A spatial study of aspects of the Roman settlement of Spain through the use of GIS

by

Federica Ada Nelda Marina Massagrande

a thesis submitted for the degree of

Doctor of Philosophy

October 1995

in the

Institute of Archaeology

University College London

Abstract

This thesis explores the use of computer, and in particular GIS (Geographic Information Systems), techniques to study the development of the Roman rural settlement in three areas of Spain: the Guadalquivir Valley, the region of Tarragona and the Maresme. The project has two aims, the first consists of finding a viable way to study the archaeological data from non-systematic surveys, which have been collected over a period of years and stored in archaeological units. The second aim of the project is to assess to what extent GIS can be used to perform such a sort of archaeological study.

Firstly, the data used in the study are presented and the way in which they were stored and manipulated in digital format is described and discussed. The computer hardware and software used are described. Some of the problems encountered with this approach are highlighted and the solutions implemented are presented.

Secondly, the analysis on the archaeological data about the Roman rural sites in the three study areas, carried out using GIS and statistical software, is presented in detail. For each area the shape of rural settlement pattern at different periods is investigated in relation to environmental and socio-economic factors. Multivariate statistical techniques are used to study the pottery assemblage of the rural sites. The information derived from the analysis is then used to create models of the development of each area.

In the concluding chapter the resulting models are compared and the differences and similarities in the development of the three areas observed. The advantages and disadvantages of using GIS with archaeological non-systematic survey data are discussed and assessed. Finally, a series of considerations on the main differences between the approach described in this thesis and earlier comprehensive works dealing with data from several separate areas are presented.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my supervisor, Mr Clive Robert Orton of the Institute of Archaeology, University College London, and my advisor, Dr Simon J. Keay of the Department of Archaeology, University of Southampton, for their guidance and support throughout this programme of research. Clive offered me his help on the computing and statistical part of the project and ensured that I did not lose sight of the main scope of the thesis, while Simon shared his knowledge of Roman Spain and most of the content of his bookshelves with me.

I would also like to thank the staff of the Department of Archaeology, University of Southampton, for allowing me to use the departmental equipment, and the staff at the Archaeological Unit of Seville and at the *Generalitat de Catalunya* in Barcelona, for giving me access to their data.

A special thank goes to my parents for their support and their contribution where SERC (Science and Engineering Research Council) and then NERC (Natural Environment Research Council) could not reach.

Several other people helped and supported me along the way. In particular, I would like to thank Dr Matthew J. Carter for putting up with me in the final stages of the project, Anja-C. Wolle and Maria J. Fernández Fonseca for sharing an office with me without ever losing their temper, Timothy J. T. Sly for looking after the computer equipment, often at the risk of his own mental health, Kathryn Knowles for keeping my spirit high with her email messages, John G. Glenn and Peter Durham for being often ready to go down the pub, Kris Lockyear for his (occasionally constructive) criticism of GIS and Dr David Wheatley for his advice on several GIS issues. I would also like to thank Dr César Carreras Monfort and Leonardo Garcia Sanjuan for helping me with accommodation during my permanence in Spain to gather data. My apologies go to all those who deserve to be mentioned here, but whose names presently escape my mind.

Chapter 1

Introduction

This research attempts to use modern computer techniques to study the distribution of Roman sites in separate areas of the Mediterranean region and to compare the development of the settlement patterns in the different regions. The research was started with two main objectives:

1 To investigate to what extent GIS could be used to perform landscape analysis of archaeological settlement and in what way computer techniques could help to handle and analyse non-systematically collected archaeological data.

2 To investigate the development of Roman settlement patterns in parts of Spain and try to relate the observed variation to social, political and economic transformations taking place between the 2nd century BC and the 7th century AD.

This chapter describes what systematic and non-systematic archaeological field surveys are, then it provides a brief introduction to computing in archaeology and, in particular, GIS. More advanced functions of GIS are briefly presented where their application to the available data is discussed in later chapters.

1 Archaeological surveys

Excavation is necessarily a destructive, expensive and highly time consuming means of conducting archaeological research, therefore quicker, cheaper and non-destructive methods of investigation are used when it is possible to do so. The archaeological map of a region usually consists of the information that can be gathered by activities such as field walking and geophysical survey. Often, a field survey is conducted in an area to determine where the most interesting features are located, before excavation is planned and carried out. Surveying is not simply a complementary technique to excavation, it is possible that surveys are carried out without an excavation following, when the main focus of the research is the investigation of the past use of a particular territory and its natural resources, or when just a map of possible site locations is required. The two basic kinds of archaeological surface survey are the systematic survey and the non-systematic survey. Surveying in the Roman Mediterranean is not as widespread as it is in other branches of archaeology (such as prehistory), because the Roman remains in the Mediterranean region are monumental in scale while, where the archaeological record is more limited, the archaeologists tend to concentrate more on methods to optimise results from the data (Keay and Millett 1991, 129). The adoption of surveying in archaeology lead to the question being asked of what a site actually is and how surface scatters can be meaningfully interpreted and related to ancient societies (Haselgrove *et al.* 1985). In the last 15 years several surveys have been carried out in the Mediterranean (Macready and Thompson 1985; Keller and Rupp 1983), partially redressing the balance between Roman monumental archaeology and landscape archaeology. Various papers in Schofield (1991) deal with the problems of surface survey method and theory, landscape processes and the interpretation of survey data.

1.1 Systematic surveys

Systematic surveys are usually planned in advance and carried out over a number of seasons. Usually there is a well defined aim of the survey and one particular aspect of the region is investigated as far as it is possible. Systematic surveys are carried out by dividing the landscape into a number of fixed units and by studying these in detail by means of intensive field walking. Usually, when a region is surveyed, effort is put into trying to cover as many of the different landscape characteristics as possible. If the study area consists of varied self-excluding landscape zones such as valley bottoms and mountain sides, flat plains and hills, the survey will be designed to cover all the different geo-environmental zones. Systematic surveys can occasionally cover the same area in following seasons to counteract the problems arising from the different visibility of archaeological surface material in different seasons and years.

There are two basic ways to subdivide the landscape into sections to carry out a systematic survey: the *transect* and the *quadrat*. The transect is more appropriate for large-scale surveys as it cuts across landscapes and allows samples to be taken from all the geo-environmental zones present in the study area. Rectilinear paths are selected in the region under investigation and these are intensively field walked recording the position of archaeological find, while the area outsides the transects is largely ignored. The percentage of the study area to be surveyed depends on the number, spacing and size of the transects. The quadrat is more appropriate for smaller scale surveys. This method involves dividing the study area into squares and then studying in detail some of them. Which squares to investigate is usually chosen by means of different sampling strategies. The percentage of the total area surveyed depends on the size of the quadrats and on the number of these which are studied. The classic example of the application of the quadrat is the systematic survey of the Formative-period hamlet of Tierras Largas in the Oaxaca highlands in Mexico (Winter 1976), while Flannery (1976a) gives an example of the use of transect in the same area.

It is possible to combine the two methods and use transects to study large parts of the landscape and then use quadrats to study the areas with particularly high scatters of archaeological material revealed by field walking the transects (Renfrew and Bahn 1991, 62-69). The relative efficiency of the sampling techniques for archaeological surveys is discussed by Plog (1976; 1978) and Flannery (1976c).

In recent years systematic surveys have started being employed in the Mediterranean region. The Ager Tarraconensis region has been systematically surveyed since 1985. In this area the sampling strategy involved walking intensively a number of transects 1 km wide and 5 km apart collecting all the surface pottery and other types of finds. The transects were cutting the territory from west to east, roughly following the coast. The whole area under study covered some 1570 km^2 and the systematic survey examined in detail about 3% of the total (Keay 1991).

Prevosti Monclús (1991) was faced with the problem of studying the archaeology of the Maresme when a large part of its territory had undergone development and was therefore no longer available for archaeological research. Instead of imposing pre-defined transects as in the case of the Ager Tarraconensis survey, she selected a band having as borders the river Argentona on one side and the modern town of Mataró on the other. The northern edge of the transect was determined by the presence of hills, and the southern one by the sea. This way the transect stretched from the hills to the sea providing a sample of the whole range of the region's landscape.

1.2 Non-systematic surveys

As opposed to systematic surveys, non-systematic surveys are not always planned in advance and do not usually have a specific aim (*ie* the surveyors would not be trying to answer a specific question, but would rather try and collect whatever they might find). Often, a non-systematic survey is carried out when a particular location known or suspected to contain archaeological material is threatened with destruction by development, intensive cultivation or natural causes (floods, erosion). Non-systematic field walking can be carried out by members of archaeological units trying to create a distribution map of local archaeological remains. In these cases, the main interest of the surveyors is that of retrieving as many archaeological finds as possible in the shortest time, thus often sacrificing the less visible archaeological items. Pottery is the most commonly retrieved material in field surveys and the percentage of the more eye-catching fine wares to the total tends to be higher than in systematic surveys and excavations (Hodder and Orton 1976, 105). As Renfrew and Bahn (1991, 64-68) point out, walkers have an inherent desire to find material, therefore they tend to concentrate on the areas which are seen as being more likely to yield archaeological finds. This procedure biases the sample as the parts of the study areas considered poorer in archaeological material are disregarded. Non-systematic surveys yield every season a large amount of non-quantified archaeological data which are catalogued and often stored in archaeological units. Near to none of these data ever get published, but they are usually available for research to the archaeologists who may need to look at them. Though a certain degree of planning exists when an archaeological unit is trying to cover a certain area, detail examination of selected sections is not carried out, marring the final quality of the assemblage.

2 Field Survey in Spanish Roman archaeology

Most of the survey work carried out in Spain in the last 30 years concentrated in the areas of southern Spain, eastern Spain and the Ebro valley and was mainly aimed at trying to improve existing knowledge of the archaeology of these areas (Keay and Millett 1991, 131). The first large scale mapping of sites in the Lower Guadalquivir Valley was carried out by Bonsor (1931), whose work was then expanded by Collantes de Terán (1939; 1943; 1951; 1955). A major series of surveys in the Guadalquivir Valley were carried out from the 70's onwards by Ponsich (1974; 1979; 1987; 1991), who also developed a site hierarchy based on the density of pottery found at each location. For the first time, the existence of scatters of archaeological material which could not be directly identified with continuous occupation was recognised in Spanish archaeology. More recent work in the same area has been carried out by Amores Carredano (1982), Ruiz Delgado (1985) and Escacena Carrasco and Padilla Monge (1992).

In Catalunya, amateurs have been finding and publishing archaeological sites since the later nineteenth century, building up substantial lists of various classes of ancient monuments (Keay and Millett 1991, 132). The first attempt to create a regional summary of the archaeological evidence was done by Serra Ràfols (1928) and then expanded by Almagro Basch *et al.* (1945). More large scale coverages were produces for the Vallès (Estrada 1955; 1969) and the Penedès by Romeu who, rather than publishing monographs, published a series of individual finds, such as Romeu 1959. Other archaeologists carried out survey work in Catalunya, such as Miret *et al.* 1984; Sanmartí and Santacana 1986; Sanmartí *et al.* 1984; Ribas 1952; Cuyas 1977. In 1981 Prevosti Monclús (1981a; 1981b) produced two large and detailed site lists for the Catalan region of Maresme, while M. Oliva did the same thing for the area of Girona. His work was published after his death by Nolla and Casas (1984), who also attempted an interpretation of the sites after having divided them into chronological bands.

The other area in Spain where survey work has been carried out is the Ebro valley. From 1979 an important project has been carried out in the province of Teruel (Burrillo 1984). More work was carried out in the Ega valley by Oña González (1984), who also took into account the environmental aspects of the area.

With the increase in availability of survey information, general overviews of the whole of the Iberian peninsula were attempted. One such work is the listing of the Roman villae known in Spain by Gorges (1979). He also attempted to study the settlement of the whole of Spain based on his summary, but his study is seriously flawed by the concentration of the surveys in specific areas, so that his distribution maps just reflect the intensity of survey in various areas. Despite the locational bias of the data, Gorges described the distribution of Roman villae chronologically and spatially, even attempting to calculate the density of population in Roman Spain. Another comprehensive study of Roman villae in Spain was carried out by Fernández Castro as a PhD project and published in 1982 (Fernández Castro 1982). Though she used information from all over Spain, she was only interested in the architectural and cultural aspect of the Spanish villae, therefore only used information about well preserved and well documented villae and attempted no spatial analysis on the data. Lewit (1991) used data from Spain as well as from other parts of Europe to study the Roman economic and settlement development across the Mediterranean. She used 201 rural archaeological sites, according to Fulford's definition of rural site (Fulford 1982, 404), from 7 different areas (including north Spain and south Spain), choosing sites which had been well excavated and documented. She did not use statistical methods to study these data as she argued that the addition of the kind of assumptions required for mathematical analysis would only compound the many methodological problems caused by the uncertain interpretation of the data and confuse the untrained reader (Lewit 1991, 24). This work is valuable in trying to assess change in the Roman world on a large scale, but its main drawback is the inclusion of only very well studied sites in the analysis, excluding a large amount of available data. All these authors fail to use statistical methods to study the data and their conclusions are based on the observation of variation in site distribution maps referring to different periods.

3 Introduction to GIS

This section reviews the development of the use of computers in archaeological studies from its very first applications to the modern day. The application of GIS to archaeology is reviewed in detail and a short introduction to the main features and types of GIS is given.

3.1 Computing in archaeology

The first reference to the use of computers in a subject closely related to archaeology is a paper on the use of IBM machines on anthropological data (Griffin 1951) presented at the conference on archaeological methods held in New York in 1950. Computers have been used in archaeology since the 1960's (Hodson *et al.* 1966; Hole & Shaw 1967), when the university mainframes started to be employed by archaeologists for tasks such as seriation and classification (see also Doran 1971). Almost as soon as computers started being used in archaeology, the problem of using archaeological data with computers emerged. The issue of describing archaeological data in an appropriate way for cataloguing in a computer was addressed by Chenhall as early as 1967 (Chenhall 1967), while in 1968 (Chenhall 1968) he discussed the impact of computer techniques on archaeological theory. Other early applications of computing to archaeology came as a `second wave' in the 1970's and were mainly limited to database storage and management (Scholtz and Chenhall 1976) and the word processing of site reports or articles, together with the cataloguing and management of museum data. Wilcock (1971) attempted to produce guidelines for an overall computer-aided system for archaeologists, including information retrieval, graphics, routine reduction of instrument survey reading, objective classification of profiles and statistical data analysis. Ammerman (1971) used an Atlas computer to group geographically assemblages from the Italian epipalaeolithic. The early issues of the Proceedings of the CAA (Computer Applications in Archaeology) conference[1], which has been held annually since 1973, are a good source of information on the early usage of computers in archaeology. The 80's witnessed the spread of computing to almost all archaeological activities and groups, from universities to archaeological units and workgroups, so that

[1] The proceedings from the first CAA conference were published in *Science and Archaeology*. Subsequent proceedings were published by the University of Birmingham Computer Centre and since 1987 they have been published in the British Archaeological Reports, International Series, with the exception of the proceedings of CAA92, which were published in Andresen *et al.* 1993.

by the early 90's Reilly and Rahtz (1992, 1) say that quantitative methods are so much part of the archaeological method that they no longer appear in the literature as distinct research papers. This wide distribution of computing in archaeology is largely due to the appearance of the microcomputer in the early 80's, such as the Commodore Pet, the BBC, the Tandy and the Apple II, followed by the rise of the Personal Computer in the late 80's. The applications used in archaeology on the microcomputers at this stage included computer graphics (Upham 1979) and database management (Gaines 1981), while specialised seminars started being held at major universities (Stewart 1980). In the 90's it is hard to believe than any archaeologist (at least the British and American ones) has not been exposed to computers at some stage of his/her work, even if only for the word processing of articles. Nowadays computers are being used in archaeology almost universally, though it is still true that the majority of archaeologists are happy to know just enough computing to perform their tasks, asking their `expert' colleagues every time something unexpected happens.

The natural development of the application of computer techniques to various archaeological problems since the 60's has brought about the appearance of what can be defined as a new branch of archaeology: Archaeological Computing. This includes not only the `traditional' uses of computing such as database management, word processing and statistical analysis, which are generally accessible to the majority, if not the totality, of archaeologists who wish to use them, but also a few other applications which do require a rather high degree of expertise and specialised hardware, such as Artificial Intelligence (AI) and Expert Systems. Somewhere in between these two extremes lay some applications of computing that do require more than just a basic understanding of computing, but that at the same time can be approached by people whose primary interest lies in archaeology and not computer science. To this group belong applications such as Multimedia and GIS.

3.2 What is GIS?

Geographical Information Systems (GIS) is the name given to a set of programs aimed at storing, transforming, manipulating and analysing spatially distributed data. Generally, a GIS package would consist of a number of independent modules which share the same data structure but which can quite easily run independently from each other, except in the cases where the output of one module is required as input to another. The modules are usually unified under a common front-end to make the data transfer from one to the other easier. Advanced statistical capabilities are still lacking from most GIS packages, which makes it necessary to transfer the data to independent statistical packages or, in extreme cases, write one's own analysis routines. Despite recent conferences (Fotheringham and Rogerson 1994) and workgroups (GISDATA) stressing the point that better analysis facilities are required from standard GIS software, no commercial GIS package is still completely satisfactory in this respect.

GIS was originally developed by geographers for geographers and, though researchers from various disciplines have since added new features, GIS still remains very much a geographer's tool. Historically, GIS evolved from computer assisted cartography when this started to be used for resource assessment and land planning in the 1970s (Rhind 1977, Nagy and Wagle 1979). The development of digital cartography towards integrated geographical software systems was led by the idea that a certain event could be tested in advance on a computer model so that its effects were evaluated before the event was imposed upon the environment. The need for accurate predictions, coupled with the requirements of the military applications, led to the development of more and more complex systems and analysis modules which ultimately evolved into GIS software (Burrough 1986, 4-7). The first `true' integrated GIS was the Canada Geographic Information System (CGIS), which in 1972 became distinct from cartographic systems by having the capacity to overlay two or more coverages for a region and calculate the area of simple or compound coverages (Lock and Harris 1992, 89; Nagy and Wagle 1979, 171).

The very first applications of GIS to archaeology took place in the early 1980s in North America (Brown and Rubin 1982), followed by conferences with specific sections on the subject, in which papers were presented covering both methods and principles (Kvamme 1985; Ferguson 1985) and regional applications (Bailey *et al.* 1985; Creamer 1985). GIS was applied to archaeology later in Britain than it was in the United States and the main input came from geographers rather than archaeologists (Harris and Lock 1990, 35-36; Lock and Harris 1992, 89-90).

There is a good deal of publications about GIS, but, up to date, the best and most complete introduction to GIS principles and techniques is Burrough 1986 and subsequent reprints. The topic of GIS application in archaeology has been covered, apart from an increasing number of papers in the proceedings of the CAA conferences, in specialised volumes such as Allen *et al.* 1990 (a good selection of case studies mainly from the States) and the recent Lock & Stančič 1995 (case studies from Europe). The proliferation of papers dealing with GIS in archaeology is explained by the strong visual component characteristic of this software, in fact GIS has been used extensively in archaeology as a mapping and presentation tool. More recently GIS has been used to perform specific types of analysis on archaeological data. Gaffney and Stančič (1991 and 1992) used cost distance surfaces to study the extent of the catchment areas of archaeological sites on the island of Hvar, in Dalmatia. The papers on the application of GIS to archaeology presented at the CAA95 conference held in Leiden (the Netherlands) showed that the stage of using GIS simply to map archaeological data is over and that the techniques of GIS are now being used for tasks such as the exploration and analysis of spatial and chronological data (see for example Massagrande forth).

The natural process of using GIS in archaeology can be viewed as one in which the first few years were devoted by the archaeologists to becoming familiar with the tool, followed by the first attempts of using the tool for more than just presentation and, finally, by the stage in which the tool has become largely transparent and it is being used as an aid for archaeological studies. This process has not been a simple and linear one and there has been much discussion of whether GIS really is suitable for archaeology and whether it is so easy to use that it becomes too easy to misuse. It is true that the ease of data manipulation offered by GIS invites the user to experiment with a variety of possibilities, occasionally losing sight of the original scope of the project, however, it should be borne in mind that GIS is only a means to achieve an end and it is up to the user to define precise goals and devise a viable strategy to pursue them. Data exploration is an integral part, and often the first step of, data analysis and in a spatial context it makes sense to try to visualise the site distribution before the relationship between archaeological sites and the landscape is studied. It has been argued (Kvamme 1994, 1) that the abuse of the plotting facilities in GIS easily leads the user to detect spatial patterns where there actually are none. For this reason it is important for the act of data visualisation to be preceded or accompanied by statistical analysis. Statistical techniques can also highlight the existence of relationships and trends which are impossible to visualise. The other big accusation applied to GIS is that its use in archaeology simply hides a return to environmental determinism (Gaffney and van Leusen 1995). This is true where only environmental variables are being used to try and explain the processes which lead to settlement formation. Other types of data, such as political boundaries, relative wealth of areas, location of communication routes and of administrative centres can be included in a GIS as well as environmental data. Recently, the possibility to include the way ancient people perceived their landscape in a GIS study has been investigated by Wheatley (1993), thus reintroducing the element of `culture' which was missing from the purely mechanical study of the relationship between the archaeological site and the landscape. Again, the criticism should be directed to the use that is made of the tool, not to the tool itself. GIS can be a powerful tool in landscape analysis, but a certain level of knowledge is required before it is applied to archaeological research. The user should have a clear idea of the type and nature of the archaeological data under study, the goals of the study, together with some knowledge of GIS procedures, general computing and statistical analysis.

Another issue concerning the use of GIS in archaeology is the `theoretical neutrality' of the technique itself (Wheatley 1993, 134). No technique can ever be totally neutral, because it is the person who is using that specific technique to decide which data to use, how to present them, what questions to ask. This is true of all the techniques applied in archaeology as well as all other disciplines to specific problems. The nature of the problem can be perceived differently by different people, so that the problem is approached in different ways. Even using GIS, it is impossible to look at all the available information in all the possible ways, or check for correlation between all the combinations of elements in a data set. This is particularly true when it is possible to use the available data to create new information, like GIS allows to do (*eg* by calculating distances from specific features). Before the technique is applied, the aspects of the problem to be investigated must first be selected by the human brain, which uses certain types of filters according to the type of school the owner of the brain belongs to. It is also true that a specific technique can influence the way a certain problem is looked at. GIS makes it very easy to investigate information that can be classified into specific categories, but it makes it more difficult to study aspects of the archaeological assemblages which have been created by subjective human behaviour such as ritual (Wheatley 1993, 134). The data used to ask the questions exist as a combination of the intrinsic essence of the information and the view of the data the person who is studying them has, therefore GIS is not theoretically neutral, just like no other technique can be, for as long as the human input is required.

3.3 GIS software

GIS only provides the tools to study and manipulate the spatial data. Non-spatial information needs to be attached to the entities in the analysis by some other means. In archaeology, more often than not, the data used in a GIS system consist of information about sites distributed across a landscape, thought there are examples of GIS having been used with other archaeological elements, such as different pottery types (Massagrande 1991) and in intra-site studies (Carreras and Massagrande forth; Csáki *et al.* 1995, Meffert 1995). These data need to be organised and stored within a database management system, which is much more powerful then a simple GIS in terms of handling of multiple attributes (a GIS can only handle one attribute at a time) and retrieval of data according to particular characteristics (`querying' the database). The GIS software which was used for this research, Idrisi, does not support directly a database management system, so that an external piece of software (dBase III+) was used and linked to Idrisi via a simple data conversion program (see discussion in chapter 3, *The system*). The database management system is one of the most vital parts of a GIS, though it is not part of GIS software proper. Without a link to a database system, the potential of using GIS software in archaeology, as well as in any other discipline, is dramatically reduced. For the scope of this thesis, GIS will be considered as including a database management system and statistical software, though separate packages had to be used for these functions.

3.4 GIS data types

All the geographical features present on a paper map can be reduced to three basic representational concepts: the point, the line and the area. On a map, some other element associated to the point, line or area defines its non-spatial attributes. So, a line representing a road can be distinguished from a line representing a river by its colour, a label next to a point can indicate that the point represents a town or an archaeological feature. In GIS the data are organised in a similar way, that is, the basic types are still the point, the line and the area (or polygon), only the way the non-spatial information is associated with each feature changes. Any information inherent to a particular feature is stored in GIS as a value, which can either be a measurement (*ie* the value of one feature might represent its elevation over sea level) or a label for discrete categories (in a soil map, 1 might represent clay, 2 chalk and so on). On the visual level, different values can be displayed as different colours or, if required, a set of different values can be displayed in the same colour and another set in another colour, creating visual groupings of separate features.

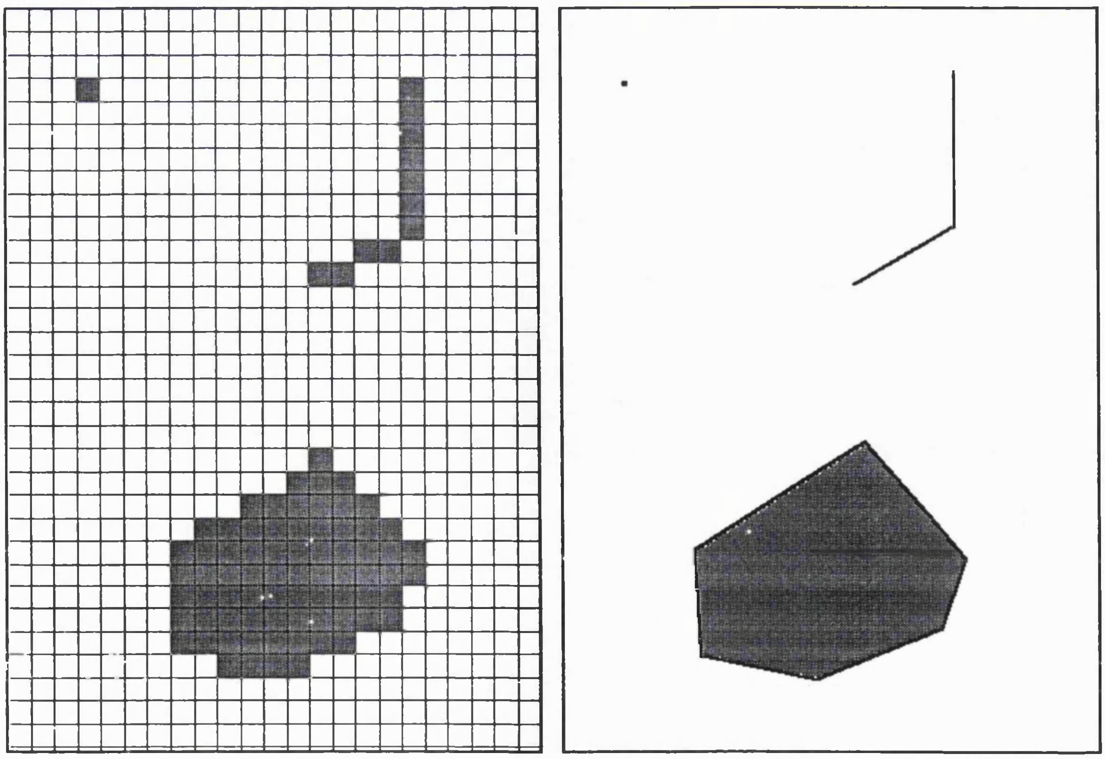

Figure 1.1 The representation of a point, a line and an area in raster (left) and vector (right) format.

There are two basic ways of storing spatial data in a computer, which are contrasting and which have different advantages and disadvantages; these are the *raster* and the *vector* format. Idrisi uses principally the raster format, so this will be described in more detail. The basic difference between the two formats is shown in figure 1.1. The spatial information is stored in raster format as a grid in which each cell has a value. A set of contiguous cells with the same value makes up a feature. Conversely, data in vector format are stored as a set of points joined together.

The vector representation is aesthetically more pleasing and, in computer terms, it has the advantage of requiring a much smaller amount of disk space than the raster representation. The vector representation, in fact, only stores information about the location of the points where the lines change direction, together with the information on the nature of the feature itself, while the raster representation always needs to store information about each individual cell, whether this is part of a feature or of the background (Idrisi does not distinguish between a cell without a value and a cell with a value of 0)[2]. The size of the raster grid is usually defined by the user before the data are input into the GIS. Obviously, the finer the grid, the better the resolution, but also the more disk space is used. It is a good policy to evaluate carefully the pros and cons of different raster grid sizes for the type of data available before the system is set up. This issue is discussed in more detail in chapter 2, *The data*. The raster and vector format also have other characteristics which make them best suited for different types of GIS analyses. Because features expressed in vector format are essentially lines with a starting and ending point joined by means of some sort of connectivity, this representation is very well suited for network analysis (such as the study of transport along certain routes), which is impossible if data are stored in raster format, where no implicit connectivity exists. On the other hand, it is virtually impossible to represent in vector format continuous features with variable values, such as an elevation map. Conversely, the raster format is perfectly suited for this type of data, as the value of each individual cell (representing a specific location on the ground) is stored separately. As this research is concerned with the study of the spatial distribution of Roman settlement

[2] In reality, there is a number of methods to store raster format data in a more space-efficient way, however, it was felt that the subject was too technical and not entirely relevant to the research to be discussed in detail here. A description of such methods can be found in Burrough 1986, 25.

over different types of soil and terrain, the vector format would not be adequate for the needs of the study. In raster GIS each cell in a map can only contain one value, therefore the values of different variables relative to the same area on the ground must be kept in separate maps, which are called *overlays.* The use of the raster format for the scope of this research also has the other advantage of allowing the combination of different overlays to create new maps containing derived data (such as is described in chapter 4, *The Guadalquivir Valley*, section 4.2).

3.5 GIS features

This section briefly presents some of the most commonly used features of raster-based GIS, more detailed discussion can be found in Burrough 1986. One of the first steps when working with a raster GIS is often that of creating a Digital Elevation Model (DEM) of the study area (see chapter 6, *The Maresme*, figure 6.9). The DEM is a digital representation of the of the continuous variation of the relief over space (Burrough 1986, 39). DEMs are used in archaeological studies to gain a visual understanding of the land form of the study area and to get a first general idea of where in the landscape archaeological sites occur. DEMs are produced by interpolating a continuous surface from a set of digitised and rasterised elevation contours, and can be used to produce derived data such as the slope, the orientation and the sun exposure. In turn, these derived maps can be used to create friction surfaces needed to generate cost distance surfaces (see below), watersheds, viewsheds (lines of sight) or least cost pathways.

Another feature of GIS which can be of use to landscape archaeology is the cost distance surface. This feature allows to calculate distances from any point taking into account constraints such as the land form, natural and artificial barriers or vegetation (see chapter 5, *The region of Tarragona*, figure 5.1). Cost distance surfaces were used extensively in the project presented in this thesis. The use of cost distance surfaces to determine the extent of the hypothetical territories of Roman towns in the Guadalquivir Valley is described in chapter 4, *The Guadalquivir Valley*, section 5.1).

Map algebra is the basic process of combining two or more overlays together. It is possible to combine two maps by adding them, in which case each cell in the resulting map will have the value corresponding to the sum of the values of the corresponding cells in the two input maps, by subtracting one from the other, by multiplying one by the other, by diving one by the other, or by having one map covering the other except where the value of the cells in the first map is 0. Map algebra is the basic set of operation in the manipulation of the overlays (see chapter 4, *The Guadalquivir Valley*, section 4.2).

Other operations which can be performed on the overlays include the possibility to perform the same operation on each cell in a map, such as multiplying all the cells in a map by a specific number, or raising them to a certain power. Related to this is application of certain mathematical operators to each cell in the map, such as square rooting, extracting the natural logarithm and so on. It is also possible to operate direct reclassification of the values of the cells in a map, such as changing all cells with a certain value to another value. GIS will also allow the calculation of the total area covered by specific features with a given value, which is useful when the percentage of (*eg*) different soil types in an area needs to be calculated.

4 Bringing archaeological surveys and GIS together

The simplest way to use GIS with survey data (systematic or not) is that of plotting the information to see whether any patterns can be identified visually, very much like archaeologists have been doing with distribution maps for years. A GIS data plot is actually not different in any way from a plot produced on paper, with the advantage that it is much easier and quicker to make variations to the way the plot is displayed. If a few hundred sites are plotted by hand on paper and only at the end it appears appropriate to use different symbols or colours, the whole job must be restarted from scratch. With data stored in a GIS the same effect can be achieved just with a few keystrokes. Since it is so simple to produce plots of the data, it is easier for the archaeologist to obtain a large number of these, with different combinations of site types, periods or any other information is thought to be relevant. While this ease of data manipulation allows the same set of information to be looked at from different points of view and angles, at the same time it carries the implicit danger of creating a large amount of redundant and non-informative distribution maps. Worse still, sometimes it is so easy to create convincing looking maps that the archaeologist simply `forgets' to interpret the data, relying on the information in the plots to be immediately self evident to the audience. While a large number of the first applications of GIS in archaeology often resulted in nice looking maps with very little analysis (if any) attached to them, it cannot be denied that there are evident advantages in using GIS to look at the distribution of archaeological sites in their landscape. Though a GIS can immediately display the result of a query on a site database, it is up to the user to make sure that the query formulated is archaeologically meaningful. The theory behind using GIS software in archaeological research is no different from answering the questions that archaeologists have been investigating for years. GIS only offers a quicker way to obtain an answer, but it is up to the archaeologist to ask the right question.

The clear advantage of using computer technology is that much larger data sets than ever before can be handled, allowing the user to obtain a wider picture of the situation. In terms of survey data, the information from several surveys carried out in the same region can be brought together and studied at the same time. Data derived from different sources need to be standardised before it is used in one system, increasing the reliability of the result by eliminating some of the recovery bias. The standardisation process would be very long and time consuming if performed by hand, but can be automated and performed by the computer. This issue is discussed in chapter 4, *The Guadalquivir Valley*, section 4.2. Database querying is not the only aspect of landscape archaeological studies which is made easier and quicker by computer techniques, a larger number of statistical tests can also be carried out. This is due to the fact that when the data are stored in digital format, it is easy to manipulate them and adapt them to the requirements of specific tests and/or specific statistical packages. Unluckily, the different components of the system used in this research (*ie*, the database management system, the GIS, the statistical packages etc), could not exchange data directly with each other, therefore it was necessary to produce a number of conversion programs. Though computing in general is moving towards integrated systems (such as Windows), in practice it will still be a few years before all the pieces of software running on one particular platform will be able to communicate and exchange data directly.

There also are problems with using raster GIS for the study or archaeological sites in a landscape. The main problem is that sites are usually reduced to single cell, disregarding the size of the actual sites, which are then studied as if they all were the same size, corresponding to the size of the cell (see chapter 2, *The data*, section 4 for a discussion of the size of cells). Van Leusen (Gaffney and van Leusen 1995, 382) suggests the use of continuous layers of some density measure or `activity index' as a possible solution. In the case of the data used in this project, this problem does not occur, as the data were collected during non-systematic surveys and the only information available is the (probable) location of the site and a list of archaeological material found at that location. The size or status of the site are unknown elements, therefore the sort of problem outlined above exists intrinsically in the data, whether or not these are studied by means of GIS. Another problem with the application of GIS to archaeological data is pointed out by Harris and Lock (1995, 356), who say that the time dimension, at present, can only be included into a GIS as a series of static `snapshots' of a specific period. Again, the data used in this project were dated to broad chronological bands because no detailed study was ever carried out on the pottery to allow smaller periods of existence or use of specific sites to be identified. Even without using a GIS, these data would have been studied as belonging to large time bands, indeed, the database management system, which an integral part of the GIS, allowed the chronological standardisation, as well as the type standardisation, of some of the data (chapter 4, *The Guadalquivir Valley*, section 3).

5 Outline of the thesis structure

Chapter 1. Introduction — This chapter explains the nature of systematic and non-systematic survey in archaeology, then it presents a review of archaeological surveying of Roman Spain. It then briefly reviews the adoption of computer techniques and GIS in archaeology and introduces the fundamentals of GIS Finally the advantages and disadvantages of the use of GIS in landscape archaeology are briefly discussed.

Chapter 2. The data — This chapter describes the type and format of the data used in the research and the way the information is organised and stored. Then the areas where the archaeological field survey has been carried out are presented and the resolution of the images in Idrisi for each area is discussed.

Chapter 3. The system — This chapter provides a description of the computer system used in the research. It describes the hardware platform and both the commercial software and, in more detail, the custom software used. It then explains what are the limitations of the system implemented and in what way they influenced the research.

Chapter 4. The Guadalquivir Valley — This chapter gives a brief historical background of the Valley of the Guadalquivir, then it discusses how the data for this region were standardised and classified chronologically. Then the characteristics of the settlement pattern in different periods are presented and the analysis of the development of settlement is discussed.

Chapter 5. Tarragona — This chapter presents the historical background of the region of Tarragona, then it details the analysis carried out on the data from this area and discusses the development of the Roman settlement pattern in time.

Chapter 6. The Maresme — This chapter briefly presents the historical background of the region of Maresme, than it describes the information available from the area and the analysis that was carried out on it in the scope of the present research.

Chapter 7. Conclusions — The results obtained in the study are presented and discussed, and a brief comparison of the three study area is made. This chapter also discusses whether the use of GIS in landscape archaeology analysis really has advantages when compared to more traditional non-computing methods. The question of whether the same study could have been carried out without using GIS software is addressed.

Chapter 2

The data

This chapter describes the different types of data used in the project presented in this thesis, their sources, the way they were collected and the method used to input and store them in digital format. Some of the technical considerations apply not only to GIS-specific archaeological data, but also to other types of data used in archaeology, such as any sort of information which needs to be stored in a database, or represented in graphic form. The nature of the available data influenced the way the information was collected and manipulated for the scope of the present research.

The data, as they were used for the GIS analysis, can be divided into two basic types: the site data and the background data, which require different means of collection, input and storage. Knowledge of the nature and form of the data is required in order to be able to appreciate fully the advantages, problems and requirements of the application of GIS in landscape archaeology, as well as in other branches of archaeology which make use of computer systems.

1 Data description

This section describes the data used in the project presented in this thesis and how they were stored before being put into digital format.

1.1 The site data

The site data consist of the information about known archaeological sites in each of the study areas. The majority of these data have been collected in non-systematic surveys and stored in card catalogues in archaeological units in different parts of Spain. Other site data were collected by people other than those working in archaeological units and published in reports by the people who carried out the surveys; part of this information was also occasionally integrated in the catalogues of the archaeological units. Some of the published data were collected during systematic surveys as well as non-systematic ones, as it is discussed in chapter 4, *The Guadalquivir Valley*. The information available for each site includes the site coordinates, the site contents and the site date and function (as inferred by the surveyors). The data stored on the card catalogues in archaeological units had to be copied on paper first, and then manually input into a database system.

1.2 The background data

The background data consist of the information about the physical characteristics of the area for which the archaeological data are available. The background data can be divided further into original data, such as the geology, the soil types, the hydrology, the coastline, the elevation and the position of modern features (such as modern towns and roads), and derived data, that is the data that can be derived from these with standard GIS modules. The derived data include information such as the slope and orientation of the land (derived from the elevation), or the distance from certain starting points (such as modern towns, water sources etc). As such data were produced only when they were required by the research, they will be discussed in the sections dealing with the analysis of the data.

2 Data storage

This section describes the structure of the database used to store the site data and the procedure employed to put the geographic data in a format suitable for use within Idrisi.

2.1 The dBase III+ data

The site data for each of the survey areas are contained in a separate database, managed with dBase III+. The information stored in these tables includes the sites coordinates on the country's national grid (in the UTM[1] system), the site type, the types of pottery and other materials recorded at each site and the period(s) in which the site

[1] Universal Transverse Mercator

was in use (as inferred by the surveyors). As Idrisi requires the identifiers of each entity or cell to be a number, an unique integer value is associated with each site type. These value can be thought of as labels representing the different site types and are used when exporting the data from dBase III+ to Idrisi to create the site types distribution maps.

The main tables of the databases containing the site data for the four survey areas all have the same basic structure, which is outlined in table 2.1. The field NUMERO contains a code which uniquely identifies the site. This is usually the code given to each site in the card catalogues in archaeological units, or in the site lists in published sources. The second and third fields, COOR_1 and COOR_2, contain the easting and northing of the site, expressed on the UTM grid. The UTM grid is the standard used in the Spanish archaeological units to record the position of archaeological sites, but some of the data from other sources were referenced to the Lambert grid, so that they had to be converted to the UTM system before they could be used in the study (see chapter 3, *The system*, section 2.21).

NAME	TYPE	SIZE
NUMERO	Numeric	4
COOR_1	Numeric	4
COOR_2	Numeric	4
NOME	Character	51
VALORE	Numeric	2
TIPO	Character	35

Table 2.1 The basic structure of the site database. To this more boolean fields were added for each of the study areas.

The UTM grid was chosen as standard in this project because the majority of the data were already referenced to it and most of the maps used had the lines of the UTM grid but not the Lambert grid, though some had both. The NOME field contains the name of the site. This is the name as it appears on the catalogues the information was taken from. The VALORE field contains the unique value used as identifier for the site type, which is also described in words in the next field, TIPO. Even though it might look like redundant data exist in the database, as the same information is entered twice in different forms (in words and in identifier values), in reality the same code can be used for a variety of slightly different types. Therefore, the identifier value 5 can indicate an inhumation grave, a cremation burial or a mausoleum; the identifier value 10 can indicate a large town (such as Seville), a smaller town (such as Écija) or a small rural agglomeration. The identifier values can be used in the analysis of the distribution of site types, while the detailed description can be used when one particular site is shown to be in some way different from the rest and has to be looked at in detail. Moreover, getting rid of the description of each site in favour of just the value identifier would imply ignoring information which is available and which might be relevant. We might want, for example, look at the distributions of the different types of burials and see if any particular type (such as inhumation or cremation) seems to be associated with other specific site types, or with sites of a particular age. We could then give a different identifier value to the each types of burial, rather than sticking to just one value to describe the whole category. On the other hand, when we want to look at the whole of the site types distribution, too many categories would be confusing.

Given the qualitative (as opposed to quantitative) nature of the data, it was decided that the best method to design the database structure was to have a separate boolean field for each of the materials that could be expected to be found in the survey areas. This meant going through the data beforehand and identifying what categories of materials were likely to be relevant for the type of analysis to be carried out. The questions which had to be answered by the material evidence are the date and function of the site, therefore elements such as the different pottery types, the presence of kilns or quernstones, evidence of use of marble or mosaic and architectural elements and others were selected as being relevant. However, since different types of material were recovered in the four survey areas, it was not possible to create a standard database table layout containing a boolean field for each of the elements expected in all the survey areas. If this was done, the result would have been an enormous database table with most of the fields utilised only when the data were entered for the particular survey area for which those fields were relevant. Such a table layout would have been more difficult to query because of the presence of a large number of fields, of which only a few would be relevant in any one occasion, and would have caused a considerable waste of disk storage space. Instead, a separate site database was created for each of the survey areas, with each main table containing the basic fields as they appear in table 2.1 plus a unique set of other fields relevant to the area of which site information is stored.

The data from the four surveys do not differ only on what materials were found in each region, but also on the degree of detail to which the same materials were recorded in each of the survey areas. The most important fine wares were recorded in all the survey areas, while others, such as coarse ware, were not. In the database of the Guadalquivir Valley the three subtypes of Terra Sigillata Chiara[2] (A, C and D) are recorded, but so is occasionally the presence of Terra Sigillata Chiara without specifying the subtype, so a separate boolean field had to be created to accommodate this. In the Maresme, the data come from two sources: the archaeological unit and the publication by Prevosti Monclús (1981b). Prevosti Monclús identifies chronological subdivisions in her data, so that the table containing the data from her publication has extra fields for Late Spanish Terra Sigillata, as well as Spanish Terra Sigillata, and for Palaeochristian pottery. The database tabes for the Maresme and the region of Tarragona have separate entries for amphorae of different origin, while the database for the Guadalquivir Valley only has a generic field for amphora, whatever they origin. This is so because the amphora type was specified too rarely in the Guadalquivir Valley card catalogue to make a study on the different amphorae possible. Other pieces of information were entered in character fields in the database and are descriptive in nature. One such is the presence of floors. Where a floor was identified, the information whether it was a mosaic floor or other types *(eg opus signinum)* was recorded in the database.

Each database table also contains a number of boolean fields representing periods of occupation. If a period field is set to true, the site was occupied in that period. This makes it quite simple to extract information about the sites according to when they were occupied. This information was included in the published material or in the paper cards, but it is possible to correct and standardise it using the information about the pottery types found at the sites.

2.2 The graphics data

The background data consist of all the environmental, human geographical and the archaeological (other than site data) variables used in the analysis. This information was derived from a number of standard military maps for each of the survey areas. For

[2] African Red Slip

the valley of the Guadalquivir the soil map was digitised as well. Different resolutions were used according to the size of the area taken into consideration. To digitise the whole of the province of Seville at a resolution of 1:25.000 would have involved using 384 maps, so the 1:100,000 maps were used instead. The soil map for the valley of the Guadalquivir only exists at a resolution of 1:400,000. For the other areas maps with a resolution of 50,000 were used. The information digitised from the maps is:

- the elevation contour lines
- the hydrology (rivers and lakes)
- the position of modern towns and roads
- the coastline (for the Tarragona and Maresme regions)
- the soil map (for the Guadalquivir Valley)

This information was originally digitised using AutoCAD v.11 and 12 (for DOS and for Windows) and then transferred into the Idrisi format with the Idrtools program (see chapter 3, *The system*, section 2.2.3). In Idrisi the files containing the elevation contour lines were used to interpolate continuous elevation surfaces (Digital Elevation Models or DEMs), creating computer-held representations of the land form of the three survey areas. The DEMs were also used to create images containing derived data, such as slope and orientation, and to produce friction surfaces to be used in the generation of cost distance surfaces.

The soil types were digitised as closed polygons. The problem of shared border lines was solved by duplicating them by digitising each line twice[3]. This caused marginal errors which lead to the existence of tiny gaps between bordering soil type areas in the resulting Idrisi images (see figure 2.1). The gaps were later filled in manually and assigned to one of the neighbouring areas. Though this method does not solve the marginal error, it is still more accurate than preserving the non-data cells, as these were originally created by small digitisation errors and do not truly reflect the position of areas for which soil data are not available. Towns and lakes were entered

[3] Though there is an edge tracing facility in AutoCAD v.12 for Windows (available indirectly though the options of the ***bhatch*** command), this failed on a number of occasions and was deemed not to be reliable enough to be used.

as closed polygons too but, obviously, there was no border line problem for these. The elevation contours were entered as lines and care was taken to prolong the lines over the border of the area which was selected to be used in the analysis, so as to avoid the interpolation errors in Idrisi which are caused when contour lines do not cross the edge of the image (the interpolated value `leaks' because the contour line is interrupted before the edge of the image).

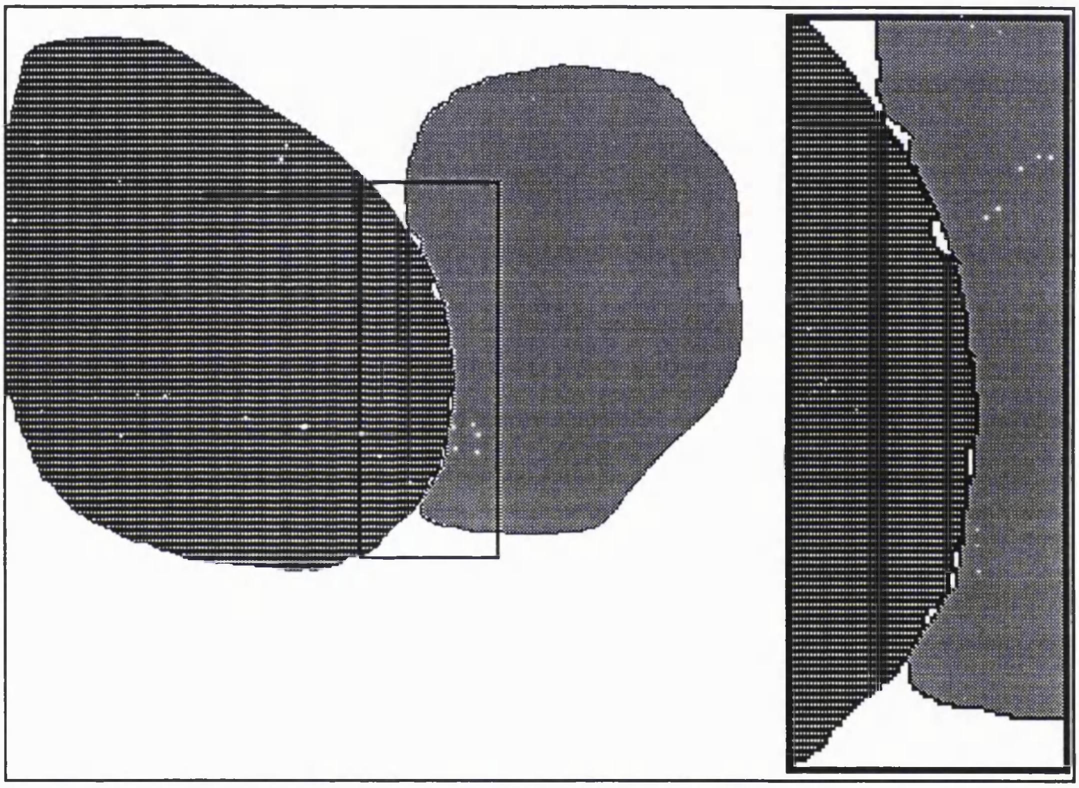

Figure 2.1 When two polygons (areas) share a common border, this has to be digitised twice, leaving non-data gaps in between

3 The three survey areas

Four criteria were used to choose a selection of study areas from the Mediterranean region. The four criteria are:

1 that the non-systematic survey had been carried out for enough seasons to cover a large enough region;

2 that the data recorded in the surveys was available to the public either as published material or from the archives of local archaeological units;

3 that enough information was recorded about each site (*ie.* site contents, not just site location);

4 that the survey areas, taken as a whole, should offer a good sample of the different geographical, geological and environmental conditions occurring throughout Spain;

Several areas which responded to these criteria were identified. Of these, three were chosen as sample study areas. The physical characteristics of the three selected survey areas are described below.

3.1 The Guadalquivir Valley

This is the region of Seville in the Guadalquivir Valley (south-west Spain). The site data for the province of Seville were collected during surveys carried out in the Guadalquivir Valley, in south-west Spain, until 1986. The more recent data collected in the 1989 survey were not yet available at the time of gathering the data for this project. For this area, a map of the different soil types was available (this map was included in De La Rosa & Moreira 1987), in addition to the maps containing the information about the hydrology, elevation, modern roads and towns.

Part of the site data were kept at the *Dirección General de Bienes Culturales* in Seville. The data were stored in a card catalogue and covered a proportion of the province of Seville. Other site data for the Guadalquivir valley were obtained from the four books published by M. Ponsich (1974, 1979, 1986, 1992) and containing the information he collected during a number of survey seasons in the area. More data were obtained from the systematic surveys carried out by Amores Carredano (1982), Escacena Carrasco and Padilla Monge (1992) Ruiz Delgado (1985) and Padilla Monge and Durán Recio (1990).

The Guadalquivir Valley is (nowadays) one of the most important agricultural areas of Spain, with very fertile soils.

This is the largest of the three survey areas, covering a region of 143 (east-west) x 108 (north-south) km. The coordinates of the south west corner of this area are 29SQA545893, and those of the north-east corner are 30SUG450880 (UTM).

3.2 The region of Tarragona

This is the area around modern Tarragona, to the south-west of Barcelona (Spain). The site information for the Tarragona area was kept at the Archaeological Service of Catalonia (*Servei de Arqueologia*, Generalitat de Catalunya) in Barcelona and stored in form of a catalogue of cards, each containing the information about one of the sites. The sites were grouped together according to the nearest town or urban centre. The background data were digitised from 1:50,000 military maps.

Tarragona is located along the Mediterranean coast of Spain. The river Francolí divides the region into two parts, which differ in geology and land form.
The size of the area which was digitised for the computer analysis is 34 (east-west) x 23 (north-south) km, the coordinates of the south west corner of this area are 31TCF390460, and those of the north-east corner are 31TCF730690 (UTM).

3.3 The Maresme

This is the region of the Maresme, north-east of Barcelona (Spain). The information concerning the of the Maresme was stored in the archives of the Archaeological Service of Catalonia (*Servei de Arqueologia*, Generalitat de Catalunya) in the same way as the data for the region of Tarragona. The background data was obtained from Spanish army maps of the region.

The Maresme is along the north-east Mediterranean coast of Spain and is separated from the region of Tarragona by the Garraf massif. The area is not very high above sea level.
The area on the ground covered by the digitised information is 25 (east-west) x 20 (north-south) km, the coordinates of the south west corner of this area are 31TDF320850, and those of the north-east corner are 31TDG570050 (UTM).

4 The resolution of the Idrisi image files

Since the three survey areas cover different extensions on the ground, the resolution of the Idrisi images containing the digital representation of the physical characteristics of the landscape is different for each area. The resolution of the Idrisi images was chosen by taking the size of the survey areas on the ground and then finding an appropriate multiplier. The multiplier was selected so that the resolution of each cell in the image would be a simple ratio. This is better explained with an example. If we have an area of 10 by 20 km on the ground and we want to choose a resolution for the image files which contain the data for that area, we could choose the value 40 as a multiplier. This would lead to the creation of an image of 400 by 800 cells, with each cell representing a square with a side of 25 metres. On the other hand, if we chose the value of 30 as multiplier, we would end up with an image of 300 by 600 cells, in which each cell represented a square with a side of 33.333333...... metres (which can also be expressed as 100/3, but would make the calculation of measurements more difficult). Obviously, since we are dealing with physical distances and measurements, it makes much more sense to stick to simple ratios which can be expressed with integers, therefore the images with the resolution described in the first of the two examples are a better representation of the real world and would make any analysis of the data based on distance values much easier than the images with the resolution described in the second example. The other important element in choosing the size of GIS images is that we want the resolution to be good enough not to lose important information but, equally, we want to take into consideration the limitations of the system we are working with. An unnecessarily large image would imply loss of disk space and longer processing time, while not necessarily containing more information than a smaller image.

Taking these points into consideration, the resolution of the images for the four survey areas were chosen. The images for the province of Seville are 715 cells wide by 540 cells high, with a cell resolution of 200 metres on the ground. The images for the region of Tarragona are 680 cells wide by 460 cells high, with a cell resolution of 50 metres on the ground. The images for the Maresme are 500 cells wide by 400 cells high. The cell resolution is 50 metres. However, if particular aspects of any one area have to be studied in detail, it is possible to isolate the relevant portion of the study area and produce images of that particular bit in greater detail, starting from the original data stored in vector format (see chapter 1, *Introduction*, section 3.4, for a discussion of the properties of data stored in vector and raster format). This was done especially for the Guadalquivir Valley, due to its size.

Chapter 3

The system

This chapter describes the system which was used in the research project presented in this thesis. First, a brief description of the hardware is given, then the software used is discussed. While generic commercial software was often adequate for the requirements of the different stages of the project, a few problems encountered during the preparation and analysis of the data required that specific software be produced to solve them.

1 The hardware

The whole research project was carried out on a 486DX computer, running at 33 MHz and equipped with a 8514/a graphic monitor and a 200Mb hard drive. Though the hard disk size is rather small, most of the software required (see section 2.1) was stored on a server of a LAN (Local Area Network) of which the machine used for the project is part. The computer also has a 120 Mb tape streamer for the storage of the data not being used for a particular part of the project. The digitisation was carried out on a 1812 Genitiser digitising tablet. Hardcopy output was obtained through a Panasonic LQ 1050 printer and a Hewlett-Packard 4MP laser printer. Colour output was produced on a colour bubble-jet printer.

This description of the hardware used, like the software section below, is intended as a rough guide to whoever may be interested in carrying out a similar sort of study as that described here. Other configurations would be equally acceptable, though it should be noticed that given the ever increasing power requirement of new software, a Pentium processor will soon become a minimum requirement.

2 The software

This section describes the software which was used for the research project. Two types of software were required to complete the project. These are the commercial software described in section 2.1 and the specific custom software which had to be produced to integrate and supplement the commercial software. This is described in section 2.2, with the most important programs being explained in detail.

2.1 The commercial software

The GIS package used was Idrisi, produced by Clark University, Massachussets. This runs under DOS, though the latest release runs under Windows and only became available at the very last stage of this project. When the project was started, Idrisi version 3.0 was available, but later versions 4.0 and 4.1 became available and were used.

The input of map data was carried out with AutoCAD. When the project was started, version 11 was available and later AutoCAD 12 for Windows was acquired. This involved the modification of some of the AutoLISP routines, originally produced for AutoCAD 11, used to export the data to Idrisi. This is discussed in section 2.2.3.

The software used for the input of site data and the management of the database of sites is dBase III+. At the time of starting the research this was the most suitable database software available and, though Access for Windows became available later, it was felt that transferring all the data to a new management system would have created useless complications without actually providing any benefits. The situation would have been different, had Idrisi for Windows been available already at the time, and anybody wishing to conduct a similar research should consider the advantages of running different pieces of software with data formats which could be easily integrated, such are the applications running under the Windows environment.

Statgraphics 4.0 and then 5.0 was used for some of the statistical analysis of the data. Some of the plots were produced using Excel, taking advantage of the fact that this package can import dBase III+ files directly. The multivariate analysis of the pottery assemblage was carried out using MV-ARCH and WinBASP.

2.2 The custom software

Due to the specific requirements of the research project, the commercial software described in the preceding section was not completely adequate. It was necessary to produce new programs to supplement the commercial packages and increase the power of the software tools used. In particular, a generic program for the transformation of point coordinates had to be produced to implement the data set from the Guadalquivir Valley in the GIS (section 2.3). Other programs which had to be developed are a set of AutoLISP routines to automate the export of AutoCAD maps to Idrisi (section 2.5) and a new Idrisi module (compatible with the DOS version and using the same front-end) to extract a background variable value in correspondence with the occurrence of archaeological sites. In addition to these, a number of programs in dBase III+ language (section 2.6) were implemented to carry out basic statistical tests on data held in the database.

2.2.1 The AxisConvert program

The data from the Guadalquivir valley present a serious problem which makes the direct implementation into the GIS impossible.

The area of the province of Seville for which information is available spans two UTM[1] zones (Figure 3.1). The line representing the border between the two UTM zones cuts across the city of Seville, therefore it is not feasible to exclude one of the two zones from the analysis, though the area under study falling within the QB zone, on the west side of the city, is much smaller than that falling in the TG zone, to the east of Seville. As Idrisi (and all other GIS packages) requires that all the elements are referenced to the same coordinate system (while here two are present for contiguous areas), it is impossible to use these data directly. The sites coordinates contained in the database are referenced to the grid covering the area in which each site occurs and can not, consequently, be used together.

The most straightforward solution is to create two different GIS sets of maps, each containing the information about the background and the sites in the two UTM zones, but this would then make it impossible to use the two sets of data together in the

[1] Universal Transverse Mercator projection.

analysis.

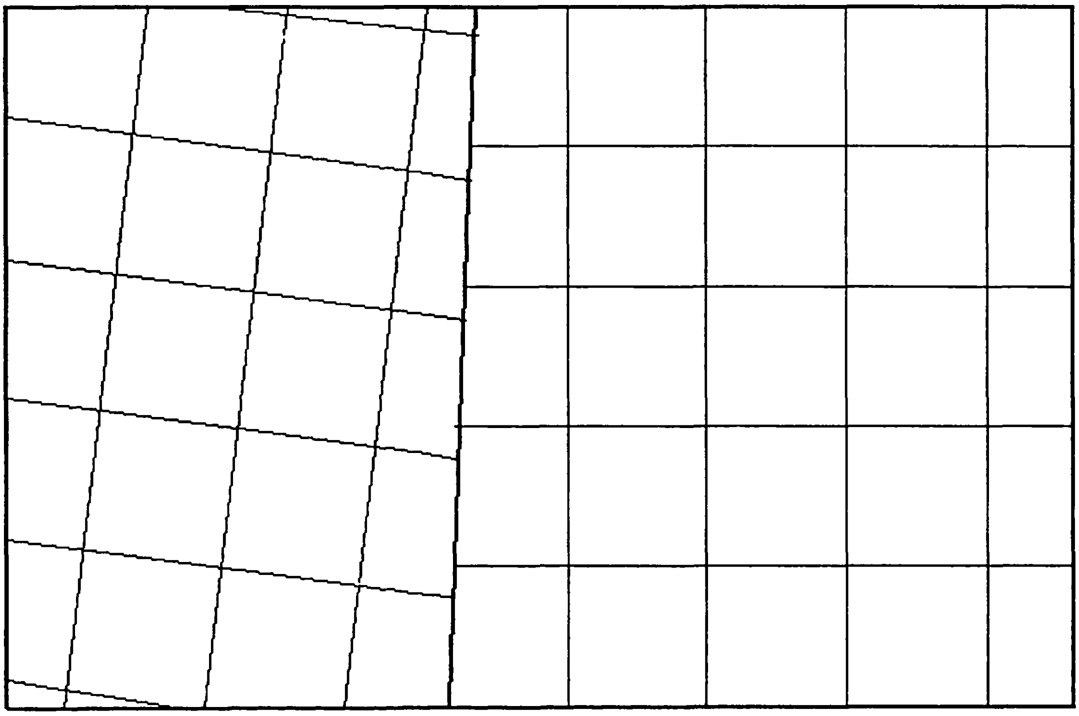

Figure 3.1 The problem with the two UTM zones in the Guadalquivir Valley

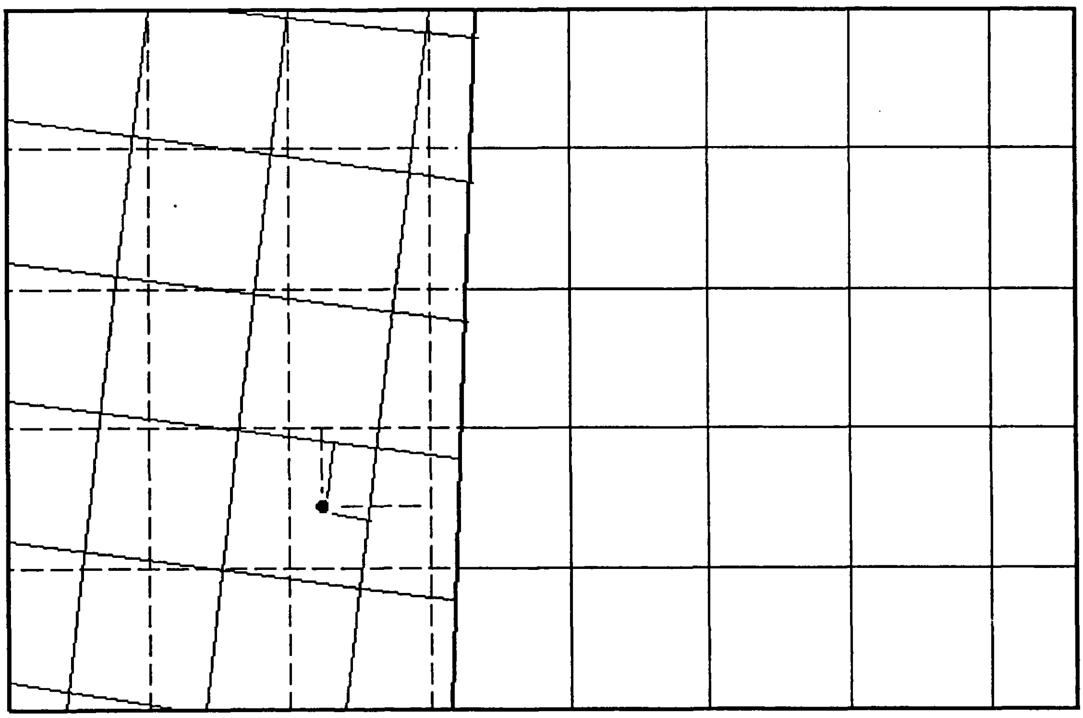

Figure 3.2 The solution to the problem of the two UTM zones in the Guadalquivir Valley, showing how the coordinates of the same point can be shifted from one system of reference to another

A more laborious but better alternative solution is to convert the coordinate system of one of the two zones, so that it uses the same origin and orientation as the coordinate system in the second zone. As the relevant area in the QB zone to the west of the border line across Seville is much smaller that the TG zone to the east of Seville, it makes sense to convert the coordinates of the database items which occur in the QB zone to the reference system used in the TG zone (see figure 3.2). The background data can easily be referenced to the TG zone system, as this process only involves calibrating the digitising tablet for the TG zone system and then using the same calibration to digitise the map data included in the QB zone.

With the site data contained in the database, however, a different approach is required. The easiest, but very time consuming, solution is to plot manually all the sites occurring in the QB zone on the maps and then either digitise them in AutoCAD and export them to the database format, or extend the grid of the TG zone with pencil and ruler over the QB zone, calculating the position of each point according to the new grid and then entering the corrected coordinate information in the database. The first solution has the disadvantage that the same site would then have different coordinates in the GIS system and in the database. This means that the sites in the GIS maps would no longer be linked to the database records containing the information about them and they could not be traced back with the ***Findsite*** program (see below, section 2.5), or exported directly from the database according to some chosen characteristics, which defies the aims of the exercise. The second solution does not suffer from this flaw, but it is very time consuming and error-prone, especially when dealing with quite a large database (302 sites in the Seville Archaeological Unit database occur in the QB zone).

A quicker and more reliable solution, in computer terms, is to write a program that uses geometrical transformations to recalculate the coordinates of the sites in the database. In a Cartesian plane (such as the UTM grid can be considered) the rotation of a point through an angle around the origin is achieved with equation (1):

$$x' = x \cdot \cos\theta - y \cdot \sin\theta, \quad y' = x \cdot \sin\theta + y \cdot \cos\theta \qquad (1)$$

where x and y are the coordinates of each point in the original grid (the grid to the left of the UTM zone border line in figure 3.2), x' and y' are the coordinates of the point in the new grid (the grid represented by the dotted lines on the left in figure 3.2) and θ is the angle of rotation between the two grids (Foley *et al.* 1990: 203).

Because the two coordinate grids do not have the same origin, as it is required by the equation, the site coordinates need to be translated (*ie* moved horizontally and vertically) as well. The equation to translate points is (2):

$$x'=x+d_x,\ y'=y+d_y \tag{2}$$

where d_x and d_y are the number of units on the x and y axes that must be added to a point's x and y coordinates to move it to its new location in a grid (Foley *et al.* 1990, 201). Because the points in the site database need to be both rotated and translated, the two equations can be combined into another (3):

$$x'=x\cdot\cos\theta-y\cdot\sin\theta+d_x,\ y'=x\cdot\sin\theta+y\cdot\cos\theta+d_y \tag{3}$$

Though it would have been very convenient to implement the equation in dBase language and operate directly on the dBase tables containing the database of sites, this was not possible as the dBase language does not support trigonometric functions such as *sin* and *cos*. Instead, the ***AxisConvert*** program was written in Turbo Pascal and then run on the coordinate values exported to ASCII format from the database. The resulting ASCII file, containing the transformed coordinate values, was then imported back into the dBase table the original coordinate values came from using a program, written in dBase language, which made sure that the transformed values were put in the right record using a unique identifier associated with each set of coordinates as a controller.

Equation (3) requires three parameters to run, that is, the angle of rotation (θ), the translation value on the x axis (d_x) and the translation value on the y axis (d_y). The rotation angle can be calculated simply by extending one of the lines of one grid over the other and then either use a protractor, or work it out using trigonometry. The translation values on the two axes were worked out by extending part of the main grid (*ie* the one to the right of the UTM zone border line) onto the other grid and plotting a few points there. The coordinate values of these points in the untransformed grid (to the left of the section border line) and in the extended grid could then easily be worked out manually. The rotation equation (1) was then applied to the values of these points in the untransformed grid, obtaining their values in a rotated, but not translated, grid. The differences on the x and y axes between the values of the same point in the rotated grid and in the extended grid are the values of d_x and d_y. To obtain a more reliable value for the translation parameters, several points were used. The values thus found were then averaged and the resulting values were used as d_x and d_y in equation (3).

It should be pointed out that Idrisi version 4.1 has a module for the conversion between different sets of coordinates. This was not available at the time the problem with the two UTM zones in the valley of the Guadalquivir was tackled. Moreover, the conversion module comes with a large number of geographical reference files which make it possible to convert from each of the geographical reference systems to any other of those provided in the files. However, these reference definition files are very biased towards the American Continent and the ones provided for Europe were not suitable for the Guadalquivir Valley, so that a transformation like the one described above could not have been performed. Despite the fact that it might have been possible to try and obtain the right geographical reference definition file from Clark University, this would have required time and would only have been possible if Clark University were prepared to produce the file. On the contrary, the coordinate system transformation procedure outlined above is a simple and quick way to transform from any system of coordinates to any other, provided that the two systems are expressed in the same units of measurement (*ie* the procedure would not work if the distances in one grid were expressed in metres and in yards in the other).

A similar problem was presented by the data obtained from the published sources. The data published by Ponsich (1974, 1979, 1987, 1991), Amores Carredano (1982), Escacena Carrasco and Padilla Monge (1992), Ruiz Delgado (1981) and Padilla Monge and Durán Recio (1990) are all referenced to the Lambert grid. Again, the direct implementation of these data into the GIS database was not possible and the same sort of transformation had to be performed. The values of the parameters were obtained following the method described above and the transformed data were then loaded into the main Seville site database.

There is an unavoidable error in the process of coordinate transformation from one system to another, brought about by number rounding during the calculation. The error in the transformation is greater as the point is further away from the starting point of the coordinate grid. For the standardised system of coordinates which was applied to the different sets of data for the Guadalquivir Valley, the origin of the grid lies at 29SQA353773 (UTM). The error in the final site distribution map should be greatest in the top right corner of the transformed and translated grid.

2.2.2 The Idr_Vals program

A different sort of problem was presented by the necessity to extract the background information at locations where archaeological sites occur. This is the case when we are interested in studying the relationship between site distribution and variables such as elevation or distance from water sources. Idrisi provides the ***crosstab*** module to do so, but, because of the limitations of the Idrisi package, this module would produces a paper output of a crosstabulation between two image files. The end result would therefore be a large table containing the number of cells for which each value in one of the images corresponds to each value in the other image. The problem is best highlighted with an example: if we wanted to know the number of sites occurring on one particular soil type and we ran the ***crosstab*** module on two images, one containing the site data and the other the soil type map, we would obtain a result as shown in table 3.1 (note: this is only an example and does not refer to any of the data sets analysed for the research project):.

Site type	Soil type 1	Soil type 2	Soil type 3	Soil type 4	Soil type 5	Soil type 6
1	3	7	12	1	0	4
2	5	4	1	7	9	1
3	14	3	5	12	0	7
4	7	5	9	8	15	17
5	6	4	2	8	13	8

Table 3.1 An example of the output of the ***crosstab*** module

This means that on soil type 1 there are 3 sites type 1, 5 sites type 2, 14 sites type 3 and so on. If we are interested in knowing the total number of sites (regardless of their type) on any one soil type, the only solution is to add them up manually. This approach is viable provided that the number of sites and of soil types is limited (see Massagrande 1991 for an example of the manual approach applied to pottery distribution) but, when we are dealing with a large number of background categories and a large number of point data, this method is definitely too slow, paper-consuming and error-prone, especially because the data also have to be manually typed into Statgraphics or any other statistical package before any further analysis can be carried out. It is actually possible to redirect the printer output to a file by specifying the file name instead of the printer port in the Idrisi environment settings, but the resulting text file is then difficult to process because it contains the character formatting used by Idrisi to make it look good on a printer, including horizontal and vertical bar characters to separate the fields.

An alternative to the ***crosstab*** module would be to use the ***extract*** module, which very conveniently gives the user the choice to select either a paper output or a file output. This module scans two image files, one containing the features and the other the background, giving as output a table of a chosen background statistic (such as minimum, maximum, average and so on) for each identifier value. Therefore, if we select the *minimum* statistic, it would tell us that the lowest elevation band (for example) site type 7 occurs on is 4. For this module to achieve the result of extracting the background data for each site location, it is necessary that each site is identified uniquely, rather than having a category value for all sites of a certain type. Even if each site was given a unique identifier to use the module, sites of the same type would have to be grouped together again after performing the operation, which would be very time consuming and would again involve a large amount of manual work, especially when dealing with large databases such as the Guadalquivir Valley one.

The ***Idr_Vals*** program, written in Turbo Pascal 6.0, was produced to allow the easy extraction of background data in the fashion required by the aims of the project.

The program takes two Idrisi image files as input[2], a continuous or categorised data background file and a point data site file. It automatically detects the data type and file format of the background data file, while it expects the site image file to contain integer identifiers of points on a background of cells with value 0. The program first scans the site image and searches for all cells with a non-zero value and records the row/column position of each site, together with the site identifier. The program then scans the background file and retrieves the value of the background variable at the same row/column location where a site occurs. The program output consists of a text file of four columns separated by a comma. The first column contains the identifier of the point, that is the site type code, the second column contains the value of the background variable at the corresponding location and the third and fourth columns contain the column and row numbers in the Idrisi image grid to uniquely identify each site. The data in this format can easily be imported into dBase III+ for further processing or into a statistical package for more extended statistical analysis than can be carried out with Idrisi alone.

Another advantage of the program over the manual method is that real numbers can be extracted and subsequently used for the analysis, while the crosstabulation can only be performed on integer numbers, which means that continuous data such as altitude, slope or distance have to be classified in a number of bands, thus losing some of the accuracy that had been gained over traditional (*ie* non-computing) methods.

2.2.3 The Idrtools routines

An AutoLISP routine, ***acdtoidr***[3] is provided with the Idrisi package for the export of map data from AutoCAD to Idrisi. The program exports the coordinates of open and closed polygons (AutoCAD polylines) taking the elevation, or z value, of each entity as the value to be given to the feature in the Idrisi vector file. This utility, however, is somehow limited as there are several ways of specifying the z value of an

[2] There also is the option to enter as input the name of a file containing the required parameters. This way, it is possible to set up a batch process which will deal with up to four sets of files (a set consists of two files for input and one for output).

[3] This is a user-provided module for which no support is offered by the Idrisi technical support team. The ***acdtoidr*** program was produced and made available by Ed Riengelman.

entity in AutoCAD, while the ***acdtoidr*** routine only recognises one type of z values as attached to AutoCAD entities. This means that in some cases the ***acdtoidr*** will export Idrisi files with identifier 0 for all drawing entities, which causes the loss of all information attached to the map data. Though the ***acdtoidr*** utility could have been modified to take into account different ways of attaching z values to AutoCAD entities (as it was in fact done in a later release of Idrisi), it was also felt that more freedom of manipulating the AutoCAD data prior to export was required. A program consisting of a set of separate and independent routines, the ***Idrtools*** program was thus created. The routines all use extended data as opposed to simple elevation (z value), this allows as many sets of extended data to be attached to any entity. This is useful if we want to have the same value for all rivers (for example) in certain types of studies, but want to be able to discriminate between navigable and non navigable rivers in others. The routines contained in the program are described below[4].

Addalt creates an AutoCAD application called ***appveiielev***[5] and uses it to store an integer number as part of the extended data set of a drawing entity selected by the user. An AutoCAD application can be thought of as a folder containing an extended data set. Any number of applications, each with a unique name, can be attached to the same drawing entity. The extended data set is a AutoCAD data type which has been included only from release 11. The extended data set can consist of strings, reals, integers or lists, or any combination of these and can be associated to any AutoCAD entity as part of the entity database. In the case of the ***Addalt*** routine only an integer is attached to the AutoCAD entity for each application. The integer attached will then be used as the object value in Idrisi. If a value is already attached to the selected drawing entity, the program notifies the user before a new value is attached.

[4] The reader should refer to the AutoCAD Reference Manual and the AutoLISP programming Reference Manual for a complete description of the AutoCAD drawing and entity database, the drawing database applications and the extended data set.

[5] All these routines automatically operate on an application called ***appveiielev***, however, this can easily be changed by editing the ***Idrtools*** program, which is stored in ASCII format, so that any number of applications as required can be created for the same drawing.

Elevall is very similar to the ***addalt*** routine but, instead of acting on one entity only, it lets the user select a set of entities, either by choosing all the entities contained in a layer, or all those contained or crossing the limits of a user-defined window. This is useful when a number of entities must be given the same value, such as the hydrology of a region, or its road network, thus saving the user the problem of having to select each single entity in turn with the risk of forgetting some.

Idrmake is the most important routine in the ***Idrtools*** program. It creates a vector file and a documentation file in the format used by Idrisi (see the Idrisi manual, Eastman 1992, for details). The program first prompts the user to enter a file name, then it checks whether this file already exists in the default directory. If it does, the user is warned, otherwise the programs gives the user the choice of how to select the set of drawing entities to be converted into the Idrisi file. The parameters required to produce the documentation file (see Idrisi manual) are also prompted for. According to these parameters the program produces either an Idrisi line vector file (open polygons, *eg* contour lines), or an Idrisi area vector file (closed polygons, *eg* soil types). This program also checks for drawing entities with no value attached and excludes them from the Idrisi file as these, if included, would cause Idrisi to crash.

This routine was originally produced for AutoCAD 11 for DOS. When it was tested on AutoCAD 12 for Windows, the program failed to work, producing a number of extra lines running from random polygon vertex points to the origin. The problem seems to be the fact that in AutoCAD 12 for Windows some extra points with x and y coordinate with empty value fields (set to 0) were introduced in the drawing database at the start of the list of vertexes of a polygon. This is not reported in the AutoCAD manual, but introducing in the ***Idrmake*** routine a check for vertexes with empty coordinate values solved the problem. This only affects the ***Idrmake*** routine, as all the others do not need to operate on the coordinates of the entities they process.

Indent highlights all the entities with no value attached to them. This is to be used if the ***Makeidri*** routine reports the presence of entities excluded from the export process to Idrisi because without a value.

Askval lets the user find out the value attached to a chosen entity. This is useful when editing large drawings where it is possible to lose track of the entities to which extended data have already been associated.

Getval highlights all entities in a drawing with a given value. This routine, together with the ***Askval*** one, are useful to identify the right entity when the lines from two paper map sheets have to be joined together in AutoCAD.

Delins deletes from the drawing any type of data except polylines, which are the only AutoCAD entity that can be successfully processed by the programs. If other types of drawing data were left in the drawing, they could interfere with the proper functioning of some of the other routines.

Newval searches for all the entities in a drawing with a specific value and replaces this value with another one. This is used to operate a reclassification on the data before these are exported to Idrisi. A reclassification module is available in Idrisi, but it was found that sometimes it was more convenient to reclassify certain types of data in AutoCAD, especially when further editing of the AutoCAD drawing was required, as there is no simple way to export an image from Idrisi back to AutoCAD[6].

2.2.4 The dBase III+ programs

A few programs were produced in dBase III+ language to achieve a better integration between this package and Idrisi and to perform some simple statistical tests. The programs are described below.

Findsite takes as an input two coordinates, then it searches the active database and returns the complete record of the site found at that location. This programs is a way to link back the sites as they are stored in Idrisi to the original database the information came from. The coordinates of a site of interest can be obtained in Idrisi by using the

[6] The ***dxfidris*** module allows the conversion of an Idrisi vector file to AutoCAD through the DXF format, but there is no way to import an Idrisi raster image to AutoCAD.

x or *c* options while displaying an image with the ***Color*** or ***Color85*** modules. These coordinates are then fed to the ***findsite*** program, which thus allows to relate a point on the screen to a specific archaeological site in the database. This program is particularly important when the object of interest is a single site, rather than the whole distribution.

Chi2 performs a χ^2 test on data imported directly into a dBase III+ table from the output of the ***Idr_Vals*** program described in section 2.2.2. The formula is implemented as it appears in Shennan 1988, 65-70. It also calculates the value of ϕ^2 (Shennan 1988, 78-79) and of Cramer's *V* (Fletcher & Lock 1991, 119-120).

KS performs a Kolmogorov-Smirnov 1-sample test on data imported into a dBase III+ table from the output of the ***Idr_Vals*** program. The values of the hypothetical distribution (the size of the areas of the variables of interest) are obtained with the ***area*** module within Idrisi. The formula is implemented as it appears in Hodder and Orton 1976, 226). The format of the formula is such that the threshold value is calculated by the program itself at the 5% and 1% levels of significance, eliminating the need to compare the results to a table of scores.

KS2 performs a 2-sample Kolmogorov-Smirnov test on data imported into a dBase III+ table from the output of the ***Idr_Vals*** program. The formula was implemented as it appears in Shennan 1988, 53-61. This program, like the ***KS*** program described above, does not produce a graphic representation of the shape of the two curves, but this can easily be achieved with Excel for Windows or Statgraphics, which were available for the project, so that it was deemed not necessary to write a specific program to obtain this result.

Though there is a number of statistical packages available on the market which would perform the same sort of tests, these dBase III+ programs were produced mainly for the reason that by working on dBase III+ tables they eliminate the need to export the data to other packages. This is only an advantage when the same test or tests must be run on a large number of data sets, as is the case when studying the distribution of sites per type and per period. The site data were exported from dBase III+ to Idrisi using the ***makeidri*** program, which was written by Dr Wheatley of Southampton University. Since this program was already available, though it is not a commercial program, producing another program to perform the same task would have meant duplicating the function of an existing program and would have been a waste of time.

3 The limitations of the system

The system described in this chapter is well suited for the sort of study which was carried out on the data sets, however, there still is scope to improve it. One of the main disadvantages with Idrisi, and with most other GIS packages, is that there is no direct link between the database and the GIS software itself. In section 2.2.4 the ***findsite*** program was discussed. Though this program does allow a link from the GIS image back to the site database, it does require a considerable amount of user interaction and input. The user has to find out the coordinate of the site of interest with the *c* or *x* options while the image is displayed, these coordinates must be written down on a piece of paper, then the user must start dBase III+ and run the ***findsite*** program before the right record is found in the database. No pointer or relationship exist any longer between the site image and the site database. Similarly, when a query is run on the database data, the results still have to be exported to Idrisi before they can be used. Idrisi does not allow any queries to be made on its graphical database. In general, Idrisi does not allow any direct link to be made with other packages such as database management software or statistical software. Though this is the case for most (if not all) packages running under the DOS environment, when there must be so much interaction between the packages as is the case for the research presented in this thesis, it is desirable that a system with a good degree of integration between its component parts be used.

It appears at the moment that the Windows environment might allow a higher degree of integration between different packages, not least for the possibility to store data in the clipboard and then read the information into another piece of software, when no other more direct link is available. In fact, the possibility to transfer data between two or more DOS applications running under Windows was exploited to a large extent during the research project. Ideally, it should be possible to exchange data forwards and backwards directly between the database management system, the GIS package and the statistical software. Alternatively, a GIS package should have its own database management facility and statistical capability. This would in the end lead to a huge monolithic package incorporating all those elements required by a working GIS analytical tool (as opposed to a simple GIS package like Idrisi, where only the management of spatial data and some very basic statistic analysis facilities exist, but no direct database management system). This idea of a all-in-one GIS tool contradicts one of the starting aims of this project, which was to use a inexpensive GIS package which would not require a specialised system to run on. Despite the fact that the computer industry is constantly progressing and creating more and more powerful machines, at the moment, a system like the one hypothesised here might require more computational power than most Archaeological Units and Institutions have. A good example of this is ARCH/INFO, possibly the best integrated GIS software available on the market at present, but rather taxing in terms of hardware. The second drawback would be that the user would be limited to the facilities offered by the package itself and would not be able to choose the database management and statistical software. Generally, specialised software works better and more efficiently than individual items making up a large, monolithic piece of software.

A new GIS package, VB.GIS 3D (Reynoso 1994), which has been presented recently seems to have all the features to make it a successful integrated GIS piece of software. Though the authors of VB.GIS 3D are still developing new functions and features for their package, so far it is has not been possible to obtain a copy of it to assess its performance. The new Windows release of Map-Info also supports raster format data and has both its own database management system and a direct link to dBase.

Apart from an advanced integration between its component parts, an effective GIS package should also have facilities such as point and click on some displayed features to retrieve the information relative to it, or the possibility to select the data to be used in statistical analysis interactively. These features are not desirable just for a

GIS package to be used in archaeology, but also for generic GIS packages or GIS packages produced for a specific field other than archaeology. Nor should this sort of features be restricted to GIS alone. These are very much the basic concepts of the field of Multimedia, which has recently made its appearance in archaeology for a variety of applications, ranging from presentation to interactive teaching (Wolle and Gamble 1995; Ryan 1995; Banning 1993) to the preparation of archaeological reports. The advantage of most Multimedia application creation tools is that it is possible to select not only which packages to use for a specific task (such as statistical analysis or map display), but also which functions need to be implemented. If a specific project involves little or no multivariate statistical analysis, implementing one or more multivariate statistical module would represent a waste of resources such as speed or disk space, which is very important in GIS because of the number of large graphics files which need to be used concurrently, or at least stored permanently on a hard drive for easy and immediate access, for which purpose a tape streamer would not be suitable. Ideally, only the facilities which are actually required would be present on the system, which would give the user a choice both qualitatively (which package to use for a specific task) and quantitatively (which features to implement in the system). When Multimedia capabilities will be available in a GIS package, the sort of analysis presented in this thesis will be much easier to perform.

Chapter 4

The Guadalquivir Valley

Before Romanisation the Iberian settlement pattern in the Guadalquivir Valley was fairly dense and organised in a network of major sites, probably urban in character. These sites controlled large territories either directly or through a network of lesser walled settlements. This settlement pattern suggests the existence of a complex social organisation, in which each major urban settlement was the focus of a chiefdom and controlled its territory through the second-order sites (Ruiz Rodríguez *et al.* 1991, 31-32; Keay 1992, 278-285; see also Ruiz Rodríguez 1988 for a study of the spatial distribution of Iberian *oppida* prior to the arrival of the Romans).

After the First Punic War (264-241 BC) Carthage expanded its presence in south-east Spain, thus causing Rome to intervene in this area to gain control of the Mediterranean. The Second Punic War (218-202 BC) saw the defeat of the Carthaginians by the hands of the Romans, who settled in the Guadalquivir Valley to prevent Carthage from re-entering the area. During the campaign against the Carthaginians, the Romans tried to gain the support of the native Iberian population through gifts and gestures of goodwill by individual commanders (Knapp, 1977, 37, ff). The personal contacts between the Roman commanders and the local elite later played an important role in the Romanisation of the various Iberian communities (Curchin 1990, 86, ff). In the year 197 BC the Romans divided the Iberian territories into two separate provinces: *Hispania Citerior* in the north-east and *Hispania Ulterior* in the south-west (Knapp 1977, 61-5; Keay 1988, 31; Albertini 1923, 12-15), this latter corresponding roughly to modern Andalucía to the east of the lower Guadalquivir with capital at Corduba (Córdoba; see Keay 1988, 32; 47). Because of the alliances the Roman commanders had already struck with the local elite, little military intervention was required to gain control of the region, with a few military camps manned by Italians sited at Iberian settlements between the later 3rd and middle 2nd centuries BC, such as the one at Italica (Knapp, 1977, 111 ff; Peña 1984, 50-53; Keay 1992, 287). The Roman camps in the Guadalquivir Valley were used both to keep stability in the area and as starting point for several campaigns of the Roman army against the *Lusitani*, an Iberian population living in neighbouring modern Portugal (Keay 1992, 287), who had been constantly threatening *Hispania Ulterior* and whom the Romans finally defeated in 139 BC (Keay 1988, 32-35). The danger of the Lusitanian raids re-emerged between the years 72 and 62 BC, when Caius Julius Caesar managed to dislodge the Lusitanians from the safety of their hillforts and keep them under control (Keay 1988, 44).

The Romans did not impose strict adminstrative arrangements in the Guadalquivir Valley in the second century BC (Richardson 1986, 109 ff; Knapp 1977, 61 ff), but seem to have worked through the pre-existing Iberian political system until the middle of the 1st century BC (Keay 1992, 287). Soon after the creation of the province of *Hispania Ulterior*, the Romans started creating new urban sites, such as the Latin *colonia* at Carteia and a mixed settlement of Roman and local elites at Corduba in the middle of the 2nd century BC (Keay 1992, 288, Knapp, 1977, 116 ff; Richardson 1986, 118 ff, Marín Díaz 1988, 129 ff, Rodríguez Neila 1988, 209 ff; Peña 1984, 60-62). The pre-existing Iberian *oppida* urban settlement system survived throughout the 2nd century BC and recent research shows that Roman municipalization at this time was very subsidiary in nature (Ruiz Rodríguez *et al.* 1991, 32; Ruiz Rodríguez and Molinos 1985). The very first Roman town in Iberia was founded in 206 BC at Italica, near Seville, in the Guadalquivir Valley (Knapp 1977, 27), while other major settlements had been created in the valley during the Carthaginian occupation, such as Carmo (Carmona) and Carteia (Keay 1988, 50). Several Roman *coloniae* (formal settlements of Roman citizens) were created in the Guadalquivir Valley under Caesar and Augustus (Galsterer 1971), while a number of Iberian towns were given the status of *municipium*. Between the deaths of Julius Caesar (44 BC) and Augustus (14 AD) nine *coloniae* were founded in *Baetica* and settled by veteran legionaries (Keay 1988, 55). A formal administrative structure was imposed upon *Hispania Ulterior* under Augustus (Albertini 1923, 3 ff), who intended to make a statement about the supremacy of Rome over the subject communities of *Ulterior* (Keay 1992, 295). The reorganisation of *Hispania Ulterior* was based on a idealised political landscape which recognised the supremacy of Rome, as it was the case for other provinces (Purcell 1990). Between 16 and 13 BC Augustus split *Hispania Ulterior* into two parts: *Hispania Baetica* to the south and *Hispania Lusitania* to the north (Keay 1988, 49; Albertini 1923, 33 ff.; Mackie 1983, 16 ff.). At the same time the province was divided into four judicial *conventus*, with Hispalis (Seville) as capital of one of them and Astigi (Écija) as capital of another (Keay 1992, 295-297).

Between the late 1st century BC and the 1st century AD, the Guadalquivir Valley had started producing both oil and wine, which were exported to various provinces of the western Empire. By the middle of the 1st century AD, the production and export of wine from the Guadalquivir Valley diminished dramatically, leaving olive oil, which was being produced on large scale, as the main cash crop (Keay 1992, 305-306). Baetican olive oil was exported to the city of Rome and the legionary camps along the British, Rhine and Danube frontiers during the first two centuries AD. The focus of export shifted from the 3rd century onwards to the northern frontiers due to competition from the African production (Remesal 1986; Keay 1988, 176). To the north the rich silver mines of the Sierra Morena, which had been exploited since prehistoric times (Domergue 1990, 164), were easily accessible (Keay 1988, 63). The Romans did not start exploiting the silver mines of the Sierra Morena until the beginning of the 1st century BC (Keay 1992, 289), probably because their control over this area was not secure until this time (Domergue 1990, 1884 ff).

1 Physical description of the land

The study area covers a large part of the modern province of Seville, spanning 143 (east-west) x 108 (north-south) km. The coordinates of the south west corner of this area are 29SQA545893, and those of the north-east corner are 30SUG450880 (UTM[1]). Most of the study area consists of the fertile valley of the river Guadalquivir, with the first foothills of the Sierra Morena to the north-west and hills to the south-east. Apart from these two areas of higher ground, the rest of the study area is almost completely flat and well-drained. The Sierra Morena has rich mineral resources which have been continuously exploited since prehistoric times. To the north are the diluvial terraces of

[1] Universal Transverse Mercator

the Guadalquivir, which formed during the quaternary, while to the south is the Vega, a flat erosion area. These two landscapes are separated by the line of the Alcores, a series of hills which formed during the tertiary. A number of streams flow from the Alcores towards the Vega, eroding parts of the hills which are used for passage between the terraces and the Vega (Amores Carredano 1982, 48).

2 The classification of rural sites

The classification of the sites was standardised according to the characteristics of the data stored in the database. Despite the fact that no quantitative data were available in the database, it was still possible to discriminate between the sites according to the materials contained in them. The data allowed a distinction to be made between high-status sites and low-status sites. While at the beginning of the project theses two categories had been labelled as `villae' and `farms', the term `villa' carries a strong cultural baggage so that it was decided to drop this terminology to avoid confusion. It would have been inappropriate to use the term `villa' in a study which spans a large area such as the province of Seville and several centuries, during which the concept of `villa' itself changed to reflect different realities. The problem of defining what a `villa' was and how to see it on the ground has been discussed by various authors (Prevosti Monclús, 1981b, 21-26; 1984, 164-168; Percival 1976, 14-15; Collingwood 1930, 113; Rivet 1969). For these reasons it was felt that it would be better to use generic terms such as `high-status site' and `low-status site' instead. Sites containing combinations of certain diagnostic elements were labelled as `high-status sites' in the database, while those missing them were labelled as `low-status sites'. If only one diagnostic element was present in a site, that was not considered enough to regard the site as a high-status one. This classification was only applied to those sites in the database which the surveyors had classified as `villa', `farm' or `rural site', while the other site types, such as kilns and cemeteries were not standardised, as it was felt that there was less risk of personal bias in the recognition of these categories of site. The classification of urban sites by the surveyors also posed some problems, so that a selection had to be operated on the sites marked as `urban' to ensure that towns of similar size and importance were being used in the analysis. These are discussed in section 4.3.

A rural site was classified as a high-status one if it contained:

- a floor (mosaic or *opus signinum*), or
- structures such as cisterns or visible complex room plans, or
- marble elements (statues, architectural parts), or
- a kiln, or
- dolia, or
- mill or quernstones,

— and —

- Arretine Terra Sigillata, or
- Terra Sigillata Hispanica, or
- South Gaulish Terra Sigillata, or
- Terra Sigillata Chiara (any subtype), or
- Thin-walled ware.

On the other hand, if a kiln and only local pottery were present, but no other elements such as imported pottery or floors, that was not considered enough to classify the site as a high-status one. Any site which had been classified as a high-status one (villa) by the surveyors (systematic and non-systematic) was classified as a low-status one if it did not meet the specifications outlined above. As a result, according to this classification, 476 sites were classified as high-status. Of these, 285 had been originally classified as villae by the surveyors, while 191 were otherwise classified. The total number of sites which had been classified as villae by the surveyors was 642, which reflects the tendency for this methodology to stay on the `low-status' side when classifying rural sites. The total of sites which were classified as low-status rural sites following this method is 953.

Carrillo Díaz-Pines and Hidalgo-Prieto (1990) carried out a study on the Roman settlement pattern in the parish of Palma del Río, near Córdoba, close to the border of the province of Seville. They used a similar method to standardise sites according to their contents but had a larger number of categories (*ibid.* 44-47).

3 The chronological division of the sites

The sites dating to the Roman period found in the Guadalquivir Valley were dated into three broad chronological categories according to the diagnostic pottery found in them[2]. The three categories are:

• Period 1 - Republic

The sites included in this category were those in which Black Glaze pottery was found. The chronological limits of this period for the purposes of the data dealt with in the project are from the 3rd century BC to end of the 1st century BC.

• Period 2 - Early Empire

To be included in this category sites had to have one or more of the following pottery types: Arretine Terra Sigillata, South Gaulish Terra Sigillata, Terra Sigillata Hispanica, Thin-Walled ware or Terra Sigillata Chiara A[3]. This period spans from the beginning of the 1st century AD to the mid 3rd century AD.

• Period 3 - Late Empire

Sites dated to this period had either Terra Sigillata Chiara C or Terra Sigillata Chiara D. It should be noted that a large part of the Terra Sigillata Chiara D found and identified in local excavations is actually imitation pottery produced locally. This fact, however, has no effect as to the value of this type of pottery to the dating of the sites, as the imitation pottery must be contemporaneous to, or later than, the type it imitates. Interestingly, the fact that local imitation of Terra Sigillata Chiara D is so widespread in the area is a positive factor in terms of dating because, generally, sites are likely to

---

[2]For the chronology of the pottery types see Hayes 1972 & 1980; Sanmartí-Grego 1978; Johns 1971; Fernández Fonseca 1995; de la Bédoyère 1988.

[3] Terra Sigillata Chiara is better known to British archaeologists as African Red Slip (ARS).

have local fine pottery than import fine pottery. The chronological limits are from the mid 3rd to the 6th century AD.

A number of sites contained a type of pottery which the surveyors just classified as Terra Sigillata Chiara without specifying the sub-type. These sites were included in both the period 2 and period 3 groups, but were flagged to signify that the dating is not precise. A large number of sites (1007) were classed as dating to the Roman period by the surveyors, but no diagnostic pottery was recovered from these. These sites were again flagged to distinguish them from those which have been classed into any of the three chronological groups.

Even excluding the sites which only contained non-identified Terra Sigillata Chiara, the number of sites increased dramatically during the Early Empire (period 2) and then dropped again, though not so sharply, in the Late Empire (period 3). The number of sites which were certainly in use during the Republic is 42, while the number increases to 618 in the Early Empire (721 including the sites where generic Terra Sigillata was recorded) and then decreases to 283 in the Late Empire (467 including those sites with generic Terra Sigillata).

4 The analysis of the data

The distribution of high and low-status sites was tested against several landscape features[4] to establish whether a relationship between these features and the rural site distribution of the Guadalquivir Valley existed. Both environmental and socio-economic features were used in the tests, to try to avoid the danger of falling into environmental determinism. The environmental features tested for relationship with the rural site distribution are the soil types and the agricultural potential of the valley, while the socio-economic elements are the distance from Roman towns, the distance from waterways and the distance from Roman roads. Why the waterways are considered such an important feature and treated separately from the roads will become clear later. To

[4] In this context the term `landscape feature' refers to any element present on the landscape, whether it is environmental in nature, such as the soil type, or artificial, such as the road network.

make the discussion of the results of the various tests easier, a summary of these is presented in tabular format in tables 4.3 for the low-status sites and 4.4 for the high-status sites.

4.1 The soil map

The first series of tests to be carried out was on the relationship between the soils and the location of rural sites. The soil maps was included in De la Rosa and Moreira 1987, who also carry out a study on the soils of Andalucía. The χ^2 (chi-squared) one-sample test was used to determine whether there is a significant relationship between the location of high and low-status sites and the soil type on which the sites occur. The χ^2 test is used to determine whether a statistically significant relationship exists between two distributions (two-sample test), or between a distribution and a theoretically-based model (one-sample test). The one-sample χ^2 test, as used in this study, derives a theoretical distribution from the area of each of the soil types (or other variable of interest) and compares this to the count of sites occurring on each soil type. If the difference between the distribution of site counts and the soil type area sizes exceeds a set value, then the relationship between the two is said to be *statistically significant*, that is, the spatial distribution of sites is not random but influenced by the soil type. The value against which to compare the difference between the two distributions is determined by the *level of significance*, *ie* the probability that the difference between two distributions can be as large as it is under the assumption that there is no difference between them (traditionally the levels of significance mostly used are the 5% and 1% levels). An introduction to the one-sample χ^2 test can be found in Cohen and Holliday (1982, 132-135) and Shennan (1988, 65-70), who also provides an archaeological example. The use of the one-sample χ^2 test to study the relationship between sites and other features can be found in Hodder and Orton (1976, 224-226). An introduction to statistical inference and significance testing in archaeology is in Shennan (1988, 48-61).

High-status sites were tested separately from the low-status ones. Since the χ^2 test compares the area covered by a particular soil type with the proportion of sites occurring on that soil type, it was thought that an unnecessary error would be introduced in the analysis by including areas for which data were not available. Idrisi images are necessarily rectangular, while the archaeological data had only been collected within the borders of the province of Seville, which is irregularly shaped. To reduce the error, the area outside the province of Seville was masked and excluded from the χ^2 test (see figure 4.1). The test was significant both in the case of the high and the low-status sites at the 5% level of significance, indicating that the sites are distributed differently on different soil types. One interesting feature emerging from the detailed analysis of the observed versus the expected number of sites on each soil type is that there are more sites than expected on the Entisols and Vertisols, which are not very good agriculturally, while the Alfisols, an agriculturally good soil type, had fewer sites than expected.

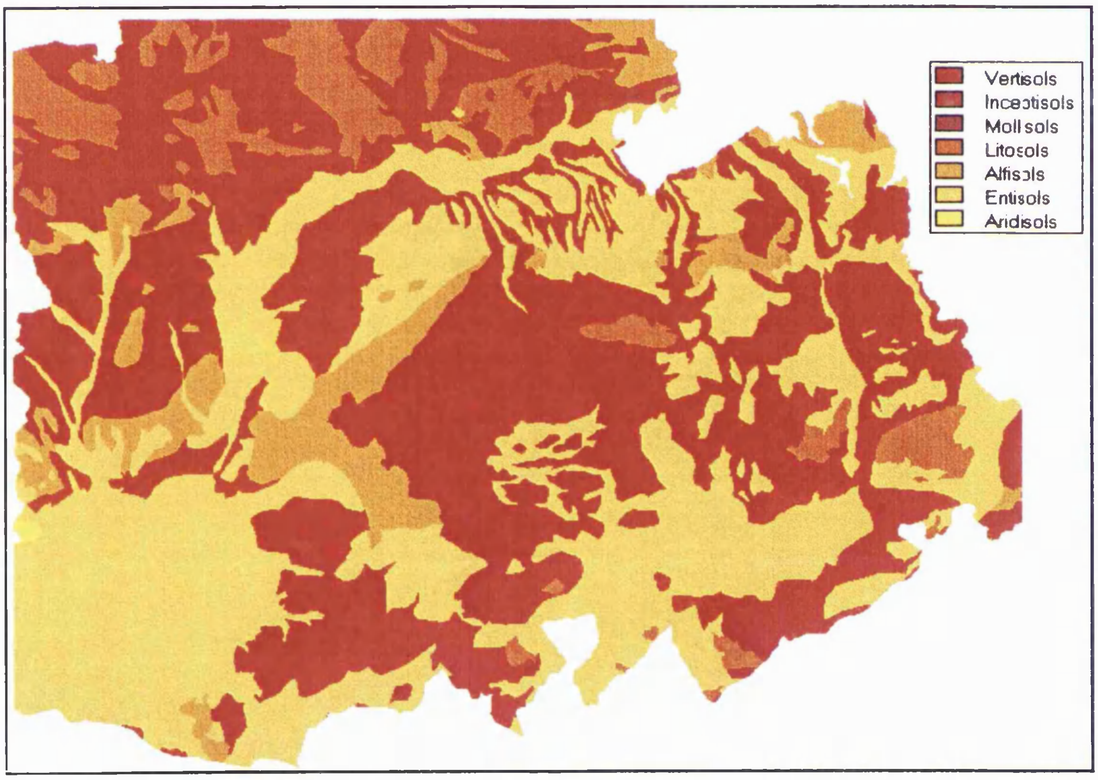

Figure 4.1 The soil map masked to exclude the area outside the province of Seville, for which no archaeological data are available.

The other important soil type in the area, the Inceptisols, did not show a significant difference between the expected and the observed number of sites. In practice, this means that rural sites, both high and low-status, were apparently located preferentially on bad soils, while good soils were chosen less for settlement. The tests carried out on the site distributions for the different chronological bands, Republic, Early Empire and

Late Empire, yielded very similar results. The only period for which the χ^2 test proved not significant (at the 5% level of significance) is the Republic (both for high and low-status sites).

A few hypotheses can be formulated to explain this pattern:

i Bad soils were preferred for settlement so that good soils were left available for cultivation.

ii Sites could be more visible on `bad' soils than on `good' soils because of intrinsic characteristics of the soil itself.

iii Since there has been continuous settlement and agricultural exploitation of the Guadalquivir Valley until modern times, the archaeological evidence on the best soils has been destroyed, while that on the worse soils has been disturbed to a lesser extent.

iv The best soils were exploited so much in the past, that they have been impoverished and now they appear as the least fertile ones in the valley.

v The archaeologists could gain access to areas of bad soil because they not being used for agriculture, therefore the sample is biased.

At present it is difficult to see which of these hypotheses, if any one at all, can help explain why bad soils have more sites than expected, while good soils have less sites than expected. There is another important element which must be taken into account when considering the results of this analysis, which is the fact that the χ^2 test is extremely sensitive to sample size. This generally means that if the number of sites in the test is high enough, the test will be significant. Shennan (1988, 77-78) discusses this point and shows that if the same proportional distribution is kept in each category within the χ^2 test, but the sample size is halved, the resulting value of the calculated χ^2 is halved and, similarly, if the sample size is doubled, the value of the calculated χ^2 doubles as well.

> Thus, in general, if we keep the proportions in the cells constant and simply multiply the numbers by some factor k, then we

> multiply the resulting chi-squared by *k*.
>
> (Shennan 1988, 78)

> ... you can practically always obtain a significant relationship by making the sample size large enough.
>
> (Shennan 1988, 74)

When the number of observations is very large very small relationships will be picked up by the χ^2 test, while, if the sample size is very small, the relationship must be very strong to be identified by the χ^2 test. While very strong relationships in very small samples are bound to be extremely interesting, the same can not be said for very weak relationships in very large samples (*ibid*).

> ... a very slight relationship or difference may be `real', but does it matter?
>
> (Shennan 1988, 78)

In the analysis just carried out, each of the data sets contains a large number of sites, except from the samples dating from the Republic, which are the only sets to have been shown not be significantly similar to the background distribution (*ie* the soil type). For this reason, it was felt that the results of these tests should be treated with some caution.

4.2 The Agricultural potential

Given the limitations listed above, it was felt that a different approach should be used as a control for the results of the χ^2 test on the relationship between rural site location and soil type. Moreover, while the quality of the soil certainly is an important element in agriculture, there are other factors which determine whether an area is suitable for cultivation or not. Two important factors in agriculture in the Guadalquivir Valley, besides soil type, are the slope and the distance from water. These elements were combined to create a model representing the agricultural suitability of different areas of the valley. The original data used to achieve this result were stored as maps within Idrisi. The information for each of the three elements was weighted according to their relative importance and then combined with the others through the use of map algebra. The weighting of the variables is outlined below.

Possibly the most important element in the agricultural suitability map is the soil map. This was reclassified so that each soil type area within the Idrisi map was given a weighted value corresponding to the relative fertility of the soil. The best soils for cultivation were given a score of 4 while soils completely lacking nutrients were given a score of 0 (for soil fertility classification see Cruickshank 1972, 109-110; see also Gerrard 1981). Table 4.1 contains the scores given to each soil type. It should be noticed that while Alfisols are generally regarded as being more fertile than the Inceptisols, this relationship was reversed in the weighting because the Alfisols in the Guadalquivir Valley are subject to a strong erosion causing the underlying clay, which hinders cultivation, to come to the surface (De la Rosa and Moreira 1987, 22).

Soil Type	Value
Inceptisols	4
Alfisols	3
Vertisols	2
Entisols	1
Aridisols	0
Litosols	0

Table 4.1 The fertility score of the soil types

The slope gradient was derived from the DEM of the Guadalquivir Valley and then classified into 5 categories, from 4 (almost flat ground) to -10 (cultivation impossible). Table 4.2 details the categories limits and the values attributed to each.

Slope (%)	Value
0-7	4
8-15	3
16-23	2
24-30	1
30+	-10

Table 4.2 The weighted values of the slope gradient

A slope gradient higher than 30% would make any sort of cultivation impossible whatever the values of the other two variables used to create the agricultural suitability map in those location, therefore a value of -10 was given to the cells where the slope is above this value, to make sure that the final agricultural potential map would have a negative value in those location, even if the optimal soil and water condition occur there. The ranges used to weight the slope gradient were derived from De La Rosa and Moreira (1987, 90, table 6.4). Despite the fact that the largest part of the study area consists of flat valley bottom, the higher ground to the north-west and south-east of the area features high slope gradients at points, so that cultivation in these locations would be impossible. Moreover, since the majority of the valley is flat, areas with a certain slope would only be put under cultivation if flatter land was not available, so that locations where cultivation is still possible but with rather high slope gradient can be considered as marginal land.

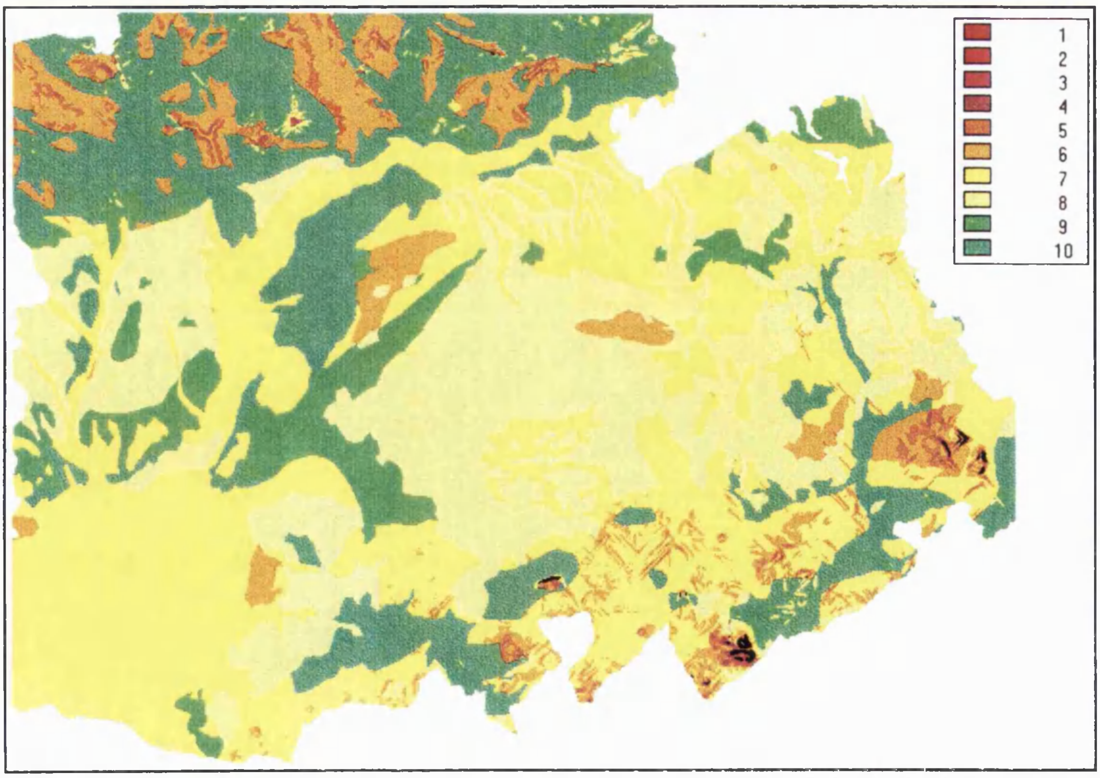

Figure 4.2 The agricultural potential prediction map, masked to exclude the area outside the province of Seville. 1 = lowest agricultural potential, 10 = highest agricultural potential

The last variable used to create the agricultural suitability map was the distance from water. This map was obtained by calculating a cost distance from water sources. Despite the fact that the Guadalquivir Valley is extremely well drained, the area can be very dry in summer, therefore an easy access to water would have been a requirement for people wishing to take up cultivation in Roman time. Though the Romans created a system of artificial canals in the area, the further away they were from the water source, the smaller the amount of water that would be carried by the canals, so that being close to major water sources which would not dry up completely in summer would still have been an important issue. The climate in Roman times was similar to the modern and the severe droughts which plague the Guadalquivir Valley nowadays would have been a problem in the past as well. The distance from water map was reclassified to have a value of 2 where the water was easily accessible and 1 where the water was not so easily accessible. Note that since the distance from water sources was calculated using a cost surface, it does not make sense to give a linear distance before or after which it is easy or difficult to have access to water.

The three overlays were combined by adding them together using map algebra. The map obtained from the overlay operation ranged from -9 (where cultivation is impossible due to the presence of a slope gradient greater than 30%) to 10 (the best possible location for cultivation). The map was reclassified so that any cell with a negative value was given a new score of 0. The result was a map in which each cell has a value ranging from 0 (worst agricultural land/cultivation impossible) to 10 (best agricultural land). The resulting map is shown in figure 4.2. This map was used to test whether agricultural potential was one of the reasons for choosing any specific location for settlement.

Since the agricultural capability prediction classes can be thought of as an ordinal scale, it is possible to use the Kolmogorov-Smirnov one-sample test to test the relationship between the site distribution and the agricultural potential of the land. The Kolmogorov-Smirnov test (Massey 1951) is a test for goodness of fit and it is generally considered to be more sensitive than the χ^2 test because it takes into account the additional information contained in the ordering of the categories (Hodder and Orton 1976, 226). Like the χ^2 test, the Kolmogorov-Smirnov one-sample test compares a real distribution to a reference distribution. Fletcher and Lock (1991, 91-100) use the Kolmogorov-Smirnov one-sample test to test for normality in a distribution by comparing it to the normal distribution. In the same way, the distribution of site counts can be tested against an hypothetical one derived from the background distribution (in this case the agricultural potential). Instead of comparing the actual values to the expected values the Kolmogorov-Smirnov test compares the cumulative value of all the categories below (or contained in) each category in turn. So, agricultural potential 5 can be though of as containing agricultural potential 4, 3, 2 and 1 and the counts of sites occurring in each of the areas are summed to give the cumulative count for agricultural potential 5. This is then compared to the cumulative hypothetical value for the same agricultural potential. An introduction to the Kolmogorov-Smirnov test one-sample test can be found in Cohen and Holliday (1982, 138-141), while Hodder and Orton (1976, 226-229) present an example of its application to archaeological data. While it can be argued that the soil fertility index as weighted in table 4.1 can be considered measured on an ordinal scale and would therefore have been suitable just on its own to carry out the Kolmogorov-Smirnov one-sample test, the information added by the other elements of the model used to create the agricultural suitability map add extra dimension to the analysis.

The Kolmogorov-Smirnov one-sample test was carried out for the distributions of high-status and low-status sites for all periods first and then for the Republic, the Early Empire and the Late Empire separately. The results of the tests are summarised below.

Low-status sites: the test was significant for the sites from all the periods considered together at the 5% level of significance. However, when the three periods were tested separately, none of the test was significant. All tests were carried out at the 5% levels of significance.

High-status sites: the test was significant for the sites from all periods considered together at the 5% level of significance. However, when the three periods were tested separately, none of the test was significant. All tests were carried out at the 5% levels of significance.

These results of the tests on the relationship between rural sites and the agricultural potential of the land contradict the tests on the relationship between the rural sites and the soil type, thus generating the problem as to which of these two sets of results are to be trusted. An important consideration is that the soil data are included in the agricultural potential map, so that all the variation in the result must be due to the other two elements which were used in the model for the agricultural suitability, the distance from water and the slope gradient. These two elements have a limiting effect on the soil agricultural potential, in fact very good soil can occur very far from water or on slope gradients too high to allow cultivation. Faced with the choice between very good agricultural soil under these condition and less good soil where the other two elements have a high enough score, the Roman farmer would have chosen those areas where cultivation was easier or at all possible. For these reasons, the results of the tests on the relationship between agricultural potential and the position of Roman sites were preferred over the results of the tests on the relationship between rural sites and just soil types.

These results seem to indicate that there is a significant relationship between rural site (both high and low-status) and areas particularly suited for agriculture, but the significance of this relationship diminishes when the chronological subsets of the site data set are taken individually. None of the subsets present a significant relationship between rural sites and the agricultural potential of the land at the 5% level of significance. Overall, the results of the analysis suggest that the intrinsic agricultural characteristics of the land were not a major influence in the Romans' choice of location for new rural sites. The reasons for these can be varied, but it is known that the first Roman settlement in the area was military in nature, therefore there would have been factors other than agricultural exploitation to influence site location. When the first `true' Roman settlers (*ie* civilians as opposed to soldiers) arrived in the area in large enough number to have an impact of the landscape, the land was redistributed according to the scheme of centuriation of which, unluckily, very little example is left in the Guadalquivir Valley nowadays. The centuriated land was based around the Roman colonies which we know were created on or by former Iberian urban centres to take full advantage of the pre-Roman Iberian organisation, so that the choice of land was based on the factors which had caused the local Iberian population to select certain areas rather than others. There also is the possibility that the Romans did not actually have complete choice of where to settle, as the pre-existing Iberian settlement was still using large part of the land.

If agricultural potential did not have a direct influence in the creation of the Roman settlement pattern, there must have been other elements to influence site location. Other possible factors which were investigated are the distance of rural sites from towns, the distance from navigable rivers and the distance from Roman roads.

4.3 The distance from towns

The relationship between the location of rural sites and that of known Roman towns was tested by calculating a cost distance surface having the Roman towns in the Guadalquivir Valley as starting points. Before this could be done, however, it became necessary to identify which of the urban sites present in the database were actual towns. An approach like that used to identify high-status and low-status rural sites described in section 2 is not possible here because qualitative data alone can not be used to define the urban nature of a site. In the site database for the Guadalquivir Valley, there are a number of sites which are classified as \`town', \`city', \`urban centre', \`urban agglomeration', \`village' or \`*oppidum*'. Because of the lack of a better method, only the urban centres which had been identified as those mentioned in classical sources and those which have been accepted as such by the scholars (such as Sillières 1990; Knapp 1977; Galsterer 1971) working on the Roman archaeology of the area were selected. The cost distance map derived from these towns was used to carry out a series of Kolmogorov-Smirnov one-sample tests to check for the existence of a relationship between the position of rural sites and the location of Roman towns. Again, high-status and low-status sites were tested separately and so were the sites dated to specific chronological bands.

The cost distance surface was divided into a number of distance bands, each corresponding to the cost of moving over 1.5 km of flat (non-friction) land. See below, section 5.1, for a discussion of cost equivalent units and definition of terms. The count of rural sites on the cost distance bands was used to perform a series of Kolmogorov-

Smirnov one-sample tests for all the site distributions, divided by status and by period. All tests were significant at the 5% level of significance, with the exception of the test for the relationship between the Roman towns and the Republican low- and high- status sites. The results of the tests indicate that the location of low and high-status sites was likely to be influenced by the location of the urban centres. Having established that there is a significant relationship between Roman towns and Roman rural sites, this relationship was investigated further. The graphs of the percentages of sites in distance bands from the nearest town for each of the three chronological bands are shown in figures 4.3-4.5. The percentages are calculated on the total number of sites of the same status dating from the specific period in question.

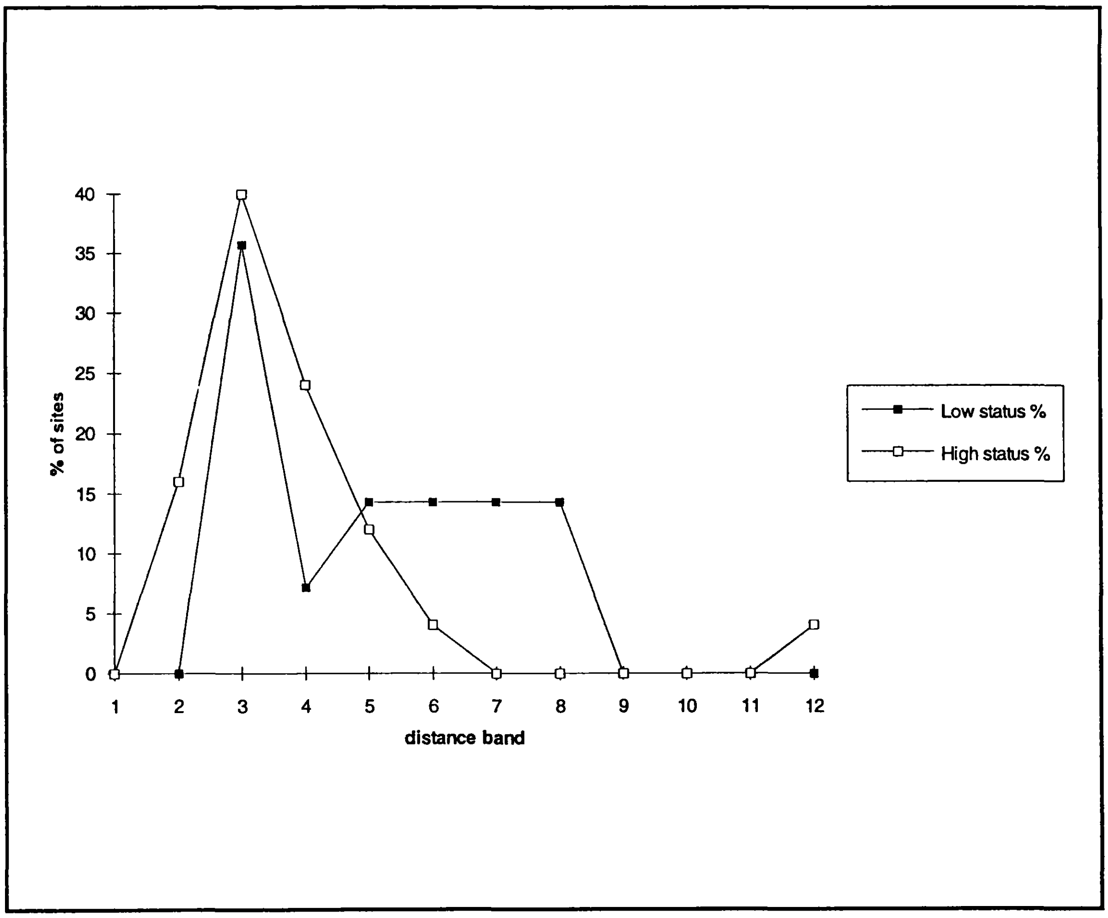

Figure 4.3 The percentage of high and low status sites dating to the Republic on 1.5 km eq cost distance bands from the Roman towns

Figure 4.3 shows the percentages of low-status and high-status sites dating from the Republic. As it can be seen from the graph, high-status sites in the Republic were much closer to towns than the low-status sites and, while the percentage of high-status sites drops dramatically from band 4 onwards, the low-status sites seem to have had a rather dispersed distribution in the countryside until their percentage drops to 0 in band 9. The patterns of distribution of high-status and low-status sites in the Republic are rather different, with high-status sites being clustered around the towns while low-status sites are scattered around the countryside and do not occur at all close to towns. Why the towns should be attractive to the high-status sites but not the low-status ones at this time is not clear. It is possible that wealthier colonists had first choice of were to settle, so that they built their (high-status) rural sites closer to the towns to have easier access to resources, leaving the poorer colonists to settle in areas disadvantaged for access to towns. On the other hand, the opposite argument can be brought forward, that is, the sites which were founded closer to the towns received more diagnostic ceramic material, are more visible archaeologically and were classified as high-status ones as a consequence. Variation in assemblages caused by varying pottery supply at different times has been discussed by Marsh (1981) for the Gaulish Terra Sigillata pottery found in London. Fentress and Perkins (1987) make the same point about the production of African Red Slip and its distribution in the Western Mediterranean. Despite the fact that these authors deal with variation in time, it can be argued that variation can also occur in space.

> ...does the lack of samian indicate the lack of occupation, lack of supply, or both?
>
> (Marsh 1981, 214)

During the Republic, very few Roman sites were occupied, compared to the following periods, and it is likely that at this time a redistribution system had not yet completely developed, so that the sites closer to the towns would have been more likely to receive diagnostic pottery than those further away. The majority of the Roman towns found in the Guadalquivir Valley were continuation of earlier Iberian urban settlements, which also acted as redistribution centres for the Iberian sites in their countryside, by exchanging imported foreign pottery for agricultural surplus and metal products (Keay 1992, 285). It is possible that the towns retained this redistribution function under the

Romans, through markets and fairs.

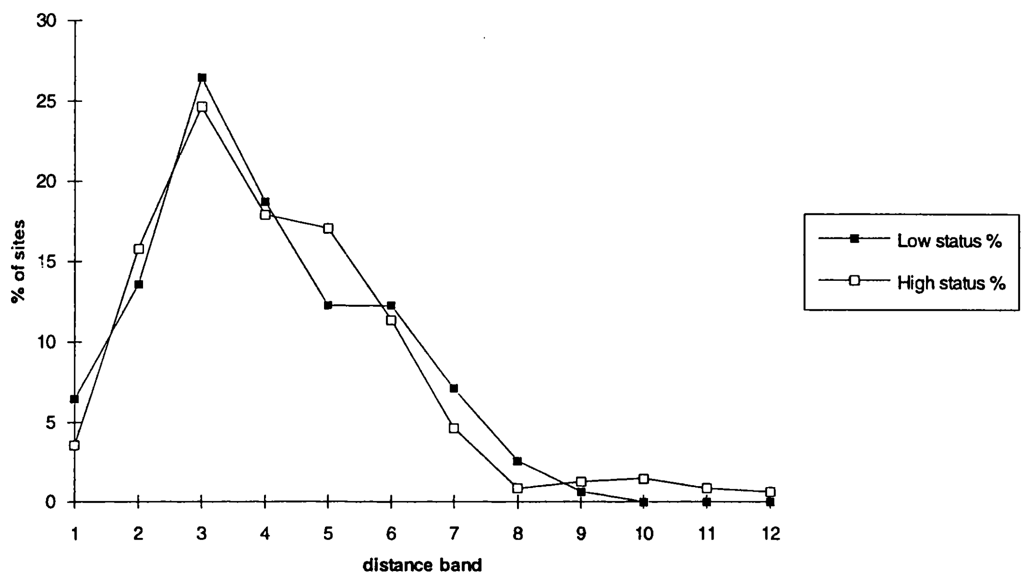

Figure 4.4 The percentage of low and high-status sites dating to the Early Empire on 1.5 km eq cost distance bands from the Roman towns

Figure 4.4 shows the percentages of low and high-status sites dating from the Early Empire. The situation is different from that of the Republic, the two distributions are similar and it appears that while the total number of rural sites (both low and high-status) increased, new low-status sites were being created near the towns, where there were not any before. Apart from the largely increased number of rural sites in occupation compared to the Republic, the distribution of low-status sites is now very similar to that of the high-status ones, concentrated around the towns, rather than dispersed in the countryside. The trend in the settlement pattern seen in the Republic changes in the Early Empire and it appears that a more extensive rural site distribution is achieved throughout the Guadalquivir Valley. This development corresponds with the creation of several *coloniae* (Keay 1992, 297-299) in the valley, with the consequent redistribution of the land formerly owned by the Iberians to the Roman settlers. The extensive Roman presence throughout the valley indicates that by the end of the Early Empire the Roman landscape dominated over the Iberian one, though it would be wrong to assume that the Iberian system had been completely replaced. The majority of the rural sites which came into existence in the Early Empire had enough material wealth to cause them to be classified as high-status ones, reflecting the economic effort Rome put into the reorganisation of the Guadalquivir Valley. According to Ruiz Rodríguez *et al.* (1991, 35), the late 1st century BC is the time of the first appearance of the dispersed settlement focused upon sites which can be identified with `villae' (see also Hornos *et al* 1985a; 1985b). The centralisation of low-status sites (`farms') around the high-status ones (`villae') noticed by these authors is reflected in the shape of the graph, in which the two curves are similar, indicating that similar proportions of high and low-status sites are found across the countryside, while this is not the case during the preceding period (see figure 4.3) and the following one (see figure 4.5).

Figure 4.5 shows the percentages of low and high-status sites dating from the Late Empire. The most noticeable feature of this graph is the fact that the distribution of high-status sites is bimodal, showing a net increase in the high-status sites away from towns, while other high-status sites are still clustered around the towns. This is consistent with the creation of larger estates in the open countryside, while the high-status sites whose welfare depended on the towns were still in existence. Overall, the total number of sites dating from the Late Empire is lower than the total number of sites dating from the Early Empire, therefore rather than the creation of new sites, the graph reflects the disappearance of those sites which were not close enough to the towns to survive or which did not manage to evolve into larger estates. It is also possible that this pattern is again influenced by the supply of diagnostic material to the sites further away from the towns, as discussed for the distribution of low-status sites dating to the Republic. If the administrative role and importance of the towns declined in the Late Empire, it is possible that their function as redistribution centres of pottery was lost, so that the sites received diagnostic material directly, rather than going through the markets and the towns.

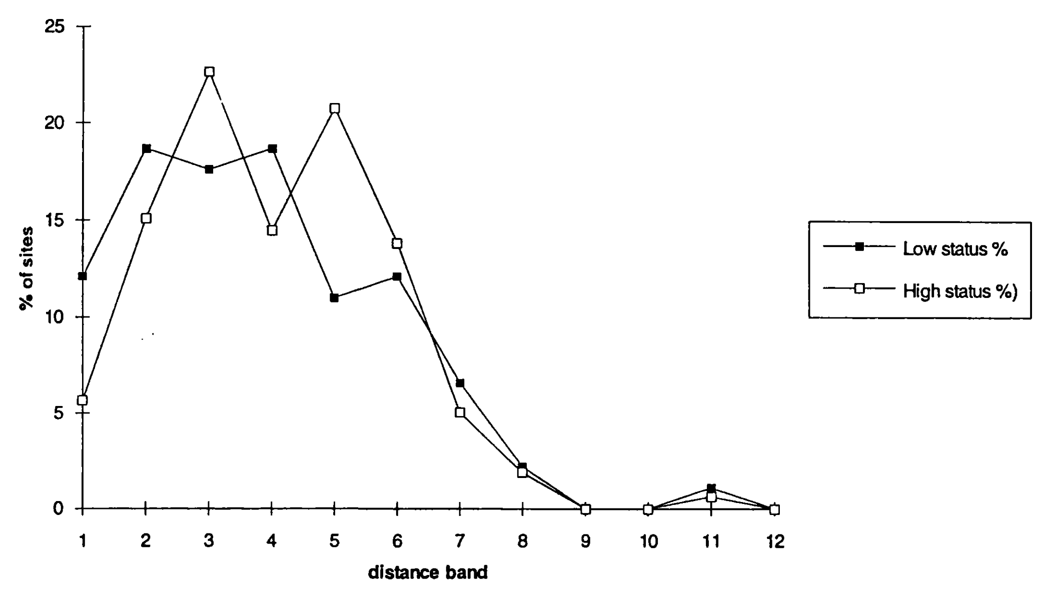

Figure 4.5 The percentage of low and high-status sites dating to the Late Empire on 1.5 km eq cost distance bands from the Roman towns

4.4 The distance from navigable rivers

One of the main characteristics of the Guadalquivir Valley in Roman times was the presence of two major waterways, the Guadalquivir and the Genil, which had been adapted for river transport under Augustus (Escacena Carrasco and Padilla Monge 1992, 83; Chic García 1978). It was possible for smaller boats to move between Cordoba and Hispalis (Seville) on the Guadalquivir (Strabo III, 2,3), while west of Seville the river was suitable for larger vessels, so that Hispalis was actually considered a sea port (Silius Italicus, book III, 390). The river Genil, connected to the Guadalquivir, was navigable up to Astigi, modern Écija (Pliny, Naturalis Historia, book III, 2, 10). The two navigable rivers were used as starting points to create a cost distance surface to test the relationship of sites to the waterways of the Guadalquivir Valley. Again, the relationship was tested by means of the one-sample Kolmogorov-Smirnov test run on the low and high-status sites and the three chronological bands separately.

All the relationships tested significant, with the exception of both the low-status and the high-status site distribution dating from the Republic. This is hardly surprising, as the navigable rivers were the main route into the Guadalquivir Valley and the main towns, which have been proved to influence the location of rural sites (section 4.3), were located in relation to the rivers. The different distribution of the Republican sites with regards to the waterways reflect the dispersed pattern in the countryside already discussed.

Except for the case of the Republican low-status sites, all the other settlement distributions (high-status sites in the Republic, Early Empire and Late Empire and low-status sites in the Early and Late Empire) show concentrations of sites close to the navigable rivers. In the case of the high-status sites, the percentage of sites near the waterways is higher in the Republic and in the Late Empire than in the Early Empire, despite the fact that the test was not significant. This pattern possibly reflects the even spread of high-status rural sites throughout the countryside in the time of maximum economic exploitation of the area, when land further away from the main communication routes was used to maximise production for trade, while when the number of sites in existence was smaller, such as immediately after the beginning of the Roman rule during the Republic or after the end of the period of economic growth in the Late Empire, the sites in existence tend to be closer to the communication routes. It is possible that the sites which were further away from the waterways could not survive after the 3rd century AD, when a generalised decline of trade and development occurred throughout the Roman Empire. A similar pattern is visible in the distribution of low-status sites. Leaving aside the sites dating from the Republic, the percentage of sites close to the waterways is much higher in the Late Empire than in the Early Empire, despite the fact that many more sites were in existence in the Early than in the Late Empire. Again, since the majority of the kilns in the area were located along the river, and since Terra Sigillata Chiara A and D were carried along the Guadalquivir (see section 5.4), it is possible that the sites closer to the Guadalquivir in the Late Empire were more visible than those further away from it, if we accept that the local exchange network had collapsed by this time.

4.5 The distance from the Roman roads

The influence of the Roman roads on the distribution of rural sites was also tested. The most important road in the Guadalquivir Valley was the Via Augusta, which had originated in the late 2nd century BC as the Via Herculea as a link from Rome to *Gallia Narbonensis*, *Hispania Citerior* and *Ulterior*. Under Augustus the road was enhanced and a arch to Janus Augustus was built where it enters *Baetica*, at Iliturgis (Keay 1992, 298) and had its name changed to Via Augusta. The road linked the most important centres of the region: Corduba, Astigi and Hispalis (Sillières 1976), emphasising their primary role in the province. The importance of these centres is also highlighted by the fact that the course of the Via Augusta diverged from that of the earlier road to go through them. The Via Augusta was the axis of a complex network of roads, which was based in part on the earlier Iberian road network and which was also developed under Augustus (Keay 1992, 298).

A cost distance surface was created having the known Roman roads in the area as starting points and the relationship between the roads and the Roman rural sites was again tested by means of the Kolmogorov-Smirnov one-sample test. The high and low-status sites distributions were tested separately and so were the site data sets divided into the three chronological bands.

The results of the tested showed that there is a significant relationship between the Roman roads and the rural sites, both of high and low-status, dating to the Early and Late Empire, but not those dating to the Republic. This result is not surprising, since most of the Roman roads were not in existence until the Early Empire, when an extensive road network was created to connect the new colonies which had been founded by Caesar and Augustus (Keay 1992, 297-299).

The pattern of site distribution on distance from the Roman roads reflects that of the site distribution on distance from the navigable rivers discussed above. Despite the fact that many more sites were in existence during the Early Empire than the Late Empire, the percentage of both high and low-status sites near the Roman roads dating from the Late Empire is much higher than the percentage of those dating from the Early Empire. There appears to be a trend for Late Imperial sites to be located very close to communication routes, be they the navigable rivers or the Roman roads. It makes little sense to talk about the Republican site distribution in relation to the Roman roads because almost all of these were not built until the Early Empire.

4.6 Conclusions

To make the discussion of the results of the tests outlined in the preceding sections easier, a summary of these is available in tabular format in tables 4.3 for the low-status sites and 4.4 for the high-status sites.

The results of the tests outlined above suggest the location of rural sites was influenced more by socio-economic landscape factors, rather than just the needs of agricultural production. The Romans might have preferred to settle in the best agricultural areas when they first started to colonise the valley, but they might not have had the choice because such land was being used by the Iberian population at the time. With the creation of an extensive network of Roman or Romanised towns (the *coloniae* and the *municipia*), linked together by Roman roads in the Early Empire, it became more and more important to be part of the system and have easy access to the communication routes and the towns which were acting as economic and political as well as cultural centres. Almost all the Roman towns in the Guadalquivir Valley were built by, or on top of, earlier Iberian settlement and indeed the Romanisation of the area seems to have been more of an integration of the Roman settlers than a take-over, as it has already been argued by Keay (1992). The good performance of high-status sites away from towns in the Late Empire might also be related to a loss of importance and power of the towns themselves, so that large estates could do well without being located very close to towns. It can also be argued that being too close to a town would have hampered the development of large estates because there would have been much more competition for space. The argument has also been brought forward that different sites at different times might have been more visible archaeologically because of the supply of diagnostic material used to date them and to identify them on the ground. If this hypothesis is true, the direct consequence is that more sites were in use during time of low pottery supply than it is possible to see nowadays. This is true whether the supply varied because of differences in volume of pottery production (see Marsh 1981 and Fentress and Perkins 1987), or because the pottery could not get to sites further away from transportation routes (the roads and the navigable rivers) or the redistribution centres (the towns). It is also possible that a combination of the two factors was at play, that is, fewer sites were in existence during the Republic and the Late Empire than in the Early Empire, and they were receiving less pottery because the large scale exchange network was still not fully developed (in the Republic) or had already collapsed (in the Late Empire). In all cases, however, the different amounts of diagnostic pottery found in different sites at different times indicates that the sites which were receiving ceramic material were more important or better located than the ones which are not visible archaeologically. If a site was not receiving diagnostic ceramic in the Late Empire because it was too far from the navigable rivers, it is evident that that site was disadvataged compared to those which were located within easy access to the communication routes. It is no surprise that the new sites which come into existence in the Late Empire (or whose presence was not recorded in the preceeding periods) are all located along the banks of the Guadalquivir, which was the main communication route in the Guadalquivir Valley. Interestingly, this fact also suggests that at this time the navigable rivers were more important than the Via Augusta as communication routes.

TEST NAME	ALL Chi2	ALL KS 5	ALL KS 1	REP Chi2	REP KS 5	REP KS 1	EE Chi2	EE KS 5	EE KS 1	LE Chi2	LE KS 5	LE KS 1
SOILS	S			N			S			S		
AGRICULTURAL POTENTIAL		S	S		N	N		N	N		N	N
DISTANCE FROM TOWNS		S	S		N	N		S	S		S	S
DISTANCE FROM RIVERS		S	S		N	N		S	N		S	S
DISTANCE FROM ROMAN ROADS		S	S		N	N		S	S		S	S

Table 4.3 Summary of the results obtained in the tests on the relationship between low-status sites and various landscape factors.

TEST NAME	ALL Chi2	ALL KS 5	ALL KS 1	REP Chi2	REP KS 5	REP KS 1	EE Chi2	EE KS 5	EE KS 1	LE Chi2	LE KS 5	LE KS 1
SOILS	S			N			S			S		
AGRICULTURAL POTENTIAL		S	N		N	N		N	N		N	N
DISTANCE FROM TOWNS		S	S		N	N		S	S		S	S
DISTANCE FROM RIVERS		S	S		N	N		S	S		S	S
DISTANCE FROM ROMAN ROADS		S	S		N	N		S	S		S	S

Table 4.4 Summary of the results obtained in the tests on the relationship between high-status sites and various landscape factors.

Key to the tables: Chi2 = χ^2 at the 5% significance level, KS 1= Kolmogorov-Smirnov one-sample test at the 1% level of significance, KS 5 = Kolmogorov-Smirnov one-sample test at the 5% level of significance, ALL = Sites from all periods, REP = sites which existed during the Republic, EE = sites which existed during the Early Empire, LE = sites which existed during the Late Empire, S = significant, N = not significant, blank = test not carried out

5 The analysis of the trade patterns

Another area of interest in the study of Roman settlement and society in the Guadalquivir Valley is the pattern of movement of goods throughout the region. There are serious problems with this type of study because the data are not quantified, so that where a certain material is reported as present, the actual amount of that material is not known. A site flagged as containing Black Glaze pottery may have had only a small sherd, or a whole deposit of pots. Given this limitation, traditional approaches based on the study of the sherd count, weight or the vessel equivalent measure (Orton *et al.* 1993, 21-22, 166-171; Orton 1993, 169-174) are not possible. As for the site data used in the analysis of the relationship between site location and landscape features described above, the only information available is whether a certain type of material was found at a specific site or not. Despite the fact that a quantitative study of the distribution of different materials can not be carried out, it is still possible to do a qualitative analysis of the data. If a certain type of pottery was imported in larger quantities in some areas rather than others, it can be expected that where that pottery type was more common, a higher proportion of the sites would be flagged as containing it than in the areas where that pottery type was not a preferred import item. The problem of defining areas in which to group the sites was approached by calculating the extent of the territories of influence of the Roman towns in the valley.

5.1 The territories of the Roman towns

Before a study of the pottery distribution pattern around individual towns is carried out, it is necessary to determine the size and shape of the region which might have been influenced (or controlled?) by the town. Frayn (1993, 77) argues that rural people would want to live at such a distance from the nearest market town so that they could travel to the market and back in the same day. She also (1993, 77) mentions how Varro (i. 16. 5) stresses that people went back home to the countryside in the evening of the same day they went to the market and would not spend the night in the town. Rural sites had therefore to be located that it was possible to go to the market and back in a day. Products which had to be sold fresh, such as fresh cut flowers, could only be produced within a radius of no more than 5 km from the town, or they would be ruined by the time they reached the market. Other products which would withstand a journey, such as animals and cheese, could come from further away, but in no case from more than 10-15 km away. Frayn refers to markets in Roman Italy, but her considerations are very general and are also valid for the Spanish situation.

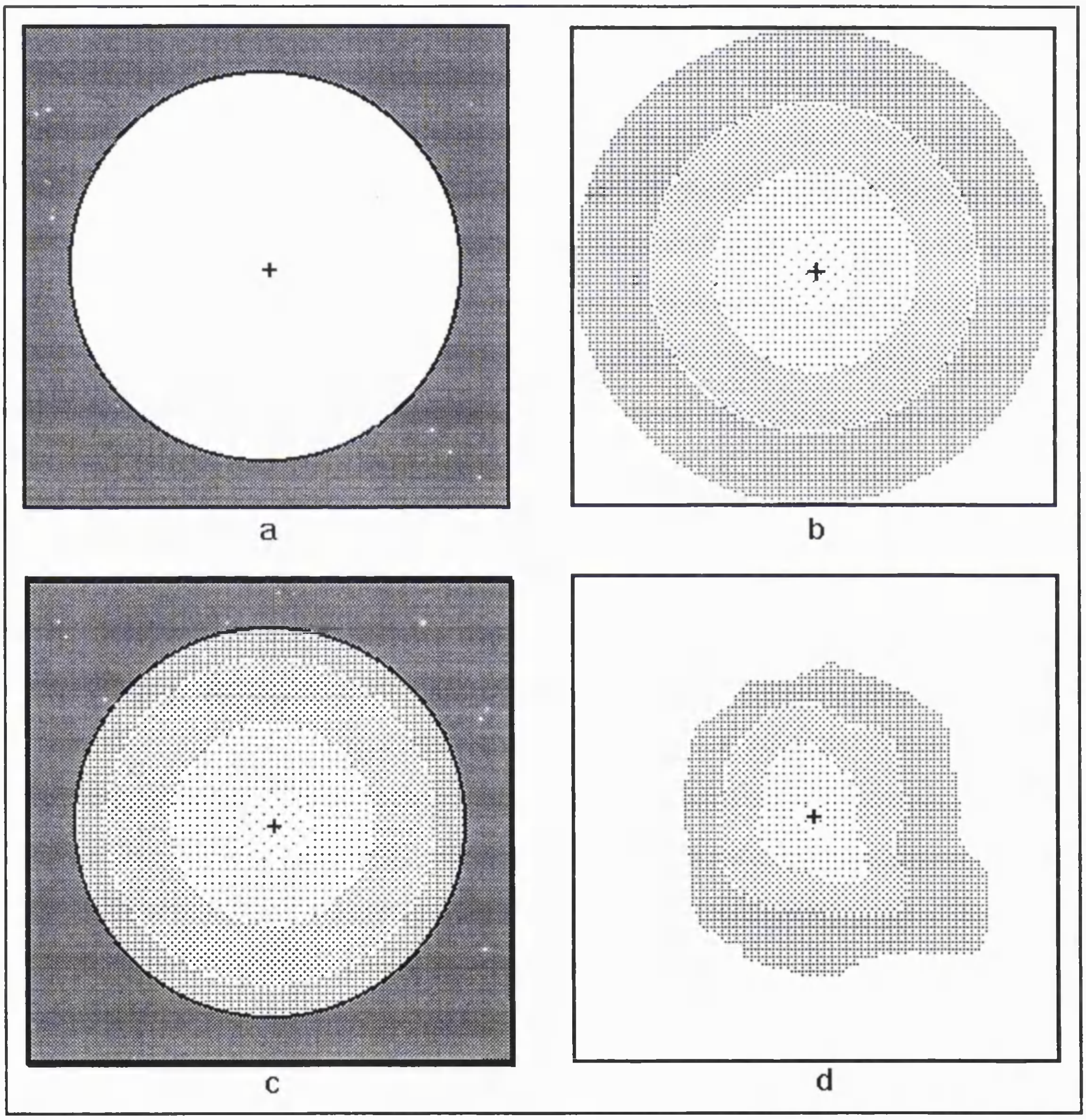

Figure 4.6 a) The linear distance binary mask corresponding to 15 km around an ideal town. b) The cost distance calculated with a constant friction surface of 1.0 around an ideal town. c) The 15 km binary mask is superimposed on the cost distance and the cost limit value is measured. d) The real cost surface around a real Roman town is reclassified to be no higher than the cost limit value. This defines a town territory corresponding to a catchment area of 15 km equivalent units.

It is possible to go one step further along these lines and try to compensate for the influence of the land form on the ease of travel along certain routes. The nature and utility of cost distance surfaces have already been discussed (see chapter 1, *Introduction*, section 3.5). The problem with cost surfaces is how to work out what is the `cost equivalent' to a given linear distance, which is a necessary prerequisite to apply them to archaeological research of this sort. When a cost surface is calculated, the resulting `distance' is no longer a linear distance as usually employed in archaeological studies, but a new measurement which expresses the cost involved in moving from one point to another over certain types of constraints. Expressed in this way, it is quite difficult to see what this `distance' actually means, also considering that majority of the other authors who have been working in location analysis have always used known and easily understandable units of distance such as km or miles.

The cost surface had to be reduced to something that would be easier to adapt to territorial studies as they are known to the wider archaeological world before it could be used to create the map of town territories. The value of the cost distance units corresponding to moving linearly over 15 km was calculated by using ideal conditions as described below.

First, an Idrisi image the same size and resolution as the one containing the Roman towns was created, but with only one point in the centre, representing an hypothetical, ideally located, town. This was used to create a linear distance image using the ***distance*** module. The linear distance image was then used to create a mask of linear distance of 15 km from the starting point, in which the area inside the 15 km band was classified as 0 and everything else as 1000 (see figure 4.6a). The same image with the hypothetical town was used to create a cost distance surface using the ***cost*** module. The friction surface used to create the cost surface had an identical value of 1.0 in each cell. In a friction surface, a value of 1.0 is the base cost of moving through a cell, therefore a cell with a value of 2.0 indicates that the cost of moving through that cell is twice the base cost while a cell with a value of 0.5 indicates half the base cost (Eastman 1992, 39). The resulting cost distance surface represented the cost of moving through an ideal area of identical friction in every point (figure 4.6b). The 15 km binary mask was then superimposed upon the cost distance surface, masking out everything which was at more than 15 km linear distance from the hypothetical town (figure 4.6c). The masking value, 1000, was then reclassified to 0, leaving a cost surface with equal values in each direction and corresponding to the cost of moving 15 km over ground with equal friction of 1.0. The highest value in this image (*ie* the one occurring at the border of the reclassified images) was noted and will be referred to as `cost limit value'.

The next step was to create a real cost distance surface starting from the Roman towns and using a real friction surface derived from the slope map, then reclassify it so that any value higher than the cost limit value would be reclassified as 0[5]. The result was a cost surface image with cost distance catchments around towns equivalent to the cost of moving over 15 km (see figure 4.6d). Since the units used in the cost surface are not linear measures, they can be referred to as 1 km equivalent units, so we can say that every location in the town catchments thus derived lies within 15 km equivalent units (km eq) from the nearest town. Where two towns were too close together to have 15 km eq catchments, the mid point between the two cost catchments was used to define the border between the two territories. This is effectively a variation of Thiessen polygons (Burrough 1986, 148-149) taking into account the land form and imposing a maximum distance for the size of the catchment. Some of the catchments thus derived did not contain any rural sites, therefore these were excluded from the analysis. The most noticeable exclusion is the territory of the Roman town of Hispalis (modern Seville), due to a combination of a very small territory because of the closeness to Italica and Osset and intensive development in modern times. This division of the territories also assumes that the towns were all equally important and that they were so at the same time. In reality, it is likely that the importance of the various centres varied in time and that there was a considerable overlap in the area of influence of neighbouring towns, but at present, because of the limitations of the software, it is almost impossible to include this sort of considerations in the analysis.

5.2 The Correspondence Analysis

Having created a theoretically plausible approximation to the ancient town

[5] Since the reclassification module within Idrisi, ***reclass***, does not allow the output of images containing real data, the freeware ***reclassr*** module, created by Eric J. Lorup, which does not have this limitation was used.

territories, the number of rural sites containing specific pottery types occurring in each town territory was counted. This process produced a table containing the information about the number of locations within each town territory where each pottery type in turn was found. This method does not give a measure of the volume of a certain pottery type present within a territory, but rather an indication of which pottery type was more common in the area. These data were particularly suitable to be studied by means of Correspondence Analysis. Correspondence Analysis is a technique of multivariate analysis which reduces large numbers of variables to two or three `mean' variables, called *components*, which are much easier than the original variables to study in terms of how they influence the variation in the data sets. The components have a percentage of inertia associated with them, which identifies what percentage of the total variation is `explained' by each component. Correspondence Analysis can be regarded as a form of Principal Components Analysis appropriate for discrete rather than continuous data, but it has the advantage over Principal Components Analysis that the First Component is not heavily influenced by the columns or rows containing higher scores than the rest in the raw data table. Correspondence Analysis also has the characteristics of the biplot, in that both the rows and the columns of the data matrix (*ie* the variables and the objects) can be represented as points on the same plot, so that the relationship between certain variables and objects can be assessed visually. The main disadvantage of Correspondence Analysis lies in the fact that it is only suitable for tables in which the data consist of counts or binary (presence/absence, see chapters 5 and 6) data, while Principal Components Analysis can be used with mixed data (see Baxter 1994, 65-6). Because of the complexity of the subject, some background reading is strongly reccomended. Shennan (1988, 241-297) discusses various techniques of multivariate analysis, but concentrates on Principal Components Analysis,rather than Correspondence Analysis. Baxter (1994, 100-139) gives an in-depth discussion of Correspondence Analysis and presents archaeological examples. Examples of the application of Correspondence Analysis in archaeology can be found in Madsen 1988, though the case studies all use prehistoric data.

The percentage of inertia and other information on the first three components is shown below. The cummulative percentage of inertia on the firts two components is over 74%, which is quite a good proportion of the total of variation, and it increases to over 88% when the third component is considered. The table of component loadings for the objects and the variables can be found in Appendix B.

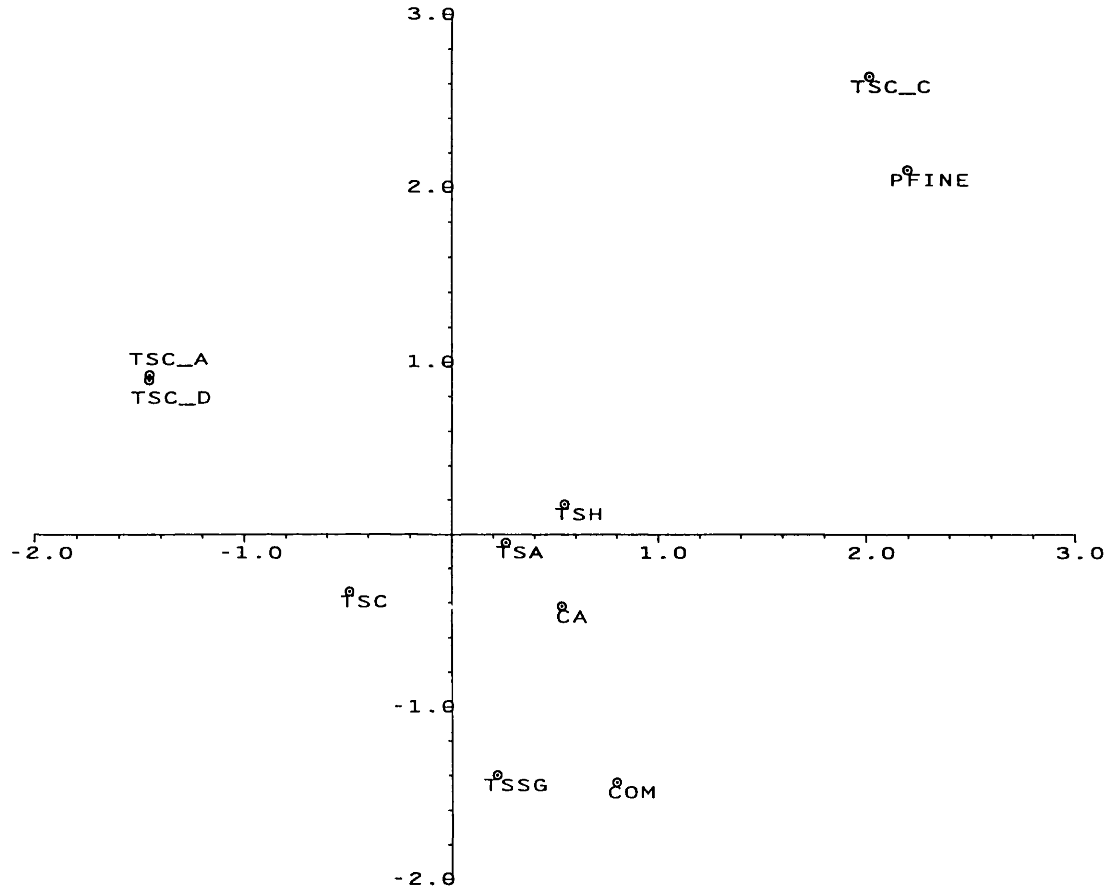

Figure 4.7 The Correspondence Analysis variable plot (pottery types) on Components I (horizontal) and II (vertical).

Component	Iterations	Norm	Eigenvalue	% Inertia	Cummulative
I	13	0.050	0.294260	53.5	53.5
II	30	0.059	0.113779	20.7	74.2
III	11	0.035	0.078886	14.3	88.5

Table 4.5 The Correspondence Analysis information for the data set of count of sites with specific pottery types in the territories of the Roman towns

The plots of the variables and objects components calculated with the Correspondence Analysis are shown in figures 4.7 and 4.8. The plot of the variables on the first (I) and second components (II) shows the pottery types divided into three groups: at one extreme of Component I are Terra Sigillata Chiara A (TSC_A) and Terra Sigillata Chiara D (TSC_D), at the other extreme are Terra Sigillata Chiara C (TSC_C) and Thin Walled ware (PFINE), while all the other wares occur roughly in the centre of the plot. Terra Sigillata Chiara A and Terra Sigillata Chiara D are very close to each other in the plot, while Thin-Walled ware and Terra Sigillata Chiara C are at the opposite extreme of Component I. The closeness of Terra Sigillata Chiara A and D (the two data points are almost superimposed) suggests that these two pottery types are distributed in a very similar way across the Guadalquivir Valley, and that where one type occurs, the other is likely to occur as well. At the opposite extreme of Component I are Terra Sigillata Chiara C (TSC_C) and Thin Walled ware (PFINE). Component I appears to explain the variation between these two groups of pottery. While it is plausible that Thin-Walled ware could be distributed differently from Terra Sigillata Chiara A and D, as it is a completely separate type of ware, it is striking that Terra Sigillata Chiara C would have a similar distribution to Thin-Walled ware and be totally different from Terra Sigillata Chiara A and D, which are subtypes of the same ceramic type and are located chronologically just before and after it, with a good deal of chronological overlap. Generic Terra Sigillata Chiara (*ie* Terra Sigillata Chiara which was not further identified as belonging to a specific subgroup) is on the edge of the central group in the Correpondence Analysis plot and leans towards the group containing Terra Sigillata Chiara A and D, well away from Terra Sigillata Chiara C. This might indicate that the Terra Sigillata not identified behaves much more like the two subgroups A and D rather than subtype C. The three subtypes of Terra Sigillata Chiara were all produced in modern Tunisia, but while subtypes A and D were produced in the workshops in northern Tunisia, in the region of Carthage (Hayes 1972, 298; 1980, 518), subtype C was produced in the workshops of central Tunisia (Hayes 1972, 297; Peacock *et al.* 1990, 83). We know that Terra Sigillata Chiara A and D were very probably exported from northern Tunisia from the port of Carthage (Hayes 1980, 518), but there is no evidence to suggest that subtype C was exported from Carthage as well. If subtype C was imported into the Guadalquivir Valley from its place of production following different routes from subtypes A and D, it might have been traded to different places. This situation would explain the different distribution of subtype C as opposed to subtypes A and D.

The fact that the distribution and frequence of Terra Sigillata Chiara C differ from those of Terra Sigillata Chiara A and D has been noticed by other people working with data from the Guadalquivir Valley. Escacena Carrasco and Padilla Monge (1992) carried out a study in the area immediately surrounding Seville. They calculated the percentage of rural sites within the area they surveyed containing any of the three subtypes of Terra Sigillata Chiara. They only used the sites yielding Terra Sigillata Chiara in their study, that is, the sites where only pottery types other than Terra Sigillata Chiara were found were excluded from the calculation of the percentages. In the area they surveyed the percentage of sites containing Terra Sigillata Chiara C is 14%, while the Terra Sigillata Chiara A, which is earlier, is found in 32.5% of the sites and the Terra Sigillata Chiara D, which is later than the Terra Sigillata Chiara C, is found in 53.5% of the sites (these percentages are calculated out of the number of sites which contain one or more subtypes of Terra Sigillata Chiara, not out of the total number of sites in the database). Taken at face value, these percentages would indicate that the sites without Terra Sigillata Chiara suffered a hiatus in occupation in the second half of the 3rd century AD and were then reoccupied at the early 4th century AD. On the other hand, there is evidence from other parts of the western Mediterranean, such as *Turris Lisbonis* (Cerdeña, see Villedieu 1986, 154) and Sperlonga (Campania), that the relative percentage of Terra Sigillata Chiara C to Terra Sigillata Chiara A and Terra Sigillata Chiara D tends to be rather low (Escacena Carrasco and Padilla Monge 1992, 14-15). When we compare these results with the percentage of sites in the whole Guadalquivir Valley, it is clear that the trend observed by Escacena Carrasco and Padilla Monge is reflected in the total distribution. In the whole of the Guadalquivir Valley, the sites with Terra Sigillata Chiara A are 150, those with Terra Sigillata Chiara D are 261, while only 28 sites contain Terra Sigillata Chiara C. The number of sites which contain unidentified Terra Sigillata Chiara is 432. In percentage, Terra Sigillata Chiara A is found in 34.2% of the sites containing Terra Sigillata Chiara, subtype C in 6.3% and subtype D in 59.5%. These percentages were calculated out of the number of sites which contained any of the three subtypes of Terra Sigillata Chiara, rather than out of the total number of sites known from the Guadalquivir Valley, to be able to compare them with Escacena Carrasco and Padilla Monge's results. The percentages for the whole Guadalquivir Valley are very similar to those of Escacena Carrasco and Padilla Monge, therefore their argument that the relative scarcity of Terra Sigillata Chiara C compared to subtypes A and D is due to variation in supply, rather than to a hiatus in occupation, can be extended to cover the whole of the Guadalquivir Valley.

Another interesting feature in the variables plot is the fact that Black Glaze Pottery (CA) and Arretine Terra Sigillata (TSA), which are similar in date, cluster together but so also does Terra Sigillata Hispanica (TSH) which was produced locally and was in use until the 3rd century AD. The workshop at Andujar which produced most of the Terra Sigillata Hispanica found in the province of Seville disappeared by the beginning of the 3rd century AD (Fernández Fonseca 1995, 64) and the little later Terra Sigillata Hispanica found was produced in the workshop of Tritium Magallum (La Rioja). The Terra Sigillata Hispanica production was influenced by both Terra Sigillata Aretina and South Gaulish Terra Sigillata (Mayet 1984, 16), but South Gaulish Terra Sigillata influenced the southern production, while Arretine Terra Sigillata influenced the northern production (Fernández Fonseca 1995, 62; Romero Carnicero 1983, 132). South Gaulish Terra Sigillata (TSSG) appears to be close in the plot to coarse ware (COM), however, coarse ware is very evenly distributed throughout the valley and can not be used for dating purposes. As a test of the importance of coarse ware, the Correspondence Analysis was carried out without including it in the pottery types count. The result was that all the other types mantained the same relative position in the plot and so did the towns in the object plot, therefore it can be concluded that coarse ware was so common and so evenly distributed that it can not give any valuable information on pottery distribution and trade patterns.

The object plot (figure 4.8) shows a group of towns in the region of the plot corresponding to the Terra Sigillata Chiara A and D. In a Correspondence Analysis, when one or more variables are closely related to one or more cases, it is common to say that those variables `explain' those cases. In this study, Terra Sigillata Chiara A and D `explain' the towns present in the same area of the plot, *ie* these two pottery types have a much stronger influence in the assemblages of the rural sites in the territories of those towns than the other pottery types do. Similarly, the towns which occur at the other extreme of Component I are `explained' by Thin-Walled ware and Terra Sigillata Chiara

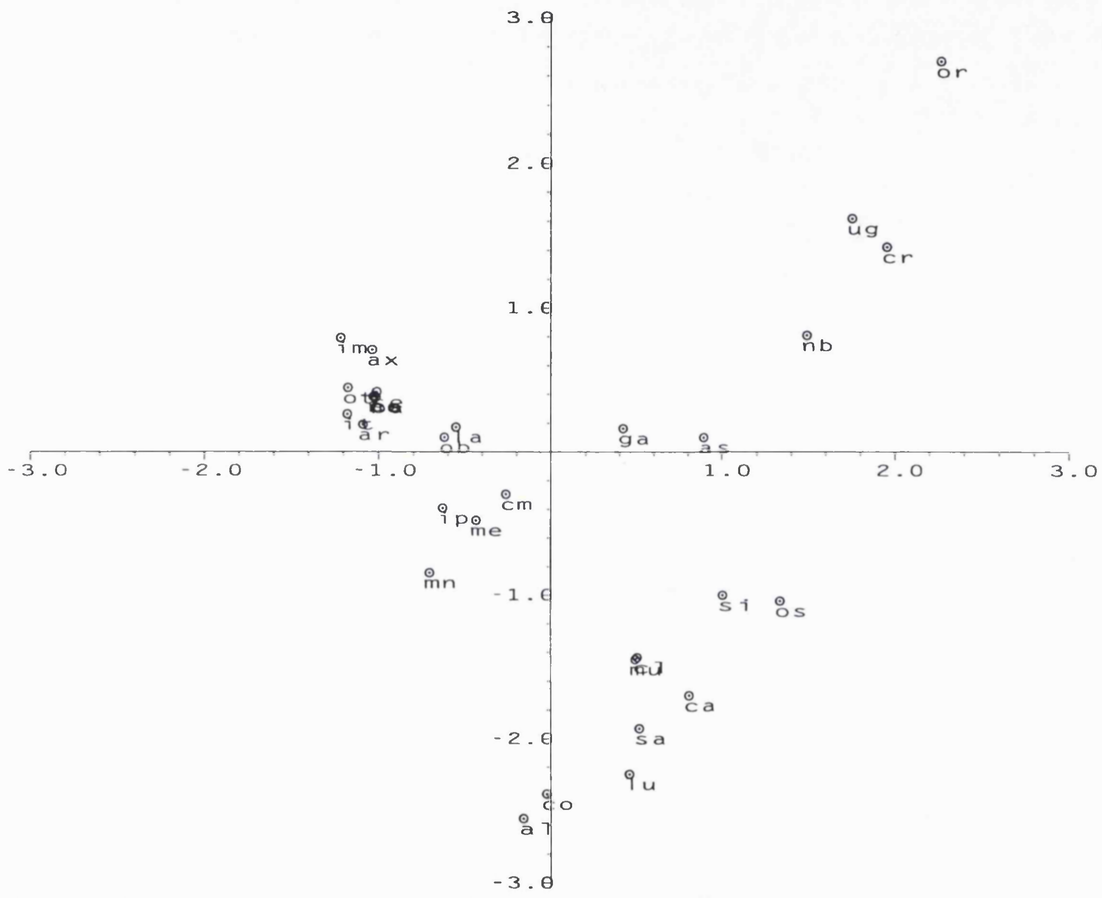

Figure 4.8 The Correspondence Analysis object plot (towns) on Components I (horizontal) and II (vertical).

C. The trend seems to be one of self-exclusion, where Terra Sigillata Chiara A and D are present, Thin-Walled ware and Terra Sigillata Chiara C are absent and the other way round. This might suggest that there were areas where some wares were easily available while others were not. Some of the towns occur loosely scattered in the centre of the plot, where Black-Glazed pottery, Arretine Terra Sigillata, Terra Sigillata Hispanica and South-Gaulish Terra Sigillata are. These towns do not seem to have been heavily influenced by one specific pottery type, but rather seem to have a combination of different types quite evenly distributed in similar proportions in their territories.

At this stage it became important to investigate whether there are similarities in the geographical position of the towns with similar catchment assemblages according to the Correspondence Analysis. The towns were therefore grouped together in five clusters (numbered from 1 to 5) in the object plot (figure 4.9). These groups were defined arbitrarily according to the position of the cases in the plot and it would have been possible to have more groups with a stricter definition of group membership, or fewer groups with a more relaxed approach to maximum distance between cases in the plot. The groups were only created to make interpretation of the data easier, but the limitations and arbitrary nature of group membership were kept in mind during the analysis of the spatial characteristics of the territories occurring in each group. To make group membership recognition easier in the Idrisi maps, the territories of the towns in each of the 5 clusters were differentiated by giving them a different colour (figure 4.10). Some spatial patterns are immediately apparent in the group map and are discussed in detail below.

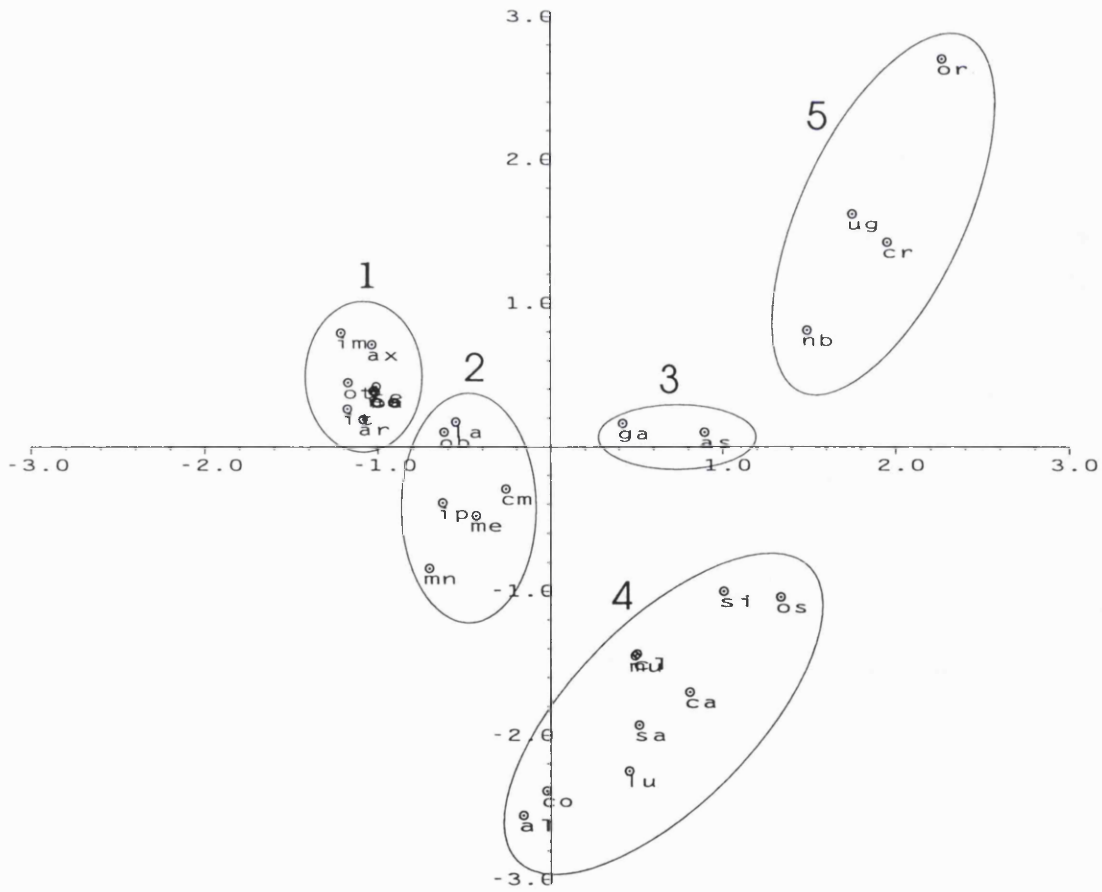

Figure 4.9 The towns in the object plot grouped together to study the spatial characteristics of territories with similar assemblages.

The object plot shows a concentration of cases in the position corresponding in the variable plot to the Terra Sigillata Chiara A and Terra Sigillata Chiara D pottery types (group 1). When the towns in this group are checked against their geographical position, the relative closeness of their territories to the river Guadalquivir is striking. It is immediately evident that almost all of them are either crossed by the river Guadalquivir itself, or were on major Roman roads which lead directly to Hispalis (Seville). The majority of the territories of the towns which correspond in the object plot to the position of Terra Sigillata Chiara (group 2) in the variable plot are found along the Via Augusta, the main communication route which connected south-west Spain to north-east Spain (going through the region of Tarragona and the Maresme), and to the north of the territories linked to the Terra Sigillata Chiara A and D. Terra Sigillata Chiara A and Terra Sigillata Chiara D appear to be dominant in the northern half of the study area, close to the communication routes (Roman road network and navigable rivers).

On the contrary, the territories explained by Thin Walled ware and Terra Sigillata Chiara C (group 5) tend to occur to the south-west of the region, or at least to

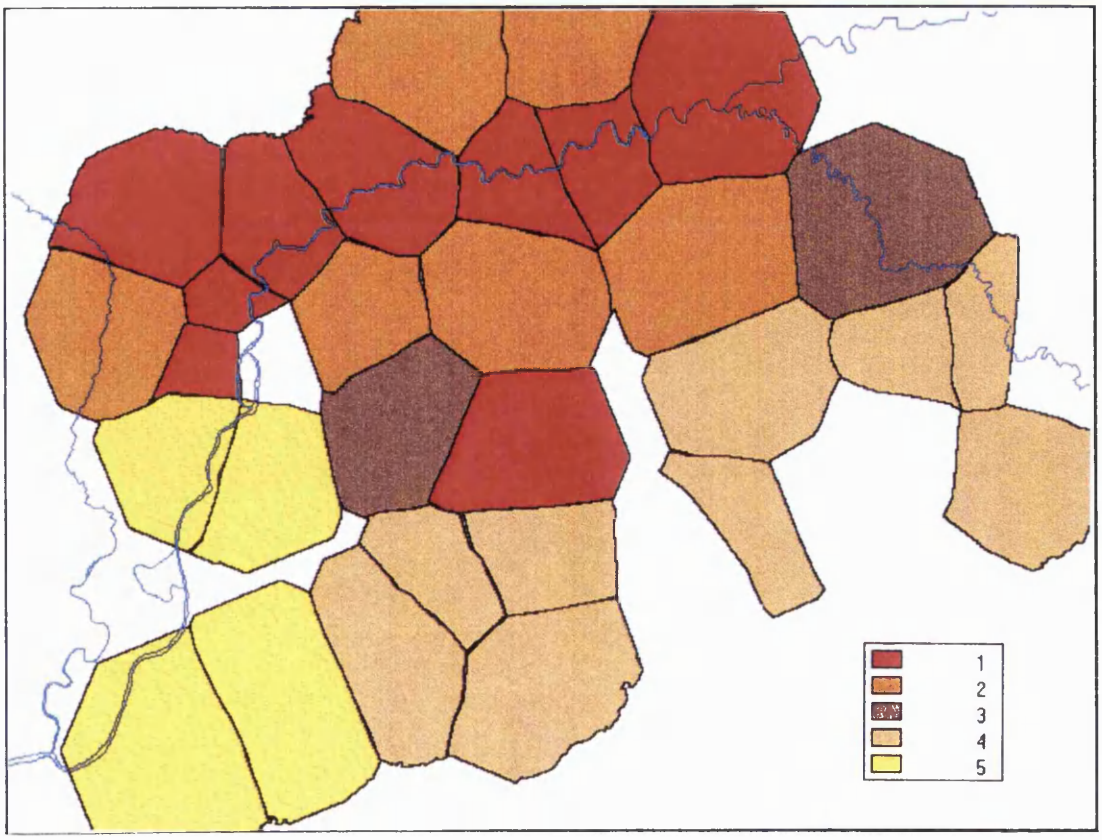

Figure 4.10 The town territories colour-coded to show the groupings in the object plot (figure 4.9). The position of the river Gudalquivir is also shown.

the south-west of the distribution of Terra Sigillata Chiara A and Terra Sigillata Chiara D. When the total of sites in each town territories with the wares of either of the two groups 1 and 5 is checked against their geographical positions, the two distributions clearly appear to be self-exclusive on the ground, as it is strongly suggested by the Correspondence Analysis plots, where Terra Sigillata Chiara A and D are present, Thin Walled ware and Terra Sigillata Chiara C are unlikely to be present and the other way round.

The other pottery types, Black Glaze pottery, Common (coarse ware) pottery, Arretine Terra Sigillata, Terra Sigillata Hispanica and South Gaulish Terra Sigillata[6] (groups 3 and 4) occur throughout the study area and do not seem to be limited to specific locations as it happens for the Terra Sigillata Chiara subgroups and Thin-Walled ware. The fact that the distribution of Arretine Terra Sigillata is similar to that of Black Glaze pottery, as can also be seen from the Correspondence Analysis variable plot, might suggest that Terra Sigillata Aretina was distributed along the same routes as Black Glaze pottery, which was imported in the Guadalquivir Valley before Romanisation and was redistributed along the Iberian exchange network.

5.3 Amphorae

The sites database also contains information about the presence/absence of amphora, though without specifying the type of amphora. Because of the fact that amphorae were containers of goods rather than goods themselves as the pottery discusses above, they were not included in the Correspondence Analysis. The main product of the Guadalquivir Valley was olive oil, which was being traded in Dressel 20 amphorae (Ponsich 1974; 1979; 1987; 1991). These amphore were produced in kilns located along the banks of the Guadalquivir and the Genil (Keay 1988, 98), which stringly suggests that they were being exported via the waterways. The high-status sites containing amphorae are found throughout the Guadalquivir Valley, while the low-status sites with amphorae are only found close to the Guadalquivir and the Genil. It can be assumed that the majority of the sites flagged as containing amphorae in the database actually contained Dressel 20, as this was the type produced in largest quantity in the

[6] South-Gaulish Samian ware

valley. The differential distribution of high and low-status sites containing amporae shows that when the effort was being made to transport these within the valley itself, rather than just using them for export, it was only the high-status sites which could afford to pay for them. This could be explained by the fact that the high-status sites were producing agricultural surplus for export and needed the containers to do so. Exporting high-status sites which did not produce their own amphorae were buying them from the kilns along the Guadalquivir. Since no low-status sites away from the rivers contained amphorae, it is possible that either low-status sites did not produce surplus for export, or their production was carried to high-status sites and exported together with the production of the high-status sites. This system would reflect the 3-tier Iberian system which was in existence before Romanisation and would act as a prerequisite for the development of large estates in the southern half of the Guadalquivir Valley in the Late Empire, in which centralised high-status sites controlled and marketed the production of dependent low-status sites, while the towns had lost their administrative and commercial role. Interestingly, the high-status sites away from the navigable rivers with amphorae all but one date to the Early Empire and ceased to exist in the Late Empire. There are not enough sites (high or low-status) with amphorae dating from the Republic to allow any general trends to be inferred.

5.4 Conclusions

The Correspondence Analysis described in this chapter was aimed at throwing some light on the way the different pottery types are distributed throughout the Guadalquivir Valley and what mechanisms might have shaped the present distribution.

Black Glaze pottery was imported into the region since the earlier 2nd century BC (Amores Carredano 1982, 245), well before Romanisation, and was redistributed along the Iberian exchange network, which was used by the Iberian aristocracy to obtain agricultural surplus and metal supply from their dependent sites in exchange for fine imported ware (Keay 1992, 285). The similar distribution of Arretine Terra Sigillata and Black Glaze pottery suggests that the Iberian exchange networks was still being used after the Romans first settled the area. Terra Sigillata Hispanica, which was a ware produced in workshops in Spain imitating the foreign forms, mainly Arretine Terra

Sigillata and South Gaulish Terra Sigillata, also has a very similar distribution to Black Glaze pottery and Arretine Terra Sigillata, showing that the original Iberian exchange network might still have been in use as late as the 2nd century AD while, at the same time, a different redistribution network was being used for Terra Sigillata Chiara A and then Terra Sigillata Chiara D, as shown by the relationship between these two types and the towns along the Guadalquivir.

In the Late Empire the only evidence we have comes from the Terra Sigillata Chiara D, which follows the Guadalquivir network, so that the first two centuries AD can be seen as a period of transition from the Iberian system to the Roman. There is no information on whether Black Glaze and Arretine Terra Sigillata were being traded along the Guadalquivir, but if they were, they were then distributed more extensively than the Terra Sigillata Chiara A and D. It is also important to notice that Spanish Terra Sigillata was being produced in workshops at Andujar, Granada and Málaga (Fernández Fonseca 1995, 64) and was probably imported in the Guadalquivir Valley by land, rather than along the navigable rivers, as such a route would be more direct. Arretine Terra Sigillata and South Gaulish Terra Sigillata had a strong influence on the form of Spanish Terra Sigillata, at least in the first two centuries AD, after which Spanish Terra Sigillata established its own personality (Fernández Fonseca 1995, 63) and it is likely that this latter ware was traded to the same areas of the pottery it ended up replacing. Not surprisingly, the variable plot (figure 4.7) shows that Spanish Terra Sigillata is very close to Arretine Terra Sigillata, which it imitated and whose exchange routes probably followed.

Terra Sigillata Chiara was clearly almost only available in the town territories which were directly connected to Hispalis, either by road or by river. The majority of the territories in which the Terra Sigillata Chiara subtypes are found depended on towns which were either directly located on the Guadalquivir (such as Axati), or on a major road leading directly to Seville (such as Segida). This also suggests that Terra Sigillata Chiara was brought to the towns and then redistributed to the sites in the town territories from the centre, rather than being brought to the rural sites directly. This is consistent with the creation of influential Roman centres (*coloniae*) under Caesar and Augustus linked by an integrated road network. Since the sites which received the Terra Sigillata

Chiara A and D are on the Via Augusta or the Guadalquivir, these were the main routes to redistribute this type of pottery. The pots were brought into the valley via the navigable rivers and then carried along the Via Augusta, probably from Hispalis. Though there are other Roman roads leading south from these towns, the analysis on the distribution of pottery types shows that these were not used to redistribute the Terra Sigillata Chiara A and D to the other centres to the south, suggesting a differentiation between the northern and southern parts of the Guadalquivir Valley.

Thin Walled ware had a completely different distribution from terra Sigillata Chiara A and D and it is found where these two types are not present. It is likely that there was not an interest in carrying types A and D further from the Via Augusta, so that in the territories where A and D were not received Thin Walled ware was predominant. Very little is known about Thin Walled ware, except from the fact that it was produced in Spain, possibly in some of the workshops which also produced Spanish Terra Sigillata.

Most of the territories in the southern part of the Guadalquivir Valley were dominated by a combination of South Gaulish Terra Sigillata, Spanish Terra Sigillata and the early types. Spanish Terra Sigillata imitates Arretine Terra Sigillata and South Gaulish Terra Sigillata, so that it is hardly surprising to find it distributed in the same locations. South Gaulish Terra Sigillata can be seen as the counterpart of Terra Sigillata Chiara A and D in the south/south-west of the Guadalquivir Valley. This ware was produced at the La Graufesenque centre and might have arrived to the Guadalquivir Valley through Tarraconensis, along the Via Augusta.

6 Conclusions

All the information presented in this chapter can be put together to create a model of the development of site and exchange patterns oin the Guadalquivir Valley from the 1st century BC to the 6th-7th centuries AD.

6.1 The model

The Romans already had the very first trade contacts with the Iberians in the Guadalquivir Valley during the 3nd century BC, as proved by the presence of Black Glaze pottery. There is archaeological and historical evidence that at this time the Iberian society consisted of a centralised network of towns with dependent rural settlement (Keay 1992, 278-285). After defeating the Carthaginians, in the 1st century BC the Roman first settled the area, but the old Iberian trade network and organisation were still strong, as proved by the persistence of the Iberian exchange network into the 1st century AD, and the Romans' impact on the local society was not too strong. The Roman rural settlement pattern at this time was already differentiated between high-status and low-status sites, with the high-status sites clustering around the Romano-Iberian towns, while the low-status sites were spread across the countryside.

The situation changed in the 1st to 3rd century AD. While right at the beginning of the Early Empire it looks likely that the old Iberian exchange network was still being used for Arretine Terra Sigillata, a new distribution network appears for the Terra Sigillata Chiara A and, later, the Terra Sigillata Chiara D, favouring those sites which were located along the Guadalquivir and the Genil (which could be used for transport) or on the Via Augusta. At this time a differentiation occurs in the way wares were distributed across the valley, with the Terra Sigillata Chiara A and D being found in the northern part of the Guadalquivir Valley and a selection of other pottery types to the south. The density of rural sites increased dramatically with new high and low-status sites appearing centred around the Roman or Romanised towns. Several of the Iberian towns either had a *colonia* of Roman citizens founded on their territory, or were granted status of *municipium.* Under Augustus most of the towns which became *coloniae* or *municipia* were located on the Guadalquivir and seem to have monopolised the distribution of fine Terra Sigillata Chiara A and then D which is hardly found outside their territories. These towns also were the key sites in the trade of oil and corn produced in the Guadalquivir Valley.

Between the 3rd and 6th centuries AD the situation changed again and several high and low-status sites disappeared. The pattern of disappearance of the high-status sites seems to have been different in different parts of the valley. Around the

Guadalquivir the situation stayed rather similar to what it was in the Early Empire, with high density sites clustered around the towns, while elsewhere a dispersed pattern appeared. This is consistent with the appearance of larger estates in the southern part of the Guadalquivir Valley, while the sites along the Guadalquivir were still depending from the towns for their wealth. This suggests a different pattern of decline for the towns in the northern and southern parts of the valley, probably linked to the easier access to communication routes in the north. The low-status sites which disappeared were mainly those away from the towns, showing that there might have been a collapse of the large scale exchange networks with smaller sites having to rely on the services offered by the towns to survive, despite the fact that a generalised decline in the importance of towns as administrative, political and economic centres had occurred by the late 3rd century AD.

Chapter 5

The region of Tarragona

The Romans entered Iberia as a result of the conflict between Rome and Carthage in the late 3rd century BC (second Punic war). By 195 BC they had already defeated the Carthaginians and sedated a rebellion of the Iberians. In 197 BC the Romans divided Iberia into two separate provinces, *Hispania Ulterior* and *Hispania Citerior*, with Tarraco being the main military base of the latter province. In the middle of the 2nd century BC the defensive walls of thc garrison at Tarraco were extended to include the Iberian settlement and provided with towers at 50 metre intervals. It is unlikely that this was a true urban development of the Hispano-Roman town, as Roman-style building techniques only appeared in the last quarter of the second century BC (Keay 1990, 127; Miró 1987). By the later 2nd century BC Tarraco had also become the focus of an important road network, which linked it to the lower Ebro Valley, the Vallès and the Maresme (Keay 1990, 128). By 5 BC the province of *Hispania Citerior* had been enlarged with the additions of inland territories and had been renamed *Hispania Tarraconensis*. Between 45 and 27 BC Caesar founded a *colonia* at Tarraco (Keay 1990, 137; Alföldy 1978, V.1.a), after the town had already been a centre of Roman power for 200 years. Instead of imposing the traditional rectangular plan, a Roman forum and a theatre were built on the site of the pre-Roman settlement (Keay 1988, 56) and it appears that the granting of colonial status to Tarraco did not involve either a resettlement of Roman citizens or a replanning of the layout of the town, though Caesar did install a contingent of veterans here (Keay 1990, 137). Augustus designated Tarraco as capital of *Hispania Tarraconensis* and the town kept this role until the Visigothic period (Keay 1991, 79).

In the 3rd century AD Tarraco had ceased to develop and expand, while several large early imperial mansions along the river Francolí were abandoned by the end of the century. By 262 a body of Frankish tribesmen crossed into north-east *Tarraconensis* from southern *Narbonensis*, sacked Tarraco and then crossed over to north Africa by sea as reported by the Latin author Aurelius Victor (Keay 1988, 177). In AD 298, Tarraconensis was subdivided into a lesser *Tarraconensis* (roughly corresponding to the *conventus Tarraconensis, Caesaraugustanus* and *Cluniensis*), with its capital at Tarraco, *Carthaginiensis* (roughly, the *conventus Carthaginiensis*) with its capital at Carthago Nova, and *Gallaecia* (roughly the *conventus Bracaraugustanus, Lucensis* and *Asturum*) with its capital at Bracara Augusta (modern Braga).

1 Physical description of the land

The study region consists of the ancient *Ager Tarraconensis*. This is divided into two parts by the Francolì river, which runs north to south almost perpendicular to the coast line. The half of the *Ager Tarraconensis* to the west of the Francolì contains the most Roman and pre-Roman sites, while only a few were found in the eastern part. The western part consists of a flat area with low mountains to the north west, while the eastern part is dominated by hills, with the Francolì running along the foot of these. Despite the fact that 90% of the whole area lies below 200 metres above sea level, the slope can be quite steep at places and can be a serious hindrance to movement.

2 The analysis

Since the region of Tarragona is considerably smaller than the Guadalquivir Valley, the sort of extensive study carried out on the archaeological data set from the province of Seville could not be repeated in the same fashion. Generally, it was felt that the data collected in the *Ager Tarraconensis* were slightly better integrated than those from the province of Seville, as a much smaller area had been covered and most of the collection had been carried out by the same people, so that the bias introduced by personal preference of the surveyor did not create a differentiation between zones. Despite the fact that the data were not quantified, the qualitative information available was superior to that concerning the archaeological sites from the Guadalquivir Valley, as more detailed classification of the ceramic material had been carried out. The amphorae were not simply recorded as amphorae, but their origin was stated, therefore the database contains information on whether any specific amphora was locally produced or came from Italy or Africa. Since the data had already been classified chronologically in several time bands, this division was preserved as it allowed a more detailed chronological study than it was possible for the Guadalquivir Valley.

The other major difference between the Guadalquivir Valley and the Tarraco data sets was that it was not possible to obtain a soil map of Catalunya, so that the study of the relationship between archaeological sites and soil types or agricultural potential was severely limited. The only soil map available was incomplete and poorly printed (in Cobertera 1986). The relationship of sites classed as `villa' and as `rural site' by the surveyors to the Roman town of Tarraco was investigated, and so was the difference in the elevation at which rural sites dating from different periods are found.

2.1 The classification of the sites

While the rural site data from the Guadalquivir Valley had to be reclassified according to pre-defined criteria because of the large area covered and the number of people who had carried out the surveys and identified the site type, this problem does not really occur in the region of Tarragona. This depends on the fact that the materials found were studied in more detail and by a smaller number of people, so that the classification is more reliable in terms of site type and the chronological life span of each site had already been worked out and was included in the site cards.

2.1 The distance of rural sites from ancient Tarraco

The distance from ancient Tarraco was calculated as a cost surface using a friction surface derived from the slope map of the region (figure 5.1). The cost surface was then reclassified into 11 bands whose width corresponds to the cost of moving 5 km over a completely non-friction surface[1]. This measure was achieved in the same way as described for the calculation of the 15 km areas around the towns of the Guadalquivir Valley, by creating a binary mask corresponding to 5 km linear, superimposing it onto an ideal cost surface and thus finding the cost limit value which was then used to

[1] The cost distance units equivalent to the cost of moving over a km in linear distance are referred to as km equivalent (km eq), therefore the bands are said to be 5 km eq wide, see chapter 4, *The Guadalquivir Valley*, section 5.1.

reclassify the true cost surface. Notice that since the increment in the cost surface units is not the same as the linear units, the cost limit value was calculated for each distance band (*ie* 10 km, 15 km etc), rather than calculating a cost limit value for the 5 km band and multiplying it by the required value.

The relationship between the rural sites and Roman Tarraco was tested by means of the Kolmogorov-Smirnov one-sample test, using the area of each cost distance band as the hypothetical distribution against which to compare the number of rural sites

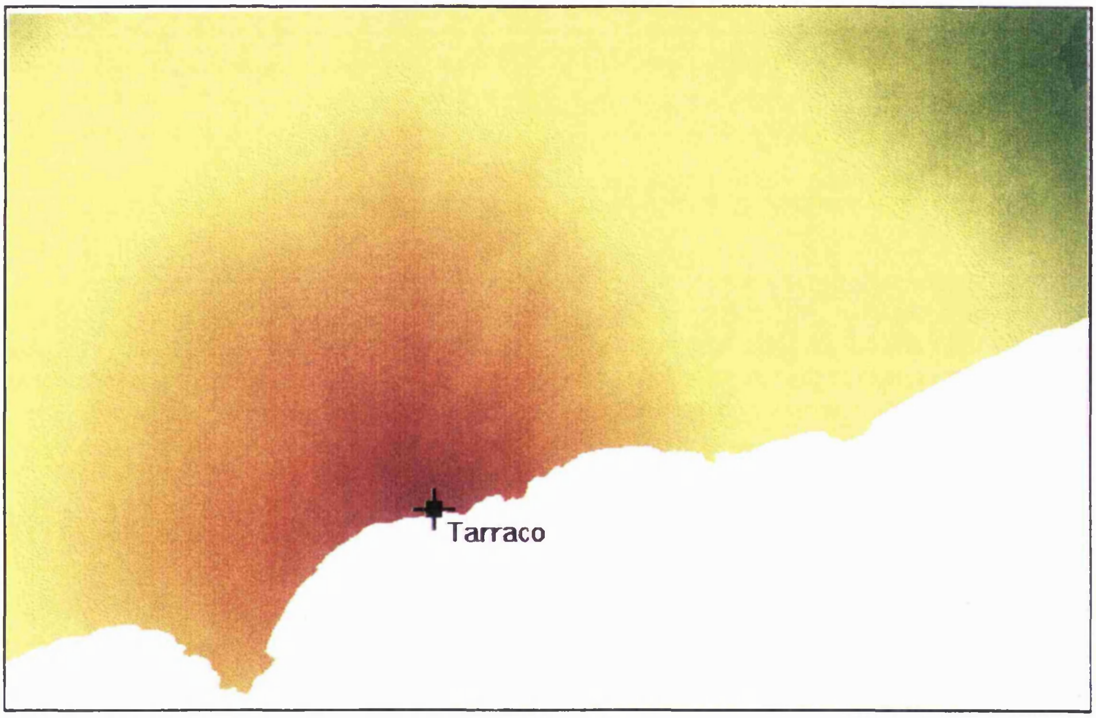

Figure 5.1 The cost distance surface calculated from the Roman town of Tarraco

occurring on each band. Of the 11 distance bands, only the area covered by the 6 nearest to Roman Tarraco had been surveyed, therefore only these were used in the Kolmogorov-Smirnov test[2], to avoid the error which would be introduced by the influence of the parts of the region for which data were not available because no survey had been carried out there. A Kolmogorov-Smirnov test was carried out for each period identified in the database, the Republic, the reign of Augustus, the Early Empire, the Third century and the Late Empire. The results of the Kolmogorov-Smirnov tests, at the

[2] For an introduction to the Kolmogorov-Smirnov one-sample test see chapter 4, *The Guadalquivir Valley*, section 4.2

5% level of significance, show that there is a significant relationship between the position of the Roman rural sites and the Roman town of Tarraco in all periods except from the Late Empire, for which the test proved not significant. This relationship can be further investigated by looking at the way the distribution of rural sites with respect to the position of Tarraco changed in time.

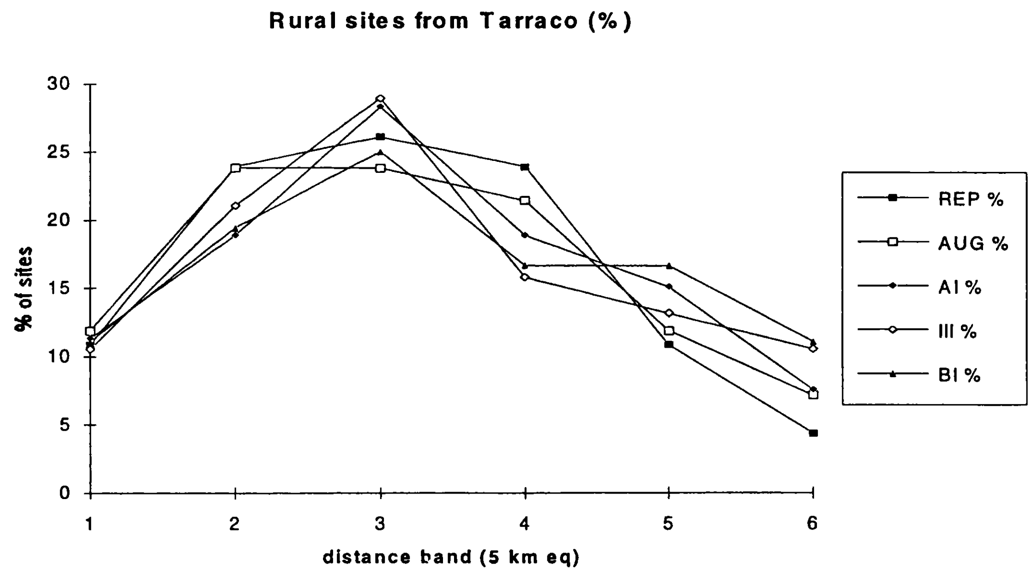

Figure 5.2 The percentage of rural sites on cost distance bands (5 km eq) from Roman Tarraco

Figure 5.2 shows the percentages of sites from each period on the 5 km eq. distance bands from Roman Tarraco. The two curves detailing the percentage of rural sites dating from the Republic and the reign of Augustus on cost distance bands from Roman Tarraco are similar. The three graphs of the percentages of rural sites dating from the Early Empire, the 3rd century and the Late Empire are also similar to each other, but different from the other two. These graphs suggest that the pattern of distribution of rural sites around Tarraco changed in the middle of the 1st century AD.

The two patterns (before and after the mid 1st century AD) can be looked at separately by splitting the graph into two parts, as shown in figures 5.3 and 5.4.

The distribution of rural sites in existence during the Republic and the reign of Augustus shows that the majority of the sites tend to occur in the cost distance bands closest to Tarraco, decreasing with distance. In the Early Empire this pattern changes and the largest concentrations of rural sites are further away from Tarraco, in the distance band corresponding to 10-15 km eq, with another peak in band 5, corresponding to 20-25 km eq. The graph of rural sites dating to the Late Empire shows a sort of flattening of the curve, indicating further disappearance of rural sites close to Tarraco, but not away from the town. This is consistent with the decrease in importance and power of Tarraco itself clearly visible after the 3rd century (Keay 1988, 175). The decrease in the total number of rural sites in the Late Empire also suggests that the

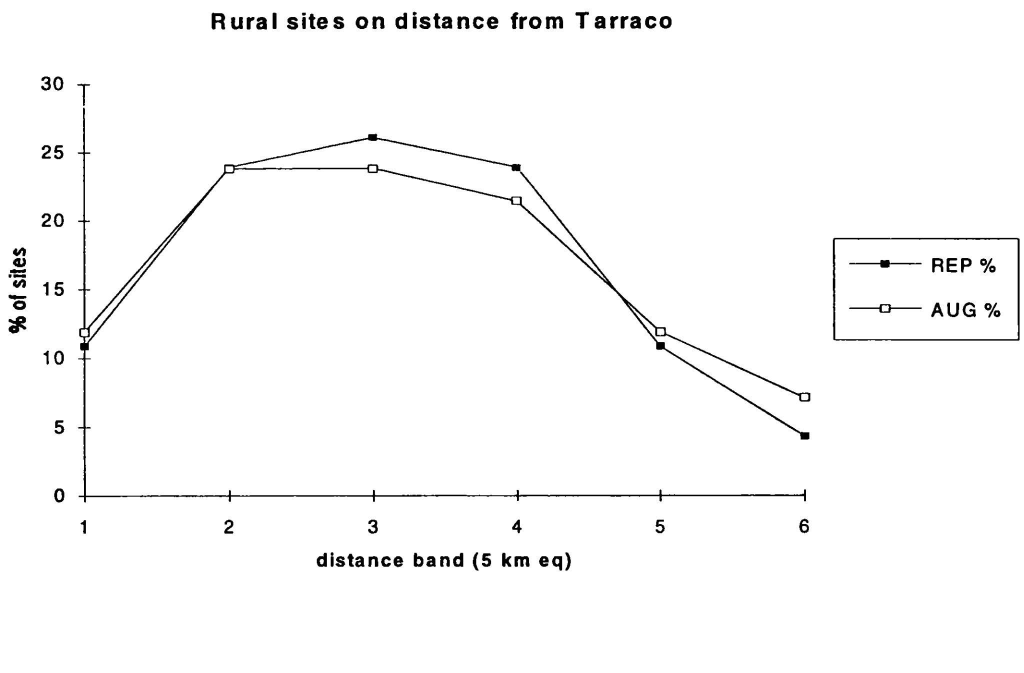

Figure 5.3 The percentage of rural sites dating from the Republic and the reign of Augustus on cost distance bands (5 km eq) from Tarraco

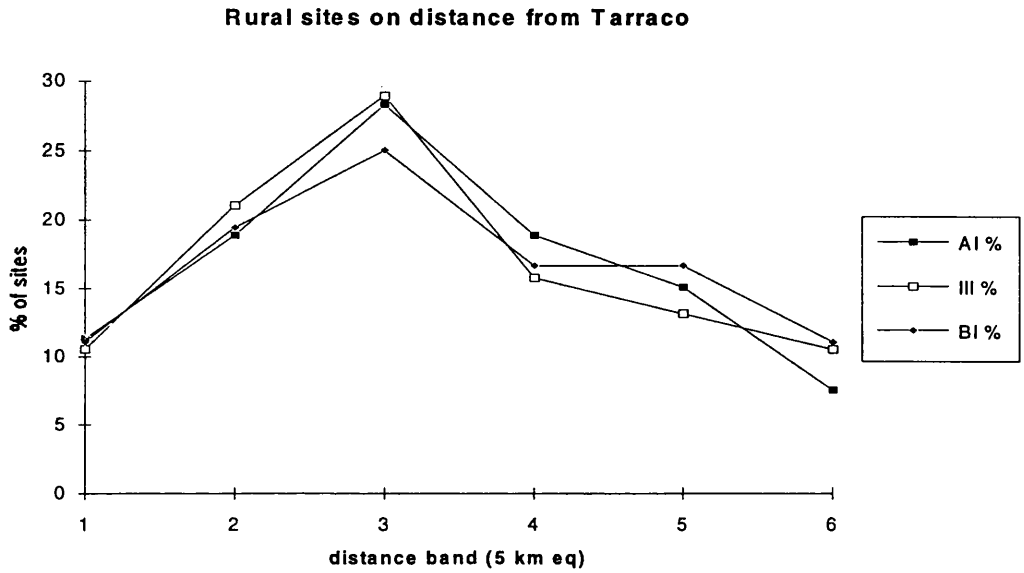

Figure 5.4 The percentage of rural sites dating from the Early Empire (AI), 3rd century (III) and Late Empire (BI) on cost distance bands (5 km eq) from Tarraco

remaining sites turned into larger estates controlling larger areas. These results contradict Millett (1992, 179), who argues that the rural sites were not originally centred on Tarraco but came closer to the town in later periods. It is not clear exactly what periods Millett refers to in his paper, as no chronological information is given for this switch from a non-town-centred to a town-centred rural settlement pattern in the *Ager Tarraconensis*, but from the above discussion it appears that the opposite is true.

The results of the Kolmogorov-Smirnov tests discussed above confirm that there is a significant relationship between the position of the rural sites and Roman Tarraco in the earlier periods, but this relationship is not statistically significant in the Late Empire. This trend is further confirmed by the study of the distribution of sites in the *Ager Tarraconensis*, which shows a settlement pattern centred around Tarraco in the Republic and the reign of Augustus, followed by a shift away from the town from the

Early Empire onwards, culminating with a fully distributed pattern in the Late Empire. Despite the fact that quarries in the eastern part of the *Ager Tarraconensis* had been exploited since before the arrival of the Romans, other quarries further away from Tarraco itself come into use in the Early Empire, suggesting an increased need of stone for the creation of new rural sites and for the urban development of Tarraco. These quarries will continue being in use into the Middle Ages, when new ones further away from Tarraco will start being exploited.

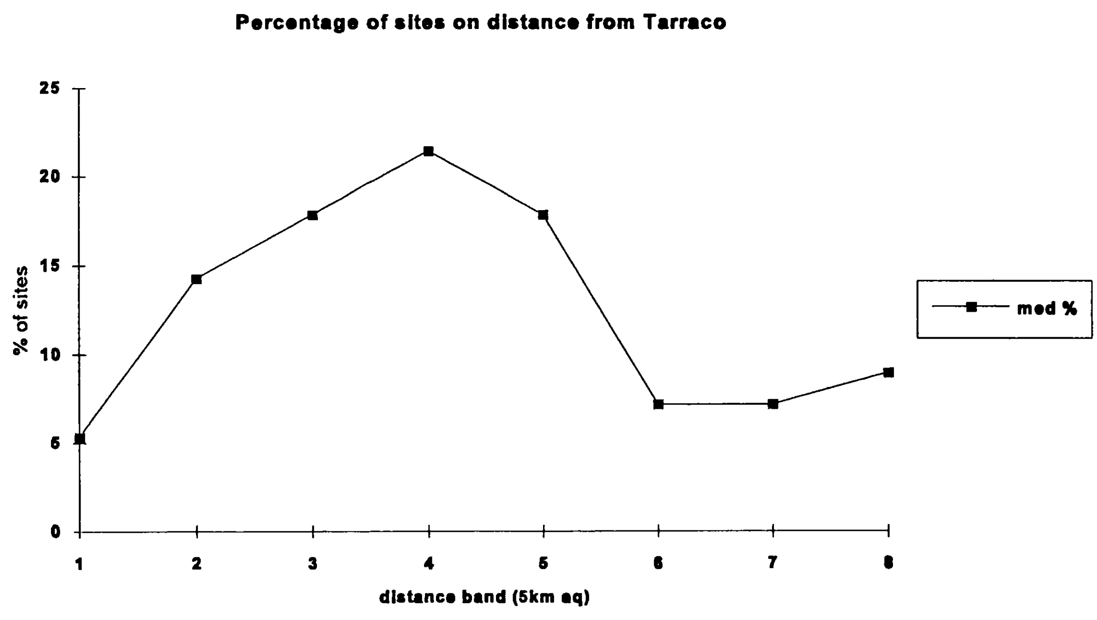

Figure 5.5 The percentage of rural sites dating from the Medieval period on distance from Tarraco

Tarraco was founded as a military settlement and it is likely that in the very early period the very first Roman settlement around the town occurred quite close to it for reasons of protection. This pattern was maintained until the colony was founded by Caesar, which involved a reorganisation of the countryside around the town itself, visible from the Early Empire onwards.

In the medieval period, new sites were founded further away from Tarraco, reaching as far as distance band 8 (40 km eq). This pattern continues the trend already evident in the Late Empire of movement away from Tarraco, probably as a consequence of a loss of power of the town itself, together with a reorganisation of the countryside brought about by the end of the Roman rule. The fact that new sites are found so far away from Tarraco itself suggests that they were not influenced by the location of the town at all, which is also confirmed by the fact that the Kolmogorov-Smirnov test for this period was not significant. At this stage Tarraco had completely lost its role as the political, economic and administrative centre of the region, a fact that is reflected by the dispersed pattern of rural settlement distribution.

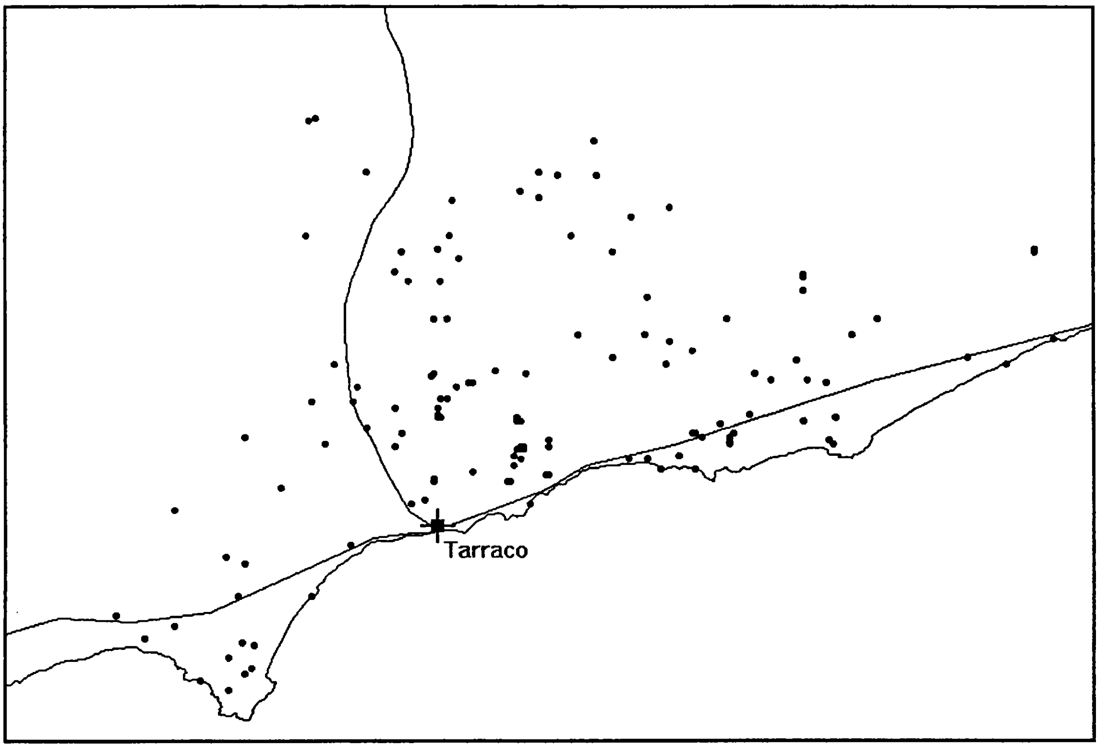

Figure 5.6 The position of the Via Augusta (along the coast) and the road to Ilerda (going north) in relation to Tarraco and the rural sites

2.2 The rural sites on distance from the Via Augusta and the road to Ilerda

The main communication route in the region of Tarragona was the Via Augusta, which ran parallel to the coast, connection the Maresme, over the Garraf massif to the north, to the south of Spain (Pallí Aguilera 1985, 207), eventually leading to Hispalis

(Seville). The other important road in the area was the one leading north to Ilerda and Caesaraugusta from Tarraco (Keay 1990, 139).

These two roads were used as starting points to calculate a cost distance map, which was then reclassified into a number of 2 km eq cost distance bands and used to perform a series of Kolmogorov-Smirnov tests for each of the periods the site data are dated to. Since rural sites are only found in the first 6 cost distance bands away from the roads, the bands with no sites on them were excluded from the analysis to limit the bias introduced by the areas which were not surveyed. The results of the tests were significant in all cases at the 5% level of significance except for the data set from the Late Empire, indicating that a statistically significant relationship existed at almost all times between the location of the rural sites and the position of the two main Roman roads in this area.

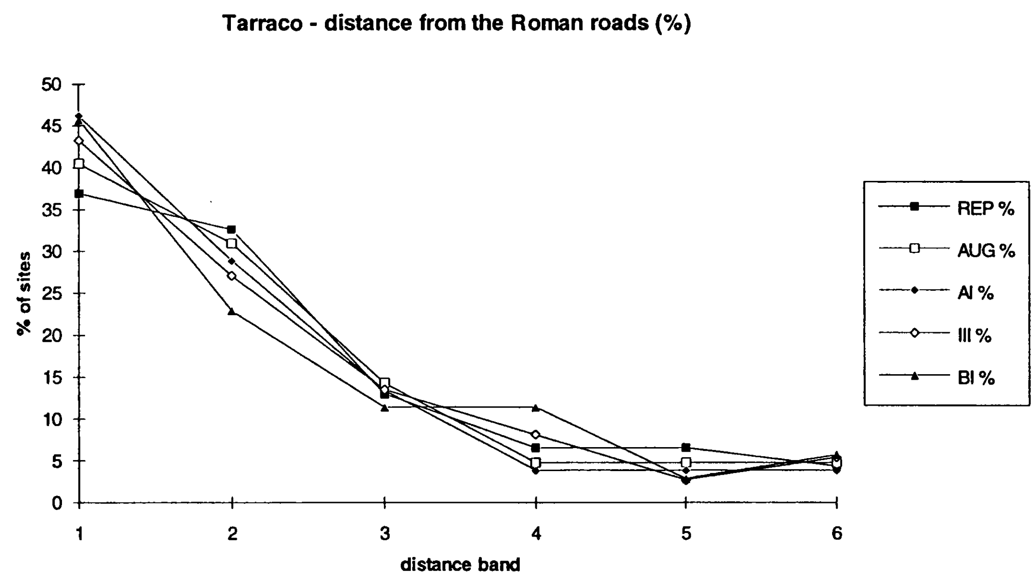

Figure 5.7 The proportion of rural sites on 2 km eq distance bands from the Roman roads

Further insights can be gained by studying the graph of the percentages of sites occurring on each distance band (figure 5.7). The curve of the proportion of rural sites dating to the Republic shows that a lower proportion of sites existed very close to the roads (in band 1) than in later periods. The proportion of sites dating to the Republic in the second band, however, is higher than in the following periods. This suggests a site pattern centred around the roads in the early period of Romanisation, though at this time the communication routes were not as important as they were in the later periods to influence settlement. In the later periods, the proportion of rural sites located close to the roads increased and the pattern of distribution around the roads stayed the same until the Late Empire, when the proportion of sites away from the roads (in band 4) increased, while the proportion of sites in bands 2 and 3 decreased and the proportion of sites in band 1, very close to the roads, stayed the same as the preceding periods.

These variations in the way rural sites were distributed along the communications routes at different times suggest that in the first period of Roman settlement it was important to locate the new sites near the roads, but it was not so important as it was in the following periods, possibly because import/export was not such an important feature as it was later. It is also possible that some of the new Roman sites were continuations of pre-existing Iberian farms, so that the Republican site distribution pattern reflects in part the pre-Roman settlement pattern. In the following periods, new sites were created along the roads to minimise the effort of reaching the communication and commercial routes (the Via Augusta led to Tarragona, which was the main port of the region), showing that trade had become more important. In the Late Empire the proportion of sites away from the roads increased suddenly, perhaps as a consequence of the decline of the large-scale exchange system which had played such an important role in the area, so that sites no longer needed to be located within easy access of the communication routes.

2.3 The rural sites on elevation

The few Iberian sites known from the *Ager Tarraconensis* were located in protected locations, often on high points, though in no case at points higher than 200 metres above sea level, as 90% of the study area falls below this altitude.

Archaeological evidence has shown that most hill-top settlements were gradually abandoned from the 2nd century BC onwards (Keay 1990, 130). When new Roman sites appeared during the Republic, the majority of these were created on lower land than the Iberian settlement had been, probably to exploit the agriculturally better low lands, determining the shift in the settlement pattern visible in the graph in figure 5.8. The rural site located at the highest elevation, Vil.la `Les Domengs'), ceased to be used in the Early Empire. This probable villa was located on considerably higher ground than all the others at 180 metres above sea level. Generally it seems that there was a slightly higher concentration of rural sites on lower ground than on higher ground at any one time. It should be pointed out, though, that the largest part of the land lies below 200 metres above sea level.

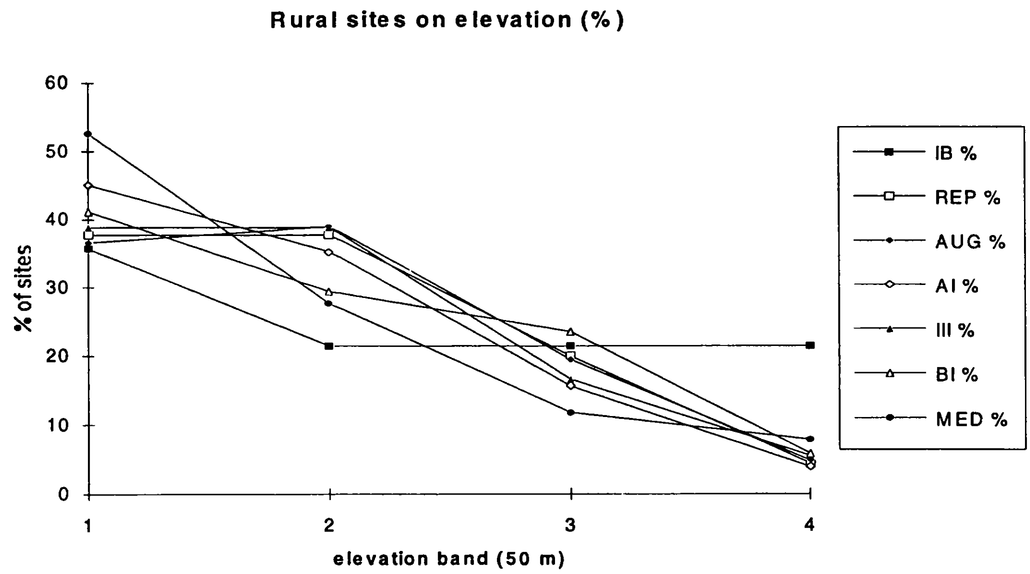

Figure 5.8 The proportion of rural sites on elevation bands (50 metres)

The relationship between the rural sites and the elevation they occur on was tested by dividing the altitude into a series of bands 50 metres wide. The sites occurring on each elevation band were counted and this information was used to carry out a series of Kolmogorov-Smirnov one-sample tests, comparing the distribution of sites on each band to a hypothetical distribution derived from the area of the elevation band. Of the bands, only the first 4 (0 to 50 metres, 50 to 100 metres, 100 to 150 metres and 150 to 200 metres) contained rural sites, therefore the others were excluded from the analysis to try and reduce the bias introduced by the archaeologically empty large proportion of the study area. It might be argued that the information of where rural sites *were not* is as important as the information of where rural sites *were*, but it should be borne in mind that a large part of the archaeologically empty area is so because no surveys or excavations have been carried out, rather than because no archaeological material was found there. Including these areas in the study would introduced a larger error in the analysis than excluding areas with no archaeological material would. A finer distinction between the areas which have not been surveyed and those which have been surveyed but yielded no archaeological material can not be done at present because, apart from Keay and Millett's survey (Keay *et al.* 1989; Keay and Millett 1991; Millett 1991), it is not known exactly to what extent the area has been investigated. The selection of the elevation bands on which rural sites occur will give a good picture of the type of site distribution where we know rural sites were present.

The results of the Kolmogorov-Smirnov tests showed a significant relationship between the elevation above sea level and the distribution of rural sites dating to the Republic, the Early Empire, the 3rd century and the Middle Ages, while the relationship between the elevation and the distribution of rural sites dating to the Iberian period, the reign of Augustus and the Late Empire was not significant.

Figure 5.8 shows the proportion of rural sites on 50 meters elevation bands for all periods. The graph shows that in the Iberian period more sites existed on high ground than it was the case in the Roman period, while the largest proportion of sites in the first band (0 to 50 metres above sea level) dates from the Middle Ages. The sites classified as medieval, on the other hand, encompass several centuries and there might be several settlement formation processes at work, which are reflected in the resulting medieval settlement pattern, so that no clear conclusions can be safely drawn, except from a tendency for rural sites to be located on lower ground than before.

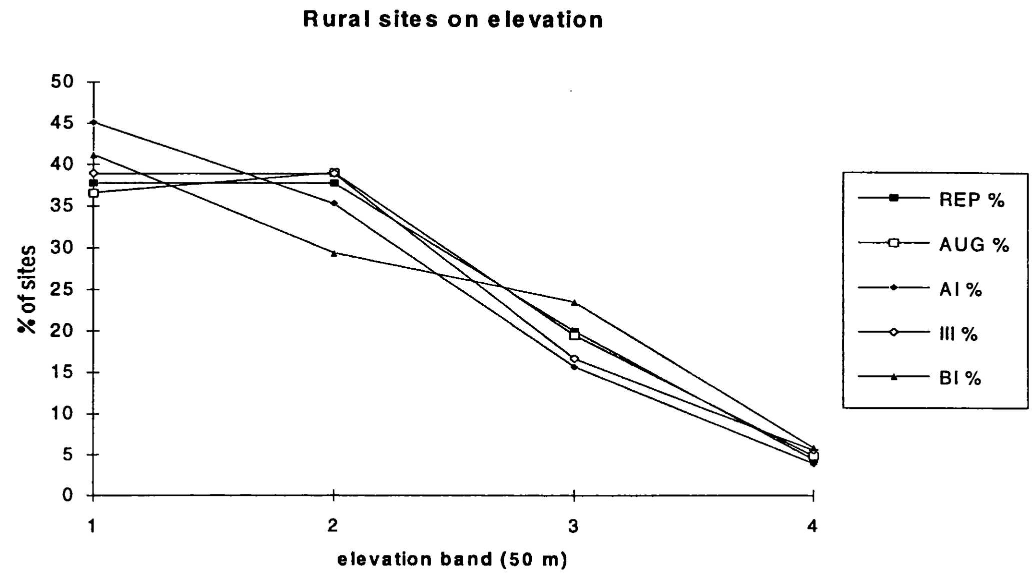

Figure 5.9 The proportion of rural sites on elevation bands (50 metres). Only the sites dating to any of the Roman periods are shown

Figure 5.9 shows the proportion of the rural sites dating from the Roman period only on the four elevation bands. During the Early Empire several new sites were created and the curve for this period shows that these were located preferentially on low ground. This is consistent with the reorganisation of the land following the creation of the *colonia* at Tarraco, as the new sites were located on the better soils of the lowlands. In the 3rd century the graph shows that a higher proportion of sites existed on elevation band 2 (50 to 100 metres above sea level). The curve for the Late Empire shows that at this time the proportion of sites on high ground increased compared to the earlier periods, showing that a movement of sites towards high ground occurred. Less sites existed in the Late Empire than in the Early Empire and in the 3rd century, so it seems that sites disappeared from low ground, rather than new sites being created on high ground. This fact might be linked to incursion of the Frankish tribesmen who sacked

Tarraco (Keay 1988, 177). It is also possible that the higher proportion of sites on high ground is linked to the distributed pattern of rural sites already noticed in the study of the relationship between rural sites and ancient Tarraco (section 2.1). If this pattern was due to the appearance of large estates controlling larger areas, the main rural site (a `villa') would no longer be required to be close to the territory it exploited, but could be located on higher ground to offer better protection against incursions and to give a better view over the Francolì valley.

Despite the fact that very few Iberian sites are present in the database the pattern of distribution of rural sites in the Iberian period is markedly different from the pattern of distribution of rural sites in all of the Roman and Medieval periods (figure 5.8). There is a definite shift in the site location after Romanisation, though available evidence suggests that this shift might have started before the foundation of Tarraco. The new Roman rural sites were created preferentially on lower ground. The results of the Kolmogorov-Smirnov test, which was significant in the case of the Republic, the Early Empire and the 3rd century shows that in these periods there was a marked preference for the agriculturally better low lands. The test for the reign of Augustus was not significant and in fact a careful study of the sites which ceased to be used in this period shows that these were mainly those on the lower ground, possibly due to a reorganisation of the land following the creation of the colony at Cissis-Tarraco. In the Early Empire the new foundations occur again preferentially on lower ground. In the 3rd century sites started to be abandoned and a shift towards a higher concentration of sites on higher ground than in the preceding period is visible. The 3rd century shows the beginning of a generalised transformation in the settlement pattern, which might be linked with incursions in the *Ager Tarraconensis* such as that of the Frankish tribesmen. Fewer sites existed in the Late Empire than in the other Roman periods, but the pattern is different from the earlier period. While a few low land sites disappeared altogether, new foundations tended to occur preferentially on higher ground, reversing the trend which had existed in the Republic and Early Empire. This more even distribution of rural sites on elevation is reflected by the result of the Kolmogorov-Smirnov test for this period, which shows that there is not a significant relationship between the rural sites and the elevation they occur on. On the other hand this relationship was significant in the early Empire and the 3rd century, showing a preference for lower land for the location of rural sites. While a higher proportion of sites than before occurred on band 2 (50-100 metres above sea level) in the 3rd century, in the Late Empire the increase in the proportion of sites occurred on band 3 (100-150 metres above sea level), possibly indicating a slow continuous process of movement towards high ground in the periods following the Early Empire. The medieval settlement pattern shows a high proportion of rural sites occurring on low ground and quite a few occurring on rather high ground (though still below 200 metres above sea level).

3 The pottery assemblage

The distribution of pottery types across the sites in the *Ager Tarraconensis* differs somehow from the pattern found in the Guadalquivir Valley. The wares most commonly found in the sites, after unspecified local common pottery, are Spanish Terra Sigillata and South Gaulish Terra Sigillata, followed closely by Terra Sigillata Chiara A. Black Glazed pottery is found in 35% of the sites and it appears that this ware was more commonly available in the *Ager Tarraconensis* than in the was in the Guadalquivir Valley. Conversely, Arretine Terra Sigillata is found in only 12% of the sites. The other major difference between the Guadalquivir Valley and the *Ager Tarraconensis* lies in the fact that Terra Sigillata Chiara D, which dominated the assemblages of the sites along the Guadalquivir river together with Terra Sigillata Chiara A, is found in 12% of the sites only. This preliminary analysis suggests that the pattern of pottery distribution in south-west Spain differed from that of the north-east. Despite the fact that Common local pottery was found in over 63% of the sites, this category encompasses several different types of local coarse wares which were not better identified and is therefore very difficult to use for any sort of meaningful analysis. A table of the number of sites in which each ware was found is presented below. The percentage in the table was calculated out of the number of sites in the database for which the presence/absence of the wares was recorded, not the total number of sites.

Ware	Number of sites	Percentage of sites
ANF	11	16.9
ANF_I	44	67.7
ANF_IB	33	50.8
ANF_A	12	18.5
CA	23	35.4
COM	41	63.1
TSA	8	12.3
TSH	33	50.8
TSSG	33	50.8
TS	7	10.8
TSC_A	31	47.7
TSC_C	4	6.2
TSC_D	8	12.3
PFINE	7	10.8

Table 5.1 Total number and percentage of sites containing a specific ware. ANF = generic amphora, ANF_I = Italian amphora, ANF_IB = Iberic amphora, ANF_A = African amphora, CA = Black Glazed pottery, COM = common pottery (local coarse ware), TSA = Arretine Terra Sigillata, TSH = Spanish Terra Sigillata, TSSG = South Gaulish Terra Sigillata, TS = generic Terra Sigillata, TSC_A = Terra Sigillata Chiara A, TSC_C = Terra Sigillata Chiara C, TSC_D = Terra Sigillata Chiara D, PFINE = Thin Walled ware.

3.1 Correspondence Analysis of the site assemblages

A study of the distribution of pottery types was carried out by means of Correspondence Analysis, which is one of the few multivariate statistical methods to allow analysis of presence/absence data (Baxter 1994, 104-107). This type of analysis is not as powerful as that applied to quantified data, but it can nonetheless be used to observe general trends in the data. A table of the presence/absence of different pottery types was produced, indicating the presence of a specific pottery type in a site with a 1 and its absence with a 0. Of all the sites in the database, only 65 had at least one pottery type and were included in the analysis. It was decided to exclude from the analysis the ANF (generic amphorae), TS (generic Terra Sigillata) and COM (common ware) pottery types because these three groups were miscellaneous collection of pottery types which had been no better identified and could have biased the analysis.

Component	Iterations	Norm	Eigenvalue	% Inertia	Cumulative
I	57	0.072	0.316546	24.0	24.0
II	37	0.098	0.258100	19.6	43.6
III	31	0.078	0.194179	14.7	58.4

Table 5.2 The Correspondence Analysis information for the analysis of the pottery assemblage of the rural sites in the region of Tarragona

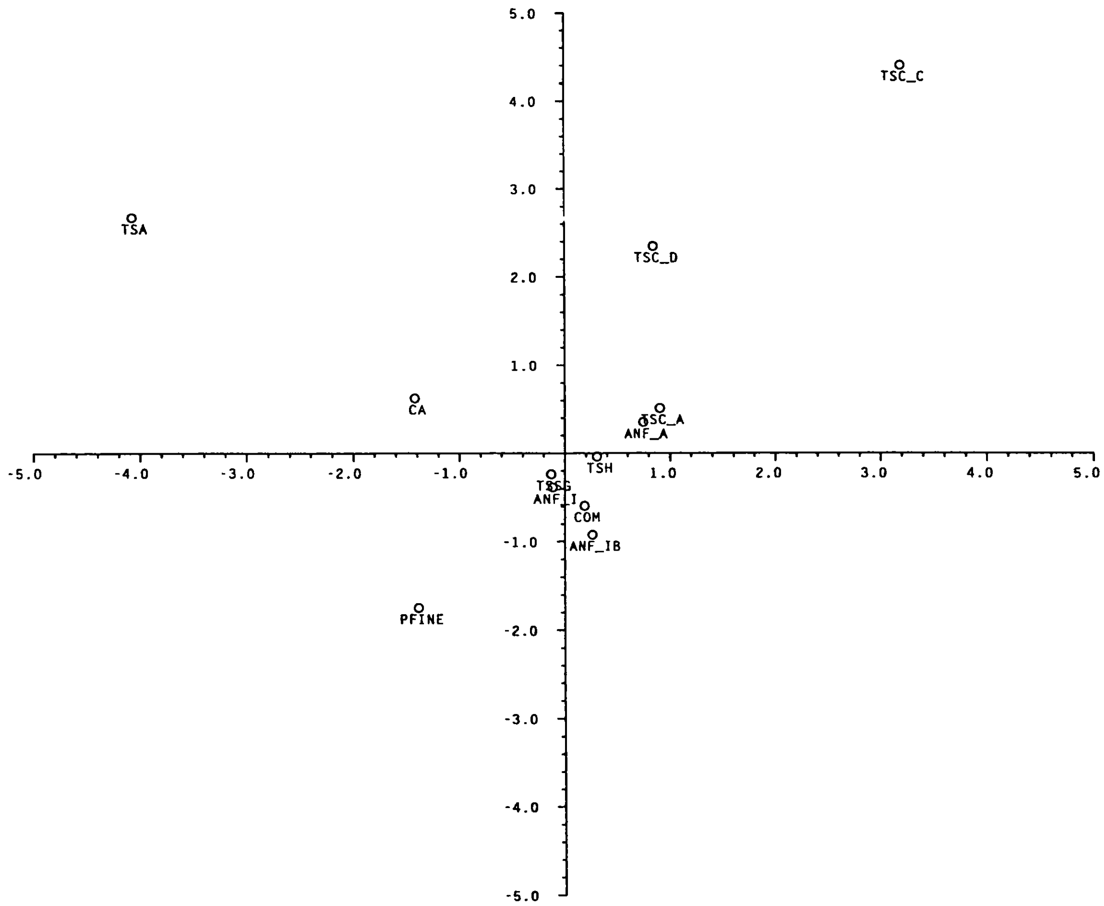

Figure 5.10 The variable plot of the Correspondence Analysis of pottery presence/absence of the sites. Component I on horizontal axis, Component II on vertical axis

Table 5.2 contains the percentages of inertia and other information for the first three components. A study of the components loadings shows that the first three components

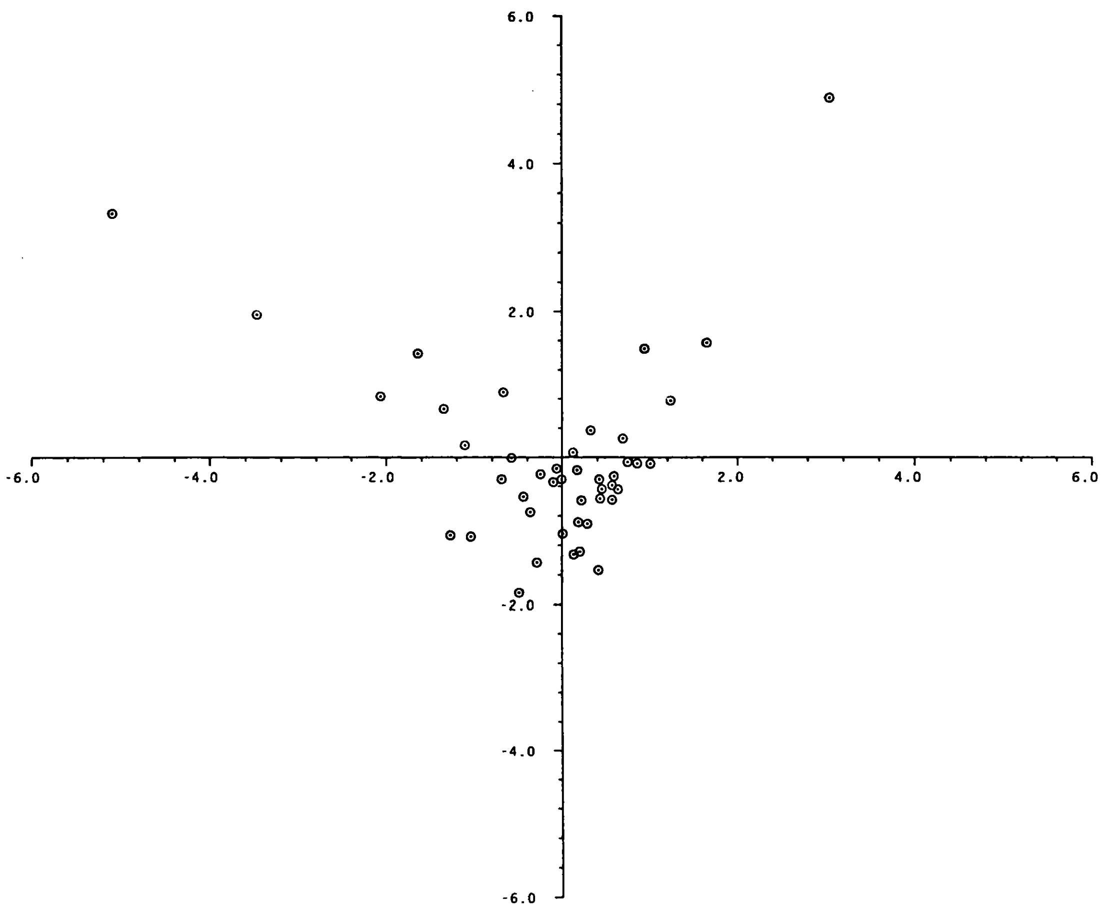

Figure 5.11 The object plot of the Correspondence Analysis on the pottery presence/absence in sites. Component I on the horizontal axis, Component II on the vertical axis

of the correspondence analysis all are strongly influenced by Arretine Terra Sigillata, Terra Sigillata Chiara C and Thin Walled Ware. The variable plot (figure 5.10) confirms this pattern: Arretine Terra Sigillata dominates the top left quadrant, Terra Sigillata Chiara C the top right quadrat and Thin-Walled ware the bottom left one, while all the other wares tend to cluster towards the centre of the plot. The clear break between the Thin-Walled ware and the fine imported wares which was already apparent in the pottery distribution of the Guadalquivir Valley (with the exception of Terra Sigillata Chiara C) appears in the archaeological data set of the region of Tarragona. The separation of Arretine Terra Sigillata from the other wares is interesting as well.

Another trend which can be noticed is that the imported wares tend to be in the top half of the graph, while Thin-Walled ware, a local production, and Iberian amphorae are in the bottom half of the graph, with Spanish Terra Sigillata in the centre of the graph, close to South Gaulish Terra Sigillata which it imitated. It seems that the second component explains the variation between local and imported wares. Contrary to what happens in the Guadalquivir Valley, the Terra Sigillata Chiara C is not completely separated from the subtypes A and D, in fact the three Terra Sigillata Chiara subtypes occur in the same quadrat of the graph, showing that their patterns of occurrence in sites are not very different, while they are definitely separated from the pattern of occurrence of Thin-Walled ware (which was closely related to Terra Sigillata Chiara C in the Guadalquivir Valley) and of Arretine Terra Sigillata.

The plot of sites on the first two components of the Correspondence Analysis (figure 5.11) shows a concentration of sites in the centre of the graph, where most of the wares also cluster in the variable plot. The sites which are separated from the central cluster and are in the quadrant of the plot where Arretine Terra Sigillata dominates all but a few tend to be very close to the Roman road leading from Tarragona to Ilerda. These sites also contain Black-Glaze pottery, which is nonetheless found in other locations as well. If this trend is true it would appear that Arretine Terra Sigillata had a less widespread distribution in the *Ager Tarraconensis* than the other wares. There also appears to be a tendency for sites which were further away from the Roman roads to be in the same area of the graph as imported Terra Sigillata Chiara D and C, while sites closer to the Roman roads (both the Via Augusta and the road leading to Ilerda) are in the area of the graph where the Iberian amphorae and Thin-Walled ware are. The sites closest to the roads are equally likely to occur in the areas of the graph where imported or local productions are.

The horse-shoe shape of the graph is very often associated with seriation of the objects, however, the sites in the Tarragona data set do not seem to be ordered chronologically as a result of the Correspondence Analysis. The only clear chronological pattern visible in the Correspondence Analysis plots is that the sites in the centre of the plot, corresponding to the main concentration of pottery types, were in occupation longer than those found in the top right corner (close to Terra Sigillata

Chiara C and D), which were all occupied only between the Early Empire and the Late Empire.

The widespread distribution of Black-Glaze pottery is in contrast to the distribution of Arretine Terra Sigillata, which features strongly in the assemblages of the sites close to the road from Tarraco to Ilerda. While this fact suggests that this type of pottery arrived at Tarraco and was then carried to Ilerda along the Roman road, it seems that the same was not true for Black-Glaze pottery, which was imported until the 1st century BC. Imported pottery types immediately later than Arretine Terra Sigillata, such as South Gaulish Terra Sigillata and Terra Sigillata Chiara A are found more widely distributed than Arretine Terra Sigillata and do not appear to have a strong influence on the assemblage of the sites near the road to Ilerda. Spanish Terra Sigillata, a local ware which imitates Arretine Terra Sigillata and South Gaulish Terra Sigillata follows a similar distribution as South Gaulish Terra Sigillata and it tends to be found at the same sites as this ware, but has a rather separate distribution from Arretine Terra Sigillata. Both Terra Sigillata Chiara C and D have influence on the assemblage of sites rather far away from Tarraco, as well as close to it.

The Iberian and Italian amphorae are found towards the centre of the graph, near South Gaulish Terra Sigillata and Spanish Terra Sigillata. Interestingly the Iberian Amphorae variable is the type closes to Thin-Walled ware, a local production, while the African amphorae variable is close to Terra Sigillata Chiara A, which was being produced in Tunisia like the African amphorae imported to the *Ager Tarraconensis.* Italian amphorae are close to South Gaulish Terra Sigillata in the plot. Idrisi was used to compare the distribution of the sites with the three different types of amphorae. The Iberian and Italian amphorae are found in sites which are quite evenly distributed around the Ager Tarraconensis, while the sites with African amphorae tend to be found more distant from the coast and have a similar distribution to the sites containing Terra Sigillata Chiara C and D. While a Idrisi plot of the sites is required to look in detail at the distribution of the types occurring in the centre of the object plot because they are close to each other, the subtle difference between the Italian and Iberian amphorae, with the Iberian amphorae being closer to the local Thin-Walled ware, can only be picked up by the correspondence analysis.

3.2 Discussion

The results of the analysis suggest that at very early times of Romanisation, in the 2nd to 1st century BC the early imports of Roman ware (Black-Glaze pottery) arrived to Tarraco and were carried to Ilerda along the Roman road, while the same ware was being also distributed in the immediate countryside. Since at this time the majority of the Roman rural sites were centred around Tarraco, it is likely that Iberian sites were also receiving the pottery. The military centre of Tarraco had been built on or by Iberian Cissis (or Kesse) and this might have served as administrative or political centre for the Iberian country settlement before the arrival of the Romans. The amount of imported Roman pottery seems to have decreased at the end of the 1st century BC to the mid 1st century AD, since Arretine Terra Sigillata is found in fewer sites than Black-Glaze pottery, but this might just be due to a decrease in the total amount of pottery imported. It is indicative, though, that the majority of the sites whose assemblage is influenced by this ware, according to the correspondence analysis, are on the road leading to Ilerda, as this fact shows that the pottery was still travelling along this road at this time and that it was being carried from Tarraco. At the time Arretine Terra Sigillata ceases to be produced, the mid 1st century AD, there also is a shift in the rural settlement pattern, showing a clear break from the Republic and the reign of Augustus (section 2.1), possibly due to the creation of the colonia and the reorganisation of the land around the town.

The apparent trend for later wares is that of a distribution not centred on the main roads but scattered in the countryside, increasingly away from Tarraco and the coast. This ties in with the evidence of a movement of rural sites towards a distributed pattern culminating in the late Empire. This pattern reflects the creation of new rural sites and satellite sites which were now the recipients of the imported fine wares and the local fine production, whose focus shifts away from the road to Ilerda. From the 3rd century the increase in the relative number of rural sites away from the town coupled with the fact that late fine wares (Terra Sigillata Chiara C and D) predominate their assemblages shows that the importance of Tarraco itself was diminishing and that the country estates had reached a certain degree of autonomy and independence from the centre.

4 Conclusions

The available evidence, based on the study of material non-systematically collected, was used to assess the chronological variation in the archaeological settlement pattern in the *Ager Tarraconensis*. The situation present prior to Romanisation was also looked at and so was the period immediately following the Late Empire.

4.1 The model

In the period immediately preceding the Roman settlement in the area, when the Romans were already present militarily in the area, the Iberian communities of the *Ager Tarraconensis* had a network of main larger centres often located on high ground, which was easier to defend, and overlooking important resources such as rivers and agricultural land. These sites were also centres of exchange and manufacture. Smaller, less important Iberian sites were found equally on higher and lower ground (Keay 1990, 122-123). This pattern is reflected in the distribution of the Iberian sites contained in the database on elevation. The sites are found on both high and low ground, but when this pattern is compared with the succeeding Roman settlement pattern, the difference is clear, with the Roman sites tending to be preferentially located on lower ground. At this time Roman pottery was being imported in large quantities to Tarraco itself and was the redistributed to inland sites and other towns along the network of roads starting from Tarraco, such as the road to Ilerda, which was the main connection to the inland area. With the establishment of the Roman presence in Iberia, after the defeat of the Carthaginians and the suppression of the Iberian revolts, a dramatic change is visible in the archaeological landscape. The new rural sites were organised around the town and the high ground locations are gradually abandoned in favour of the better agricultural lowland. This situation culminates under Caesar with the creation of the colonia of Tarraco on the site of Iberian Cissis. While this did not involve the actual creation of a new town and the installation of new settlers, it does appear that a transformation occurred in the organisation of the countryside at the time of the reign of Augustus. The development of the Roman landscape can be seen as consisting of two main stages, the first one of which begins with the imposition of the Roman settlement pattern over the Iberian one and ends with the reorganisation of the countryside at the time of Augustus.

The reign of Augustus also possibly saw a decrease in the amount of Italian pottery imported, although at this time Spanish Terra Sigillata, which imitates the late forms of Arretine Terra Sigillata started being produced and it might have replaced some of the foreign imports, although the correspondence analysis shows that the distribution of Spanish Terra Sigillata is quite different from that of Arretine pottery.

The second stage of the development of the Roman landscape in the *Ager Tarraconensis* goes from the Early Empire to the Late Empire, after which time the rural settlement pattern undergoes another transformation. Some of the rural sites close to Tarraco disappeared, while new ones were founded further away from the town. A large number of new rural sites were created leading to an increased requirement of building material, which was obtained by starting to exploit new quarries in the eastern part of the *Ager Tarraconensis*. At the same time the distribution of pottery changes as well. The sites with imported South Gaulish Terra Sigillata and Spanish Terra Sigillata, which imitates this type as well as Arretine Terra Sigillata, no longer show a concentration along the road to Ilerda. This can be due to the reorganisation of the land, with pottery being carried from Tarraco directly to the new and old surviving rural sites, but also to a decline in the use of the road itself. It is likely that Tarraco still was the centre to which foreign pottery arrived to be redistributed, though the redistribution process might have changed, with the sites in the *Ager Tarraconensis* being the recipients of the imported pottery, rather than the other towns further away. The correspondence analysis shows that the distribution of Spanish Terra Sigillata is very similar to that of South Gaulish Terra Sigillata, but very different from that of Arretine Terra Sigillata (which was shown to influence the assemblages of the sites near the road to Ilerda). This fact suggests that the decline in the transport of pottery to Ilerda and the other towns occurred towards the middle of the 1st century AD. The same process was still at work in the 3rd century AD, when more rural sites close to Tarraco ceased being used and new ones were created further away. It is possible that this transformation of the country settlement pattern was the consequence of the disappearance of smaller sites or their inclusion in larger complexes and estates controlling larger pieces of land. The estates became increasingly independent from Tarraco, diverting the movement of fine imported pottery away from the roads towards the countryside. This trend is also shown by the three subtypes of Terra Sigillata Chiara. The earliest type, A, which started being produced in the late 1st century, is found in similar locations as Spanish Terra Sigillata and South Gaulish Terra Sigillata, though it already shows a tendency to being more evenly scattered across the countryside. The subtype D follows a similar pattern, while subtype C tends to be found in rural sites away from the road to Ilerda and Tarraco itself. Though all the Terra Sigillata subtypes were produced in Tunisia, A and D were being produced in northern Tunisia, while C was being produced in central Tunisia. The difference in the place of production might account for a different pattern of import and redistribution, as it has already been discussed for the pottery assemblage of the Guadalquivir Valley.

It is not clear whether it was an initial decline of the importance of the town which caused the initial breakdown between the town and the rural sites, or whether a few landlords just bought off the other rural sites and became powerful enough to break out of the administrative and economic control of Tarraco, or, more likely, a combination of these two factors, but once the process had been started it simply carried on until the relationship between the town and the countryside was completely broken and any exchange would occur between the estates, bypassing the town. This process is completed in the Late Empire, when the rural sites no longer bear any relation to the position of Tarraco or the main roads, showing that the exchange occurred within the local area, rather that being geared towards export. A good indication of this situation is the distribution of sites where late fine pottery, Terra Sigillata Chiara C and D occur, away from the main roads rather than close to them. In this period the distance between rural sites increases as well and most new sites appear on higher ground, suggesting that defence had become a factor influencing site location.

Chapter 6

The Maresme

The Maresme is the area found along the coast of Catalunya north of Barcelona. Before the Romanisation, the Iberian settlement pattern consisted of small nucleated villages or hamlets, found on the peaks and flanks of the major hills and ridges The type of settlement pattern suggests that each village exploited a narrow and elongated band of land perpendicular to the coast and the mountains, comprising all possible types of terrain and resources available in the region: fishing, shellfish gathering, arable land, woodland and pasture. Smaller sites (second and third order sites) were scattered around the countryside to take full advantage of the possibilities of resources exploitation. Available evidence suggests that the arrival of the Romans in this area was rather peaceful, as no Iberian site in this area suffered violent destruction at the beginning of the 2nd century BC (Miret *et al.* 1991, 47-50). Two Roman towns existed in this area: Baetulo (modern Badalona), which was founded towards the end of the 2nd or the beginning of the 1st century BC, and Iluro (modern Mataró), probably founded at roughly the same time (Prevosti Monclús 1991, 135). Excavations carried out at these two sites have shown that both towns were effectively abandoned by the late 3rd century, by which time their administrative role had been taken over by Barcino (modern Barcelona, see Keay 1988, 176).

1 The geography

The Maresme comprises the coastal strip running north from Barcelona and is defined to the north-west by a chain of low and highly eroded granite hills running parallel to the coast known as the Serralada Litoral. The hills flanking the thin coastal strip are characterised by a number of small valleys created by the action of the several streams which run to the sea from the highlands. It differs somehow from the nearby region of Tarragona in that a larger proportion of the land is found above 100 metres above sea level, causing the coastal strip to be rather thin and favouring a pattern of settlement constrained between the foothills and the coast. The Maresme is separated from the region of Tarragona by the Garraf massif, which effectively splits Catalunya into two distinct zones distinguished by climate, geology and vegetation (Solé Sabarís 1958-68).

2 The archaeology

The archaeology of the Maresme is as yet fragmentary compared to that of the neighbouring region of Tarragona (Miret *et al.* 1991). In 1952 Ribas (Ribas 1952) produced one of the first works on the settlement around the Roman town of Iluro, but the largest project carried out in this area so far is that of Marta Prevosti Monclús (1981a and 1981b), who also attempted some detailed study of the Roman settlement of the Maresme.

2.1 The archaeological data

The site data were obtained partly from the Generalitat de Catalunya and partly from Prevosti Monclús' published survey and excavation data (1981b). Since Prevosti Monclús only produced distribution maps but did not publish the exact coordinates at which each site was found, it was necessary to use the map included in her book to obtain the coordinates of the sites. This was done by digitising the points representing the sites in AutoCAD after having calibrated the digitising tablet on a few known points on the map. This method is not as accurate as the input of the exact coordinates could be, especially because of the large scale of the symbols used to identify the sites on the map, but it was felt that excluding the data published by Prevosti Monclús from the study would have biased the analysis by withdrawing important available information. The sites were dated into six chronological bands by the people who produced the records in the archaeological unit of the Generalitat de Catalunya. These bands are the Iberian period (IB), the Republic (REP), the reign of Augustus (AUG), the Early Empire (AI), the 3rd century (III) and the Late Empire (BI). Despite the fact that the presence of Medieval sites was recorded in the region of Tarragona, this information was not available for the Maresme at the time of collecting the data, so that the period following the Roman time was not included in the study. The sites from Prevosti Monclús' catalogue were dated on the basis of the pottery found in them. Prevosti Monclús (1981b, 21) redefines the term villa and defines all the rural sites she found as `villae' or as `probable villae'. To avoid confusion and to integrate her data with the data from the catalogues of the Generalitat de Catalunya, it was decided to drop the term `villa' altogether and use the term `rural sites' instead. Information on the position of the Via Augusta, which ran parallel to the coast of Maresme was obtained from Pallí Aguilera 1985.

3 The analysis

The archaeological site data were used to study whether any variation occurred in the settlement pattern of the Maresme after Romanisation and at different times during the Roman period. The relationships between the location of known rural sites and the elevation they occur on, the distance from the two main Roman cities and from the Via Augusta were tested for all the different period spans the data are divided into. The information about the presence/absence of the different pottery types in each site was used as raw data to perform a Correspondence Analysis to study the pattern of distribution of the ceramic groups.

3.1 The rural sites on elevation

The Maresme has higher hills than those found in the region of Tarragona. While in the region of Tarragona the maximum elevation is 700 metres and 90% of the territory lies below 200 metres above sea level, the situation in the Maresme is slightly different. While there is a marginally higher proportion of land on higher ground, (83% of the land lies below 200 metres above sea level), a number of sites is found at higher elevation than this. Prevosti (1981b, 264) concludes from the study of her data that the majority of the `villae' (rural sites) on her study area occur around 100 metres above sea level, with a group of `villae' occurring on much lower ground along the coast. The elevation map was divided into bands of 50 metres and a Kolmogorov-Smirnov one-

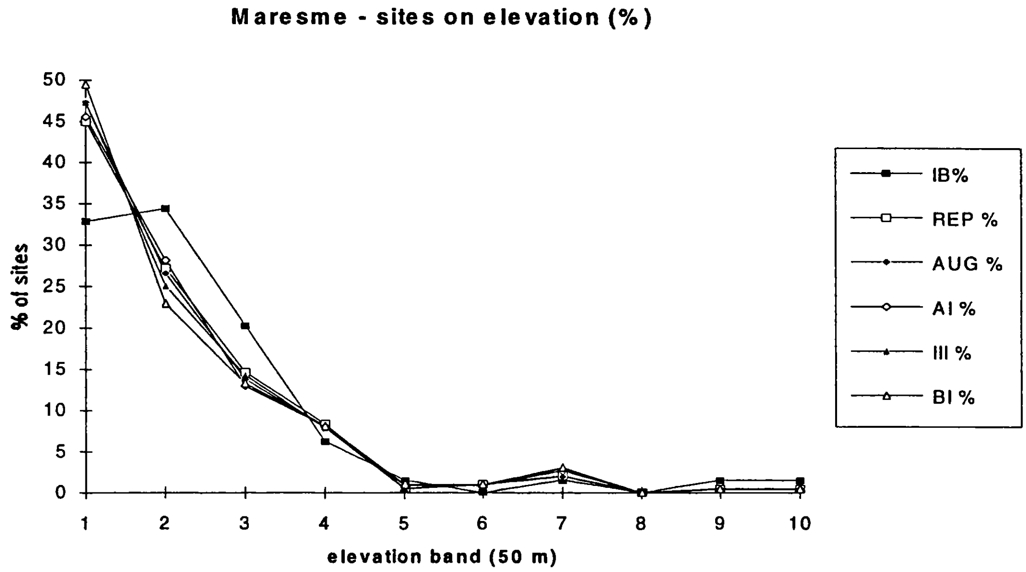

Figure 6.1 The percentage of sites dating from different periods on 50 metres elevation bands. IB = Iberian period, REP = Republic, = AUG = the reign of Augustus, AI = Early Empire, III = the 3rd century, BI = Late Empire

sample test was carried out to see whether there is a significant relationship between the rural sites in existence at any one period and the altitude they are found at. In all cases the Kolmogorov-Smirnov test was significant at the 5% level of significance, but this was expected as fewer sites were built on higher ground which was less easily accessible due to the steep slopes. Most of the settlement occurs in the thin coastal strip at the foot of the mountains. Figure 6.1 shows the proportion of rural sites on 50 metres elevation bands at different times. The graph shows that in Iberian time there was a smaller proportion of sites on low ground and a higher proportion of sites on higher ground than in the Roman period, though very few rural sites were found above elevation band 5 (200 to 250 metres above sea level). The graph also shows that the proportion of rural sites dating to the Roman period on elevation differs from that of the Iberian sites. In Roman times a higher proportion of sites is found on low ground than in the earlier period. There does not appear to be any variation in the proportion of sites on elevation throughout the Roman period from the Republic to the Late Empire, despite the fact that some sites disappeared and others came into existence at different times.

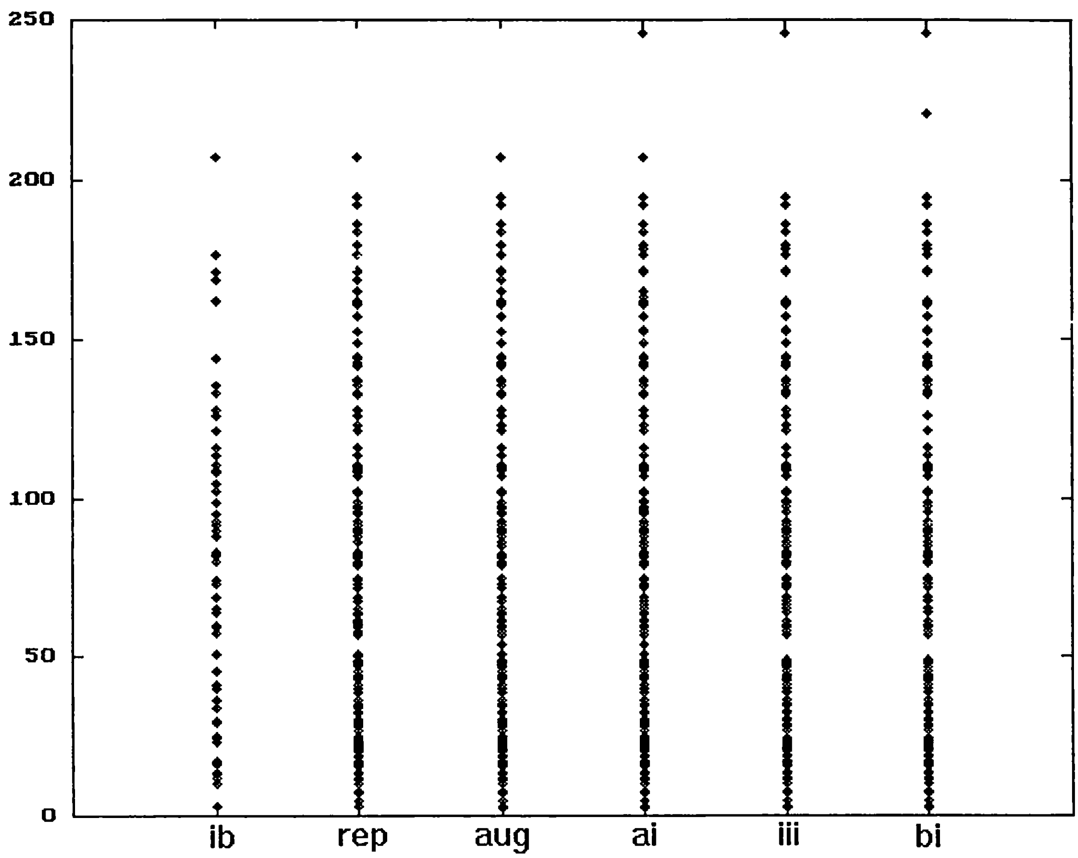

Figure 6.2 The distribution of rural sites on elevation, up to 250 metres above sea level (IB = Iberian, REP = Republic, AUG = the reign of Augustus, AI = Early Empire, III = 3rd century, BI = Late Empire)

Prevosti Monclús (1981b, 264) states that the preferred location for villae in Roman times was around 100 metres above sea level with a few luxurious villae located by the seaside. Some villae were found at higher ground, but these were exceptions. Since Prevosti Monclús talks about 'villae', it might be argued that she only studied high status sites, while the sites contained in the Maresme site database are labelled 'rural sites' and could be high status or low status. It was decided at the beginning of this project not to use the term 'villa' because of the problems of defining exactly what a Roman villa was, while Prevosti Monclús refers to all the rural sites in her survey area as 'villae' or 'probable villae', for the same reason (Prevosti Monclús 1981b, 21). The difference simply lies in the name used to designate locations in the countryside where people lived in Roman times. Prevosti Monclús' data also were integrated in the Maresme site database and the results from this study can be compared her conclusions. Prevosti Monclús' statement that rural sites are found primarily around the elevation of 100 metres is contradicted by the results of this analysis. The graph shows that in Roman times the largest concentration of rural sites was on elevation band 1 (0 to 50 metres above sea level) with 45% to 50% of the cases in all the Roman periods. Around 25% of the sites from each of the Roman period are found on band 2 (50 to 100 metres), while over 35% of the Iberian sites are found at this elevation. Prevosti Monclús notices that some sites are found at low elevation, near the coast and hypothesises that these sites are so located because they were involved in overseas exchange and that sites occurring along the coast and on higher ground had different functions (*ibid.* 264).

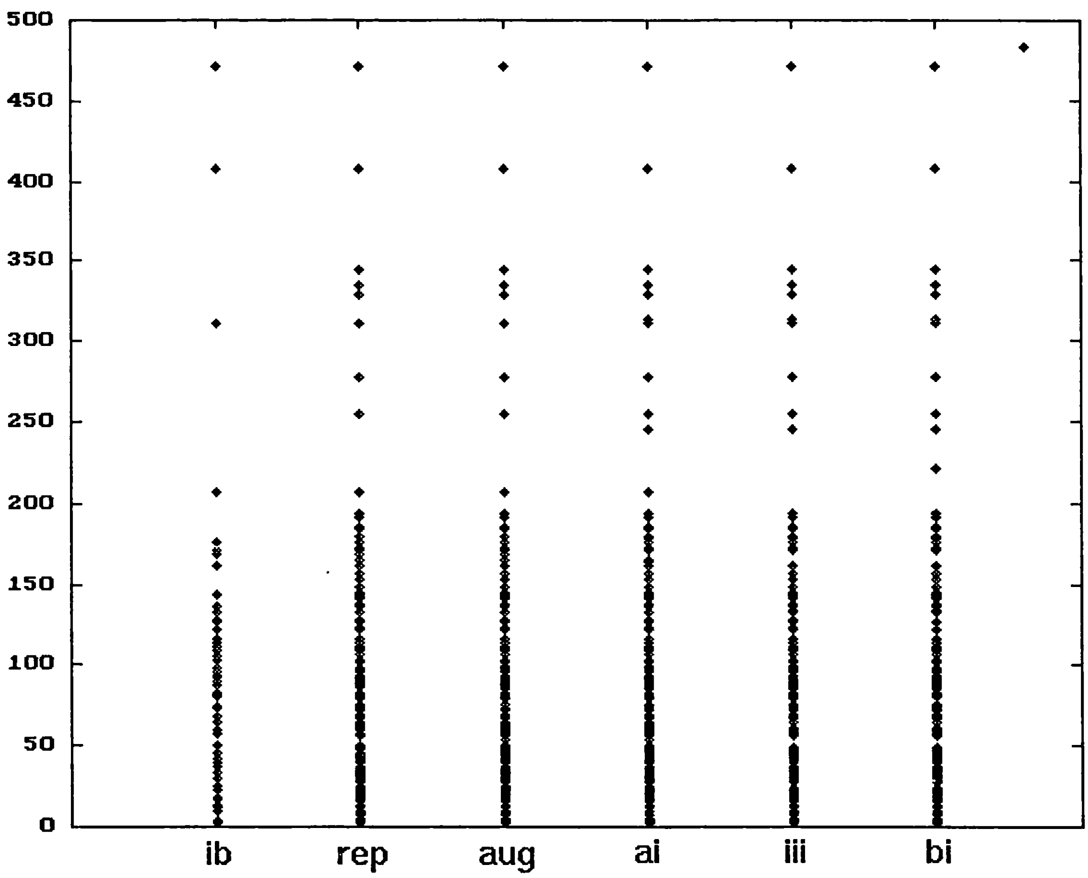

Figure 6.3 The distribution of rural sites on elevation (IB = Iberian, REP = Republic, AUG = the reign of Augustus, AI = Early Empire, III = the 3rd century, BI = Late Empire). The scale on the vertical axis is the elevation in metres above sea level.

Figure 6.2 shows the elevation (the real elevation, not divided in bands) at which sites occur in each period (each dot represents a site). To make eventual pattern easier to see only the sites occurring below 250 metres above sea level were included in the graph. There does not appear to be a preference for the rural sites to occur around 100 metres or at sea level, as Prevosti Monclús states, in fact, the cutting point seems to be around 200 metres. Above this elevation only a few sites existed, as it can be seen in figure 6.3, which also shows that despite the fact that the proportion of Iberian sites on high ground (figure 6.1, bands 9 and 10) is higher than in Roman times, when the actual count of sites is checked, more Roman than Iberian sites existed on high ground. The number of rural sites found at higher elevation than 250 metres above sea level shows that in a few cases some high locations where especially selected for settlement and continued in occupation until the Late Empire. These sites on higher ground had very limited access to good agricultural land, so it is possible that they had a different function from the lowland sites. As all of the high land sites occur on crests and ridges, it appears that the sites were located to have a good visibility of the surrounding area. This might be linked with reasons of security, or to control the maritime traffic along the Catalan coast. If this is the case, however, it is odd that similarly located sites were not found in the region of Tarragona.

3.2 The distribution of rural sites on distance from towns

During the Roman period two main towns existed in the Maresme, Baetulo (modern Badalona) to the west, and Iluro (modern Mataró) to the east. In the Republican period both towns had the status of *oppida civium Romanorum*, and were later given the status of *municipia* (Prevosti Monclús 1991, 135). The ceramic evidence shows that Baetulo was very probably founded at the end of the 2nd or beginning of the 1st century BC (Guitart 1976), and it is very likely that Iluro was founded at roughly the same time, though it has not been studied as well as Baetulo. The similar date of foundation of the two towns is suggested by the similar topographical situation along the coast and the identical status attributed to the two towns by the Latin sources (Prevosti Monclús 1991, 135).

The distribution of rural sites in relation to the position of Baetulo and Iluro was studied. A cost distance surface was created starting from the two towns using a friction surface derived from the slope of the Maresme. The resulting image was then divided into a series of distance bands corresponding to 5 km eq[1] and the number of rural sites dating to each period (Iberian, Republic, the reign of Augustus, the Early Empire, the 3rd century and the Late Empire) used in the analysis was extracted using the ***Idr_Vals*** program and these data were then used to calculate a series of Kolmogorov-Smirnov one-sample tests. The test was significant in all cases at the 5% level of significance except for the Iberian rural site distribution, showing that the new rural sites founded after Romanisation were centred on the two towns rather than having a dispersed pattern in the countryside as it was the case in the Iberian period.

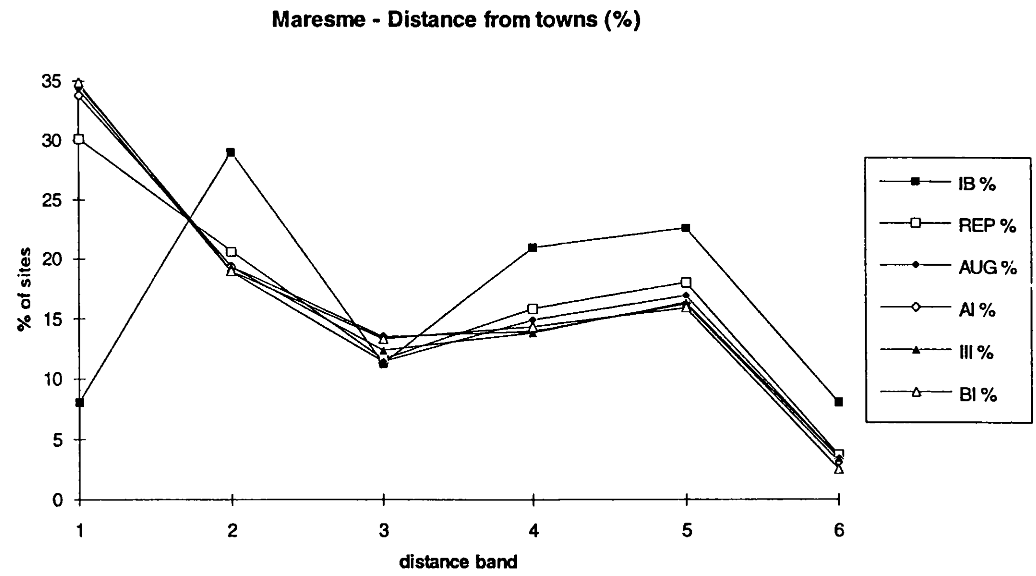

Figure 6.4 The distribution of rural sites on 5 km eq cost distance bands from the Roman towns

[1] See chapter 4 (The Guadalquivir Valley), section 5.1 for an explanation of km equivalent (km eq) cost units.

The relationship between the rural sites and the towns can be visualised and studied further by plotting the proportion of sites on cost distance bands from the towns. Figure 6.4 shows the graph of the percentages of rural sites on 5 km eq distance bands from the towns for each chronological band considered in the analysis. The curve of the graph after Romanisation shows that the distribution of rural sites in relation to the towns stayed pretty much the same until the Late Empire. The Iberian pattern present before Romanisation, on the other hand, is markedly different from the pattern for the Roman periods, showing a much smaller number of sites close to the towns and a more dispersed pattern away from the town (or rather, the location where the Roman towns were to be built). This suggest that the two Roman towns were not built over pre-existing important Iberian settlement or locations that were of high importance to the Iberian organisation of the landscape. It is possible to remove from the graph the curve referring to the distribution of Iberian sites and compare the proportions of rural sites in relation to the towns at different times of the Roman control of the area.

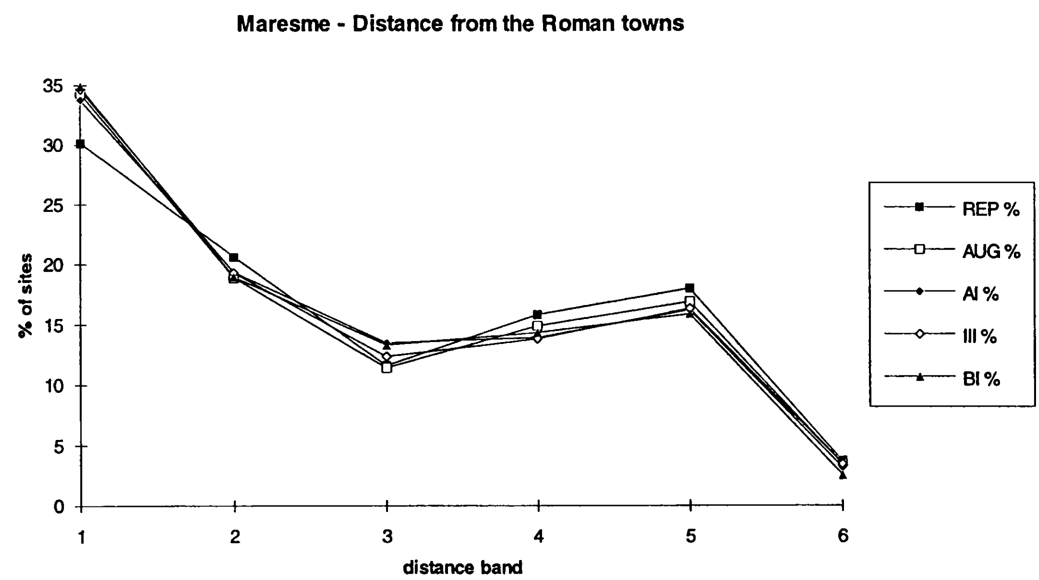

Figure 6.5 The distribution of the rural sites dating from any of the Roman periods on 5 km eq cost distance bands from the Roman towns

Figure 6.5 shows the graph with only the curves referring to the sites in existence during the Roman periods. Remarkably little variation is visible between the curves, with the possible exception of the distribution of sites dating to the Republic, which appear to be concentrated closer to the towns than in the following periods. This is consistent with the hypothesis that a large number of new rural sites were created in the Maresme after the foundation of the two towns. This pattern of no variation is confirmed by the results of the Kolmogorov-Smirnov one-sample tests discussed above: apart from the Iberian period, in all the other periods the rural sites pattern was significantly distributed around the towns. These results differ from the picture which had emerged for the region of Tarragona, where a clear shift in the organisation of the countryside immediately following the reign of Augustus was visible from both the results of the Kolmogorov-Smirnov significance tests and the detailed study of the graphs of the distribution of proportion of sites on distance from Tarragona at different times. This is consistent with the view that Tarragona was an important centre and the development of its countryside reflected the fortune of the town, which became a *conventus capitalis*, while the Maresme was a marginal area which received input right at the time of the first large scale Roman settlement in the area but then froze and no transformation occurred in its countryside over the different periods of Roman rule. Interestingly, the shape of the curves in the graph of the proportions of sites from all periods on cost distance from the towns in the Maresme resembles the curves of the proportions of rural sites on distance from Tarragona for the periods following the reign of Augustus, showing that despite the fact that these two areas (the Maresme and the *Ager Tarraconensis*) are so close geographically, the early stages of the Romanisation of the countryside in the two areas were different. It is possible that the different development of the countryside of Tarragona was linked to the fact that the town was created over the site of a pre-existing important Iberian site and the organisation of the countryside until the mid 1st century AD still reflected the Iberian organisation of the land, while this does not appear to be the case in the Maresme.

3.3 The distribution of rural sites on distance from the Via Augusta

The Via Augusta ran parallel to the coast from north to south (Pallí Aguilera 1985). The position of the road is also confirmed by two milestones found at Vilassar de Mar and Arenys de Munt (Prevosti Monclús 1991, 135). The importance of the Via Augusta to the placement of rural sites at different times was tested by creating a cost distance surface using a friction surface derived from the slope gradient map of the area and a digitised map of the Via Augusta as starting point. The cost distance surface was then reclassified into a number of cost distance bands 1 km eq wide and the number of rural sites occurring on each band in every one period counted. These data were used to carry out a series of Kolmogorov-Smirnov tests to see whether there is a significant relationship between the position of the Via Augusta and the rural sites at any one time. The tests for all periods were significant at the 5% level of significance, showing that the rural sites were distributed in relation to the position of the Via Augusta.

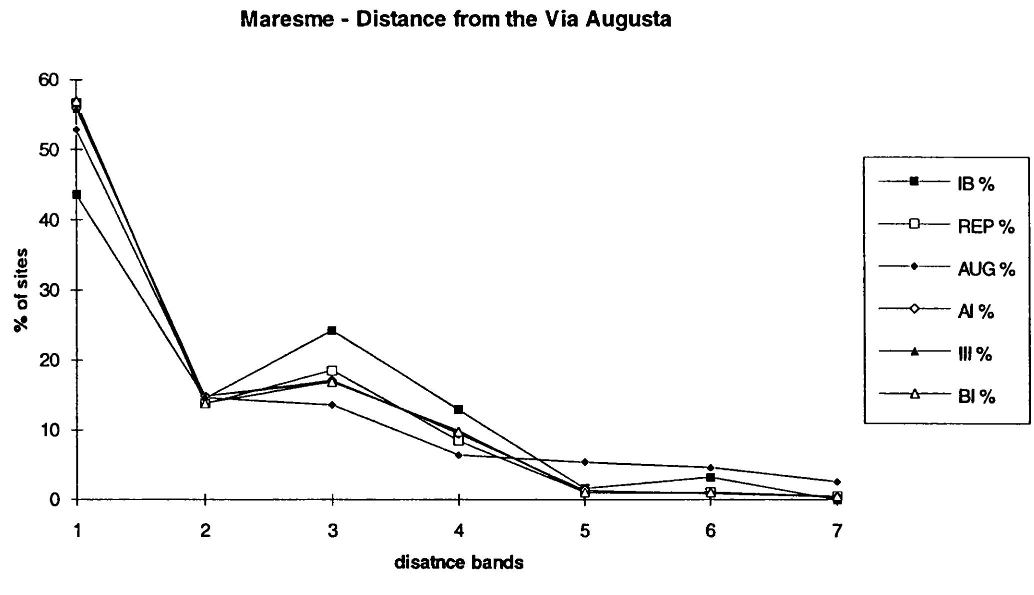

Figure 6.6 The distribution of rural sites on 1 km eq cost distance bands from the Via Augusta

These results, however, must be treated with some caution, as almost the whole of the Roman settlement in the Maresme is found in the thin and long coastal strip, which was followed by the Via Augusta. The fact that the Iberian rural site distribution tested significant as well can be interpreted in two ways:

i The Via Augusta ran along the track of an ancient path which was used well before the arrival of the Romans.

ii The Iberian sites were distributed along the coast, like the Roman ones, and the results of the Kolmogorov-Smirnov tests are biased by the land form, which influences both where the road could go and where the sites could be located.

It is very likely that a road existed along the track of the Via Augusta before the arrival of the Romans, as the coastal Maresme represents the ideal corridor along which a road connecting the north to the south would run, keeping in mind that Cissis, to the south of the Garraf massif, was an important Iberian site and was located along the course of the Via Augusta. It is therefore likely that a combination of the two elements listed above would be the best explanation for the results of the Kolmogorov-Smirnov tests. The graph in figure 6.5 shows the curves of the proportions of rural sites on distance bands from the Via Augusta. Apart from the fact that a slightly smaller proportion of Iberian sites appear to be close to the Via Augusta and a higher proportion away from it than it is the case for the Roman sites of all periods, very little variation seems to occur overall. This pattern is very likely to be biased by the shape of the land with the thin coastal strip along which the Via Augusta ran and the sites were located, so it would be dangerous to try and interpret these results from an archaeological point of view.

4 The pottery assemblage

Since the presence of some pottery types was recorded in the area surveyed by Prevosti Monclús, but not in the area the information for which was made available by the Generalitat de Catalunya, and the other way round, it was necessary to reorganised the data and only take into account the information about the pottery types whose presence was recorded in the two data sets. Table 6.1 shows the total count of sites with a pottery type in the area surveyed by Prevosti Monclús[2] (only those sites included in the map considered), in the area surveyed by the Generalitat de Catalunya and the total where the presence/absence of the pottery type was recorded for both surveys.

GENERALITAT	
Pottery type	Number of sites
ANF	68
ANF_I	8
ANF_IB	6
CA	44
COM	89
TSA	29
TSH	45
TSSG	41
TS	15
TSC_A	34
TSC_C	8
TSC_D	24
PFINE	17

PREVOSTI	
Pottery type	Number of sites
AP	4
CA	17
CB	27
TSA	21
TSSG	34
TSH	35
TSC_A	26
TSC_C	13
TSC_D	18
EST	11
TSH_T	2
PFINE	9

TOTAL	
Pottery type	Number of sites
CA	73
TSA	50
TSH	80
TSSG	75
TSC_A	60
TSC_C	21
TSC_D	42
PFINE	26

Table 6.1 The count of sites where a certain pottery type was found. ANF = generic amphorae, ANF_I = Italian Amphora, ANF_IB = Iberian amphora, CA = Black Glaze A, CB = Black Glaze B, COM = Common coarse ware, TSA = Terra Sigillata Aretina, TSH = Spanish Terra Sigillata, TSSG = South Gaulish Terra Sigillata, TS = Generic Terra Sigillata, TSC_A = Terra Sigillata Chiara A, TSC_C = Terra Sigillata Chiara C, TSC_D = Terra Sigillata Chiara D, PFINE = Thin-Walled ware, AP = Pre-Black Glaze imported pottery, EST = Stamped Terra Sigillata, TSH_T = Late Spanish Terra Sigillata. Note that in the third table category CA includes both CA and CB from the second table.

[2] This information was derived from the tables in Prevosti Monclús 1981b, 285-289.

It is immediately evident that the proportion of sites with early wares is much higher than in the Guadalquivir Valley. The other interesting difference is that the number of sites with Terra Sigillata Chiara A is higher than the number of sites with Terra Sigillata Chiara D, as it was the case in the region of Tarragona, but not in the Guadalquivir Valley. This evidence seems to suggest that the early types of pottery were more widespread in Catalunya than in the south-west, probably as a consequence of contacts occurring at different times, earlier in the north-east than in the south-west. The fact that the dominant type in the Guadalquivir Valley is the Terra Sigillata Chiara D, while in the region of Tarragona and in the Maresme Terra Sigillata Chiara A is present in higher quantities than this ware and the dominant types are South Gaulish Terra Sigillata and the very early types, suggests that the pattern of import was different in the south-west and the north-east. It appears that foreign (Roman and Italian) pottery started being imported in Catalunya earlier than it was in the Guadalquivir Valley, but it also seems that imports were more common in the Guadalquivir Valley in later periods than they were in Catalunya.

Component	Iterations	Norm	Eigenvalue	% Inertia	Cumulative
I	76	0.085	0.186162	23.2	23.2
II	39	0.078	0.165210	20.6	43.8
III	53	0.075	0.127085	15.8	59.6

Table 6.2 The Correspondence Analysis information for the analysis of the pottery assemblage of the rural sites in the Maresme

The information about pottery types recorded in the two surveys was used to create a binary presence/absence table of pottery types in sites. This table was then used to perform a Correspondence Analysis[3]. Table 6.2 contains the percentages of inertia and other information about the analysis of the assemblages of the rural sites of the Maresme. Figure 6.7 shows the variable plot for the presence/absence pottery data on the first two components. Spanish Terra Sigillata and South Gaulish Terra Sigillata are close together in the plot, while Arretine Terra Sigillata is away from them, as it was the

[3] See chapter 5 (*The region of Tarragona*) for a discussion of the use of Correspondence Analysis with binary data.

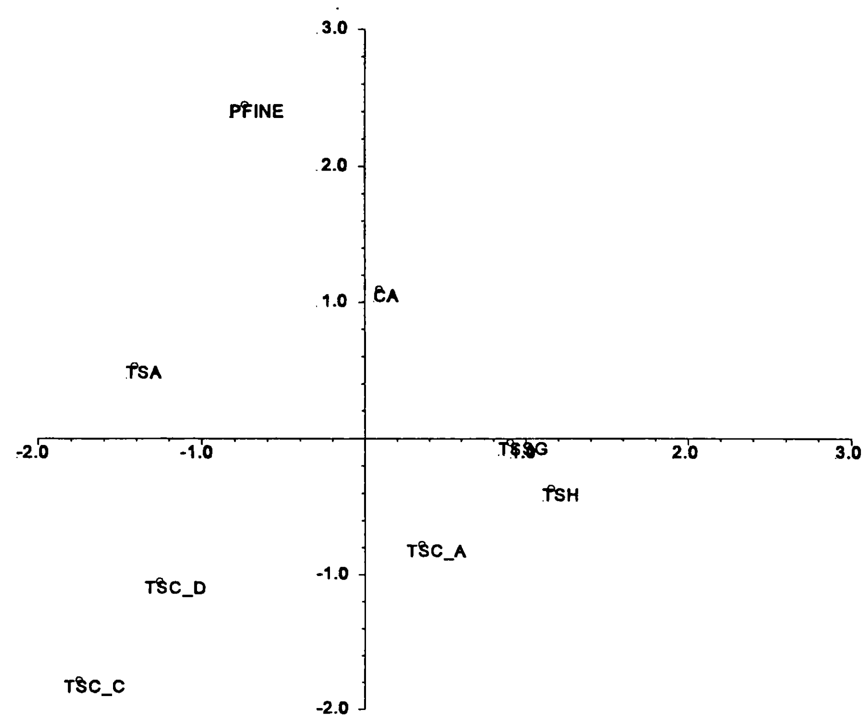

Figure 6.7 The variable plot (pottery types) of the Correspondence Analysis carried out on the assemblage data of the rural sites in the Maresme. Component I is on the horizontal axis, Component II on the vertical axis.

case in the Correspondence Analysis of the data from Tarragona. This indicates that despite the fact that Spanish Terra Sigillata was a local imitation of both Arretine Terra Sigillata and South Gaulish Terra Sigillata, the Spanish Terra Sigillata found in the Maresme is distributed in a way similar to the South Gaulish Terra Sigillata, rather than the Arretine Terra Sigillata. This pattern is also visible in the Correspondence Analysis of the pottery data from the region of Tarragona, indicating that a similar process of distribution of pottery types was at work in the two areas of Catalunya.

The other feature emerging from the analysis of the variables on the first two components is that the Thin-Walled ware is separate from all the other pottery types in terms of its distribution. This feature is consistent with the results of the Correspondence Analysis carried out on the data sets from the region of Tarragona and the Guadalquivir Valley, showing that in Spain the locally produced Thin-Walled ware tends to have different distribution from the other (both imported and local) pottery types. Not enough work has been carried out on Thin-Walled ware at present to try and assess why this might be the case. The three Terra Sigillata Chiara subgroups are found in the same area of the graph, though subgroups C and D are closer together than they are to subgroup A. This pattern is similar to the one observed in the region of Tarragona, but it is completely different from the one observed in the Guadalquivir Valley, where a close relationship between types A and D and their clear separation from type C was noticed.

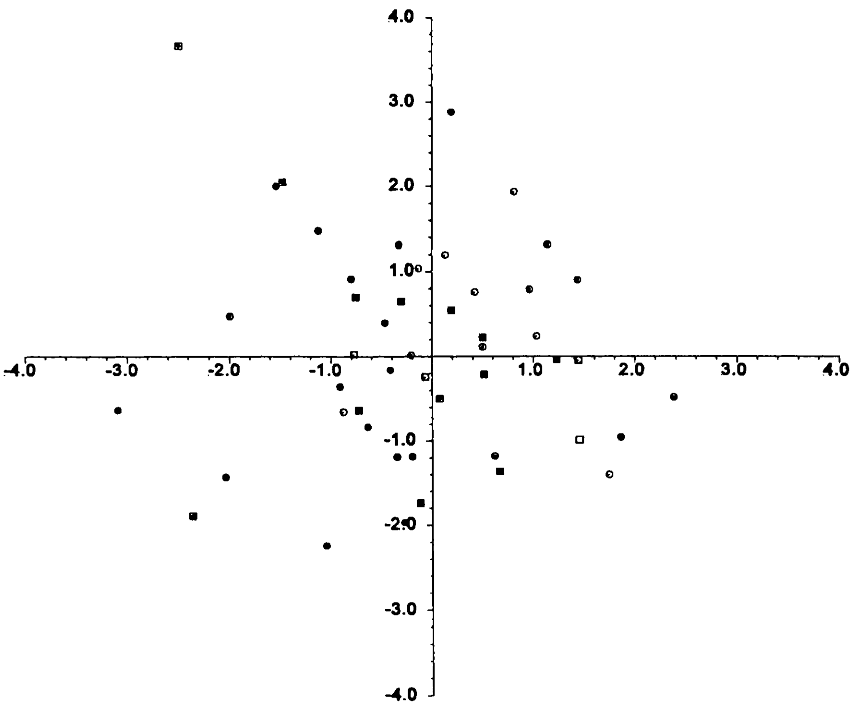

Figure 6.8 The object plot (sites) of the Correspondence Analysis carried out on the rural sites from the Maresme. Component I is on the horizontal axis, Component II on the vertical axis

The plot of objects on the first two components (figure 6.8) does not suggest any direct relationship between the location of the sites and the type of pottery assemblage they contain. The distribution of the pottery types in the Maresme appears to be uniform, without specific concentrations in any place. This pattern of no variation ties in very well with the results of the analysis of the relationship of rural sites to elevation, distance from the Via Augusta and from the Roman towns, which showed a constant settlement pattern throughout the Roman periods. Though there appear to be some sort of covariation between certain pottery types within the assemblages, such as South Gaulish Terra Sigillata and Spanish Terra Sigillata or Terra Sigillata Chiara C and Terra Sigillata Chiara D, the sites containing the assemblages with these characteristics do not appear to be located in specific positions with regards to elements such as the administrative centres and the main communication and import/export route.

5 Conclusions

Miret *et al.* (1991) state that the Garraf massif does not just divide Catalunya into two separate geographical parts, but into two separate cultural units as well, with visible different characteristics before Romanisation occurred. To the south of the Garraf massif (the region of Tarragona) was the land of the Cessetani, with capital Kesse (or Cissis), on the immediate small area of which Tarraco was founded, while to the north of the massif (the Maresme) was the land of the Laietani. One of the main differences between the two areas noticed by Miret *et al.* (1991) is that Iberian villages occur at much higher elevation in the Maresme than in the region of Tarragona, despite the fact that the difference in the proportion of high land in the two areas is marginal. This pattern of sites on high elevation continued in the Roman period, with Iberian sites on high ground surviving and new ones (though very few) being created under the Romans. If the reason for villages to be on high ground was one of defence, as it seems reasonable to assume since this type of location would make access to resources more difficult, such a situation seems to continue during the Roman time. It can be argued that the reason for new sites to be created on high ground was that the Roman colonists had taken over the best low lands and the Iberian population found itself confined to less desirable locations. Overall, it seems that the Iberian pattern of settelment on high ground continued in Roman times. Overall, Prevosti's idea that there were two clusters of sites with different functions, one occurring around 100 metres of elevation and th eother on the coast, was not reflected in the results of the analysis. Conversely, the sites seem to be concentrated in the very low lands and decrease gradually with the increase in elevation. Apart from the sites located at vantage high points, the elevation at which the other rural sites occurred appears to have been dictated by the land form.

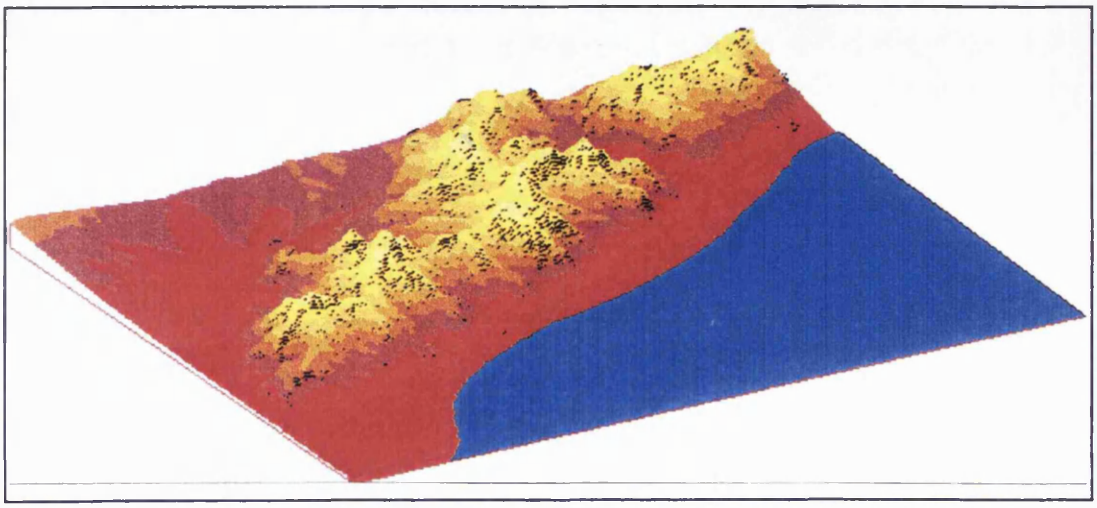

Figure 6.9 A Digital Elevation model (DEM) showing the land form of the Maresme. The human settlement was concentrated along the thin coastal strip, along which the Via Augusta also ran. The elevation has been exagerated by a factor of 2 to emphasise the distinction between the coastal low land and the hills.

The graph of proportion of sites in distance bands from the Roman towns shows a difference in the shape of the curve representing the distribution of Iberian sites and those representing the settlement patterns for the following Roman periods. The two towns of Baetulo and Iluro were both created at the end of the 2nd century/beginning of the 1st century BC and at this time a reorganisation of the landscape occurred. Several new sites were created centred around the two towns. The curve of the proportion of sites on cost distance bands dating from the Republic shows that at that time the settlement pattern was slightly less centred on the towns than it was in the following periods, marking the point of transition between the Iberian settlement pattern, which was scattered rather than centred, to the Roman one. The curves for the following periods all have basically the same shape, indicating that no further variation occurred in the pattern of distribution of rural settlement around the towns after the transition from the Iberian to the Roman organisation of the landscape. One interesting feature is that the settlement pattern does not show any variation after the 3rd century AD, when both Baetulo and Iluro lost their administrative role to Barcino (Barcelona), to the south-west. This seems to indicate that the role of the towns in influencing the organisation of the landscape always was very marginal in the Maresme. In the region of Tarragona, after the decline of the town, the Late Empire is characterised by the appearance of a more widespread pattern of rural settlement, though still based on the pattern from the preceeding period, while no such thing seems to occur in the Maresme. Tarragona was a more important Roman centre than either Baetulo or Iluro and its decline was more likely to have an influence on its hinterland than it would be the case for the two towns in the Maresme. The reorganisation of the countryside visible after the reign of Augustus in the region of Tarragona is completely absent from the Maresme. This reinforces the hypothesis that such reorganisation was initiated by the creation of the colonia at Tarragona under Caesar with subsequent redistribution of land to the new settlers, while no coloniae were ever created in the Maresme.

As it happens for the distribution of sites on cost distance bands from the two Roman towns, very little variation is visible in the graphs of the proportions of sites dating to different periods on cost distance bands from the Via Augusta, the main communication route along the coast of Catalunya. It is worth pointing out that other local routes existed in the Maresme in Roman times but, as very little work has been done on this subject, no information is available on where they might have run, except from the possible route leading north to Semproniana from Iluro (Prevosti Monclús 1991, 135). Since the Via Augusta ran parallel to the coast in the thin coastal strip which contained most of the sites used in the study, it is not surprising that the Kolmogorov-Smirnov tests show that a significant relationship exists between the position of rural sites at every time and the Via Augusta. The shapes of the curves for the different periods in the graph only suggest that a smaller proportion of the Iberian sites were found close to the Via Augusta and a larger proportion were found away from it than it was the case for the rural sites dating to any of the Roman periods. This pattern can be reconducted to the pattern of distribution of the Iberian sites, found more widespread in the countryside than the Roman ones, which were centred on the towns.

The fact that no variation seems to occur in the Maresme after the first Roman settlement suggest a stable situation in all periods. The graphs show a clear break with the preceeding Iberian rural settlement pattern and the establishment of a new organisation of the countryside at the time of foundation of the two towns. From this point onwards until the Late Empire, the Roman organisation of the countryside survived without variation. Even when the two towns of Baetulo and Iluro were given the status of *municipia*, the increased importance of the towns (reflected in architectural redevelopment at Baetulo, see Keay 1988, 59) is not mirrored in the organisation of the countryside, as it was instead the case in the countryside of Tarragona when this town acquired the status of *colonia*. The creation of a *colonia* involved in most cases the division of the land into allotments for the colonists (*centuriation*), while this was very rarely the case when a *municipium* was created (*ibid.* 68). If the shift in the organisation of the countryside visible in the region of Tarragona was due to *centuriation*, it appears that *centuriation* never happened in the Maresme. Moreover, after the decline of Roman Tarraco the rural settlement acquired a more dispersed nature, as the main centre of gravity of the area had caesed its function, while this dispersion of the rural sites in the Late Empire is not visible in the Maresme, despite the fact that the two towns of Baetulo and Iluro had effectively lost their role to Barcino (modern Barcelona) by the late 3rd century AD. Overall, it seems that the two Roman towns of the Maresme never had on their countryside the impact that Tarraco had on its, perhaps because they never became as important as Tarraco was. The Ager Tarraconensis was one of the most important areas of Roman Spain in terms of production and exports, while the Maresme never quite acquired such an importance, which probably influenced the way the two neighbouring regions developed. The circumstances of Tarraco had a much stronger influence on the rural settlement pattern of its hinterland than it was the case for Baetulo and Iluro, which is why fluctiations are visible in the rural settlement of the Ager Tarraconensis while the Roman country settlement of the Maresme was stable from the Republic to the Late Empire.

5.1 The model

In the 2nd century BC the Iberian population of the Laietani inhabited the Maresme to the north of the Garraf massif and their settlement was organised with a network of villages, often located on high ground. Second order sites were found on lower gound, positioned to exploit the resources of the area, while smaller third order sites were centred around them. The economy was one of peasant production, but generating enough surplus to import foreign pottery such as the Black-Glaze type and Punic amphorae (Miret *et al.*, 1991, 47-50). In the 2nd century BC the Romans entered this area as a consequence of the campaign of M. Porcius Cato, though it does not appear that the conquest was violent. The pre-existing Iberian organisation survived well into the Republican period, despite the fact that at the end of the 2nd century BC the two Roman towns of Baetulo and Iluro were created, together with a number of new rural sites centred around the two Roman towns. The lowland locations were preferred by the Romans, as these enabled them to exploit the better land and to have easy acces to the Via Augusta, the main coastal communication route leading to Tarragona and to the south-west on one side, and to France and Italy on the other.

With the appearance of the new Romanised landscape, the shift towards a pattern centred on the towns shows that a different type of economy developed. While the distributed pattern of the Iberian period is typical of the peasant mode of production, characterised by groups of small farmers practising subsistence agriculture (Miret *et al.* 1991, 50), the villa system was geared towards producing surplus that could be exchanged or sold in the towns, which acted as the main markets. It is possible that the pre-existing Iberian system and the Roman system co-existed for some time at the beginning of Romanisation. Several of the Roman rural sites found in this area were built over a pre-existing Iberian settlement, showing a continuity in the use of the land which also suggests an integration of the Iberian population and possibly their production strategies within the Roman economy.

The rural settlement pattern continued unchanged after the decline of the two towns in the 3rd century AD, indicating that by this stage a fossilisation of the countryside had occurred. Fewer rural sites existed in the Late Empire than they did in the earlier periods, but the decrease in the number of sites was a very small one, as only a few disappeared, while no new ones were created. While something similar happens in the region Tarragona (*ie* there are marginally less sites in the Late Empire than there are in the earlier period), it seems that in this area a few sites disappeared and new ones were created away from the towns, achieving a more distributed pattern in the countryside, which is not visible in the Maresme.

Chapter 7

Conclusions

This chapter reviews the information and the analyses which have been presented in the preceding chapters and provides a possible interpretation of the results of the study carried out on the archaeological data. The research project presented in this thesis had two main aims:

1. To investigate to what extent GIS could be used to perform landscape analysis of archaeological settlement and in what way computer techniques could help to handle and analyse non-systematically collected archaeological data.

2. To investigate the development of Roman settlement patterns in parts of Spain and try to relate the observed variation to social, political and economic transformations taking place between the 2nd century BC and the 7th century AD.

This chapter shows that the two goals listed above have been fully met.

The first part of this chapter summarises the results obtained in the analysis of the archaeological data described in chapters 4 to 6. The results of the studies carried out on the distribution of ceramic types in the three study areas are also discussed. Finally, a comparison of the development of the rural settlement in the study areas is made and general trends in the data sets are discussed.

The second part of this chapter summarises the various uses GIS in general and Idrisi in particular have been put to during the course of the research project. The advantages of applying GIS to landscape archaeology are listed and discussed and so are the disadvantages. The Idrisi package is then examined in detail and its performance for the specific tasks it has been put to is assessed. Finally, the concluding section lists a few consideration on whether there is scope for continuing to use GIS in landscape archaeology. The approach used in this project is also compared with that of earlier large scale studies.

1 The summary of results

This section reviews the analysis carried out on the site data from the three Spanish data sets and summarises the results obtained for each of the study regions. Each subsection reviews the results of one of the study area. The following section reviews the results of the analysis of the pottery assemblages and the final section of this chapter compares the three study areas and tries to detect general trends and differences in the Romanisation and Roman control of different areas of Spain. The information and the analysis discussed in this section are covered in greater detail in chapter 4, *The Guadalquivir Valley*, chapter 5, *The region of Tarragona* and chapter 6, *The Maresme.*

1.1 The Guadalquivir Valley

Because the data from the Guadalquivir Valley came from a number of different sources and had been classified and stored in a variety of different ways, a standardisation procedure had to be worked out and applied to the data before they could be analysed. The standardisation method used produced a site database in which all the rural sites had been marked as either low status sites or high status sites and had been divided into three chronological bands, the Republic, The Early Empire and the Late Empire. The site data were split into subsets by status and by period and were tested against information about the landscape of the Guadalquivir Valley such as the soil type and the agricultural potential. The data subsets were also tested against the distance from the navigable rivers, which represented one of the main communication routes inside the Guadalquivir Valley in Roman times, the Roman road network and the distance from the Roman towns, which were political, economic and administrative centres. The results of these tests showed that socio-economic factors such as the easy access to the urban centres and the communication routes had a bigger influence on the position of rural settlement in the countryside than environmental variables. The position of the waterways and the urban centres proved to be relevant to the location of both low and high status rural sites in the Early Empire and Late Empire, but not in the Republic Similarly, the high and low status sites did not prove to have a significant relationship with the position of the Roman roads in the Republic, which was to be expected, as the Roman road network only became fully operative during the Early Empire.

During the Republic, the new Roman high status sites were built close to the urban centres, while low status sites appeared to divided in two groups, a group concentrating around the towns together with the high status centres, and others scattered in the countryside (figure 4.3). At this time very few Roman sites existed, showing that the Romanisation of the Guadalquivir Valley was a slow process at the beginning and only reached its full peak in the first years AD. The Early Empire is the period in which the largest number of sites were in existence at the same time. The curve of the proportion of sites on distance from the Roman towns for this period (figure 4.4) is similar for the high status sites and the low status sites, showing that at this stage the distribution of both low status and high status sites was centred around the towns. High status sites in this period are found further away from urban centres in the countryside than they were in the Republic, indicating that a double process had been at play since the preceding period. The high status rural sites, which were concentrated around the towns during the Republic, spread in the countryside, while the low status sites, which were scattered in the countryside during the Republic, show a proportional increase in the cost distance bands closest to the urban centres. This process of expansion of high status sites into the countryside away from the administrative, political and economic centres reflects the appearance of a large scale organisation and exchange network covering the whole of the Guadalquivir Valley and probably centred on the towns on the navigable rivers and the waterways, as these were in ideal position to act as export centres of the goods produced in the valley to the external world.

In the Late Empire (figure 4.5), a new transformation of the high and low status settlement pattern takes place. The proportion of high status sites away from the towns increases, with the distribution curve showing a bimodal distribution, one group of sites close to the urban centres and one away from them. This pattern probably reflects the creation of large estates in the countryside, as the distance between the high status sites away from the towns increases as well. On the contrary, the high status sites close to the towns, especially those on the navigable rivers, are as close to each other as they were during the Early Empire, indicating that the variation in high status rural settlement from the earlier period interests mainly the southern part of the Guadalquivir Valley. Almost all of the sites which came into existence during the Late Empire are located along the navigable rivers. This also spells out a differentiation between the two halves of the valley in terms of their development in the later Roman period. The two peaks in the graph of the proportion of high status rural sites away from towns (figure 4.5) can be interpreted as the sites occurring around the towns along the Guadalquivir and the Genil rivers and those found to the south, away from the navigable rivers. The graph of the low status sites reflects roughly that of the preceding period, with the difference that a smaller proportion of sites are found closer to the urban centres. The division between rural sites close to the towns along the navigable rivers and those scattered in the countryside in the south is not visible in the case of the low status sites. This pattern shows that in the Late Empire the large scale exchange system based on the towns which was at work during the Early Empire had collapsed in the south and a large number of small (low status) rural sites had disappeared as a consequence or had been incorporated into the large estates. The majority of the low status sites which disappeared in this period were located in the south of the Guadalquivir Valley. In the north, along the waterways, a number of low status sites disappeared, showing a generalised decrease in site number, but not as many as in the south. The survival rate of low status sites reflects that of high status sites, in terms of geographical differentiation, with the main difference that a few high status sites survived in the south of the Guadalquivir Valley, while very few low status sites did.

More information about the development of the Roman society in the Guadalquivir Valley was obtained from the study of the pottery assemblage. The number of sites with a specific pottery type occurring in the territory of each Roman town were counted and this information was used to carry out a Correspondence Analysis to try to detect spatial patterns of ceramic distribution within the province of Seville (see section 5). The results of the Correspondence Analysis allow inferences to be made about the way different pottery types were distributed and also gave insights into a possible continuation of the Iberian way of life after Romanisation.

The position of Terra Sigillata Chiara A and D in the variable plot (figure 4.7) suggests that these two types were being imported along the same routes and quite separately from Terra Sigillata Chiara C, which was being produced in central Tunisia, while subtypes A and D were being produced in northern Tunisia. This situation reflects separate commercial contacts between the Guadalquivir Valley and different areas of Tunisia. The closeness of Black Glaze pottery and Arretine Terra Sigillata in the plot can be a reflection of the persistence of the Iberian exchange network inside the Guadalquivir Valley until at least the second half of the 1st century AD. Spanish Terra Sigillata occurs in the plot close to Arretine Terra Sigillata, whose form it imitated in this region and if it was redistributed along the Iberian network, as seems possible given the closeness of this type to Arretine Terra Sigillata and Black Glaze in the plot, the Iberian network might still have been in use as late as the 2nd century AD.

The geographical position of the towns `explained' by the different pottery types reveals more clues as to the different distribution of pottery types in the Guadalquivir Valley (figures 4.9 and 4.10). The towns whose countryside assemblage was heavily influenced by Terra Sigillata Chiara A and D are found along the Guadalquivir river, therefore it is very likely that these wares reached the valley along the waterways rather than along the roads. Terra Sigillata Chiara A was being imported in the Guadalquivir Valley at the same time as the Arretine Terra Sigillata, so it appears that at this time two redistribution networks were in existence at the same time: the original Iberian one along which Black Glaze pottery and then Arretine and Spanish Terra Sigillata were being imported, and the new Roman one along which Terra Sigillata Chiara A and D were imported and redistributed. The distribution of terra Sigillata Chiara A and D is biased towards the northern part of the Guadalquivir Valley and is concentrated along the Guadalquivir and the wealthier towns found on the banks of the river, while the other pottery types are found in the southern part of the valley, which was characterised by a more dispersed rural settlement pattern and evolved into large estates in the Late Empire. These results reflect a process of colonisation of the countryside the first stage of which consisted of a spatial division between the high and low status sites, possibly due to different privileged access to the administrative and commercial centres (the Roman towns) for the high status sites. At this early stage, the Roman and Romanised Iberian towns of the Guadalquivir Valley offered the basic services the rural sites needed to prosper, so it is also possible that the Roman rural sites close to the towns were classified as high status because they were more likely to receive diagnostic material from the urban centres. The Iberian towns had served as redistribution centres before Romanisation and it is likely that they retained this role after Romanisation, serving the local Iberian population and probably the new Roman rural sites as well, while the Roman towns acted as the reference point of `Romanity' in the valley to the newly founded Roman rural sites. It is therefore understandable that the sites closest to the urban centres would do better than those away from them. The question remains open whether the wealthier farmers had the choice to settle close to the towns at this stage, or whether the sites closest to the towns were wealthier because of their privileged access to the resources and services offered by the towns. It is not unlikely that a combination of these two factors was at play. The Iberian population at this time still retained much of its organisation from before the time of Romanisation and it is likely that the local elite used their contacts with the Romans as a point of prestige. Since the Iberian elites were living in the towns, the concentration of high status Roman sites around the urban centres (which were almost all of Iberian origin) might have been promoted by members of the Iberian elite themselves, who wished to be seen to be associated with the Romans.

In the early Empire the similar and more widespread distribution of the high and low status sites shows that the initial impetus to the colonisation of the countryside by larger farm holders started from the towns and then expanded, while smaller rural sites, which had been scattered in the countryside before, started to concentrate around the towns, possibly as satellites of high status sites. This period witnesses the peak of Roman control of the Guadalquivir Valley, with a number of Roman and Romanised towns serving as political, economic and administrative centres, a well developed network of roads and a fully functional access to the waterways. The Iberian population had become integrated into the Roman society, as it is also shown by the granting of municipal status to a number of Iberian towns, at the same time as several *coloniae* were being founded. In this period the local economy was producing enough surplus to be one of the main export centres of olive oil of the Roman Empire, serving mainly the city of Rome and the military camps along the British, Rhine and Danube frontiers (Keay 1988, 176). The olive oil was probably produced in large scale in the high status sites and also in the (smaller) low status sites. It is possible that the process of transporting the oil produced for export to the urban centres on the communication routes, from where it was then exported, was carried out in stages. First the oil was carried from the low status sites to the high status sites. This could easily have been the case if the low status sites had a relationship of dependence on the high status sites, as it appears to be the case in the Late Empire. At the high status sites the production of the low status sites was put together with the production of the high status site and carried to the nearest export town. This method would have reduced the overall cost of transporting the oil, the low status sites might have received immediate payment for their production from the high status sites (though at a lower price than what the high status sites would get from selling it directly), or they would at least have been able to have their production transported to the towns from the local high status sites at a lower price than if they had to organise the transport themselves, and the high status sites would have increased their margin of profit by being able to sell a larger amount of oil with each transport. It would have been very uneconomical for the low status sites to organise their own transport of their oil production to the towns from which it was being exported, so it is more likely that they sold it locally to the high status sites. This model allows for a local small scale exchange to be integrated into the larger scale exchange network of the Guadalquivir Valley. The towns along the Guadalquivir also served as centres for the import of fine foreign pottery. While the fine Terra Sigillata Chiara A was being distributed mainly to the towns along the Guadalquivir, other types were transported to the southern part of the valley along the network of Roman roads. The differentiation in the settlement and organisation of the valley clearly visible in the Late Empire (see below) had its roots in the different role of the towns and their influence on their hinterland dating to this period.

In the Late Empire the large scale exchange network within the Guadalquivir Valley collapsed, probably as a direct consequence of most of the Roman towns losing their central role of political, administrative and economic centres. The settlement pattern of the valley split into two separate trends. In the north, along the navigable rivers and the main roads, the rural sites continued in existence still centred on towns, showing that the urban centres had not completely lost their role, at least those close to the communication and export routes. Terra Sigillata Chiara D was the main pottery import in the valley at this time, and its distribution followed very closely that of Terra Sigillata Chiara A which was being imported during the Early Empire. Export of Baetican olive oil still continued at least at the beginning of the Late Empire, though the focus of export had shifted to the northern frontiers, due to the competition from the African production. A generalised decline of site numbers is visible in the south of the Guadalquivir Valley (chapter 4, *The Guadalquivir Valley*, section 4.3), in contrast to what was happening in the north. It is likely that at this stage the production of oil in the south had been concentrated in the large estates which had absorbed the low status sites, thus also determining the collapse of the small-scale oil trade. The take-over of the low status sites by the high status ones might have its roots in the relationship of dependence between the two classes of sites in the earlier period, with the weaker part giving in at the time of the economic slow-down of the 3rd century AD.

1.2 The region of Tarragona

The rural sites from the region of Tarragona were tested for their relationship to the Roman town of Tarraco, which was the main centre in the area. The site data were divided into chronological subsets corresponding to the Republic, the reign of Augustus, the Early Empire, the 3rd century and the Late Empire. The tests showed that a significant relationship existed between the rural sites and the position of Roman Tarraco in all periods except from the Late Empire. The graphs of the percentage of sites on distance from Roman Tarraco (figure 5.1) shows that a reorganisation of the countryside took place after the reign of Augustus. During the Republic and the reign of Augustus the rural settlement pattern was centred on the town, though this is more visible during the Republic, as under Augustus a movement towards the pattern which emerged in the Early Empire was already visible (figure 5.2). By the Early Empire the proportion of sites close to the town had decreased, while a new peak in the curve (figure 5.3) shows that a group of sites based away from the town has made its appearance, resulting in a more dispersed settlement pattern in the countryside. The shape of the graph remains similar for the settlement of the 3rd century and of the Late Empire, though in this last period the proportion of sites away from Tarraco increased, as it is also reflected in the result of the Kolmogorov-Smirnov test. A visible trend of movement away from Tarraco took place, which culminated in the almost scattered pattern of the Late Empire. A separate graph produced for the sites known to have survived in the Middle Ages shows a completely different pattern, with the sites being located preferentially at a certain distance from the town, followed by a sharp drop after this point. The Kolmogorov-Smirnov test for this period was not significant, showing that the process of movement away from the town which had started in Roman times had reached its culmination by this stage. Enough Medieval sites were recorded in the site database to allow a Kolmogorov-Smirnov test to be performed, therefore the result of the test is not influenced by sample size.

Two main Roman road existed in the region of Tarragona, the Via Augusta and the road leading from Tarraco to Ilerda to the north. The influence of the communication routes to the rural settlement pattern was tested by calculating a Kolmogorov-Smirnov tests for each of the chronological bands. The tests were significant in all cases, indicating that the roads always were very important to the shape of the rural settlement pattern in this area. The study of the graph of the proportion of rural sites on distance bands from the roads (figure 5.5) shows that the sites were less concentrated around the roads in the Republic than they were in later periods, indicating that the importance of the communication routes grew in time, possibly as a consequence of an expansion in the import/export from and to the region of Tarragona. The smaller concentration of rural sites close to the roads during the Republic can also be linked to the survival of the pre-Roman Iberian settlement. The curve for the site data dating to the Late Empire shows a higher concentration of sites away from the roads than in the earlier periods, reflecting the dispersed settlement pattern visible in the graph for the proportion of sites on distance from Tarraco (figure 5.3). The movement away from the roads might indicate a collapse in the export trade at this time, probably linked to the crisis of Tarraco reflected by the change in the settlement pattern around the town as well as evidence from the town itself (Keay 1988, 175). Despite the transformation in the distribution of sites around Tarraco after the reign of Augustus discussed above, a similar clear break between two periods in the distribution of the sites along the roads does not seem to take place, which reinforces the hypothesis that such a transformation was brought about by the acquisition of the status of *colonia* by Tarraco.

The elevation at which rural sites occur in the region of Tarragona was studied. While the distribution of Roman rural sites on elevation did not vary much across the Roman periods, a difference is visible between the Roman distribution and those of the Iberian period and of the Middle Ages (figure 5.6). A higher proportion of Iberian rural sites occurred on higher ground than it was the case for the Roman sites. A similar trend is visible for the distribution of sites dating to the Middle Ages, though a higher concentration of sites occurred on low ground than it was the case for the Iberian site distribution. There also are differences in the elevation at which sites occurred in the separate Roman periods, despite the fact that the differences between the Roman chronological bands are not as big as those between the Roman distribution for any period and the Iberian and Medieval distributions. A number of sites located on lower ground ceased to be used at the time of the reign of Augustus, but new ones came into use on low ground in the Early Empire. This process is probably linked to the reorganisation of the countryside of the *Ager Tarraconensis* following creation of the *colonia*. The fact that almost none of the sites on higher ground were touched by this process emphasises the agricultural reasons behind the reorganisation of the land, as cultivation was much easier on lower ground and the best soils were located on the low lands. In the Late Empire a number of sites disappeared and new ones were created indifferently on high and low ground, showing that cultivation and access to the town were no longer the primary influences to the location of rural sites, but other factors were at play, such as the necessity to protect the isolated country sites from the incursions of the Frankish tribesmen.

The study of the pottery assemblage of rural sites of the region of Tarragona by means of Correspondence Analysis showed that Black Glaze pottery, which was imported into the area until the 1st century BC, was quite widely distributed in the area, while Arretine Terra Sigillata, which was being imported in the immediately succeeding period was concentrated in the sites along the road to Ilerda. Black Glaze pottery was redistributed to the local Iberian population, while Arretine pottery was imported to Tarraco from which it was then carried to Ilerda and the other towns to the north rather than being redistributed in the sites of the immediate countryside. This fact reflects the role and importance of Tarraco as a large scale redistribution centre of imported goods to the Romanised areas immediately after its foundation. The local production of Spanish Terra Sigillata had a similar distribution to that of South Gaulish Terra Sigillata, which form it imitates in this area. Terra Sigillata Chiara A and South Gaulish Terra Sigillata, which followed Arretine Terra Sigillata chronologically, had a much wider distribution in the region of Tarragona than this type, indicating that after the reorganisation of the countryside in the Early Empire, Tarraco also acted as a redistribution centre to its immediate hinterland rather than just serving as a station in the process of importing goods and carrying them to the northern regions. Terra Sigillata Chiara C and D, which are later in date than Terra Sigillata Chiara A, South Gaulish Terra Sigillata and Spanish Terra Sigillata, are found in rural sites rather away from the town. Very few sites contained Terra Sigillata Chiara C and D, but a trend can be seen for sites away from the town in the Late Empire not only to be created or surviving better than those close to Tarraco, but also to be receiving diagnostic pottery while the sites close to the town were not. If the Late Empire witnessed the appearance of large estates, it seems that those away from the town were doing better than those close to it. This might be due to the fact that the denser concentration of sites close to the town which had characterised the preceding periods was a disadvantage after the town had lost its role. The sites which had to struggle to survive without being able to rely on the services of the administrative and economic centre were faced with a fiercer competition for space if they were located in the more densely settled are, while the sites which existed in the countryside further from Tarraco had more space over which to expand. The distribution of the amphorae roughly reflects that of the pottery types which were produced in the locations the amphorae come from. African amphorae have a similar distribution to Terra Sigillata Chiara C and D, while the Iberian amphorae are found roughly in the same places where Thin walled ware, a Spanish pottery type occurs. This suggests that the production of a specific area was imported to the region of Tarragona and distributed along the same redistribution routes, though it is also possible that the chronological factor is at play. Different areas might have exported to the region of Tarragona at different times, so that the pattern of distribution of goods coming from the same area is dictated by the location of the sites which were receiving imported goods at any one time.

During the first stage of the Roman control of the region of Tarragona, Tarraco was partially carrying out the function of Iberian Cissis-Kesse, over which it had been founded. Despite the creation of new Roman sites, the rural settlement still partially reflected the Iberian one and the redistribution network to the Iberian country sites was still at work, though this was a time of change, as it can be seen in the distribution pattern of the rural settlement during the reign of Augustus. At this time a movement of the sites towards Roman Tarraco took place, some sites disappeared away from the town and a few were created close to the town, showing that Tarraco was growing in importance at this time. The import of Arretine Terra Sigillata was carried from Tarraco to the northern area along the road to Ilerda, indicating that the town had acquired a new role as one of the stations in the international trade. In the following period, the Early Empire, Tarraco was established as an important centre of import/export and, at the same time, an organised rural settlement pattern centred around it appeared in the countryside, as a consequence of the redistribution of land following its elevation to the status of *colonia*, another testimony of the importance it had acquired by this period. At this stage Tarraco was acting not only as a station in the overseas import of foreign good, but it also was the centre of a local exchange network on which its countryside was dependent. Archaeological evidence shows that the town lost its function in the 3rd century AD. After this time a new settlement pattern appeared, no longer centred on the town and characterised by fewer sites and greater distances between them. Overall, the sites away from the towns seem to have feared better than those close to the town, suggesting that those sites further from Tarraco had the space to expand and evolve into larger estates, possibly incorporating smaller rural sites, while the sites close to the town found themselves competing for space and not being able to rely any longer on the services the town had formerly offered. Paradoxically, it is possible that the rural sites further away from Tarraco were already better equipped to survive without the services of the urban centre before the crisis of the 3rd century because they had more limited access to the town due to their distance from it during the Republic and the Early Empire.

1.3 The Maresme

The archaeological site data from the Maresme were tested for their relationship to the elevation the sites occurred on, the two Roman towns found in the area and the main Roman road crossing the region, the Via Augusta.

The elevation at which rural sites were found in the Maresme was studied by means of Kolmogorov-Smirnov one-sample test. The study area was subdivided into a series of elevation bands, each 50 metres wide. The data were tested separately for each of the chronological bands considered in the analysis and a significant relationship was found for all periods between the elevation and the location of the sites. The graph in figure 6.1 shows that a higher proportion of Iberian than Roman sites existed on high ground and that the vast majority of Roman sites were found on low ground, in the thin coastal strip at the foot of the hills. A few Roman sites did exist on high ground, as shown in figure 6.3, but these were located on vantage points on crests and ridges and were obviously positioned to gain the best possible view of the coast below. it is possible that these sites existed for protection, but almost all of them first appeared in the Republic and carried out throughout the Late Empire, spanning several centuries in which it is unlikely that the Maresme might have been threatened by external enemies. The position of the high sites facing towards the sea makes it possible that these had the function of lookout posts to keep an eye over the maritime transport and trade, perhaps to spot pirate vessels which might have endangered the merchant ships.

The Kolmogorov-Smirnov one-sample test carried out on the proportion of sites from each chronological band and the distance from the two Roman towns, Baetulo (modern Badalona) and Iluro (modern Mataró), was significant in all cases, showing that in all periods the location of the rural sites was linked to the two urban centres. The graph of proportion of sites on distance bands from the towns (figure 6.4) shows that there was a difference between the Iberian settlement pattern and the Roman one, but that after the Romanisation no transformations occurred in the rural settlement pattern with respect to the distance from the Roman towns. The graph in figure 6.5 emphasises this by only showing the sites dating to the Roman period: the curves for all period are very similar. In the Iberian period fewer sites existed close to the location where the Roman towns were to be founded than in the Roman period. Before Romanisation, a small proportion of the rural sites are found close to the location where the two towns were to be built, while a scattered settlement pattern is visible. After Romanisation, a high proportion of the sites (30% to 35%) occurs within 5 km eq (band 1) of Baetulo and Iluro, with a smaller peak further away from the towns, showing that some sort of scattered settlement also existed in the countryside. It is possible that the first peak in the graph represents the Roman sites which were created anew close to the towns when these were founded, while the second, smaller, hump in the graph represents a continuation of the pre-Roman scattered Iberian settlement. Though this pattern shows that sites were highly attracted by the urban centres in Roman times, it is worth pointing out that quite intensive survey work has been carried out in the immediate hinterland of Mataró, which might be biasing the results of the analysis. During the Republic a slightly smaller proportion of the sites were located close to the towns than in the following periods, but the shape of the curves is almost identical from the reign of Augustus until the Late Empire, indicating that no variation in the countryside occurred, despite the acquisition of the municipal status by the two towns and their decline in the 3rd century AD.

The relationship between the rural sites and the position of the Via Augusta was also tested by means of Kolmogorov-Smirnov tests for each of the chronological bands. The test was significant for all the data subsets, including the Iberian one, but it is likely that this result is strongly biased by the shape of the study area, a long and thin corridor which really dictated where the road could go and where the sites could be located, unless a specific high location was sought. The fact that the shape of the graph for the proportion of sites dating to the Iberian period is so similar to those of the sites dating to the Roman periods (figure 6.6) indicates that it is likely that the analysis is picking up an environmental constraint, rather than an archaeological pattern. It is also possible that an Iberian road existed on the path later followed by the Via Augusta, as the coastal corridor offers the only viable track for a north east-south west road.

The pottery assemblage of the rural sites of the Maresme was studied by means of Correspondence Analysis. Very little patterning is visible from the results of the Correspondence Analysis. Black Glaze pottery and Arretine Terra Sigillata, the earliest Roman types found in this area, do not seem to be related in the composition of the assemblages. Spanish Terra Sigillata is close to South Gaulish Terra Sigillata, which it imitated. Terra Sigillata Chiara A occurs in the plot close to Spanish Terra Sigillata, to which it is similar in date, while Terra Sigillata Chiara C and D are reasonably close to each other in the plot, showing a similar pattern to the one seen in the region of Tarragona, but a completely different one from the data set of the Guadalquivir Valley. No pottery type appears to be distributed in any specific way within the Maresme, so that it is difficult to make inferences about differences in the import and redistribution of foreign ceramics.

The first Roman settlement in the Maresme did not involve violent destruction of the Iberian sites, which can possibly be seen surviving in later periods as a few of the sites on high ground and as the sites further away from the towns, preserving the dispersed settlement pattern which had characterised their countryside before Romanisation. The creation of the two Roman towns coincided with the reorganisation of the countryside, creating a settlement pattern centred on the towns. Some sites were created during this first period on the high lands, around 250 to 500 metres above sea level, on ridges facing towards the sea, possibly to allow an early sighting of pirate ships which might have attacked the merchant vessels, or to control the coastal strip, which was a compulsory passage for all the traffic going from the north-east to the south-west along the coast. The two towns had been given the status of *municipia* by the Early Empire, which lead to architectural development within the towns themselves, but, remarkably, did not seem to have any effect to the organisation of the countryside, which continued as before. Even when the two towns declined and lost their administrative, political and economic role, by the late 3rd century AD, the rural settlement pattern does not appear to be influenced by the events of the two towns. Throughout the Roman periods sites disappeared and new ones appeared, but always preserving more or less the same proportions of sites on distance bands from the towns and the Via Augusta and on the elevation bands. Overall it seems that one initial effort was put into the foundation of the two towns and the organisation of the landscape in the first stage of Romanisation, but that afterwards the development of the towns and of the countryside proceeded separately, with the countryside being almost crystaliised in the organisation which had been set during the Republic.

2 The pottery assemblages of the three study areas

The technique of Correspondence Analysis was used to study the pottery assemblages of the rural sites contained in the site databases for the three study areas. Because of the large size of the Guadalquivir Valley when compared to the other two and the fact that a large number of Roman rural sites were known from this area, while only one Roman town was present in the region of Tarragona and two in the Maresme, a different approach was used for the data from the province of Seville. In this area, hypothetically derived town territories were constructed using cost distance surfaces and the number of rural sites falling within each territory and containing a certain pottery type was counted. This produced a raw data table consisting of counts of the locations within each town territory where a certain pottery type was found. Conversely, the data from the other two study areas could not be quantified in any way and Correspondence Analyses of presence/absence data were carried out. Despite the fact that quantified data were analysed in one case and binary data in the other two, it is still possible to compare the results obtained from these analyses and see whether there are any recognisable large scale trends in the composition of Roman pottery assemblages in Spain.

2.1 Black Glaze and Arretine Terra Sigillata

The proportion of Black Glaze pottery, which is the earliest Roman type studied in this project, varies in the three study areas. The proportion of sites containing Black Glaze pottery is much higher in the region of Tarragona and in the Maresme than it is in the Guadalquivir Valley, showing that the south-west and the north-east received different ratios of imported pottery during the Republic. Black Glaze pottery was imported into the Guadalquivir Valley before the arrival of the Romans, and its closeness in the Correspondence Analysis variable plot to Arretine Terra Sigillata shows that the Iberian network along which Black Glaze was being redistributed was still being used for Arretine Terra Sigillata, and possibly Spanish Terra Sigillata, until the 2nd century AD. The closeness between Black Glaze and Arretine Terra Sigillata is not visible in the Correspondence Analysis plots in the region of Tarragona and in the Maresme.

2.2 Terra Sigillata Chiara

The position of Terra Sigillata Chiara A and D in the variable plot of the Correspondence Analysis on the data from the Guadalquivir Valley is remarkable, in that the two points are extremely close together on the first two components. It appears that these two subtypes of Terra Sigillata Chiara occur in very similar proportions in the same areas. The analysis of the town territories most closely influenced by these two subtypes showed that they all occur along the Guadalquivir river, suggesting that these wares reached the Guadalquivir Valley and were redistributed within it along the waterways. The position of Terra Sigillata Chiara C in the plot is striking, as this ware is at the opposite extreme of Component I and appears to have had a completely different distribution from subtypes A and D. Another difference is the very low number of sites in which Terra Sigillata Chiara C occurs, compared to subtypes A and D. It is known that subtypes A and D were both produced in the same workshops in northern Tunisia, while subtype C was produced in central Tunisia. The different place of origin of the subtypes suggests that they were marketed separately and that much less subtype C was imported into Spain than it was the case for subtypes A and D. The fact that subtype A was imported in the Guadalquivir Valley at the same time as Arretine Terra Sigillata, which was being traded along the pre-Roman Iberian exchange network, suggests that at this time two exchange mechanisms were in operation at the same time. The old Iberian system was still in existence and was being used mainly in the southern part of the Guadalquivir valley, while the newly formed Roman exchange network was in use in the north, serving the sites along the Guadalquivir which were involved in the trade of the Baetican olive oil.

In the Correspondence Analyses of the Catalan data sets this division between subtypes A and D and subtype C is not apparent. The proportion of sites with Terra

Sigillata Chiara C in the region of Tarragona and in the Maresme still is quite low compared to the proportion of sites containing the other two subtypes, but the difference is not as large as it is in the Guadalquivir Valley. In the variable plot of the region of Tarragona, subtypes A and D are not too distant, though they are not as close as in the Guadalquivir Valley plot, and Terra Sigillata Chiara C is nearby as well. In the variable plot of the Maresme, Terra Sigillata Chiara D is actually closer to subtype C than it is to subtype A. If the difference in the area of origin of the three subtypes is the correct explanation for their position in the Guadalquivir Valley plot, it seems that this factor did not have an influence on the way the three subtypes were imported in north-east Spain. All the same, it is interesting to notice that Terra Sigillata Chiara C always have very high scores on the first two components in all three data sets (see table of component loadings in appendix B).

2.3 Spanish Terra Sigillata

One interesting feature which emerges from the plots on the first two components in the Correspondence Analyses is the relative position of Spanish Terra Sigillata to Arretine Terra Sigillata and South Gaulish Terra Sigillata in all three data sets. It is an accepted fact that Spanish Terra Sigillata was a ware produced in several different workshops in Spain and that it first appeared as imitation of the imported types Arretine Terra Sigillata and South Gaulish Terra Sigillata, at least until the second century AD. Generally, it is accepted that the Gaulish forms influenced the production of Spanish Terra Sigillata in the northern workshops, while Arretine ware influenced the southern production centres (Fernández Fonseca 1995, 62). These regionally differentiated influences are reflected in the Correspondence Analyses plots. In southern Spain, Spanish Terra Sigillata production was influenced by the Arretine forms and Spanish Terra Sigillata occurs very close to Arretine Terra Sigillata in the variable plot for the Guadalquivir Valley data set, while the South Gaulish Terra Sigillata is further away. In northern Spain, the influence of South Gaulish Terra Sigillata on Spanish Terra Sigillata prevailed over Arretine's and in both the Tarragona and the Maresme Correspondence Analysis variable plots Spanish Terra Sigillata occurs close to South Gaulish while Arretine Terra Sigillata is further away. If Spanish Terra Sigillata tends to be distributed in assemblages similarly to the types it is most similar to, it can be argued that the locally produced ware slowly took the place of the similar imported forms and simply substituted them, taking up their place in the market exchange network of Roman Spain. Once a redistribution system was in existence, the locally produced pottery would have been easier and cheaper to obtain and would have been preferred over the more expensive imported forms.

2.4 Thin walled ware

One interesting recurrent feature in the variable plots of the three data sets is that Thin Walled ware occurs away from the other wares. In the Guadalquivir Valley plot it occurs very close to Terra Sigillata Chiara C, in the Tarragona plot it is separated from the group of wares in the centre of the plot, though other wares occur in isolation in this data set as well, while in the Maresme plot it is isolated from the rest and has the highest loading on component II. Thin Walled ware was produced in some of the same workshops where Spanish Terra Sigillata was produced. the fact that the plots show no association between these two wares indicates that they were serving different purposes and had a different distribution and market. This element is constant in the three data sets studied.

3 The comparison of the three study areas

This section discusses the problems with comparing the information coming from different sources, as well as presenting the similarity and differences between the results of the analysis carried out on the three data sets.

3.1 Problems with comparing data from different sources

The comparison of the three study areas must be made with some caution. The data used in the three studies came from different sources, had to be processed in different ways before being analysed and, finally, the type of analysis that could be carried out on the three data sets was different in each case. The Guadalquivir Valley is much larger than the two study areas in Catalunya, therefore the results from the province of Seville tell us more about larger scale settlement patterns and pottery distribution and less about small scale regional developments. The analysis of the data from the Guadalquivir Valley showed that a different development had occurred in the northern and southern parts of this region, while geographical differentiation within the other two study areas could not be seen because they simply are too small to reveal such patterns from the unquantified non-systematic data. Because of the large size of the Guadalquivir Valley it was also possible to quantify the relative proportion of pottery data in the territories of the Roman towns, while this was not possible in the region of Tarragona (only one Roman town) and in the Maresme (two Roman towns). As a direct consequence of this difference, the analysis of the pottery data from the Guadalquivir Valley has an element of pottery frequency in it, which is lacking from the analysis of the pottery assemblages from the other two study areas.

The chronological division of the sites in three areas is different as well. The sites from the Guadalquivir Valley, the information on which had been copied from the records of the archaeological unit, had been dated, but often the material coming from other sources and integrated in the site database had not. Moreover, the great variability in the type classification of the sites in the catalogue of the archaeological unit lead to the suspicion that a similar variability existed in the chronological classification of the sites. The site data from the Guadalquivir Valley were therefore standardised in terms of chronology and function on the basis of the diagnostic material which was found in the sites. The standardised chronological division of the sites in the Guadalquivir Valley only has three time bands, the Republic, the Early Empire and the Late Empire. The site data from the region of Tarragona and from the Maresme, on the other hand, came all from the same source (with the exception of some site data from Prevosti's publication) and did not need to be standardised. As a result, the site data from these two areas were classified into Republic, the reign of Augustus, the Early Empire, the 3rd century, and the Late Empire, so that a more accurate study of the chronological development of the settlement pattern in these two areas could be carried out. Similarly, the sites from the Guadalquivir Valley had been divided into low status and high status ones, while all the rural sites in the region of Tarragona and the Maresme had been classified as `villae' (see discussion on the use of the term `villa' in chapter 4, *The Guadalquivir Valley*, section

2). The sites in the Catalan study areas were not reclassified into low and high status sites, so that only the general development of the rural settlement pattern was studied. The sites from these two study areas were not reclassified into high and low status ones because the small size of the two study areas would have meant that the data subsets would have become too small to carry out any sort of statistical study on them.

Other differences between the data sets are not so immediately apparent, but they exist all the same, so that it is necessary to be aware of them before the results of the analysis on the data from the three study areas are compared. It has already been said that the data from the Guadalquivir Valley came from different sources and that different surveyors had imposed their own personal interpretation on the finds. The same is true for the data from the three data sets taken together. When the results from the three study areas are compared, in practice the data sets are equated and considered equivalent, almost as if they were put together into one large site database. The three data sets are not necessarily equivalent and, since no information is available on the methods used to collect the data in the different areas, it is safer to assume that different methods were employed, introducing different biases in the recovery, classification, analysis and interpretation of the data. Even within the same study area it is possible to see the effect of different survey activity. The site data from the Maresme shows a concentration of sites close to the town of Iluro (modern Mataró), which is much more likely to be due to extensive survey carried out in this area rather than being a `true' archaeological pattern. This is also confirmed by the empty gap in between the concentration of sites around Mataró and the scattered site data to the south-west.

This discussion on the intrinsic differences between data from different sources is only aimed at highlighting the fact that it is easy to fall into the trap of considering equivalent data sets which have been collected in different ways and comparing them directly extracting general conclusions from them. If these limitations are kept in mind, the comparison of studies from different areas can help build up a large-scale picture from regional studies.

3.2 Similarity and differences in the development of the three study areas

The most noticeable difference between the settlement pattern from the Guadalquivir Valley and the two Catalan study areas is the much higher proportion of rural sites in existence during the Republic in the region of Tarragona and in the Maresme. This difference is even greater when the fact is considered that the three Roman chronological bands in which the data from the Guadalquivir Valley were divided are wider than the five Roman chronological bands in which the data from the two Catalan study areas were divided. Of course, the apparent lack of early Roman sites in the Guadalquivir Valley might be connected to a lower visibility due to a more limited import of diagnostic ceramic material, but then the question arises of why less pottery was imported into the Guadalquivir Valley during the Republic than in the other two areas. There does seems to be little difference between the proportion of rural sites in existence in the region of Tarragona and the Maresme during the Republic and the reign of Augustus. This evidence suggests that during the Republic the north-east of Spain was already an organised part of the Roman world, while society in the Guadalquivir Valley was still very much based upon the pre-Roman Iberian organisation. The survival and continuation of the Iberian organisation in the Guadalquivir Valley is also reflected in the survival of the Iberian exchange network well into the Early Empire, while in the region of Tarragona the imposition of the Roman organisation of the countryside is well visible immediately after the *colonia* was created at Tarraco. In the Maresme, the Roman settlement pattern is imposed immediately as soon as the area is Romanised and the two towns of Baetulo and Iluro are founded. The fact that no transformations occur in this area after the first stage of Romanisation suggests that the pre-existing Iberian organisation was totally replaced, otherwise subsequent reorganisations of the rural settlement pattern would have been expected. In the Guadalquivir Valley the high status sites were concentrated around the towns and the low status sites had a more dispersed pattern. In the region of Tarragona the rural sites (not divided into high and low status ones) are found evenly distributed away from the town, showing that the first Roman settlement had a different character in the two areas. The rural settlement of the Maresme presents a concentration of sites very close to the towns, followed by a decline in proportion of sites, followed by an increase, presenting a distribution pattern which reflects in part that of the Guadalquivir Valley (sites close to the towns) and in part that of the region of Tarragona (sites dispersed in the countryside). It is worth pointing out that the much higher proportion of sites in band 1 of the Maresme site data compared to the Guadalquivir Valley data is influenced by the difference in the width of the cost distance bands (5 km eq in the case of the Maresme and 1.5 km eq in the case of the Guadalquivir Valley). It was necessary to use different band widths in different data sets to have enough cases in each band to make the tests possible at all.

The Early Empire witnessed a dramatic settlement development in the Guadalquivir Valley, coupled with the increased import of foreign pottery and the growth of the towns located along the navigable rivers and the main roads, whose fortune was linked to the expansion in external trade. The rural settlement patterns of high and low status sites developed into similar distributions, showing that a thorough organisation of the countryside had been achieved. At the same time the countryside in the region of Tarragona was reorganised and the resulting settlement pattern resembles roughly that visible in the Maresme, though a smaller proportion of sites occurred very close to the town. Since the width of the cost distance bands in the region of Tarragona and in the Maresme are the same (5 km eq) it is possible that this difference is due to the fact that intensive surveys were carried out in the immediate hinterland of Iluro, biasing the distribution. Both in the region of Tarragona and in the Guadalquivir Valley this period witnesses the increase in the importance of the town as political and economic centre, while in the preceding centuries their role was more likely to have been one of reference point of `Romanity' in a otherwise still largely (in the Guadalquivir Valley) or partially (in the region of Tarragona) Iberian countryside. At this time a number of towns were given the status of *colonia* or *municipium* in the Guadalquivir Valley, while Tarraco acquired the status of *colonia* and Iluro and Baetulo that of *municipium*. The process of creating Roman towns or Romanising those already existing on a large scale had been started by Caesar and then carried on by Augustus and the other emperors. The large number of towns which were give the status of either *colonia* or *municipium* in the Guadalquivir Valley at this time shows that this area had become a full part of the Roman world by this stage, as it is also reflected by its increasing importance in the international trade of Baetican olive oil. The elevation of the two Roman towns in the Maresme to the status of *municipium* did not influence the way the countryside was organised, in the same way as their decline in the Late Empire is not reflected in a variation in the rural organisation. Despite the fact that the rural sites did appear to be concentrated around the towns when the area was Romanised, the fortunes of the two towns does not seem to have any influence on the rural settlement pattern.

In the Late Empire a similar process can be seen to take place in the Guadalquivir Valley and in the region of Tarragona: large estates appear, while the role of the towns diminished. There are important differences in this process in the two areas.- In the Guadalquivir Valley a division between the area along the rivers Guadalquivir and Genil and the area to the south west is visible. The towns along the navigable rivers continued to have some sort of commercial role and the rural sites in their hinterland survived the crisis of the 3rd century, in fact new rural sites appeared at this time along the navigable rivers. The area to the south, away from the navigable rivers, saw the decline of the towns and the appearance of estates which did not need the services offered by the urban centres to survive. The settlement pattern was much denser along the Guadalquivir and the Genil than it was in the other areas. The different fortunes of the towns and the rural sites in their hinterlands appears to have been linked to the easy access to the navigable rivers, the main route to the open sea and therefore to the export of the locally produced goods. In contrast to what happens in the Guadalquivir Valley, in the region of Tarragona the position close to the town proved to be a disadvantage for the rural sites, leading to their decline while the sites away from the town had better chances of survival.

4 `That really made my life easier': advantages of using GIS

This section examines in what ways GIS aided the study of the archaeological data presented in this thesis. For this purpose, the database management system and the statistical software are considered to be an integral part of the GIS, though these were separate packages and had to be linked to Idrisi via custom programs. Rather than presenting a generic discussion of the advantages of using GIS in landscape archaeology, specific examples from the analysis of the data from the three Spanish study areas are discussed.

There are clear and immediate advantages in applying GIS techniques to landscape archaeological studies. The main advantage is that very large database of sites can be used and that all the information that can be derived from maps and/or satellite images, traditionally used in landscape archaeology, can be directly inputted into a GIS and used in conjunction with the archaeological data. The integration of site data and background information makes the visualisation of the relative importance of locations with specific characteristics (*ie* proximity to water sources, on low land and so on) to settlement immediately possible, while without a GIS this process would involve the production of several distribution maps, which might require several days or even weeks to produce in the case of large databases. The site database tables for the region of Tarragona and the Maresme were rather small (219 records and 294 records respectively), but the site database table for the Guadalquivir Valley contained over 2,300 records and the extraction of information from it without a database management system would have been almost impossible to do manually.

The statistical analysis of the data was made easier by already having the information stored in digital format, so that the process of carrying out the analysis only involved extracting the data and transferring them between packages, rather than having to enter selected information into a statistical package every time a new analysis was required. Without using a GIS, information such as the elevation at which each site occurs, or the total number of sites occurring on specific soil types, or at a given distance from water sources or from administrative centres, would have had to be extracted by hand and then inputted into digital format for analysis. This would not have been very time and effort effective, because it would have involved creating several distribution maps of different site groups (such as sites dating to the same period, or of the same type) on the same background maps before the relevant information could be extracted, while with a GIS this process can be almost completely automated. Despite the fact that even in a GIS separate maps still have to be produced before the information can be extracted, selecting sites corresponding to specific characteristics can be done simply by running standard database queries. The process of extracting information such as counts of sites occurring in certain places, or the value of a background variable where sites occur, can then be performed automatically. The ease of extracting data according to certain specifications and being able to carry out any possible type of analysis on them also means that more tools are available to the archaeologist to study the data and identify trends. This point is better illustrated by looking at the sort of analyses which have been carried out on spatial archaeological data in Spain in the past. Most of the studies involved just plotting the data on maps (often very coarse ones) and trying to `see' trends without actually testing whether they really existed. One example is Prevosti Monclús' statement (1981b, 264) that the majority of the Roman villae in the Maresme occurred around 100 metres above sea level, with a smaller concentration of sites along the coast. This statement was proved to be untrue when the elevation at which sites occur in the Maresme was studied by extracting the information of where sites occur from the DEM (Digital Elevation Model) and then statistically testing these data (see chapter 6, section 3.1). Another example is from Carrillo Díaz-Pines and Hidalgo Prieto (1991), who studied the distribution of sites in an area neighbouring to the province of Seville and concluded that the Roman sites were preferably located on areas with a certain land capability rather than others. Carrillo Díaz-Pines and Hidalgo Prieto do admit that the land capability on which the majority of the sites were located and on which they based their conclusions covered a very large proportion of their total study area *(ibid.* 63), but the point is that he did not carry out even a statistical test of significance on his data. The study carried out on the Guadalquivir Valley data set on the relationship between the position of known rural sites and the soil type proved to be too unreliable because of the limitations of the χ^2 test and the large size of the data set, but with map reclassification and map algebra it was possible to create an agricultural potential map which allowed consideration of more elements together with the soil type and to use the Kolmogorov-Smirnov one-sample test, generally considered to be more sensitive than the χ^2 test (see chapter 4, *The Guadalquivir Valley*, sections 4.1 and 4.2).

Another advantage of GIS is the possibility to use the existing data to generate new information. A set of digitised contours can be used to interpolate a DEM to get a more direct feel of the location of sites in the landscape than it is possible to get from a flat 2D map. Similarly, a GIS can create cost distance surfaces to study the relationship between archaeological sites and other features, both natural such as water sources, and artificial such as communication routes and administrative centres. This derived information can greatly improve the precision of archaeological locational analysis, as sites catchments can be created allowing for the influence of land form rather than using linear `as the crow flies' distances. Similarly the study of elevation at which sites occur can be carried out on interpolated DEMs rather than just by classifying sites in elevation bands whose width is determined by the resolution of the maps used. The possibility to create cost distance surfaces was used as starting point to generate a series of hypothetical territories around the Roman towns in the Guadalquivir Valley. Since practically no information exists as to the extent of each town territory on Roman times, the model derived from calculating cost distances corresponding to the cost of moving 15 km over flat ground (a measure derived from ancient sources) can be seen as being more accurate than the alternative of using linear catchments or Thiessen polygons (see chapter 4, section 5.1). The creation of the hypothetical town territories also allowed the study of the number of locations within each town territory where a certain pottery type was found, giving some indication of the density of occurrence of each type in different parts of the Guadalquivir Valley (see chapter 4, section 5.2). Without a GIS, the calculation of the hypothetical town territories would have been impossible, though, in theory, it would have been possible to input the data in a statistical package and carry out the Correspondence Analysis on the pottery table, provided that the extent of the Roman town territories had been calculated in some way beforehand and that the number of sites with specific pottery types in each territory had been counted.

To summarise, the three main contributions of GIS to the study of the data sets from Spain have been:

1. the capability to manage a large database and extract data easily according to specific criteria
2. the capability to extract specific information from the background maps and transfer these data directly to statistical packages to perform analysis
3. the capability to generate new data from the existing information to explore further certain points of interest

It can be argued that point 1 is true of database management systems, rather than GIS. As it is explained in the *Introduction* chapter (section 3.3) a database management system is an integral part of a GIS, though Idrisi (at least the DOS version which was used for this project) itself does not provide one, or a direct link to an external one. The way the data needed to be extracted and organised for the type of work carried out for this research project required the extensive use of querying facilities in a database management system, which therefore became a very important part of the whole GIS analysis system. Similarly, point 2 refers to the use of a GIS as a means to store spatial data (background data and site data organised by the database management system) and extract them in a variety of fashions suitable for analysis by an external statistical package. In this form, Idrisi can be seen simply as a spatial database which stores spatially-referenced information and allows their organisation and retrieval according to what can be defined as `spatial querying'. Point 3 is the only one which really refers to something that can only be done with GIS software alone.

5 Look back in anger: disadvantages of using GIS

This section discusses the problems encountered in using GIS with archaeological data to carry out the study described in the preceding chapters. Some of these considerations refer to GIS in general, rather than being directly related to the type of use GIS has been put to for the purposes of the project. This section is also intended as a warning on the sort of problems involved in this type of work to anybody who might want to carry out a similar study.

Just as there are advantages in using GIS in landscape archaeology, similarly there also are disadvantages. The main disadvantage of using GIS to carry out a landscape archaeological study lies in the amount of time and computer expertise required to prepare and format the data for use in the system. Since maps in digital format were not available at the beginning of the project, a large proportion of the first two years were spent digitising paper maps in AutoCAD and exporting them to the Idrisi format. To achieve this goal specific programs had to be produced in AutoLISP, which involved learning the AutoLISP programming language. After the basic maps had been produced in Idrisi format, these had to be further processed to create digital elevation models of the three study areas and to smooth out errors, such as the gaps between polygons (see description of the problem in chapter 2, *The Data*, section 2.2).

Another sort of problem was presented by the sites database tables. The data had been collected by different people, often using different techniques and storing them in a variety of different formats. The nature of the information recorded also varied greatly among surveyors and study areas. Some effort had to be put in the detailed study of the raw survey data before a suitable database structure could be designed. Not only the problem existed of integrating data coming from different sources and recorded in a variety of different formats, but there also was the problem of deciding which pieces of information were relevant for the study and which were not and could be discarded. The data then had to be typed in and crosschecked to reduce the magnitude of the unavoidable copying error. Another major problem with integrating site data from different sources was that different coordinate systems had been used by different surveyors to record the position of the sites, so that a coordinate transformation routine had to be produced and applied (see chapter 3, *The system*, section 3.2.2.1). The same routine was also used to transform the coordinates of the site data of the Guadalquivir Valley falling within the UTM QB zone into a prolongation of the neighbouring UTM TG zone.

Once both the background data and the site data had been inputted and formatted to be used within Idrisi, it was necessary to develop new tools to extract the information in a format suitable for the sort of statistical analysis which was to be carried out. The particular problem encountered in doing so was more a deficiency of Idrisi itself, rather than being specific to GIS (see below, section 3.2).

Overall, it would have been very difficult to carry out the project without previous extensive knowledge of computing, programming, graphics data manipulation, GIS techniques and statistical analysis, and one extra year would probably have been required while the relevant skills were being learnt. It is important to realise that GIS software on its own is rather useless and it is necessary to integrate it with a variety of other pieces of software, mainly CAD packages, database management systems and statistical software. It is also necessary to be able to program and produce specific routines to extract data in certain ways which might not be provided by the GIS package itself, and to transfer the data between applications. Knowledge of the GIS package Idrisi alone and of GIS techniques would not have been enough to carry out the study of the three Spanish data sets. Though it is to be hoped that `complete' GIS packages[1] (including extensive database management and advanced statistical capabilities) will be produced and become available to be used in archaeology in the future, at present this is not the case and anybody willing to undertake a study similar to the one described in this thesis would need to have more than just a basic understanding of computing.

Another type of problem with using GIS with landscape archaeological data is that the ease of visualising any chosen type of site on any chosen type of background (discussed in section 1), can easily lend itself to abuse. If a certain type of pattern is looked for, it is easy to make it `happen' by manipulating the data, even unconsciously. For this reason it is very important to test statistically the reality of any relationship that becomes visible in distribution maps. This problem is not specific to the research project described in this thesis, but it was nonetheless necessary to be aware of it to avoid falling into the trap of `seeing' patterns where there were none.

> We must realise that the human eye is easily fooled — it might perceive patterns in an image which really do not exist or fail to recognise trends that do.
>
> (Kvamme 1994, 1)

When using GIS with archaeological data there also is the risk of putting too much trust on the data derived from the basic maps. In a DEM the only elevation which is likely to reflect faithfully enough the actual elevation of the land is where the contours used to interpolate the DEM are. All the other cells have been filled by means of mathematical formulae and they are just possible approximations of the real elevation at that point. Different formulae and algorithms will generate slightly different DEMs starting from the same set of digital contours. It is also true that elevation contours in

[1] VB.GIS (Reynoso 1994) seems to have all the characteristics of the `complete' GIS package, but at present it has made very little impact on archaeological GIS.

traditional paper maps are themselves interpolations of a few recorded points, so that there already is an error in the basic information used to interpolate smooth 3D surfaces reproducing the landscape. If these limitations are not kept in mind, there is a risk of putting too much emphasis on information which is not necessarily as precise as it looks. This problem can be exacerbated if the resolution of the images is too low, increasing the amount of land represented by each single cell which can only ever have just one value.

6 Idrisi: a cheap and cheerful tool for landscape archaeology?

This section discusses the strengths and weaknesses of the PC-based Idrisi GIS package, which was used for the research project presented in this thesis. Despite the fact that recently a Windows version of Idrisi has become available, the whole of the study was carried out using the DOS versions 4.0 and then 4.1 of the package, so only these will be discussed. Reference to the Windows version is made in passing at points were particular features of the DOS version discussed were improved.

6.1 What Idrisi could do

Idrisi proved to be adequate to perform directly most of the tasks which were required of it. Since Idrisi is a small package by GIS software standards, some operations which can be performed directly in other packages such as Arc/Info or Spans had to be split up in separate steps which were carried out individually. One such example is the use of map reclassification to weight the variables before they were combined by means of map algebra in the production of the agricultural potential map of the Guadalquivir Valley (see chapter 4, *The Guadalquivir Valley*, section 4.2). Other GIS packages allow a weighting of the variables to be specified before the overlays are combined together, thus bypassing the intermediate stages and the creation of temporary maps.

The display of 2D images is good, though the plotting facilities over displayed raster images is not very good, only allowing the last superimposed vector file to survive operations such as zooming and restoring full view, which means that if a map detail was zoomed upon, then all the vector files but one had to be superimposed on the image again. In the Windows version this feature has been improved and the display of vector files over a base raster map and its manipulation has been made more flexible. When an image is viewed with the 3D viewer (***ortho***), there are no facilities to manipulate it, so that the whole image has to be regenerated if any of the display parameters are to be changed.

The interpolation of Digital Elevation Models from digitised contours is very time consuming but gives satisfactory results, despite the presence of a number of star-shaped interpolation artifacts which depend on the basic search routine used by Idrisi to find the nearest contour. Idrisi also offers a number of filters, including the option to use a user-defined filter, to smooth the final DEM. The DEM can be used as the starting image to produce maps of other characteristics such as the slope and the aspect, as well as friction surfaces based on land form by means of operations such as logarithm, square rooting or exponentiation applied to each data cell.

The calculation of cost distance surfaces is very time consuming but gives satisfactory results, despite the fact that it is not easy to relate the resulting cost units to `real' distances, so that the method described in chapter 4, section 5.1, had to be employed.

6.2 What Idrisi could not do

Idrisi is a completely raster-based GIS, so that a number of functions, which work better on vector systems, are not implemented. Luckily none of the functions not implemented in Idrisi were required in the study of the three data sets and it was possible to achieve almost all results that were required by going through a series of steps (see above, section 3.1).

Idrisi lacks completely a direct link to a database management system to handle point (site) data. This means that a completely separate software must be used for the storage, management and retrieval of the site data. In the course of the project described in the preceding chapters, dBase III+ was used, which meant that a specific routine had to be produced to export the dBase data to Idrisi vector format. Any query on the site data had to be carried out *outside* Idrisi and the information resulting from the query had to be extracted from the site database and transferred to Idrisi to be used with the rest of the spatial database. There is a module in Idrisi (***dBidris***) to read a dBase III+/IV data file and use the information contained in it to reclassify the values of the cells in an integer image file by means of an Idrisi value file, but this is not very effective and requires the dBase data file to have a specific format which is not necessarily the most efficient one to store the data used in this project. Moreover, all the database querying and data sorting still have to be performed using an external program. The Windows version of Idrisi largely improves on the links between the external database management system and Idrisi itself, even providing its own front end to the foreign format tables, but it still does not support its own complete database management system. Database tables still have to be created with a separate program and despite the fact that some attempt has been made to create direct links between the data being displayed on a map and the database table the data come from, it still is not possible to retrieve the database information about a data point selected on a map.

Another major problem with Idrisi, in terms of what was required for the analysis of the site data, is the fact that it is extremely difficult to extract the information about a background variable at locations where sites occur. The only way to do so in Idrisi is to use the ***extract*** module but, though this is cumbersome but feasible in small databases, as each data point must be given a unique identifier, with large databases such as the one for the Guadalquivir Valley this becomes almost impossible. A custom program (***Idr_Vals***, see chapter 3, *The system*, section 2.2.2) had to be written to read two images, containing the site information and the background variable, writing a text file with the list of extracted values. The process of extracting information about the environment (be that the elevation, the soil type or the value of a cost distance surface from a given source) at locations where sites occur is likely to be one of the most needed by landscape archaeologists using GIS, and the difficulty on performing such an operation is a serious blow to the usability of Idrisi in this type of study.

6.3 Idrisi: overall performance

In short it was possible to get Idrisi to perform all the GIS functions which were required, though in some cases it was necessary to go through a series of steps to obtain the final result. The graphics display was adequate, despite the fact that it was almost impossible to get a hardcopy output from Idrisi and a screen grabber had to be used. The main strength of Idrisi lies in its small size, its modest hardware requirements compared to most other GIS packages and the amount of modules which have been included into the small package, enabling it to carry out virtually all of the standard raster GIS functions. The main weakness of Idrisi lies in its being a monolithic piece of GIS software with very little capacity to communicate with other applications (such as AutoCAD, dBase III+, the statistical packages) without having to create custom import/export programs. This closed system, however, tends to be the case with most software designed to run under the DOS operating system and the new Windows version of Idrisi goes a long way to implement a link between the GIS and external database management systems, though the extraction of data still needs to be improved upon. The only major fault was the difficulty of extracting the information about the landscape at locations where archaeological sites occurred, as this probably was the operation which had to be carried out most often during the course of the project.

7 Conclusions. Would I use GIS again?

Using GIS in the landscape archaeological study described in the preceding chapters allowed improvement to be made on the sort of analysis which could be carried out. The majority of the tests carried out would have been possible to perform in principle without using a GIS, though they would have required a large amount of manual work. Earlier works (Gorges 1979; Fernández Castro 1982; Lewit 1991) carried out on similar data failed to include statistical analysis and even the study of particular subsets of data (such as sites of a certain type dating to a certain period) was very limited. Other operations which were performed in the study would be absolutely impossible without the aid of GIS software, such as the creation of the hypothetical territories of the Roman towns of the Guadalquivir Valley. The increase in time and effort which had to be put into inputting and preparing the data was counterbalanced by the increased ability to perform detailed analysis of the data and carry out operations which would not have been possible otherwise. Despite the fact that there are problems in using GIS with landscape archaeological data, especially for someone without a strong computing background, the advantages outnumber the disadvantages by far.

7.1 Would I use Idrisi again?

As for the specific GIS software used in this study, Idrisi for DOS, the verdict is slightly different. Idrisi is a very effective package given its size, but it also is a old one and the problems with its age are becoming apparent. More and more of the latest PC application software is being produced to run under Windows, allowing better data exchange between applications and better memory management, which gives more computational speed. Clark University, which produces Idrisi, has become aware of this and has produced a Windows version of Idrisi, which still needs considerable debugging before it becomes reliable. As most of the complementary software landscape archaeologists are likely to use (CAD packages, database management systems, statistical software) nowadays runs under Windows, it is likely that the DOS version of Idrisi, which also is much more limited compared to the Windows version, will soon become obsolete. At present, given the option , a different GIS package running under an integrated environment, such as OS/2 or Windows, would probably be a better choice than Idrisi, at least until the Windows version is completely debugged and better linking to external programs are added to it.

8 Final remarks

This thesis has shown that GIS software, coupled with a database management system and statistical software, can be very effective in studying survey archaeological data. The system allows the application of standardisation procedures to large databases, which are required if information coming from different sources or collected with different methods, as it is very often the case with survey data, is to be used as one data set. Large integrated data sets can be studied and the relationship of sites with spatial information concerning the landscape and the socio-economic factors present in an area can be studied in detail. Despite the fact that the preparation of the data for use in a GIS is very labour-intensive, the analytical and data integration possibilities of the system outnumber the problems which have to be faced in the process of putting the information in a suitable format. A large scale study as the one presented in this thesis would not have been possible to carry out without a GIS. Not only significantly more time and effort would have had to be put in the analysis of the data, but some of the processes could not have been carried out at all, such as the creation of the hypothetical Roman town territories in the Guadalquivir Valley and the use of cost distances to study the relationship between rural sites and the background information.

There are drawbacks with GIS, as it makes it very easy to carry out several types of test, some perhaps not suitable for the data under study, and there always is the risk of spending extra time preparing data which will not then be used in the analysis, but if the aim of the research is clear, the use of GIS can be a real advantage to landscape archaeological studies. The most important step in an analysis of the kind presented in here is to have a clear understanding of the aims of the project and, if this is present, then GIS can be a really effective tool in the study of spatial archaeological data.

The study presented in this thesis differs from earlier comprehensive works of the same sort in a number of ways. As it was shown in chapter 1, *Introduction,* section 2, other works which aimed at studying large amount of data from several areas fell into the trap of ending up as being just large descriptive catalogues with very little analysis in them. Gorges (1979) did try to study the villa data from the whole of Spain by plotting pie-charts of various types of information over a map of Spain, but his study was seriously flawed by the different intensity of the survey work which had been carried out in different parts of Spain. This bias is very evident in his work and is reflected in the patterns of settlement which he detects in his distribution maps. The problem of the different intensity of surveys in different parts of Spain was tackled for the projects presented in this thesis by studying the chronological changes in settlement pattern in each area in turn, rather than trying to look at the data from the three areas together.

Other comprehensive works such as Fernández Castro (1982) and Lewit (1991) only look at a small number of very well studied sites from Spain (Fernández Castro) and various European countries (Lewit). Lewit does try to assess differences and variation in the development of the Roman settlement pattern in various parts of Europe, though just by looking at the development of 201 excavated sites and generalising the fortune of each site to the area where it occurs. Fernández Castro, on the other hand, is more concerned with the study of the development of each site in isolation, rather than to the wider pattern of settlement. Both these authors study large rich 'villae' which had been well excavated (though Lewit does look at some survey data from Italy), but by doing so they ignore the information about the smaller sites, which were much more numerous than the rarer large rich villae and more likely to reflect the fluctuations of the settlement pattern in time. Fernández Castro's aim is that of studying the individual villae and does not try to make inferences about variation in settlement pattern and development, while Lewit does draw conclusions from her research. The main problem with Lewit's work is that she does not use any statistical method to study her data, basing her conclusions simply on the observation of unmeasured variation in the samples from different areas.

There are major differences between these earlier works and the project presented in this thesis. One such difference is the use of all the archaeological data available for a specific area, rather than just a sample of sites selected because better studied (a technique which also tends to introduce a high-status-only bias in the sample). Though survey data are not so reliable and precise as excavation data can be, nonetheless they provide a much better coverage of an area and are more likely to reflect generalised trends than the study of a small sample of individual sites. Another major difference between earlier works and this project is the use of statistical techniques to study the data. Tests of significance were used to verify the existence of relationships between the location of sites at different times and a number of variables, which could be simply environmental (such as the agricultural potential) or socio-economic (such as the accessibility to urban centres and communication routes). Correspondence Analysis was used to study the distribution of pottery types within the assemblages of individual sites (in the region of Tarragona and in the Maresme) and in the territories of the Roman towns (the Guadalquivir Valley). A relationship between the location of sites and the soil type was noticed in the area of Cordoba (neighbouring the province of Seville in the Guadalquivir Valley, but actually outside the study area) by Carrillo Díaz-Pines

(1991), but he did not test this relationship by any means, though he does admit that the soil type on which most sites occur covers a very large part of his study area.

In general, studies carried out on the Roman settlement pattern in Spain tend to use only small samples of the available settlement data, giving priority to well studied large rich sites. They also tend to be descriptive and not to use statistical techniques to analyse the data, and to generalise the observations of small proportion of sites to whole areas. When larger sets of archaeological data are studied, the problem of the bias introduced by the different intensity of survey work in different areas is not addressed. With the project presented in this thesis and attempt was made to study large sets of settlement data from three areas of Spain using GIS to store, organise, manipulate and analyse the data, and statistical techniques to test the data and detect patterns in them. The aim was to provide an overview of the similarities and differences in these areas, rather than to obtain a general view of Romanisation and the development of Roman settlement in the whole of Spain. The author does not imply that the picture presented in this thesis is definitive, and it is very likely that future archaeological work in the areas studied will uncover data which will modify some of the conclusions presented here.

Bibliography

ALBERTINI E. 1923, *Les divisions administratives de l'Espagne Romaine.* (Paris)

ALFÖLDY G. 1978, `Tarraco' in *Pauly-Wissowa Real Enzyclopädie der Klassischen Altertumswissenschaften* Sup. XIV 570-644. J. B. Metzler (Stuttgart)

ALLEN K. M., GREEN S. W. & ZUBROW E. B. W. (eds.) 1990, *Interpreting Space: GIS and Archaeology.* Taylor & Francis (London)

ALMAGRO BASCH M., SERRA RÀFOLS J. DE C. & COLOMINAS ROCA J. 1945, *Carta arqueológica de España I: Barcelona.* (Barcelona)

AMMERMAN A. J. 1971, *A computer analysis of epipalaeolithic assemblages in Italy.* In Hodson, Kendall & Tautu (eds.) 1971, 133-137

AMORES CARREDANO F. 1982, *Carta arqueológica de los Alcores (Sevilla).* Excma. Diputación Provincial de Sevilla (Seville)

ANDERSON A. C. & ANDERSON A. S. 1981, *Roman pottery research in Britain and north-west Europe.* British Archaeological Report International Series 123 (2 vol). Tempus Reparatum (Oxford)

ANDRESEN J., MADSEN T. & SCOLLAR I. (eds.) 1993, *Computing the past - CAA92.* Aarhus University Press (Aarhus)

BAILEY R., HOWES D., HACKENBERGER S. & WHERRY D. 1985, `Geographic information processing in land use modelling and testing in the Columbia River basin'. Paper presented at the 50th Annual Meeting of the Society for American Archaeology (Denver)

BANNING E. B. 1993, `Hypermedia and archaeological publication: the Wadi Ziqlab project'. In Andresen, Madsen & Scollar (eds.) 1993, 441-447

BARKER G. W. W. & LLOYD J. (eds.) 1991, *Roman landscapes. Archaeological survey in the Mediterranean region.* British School at Rome (London)

BAXTER M. J. 1994, *Exploratory multivariate analysis in archaeology.* Edinburgh University Press (Edinburgh)

BLAGG T. & MILLETT M. 1990, *The early Roman Empire in the west.* Oxbow (Oxford)

BONSOR, E. G. 1931, *The archaeological expedition along the Guadalquivir (1889-1901).* (New York)

BROWN P. E. & RUBIN B. H. 1982, `Patterns of desert resources use: an integrated approach to settlement analysis'. In Brown & Stone (eds.) 1982, 267-305

BROWN P. E. & STONE C. L. (eds.) 1982, *Granite reef: a study in desert archaeology.* Anthropological Research Papers 28. Arizona State University (Arizona)

BURRILLO F. 1984, `Un estudio sincrónico y diacrónico del poblamiento y del territorio: el proyecto interdisciplinar de Mora de Rubielos (Teruel)'. *Arqueología Espacial* 1, 187-205

BURROUGH P. A. 1986, *Principles of Geographical Information Systems for land resources assessment.* Clarendon Press (Oxford)

CARRERAS C. & MASSAGRANDE F. A. forth, `Peñaflor — intrasite analyses'. In Keay forth, *The excavation at Peñaflor.*

CARRILLO DÍAZ-PINES J. R. & HIDALGO PRIETO R. 1990, `Aproximación al estudio del pobliamento romano en la comarca de Palma del Rio (Córdoba): la implantacion territorial'. *Ariadna* 8 (June 1990), 37-68

CHENHALL R. G. 1967, `The description of archaeological data in computer language'. *American Antiquity* 32, 161-7

CHENHALL R. G. 1968, `The impact of computers on archaeological theory: an appraisal and projection', *Computers and the Humanities* 3 (1), 15-24

CHIC GARCÍA G. 1978, `Consideraciones sobre la navegabilidad del Guadalquivir en época romana'. *Gades* 1, 7-ss

COBERTERA E. 1986, *Los suelos cultivados de la provincia de Tarragona.* L'Escma. Diputació de Tarragona (Tarragona)

COHEN L. & HOLLIDAY M. 1982, *Statistics for social scientists.* Paul Chapman Publishing Ltd. (London)

COLLANTES DE TERÁN F. 1939, *Catálogo arqueológico y artistico de la provincia de Sevilla*, Tomo I (Seville)

COLLANTES DE TERÁN F. 1943, *Catálogo arqueológico y artistico de la provincia de Sevilla*, Tomo II (Seville)

COLLANTES DE TERÁN F. 1951, *Catálogo arqueológico y artistico de la provincia de Sevilla*, Tomo III (Seville)

COLLANTES DE TERÁN F. 1955, *Catálogo arqueológico y artistico de la provincia de Sevilla*, Tomo IV (Seville)

COLLINGWOOD R. G. 1930, *The archaeology of Roman Britain.* Menthuen and Co. Ltd. (London)

CREAMER W. 1985, `The upper Klethia Valley: computer generated maps of site location'. Paper presented at the 50th Annual Meeting of the Society for American Archaeology

CRUICKSHANK J. G. 1972, *Soil geography.* David & Charles Ltd. (Newton Abbot, Devon)

CSÁKI GY., JEREM E. & REDÖ F. 1995, `Data recording and GIS applications in landscape and intra-site analysis: case studies in progress at the Archaeological Institute of the Hungarian Academy of Sciences'. In Lock & Stančič (eds.) 1995, 85-99

CURCHIN L. A. 1990, *The local magistrates of Roman Spain* (Toronto)

CUYAS J. M. 1977, *Historia de Badalona: Badalona romana y visigótica,* vol. III. (Badalona)

DE LA BÉDOYÈRE G. 1988, *Samian ware.* Shire Publications Ltd. (Aylesbury)

DE LA ROSA D. & MOREIRA R. 1987, *Evaluación ecológica de recursos naturales de Andalucía.* Agencia de Medio Ambiente — Junta de Andalucía (Seville)

DOMERGUE C. 1990, *Les mines de la peninsule Ibérique dans l'antiquité romaine.* Collection de l'École Française de Rome (Rome)

DORAN J. E. 1971, `Computer analysis of data from La Tène cemetery at Müsingen-Rain'. In Hodson, Kendall & Tautu (eds.) 1971, 422-431

EASTMAN J. R.,1992, *Idrisi version 4.0 technical reference* (Worcester, Massachusetts)

ESCACENA CARRASCO J. L. & PADILLA MONGE A. 1992, *El pobalmiento romano en las margenes del antiguo estuario del Guadalquivir.* Editorial Graficas Sol (Écija, Seville)

ESTRADA J. 1955, *Síntesis arqueológica de Granollers y sus alrededores* (Granollers)

ESTRADA J. 1969, *Vías y pobliamento en el territorio del area metropolitana del urbanismo de Barcelona.* Comisión Técnica revision PC-53 y para uso interno (Barcelona)

FENTRESS E. & PERKINS P. 1987, `Counting African Red Slip ware'. In *L'Africa Romana* 5. Atti del V Convegno di Studi, 205-214 (Sassari)

FERGUSON T. A. 1985, `Use of Geographic Information Systems to recognise patterns of prehistoric cultural adaption'. Paper presented at the 50th Annual Meeting of the Society for American Archaeology (Denver)

FERNÁNDEZ CASTRO M. C. 1982, *Villas romanas in España.* Ministerio de Cultura (Madrid)

FERNÁNDEZ FONSECA M. J. 1995, `Spanish Samian ware: fundamentals and references'. *Papers from the Institute of Archaeology* 6, 57-71

FLANNERY K. V. 1976a, `Excavating deep communities by transect samples'. In Flannery (ed.) 1976b, 68-72

FLANNERY K. V. (ed.) 1976b, *The early Mesoamerican village.* American Press (New York, San Francisco, London)

FLANNERY K. V. 1976c, 'The trouble with regional sampling'. In Flannery 1976b, 159-160

FLETCHER M. & LOCK G. 1991, *Digging numbers: elementary statistcs for archaeologists*. University of Oxford. Committee for Archaeology (Oxford)

FOLEY J., VAN DAM A., FEINER S. AND HUGHES J. 1990, *Computer graphics: principles and practice* (second edition). Addison-Wesley Publishing Company (Reading, Massachusetts)

FOTHERINGHAM S. & ROGERSON P. 1994, *Spatial analysis and GIS*. Taylor & Francis (London)

FRAYN J. M. 1993, *Markets and fairs in Roman Italy*. Clarendon Press (Oxford)

FULFORD M. 1982, 'Town and country in Roman Britain — a parasitical relationship'. In Miles (ed.) 1982, 403-419

GAFFNEY V. & VAN LEUSEN M. 1995, 'Postscript — GIS, environmental determinism and archaeology: a parallel text'. In Lock & Stančič (eds.) 1995, 367-382

GAFFNEY V. & STANČIČ Z. 1991, *GIS approaches to regional analysis: a case study of the island of Hvar*. University of Ljubljana (Ljubljana)

GAFFNEY V. & STANČIČ Z. 1992, *Diadorus Siculus and the island of Hvar, Dalmatia: testing the text with GIS*. In Lock & Moffett (eds.) 1992, 113-125

GAINES S. W. (ed.) 1981, *Data bank applications in archaeology*. The University of Arizona Press (Tucson)

GALSTERER H. 1971, *Untersuchungen zum römischen Stadtwesen auf der iberischen Halbinsel* (Berlin)

GERRARD A. J. 1981, *Soils and landforms: an integration of geomorphology and pedology*. Gerge Allen & Unwin (London)

GIARDINA A. (ed.) 1986, *Società romana e impero tardoantico. III. Le merci e gli insediamenti*. (Roma, Bari)

GORGES J.-G. 1979, *Les villas hispano-romaine. Inventaire et problématique archéologiques*. E. de Boccard (Paris)

GRIFFIN J. B. (ed.) 1951, `Essays on archaeological methods'. *Anthropological Papers* 8, University of Michigan Press (Ann Arbor)

GUITART J. 1976, *Baetulo. Arquelogía Urbana.* Monografías Badalonesas 1. Museu de Badalona (Badalona)

HARRIS T. & LOCK G. 1990, `The diffusion of a new technology: a perspective on the adoption of geographic information systems within UK archaeology'. In Allen, Green & Zubrow (eds.) 1990, 33-5

HARRIS T. & LOCK G. 1995, `Towards an evaluation of GIS in European archaeology: the past, present and future of theory and application'. In Lock & Stančič (eds.) 1995, 349-365

HASELGROVE C., MILLETT M. & SMITH I. (eds.) 1985, *Archaeology from the ploughsoil.* Dept of Archaeology and Prehistory, University of Sheffield (Sheffield)

HAYES J. W. 1972, *Late Roman pottery.* The British School at Rome (London)

HAYES J. W. 1980, *A supplement to Late Roman pottery.* British School at Rome (Rome)

HODDER I. R. & ORTON C. 1976 *Spatial analysis in archaeology.* Cambridge University Press (Cambridge)

HODSON F. R., SNEATH P. H. & DORAN J. E. 1966, `Some experiments in the numerical analysis of archaeological data'. *Biometrica* 53 (3 and 4)

HODSON F. R., KENDALL D. G. & TAUTU P. (eds.) 1971, *Mathematics in the archaeological and historical sciences.* University Press (Edinburgh)

HOLE F. & SHAW M. 1967, *Computer analysis of archaeological seriation.* Rice University Studies 53 (3). Rice University Press (Houston)

HORNOS F., CASTRO M. & CRESPO, J. 1985a, `Excavación de urgencia de la villa romana de Baños (Arroyo del Ojanco, Beas de Segura, Jaén)'. *Anuario Arqueológico Andaluz* 1, 210-217

HORNOS F., CHOCLAN C., CRUZ, J. T. 1985b, `Excavación de urgencia de la villa romana del Campillo, Castellar (Jaén)'. *Anuario Arqueológico Andaluz* 1, 217-222

HUGGETT J. & RYAN N. (eds.) 1995, *Computer applications and quantitative methods in archaeology 1994.* British Archaeological Report International Series 600. Tempus Reparatum (Oxford)

JOHNS C. 1971, *Arretine and samian pottery.* The Trustees of the British Museum (London)

KEAY S. J. 1988, *Roman Spain.* Exploring the Roman World Series. British Museum Publications (London)

KEAY S. J. 1990, `Processes in the development of coastal communities of Hispania Citerior in the Republican period'. In Blagg T. & Millett M. (eds.) 1990, 120-150

KEAY, S. J. 1991, `The Ager Tarraconensis in the Late Empire: a model for the economic relationship of town and country in easter Spain?'. In Barker & Lloyd (eds.) 1991, 79-87

KEAY S. J. 1992, `The "Romanisation" of Turdetania', *Oxford Journal of Archaeology* 11(3), 275-315

KEAY S. J., CARRETÉ J. M. & MILLETT M. 1989, `Ciutat i camp en el món romà: le prospeccions a l'Ager Tarraconensis'. *Tribuna d'Arqueologia* 1988-89, 121-129

KEAY S. J. & MILLETT M. 1991, `Surface survey and site recognition in Spain: the Ager Tarraconensis survey and its background'. In Schofield (ed.) 1991, 129-139

KELLER D. R. & D. W. RUPP (eds.) 1983, *Archaeological survey in the Mediterranean area.* British Archaeological Report International Series 155. Tempus Reparatum (Oxford)

KNAPP R. C. 1977, *Aspects of the Roman experience in Iberia, 206-100 BC.* Anejos de Hispania Antiqua, 9. Universitad de Valladolid (Valladolid and Alava)

KVAMME K. L. 1985, `Fundamentals and potentials of Geographic Information System techniques for archaeological spatial research'. Paper presented at the 50th Annual Meeting of the Society for American Archaeology (Denver)

KVAMME K. L. 1994, `GIS graphics vs. spatial statistics: how do they fit together?'. *Archaeological Computing Newsletter* 1994, 1-2

LEWIT T. 1991, *Agricultural production in the Roman economy*, AD 200-400. British Archaeological Reports International Series 568. Tempus Reparatum (Oxford)

LOCK G. & HARRIS T. 1992, `Visualising Spatial Data: the Importance of Geographic Information Systems'. In Reilly & Rahtz (eds.) 1992, 82-96

LOCK G. & MOFFETT J. 1992, *Computer applications and quantitative methods in archaeology 1991*. British Archaeological Report International Series 577. Tempus Reparatum (Oxford)

LOCK G. & STANČIČ Z. (eds.) 1995, *The impact of geographic information systems on archaeology: a European perspective*. Taylor and Francis (London)

MACKIE N. 1983, *Local administration in Roman Spain AD 14-212*. British Archaeological Report International Series 172. Tempus Reparatum (Oxford)

MACREADY S. & THOMPSON F. H. (eds.) 1985, *Archaeological field survey in Britain and abroad*. Society of Antiquaries, Occasional Papers No 6 (London)

MADSEN T. (ed.) 1988, *Multivariate archaeology*. Aarhus University press (Aarhus)

MARSH G. 1981, `London's samian supply and its relationship to the development of the Gallic samian industry'. In Anderson A. C. & Anderson A. S. (eds.) 1981 (i), 173-238

MARÍN DÍAZ A. 1988, *Emigración, colonización y municipalización en la Hispania republicana*. (Granada)

MASSAGRANDE F. A. 1991, *The Raunds area project field survey: a GIS analysis on the pottery distribution.* Unpublished M.Sc. Dissertation. Department of Archaeology, University of Southampton (Southampton)

MASSAGRANDE F. A. forth, `The Romans in south-west Spain: total conquest or partial assimilation? can GIS answer?'. In Computer applications and quantitative analysis in archaeology 1995. Proceedings of the CAA95 conference.

MASSEY F. J. JR. 1951, `The Kolmogorov-Smirnov test for goodness of fit'. *Journal of the American Statistical Association* 46, 68-78

MAYET F. 1984, *Les céramiques sigillées hispaniques: contribution à l'histoire économique de la Péninsule Ibérique sous l'Empire romain* (2 vols.). Publications du Centre Pierre Paris (Paris)

MEFFERT M. 1995, `Spatial relations in Roman Iron Age settlements in the Assendelver Polders, The Netherlands'. In Lock & Stančič (eds.) 1995, 287-299

MILES D. (ed.) 1982, *The Romano-British countryside. Studies in rural settlement and economy.* British Archaeological Report International Series 103. Tempus Reparatum (Oxford)

MILLETT, M. 1991, `Pottery: population or supply patterns? The Ager Tarraconensis approach'. In Barker & Lloyd (eds.) 1991, 18-26

MILLETT M. 1992, `Roman towns and their territories: an archaeological perspective'. In Rich & Wallace-Hadrill (eds.) 1992, 169-189

MIRET M., SANMARTÍ J. & SANTACANA J. 1984, `Distribución espacial de núcleos Ibericos: un ejemplo en el litoral Catalan'. *Arqueológia Espacial* 4, 173-186

MIRET M., SANMARTÍ J. & SANTACANA J. 1991, `From indigenous structures to the Roman world: models for the occupation of central coastal Catalunya'. In Barker & Lloyd (eds.) 1991, 47-53

MIRÓ M. 1987, `El nucli Ibèric de Tarraco: dels seus inicis a la integracio dins la ciutat Romana'. In *De Les Estructures Indigenes a L'Organizacio Provincial Romana de la Hispania Citerior. Jornales Internacionals d'Arqueologia* 53-8, 284-290

NAGY G. & WAGLE S. 1979, `Geographic data processing'. *Computing Survey* 11 (2), 139-181

NOLLA J. M. & CASAS J. 1984, *Carta arqueológica de les comarques de Girona. El pobalment d'època Romana al N.E. de Catalunya.* (Girona)

ORTON C. 1993, `How many pots make five? — an historical review of pottery quantification'. *Archaeometry* 32 (2), 169-184

ORTON C., TYERS P. & VINCE A. 1993, *Pottery in archaeology.* Cambridge Manuals in Archaeology. Cambridge University Press (Cambridge)

OÑA GONZÁLEZ J. 1984, `El poblamiento rurai de epoca Romana en una zona de la ribera de Navarra', *Arqueológia Espacial* 5, 71-93

PADILLA MONGE A. & DURÁN RECIO V. 1990, *Evolución del pobliamento antiguo en el término municipal de Écija.* Editorial Gráficas el Sol (Écija)

PALLÍ AGUILERA F. 1985, *La Via Augusta en Cataluña.* Universitad Autónoma de Barcelona (Bellaterra)

PEACOCK D. P. S., BEJAOUI F. & BEN LAZREG N. 1990, `Roman pottery production in central Tunisia'. *Journal of Roman Archaeology* 3, 59-84

PERCIVAL J. 1976, *The Roman villa.* B. T. Batsford (London)

PEÑA M. J. 1984, `Apuntes y observaciones sobre las primeras fundaciones romanas en Hispania'. *Estudios de la Antegüedad* 1, 49-85

PLOG S. 1976, `Relative efficiencies of sampling techniques for archaeological surveys'. In Flannery (ed.) 1976b, 136-158

PLOG S., PLOG F. & WAIT W. 1978, `Decision making in modern surveys'. In Schiffer (ed.) 1978 vol. 1, 383-421

PONSICH M. 1974, *Implantation rurale antique sur le Bas Guadalquivir* vol. I (Paris)

PONSICH M. 1979, *Implantation rurale antique sur le Bas Guadalquivir* vol. II (Paris)

PONSICH M. 1987, *Implantation rurale antique sur le Bas Guadalquivir* vol. III (Paris)

PONSICH M. 1991, *Implantation rurale antique sur le Bas Guadalquivir* vol. IV (Paris)

PREVOSTI I MONCLÚS M. 1981a, *Cronologia i poblament a l'area rural d'Iluro* (2 vols). Rafael Dalmau (Barcelona)

PREVOSTI MONCLÚS M. 1981b, *Cronologia i poblament a l'area rural de Baetulo.* Monografies Badalonines N°3. Museu de Badalona (Badalona)

PREVOSTI MONCLÚS M. 1984, `L'estudi del món rural romà. Un programa metodològic'. *Fonaments* 4, 161-211

PREVOSTI MONCLÚS M. 1991, `The establishment of the villa system in the Maresme (Catalonia) and its development in the Roman period'. In Barker & Lloyd (eds.) 1991, 135-141

PURCELL N. 1990, `The creation of provincial landscape: the Roman impact on Cisalpine Gaul'. In Blagg & Millett (eds.) 1990, 6-29

REILLY P. & RAHTZ S. 1992, *Archaeology and the information age.* Routledge (London & New York)

REMESAL J. 1986, *La annona militaris y la exportación de aceite Bético a Germania.* Editorial de la Universidad Complutense (Madrid)

RENFREW C. & BAHN P. 1991, *Archaeology. Theories, methods and practice.* Thames and Hudson (London)

REYNOSO C. 1994, `VB.GIS 3D — A development report on a GIS model for archaeology'. *Archaeological Computing Newsletter* 38 (Spring 1994), 3-7

RHIND D. 1977, `Computer-aided cartography'. *Transactions of the Institute of British Geographers* 2, 71-97

RIBAS M. 1952, *El poblamento d'Iluro.* (Barcelona)

RICH J. & WALLACE-HADRILL A. (eds.) 1992, *City and country in the ancient world.* Routledge (London, New York)

RICHARDSON J. 1986, *Hispaniae. Spain and the development of Roman imperialism 218-82 BC.* (Cambridge)

RIVET A. L. F. 1969, *The Roman villa in Britain.* Routledge & K. Paul (London)

RODRÍGUEZ NEILA J. F. 1988, *Del almanecer prehistórico al ocaso visigodo.* (Córdoba)

ROMERO CARNICERO M. V. 1983, `Aspectos formales de la sigillata hispánica'. *Boletín del Seminario de Arte y Arqueología* 49, 105-34

ROMEU G. 1959, `Una lápida romana en San Pere Molents'. *Ampurias* 20, 302

RUIZ DELGADO M. M. 1985, *Carta Arqueológica de la Campiña Sevillana — Zona Sureste I.* Publicaciones de la Universitad de Sevilla (Sevilla)

RUIZ RODRÍGUEZ A. 1988, `Reflexiones sobre algunos conceptos de la arqueología espacial a partir de una experiencia: Iberos en el Alto Guadalquivir'. *Arqueología Espacial* 12, 157-172

RUIZ RODRÍGUEZ A. & MOLINOS M. 1985, `Informe de la campaña de excavaciones de 1985 en el Cerro de la Plazade Armas de Puente Tablas (Jaén)'. *Anuario Arqueológico Andaluz* 2, 345-53

RUIZ RODRÍGUEZ A., MOLINOS M. & CASTRO LÓPEZ M. 1991, `Settlement and continuity in the territory of the Guadalquivir Valley (6th century BC-1st century AD)'. In Barker & Lloyds (eds.) 1991, 29-36

RYAN N. 1995, `The excavation archive as hyperdocument?'. In Huggett & Ryan (eds.) 1995, 211-220

SANMARTÍ J. & SANTACANA J. 1986, `La jerarquia de nuclis en el pobalment Ibèric de la costa del Penedès. Protohistoria Catalana'. *6e Col.loqui Internacional d'Arqueologia de Puigcerdà*, 227-243

SANMARTÍ J., SANTACANA J. & SERRA R. 1984, *El jaciment Ibèric de l'Argilera i el poblament protohistòric al Baix Penedès.* Quadern de Treball 6 (Barcelona)

SANMARTÍ-GREGO E. 1978, *La cerámica campaniense de Emporion y Rhode.* Monografies Emporitanes IV, Tom I (Barcelona)

SCHIFFER M. B. (ed.) 1978, *Advances in archaeological method and theory.* Academic Press (New York)

SCHOFIELD A. J. (ed.) 1991, *Interpreting artefact scatters. Contributions to ploughzone archaeology.* Oxbow Monograph 4, Oxbow Books (Oxford)

SCHOLTZ S. & CHENHALL R. G. 1976, `Archaeological data banks in theory and practice'. *American Antiquity* 41 (2), 89-96

SERRA RÀFOLS J. DE C. 1928, *Forma Conventus Tarraconensis I: Baetulo-Blanda.* (Barcelona)

SHENNAN S. J. 1988, *Quantifying archaeology.* Edinburgh University Press (Edinburgh)

SILLIÈRES P. 1976, `La Via Augusta de Cordue à Cadix'. *Mélanges de la Casa de Velázquez* 12, 27-65

SILLIÈRES P. 1990, *Les voies de communication de l'Hispanie méridionale.* Diffusion de Boccarid (Paris)

SOLÉ SABARÍS L. 1958-68, *Geografía de Catalunya* (3 vols). Aedos (Barcelona)

STEWART J. D. (ed.) 1980, *Microcomputers in archaeology: proceedings of a seminar held in the Institute of Archaeology, 18 June 1980.* Museum Documentation Association occasional paper 4 (Duxford, Cambridgeshire)

UPHAM S. (ed.) 1979, *Computer graphics in archaeology. Statistical cartographic applications to spatial analysis in archaeological contexts.* Anthropological Research Papers No 15. Arizona State University (Tucson)

VILLEDIEU F. 1986, *Turris Lisbonis-Porto Torres (Sardegna). Il contesto delle mura.* In Giardina (ed.) 1986, 145-ss

WHEATLEY D. 1993, `Going over old ground: GIS, archaeological theory and the act of perception'. In Andresen, Madsen & Scollar (eds.), 133-138

WILCOCK J. D. 1971, `Non-statistical applications of the computer in archaeology'. In Hodson, Kendall & Tautu (eds.) 1971, 470-481

WINTER M. C. 1976, `Excavating a shallow community by random sampling quadrats'. In Flannery (ed.) 1976b, 62-67

WOLLE A.-C. & GAMBLE C. S. 1995, `The ENVARCH project'. In Huggett & Ryan (eds.) 1995, 35-41

Appendix A

The χ^2 and Kolmogorov-Smirnov tests

This appendix lists all the tables of data for the χ^2 and Kolmogorov-Smirnov test discussed in the text. The area of the bands (or categories) is expressed in hectares. When a band contained no sites, this was excluded from the analysis. The sites count for the `all periods' data sets in the tests carried out on the data from the Guadalquivir Valley includes sites which did not contain enough diagnostic material to date them, therefore some of the sites in these data sets were not used in the data sets defined by period.

1 The Guadalquivir Valley

1.1 Low status sites on soil type

All periods

Soil type	Area	Area %	Numer of sites	Expected number of sites
Vertisols	432012	28.01	561	431.63
Inceptisols	323608	20.98	177	323.30
Litosols	89668	5.81	25	89.53
Alfisols	100268	6.50	231	100.17
Entisols	596736	38.69	547	596.21

Chi-squared value is: 326.4486
Phi-squared value is: 0.2118
Cramer's V value is: 0.4603

Republic

Soil type	Area	Area %	Numer of sites	Expected number of sites
Vertisols	432012	31.95	9	4.47
Inceptisols	323608	23.93	1	3.35
Entisols	596736	44.13	4	6.18

Chi-squared value is: 6.9983
Phi-squared value is: 0.4999
Cramer's V value is: 0.7070

Early Empire

Soil type	Area	Area %	Numer of sites	Expected number of sites
Vertisols	432012	28.01	86	63.02
Inceptisols	323608	20.98	24	47.20
Litosols	89668	5.81	7	13.07
Alfisols	100268	6.50	21	14.62
Entisols	596736	38.69	87	87.05

Chi-squared value is: 25.3842
Phi-squared value is: 0.1128
Cramer's V value is: 0.3359

Late Empire

Soil type	Area	Area %	Numer of sites	Expected number of sites
Vertisols	432012	29.74	44	34.20
Inceptisols	323608	22.28	20	25.62
Alfisols	100268	6.9	10	7.93
Entisols	596736	41.08	41	47.24

Chi-squared value is: 5.4033
Phi-squared value is: 0.0470
Cramer's V value is: 0.2168

1.2 High status sites on soil types

All periods

Soil type	Area	Area %	Numer of sites	Expected number of sites
Vertisols	432012	28.01	163	133.33
Inceptisols	323608	20.98	55	99.86
Litosols	89668	5.81	5	27.66
Alfisols	100268	6.50	55	30.94
Entisols	596736	38.69	198	184.16

Chi-squared value is: 65.0683
Phi-squared value is: 0.1367
Cramer's V value is: 0.3697

Republic

Soil type	Area	Area %	Numer of sites	Expected number of sites
Vertisols	432012	28.01	6	7
Inceptisols	323608	20.98	5	5.25
Litosols	89668	5.81	1	1.45
Alfisols	100268	6.50	1	1.62
Entisols	596736	38.69	12	9.67

Chi-squared value is: 1.0964
Phi-squared value is: 0.0439
Cramer's V value is: 0.2094

Early Empire

Soil type	Area	Area %	Numer of sites	Expected number of sites
Vertisols	432012	28.01	139	114.28
Inceptisols	323608	20.98	47	85.60
Litosols	89668	5.81	4	23.70
Alfisols	100268	6.50	53	26.52
Entisols	596736	38.69	165	157.86

Chi-squared value is: 65.8950
Phi-squared value is: 0.1615
Cramer's V value is: 0.4019

Late Empire

Soil type	Area	Area %	Numer of sites	Expected number of sites
Vertisols	432012	28.01	45	44.54
Inceptisols	323608	20.98	34	33.36
Litosols	89668	5.81	3	9.24
Alfisols	100268	6.50	13	10.34
Entisols	596736	38.69	64	61.52

Chi-squared value is: 5.0167
Phi-squared value is: 0.0316
Cramer's V value is: 0.1776

1.3 Low status sites on agricultural prediction

All periods

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
5	25896	25896	0.0169	4	4	0.0026	0.0143
6	120760	146656	0.0956	58	62	0.0402	0.0554
7	560552	707208	0.4611	516	578	0.3751	0.086
8	417560	1124768	0.7334	558	1136	0.7372	0.0038
9	141080	1265848	0.8253	237	1373	0.891	0.0657
10	267876	1533724	1	168	1541	1	0

Maximum distance (D) is 0.0860
Critical value at 5% is 0.0346
Critical value at 1% is 0.0415

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
7	560552	560552	0.4499	4	4	0.2857	0.1642
8	417560	978112	0.785	9	13	0.9286	0.1436
10	267876	1245988	1	1	14	1	0

Maximum distance (D) is 0.1642
Critical value at 5% is 0.3635
Critical value at 1% is 0.4356

Early Empire

Cate gory	Area	Cum. Area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
5	25896	25896	0.0169	3	3	0.0133	0.0036
6	120760	146656	0.0956	11	14	0.0622	0.0334
7	560552	707208	0.4611	80	94	0.4178	0.0433
8	417560	1124768	0.7334	87	181	0.8044	0.071
9	141080	1265848	0.8253	21	202	0.8978	0.0725
10	267876	1533724	1	23	225	1	0

Maximum distance (D) is 0.0725
Critical value at 5% is 0.0907
Critical value at 1% is 0.1087

Late Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
6	120760	120760	0.0801	1	1	0.0087	0.0714
7	560552	681312	0.4518	40	41	0.3565	0.0953
8	417560	1098872	0.7288	45	86	0.7478	0.019
9	141080	1239952	0.8223	9	95	0.8261	0.0038
10	267876	1507828	1	20	115	1	0

Maximum distance (D) is 0.0953
Critical value at 5% is 0.1268
Critical value at 1% is 0.152

1.4 High status sites on agricultural prediction

All periods

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
5	25896	25896	0.0169	3	3	0.0063	0.0106
6	120760	146656	0.0956	14	17	0.0357	0.0599
7	560552	707208	0.4611	189	206	0.4328	0.0283
8	417560	1124768	0.7334	162	368	0.7731	0.0397
9	141080	1265848	0.8253	55	423	0.8887	0.0634
10	267876	1533724	1	53	476	1	0

Maximum distance (D) is 0.0634
Critical value at 5% is 0.0623
Critical value at 1% is 0.0747

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
5	25896	25896	0.0169	1	1	0.04	0.0231
6	120760	146656	0.0956	1	2	0.08	0.0156
7	560552	707208	0.4611	11	13	0.52	0.0589
8	417560	1124768	0.7334	6	19	0.76	0.0266
9	141080	1265848	0.8253	1	20	0.8	0.0253
10	267876	1533724	1	5	25	1	0

Maximum distance (D) is 0.0589
Critical value at 5% is 0.2720
Critical value at 1% is 0.3260

Early Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
5	25896	25896	0.0169	3	3	0.0074	0.0095
6	120760	146656	0.0956	12	15	0.0368	0.0588
7	560552	707208	0.4611	157	172	0.4216	0.0395
8	417560	1124768	0.7334	138	310	0.7598	0.0264
9	141080	1265848	0.8253	53	363	0.8897	0.0644
10	267876	1533724	1	45	408	1	0

Maximum distance (D) is 0.0644
Critical value at 5% is 0.0673
Critical value at 1% is 0.0807

Late Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
5	25896	25896	0.0169	1	1	0.0063	0.0106
6	120760	146656	0.0956	7	8	0.0503	0.0453
7	560552	707208	0.4611	59	67	0.4214	0.0397
8	417560	1124768	0.7334	46	113	0.7107	0.0227
9	141080	1265848	0.8253	14	127	0.7987	0.0266
10	267876	1533724	1	32	159	1	0

Maximum distance (D) is 0.0453
Critical value at 5% is 0.1079
Critical value at 1% is 0.1293

1.5 Low status sites on distance from Roman towns

All periods

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	34368	34368	0.0256	82	82	0.0532	0.0276
2	102632	137000	0.102	243	325	0.2109	0.1089
3	162192	299192	0.2227	289	614	0.3984	0.1757
4	194056	493248	0.3672	298	912	0.5918	0.2246
5	189632	682880	0.5084	229	1141	0.7404	0.232
6	162096	844976	0.6291	182	1323	0.8585	0.2294
7	122140	967116	0.72	114	1437	0.9325	0.2125
8	102296	1069412	0.7962	51	1488	0.9656	0.1694
9	87908	1157320	0.8616	25	1513	0.9818	0.1202
10	71448	1228768	0.9148	18	1531	0.9935	0.0787
11	59984	1288752	0.9595	8	1539	0.9987	0.0392
12	54440	1343192	1	2	1541	1	0

Maximum distance (D) is 0.2320
Critical value at 5% is 0.0346
Critical value at 1% is 0.0415

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
3	162192	162192	0.1739	5	5	0.3571	0.1832
4	194056	356248	0.3821	1	6	0.4286	0.0465
5	189632	545880	0.5854	2	8	0.5714	0.0140
6	162096	707976	0.7593	2	10	0.7143	0.0450
7	122140	830116	0.8903	2	12	0.8571	0.0332
8	102296	932412	1	2	14	1	0

Maximum distance (D) is 0.1832
Critical value at 5% is 0.3635
Critical value at 1% is 0.4356

Early Empire

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	34368	34368	0.0280	15	15	0.0667	0.0387
2	102632	137000	0.1115	32	47	0.2089	0.0974
3	162192	299192	0.2435	52	99	0.4400	0.1965
4	194056	493248	0.4014	43	142	0.6311	0.2297
5	189632	682880	0.5557	28	170	0.7556	0.1999
6	162096	844976	0.6877	30	200	0.8889	0.2012
7	122140	967116	0.7871	17	217	0.9644	0.1773
8	102296	1069412	0.8703	4	221	0.9822	0.1119
9	87908	1157320	0.9419	2	223	0.9911	0.0492
10	71448	1228768	1	2	225	1	0

Maximum distance (D) is 0.2297
Critical value at 5% is 0.0907
Critical value at 1% is 0.1087

Late Empire

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	34368	34368	0.0304	12	12	0.1043	0.0739
2	102632	137000	0.1213	24	36	0.3130	0.1917
3	162192	299192	0.2649	18	54	0.4696	0.2047
4	194056	493248	0.4367	19	73	0.6348	0.1981
5	189632	682880	0.6046	15	88	0.7652	0.1606
6	162096	844976	0.7482	16	104	0.9043	0.1561
7	122140	967116	0.8563	8	112	0.9739	0.1176
8	102296	1069412	0.9469	2	114	0.9913	0.0444
11	59984	1129396	1	1	115	1	0

Maximum distance (D) is 0.2047
Critical value at 5% is 0.1268
Critical value at 1% is 0.1520

1.6 High status sites on distance from towns

All periods

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	34368	34368	0.0256	17	17	0.0357	0.0101
2	102632	137000	0.102	75	92	0.1933	0.0913
3	162192	299192	0.2227	118	210	0.4412	0.2185
4	194056	493248	0.3672	85	295	0.6197	0.2525
5	189632	682880	0.5084	81	376	0.7899	0.2815
6	162096	844976	0.6291	54	430	0.9034	0.2743
7	122140	967116	0.72	22	452	0.9496	0.2296
8	102296	1069412	0.7962	4	456	0.9580	0.1618
9	87908	1157320	0.8616	6	462	0.9706	0.1090
10	71448	1228768	0.9148	7	469	0.9853	0.0705
11	59984	1288752	0.9595	4	473	0.9937	0.0342
12	54440	1343192	1	3	476	1	0

Maximum distance (D) is 0.2815
Critical value at 5% is 0.0623
Critical value at 1% is 0.0747

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
2	102632	102632	0.1186	4	4	0.1600	0.0414
3	162192	264824	0.3061	10	14	0.5600	0.2539
4	194056	458880	0.5305	6	20	0.8000	0.2695
5	189632	648512	0.7497	3	23	0.9200	0.1703
6	162096	810608	0.9371	1	24	0.9600	0.0229
12	54440	865048	1	1	25	1	0

Maximum distance (D) is 0.2695
Critical value at 5% is 0.2720
Critical value at 1% is 0.3660

Early Empire

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	34368	34368	0.0256	14	14	0.0343	0.0087
2	102632	137000	0.102	61	75	0.1838	0.0818
3	162192	299192	0.2227	100	175	0.4289	0.2062
4	194056	493248	0.3672	73	248	0.6078	0.2406
5	189632	682880	0.5084	69	317	0.7770	0.2686
6	162096	844976	0.6291	48	365	0.8946	0.2655
7	122140	967116	0.72	20	385	0.9436	0.2236
8	102296	1069412	0.7962	4	389	0.9534	0.1572
9	87908	1157320	0.8616	6	395	0.9681	0.1065
10	71448	1228768	0.9148	7	402	0.9853	0.0705
11	59984	1288752	0.9595	3	405	0.9926	0.0331
12	54440	1343192	1	3	408	1	0

Maximum distance (D) is 0.2686
Critical value at 5% is 0.0673
Critical value at 1% is 0.0807

Late Empire

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	34368	34368	0.0304	9	9	0.566	0.0262
2	102632	137000	0.1213	24	33	0.2075	0.0862
3	162192	299192	0.2649	36	69	0.4340	0.1691
4	194056	493248	0.4367	23	92	0.5786	0.1419
5	189632	682880	0.6046	33	125	0.7862	0.1816
6	162096	844976	0.7482	22	147	0.9245	0.1763
7	122140	967116	0.8563	8	155	0.9748	0.1185
8	102296	1069412	0.9469	3	158	0.9937	0.0468
11	59984	1129396	1	1	159	1	0

Maximum distance (D) is 0.1816
Critical value at 5% is 0.1079
Critical value at 1% is 0.1293

1.7 Low status sites on distance from the navigable rivers

All periods

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	253896	253896	0.1839	384	384	0.2492	0.0653
2	188032	441928	0.3201	229	613	0.3978	0.0777
3	158804	600732	0.4352	235	848	0.5503	0.1151
4	119932	720664	0.522	144	992	0.6437	0.1217
5	103132	823796	0.5967	162	1154	0.7489	0.1522
6	91656	915452	0.6631	140	1294	0.8397	0.1766
7	77056	992508	0.719	96	1390	0.902	0.183
8	67852	1060360	0.7681	71	1461	0.9481	0.18
9	60692	1121052	0.8121	45	1506	0.9773	0.1652
10	53832	1174884	0.8511	11	1517	0.9844	0.1333
11	46324	1221208	0.8846	6	1523	0.9883	0.1037
12	43228	1264436	0.9159	6	1529	0.9922	0.0763
13	40880	1305316	0.9455	4	1533	0.9948	0.0493
14	38468	1343784	0.9734	7	1540	0.9994	0.026
15	36700	1380484	1	1	1541	1	0

Maximum distance (D) is 0.1830
Critical value at 5% is 0.0346
Critical value at 1% is 0.0415

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
2	188032	188032	0.256	3	3	0.214	0.0417
3	158804	346836	0.4721	3	6	0.4286	0.0435
4	119932	466768	0.6354	2	8	0.5714	0.0640
6	91656	558424	0.7601	2	10	0.7143	0.0458
7	77056	635480	0.8650	2	12	0.8571	0.0079
9	60692	696172	0.9476	1	13	0.9286	0.0190
14	38468	734640	1	1	14	0.9286	0

Maximum distance (D) is 0.0640
Critical value at 5% is 0.3635
Critical value at 1% is 0.4356

Early Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	253896	253896	0.1889	55	55	0.2444	0.0555
2	188032	441928	0.3289	33	88	0.3911	0.0622
3	158804	600732	0.4470	31	119	0.5289	0.0819
4	119932	720664	0.5363	18	137	0.6089	0.0726
5	103132	823796	0.6130	24	161	0.7156	0.1026
6	91656	915452	0.6812	15	176	0.7822	0.1010
7	77056	992508	0.7386	8	184	0.8178	0.0792
8	67852	1060360	0.7891	11	195	0.8667	0.0776
9	60692	1121052	0.8343	9	204	0.9067	0.0724
10	53832	1174884	0.8743	7	211	0.9378	0.0635
11	46324	1221208	0.9088	4	215	0.9556	0.10468
12	43228	1264436	0.9410	3	218	0.9689	0.0279
13	40880	1305316	0.9714	1	219	0.9733	0.0019
14	38468	1343784	1	6	225	1	0

Maximum distance (D) is 0.1026
Critical value at 5% is 0.0907
Critical value at 1% is 0.1087

Late Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	253896	253896	0.2161	41	41	0.3565	0.1404
2	188032	441928	0.3761	20	61	0.5304	0.1543
3	158804	600732	0.5113	18	79	0.6870	0.1757
4	119932	720664	0.6134	4	83	0.7217	0.1083
5	103132	823796	0.7012	10	93	0.8087	0.1075
6	91656	915452	0.7792	5	98	0.8522	0.0730
7	77056	992508	0.8448	3	101	0.8783	0.0335
8	67852	1060360	0.9025	7	108	0.9391	0.0366
9	60692	1121052	0.9542	6	114	0.9913	0.0371
10	53832	1174884	1	1	115	1	0

Maximum distance (D) is 0.1757
Critical value at 5% is 0.1268
Critical value at 1% is 0.1520

1.8 High status sites on distance from the navigable rivers

All periods

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	253896	253896	0.1839	104	104	0.2185	0.0346
2	188032	441928	0.3201	51	155	0.3256	0.0055
3	158804	600732	0.4352	62	217	0.4559	0.0207
4	119932	720664	0.522	76	293	0.6155	0.0935
5	103132	823796	0.5967	46	339	0.7122	0.1155
6	91656	915452	0.6631	40	379	0.7962	0.1331
7	77056	992508	0.719	22	401	0.8424	0.1234
8	67852	1060360	0.7681	8	409	0.8592	0.0911
9	60692	1121052	0.8121	7	416	0.8739	0.0618
10	53832	1174884	0.8511	17	433	0.9097	0.0586
11	46324	1221208	0.8846	8	441	0.9265	0.0419
12	43228	1264436	0.9159	10	451	0.9475	0.0316
13	40880	1305316	0.9455	17	468	0.9832	0.0377
14	38468	1343784	0.9734	6	474	0.9958	0.0224
15	36700	1380484	1	2	476	1	0

Maximum distance (D) is 0.1331
Critical value at 5% is 0.0623
Critical value at 1% is 0.0747

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	253896	253896	0.2685	9	9	0.3600	0.0915
2	188032	441928	0.4674	2	11	0.4400	0.0274
3	158804	600732	0.6354	4	15	0.6000	0.0354
4	119932	720664	0.7622	6	21	0.8400	0.0778
6	91656	812320	0.8591	1	22	0.8800	0.0209
10	53832	866152	0.9161	1	23	0.9200	0.0039
13	40880	907032	0.9593	1	24	0.9600	0.0007
14	38468	945500	1	1	25	1	0

Maximum distance (D) is 0.0915
Critical value at 5% is 0.2720
Critical value at 1% is 0.3260

Early Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	253896	253896	0.1839	86	86	0.2108	0.0269
2	188032	441928	0.3201	40	126	0.3088	0.0113
3	158804	600732	0.4352	57	183	0.4485	0.0133
4	119932	720664	0.522	66	249	0.6103	0.0883
5	103132	823796	0.5967	40	289	0.7083	0.1116
6	91656	915452	0.6631	37	326	0.7990	0.1359
7	77056	992508	0.719	16	342	0.8382	0.1192
8	67852	1060360	0.7681	8	350	0.8578	0.0897
9	60692	1121052	0.8121	6	356	0.8725	0.0604
10	53832	1174884	0.8511	16	372	0.9118	0.0607
11	46324	1221208	0.8846	6	378	0.9265	0.0419
12	43228	1264436	0.9159	9	387	0.9828	0.0326
13	40880	1305316	0.9455	14	401	0.9828	0.0373
14	38468	1343784	0.9734	5	406	0.9951	0.0217
15	36700	1380484	1	2	408	1	0

Maximum distance (D) is 0.1359
Critical value at 5% is 0.0673
Critical value at 1% is 0.0807

Late Empire

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	253896	253896	0.2185	54	54	0.3396	0.1211
2	188032	441928	0.3803	28	82	0.5157	0.1354
3	158804	600732	0.5170	29	111	0.6981	0.1811
4	119932	720664	0.6202	24	135	0.8491	0.2289
5	103132	823796	0.7090	10	145	0.9119	0.2029
6	91656	915452	0.7879	6	151	0.9497	0.1618
7	77056	992508	0.8542	2	153	0.9623	0.1081
8	67852	1060360	0.9126	3	156	0.9811	0.0685
9	60692	1121052	0.9648	2	158	0.9937	0.0289
13	40880	1161932	1	1	159	1	0

Maximum distance (D) is 0.2289
Critical value at 5% is 0.1079
Critical value at 1% is 0.1293

1.9 The low status sites on distance from the Roman roads

All periods

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	321940	321940	0.2268	614	614	0.3984	0.1716
2	266316	588256	0.4143	358	972	0.6308	0.2165
3	205432	793688	0.559	215	1187	0.7703	0.2113
4	159672	953360	0.6715	167	1354	0.8787	0.2072
5	112508	1065868	0.7507	86	1440	0.9345	0.1838
6	90368	1156236	0.8144	56	1496	0.9708	0.1564
7	75936	1232172	0.8679	20	1516	0.9838	0.1159
8	60524	1292696	0.9105	9	1525	0.9896	0.0791
9	48756	1341452	0.9449	8	1533	0.9948	0.0499
10	42544	1383996	0.9748	4	1537	0.9974	0.0226
11	35752	1419748	1	4	1541	1	0

Maximum distance (D) is 0.2165
Critical value at 5% is 0.0346
Critical value at 1% is 0.0415

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	321940	321940	0.2858	5	5	0.3571	0.0713
2	266316	588256	0.5222	3	8	0.5714	0.0492
3	205432	793688	0.7046	1	9	0.6429	0.0617
4	159672	953360	0.8464	1	10	0.7143	0.1321
5	112508	1065868	0.9463	3	13	0.9286	0.0177
8	60524	1126392	1	1	14	1	0

Maximum distance (D) is 0.1321
Critical value at 5% is 0.3635
Critical value at 1% is 0.4356

Early Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	321940	321940	0.2338	82	82	0.3644	0.1306
2	266316	588256	0.4271	43	125	0.5556	0.1285
3	205432	793688	0.5763	29	154	0.6844	0.1081
4	159672	953360	0.6922	25	179	0.7956	0.1034
5	112508	1065868	0.7739	12	191	0.8489	0.0750
6	90368	1156236	0.8396	16	207	0.9200	0.0804
7	75936	1232172	0.8947	5	212	0.9422	0.0475
8	60524	1292696	0.9386	5	217	0.9644	0.0258
9	48756	1341452	0.9740	4	221	0.9822	0.0082
11	35752	1377204	1	4	225	1	0

Maximum distance (D) is 0.1306
Critical value at 5% is 0.0907
Critical value at 1% is 0.1087

Late Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	321940	321940	0.2613	62	62	0.5391	0.2778
2	266316	588256	0.4774	23	85	0.7391	0.2617
3	205432	793688	0.6441	10	95	0.8261	0.1820
4	159672	953360	0.7737	8	103	0.8957	0.1220
5	112508	1065868	0.8650	6	109	0.9478	0.0828
6	90368	1156236	0.9384	5	114	0.9913	0.0529
7	75936	1232172	1	1	115	1	0

Maximum distance (D) is 0.2778
Critical value at 5% is 0.1268
Critical value at 1% is 0.1520

1.10 High status sites on distance from Roman roads

All periods

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	321940	321940	0.222	159	159	0.3340	0.112
2	266316	588256	0.4057	116	275	0.5777	0.172
3	205432	793688	0.5474	58	333	0.6996	0.1522
4	159672	953360	0.6575	50	383	0.8046	0.1471
5	112508	1065868	0.7351	25	408	0.8571	0.122
6	90368	1156236	0.7974	19	427	0.8971	0.0997
7	75936	1232172	0.8498	11	438	0.9202	0.0704
8	60524	1292696	0.8915	14	452	0.9496	0.0581
9	48756	1341452	0.9252	6	458	0.9622	0.037
10	42544	1383996	0.9545	11	469	0.9853	0.0308
11	35752	1419748	0.9792	6	475	0.9979	0.0187
12	30232	1449980	1	1	476	1	0

Maximum distance (D) is 0.1720
Critical value at 5% is 0.0623
Critical value at 1% is 0.0747

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	321940	321940	0.2486	10	10	0.4	0.1514
2	266316	588256	0.4542	4	14	0.56	0.1058
3	205432	793688	0.6129	2	16	0.64	0.0271
4	159672	953360	0.7362	2	18	0.72	0.0162
5	112508	1065868	0.823	3	21	0.84	0.017
6	90368	1156236	0.8928	1	22	0.88	0.0128
8	60524	1216760	0.9395	1	23	0.92	0.0195
10	42544	1259304	0.9724	1	24	0.96	0.0124
11	35752	1295056	1	1	25	1	0

Maximum distance (D) is 0.1514
Critical value at 5% is 0.272
Critical value at 1% is 0.326

Early Empire

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	321940	321940	0.222	129	129	0.3162	0.0942
2	266316	588256	0.4057	103	232	0.5686	0.1629
3	205432	793688	0.5474	50	282	0.6912	0.1438
4	159672	953360	0.6575	44	326	0.799	0.1415
5	112508	1065868	0.7351	24	350	0.8578	0.1227
6	90368	1156236	0.7974	17	367	0.8995	0.1021
7	75936	1232172	0.8498	8	375	0.9191	0.0693
8	60524	1292696	0.8915	14	389	0.9534	0.0619
9	48756	1341452	0.9252	4	393	0.9632	0.038
10	42544	1383996	0.9545	9	402	0.9853	0.0308
11	35752	1419748	0.9792	5	407	0.9975	0.0183
12	30232	1449980	1	1	408	1	0

Maximum distance (D) is 0.1629
Critical value at 5% is 0.0673
Critical value at 1% is 0.0807

Late Empire

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	321940	321940	0.2905	74	74	0.4654	0.1749
2	266316	588256	0.5303	42	116	0.7296	0.1989
3	205432	793688	0.7161	20	136	0.8553	0.1392
4	159672	953360	0.8601	20	156	0.9811	0.121
5	112508	1065868	0.9616	2	158	0.9937	0.0321
10	42544	1108412	1	1	159	1	0

Maximum distance (D) is 0.1989
Critical value at 5% is 0.1079
Critical value at 1% is 0.1293

2 The region of Tarragona

2.1 Rural sites on distance from Roman Tarraco

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
1	1499.2	1499.2	0.0356	6	6	0.1277	0.0921
2	5148.6	6647.8	0.1577	11	17	0.3617	0.2040
3	8866.8	15514.6	0.3681	12	29	0.6170	0.2489
4	10695.8	26210.4	0.6219	11	40	0.8511	0.2292
5	9574.5	35784.9	0.8491	5	45	0.9574	0.1083
6	6360.8	42145.7	1	2	47	1	0

Maximum distance (D) is 0.2489
Critical value at 5% is 0.1984
Critical value at 1% is 0.2378

The reign of Augustus

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
1	1499.2	1499.2	0.0356	6	6	0.1395	0.1039
2	5148.6	6647.8	0.1577	10	16	0.3721	0.2144
3	8866.8	15514.6	0.3681	10	26	0.6047	0.2366
4	10695.8	26210.4	0.6219	9	35	0.8140	0.1921
5	9574.5	35784.9	0.8491	5	40	0.9302	0.0811
6	6360.8	42145.7	1	3	43	1	0

Maximum distance (D) is 0.2366
Critical value at 5% is 0.2074
Critical value at 1% is 0.2486

Early Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
1	1499.2	1499.2	0.0356	6	6	0.1132	0.0776
2	5148.6	6647.8	0.1577	10	16	0.3019	0.1442
3	8866.8	15514.6	0.3681	15	31	0.5849	0.2168
4	10695.8	26210.4	0.6219	10	41	0.7736	0.1517
5	9574.5	35784.9	0.8491	8	49	0.9245	0.0754
6	6360.8	42145.7	1	4	53	1	0

Maximum distance (D) is 0.2168
Critical value at 5% is 0.1868
Critical value at 1% is 0.2239

The 3rd century

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
1	1499.2	1499.2	0.0356	4	4	0.1053	0.0697
2	5148.6	6647.8	0.1577	8	12	0.3158	0.1581
3	8866.8	15514.6	0.3681	11	23	0.6053	0.2372
4	10695.8	26210.4	0.6219	6	29	0.7632	0.1413
5	9574.5	35784.9	0.8491	5	34	0.8947	0.0456
6	6360.8	42145.7	1	4	38	1	0

Maximum distance (D) is 0.2372
Critical value at 5% is 0.2206
Critical value at 1% is 0.2644

Late Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Difference
1	1499.2	1499.2	0.0356	4	4	0.1111	0.0755
2	5148.6	6647.8	0.1577	7	11	0.3056	0.1479
3	8866.8	15514.6	0.3681	9	20	0.5556	0.1875
4	10695.8	26210.4	0.6219	6	26	0.7222	0.1003
5	9574.5	35784.9	0.8491	6	32	0.8889	0.0398
6	6360.8	42145.7	1	4	36	1	0

Maximum distance (D) is 0.1875
Critical value at 5% is 0.2267
Critical value at 1% is 0.2717

2.2 Rural sites on distance from the Roman roads

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12628.8	12628.8	0.2764	18	18	0.383	0.1066
2	9302.45	21931.25	0.4801	15	33	0.7021	0.222
3	7662.14	29593.39	0.6478	6	39	0.8298	0.182
4	6754.16	36347.55	0.7957	3	42	0.8936	0.0979
5	5318.49	41666.04	0.9121	3	45	0.9574	0.0453
6	4016.11	45682.15	1	2	47	1	0

Maximum distance (D) is 0.2220
Critical value at 5% is 0.1984
Critical value at 1% is 0.2378

The reign of Augustus

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12628.8	12628.8	0.2764	18	18	0.4186	0.1422
2	9302.45	21931.25	0.4801	13	31	0.7209	0.2408
3	7662.14	29593.39	0.6478	6	37	0.8605	0.2127
4	6754.16	36347.55	0.7957	2	39	0.907	0.1113
5	5318.49	41666.04	0.9121	2	41	0.9535	0.0414
6	4016.11	45682.15	1	2	43	1	0

Maximum distance (D) is 0.2408
Critical value at 5% is 0.2074
Critical value at 1% is 0.2486

The Early Empire

Category	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12628.8	12628.8	0.2764	25	25	0.4717	0.1953
2	9302.45	21931.25	0.4801	15	40	0.7547	0.2746
3	7662.14	29593.39	0.6478	7	47	0.8868	0.239
4	6754.16	36347.55	0.7957	2	49	0.9245	0.1288
5	5318.49	41666.04	0.9121	2	51	0.9623	0.0502
6	4016.11	45682.15	1	2	53	1	0

Maximum distance (D) is 0.2746
Critical value at 5% is 0.1868
Critical value at 1% is 0.2239

The 3rd century

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum number of sites	Cum. prop. of sites	Diff.
1	12628.8	12628.8	0.2764	17	17	0.4474	0.171
2	9302.45	21931.25	0.4801	10	27	0.7105	0.2304
3	7662.14	29593.39	0.6478	5	32	0.8421	0.1943
4	6754.16	36347.55	0.7957	3	35	0.9211	0.1254
5	5318.49	41666.04	0.9121	1	36	0.9474	0.0353
6	4016.11	45682.15	1	2	38	1	0

Maximum distance (D) is 0.2304
Critical value at 5% is 0.2206
Critical value at 1% is 0.2644

The Late Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12628.8	12628.8	0.2764	17	17	0.4722	0.1958
2	9302.45	21931.25	0.4801	8	25	0.6944	0.2143
3	7662.14	29593.39	0.6478	4	29	0.8056	0.1578
4	6754.16	36347.55	0.7957	4	33	0.9167	0.121
5	5318.49	41666.04	0.9121	1	34	0.9444	0.0323
6	4016.11	45682.15	1	2	36	1	0

Maximum distance (D) is 0.2143
Critical value at 5% is 0.2267
Critical value at 1% is 0.2717

2.3 Rural sites on elevation

Iberian

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12067.9	12067.9	0.2844	5	5	0.3571	0.0727
2	11693.7	23761.6	0.56	3	8	0.5714	0.0114
3	10766.4	34528	0.8138	3	11	0.7857	0.0281
4	7901.02	42429.02	1	3	14	1	0

Maximum distance (D) is 0.0727
Critical value at 5% is 0.3635
Critical value at 1% is 0.4356

Republic

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12056.2	12056.2	0.2845	18	18	0.3830	0.0985
2	11686.1	23742.3	0.5602	18	36	0.7660	0.2058
3	10774.5	34516.8	0.8144	9	45	0.9574	0.1430
4	7864.9	42381.7	1	2	47	1	0

Maximum distance (D) is 0.2058
Critical value at 5% is 0.1984
Critical value at 1% is 0.2378

The reign of Augustus

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12056.2	12056.2	0.2845	16	16	0.3721	0.0876
2	11686.1	23742.3	0.5602	17	33	0.7674	0.2072
3	10774.5	34516.8	0.8144	8	41	0.9535	0.1391
4	7864.9	42381.7	1	2	43	1	0

Maximum distance (D) is 0.2072
Critical value at 5% is 0.2074
Critical value at 1% is 0.2486

Early Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12056.2	12056.2	0.2845	24	24	0.4528	0.1683
2	11686.1	23742.3	0.5602	19	43	0.8113	0.2511
3	10774.5	34516.8	0.8144	8	51	0.9623	0.1479
4	7864.9	42381.7	1	2	53	1	0

Maximum distance (D) is 0.2511
Critical value at 5% is 0.1868
Critical value at 1% is 0.2239

The 3rd century

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12056.2	12056.2	0.2845	15	15	0.3947	0.1102
2	11686.1	23742.3	0.5602	15	30	0.7895	0.2293
3	10774.5	34516.8	0.8144	6	36	0.9474	0.1330
4	7864.9	42381.7	1	2	38	1	0

Maximum distance (D) is 0.2593
Critical value at 5% is 0.2206
Critical value at 1% is 0.2644

Late Empire

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12056.2	12056.2	0.2845	15	15	0.4167	0.1322
2	11686.1	23742.3	0.5602	11	26	0.7222	0.1620
3	10774.5	34516.8	0.8144	8	34	0.9444	0.1300
4	7864.9	42381.7	1	2	36	1	0

Maximum distance (D) is 0.1620
Critical value at 5% is 0.2267
Critical value at 1% is 0.2717

Medieval

Cate gory	Area	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	12056.2	12056.2	0.2845	33	33	0.5238	0.2393
2	11686.1	23742.3	0.5602	14	47	0.7460	0.1858
3	10774.5	34516.8	0.8144	11	58	0.9206	0.1062
4	7864.9	42381.7	1	5	63	1	0

Maximum distance (D) is 0.2393
Critical value at 5% is 0.1713
Critical value at 1% is 0.2054

3 The Maresme

3.1 Rural sites on elevation

Iberian

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum prop. of sites	Diff.
1	5639.25	5639.25	0.1989	21	21	0.3281	0.1292
2	6678.75	12318	0.4344	22	43	0.6719	0.2375
3	6885.5	19203.5	0.6772	13	56	0.875	0.1978
4	3992.5	23196	0.818	4	60	0.9375	0.1195
5	2962.5	26158.5	0.9225	1	61	0.9531	0.0306
7	1788	27946.5	0.9855	1	62	0.9688	0.0167
9	346.25	28292.75	0.9978	1	63	0.9844	0.0134
10	63.75	28356.5	1	1	64	1	0

Maximum distance (D) is 0.2375
Critical value at 5% is 0.1700
Critical value at 1% is 0.2037

Republic

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	5639.25	5639.25	0.1829	86	86	0.4503	0.2674
2	6678.75	12318	0.3994	52	138	0.7225	0.3231
3	6885.5	19203.5	0.6227	28	166	0.8691	0.2464
4	3992.5	23196	0.7521	16	182	0.9529	0.2008
5	2962.5	26158.5	0.8482	1	183	0.9581	0.1099
6	2483.75	28642.25	0.9287	2	185	0.9686	0.0399
7	1788	30430.25	0.9867	4	189	0.9895	0.0028
9	346.25	30776.5	0.9979	1	190	0.9948	0.0031
10	63.75	30840.25	1	1	191	1	0

Maximum distance (D) is 0.3231
Critical value at 5% is 0.0984
Critical value at 1% is 0.1179

The reign of Augustus

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	5639.25	5639.25	0.1829	96	96	0.4729	0.29
2	6678.75	12318	0.3994	54	150	0.7389	0.3395
3	6885.5	19203.5	0.6227	28	178	0.8768	0.2541
4	3992.5	23196	0.7521	16	194	0.9557	0.2036
5	2962.5	26158.5	0.8482	1	195	0.9606	0.1124
6	2483.75	28642.25	0.9287	2	197	0.9704	0.0417
7	1788	30430.25	0.9867	4	201	0.9901	0.0034
9	346.25	30776.5	0.9979	1	202	0.9951	0.0028
10	63.75	30840.25	1	1	203	1	0

Maximum distance (D) is 0.3395
Critical value at 5% is 0.0955
Critical value at 1% is 0.1144

The Early Empire

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	5639.25	5639.25	0.1829	102	102	0.4554	0.2725
2	6678.75	12318	0.3994	63	165	0.7366	0.3372
3	6885.5	19203.5	0.6227	29	194	0.8661	0.2434
4	3992.5	23196	0.7521	18	212	0.9464	0.1943
5	2962.5	26158.5	0.8482	2	214	0.9554	0.1072
6	2483.75	28642.25	0.9287	2	216	0.9643	0.0356
7	1788	30430.25	0.9867	6	222	0.9911	0.0044
9	346.25	30776.5	0.9979	1	223	0.9955	0.0024
10	63.75	30840.25	1	1	224	1	0

Maximum distance (D) is 0.3372
Critical value at 5% is 0.0909
Critical value at 1% is 0.1089

The 3rd century

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	5639.25	5639.25	0.1829	96	96	0.4729	0.29
2	6678.75	12318	0.3994	51	147	0.7241	0.3247
3	6885.5	19203.5	0.6227	29	176	0.867	0.2443
4	3992.5	23196	0.7521	16	192	0.9458	0.1937
5	2962.5	26158.5	0.8482	1	193	0.9507	0.1025
6	2483.75	28642.25	0.9287	2	195	0.9606	0.0319
7	1788	30430.25	0.9867	6	201	0.9901	0.0034
9	346.25	30776.5	0.9979	1	202	0.9951	0.0028
10	63.75	30840.25	1	1	203	1	0

Maximum distance (D) is 0.3247
Critical value at 5% is 0.0955
Critical value at 1% is 0.1144

The Late Empire

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	5639.25	5639.25	0.1829	97	97	0.4949	0.312
2	6678.75	12318	0.3994	45	142	0.7245	0.3251
3	6885.5	19203.5	0.6227	26	168	0.8571	0.2344
4	3992.5	23196	0.7521	16	184	0.9388	0.1867
5	2962.5	26158.5	0.8482	2	186	0.949	0.1008
6	2483.75	28642.25	0.9287	2	188	0.9592	0.0305
7	1788	30430.25	0.9867	6	194	0.9898	0.0031
9	346.25	30776.5	0.9979	1	195	0.9949	0.003
10	63.75	30840.25	1	1	196	1	0

Maximum distance (D) is 0.3251
Critical value at 5% is 0.0971
Critical value at 1% is 0.1164

3.2 Rural sites on distance from the Roman towns

Iberian

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	1879.25	1879.25	0.0734	7	7	0.1094	0.036
2	4445.25	6324.5	0.2469	18	25	0.3906	0.1437
3	4146	10470	0.4087	7	32	0.5	0.0913
4	5334	15804.5	0.6169	13	45	0.7031	0.0862
5	4977	20781.5	0.8111	14	59	0.9219	0.1108
6	4838.5	25620	1	5	64	1	0

Maximum distace (D) is 0.1437
Critical value at 5% is 0.17
Critical value at 1% is 0.2037

Republic

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	1879.25	1879.25	0.0734	59	59	0.3089	0.2355
2	4445.25	6324.5	0.2468	39	98	0.5131	0.2662
3	4146	10470.5	0.4087	22	120	0.6283	0.2196
4	5334	15804.5	0.6169	30	150	0.7853	0.1684
5	4977	20781.5	0.8111	34	184	0.9634	0.1523
6	4838.5	25620	1	7	191	1	0

Maximum distace (D) is 0.2662
Critical value at 5% is 0.0984
Critical value at 1% is 0.1179

The reign of Augustus

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	1879.25	1879.25	0.0734	71	71	0.3498	0.2764
2	4445.25	6324.5	0.2468	38	109	0.5369	0.29
3	4146	10470.5	0.4087	23	132	0.6502	0.2415
4	5334	15804.5	0.6169	30	162	0.798	0.1811
5	4977	20781.5	0.8111	34	196	0.9655	0.1544
6	4838.5	25620	1	7	203	1	0

Maximum distace (D) is 0.29
Critical value at 5% is 0.0955
Critical value at 1% is 0.1144

Early Empire

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	1879.25	1879.25	0.0734	77	77	0.3338	0.2704
2	4445.25	6324.5	0.2468	43	120	0.5357	0.2888
3	4146	10470.5	0.4087	30	150	0.6696	0.2609
4	5334	15804.5	0.6169	31	181	0.808	0.1911
5	4977	20781.5	0.8111	36	217	0.9688	0.1577
6	4838.5	25620	1	7	224	1	0

Maximum distace (D) is 0.2888
Critical value at 5% is 0.0909
Critical value at 1% is 0.1089

The 3rd century

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	1879.25	1879.25	0.0734	71	71	0.3498	0.2764
2	4445.25	6324.5	0.2468	39	110	0.5419	0.2950
3	4146	10470.5	0.4087	25	135	0.665	0.2563
4	5334	15804.5	0.6169	28	163	0.803	0.1861
5	4977	20781.5	0.8111	33	196	0.9655	0.1544
6	4838.5	25620	1	7	203	1	0

Maximum distace (D) is 0.295
Critical value at 5% is 0.0955
Critical value at 1% is 0.1144

Late Empire

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Cum. prop. of sites	Diff.
1	1879.25	1879.25	0.0734	69	69	0.352	0.2786
2	4445.25	6324.5	0.2468	37	106	0.5408	0.2939
3	4146	10470.5	0.4087	26	132	0.6735	0.2648
4	5334	15804.5	0.6169	28	160	0.8163	0.1994
5	4977	20781.5	0.8111	31	191	0.9745	0.1634
6	4838.5	25620	1	5	196	1	0

Maximum distace (D) is 0.2939
Critical value at 5% is 0.0971
Critical value at 1% is 0.1164

3.3 Rural sites on distance from the Via Augusta

Iberian

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Prop. number of sites	Diff.
1	4491.75	4491.75	0.2868	29	29	0.4531	0.1663
2	2902	7393.75	0.4721	9	38	0.5938	0.1217
3	2033.5	9427.25	0.6019	15	53	0.8281	0.2262
4	2008.75	11436	0.7302	8	61	0.9531	0.2229
5	2117.5	13553.5	0.8654	1	62	0.9688	0.1034
6	2108.25	15661.75	1	2	64	1	0

Maximum distance (D) is 0.2262
Critical value at 5% is 0.17
Critical value at 1% is 0.2037

Republic

Category	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Prop. number of sites	Diff.
1	4491.75	4491.75	0.2538	109	109	0.5707	0.3169
2	2902	7393.75	0.4177	26	135	0.7068	0.2891
3	2033.5	9427.25	0.5326	35	170	0.8901	0.3575
4	2008.75	11436	0.6461	16	186	0.9738	0.3277
5	2117.5	13553.5	0.7657	2	188	0.9843	0.2186
6	2108.25	15661.75	0.8848	2	190	0.9948	0.11
7	2039	17700.75	1	1	191	1	0

Maximum distance (D) is 0.3575
Critical value at 5% is 0.0894
Critical value at 1% is 0.1179

The reign of Augustus

Category	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Prop. number of sites	Diff.
1	4491.75	4491.75	0.2538	120	120	0.5911	0.3373
2	2902	7393.75	0.4177	27	147	0.7241	0.3064
3	2033.5	9427.25	0.5326	35	182	0.8966	0.364
4	2008.75	11436	0.6461	16	198	0.9754	0.3293
5	2117.5	13553.5	0.7657	2	200	0.9852	0.2195
6	2108.25	15661.75	0.8848	2	202	0.9951	0.1103
7	2039	17700.75	1	1	203	1	0

Maximum distance (D) is 0.364
Critical value at 5% is 0.0955
Critical value at 1% is 0.1144

Early Empire

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Prop. number of sites	Diff.
1	4491.75	4491.75	0.2538	126	126	0.5625	0.3087
2	2902	7393.75	0.4177	33	159	0.7098	0.2921
3	2033.5	9427.25	0.5326	38	197	0.8795	0.3469
4	2008.75	11436	0.6461	21	218	0.9732	0.3271
5	2117.5	13553.5	0.7657	3	221	0.9866	0.2209
6	2108.25	15661.75	0.8848	2	223	0.9955	0.1107
7	2039	17700.75	1	1	224	1	0

Maximum distance (D) is 0.3469
Critical value at 5% is 0.0909
Critical value at 1% is 0.1089

The 3rd century

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Prop. number of sites	Diff.
1	4491.75	4491.75	0.2538	114	114	0.5616	0.3078
2	2902	7393.75	0.4177	30	144	0.7094	0.2917
3	2033.5	9427.25	0.5326	34	178	0.8768	0.3442
4	2008.75	11436	0.6461	20	198	0.9754	0.3293
5	2117.5	13553.5	0.7657	2	200	0.9852	0.2195
6	2108.25	15661.75	0.8848	2	202	0.9951	0.1103
7	2039	17700.75	1	1	203	1	0

Maximum distance (D) is 0.3442
Critical value at 5% is 0.0955
Critical value at 1% is 0.1144

Late Empire

Cate gory	Area in hectares	Cum. area	Cum. area prop.	Number of sites	Cum. number of sites	Prop. number of sites	Diff.
1	4491.75	4491.75	0.2538	112	112	0.5714	0.3176
2	2902	7393.75	0.4177	27	139	0.7092	0.2915
3	2033.5	9427.25	0.5326	33	172	0.8776	0.345
4	2008.75	11436	0.6461	19	191	0.9745	0.3284
5	2117.5	13553.5	0.7657	2	193	0.9847	0.219
6	2108.25	15661.75	0.8848	2	195	0.9949	0.1101
7	2039	17700.75	1	1	196	1	0

Maximum distance (D) is 0.345
Critical value at 5% is 0.0971
Critical value at 1% is 0.1164

Appendix B

The Correspondence Analysis Component loadings

1 The Guadalquivir Valley data set

1.1 The variable loadings on the first three components

Types Name	Qlt	Mas	Inr	Comp1	Cor	Ctr	Comp2	Cor	Ctr	Comp3	Cor	Ctr
CA	489	30	41	-288	108	8	-142	26	5	520	355	102
COM	959	109	134	-432	276	69	-486	349	227	-474	333	311
TSH	845	215	61	-296	556	64	59	22	6	-205	267	114
TSA	33	24	31	-141	28	2	-17	0	0	-54	4	1
TSSG	948	98	100	-118	25	5	-471	397	191	542	526	365
TSC	908	245	47	267	671	59	-114	122	28	111	115	38
TSC_A	889	74	112	790	748	156	312	116	63	-143	25	19
TSC_C	792	14	63	-1093	467	55	892	311	95	196	15	7
TSC_D	944	129	178	791	823	273	301	119	102	-38	2	2
PFINE	984	64	233	-1192	709	308	709	251	282	223	25	40

Average Type QLT: 779

1.2 The object loadings on the first three components

Units Name	Qlt	Mass	Inr	Comp1	Cor	Ctr	Comp2	Cor	Ctr	Comp3	Cor	Ctr
se	868	34	23	548	815	35	140	53	6	-5	0	0
la	840	8	7	300	187	3	58	7	0	-557	646	33
cr	900	9	24	-1061	747	34	480	153	18	-9	0	0
it	890	16	13	638	872	22	88	17	1	-26	1	0
ot	627	19	24	636	575	26	150	32	4	-120	20	3
or	919	24	114*	-1232	592	126	912	324	178	-89	3	2
im	959	67	65	659	811	99	267	133	42	-91	16	7
me	872	20	11	236	192	4	-161	90	5	-414	590	43
ga	631	57	10	-229	550	10	54	31	1	-69	50	3
cm	229	89	21	142	152	6	-100	75	8	17	2	0
na	897	54	36	558	843	57	129	45	8	55	8	2
mn	298	1	5	384	74	1	-285	41	1	-602	183	7
ip	574	6	4	342	370	3	-132	55	1	217	149	4
ar	904	37	27	589	859	43	64	10	1	119	35	7
ax	903	55	41	561	760	58	239	138	27	-45	5	1
ce	893	25	16	555	846	26	131	47	4	-5	0	0
ob	770	73	20	337	734	28	34	7	1	-66	28	4
as	992	100	99*	-485	432	80	33	2	1	551	558	385
al	800	9	28	87	4	0	-862	420	57	816	376	74
co	942	15	34	14	0	0	-805	500	83	756	442	106
mu	559	14	17	-268	105	3	-489	352	30	264	102	12
os	950	18	53	-725	328	32	-352	77	20	-934	544	200
ca	955	2	6	-439	140	2	-574	240	7	-889	575	24
ba	915	23	15	550	837	24	127	45	3	110	33	3
lu	910	22	30	-249	85	5	-759	790	114	-161	36	7
cl	799	30	22	-274	188	8	-485	587	61	-99	25	4
sa	984	52	55	-281	136	14	-651	733	194	-257	114	44

si	839	27	25	-546	575	27	-338	221	27	-148	42	7
ug	953	18	41	-953	712	54	546	234	46	-89	6	2
nb	926	78	113*	-810	817	173	273	93	51	-116	17	13

Average Unit QLT: 817

2 The region of Tarragona data set

2.1 The variable loadings on the first three components

Types Name	Qlt	Mass	Inr	Comp1	Cor	Ctr	Comp2	Cor	Ctr	Comp3	Cor	Ctr
ANF_I	176	175	48	-61	10	2	243	164	40	-22	1	0
ANF_IB	441	138	76	113	18	6	536	395	154	-144	29	15
ANF_A	119	55	74	453	116	36	-53	2	1	-54	2	1
CA	542	101	87	-706	440	160	-316	88	39	124	14	8
TSA	886	32	165*	-1890	529	364	-1501	334	282	-397	23	26
TSH	172	143	59	207	79	19	1	0	0	-224	93	37
TSSG	162	147	44	-57	8	2	143	52	12	-201	102	31
TSC_A	473	120	66	515	366	100	-212	62	21	-181	45	20
TSC_C	810	18	145	1798	312	188	-1666	268	198	1542	230	226
TSC_D	439	37	89	485	74	27	-888	249	113	606	116	70
PFINE	921	32	146	-971	158	96	1061	189	141	1846	573	566

Average Type QLT: 467

2.2 The object loadings on the first three components

Units Name	Qlt	Mass	Inr	Comp1	Cor	Ctr	Comp2	Cor	Ctr	Comp3	Cor	Ctr
al3	577	37	17	-310	157	11	-490	394	34	-124	25	3
al9	605	18	52	1093	322	70	-804	174	46	633	108	38
cal	108	28	12	104	20	1	-63	7	0	-211	81	6
ca5	861	14	44	-1574	593	108	-1033	255	57	-223	12	4
ca6	373	14	13	-3	0	0	606	308	20	-278	65	6
ca8	310	28	5	3	0	0	130	68	2	-245	242	9
ca14	166	14	16	-488	161	10	46	1	0	-75	4	0
co10	777	23	27	-687	310	34	-742	362	49	-399	105	19
co14	718	32	35	-760	407	59	-353	88	16	562	223	52
co16	869	18	33	-1206	623	85	-704	212	35	-281	34	8
co18	280	9	16	46	1	0	768	263	21	-189	16	2
co19	262	32	7	118	49	1	96	33	1	-228	181	9
cr2	661	23	7	255	161	5	280	195	7	-350	305	15
no2	155	37	11	211	109	5	-134	44	3	-27	2	0
pe2	926	23	25	-598	251	26	656	303	38	728	372	63
pe3	193	14	14	392	114	7	21	0	0	-323	78	7
pe4	495	28	19	-318	111	9	302	100	10	508	284	37
pe6	155	37	11	211	109	5	-134	44	3	-27	2	0
pe19	661	23	7	255	161	5	280	195	7	-350	305	15
ma1	935	9	64*	-2307	577	155	-1789	347	114	-309	10	5
ma7	393	18	11	344	154	7	280	102	6	-324	137	10
re1	468	23	23	-600	275	26	-227	39	5	-449	154	24
re2	661	23	7	255	161	5	280	195	7	-350	305	15
rg2	868	14	41	-544	76	13	1208	376	78	1271	416	115
rg5	155	37	11	211	109	5	-134	44	3	-27	2	0
rg7	603	18	32	-315	44	6	857	326	52	724	233	50
rg9	280	9	16	46	1	0	768	263	21	-189	16	2
sa1	875	14	44	-1029	254	46	648	101	23	1474	520	155
se1	589	32	29	754	476	58	-283	67	10	232	45	9
se3	845	41	25	543	376	39	-484	299	38	365	170	28

se4	82	23	16	-92	9	1	218	52	4	-135	20	2
se5	555	28	10	347	252	10	216	98	5	-313	205	14
ta83	111	9	17	133	7	1	142	8	1	-482	95	11
ta85	235	9	20	641	145	12	-208	15	2	-459	74	10
ta86	247	14	16	338	77	5	307	63	5	-398	107	11
ta87	279	18	10	268	97	4	86	10	1	-356	171	12
ta88	489	18	9	90	12	0	455	309	15	-336	168	11
ta95	128	23	9	-70	9	0	155	47	2	-192	72	4
ta96	291	32	16	420	277	18	-64	6	1	-71	8	1
ta102	310	28	5	3	0	0	130	68	2	-245	242	9
ta115	221	18	20	495	170	14	-10	0	0	-273	51	7
ta117	95	9	23	-444	61	6	-310	30	3	-114	4	1
ta125	612	32	18	-259	90	7	454	276	26	429	246	31
vs1	280	9	16	46	1	0	768	263	21	-189	16	2
vs6	307	14	14	155	18	1	447	149	11	-431	139	13
vs56	280	9	16	46	1	0	768	263	21	-189	16	2
vc3	921	14	94*	1658	306	120	-1815	367	177	1489	247	158

Average Unit QLT: 452

3 The Maresme data set

3.1 The variable loadings on the first three components

Types

Name	Qlt	Mass	Inr	Comp1	Cor	Ctr	Comp2	Cor	Ctr	Comp3	Cor	Ctr
CA	536	157	115	-37	2	1	-445	338	188	-340	196	142
TSA	733	114	132	609	400	228	-216	50	32	-512	283	236
TSSG	374	181	96	-388	354	147	12	0	0	91	20	12
TSH	593	181	104	-498	539	241	148	48	24	-57	7	5
TSC_A	212	144	108	-152	38	18	315	164	86	76	10	7
TSC_C	565	53	136	755	277	163	725	255	169	-259	32	28
TSC_D	676	106	147	542	265	168	426	164	117	525	248	231
PFINE	869	64	162	319	50	35	-995	486	383	823	333	340

Average Type QLT: 570

3.2 The object loadings on the first three components

Units

Name	Qlt	Mass	Inr	Comp1	Cor	Ctr	Comp2	Cor	Ctr	Comp3	Cor	Ctr
a1	673	16	7	-57	9	0	-484	660	23	38	4	0
a4	329	5	13	-493	123	7	-533	144	9	-348	61	5
a5	696	8	13	58	3	0	-421	137	9	-849	557	45
a6	717	11	8	-182	52	2	-308	149	6	-573	516	27
a7	902	19	6	275	301	8	339	457	13	-190	143	5
a8	114	16	4	29	5	0	99	53	1	-101	56	1
a9	902	19	6	275	301	8	339	457	13	-190	143	5
a10	621	5	12	-1027	598	30	196	22	1	49	1	0
a11	621	5	12	-1027	598	30	196	22	1	49	1	0
a15	621	5	12	-1027	598	30	196	22	1	49	1	0
a16	803	8	10	-802	642	28	389	151	7	103	11	1
a17	803	8	10	-802	642	28	389	151	7	103	11	1
a18	621	5	12	-1027	598	30	196	22	1	49	1	0
a22	452	8	13	-266	55	3	481	181	11	523	215	17
a23	803	8	10	-802	642	28	389	151	7	103	11	1
a24	902	19	6	275	301	8	339	457	13	-190	143	5
a25	505	8	17	861	430	32	-192	21	2	-305	54	6
a29	118	11	9	-221	72	3	87	11	0	154	35	2
a38	418	5	13	-620	194	11	-366	68	4	-556	156	13

a39	621	5	12	-1027	598	30	196	22	1	49	1	0
a43	902	19	6	275	301	8	339	457	13	-190	143	5
a50	803	8	10	-802	642	28	389	151	7	103	11	1
a51	552	16	10	380	283	12	267	139	7	-257	130	8
a55	673	16	7	-57	9	0	-484	660	23	38	4	0
a56	789	11	7	-623	740	22	18	1	0	-161	49	2
a57	902	19	6	275	301	8	339	457	13	-190	143	5
a58	621	21	5	333	547	13	-9	0	0	122	74	3
a60	188	19	9	201	107	4	-160	68	3	-71	13	1
a61	188	19	9	201	107	4	-160	68	3	-71	13	1
a62	565	5	20	452	66	6	912	269	27	843	230	30
a68	625	8	13	142	16	1	-532	219	14	-711	390	32
a69	339	8	11	-530	257	12	15	0	0	-299	82	6
a74	696	8	13	58	3	0	-421	137	9	-849	557	45
a77	824	13	14	120	17	1	800	754	52	212	53	5
a80	395	19	9	179	82	3	65	11	0	345	303	17
a83	621	21	5	333	547	13	-9	0	0	122	74	3
a87	781	19	4	131	91	2	-265	375	8	243	315	9
em1	699	5	34*	1076	227	33	-1490	435	71	437	37	8
em3	781	19	4	131	91	2	-265	375	8	243	315	9
em4	789	11	7	-623	740	22	18	1	0	-161	49	2
em6	909	5	18	663	158	13	-813	238	21	-1194	513	60
em7	902	19	6	275	301	8	339	457	13	-190	143	5
ma2	523	5	23	1334	503	51	259	19	2	19	0	0
ma3	452	8	13	-266	55	3	481	181	11	523	215	17
ma7	480	13	14	151	28	2	484	285	19	-370	167	14
ma15	673	16	7	-57	9	0	-484	660	23	38	4	0
ma17	781	19	4	131	91	2	-265	375	8	243	315	9
ma24	420	5	14	-753	267	16	570	153	10	27	0	0
ma28	803	8	10	-802	642	28	389	151	7	103	11	1
ma34	761	16	9	345	257	10	-370	296	13	310	208	12
ma35	418	5	13	-620	194	11	-366	68	4	-556	156	13
ma36	809	8	20	-82	3	0	-1172	666	66	538	140	18
ma37	699	13	5	-216	143	3	-92	26	1	-416	530	18
ma52	418	5	13	-620	194	11	-366	68	4	-556	156	13
ma56	621	5	12	-1027	598	30	196	22	1	49	1	0
ma57	680	11	22	880	468	44	583	205	22	-108	7	1
ma58	525	16	9	86	16	1	484	509	23	18	1	0
ma59	329	5	13	-493	123	7	-533	144	9	-348	61	5
ma62	848	13	14	484	283	17	-599	434	29	330	131	11
ma68	902	19	6	275	301	8	339	457	13	-190	143	5
ma70	571	19	10	393	350	15	146	48	2	276	172	11
ma75	329	5	13	-493	123	7	-533	144	9	-348	61	5
ma92	819	11	14	-350	115	7	-788	581	40	363	124	11
ma93	214	8	11	-446	182	9	-97	9	0	-161	24	2
ma96	206	8	12	-214	38	2	-46	2	0	-446	166	12
ma97	55	8	14	90	6	0	-6	0	0	259	49	4
ma111	637	11	15	-417	157	10	-320	93	7	655	388	36
vr4	287	5	14	-626	185	11	402	76	5	235	26	2
vr11	797	8	25	637	163	17	-832	278	33	943	357	56
mo5	824	11	9	-287	116	5	554	430	20	446	278	17
pm2	631	13	14	314	115	7	260	79	5	-612	437	39
pm3	789	11	7	-623	740	22	18	1	0	-161	49	2
pm6	187	8	14	-31	1	0	203	30	2	-461	156	13
pd1	489	8	14	325	73	5	-283	55	4	-725	362	33
pd8	699	13	5	-216	143	3	-92	26	1	-416	530	18
pd9	342	5	18	52	1	0	706	183	16	657	158	18
pd10	118	11	9	-221	72	3	87	11	0	154	35	2
pd11	114	16	4	29	5	0	99	53	1	-101	56	1
pd15	940	11	23	1017	593	59	769	339	38	-119	8	1
pd16	881	16	7	-83	18	1	-221	131	5	523	732	34
pd23	339	8	11	-530	257	12	15	0	0	-299	82	6
te2	781	19	4	131	91	2	-265	375	8	243	315	9

te11	819	11	14	-350	115	7	-788	581	40	363	124	11
ti7	881	16	7	-83	18	1	-221	131	5	523	732	34
ti8	625	8	13	142	16	1	-532	219	14	-711	390	32
ti9	781	19	4	131	91	2	-265	375	8	243	315	9
ti10	339	8	11	-530	257	12	15	0	0	-299	82	6
ti13	824	11	9	-287	116	5	554	430	20	446	278	17
ti14	565	5	20	452	66	6	912	269	27	843	230	30

Average Unit QLT: 605

* = Inertia Outliers

Appendix C

The accompanying diskette

This appendix lists the files contained in the diskette enclosed with this thesis. Two types of files are included: the site database tables and the Turbo Pascal programs described in chapter 3, *The System.*

1 The site database tables

The following database files can be found on the diskette:

- sevilla.dbf — The site database of the Guadalquivir Valley
- tarraco.dbf — The site database of the region of Tarragona
- maresme.dbf — The site database of the Maresme
- baetulo.dbf — The site database from Prevosti Monclús 1981b

1.1 The site database of the Guadalquivir Valley

The structure of the database table containing the site data of the Guadalquivir Valley is as follows:

Field name	Type	Size	Content
CODICE	Character	2	site code
NUMERO	Numeric	4	site number
COOR_1	Numeric	4	easting
COOR_2	Numeric	4	northing
NOME	Character	45	site name
VALORE	Numeric	2	value
TIPO	Character	35	site type
STRUT	Logical	1	pres/abs structures

PAVIMENTO	Character	16	pres/abs floor
TEGOLE	Logical	1	pres/abs tiles
MOLA	Logical	1	pres/abs quern or mill
FORNO	Logical	1	pres/abs kiln
SECOLI	Character	20	date span (according to archaeological unit)
DOLIA	Logical	1	pres/abs dolia
ANF	Logical	1	pres/abs amphorae
CA	Logical	1	pres/abs Black Glaze pottery
COM	Logical	1	pres/abs coarse ware
TSH	Logical	1	pres/abs Spanish Terra Sigillata
TSA	Logical	1	pres/abs Arretine Terra Sigillata
TSSG	Logical	1	pres/abs South Gaulish Terra Sigillata
TSC	Logical	1	pres/abs generic Terra Sigillata Chiara
TSC_A	Logical	1	pres/abs Terra Sigillata Chiara A
TSC_C	Logical	1	pres/abs Terra Sigillata Chiara C
TSC_D	Logical	1	pres/abs Terra Sigillata Chiara D
PFINE	Logical	1	pres/abs Thin Walled ware
TOMBA	Logical	1	pres/abs burial
VETRO	Logical	1	pres/abs glass artifacts
MARMO	Logical	1	pres/abs marble
MONETA	Logical	1	pres/abs coins

1.2 The site database of the region of Tarragona

The structure of the database table containing the site data of the region of Tarragona is as follows:

Field Name	Type	Size	Content
CODICE	Character	2	site code
NUMERO	Numeric	3	site number
COOR_1	Numeric	3	easting

COOR_2	Numeric	3	northing
NOME	Character	51	site name
VALORE	Numeric	2	value
TIPO	Character	17	site type
STRUT	Logical	1	pres/abs structures
PAVIMENTO	Character	16	pres/abs floor
TEGOLE	Logical	1	pres/abs tiles
IB	Logical	1	site existed in Iberian period
REP	Logical	1	site existed during Republic
AUG	Logical	1	site existed under Augustus
AI	Logical	1	site existed in Early Empire
III	Logical	1	site existed in 3rd century
BI	Logical	1	site existed in Late Empire
MED	Logical	1	site existed in Middle Ages
DOLIA	Logical	1	pres/abs dolia
ANF	Logical	1	pres/abs generic amphorae
ANF_I	Logical	1	pres/abs Italian amphorae
ANF_IB	Logical	1	pres/abs Iberian amphorae
ANF_A	Logical	1	pres/abs African amporae
CA	Logical	1	pres/abs Black Glaze pottery
COM	Logical	1	pres/abs coarse ware
TSA	Logical	1	pres/abs Arretina Terra Sigillata
TSH	Logical	1	pres/abs Spanish Terra Sigillata
TSSG	Logical	1	pres/abs South Gaulish Terra Sigillata
TS	Logical	1	pres/abs Generic Terra Sigllata
TSC_A	Logical	1	pres/abs Terra Sigillata Chiara A
TSC_C	Logical	1	pres/abs Terra Sigillata Chiara C
TSC_D	Logical	1	pres/abs Terra Sigillata Chiara D
PFINE	Logical	1	pres/abs Thin Walled ware
VETRO	Logical	1	pres/abs glass artifacts

MARMO	Logical	1	pres/abs marble

1.3 The site database of the Maresme

The site data of the Maresme were kept in two separate tables, one containing the data from the archaeological unit and the other the data from Prevosti Monclús 1981b. The structure of the database table containing the site data from the archaeological unit is as follows:

Field Name	Type	Size	Content
CODICE	Character	2	site code
NUMERO	Numeric	3	site number
COOR_1	Numeric	3	easting
COOR_2	Numeric	3	northing
NOME	Character	51	site name
VALORE	Numeric	2	value
TIPO	Character	17	site type
STRUT	Logical	1	pres/abs structures
PAVIMENTO	Character	16	pres/abs floor
TEGOLE	Logical	1	pres/abs tiles
IB	Logical	1	site existed in Iberian period
REP	Logical	1	site existed during Republic
AUG	Logical	1	site existed under Augustus
AI	Logical	1	site existed in Early Empire
III	Logical	1	site existed in 3rd century
BI	Logical	1	site existed in Late Empire
MED	Logical	1	site existed in Middle Ages
DOLIA	Logical	1	pres/abs dolia
ANF	Logical	1	pres/abs generic amphorae
ANF_I	Logical	1	pres/abs Italian amphorae
ANF_IB	Logical	1	pres/abs Iberian amphorae

CA	Logical	1	pres/abs Black Glaze pottery
COM	Logical	1	pres/abs coarse ware
TSA	Logical	1	pres/abs Arretina Terra Sigillata
TSH	Logical	1	pres/abs Spanish Terra Sigillata
TSSG	Logical	1	pres/abs South Gaulish Terra Sigillata
TS	Logical	1	pres/abs Generic Terra Sigllata
TSC_A	Logical	1	pres/abs Terra Sigillata Chiara A
TSC_C	Logical	1	pres/abs Terra Sigillata Chiara C
TSC_D	Logical	1	pres/abs Terra Sigillata Chiara D
PFINE	Logical	1	pres/abs Thin Walled ware
VETRO	Logical	1	pres/abs glass artifacts
MARMO	Logical	1	pres/abs marble

The structure of the database table containing the data from Prevosti Monclús 1981b is as follows:

Field Name	Type	Size	Dec	Content
NUMERO	Numeric	3		site number
NOME	Character	45		site name
COOR_1	Numeric	8	4	easting
COOR_2	Numeric	8	4	northing
TIPO	Character	10		site type
VALORE	Numeric	2		value
AP	Logical	1		Pre-Balck Glaze pottery
CA	Logical	1		Black Glaze A pottery
CB	Logical	1		Black Glaze B pottery
TSA	Logical	1		Arretine Terra Sigillata
TSSG	Logical	1		South Gaulish Terra Sigillata
TSH	Logical	1		Spanish Terra Sigillata
TSC_A	Logical	1		Terra Sigillata Chiara A

TSC_C	Logical	1	Terra Sigillata Chiara C
TSC_D	Logical	1	Terra Sigillata Chiara D
EST	Logical	1	Stamped Terra Sigillata
TSH_T	Logical	1	Late Spanish Terra Sigillata
PFINE	Logical	1	Thin Walled ware
IB	Logical	1	site existed during Iberian times
REP	Logical	1	site existed during Republic
AUG	Logical	1	site existed under Augustus
AI	Logical	1	site existed in Early Empire
III	Logical	1	site existed in 3rd century
BI	Logical	1	site existed in Late Empire

2 The program files

The following executable program files are included in the diskette. See chapter 3, *The system*, sections 2.2.1 and 2.2.2 for an examplanation of their function.

- axiscnvr.exe — The AxisCovert program
- idrvals.exe — The Idr_Vals program