

# Geophysical Prospection at Portus: An Evaluation of an Integrated Approach to the Interpretation of Subsurface Archaeological Features

Jessica Ogden,<sup>1</sup> Simon Keay,<sup>1,2</sup> Graeme Earl,<sup>2</sup> Kristian Strutt,<sup>2</sup> and Stephen Kay<sup>1</sup>

<sup>1</sup> British School at Rome. Italy.

<sup>2</sup> University of Southampton. United Kingdom.

---

## Abstract

Recent work on geophysical data analysis suggest that in addition to a multi-method approach, “data fusion” techniques can offer meaningful insights into archaeological features, as well as allow for researchers to establish patterns between multivariate data sets that might otherwise go unnoticed. Extensive and intensive geophysical prospection has been employed at the site of Portus in recent years, playing an integral role in discerning the nature and extent of the archaeological record of the port complex. Excavations at the site have allowed for a reciprocal relationship to exist between geophysical and archaeological research, and have paved the way for a regime of meaningful, integrated geophysical analysis. Many types of geophysical and archaeological survey methods have been employed to interpret the archaeological record, as well as to provide an immense volume of data to be compared and contrasted to the excavation data. The sheer quantity of data, in addition to the nature of the archaeology at Portus, have provided an ideal site for the exploration of spatial data and remote sensing analysis techniques, as well as the assessment of their utility within archaeo-geophysical research as a whole. This research attempts to critically assess the data processing methodologies used, and to examine the applicability of a variety of mathematical and multivariate analytical approaches to the prospection results at Portus.

**Keywords:** *geophysical prospection, quantitative methods, integration, GIS, survey*

---

## 1 INTRODUCTION



**Figure 1.** *Location Map of the site of Portus.*<sup>1</sup>

The site of Portus, located to the north of the mouth of the River Tiber, served as the maritime port of Rome during the Imperial and Late Antique periods of the Roman Empire (see fig. 1). The initial construction of a

harbor at Portus is believed to have begun around 42 AD under the reign of the emperor Claudius.<sup>2</sup> This construction involved linking the harbor basin to the River Tiber through a series of canals and an aqueduct. Later, under the reign of the emperor Trajan, Portus was expanded with the construction of a hexagonal inner basin, potentially to absorb the increased economic traffic occurring between Rome and the rest of the Empire.

One of the structures of interest erected around the hexagonal Trajanic basin was an extensive complex now known as the “Palazzo Imperiale.” This structure and the surrounding area, including a sub-circular structure situated between the Trajanic and Claudian basins, form the focus of current excavations as part of the Portus Project; this is the study area for the integrated geophysical survey presented in this paper.

## 2 THE INTEGRATED SURVEY METHODOLOGY AT PORTUS

Various approaches to archaeological survey have been utilized at Portus and in the surrounding area that reflect the research goals of the project, as well as the nature and scale of the archaeological deposits on site. Emphasis has been placed on an integration of methods from the outset, with particular attention focused on multi-scaled methods for surveying the site. As part of the Roman Towns in the Middle and Upper Tiber

---

<sup>1</sup>Simon Keay et al., *Portus: An Archaeological Survey of the Port of Imperial Rome* (London: The British School at Rome, 2005) 63–64.

<sup>2</sup>Simon Keay et al., (p. 273n1) 11.

Valley Project, from 1998 to 2004 an extensive magnetometer survey was conducted throughout the region of the port complex by the British School in Rome (BSR), the University of Southampton and the Soprintendenza per i Beni Archeologici di Ostia.

As part of the extensive survey, magnetometry was chosen for its quick and efficient data capturing capabilities, as well as the good response to the buried brick structures, which showed a high magnetic contrast to the surrounding soil types at Portus.<sup>1</sup> The magnetometry, combined with extensive field walking, topographic survey and air photographic interpretation, has proven successful in locating and mapping large landscape archaeological and geomorphological features at Portus and in the surrounding area. The magnetometry was conducted using Geoscan Research FM36 fluxgate gradiometers with automatic data-logger. A 30m x 30m grid was established using a Total Station, and data was collected at 0.5m intervals along 1m parallel traverses. In total, the magnetometer survey covered an area of c.220 hectares and revealed considerable new evidence about the buildings and canals constructed around the Trajanic and Claudian harbors (see fig. 2).

### 3 INTENSIVE GEOPHYSICAL SURVEY IN THE AREA BETWEEN THE CLAUDIAN AND TRAJANIC BASINS

The Portus Project is directed by Simon Keay and is the current focus of work at Portus. It is funded by the Arts and Humanities Research Council (AHRC) in collaboration with the *Soprintendenza per i Beni Archeologici di Ostia*, and the University of Southampton, and is a flagship project of the British School at Rome (BSR) ([www.portusproject.org](http://www.portusproject.org)). In 2007, as the first phase of the Portus Project, intensive geophysical prospection began in the area between the Trajanic and Claudian basins near the “Palazzo Imperiale” with the aim of assessing the depth of overburden and features prior to the commencement of excavation.<sup>2</sup> This area of interest was targeted for many reasons, one of which lies in its unique position between the Trajanic and Claudian basins.<sup>3</sup> Within this area the magnetometry revealed a complex series of east-west linear features, and a large sub-circular feature on Side VI of the hexagon.<sup>4</sup> However, it is worth noting that in this

topographically and archaeologically complex area, the prior results of the magnetometry revealed a substantial amount of near surface rubble, complicating possible interpretations.<sup>5</sup> Thus a targeted Electrical Resistivity Tomography (ERT) survey, in conjunction with 28 shallow hand auger samples, was conducted in May and June of 2007 to complement the magnetometer survey. For this work a Geoscan RM15 Resistance Meter with a PA3 probe system was used. Four separate probes were arranged in an expanding Wenner array with 1m probe separation, and readings taken at the center point of the array. Eight traverses of data were collected along each of the seven profiles, each of varying length. By expanding the probe separation by 1m for each traverse, readings were increased by a depth of 0.5m for each traverse to build a three dimensional profile of subsurface resistance readings. This survey revealed the location of at least one harbor mole, as well as extensive structural remains to the west of the “Mura Costantiniane” (see fig. 3).<sup>6</sup>

The magnetometer and ERT surveys have revealed a great deal about the archaeological remains on Side VI of the Trajanic basin. Many questions remain unanswered, however, particularly those concerning the “Palazzo Imperiale” and the massive “warehouses” adjacent to it. The modern trackway bisecting the “Palazzo Imperiale,” as well as limitations for magnetic survey in this area, limited the interpretation and construction of a chronological sequence for this area of the port.<sup>7</sup> A core excavation area of 3000 m<sup>2</sup> was opened in 2007 on Side VI of the Trajanic hexagon, with the aim of understanding the complicated relationship of structures and deposits associated with the pre-Trajanic and Trajanic harbor structures in a key area of the port. Two seasons of excavation (2007 and 2008) have facilitated the development of a reciprocal relationship, in which light is shed upon the nature of geophysical anomalies on the one hand and, on the other, archaeologists are better able to understand features in the process of excavation. Taken together, both techniques are making a significant contribution to understanding the layout and development of the port complex as a whole.<sup>106</sup> To further complement the magnetic and resistance tomography surveys, an area resistance and Ground-Penetrating Radar (GPR) survey were conducted in May 2008, with some supplementary data capture that occurred throughout the excavation season of September 2008. The targeted survey area used within this portion of the project’s research was just under 1 hectare, on top of the eastern edge of the “Palazzo Imperiale” in the area between the excavations and the Trajanic basin.

<sup>1</sup>Simon Keay et al. (p. 273n1) 63–64.

<sup>2</sup>Work was undertaken by the Archaeological Prospection Services of Southampton (APSS) and The British School at Rome (BSR).

<sup>3</sup>Elizabeth De Gaetano and Kristian Strutt, *Report on the Geophysical Survey at Portus May-June 2007*. Archaeological Prospection Services of Southampton and the British School at Rome (Unpublished report, 2007) 4–7.

<sup>4</sup>Simon Keay et al. (p. 273n1) 99–103.

<sup>5</sup>Simon Keay et al. (p. 273n1) 9–14.

<sup>6</sup>Elizabeth De Gaetano and Kristian Strutt, *Report on the Geophysical Survey at Portus* (p. 274 n3) 4–7.

<sup>7</sup>Simon Keay et al. (p. 273n1) 9–14.

Geophysical Prospection at Portus

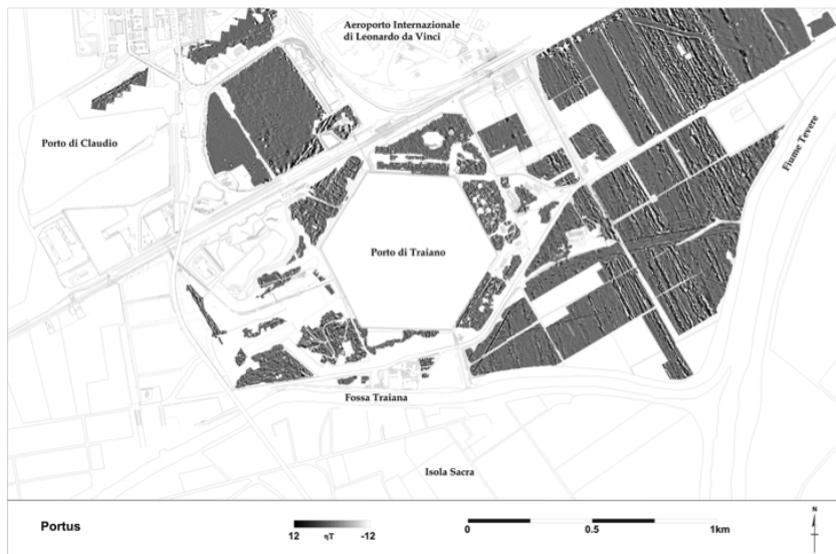


Figure 2. Results from extensive magnetometry survey of Portus and surrounding region.

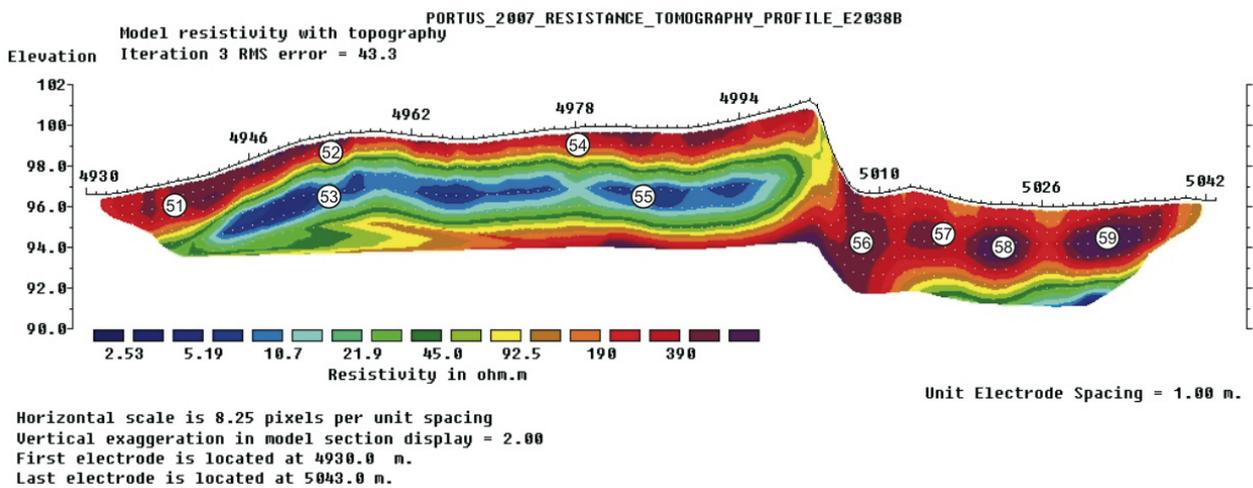
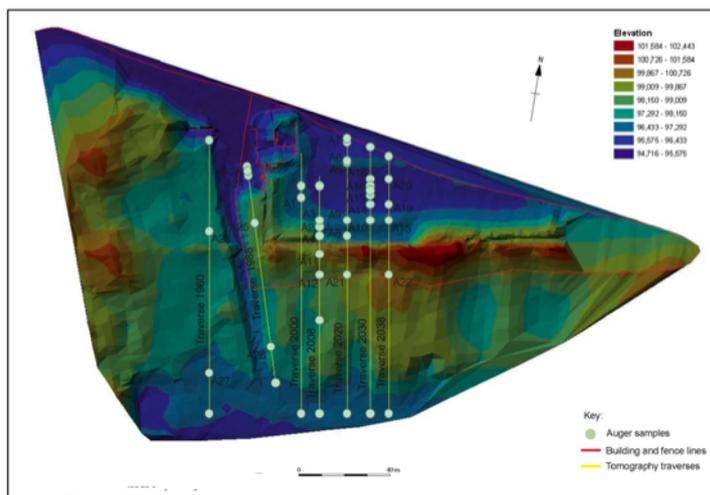


Figure 3. Location and example profile from 2007 Electrical Resistivity Tomography (ERT) survey.

For the resistance survey, a Geoscan RM15 Resistance Meter was used to survey with a 0.5m probe separation on a 30m x 30m grid. A multiplexer with twin probe array was used to survey two 0.5m transects simultaneously, doubling the rate of data collection. The resistance data sets were processed using Geoplot 3.0, and were exported to the GIS for integration with other data. The results of the resistance survey seemed to indicate north-south linear anomalies, as well as extensive disturbances from collapse, vegetation, and potential “wall-robbing” which may have taken place in the north-eastern portion of the “Palazzo Imperiale” (see fig. 4). Extremely high resistance readings in this area may have also indicated the presence of air pockets, or voids between first storey rooms which may still be preserved intact beneath the surface.

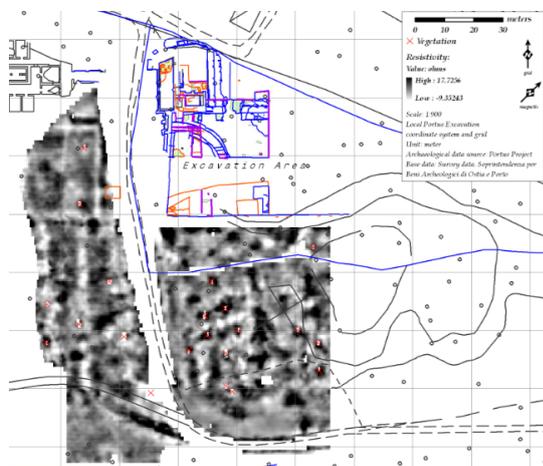


Figure 4. Results from resistance survey.

The GPR survey was completed using a Sensors and Software 500 MHz antenna configured with a Noggin SmartCart. The radar antenna has an estimated ground penetration of 3.5m. Zigzag traverses were collected at 0.5m traverses, with traces of data collected at 0.025m intervals at 512 samples per scan, with a setting of 4 stacks. All GPR data was processed using GPR-slice, before being exported for integration in the GIS. Fifteen timeslices at twenty-five centimeter intervals were created at varying subsurface depths and geo-referenced to the site grid. An “overlay” grid containing high amplitude reflections of interest was created and used as the input for the data integration (see fig. 5). The GPR results presented a number of challenges for interpretation and digitization. With the presence of a considerable amount of near surface rubble, as observed in the magnetometer data, it was often difficult to differentiate collapse and random noise from intact archaeological features. An integrated approach to interpretation was essential to untangling and calibrating the high amplitude responses.

Hand auger samples were taken in conjunction with the resistance tomography survey in May and June 2007 to determine the depth of overburden in preparation for the

excavation survey that year. Twenty-eight auger samples were taken along the ERT profiles to depths up to 3m, with a concentration of samples located within the excavation area.<sup>1</sup> In September 2008, nine mechanical augers were conducted up to depths between 10–13m throughout the excavation area and the adjacent archaeological park.<sup>2</sup> The results of the augering have given an additional mechanism for “ground-truthing” and the verification of the geophysical signatures of features of interest within the port complex.

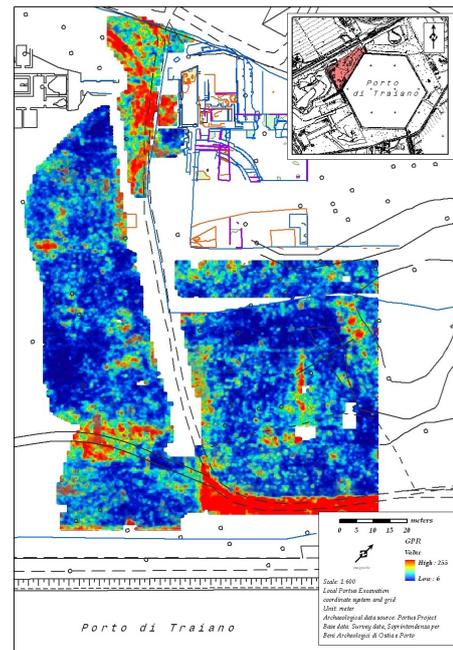


Figure 5. Overlay image of high amplitude reflections in GPR results.

#### 4 MECHANISMS FOR INTEGRATED DATA ANALYSIS

An integrated approach to data analysis has been applied in a North American historic archaeological context by Kvamme,<sup>3</sup> and in classical Roman archaeology in Austria and Italy by Neubauer et al.<sup>4</sup> and

<sup>1</sup>Elizabeth De Gaetano and Kristian Strutt, *Report on the Geophysical Survey at Portus* (p. 274 n3).

<sup>2</sup>In collaboration with Jean-Philippe Goiran and Ferreol Salomon (Université de Lyon) and the Portus Project.

<sup>3</sup>Kenneth L. Kvamme, “Geophysical Surveys as Landscape Archaeology,” *American Antiquity* 68 (3) (2003): 435–456; Kenneth L. Kvamme, “Integrating Multidimensional Geophysical Data,” *Archaeological Prospection* 13 (2006): 57–72; Kenneth L. Kvamme, “Data Processing and Presentation,” in *Remote Sensing in Archaeology: An Explicitly North American Perspective*, ed. by Jay K. Johnson (Tuscaloosa: University of Alabama Press, 2006) 235–50.

<sup>4</sup>Wolfgang Neubauer and Alois Eder-Hinterleitner,

and Piro et al.,<sup>1</sup> respectively. In each case different approaches were applied (sometimes using multiple analysis techniques) to extract the maximum level of interpretation and analysis from the archaeo-geophysical record. As outlined in Kvamme's 2006 publication,<sup>2</sup> a range of data analysis methods exists for the integration of geophysical survey results. The various categories for analysis outlined in that publication were used as a baseline for the data integration methodologies used within this research.

The researcher must first "establish the hypothesis that each geophysical method investigates one event, i.e. the presence of anomalous volumes underground," to allow for the quantification and integration of each set of geophysical results.<sup>3</sup> Integrated geophysical data analysis allows the geophysicist to establish interrelationships and patterns between multi-dimensional data sets, and therefore improve the identification and interpretation of subsurface anomalies, that may otherwise go unnoticed.<sup>4</sup> As demonstrated in recent publications, the integration of geophysical survey results allows the geophysicist to "better define position, extension, depth, thickness, and physical characteristics of any anomalous body within its geological context."<sup>5</sup>

Several types of data integration were produced as part of this research, all of which can be divided into three categories: Graphical, Discrete, and Continuous data sets. All types of integration used within this research were performed within ArcGIS and ERDAS IMAGINE.

#### 4.1 GRAPHICAL INTEGRATION

Integration using graphical overlays and composite images is a simple and easy mechanism for viewing separate geophysical data sets together in their spatial context. These techniques are often used in archaeo-geophysics as a way of visualizing and interpreting separate data sets, but are often overlooked as a means for data integration.

#### Overlays

Several two dimensional overlays were created to visualize the geophysical anomalies within each data set. Contour lines were generated for both the magnetometry and the resistance data and overlaid on the relative data sets. This mechanism is particularly helpful in discerning sharp differences in geophysical signatures, and clearly defines linear archaeological features. To best represent each data set, a variety of resolutions and intervals were utilized.

Overlaying one to two data sets with different transparencies on top of an opaque data set produced an additional mechanism for visualizing multiple methods. One criticism of this technique is that the overlays often produce a "muddy" effect, masking the viewers ability to make out which features relate to which geophysical survey method.

Although the production of such visualizations is not grounded in any particular theoretical approach, the results proved helpful in emphasizing and visualizing positive and negative anomalies within each data set. It was, however, easy to become confused with too many color combinations (see fig. 6a).

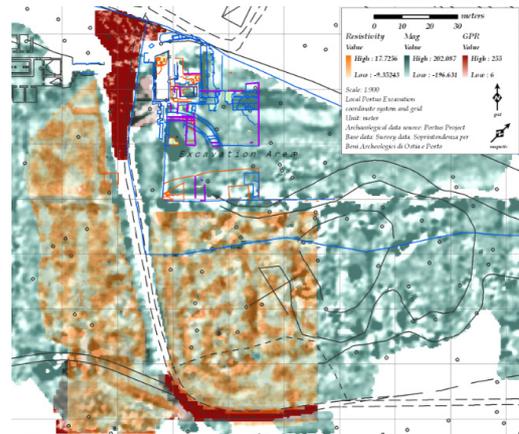


Figure 6a. Example of transparent overlays of magnetometry, resistance, and GPR datasets.

#### RGB Color Composite

The three normalized data sets (magnetometry, resistivity, and GPR) were assigned to each of the three bands, red, green, and blue, respectively. This model was the first multi-banded raster created from the geophysical data, and provided a simple and easy format for manipulating and visualizing the different survey results. Though this particular combination emphasizes positive features, manipulating and inverting the band assignments can achieve a number of color combinations, therefore emphasizing different types of positive and negative features (see fig. 6b).

The RGB model was potentially the most effective in utilizing all aspects of each geophysical data set, and integrating them in a meaningful way. Through

"Resistivity and Magnetics of the Roman Town Carnuntum, Austria: An Example of Combined Interpretation of Prospection Data," *Archaeological Prospection* 4 (1997): 179–189; Wolfgang Neubauer et al., "Georadar in the Roman Civil Town Carnuntum, Austria: An Approach for Archaeological Interpretation of GPR Data," *Archaeological Prospection* 9 (2002): 135–156.

<sup>1</sup>Salvatore Piro et al., "Quantitative Integration of Geophysical Methods for Archaeological Prospection," *Archaeological Prospection* 7 (2000): 203–213.

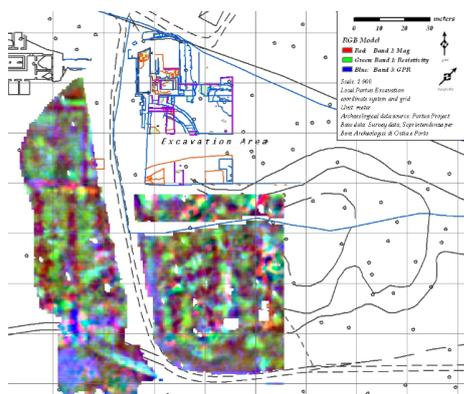
<sup>2</sup>Kenneth L. Kvamme, "Integrating" (p. 276n3): 57–72.

<sup>3</sup>Salvatore Piro et al. (p. 277n1).

<sup>4</sup>Kenneth L. Kvamme, "Integrating" (p. 276n3).

<sup>5</sup>Salvatore Piro et al. (p. 277n1).

manipulation of the band assignments, the RGB image proved to be a simple mechanism for interpreting the positive and negative features, particularly in the area to the west of the modern access path, where it was challenging to assess the precise feature boundaries using the two dimensional overlays. The RGB composite emphasized robust features which were observed in all methods. It also allowed for the visualization of more subtle features that might have otherwise gone undetected. The RGB composite was the most effective data set produced in this analysis, and given its theoretical grounding within remote sensing techniques, there is much potential for further exploration of this data set.



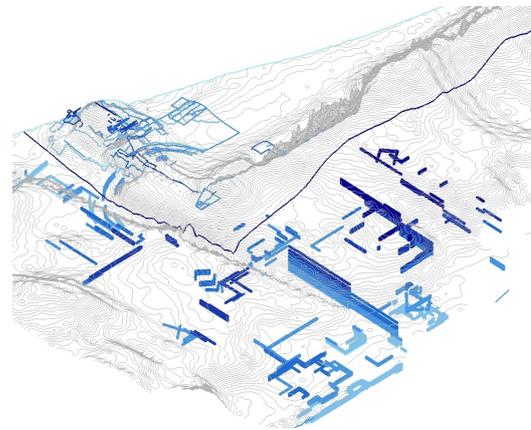
**Figure 6b.** RGB composite of magnetometry (Red-Band 2), resistance (Green-Band 1), and GPR (Blue-Band 3) datasets.

### 3D Vector Integration

The limitations of two dimensional platforms such as Geographic Information Systems sometimes prevent the true integration of three-dimensional data volumes such as GPR and resistance tomography data sets. Consequently, a simple method was developed for viewing the GPR vector data in three dimensions, using ArcGIS and basic feature class editing tools (see fig. 7). This method involved first extracting the surface elevations of each feature from the digital elevation model (DEM) produced from a micro-topographic survey of the site. Subsurface elevations were then approximated for each feature, based upon the velocity calculations achieved during data processing and subtracted from the surface elevation. The new “z-enabled” shapefile was then added to ArcScene, and plotted using the subsurface (z) value for integration analysis with the detailed micro-topographic, excavation, and standing building survey of the site.

The benefits of visualizing the 3D GPR shapes within their subsurface locations in relation to the excavation and topographic data are apparent, although the greatest strength of this method may be in its potential for a platform which also facilitates interactive querying of the results. The selection and display of only features at corresponding depths of “key horizons” at Portus could potentially allow for a clearer integration of survey methods, as well as a clearer understanding and

interpretation of the chronological sequence of structures in this area.



**Figure 7.** Three-dimensional GPR interpretation vector data with excavation features, color coded according to subsurface elevation.

## 4.2 DISCRETE DATA ANALYSIS

Data is said to be discrete if the data values are distinct, separate, and can be categorized.<sup>1</sup> Dividing data into discrete classes with definitive boundaries has the theoretical advantage of removing ambiguity about the location and nature of geophysical anomalies.<sup>2</sup> In this analysis, discrete data formed the input and the output for the operations described in this section.

### Binary Data Analysis

Binary data was generated for each geophysical data set for use as data inputs for the Boolean and Binary Sum calculations. The reclassification values were obtained through examination of known anomaly data ranges before generating value ranges which were representative of the presence (1) and absence (0) of archaeological features. A variety of logical, or Boolean operations, and simple arithmetic operators was performed to analyze the geophysical data. In general, Boolean operators result in grids with cells coded as either TRUE (1) or FALSE (0).<sup>3</sup> Boolean operators are “a class of operations that use Boolean logic to define a selection through the actions of union, intersection, difference, and exclusion.”<sup>4</sup>

A Boolean Union (Boolean OR) is said to be True when

<sup>1</sup>[www.isixsigma.com/dictionary/Discrete\\_Data-226.htm](http://www.isixsigma.com/dictionary/Discrete_Data-226.htm) (seen November 5, 2008).

<sup>2</sup>Kenneth L. Kvamme, “Integrating” (p. 276n3): 63–64.

<sup>3</sup>David Wheatley and Mark Gillings, *Spatial Technology and Archaeology: The Archaeological Applications of GIS* (London: Taylor & Francis Press, 2002) 105.

<sup>4</sup>James Conolly and Mark Lake, *Geographic Information Systems in Archaeology* (Cambridge: Cambridge University Press, 2006) 114.

at least one method detected a geophysical event. The output, due to the overall coverage of the geophysical responses, resulted with a grid with almost 60% of the total cells classified as TRUE. The overall spread of “TRUE” values in the output was extensive, making it difficult to delineate individual features (see fig. 8a).

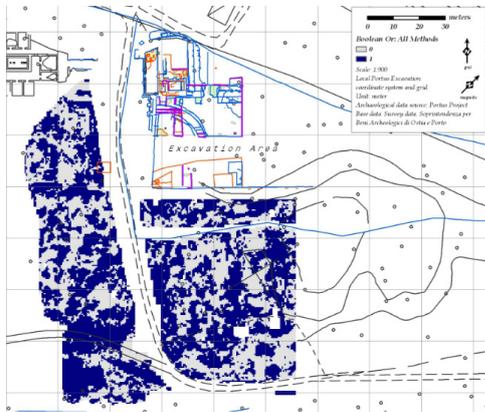


Figure 8a. Results of Boolean OR function, 0 (grey), 1 (blue).

A Boolean Intersection (Boolean AND) occurs where the positive values intersect. In this analysis, the output is a raster with cell values of True where all 3 methods detected a geophysical event. The results produced a very small number of TRUE cells, accounting for less than 3% of the total number of cells (see fig. 8b). As one might expect with using only three input data sets that measure different geophysical elements, this function produced a binary output with very limited analysis capabilities. The results convey very little about the nature of the geophysical anomalies, as the only observation that can be made is “presence” or “absence” of positive anomalies in all methods.

A simple Binary Sum was also performed to produce a summation of the values within each binary data set. This essentially produced a “confidence map” of the number of geophysical methods which observed a single “event” or anomaly. The resulting raster image displayed cell values ranging from 0 (no event observed with any method) to 3 (event observed with 3 survey methods); see fig. 8c. This output of the Portus data sets produces an interpretable map, which researchers can use to assert some degree of “objectivity” when making interpretations of anomalies. However, it still only verifies the existence of anomalies detected by ‘x’ methods, leaving the viewer with the task of relating the image back to the original individual results.

The data analyses that used the binary data as input variables (including the Boolean calculations and mathematical functions) produced the weakest output, in terms of the level of meaningful interpretations which could be made from them. The outputs failed to convey any information about the nature of the anomalies, and only indicated presence, absence, and the number of methods which detected an anomaly at a particular spatial location. Caution was exercised while examining

these data outputs, because if three methods observe an anomaly, this does not necessarily indicate a feature of interest, particularly when the classification of the initial thresholds was the result of a subjective, rather than objective, means of choosing the data ranges.



Figure 8b. Results of Boolean AND function.

#### Cluster Analysis

The goal of classification investigations is to discover patterns in groupings of values within a set of data.<sup>1</sup> With this aim in mind, cluster analysis was used as an unsupervised<sup>2</sup> mechanism for establishing natural spectral groupings between each band of geophysical data.<sup>3</sup> For this research an ISODATA algorithm was used, a variant on the commonly used K-means method for unsupervised clustering.<sup>4</sup> As noted by Kvamme, cluster analysis works well with large data sets, and allows the user to define the number of classes anticipated within the resulting data set.<sup>5</sup> This is a “partitioning cluster technique” which divides the group of values, or attributes, into a specified number of clusters as defined by the user.<sup>6</sup> The center of each cluster is initially determined by a random selection of “seeds” and the remaining objects are added to the nearest cluster. As new objects are added to the clusters, the cluster centers are recalculated. After all objects have been assigned to a cluster, the sum of squared

<sup>1</sup>Stephen Shennan, *Quantifying Archaeology* 2nd ed. (Edinburgh: Edinburgh University Press, 1997) 220.

<sup>2</sup>Unsupervised classification is a system of algorithms which examines unknown pixel values, and aggregates them into a user defined number of spectral classes based upon natural groupings or clusters; see Thomas M. Lillesand et al., *Remote Sensing and Image Interpretation*, 6<sup>th</sup> edition (Hoboken, NJ: John Wiley and Sons, 2008) 568.

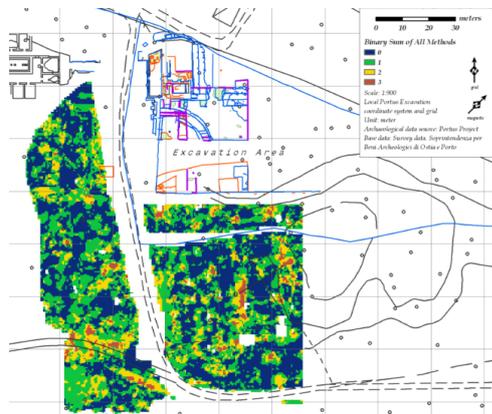
<sup>3</sup>Kenneth L. Kvamme (p. 276n3): 66.

<sup>4</sup>Thomas M. Lillesand et al. (p. 279n2) 570–573.

<sup>5</sup> Kvamme (p. 276n3): 66.

<sup>6</sup>James Conolly and Mark Lake, *Geographic Information Systems in Archaeology* (Cambridge: Cambridge University Press, 2006) 171.

distances (the distance between the object and the cluster center) are calculated and provided for user assessment of the cluster allocation.<sup>125</sup>



**Figure 8c.** Results of Binary Sum of magnetometry, resistance, and GPR data.

Clusters were created using the normalized magnetometry, resistance, and GPR data sets. Three classes were specified, presuming the location of positive, negative, and background events within the 3 bands of data. This function produced a signature file outlining the layers (each band of data input), mean vectors (the average spectral value in each layer), and covariances (the tendency for values to vary similarly in two bands)<sup>1</sup> of the data.

Next, the clusters were used to classify the remainder of the geophysical data within each raster. In this analysis, the Maximum Likelihood Classifier was used to produce a statistical probability that a specified pixel value belonged to a discrete cluster or class.<sup>2</sup> Each class or cluster was given equal weight, and a confidence raster of the classification certainty, in addition to the maximum likelihood classification, was outputted.

After evaluating the probability of each pixel occurring within each class, the pixel was assigned to the class with the highest probability, given its attribute values.<sup>3</sup> This grid file was then filtered using a majority filter to smooth the output and accentuate the dominant classification.<sup>4</sup> Due to the nature of the geophysical data, the appropriate or optimal number of classes to assign the cluster analysis may not be known.<sup>5</sup> Consequently, the output for the maximum likelihood classification was 3 rasters classified into 2, 3, and 4

classes respectively (see fig. 9). The first cluster analysis was performed using a setting of 2 classes, intended to represent anomaly “presence” or “absence.” The filtered output produced a classification that corresponded to interpreted high amplitude GPR features and, to a lesser extent, positive magnetic and resistance features (2), while class (1) corresponded to negative anomalies and “background data.” The cluster analysis was then performed with a setting of 3 classes, representing positive, negative, and background data. The three-class analysis produced a classification that corresponded to more “robust” positive features (i.e. features which were detected by 2–3 methods) (3), positive magnetic features (2) that do not correspond to anomalies detected by other methods, and negative features with background data as (1). Lastly, the cluster analysis was performed using 4 classes, as an attempt to successfully extract and classify the negative features from the background data. The four-class analysis again created a classification corresponding to the robust features detected by all methods (4), with classes (3) and (2) corresponding to progressively more subtle positive features and (1) corresponding to negative features and background data.

#### 4.3 CONTINUOUS DATA ANALYSIS

The previous sections dealt with the classification of discrete and continuous data with the aim of producing defined classes which combined and integrated each of the geophysical data sets. Continuous data is “information that can be measured on a continuum, or scale.”<sup>6</sup> Unlike discrete data, continuous data can be broken down into smaller increments and can represent any number between the minimum and maximum values within the data set. “Continuous data are naturally richer than categorized information, potentially enabling superior data integrations.”<sup>7</sup> In this case, the continuous data input is the real number, normalized measurements from the geophysical survey results.

##### *Data Sum, Product, Max, and Min*

A variety of functions were performed using basic map algebra on the three standardized geophysical data sets. These mathematical functions involved adding and multiplying the cell values of each raster together to produce a raster output containing the new values. These functions should theoretically emphasize existing anomalies, particularly those closer to 1. Different sum combinations were made, which seemingly emphasized

<sup>1</sup>Thomas Lillesand et al (p. 279n2) 570–573.

<sup>2</sup>Thomas Lillesand et al. (p. 279n2) 554–555.

<sup>3</sup>Thomas Lillesand et al (p. 279n2) *ibid*.

<sup>4</sup>Thomas Lillesand et al. (p. 279n2) 580.

<sup>5</sup>Kvamme, “Integrating” (p. 276n3) 66.

<sup>6</sup>[www.isixsigma.com/dictionary/Discrete\\_Data-226.htm](http://www.isixsigma.com/dictionary/Discrete_Data-226.htm) (accessed November 5, 2008).

<sup>7</sup>Kenneth L. Kvamme, “Integrating” (p. 276n3) 66.

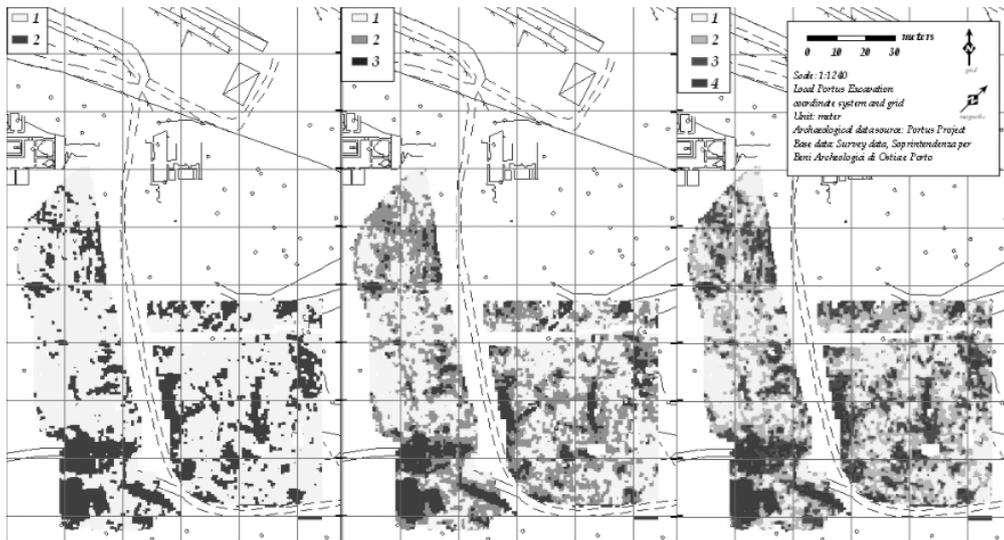


Figure 9. Results from cluster analysis using 2, 3, and 4 classes.

positive and negative anomalies, making the boundaries of some more definitive than others.

As one might expect, the Data Sum output emphasized robust anomalies, yet also included more subtle positive anomalies that were not particularly apparent in the previous data outputs. In addition, there seemed to be an absence of a strong correlation between negative features in all 3 data sets. The Data Product was particularly useful for emphasizing and exaggerating robust anomaly boundaries, and masking subtle ones (see fig. 10a).

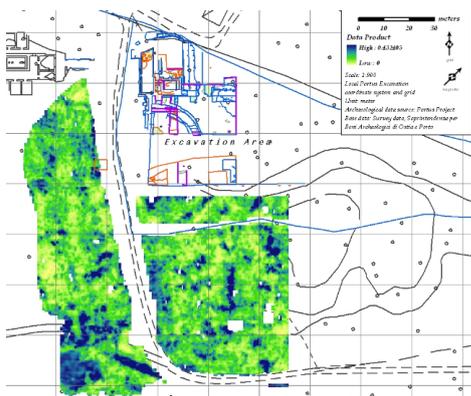


Figure 10a. Results of Data Product of magnetometry, resistance, and GPR data sets.

The “Maximum” and “Minimum” values were also calculated to create raster outputs containing the maximum and minimum cell values contained in each input geophysical data set. The resulting MAX grid emphasized the positive features in each survey method, including potential structural remains and near surface

rubble (see fig. 10b). The MIN grid seemed to correspond to “negative” anomalies within each data set, including proposed “voids” between structural remains. This is one of the first functions performed on the data that has resulted in an output which has examined the negative anomalies within the geophysical data sets.

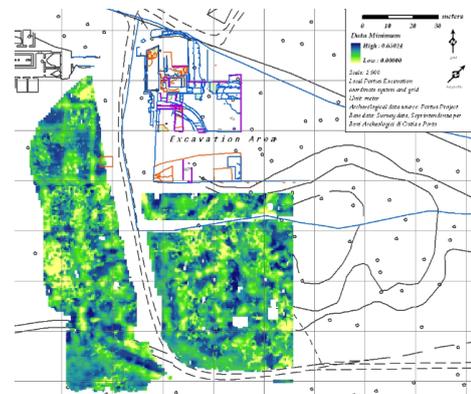


Figure 10b. Results of Data MIN of magnetometry, resistance, and GPR data sets.

### Principal Components Analysis

In essence, Principal Components Analysis (PCA) is “designed to reduce redundancy in multispectral data.”<sup>1</sup> As one might expect, input variables must be highly correlated for there to be a significant reduction in redundancy.<sup>2</sup> The closer the original variables are

<sup>1</sup>Lillesand et al. (p. 279n2) 527.

<sup>2</sup>Stephen Shennan, *Quantifying Archaeology*, 2nd ed. (Edinburgh: Edinburgh University Press, 1997) 269–70.

correlated, the more meaningful the new bands of data will be, and thus the more information one can retrieve from the reclassification. One might suspect that the use of PCA in the context of geophysical prospection is theoretically applicable, particularly in cases where survey methods are highly correlated (whether positively or negatively), such as the correlation between electrical resistivity and electrical conductivity.<sup>1</sup>

The PCA was performed using the normalized results for each geophysical method as input: resistivity, magnetometry, and GPR. The correlation coefficients were plotted on a scale from -1 to +1, where -1 equaled a negative correlation, +1 equaled a positive correlation, and 0 equaled the absence of correlation.<sup>2</sup> However, as with Kvamme's analysis at Army City, the overall correlation between the input data variables, or Pearson correlation coefficient,  $r$  remains relatively low with the highest value at 0.2135. The applicability of the Portus geophysical results in this type of analysis is questioned, as an examination of the scatter plots of each method does not indicate extensive overlap between the normalized values. As a result, the first principal component contains minimal contrast, and the second and third components are the input variables, resistivity and magnetometry respectively (see fig. 11).

## 5 CRITICISMS CONCERNING CLASSIFICATION

“The quality of the training process determines the success of the classification stage, and therefore, the value of the information generated from the entire classification effort.”<sup>3</sup> Clearly, Lillesand et al. are referring to the act of determining training data sets for use in a supervised classification. Nevertheless, the same point may be made about the selection of anomaly thresholds for the binary data classification. These thresholds, though based on a cautious examination of the range of anomaly values within each data set, were a subjective selection of values based on inductive reasoning and knowledge of the results. The “goodness of fit” of the chosen anomaly ranges will never be determined unless extensive ground truthing of every anomaly takes place, which in turn, defeats the purpose of the non-invasive, inductive nature of geophysical prospection.

Both multivariate classifications performed in this research (cluster analysis and principal components analysis) are unsupervised and result from algorithms which “examine the unknown pixels in an image and aggregate them into a number of classes based on

natural groupings.”<sup>4</sup>

One criticism of this method, though clearly useful for recognizing patterns which may not be readily apparent in a data set, is that the output of such classifications may emphasize or understate relationships between data values that may not be useful for their applications in the relative research. In contrast, supervised classifications require the user to define “useful information categories”<sup>5</sup> to be compared to the spectral signatures of other cells within the data set. Where in unsupervised approaches results should be compared and contrasted with real data distributions, supervised classes allow for the immediate association of results based on initial training categories. However, a critique of supervised classifications may be made of the inherent bias ingrained within the data output, as defined by the training process. In the end, it is clearly ideal to utilize both strategies for determining patterns in one's data, as both classification types act as complementary analysis techniques, where the limitations of one are compensated by the strengths of the other.

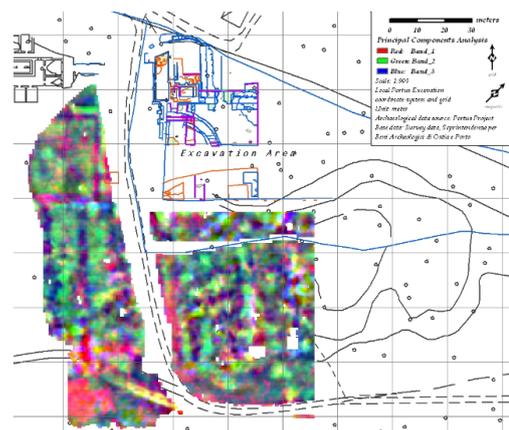


Figure 11. Composite of principle components analysis.

## 6 CONCLUSIONS AND FUTURE PROSPECTS

A major distinction between recent examples of geophysical data integration in archaeology<sup>6</sup> and the analyses completed for this research is the difference in the level of assumptions that can be made about the results. With recent historical archaeological sites such as Army City,<sup>7</sup> researchers have the benefit of historic

<sup>1</sup>Kenneth L. Kvamme, “Integrating” (p. 276n3) 68.

<sup>2</sup>Lillesand et al. (p. 279n2) 557.

<sup>3</sup>Lillesand et al., *ibid.*

<sup>4</sup>Lillesand et al. (p. 279n2) 569.

<sup>5</sup>Lillesand et al. (p. 279n2) 557.

<sup>6</sup>Kenneth L. Kvamme, “Integrating” (p. 276n3) 57–72; Kenneth L. Kvamme, “Data Processing” (p. 276n3) 235–50; Wolfgang Neubauer and Alois Eder-Hinterleitner, “Resistivity and Magnetism” (p. 276n4); Wolfgang Neubauer et al., “Georadar in the Roman Civil Town Carnuntum” (p. 276n4); Salvatore Piro et al. (p. 277n1) 203–213.

<sup>7</sup>Kenneth L. Kvamme, “Geophysical Surveys” (p. 276n3)

records, including plans and photographs, and even oral accounts of the nature of the subsurface features being detected. Though antiquarians have conducted extensive research at Portus for some time, many questions regarding the chronological sequence of the port, as well as its relationship with other ports in Italy and elsewhere are still under debate, not least in the context of the Portus Project. Establishing a chronological sequence and overall plan of the structures, including the “Palazzo Imperiale,” the “sub-circular wall,” and the “warehouses” have proved to be challenging, and a continuous forum for archaeological dialogue. Though the geophysical results have made a substantial contribution to the discussion about the nature of the structures at Portus, a certain level of uncertainty still remains about the nature of the anomalies. Much of this may be attributed to the state of remains within the area in question. As stated previously, the surface of the portion of the site being investigated here has been obscured by demolition and collapse, making interpretations of the geophysical anomalies difficult. The prospect of determining “four types of floors”<sup>1</sup> remains unlikely for the foreseeable future. However, in this case a successful data fusion is not judged on the basis of one’s ability to discern the minute details of archaeological features; those are merely by-products of a series of optimal conditions which allow for exciting, innovative finds. Here, the authors have chosen to focus on the mere creation of a type of data fusion that champions exploratory data analysis and emphasizes positive and negative correlation of feature existence.

A potential limitation of the more sophisticated methods of cluster analysis and principal components analysis techniques may be the number of input variables required to create a meaningful output. A second season of intensive resistance tomography was conducted at Portus in February and May 2009, which may provide an additional three dimensional data input for future data fusion research. With the addition of a fourth data input, additional analyses may be performed, including supervised classifications using training data sets derived from the excavation data recovery.

Future prospects for the use of data fusion in general certainly include the incorporation of the third dimension in data analysis techniques. The three dimensional vector data created for the GPR data provides an accessible interface for visualizing and interpreting the relationships between the GPR results and the excavation data. The addition of the resistance tomography data, as well as models of the standing building survey, will greatly increase the researcher’s

ability to correlate and interpret the features of interest based upon their elevations. Though elsewhere alternative software has also been used,<sup>2</sup> such as Amira,<sup>3</sup> to visualize three dimensional geophysical data sets in their context, the strength of the 3D vector data created for this research lies in its simplicity. This shapefile can be imported and exported to any 3D viewer or drawing package for interpretation, whereas using expensive proprietary software often limits the full realization of the data’s potential.

New data fusion software is in production which imports, processes, analyzes, and essentially fuses geophysical data within a single user interface.<sup>4</sup> In addition, recent success with visualizing topographically corrected resistance tomography data with GPR volumes in GPR-slice has proven to be a new and exciting potential platform for the integration of three dimensional geophysical data sets.<sup>5</sup> These types of interfaces will not only encourage the increased use of data fusion techniques but will also increase the level of meaningful, progressive research within geophysical prospection, and permit an opportunity for a wider understanding of the archaeology in question.

Though the interpretations of the research conducted here have not yet been fully realized, the types of methodology which were used have provided a much more holistic view of the subsurface anomalies at Portus. The combination of integrated survey methodologies and integrated data analysis has provided a wealth of different types of data, including resources with both analysis and visualization capabilities, increasing the potential for future interpretations of archaeological and geophysical features at Portus. Though each method used in this research contains strengths and weaknesses, of all of the analysis methods used, the RBG model, cluster analysis, and 3D vector exploration have been the most insightful and visually pleasing results of this analysis.

The process of archaeological data integration, in general, is a process that is comprised of multiple phases, including data collection, data analysis, and interpretation. A perpetual cycle of reevaluation is required as new data is gathered, analyzed, or

---

435–56; Kenneth L. Kvamme, “Data Processing and Presentation” (p. 276n3) 235–50.

<sup>1</sup>Neubauer and Eder-Hinterleitner, “Resistivity and Magnetism of the Roman Town Carnuntum, Austria: An Example of Combined Interpretation of Prospection Data,” *Archaeological Prospection* 4 (1997): 185.

---

<sup>2</sup>Meg S. Watters, “Geovisualization: An Example from the Catholme Ceremonial Complex,” *Archaeological Prospection* 13 (2006): 282–290.

<sup>3</sup>Amira is a three dimensional imaging software originally developed for the medical field (Watters [p. 283 n2] 285).

<sup>4</sup>Amira is a three dimensional imaging software originally developed for the medical field (Watters, *ibid.*).

<sup>5</sup>The University of Arkansas Center for Advanced Spatial Technologies: Geophysical Data Analysis Toolkit, [www.cast.uark.edu/home/research/geophysics/geophysical-data-analysis-toolkit.html](http://www.cast.uark.edu/home/research/geophysics/geophysical-data-analysis-toolkit.html).

interpreted, ideally forming a continuous progression towards a better understanding of the archaeology. Portus is no different, in that each phase of research, from classical texts to excavation through to geophysical prospection, is never complete, and as new data sets are acquired, additional groundwork is laid to interpret and reinterpret the history of the port complex.

Despite the limitations of individual methods performed in the integration data analysis, it is strongly believed

that the results of the foregoing methodology have considerably increased the potential for using geophysical prospection as a means for understanding the uncertainties inherent to archaeological and geophysical research. The archaeological interpretations of the integration data analysis has by no means provided a comprehensive list of conclusions, but rather provided the framework for continued discussion, analysis, and interpretation.

#### ACKNOWLEDGEMENTS

The authors would like to acknowledge a major part of the work presented in this paper, which was undertaken by Jessica Ogden as part of her MSc in Archaeological Computing at the University of Southampton. In addition, as portions of the data used in this research were collected over several years and as part of multiple projects, all of those involved with the geophysical data collection and processing on the Roman Towns Project, as well as the Portus Project, including Rose Ferraby, Leonie Pett, Elizabeth De Gaetano, Giles Richardson, and Gregory Tucker are gratefully acknowledged. The continued support of the Soprintendenza per i Beni Archeologici di Ostia, as well as the various institutions involved with the Portus Project including the University of Southampton and the British School at Rome, has facilitated this ongoing research. Dean Goodman is also gratefully acknowledged for his continued involvement with the integration of resistance tomography data in GPRslice and for his support in the processing and visualization of the radar and tomography data sets for this project.

#### BIBLIOGRAPHY

- Conolly, James, and Mark Lake. *Geographic Information Systems in Archaeology*. Cambridge: Cambridge University Press, 2006.
- De Gaetano, Elizabeth and Kristian Strutt. *Report on the Geophysical Survey at Portus May-June 2007*. Archaeological Prospection Services of Southampton and the British School at Rome. Unpublished report, 2007.
- Keay, Simon, Martin Millett, Lidia Paroli, and Kristian Strutt. *Portus: An Archaeological Survey of the Port of Imperial Rome*. London: The British School at Rome, 2005.
- Keay, Simon, Graeme Earl, Sophie Hay, Stephen Kay, Jessica Ogden, and Kristian Strutt. "The Potential of Archaeological Geophysics, The Work of the British School at Rome in Italy." *Geofisica per l'archeologia: Possibilita e Limiti: Roma 10 Dicembre 2008. CISTEC* (2008): 25–44.
- Kvamme, Kenneth L. "Geophysical Surveys as Landscape Archaeology." *American Antiquity* 68 (3) (2003): 435–456.
- Kvamme, Kenneth L. "Integrating Multidimensional Geophysical Data." *Archaeological Prospection* 13 (2006): 57–72.
- Kvamme, Kenneth L. "Data Processing and Presentation," in *Remote Sensing in Archaeology: An Explicitly North American Perspective*, edited by Jay K. Johnson, 235–250. Tuscaloosa: University of Alabama Press, 2006.
- Lillesand, Thomas M., Ralph W. Kiefer, and Jonathan W. Chipman. *Remote Sensing and Image Interpretation*, 6<sup>th</sup> edition. Hoboken, NJ: John Wiley & Sons, 2008.
- Neubauer, Wolfgang, and Alois Eder-Hinterleitner. "Resistivity and Magnetism of the Roman Town Carnuntum, Austria." *Archaeological Prospection* 4 (1997): 179–189.
- Neubauer, Wolfgang, Alois Eder-Hinterleitner, Sirri Seren, and Peter Melichar. "Georadar in the Roman Civil Town Carnuntum, Austria." *Archaeological Prospection* 9 (2002): 135–156.
- Piro, Salvatore, Paolo Mauriello, and Fabio Cammarano. "Quantitative Integration of Geophysical Methods for Archaeological Prospection." *Archaeological Prospection* 7 (2000): 203–213.
- Shennan, Stephen. *Quantifying Archaeology*. 2nd ed. Edinburgh: Edinburgh University Press, 1997.
- Meg S. Watters. "Geovisualization: An Example from the Catholme Ceremonial Complex." *Archaeological Prospection* 13 (2006): 282–290.
- Wheatley, David, and Mark Gillings. *Spatial Technology and Archaeology: The Archaeological Applications of GIS*. London: Taylor & Francis Press, 2002.