

# Assessment of Results Produced by an Algorithm Delivering Earthquake Epi- & Hypocenters Concurrently and in Real-Time

George R. Daghish<sup>1</sup> & Yurii P. Sizov<sup>2</sup>

<sup>1</sup>Independent Analysis & Computation, Partridge Close, Frimley, UK

<sup>2</sup>Schmidt Institute of Physics of the Earth, Troitsk, Moscow, Russian Federation

Correspondence: George R. Daghish, Independent Analysis & Computation, Partridge Close, Frimley, CAMBERLEY. GU16 8PD, UK. Tel: 44-012-762-6942. E-mail: georgerdaghish@gmail.com

Received: November 16, 2014 Accepted: December 4, 2014 Online Published: December 19, 2014

doi:10.5539/esr.v4n1p47

URL: <http://dx.doi.org/10.5539/esr.v4n1p47>

## Abstract

This paper is to suggest a means by which confidence limits may be placed around the Epi- & Hypocenter values that are evolved, concurrently and in real-time, on the detection of an increasing number of P-wave first arrivals, by the present algorithm.

As described in previous papers, this algorithm is table driven in that it uses an “interpolative tabular scanning process” to deliver its results. For this purpose a set of three main tables are provided: a table of travel times for rays originating from a graduated set of depth points to a given set of colatitudes; a table for a set of take-off angles corresponding to the traveltimes and a set of calibrating tables to correct numerical error found in the table generating process itself. These tables are generated by any from a set of point-to-point (P2P) ray tracers parameterized by any from a set of radial Earth velocity models {PREM, iasp91, ak135}.

The production of the confidence limits is concomitant on considering that each value (i.e. discovered Epi- & Hypocenter) is stationary, and subject to perturbation by error.

In brief, the error is considered ultimately to be normal. Therefore normal theory is used, as in Chauvenet’s test to screen the production of individual localizations for outlying data inputs. Subsequently it is used to monitor for outliers in the set of localizations themselves.

Use is then made of the t-variate (“Student’s t”) to dynamically establish confidence limits, of varying levels of significance, about regressions on the set of localizations as this set increases in real-time.

The upshot being that the algorithm can produce successive localizations of the Earthquake as data input (P-wave onset times) arises and at the same time monitor the accuracy and integrity of the solutions.

**Keywords:** earthquake, epicenter, hypocenter, concurrent localization, real-time tracking, student’s t, normal theory

## 1. Introduction

### 1.1 Aims, Uses and Properties

The aim of this paper is to demonstrate that

- The algorithm structure described below will produce results in close accord with the received values
- This overall structure is capable of assessing the error (to given levels of significance) in this output.

The authors feel that the algorithm shall be employed as a rapid and lightweight front-end for possible Earthquake Early Warning (EEW) systems. By using incoming data to rapidly evolve solutions for

- Epicenter and Hypocenter coordinates
- Take-off angles

for events in real-time, there can evolve, from an initial small set of onsets, localizations of increasing accuracy. It must be emphasized that only the timings of P-wave first arrivals are initially required from the activated stations, and this, being one of the clearest observations an automatic energy-onset detection system can make, would make it a prime candidate for inclusion in the overall software architecture of a centralized facility

supporting EEW.

The build-up to this algorithm is given in detail in references from (Daglish G. R. & Sizov Yu. P. 2011), to reference (Daglish G. R. & Sizov Yu. P. 2014. 2ECEES). In (Daglish G. R. & Sizov Yu. P. 2013) there is given a short explanation of how the algorithm appears to branch out from the main progression, being table-driven and dependent on pre-calculated tables for use in an “interpolative tabular scan”

As stated these tables are 5-fold

1. Travel times for rays from each depth point in a set of depth points to a set of specific colatitudes
2. Take-off angles for each ray so treated
3. Calibration data for later interpolation to correct error due to the ray tracer’s occasional non-convergence
4. Accuracy data monitoring the degree of convergence of each ray traced
5. Path length of each ray

These five tables are also accompanied by a Log of non-convergence which lists those points at which the P2P ray tracers did not achieve the required accuracy.

The timing for the primary (see below) algorithm is ~0.7 s/station for a set of 4 to 8 stations; ~1.3 s/station for a set of up to 30 stations and ~2.15 s/station for a set of up to 45 stations, running on a 3.2 GHz processor.

### 1.2 Data Structures and Modus Operandi

The theory alluded to in the Abstract, above, is implemented in the routine:

#### ETTST 16.cpp

This routine will evolve a set of sequential solutions or localisations as it accrues data from elements in the network surrounding the event. Each localization is accompanied by the following vector of information:

1. Number of Stations in Localisation
2. Ordinal Number for Localisation
3. Time in Main Algorithm [s]
4. Running Total of 3.
5. Delta of Total Processing Times for main algorithm, secondary algorithm , (vid. inf. 28), and statistical processing:  $(T_{i+1} - T_i)$
6. Running Total of 5.
7. 3. + 6.
8.  $\chi^2$  value
9. Significance flag: (-1 or +1)
10. Level of probability
11. Threshold for RMS Comparison on Calculation of  $\underline{\epsilon}$  (Epicenter coordinates)
12. Actual RMS from Final Calculation of  $\underline{\epsilon}$
13. Radial Discrepancy (as percentage)
14. Primary Hypocenter,  $H_d$
15. Tracked  $1^\circ H_d$  values
16.  $\delta\epsilon = |\underline{\epsilon} - \underline{\epsilon}_I|$  (Calculated Epicentre,  $\underline{\epsilon}$ , and reference Epicentre,  $\underline{\epsilon}_I$ , from IRIS)
17. Tracked  $\delta\epsilon$  value
18. Radial Discrepancy,  $\delta\rho$ , as percentage (again)  $= \frac{|\rho - R_e|}{R_e} \times 100$ .  $\rho$  is the derived Earth radius, while  $R_e$  is an Earth radius, local to the Latitude given at 20.
19. Tracked  $\delta\rho$  value
20. Latitude

21. & Longitude
22.  $\rho$  (Calculated Earth radius) - 20, 21, 22 are for the calculated  $\underline{\epsilon}$ , as in 16.
23.  $x_i$
24.  $y_i$
25.  $z_i$  - 23, 24, 25 are IRIS Cartesian coordinates in Earth space-frame for  $\underline{\epsilon}_i$  as reference.
26. Latitude
27. & Longitude for IRIS Epicentre -  $\underline{\epsilon}_i$
28. Consensus on secondary Hypocenter,  $H_d$ , from 29,30, 31 and 32, below, all produced by the secondary algorithm which runs these distinct interpolation processes in tandem
29.  $2^\circ H_d$  by linear/linear interpolation
30.  $2^\circ H_d$  by linear/cubic interpolation
31.  $2^\circ H_d$  by Cubic/Lagrange interpolation
32.  $2^\circ H_d$  by Lagrange/Lagrange interpolation
33.  $\theta$ : an angle between the current Tracked version of  $\underline{\epsilon}$  and  $\underline{\epsilon}_i$  or the prior  $\underline{\epsilon}$ .
34. Length of the spherical arc subtended by  $\theta$
35. A similar arc distance to the above, at 33 and 34, but between  $(\underline{\epsilon}_{i+1} - \underline{\epsilon}_i)$
36. Upper
37. & Lower Confidence Limits for the arc at 35.
38. Average or "Best Estimate" for 35.,  $\langle \Delta \underline{\epsilon} \rangle$
39. Upper
40. & Lower bounds for the Confidence Interval on 38.,  $\langle \Delta \underline{\epsilon} \rangle$
41. Primary  $H_d$  for this localisation (stated again)
42. Upper
43. & Lower bounds for the Confidence Interval on 41.,  $1^\circ H_d$
44. Average or "Best Estimate" for 41.,  $\langle H_d \rangle$
45. Upper
46. & Lower bounds for the Confidence Interval on 44.,  $\langle H_d \rangle$

This set of vectors, produced sequentially, is destined for the output file:

#### **ABTrackedData.txt**

As stated the main point for the Algorithm is to produce an Epicenter/Hypocenter pair together with an error assessment. In the above vector definition we can see that the output fulfills this. The program **ETTST 16.cpp** incorporates two scanning algorithms in tandem:

1. A primary algorithm which produces the primary Epicenter/Hypocenter pair.
2. A subsequent algorithm, which takes the Epicenter from the afore-mentioned pair and, knowing this, performs a scan for a secondary Hypocenter. The two Hypocenters thus found tend to agree closely.

This program performs a restricted emulation of the output of an Earthquake event in so far as it takes in data (P-wave onset times) from stations (with corresponding station coordinates) in strict sequence but does not attempt to emulate the real time buffering that would be necessary if the emulation stuck to the actual time differences between the incoming data items.

The file which contains the input data for the entire process associated with one Earthquake is:

#### **HypoCentreData.txt.**

The output from **ETTST 16.cpp**, apart from a narrative to the monitor, goes in the following directions, (these Data tables apply to a single, unique earthquake):

<b>HypoCentreResultSQ.txt</b>	(This contains all output relating to all the localisations performed, including Take-off angles)
<b>StationPolarTransformSQ.txt</b>	(This is a resume of the input subsets and traces of each scan)
<b>Combinations02Sequence.txt</b>	(This is a set of traces of the localisations made on all subsets for this Earthquake)
<b>ABTrackedData.txt</b>	(This is the set of data vectors for the localisations realised for each subset, structured as above)
<b>EpiCentrePolar04.txt</b>	(This is the set of Epicentre localisations for each subset in this Earthquake)

The above output is created from the following input, ancillary to the main table, **HypoCentreData.txt**, referred to above:

<b>EpicentrePolar03.txt</b>	(This contains the received value for the Epicentre of this Earthquake)
<b>StationLatLong.txt</b>	(Dimensions of Depth and Colatitude for the following Matrix tables)
<b>PSArrivalTabulation00.txt</b>	(Matrix of P-wave Travel-times [s] from source at Depth to a surface Colatitude)
<b>TakeOffAngleGrid.txt</b>	(Matrix of Take-off angles corresponding to the P-wave rays in the above Matrix)
<b>RayPathDistance.txt</b>	(Matrix of Ray-path lengths from source to target surface Colatitude).

## 2. Method

### 2.1 Overview of Software Action

The processing of each localization proceeds in 5 main phases (for symbols, see above):

1. The production of the next Epicenter/Hypocenter pair  $(\underline{\varepsilon}, H_d)$  by the primary algorithm
2. Tracking the  $\underline{\varepsilon}$  and the  $H_d$ ; generating the  $\delta\varepsilon$ . Also generating the parameters for the secondary algorithm; i.e. the set of relative colatitudes for each active station based on the position of the primary  $\underline{\varepsilon}$
3. The production of the secondary and consensus  $H_d$  from the four interpolation regimes (see above at 1.1)
4. The application of the external Chauvenet procedure to the series  $\{\delta\varepsilon_i\}$  and  $\{H_{d_i}\}$
5. Incorporating the results of 4. (i.e. taking into account the temporarily eliminated elements of these two latter series), confidence limits are generated for:
  - Last value of  $\delta\varepsilon$
  - “Best estimate”  $\langle\delta\varepsilon\rangle$
  - Last value of  $H_d$
  - “Best estimate”  $\langle H_d\rangle$

The application of the Chauvenet test (Neville A.M. & Kennedy J.R.1964) takes place in two situations at two levels:

In **phase 1**: to screen and eliminate if necessary data used for the primary algorithm, within that algorithm

In **phase 4 and 5**: to screen and eliminate if necessary elements of the series of localized:

$$\{\delta\varepsilon_i\}; \{H_{d_i}\}$$

and to use the resulting map of eliminations in generating:

$$\langle\delta\varepsilon\rangle; \langle H_d\rangle.$$

The means whereby the “best estimates” are produced and how the Chauvenet procedure is implemented are described at Appendices B and C.

The system is tested by taking each of 7 known Earthquakes and passing their data through the routine **ETTST 16.cpp** (see above).

Summaries of these known Earthquakes are given at Table 1.

The data used for these trials is taken, using WILBER III (supported by Incorporated Research Institutes for Seismology - IRIS). Their parameters are tabulated at Table 1.

Table 1. Earthquake parameters

Number	Region	Magnitude ( $M_w$ )	Epicentre Lat/Long (degrees)	$H_d$ (km)
1	West of North American Coast	5.4	41.7136N/126.8446W	9.9
2	Fiji Islands region	4.5	17.8997S/ 178.5269W	558.9
3	Just East of Honshu	5.4	35.621N/ 140.6862E	36.0
4	South Mid Atlantic Trench	5.2	47.1861S/13.4304W	10.0
5	Celebes Sea	4.1	3.9999N/ 123.3981E	495.1
6	Yunnan (China)	6.1	27.446N/103.427E	10.0
7	Santa Cruz Islands	4.7	12.23985S/167.135E	259.3

## 2.2 Algorithm Fine Detail

This section deals with the location procedures as mentioned at the outset and is largely paraphrased and quoted from Daghli G. R. & Sizov Yu. P. 2014. We begin with the scanning process. Although there appear to be two possible ways of scanning within this context:

1. that which makes use of a prior estimate of the coordinates of the event epicenter, here the **secondary process**
2. that which makes no such use of prior estimation, here the **primary process**.

space dictates that only the second, and more important, type will be dealt with in detail here.

The co-latitudes and longitudes of the set of active stations are formed into a set of Cartesian co-ordinates within the Earth space-frame. These are to be used later in the localisation calculation which is used to derive the Epicentral position.

The next step is the organisation of the set of P-wave first arrival times into a set of differences forming a fixed "Timing Template". Having formed these two sets of information the scan commences by:

- Interpolating an entire co-latitudinal row of timings (P-wave 1<sup>st</sup> arrivals from the tabular Structure referred to above) for the next depth-point reached in the scan.
- Laterally scanning the fixed "Timing Template" along this interpolated row to generate co-latitudes corresponding to its elements by reverse interpolation.
- The above two processes are repeated for each of the set of depth points which form the scan and the smallest local minimum of the indicator so found is taken to define the Epicentral co-ordinates,  $\underline{x}$ , and the Hypocentre Depth,  $H_d$ .

To repeat this in plainer language: an actual fixed "Time Template" for the lateral scanning procedure consists of a set of differences:

$$\delta t_j = t_j - t_0; \quad j = 0, n-1$$

The base in time for the lateral scan is defined as  $T_0$ , and the template is shifted across the depth-interpolated time row as:

$$\tau_i = T_0 + i \cdot \Delta t; \quad i = 0, N-1$$

$$t_j = \tau_i + \delta t_j; \quad j = 0, n-1$$

Here  $N$  is the granularity of the scan and  $\Delta t = \frac{T_{\max} - T_0}{N}$ .  $T_0$  and  $T_{\max}$  are the limiting values of the

depth-interpolated row. These new  $t_j$  are used to reverse interpolate to a set of values for colatitude. At each point,  $\tau_i$ , in the scan the indicator is formed. This indicator is

$$\sigma = \ln \left( \sum_{i=0}^{n-1} \frac{(v_i - c_i)^2}{c_i} \right)$$

$\sigma$  is the logarithm of a  $\chi^2$  variate and is to be minimised. The  $v_i$  are the colatitudes calculated for the trial epicenter, while the  $c_i$  are those formed from the time-template. There are  $n$  stations. The calculation for the trial epicenter is as follows: at each point in the scan, the colatitudes  $c_i$  are used to form radii (as depicted in Figure 1, below). These radii are translated and then subtended from the known station locations. A least-squares calculation for the trial epicentral co-ordinates ensues.

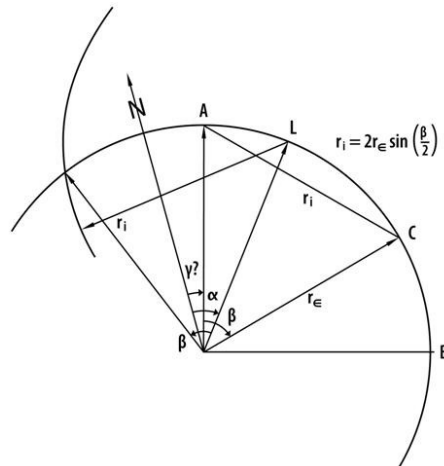


Figure 1. Geometrical basis for within-scan Epicentre location

In the figure the arc AB represents the latitudinal extent, starting at zero degrees at A, of the P-wave 1st arrival Tables. C represents a point of match for the object time position from the moving time-template, now associated with a possible minimum. The angle “ $\gamma$ ” is the unknown angle giving the immaterial orientation of the frame of the Table to the Earth space-frame. AC is a chord subtended from the point of match, C, to the “epicenter pole” of the Table at A. This chord is then translated and placed with one end at L which now subtends it as the spherical radius,  $r_i$ , of the  $i^{th}$  sphere centered at the  $i^{th}$  station colatitude, L.

$$r_i = 2r \cdot \sin\left(\frac{\beta}{2}\right)$$

Once all these radii have been assembled from the matched positions of timings within the template, then the following system of equations is solved for  $(x, y, z)_\epsilon$ , which will be a possible location for the epicenter within the Earth space-frame:

$$a_i x_\epsilon + b_i y_\epsilon + c_i z_\epsilon = \frac{1}{2} (r_i^2 - r_\epsilon^2 - (a_i^2 + b_i^2 + c_i^2)); \quad i = Zero, (n-1)$$

In the above, the  $(a, b, c)_i$  are the coordinates of the station positions within the Earth frame.  $r_\epsilon$  is a value for the Earth radius, while  $n$ , ( $\geq 3$ ), is the number of stations with which the scan is undertaken. The  $(x, y, z)_\epsilon$  is converted to Latitude and Longitude.

### 3. Results

Although there are 7 Earthquakes in the process, only the results of the 7<sup>th</sup> will be fully presented, since space does not allow for all 7. However the results for the entire set can be made available on request.

All subsequent figures 2 to 7 apply to Earthquake 07, Santa Cruz Islands. The sequences of localisations found by this algorithm generally are characterized by an initial “erratic” or “transient” period followed by a phase when the output is stabilized around a more or less constant set of values.

This is shown in Figure 3 which displays the output from the first localization using 4 stations to 12<sup>th</sup> localization, using 15 inputs.

Figure 2 shows a comparison between the RMS for residuals arising from the final Epicenter calculations and threshold values formed from a percentage of the mean of the radii also involved in the same calculation. It can be seen that the first points (5 to 15) are more erratic than those subsequent to point 15, which remain on a more regular “trajectory”. This characterizes the transient phase for this set of localizations.

The zones 5 to 15 in figures 2 and 3 are complementary with the erratic behaviour in 3 corresponding to the disturbed zone in 2.

Hereafter most of the graphs begin at about the level of 15 inputs (i.e. the 12<sup>th</sup> localization). In fact the calculations establishing the confidence intervals (see Appendix C) do not output until they have accumulated 10 points of data to form mean- and sigma-values. Because of this initial wildness of the output, the confidence intervals tend to be wider at the start.

The effect of the Chauvenet rejections, (also see Appendix D), is discussed in the next section, “Results”.

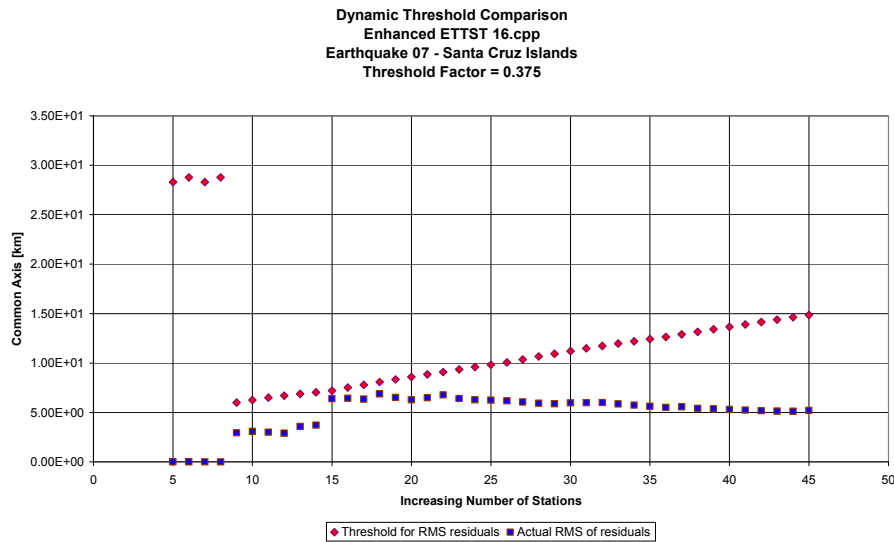


Figure 2.

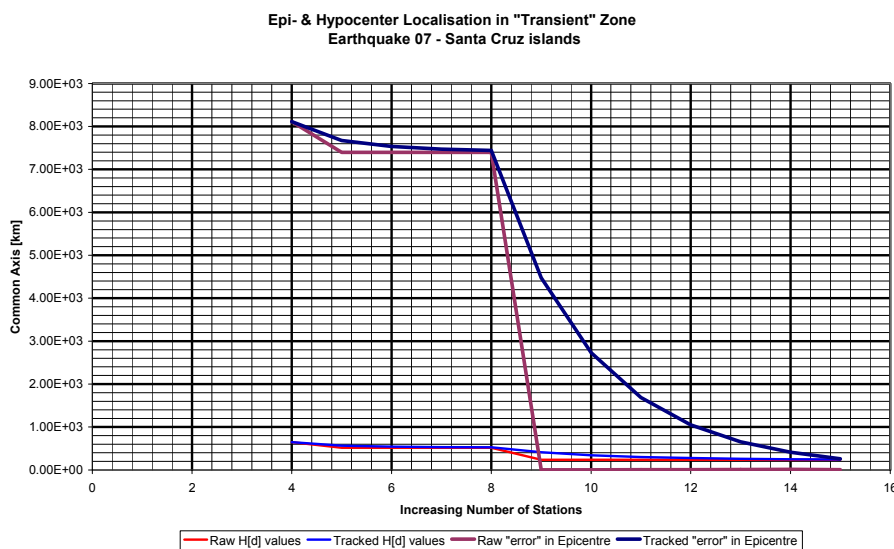


Figure 3.

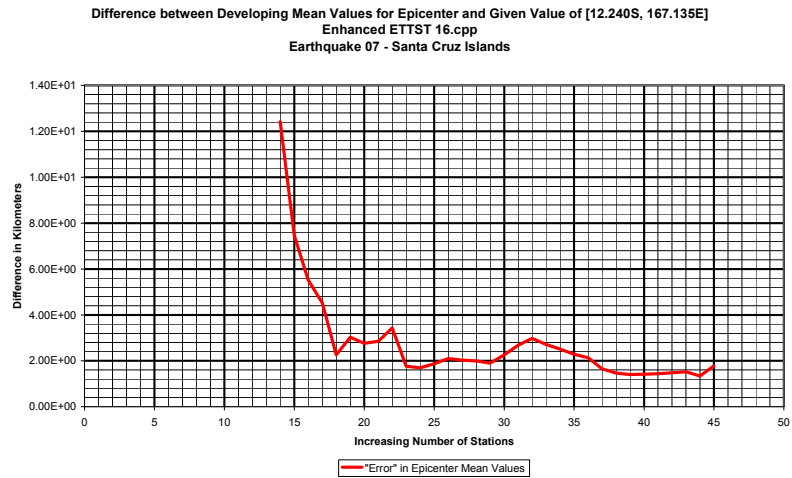


Figure 4.

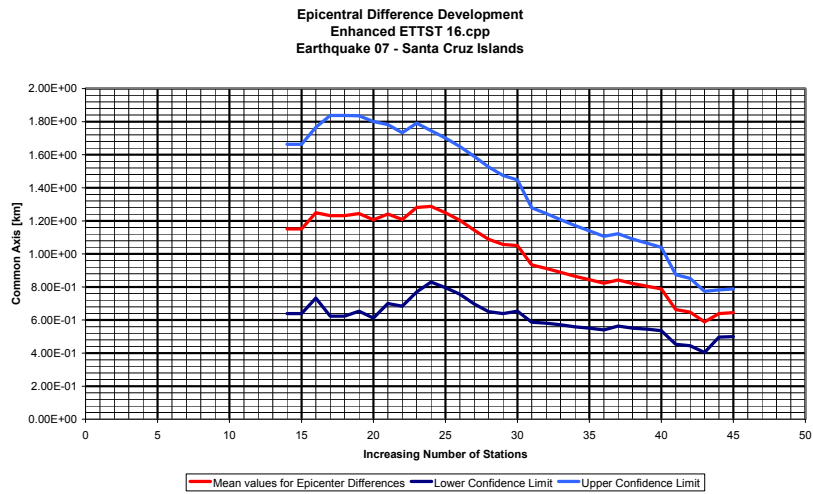


Figure 5. Arc distance between successive values for epicenters,  $\langle \Delta \underline{\epsilon} \rangle$ .

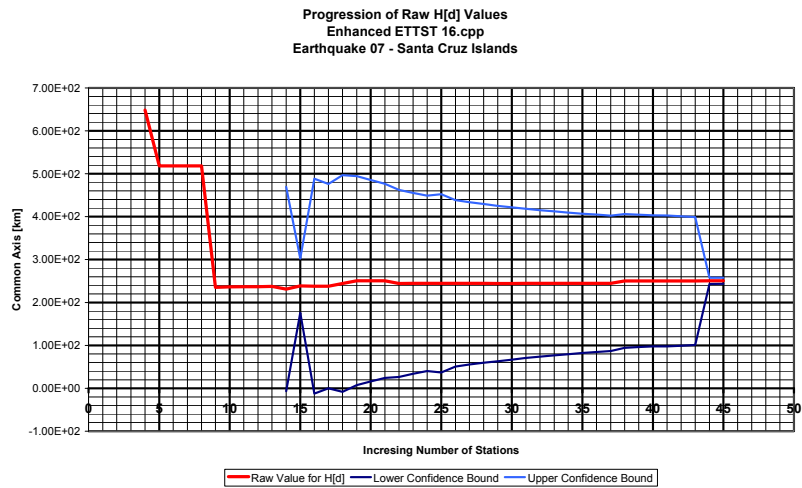


Figure 6. Series of raw  $H_d$  values



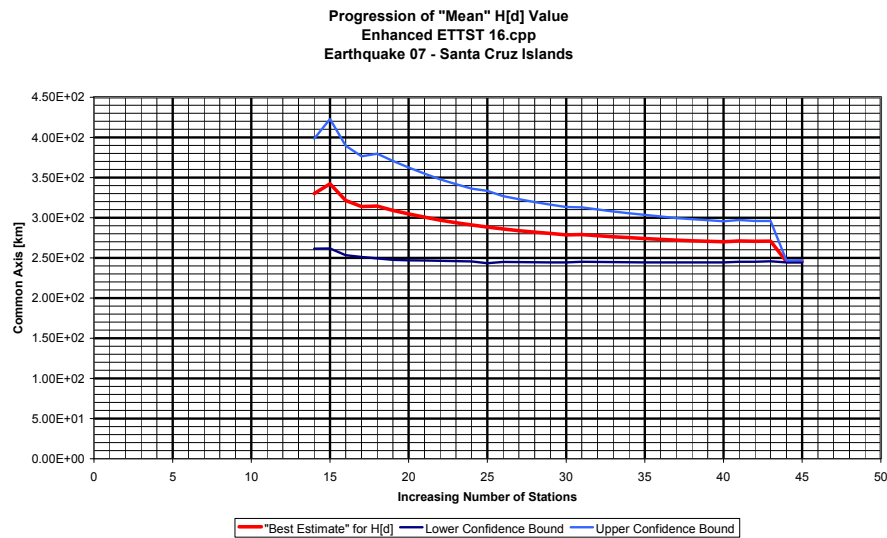


Figure 7. Series of "best estimates" of  $H_d$

#### 4. Discussion of Results

The main features to note are that:

- The Hypocenter was delivered at  $251.0 \pm 8.0$  kilometers
- The Epicentre was delivered at  $12.2266S/167.144E$ .
- The distance between the delivered Epicenter and the the received value was  $1.78 \pm 7.89_{10} - 01$  km. (This delivered Epicenter was estimated by the method given in Appendix A).
- The confidence intervals are seen to severely contract when the number of rejected localizations peaked at the end of the run (see figures 6 & 7).

These values are very close to the received values and their Confidence Intervals are narrow.

By consulting Appendix D it can be seen that the evolution of the Confidence Intervals follows more or less closely the gradually increasing number of rejection of localizations by the Chavenet test.

The results arising from the processing of Earthquake 07 (Santa Cruz Islands), and given here, would appear to show that this algorithm which concurrently delivers Epi- & Hypocentres by "interpolative tabular scanning" is capable of monitoring the accuracy of its output reliably and in real-time.

#### Acknowledgments

Acknowledgement is made to IRIS (Incorporated Research Institutes for Seismology) for providing the Earthquake Data Base handled by WILBER III, without which this analysis would not have been possible.

#### References

- Daglish, G. R., & Sizov, Y. P. (2013). System Suggestions for Hypocentre Location using Tabular Data with some Preliminary Illustrative Results. [unpublished report, available on request].
- Daglish, G. R., & Sizov, Y. P. (2013). System Suggestions for Hypocentre Location using Tabular Data with Some Illustrative Results. *Journal of Computer Engineering and Informatics*, Oct, 1(4), 96-111. World Academic Publishing (WAP).
- Daglish, G. R., & Sizov, Y. P. (2011). Suggestions for the Structure of Kernel Software in Intelligent Seismic Nets. World Congress on Engineering [WCE], London.
- Daglish, G. R., & Sizov, Y. P. (2012). Suggestions for Tele-Seismic Hypocentre Location using Table-driven Scanning, World Congress on Engineering [WCE]. London.
- Daglish, G. R., & Sizov, Y. P. (2013). A Hypocentre Determination using P- & S-Wave Interpolative Tabular Scans. [unpublished report, available on request].
- Daglish, G. R., & Sizov, Y. P. (2014). An Algorithm for Concurrent Location of Earthquake Epi- & Hypocentres

using P-arrivals in an Interpolative Tabular Scan. *Journal of Earth Science Research*, World Academic Publishing (WAP).

Daglish, G. R., & Sizov, Y. P. (2014). An Algorithm for Tracking Earthquakes upon their Initial Detection by Concurrent Localisation of Epi- & Hypocentres and Using Incremental P-wave First Arrivals. *Journal of Earth Science Research*, Aug, 2(1), 102-113. World Academic Publishing (WAP).

Daglish, G. R., & Sizov, Y. P. (2014). Exploration into the Response of an Algorithm Cluster performing Concurrent Epi- & Hypocentre Localisation: a Sequel. *Journal of Earth Science Research*, May, 2(2). World Academic Publishing (WAP).

Daglish, G. R., & Sizov, Y. P. (2014). Rapid and Concurrent Epi- & Hypocentre Location using Tabular Data Structures, Proceedings of Second European Conference on Earthquake Engineering and Seismology [2ECEES]. Istanbul.

Kennett, B. L. N., Engdahl, E. R., & Buland, R. (1995). Constraints on seismic velocities in the earth from travel-times. *Geophysical Journal International*, 122(1), Blackwell. <http://dx.doi.org/10.1111/j.1365-246X.1995.tb03540.x>

Neville, A. M., & Kennedy, J. B. (1964). *Basic Statistical method for Engineers and Scientists*. Intertext Books.

## Appendix A

### Best Estimate of Epicentre from the Accumulating Set $\{\underline{x}_i\}$

This forms a calculation for a point on a sphere,  $\underline{\alpha}$ , as “centroid” to a set of other points,  $\underline{x}_i$ , on a sphere. We have:

$$\alpha x_i + \beta y_i + \gamma z_i = r_e^2 \cos \vartheta_i$$

Here,  $\underline{\alpha} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$  and  $\underline{x}_i = \begin{pmatrix} x \\ y \\ z \end{pmatrix}_i$ .  $r_e$  is the local Earth radius.  $\vartheta_i$  is the angle between  $\underline{\alpha}$  and  $\underline{x}_i$ , at Earth

centre. Considering the arc-lengths subtended by  $\vartheta_i$ , namely  $S_i$ , we have:

$$S_i = r_e \vartheta_i = r_e \cos^{-1} \left( \frac{1}{r_e^2} \cdot (\alpha x_i + \beta y_i + \gamma z_i) \right)$$

We should seek to minimize the sum:

$$\frac{F_n}{r_e^2} = \frac{1}{r_e^2} \sum_i S_i^2 = \sum_i \vartheta_i^2 = \sum_i \left\{ \cos^{-1} \left( \frac{1}{r_e^2} (\underline{\alpha} \cdot \underline{x}_i) \right) \right\}^2$$

on behalf of the elements of  $\underline{\alpha}$ .

$$f_i = \cos^{-1} \left( \frac{1}{r_e^2} (\underline{\alpha} \cdot \underline{x}_i) \right); \underline{f} = [f_0 f_1 f_2 \dots f_{n-1}]$$

$$\frac{\partial f_i}{\partial \alpha, \beta, \gamma} = -\frac{1}{r_e^2} \cdot \frac{x_i, y_i, z_i}{\sqrt{1 - \frac{(\underline{\alpha} \cdot \underline{x}_i)^2}{r_e^4}}}$$

Let  $J_{\underline{f}}$  be the Jacobean of  $\underline{f}$ , then, applying the Gauss-Newton iterative method, we have:

$$\underline{\alpha}_{i+1} = \underline{\alpha}_i + \left( \underline{J}_f^T \cdot \underline{J}_f \right)^{-1} \cdot \underline{J}_f^T \cdot \underline{f}^T \Big|_i$$

This is considered to have converged when the modulus of the adjustment vector:

$$-\underline{h} \Big|_i = \left( \underline{J}_f^T \cdot \underline{J}_f \right)^{-1} \cdot \underline{J}_f^T \cdot \underline{f}^T \Big|_i$$

is less than a given small value, e.g. one meter.

However, if the convergence is not satisfactory, then it can be accelerated by including the Hessian matrix of the cost function in the calculation of the adjustment vector:

$$\underline{\alpha}_{i+1} = \underline{\alpha}_i + \left( \underline{J}_f^T \cdot \underline{J}_f + \sum_{j=0}^{n-1} f_j \cdot H_{f_j} \right)^{-1} \cdot \underline{J}_f^T \cdot \underline{f}^T \Big|_i$$

Here the Hessian matrices,  $H_{f_j}$ , are constructed as:

$$-\frac{1}{R_e^4} \begin{bmatrix} \frac{x_0^2 \cdot (\underline{\alpha} \cdot \underline{x}_j)}{(\sqrt{A_j})^3} & \frac{x_0 x_1 \cdot (\underline{\alpha} \cdot \underline{x}_j)}{(\sqrt{A_j})^3} & \frac{x_0 x_2 \cdot (\underline{\alpha} \cdot \underline{x}_j)}{(\sqrt{A_j})^3} \\ \frac{x_1 x_0 \cdot (\underline{\alpha} \cdot \underline{x}_j)}{(\sqrt{A_j})^3} & \frac{x_1^2 \cdot (\underline{\alpha} \cdot \underline{x}_j)}{(\sqrt{A_j})^3} & \frac{x_1 x_2 \cdot (\underline{\alpha} \cdot \underline{x}_j)}{(\sqrt{A_j})^3} \\ \frac{x_2 x_0 \cdot (\underline{\alpha} \cdot \underline{x}_j)}{(\sqrt{A_j})^3} & \frac{x_2 x_1 \cdot (\underline{\alpha} \cdot \underline{x}_j)}{(\sqrt{A_j})^3} & \frac{x_2^2 \cdot (\underline{\alpha} \cdot \underline{x}_j)}{(\sqrt{A_j})^3} \end{bmatrix},$$

and:

$$A_j = \text{Unity} - \left( \frac{\underline{\alpha} \cdot \underline{x}_j}{R_e^2} \right)^2$$

## Appendix B

### The Chauvenet test for Outliers

This simple test consists in comparing a normalized variate against a value taken from a table of values which establishes outlying limits of normal distribution according to the order of the set from which the candidate normalized variate is generated. If the normalized variate is greater than the table value then it may be judged an outlier.

The value from the table is indexed by  $n$  the order of the set from among whose elements outliers are required to be identified. This is a single tails test and the normalized variate builds up as:

$$\frac{|h_i - \bar{h}|}{\sigma_h}$$

Here the set is  $H = \{h_i\}$  with  $\bar{h}$  the mean and  $\sigma_h^2$  the variance.  $h_i$  is a general set element to be screened for disassociation from  $H$ .

A more in depth way of describing this process of disassociation, which is Chauvenet's test, is to state as follows:

“If the population from which the sample is drawn is normal then:

$$P_{>|z_i|} = 1 - \sqrt{\frac{2}{\pi}} \int_0^{z_i} \exp\left(-\frac{z^2}{2}\right) dz; \quad z_i = \frac{h_i - \bar{h}}{\sigma_h}$$

represents that proportion of the population that is equal to or exceeds the value  $h_i$ . This proportion can be forced to be:

$$n \cdot \left( 1 - \sqrt{\frac{2}{\pi}} \int_0^{z_i} \exp\left(-\frac{z^2}{2}\right) \cdot dz \right) \leq \frac{1}{2} \text{ or } \frac{1}{q_0}$$

So that, now  $h_i$  represents a margin of probability at and beyond which only smaller and smaller non-integral sample-elements can exist, then we can decide to call  $|h_i|$  and all smaller values – outliers,

We can say that the expected or probable relative frequency for the occurrence of the value  $\langle h_i \rangle$  is at the most

$$\frac{1}{n \cdot q_0}.$$

Further, we can solve for  $z_0$  in the inequality below for a set of  $n$  elements and parameterized by values of  $q_0$ :

$$\left( 1 - \sqrt{\frac{2}{\pi}} \int_0^{z_0} \exp\left(-\frac{z^2}{2}\right) dz \right) \leq \frac{1}{n \cdot q_0}$$

$n$  is the number of elements in the sample set  $H = \{h_i\}$  and  $q_0 \geq 2$  is a set of arbitrary values..

Then:

$$z_0 = \frac{h_i - \bar{h}}{\sigma_h}; \quad h_i = \sigma_h \cdot z_0 + \bar{h}$$

defines outliers at the relative frequency level  $\frac{1}{n \cdot q_0}$ .”

### Appendix C

#### “Best estimation” and Confidence Interval

We are dealing with a set of values, say Hypocenter Depths, collectively  $H_d$ , which correspond to a set of ordinal numbers, integers, which have no error. The readings of  $H_d$  are deemed constant, it could be the same value, and are therefore only perturbed by error in the data and the self-noise in the algorithm. The overall error distribution affecting the values  $H_d$  is assumed normal, invoking the Central Limit theorem.

Therefore a line of regression is fitted to the row of perturbed  $H_d$ . For clarity we say that the set of hypocenter depths is denoted by  $H = \{h_i\}$ , whose order is  $n$ , the index of the last localization.

The line of regression will have been found as:

$$\langle h_i \rangle = a_n + b_n \cdot i; \quad i \in [1, n]$$

This regression will be performed anew for each localization as  $n$  increases

A “best estimate” of  $H_d$  from the set  $H$ , representing a value assumed stationary, would be:

$$\langle h \rangle = \frac{1}{n} \sum_i h_i$$

While:

$$s_h^2 = \frac{\sum_i \epsilon_i^2}{\nu}; \quad \nu = n - 2$$

Where

$$\epsilon_i = h_i - (a_n + b_n \cdot i)$$

Then

$$s_{\bar{h}}^2 = \frac{s_h^2}{n}$$

By which the Confidence Interval is given as:

$$I = [\langle h \rangle - t \cdot s_{\bar{h}}, \langle h \rangle + t \cdot s_{\bar{h}}]$$

Here the  $t$ -variate is indexed from tables by the degree of freedom,  $\nu$ , and the level of significance (probability level) required. In these tests a level of 5% is chosen. Thus the value  $\bar{h}$  has a 95% chance of finding itself within the interval  $I$ .

The above refers to the “best estimate” and its Confidence Interval. However, we also would like to put a Confidence Interval about the latest entry to the set  $H$ . We write:

$$s_{h_i}^2 = s_{h_i}^2 + s_{\bar{h}}^2$$

$$s_{h_i}^2 = s_{\bar{h}}^2 \left[ \text{Unity} + \frac{1}{n} + \frac{(h_n - \bar{h})^2}{\sum_i (h_i - \bar{h})^2} \right]$$

The required Confidence Interval is:

$$I = [\langle h_n \rangle - t \cdot s_{h_n}, \langle h_n \rangle + t \cdot s_{h_n}]; \quad \langle h_n \rangle = a_n + b_n \cdot n$$

### Appendix D

#### Rejections by Applying the Chauvenet Test

This Appendix lists those Localisations which have been rejected as outliers by applying the Chauvenet principle. The first line, for instance, reads “at localization 11, the results of localizations 1 and 5 were excluded from the calculations leading to the “best estimates” of whatever variables and their corresponding Confidence Intervals”. Such variables would be:  $\langle H_d \rangle$  and  $\langle \Delta \mathcal{E} \rangle$  (see section 1 for symbols).

[NB Localisation 1 corresponds to the first data set of order 4 and so on.]

11:	1	5: Total	2		
12:	1	5	11	12: Total	4
13:	1	5	11: Total	3	
14:	1	5	11: Total	3	
15:	1	5	11	14: Total	4
16:	1	5	11	14: Total	4
17:	1	5	11	14: Total	4
18:	1	5	11	14: Total	4
19:	1	5	11	14: Total	4
20:	1	5	11	14: Total	4
21:	1	5	11	14: Total	4
22:	1	5	11	14: Total	4
23:	1	5	11	14: Total	4

24:	1	5	11	14: Total	4							
25:	1	5	11	14: Total	4							
26:	1	5	11	14: Total	4							
27:	1	5	11	14: Total	4							
28:	1	5	9	11	14: Total	5						
29:	1	5	9	11	14: Total	5						
30:	1	5	9	11	14: Total	5						
31:	1	5	9	11	14: Total	5						
32:	1	5	9	11	14: Total	5						
33:	1	5	9	11	14: Total	5						
34:	1	5	9	11	14: Total	5						
35:	1	5	9	11	14: Total	5						
36:	1	5	9	11	14: Total	5						
37:	1	5	9	11	14: Total	5						
38:	1	5	9	10	11	14	19: Total	7				
39:	1	5	9	10	11	14	19: Total	7				
40:	1	5	9	10	11	12	14	19: Total	8			
41:	1	2	3	4	5	9	10	11	12	14	19: Total	11
42:	1	2	3	4	5	9	10	11	12	14	19: Total	11

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).