Solving the Grid-to-Ground Problem when Using High Precision GNSS in Archaeological Mapping

William (Fred) Limp and Adam Barnes

The techniques used in archaeological mapping follow the lead of the techniques used across the traditional surveying community, but usually lag in time. Increasingly, surveying is now leaving EDM (Electronic Distance Measurement) technology behind and moving to survey-grade GPS/GNSS (Global Positioning System/Global Navigation Satellite System). The particular approach is usually called real-time kinematic, or RTK, and it provides both rapid (often in a few seconds per observation) and precise measurement (typically 1- to 2-cm precision for horizontal and 2 to 4 cm for the vertical at a 95 percent level) over very extensive areas (multiple kilometers). We can anticipate that this technology will become increasingly used in archaeological projects. However, when utilizing high-precision satellite-based methods for archaeological location purposes, there is a significant but usually unanticipated problem that must first be addressed if accurate measurements are to be made. Simply put, unless proper correction methods are used, horizontal distances between two locations determined by the GNSS method will differ from the measurements made by an EDM (or a traditional measuring tape). This is often referred to as the grid-to-ground problem.

ABSTRACT

Increasingly, high-precision GPS/GNSS (global positioning system/global satellite navigation system) based real-time-kinematic methods are being used in the mapping of archaeological sites. However, when utilizing high-precision satellite-based methods for archaeological location purposes, there is a significant but usually unanticipated problem that must first be addressed if accurate measurements are to be made. Simply put, unless proper correction methods are used, horizontal distances between two locations determined by the GNSS method will differ from the measurements that are made by traditional survey methods. This difference between the two measurements is often referred to as the grid-to-ground problem. This article provides a process to address this problem.

Cada vez más GPS/GNSS de alta precisión (sistema de posicionamiento global/sistema global de navegación por satélite) basado en métodos cinemáticos de tiempo real se están utilizando en el mapeo de sitios arqueológicos. Sin embargo, cuando se utilizan métodos de alta precisión basados en satélites, hay un problema importante y generalmente no previsto que debe de resolverse antes de medir ubicaciones arqueológicas con exactitud. En pocas palabras, a menos que se utilicen métodos adecuados de corrección horizontal, las distancias medidas entre dos ubicaciones determinadas por el método GNSS diferirán con las realizadas por métodos tradicionales de prospección. Esta diferencia entre las dos medidas se refiere a menudo como el problema de la cuadrícula a la tierra. Este artículo proporciona un proceso para solucionar este problema.

The problem derives from the way in which distances are computed by the two methods. In an EDM, the phase difference created by the transmission of light from the instrument to the reflector and back allows direct measurement of the distance through calculation of the time of transmission. This calculation (with some adjustment for atmospheric properties) corresponds exactly to the traditional tape distance. GNSS however, does not directly measure distances; instead, it reports observed (measured) positions in space. From these positions, software can calculate distances. These can be horizontal (2D) distances or slant (3D) distances.

Using GNSS, the precise location of each observation is determined by satellite trilateration. The locations reported are in

https://doi.org/10.7183/2326-3768.2.2.138 Published online by Cambridge University Press

the specific form of latitude and longitude coordinates referred to as WGS84 (G1674). The angular, geographic coordinates (latitude and longitude) are typically converted to Cartesian coordinates of grid systems such as the Universal Transverse Mercator (UTM), state plane coordinate systems (SPCS), or other similar grid systems around the world. Horizontal distances are then calculated based on the Cartesian differences between these grid coordinates. The Cartesian distances measured from GNSS positions can initially differ significantly from the taped or EDM distances. Specialized techniques are needed to ensure that taped/EDM and GNSS measurements are the same. How to accomplish this is the focus of this article. Note that the following discussion simplifies a number of the more complex geodetic and map projection processes but does preserve the practical essences of them. For more details, see the suggested readings at the end of the article.

In order to understand how the difference between EDM and GNSS measurements arises and how to correct for it, it is first necessary to become familiar with some aspects of the method used for location reporting in GNSS and its relationship to standard mapping grid systems. The GNSS provides three coordinate values: latitude, longitude, and the height (h) above the ellipsoid (also called ellipsoid height or geodetic height). These coordinates are defined with respect to a particular mathematical characterization of the shape of the earth. This characterization, referred to as a datum, is defined by numeric parameters (there can be from 3 to 14 parameters) that locate the center of the earth and define the shape of the ellipsoid and (in some versions) the effects of crustal movement—called velocities. These parameters change through time as additional information becomes available to the geodetic community. Each revision of a datum is referred to as a realization. The current realization of the shape of the earth that is used by the GPS is WGS84 (G1674). WGS84 refers to a datum—called the World Geodetic System—first released in, not surprisingly, 1984. Since that time, the parameters have seen increased refinement, creating multiple realizations of WGS84, with the most current one being G1674 ("G" stands for GPS observations used to create the realization and "1674" for the realization's operational start on GPS Week 1674).

In order to use GNSS positions within common coordinate systems such as UTM or SPCS, the geographic coordinates must be converted to Cartesian coordinates. This process, referred to as map projection, addresses the age-old problem of displaying three-dimensional global coordinates on a flat surface. To accomplish this conversion, it is first necessary to define the position and orientation of a map surface (a Cartesian plane) with respect to the specific ellipsoidal characterization of the shape of the earth. With this, the latitude and longitude values are then "projected" on to the surface. Many different mapping systems are used, but the most common systems wrap the ellipsoid with a cylinder or a cone-shaped surface and use a conformal projection (which preserves shapes and angles, but distorts distances). In the most widely used map projections, the cylinder or cone intersects the surface of the ellipsoid at two "lines of intersection" and is referred to as a secant projection. The cylinder or cone is then unwrapped to reveal the flat Cartesian surface. Remembering that the positions reported by the GNSS are in latitude, longitude, and ellipsoid height, we can see that these are, effectively, coordinates defined by a line projecting from near the center of the earth outward. Therefore, a horizontal line connecting two coordinates will have a different length depending upon its distance from the center of the earth. This is shown in Figure 1.

This is one of the mathematical reasons why GNSS distances and tape distances differ. The greater the distance that the map surface is from the ellipsoid (above or below), the greater the difference is between a map (grid) distance measurement and an ellipsoidal distance measurement. This is, in effect, a type of map projection distortion. On the other hand, where the map surface intersects the ellipsoid, there is no distortion and a one-to-one relationship exists. When the map surface is below the ellipsoid, the grid distance is less than the ellipsoidal distance. And when the map surface is above the ellipsoid, the grid length is greater than the ellipsoidal length, as shown in Figure 2.

The mathematical relationship between the grid length and the ellipsoidal length is called the scale factor. This factor can be calculated for any location based on the map projection parameters. At those locations where the mapped surface intersects the ellipsoid, the scale factor has the value of one. Where the map surface is below the ellipsoid, the scale factor is less than one and, inversely, when the map surface is above, the scale factor is greater than one (Figure 3).

Except along the line(s) of intersection, an ellipsoidal distance will differ from the equivalent grid distance. This is demonstrated by a UTM map where the scale factor varies continuously across the zone from .9996 in the center (north and south along the central meridian), 1.0 along the lines of intersection (which run north and south roughly 180 km east and west of the center), and greater than 1.0 along the eastern and western parts of the zone. By way of a specific example, let us assume you converted two GNSS coordinates to UTM and they happen to fall in the center of your UTM zone. If the coordinates were 1,000 m apart on the ellipsoidal surface, their map (grid) distance apart would be 999.6 m—a difference of 40 centimeters. Conversely, if a map indicated that two points were 1,000 m apart, the ellipsoidal distance would be 1000.4 m. Scale factors for any UTM coordinate can be calculated using the National Geodetic Survey's (NGS) online Geodetic-to-UTM calculator (NGS 2013a). Note that the NGS calculations use NAD83 rather than WGS84 (see below).

The ellipsoidal surface, however, is not the actual (ground/earth/topographic) surface. The ground surface may be above or below the ellipsoid surface. A similar relationship occurs between ground measurements and grid measurements. The greater the distance between the grid surface and the ground surface, the greater the difference between the mapped distance and the "true" distance (Figure 4).

A measure of the vertical difference between the ellipsoid surface and ground surface is frequently referred to as the elevation factor. The elevation factor and the scale factor are multiplied together to create the combined factor. The combined factor is used to convert grid distances to ground distances. These factors are made part of all National Geodetic Benchmark reports that can be examined using the NGS's Data Explorer (NGS 2012). Table 1 provides an example from one report.

https://doi.org/10.7183/2326-3768.2.2.138 Published online by Cambridge University Press

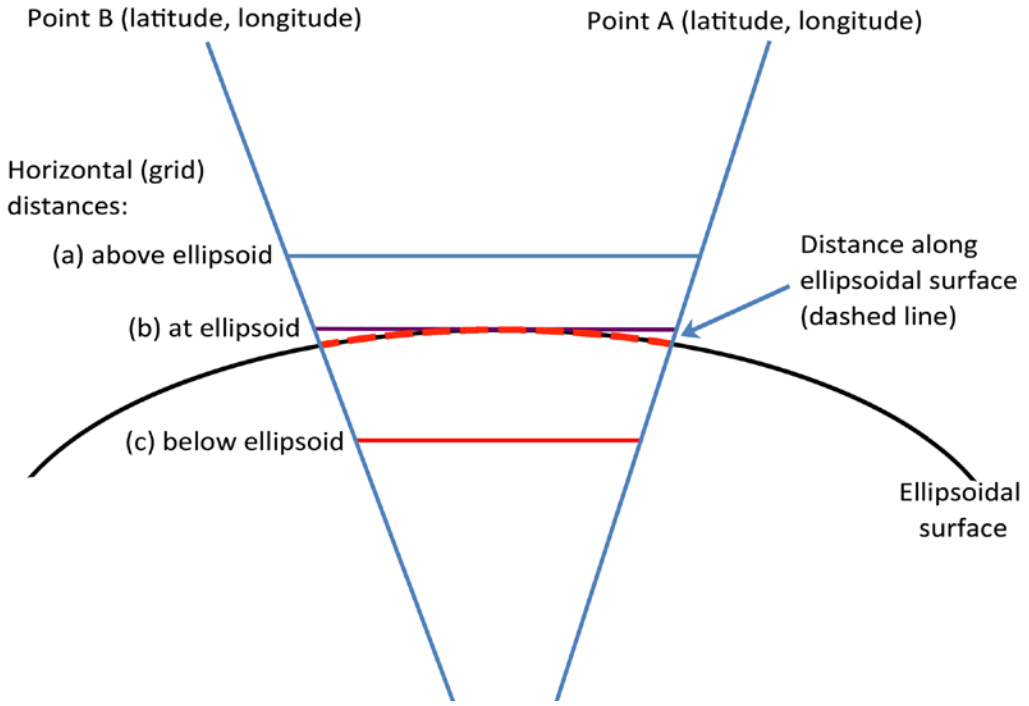


FIGURE 1. Relationship between horizontal (grid) distances and ellipsoidal surface.

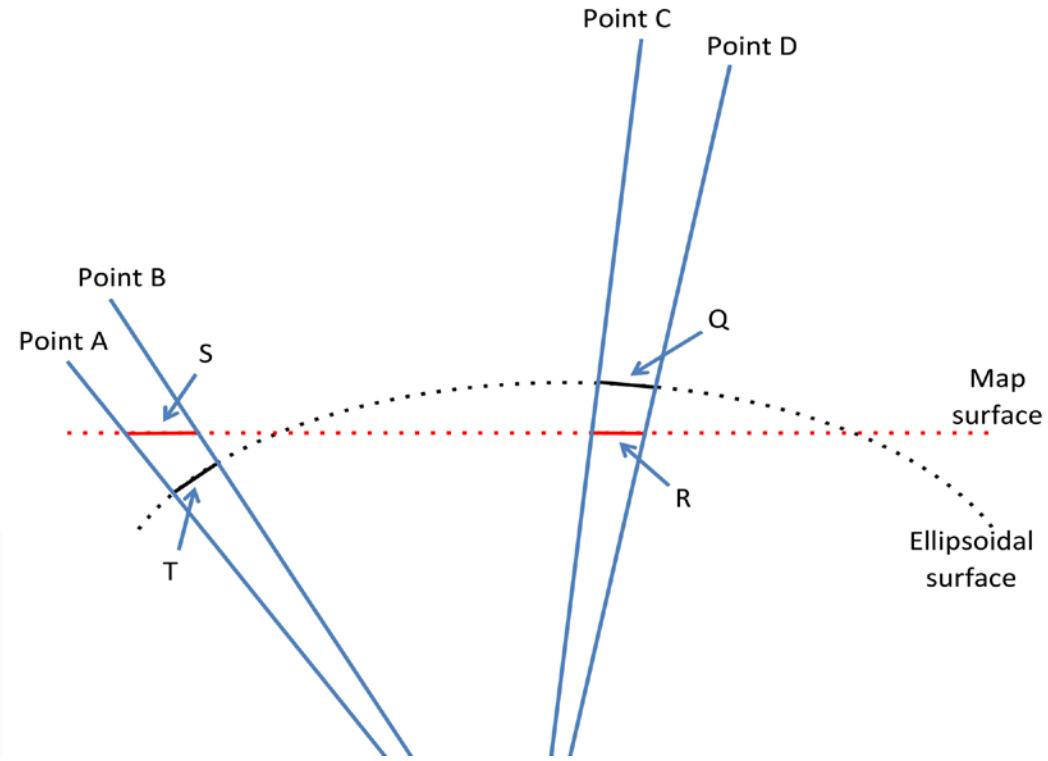


FIGURE 2. Relationship between ellipsoidal distances and map distance for locations on the map surface above the ellipsoid and below the ellipsoid. Distance Q (on ellipsoid) is greater than the mapped distance R. In contrast, distance T (on ellipsoid) is less than the mapped distance - S.

https://doi.org/10.7183/2326-3768.2.2.138 Published online by Cambridge University Press

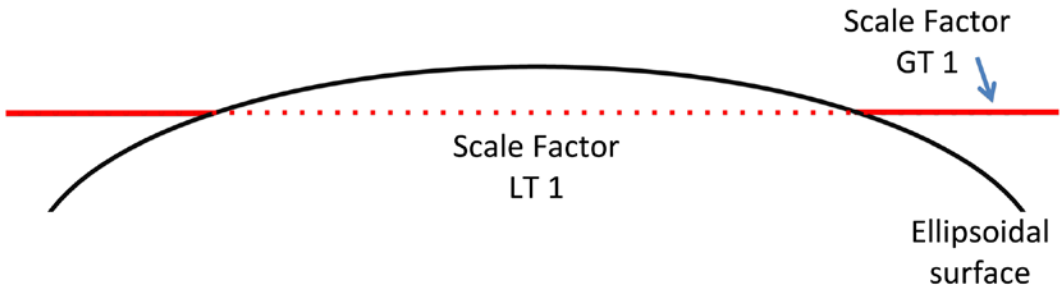


FIGURE 3. Scale factor and relationship of mapped (grid) surface to ellipsoid.

Calculation of the elevation factor is straightforward. The amount of the distortion increases as the vertical difference between the ground surface and ellipsoidal surface increases. Coordinates based on SPCS coordinates usually (not always) have smaller scale, elevation, and combined factors than do UTM coordinates.

While it is possible to use these factors to adjust grid-to-ground measurements for a particular distance measurement, a much more effective strategy to ensure that GNSS-derived measurements correspond to "true" measurements is to create a site or local area-specific low distortion projection (LDP). This may, at first, appear to be the old archaeological field strategy of creating a "floating" site grid, but it is actually quite different. A properly defined LDP can be entered into most GNSS software, allowing it to automatically convert GNSS values to accurate ground values and provide for future automatic conversion and integration of the local data to other smaller scale regional data based on traditional mapping systems. It is commonly the case that an area of up to 25 km wide can have an LDP that yields distortions of fewer than 1 mm in 1 km. Larger areas, or areas with large variation in elevations, will have larger distortions, but they will still be quite small.

In order to create an LDP the following steps are needed:

(1) Choose a new representative ellipsoid height (h_0) for the study area. This is not the elevation of the locale but considers it. The objective is to define an ellipsoidal surface that closely approximates the current surface. A good starting point is the average height above mean sea level (MSL) for the study area. To make this more accurate, you can convert digital elevation model (DEM) values from height above MSL (more correctly orthometric height) to ellipsoid height by accounting for geoidal separation (also known as geoid undulation). All GNSS software and some GIS software packages can store a geoid undulation raster on your computer. See the notes for one example of this.

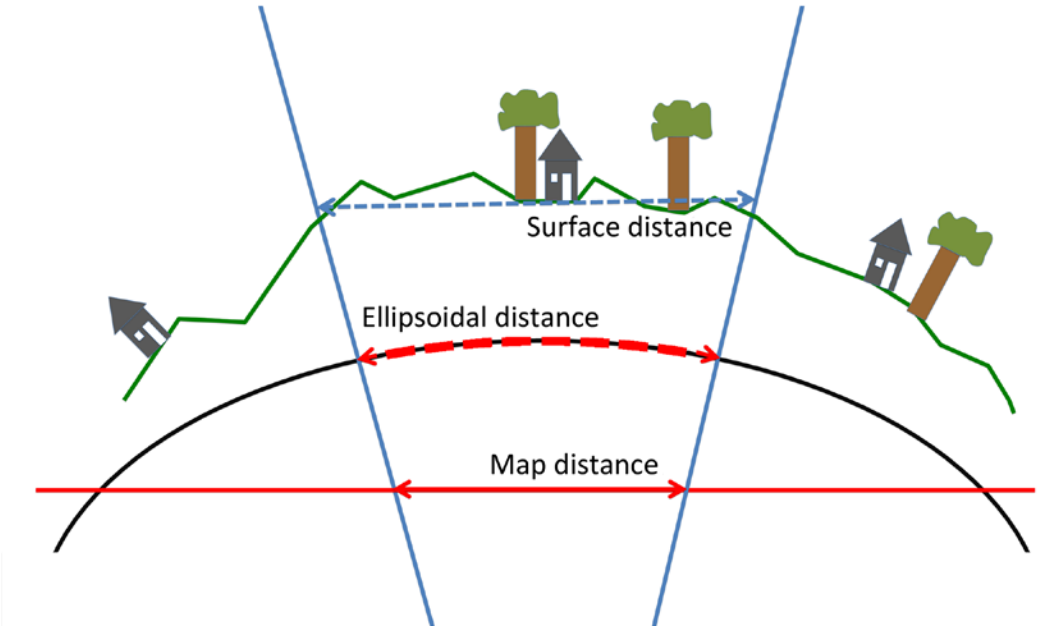


FIGURE 4. Surface, ellipsoidal, and map (grid) distances when surface is above ellipsoid.

https://doi.org/10.7183/2326-3768.2.2.138 Published online by Cambridge University Press

Table 1. A Section of the National Geodetic Survey's Data Sheet for the Cooperative Base Station Located at Old Main on the University of Arkansas Campus, Position ID AE2999.

```
AE2999. The following values were computed from the NAD 83(2011)
position.
AE2999
AE2999;                North          East    Units Scale Factor Converg.
AE2999;SPC AR N -   194,614.128   204,534.931   MT  0.99997151 -1 15 45.8
AE2999;SPC AR N -   638,496.52    671,045.02   sFT  0.99997151 -1 15 45.8
AE2999;UTM  15  - 3,992,133.643   394,639.551   MT  0.99973678 -0 41 20.0
AE2999
AE2999!           -  Elev Factor  x  Scale Factor =  Combined Factor
AE2999!SPC AR N   -   0.99993630  x   0.99997151  =   0.99990781
AE2999!UTM  15    -   0.99993630  x   0.99973678  =   0.99967310
```

Note: Retrieved March 6, 2014, from the NGS Data Explorer https://www.ngs.noaa.gov/NGSDataExplorer/.

(2) Determine the location of the grid projection axis. This will be the central meridian (longitude) of the LDP area and is usually located near the centroid of the study area. It is recommended that this value be a clean, round value (e.g., to the nearest arc minute) and distinct from other local coordinate systems (i.e., state plane) to reduce the possibility of confusion.

(3) Determine the type of projection. The LDP should be based on either Transverse or Oblique Mercator, or on a one-parallel Lambert Conformal projection. For most applications, the Transverse Mercator is probably the best choice.

(4) Determine the map projection scale factor (k_0) to ground at the central meridian. This is calculated (for the Transverse Mercator projection) as:

$$k_0 = 1 + \frac{h_0}{R_G\,(\text{e.g. earth radius at mid-latitude of LDP})}.$$

The average Earth radius of 6,371,000 m can be used. However, the geometric mean radius (or radius of curvature in the meridional plane) can be precisely calculated using the formula in the notes or approximated for certain latitudes from publicly accessible tables.

(5) Define a false easting and northing origin. This is done to ensure that all values are positive. It is often useful to set these to values such that the easting and northing are different orders of magnitude to avoid confusion and to set them to be different than the prevailing coordinate systems (e.g., UTM or SPCS).

(6) Explicitly declare the measurement unit (e.g., international foot, U.S. foot, meter, etc.).

(7) Explicitly declare the datum—WGS84 (G1674).

(8) Document all parameters used to define the LDP.

With these parameters defined, you can enter them into your GPS/GNSS controller. For example, with the Trimble Pathfinder Office™ and/or Terrasync ™ software the Coordinate System Manager provides an easy way to enter a new LDP. The parameters defined in the steps above are simply entered into the appropriate fields in the software. After you have done this, distance measurements produced between any two GNSS coordinates will closely correspond to "taped" measurements. In addition, these parameters can be entered into the great majority of software mapping solutions, which can then automatically transform the LDP values to regional, smaller scale coordinate systems such as traditional SPCS or UTM. In the popular ArcGIS™ software package, this is as simple as clicking on "New" in the "Coordinate System" dialogue and entering the parameters you defined for your LDP. In the open-source GIS package, QGIS, new projections are added using the "Custom Coordinate Reference System" dialogue. For QGIS you will need to organize your parameters as they are defined in the widely accepted "proj4" format. In both software packages, data entered from your LDP can then be projected "on the fly" to other data sets you have. This provides you with the ability to have your cake (local coordinates that correspond to taped measurements) and eat it too (easily integrate with other existing data).

There are many additional complications that may arise when using modern GNSS strategies. Conversion of the GNSS elevation from height above ellipsoid to "mean sea level" (more correctly orthometric height) requires use of a proper geodial correction. Additionally, the most recent realization of the datum used by the GPS/GNSS (e.g. WGS84 (1674)) may not be fully represented in the datum transformation capabilities in many GIS software. An earlier version may be there instead. The mapping datum used for most U.S. mapping is NAD83, and, depending on which realization is involved, there may be substantial differences (2 m or more) between it and WGS84 coordinates. Going into the issue of WGS84 to NAD83 transformations is beyond the scope of this article, but the reader may wish to read Ghilani (2013a).

https://doi.org/10.7183/2326-3768.2.2.138 Published online by Cambridge University Press

Acknowledgments

This article is a contribution of the NSF-funded Spatial Archaeometry Research Collaboration, SBE 1321443. The Spanish abstract was originally translated by Jason Tullis.

Data Availability Statement

No original data were developed or used in the paper.

Notes for Further Reading and Resources

1. For more details on the LDP and worked examples, see Dennis (2010).
2. A short article on the topic is Ghilani (2013b).
3. A discussion with an Australian flavor on "local ground-based plane coordinates" can be seen in Pickford and Gibbings (2009).
4. A British view and British Ordnance Survey examples can be found in Ordnance Survey (2013).
5. Scale factors for any UTM coordinate can be calculated using the NGS online Geodetic to UTM calculator (NGS 2013a) and a similar one for U.S. state plane coordinates (NGS 2013b). The calculations use NAD83 rather than WGS84. The downloadable program CORPSCON (U.S. Army Corps of Engineers 2004) performs conversions of many types and calculates scale factors for both UTM and SPCS.
6. The formula for precise calculation of the earth radius (R_G) is:

$$R_G = \frac{\alpha\sqrt{1-e^2}}{1-e^2\sin^2\varphi}$$

where

φ = the geodetic latitude of the point for the WGS84 ellipsoid

α = semi-major axis (6,378,137.0 m)

e^2 = first eccentricity squared = $2f - f^2$

f = ellipsoidal flattening (298.257223563)

7. An online ArcGIS ™ help file describes the process of converting raster values for orthometric heights to ellipsoidal heights (ESRI 2013).

References Cited

Dennis, Michael

2010 Low Distortion Projections (Appendix E). In *National Geodetic Survey User Guidelines for Single Base Real Time GNSS Positioning. Version 1.1*, by William Henning (lead author). Electronic document, http://www.ngs.noaa.gov/PUBS_LIB/NGSRealTimeUserGuidelines.v1.1.pdf, accessed March 25, 2014.

ESRI

2014 *Converting from Orthometric to Ellipsoidal Heights.* Electronic document, http://resources.arcgis.com/en/help/main/10.2/009t/009t000001vm000000.htm, accessed March 25, 2014.

Ghilani, Charles

2013a Where Theory Meets Practice: Transformations and Reference Frames, *Professional Surveyor Magazine*. Electronic document, http://www.profsurv.com/magazine/article.aspx?i=71455, accessed March 25, 2014.

2013b Where Theory Meets Practice: Grid versus Ground. *Professional Surveyor Magazine*. Electronic document, http://www.profsurv.com/magazine/article.aspx?i=71419, accessed March 25, 2014.

National Geodetic Survey (NGS)

2012 *National Geodetic Survey Data Explorer.* Electronic document, https://www.ngs.noaa.gov/NGSDataExplorer/, accessed March 25, 2014.

2013a *UTM Utilities.* Electronic document, http://www.ngs.noaa.gov/TOOLS/utm.shtml, accessed March 25, 2014.

2013b *SPC Utilities.* Electronic document, http://www.ngs.noaa.gov/TOOLS/spc.shtml, accessed march 25, 2014.

Pickford, Jeffrey, and Peter Gibbings

2009 Local Ground-Based Plane Coordinates. *Spatial Science Queensland.* Electronic document, http://eprints.usq.edu.au/5212/1/Local_Ground-Based_Plane_Coordinates_-_Submitted.pdf, accessed March 25, 2014.

Ordnance Survey

2013 *A Guide To Coordinate Systems in Great Britain.* Electronic document, http://www.ordnancesurvey.co.uk/docs/support/guide-coordinate-systems-great-britain.pdf, accessed March 25, 2014.

U.S. Army Corps of Engineers

2004 *CORPSCON Version 6.0.* Electronic document, http://www.agc.army.mil/Missions/Corpscon.aspx, accessed March 25, 2014.

Authors

William (Fred) Limp ■ Department of Geosciences and Center for Advanced Spatial Technologies, University of Arkansas, Fayetteville. 304 JB Hunt University of Arkansas, Fayetteville AR 72701 (flimp@uark.edu)

Adam Barnes ■ Center for Advanced Spatial Technologies, University of Arkansas, Fayetteville. 304 JB Hunt University of Arkansas, Fayetteville AR 72701 (adam@cast.uark.edu)

https://doi.org/10.7183/2326-3768.2.2.138 Published online by Cambridge University Press