. (1960), "The Revision of Earthquake Epicentres, Focal Depths and Origin Times using a High Speed Computer", *Geophys. Jnl. Roy. Astr. Soc.* 3, pp. 433-440.

Bolt, B.A. and Freedman, H.V. (1968), "Group Analysis of Variance for Earthquake Location and Magnitude", *Nature* 217, pp. 47-48.

Buland, R. (1976), "The Mechanics of Locating Earthquakes", *Bull. Seism. Soc. Am.* 66, pp. 173-187.

Bullen, K.E. (1965), *An Introduction to the Theory of Seismology*, 3rd ed., Cambridge University Press.

Crosson, R.S. (1976a), "Crustal Structure Modelling of Earthquake Data, 1, Simultaneous Least Squares Estimation of Hypocentre and Velocity Parameters", *J. Geoph. Res.* 81, pp. 3036-3047.

------------ (1976b), "Crustal Structure Modelling of Earthquake Data, 2, Velocity Structure of Puget Sound Region, Washington", *J. Geoph. Res.* 81, pp. 3047-3055.

Dahlquist, G. and Bjorch, A. (1974), *Numerical Methods*, Prentice-Hall.

Dewey, J.W. (1971), "Seismicity Studies with the Method of Joint Hypocentre Determination", Ph.D. thesis, University of California, Berkeley.

Douglas, A. (1967), "Joint Epicentre Determination", *Nature* 215, pp. 47-48.

Douglas, A. and Lilwall, R.C. (1968), "Does Epicentre Source Bias Exist?" *Nature* 220, pp. 469-470.

---------------------------- (1972), "Methods of Estimating Travel Times and Epicentres", *Geophys. Jnl. Roy. Astr. Soc.* 30, pp. 187-197.

Eiby, G.A. (1971), "Seismic Regions of New Zealand", Recent Crustal Movements, Royal. Soc. N.Z., *Bulletin 9*, pp. 153-160.

Engdahl, E.R. (1973), "Relocation of Intermediate Depth Earthquakes in the Central Aleutians by Seismic Ray Tracing", *Nature Physical Science* 245, pp. 23-25.

Evernden, J.F. (1969a), "Precision of Epicentres, Obtained by Small Numbers of World-Wide Stations", *Bull. Seism. Soc. Am.* 59, pp. 1365-1398.

------------- (1969b), "Identification of Earthquakes and Explosions by Use of Teleseismic Data", *J. Geophys. Res.* 74, pp. 3828-3856.

Fitch, T.J. (1975), "Compressional Velocity in the Source Regions of Deep Earthquakes", *Earth Pl. Sci. Letters* 26, pp. 156-166.

Fitch, T.J. and Muirhead, K.J. (1974), "Depths to Larger Earthquakes Associated with Crustal Loading", *Geophys. Jnl. Roy. Astr. Soc.* 37, pp. 285-296.

Fitch, T.J. and Rynn, J.W. (1976), "Inversion for V_P/V_S in Shallow Source Regions", *Geophys. Jnl. Roy. Astr. Soc* 44, pp. 253-267.

Flinn, E.A. (1965), "Confidence Regions and Error Determinations for Seismic Event Locations", *Rev. Geophys.* 3, pp. 157-185.

Freedman, H.W. (1966a), 'The "Little Variable Factor": a Statistical Discussion of the Reading of Seismograms', *Bull. Seism. Soc. Am.* 56, pp. 593-604.

Freedman, H.W. (1966b), "A Statistical Discussion of P_n Residuals from Explosions", *Bull. Seism. Soc. Am.* 56, pp. 677-695.

-------------- (1967), "A Statistical Discussion of P Residuals from Explosions", *Bull. Seism. Soc. Am.* 57, pp. 545-561.

-------------- (1968), "Seismological Measurements and Measurement Error", *Bull. Seism. Soc. Am.* 58, pp. 1261-1271.

Geiger, K. (1910), "Herdbestimmungen bei Erdbeden aus den Ankunftzeiten", *K. Ges. Wiss. Götingen* 4, pp. 331-349.

Haines, A.J. (1977), "P and S Velocities in the Upper Mantle Beneath New Zealand", M.Sc. thesis, Victoria University of Wellington, New Zealand.

Hamilton, R.M. and Gale, A.W. (1968), "Seismicity and Structure of North Island, New Zealand", *J. Geophys. Res.* 73, pp. 3859-3876.

Hamilton, R.M. and Gale, A.W. (1969), "Thickness of the Mantle Seismic Zone beneath the North Island of New Zealand", *J. Geophys. Res.* 74, pp. 1608-1613.

Hamilton, R.M. (1969), "Seismological Studies of the Gisborne Earthquake Sequence, 1966", in *Gisborne Earthquake, New Zealand, March 1966*, N.Z. Dept. of Scientific and Industrial Research Bulletin 194.

Harris, F. (1975), "Focal Mechanisms of Intermediate and Deep Earthquakes in North Island, New Zealand", M.Sc. Thesis, Victoria University of Wellington, New Zealand.

Hartley, H.O. (1961), "The Modified Gauss-Newton Method for Fitting of Non-Linear Regression Functions by Least Squares", *Technometrics* 3, pp. 269-280.

Hatherton, T. (1970a), "Upper Mantle Inhomogeneity Beneath New Zealand: Surface Manifestations", *J. Geophys. Res.* 75, pp. 269-284.

------------- (1970b), "Gravity, Seismicity, and Tectonics of North Island, New Zealand", *N.Z. J. of Geol. and Geophys.* 13, pp. 126-144.

Herrin, E., Tucker, W., Taggart, J., Gordon, D.W., Lobdel, J.L. (1968), "Estimation of Surface Focus P Travel Times", *Bull. Seism. Soc. Am.* 58, pp. 1273-1291.

Householder, A.S. (1953), *Principles of Numerical Analysis*, McGraw-Hill.

Jacob, K.H. (1972), "Global Tectonic Implications of Anomalous Seismic P Travel Times from the Nuclear Explosion Longshot", *J. Geophys. Res.* 77, pp. 2556-2573.

James, D.E., Sacks, I.S., Lazo, E.L., Aparico, P.G. (1969), "On Locating Local Earthquakes Using Small Networks", *Bull. Seism. Soc. Am.* 59, pp. 1201-1212.

Jeffreys, H. (1939), "Seismological Tables", *M.N.R.A.S.* 99, pp. 397-408.

------------ (1970), *The Earth*, 5th ed., Cambridge University Press.

Jeffreys, H. and Bullen, K.E. (1948), *Seismological Tables*, British Assn. Advancement Sci., Burlington House, London.

Keilis-Borok, V.I., Pavlova, L.G., Pyatetskii-Shaprio, I.I., Rezuyakouskii, P.T., Zhelankina, T.S. (1972), "Computer Determination of Earthquake Epicentres", in *Computational Seismology*, ed. Keilis-Borok, V.I., Consultants Bureau, New York.

Lee, W.H.K. and Lahr, J.C. (1975), "HYPO 71 (Revised). A Computer Program for Determining Hypocentre, Magnitude and First Motion Pattern of Local Earthquakes", U.S. Geol. Survey Open File Report, 75-311.

Leeds, A.R., Knopoff, L., Kausel, E.G. (1974), "Variations of Upper Mantle Structure Under the Pacific Ocean", *Science*, Vol. 186, pp. 141-143.

Lilwall, R.C. and Douglas, A. (1969), "Quest for a P travel-time standard", *Nature* 22, pp. 975-977.

Lindquist, E.F. (1971), "The Norton Study of the Effects of Non-Normality and Heterogeneity of Variance", in *Contemporary Problems in Statistics*, ed. Lieberman, B., Oxford University Press, New York.

Lomnitz, C. (1971), "What is the Accuracy of P Travel-Time Tables?" *Earth Planetary Sci. Letters* 10, pp. 193-195.

---------- (1977), "A Fast Epicentre Location Program", *Bull. Seism. Soc. Am.* 67, pp. 425-431.

Marquardt, D.W. (1963), "An Algorithm for Least Squares Estimation of Non-Linear Parameters", *J. Soc. Indust. Appl. Math.* 11, pp. 431-441.

Marsh, B.D. and Carmichael, I.S.E. (1974), "Benioff Zone Magmatism", *J. Geophys. Res.* 79, pp. 1196-1206.

Mooney, H.M. (1970a), "Upper Mantle Inhomogeneity Beneath New Zealand: Seismic Evidence", *J. Geophys. Res.* 75, pp. 285-309.

---------- (1970b), "Theoretical and Observed Travel Times for New Zealand Deep Earthquakes", *N.Z. J. Geol. and Geophys.* 13, pp. 703-717.

Reilly, W.I. (1965), "Gravity Map of New Zealand 1:4,000,000 Isostatic Anomalies", *Dept. Sci. Indus. Res.*, Wellington, New Zealand.

Ringwood, A.E. and Green, D.H. (1966), "An Experimental Investigation of the Gabro-Eclogite Transformation and Some Geophysical Implications", *Tectonophysics* 3, pp. 383-427.

Robinson, R. (1976), "Relative Teleseismic Residuals, North Island, New Zealand, and their Relation to Upper-Mantle Structure" (Letter), *Tectonophysics* 31, T41-T48.

Scheffé, H. (1959), *The Analysis of Variance*, Wiley.

Smith, E.G.C. (1976), "Scaling the Equations of Condition to Improve Conditioning" (Letter), *Bull. Seism. Soc. Am.* 66, pp. 2075-2076.

Smith, W.D. (1973), "Upper Mantle Velocities Determined from Local Observations of Deep Earthquakes", *Bull. Seism. Soc. Am.* 63, pp. 807-817.

Snedecor, G.W. and Cochran, W.G. (1967), *Statistical Methods*, 6th edition, Aimes, Iowa State University.

Sutton, D.J. (1974), "A Fall in P-Wave Velocity before the Gisborne, New Zealand, Earthquake of 1966", *Bull. Seism. Soc. Am.* 64, pp. 1501-1508.

Veith, K.F. (1975), "Refined Hypocentres and Accurate Reliability Estimates", *Bull. Seism. Soc. Am.* 65, pp. 1199-1222.

Wyss, M. (1973), "The Thickness of Deep Seismic Zones", *Nature* 242, pp. 255-256.

Zelen, M. (1962), "Linear Estimation and Related Topics", in *A Survey of Numerical Analysis*, ed. Todd, J., McGraw-Hill.