The handle http://hdl.handle.net/1887/65569 holds various files of this Leiden University dissertation.

**Author:** Casarotto, A.  
**Title:** Spatial patterns in landscape archaeology : a GIS procedure to study settlement organization in early Roman colonial territories  
**Issue Date:** 2018-09-18
Material remains of ancient settlements in and on the soil provide important empirical evidence of activities carried out in the past by the ancient communities that settled and transformed the landscape. Through field survey, archaeologists can detect a surface sample of what once existed. However, these archaeological traces are exposed to the elements and constantly run the risk of being destroyed or displaced. Also their visibility may be changed by post-depositional processes (Schiffer 1987; Leonardi 1992; De Guio 1995). The number, density and pattern of sites as documented by surveyors in the field is therefore inherently biased, and not directly representative of ancient human behavior. Before the collected survey data can be used to reconstruct ancient settlement patterns, thorough checks and balances need to be made. The aim of this book is to provide an approach to these checks and balances by developing a set of interrelated GIS analyses that form a research procedure for the use of legacy survey data in settlement pattern analysis (Parsons 1972; Balkansky 2008; Drennan 2008).

The GIS procedure presented in this book is devised to be widely applicable, and thus may be useful to several landscape archaeology projects that deal with (legacy) site-based survey datasets. However, to describe the steps which make up the method, and to substantiate the theory, a case-study is used. Specifically, this book focuses on a recent landscape archaeology project that investigates patterns in early Roman colonial territories of Italy (a description of this project is offered in section 1.5). The available legacy survey data and the research question posed by this project (described in section 1.5) form a key factor for the development and application of the methodological procedure proposed in this book.

Landscape archaeologists concur that GIS is the ideal platform for the analysis of settlement patterns inferable from survey data (Kvamme 1989; Allen et al. 1990; Lock & Stančič 1995; Gillings et al. 1999; Wheatley & Gillings 2002; Kay & Witcher 2005; Conolly & Lake 2006; Kantner 2008; but see Sharon et al. 2004). GIS facilitates the identification of correlations and associations between the environment and the spatial distribution of archaeological data by combining them in an overlay format. This enables significant data patterns to be identified, which helps providing a better understanding of the archaeological problem under study. Since its emergence in archaeology in the early 1990s, GIS has facilitated the integration and comparison of large survey datasets on an inter-regional level, by allowing for integrated digitalization, dynamic spatial visualization, and efficient comparative testing of data. This book shows the crucial role of GIS in the development of methods that use (legacy) survey data in a proper way, and also shows how these methods are key for the future of settlement pattern archaeology. GIS-based procedures will undoubtedly continue to be developed further and gain importance in archaeological pattern analysis, especially as the huge volume of available data resulting from recent fieldwork will surely trigger more GIS experimentation.

1.1 Survey Archaeology and Regional Analysis: A Focus on Italy

From the 1950s onwards, a widespread exploitation of land for agriculture across the Mediterranean has brought a large amount of archaeological material to the surface (see discussion in Ammerman 1981;
Barker & Llyod 1991; Barker & Mattingly 1999 – 2000; Barker 1996; Bintliff & Sbonias 1999; Pasquinucci & Trenten 1999; Piccarreta & Ceraudo 2000; Alcock & Cherry 2004a; Witcher 2008; Attema et al. 2010, 1 – 26; Campana 2018, 1-26). Over the last decades, data have been collected in several regions by systematically line-walking (mainly) arable fields. Italy is a prime example of a Mediterranean country where survey has been widely applied to collect archaeological data (see e.g. Ward-Perkins 1961; Potter 1979; Yntema 1993; Barker 1995; Marchi & Sabatini 1996; De Guio et al. 1999; Carandini et al. 2002; Stek et al. 2015; Travigia 2016; Vermeulen et al. 2017; for other Mediterranean countries see Wilkinson 2000; Cherry 2003; Alcock & Cherry 2004a). This has resulted in impressively large datasets that can be used to address a whole range of different archaeological and historical issues. Despite the persistent debate on the validity of field survey as a prospection technique, it is still the most widely applied approach to the study of regional landscapes (Alcock & Cherry 2004a; Galaty 2005; Kantner 2005; Kowaleski 2008; Peterson & Drennan 2011; Feinman 2015).

Conventionally, surface concentrations of artifacts with recognizable limits or sufficiently high sherd density are translated by surveyors into units of analysis, that is archaeological sites (for a discussion on site definition see Plog et al. 1978, 365; Gallant 1986; Cherry et al. 1991, 45 – 47; Alcock & Rempel 2006). The recorded archaeological sites are subsequently described in databases and represented as points or polygons on topographic maps. The inventories of archaeological sites compiled during field surveys (i.e. site-based survey datasets) have an important role in heritage preservation strategies. In Italy for example, the Forma Italiae topographic survey program has been producing large registers of sites and monuments that can be visualized in comprehensive archaeological distribution maps (Castagnoli 1978; Azzena & Tascio 1996; Guaitoli 1999; Quilici & Quilici Gigli 2004, 64-80; Sommella 2009; Marchi 2016). These maps can be consulted for several regions across the Italian peninsula and provide valuable tools to modern landscape planners in their decision-making process (for a summary list of survey projects in Italy see Cambi & Terrenato 1994, 21 – 30; Pelgroma 2012; Corazza et al. 2014).

In addition to the cultural resource management function, site-based survey datasets have been used to address historical questions. The availability of legacy survey data from different regions encourages integration and cross-comparison of settlement patterns on an inter-regional level, which is uniquely valuable for research focusing on large-scale cultural processes (e.g. Patterson et al. 2000; 2004; Attema & van Leusen 2004; Patterson 2004; Thompson 2004; Palmisano et al. 2018; for Mediterranean examples outside Italy, e.g. Alcock 1993; Stone 1997; Blanton 2000; Wilkinson 2000; 2003; Diaconopoulos 2004; Bintliff et al. 2007; Farinetti 2011). With regard to Italy, and Italian Roman contexts in particular, these data have been used to study ancient settlement patterns (e.g. Pelgroma 2008; Witcher 2008; Kay & Witcher 2009; Stek et al. 2015; 2016, 36 – 38), demographic trends, and economic strategies of past societies (e.g. Goodchild 2007; 2013; Witcher 2005; Wilson 2008; Fentress 2009; Launaro 2011; Pelgroma 2012; 2013) in a diachronic perspective.

The GIS research procedure proposed in this book is devised to use site-based survey datasets for research purposes (although it can also be useful for cultural resource management). Specifically, it can be used to implement a sound regional-scale investigation of settlement patterns based on a site dataset. Additionally, this method can be useful for the comparative analysis of site datasets from different regions. To show its applicability for both regional pattern analysis and inter-regional comparative pattern analysis, a case-study is presented in this book which is based on legacy site-based survey datasets of regions in central-southern Italy. This analysis aims to improve the understanding of the settlement developments triggered by an important large-scale cultural phenomenon that affected the Italian peninsula during the 3rd - 1st century B.C., namely Roman conquest and colonization.
1.2 THE IMPORTANCE OF LEGACY SURVEY DATA

Legacy survey data can be defined as the archaeological information registered in the early decades of archaeological survey (1950s – 2000s), when survey methodology was in development and had not yet reached the methodological sophistication of today’s surveys and techniques (Witcher 2008). Usually these legacy datasets consist of registers of archaeological sites which are only roughly described and represented as symbols (i.e. points or polygons) on topographic maps. Devising a method for realizing the potential of these legacy survey data is extremely important, especially as recent studies have convincingly argued that the archaeological surface record is disappearing very rapidly (Knudson 1989; Cameron 1994; Wilkinson 2000; Cherry 2003, 157; Kowalewski 2008, 231 – 232 with further references; Renfrew & Bahn 2008, 559-570; Witcher 2008; Campana 2018). Modern anthropogenic transformations (triggered by the intensification of farming practices and urban expansion) and related geomorphological processes (erosion, deposition) are acutely affecting the soil and thus pose major threats to the preservation of the archaeological record contained on and under the surface (Boardman & Poesen 2006; Witcher 2008). It is likely that legacy data will soon be the only available source of information to study past settlement dynamics for many regions.

Survey has been traditionally conceived as a repeatable experiment (but see discussion in Ammerman 1981; Lloyd & Barker 1981). Ideally, archaeologists could come back into the field if necessary and walk it again to collect more or better data. Especially for Italy, this scenario has turned out to no longer be possible in several areas. The increased deforestation and urbanization, and the intensive plowing and land leveling for agriculture, endanger the soil layer by causing extensive soil loss in many farmed and urban areas of the world (see discussion in Borselli et al. 2006; Torri et al. 2006; Van Oost & Govers 2006; WWF 2017). The extent to which this landscape transformation is taking place is remarkable, and for Italy previously unparalleled (see discussion in Torri et al. 2006; ISPRA 2017).

This situation, of course, also has a detrimental effect on the survival, identification, quality, and reliability of the archaeological material recordable nowadays at the surface level. In many places, material scatters have disappeared or have been dramatically damaged. During our recent survey campaigns in southern Italy (in 2013 – 2016, see Chapter 5) we had the opportunity to assess directly in the field how dramatic this situation is, and we concluded that a reduced quality of the material collectable at the surface by today’s surveys is to be expected. As this trend in agriculture is unlikely to stop in the near future, and its changes to the landscape geomorphology are irreversible, archaeologists will need to depend more and more heavily on surveys that have been conducted in the past. For this reason, there is a growing interest in formal procedures to use the available legacy data systematically.

In addition to the threat on the surface record, interest in systematic procedures has also been stimulated by the rapid and widespread sharing of digitalized legacy survey data. Thanks to open access web repositories for data dissemination1, survey data are becoming more and more accessible. This increased availability of digital datasets was already noted by Alcock and Cherry in their seminal introduction to the book Side-by-side survey (2004a). They also stressed the fact that “[...] we are faced with an influx of regionally based archaeological data of unprecedented quantity, quality, and diversity from hundreds of individual survey projects – but no widely agreed procedures for juxtaposing, combining, or synthesizing individual survey datasets” (Alcock & Cherry 2004b, 4). This book helps filling in this gap by providing a GIS-based procedure for the systematic and comparative analysis of existing regional survey datasets to tackle specific large-scale research questions.

---

1 E.g., MAGIS project (http://cgma.depauw.edu/MAGIS/), ARIADNE project (http://www.ariadne-infrastructure.eu/), DANS project (https://dans.knaw.nl/nl), MGD project (http://mappaproject.arch.unipi.it/mod/index.php), Fasti Online Survey project (http://www.fastionline.org/survey/). At Leiden University, the Leiden Centre of Data Science (LCDS, https://www.universiteitleiden.nl/en/science/leiden-centre-of-data-science) is a network of researchers from different disciplines (e.g. computer science, statistics, mathematics, law, natural sciences, and also archaeology), who explore methods to manage the large amounts of data that are becoming available (Big Data) also in archaeology (see discussion in Gattiglia 2015).
1.3 THE INTEGRATION OF LEGACY SURVEY DATASETS FOR INTERREGIONAL COMPARATIVE ANALYSIS

There are two main problems when using legacy survey data for inter-regional analysis of settlement patterns. One problem, which affects both regional and inter-regional investigations, regards the fact that these data are often not accurate enough to perform advanced modeling and formal statistical analysis (see discussion in Crema 2015). The second issue concerns the many methodological differences between individual surveys and the results they produce. The variegated nature of legacy survey datasets challenges integration and thus has significant limitations for inter-regional comparison of settlement patterns (see discussion in Mattingly & Witcher 2004; Witcher 2008; Launaro 2011, 85-88; Fracchia 2013, 185 – 190; Witcher & Craven 2013).

The need is pressing for the development of methods that enable comparative analysis of existing and new datasets (Cherry 1983; Kintigh 2006; Witcher 2008; Kintigh et al. 2014; Feinman 2015). However, despite the recent intense theoretical discussions within the archaeological community on the unexploited potential of legacy data on the one hand, and their many weaknesses on the other, there is a lack of practical methodological solutions which can accommodate the innate problems and eventually enable correct integration for inter-regional comparisons (but see discussion and possible solutions in Van Leusen 2002; Attema & van Leusen 2004; Allison 2008; Witcher 2008; Aloia et al. 2017).

A significant obstacle for the comparison of these datasets is the lack of sufficient metadata, that is data describing data properties regarding, for instance, site definition criteria or survey strategy adopted by survey teams to collect these data (Wise & Miller 1997; Witcher 2008). Many of these datasets were collected using different methods that were often only generically described. This lack of detail in methodological information limits the possibilities to compare different regional datasets. Additionally, the data quality varies a good deal between surveys. These qualitative and methodological differences between surveys are extremely difficult to overcome and significantly challenge systematic integration and comparison of data properties across regions (Ammerman 1981; Witcher 2008; Attema et al. 2010, 19-29). Hence, it is important to underline that a fundamental condition for the correct application of the procedure presented in this book lies in the comparability of the site-based survey datasets selected for the comparison (see Chapter 3).

Thus far only few solid research procedures to use these datasets appropriately have been devised (e.g. Van Leusen 2002; Attema et al. 2010; Feilken 2014). One of the reasons for this is the lack of comparable regional datasets, namely datasets that are comparable because they were collected using the same (or sufficiently similar) survey methodologies and described with a common terminology. To overcome this lack of compatibility, scholars are investigating solutions to make more datasets comparable. Several researchers are currently building formal ontologies and standard vocabularies that will allow integration of a higher number of (legacy) survey datasets. These specialists make a serious effort to try to arrive at a shared consensus on standardization procedures for survey methods and survey datasets. If successful, such ambitious projects will greatly impact inter-regional comparative pattern analysis, and will surely trigger the development of new procedures to formally use survey datasets in settlement pattern

2 Scholars are well-aware that the production of concrete methods should be undertaken now. This is testified by the high number of recent conferences and interdisciplinary collaborations that aim to stimulate discussion on methods for data preservation and (re)use. Just to name an example, the most important and well-known international conference on computer applications and quantitative methods in archaeology (CAA) activated in 2012 the CAA Recycle, a yearly award to encourage scholars using legacy data to share fruitful ideas about the best-practice on data reuse methods and technologies (http://caa-international.org/bursaries/recycle-award/).

3 An interesting project aiming at integration of legacy and new datasets of the suburbium of Rome using standard ontology, was presented in a session organized by Peter Attema, Paolo Carafa, Willem Jongman, and Christopher Smith at the RAC conference in Rome, March 17th 2016 (session 1). Another interesting roundtable session on Conceptual Reference Models (CIDOC CRM approach) was organized by Martijn van Leusen, Tymon de Haas, and Achille Felicetti at the EAA conference in Maastricht, September 1st 2017 (roundtable session 251). Tymon de Haas and Martijn van Leusen (De Haas & Van Leusen 2016) also directed a pilot project funded by DANS (Data Archiving and Networked Services) and the Groningen Institute of Archaeology (GIA, University of Groningen) that aimed to use the CRM approach for integrating archaeological survey datasets.
investigations. The current research shows that when comparable survey datasets are available it is possible to successfully apply the GIS procedure described below (section 1.4) and to shed light on ancient settlement patterns both on a regional and on an inter-regional level. Thus, this book aims to provide a solid research procedure for the use of comparable (legacy) survey datasets (section 1.4).

1.4 OUTLINE OF THE PROPOSED RESEARCH PROCEDURE

1.4.1 PREMISES

The proposed research methodology is based on two fundamental premises. Firstly, (legacy) survey data are intrinsically biased and incomplete, and thus a formal research procedure is required to cope with the limitations and fragmentary nature of this particular type of data. Another fundamental problem is that legacy survey data are usually not collected and described with the detail that sophisticated modeling techniques require to perform sound statistical analysis. A procedure must be developed that can deal with these uncertainties and shortcomings in the source datasets (see also discussion in Orton 2004, esp. section 4.2). Most importantly, the research procedure cannot be technologically-driven, using for instance readymade GIS packages and sophisticated statistical models borrowed from other disciplines, but should, instead, be data-driven. If the data type does not allow for the application of sophisticated statistical tools, we should not use these tools. It is always tempting to apply innovative quantitative tools for the analysis of survey data. However, as these tools are often developed to analyze datasets of very different quality and nature, there is a real risk that such advanced modeling techniques lead to misguided results when it comes to archaeological survey data.

Therefore, it is crucial to develop an analysis strategy that is based on the survey data and their particular properties. As a rule, this requires the development of tailor-made GIS-based strategies, in which only a selection of suitable quantitative and qualitative methods can be applied. This process requires a lot of preliminary experimentation before the best suited methodology surfaces for answering specific research questions. Cutting-edge modeling techniques, when performed solely by virtue of their level of technological or statistical innovation, do not necessarily guarantee that a research question can be tackled correctly (see discussion in Stek 2016). As a preliminary step, therefore, a careful examination of both data quality and research goals is needed, followed by the selection of suitable methods that do justice to the potential of the data as well as addressing their problems.

The second premise regards the issue of comparability between surveys and the data they produce. In this book legacy site datasets have been selected that were collected using comparable recording methods (e.g. survey methodology, space interval between surveyors) and site criteria (e.g. sherd threshold for site detection; scatter size; dating method). In this way, the effect of methodological differences in survey strategies is limited and can be assumed (at least theoretically) not to be responsible for possible inter-regional variations in the spatial patterns inferable from the site distributions of different regional datasets. Still, there are certain site criteria that are subjectively defined by survey teams and, as a rule, are uncertain and vary considerably from survey to survey, and thus from region to region. For example, functional classifications of site types (e.g. farm, hut, villa, hamlet, village) are notoriously difficult to establish and metadata describing type definition criteria convincingly does not as yet exist. Differences in definitions used between surveys are thus very common. These differences are extremely difficult to resolve and constitute an obstacle for data

---

4 Experiments with inductive predictive modeling were carried out at the beginning of this project. This technique predicts zones of expected high and low settlement suitability based on the correlation between known sites and landscapes (Kohler & Parker 1986; Judge & Sebastian 1988; Kvamme 1990; Kamermans & Wansleeben 1999; Verhagen 2007). Since this technique highlights attractive and repulsive ecological units for settlement within the landscape, I initially used it as a means to investigate ancient settlement strategy (for this approach in another Roman context see e.g. Goodchild 2007, 121 - 178). However, because of the nature of the available source datasets, I realized that in doing so I was providing hypothetical reconstructions that worked only under the assumption of representative samples and environmental determinism in ancient location preferences. To put forward sounder reconstructions a testing of these assumptions was in order. Therefore, I decided to use a different strategy and to develop a procedure that incorporates such a testing.
integration and comparison on the level of functional site types. Therefore, in this book such qualitative functional categories are disregarded. Instead, only the more formally established site properties such as site position, size, and chronological range of occupation have been included in the various performed GIS analyses.

1.4.2 PROCEDURE

As will become clear in the following chapters, it is possible to produce useful results through GIS strategies that are based on the above described premises, especially in the case of problem-oriented regional investigations. The best results are attained when using a research methodology that consists of two parts. One part regards the assessment of the degree of distortion characterizing our data (i.e. the application of bias-testing approaches). The other part aims to extract historical meaning from these data. This part constitutes the problem-solving phase of the procedure: through the application of different approaches, it aims to shed light on the archaeological question that is posed by a project. It is important to stress that the order in which these two operations (the bias-testing and problem-solving part) are carried out is flexible and depends very much on the case-study and on the question being asked.

The problem-solving part of the procedure consists of three subsequent, but interrelated analytical steps. The first step is a deductive analysis in which various existing theoretical models on settlement patterns and location preferences are used as testable hypotheses. These models are confronted with a survey dataset in order to establish which one is the most reliable for that particular territory (see Chapter 3). As a second step, a bottom-up, that is inductive, analysis is performed to establish if unexpected patterns emerge, which cannot be explained by the initial working hypothesis (Chapter 6). The analysis starts from the data and highlights patterns in site distributions with respect to the local environment. As a third step, the results of the deductive analysis are compared to the results of the inductive analysis5. It is at this stage that new insights can be gained about the initial theoretical settlement models, namely on how these models functioned in specific regions (see Chapter 7).

There is a long history of discussion in archaeological theory on which type of scientific reasoning is the most appropriate to derive knowledge: deductive or inductive (Salmon 1976; Kamermans & Wansleeben 1999; Wheatley 2004; Verhagen 2007; Kamermans et al. 2009; Bintliff & Pearce 2011; Verhagen & Whitley 2012; Citter 2012; Casarotto 2015). This opposition between deductive and inductive reasoning, however, is only a construct; both approaches are equally valid, and should be combined to achieve a solid interpretation (Glaser & Strauss 1967; Kamermans 2000; Löwenborg 2010; Verhagen & Whitley 2012, 50 – 55; De Guio 2015, 301; Arnoldus et al. 2016). This book, indeed, is a further demonstration that such approaches do not oppose one another, but on the contrary, are complementary. Through the combined application of deductive and inductive methods, and the consequent dialectic argumentation on the results, it was possible to acquire new information from the initial theory and the initial data. What follows now is a more detailed explanation of each of the three steps composing the problem-solving part of the research procedure.

The first step is defined as a theory-driven approach (deductive analysis). It consists of the testing of existing theories on ancient settlement patterns using legacy survey data. In this phase, these data are used simply as a touchstone for testing the tenability of theories. In order to do so, the settlement theories must first be properly described and translated into archaeological surface scenarios. This can be achieved through an analytical exercise. To do this, the archaeological evidence (detectable at the surface by means of field survey) associated with the theoretical models under consideration must be established. In other words, the theoretical models and the site patterns projected by them (i.e. the expected empirical output), are retraced to the

---

5 Besides its utility for settlement pattern analysis, this deductive-inductive procedure has also proved to be useful for archaeological predictive modeling (Verhagen & Whitley 2012; Arnoldous et al. 2016; Patacchini & Nicatore 2016).
archaeological correlates upon which they are based. Then, those correlates that are retrievable from the available legacy survey data, and therefore are testable, can be selected. These correlates play a crucial role as potential indicators for the existence of certain models of settlement organization in antiquity. Especially those empirical correlates that, according to competing settlement theories, would manifest themselves differently, or in total opposition, in the archaeological surface record play a key role in verifying the models’ reliability (e.g. regular or clustered site distribution; low or high site density). By systematically analyzing the selected empirical correlate using suitable GIS quantitative and qualitative tools we can determine which model(s) most likely occurred in the past (for more details on this indicator-based approach see Stek 2016, 151 - 158).

As a second step, the data are tested for patterns using an inductive approach. Whereas the deductive approach is mainly descriptive with respect to patterns exhibited by site-based data, the inductive analysis is, instead, exploratory. The goal in this case is to explore whether relationships can be found between site-based survey data and a set of local landscape characteristics (i.e. environmental and cultural factors characterizing the local context). This eventually enables the gaining of precise insights into the local settlement system behind settlement patterns of a particular region. Significant local correlations spark further understanding on the settlement rationale underlying survey data, and thus enrich the initial theory on ancient settlement systems. By indicating possible variants in settlement location strategy, this inductive analysis prompts researchers to calibrate or refine existing general starting models (namely, those tested in the previous deductive analysis), and to assess how they functioned with respect to the specific historical and ecological region under consideration (see below).

Such a calibration constitutes the third phase of the problem-solving part of the procedure. The calibration occurs in at least three different ways. When there is accordance between the results from the deductive and inductive analyses, the existing settlement theory acquires more credibility. If, instead, discordance exists, partial or total reconsideration of previous theories is in order. In the case in which the inductive analysis discloses patterns that are not in opposition with the general patterns expected by a previous theory, but that need further explanation, researchers may be prompted to fine-tune such a theory in more detailed terms, and for the specific context under consideration. Moreover, the comparison between the results from the deductive and inductive analyses potentially enables moving from observations of patterns to contextual interpretations of these patterns (i.e. aspects of past settlement strategies in specific regions) (Kantner 2005, 1182). Through a process of validation and refinement of an existing theory, this procedure may eventually advance previous knowledge about ancient settlement strategies (see Chapters 6 and 7).

From the above, it is argued that this deductive-inductive research method allows the gaining of new knowledge from a previously existing theory. When the initial theory is disproved because unexpected patterns are found in the data, this triggers the development of new theories. The same goes for the opposite case. If the initial theory has been proven to be reliable, but is still underdeveloped since there is insufficient insight into certain aspects, this approach can enrich it, possibly filling some gaps with supplementary correlations inferred from the data (Verhagen & Whitley 2012, 90). Eventually, these additions to the extant knowledge lead to new theory, and the deductive-inductive cycle can be activated again: the theory fine-tuned with newly acquired insights about certain aspects can be tested, in other regions as well, and if proven to be reliable and leading to new correlations, the researcher may try to feed it back again, with other significant observations inferred from the data (this is seen as the cycle of research, Tashakkori & Teddlie 2003).

However, the deductive-inductive research method described above must be supplemented with bias testing. Survey data are intrinsically biased, and to propose an accurate territorial reconstruction it is necessary to test whether these data are representative. That is to say, it is necessary to test whether (legacy) survey data, from which ancient settlement patterns can be assessed both deductively and inductively,
are affected by distorting factors such as surface visibility and geomorphological processes of erosion and deposition. The research procedure proposed in this book incorporates methods for the evaluation of biases that may alter the regional configuration of archaeological sites in a landscape. The position of this bias testing within the procedure is flexible; we will see that for the specific case-study analyzed in this book, it was performed after the deductive and before the inductive analysis in order to follow the logical flow of the argumentation (Chapters 4 and 5). However, this is not a rule, and researchers may decide, according to the questions they are tackling, when it is the proper moment to carry out the test.

The full integration within the research procedure of GIS methods to assess visibility and geomorphological biases involved in regional field surveys, is the greatest strength of the proposed methodology. Through these integrated methods archaeologists can assess, while performing regional analysis, the degree of incompleteness of their datasets and thus further calibrate their interpretations about settlement patterns. In this way, sound historical reconstructions can hopefully be put forward.

1.5 CASE STUDY

To show how the proposed research procedure works in practice, a particular case-study is considered in this book. This case-study regards a landscape archaeology project that aimed to test two different settlement theories by using (mainly) survey data. These two competing theories describe the impact of early Roman colonization on rural settlement patterns in central and southern Italy in a very different way.

The LERC archaeological project (Landscapes of Early Roman Colonization, led by Dr. Tesse D. Stek and Dr. Jeremia Pelgrom), which this PhD research is part of, aimed to test a new conception of Roman Republican colonization. This project targeted non-urban settlements, such as villages and hill forts, to gain understanding of their societal role during the early Roman expansion in Italy (Pelgrom 2008; Stek & Pelgrom 2013; Pelgrom et al. 2014; Stek et al. 2015; Stek et al. 2016). Traditionally, the settlement organization of colonial landscapes is imagined to have been characterized by the presence of small colonists’ farms regularly and densely distributed across the countryside (in a centuriated area) around a colonial town (e.g. Brown 1980; Rathbone 1981; 2008; Celuzza & Regoli 1982; Settis 1984). The existence of such an agriculture-based city state model has long been conventionally accepted in discourses about Roman colonization (Salmon 1969; see discussion in Pelgrom & Stek 2014; Pelgrom 2014; Stek 2014). LERC challenges this view by testing survey data for the existence of nucleated colonial settlement systems (i.e. villages) in several territories affected by Roman colonization prior to the Second Punic War (before 218 – 201 B.C.). In this context, the PhD research aimed at assessing the viability of these two competing settlement theories. The research question to be tackled using legacy site-based survey data was: which is the most probable settlement model for the organization of colonial rural landscapes?

With this research goal in mind, the GIS-based procedure described in section 1.4 was developed and applied to this case study. As a first step, a testable archaeological correlate for which the two settlement theories expected different spatial patterns was selected (see also Stek 2016). To be suitable for rigorous comparative testing, this archaeological correlate had to have been measured by the previous surveys using comparable methods and criteria (i.e. similar survey methodology, sherd threshold for site detection, space interval between surveyors, and dating method). Based on these parameters, point distributions representing settlement sites

6 The theories tested in the deductive analyses were based on the same biased legacy datasets; in the context of testing their reliability it was therefore crucial to initially use the datasets as they stand, without any bias filtering.

7 I conducted the PhD project at the Faculty of Archaeology of Leiden University from 2013 to 2018 thanks to the LERC funds deriving from a NWO free competition grant (Netherlands Organization for Scientific Research, LERC project number 360-61-040, PI Dr. Tesse D. Stek). For further information on the LERC project: https://landscapesofearlyromancolonization.com/ https://www.universiteitleiden.nl/en/research/research-projects/archaeology/landscapes-of-early-roman-colonization
were chosen as a suitable correlate. Specifically, site distributions registered on topographic maps by survey projects carried out in the colonial landscapes of Venusia, Cosina and Aesernia (central-southern Italy) were considered. The main reason these three colonial territories were selected for a comparative analysis of settlement patterns was because the field surveys conducted there were comparable in terms of survey methodology. This way, we could exclude major differences in survey methodologies as the cause for differences in regional patterns between the territories under consideration. Additionally, the selection of sample areas also depended on the fact that these territories historically played a key role in the Roman colonization of the Italian peninsula (3rd – 1st century B.C.) (see Chapter 3).

First, the research question was addressed deductively. That is, the feasibility of existing settlement theories was assessed by noting to what extent survey data, and specifically the selected testable archaeological correlate (i.e. point distributions of settlement sites), fitted within the patterns expected from these theories. Specifically, the reliability of the two competing settlement models of Roman colonization was tested against the selected archaeological correlate using GIS-based density and pattern analyses. The first model envisages large colonies of people departing from Rome to occupy wild indigenous territories through the establishment of a fortified urban center, regular land division systems for agricultural production in the surrounding territory, and the construction of numerous farms in the allotted holdings resulting from these partitions. Such an impressive re-organization of the landscape would have, in most cases, destroyed previous settlements or driven off native people into marginal zones where they continued settling in villages (see also Attolini et al. 1991, 144 - 145; Marchi & Sabbatini 1996; Carandini et al. 2002, 108 – 110). Therefore, scholars supporting this traditional view would expect survey to have detected in these colonial landscapes a high and even distribution of sites, which ought to have been plotted on maps as evenly dispersed points (but see discussion in Chapter 2 and results in Chapter 3).

Due to a lack of archaeological and epigraphic evidence in support of such a scenario, scholars have lately started to question this view, and an alternative settlement model has been proposed. This new model puts forward a revisionist understanding of early colonial landscapes, rejects the conventional view of radical changes in settlement strategies from pre-colonial to colonial periods, and proposes, instead, the presence of a dominant clustered village-based settlement system both in pre-colonial and early Roman colonial periods (4th - 3rd century B.C.) (Pelgrum 2008; Stek 2009; Stek & Pelgrum 2013; Stek 2018). According to this theory, survey in colonial territories would not record regular site distributions, but rather clusters of archaeological sites. The first analysis presented in this book (Chapter 3) shows that there is a remarkable discrepancy between survey data and the conventional theory that assumes high site density and dispersed patterning. Instead, the archaeological site dataset is better compatible with the clustered configuration proposed more recently. In conclusion, according to a deductive analysis of the raw survey data, the nucleated settlement theory is more probable.

As a second step, however, the question is raised whether the observed nucleated patterns in the datasets may be a result of post-depositional biasing factors, rather than of ancient settlement strategies (Chapters 4 and 5). We investigated erosion and depositional geomorphological processes along with surface visibility conditions of the walked fields, and tested whether these factors could have been responsible for the recording of empty or low site density zones in between clusters of archaeological sites. Testing for distorting factors involved in survey procedures (i.e. biases) is crucial to tackle the research question addressed in this case-study. For instance, the low site density documented in a majority of colonial landscapes in central-southern Italy has been often explained by a supposed poor surface visibility of colonial sites (Rathbone 1981; 2008; see discussion in Pelgrum 2008; 2012). The bias-testing could eventually shed light on whether the smoothly dispersed and densely-settled farm landscape expected by conventional views was invisible to modern archaeological surveys (thus it could not be recorded) or, conversely, whether it is unlikely to have ever existed in the early colonial period (3rd century B.C.).
Field survey methods have their implicit limitations, and can offer only a fragmentary image of ancient landscapes. However, as shown in Chapters 4 and 5, the settlement data recorded in two field surveys actually display representative evidence to detect significant regional patterns and large-scale trends in settlement location preferences (Haggett et al. 1977; Kantner 2005). These patterns and trends can be highly informative of the social use of space in ancient times (see also the discussion in Barker 1991). The reliability of the original survey data was not only confirmed by modeling techniques performed in GIS, but also by new survey campaigns carried out recently in the context of our project (Pelgrom et al. 2014; 2016; Stek et al. 2015; 2016). The direct testing of the legacy data and models in the field with new surveys lent convincing support to the conclusions drawn for this case-study (see also van Leusen 2002; Patterson 2004; Patterson et al. 2004; Kay & Witcher 2005; Attema et al. 2010; De Haas 2011).

In order to gain further understanding on the logic behind the detected settlement clustering, as a next step, a bottom-up, inductive analysis was performed on the legacy survey data (Chapter 6). This not only permitted to further test the highlighted patterns (Chapter 3) but also provided the opportunity to assess whether there were other local correlations which could help explain them for one specific historical and ecological context (Chapter 6). The comparison between deductive and inductive results adds a new explanatory perspective to the lively debate on the interpretation of regional settlement patterns in Roman colonial landscapes. By confronting the results from the array of approaches used in the deductive and the inductive analyses we could note whether or not they corroborated and/or complemented each other. It is especially through this feedback between the deductive and the inductive reasoning that further understanding on settlement patterns in colonial landscapes could be gained.

The comparison between the results from the first-deductive (Chapter 3) and the last-inductive analysis (Chapter 6) showed that the impact of Roman colonization on a newly conquered area of Southern Italy may not have been as harsh as previously imagined. It is possible that, as supported by the LERC nucleated model, the establishment of a colonial population coming from Rome did not immediately lead to radical transformations in the organization of previous landscapes. The colonial-period site clustering displayed in archaeological site distribution maps, whether representing villages or other types of settlement agglomerations, grew organically and complemented, rather than replaced, the pre-colonial settlement system in the Republican period (3rd – 1st century B.C.).

1.6 STRUCTURE OF THE BOOK

This book consists of a collection of articles. These articles are the output of my PhD project, which is part of the larger Landscapes of Early Roman Colonization project introduced earlier. A majority of the articles has already been published in peer-reviewed journals during the course of the PhD, and some others are awaiting publication. Each chapter contains one article. Articles are not presented in chronological order of publication but instead follow the logical flow of the proposed research procedure. This basically consists of the methodological steps a researcher may follow in order to implement a regional pattern analysis based on an existing survey site dataset(s). All chapters, except the second, tackle one step. Chapter 2, instead, provides background about the status quaestionis on settlement theory regarding the organization of colonial landscapes, the case study presented, and the historical debate followed throughout the book.

1.6.1 CHAPTER 2: FAMILIARIZING WITH THE DEBATE

What would the conventional settlement model for Roman colonial landscapes, in theory, have looked like on a site distribution map?

---

8 From site point distributions it is impossible to assert whether the detected clustering corresponded to villages, other types of nucleated settlements or loose groups of sites (see discussion in Roberts 1996). Only excavation and invasive prospection techniques (see discussion in Verhagen 2013) could potentially unlock the characteristics of these clusters, thus permitting precise evaluations about their nature and chronology (Pelgrom et al. 2014; 2016; Stek et al. 2016).
It is difficult to picture how a colonized territory may have appeared using a conventional view without a drawing or a map to illustrate it. It is even more difficult, then, to imagine how archaeological sites representing the surface evidence of colonial-period settlements should have been plotted as dots on a topographic map during pedestrian surveys (i.e. our testable archaeological correlate). Since dots-on-a-map is the empirical correlate used in the case-study presented in this book, in Chapter 2, a quantitative GIS method is proposed to simulate the dispersed distribution of archaeological sites expected by the conventional settlement theory. The visualization of this hypothetical landscape permits to visually assess the huge difference in site number and pattern between what was actually recorded during a field survey in the area of Venosa in southern Italy and what was expected to be recorded. The chapter ends by opening up the possibility that, rather than visibility impediments for site detection, other reasons may have determined the recorded site arrangement, such as a different settlement strategy from the one expected.

This topic is treated in Chapter 2 that corresponds to the article:


1.6.2 CHAPTER 3: DEDUCTIVE ANALYSIS

Which theoretical model fits the colonial site densities and patterns best?

The dispersed and the village settlement models expect radically different settlement densities and patterns in point distributions representing archaeological sites. These expectancies can be tested through GIS point density and pattern tools. In this chapter three colonial landscapes of central-southern Italy are systematically compared, to assess with which settlement model the survey data accord best. This theory-laden analysis indicates which scenario is more likely to have actually existed in the past, and clearly points out which one is improbable. In addition to that, this analysis also shows that the dichotomous interpretative framework (dispersed versus clustered settlement patterning) adopted in this deductive analysis to investigate settlement systems in colonial landscapes may be too narrow: cultural and environmental factors of the local context may have led to the establishment of a more flexible model of settlement occupation.

This topic is treated in Chapter 3 that corresponds to the article:


1.6.3 CHAPTERS 4 AND 5: ASSESSING BIASES

Are the patterns and densities of survey sites the result of biasing factors rather than real location preferences?

It is possible that, to a certain extent, the site point distribution may be patterned by geomorphological and surface visibility biases (e.g. Cherry et al. 1991; Terrenato & Ammerman 1996; Attema 2017). For example, there may exist an association between favorable surface visibility conditions and the detection of localized high site densities (i.e. point clusters). Conversely, in a zone where site density is low and erosion or deposition is active or occurred recently, the lack of archaeological sites may be explained by geomorphological processes affecting the preservation of sites rather than by past constraints against settlements. It is crucial to assess the effect of distorting factors on site discovery when searching for reliable patterns of ancient settlement systems. In these chapters, survey data of the colonial landscapes of Aesernia and Venusia are tested for possible regional biases using two different methodological approaches. For the case of Cosa, this testing is not carried out because the available information on survey methods and visibility conditions is insufficient for the types of analyses applied in these chapters.
Chapters 4 and 5 correspond to the articles:


1.6.4 CHAPTER 6: INDUCTIVE ANALYSIS AND CALIBRATION

What is the influence of local environmental and cultural factors on early colonial settlement?

In Chapter 6 the location preferences and patterns inferred from the legacy data by means of inductive descriptive and quantitative analyses are described. The clustering displayed by the survey data is compared to the natural and cultural local environment to shed further light on the rationale behind its spatial arrangement. Specifically, we evaluate to what degree the settlement patterning is influenced by local environmental and cultural factors. This is done only for the colonial landscape of Venusia because for the other two colonial territories the quantity and/or quality of the available data and/or metadata were insufficient to allow the type of analyses implemented in this chapter. However, the regional patterns of the landscape of Venusia could trigger further comparative testing in other colonial landscapes. The insights gained from the settlement patterns in Venusia can thus contribute to the debate on the rationale of colonial settlement (cf. section 1.4). After having assessed the effect of biases (Chapter 5) and having established that survey data retain substantial representative evidence for inferring ancient patterns and location preferences, in the final part of the book a historical interpretation of the settlement behavior behind the patterns is provided. This is done by combining the results obtained in the previous deductive investigation (Chapter 3) with the results of the inductive analysis.

Chapter 6 corresponds to the article:


1.7 REFERENCES


Attema, P.A.J., Burgers, G.-J. & van Leusen, P.M. 2010. Regional pathways to complexity: Settlement and land-use dynamics in early Italy from the Bronze age to the Republican period. Amsterdam: Amsterdam University Press.


Diacopoulos, L. 2004. Investigating social complexity through regional survey:
“second-generation” analysis of Bronze Age data from the Canadian Palaipaphos survey project, Southwestern Cyprus. *Journal of Mediterranean Archaeology* 17, 59-85.


Kantner, J. 2008. The archaeology of regions: from
discrete analytical toolkit to ubiquitous spatial perspective. *Journal of Archaeological Research* 16(1), 37-81.


Wheatley, D. 2004. Making space for an archaeology of place. *Internet archaeology* 15. doi: 10.11141/ia.15.10


