EMANUEL SAVINI

The topographical survey and the Geographical Information System (GIS)

Methodology and documentation strategies

The main objectives for the documentation strategies can be summarised in three words: computerisation, accuracy and interoperability. At an early stage it was decided that all the information would be digitised. All topographical data and the boundaries of the survey transects were collected in digital form using a total station bringing the accuracy of the measurements down to the centimetre. All the measurements were *georeferenced*, i.e. related to a fixed trigonometric point, ensuring for the anchoring of the information.[1] A great effort was given to structuring and documenting all the parameters of our data (see the metadata section) to make sure that it can be operable within other systems.

Several circumstances pertaining to the Mastos survey project made possible the achievement of those goals: the scale of the project—the dimensions of the study area being very small—made possible the full use of total stations; the fact that a trigonometric point, part of the Greek National Grid existed on the top of the hill, allowed us to anchor the model; finally the incorporation of ArcView GIS, a geographic information system software package, was used as a tool for data management and visualisation.

Geographical Information System (GIS)

There are two ways in which a GIS handles the spatial data. The vector data structure uses Cartesian co-ordinates to store data in the form of points, lines and polygons. Additionally it registers the spatial relations between objects, referred to as topology, reorganising the data in nodes, links and polygons. The raster data structure, on the other hand, describes reality with the help of lattices. A surface is divided into a matrix of grid cells (usually square) forming rows and columns. Every cell has a given position and at the same time it stores a numerical value, an attribute, representing a characteristic of the feature. The information is stored in superimposed, consecutive layers. Each structure has its advantages and disadvantages and the combination of the two, as in the ArcView GIS software gives the best results.

In the Mastos survey project the actual software was used in this first stage as a data management tool for the integration and structuring of the data, both archaeological and environmental, gathered in various ways and in different media. Also, it was used as a visualisation tool for image processing, thematic map creation and the presentation of the distribution and density of the finds. One of the main concerns was the production of an accurate Digital Terrain Model (DTM). To provide the appropriate data terrestrial and topographical surveys were carried out.

The terrestrial survey

A Leica total-station (Wild TC 600) was used for the measurement of 49 base station co-ordinates, which, besides base station 1, were used for the topographical survey. The estimated error was less than 2 cm. For the purpose of the archaeological survey, the area under study was divided in 64 units following the natural environment. Their boundaries were measured along with some architectural features. The aim of the project was the creation of a Digital Terrain Model (DTM) as well as distribution maps of the collected and counted sherds.

At the end of each day the total station was connected with a laptop computer in order to transfer the data in the form of co-ordinate triples with two attached attributes. The attached attributes named "ID" and "Code" stored information about the kind, type and relative position of the mea-

[1] For the term georeferencing and anchored as opposed to floating or divorced surveys see Wheatley & Gillings 2002, 26 f.

Software	*Action*
Liscad plus Primer (ESP) 2.04a	Transfers data from the total station and stores it as RAW-files
MS-DOS Editor	Converts RAW-files to GSI-files
File	Converts GSI-files to PXY-files
Microsoft ® Notepad	Stores data in TXT-format
Geodos-PC	Converts PXY-files to ODB-files
Geodos-PC	Converts ODB-files to DXF-files

Fig. 31. Transfer and processing of raw data from the total station.

sured points. "ID" is a descriptive identification number that usually consists of an eight-digit integer where the first two positions refer to the type of the data (10 for a geographical point, 20 for a part of the circumference of a unit and 31–35 for a part of all other kind of lines), the following three refer to its position with regard to the units and the last three positions refer to the measurements themselves in consecutive order. "Code" refers to the vector form of the data (point, line or polygon) and is used through the translation process for the arrangement of the thematic structure of the CAD drawing.[2] For the conversion of the raw data from the total station, a number of steps and softwares were used (*Fig. 31*). When a compatible file format had been stored, it was imported into the GIS program (ArcView GIS 3.1 with 3D Analyst 1.0 and Spatial Analyst 1.1 extensions loaded); the elevation points in TXT-format (Tpunkt.txt) as an "Event theme" and the line features in DXF-format (Units.dxf; antmur.dxf; road-0.dxf). Several Digital Terrain Models were created using different interpolation methods (TIN, IDW, Spline) and testing various parameters within the raster methods seeking a model giving a more faithful representation of the coarse surface of the hill. The cell size of the raster was always 0.1 m.

For the distribution maps, the line features representing the boundaries of the 64 units were edited and transformed to polygons. Their theme table, storing only spatial data, was joined with the table containing the results of the archaeological survey and a theme was created where dots representing the sherds were spread randomly over the area of each unit. The representation of sherds by dots in the following chapters thus only specifies the number of sherds found in each unit and not their exact location within this bounded space. It would have been impossible and of little analytical value to record the exact position of all surface finds.

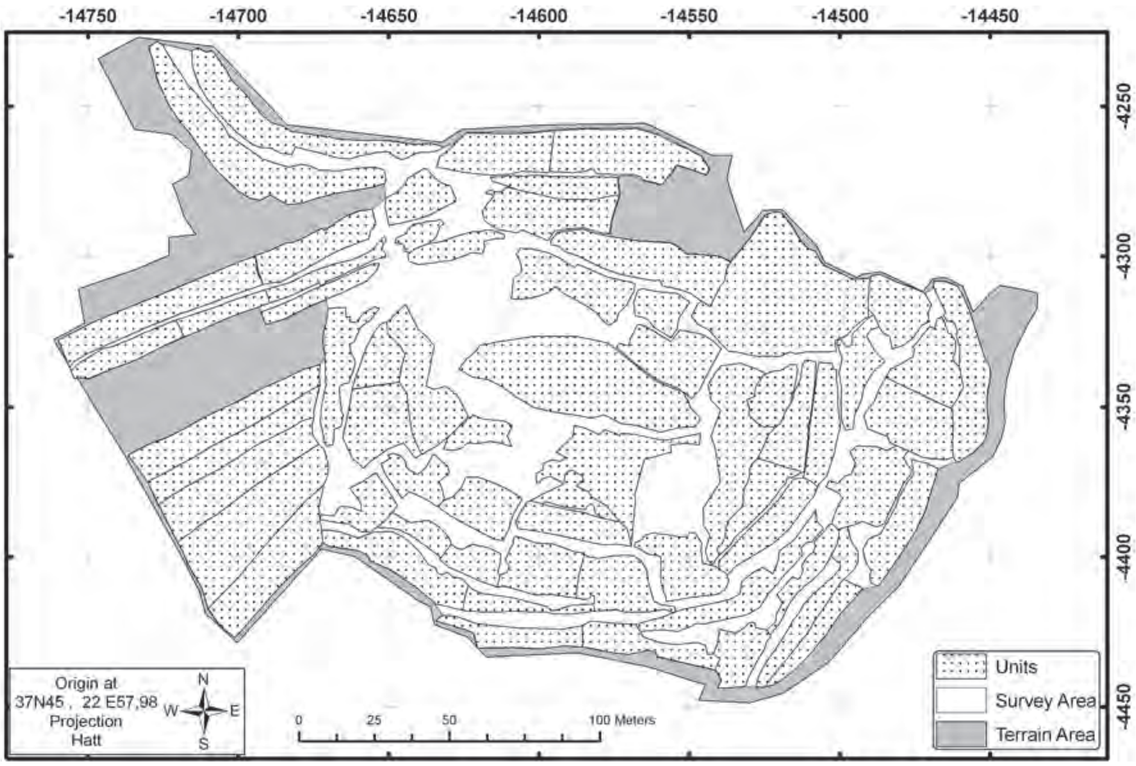

Fig. 32. The relation between the terrain, the survey area and the walked units. Illustration by E. Savini.

[2] This attribute is usually utilised under the translation/transformation process before the creation of a DXF file.

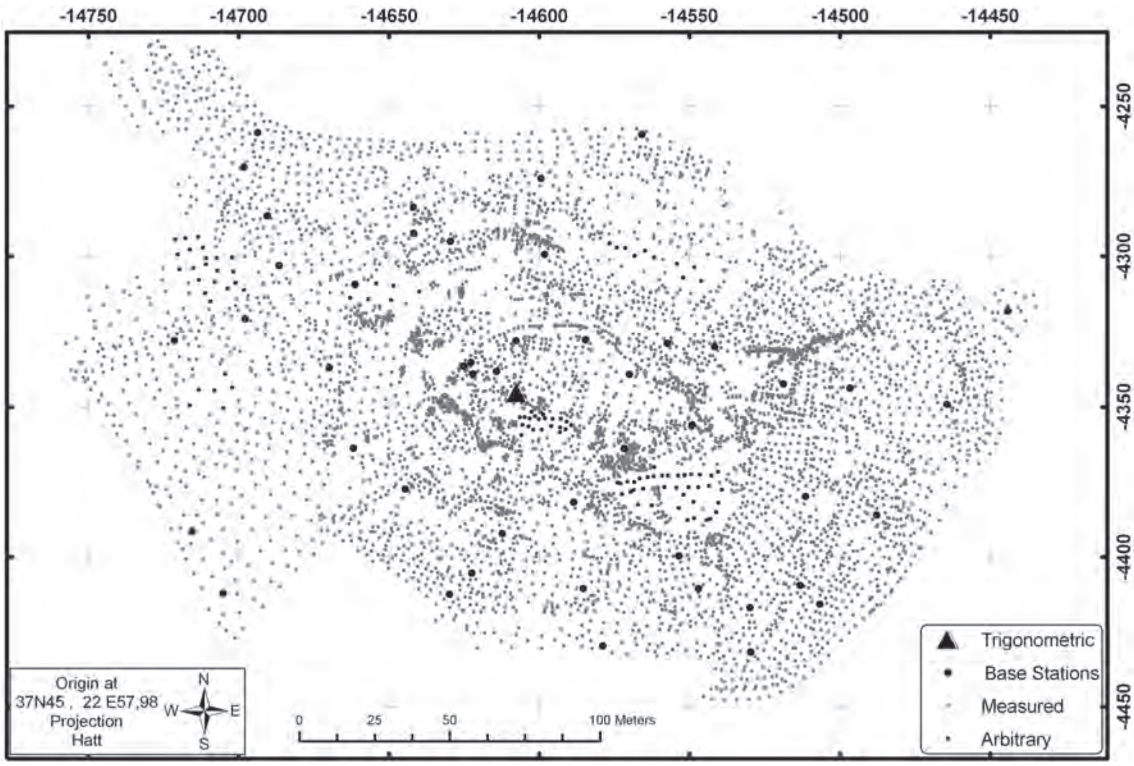

Fig. 33. Distribution of base stations, measured and arbitrary spot heights. Illustration by E. Savini.

The topographical survey

The topographical survey was conducted over an area that measures 220 × 320 m (north–south by east–west), covering a total of 44,291 m². The archaeological survey was limited in a narrower part consisting of 36,664 m² due to time limitations. Of these only 27,222 m² could be walked, the rest being inaccessible (*Fig. 32*).

Data catchment strategy

The terrain shows a significant variability in elevation with the relatively flat areas of the fields in the southwest and the agricultural terraces, mostly on the south slopes of the hill. The rough areas are concentrated mostly around the highest plateau in the centre of the hill. Because of this, irregularly spaced spot heights were preferred with higher concentrations in rough terrain (a few decimetres between height observation points) while in the areas with low variability in elevation fewer observations were made. Regularly spaced spot heights would have required much more effort and time for the same area (about seven times with a resolution of 1 metre). A total of 6,540 spot heights were gathered in the manner described using a Leica total station. Besides the terrain, a nearby road and some standing features were registered. As the terrain could not be measured in some places, leaving "white spots" in the map, 102 additional points were construed given an elevation value by approximation (*Fig. 33*).

Digital Terrain Models (DTMs)

DTMs are representations of the landscape.[3] The computation of a DTM in a GIS software like that of ArcView is quite a simple procedure. However, as all of the DTMs are approximations, an uncritical use of the software can result in unrealistic representations of the landscape.[4] The quality of the model depends on several factors related to the collection of the sample points but even more on the choice of the in-

---

[3] The terms digital elevation model (DEM) and digital terrain model are used interchangeably by many people. Wheatley and Gillings suggest that the term digital terrain model should be used when topography is described instead of the more often used but even more general term digital elevation model. Wheatley & Gillings 2002, 107. For the opposite view see Gillings *et al.* 1990, 34.

[4] See Kvamme 1990, 123; Hageman & Bennett 2000, 113 f.; Wheatley & Gillings 2002, 100.

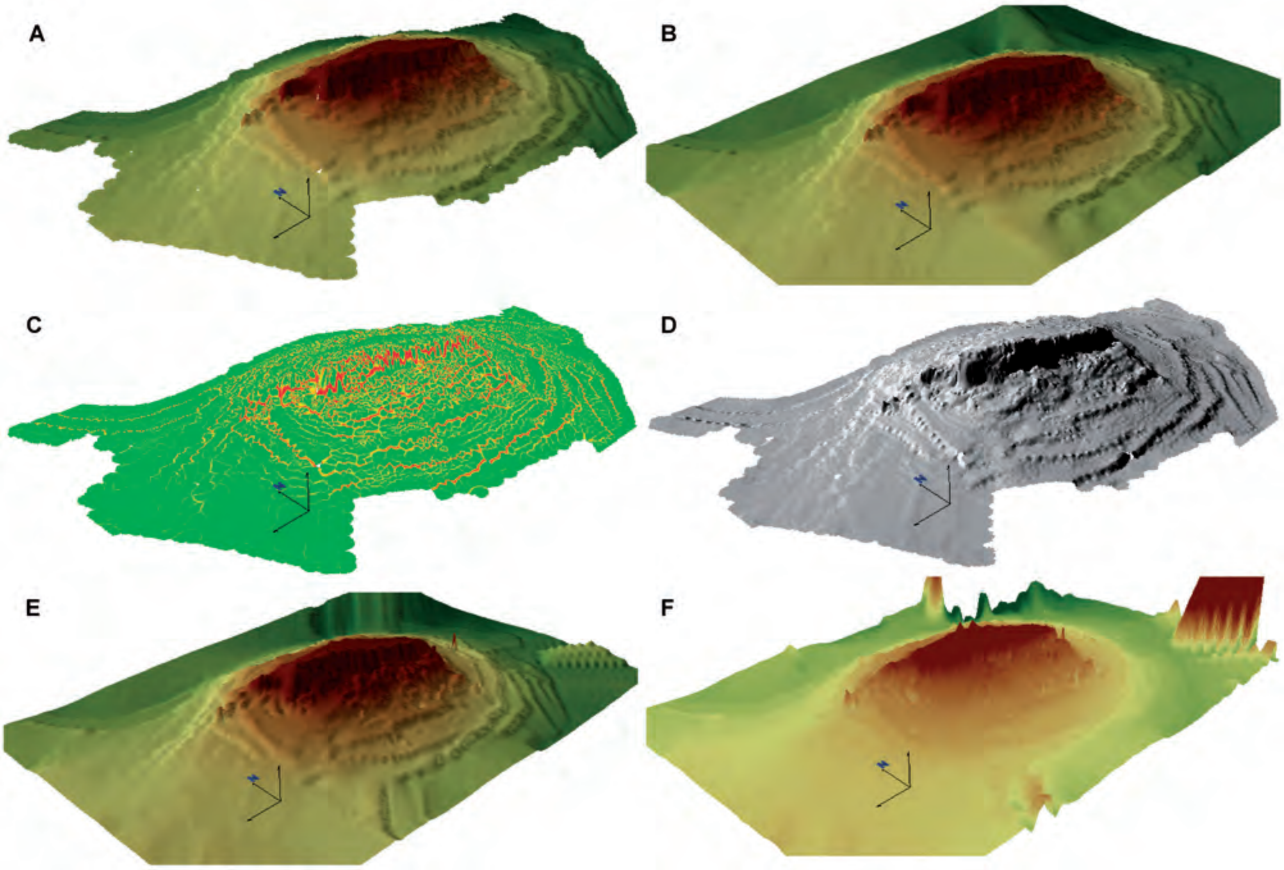

Fig. 34. Digital Terrain Models (DTMs) produced using different interpolation algorithms: A. IDW, Fixed radius 5 m.; B. IDW, Nearest Neighbours; C. Slope of A; D. Hillshade of A; E. Spline, Tension; F. Spline, Regularised. Illustration by E. Savini.

terpolation method.[5] Interpolation is, in our case, the procedure used to create the continuous surface of the DTM from a collection of point observations. The values in between are estimated using various algorithms. Given the same sample, different algorithms can produce considerably different surfaces and sometimes it is very difficult to determine which algorithm gives the best approximation of the actual elevation. Three different methods are examined bellow: IDW, Spline and TIN.

Inverse Distance Weighting (IDW)

IDW is a raster-based algorithm that assumes that each input point has a local influence that diminishes with distance. The influence that a given sample point has on an interpolated value at an unknown point is weighted by the inverse of the distance between the two.[6] Two different alternatives exist for which points can be used to determine the output value for each location; either a specified number of neighbouring points, or optionally all points within a specified radius (*Fig. 34 A–B*). A power parameter controls the significance of the surrounding points upon the interpolated value. A higher power results in less influence from distant points. The nearest neighbour option—a number of 12 was used to increase accuracy—was influenced by the edge effects mentioned by Wheatley and Gillings that distorts the model in the margin of the study area.[7] Regardless of the option, the model suffers

[5] Hageman and Bennett mention the amount and the distribution of the sample points, the accuracy of the collection device and the skill and knowledge of the collector among the factors that usually have an affect on the quality and appearance of the DTM. Hageman & Bennett 2000, 114.

[6] Wheatley & Gillings 2002, 194.

[7] Wheatley & Gillings 2002, 184.

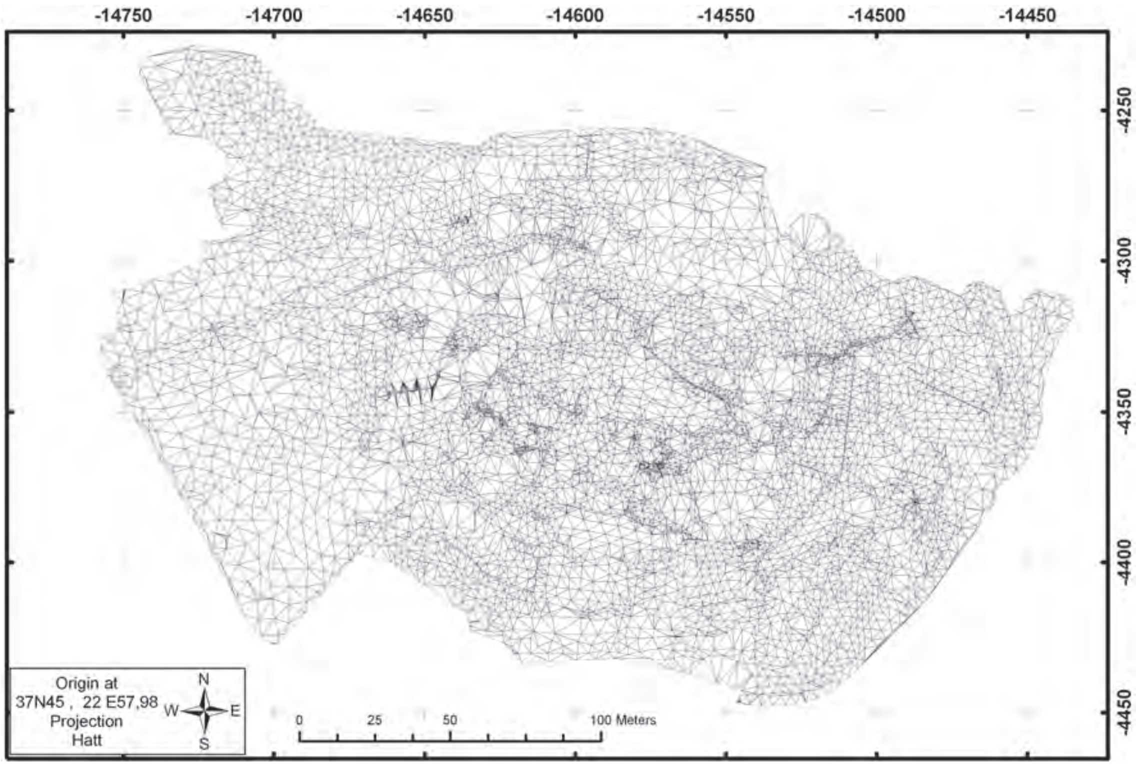

Fig. 35. The vector approach to a Digital Terrain Model (DTM): the TIN (lines). Illustration by E. Savini.

from the "stair-step" effect as can easily be seen in the slope and hill shading derivatives of the model (*Fig. 34 C–D*).[8]

Spline

The Spline interpolator, like the flexible ruler from which it borrows its name, is a method that fits a minimum-curvature surface through the input points like bending a sheet of rubber. It is a local interpolator that fits a mathematical function to a specified number of nearest input points, while passing through the spot height points.[9] In surfaces like those of the Mastos Hill, where large changes appear within a short horizontal distance, the result is not satisfactory. The two different options within ArcView, *Tension* and *Regularised*, with their parameters, allow us to control the interpolator towards a more accurate DTM (*Fig. 34 E–F*). The two models were created with the following parameters: cell size 0.1 m, weight 0.1, number of points 6 (*Tension*) and 12 (*Regularised*).[10]

Triangulated Irregular Network (TIN)

TIN is a vector-based structure that joins the height observations together with straight lines to create a set of contiguous, non-overlapping, triangles (*Figs. 35–36*). They are usually generated from a Delaunay triangulation of the irregularly distributed points.[11] Heights between nodes are calculated through bivariate interpolation of the three known elevations of each triangle thus allowing for the definition of a continuous surface. It is a flexible method that requires considerably smaller data space. The TIN model of elevation suits well the data set from the Mastos topographical survey, which is collected with the variation of the terrain in mind. It is also of

[8] The image portrays valley slopes not as the smooth declivities they actually are, but rather as the crude step-like tiers of a wedding cake. Warren 1990, 210. See also Hageman & Bennett 2000, 123.

[9] Local interpolators divide the surface in smaller parts and use different functions to calculate them. Wheatley & Gillings 2002, 184.

[10] The weight parameter defines in *Regularised* the weight of the third derivatives of the surface in the curvature minimisation expression. In *Tension* it defines the weight of tension; higher values increase the rigidity of the spline.

[11] A triangulation is considered to be a Delaunay triangulation if the circle defined by the vertices of each triangle does not contain any other point in the data set. See Hageman & Bennett 2000, 117 and Wheatley & Gillings 2002, 149 f.

fairly high accuracy being an exact interpolator and considering the overall density of the measured points.[12] Radical changes in the terrain such as terrace edges, large cliffs and features important in the micro-topography of the hill, such as the road and ancient walls, were imported as break lines and included in the calculation process since the Delaunay triangulation by itself does not take any such considerations.

Even though the terrain models in this study were created mainly for display purposes, in many cases the main reason to compute a DTM is for their derivatives. Environmental variables as slope, aspect and hill shading are very useful to various analytical processes within archaeology. Slope is the maximum rate of change of altitude and is expressed in degrees (from 0° to 90°) while aspect is the steepest down-slope direction from each cell to its neighbours. It is calculated in degrees and the values are regrouped, within ArcView, into the eight main compass directions. Slope and aspect are often used within archaeological predictive modelling. Hill shading can be used for analytical purposes to determine the length of time and intensity of the sun in a given location. The model produced by the TIN, in spite of its somehow "cubistic" appearance, is the most accurate. For that reason it was decided that this data structure would be the most suitable for obtaining the contour lines, and the slope, aspect and hill shading measurements (*Fig. 37*).

Visualisation and display

The described technology provides the means for a powerful illustration of the information on the screen. It enables the almost unlimited use of colour variations and dynamic representation in 3D. This is not unproblematic *per se* and considerations have to be made to provide results that are both aesthetically pleasing and informative.[13] However, traditional paper publications enforce a number of restrictions; the images have to be monochrome or can only display a limited colour scheme. They look flat and static. To address those issues a number of principles had to be followed.[14] Some of those principles are listed here:

(1) Scale is indicated in two ways as a measured grid on the image indicating the distance in metres from the origin of the map sheet used (see the metadata section) and as a scale bar in the lower part of the layout.

(2) An indication of north, the origin of the map sheet in geographical co-ordinates and the projection used are included.

(3) A key to the features is incorporated in the layout.

(4) The graphical elements on the page are arranged in an effort to create the appropriate amount of balance between them.

The representation of the DTM could be improved. As it is now, artificial colouring is used to give an impression of distinct altitude zones thus facilitating the perception of elevation. By adding a synchronised, aerial photograph of the area on the DTM, a more realistic impression could be provided and at the same time the information the images convey would be increased. Also, an application that combines GIS with Virtual Reality software can improve the visualisation process and make possible an interaction with the archaeological landscape.[15]

Metadata

Metadata has in simple terms been defined as "the information that you and others need to know to be able to interpret and make sensible and appropriate use of the information in the spatial database".[16] It is quite essential, in the growing world of GIS applications with the lack of common standards to follow, to document the spatial database correctly; to describe carefully the content and format of the database; and to ensure that the system can be inter-operable with other systems. The section below follows the guidance and terminology of Archaeology Data Service (ADS) as it is expressed in their guide.[17]

Projections and Co-ordinate Reference Systems

The starting point for all measurements on the Mastos was a trigonometric control point that is part of the Greek National Trigonometric Network. Its position, indicated in *Fig. 33*, as well as the publisher and copyright owner of the Greek National Grid and its components, is listed below.

[12] An interpolator is called *exact* if it requires that the result passes through the observed data points.
[13] Miller 1995, 319–324.
[14] Mostly outlined by the York Archaeological Assessment (YAA). Miller 1995, 327.
[15] Forte 2000, 210–212.
[16] Wheatley & Gillings 2002, 87.
[17] Gillings *et al.* 1990.

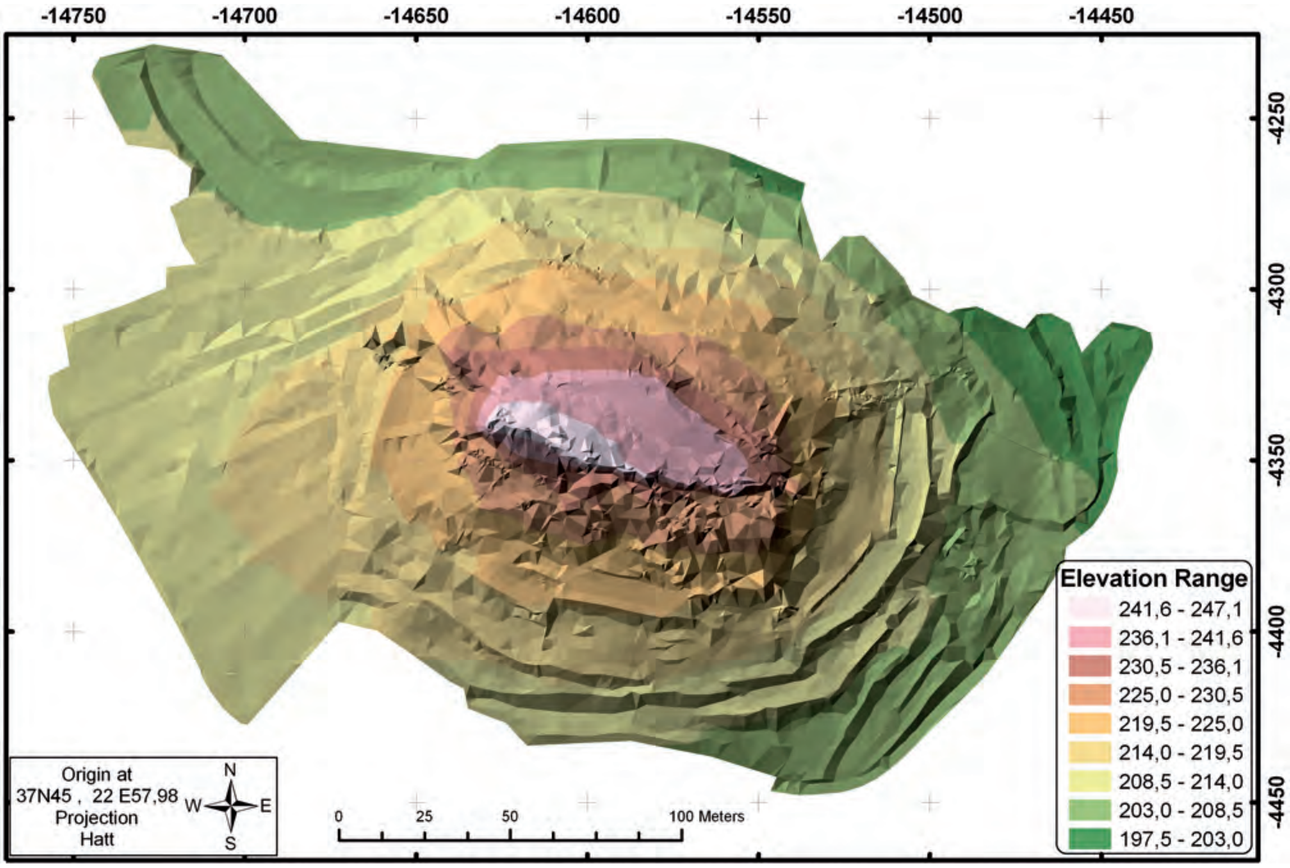

Fig. 36. The vector approach to a Digital Terrain Model (DTM): the TIN (facets). Illustration by E. Savini.

National Mapping Agency: Hellenic Military Geographical Service (HMGS)
Position: *Greek Datum*
Ellipsoid: Bessel 1841
Map projection: Hatt (Azimuthal Equidistant)
Central Point/Datum: Observatory Athens:
$\lambda 0=23° 42' 58.815''$
Height: *Piraeus tidal gauge in Piraeus harbour*
Starting point = Base station 1:

No.	Name	X=Easting	Y=Northing	Z=H
234117	ΜΕΓ. ΜΑΣΤΟΣ	-14607.67	-4345.25	246.02

Map sheet: 1:50 000 ΝΑΥΠΛΙΟΝ
Origin: $\phi_0= 37° 45'; \lambda_0=22° 57' 58.815''$

Comments

Both the precision of the method employed (a highly detailed topographical survey undertaken using a total station survey instrument), and the level of accuracy of the data (the starting point purchased through HMGS) were very high (to the nearest cm). As the study area is very small (250 × 300 m) and only one trigonometric point was used, no considerations to the curvature of the globe, different projections, or transformations between adjacent map sheets had to be made. A Cartesian co-ordinate system was utilized. The distance unit is the metre.

Hatt is the old projection used by the main provider of cartographical material in Greece, the Hellenic Military Geographical Service (HMGS). All the topographical maps of 1:5000 scale in analogue format are still in Hatt as well as the network of the trigonometric control points. While the projection is azimuthal and equidistant, i.e. preserving directions and distances from the point of origin, it is a local projection and thus problematic for the creation of maps covering large areas. The last decades a new datum, the Hellenic Geodetic

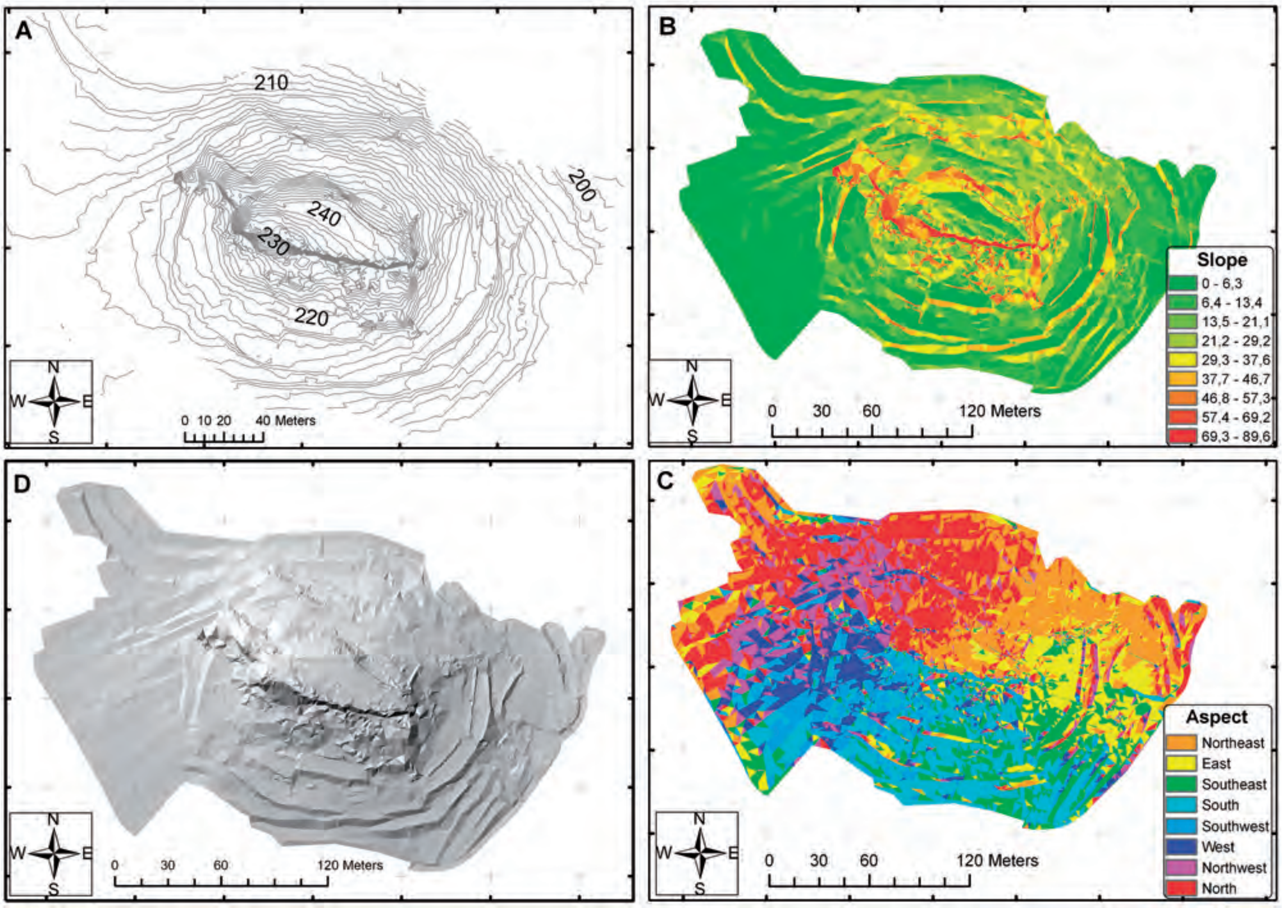

Fig. 37. A. Contours of TIN. Interval 1 m.; B. Slope of TIN. From 0% (flat) to 90% (vertical); C. Aspect of TIN; D. Hillshade of TIN. Sun position: Azimuth 315°; Altitude 45°. Illustration by E. Savini.

Reference System 1987 (HGRS87) is gaining ground and the service gradually transforms and digitalise all of their products.[18]

The comprehensive map of the Berbati Valley in *Fold-out 1* is part of a separate project to create a digital geographical database with the results of the various archaeological field projects in the area. The topographical information derives from six HMGS 1:5000 scale map sheets (63756, 63758, 63852, 63765, 63767 and 63861) which were purchased, scanned and geo-referenced in HGRS87. The contour lines, drainage system and road network were vectorised and processed in ArcInfo to produce the DEM and the final illustration.

[18] HGRS87 uses a Transverse Mercator projection, with a single zone. Central meridian is located at λ_0=24°. Scale factor is 0.9996 throughout the zone. In order to avoid negative values in coordinates along X axis, central meridian is assigned a false easting of 500,000 m. Coordinate system origin point is considered to be the intersection of the equator (ϕ_0=0°) with the central meridian. The coordinates of the Mastos trigonometric point are in this projection X=394301.18, Y=4174219.34 and Z= 245.67.