Where on Earth Are We?

The Role of Global Navigation Satellite Systems (GNSS) in Archaeological Field Survey

Summary

This is a revised version of the 2003 edition of this guidance document. It takes into account developments in technology that have occurred during the past decade drawing upon the practical experience of Historic England field teams. The guidance begins with a brief, non-specialist introduction to how the technology works, followed by more detailed consideration of scale and accuracy and the importance of transforming coordinates to the Ordnance Survey National Grid.

There is a wide variety of different satellite receivers available for use in survey and the guidance gives a brief overview of the main types and advises on the appropriate use of different receivers for archaeological mapping. The guidance gives advice on how Historic England field teams undertake landscape survey using this technology, aspects of which are further explored through several case studies. The publication concludes with a brief look at future developments including the growing use of smartphone technology for mapping and data collection in the field.

First published by English Heritage September 2003

Second edition published by Historic England June 2015; reissued October 2016 and dedicated to the memory of Bernard Thomason (1949-2016), co-author of the first edition and a pioneer in the use of GPS for archaeological survey.

HistoricEngland.org.uk/advice/technical-advice/

Cover image

Surveying with GNSS.

Contents

1 Introduction

The first edition of this guidance appeared in 2003 (Ainsworth and Thomason 2003). It appeared just as the use of satellites for surveying was becoming more commonplace and it provided a useful introduction to deploying the technology for archaeological fieldwork. This edition of the paper takes into account the advances that have occurred since 2003 and looks at how the technology is used now by Historic England for landscape survey. It is hoped that this guidance will support the work of both professional and volunteer groups that wish to develop their use of global navigation satellite system (GNSS) technology for archaeological landscape survey, whether the aim is to produce detailed plans of earthwork sites, to collect data for use in geographical information systems (GIS) or reconnaissance work. The paper begins with a brief overview of how the technology works. For those who wish to find out more about the science behind satellite navigation, there are a number of publications and websites available (eg El-Rabbany 2006). In particular, the websites of the Ordnance Survey (OS), The Survey Association (TSA) and the Royal Institute of Chartered Surveyors (RICS) contain much that is relevant and also have very useful technical papers that can be downloaded (eg Royal Institute of Chartered Surveyors 2010).

The paper then provides an overview of what is meant by accuracy and the use of coordinate systems. After describing in more detail different types of survey equipment, it goes on to describe the different uses of the technology in archaeological field survey as practised by Historic England. A short section at the end looks at future developments.

One of the main advances during the 10 years or so since the first edition of this paper is that the United States Department of Defense GPS constellation has now been rejoined by the Russian Global Orbiting Navigation Satellite System (GLONASS) constellation, after more than a decade during which it was not fully operational, and further systems are in development. These systems are collectively referred to as GNSS, and this term is used throughout this publication in preference to the term GPS, which refers specifically to the American constellation.

A metrically accurate survey underpins the process of analysis and, for many archaeological projects, a survey is the first stage in understanding and interpreting the evidence. GNSS technology has revolutionised ground-based survey because it can fix a position in real time anywhere in the world. The accuracy of any point recorded by a receiver is governed by a number of factors relating to the sophistication of the equipment used, the surveying techniques used, the environment and, if the position is to be related to background mapping, the coordinate transformation used. GNSS survey is now the principal technique used by surveyors, mapmakers and many other bodies and organisations that require accurate positioning of features, objects or assets in a plan or map-based form at various scales. Whereas previously surveying with optical instruments was governed by line of sight along the ground, whether across the landscape or among features in excavation trenches, GNSS does not have that constraint, and this has made it a much more flexible tool than survey using a total station theodolite (TST). It is important to emphasise, however, that although GNSS is good at providing data that can be turned into maps, plans and digital ground models, they do not replace the analytical skills that are the foundation for the interpretation and understanding of archaeological sites and landscapes. Nor is GNSS the only method for delivering a metrically accurate survey. TSTs, laser scanners and manual recording techniques using tapes, ranging rods and plane tables can also be used for survey (Bowden 2002; Bedford et al 2011; English Heritage 2011) although, except for laser-based solutions, they tend to be slower techniques than using GNSS.

2 Background to How GNSS Works

GNSS satellites broadcast a microwave radio signal and, put simply, the time interval between the transmission of the satellite signal and its reception by the receiver enables the receiver to calculate the distance between it and the satellite (El-Rabbany 2006, 13–16). As the positions of the satellites in orbit above the Earth are known, simultaneous measurements to more than four satellites enable the unknown position of the receiver to be computed in three dimensions. The satellites also transmit an almanac of their predicted positions that is used by the receiver when it first starts tracking the broadcasts following switch-on.

GNSS receivers need to receive a signal from at least four satellites in order to perform all the required positioning calculations. It follows, therefore, that the receiver must be able to 'see' the sky. This is probably the easiest concept to grasp, although in practice it is not quite that simple. The amount of sky visible is one factor; how many satellites are in the visible parts of the sky is another; and, finally, the quality of the satellite geometry is a third key factor. Put simply, a wider distribution of satellites across the sky gives better satellite geometry, which in turn results in smaller positional errors on the ground (Fig 1). This is referred to as the dilution of precision (DOP) and is displayed on the more sophisticated types of receiver as a numerical value: the larger the value, the poorer the precision. Fortunately, the surveyors on the ground do not have to concern themselves with this too much; they only have to be aware that these factors will affect the quality of the data and resultant positioning and may cause the receiver to reject a point depending on the threshold of accuracy that has been set. Poor geometry is more likely to occur when surveying in deep valleys or close to cliffs and buildings, where sky visibility is restricted. It is advisable when working in such areas to check first the satellite visibility for the chosen location, to find out when the greatest number of satellites will be in the observable part of the sky. Most manufacturers of GNSS survey equipment provide a software program to calculate satellite visibility to assist with timetabling fieldwork. Even then, if the satellite geometry is poor the receiver may still struggle to resolve a point accurately.

The signals broadcast by the GNSS constellations are extremely weak and subject to interference from a number of sources during their passage to the receiver (El-Rabbany 2006, 43–63). Left unresolved, this interference can result in quite major errors in calculated positions on the ground. Changes within the Earth's atmosphere can affect the passage of the signal both through the upper atmosphere (ionosphere), where solar radiation can cause problems, and the lower atmosphere (troposphere), where variations in atmospheric pressure and water vapour can disturb the passage of the signal. The signals are more prone to distortion when the satellites are low on the horizon, because they have more of the atmosphere to pass through, but usually a parameter can be set on the receiver (called an elevation mask) to reject data until a satellite rises to a certain angle above the horizon. This is typically 15 degrees. Compared with single-frequency (L1) receivers, the more sophisticated (and expensive) dual-frequency (L1/L2) receivers are better able to mitigate against ionospheric errors caused by the dispersive effect of the ionosphere on GNSS signal frequencies.

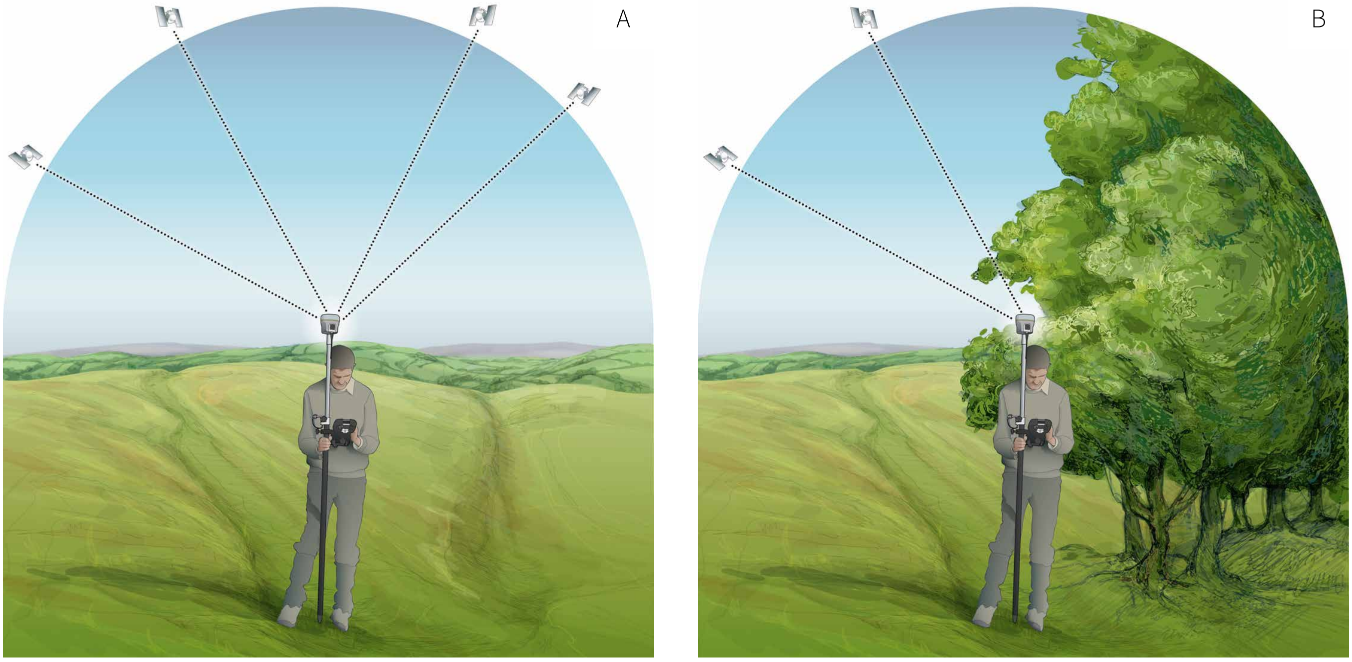


Figure 1
Diagram showing the difference between (A) strong satellite geometry, with an open view of the sky, and (B) poor satellite geometry, where sky visibility is restricted.

Once on the ground, the signal can reflect off buildings and large objects, such as passing vehicles, resulting in multi-path interference: the same signal arrives more than once at the receiver, giving rise to potential errors in the calculated position (Fig 2). Such errors can be detected by the more expensive receivers and compensated for, while higher specification receivers use differential survey to filter errors and improve accuracy, as will be described in more detail below.

Aside from physical obstructions, GNSS reception might also disappear entirely because of radio interference from masts and transmitters or from working machinery, so it is important to assess the risk of such interference when planning a survey. It is also usually very difficult to use a GNSS receiver among trees because foliage blocks reception, especially when leaves have a high moisture content. Usually it is far more efficient to use a TST to capture detail in woodland (Bedford et al 2011).

Most modern survey- and mapping-grade receivers are capable of receiving signals from both the American GPS and the Russian GLONASS satellite constellations. With more satellites available, it is possible to work in marginal areas, for example near trees or close to buildings, where formerly there were too few satellites for the receivers to calculate accurate positions. Europe is also developing its own satellite navigation system, known as Galileo. Currently under development by the European Space Agency (ESA), Galileo will be an independent, civilian-operated system, but it will be compatible and interoperable with GPS and GLONASS. It is intended that Galileo will be fully operational by 2020. A constellation is also being developed by the Chinese, called the BeiDou Navigation Satellite System (BDS). It will initially only be available to users in the Asia–Pacific region but when fully operational in 2020 it will be available to users across the globe.

There are several other satellites used for positioning that are designed to augment the major constellations described above. They are referred to as satellite-based augmentation systems (SBAS) and differ from the main constellations in that they are geostationary. The SBAS system observable from most of Europe is the European Geostationary Navigation Overlay Service (EGNOS) developed by the ESA. North America has an equivalent, called the Wide Area Augmentation

System (WAAS), and others are in development. The EGNOS system consists of three telecommunication satellites in fixed positions above the Earth. They receive correction data from six GNSS monitoring stations across Europe and then broadcast these corrections for use by any SBAS-enabled GNSS receiver. The receiver can then resolve its position more accurately than if it was relying on GPS and GLONASS alone. Most types of GNSS receiver on the market today have SBAS capability. The downside of EGNOS, particularly at higher latitudes, is that the EGNOS geostationary satellites can appear at quite low elevations in the sky and therefore can be easily blocked by obstructions or terrain.

Before discussing the equipment and applications of GNSS for archaeological surveying in the field, it is important firstly to be clear about what is meant by accuracy and secondly to understand the transformations that take place to fix a survey derived from satellite data accurately on to a national grid, which in Great Britain means relating to OS mapping. As a new user, it is easy to set up a receiver and to produce readings relating to latitude, longitude and height, or to eastings, northings and elevation, without fully understanding the accuracy of the coordinates displayed and therefore the reliance that can be put on the survey.

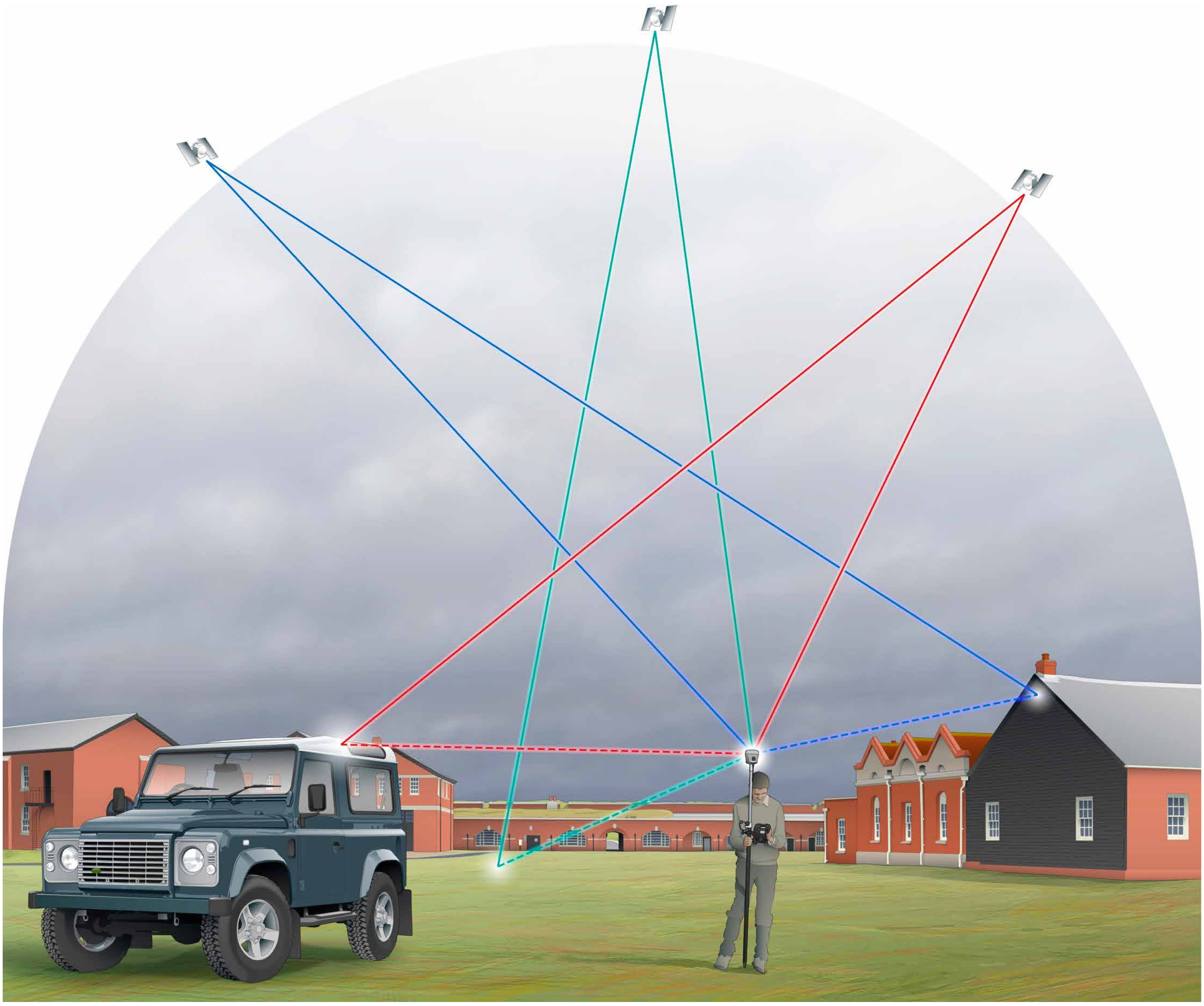


Figure 2
Diagram showing examples of how the signal broadcast from the satellite can be reflected by features on the ground, giving rise to multi-path interference and possible inaccuracies in the calculated position of the GNSS receiver.

3 Accuracy and GNSS

The term accuracy is applied in three distinct ways when used in conjunction with GNSS and survey: (1) relative; (2) map; and (3) absolute (Fig 3).

- Relative accuracy refers to the relative alignment of points in the survey and the measurements between them.
- Map accuracy refers to the accuracy of the placement of those points on a map; in Great Britain this would normally be an OS map.
- Absolute accuracy is a measure that indicates how closely the coordinates of a point agree with the 'real' coordinates of the point on the surface of the Earth. As was mentioned above, it is possible on the more sophisticated receivers to set a threshold to reject points if the calculated position does not meet a minimum level of accuracy, as will occur, for example, when satellite geometry is poor.

3.1 Relative accuracy

Underpinning any surveying technology is the concept of relative accuracy: the measurement and spatial relationship between points in a survey. This applies to all surveying equipment and methodology, and is a mark of how close the distance measured between two points with a tape, TST or GNSS receiver is to their real distance apart. The concept of relative accuracy and making choices in relation to appropriate surveying equipment and techniques for archaeologists is covered in detail elsewhere (Bowden 1999, 49–52; 2002).

3.2 Map accuracy

GNSS has a clear advantage for site survey because of the ease with which it can record points for a map using the national grid system, which in Great Britain is the OS National Grid. In the past, long traverses with electromagnetic distance measurements (EDM) and theodolites between triangulation pillars, followed by copious calculations, were necessary to fit surveys on to OS large-scale maps accurately. This was a time-consuming and complex task. Because of this, easier and less accurate methods were devised, with the position of surveys being 'best fit' on to paper maps, or on to digital maps using computer-aided design (CAD) systems.

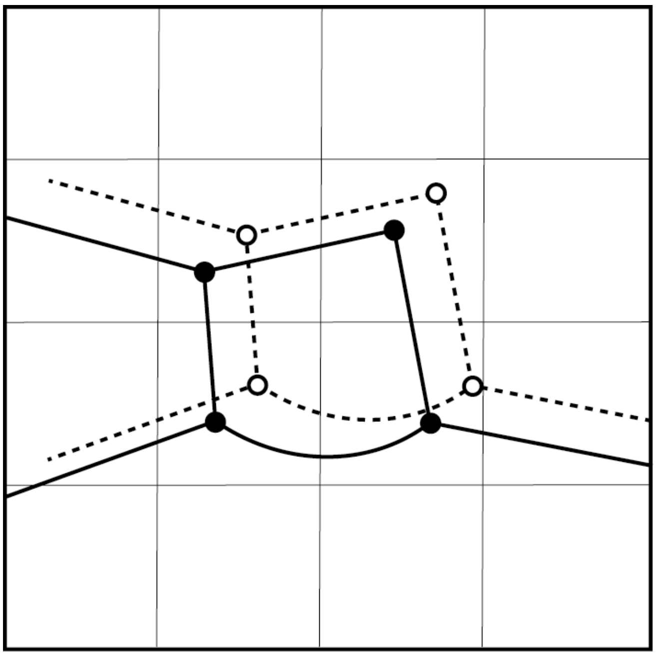


Figure 3
Diagram showing the difference between relative accuracy (dashed line) and map accuracy (solid line).

These practices were not very satisfactory solutions for maintaining map accuracy where the correct spatial relationship between features on the ground and the map is important.

The advent of GNSS has meant that the accurate positioning of surveys on to maps is no longer the realm of the specialist land surveyor, as most receivers can display points with map coordinates. However, the surveyor does have to identify the GNSS equipment that will achieve the required map accuracy for the scale of survey being undertaken. The positions of features recorded by GNSS might not agree with the same features shown on existing OS mapping; this inaccuracy might be because the wrong type of receiver is used, one that is not capable of achieving the required accuracy (see below). Alternatively it might be because the map is not as accurate as the GNSS-derived position.

3.3 Absolute accuracy

This is a measure of the closeness of a point's coordinate to its 'true' value. Satellite constellations use global coordinates defined in a system called the World Geodetic System 1984, usually abbreviated to WGS84, and because of the worldwide nature of GNSS high absolute accuracy in the same coordinate system is now possible across the surface of the Earth, although for most users the true value of a point is defined by the national map grid. Through software transformations, positions derived from GNSS in WGS84 coordinates can be projected on to different national map grids. Most GNSS manufacturers have a number of map projections built into their software to enable WGS84 positions to be transformed to national map grids, however these can be of varying quality and it is advisable to document which, if any, transformation is used.

4 Coordinate Systems, Maps and Grid References

Any map projection needs an ellipsoid, which is a mathematical surface that best approximates the curved shape of the Earth's surface. Historically, however, accurately measuring the size of the Earth has been difficult and as a result the ellipsoids defined in the past were often chosen to best fit on a regional rather than on a global scale (Iliffe 2008).

The Earth is a complex and dynamic body, and geodesists through the centuries have applied themselves to the science of defining various ellipsoids on which to base mapping projections. Map projections are necessary to allow the actual curved surface of the Earth (or, for simplicity, the ellipsoid) to be represented as a flat plane. To address local mapping requirements, many countries have devised their own map projections and referencing systems. On a more global scale, long before satellites had ever been invented, maps and charts represented the surface of the globe using latitude and longitude positions based on celestial observations. However, it is only since the advent of artificial satellites that it has been possible to measure distances on truly international scales and to construct the precise global coordinate system WGS84.

Points surveyed using a GNSS receiver need to be related to existing map detail and so require transformation from the global reference system provided by the satellite navigation system (WGS84) on to the local reference system and map projection. In the case of Great Britain, this is the OSGB36® system and OS National Grid (Ordnance Survey 2013). The process of transforming coordinates is described below, although it is worth noting that the OS provides a free model to relate GNSS-derived coordinates to the OSGB36 National Grid. This transformation model is known as OSTN02™.

The geoid is another theoretical surface that the surveyor needs to visualise in order to understand heights derived from GNSS. The geoid is an irregular gravitational equipotential surface approximating the mean sea level over the Earth and its imagined extension under the land areas. In order to give meaning to the concept of relative heights around the world, the geoid provides a theoretical figure similar to the ellipsoid, thus providing a worldwide datum to which heights can be related. This acts as a fundamental reference for all height measurement. Historically, on mainland Great Britain, heights have been related to Ordnance Datum Newlyn (ODN), thus all heights on OS maps and benchmarks are related to ODN. Just as the ellipsoid coordinates have to be transformed to local maps, so heights derived from GNSS have to be transformed to ODN or another vertical datum. The transformation from GNSS heights to the relevant national height datum is most accurately carried out using the OSGM02™ geoid model. Failure to take this transformation into account can render height values derived from GNSS inaccurate by up to 50m.

4.1 WGS84 and ETRS89

As explained above, while a grid reference is thought of in relation to a coordinate system on a map, such as the OS National Grid, a position derived from GNSS data is related to the global coordinate system WGS84. Coordinates in this system are usually expressed as either latitude, longitude and ellipsoid height, or Earth-centred Earth-fixed (ECEF) xyz Cartesian coordinates. WGS84 was designed for navigation applications throughout the world, where the required accuracy is up to 1m. A high-accuracy version of WGS84, known as the International Terrestrial Reference System (ITRS), has been created in a number of versions since 1989 and is used mostly by geoscientists. There is a problem, however, with trying to use a global coordinate system for land surveying in a particular country or region: the surface of the Earth is formed of a series of plates that are constantly in motion with respect to each other, at rates of up to 120mm/year. This phenomenon, known as continental drift, means that in reality there are no fixed points on Earth. In common with the rest of Europe, Great Britain is in motion with respect to the WGS84 coordinate system at a rate of approximately 25mm/year. Over a decade, the WGS84 coordinates of any survey station in Britain change by c 0.25m as a result of this effect, which is unacceptable for precise survey purposes.

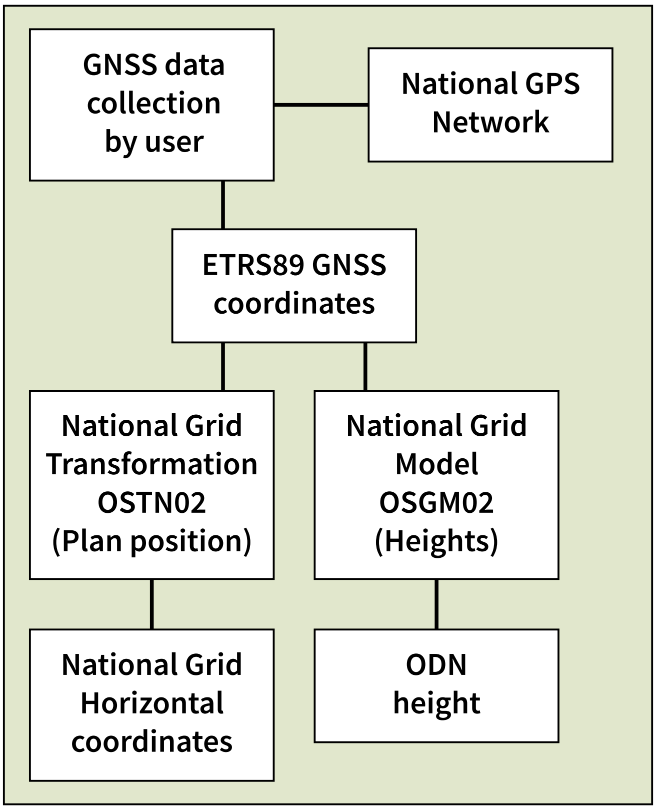


Figure 4
Diagram showing the main stages in transforming GNSS satellite data to OS National Grid coordinates and ODN heights.

For this reason, the European Terrestrial Reference System 1989 (ETRS89) is used as the standard precise GNSS coordinate system throughout Europe. ETRS89 is based on ITRS, except that it is tied to the European continental plate and fixed in time (1989). Hence it is steadily moving away from the WGS84 coordinate system with the movement of the European plate. The relationship between ITRS and ETRS89 is defined precisely at any point in time by a simple transformation of the coordinates, published by the International Earth Rotation Service (IERS).

In summary, these global coordinate systems might seem confusing but it is important to recognise that they have implications for using GNSS to relate surveys to OS maps (see below regarding the National Grid). ETRS89 has been officially adopted as a standard coordinate system for precise GNSS surveying by most national mapping agencies in Europe. By using ETRS89, it is possible to ignore the effects of continental motion because, to a high degree of accuracy, the ETRS89 coordinates of a survey station stay fixed as long as there is no local movement of the survey station. If control from the OS Net® (described below) is used, then the survey will automatically be in ETRS89 (Fig 4).

4.2 The National Grid

To use GNSS data with OS mapping, it is necessary to understand the different coordinate systems used by the OS (Ordnance Survey 2013). As the government agency responsible for national standards in spatial positioning in Great Britain, the OS has defined three national coordinate systems.

- OSGB36 National Grid. Historically this is a traditional horizontal (eastings and northings) coordinate system using the Airy 1830 ellipsoid, originally realised by a national network of triangulation pillars. It is the national coordinate system in Great Britain for topographic mapping at all scales for most maps available both in print and digital form, including MasterMap® data, and is commonly used as the reference grid for spatial records in archaeological records collections. OSGB36 National Grid is now defined by the ETRS89 positions of OS Net GNSS stations plus the OSTN02 transformation (see below)

- ODN. This is a traditional vertical coordinate system based on mean sea level tidal observations at Newlyn in Cornwall between 1915 and 1921. ODN is realised by a network of 200 fundamental benchmarks and half a million lower order benchmarks throughout mainland Britain. All observations for this network were undertaken by traditional levelling techniques. All benchmark values, spot heights and contours on OS maps throughout mainland Britain are related to the datum at Newlyn. This is the usual definition of heights above mean sea level in Great Britain. As the lower order benchmarks are no longer maintained by OS, the recommended method of realising an ODN height is via GNSS and the OSGM02 model

- ETRS89. This is the national coordinate system for three-dimensional GNSS positioning, as described above. If the scale of a survey is such that accuracy of 1m or more is acceptable, then WGS84 coordinates can be used and converted to the National Grid (OSGB36) if necessary. If the requirement for accuracy is below 1m, then ETRS89 coordinates should be derived using OS Net

4.3 The OS national GNSS base station network (OS Net)

OS has defined a new GNSS national positioning infrastructure based on ETRS89. This infrastructure has replaced the traditional network of triangulation stations still so familiar on mountain and hilltops, which formed the physical realisation of the OSGB36 National Grid. This new network of fixed points is known as OS Net and consists of a network of more than 100 GNSS receivers fixed at permanent locations throughout Great Britain, with a spatial density of 50–80km to give full coverage across the country. The precise positions of these base stations are known in relation to the ETRS89 coordinate system and they continuously log satellite data from both the GPS and GLONASS constellations. The data are used by partner organisations and communicated back to the user with a licence subscription and applied in real time to their GNSS readings, typically using a mobile phone service. The system has been rigorously tested (Penna et al 2012) and found to deliver consistently a root mean square (rms) accuracy in the order of 10–20mm in plan position and 20 –30mm in height. This means that 'for each coordinate component roughly 68% of individual solutions will have errors smaller than the rms, and 95% will have errors smaller than twice the rms' (The Survey Association 2012, 2).

If real-time corrected data are not required, or if the user cannot receive correction data because, for example, of poor mobile phone reception, it is possible to download the same data (in the Receiver-INdependent EXchange format, RINEX) for free from the OS and use post-processing software to adjust the positional data gathered in the field.

Points are transformed to eastings and northings on the National Grid (OSGB36) by applying the OS definitive transformation (OSTN02), and to heights above mean sea level that are consistent with the traditional benchmarks by using the national geoid model (OSGM02). These are computer-based transformations and are available within many GNSS software packages as well as free through the OS website. This ensures that all data captured through the use of OS Net are to a common standard and can be easily and accurately related to other map data sets that have been transformed to the OS National Grid, such as georeferenced aerial orthophotos or lidar tiles, as might be required in a GIS.

5 GNSS Equipment

Having looked at the principles that underpin GNSS and its use, the next stage is to describe the equipment itself. It is easy to become overwhelmed by the sheer volume of articles, technical reviews and advertising literature that cover the subject. To simplify this ever-expanding equipment market, the GNSS survey equipment available can be divided into three broad categories: navigation-grade (hand-held), mapping-grade (GIS data collection) and survey-grade (detail survey) (Fig 5).

These categories are principally defined by their respective levels of accuracy (as explained above), their mode of use in the field and their cost. Entry-level navigation-grade receivers cost a little more than £100 while the most sophisticated survey-grade receivers cost more than £10,000. The differences between the various grades of receiver are largely because of the different signal-processing techniques used by the receivers to determine their positions, with cheaper models only able to process one of the multiple signals available. It is important to remember, however, that the receivers are not the only pieces of equipment necessary; there will almost certainly be a need for tripods, poles, batteries, chargers, data loggers, computers and specialist survey software, especially for mapping- and survey-grade GNSS. An external GNSS antenna may also be required for use with an on-site base station, as described below. Without these peripherals, which can be costly to purchase and maintain, the receiver will be of limited use. An alternative to purchasing the more expensive receivers is to hire the equipment, especially if it is not likely to be in continuous use. Where budgets are tight, this option makes the most sophisticated and expensive survey-grade receivers viable for archaeological field survey.

5.1 Navigation-grade GNSS receivers

These receivers are compact, lightweight and comparatively cheap and are marketed to the general public as navigation aids. The same navigation facility is often found on modern smartphones and tablets. While navigation-grade receivers and smartphones are of limited use as surveying instruments, they are still useful as a means of locating or recording the approximate positions of archaeological sites or finds in areas where there is little map detail. Some more specialist receivers offered by survey companies are aimed at GIS data collection and asset management. They typically have a built-in camera that outputs georeferenced photographs, commonly with WGS84 coordinates, and a form that the user can adapt to record the attributes of each point surveyed.

Navigation-grade receivers use only a part of the GNSS signal and as a result can often work better under woodland canopy than survey-grade receivers, which rely on clear reception of a wider spectrum of signals. It must be understood, however, that navigation-grade equipment will not generally provide a horizontal position of sufficient accuracy to undertake detailed mapping. Taking repeated readings at a fixed location will result in a spread of points as the receiver calculates a slightly different coordinate each time (Fig 6), so the relative accuracy is quite low when viewed at a large scale. The technical specifications of the model of receiver being used should be checked for the stated level of accuracy normally achieved but, as Fig 6 demonstrates, it will not be sufficient for detailed planning of a site or for recording features relative to one another. The inbuilt software in navigation-grade receivers often provides positioning within a number of international and local map coordinate systems, for example the National Grid (OSGB36). It is possible with some units to export the recorded positions in WGS84 and then to use the OSTN02 and OSGM02 transformations from the OS GNSS website to obtain coordinates in the OS National Grid.

5.2 Mapping-grade GNSS receivers

These are lightweight systems that can be used for surveying and mapping purposes where accuracies between 0.5m and 5m are acceptable. This type of system usually consists of a hand-held device that combines the receiver and data logger in a single unit. Mapping-grade hand-held GNSS units are distinguished from navigation-grade units by their level of sophistication, specification and cost, which is usually in the low thousands. Mapping-grade GNSS receivers usually have the capacity to view preloaded and georeferenced digital map tiles, lidar and aerial photographs, to allow feature and attribute coding and the collection of data for use in GIS via customised forms (Fig 7) (see below regarding surveying with GNSS.

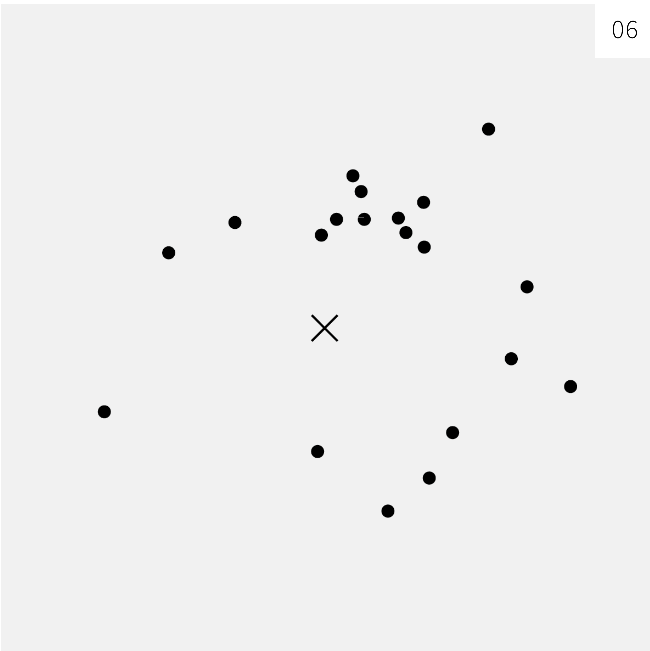
This type of receiver is ideal for mapping at scales up to 1:2 500 and for recording associated data for use in GIS systems. The majority of these receivers are capable of receiving SBAS corrections such as EGNOS, which gives them an accuracy generally of no better than 0.7m, while more sophisticated models can achieve an accuracy of 0.1m because they have the facility to use a real-time differential correction such as from OS Net. Again the manufacturer's technical specifications should be read to ascertain the accuracy of a specific model. Some data formats allow post-processing using data downloaded from OS Net via the OS website in order to improve accuracy.

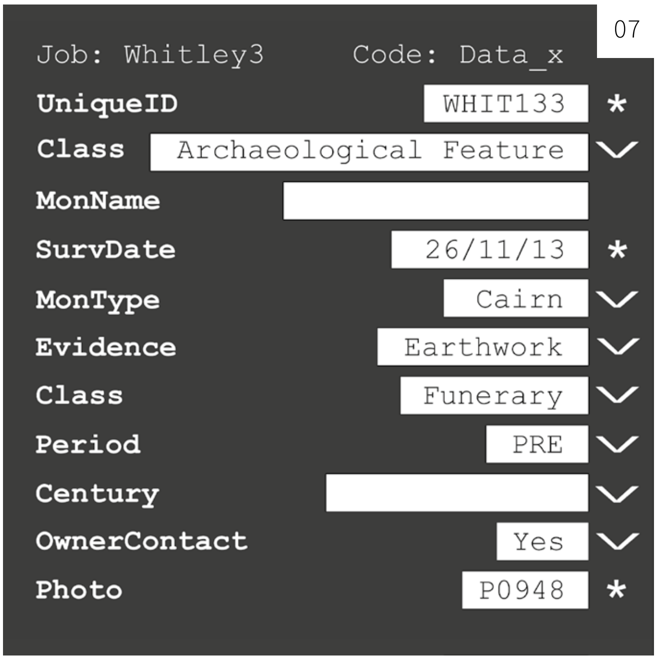
Figure 5
The three main categories of GNSS receiver. (A) Navigation-grade, (B) mapping-grade and (C) survey-grade.

5.3 Survey-grade GNSS receivers

Survey-grade equipment is capable of the greatest accuracy, so consequently is the most expensive GNSS survey option. It is the only type of GNSS receiver that can produce absolute accuracy at the centimetre level and is comparable in relative accuracy to a higher order TST survey. In order to achieve this high degree of accuracy, all survey-grade systems use a differential approach to data gathering and survey processing. One method is to use two receivers on site. One receiver is designated as the base, which remains static on a tripod, while the second is referred to as the rover. The rover is mounted on a pole and is moved by the surveyor from point to point to create the survey, ideally with a radio link to the base to receive corrections and improve relative accuracy in real time. Post-processing is then needed to adjust the data gathered by the base and rover method to achieve the highest degree of accuracy relative to the OS National Grid, or the position of the base station can be determined to the same degree of accuracy at the start of the survey by obtaining correction data in real time from OS Net via one of the subscription services developed by a number of providers. The alternative is to dispense with the base receiver and use just a rover receiver in combination with OS Net. These methods are described in more detail below.

Older rover receivers required the user to carry ancillary components, such as the power supply and a radio receiver connected by leads to the receiver, in a backpack. With modern survey-grade receivers Bluetooth technology has largely replaced the need for leads and the power supply and radio antenna are both small enough to be housed on the receiver.


06


07

Figures 6 and 7

6. Plot showing 20 points recorded at 30-second intervals using a navigation-grade receiver over a static point (represented by the cross) located using a survey-grade GNSS receiver. The bounding box represents a 10m × 10m square and the spread of points indicates the problem of using this grade of receiver for detailed survey.

7. Example of a simple form for recording attribute information on a mapping-grade GNSS unit for download into GIS. The 'down' arrow indicates entries selected from a pick list and the star indicates entries that are populated automatically.

6 Landscape Survey Using GNSS

The basics of surveying apply to GNSS as they do to any other method of survey, and choosing the correct approach and understanding the level of detail required is all part of the skill of surveying. The level of detail required should dictate the scale of the survey and vice versa (Bowden 1999, chapters 4 and 5). Sometimes it will be down to the field team to decide the scale of the survey, or it may be stipulated in a tender document or project design. With the scale of the survey established, the next decision is how to tackle the survey. While GNSS is often the technique of choice because of its versatility and speed, it is not always the best approach to use and often has to be combined with other techniques to get the best results. The terrain, site topography and geography should be assessed prior to the start of the survey to see if there are any factors that might affect clear reception of signals from satellites, such as trees, buildings, cliff faces and other obstructions. Where such obstructions exist, and depending on the scale of the survey, then a TST or manual recording using tape-and-offset or a plane table will have to be deployed alongside the GNSS to record in those areas. With this in mind, it saves time if the GNSS and TST equipment are capable of close integration so that they can share the same survey file.

It is sometimes very difficult to survey complex sites with subtle earthwork relationships using only GNSS, as even the most accurate receiver is still something of a blunt tool compared with hand drawing. In such circumstances the GNSS receiver is used to fix the location of a network of markers, such as pegs, and the resultant plot showing the pegs is then used as the basis for tape-and-offset survey with the detail added by hand drawing (Bowden 1999, chapter 4; 2002). Typically this involves a team of two people, with one measuring on the ground and the other drawing, and indeed, even when using GNSS, a collaborative approach works best when dealing with complex and subtle earthworks, as the interpretation and depiction can be discussed between the surveyors as the work progresses.

With the scale of survey decided and the decision made to use GNSS, then the question arises as to what grade of GNSS receiver to use. The preceding section has described the capabilities of the three main categories of receiver, so lower accuracy devices should not be used to try and deliver a detailed, large-scale survey, and similarly, if there is no requirement for the highest degree of accuracy, then it is a waste of effort using a survey-grade receiver that might involve setting up an on-site base station each day. It is better to consider hiring the correct grade of receiver for the project than to attempt to use one that is not suitable for the scale of survey required.

It has to be remembered that GNSS technology, or indeed any electronic survey device, only produces a series of three-dimensional points. To turn these points into an intelligible plan or map, they must be combined with feature coding: each point recorded is given a code that identifies what it represents. Software that includes feature coding is normally installed on a data logger connected to the GNSS receiver and processed after downloading, although it is more beneficial if the data logger is capable of processing the feature codes in real time to display the developing survey plan on the controller while in use. Without this facility it is easy to omit parts of the site accidentally. Ideally the surveyor should aim to create a library of feature codes that provide a clear representation of the features in the landscape, whether archaeological or modern, and that can be used on all the electronic survey devices in use, in order to create consistent maps and plans.

Feature coding involves setting up a library of the features to be surveyed; these can be points, such as survey pegs, trees, cairns or other fixed points, or linear features, such as fences, roads and the tops and bottoms of earthworks, or polygons, such as buildings or ponds (Fig 8). As each point is recorded, a feature code is attached to it by the surveyor, which determines how it will be treated when it is processed. A point may have a symbol inserted on it, or points may be joined with a particular type of line. A polygon may be given a particular colour of fill, for example to distinguish buildings from other polygon features. Control codes are used to tell the software when to stop joining a line and start a new one, or when to finish or 'close' a polygon. Feature codes might also have associated attributes defined by the user, for example the height and diameter of a cairn or the spread of a tree. Such attribute data can control the size of how a point symbol representing a specific cairn or tree is displayed on the plan. Some software is able to take the concept of attribute data much further by offering the additional facility of linking a user-defined form to a particular feature code to allow the surveyor to record a wide range of data for use in a GIS. This facility is more commonly found with software used to log data on mapping- and navigation-grade receivers than on survey-grade data loggers, where the focus is on the creation of accurate and detailed maps.

6.1 Approaches to survey

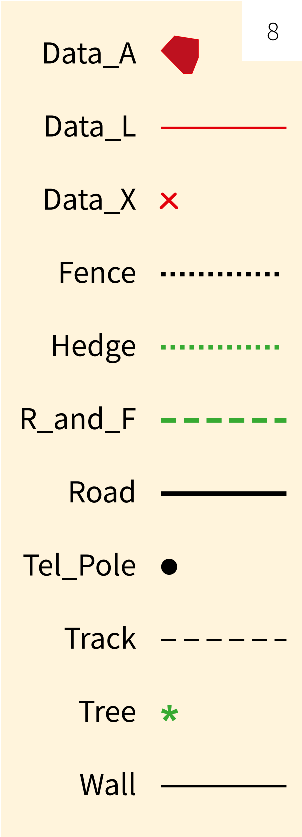
The purpose of this section is to consider in more detail best practice in tackling different types and scales of archaeological survey, including selecting the most appropriate grade of GNSS receiver for the job. It is based largely on the experience of the English Heritage field teams from across the country during the last 10 years. It is important at the end of the survey to document thoroughly, in the final report and in the archive, the methods used to survey the site, including the type of GNSS equipment and the software used for processing and how the survey was fixed to the map with what level of accuracy.

6.2 Navigation-grade GNSS survey

As this grade of receiver is primarily intended for keeping track of routes and finding locations, most of the mass-market units available have a limited facility for use as survey instruments. Some manufacturers of survey- and mapping-grade receivers also offer lower budget navigation-grade models specifically for survey, with the facility to load feature-code libraries, and background mapping and georeferenced images such as lidar and rectified aerial photographs. Such units vary considerably in the level of accuracy they normally achieve and it is advisable to check carefully what the specifications are before purchase. Normally they can only be used reliably for 1:10 000-scale mapping and above. Increasingly smartphones have in-built GNSS receivers, although for optimum results they need the additional assistance of Wi-Fi and mobile phone networks to resolve their positions. Various apps exist that can be used on smartphones for downloading pre-existing maps from the 'Cloud' and for creating forms for recording basic attribute data.

Specialist navigation-grade GNSS receivers are used by the field teams for Level 1 surveys (Fig 9), where minimum information is recorded and possibly only a single point surveyed in order to place a dot on a map to represent a site or feature (Ainsworth et al 2007). With the capability of receiving corrections from EGNOS, such receivers are able to achieve accuracies of between 2 and 4m. They are also used to find sites previously mapped from aerial photographs or shown on historic mapping, and therefore are particularly useful in upland areas where there is little map detail to guide searches. Commercial navigation-grade receivers are used outside Historic England to survey the basic components of archaeological landscapes by adapting the tracking facility to record linear features and the basic extent of settlement sites (Eastmead 2012).

8



9

Figures 8 and 9

8. Diagram showing the structure of a typical code library for use on a mapping-grade GNSS unit. The three codes Data_A, Data_L and Data_X are, respectively, for recording an area, a line and a point with a linked attribute form (as in Fig. 7), which allows the user to capture supplementary data for download into a GIS. The other codes are for recording map detail without attached attribute information (the code R_and_F refers to ridge and furrow).

9. A specialist navigation-grade GNSS unit in use to record point data for download into a GIS © Historic England, M. Bowden.

6.3 Mapping-grade GNSS survey

Mapping-grade GNSS receivers are lightweight and portable but at best are only capable of accuracies of 0.5–1m so are not capable of the relative accuracies needed for detailed earthwork survey. They are used by Historic England field teams for Level 1 surveys, and for Level 2 surveys which involve accurately mapping the location and extent of a site, usually at 1:10 000 and 1:2 500 scales, and recording attributes into a pro forma preloaded onto the device for use in a GIS.

Mapping-grade receivers generally consist of the receiver and data logger in an integrated unit designed for carrying in the hand, although there are makes where the two are separate and are mounted on a pole. This improves the accuracy and reliability of the calculated position. Older models worked with navigation transmissions from the coastguard to improve accuracy, while more recent developments have seen the introduction of receivers that can achieve a mapping accuracy of 0.1m using OS Net, which means they can be used for 1:1 000-scale surveys. In the absence of OS Net and depending upon the type of software used, it is possible to record sufficient satellite data to allow for post-processing of the results to improve map accuracy.

Case Study 1: Little Doward Landscape Park, Herefordshire

In the 18th century the steep slope below Little Doward prehistoric hillfort was turned into a landscape park overlooking the River Wye (Bowden 2009). Cliff paths, steps and grottoes were laid out for visitors to enjoy this picturesque spot, but the park fell into disrepair and has long since been abandoned. The aim of the survey was to undertake a rapid reconnaissance of the former park to identify surviving features and plot these as single points and lines on a map (Fig 10). There was no requirement to undertake any detailed survey work at this initial stage and therefore it was decided to use a navigation-grade GNSS unit designed for survey work with background mapping and a pro forma database loaded onto it to create a standardised body of information relating to each feature located with the GNSS unit.

The resulting plot showed the points and lines recorded, along with the identification numbers given to each feature in the field and forming part of the database created on the navigation-grade unit. After downloading, the data were exported in a .shp file format for use in ESRI's ArcGis software. This preserved the link created in the field between the surveyed point or line and the associated database entry. As a result it was relatively straight forward to run queries in the database and display the results on the map for interrogation on screen or for printing out for publication.

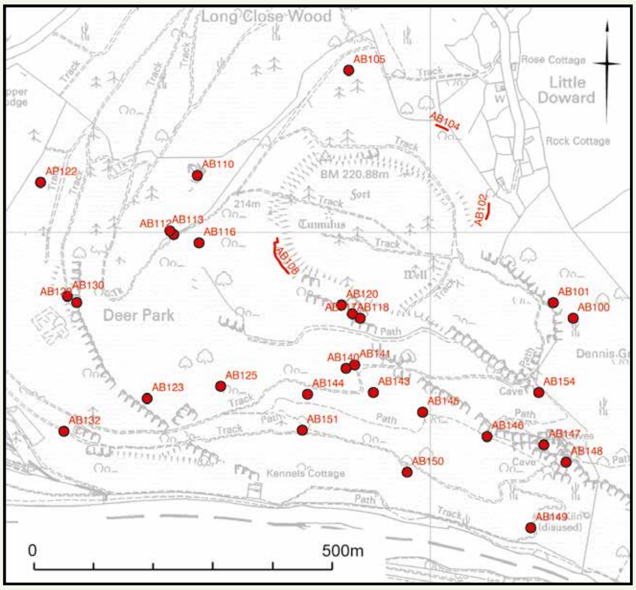


Figure 10

GIS plot of the historic garden and other archaeological features recorded at Little Doward, Herefordshire, using a specialist navigation-grade receiver on a 1:10 000-scale OS base map.

6.4 Survey procedures

A mapping-grade receiver is typically used across wide tracts of landscape where the emphasis is on creating a basic map depiction at 1:2 500 scale or smaller, rather than attempting to record the intricacies of individual earthworks. Prior to the start of fieldwork it is often helpful to load digital copies of lidar, rectified aerial photographs and historic and modern OS mapping as georeferenced image files so that they appear on the unit at the correct location as the surveyor moves about the landscape. This facility helps identify features shown on the images and saves the effort of taking a portfolio of paper maps and photographs out into the field for reference, although preparing and loading the image files will take some time prior to the start of fieldwork. The OSTN02 transformation is also preloaded on to the data logger to deliver optimum map accuracy by transforming the coordinates captured by the receiver. If the survey results are going to be fed into a GIS, then time needs to be spent before the start of fieldwork to create a feature-code library that matches the feature classes in the GIS, and to design forms for use in the field to capture attribute information about the features being mapped for use in the GIS.

6.5 Process of survey

Generally this grade of GNSS receiver is used for basic map depiction. The aim is to characterise the basic form and distribution of sites across the landscape and perhaps identify targets for more detailed survey. For example, a polygon could be used to show the extent of a settlement site and a single line the alignment of an associated field boundary or trackway. Where the survey is part of a GIS project, then more time might be spent in the field recording data about each feature as attributes, rather than on physically walking around the feature to survey it. These data could include, for example, the approximate date, the state of preservation, the nature of any threat to the feature and its relationship to other features. Some mapping-grade units also have the capability of taking georeferenced photographs that can be linked to the feature being surveyed and then downloaded into the GIS.

Case Study 2: Mendip Hills Area of Outstanding Natural Beauty, Somerset

In 2006 the former English Heritage Archaeological Survey and Investigation team embarked on a 3-year interdisciplinary project examining the historic landscape of the Mendip Hills Area of Outstanding Natural Beauty (AONB). The Mendip Hills, which lie to the south of Bristol, are largely formed by a gently undulating Carboniferous limestone plateau, and are characterised by enclosed farmland, mixed woodland and grass and heather moor. The area of the AONB covers nearly 200km2 and archaeological sites of all periods were within the scope of the project.

More than 300 Early Bronze Age round barrows were known from across the survey area, crowning the hills, ridges and valley heads, and such sites are the most common scheduled monuments. An assessment of the region's barrows and barrow cemeteries was therefore considered to be a key objective during the English Heritage project. As part of this work, a Level 2 survey was undertaken at 1:2 500 scale of some of the main barrow groups, with the aim of providing basic information about the location, form and condition of each monument. Using a mapping-grade GNSS receiver, a single point was recorded at the centre of each mound, which was linked to an attribute form stored on the receiver. Information such as the size, type and physical characteristics of each monument was recorded in the field and then downloaded into the project GIS back in the office. Georeferenced image files of OS 1:10 000-scale mapping were preloaded on to the survey device, along with point data from the National Historic Environment database representing the location of known barrows. In unenclosed areas with challenging vegetation cover, such as heather moor, the preloaded data made it quicker and easier to locate known monuments in the field because it allowed the investigator to walk directly to their immediate vicinity. Indeed, some of the low-platform cairns might have been impossible to find without the use of the mapping-grade GNSS. The unit was also light and robust enough to use over rough terrain and in all weathers.

The work added new monuments to the record as well as reinterpreting and updating information on known sites, including more accurate map positions. It also provided data in a manner that could easily be interrogated in a GIS environment alongside a range of other data sets, allowing new insights into the topographic setting, form and distribution of barrows on the Mendip Hills (Fig 11).

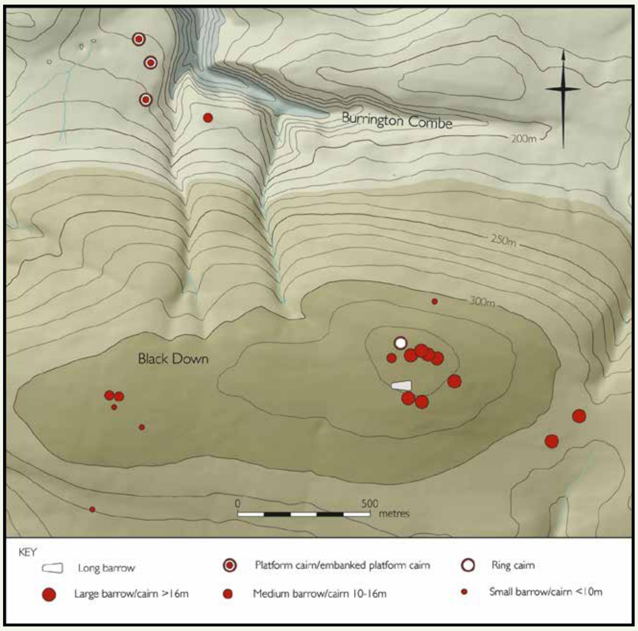


Figure 11

The Early Bronze Age barrows of Black Down, Mendip Hills, Somerset.

6.6 Survey-grade GNSS survey

Differential GNSS survey is used when the highest degree of accuracy and greatest detail is required, such as in the examples listed below.

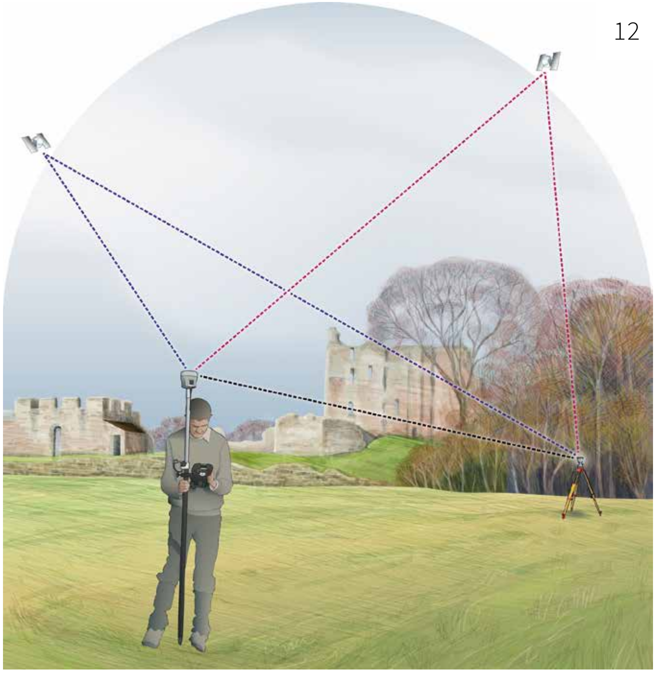
- A plan at 1:500 or 1:1 000 scale as part of a Level 3 analytical earthwork survey. As defined by English Heritage, a Level 3 survey can also include documentary research, analysis of historic views and maps, analysis of aerial photography, including lidar (if available), and assessment of past excavations (Ainsworth et al 2007, 23–4)

- Permanently marked points for survey control, fixed to the highest level of absolute accuracy on the National Grid.

- A digital ground model to create a three-dimensional representation of the ground surface or to calculate contours.

Differential GNSS survey needs at least two survey-grade GNSS receivers logging satellite data at the same time (Fig 12). One receiver (called the base) is stationary and logs satellite data at the same point (called the base station) over the duration of the survey. The second receiver is called the rover and, as the name indicates, it is carried around the site, usually mounted on a pole, by the surveyor (Fig 13). This captures satellite data every time the surveyor takes a point, but the coordinates calculated by the rover from that data will not be very accurate. The satellite data for each point surveyed is corrected from data recorded by the base receiver, thus improving the accuracy of the points recorded by the rover. The merging of the on-site base data and data logged by the rover(s) can be achieved in the office by post-processing. The closer the base and rover are to each other the better, and they must not be any more than 10–15km apart. This is to ensure that both receivers use the same area of sky and therefore log data from the same group of satellites over the same time period and encounter the same errors in satellite reception caused by the atmosphere.

12



13

Figures 12 and 13

12. Diagram illustrating a differential GNSS survey using an on-site base station.
13. A differential GNSS survey in progress at Byland Abbey, North Yorkshire, using an on-site base station and rover.
© Historic England, M. Jecock

It avoids a lot of work back in the office if the corrections can be applied to the rover in real time, which requires the base receiver to be connected to a radio antenna. This broadcasts the corrections to the rover, which applies them instantaneously to achieve the highest relative accuracy in real time. This is called real-time kinematic (RTK) differential survey. With an on-site base station, any number of rover receivers can be used with the same base station, allowing a team of surveyors to work on separate parts of the site at the same time, thus reducing the time spent in the field. RTK differential survey works best if a line of sight is kept between the base and the rover(s) in order to maintain the radio link. The link is easily lost if radio interference occurs on the same frequency, or if the line of sight is lost because of topography or the rover and base are separated by too great a distance. Radio reception between the base and rover usually only extends to a maximum of 2km. If the survey is in undulating terrain and covers a large area, then from the outset it is advisable to plan a network of base stations for RTK differential survey so that the base receiver can be moved as needed between stations to secure an uninterrupted radio link with the rover(s).

RTK survey results in a network of points all accurately fixed in relation to the on-site base station, so the resulting survey is relatively accurate but the value for each point is not absolutely accurate in terms of its National Grid reference, and consequently the result is described as a divorced survey. There may be circumstances when it is sufficient to leave the survey divorced from the National Grid, for example in some upland areas where the large-scale OS mapping is derived from legacy data. In these areas it is common to find that a feature shown on the map tile and the same feature surveyed accurately to the National Grid using GNSS do not align correctly. However, this problem is less of an issue now because of the OS Positional Accuracy Improvement (PAI) programme completed in 2006. This has produced a series of link files available to registered users that can be used to transform the position of legacy large-scale OS map tiles to improve the match with points derived from GNSS.

The alternative to a divorced survey is to locate the GNSS survey as accurately as possible on the National Grid, and this is always to be recommended as it means the survey can be compared with other georeferenced mapping and analysed alongside other geographic data in GIS systems. There are a number of ways of doing this.

- The location can be established in real time at the very beginning of the survey by fixing the position of the base station relative to OS Net using a mobile phone service. The receiver should be left for several minutes to collect data in order to calculate the most accurate value. Once the base station point is recorded, the link to OS Net can be shut down but, because the position of the base station has accurate coordinates, any rover data subsequently linked to this base station will be accurately fixed on to OS National Grid as well.

- Without a licence to access OS Net in real time, or if the link fails because of poor mobile phone reception, then the finished survey can be transformed to fit accurately on to the OS grid using post-processing software, because the OS makes all the satellite data logged by OS Net available to download for a 30-day period from its website. Post-processing against RINEX data from, for example, OS Net requires neither radio nor mobile phone coverage but has the disadvantage that the results of the survey cannot not be viewed before leaving the site and specialist software is still needed to compute the resulting positions.

- A third method, and an alternative to having a base station on site as described above, is to obtain the differential corrections for the rover receiver in real time directly from OS Net. This has the advantage of speed, as no time is spent setting up a base station each day and every point surveyed is accurately fixed on to the OS National Grid. Difficulties arise if there is poor mobile phone reception, when it then becomes a struggle to maintain a reliable link to OS Net in real time, resulting in loss of accuracy and slowing down the whole process of survey. Also, each rover that is linked to OS Net will require a separate licence from a commercial operator to access the service, which could prove expensive if two or three rovers are to be used at the same time. Experience has shown that for surveys lasting longer than a day it is usually more efficient to establish an on-site base station tied to the National Grid using OS Net than to try to work entirely using OS Net and run the risk of losing differential corrections.

6.7 RTK survey procedures

The first step is to pay a reconnaissance visit to the site to work out a strategy for the survey. This should be used to identify areas where GNSS might not work, typically next to buildings or among trees, and to work out a plan to deal with those areas, involving measures such as tape-and-offset survey to record the details graphically on to a plot of the overall GNSS survey, or the use of a TST. Many solutions now incorporate TST and GNSS measurements more or less seamlessly, allowing switching between the two point capture methods at any stage in the survey. The reconnaissance is also an opportunity to choose a position for the on-site base station. Ideally the point chosen should have a clear all-round view of the sky and be in an elevated position to facilitate real-time radio communication with the rover receiver(s). It should not cause an obstruction and should be safe from disturbance by inquisitive passers-by and livestock. The base station needs to be solidly marked as precisely the same point will be used on successive days through the duration of the survey. Finally, it is important to remember to check the mobile phone reception if the intention is to fix the location of the base station accurately using OS Net. This is even more critical if the plan is to dispense with an on-site base station all together and use a commercial network RTK solution based on OS Net for real-time corrections to the rover.

6.8 Process of survey

Whichever of the methods outlined above is used to obtain base corrections, the main focus of the survey is using the rover receiver(s) to create the survey plan, which on an earthwork site means paying close attention to the depiction of the earthworks and, most importantly, to their inter-relationships, in order to establish a relative chronology. The aims of this kind of the survey are the same irrespective of what type of survey equipment is being used and are described in detail elsewhere (Bowden 1999, chapters 4 and 5). With GNSS, the rover is used to plot out the tops and bottoms of individual earthwork slopes, using different codes from the feature code library to distinguish between the two. Other codes are used to survey 'hard' details such as roads, tracks, fences and the outline of buildings, if satellite reception is not compromised. As the survey progresses, it is immensely useful to have a data logger capable of displaying a zoomable map in real time of what has been recorded, in order to verify the accuracy of the depiction and to make sure no parts of the site are omitted. It should also be possible to review the readings taken on the data logger and to be able to delete erroneous points, rather than having to make notes to do such edits back in the office. Although it is possible to record earthworks using GNSS technology at scales larger than 1:500, experience has shown that the increase in the number of readings required for accurate map depiction and the consequent increase in time spent does not add significantly, if at all, to the understanding of a site. Also, depending on the area surveyed, it can create plans that are unwieldy to print out and too large to publish.

When it comes to areas where satellite reception is too poor to use the GNSS receiver, one technique is to lay out a network of pegs as close as possible to the 'shadow' area and survey their positions using GNSS. These can then be used as fixed points for a tape-and-offset survey or for the positions of stations for setting up and orientating a TST. In the latter case, the TST can be positioned over one peg and orientated on to a second peg as a backsight. By inputting the coordinates for both pegs as recorded by the GNSS receiver readings, the two datasets can be seamlessly integrated, especially if the same data logger is used on both devices. It is then a case of keeping the same survey file open on the data logger and swapping the logger between the GNSS receiver and the TST as required. However, when integrating GNSS and TST data it is important to bear in mind that a scale factor needs to be applied to the TST readings so that they are projected correctly to match with the OS grid coordinates logged by the GNSS receiver.

With the survey completed and the file downloaded and processed, the field teams use hachures to depict individual slopes. These can be drawn either digitally in a CAD or graphics package or by

Case Study 3: GNSS survey of Whitley Castle Roman Fort, Northumberland

Beginning in November 2007, the former English Heritage Archaeological Survey and Investigation team undertook a 1:1 000-scale survey of Whitley Castle Roman Fort near the market town of Alston in the North Pennines (Went and Ainsworth 2013). The survey formed part of the much larger landscape project called The Miner–Farmer Landscapes of the North Pennines AONB.

Because of its remote location, the fort has survived as a set of well-preserved earthworks comprising an impressive system of ditches and ramparts defending an area of about 1.25ha. The interior of the fort is covered by what at first glance looks like a bewildering pattern of slight banks, mounds and depressions within which few obvious signs of buildings survive (Fig 14). Further slight earthworks exist beyond the fort defences. The challenge was to create an accurate survey of the fort and its immediate environs, unpicking the relationship between the earthworks in order to investigate the development of the site in the Roman period and after.

The extent and complexity of the site meant this survey was going to take several weeks to complete. Therefore it was decided to establish an on-site base station rather than rely on real-time connections to OS Net for differential corrections. The point chosen was just north of the drystone wall passing through the north-west rampart, supplemented by a second base station positioned on the south-west corner tower, which was also the highest point on the site. The only drawback with these positions was that the base stations could be reached by livestock grazing the fort, so a watchful eye had to be kept on the animals. Each base station was marked with a permanent ground anchor so that the receiver could be set up exactly on whichever base station was needed to cover

Figure 14

View of Whitley Castle Roman Fort, Northumberland, showing the slight banks, mounds and depressions surviving within the fort.
© Historic England, D. Went.

hand in ink on a sheet of drawing film overlaid on the survey plot. Contours or ground models are rarely used as the primary method of depicting an earthwork site at a large scale, as neither technique delivers the same level of interpretation that can be achieved by a skilled illustrator using hand-drawn or digital hachures. Nevertheless, contours are often needed to show the detailed topographic setting of a site and ground models provide a naturalistic representation that can complement a hachure plan.

To create an accurate representation using either contours or a ground model requires a disciplined and systematic approach to data collection. Large quantities of three-dimensional points can be collected by walking with a rover receiver over a site or by driving if the receiver is mounted on a vehicle. To represent earthworks accurately, points should be recorded along the absolute top and bottom of each earthwork and along significant intermediate breaks of slope. Points should also be collected at regular intervals where there are no earthworks, in order to create an accurate model with an even surface in these areas. There is a range of software that can turn the cloud of points into a ground model. The digital ground models can then be used to generate contours to demonstrate the topographic setting of a site or to create a three-dimensional surface for visualisations and animations, or for analysis using GIS software to calculate volumes and investigate slope and drainage. It might also be possible to resurvey the same area at intervals to create comparative models for monitoring surface erosion, although TST or laser scan data will be more accurate than using GNSS.

the area being surveyed that day. They are also still available for use if further survey work is required. Because the two marked points are intervisible, they could be used as a baseline for a TST survey, which in turn would be accurately located on to the OS National Grid as they were originally recorded using GNSS and OS Net.

While the fort ramparts survive as impressive earthworks, the much shallower earthworks inside and outside the fort made little sense at first sight, but two field archaeologists, each with a survey-grade GNSS receiver, were able to unpick the pattern of earthworks to reveal the position of Roman buildings overlain by a complex of post-medieval features related to use of the fort as a farmstead and livestock pens. Disentangling the earthworks involved detailed mapping with the survey-grade GNSS receivers, following the top and bottom of each face of an earthwork. Careful attention was paid to inputting the feature codes correctly. The emerging plan on the screen of the controller linked to each receiver ensured that no part of the site was missed and that the representation of each earthwork was accurate (Fig 15a). Sufficient points were taken during the survey process to create a digital ground model of the site to be used alongside the more traditional hachure plan that was drawn-up from the GNSS plot (Fig 15b).

Figure 15

A. A plot of the interior of Whitely Castle Roman Fort, Northumberland, as recorded using a survey-grade GNSS receiver. A red line indicates the top of an earthwork and a green line the bottom.
B. The same area with the survey lines replaced by hachures to show the form and interrelationships of the earthworks more clearly.

7 The Future

It is difficult to be certain of the exact timescale, but the next few years will see vastly more satellites in orbit for use in navigation and survey, with the continuing expansion of the European Galileo and Chinese BDS systems and others still in development. In theory, having more satellites in the sky should make it possible to survey more easily in areas with restricted sky visibility, especially if more sophisticated procedures are developed to derive accurate readings from poor satellite geometry. Another practical step that is being developed to help push the use of GNSS into such shadow areas involves integrating the receiver with other sensors, such as accelerometers, gyros and laser range finding, in order to cover areas where satellite reception is lost.

Satellite technology will also advance as older satellites are replaced and new features introduced. Chief amongst these is the introduction of satellites capable of broadcasting on the new L5 frequency. Several of these are now in orbit for testing purposes in the GPS and Galileo constellations, and manufacturers are producing receivers capable of working with this new signal. The signal strength of the L5 frequency is four times stronger than that currently in use, which should mean that GNSS will be able to work in areas where the weaker signals are currently blocked. This could improve reception under tree cover, for example. It is also envisaged that L5 will bring down the overall cost of receivers, with the prospect of much better accuracy for a lower price. With this in mind, recent years have seen the development of survey and GIS apps for smartphones, and it is probable that these everyday consumer devices will be used increasingly for collecting survey data for use in GIS systems with background mapping, and the survey results stored in a cloud for access as and when required.

For those with a desire to keep up to date with developments in GNSS technology, additional material can be found in a range of subscription magazines available via the Internet aimed at the professional surveying, mapping and GIS communities. These include Geomatics World, which is the journal of the RICS Geomatics Professional Group, and GPS World (North Coast Media, USA). The websites of the main GNSS suppliers post news about imminent releases of new equipment, and there are numerous newsletters and blogs to follow online.

8 References

Ainsworth, S and Thomason, B 2003 *Where on Earth are We? The Global Positioning System (GPS) in Archaeological Field Survey.* Swindon: English Heritage

Ainsworth, S, Bowden, M, McOmish, D and Pearson, T 2007 *Understanding the Archaeology of Landscapes: A Guide to Good Recording Practice.* Swindon: English Heritage

Bedford, J, Pearson, T and Thomason, B 2011 *Traversing the Past: The Total Station Theodolite in Archaeological Landscape Survey.* Swindon: English Heritage

Bowden, M (ed.) 1999 *Unravelling the Landscape: An Inquisitive Approach to Archaeology.* Stroud: Royal Commission on the Historical Monuments of England / Tempus

Bowden, M 2002 *With Alidade and Tape: Graphical and Plane Table Survey of Archaeological Earthworks.* Swindon: English Heritage

Bowden, M 2009 *Little Doward, Ganarew, Herefordshire: Analytical Earthwork Survey of the Hillfort and Rapid Landscape Investigation.* Research Department Report Series 72-2009. Swindon: English Heritage

Eastmead, S 2012 *Landscape Surveying using Handheld GPS Receivers*. Swaledale and Arkengarthdale Archaeological Group, http://www.swaag.org/publications.htm

El-Rabbany, A 2006 *Introduction to GPS: The Global Positioning System,* 2 edn. London: Artech House

English Heritage 2011 *3D Laser Scanning for Heritage,* 2 edn. Swindon: English Heritage, https://www.historicengland.org.uk/images-books/publications/3d-laser-scanning-heritage2/

Iliffe, J 2008 *Datums and Map Projections for Remote Sensing, GIS and Surveying,* 2 edn. Caithness: Whittles Publishing

Ordnance Survey 2013 *A Guide to Coordinate Systems in Great Britain.* Southampton: Ordnance Survey

Penna, N, Clarke, P, Edwards, S and King, M 2012 *Further Testing of Commercial Network RTK GNSS Services in Great Britain.* Newark-on-Trent: The Survey Association

Royal Institute of Chartered Surveyors 2010 *Guidelines for the Use of GNSS in Land Surveying and Mapping,* 2 edn. Coventry: Royal Institute of Chartered Surveyors

The Survey Association 2012 *Best Practice Guidance Notes for Network RTK Surveying in Great Britain.* Newark-on-Trent: The Survey Association

Went, D and Ainsworth, S 2013 'Whitley Castle, Northumberland: an analytical survey of the fort and its setting'. *Britannia* **44**, 93–143

9 Glossary

Almanac A set of parameters included in the GNSS satellite navigation message that a receiver uses to predict the approximate location of a satellite. The almanac contains information about all the satellites in a constellation.

Base station A stationary GNSS receiver that serves as a reference point, providing correction data to a rover GNSS unit. Correction data can be broadcast via radio frequency, mobile phone signal or the Internet.

BDS BeiDou Navigation Satellite System. The satellite navigation system maintained by China.

CAD Computer-aided drawing/design. A term used to describe graphics packages used primarily in engineering and design. As these disciplines require a high degree of precision, they are also ideal for survey applications.

Coordinate system A pre-defined framework on to which coordinates can be related.

DOP Dilution of precision. A computed value that indicates the three-dimensional positional accuracy of every GNSS point. It can be expressed as a position DOP (PDOP) for three-dimensional accuracy or a horizontal DOP (HDOP) for two-dimensional accuracy; other variants are also used.

ECEF Earth-centred Earth-fixed xyz Cartesian coordinates.

EDM Electromagnetic distance measurement. This involves evaluating the signal returned from the target of a light beam emitted by the EDM unit. EDM is also applied colloquially to any survey instrument using this method of distance measurement.

EGNOS European Geostationary Navigation Overlay Service. EGNOS transmits a signal containing information on the reliability and accuracy of the positioning signals sent out by GPS and GLONASS.

Ellipsoid A three-dimensional geometric figure used to approximate the shape of the Earth.

ETRS89 The European Terrestrial Reference System 1989. This is the standard precise GNSS coordinate system used throughout Europe. It is a more precise definition, fixed in 1989, of the WGS84 coordinate system.

Galileo The developing European satellite GNSS constellation.

Geoid A mathematical model of the level surface that is closest to the mean sea level over the oceans. The surface is continued under the land and acts as a fundamental reference surface for height measurement.

Georeference The physical location of an object or feature according to a pre-defined coordinate system. It is used in GIS to describe the transformation of digital images such as aerial photographs or maps so that they sit in correct relation to each other within the chosen coordinate system.

GIS Geographical information system. A system for capturing, storing, checking, integrating, analysing and displaying data that are spatially referenced to the Earth. This normally comprises a spatially referenced computer database and application software.

GLONASS Global Orbiting Navigation Satellite System. A satellite navigation system maintained by Russia.

GNSS Global navigation satellite system. The generic term for satellite navigation systems, including GPS, GLONASS, Galileo and BDS.

GPS Global Positioning System. Specifically the name for the satellite constellation operated by the USA, but also a generic term used to describe surveying or navigation by reference to a satellite constellation.

Lidar Light detection and ranging. A system that uses laser pulses to measure the distance to an object or surface, typically determining the distance by measuring the time dealy between transmission of a pulse and detection of the reflected signal. Lidar is frequently deployed from a plane or helicopter to create 3-D models of the ground surface rapidly and accurately to varying degrees of resolution, depending on post spacing.

National Grid (OSGB36®) The British national standard coordinate system for OS maps. This is the grid shown on OS maps and digital data.

ODN Ordnance Datum Newlyn. This is the standard reference system for measuring heights above mean sea level across mainland Great Britain. It is measured relative to a value taken at Newlyn in Cornwall.

OSGB36® see National Grid

OSGM02® Ordnance Survey National Geoid Model 2002. A gravity-derived model used to convert ETRS89 coordinates to OS height datums, eg ODN.

OSTN02® Ordnance Survey National Grid Transformation 2002. The most recent definitive transformation used to convert ETRS89 coordinates to OSGB36 National Grid coordinates. This provides a more accurate resolution than the previous transformation (OSTN97®).

OS Net® A network of reference stations throughout Great Britain maintained by OS. It consists of more than 100 base stations. This network allows users of GNSS to carry out precise positioning within the National Grid.

RINEX A standard format for the exchange of GNSS data that can be used by any GNSS-processing software package.

rms Root mean square. A mathematical term for the statistical measure of the magnitude of a varying quantity.

SBAS Satellite-based augmentation system. A satellite system that works in conjunction with ground-based stations to supplement GNSS by broadcasting corrections to the end-user. Examples include EGNOS, which covers western Europe, and WAAS, which covers North America.

Transformation A coordinate transformation that enables a user to convert from one coordinate system to another, eg from ETRS89 to OSGB36 National Grid using the OSTN02 transformation.

TST Total station theodolite. A theodolite is a tripod-mounted calibrated optical instrument used to measure horizontal and vertical angles in order to determine relative position. On a TST the angles and distance to surveyed points are recorded digitally.

WAAS Wide-Area Augmentation System. The satellite-based augmentation system that covers North America. It is the equivalent of the European EGNOS satellite system.

WGS84 World Geodetic System 1984. The standard coordinate system for GNSS data throughout the world.

Contact Historic England

East Midlands
2nd Floor
Windsor House
Cliftonville
Northampton NN1 5BE
Tel: 01604 735400
Email: eastmidlands@HistoricEngland.org.uk

East of England
Brooklands
24 Brooklands Avenue
Cambridge CB2 2BU
Tel: 01223 582700
Email: eastofengland@HistoricEngland.org.uk

Fort Cumberland
Fort Cumberland Road
Eastney
Portsmouth P04 9LD
Tel: 023 9285 6704
Email: fort.cumberland@HistoricEngland.org.uk

London
1 Waterhouse Square
138-142 Holborn
London EC1N 2ST
Tel: 020 7973 3000
Email: london@HistoricEngland.org.uk

North East
Bessie Surtees House
41–44 Sandhill
Newcastle Upon Tyne
NE1 3JF
Tel: 0191 269 1200
Email: northeast@HistoricEngland.org.uk

North West
Suites 3.3 and 3.4
Canada House
3 Chepstow Street
Manchester M1 5FW
Tel: 0161 242 1400
Email: northwest@HistoricEngland.org.uk

South East
Eastgate Court
195-205 High Street
Guildford GU1 3EH
Tel: 01483-252000
Email: southeast@HistoricEngland.org.uk

South West
29 Queen Square
Bristol BS1 4ND
Tel: 0117 975 0700
Email: southwest@HistoricEngland.org.uk

Swindon
The Engine House
Fire Fly Avenue
Swindon SN2 2EH
Tel: 01793 414700
Email: swindon@HistoricEngland.org.uk

West Midlands
The Axis
10 Holliday Street
Birmingham B1 1TG
Tel: 0121 625 6820
Email: westmidlands@HistoricEngland.org.uk

Yorkshire
37 Tanner Row
York YO1 6WP
Tel: 01904 601901
Email: yorkshire@HistoricEngland.org.uk

Acknowledgements

The text was written by Trevor Pearson incorporating sections written by Stewart Ainsworth and Bernard Thomason for the first edition of these guidelines in 2003, incorporating editorial contributions and comments from Paul Cruddace and Colin Fane, then both with the Geodetic Section of the Ordnance Survey.

Illustrations were prepared by John Vallender, Trevor Pearson and Philip Sinton and the publication was designed by John Vallender. Photographs are by Trevor Pearson unless otherwise credited. Internal and external colleagues who commented on this new edition are thanked for their responses.

Historic England

We are the public body that looks after England's historic environment. We champion historic places, helping people understand, value and care for them.

Please contact
guidance@HistoricEngland.org.uk
with any questions about this document.

HistoricEngland.org.uk

If you would like this document in a different format, please contact our customer services department on:

Tel: 0370 333 0607
Fax: 01793 414926
Textphone: 0800 015 0174
Email: customers@HistoricEngland.org.uk

HEAG047
Publication date: First edition September 2003

Design: Historic England

Please consider the environment before printing this document