International Hydrographic Review, Monaco, LXX(2), September 1993

SOME OBSERVATIONS ON THE USE OF GPS AND CHARTS

by D. SIMPSON [1]

INTRODUCTION

The purpose of this paper is to bring to the attention of chart users some of the navigational implications which have become apparent with the growing adoption of GPS (Global Positioning System) at sea. Action taken by the UK Hydrographic Office to make users aware of weaknesses in both GPS and charting are also discussed, as are possible future trends. Much of the background to the present situation was highlighted by a colleague, Nigel GOODING, in papers published in 1991 and 1992. It is my intention to build on this base and to update and develop some of the themes he raised.

It should be noted that the views expressed herein are the result of the author's own experience and do not necessarily represent fully the current official policies of the UK HO.

PROBLEMS WITH DATUMS

In order to clarify the apparent confusion posed by the concept of a point on the Earth's surface having two or more different, but valid, positions, it is perhaps worthwhile looking at the background to the definition of horizontal datums. Those familiar with these concepts are advised to skip this section and the next.

Firstly, let us consider the different shapes which can be used to describe the surface of the Earth. The "true" shape of the solid Earth is its topographical surface which includes mountain peaks and ocean trenches. It approximates to an oblate, or flattened (at the poles) spheroid but, because of its irregularities, it departs from a spheroidal shape by as much as about 9,000 metres on land and 11,000 metres

[1] UK Hydrographic Office, Taunton, Somerset TA1 2DN, UK.

in the oceans. This topographical surface is much too complicated to be described mathematically and to use as a reference surface for positioning.

A second figure, the Geoid, provides a simpler, though still complex, surface and its departure from spheroidal form is of the order of 100 metres. It is defined as that equipotential gravity surface which equates to Mean Sea Level and can be visualised as an undulating surface which extends through the continents (Fig. 1).

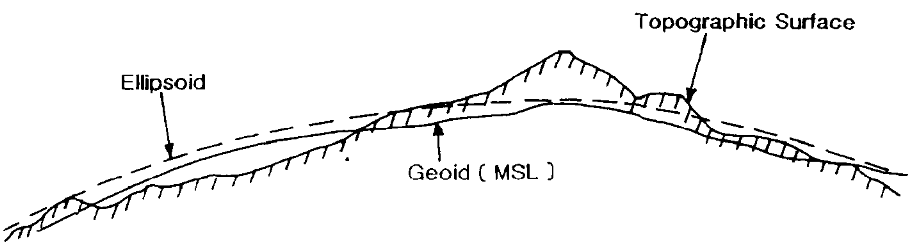

FIG. 1.- Various geodetic surfaces.

The undulating nature of the Geoid is primarily the result of the uneven distribution of mass in the Earth. Its importance lies in the fact that the direction of gravity is always perpendicular to the Geoid and thus we are defining a shape for the Earth to which surveying instruments, such as theodolites and levels, will be referred automatically.

Since the Geoid can be approximated closely by a spheroid (also referred to as an ellipsoid), this latter figure, which is mathematically defined, provides a practical Earth model upon which positional computation can be based. It is formed by rotating an ellipse about its minor axis, (XY in Fig. 2) thereby reflecting the oblate shape of the Earth.

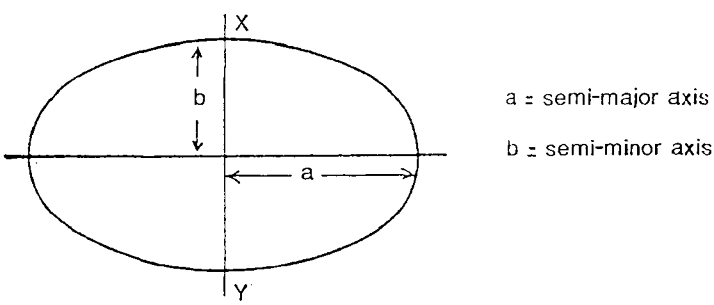

FIG. 2.- Ellipse with axes.

Many different spheroids have been developed to provide a "best fit" to the Geoid for local and more extensive areas of the Earth's surface. In order that a particular spheroid may be used for positional calculations its relationship to the Geoid must be defined. The definition of such a relationship is known as the DATUM. Thus, the OSGB 36 (Ordnance Survey of Great Britain 1936) Datum, which is used for charts covering the coastal waters of England, Scotland and Wales, relates the Airy spheroid, which has a fixed mathematical definition (ie. values of 'a' and 'b', etc), to the Geoid in that region. Similarly, ED 50 (European Datum 1950) relates the International spheroid to the Geoid, with varying degrees of success in, and beyond, Europe. This last point is the crux of the matter, since positions referred to a datum which has been defined inaccurately over all or part of its area of coverage cannot be related with confidence to another datum, which is also subject to inaccuracies.

For example, ED50 was developed for military mapping in central Europe after the Second World War. Its potential as a means to reference positions to a single datum over larger areas was soon recognised and there was pressure to extend the network. These extensions introduced errors which tend to be greatest at the limits of the area for which the datum is defined. Thus, ED50 in the Suez region is not, in effect, the same as that defined in central Europe. Similarly, in North America, extension of NAD27 (North American Datum 1927) into the Caribbean introduced further errors which became apparent when comparing the differences between satellite global datums such as WGS84 (World Geodetic System 1984) Datum on mainland USA and in the Caribbean.

The observations used to determine horizontal control take place on land and are intended for land-based systems. Such local geodetic systems are inherently weaker towards their extremities; that is, in coastal regions and offshore. This lack of homogeneity in local datums leads to additional errors if transformation from one datum to another is carried out.

THE SATELLITE ERA

In the 1960s, satellite technology provided the means of defining a global datum based on the Earth's centre of mass and this has resulted in the development of the WGS 84 Datum, which relates the satellite-derived spheroid to the Geoid. The relationship between positions referenced to this global geocentric (ie. centred on the Earth's centre of mass) system and those referenced to local datums, as are most of the charts in the British Admiralty (BA) series, has now to be addressed.

TRANSFORMATIONS - SHIFTS FROM ONE DATUM TO ANOTHER

In order to convert positions on one datum to positions on another it is necessary to establish the relationship between a number of known points, common to both datums. The quality of the transformation so determined is dependent, amongst other factors, on the number and quality of these common points; generally, the more numerous, the better. However, no transformation is perfect since both the original datum and the datum to which positions are adjusted have inherent weaknesses which vary over their areal extent, as discussed. Depending on the objective, transformation parameters may be tailored to provide best fit over limited areas, while introducing degradation in other, less important regions. The point being made is that datum transformations introduce errors which may be significant in the area of operation. GPS is referenced to WGS 84 Datum but most receivers have options to provide output positions referenced to any one of a large number of local datums in addition to WGS 84. However, the embedded parameters for the transformations are likely to be mean values for the whole area and their application, particularly in the extremities of the area of use of a given datum, might introduce errors of hundreds of metres. It is therefore preferable, from a navigational point of view, to maintain positions referenced to WGS 84 Datum whenever possible.

HORIZONTAL DATUMS AND CHARTS

There are now about 70 charts in the BA series of approximately 3300 which are on WGS 84 Datum, with a similar number on the superseded WGS 72 Datum. The shift between these datums reaches a maximum of about 17 metres at the Equator, being zero at the poles, and is therefore significant for charts of about 1:50,000 scale and larger, especially in low latitudes. There are three main reasons why so few charts in the BA series are referenced to WGS 84. The first, and most important, is that WGS 84 is a recent development and it would take a considerable period of time to replace existing charts with new editions using the new datum. A second reason is expediency. In this case, such as in the coastal waters of the UK, it is charting policy to publish charts referred to the national datum, OSGB 36, so that both charts and maps are referred to a common horizontal datum and thereby provide a seemless base for users. A third important reason is the lack of control data available to calculate a shift from the chart datum, if indeed there is a defined datum, to WGS84 Datum.

Where charting policy is to maintain reference to a national or local datum, it has been recognised that, with the rapid increase in GPS use for navigation at sea, there is a need for shift values which relate the local chart datum to WGS 84 Datum so that satellite - derived positions can be adjusted and plotted directly. To date, these have been calculated and are quoted for about a third of the BA series. They are expressed in minutes of latitude and longitude but are provided only if there is sufficient, good quality horizontal control in the area and errors would not be plottable at the scale of depiction. For example, if, in a given area the relationship between local horizontal datum and WGS 84 Datum cannot be established to an accuracy better than, say 20 metres and positions can be plotted to a precision of 0.2 millimetres, it would mean that this error would become significant at scales larger than 1:100,000. Thus, shift values could not be provided for larger scale charts, which would remain referenced to the local datum only. Similarly, if the relationship between the chart datum and WGS 84 Datum can be calculated to a high degree of accuracy and it happens to be small, then it will only be significant at the largest scales and smaller scale charts may be stated as being on WGS 84 Datum, because the shift is not plottable. This is relatively simple when scales of depiction are fixed, but becomes more problematical as we enter the era of electronic chart display systems (ECDIS) where users may increase scale inappropriately and fail to realise the significance of poor positional accuracy.

When a chart bears a note indicating that adjustments for plotting WGS 84 referenced positions cannot be determined it is because the positional control of the chart is insufficiently known to enable the calculation of a single adjustment for use throughout its extent. Internal positional discrepancies in the chart are the result of horizontal control inconsistencies within, and between, sources of data used in its compilation and MAY BE SIGNIFICANT TO NAVIGATION. Mariners are therefore advised to make greater use of classical methods of observational position fixing when closing the shore or navigating in the vicinity of dangers. The relative positions of features may be more reliable for navigation than the use of unadjusted satellite - derived position on a chart whose horizontal datum cannot be defined.

GPS - ITS STRENGTHS AND WEAKNESSES

The development of GPS (Global Positioning System) has made available a continuous, world-wide, all-weather positioning system which has the potential to provide users with greater accuracy than has been possible before.

However, there are some drawbacks to its current operation which must be recognised by those using it for navigation. NAVSTAR-GPS is a military satellite navigation system which is owned and operated by the United States Department of Defense. The system is not yet operational and the incomplete satellite constellation is subject to testing which might result in lack of availability or corruption of signals. System integrity is also downgraded by the current lack of real-time warning to users when satellites malfunction.

Most civilian users of GPS have access only to a deliberately degraded version of the system which permits horizontal positional accuracy of 100 metres or better at the 95% confidence level. Degradation is by means of Selective Availability (SA) which is a method whereby the system operator corrupts satellite position and/or timing data contained in the navigation message transmitted by the satellite.

The application of SA provided a major stimulus for the development of Differential GPS whereby corrections to the pseudoranges measured between each satellite and the receiver are transmitted to the user's receiver from a Reference Station, whose position has been determined accurately. This method can eliminate almost completely the effects of SA and enables real-time positional accuracies of 5-10 metres to be attained at sea. However, there remains the problem of integrity, whereby a user seeks assurance that his positional accuracy remains within tolerances. Large deviations from true position may occur but there are, as yet, no guarantees that these excursions can be detected and eliminated in real-time. Users must therefore be cautious of over-reliance on the system.

CHART POSITIONAL ERRORS CORRECTED BY SATELLITE IMAGERY

During the last ten years, the UKHO has made increasing use of satellite imagery such as LANDSAT and SPOT to assist in the rectification of gross positional discrepancies. It has become a routine procedure to acquire satellite imagery in support of revised or new charting in areas where conventional data is lacking or is of dubious quality. Often, the use of such imagery represents the only reliable, quick and cost -effective means of correcting coastlines and positioning visible features of navigational significance.

A current example is offered by the revised scheme of charts of the Maldive Islands. The general cover of these islands in the Indian Ocean has been provided by three fathoms charts, 66A, B and C, at a scale of approximately 1:280,000 and dating from the middle of the nineteenth century. Acquisition of LANDSAT imagery has enabled a new metric series of four charts to be compiled at a scale of 1:300,000. The satellite imagery allowed repositioning of atolls and rectification of distortions in their shapes. In some cases these positional shifts have amounted to 4 minutes of longitude while in others they have been negligible, thus indicating the lack of overall positional control in the island chain. The larger positional discrepancies would certainly be significant to a navigator placing undue trust in the relationship between charted positions and those output by his WGS 84 referenced GPS receiver.

The large areal coverage provided by satellite imagery has enabled the limited horizontal control available to be extended throughout the island group such that the new metric charts have been referenced to WGS 84 Datum. Replacement of the fathoms charts of the Maldives by metric charts 1011 - 1014 is scheduled to be completed by the end of 1993.

A recent incident in the Mediterranean provides another example of the use of satellite imagery. It was brought to the attention of the UKHO that a lighthouse marking the entrance to the Rosetta River, at the mouth of the Nile Delta, had been moved inland. Coastal alteration was suspected and this was confirmed when satellite imagery of the area indicated a 2 mile retreat of the coastline which had occurred over a period of about 5 years. This resulted in the issue of a Notice to Mariners block correction (No 259 of 1993) to chart 2681 (Fig. 3).

SOME THOUGHTS AND RECOMMENDATIONS

1. The increasing use of global navigation satellite systems such as GPS promotes the need to reference charts to WGS 84 Datum.

2. With charts referenced to WGS 84 Datum there is no need to degrade GPS positional output by datum transformation.

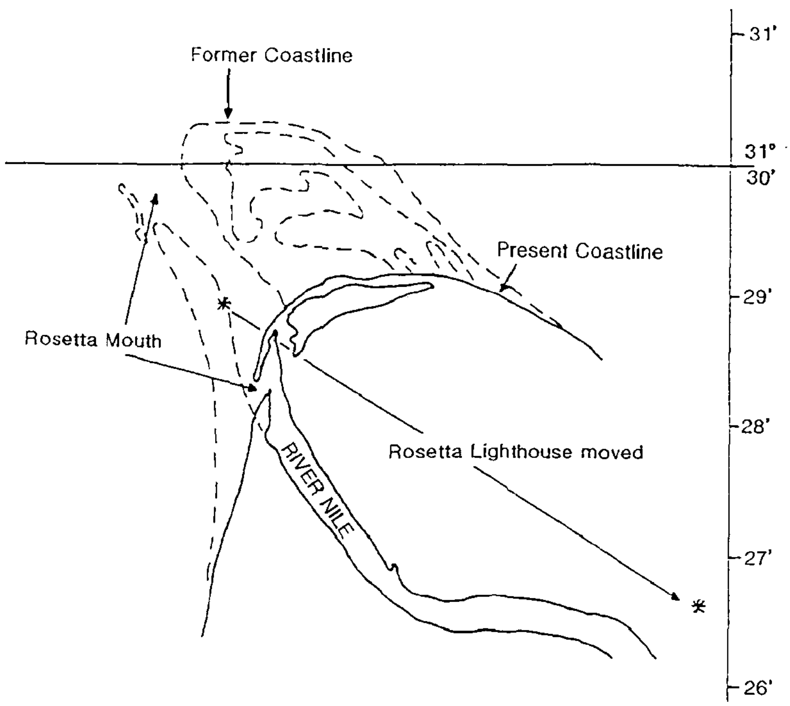

FIG. 3.- Notice to Mariners block correction to Chart 2681 (No. 259 of 1993)

3. In cases where there is limited horizontal control available to define WGS 84 in the area of chart coverage, it could be possible to estimate the likely error and indicate to the user that a chart is referred to WGS 84 Datum to an accuracy of, say, ±50 metres throughout its extent.

4. It would be an advantage if datum shift notes included a worked example applicable to the chart on which they occur. This should help to avoid the doubling of error which results if shift values are applied in the wrong direction.

5. The practice of qualifying positions by quoting a horizontal datum should be encouraged. This would serve to remind users of the possible significance of plotting a position referred to one datum on a graphic (including ECDIS) referred to another, as could happen if, for example, a Notice to Mariners issued by country X, on their national datum, was to be plotted on a BA chart referred to WGS 84 Datum.

6. Mariners to use the indicators provided by charts to assess the likelihood of positional problems. These would include positions notes which might state that horizontal datum cannot be determined, but other indicators such as the antiquity of the source material or the lack of reference to horizontal control beneath the title should also be considered. Similarly, notes may caution that sections of coastline, or islands, have been reported to be out of position. All these can provide clues to the reliability of positional control.

For example, BA 203 - Nisos Zakinthos to Nisos Paxoi - is a modern metric chart depicting part of the west coast of Greece at a scale of 1:150,000. It is on ED50 and carries notes providing shift values for satellite - derived positions and warning of positional discrepancies between it and fathoms charts of the area. Part of BA 203 is covered at a larger scale, 1:51,040, by fathoms chart 3496. This was published in 1905 and was compiled from surveys conducted in 1892 and 1904. There is no reference to a defined horizontal datum but the chart carries two notes, one of which indicates that, "longitudes should be diminished by about one minute (1928)" while the other points out discrepancies between it and BA 203 and recommends the transfer of position by distance and bearing.

Also within the limits of BA 203 is a panel on BA 1620. This is a fathoms chart published in 1866. The 1:48,500 scale plan of part of Levkas Island is ungraduated and is based on a survey of 1864. The only positional reference is the quoted latitude and longitude of the Custom House beneath the title.

7. Users must be aware of the limitations and inaccuracies of GPS.

8. When using GPS, especially in Differential mode, it should be appreciated that positional accuracy of charted features may be considerably poorer than the navigation system.

9. Reporting of GPS referenced positions would assist in the calculation of adjustments to charts.

10. User education. **Annex A** is a reproduction of a recently issued Notice to Mariners, number 943 of 1993, which summarises many of the consequences of using GPS derived positions on BA charts.

CONCLUSION

Many charts have significant inherent positional weaknesses resulting from the antiquity and inadequacies of some of the source material used in their compilation. Thus, a mariner may be tempted to make use of his apparent 10 metre navigational accuracy provided by DGPS to skirt dangers whose charted positions could be many hundreds of metres from their real location, rather than give them a suitably wide berth. The problem of having a navigation system which can be many times more accurate than the charts on which positions are plotted becomes more significant as we enter the ECDIS era. This important topic is receiving the attention of the International Hydrographic Organisation.

It is essential that the mariner is made aware of the potential inaccuracies and weaknesses in both his navigational fixing system and his charts and, whenever possible, makes use of independent navigation aids to confirm positions.

ANNEX A

943 GLOBAL POSITIONING SYSTEM (GPS) - DERIVED POSITIONS AND BRITISH ADMIRALTY CHARTS

1. GPS is referenced to the World Geodetic System 1984 (WGS 84) Datum. This datum relates positions on the Earth's surface, or in space, to a mathematically defined figure, in this case the WGS 84 spheroid (often referred to as an ellipsoid) which is used to approximate, or model, the complex shape of the Earth. Its origin is the Earth's centre of mass and it provides positional reference throughout the world. WGS 84 is thereby a global, geocentric datum.

2. Local or regional datums such as European Datum 1950 (ED 50) use different, non-geocentric spheroids which provide close approximation to the shape of the Earth over a selected area but become progressively poorer beyond that region.

It is essential, wherever possible, that the datum to which positions are referred is compatible with the datum used for the chart on which the position is to be plotted.

3. Charts referred to WGS 84 Datum are currently few in number but enable GPS - derived positions, referred to WGS 84 Datum, to be plotted directly.

4. Charts having a 'SATELLITE-DERIVED POSITIONS' note provide latitude and longitude shift values which enable GPS - derived positions to be adjusted before plotting on the chart. Shift values for some charts of UK waters are available in Annual Notices to Mariners.

5. The remaining charts, some of which carry a note stating that a satellite-derived position shift cannot be determined, are those for which insufficient details of horizontal datum are known. It is important to note that in the worst cases, such as isolated oceanic islands or charts of great antiquity, positions may be several miles discrepant from those derived from GPS. Internal positional discrepancies in such charts are the result of horizontal control inconsistencies within, and between, sources of data used in their compilation and MAY BE SIGNIFICANT TO NAVIGATION. Mariners are therefore advised to make greater use of classical methods of observational position fixing when closing the shore of navigating in the vicinity of dangers. The relative positions of features may be more reliable for navigation than the use of unadjusted satellite - derived positions on a chart whose horizontal datum cannot be defined.

6. Positions plotted on, or extracted from, a chart will contain an element of imprecision related to the scale of the chart. Thus, for example, at a scale of 1:600,000, a chart user who is capable of plotting to a precision of 0.2 millimetres must appreciate that this represents approximately 120 metres on the ground. A position shift, say from one datum to another, of this magnitude is therefore meaningless at such a scale. Similarly, at 1:25,000, the plotting error may be about 5 metres.

Thus, if WGS 84 positions could be defined only to an accuracy of 10 metres, this would be plottable at the smaller scale, in the example, (the chart could effectively be said to be on WGS 84 Datum) but would be plottable, and therefore significant, at the larger. This explains why it is not uncommon for small and medium scale approach charts to be referenced to WGS 84 Datum while the larger scale port plans have no quoted horizontal datum.

References

GOODING N R. 1991. The Navigational Chart and Marine Positioning in the 1990s "Marine Geodesy", Volume 14.

GOODING N R. 1992. Navstar GPS - Charting Aspects. "The Journal of Navigation", Volume 45, Number 3.