12 Multiple viewshed analysis using GIS and its archaeological application: a case study in northern Mull

Clive L.N. Ruggles, David J. Medyckyj-Scott & Alun Gruftydd

Summary
The North Mull project (1987–91), run jointly by Clive Ruggles and Roger Martlew, attempted an archaeological and archaeoastronomical analysis of the short (three-, four- and five-) stone rows of northern Mull through an integrated program of survey, excavation, and analysis. Part of the project has involved an analysis of the distribution of the sites in relation to various landscape parameters, a type of analysis for which Geographical Information Systems (GIS) are rapidly becoming part of the standard armoury. However, the particular demands of the archaeoastronomical analysis have led to the development of particular GIS techniques to do with viewshed analysis that are of potential relevance to a wider range of problems within archaeology. Having given a background to the archaeological problem, this paper first discusses these “multiple viewshed” techniques from a theoretical point of view, then proceeds to describe the technical considerations involved in their practical application, and finally details some of the actual difficulties encountered when they were applied in the particular context of the north Mull problem.

12.1 INTRODUCTION

Geographical Information Systems, which have been used by many branches of the scientific community, are currently breaking new technical ground in archaeological research. Because GIS integrates both cartographic/survey data and attribute data, archaeologists are able to use the technology to record and store detailed descriptions of, for example, excavated sites (Jackson 1990). Through overlay techniques, it is possible to identify trends and patterns of distribution from a micro scale, an example being occupation levels within individual sites (Werner & Brock 1992), through to the macro scale, in cases such as Gaffney & Stančić’s (1992) work on the island of Hvar, and in particular to identify distribution patterns in relation to the landscape (e.g. Kvämm 1992).

The problem that motivated the project described here relates to a group of short stone rows, a type of site of which many examples are found in western Scotland, northern Ireland and south–western Ireland, thought to date generally to the later part of the third and earlier part of the second millennia bc (Thom et al. 1990; Ó Nualláin 1988). Survey work undertaken over the last fifteen years has demonstrated that the orientations of the three-, four and five-stone rows found in western Argyll, and in particular concentrations in the Kilmartin area of mid–Argyll and the northern part of the Isle of Mull, exhibit a pattern which is related to the rising and setting positions of the moon (Ruggles 1984; 1985; 1988). The north Mull project was undertaken between 1987 and 1991 in order to investigate this result further in its fuller archaeological context through an integrated program of excavation, locational analysis, horizon survey, and statistical investigation: full interim reports are available elsewhere (Ruggles & Martlew 1989; Ruggles et al. 1991; Ruggles & Martlew 1992).

A qualitative analysis of the Mull survey data (Ruggles et al. 1991; Ruggles & Martlew 1992) suggests that there were two important considerations in placing the north Mull stone rows. The first was to have a prominent peak to the east of south in a position associated with the rising position of the moon close to its southernmost limit
or “major standstill” (see e.g., Heggle 1981). In the case of five sites in the north of the island, this function was served by Ben More, the highest mountain on Mull. The second consideration appears to have been a non-local horizon in the south and to the west of south. The two considerations taken together suggest that it may have been important that, having risen, the moon should be clearly visible until it set again. There is also evidence that at some sites the builders managed to make use of second peaks or groups of peaks, associating them with southernmost moonrise nearer the “minor standstill” or with the rise or set of the solstitial sun.

A quantitative analysis was then required in order to explore, alongside the usual landscape factors that may have acted to determine the position of sites (local topography, elevation, soil type, land use capability, proximity to water and so on), the possible relevance of symbolic factors such as the visibility of prominent horizon features and the visibility of prominent (recurrent) astronomical events. The tentative conclusions given above raise the following two general questions. First, were the stone settings placed with regard to the visibility (not necessarily on the horizon) of natural features of apparent significance in the landscape, such as prominent mountains? Second, if so, were they preferentially placed so that these prominent horizon features would align with the rising or setting points of certain celestial bodies?

This paper describes two general strategies for tackling these questions. Both involve the creation of what have been called “multiple viewsheds” using a GIS. The first is a deterministic method in which we try to identify those natural features and astronomical events that best explain the observed placing of the stone rows. The second is a predictive method in which we try to identify the most likely locations of the stone rows within areas considered most suitable on the basis of other locational factors. Both strategies rest upon the techniques of site allocation modelling and viewshed analysis.

12.2 GENERAL APPROACH

12.2.1 Fundamental operations
Three fundamental operations are involved in site allocation modelling: classification, overlay and buffering. Classification simply involves allocating occurrences of phenomena to predetermined groups in order to make the data easier to handle. These groups might be binary (e.g. good or bad) but more usually involve a number of groups or classes, e.g. land use classes.

In the general sense, overlay can be considered as the process of comparing spatial features in two or more map layers. For example, information on site suitability may involve comparing information on land use, soil and drainage capabilities. More formally, an overlay can be considered as a set of mutually exclusive contiguous regions associated with a particular area (Burrough 1986). Each overlay is defined by a given attribute. The result of an overlay transformation is the production of new regions in a new attribute layer. There are many types of overlay transformation, most of which are based on Boolean logic. For example, one may feel that there is some relationship between steep slopes and a particular soil type in terms of erosion potential. Such areas can be identified by producing binary images (by re-classing a slope map and soil map) that identify the occurrence of each condition separately and then combine them using the overlay operation (Boolean AND). The resulting image would show both conditions.

Buffering is the creation of new areas from points, lines and areas (polygons) within a database. It is a special case of the overlay operation in which one of the inputs is a user-specified object — a point, line or area. A buffer is constructed around the user-specified object, resulting in the creation of a new area, comprising only those objects that fall within the area enclosing the buffered object. The user-specified object may be a buffer around a point (e.g. distance from a ceremonial site), a buffer around a line or linear feature (e.g. the area within a certain distance of a river), or a buffer around an area (e.g. the sphere of influence of a settlement in terms of trading).

12.2.2 Viewshed analysis
Viewshed analysis is a common feature of a number of GIS. Based on single or multiple line-of-sight calculations, the viewshed operation takes as its input an attitude grid or a Triangulated Irregular Network (where a surface is represented using triangles with vertices at the sampled data points (Peucker et al. 1978)) and determines all points that are visible from one or more specified view points over the given surface. The output, in a raster-based system, is a grid in which cells within the view are given the value 1 and cells outside the view are given the value 0. In a vector-based system the output is a set of new polygons delimiting those areas that
are visible. It is possible to take account of ob-
structions such as trees and buildings and the
height of the observer/feature at a specified view
point. The major shortcoming of viewshed analy-
sis is the accuracy of the resulting viewshed
(Fisher 1991). There may be errors in the digital
terrain model (DTM) used to calculate the
viewshed, in the rounding up of elevations to in-
tegers in binary-coded viewsheds, and the failure
of the algorithms to take account of break points
in the landscape, such as ridges.

12.2.3 The deterministic method
We shall consider first a deterministic method for
identifying those natural features that best ex-
plain the observed placing of the stone rows. The
proposed steps are as follows:

1) Produce a “Site map” that denotes the posi-
tions of known stone settings.
2) Input the Site map and a DTM of the area. Use
the viewshed operation to calculate those areas
viewable from each standing stone site. (In a
vector system these will be sets of polygons, of-
some of which may contain “holes”. In a raster
system they will simply be sets of cells.)
3) Create a “multiple viewshed” as the union of
the individual viewsheds for each site. (In a
vector system this involves dissolving unneces-
sary polygon boundaries and computing the
structure of the polygons delimiting the total
viewable area from all sites. In a raster system
it involves taking the set-theoretic union of the
cells in each individual viewshed. In some
raster systems stages 2 and 3 may be automati-
cally combined in one operation.) The multiple
viewshed represents the “viewable area”
within which any feature visible from one of
the sites must fall.
4) Identify, using independent criteria, any
prominent landscape features (peaks, ridges,
etc.) falling within the viewable area. The identi-
fication may be done by eye after laying the
multiple viewshed map over the DTM map.
Alternatively, it may be possible to request the
system to generate a list of potentially signifi-
cant vertices or cells using measures such as
elevation or terrain form index (Kvamme
1989).
5) For each prominent feature in turn, use the fea-
ture as the specified view point to calculate its
viewable area over the study area. In a raster
system the resulting image has a number of ar-
eas encoded either 1 (areas which are viewable
from the prominent feature, i.e. from which the
prominent feature is viewable) or 0 (other ar-
eas). With a vector GIS, a map consisting of a
set of polygon objects is created delimiting the
total viewable area from the feature.
6) Overlay each map/image in turn from Stage 5
with the site map containing the location of
each standing stone site and determine, either
by a point-in-polygon evaluation or by inspection,
how many sites are located within that
viewable area. The natural features that best
explain the observed placing of the stone rows
are those for which this number is largest.

The actual implementation of some of these
stages will differ between vector– and raster–
based systems. This can have serious effects. For
example, in a vector system the natural course in
Stage 1 is to represent the positions of known
sites as points. However, this may create prob-
lems if the viewshed alters significantly within
the area covered by the “site”. In this case, a mul-
tiple viewshed should be created in each case. In
a raster system the problem will only arise if the
extent of a site exceeds the resolution of a cell.

12.2.4 The predictive method
This method differs from the deterministic
method in Stage 1, which is replaced by the fol-
lowing:

1) Use classification, overlay and buffer opera-
tions to determine areas that could be suitable
sites for stone settings on the basis of factors
such as local topography, elevation, soil type,
land use capability and proximity to water.
Thus produce a “Site Suitability” image.

In Stage 2, the Site Suitability Image is input to-
gether with the DTM for the viewshed analysis.
The difference between this and Stage 2 in the
predictive method is that individual point sites
are replaced by finite “Site Suitability areas”.
The area viewable from a site is replaced by “the area
viewable from a site suitability area”, in other
words the union of the viewable areas for each
point within the site suitability area; this is itself a
multiple viewshed.

Stages 3–5 proceed as before. Stage 6 is re-
placed by the following:

6) Overlay each map/image in turn from Stage 5
with the map containing areas of suitable sites
using the Boolean AND operation. In a raster
system the resulting image has a number of ar-
eas encoded either 1 (areas which are suitable
sites and are viewable from the prominent feature) or 0 (areas which are either not suitable or not viewable). With a vector GIS, a map consisting of a set of polygon objects (areas which are suitable sites and are viewable from the prominent feature) is created. Overlay this image/map with the image containing the location of each standing stone site and determine, either by a point-in-polygon evaluation or by inspection, how many sites are located within areas that are both suitable and viewable. The natural features that best explain the observed placing of the stone rows are those for which this number is largest.

12.2.5 The astronomical question

The question of whether stone rows were preferentially placed so that these prominent horizon features would align with the rising or setting points of certain celestial bodies is tackled by a simple extension to the two methods outlined above. The astronomical potential of a horizon point is determined by its declination (see, e.g., Ruggles 1984: Appendix I), which can be calculated from the azimuth (bearing) and altitude of the point concerned together with the latitude of the site. The declination of a prominent natural feature can be calculated from any point, but only has meaning at points from which it appears on the horizon, and hence astronomical bodies can be seen rising or setting directly behind it. The result of displaying declination information within the viewable area of such a feature may be seen in the “declination contour map” (actually generated by hand, without the use of a GIS) of Bein More, the tallest peak on Mull, shown in one of the north Mull project reports (Ruggles & Martlew 1992: Fig. 14).

The additional images required will consist of areas encoded 1 if the declination of a given natural feature is within an astronomically significant band, and 0 otherwise. In order to determine this image for a given feature, the GIS will need to be programmed to calculate the declination from the other parameters, which it can determine directly from the DTM data. In either the deterministic or predictive approach, the viewable area for a given feature produced in Stage 5 is overlaid with the declination bands image using the Boolean AND operation, to produce an image in which the areas encoded 1 represent those from which the feature is both viewable and correlates with a significant astronomical phenomenon. The method may be refined even further by producing separate declination images for different celestial bodies (moon, sun, etc.) and different horizon events (solstices, equinoxes, etc.).

12.3 PRACTICAL APPLICATION OF THE APPROACH

12.3.1 Technical considerations

12.3.1.1 Choice of system

There is now much evidence that, depending on the application area, one data structure approach of GIS may be more useful than the other. For example, viewshed algorithms are feasible with altitude grids (raster DEMs), not with the digitised contours often stored in vector systems.

While the types of operation described in the general model (above) can be performed using a vector-based GIS, they are more effectively carried out in GIS which make use of a raster data structure. Raster GIS divides a entire study area into a regular grid of cells in a specific sequence, which is normally row by row from the top left corner of the grid. In a raster GIS, with some exceptions, the cells each contain a single value and are usually constant in size, shape and orientation. Furthermore, each grid is space-filling since every location of a study area corresponds to a cell in the raster. One set of cells and its associated values is termed a layer and there may be many layers in a raster system, e.g. drainage, elevation, land use, and vegetation cover.

Raster data structures are appealing in this application because of the simplicity of their organisation and the speed at which certain operations can be performed. Raster data can be processed very quickly to answer most analytical questions involving overlays, proximity, and Boolean queries. This is because no calculations are required to determine relative positions between layers. This can be compared to vector-based systems where the processing times for overlaying can be excessive.

IDRISI is a raster–(grid–)based GIS developed at Clark State University. It is intended as a teaching tool and to this end is relatively inexpensive. Despite being primarily a teaching tool it has been demonstrated in a number of projects to be useful for small–scale research as well. IDRISI is modular in design and includes modules for the manipulation of raster images (data entry, storage, retrieval and display), analytical work (overlay, re–classification, area calculations, distance/proximity surfaces and watersheds) and spatial statistics (auto–correlation, Thiessen polygons and trend surfaces). Of relevance to this project is
its ability to create slope gradient and aspect images from a surface, to overlay a layer onto a perspective view, and to create viewsheds.

12.3.2 Creation of DTM
Although a set of contours is suitable for the display of terrain surfaces, it is not particularly suitable for analysis or modelling, or indeed the storage of surface information in computerised form. More appropriate methods of representing and using surface information are required, and when a digital representation is used these are known as Digital Terrain Models (DTM) or Digital Elevation Models (DEM). Digital Terrain Models have many uses (Burrough 1986). Amongst these are the three-dimensional display of landforms, analysis of inter-visibility and viewsheds, computation of slope-and-aspect maps and slope profiles, and the planning of routes of travel.

The most common form of DTM is the altitude grid. One method of deriving this is through the interpolation of irregular or regular spaced data points \((x, y, z)\). Altitude grids are used by raster systems such as GRASS and IDRISI but can also be found in hybrid GIS where most of the other spatial data are stored as vectors. Despite disadvantages such as data redundancy in areas of uniform terrain and the inability to change grid size where areas differ in relief quality (Burrough 1986), the altitude grid structure is useful for computing slope angle and aspect maps, calculating contours, viewsheds and hill shading. The most popular alternative to altitude grid structure is the Triangulated Irregular Network (or TIN), a vector topological structure, which tries to overcome the disadvantages of the altitude grid representation. However, given the requirements for effective overlay modelling found in raster systems, the current research is using the altitude grid data structure, although the possibility remains of using the TIN approach in the future.

12.3.2 Data sources
The spatial and attribute data used in this research comprised the following:

- rivers and contour data from Ordnance Survey 1:50000 Map Sheet 47 (Tobermory);
- soils and drainage data from the 1:50000 soil map of the Island of Mull;
- land capability data from the 1:250000 Land Capability map which includes Mull; and
- location and attributes of sites from previous fieldwork.

Other data such as slope and aspect images were created from the DTM using functions available within the GIS (see below).

12.3.3 Implementation
Capture of the map data was done using ATLAS*GIS, a product of Strategic Mapping Inc., USA. This is an easy-to-use, vector-based GIS which runs on DOS machines, in this particular instance a IBM PS2/80 with a high-resolution graphics screen. Spatial objects in ATLAS*GIS are organised into layers. Each layer comprises objects from the same spatial class, e.g. all rivers, all land use zones, or all roads. A disadvantage of using ATLAS*GIS is that a layer can only comprise one feature type, i.e. lines, polygons or points; thus, for example, water features that are polygons, such as lochs, have to be on a different layer from water features that are lines, such as rivers. Other GIS, such as Genemap (Genasys Inc.) do not suffer this restriction. The PS2/80 is linked to a Benson 6301 A0 digitising tablet and the digitising "module" of ATLAS*GIS was used to encode the spatial data. Like most GIS projects this was the most time-consuming aspect of the project, requiring nearly three person-months. Particularly time-consuming was the capture of the soils and drainage data. Currently, the soils layer contains 953 polygons, some comprising as many as 4000 vertices.

While it is possible to link ATLAS*GIS to an external database, dBaseII+, this was not necessary in this project as ATLAS*GIS has its own internal attribute file. Into this file attribute data relating to soil type, drainage, and site attributes were entered. Many GIS require a two-stage approach to data entry, firstly of the spatial data and then the attribute data. With ATLAS*GIS it is possible to enter attribute data at the same time as the map data, this reducing data entry time and the potential for error when linking the attribute data to the spatial objects. Furthermore, general data error problems can be avoided since an operator can quickly identify, as (s)he works, inconsistencies between spatial objects and their supposed attributes.

In this initial stage of the project we decided to focus on the stone row sites in the Glengorm area using the predictive method. One reason for doing this was to determine suitable raster grid parameters (grid size, cell resolution). Additionally, we wanted to pilot the ideas outlined in the methods above using a constrained data set size so as to determine the appropriateness of the strategy and any implementation difficulties.
ATLAS GIS permits the import and export of both attribute and geographic features. Using the ATLAS GIS Import/Export facility, geographic feature files for each of the map layers for the Glengorm area were created in a tab-delimited format suitable for importing into IDRISI. Some problems were encountered with the format, but these were solved by writing simple reformating programmes. Once the geographic files were imported into IDRISI they were converted using the IDRISI vector-to-raster conversion routines.

A general problem with raster systems is deciding upon the cell resolution. The smaller the resolution, the higher the accuracy but the larger the number the cells making up the grid, with a consequent detrimental effect on processing times. Larger cell size reduces grid size and computation times but reduces the accuracy, a significant problem as far as viewshed accuracy is concerned. A decision was made to use a cell resolution of 20 x 20 metres. For the area under study this resulted in a grid containing 100,000 (250 x 400) cells.

In the next stage of the project we intend to purchase Ordnance Survey 1:50,000 DTM data, but for the pilot stage DTM data were captured by digitising the contours from Ordnance Survey 1:50,000 maps for the area around Glengorm. There were exported from ATLAS GIS to a ASCII file and read into IDRISI. Attempts to create a DTM from these data using the IDRISI Interpol module were unsuccessful, so the DTM was instead created using the interpolation routines found in the UNIMAP software, a module of the UNIRAS package. The resulting file was imported back into IDRISI. Some concern might be raised regarding the error introduced into the data through the use of several interpolation processing routines. While this will be an important consideration in later work — e.g. it might mean the difference between a site being visible or not — given that this was a pilot exercise, the issue has been ignored.

12.3.4 Results
Some initial analysis work has been done using ATLAS GIS. It has been found useful for some site allocation modelling work and for general cartographic output, but, having no terrain modelling capabilities, its use for the main phase of the approach outlined above is limited.

Using IDRISI it has been possible to generate alternative site allocation models using operations including overlay, classification (Relclass), display (Color), distance, query, surface (creation of slope and aspect images), Ortho (to display DTM, slope images), and area (calculates areas associated with each integer category in an image). Multiple viewsheds have been generated using the viewshed operation and are displayed in 2D and 3D, the latter by using the Ortho command (which generates orthographic perspective views).

The version of IDRISI being used during this stage of the project (3.1.1) was linked to a monochrome Epson FX8 dot matrix printer. This has poor-quality output and therefore no paper output of the project is available.

The current status of the project is that we have reached stage 4 of the predictive method. The reasons why we have gone no further than this are mainly a consequence of technical constraints and data limitations. These are discussed further below.

12.4 TECHNICAL SHORTCOMINGS

12.4.1 Hardware and software limitations
As we attempted to apply the methods described above for the Glengorm area a number of hardware and software constraints emerged. Most prominent amongst these were functional limitations, speed of processing, and storage limitation. The problem of speed of processing was most apparent during calculations of viewsheds, particularly for those areas viewable from a site suitability area. The machine was often left to calculate new viewsheds overnight and it was not uncommon for it to crash near to the end of the calculation. The other problem, that of storage limitation, is a general problem of raster systems. Each IDRISI overlay, reclassification, buffer, viewshed operation produces a new image. While it is possible to delete these as one works, this does remove the ability to backtrack to an older image if a particular set of operations leads to a poor model. Running out of external memory was a frequent occurrence. The version of IDRISI used (3.1.1) was also poor as far as the production of colour hard copy was concerned, due to the fact that no allowance is made for curvature of the earth in viewshed calculations (a problem when looking at large areas), and in terms of functions for automatically determining horizon lines from the results of the viewshed analysis.

The problems outlined above occurred when only a relatively small raster grid was used. If a cell size of 20 x 20 metres is chosen for a grid covering the whole of North Mull and surrounding localities (necessary to ensure inclusion of horizon features), the total number of grid cells...
would be more than one million, with dire consequences on the time required to create viewsheds and display them. This, along with the other technical problems encountered using IDRISI, has meant that the project is now being undertaken using the GIS package GRASS on a dedicated HP750 Unix machine.

12.4.2 Data limitations
A main reservation about the data used is that it reflects the landscape as is today rather than as it was in prehistoric times. For some factors such as local topography this is unlikely to be a serious shortcoming; however, to take one example, the distribution of soil types should take account of blanket peat that has formed since the middle of the second millennium bc. Pollen analysis from core samples, such as that obtained at Glengorm (Ruggles & Martlew 1989), will give a better indication of contemporary land use. Future work will address this issue.

A further problem was that of differing resolutions. Land capability data, for example, were only available at 1:250,000. Better quality, better resolution data are rapidly coming on-line.

12.5 FUTURE DIRECTIONS

Raster-based GIS are providing an extremely useful platform for the analysis of possible symbolic associations of the north Mull stone rows with prominent landscape features and horizon astronomical events. Viewshed analyses feature prominently in such work. Multiple viewshed techniques also have clear potential in other analyses of site location that are relevant to several classic examples in British Prehistory, providing tools for tackling issues such as intervisibility and prominence. A more complex example is measuring the "spatial efficiency" of candidate view points, i.e. minimizing the number of viewing locations so as to achieve total visibility of an area.

In the meantime, the technological limitations of viewshed analysis within GIS are being tackled. Thus, for example, DTM errors may be modelled by Monte Carlo simulation so as to produce a probable viewable area (Fisher 1992). Also relevant in the current context are new techniques which are being developed for feature discrimination.

Finally, while binary viewsheds constitute a well-tested GIS functionality which is widely available, related methodologies currently being developed promise much greater subtlety and potential. An example is the "fuzzy" or "cognitive" viewshed, determined by combining the normal binary viewshed with a distance decay function, which models the degree to which a distant object, although theoretically visible in perfect conditions, might actually be discernible in average circumstances.

References
Burrough, P. A.

Fisher, P.


Gaffney, V. & Z. Stanić

Heggie, D.C.

Jackson, J.M.

Kvamme, K.L.


Ó Nualáin, Seán

Peucker, T.K., R.J. Fowler, J.J. Little & D.M. Mark

Ruggles, C.L.N.


Ruggles, C.L.N. & R.D. Martlew


Ruggles, C.L.N., R.D. Martlew & P.D. Hinge

Thom, A., A.S. Thom & H.A.W. Burl

Wemer, R. & T. Brock

Authors' addresses
Clive Ruggles
School of Archaeological Studies,
Department of Mathematics and Computer Science,
and Midlands Regional Research Laboratory
University of Leicester
GB–LE1 7RH Leicester
email: rug@le.ac.uk & c.l.n.ruggles@lut.ac.uk

David Medyckyj–Scott
Midlands Regional Research Laboratory and Department of Computer Studies
Loughborough University of Technology
GB–LE11 7TU Loughborough
email: d.j.medyckyj-scott@lut.ac.uk

Alun Gruffydd
Gwynedd Archaeological Trust
College Road
Bangor
GB–LL57 2DG Gwynedd