Sharp Tools for Optimising Navigation Sensor Arrays

Eur Ing Dr Martin J D Bishop

Emeritus Solutions Ltd

Mjd.Bishop@Emeritus-Solutions.com

www.Emeritus-Solutions.com

Biography

Eur Ing Dr Martin Bishop is a Technical Consultant and the Director of Emeritus Solutions Ltd. Martin is based in Dorchester, Dorset, UK. His current interests include positioning systems and algorithms, Digital Signal Processing, system implementation and remediation; typically for the Defence and Oil and Gas sectors.

Martin's formative years were spent with the UK Ministry of Defence and latterly QinetiQ in the Portland area. While with UK MoD he was awarded a Personal Promotion to UG6 and while with QinetiQ was a QinetiQ Fellow.

Dr Bishop has wide experience, gained over 30+ years, of:

- underwater tracking range analysis, development and grooming
- torpedo and sonar target development, calibration and analysis
- torpedo and allied systems analysis and development
- active and passive sonar design and development
- submarine and surface ship systems and constraints
- digital signal processing, statistics, computers, electronics, …
- the technical architecture of navigation, torpedo, submarine and ASW systems

Dr Bishop is a Fellow of the Institution of Electrical Engineers

Abstract

For most navigators DOP is a given, and activities are tailored to it. However some navigators, who deploy sensor networks, "make their own DOP". Undersea navigation is a domain in which this paradigm is increasingly important, similarly location services make every radio mast a navigation sensor. This increase in the volume of sensor network design requirements requires sharper tools, which can inform design decisions by less expert users. To do so, they must be better not only graphically but also in the metrics they present to inform the designer's decisions.

Significant factors in sensor network design include: the required areal and vertical coverage, the quality of coverage required, system constraints (e.g. poor clock accuracy), the topography or bathymetry, and certainly at sea the propagation medium's characteristics. Traditionally in designing sensor networks an expert navigator / surveyor has considered these explicit factors together with a substantial body of tacit knowledge, e.g. the often catastrophic reduction of position accuracy outside the sensor array. Beyond managing the data, tomorrow's tools need to offer a sharper critique of performance than a washed out spatial DOP plot.

Plots of DOP are the de facto standard for the "quality" of a position fix. However, DOP values are only as valid as the assumptions which define the problem analysed. Questionable, common assumptions in undersea navigation include: precise knowledge of the propagation velocity and the absence of clock biases. DOP plots generally indicate modest degradation at the edge of sensor arrays. However, the practical navigator knows that without a fair measure of symmetry in the user – sensor geometry a fix is, in the

absence of perfectly corrected measurements, very likely to be biased. The metrics used by tomorrow's tools must address this limitation.

This paper will approach the sensor network synthesis problem from analytic and practical perspectives, propose metrics which we consider powerful compliments to DOP, and present comparative examples. The examples presented will range from the tutorial, through the underwater realm and European Loran coverage. The presentation will use graphical tools which permit the designer to work with tens of sensors and the constraints imposed by DTMs (digital terrain maps).

Keywords

Underwater; Positioning; Trilateration; Multilateration; Dilution of Precision; OpenGL; Quaternions.

Introduction

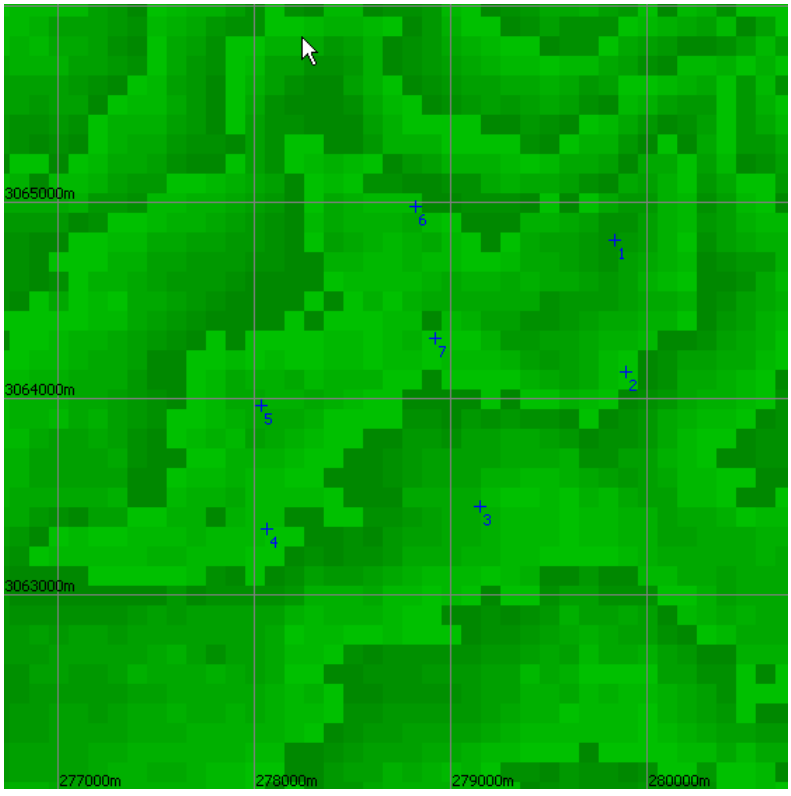
This paper aspires to be visual, therefore while it includes a quantity of mathematics, its primary focus is on presenting examples of 3-D graphical visualisation tools for 2, 3 and 4 dimensional data sets accessible to everyman. We begin with a worked example of a deepwater positioning range layout in the Gulf of Mexico, the focus of the presentation being the simultaneous presentation of the bottom bathymetry and design metrics to facilitate station location optimisation. We move on to a second worked example, the NW European Loran chain, presenting analyses of the current, possible future and might have been configurations performance. Penultimately, the paper presents its conclusions. Finally, some mathematical aspects of positioning metrics are discussed in the appendices and the references are listed.

Underwater Example : Gulf of Mexico – Bathymetry and Preliminaries

A common oil and gas problem is to lay out a station array for near bottom positioning, in 1000 to 3000 msw. With a DTM / DEM (Digital Terrain / Elevation Map) from a multibeam sonar as data and a leavening of client requirements. The surveyor / engineer's task is to provide good coverage using as few assets as possible.

For modest baselines and away from the sea bed the design task is straightforward. However, when sea bed topography obstructs the line of sound to receivers within meters of the sea bed and particularly when refraction limits propagation ranges, the task is anything but trivial.

To visualise not just what the problems are, but more importantly what the solutions might be requires sharp tools. The capabilities of such a toolset Emeritus Solutions' Plotter (ESP) are used to show that while difficult the problem is tractable.



Baseline Data – a grey scale Digital Terrain Map

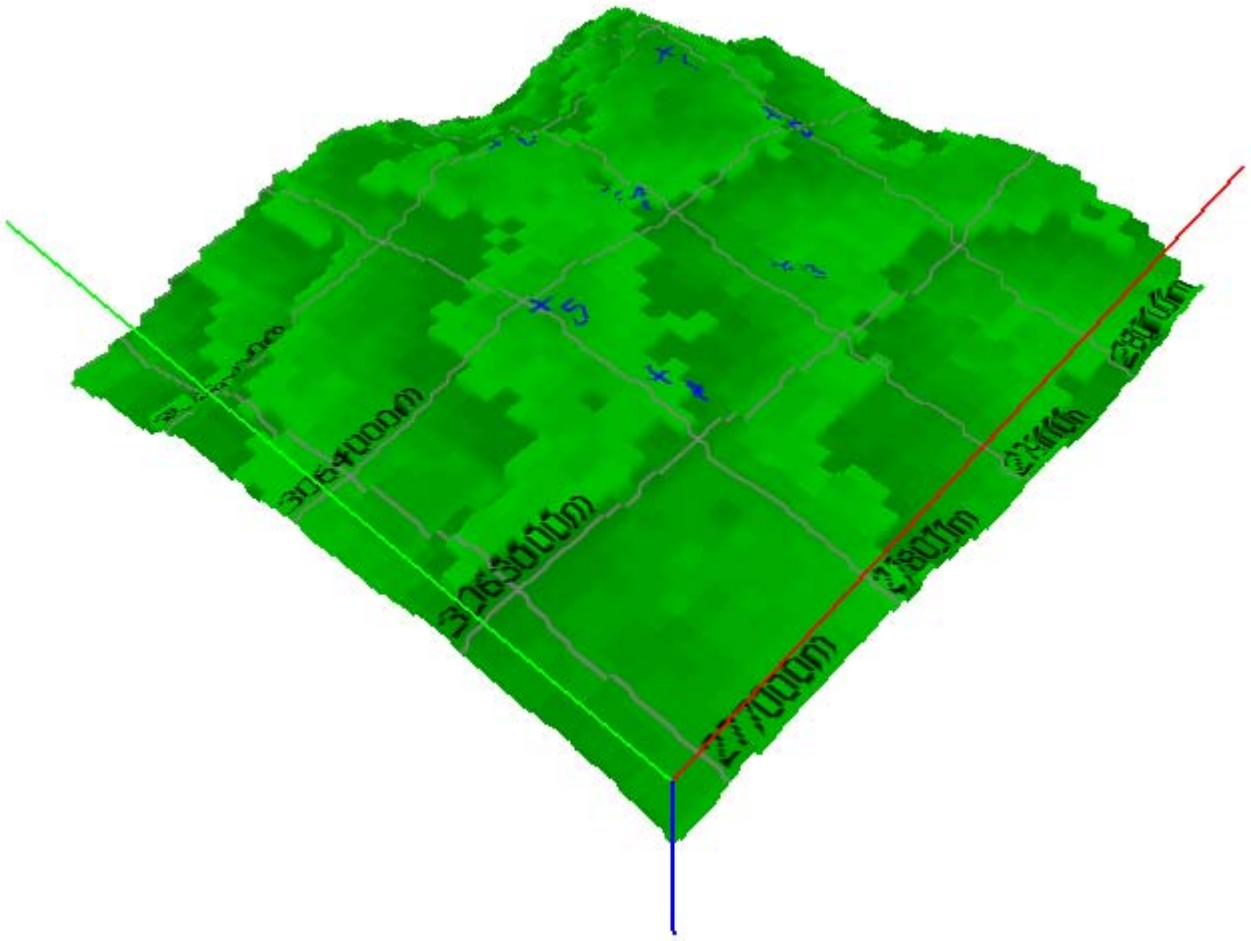
Multibeam sonar data is commonly polished, to remove track to track elevation errors, merge multiple data sets and to fill in black holes. Typically it arrives as binary or ASCII files containing X / Y / Z triples. Once loaded into a computer it can be visualised in 2-D by greyscale coding of depth or contouring.

The specimen data plotted above is a small extract from the US National Geophysical Data Centre (NGDC) 3 Arc-Second Coastal Relief Model dataset [1]. The resolution used in the computer is 100 x 100 meter cells, on a UTM grid (UTM zone 16). The XY resolution is of course much coarser than the 2 to 5 m X & Y resolution of a commercial survey. Similarly, the depth resolution and accuracy are probably no worse than a commercial survey, but the large blocks result in slopes becoming steps. In any case, as we shall see, it is good enough for a tutorial example.

The next sharp tool is OpenGL [2-4] and the computer graphics processors [5-6] on which it runs. The critical point is that Laptops and particularly desktop boxes can, not to be coy, use gaming technology to visualise 3-D datasets at modest cost in both specie and programming.

Manipulation in 3-D seems challenging, but can be straightforwardly effected with a wheel mouse. The simplest paradigm is Left-Right -> Roll, Up-Down -> Pitch, & Wheel -> Azimuth. It is also possible to program a 2-D mouse or tracker ball to provide a 3-D output [7]; all you have to do is to move it in a circle, - unfortunately in the opposite direction to the desired motion.

The sharp tools used to handle 3-D orientation are Quaternions [7-9], the Euler-Rodrigues parameterisation, rotation matrices and all the associated mathematics [10]. As complex numbers “squared” they are perhaps not everyone’s meat ; e.g. they are close relatives of Lie Groups and Algebras [11-12] ... However, quaternions are widely used in AUVs, aerospace and by the computer graphics community. And, they are well worth using, even if the implementation requires a consultant’s services.

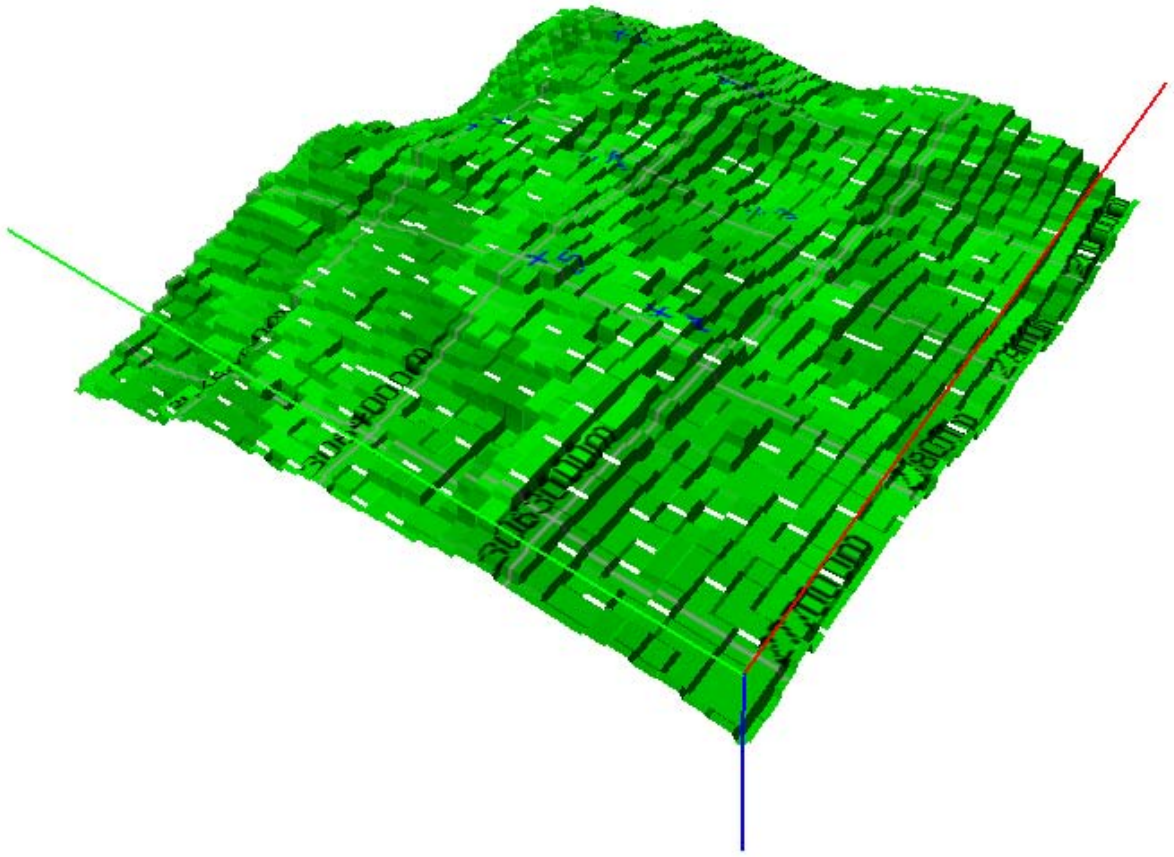


OpenGL view of a DTM : Without Lighting

Preliminaries over, our first example of what OpenGL can do. The simplest visualisation of a DTM is to render and manipulate it as an unlit surface. By comparison with a 2-D greyscale plot this is wonderful; especially when manipulated interactively rather than lying “flat” on the page. However, with lighting one can do much better.

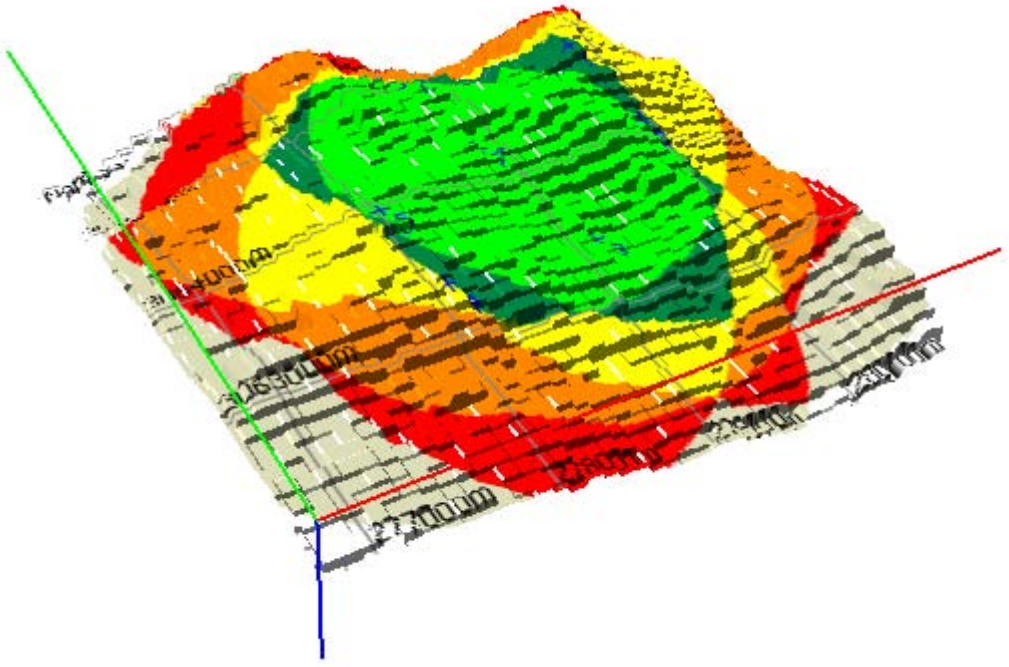
The slice is 4 km on each side, with a 1 km reticule. The Eastings and Northings are for UTM Zone 16. The location is 27° 41’ North 89° 14’ West, just South of the Latex shelf – where the oil lies. The vertical extent of

the slice is about 50 m. The red, green and blue axes correspond to the X (E), Y (N) and Z axes respectively. The blue line is 100m long, the vertical axis being exaggerated.



Digital terrain Map (DTM) with lighting

Turning the lighting on, the grainy structure of the 100 x 100 m blocks is very evident. More importantly, one can rapidly assimilate the "Lie of the Land". Henceforth, we shall primarily present lit perspective views of the terrain. The annotations and crosses in blue are the proposed station locations. These can readily be discerned in the plan views presented elsewhere in the paper.



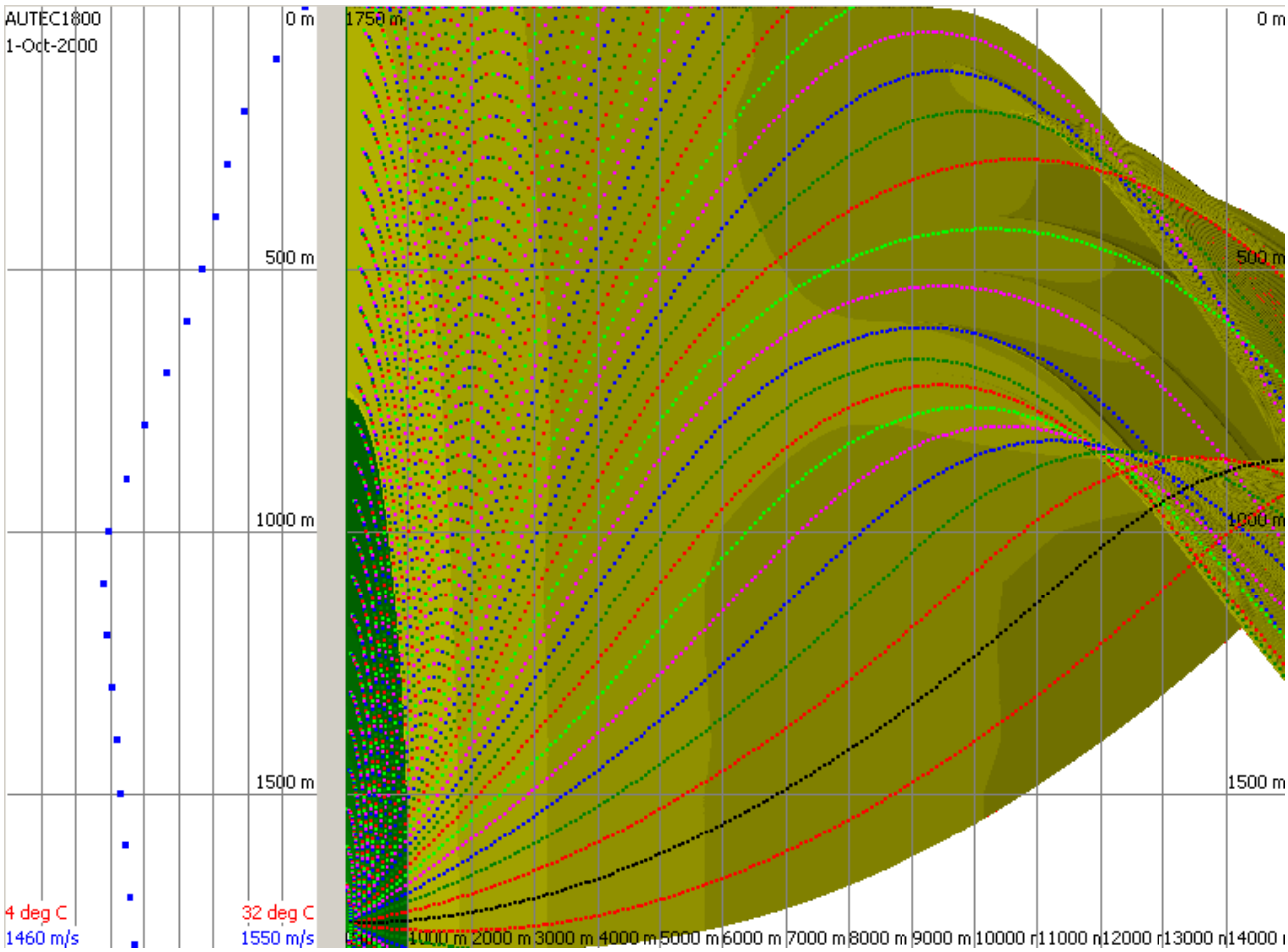
Number of Measurements (NoM) for the baseline layout

Assuming an effective 2 km range for each of the stations, a little cookie cutting, some addition and colour coding we can start to visualise the coverage on the terrain. Red = 3, Orange = 4, Yellow = 5, Dark Green = 6 and Light Green = 7 stations in range. Depending on tracking mode, accuracy and QA requirements from two to five slant ranges could be required for 2-D XY (measured depth) positioning. This is a very optimistic plot – even the eye can see that some coverage which won't work : nice toy OpenGL..

Underwater Example : Gulf of Mexico – Sound Propagation

Sound propagation and acoustics [13-17] are always central to underwater acoustic positioning. The list of potential issues is extensive, encompassing attenuation, reverberation, noise levels, multipath and most importantly refraction. The speed of sound of seawater varies vertically and to a lesser degree horizontally. The effect of the vertical (speed of sound) gradients is to refract the line of sound.

Generally, path occlusion due to refraction determines the presence of a line of sound between any two points in the water column. The specimen Sound Velocity Profile (SVP) for the AUTEC range off Andros Island in the Bahamas, taken from the internet and depicted below, is typical of a warm, deep ocean; e.g. Gulf of Mexico, offshore Nigeria, and offshore Brazil.



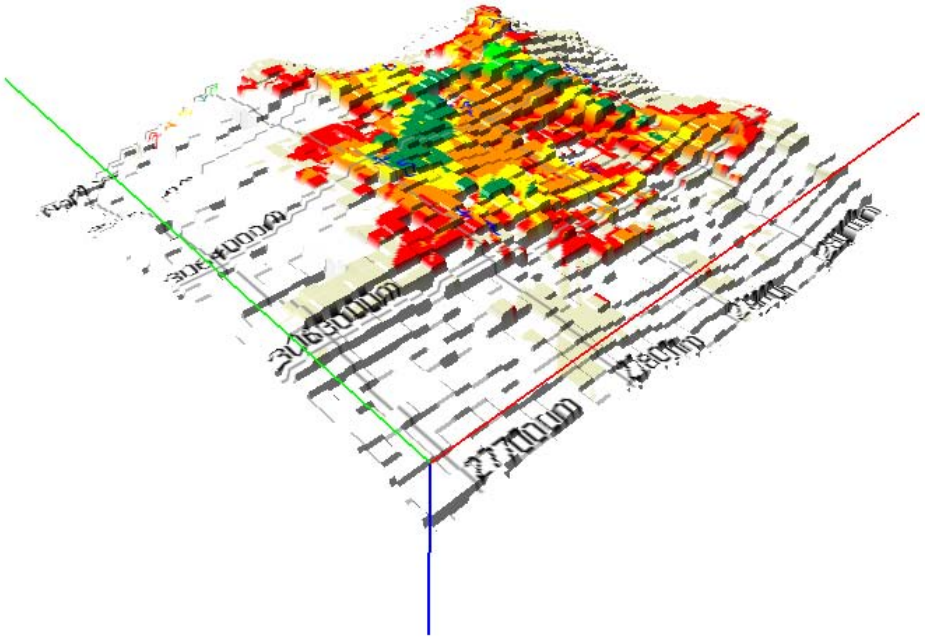
Sound Velocity Profile (SVP), Ray Traces and Transmission Loss (TL) for the deep isothermal layer

For transducers located near the sea bed and the associated well heads and production infrastructure, the dominant feature is the deep isothermal layer, located in this example below ~1100 m. The essential feature of the deep isothermal layer is the constant positive gradient of ~0.016 m/s / m. The variation with location, if the deep isothermal water assumption holds, is negligible; generally less than ~0.001 m/s / m.

The essential effect of this characteristic is upward refraction of sound. Consequently, there is a bottom shadow zone, the maximum range is proportional to the square root of station height above the topology. Which means that ranges are generally refraction not noise limited. Additionally, the sagging line of sound can be obstructed by intermediate terrain.

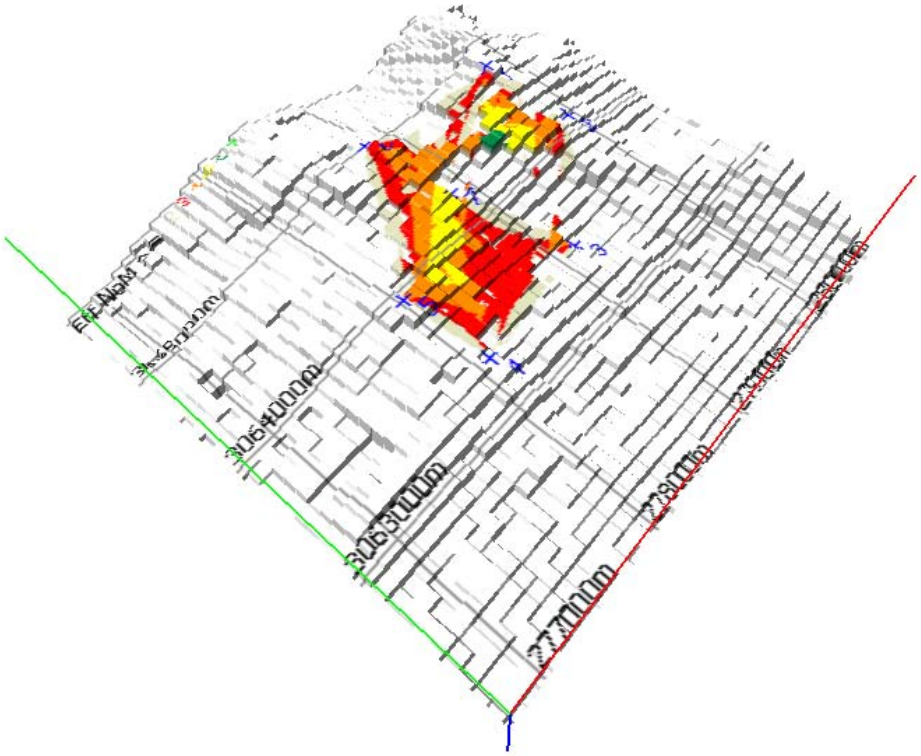
Underwater Example : Gulf of Mexico – Baseline Analysis

Now, some simple modelling permits us to re-evaluate the selected station locations. And, we find that the coverage is actually pretty marginal. The additional parameters used are: stations at 10 m elevation, above the bottom, and receiver 5 m above the bottom.



Number of Measurements : with visibility constraint

Inspection of the number of measurements with visibility constraints applied, the joint effect of terrain obstructions and sagging lines of sound, shows that the cover is appreciably worse than the cookie cutter approach suggests. Nonetheless, the availability of combined 3-D visualisation of the both the bathymetry and the performance measure permits considerable insight into the difficulties presented by the terrain and potential solutions.



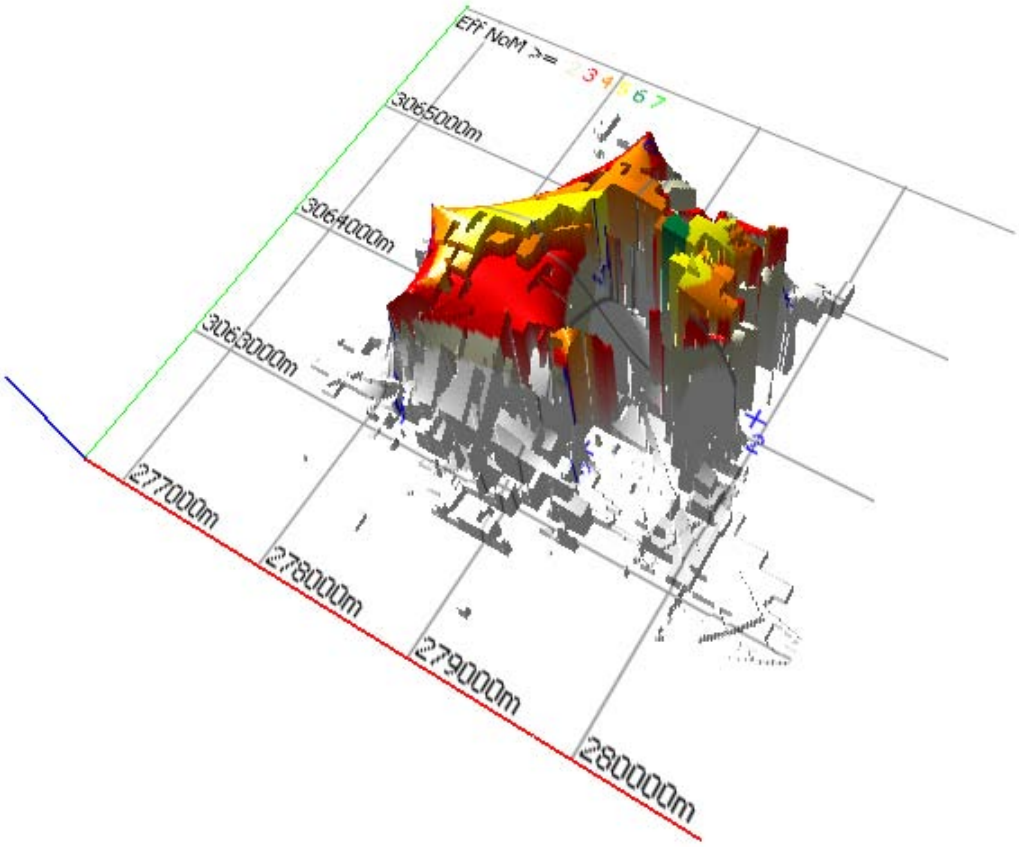
Effective Number of Measurements (Neff) : with visibility constraint

The number of measurements available at a position is however only a necessary condition for positioning. For good positioning with an accurate knowledge of time and the speed of sound, metrics such as Dilution of Precision (DOP) and Support provide a good indication of available performance. However, if clock quality is poor or the sound velocity is approximate the old sailors maxims of working inside the array become applicable.

A new metric, the Effective Number of Measurements (Neff), described later in this paper embodies these maxims in a computational measure. In a nutshell, the metric converts the direction cosines of the observations to a bias vector which it then converts to an Effective Number of Measurements. Alternatively expressed, it computes the number of (truly) independent measurements available.

The application of Neff to our example shows that a significant coverage problem exists in the triangle formed by stations 3, 4 and 7. Note the absence of optimistic coverage outside the array. Neff should be interpreted as an equivalent number of stations with good geometry. Typically for 2-D positioning between two and five stations are desirable, depending on tracking mode, accuracy and QC requirements.

Visualisation of the Neff data, scaled into units of inverse DOP for familiarity and comparison with other DOP values, can also be insightful. Neff is of course related to Tdop, Neff = 1/ $Tdop^2$; see Appendix E. The following plot depicts the preceding data set, the severity of the deficiency in coverage is again evident as is the utility of Open GL and the insight available from these plots.

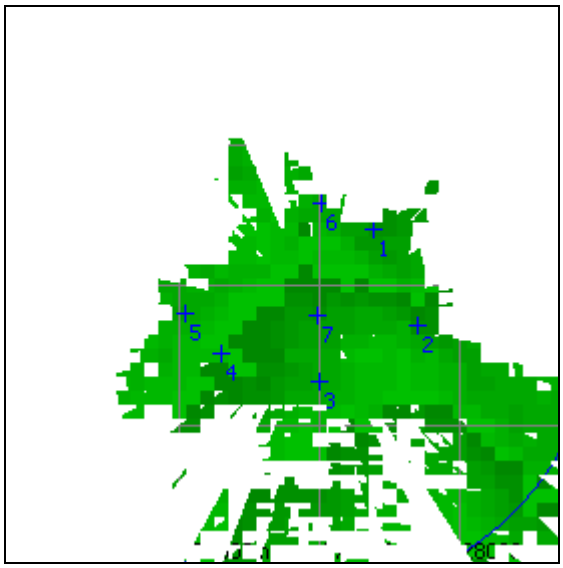


Neff value (1 / DOP units) – with visibility constraint

From the 3-D plots the features which are obstructing lines of sound are almost self evident. The movement of the stations onto the high ground using guidance from the 3-D graphics being straightforward. It should be noted that the initial locations were selected using "2-D" techniques, to demonstrate the utility of sharper tools.

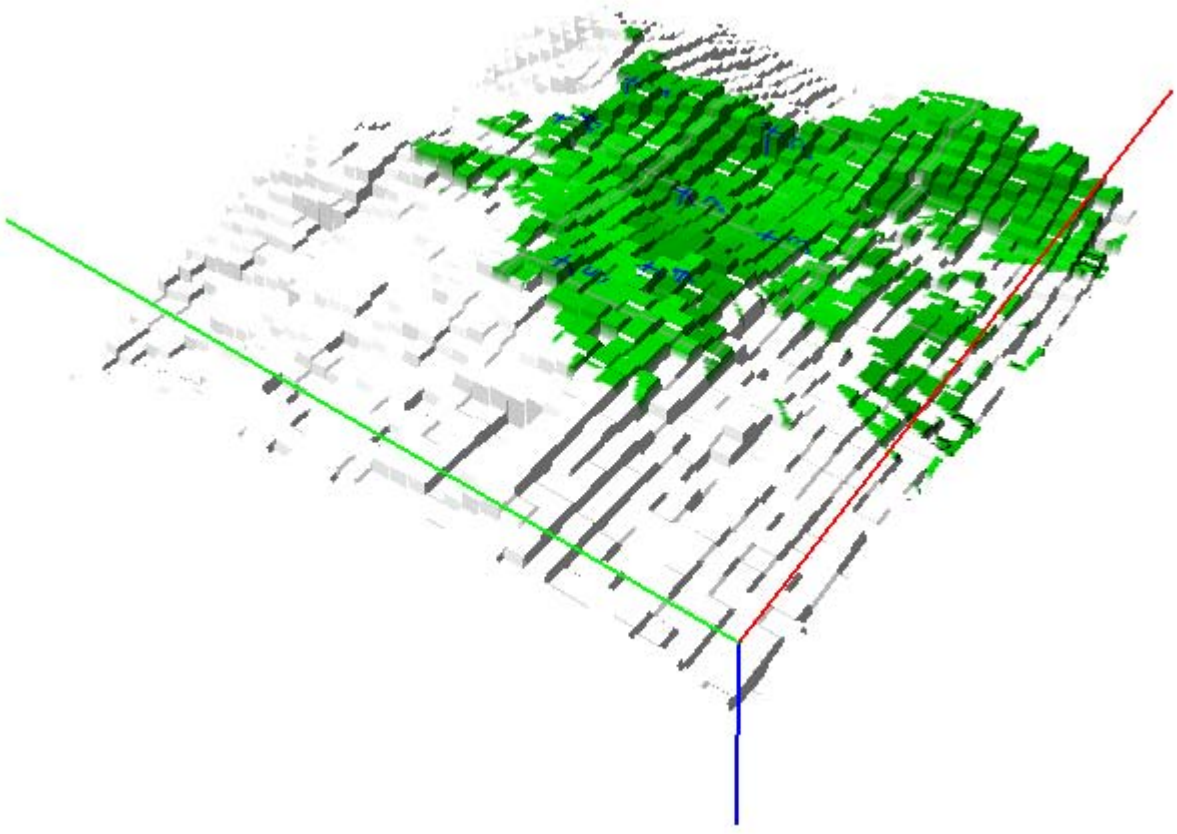
A comparative illustration of the relative power of the old 2-D plots and the new 3-D plots is now provided by using tools from each generation to review the coverage provided by a station. The task is to review the visibility of station 7.

The old approach depicted in the following plot is to shade the visible terrain, in this instance with greenscale contours. Self evidently, coverage can be verified but insight is unavailable. Why lines of sight are blocked being essentially opaque. Subsequent plots use the revised station locations.



Station 7 sight lines – all you get in 2-D

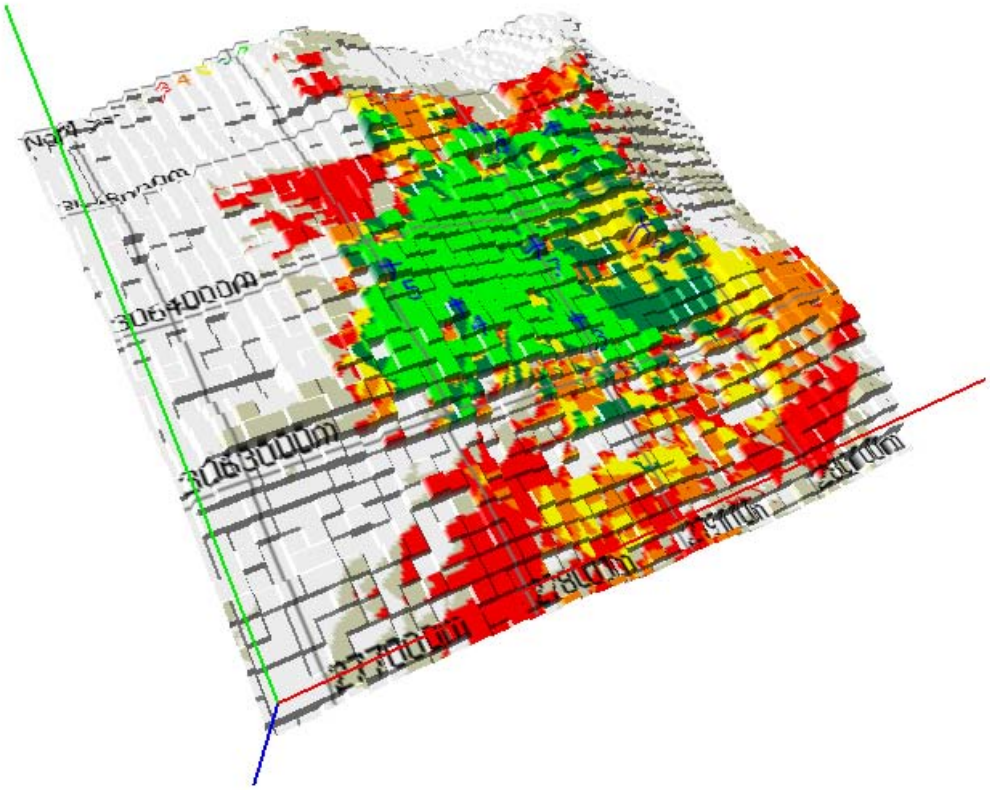
The new approach depicted below enables ready visualisation of the lie of the land. The problematic features are readily identified for appropriate mitigation. For a dataset as coarse and blocky as this one, many of the features are probably artefacts. Additionally, while a "frozen" perspective plot such as this provides considerable insight far more is available from real time interaction with the 3-D graphics.



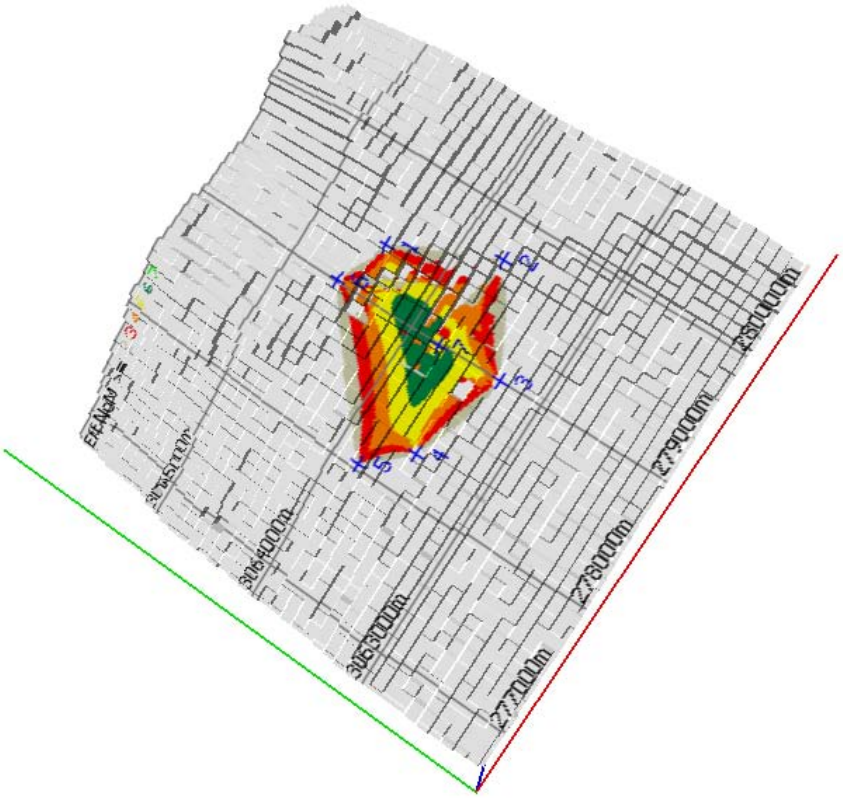
Station 7 : Visibility rendered in 3D

The OpenGL 3-D graphics capability hugely simplifies the task of assimilating and working with the bathymetry data. The task of visualising station and location visibility is significantly assisted, as is the task of optimising station locations.

Underwater Example : Gulf of Mexico – Analysis of Revised Layout



NoM for revised station locations

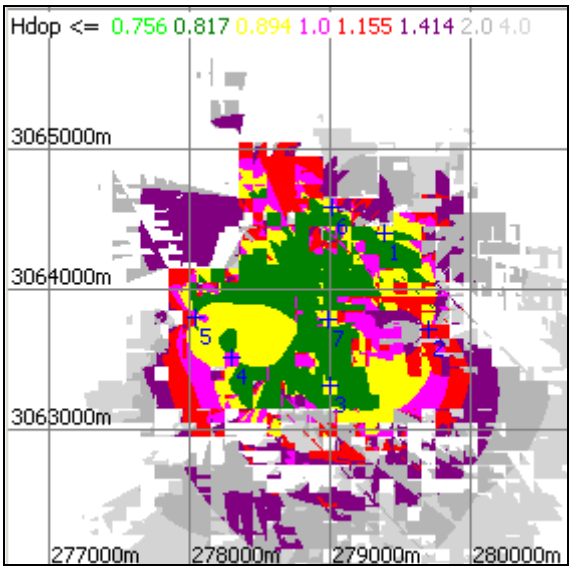


Neff for revised station locations

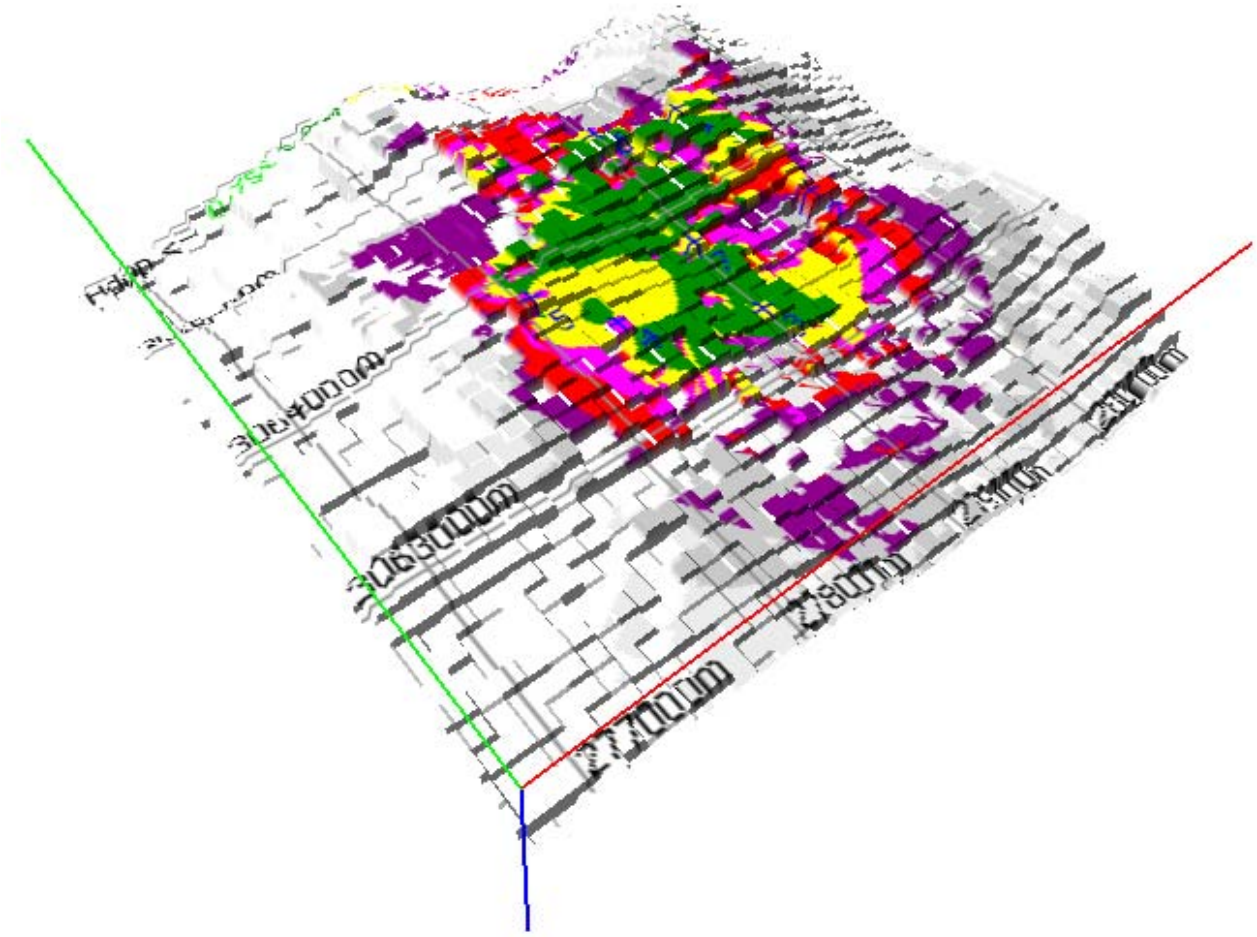
The good news is that the number of stations in contact over the area enclosed by the seven stations generally ranges from five to seven stations. The not so good news is that the number of effective stations

within the array exceeds four only over about two thirds of the enclosed area and that there are a few holes in the coverage. The really good news is that all of this information can be viewed using 3-D graphics tools and readily assimilated. And, don't forget that the Neff plot depicts coverage which is robust to biases and errors; e.g. in clocks and speed of sound.

An alternative perspective on the proposed station layout is an Hdop plot. The horizontal dilution of precision is plotted below for a synchronous system in comparative 2-D and 3-D plots. While the 2-D plot provides a good design record, the 3-D plot is unsurprisingly more insightful.



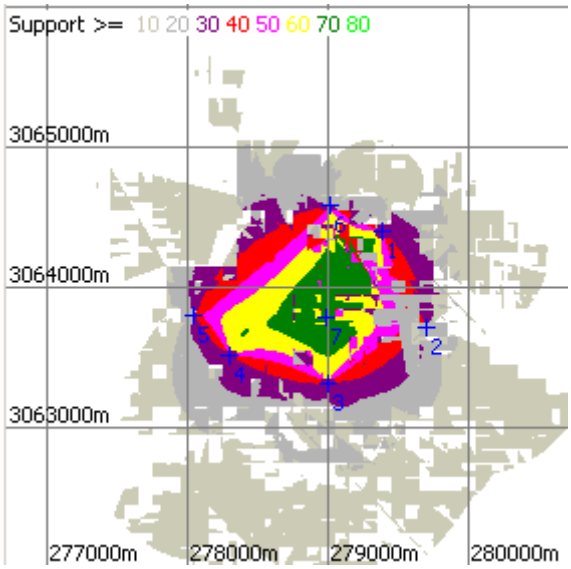
H Dop for revised station locations



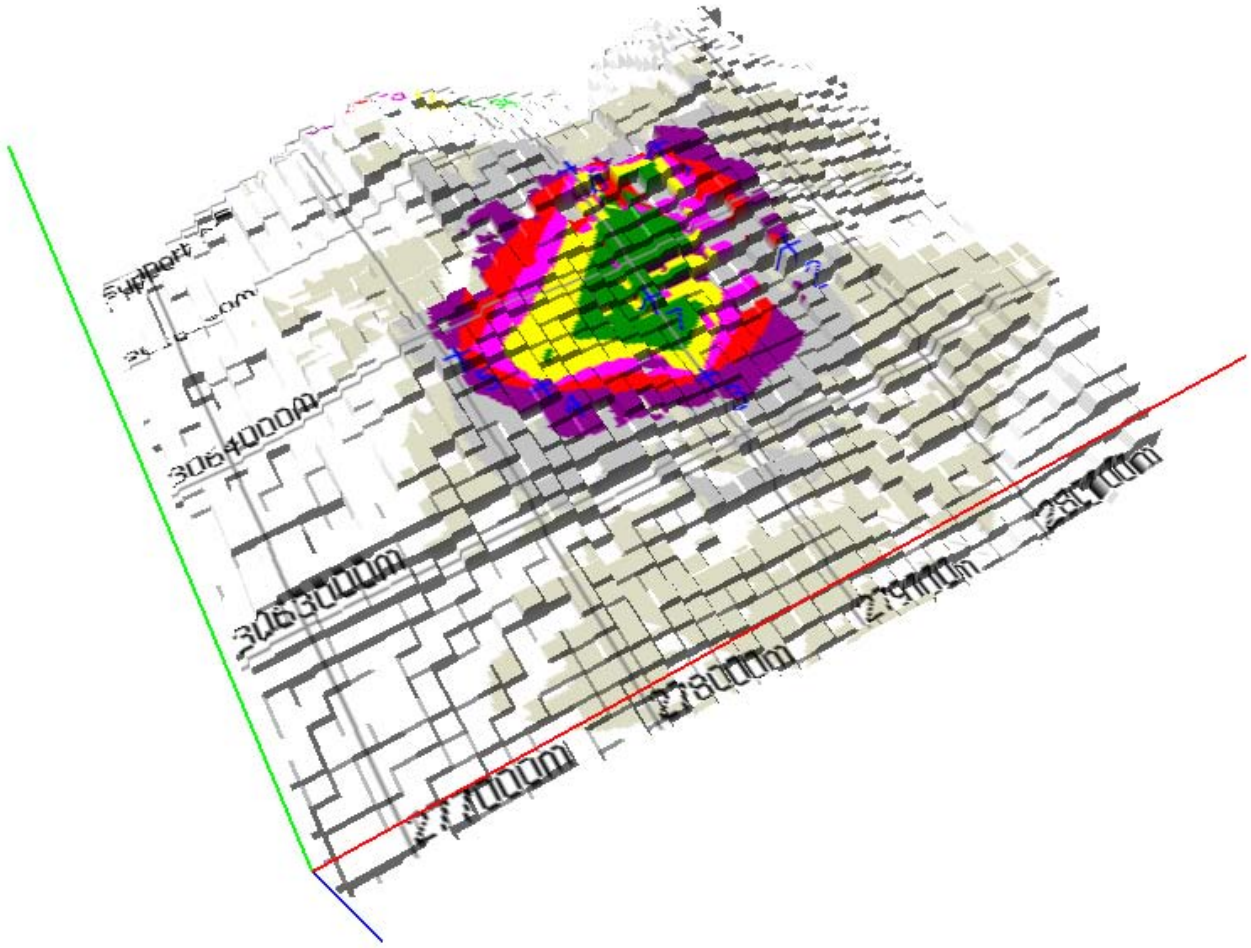
Hdop for revised station locations

A further plot which is generally valuable in evaluating tracking system layouts is the support plot [18]. This depicts how well the positioning of a point is supported by the surrounding stations. The ideal support is encirclement by a ring of stations, which would score 100%. More realistically, enclosure by a hexagon of stations provides support of ~80% and support of 60+% typically delivers good positioning.

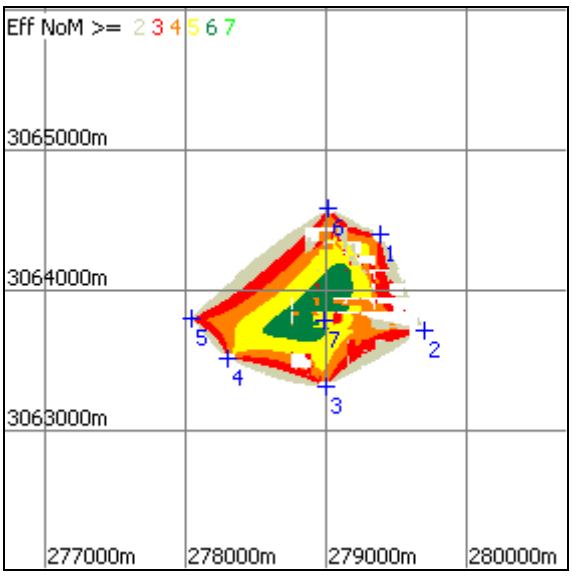
The colour coding used for support endeavours to be consistent with that used for DOP, NoM and Neff with the same colour indicating the same anticipated level of performance for each plot. The theoretical basis for support is of course the well known result that the performance of a position fixing geometry is proportional to volume (3-D) or area (2-D) enclosed by it's stations [19-20].



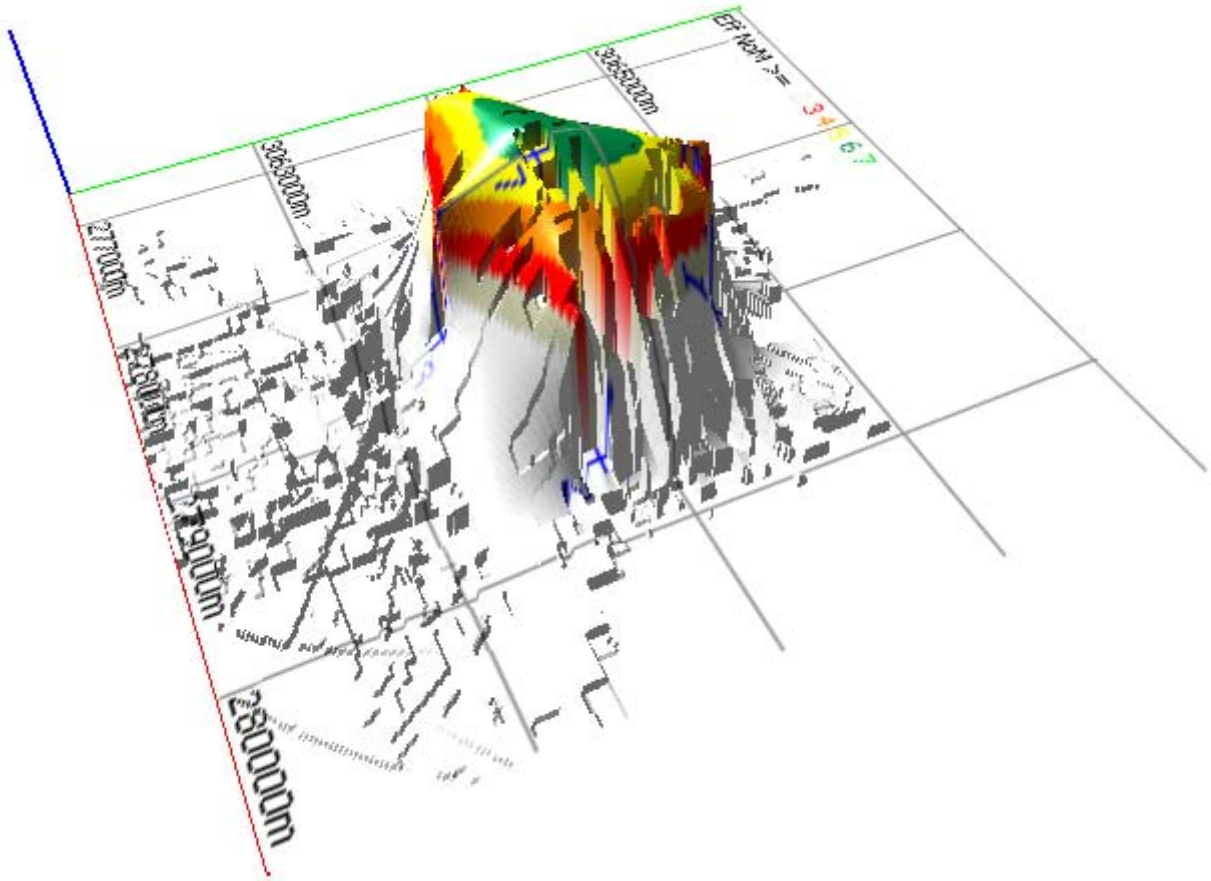
Support plot for revised station locations



Support Visualised on Bathymetry in 3D for revised locations



Neff for revised station locations

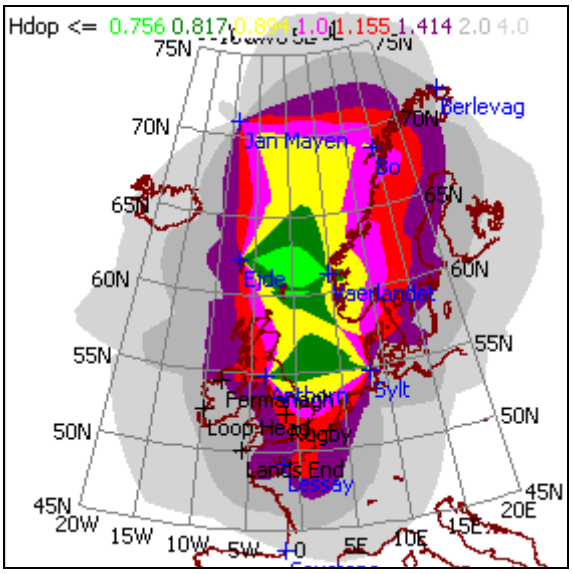


Neff for revised station locations : Neff v XY to visualise features

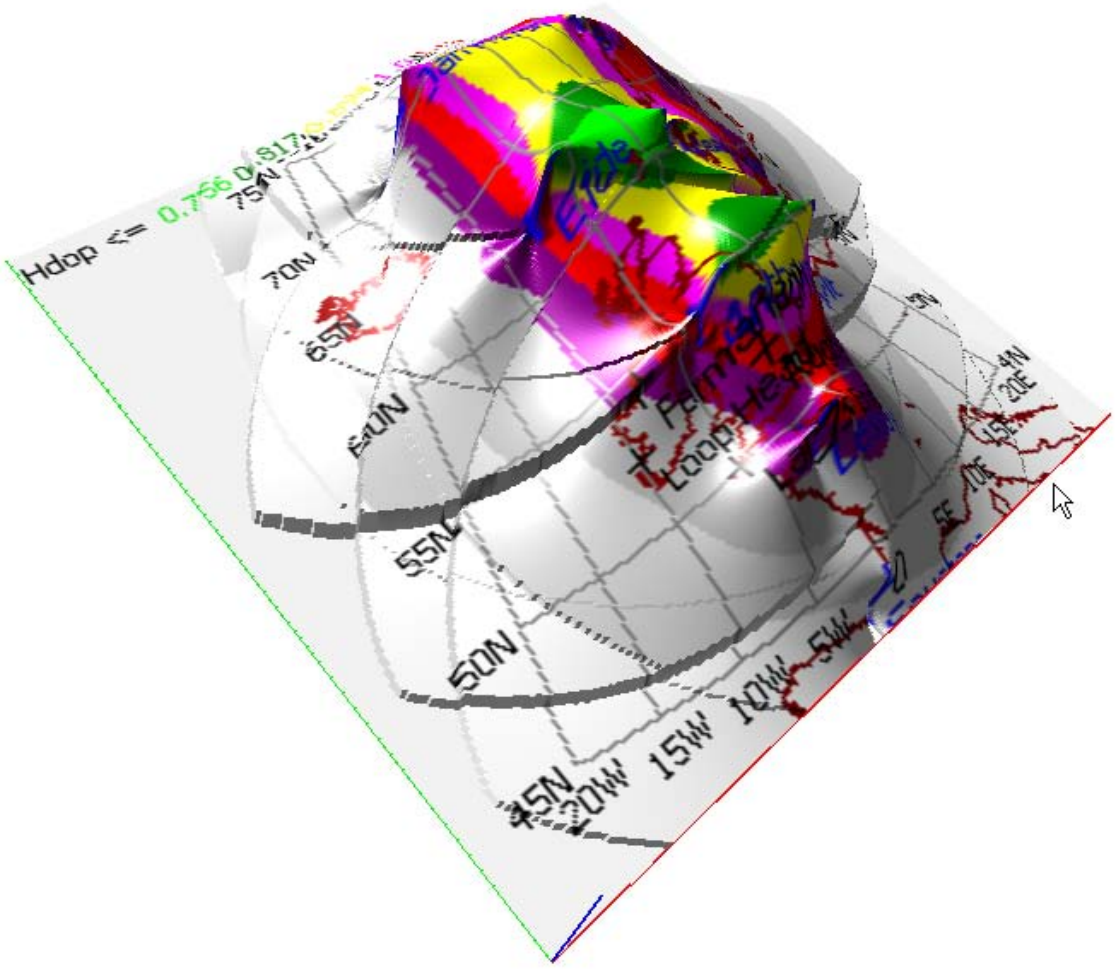
Two final Neff plots illustrate the overall smoothness of the Neff plot, the localised excisions due to masking and the second order sculpturing due to maximum range criteria (2 km for this example) being applied. Self evidently, the sharp tools provided by the NoM, DOP, support and Neff metrics and OpenGL, greatly facilitate the task of visualising, verifying and optimising station layouts to meet complex real world requirements.

Loran Example : The NW European Chain

Our second example presents an analysis, comparing a range of alternative station options, of the NW European Loran chain. The baseline configuration comprises the "French" and "Norwegian" chains: Soustons, Lessay &Sylt and Værlandet, Ejde, Jan Mayen, Bø & Berlevåg respectively. The projection used is a Lambert conformal conic projection with two standard parallels [21-23].

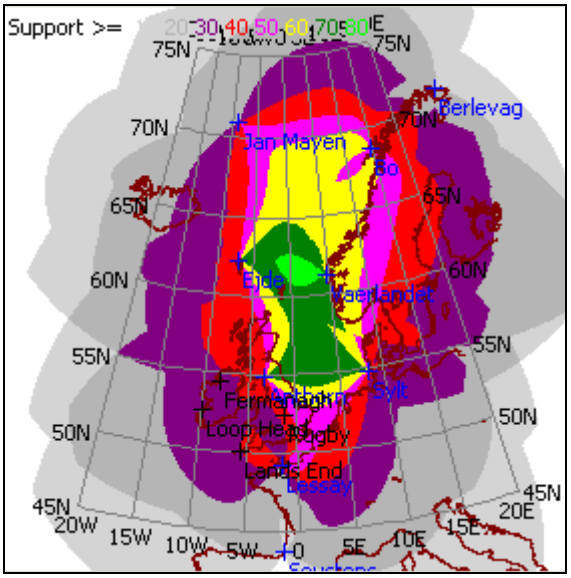


Baseline : Hdop plot

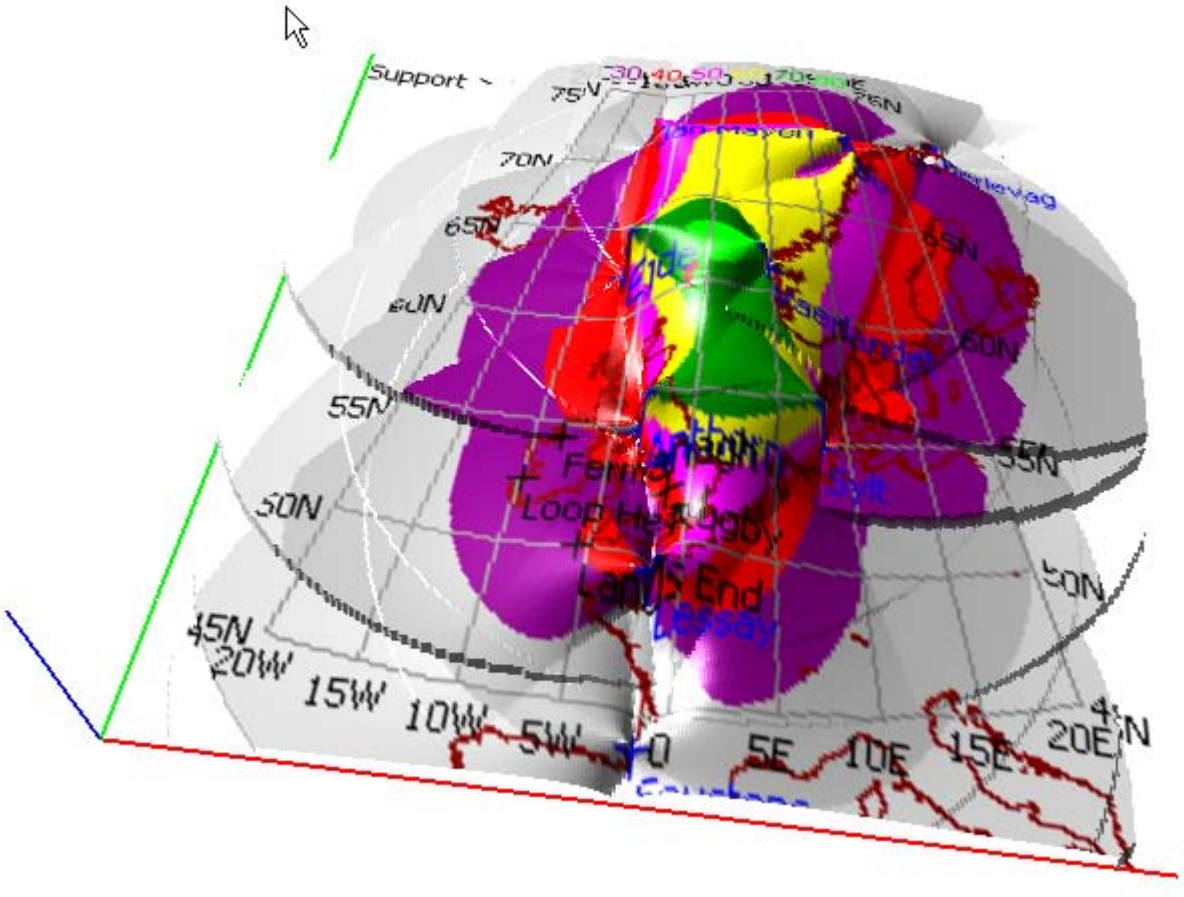


Baseline : Hdop plot in 3-D

The Hdop plots of the baseline configuration, including Anthorn, depicted above clearly show that UK coverage is (unsurprisingly) best in eastern regions. The Hdop thresholds used for colour coding broadly take values of 2 / √N, the Gdop at the centre of a symmetrically distributed N-gon [44]. A range cut off of 2000 km has been applied to the analysis – this is broadly the prudent navigator's limit of dependable ground wave coverage. The 3-D Hdop plot shows the effect of this cut-off as a slight reduction in dop which forms a circular step. That little positioning performance is lost by prudently avoiding sky wave propagation is noteworthy.

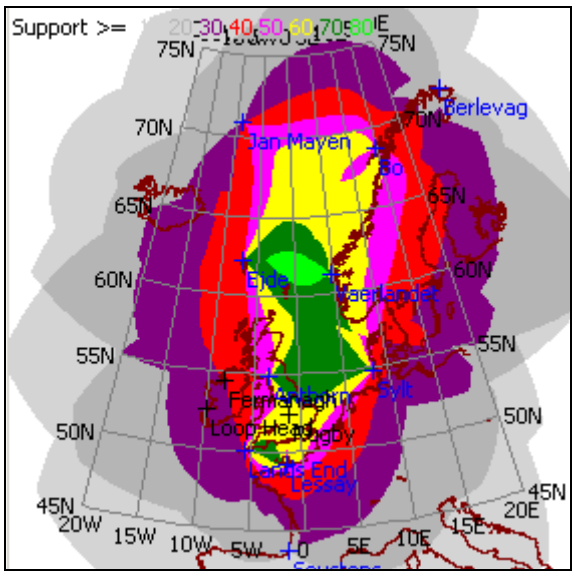


Baseline : Support metric

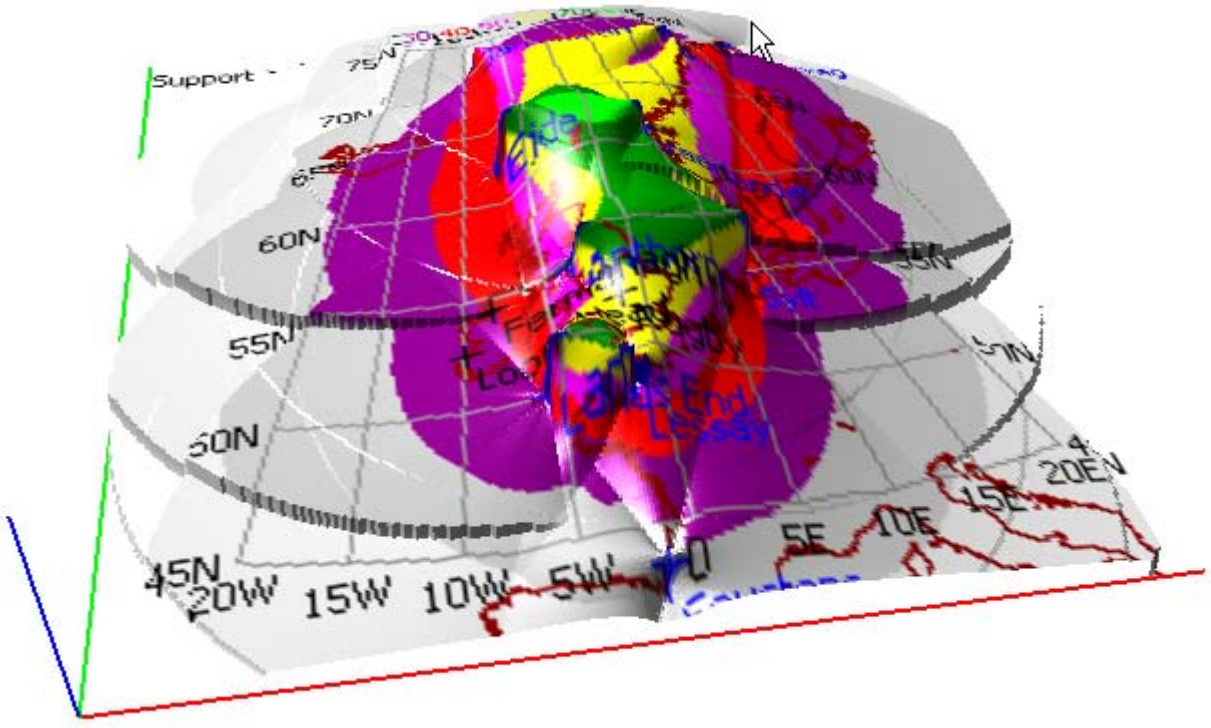


Baseline : Support Metric

The support metric provides an alternative coverage metric. Support [18] which is defined as the extent to which the user is surrounded by stations ranges from 0 to 100%. A prudent navigator would probably consider the coverage adequate for general navigation to the 50% or 40% support boundary, and beyond that for carefully observed and qualified fixes. We now use the support metric to compare alternative station configurations for the NW Europe Loran chain.

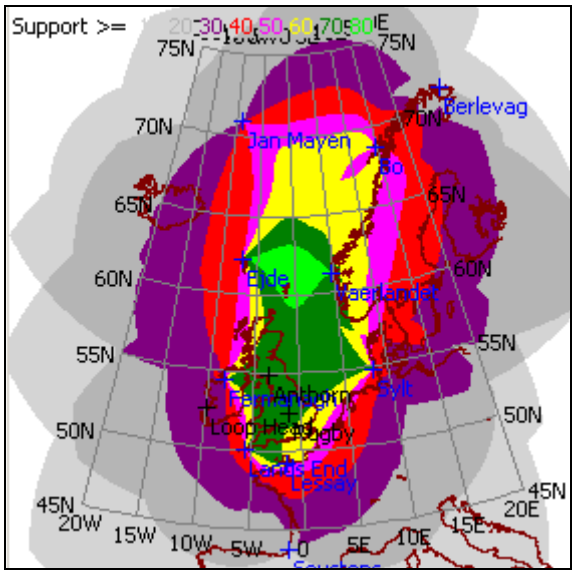


Baseline + Lands End : Support Metric

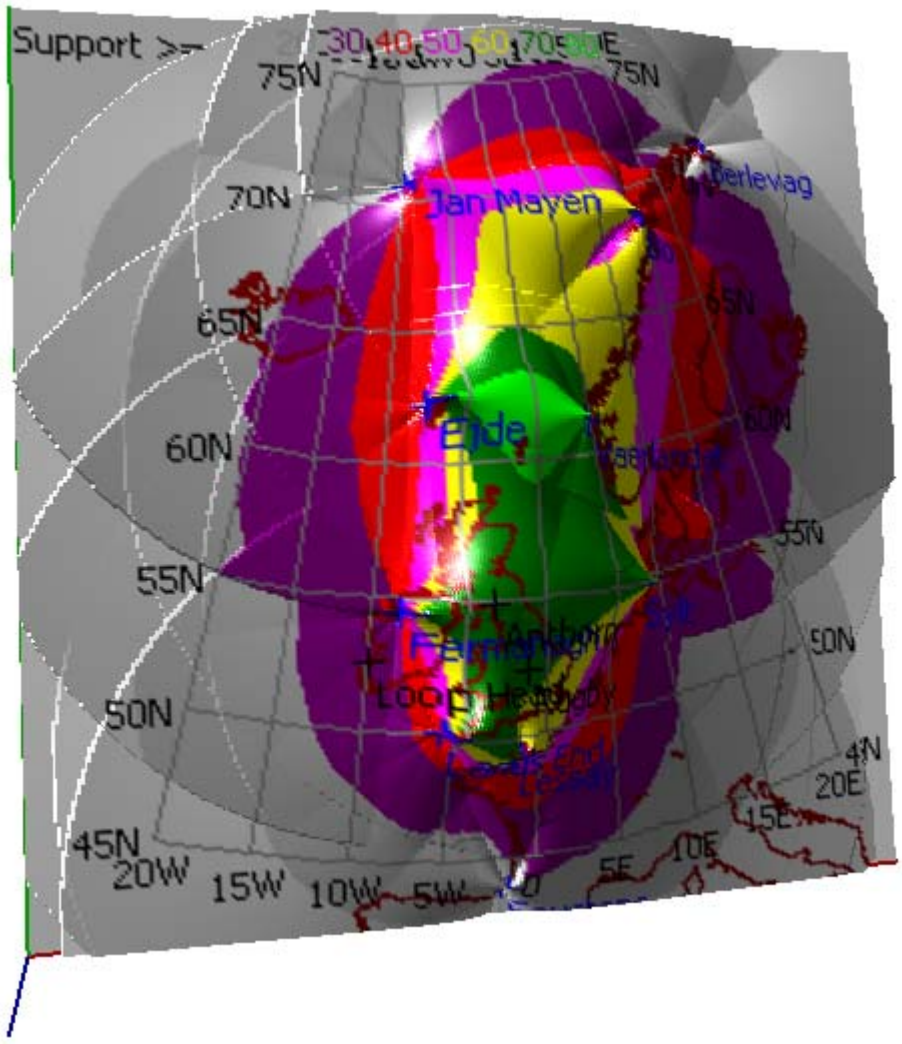


Baseline + Lands End : Support Metric

It is self evident that to enhance coverage of the UK and Ireland a "Westerly" station is required. A potential South Westerly UK location is Lands End. The preceding plots depict the significant improvement in coverage for Southern England which such a station could provide.



Baseline + Lands End + Fermanagh - Anthorn : Support

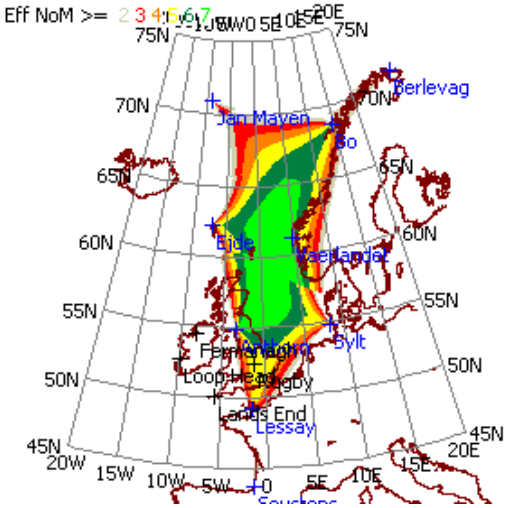


Baseline + Lands End + Fermanagh - Anthorn : Support

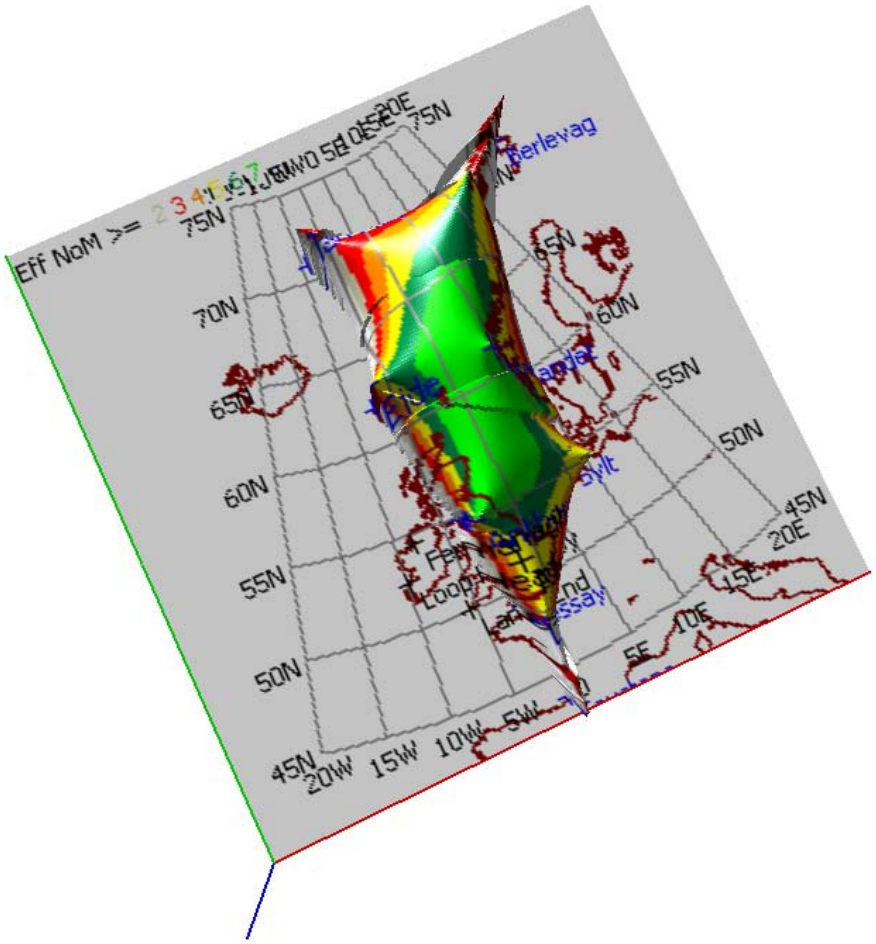
Similarly, it is evident that Anthorn is insufficiently to the West. The option of relocating this transmitter to a station is Fermanagh is depicted above. The significant improvement in coverage in Ulster, Wales, Scotland and England is self evident. The benefits of a complimentary site in Eire are self evident. Similarly, a station on the Outer Hebrides would enhance Scottish coverage.

Regrettably, it is now time to consider whether coverage really is as good as it seems. To provide robust performance in the presence of biases and errors we must “work inside the array”. An appropriate metric to evaluate system performance where time (epoch) as well as position require to be estimated is the Neff (number of effective stations) metric. Obviously estimation of X Y T requires a minimum of three valid pseudo range measurements.

Using this criterion, it is Immediately apparent that good coverage is provided in the Eastern UK. Self evidently, the accurate prior survey of receiver locations could provide some mitigation of this limitation. However, the importance of western stations is very evident.

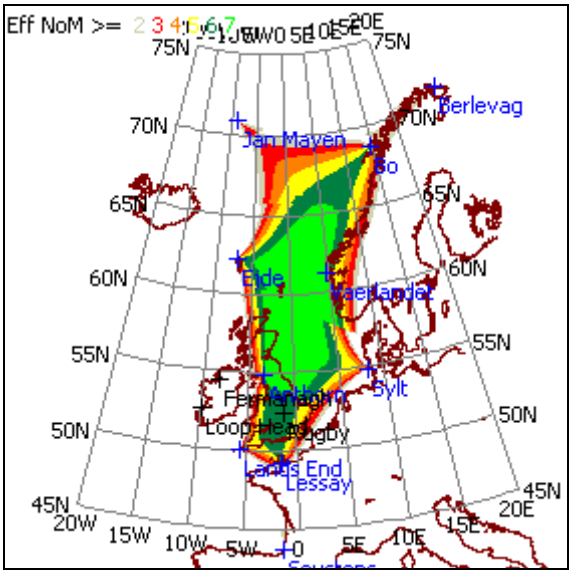


Baseline : Number of Effective Stations

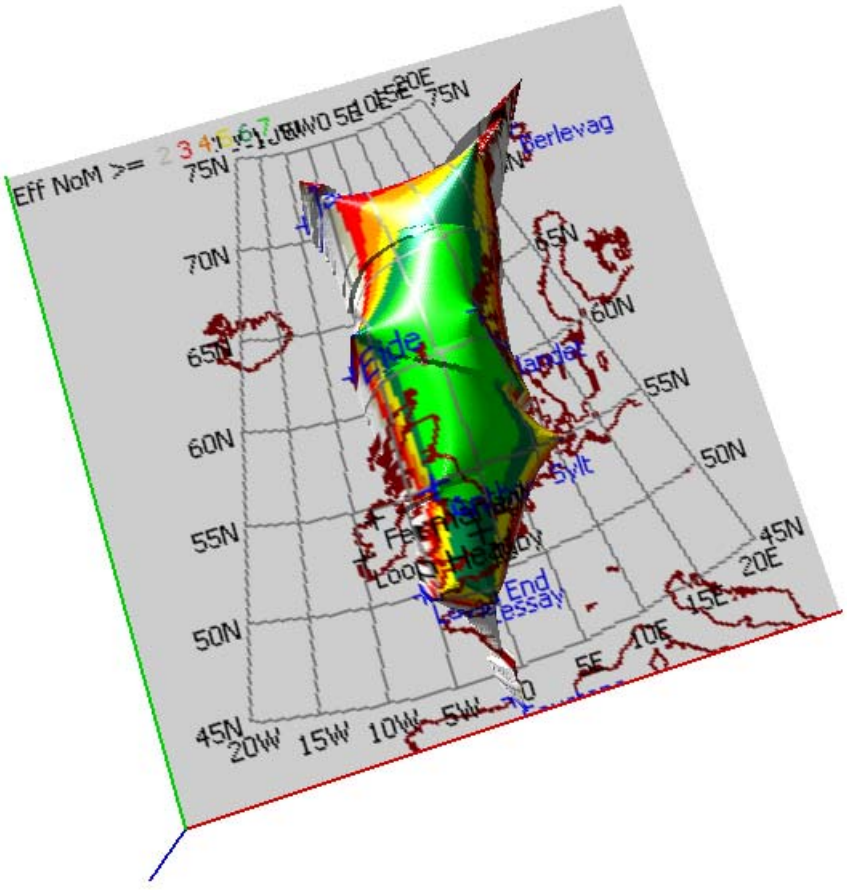


Baseline : Number of Effective Stations

The significant improvement in timing coverage in England and Wales provided by a station at Lands End is clear. The old saw about working inside the array being clearly demonstrated.

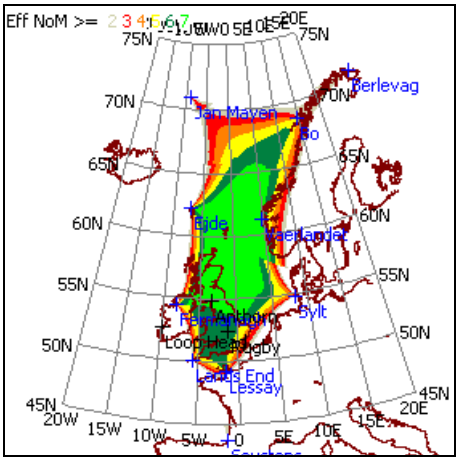


Baseline + Lands End : Number of Effective Stations

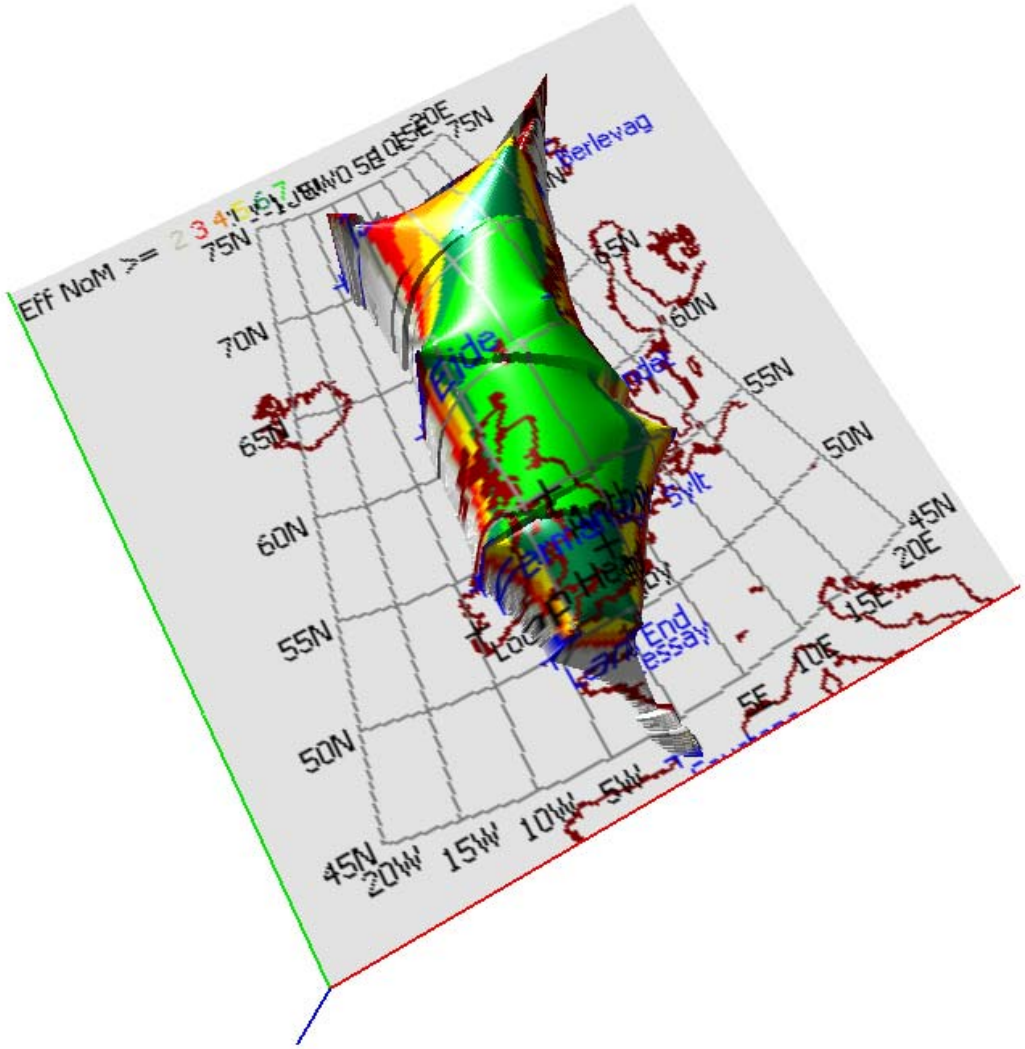


Baseline + Lands End : Number of Effective Stations

The effect of also moving the Anthorn station to Fermanagh is even more impressive. Good coverage being provided over all but the Western extremities of Scotland.

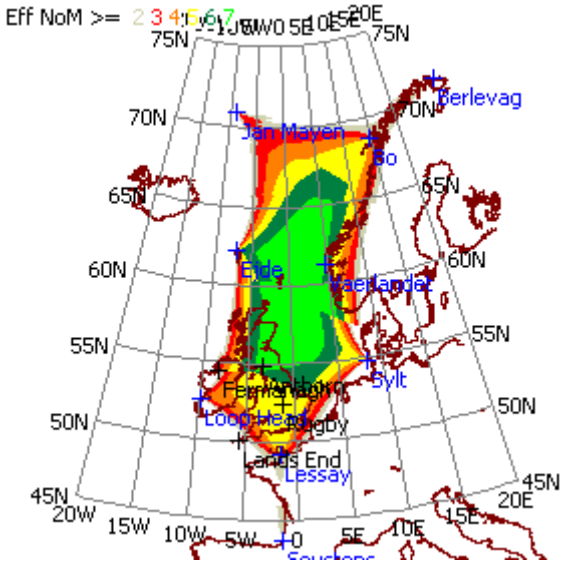


Baseline + Lands End+ Fermanagh - Anthorn : Number of Effective Stations



Baseline + Lands End+ Fermanagh - Anthorn : Number of Effective Stations

Finally, we consider what might have been. The coverage provided by a station at Loop Head without stations at Anthorn, Rugby or Lands End is depicted below. Self evidently, a station in the West of Ireland offers significant benefits in Eire as well as to navigators in the Irish Sea and in Southern England.



Baseline + Loop Head - Anthorn : Number of Effective Stations

Conclusions

Sharp tools have been described and demonstrated which significant simplify the analysis and synthesis of base station locations for tracking and positioning systems. The examples offered have ranged from the layout of deep sea stations for near bottom positioning tasks in Oil and Gas applications where locating the stations to provide adequate line of sound coverage is a major challenge, through pedagogic examples (in the appendices), to the analysis of alternate station locations for the NW European Loran system.

A metric for the number of effective measurements, derived from the dilution of precision (DOP) analysis of asynchronous systems in which emission time as well as position require to be estimated has been introduced, This metric both reports in readily assimilated units, the number of optimally spaced stations available at a location, and provides a strong indication of robustness to biases due to clock offsets and speed of sound errors. Additionally, the metric's computation from the misclosure of the direction cosines used in computing position estimates provides a strong physical insight .

The support metric is also derived from the DOP analysis, the area / volume enclosed by the direction cosines is for simple problems proportional to the denominator of the least squares iteration and therefore a good indicator of the solution's sensitivity to noise. The metric reports the degree to which a location is surrounded by stations, information which provides strong physical insight into tracking coverage. Suitably scaled and interpreted DOP values do of course provide another, classic metric for the analysis of positioning problems.

OpenGL is of course the complementary sharp tool from computer graphics which permits the visualisation of 4-D data : X, Y, Z and metric. The effectiveness of 3-D graphics in facilitating analysis and synthesis tasks which were extremely difficult using 2-D graphics cannot be overstated. The ease with which OpenGL, GPUs (Graphical Processing Units) and 3-D rotation mathematics (Quaternions) permits the manipulation of data equally cannot be overstated.

In summary, the judicious use of computer hardware (GPUs), software (OpenGL), analytic metrics (Neff, Support and DOP) and mathematical tools (e.g. Least Squares Estimation and quaternions) has enabled a new level of insight in the analysis and synthesis of station layouts for navigation and positioning.

Emeritus Solutions Ltd is experienced in the analysis and synthesis of underwater tracking range layouts, with a profound understanding of the applicable mathematics underpinning their analysis tools. Emeritus Solutions Ltd can provide (bespoke) tools for requirements ranging from range layout analysis and synthesis, through the computation of tracking solutions and sensor boxing, to the detailed analysis of range and environmental data. Additionally, Emeritus Solutions Ltd has comprehensive underwater acoustic, digital signal processing, computing and electronics competencies. These offerings can be provided on a turnkey, service or consultancy basis.

Appendices

These appendices first identify key aspects of the underwater positioning problem, which both differentiate it from the GNSS (satellite navigation) problem and make the computation of accurate positions a challenging task. The trilateration problem is then described “In a Nutshell”, a key Gdop bound result is presented and the Neff concept developed; all with the minimum of mathematics. Performance analyses using the DOP paradigm are then presented for a range of specimen tracking range geometries. Our support metric technique is then illustrated using specimen geometries. Finally, our number of effective measurements metric is also demonstrated on specimen geometries. Examples of the application of these metrics to real world problems have of course been provided in the body of the paper.

Appendix A : Underwater Positioning – Key Differences from GNSS

Almost all of the trilateration literature [see, for example, 19-20, 25-44] addresses radio navigation; from early works on Loran (and Gee), through the early days of GPS, to contemporary GNSS and mobile phone tracking. Consequently, the mode of operation and important issues for underwater positioning can differ appreciably from those of interest in radio navigation. The following paragraphs identify some of the key differences between underwater positioning and GNSS.

Clock performance is typically a significant issue for underwater positioning systems. Somewhere in the system sub-Rubidium performance will commonly “be a problem”. At a minimum this will require the system to estimate emission times and work from pseudo, rather than absolute, ranges. One consequence, is that an additional pseudo range measurement is required – which in turn requires a higher sensor density. Another, more significant, consequence is that position fixing outwith the sensor array is very imprecise – an issue which is perhaps less problematic for GNSS systems.

Propagation velocity is invariably an issue. The speed of sound in water invariably varies with depth, time of year, time of day and location. At a minimum it must be directly measured, and such is its variation with depth that its harmonic mean must be used. The most radical difference from radio navigation is that the uncertainty of propagation velocity can be sufficiently large to significantly affect position solutions. Again, working within a (preferably symmetric) sensor array is the most effective mitigation.

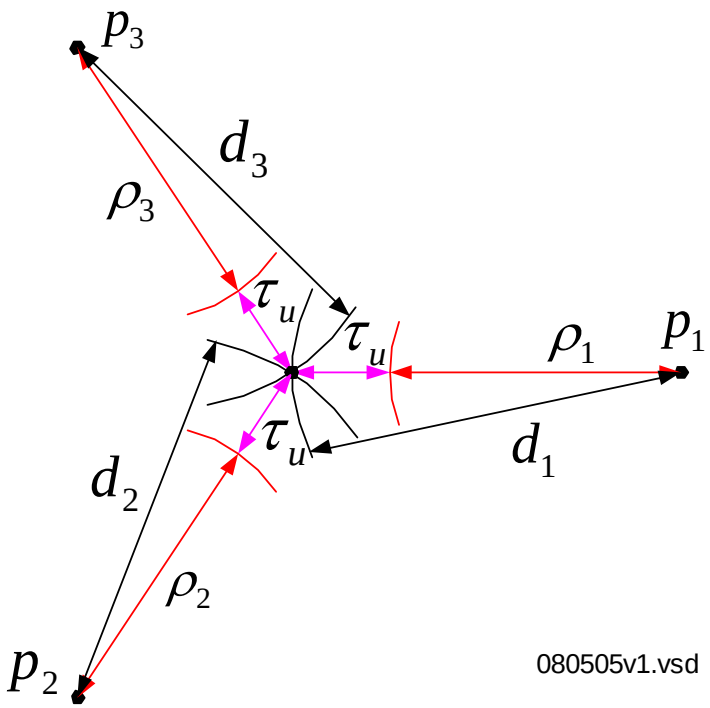
Unfortunately, many other difficulties are attributable to propagation velocity – specifically to its variation with depth. GNSS requires mitigation of ionospheric and tropospheric propagation effects at low elevation angles, and has the option of masking out low elevation satellites. As most channels of interest are longer than they are deep, underwater positioning is typically performed at low elevation angles. Consequently, shadowing and variable path lengths due to refraction are the norm, both imposing (sometimes absolute) constraints and requiring ingenuity in system design.

An important consequence of the “shallow” environment in which underwater positioning is performed is that in general depth is measured and if necessary telemetered rather than estimated by trilateration. A positive aspect of this circumstance is that the positioning problem becomes two dimensional, and its solution requires one fewer pseudo range measurement. There are of course exceptions, such as sensor boxing, where the slant range is only a few times the water depth.

Finally, we shall simply mention some of the “second order” issues which if not appropriately mitigated can comprehensively “ruin your day”, for example: anisotropic propagation velocity due to currents, multipath and surface reverberation.

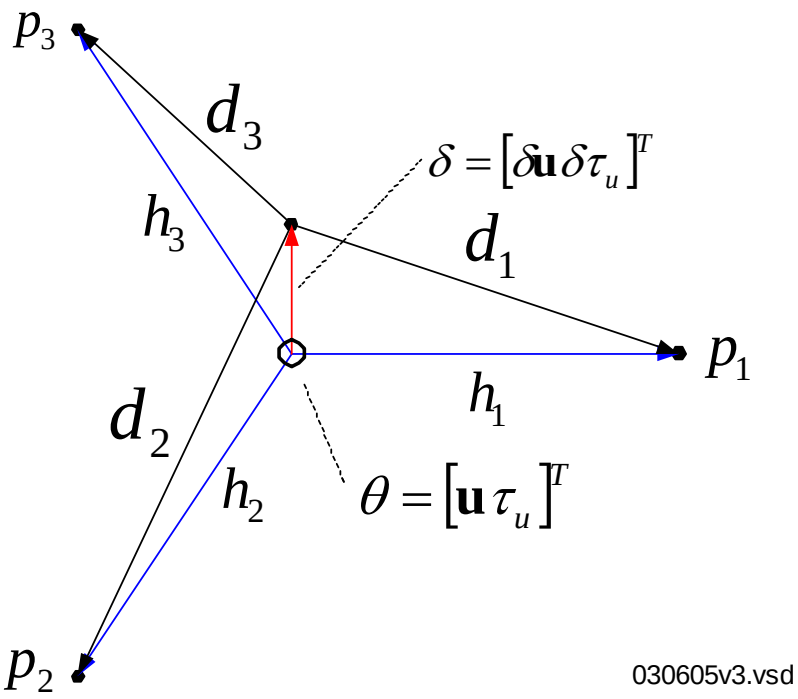
Appendix B : Trilateration - In a Nutshell

The preceding appendix indicated that enclosed geometries with significant symmetry are, if not always possible, certainly desirable. Therefore, as these are in any case the simplest trilateration geometries to analyse, we shall use them to set out the trilateration stall. General theories and analytic tools are of course available for arbitrary geometries. However, simple geometries yield both useful rules of thumb and valuable insights.



Trilateration Problem in 2-D

The preceding figure depicts the trilateration problem for three sensors deployed in an equilateral triangle. This configuration uses the minimum number of sensors for estimating 2-D position (x & y) and the user clock (epoch of emission). Further regular n-gons squares, pentagons, hexagons, etc can of course implement over-determined solutions. The observables are the pseudo ranges ρ_i which differ from the true ranges d_i by the user clock offset t_u and additive noise terms n_i.



Trilateration Iteration in 2-D

The preceding figure depicts the solution framework. The sensor geometry is described by the direction cosines $\underline{h}_i$, which are aligned with the user – sensor vectors h_i. The estimated user position is given by the column vector θ with elements x, y, z and t_u. Under the standard assumptions of independent, equal, zero

mean, Gaussian noise processes the estimator for the user state and the covariance of the estimate may be stated as

$$\varepsilon = \rho - \underline{H}.(P - \theta)$$

$$\delta\theta = \left(\underline{H}^T \underline{H}\right)^{-1} \underline{H}^T \varepsilon$$

$$\mathrm{cov}(\theta) = \sigma_r^2 \left(\underline{H}^T \underline{H}\right)^{-1} = \begin{bmatrix} \sigma_{xx}^2 & . & . & . \\ . & \sigma_{yy}^2 & . & . \\ . & . & \sigma_{zz}^2 & . \\ . & . & . & \sigma_{tt}^2 \end{bmatrix}$$

In brief, the ε is the vector of innovations or residuals, a mixture of (small) offsets and noise which the iterative estimator (for θ) will minimise. θ is the estimate of position and user clock offset x, y, t_u and potentially z which constitutes the solution to the trilateration problem. δθ is an increment in the iterative computation of θ. Cov(θ) is the covariance of the position estimates obtained for a specified sensor geometry $\underline{H}$.

DOP can now be defined, for the simple case where the ranging errors have equal variances, as the ratio of cov(θ) to σ_r^2. In particular, the diagonal terms and their aggregates provide expressions for Xdop, Ydop, etc and for Hdop, Pdop and Gdop. These DOP values provide an indication of estimator performance for a given user / sensor geometry, and are the standard method of assessing system performance [see, for example, 35, 35, 42 & 43].

$$Xdop = \sqrt{(\underline{H}^T \underline{H})^{-1}{}_{xx}} = \sqrt{\sigma_{xx}^2 / \sigma_r^2}; etc \ \ for \ Ydop \ et \ al$$

$$Hdop = \sqrt{(\underline{H}^T \underline{H})^{-1}{}_{xx} + (\underline{H}^T \underline{H})^{-1}{}_{yy}} = \sqrt{(\sigma_{xx}^2 + \sigma_{yy}^2)/\sigma_r^2}$$

$$Pdop = \sqrt{(\underline{H}^T \underline{H})^{-1}{}_{xx} + (\underline{H}^T \underline{H})^{-1}{}_{yy} + (\underline{H}^T \underline{H})^{-1}{}_{zz}} = \ldots$$

$$Gdop = \sqrt{(\underline{H}^T \underline{H})^{-1}{}_{xx} + (\underline{H}^T \underline{H})^{-1}{}_{yy} + (\underline{H}^T \underline{H})^{-1}{}_{zz} + (\underline{H}^T \underline{H})^{-1}{}_{tt}} = \ldots$$

Obviously, Gdop the geometric dilution of precision is a multiplicative factor relating the uncertainty of the position estimate to the user range error σ_r^2. Evidently, a low value of Dop is desirable – just as a good crossing angle for two lines of position is 90°. A good way of obtaining a feel for Gdop values is to examine them in 2-D space for regular n-gons. From this we can obtain an indication of minimum values and their location, the typical range of values, and for the regular n-gons which are the building blocks of most tracking ranges some insight into the interaction between sensor geometry and DOP.

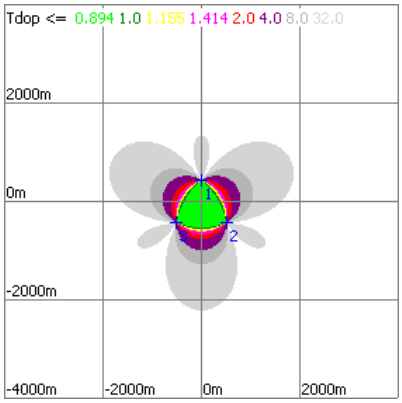
Levanon [44] derived a simple expression for Gdop at the centre of a regular polygon

$$Gdop = 2 / \sqrt{N}$$

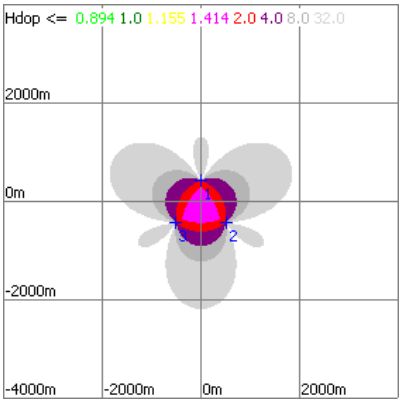
evidently Gdop reduces as the square root of N, that is to halve Gdop the number of sensors must be quadrupled. This square root relationship between the number of (independent) measurements and the accuracy of the estimate is of course unsurprising. Nonetheless, it provides a strong indication of the general relationship between the number of sensors and accuracy. Additionally, it can be shown that regular polygons minimise dop.

Appendix C : GDOP Analysis Examples

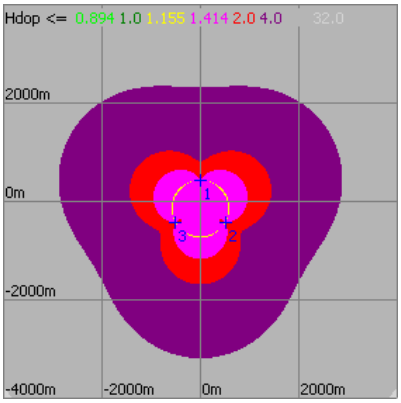
It does of course remain to partition Gdop between Pdop, the (xyz) position dilution of precision, and Tdop, the time dilution of precision. In the 2-D case one can work with either pseudo ranges, where the clock offset / emission time is estimated, or with absolute ranges where the clock offset / emission time is taken as known. As the following plots show, the use of an absolute range system promises improved coverage, especially outside the sensor array.



Tdop – pseudo ranges



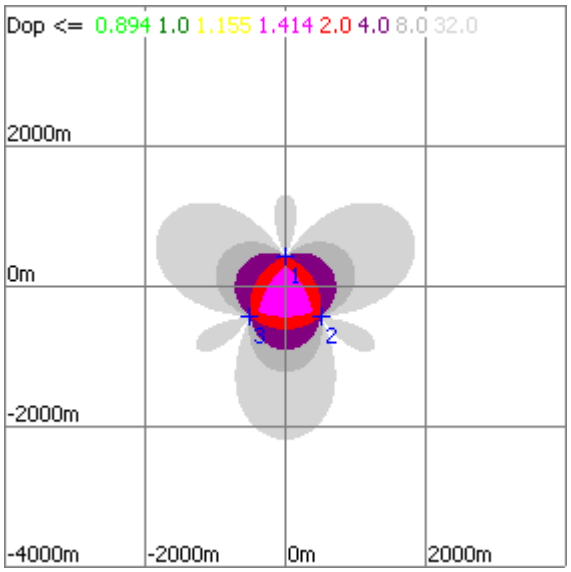
Hdop – pseudo ranges



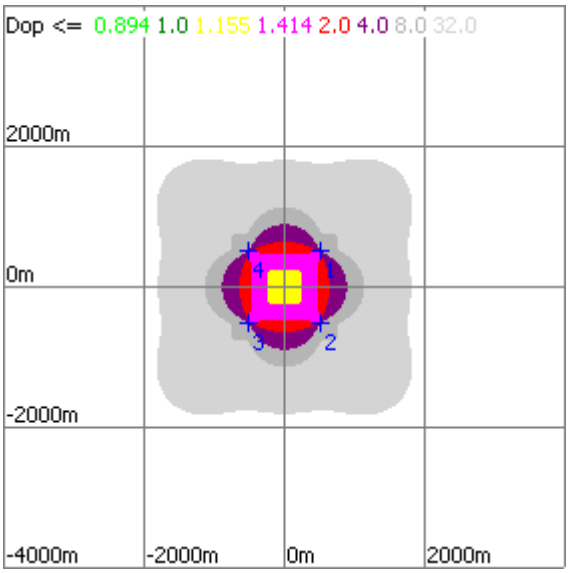
Hdop - absolute ranges

Important points to note are that in the pseudo range case Hdop and to a greater degree Tdop are only low inside the triangular sensor array. Another important point is that the improved coverage of the absolute range (synchronous) case is predicated on synchronisation; it can readily be shown that if this assumption is violated the position fixes are totally compromised. The benefits of working inside the sensor array, preferably with a balanced sensor geometry are evident.

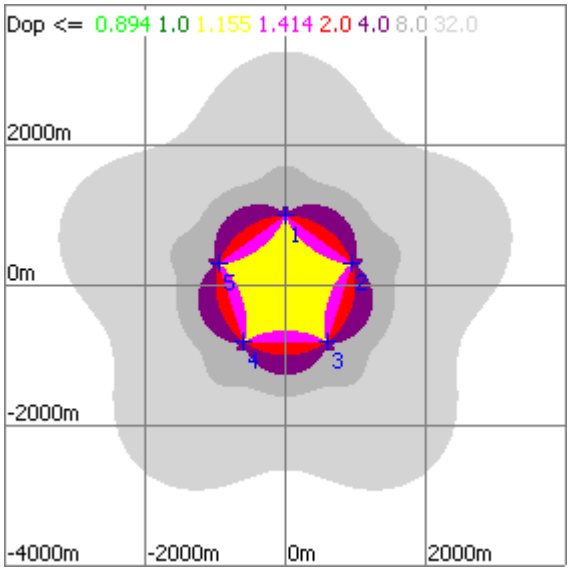
Our illustration of the trilateration performance of n-gons is completed by the following Gdop plots for a triangle, a square, a pentagon and a centre point hexagon. These figures include all the regular polygons commonly used as elements of tracking ranges. The centre-point hexagon is of course a tiling of equilateral triangles, the tiling commonly used for tracking ranges. Regular tilings of the plane are of course only available for triangles, squares and hexagons. Additionally, analysis of normalised coverage (area / sensors) for a regular array under the constraint of a maximum inter sensor spacing clearly demonstrates the optimality for many requirements of a regular tiling of equilateral triangles.



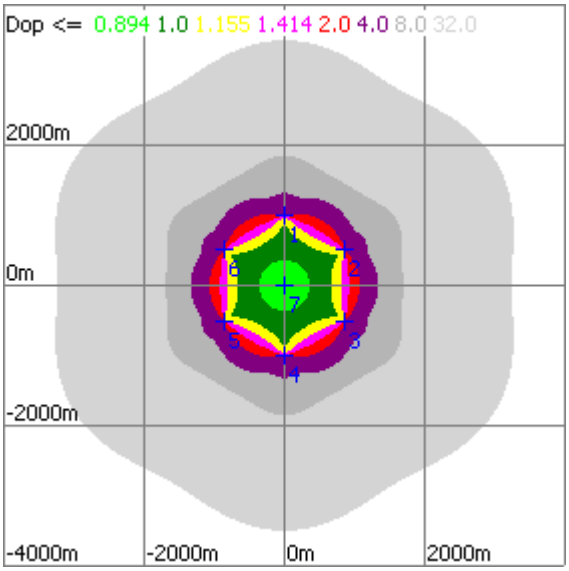
Gdop - triangle



Gdop - square



Gdop – pentagon



Gdop – centre-point hexagon

Again, the location of good coverage is within the sensor array especially where the sensor geometry is balanced. Nonetheless it is evident that Dop has limitations as an analytic metric. Firstly, it does not implicitly identify the weakness of working outside the sensor array. Secondly, it is difficult to normalise for an effective visual appreciation of support for tracking at all locations.

Appendix D : Support - An Alternative Metric to DOP

The simplest way to motivate our alternative metric is to recast the covariance expressions firstly in terms of determinants and then in terms of the lengths, areas and volumes defined by the direction cosines of the $\underline{H}$ matrix.

The expression for DOP can be reformulated in terms of the trace of the matrix $(\underline{H}^T\underline{H})^{-1}$

$$Gdop = \sqrt{(\underline{H}^T\underline{H})^{-1}_{xx} + (\underline{H}^T\underline{H})^{-1}_{yy} + (\underline{H}^T\underline{H})^{-1}_{zz} + (\underline{H}^T\underline{H})^{-1}_{tt}} = \sqrt{tr(\underline{H}^T\underline{H})^{-1}}$$

Now, if $\underline{H}^{-1}$ exists the trace can be rewritten in terms of $\underline{H}$'s adjoint matrix and determinant

$$tr(\underline{H}^T\underline{H})^{-1} = \left(1/|\underline{H}|^2\right)\sum_{ij}(h'_{ij})^2$$

where $\underline{h}'_{ij}$ is the ij th element of $\underline{H}$'s adjoint matrix.

As the denominator dominates this formulation for Gdop (we shall offer insight as to why later) maximising the discriminant $|\underline{H}|$ should maximise Gdop. Now, it is well known, that that the determinants of real n x n matrices are equivalent to volumes in n-dimensional Euclidian space; see e.g. Birkhoff and Mac Lane [45]. In particular for a 2-D problem employing 3 sensors

$$|\underline{H}| = \begin{vmatrix} \underline{h}_1^x & \underline{h}_1^y & 1 \\ \underline{h}_2^x & \underline{h}_2^y & 1 \\ \underline{h}_3^x & \underline{h}_3^y & 1 \end{vmatrix} = \begin{vmatrix} \underline{h}_2^x - \underline{h}_1^x & \underline{h}_2^y - \underline{h}_1^y \\ \underline{h}_3^x - \underline{h}_1^x & \underline{h}_3^y - \underline{h}_1^y \end{vmatrix}$$

In this example the area equivalent to the determinant is that enclosed by the three direction cosines pointing towards the sensors. The 1's in the 3 x 3 matrix are of course the terms associated with the user epoch. The notation $\underline{h}^x$ and $\underline{h}^y$ indicates the x and y components of the direction cosines.

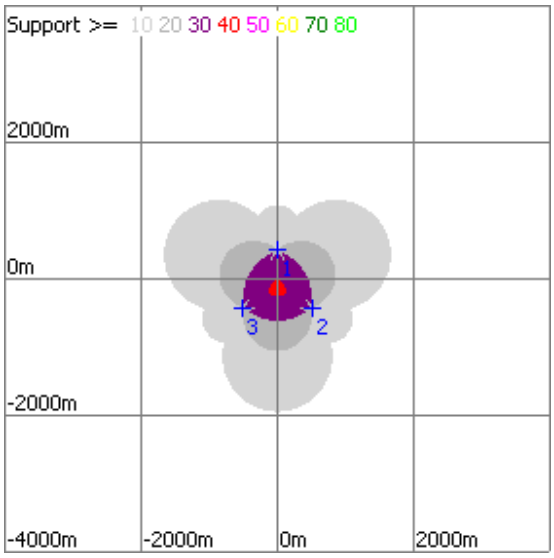
Papers by Phillips [19] and Massatt and Rudnick [20] comprehensively developed the expression of DOP in terms of lengths, areas and volumes. For example, given perfect knowledge of the z coordinates, the expression for Hdop in the 3 sensor case is

$$Hdop = \sqrt{\sum_{i=1}^{3} l_{xy_i}^2} \Big/ 2b_{xy}$$

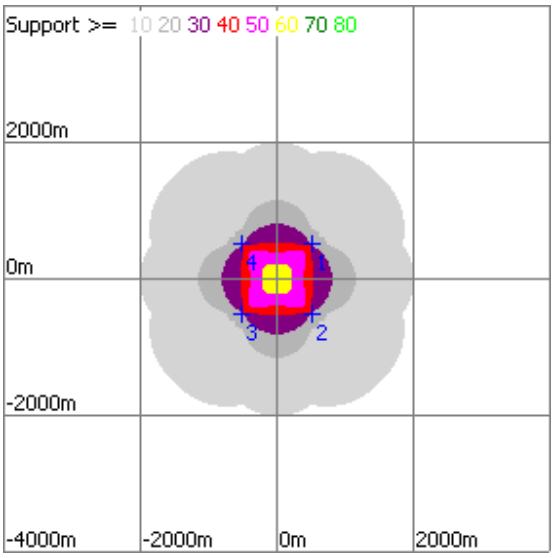
where l_{xyi} is the length of the projection on the xy plane of the side opposite the i^{th} direction cosine and b_{xy} is the area of the projection of the convex hull of the direction cosines onto the xy plane.

From this formula for Hdop we can see that the numerator, which is the root sum of squares of the lengths of the sides of the triangle, will vary much less than denominator. Consequently, as was observed earlier in this section, maximisation of the area b_{xy} or equivalently $|\underline{H}|$ effectively minimises DOP.

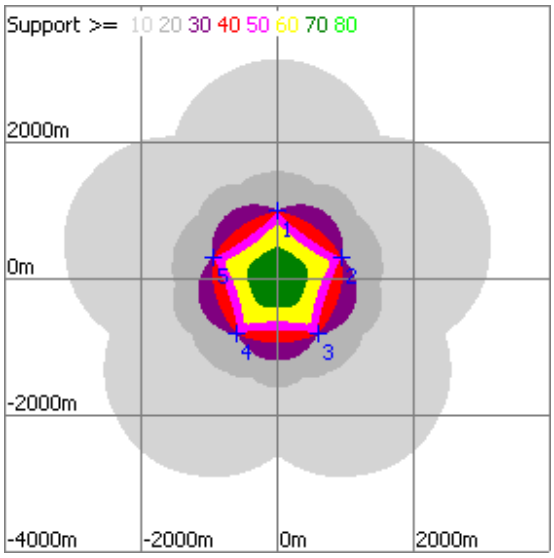
The following plots of normalised area demonstrate the sensitivity of the alternate metric to the support provided by a given sensor / user configuration for the usual sensor arrays: equilateral triangle, square, pentagon and centre point hexagon.



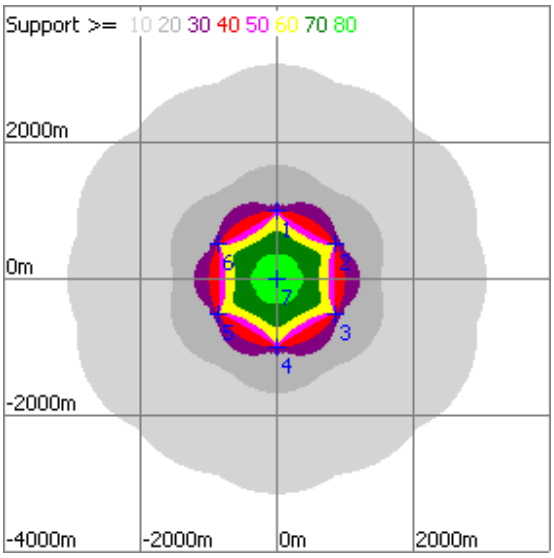
Support – equilateral triangle



Support - square



Support - pentagon



Support – centre-point hexagon

Self evidently the support metric provides a strong indication of both good and poor positioning geometries. The low values for support towards the edges and outside the sensor array do of course concur with the practitioner's expectation and experience. More importantly, as has been evidenced by the examples given in preceding sections of this paper, the support metric can meaningfully address real world problems with arbitrary sensor locations.

Appendix E : Number of Effective Measurements (Neff) Another Alternative Metric

In appendix B expressions for the covariance matrix based on direction cosines augmented with a time element were presented. That is direction cosines with x, y and z terms with unit modulus and a unit time term. We shall now reformulate the estimation framework by handling the time terms separately in a partitioned matrix, after Copps [41]. This approach yields several interesting insights into the relationship between the direction cosines, dilution of precision and common rules of thumb.

Recollect that the direction cosines, with x, y and z components, are $\underline{h}_j$ which are aligned with the user – station vectors. Assuming identical, unit measurement errors and adopting a null (i.e. identity) weighting matrix the estimated covariance for the case where user and station clocks are known is of course given by $(\underline{H}^T \underline{H})^{-1}$.

If the user clock epoch also requires to be estimated, we can write $\underline{H}$ in terms of the $\underline{h}_j$ and unit terms

$$\underline{H} = \begin{bmatrix} \underline{h}_1^T & 1 \\ \underline{h}_2^T & 1 \\ \bullet & \bullet \\ \underline{h}_n^T & 1 \end{bmatrix}$$

H^TH can be formulated as a partitioned matrix comprising a 3x3 sub-matrix equal to the covariance matrix obtained without time augmentation, a vector sum of the direction cosines and its transpose, and a scalar element equal to the number of observations n

$$\underline{H}^T \underline{H} = \begin{bmatrix} \sum_{j=1}^{n} \underline{h}_j \underline{h}_j^T & \sum_{j=1}^{n} \underline{h}_j \\ \sum_{j=1}^{n} \underline{h}_j^T & n \end{bmatrix} \equiv \begin{bmatrix} M^{-1} & l \\ l^T & n \end{bmatrix}$$

If M (the position only covariance matrix) is non-singular this partitioned matrix can readily be inverted, see for example Barnett [24], yielding

$$\left[\underline{H}^T \underline{H}\right]^{-1} = \begin{bmatrix} M^{-1} & l \\ l^T & n \end{bmatrix}^{-1} = \begin{bmatrix} M + Mll^T M/\left(n - l^T Ml\right) & -Ml/\left(n - l^T Ml\right) \\ -l^T M/\left(n - l^T Ml\right) & 1/\left(n - l^T Ml\right) \end{bmatrix}$$

Several of the terms merit discussion. ℓ is the sum of the direction cosines, it is zero when the sensor array is symmetric and small when the user is fairly well inside the array. Contrariwise ℓ is large when the geometry is poor, with a limiting case of the measurements being distributed around a line of position. The classic underwater example is attempting to estimate the depth of an underwater object using a circular boxing pattern when the user clock offset also requires to be estimated.

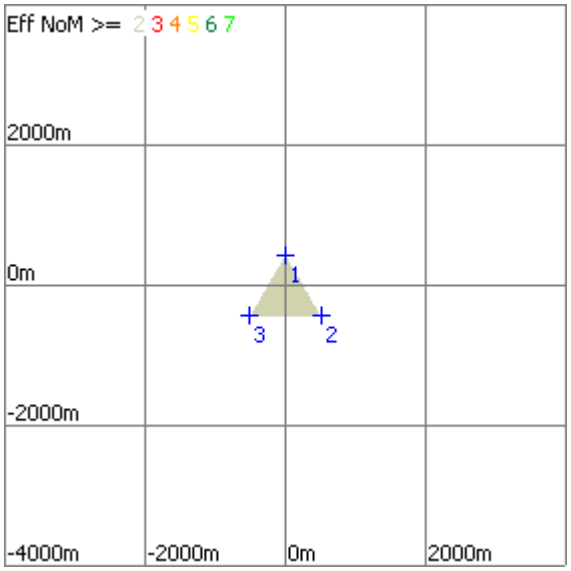
A significant practical manifestation of ℓ, which can be thought of as a measure of misclosure or lack of symmetry in the direction cosines, is (when the user clock epoch is not estimated) as the bias per unit of user clock offset. In the case where positions are being estimated outside an array of sensors the importance of accurate clocks and the impossibility of estimating the user clock are both consequences of the geometry which manifest themselves as significant values of ℓ.

The denominator $(n - \ell^T M \ell)$ is the reciprocal of Tdop squared, i.e. in an ideal geometry Tdop = sqrt (1/n). This relationship suggests the interpretation of the denominator as Neff, the effective number of observations. Obviously, this interpretation is capable of ready assimilation and application by novices and experts alike. Examples of the use of Neff to quantify the quality of coverage have been presented earlier in the paper. The quadratic form $\ell^T M \ell$ can be thought of as the direction cosines misclosure normalised by DOP into units of "worthless" observations.

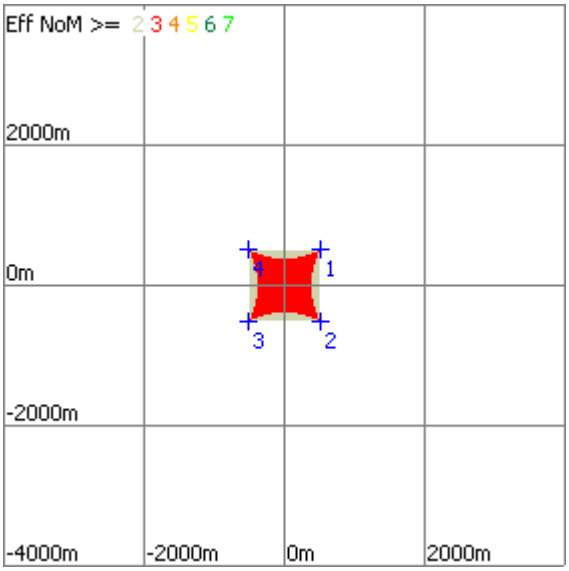
The term $M \ell \ell^T M / (n - \ell^T M \ell)$ is referred to by Copps as the bad geometry or clock penalty term. Self evidently this term will be zero when ℓ is zero, small when ℓ is negligible, and very large when poor geometry occasions a small denominator and a singularity. As Pdop, Hdop, etc are derived from the top left partition the inflation of the position DOPs due to clock estimation can be clearly identified and computed. The corollary - that poor clock quality (where it is not estimated) will feed into the solution in the same proportions - is probably valid.

In summary, by inverting the time augmented $\underline{H}^T \underline{H}$ matrix analytically we have obtained insight into how the observation geometry both determines the accuracy with which the user clock can be estimated and how the position solution is sensitive to clock accuracy. Most importantly, we have obtained an understanding of the relationship between the direction cosine misclosure, Tdop and the effective number of observations (Neff) which provides the practical navigator with a comprehensible and a useful metric.

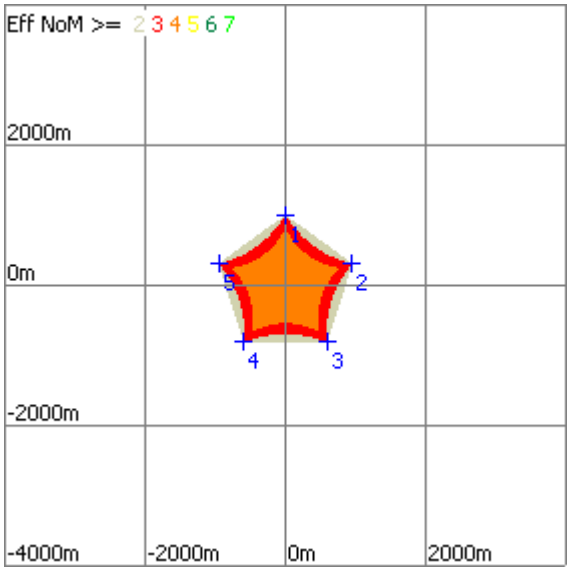
The following plots provide a pedagogic exposition of Neff for familiar, regular sensor patterns. The readers attention is drawn to the >= signs, generally Neff at the centre of the figures will be just below the next value.



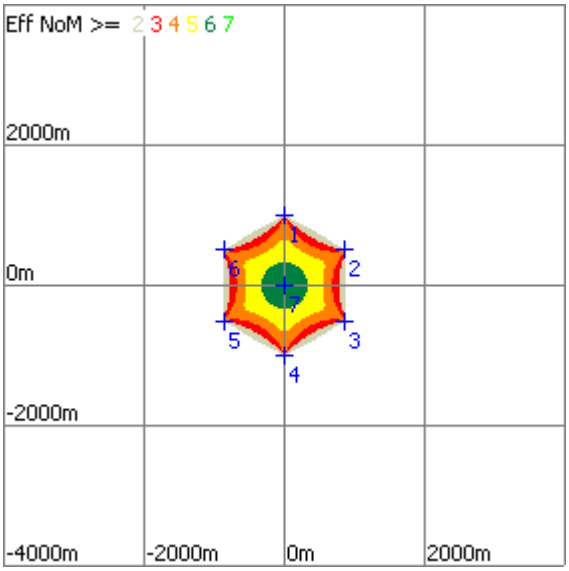
Neff – equilateral triangle



Neff - square



Neff – pentagon



Neff – centre-point hexagon

Acknowledgements

The author acknowledges many useful interactions with co-workers.

The US National Geophysical Data Centre (NGDC) for providing bathymetric data on the internet.

This paper was first presented at NAV08 and first published in the CD-ROM compilation containing the conference proceedings for NAV08.

References

[1] Divins, DL and Metzger, D; "NGDC Coastal Relief Model; http://www.ngdc.noaa.gov/mgg/coastal/coastal.html
[2] Martz, P; "OpenGL Distilled"; Addison Wesley
[3] Shreiner, D et al; "OpenGL Programming Guide"; Addison Wesley, 6th Ed, 2008
[4] Wright, R S jr et al; "OpenGL SuperBible"; Addison Wesley, 4th ed, 2007
[5] http://en.wikipedia.org/wiki/Graphics_processing_unit
[6] http://en.wikipedia.org/wiki/NVIDIA
[7] Hanson, A J; "Visualising Quaternions"; Morgan Kaufmann; 2006
[8] Kuipers, J B; "Quaternions and Rotation Sequences"; Princeton U P; 1999
[9] Stillwell, J; "Mathematics and its History"; Springer, 2nd ed, 2002

[10] Altmann, S L; “Rotations, Quaternions and Double Groups”; Dover, 1986
[11] Gilmore, R; “Lie Groups, Physics, and Geometry”; CUP, 2008
[12] Hall, B C; “Lie Groups, Lie Algebras, and Representations”; Springer, 2003
[13] Urick, R J; “Principles of Underwater Sound”; McGraw Hill, 3rd ed, 1983
[14] Clay, C S & Medwin, H; “Acoustical Oceanography”; Wiley, 1977
[15] Waite, A D; “Sonar for Practising Engineers”; Wiley, 3rd ed, 2002
[16] Brekhovskikh, L M & Lysanov, Yu P; “Fundamentals of Ocean Acoustics”; Springer, 3rd ed, 2003
[17] Jensen, F B et al; “Computational Ocean Acoustics”; Springer, 2000
[18] Bishop, M J D; “Predicting Navigation Fix Accuracy : A Realistic Alternative to Over-Optimistic DOP Values”; UDT Europe 2008, Glasgow, Scotland
[19] Phillips, A H; “Geometrical determination of PDOP”; Navigation, 31, (4), Winter 1984-85, pp329-337
[20] Massatt, P & Rudnick, K; “Geometrical formulas for dilution of precision calculations”; Navigation, 37, (4), Winter 1990-91, pp 379-391
[21] Snyder, J P; “Flattening the Earth : Two Thousand Years of Map Projections”; Chicago U P; 1993
[22] Snyder, J P; “Map Projections – A Working Manual”; US Geological Survey Professional Paper 1395; 1987
[23] Snyder, J P & Voxland, P M; “An Album of Map Projections”; US Geological Survey Professional Paper 1453; 1989
[24] Barnett, S; “Matrices : Methods and Applications”; OUP; 1990
[25] Pierce, J A et al (eds); “Loran”; MIT Radiation Laboratory Series, no 4; 1946
[26] Stansfield, R G; “Statistical theory of DF fixing”; J IEE, 14, Pt III A, (15), pp762-770, 1947
[27] Daniels, H E; “The theory of position finding”; J RSS, 13, (2), 1951, pp 186-207
[28] Ancker, C J; “Airbourne direction finding – The theory of navigation errors”; IRE Trans AN, 5, Dec 1958, pp 99-210
[29] Cooper, D C & Laite, P J; “Statistical analysis of position fixing in three dimensions”; Proc IEE, 116, 9, Sep 1969, pp 1505-1508
[30] Cooper, D C; “Statistical analysis of position-fixing theory for systems with Gaussian errors”; Proc IEE, 119, (6), Jun 1972, pp 637-640
[31] Lee, H B; “A novel procedure for assessing the accuracy of hyperbolic multilateration systems”; IEEE Trans, AES-11, (1), Jan 1975, pp 2-15
[32] Lee, H B; “Accuracy limitations of hyperbolic multilateration systems”; IEEE Trans, AES-11, (1), Jan 1975, pp 16-29
[33] Lee, H B; “Accuracy of range-range and range-sum multilateration systems”; IEEE Trans, AES-11, (6), Nov 1975, pp 1346-1361
[34] Foy, W H; “Position location solutions by Taylor series estimation”; IEEE Trans, AES-12, (2), Mar 1976, pp 187-194
[35] Milliken, R J & Zoller, C J; “Principle of operation of Navstar and system characteristics”; Navigation, 25, (2), Summer 1978, pp 95-106
[36] Swanson, E R; “Geometric dilution of precision”; Navigation, 25, (4), Winter 1978-79, pp 425-429
[37] Sturza, M A; “GPS navigation using three satellites and a precise clock”; Navigation, 30, (2),1983, pp 146-156
[38] Wax, M; “Position location from sensors with position uncertainty”; IEEE Trans, AES-19, (5), Sep 1983, pp 658-662
[39] Torrieri, D J; “Statistical theory of passive location systems”; IEEE Trans, AES-20, (2), Mar 1984, pp 183-198
[40] Kihara, M & Okada, T; “A satellite selection method and accuracy for the global positioning system”; Navigation, 31, (1), Spring 1984, pp 8-19
[41] Copps, E M; “An aspect of the role of the clock in a GPS receiver”; Navigation, 31, (3), Fall 1984, pp 233-242
[42] Parkinson, B W et al (eds); “Global positioning system : Theory and applications Vol I”; Progress in astronautics and aeronautics, 163, AIAA, 1996
[43] Parkinson, B W et al (eds); “Global positioning system : Theory and applications Vol II”; Progress in astronautics and aeronautics, 164, AIAA, 1996
[44] Levanon, N; “Lowest GDOP in 2-D scenarios”; IEEE Proc Radar, Sonar Navig, 147, (3), Jun 2000, pp 149-155
[45] Birkhoff, G & Mac Lane, S; “A survey of modern algebra”; A K Peters, 5th Ed, 1996