Proposed Solution

“Mathematical Competitive Game 2014-2015”

Alan Oxley*

27 Apr., 2015

Suggested Results

1. The expected position of the receiver is (4,343,409.09; -124,936.95; 4,653,478.56). Fig. 1 shows the plan view of the position.

   Longitude 1.64764° W of Greenwich

   Latitude 46.96205° N of equator

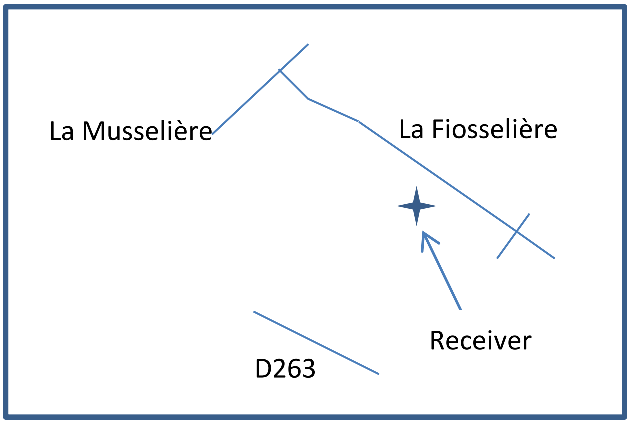


Figure 1: Plan view of expected position of receiver.

The receiver is situated in the Corcoué-sur-Logne, Pays de la Loire, region of France. The nearest city is Nantes lying far to the north. Just to the west of the map is the village of De la Baliniere.

* The author can be contacted via email at alanoxley12@gmail.com.

2. Consider a sphere whose centre is the centre of the earth and where (4,343,409.09; -124,936.95; 4,653,478.56) is a point on its surface. Now consider the plane that is tangential to this sphere at this point. The true location of the receiver has a 90% chance of lying in the region made up of four layers of rectangular cuboids, two layers lying above the plane (i.e. away from the earth's centre) and two layers lying below the plane (i.e. towards the earth's centre). The characteristics of each rectangular cuboid are shown in Fig. 2. The make-up of the layers is shown in Figs. 3 and 4.

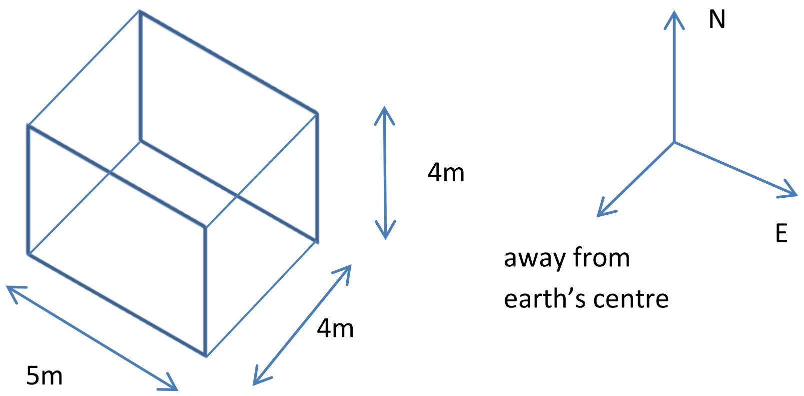


Figure 2: Dimensions and orientation of rectangular cuboid.

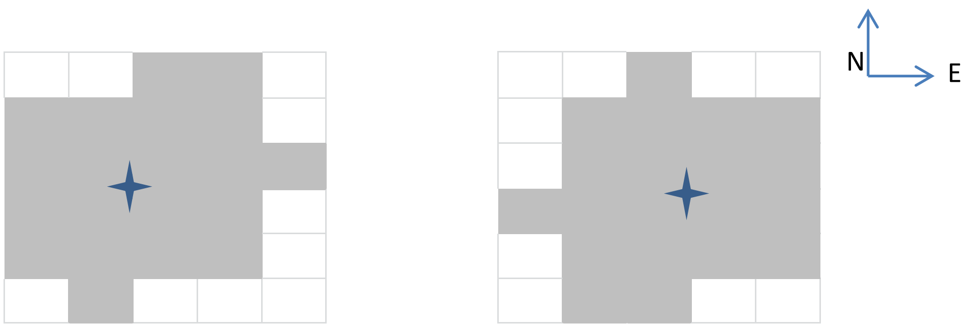


Figure 3: Plan view of first layer below the plane (left) and first layer above the plane (right). The cross marks the expected position of the receiver. There are 20 rectangular cuboids in each layer.

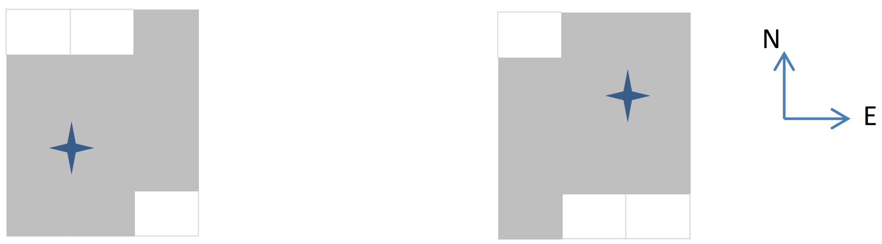


Figure 4: Plan view of second layer below the plane (left) and second layer above the plane (right). The cross marks the expected position of the receiver. There are 12 rectangular cuboids in each layer.

Supplementary Result

Due to the methods devised for problem 2, it is easy to calculate another result that is useful to those interested in GPS positioning, i.e. the region of a map (a 2D entity) within which the receiver has a 90% chance of lying. The true location of the receiver has a 90% chance of lying in the region shown in Fig. 5, where the cross is at Longitude 1.64764° W of Greenwich and Latitude 46.96205° N of equator.

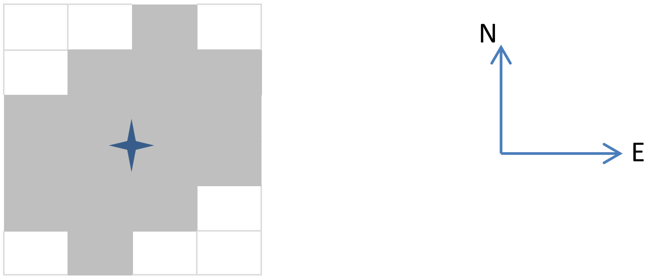


Figure 5: Region of map within which receiver has a 90% chance of lying. There are 16 rectangles in the region, each with dimensions 5m (E-W) by 4m (N-S).

Note on problems posed

Fig. 6 shows a representative view from space of GPS orbits and typical satellite positions.

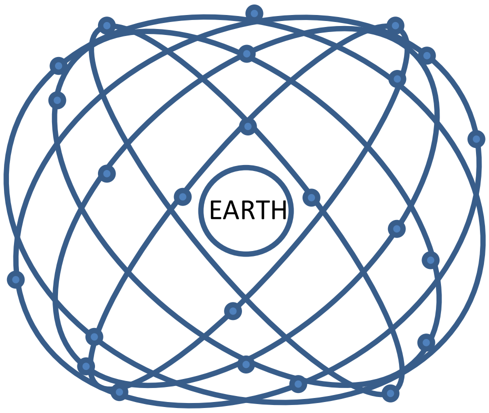


Figure 6: GPS satellite orbits.

$S_k = (x_k, y_k, z_k)$: kth 'dirty' satellite location, where 'dirty' means that it is the raw measurement data.

τ: bias of local clock in the receiver.

When a satellite signal is received and decoded at the receiver, the 'dirty' pseudo-distance δ_k to the kth satellite can be calculated. The 'dirty' geometric distance is d_k, and by taking the local clock bias into account, an estimate of the 'dirty' distance is

$$d_k = \delta_k - c\tau.$$

The estimate for the true distance is of course subject to errors. As a consequence we can write

$$d_k + c\tau = \delta_k + e_k.$$

The error term e_k accounts for measurement errors and any inaccuracies in the mathematical model, such as the effect of the ionosphere on c, the velocity of propagation. In the calculation of the expected receiver location, it is necessarily assumed that e_k is an independent random variable with zero mean.

$R = (x_0, y_0, z_0)$: expected receiver location, aka 'dirty' receiver location, where 'dirty' means that it has been calculated by assuming that e_k is an independent random variable with zero mean. The location is assumed to be stationary.

Problem 1: Methodology

Variation of Coordinates

Let the expected location and clock bias be $\mathbf{V}_0$. Given the pseudo-distances from five satellites, the objective is to estimate the components of the vector $\mathbf{V}_0 = (x_0, y_0, z_0, \tau)$. Let δ_k, $k = 1 : 5$, be a set of five measurements. Let $f_k(\mathbf{V}_0)$ be the known function which maps the expected location and clock bias $\mathbf{V}_0$ to an accurate kth measurement. Given the vector $\mathbf{V}_0$, the distance function for the kth satellite is defined as

$$\begin{aligned} f_k(\mathbf{V}_0) &= d_k + c\tau \\ &= \sqrt{(x_k - x_0)^2 + (y_k - y_0)^2 + (z_k - z_0)^2} + c\tau. \end{aligned}$$

Then with a measurement error e_k we can write $f_k(\mathbf{V}_0) = \delta_k + e_k$. With all five measurements available, and an obvious notation, the results can be expressed in vector form as $\mathbf{F}(\mathbf{V}_0) = \mathbf{R} + \mathbf{E}$. Note that the function $\mathbf{F}$ necessarily involves the locations of all five satellites.

As $\mathbf{V}_0$ is not known, to find a least-squares solution for $\mathbf{V}_0$ we create a guess $\mathbf{V}_G$, which is in error by an unknown vector $\mathbf{V}_D$ from the truth. We use a first-order Taylor expansion of $\mathbf{F}()$ in terms of $\mathbf{V}_D$ in the guess. That is, let $\mathbf{V}_0 = \mathbf{V}_G + \mathbf{V}_D$. Given the guessed location and clock bias $\mathbf{V}_G$ and the measurement vector $\mathbf{R}$, the least-squares objective is to find $\mathbf{V}_D$ such that the sum of the squares of the five measurement errors in the vector $\mathbf{E}$ is a minimum. Since for a given vector $\mathbf{V}$ the function $\mathbf{F}(\mathbf{V})$ is known, and by assumption $\mathbf{V}_D$ is small, then to the first order we can write $\mathbf{F}(\mathbf{V}_0) = \mathbf{F}(\mathbf{V}_G + \mathbf{V}_D) \approx \mathbf{F}(\mathbf{V}_G) + \mathbf{A}.\ \mathbf{V}_D$ where $\mathbf{A} = \nabla\, \mathbf{F}(\mathbf{V}_G)$ is the matrix of derivatives of $\mathbf{F}(\mathbf{V}_G)$ with respect to the vector $\mathbf{V}_G$, evaluated at $\mathbf{V}_G$. This leads to the equations

$$\mathbf{F}(\mathbf{V}_0) = \mathbf{R} + \mathbf{E} = \mathbf{F}(\mathbf{V}_G + \mathbf{V}_D) \approx \mathbf{F}(\mathbf{V}_G) + \mathbf{A}.\ \mathbf{V}_D.$$

The measurement error vector $\mathbf{E}$ can thus be expressed as $\mathbf{E} = \mathbf{A}.\ \mathbf{V}_D - \mathbf{B}$ where $\mathbf{B} = \mathbf{R} - \mathbf{F}(\mathbf{V}_G)$ is the vector of 'observed minus computed' measurements. The vector $\mathbf{V}_D$, which minimises the weighted squared error $\varepsilon = \mathbf{E}^T\mathbf{\Omega}\mathbf{E}$, where $\mathbf{\Omega}$ is a diagonal matrix of weights, is obtained by differentiating ε with respect to the vector $\mathbf{V}_D$ and setting the result to zero.

$$\mathbf{\Omega} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The result for $\mathbf{V}_D$ is given by

$$\mathbf{V}_D = (\mathbf{A}^T\mathbf{\Omega}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{\Omega}\mathbf{B}.$$

The calculated vector $\mathbf{V}_D$ is then added to the original guess $\mathbf{V}_G$ to form a new (improved) guess $\mathbf{V}_G$. The process is then repeated as often as necessary until the calculated update vector $\mathbf{V}_D$ is sufficiently small.

Problem 2: Methodology

R': true receiver location

The approach adopted is to first find an approximate bounding box within which *R'* must lie. This is followed by subdividing the bounding box into a mesh of rectangular cuboids (small boxes). Finally, the set of small boxes making up the relevant volume is identified, i.e. where there is a 90% likelihood that *R'* lies. This approach involves a number of steps:

- Deciding on the orientation of the bounding box. The initial idea was to have a bounding box whose faces are parallel to each of the original *x*-, *y*-, and *z*-axes. However, it was realised that a valuable spin-off could be achieved if the axis system was changed such that when taking a plan view of *R*, a rectangular area on the ground could represent a column of small boxes. Thus the axis system was changed by rotating the original system so that *R*, from problem 1, lies on the *x*-axis (see note 1 below). To simplify the mathematics, it was further decided to translate the axes such that the origin is at *R*, and the centre of the earth is on the negative *x*-axis.
- Finding the *x*-, *y*- and the *z*-extents of the bounding box (see note 2).
- Deciding on the dimensions of a small box and specifying the coordinates of the centroids of all small boxes (see note 3).
- Combining the probability distributions of measurement errors in S_k and d_k into a new distance error distribution (see note 4.) This involves two sub-steps:
  - Converting the 3D probability distribution of the measurement error in S_k to a 1D probability distribution (see note 5).
  - Combining the 1D probability distribution of the measurement error in S_k and the probability distribution of the measurement error in d_k into a new distance error distribution (see note 6).

The result is a piecewise constant function:

probability = *g*(new distance error)

- For each satellite, finding a mathematical expression that approximates $dist(P, S_k) - dist(R, S_k)$, where P is any point in the bounding box, and where (x, y, z) are the coordinates of P (see note 7).

  From the work here we find that

  $$dist(P, S_k) - dist(R, S_k) \approx b_1x + b_2y + b_3z.$$

  As $g()$ is symmetric about zero, this means that the value of g(distance difference) for the centroid at (x, y, z) is the same as it is for the centroid at $(-x, -y, -z)$. This reduces the workload by half!

- For each small box, calculating the probability that R' lies within the box. Two methods were devised for this step. One is an approximate method (see note 8) and the other is a more precise method. For the particular problem that is being solved here, both methods give almost the same solution. The precise method involves several sub-steps:
  - Working out the dimensions of a mesh of inspection points to be used within a small box (see note 9).
  - For each satellite, creating a 3D array of increments that describe how g(distance difference) for all of a small box's inspection points can be inferred from the value of g(distance difference) at the small box's centroid (see note 10).

    Assume that we have calculated the distance difference for a small box's centroid: $dist(C, S_k) - dist(R, S_k)$). The function $g()$ maps the interval in which the distance difference lies to the probability of it occurring. It is a piecewise constant function. We can find the 'piece' corresponding to the interval in which the distance difference of the centroid lies. The way in which the distance difference varies has been studied in an earlier step. Using this information, we can infer which 'piece' of $g()$ corresponds to each inspection point. We use '-1' to denote the 'piece' directly to the left of the centroid's 'piece.' '-2' denotes the 'piece' that is two 'pieces' to the left of the centroid's 'piece'; '+1' denotes the 'piece' directly to the right of the centroid's 'piece'; etc. For example, if $dist(C, S_k) - dist(R, S_k)$) = 2.2 the interval in which 2.2 lies is (2.0, 2.4] and so '-1' denotes (1.6, 2.0]; '-2' denotes (1.2, 1.6]; '+1' denotes (2.4, 2.8].

  - For each small box, calculating its metric value. The metric that is used, which is directly related to how likely it is for R' to be located inside a small box, is:

    $$\sum_{\text{all inspection points}} \left(\prod_{k=1}^{5} g[dist(P, S_k) - dist(R, S_k]] \right)$$

    where $g()$ is the piecewise constant function calculated in an earlier step.

The metric involves first finding g(distance difference) for the centroid (given by $g[dist(C, S_k) - dist(R, S_k)]$) and then inferring $g[dist(P, S_k) - dist(R, S_k)]$ for all of the other inspection points. (This procedure has been mentioned in the previous sub-step.)

We find the metric values for all 512 small boxes (8 X 8 X 8). The likelihood that a small box contains R', as a percentage, is calculated as follows:

$$\frac{\sum_{\text{all inspection points}} \left(\prod_{k=1}^{5} g[dist(P, S_k) - dist(R, S_k)] \right)}{\sum_{\text{all small boxes}} \left(\sum_{\text{all inspection points}} \left(\prod_{k=1}^{5} g[dist(P, S_k) - dist(R, S_k)] \right) \right)} \times 100$$

Next we create a list of the small boxes sorted by their percentages, from largest to smallest. The region of interest comprises a set of these small boxes. The set is formed by starting at the top of the list and adding each small box to the set until the cumulative probability reaches 90%.

Problem 2: Notes

1. Rotation matrices

$$R_x(\vartheta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$
$$R_y(\vartheta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}$$
$$R_z(\vartheta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2.

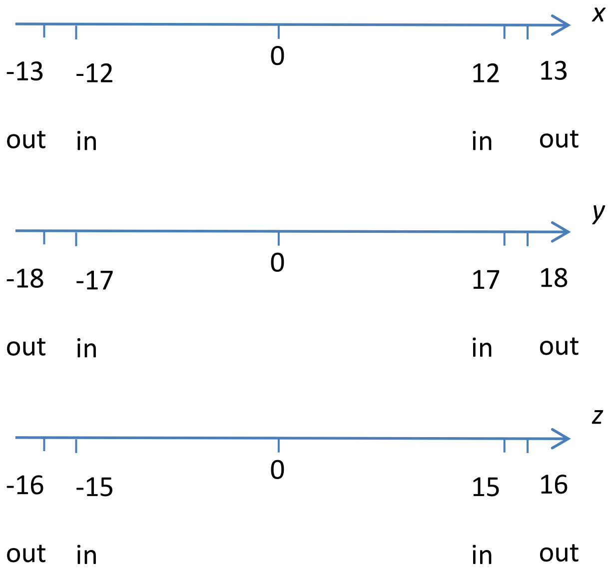


Figure 7: Delimiting the approximate bounding box.

3.

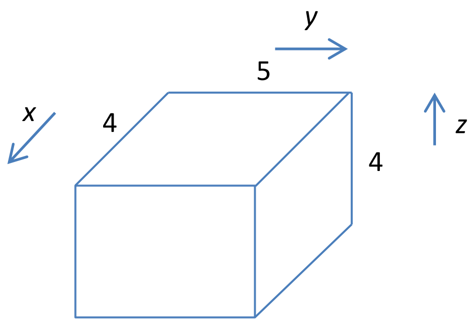


Figure 8: Dimensions of a small box.

The centroids of all the small boxes are every combination of the following *x*, *y* and *z* values:

x: -14, -10, -6, -2, 2, 6, 10, 14

y: -17.5, -12.5, -7.5, -2.5, 2.5, 7.5, 12.5, 17.5

z: -14, -10, -6, -2, 2, 6, 10, 14

4.

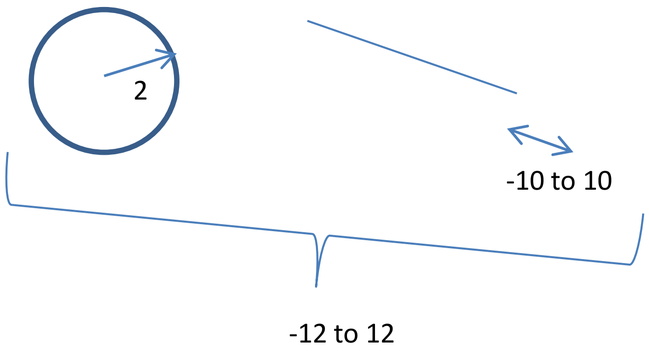
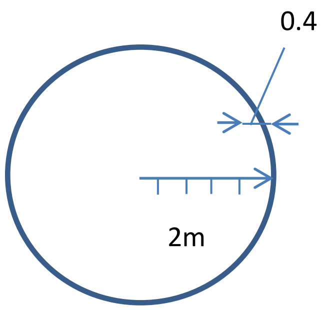
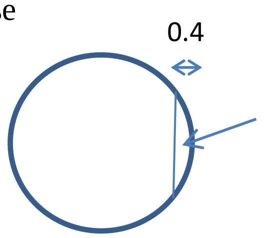
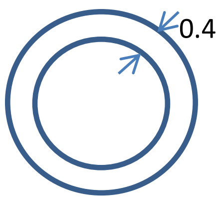


Figure 9: Combining two probability distributions into one.

5.

 Figure 10: Slices of a sphere, each with a thickness of 0.4m.	Vol. of sphere $= \frac{4}{3}\pi r^3 = \frac{4}{3}\pi 8$ $= \frac{32}{3}\pi$
1^{st} case  	Spherical cap (r = 2, h = 0.4) Vol. $= \frac{\pi h^2}{3}(3r - h) = \frac{\pi h^2}{3}(6 - h)$ $= \frac{\pi(0.4)^2}{3} 5.6$ Part of the 10% probability region. Vol. of 10% probability region = Vol. of whole sphere - $\frac{4}{3}\pi(1.6)^3$ Total probability = 0.6%
2^{nd} to 5^{th} cases	The total probability for each case is calculated.

6. In what follows we will be looking at points in the vicinity of *R*. Let us refer to a point under consideration as an 'inspection point.' For a small box, we first handle the centroid *C*. All other inspection points are considered later.

S_k is subject to error. From problem 1, we are able to calculate d_k.

We will use the sign of the error to denote whether the error contributes to *C* being nearer or further from a satellite as compared to *R*. For example, a satellite error of between 1.6 and 2 means that we are considering a satellite position that is between 1.6m and 2m further away from *R* than S_k than is. A negative satellite error means that we are considering a satellite position that is closer to *R* than S_k is. Similarly, a distance error of between 9.6 and 10 means that we are considering a distance that is between 9.6m and 10m longer than d_k. A negative distance error means that we are considering a distance that is shorter than d_k. For example,

Distance error = -10 to -8 Satellite error = -2 to -1.6

Let *x*% be the likelihood of having this distance error and this satellite error.

The combined error lies between -12 and -9.6. This spans six regions of size 0.4: -12 to -11.6, -11.6 to -11.2, …, -10 to -9.6. The probability of the combined error lying in any one of these regions is a fraction of *x*%. Table 1 shows how the component probabilities are calculated. The columns denote the six regions. The rows denote the distance errors.

Table 1: Calculating the component probabilities of the six regions.

	-12 to -11.6	**-11.6 to -11.2**	**-11.2 to -10.8**	**-10.8 to -10.4**	**-10.4 to -10**	**-10 to -9.6**
-8.4 to -8					√	√
-8.8 to -8.4				√	√	
-9.2 to -8.8			√	√		
-9.6 to -9.2		√	√			
-10 to -9.6	√	√				

Table 2 shows the probabilities for different combinations of satellite error (the rows) and distance error (the columns). For brevity, only the first two rows are shown.

Table 2: Probabilities for different error combinations.

		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
		≥-10	**>-8**	**>-6**	**>-4**	**>-2**	**>0**	**>2**	**>4**	**>6**	**>8**
1	**>1.6**	5*a*	5*b*	5*c*	5*d*	5*e*	5*e*	5*d*	5*c*	5*b*	5*a*
2	**>1.2**	5*f*	5*g*	5*h*	5*i*	5*j*	5*j*	5*i*	5*h*	5*g*	5*f*

7. For a small box, we calculate the distance difference for its centroid (C). For the other inspection points we use an indirect approach.

Table 3: $dist(C, S_1) – dist(R, S_1)$ for centroids at x = -10, to 1 d.p.

	y = -12.5	**-7.5**	**-2.5**	**2.5**	**7.5**	**12.5**
z = 14	…	…	…	…	…	8.0
10	…	…	…	8.3	8.5	8.6
6	…	…	8.8	8.9	9.1	9.3
2	…	9.2	9.4	9.6	9.8	10.0
-2	…	9.9	10.1	10.3	10.5	…
-6	…	…	10.8	10.9	…	…
-10	…	…	11.4	…	…	…

Table 4: $dist(C, S_1) – dist(R, S_1)$ for centroids at x = -6, to 1 d.p.

	y = -12.5	**-7.5**	**-2.5**	**2.5**	**7.5**	**12.5**
z = 14	…	…	…	…	3.9	…
10	…	…	4.2	4.3	4.5	4.7
6	…	4.6	4.8	5.0	5.2	5.4
2	5.1	5.3	5.5	5.7	5.9	6.0
-2	…	6.0	6.2	6.3	6.5	…
-6	…	6.6	6.8	7.0	7.2	…
-10	…	…	7.5	7.7	…	…

$dist(P, S_1) – dist(R, S_1) \approx b_1x + b_2y + b_3z$, where P is any point in the bounding box, and where (x, y, z) are the coordinates of P.

$dist(P, S_1) – dist(R, S_1)$, the distance difference for satellite 1, is in close agreement with a linear model.

Table 5: $dist(C, S_4) – dist(R, S_4)$ for centroids at x = -10, to 1 d.p.

	y = -12.5	**-7.5**	**-2.5**	**2.5**	**7.5**	**12.5**
z = 14	…	…	…	…	…	4.0
10	…	…	…	-0.2	3.1	6.4
6	…	…	-1.1	2.2	5.5	8.9
2	…	-2	1.4	4.7	8.0	11.3
-2	…	0.5	3.8	7.1	10.5	…
-6	…	…	6.3	9.6	…	…
-10	…	…	8.7	…	…	…

In summary, for each satellite k, we find an expression of the form $b_1x + b_2y + b_3z$ that approximates $dist(P, S_k) – dist(R, S_k)$.

8. The inspection points that we use are the centroids (*C*s) of the small boxes. The metric that is used, which is directly related to how likely it is for *R*' to be located inside a small box, is:

$$\prod_{k=1}^{5} g[dist(C, S_k) - dist(R, S_k)]$$

where *g*() is the piecewise constant function calculated in an earlier step. The metric involves first finding the distance difference for a centroid (given by $dist(C, S_k) - dist(R, S_k)$). Table 6 shows example distance differences for two small box centroids. The figures are correct to 1 d.p. However, where the approximation is a multiple of 0.4, the figures are given to a higher precision. This is so that when we use *g*(), we can avoid a distance difference that falls between two 'pieces' of the piecewise constant probability function.

Table 6: Distance differences for centroids of two small boxes.

	(-2,-2.5, -2)	**(-2, -2.5, 2)**
S_1	2.2m	1.5
S_2	2.3	2.82
S_3	3.8	1.57
S_4	0.42	2.04
S_5	0.3	3.5

We find the metric values for all 512 small boxes (8 X 8 X 8). The likelihood that a small box contains *R*', as a percentage, is calculated as follows:

$$\frac{\prod_{k=1}^{5} g[dist(C, S_k) - dist(R, S_k)]}{\sum_{\text{all centroids}} \left(\prod_{k=1}^{5} g[dist(C, S_k) - dist(R, S_k)] \right)} \times 100$$

Next we create a list of the small boxes sorted by their percentages, from largest to smallest. The region of interest comprises a set of these small boxes. The set is formed by starting at the top of the list and adding each small box to the set until the cumulative probability reaches 90%.

9. Let us see how the distance differences change as we move from one side of a small box to the other.

   For satellite 1: Consider $-0.98536x + 0.036312y - 0.16666z$

   A movement in the *x*-direction of 4m, keeping *y* and *z* constant, causes a change of ≈ 3.9 in the distance difference.

   Moving 5m in *y*-direction only, causes a change of ≈ 0.2.

Moving 4m in z-direction only, causes a change of ≈ 0.7.

Similarly, we study the distance difference changes for satellites 2 to 5. As a result, we decide on the dimensions of the mesh of inspection points.

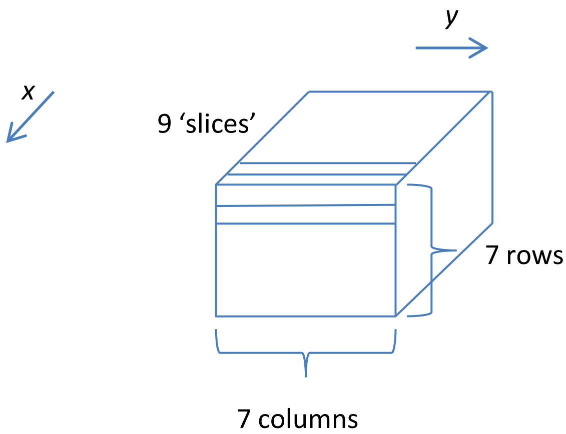


Figure 11: Mesh of inspection points within a small box.

10. $c = g(dist(C, S_k) - dist(R, S_k))$

Increments: '-1' denotes the probability is that of the 'piece' of $g()$ to the left of where c has been derived from.

Slices: '+1 slice' is the slice one away from the centre slice in the positive x direction.

Table 7: Satellite 1 increments.

Slice	**Increment for all inspection points in slice**
Middle	0 (i.e. probability is c)
+1	-1
+2	-2
+3	-3
+4	-4
-1	+1
-2	+2
-3	+3
-4	+4

Similarly, for satellites 2 to 5, we find the increments of all the inspection points.