CHAPTER 4

DEALING WITH RECENT POST-DEPOSITIONAL AND RESEARCH BIASES IN ARCHAEOLOGICAL LANDSCAPES*

1 INTRODUCTION

This chapter builds on a previous publication on bias modelling (Van Leusen 1996) and is about recent post-depositional and research biases in the kind of data that form the basis for archaeological landscape reconstruction and settlement history – site-based data coming from desktop study and older surveys, and land parcel-based data coming from modern surveys. Recent post-depositional biases are nearly exclusively related to human changes to the landscape and its use; research biases are those biases that have occurred in the past, and still occur, during the construction of the archaeological record; I specifically exclude biases occurring during the formation process.

My main point is that recent post-depositional and research biases can not only obscure, but also create patterns in the archaeological record. This has two consequences: firstly, if significant biases in the data we work with are not dealt with, then our reconstructions based on those data will be significantly flawed; secondly, comparison of the archaeological records of the three RPC study regions is predicated on the assumption that such records are, or can be made, comparable.

1.1 AIMS & DEFINITIONS

When archaeologists collect data, they nearly always do so non-randomly. In many cases this is intentional, but it also happens under circumstances where the archaeologist is unaware of, or underestimates, the selectiveness of his or her sample. Even in modern field surveys, where the data collection method is intended to give the archaeologist a representative and un-biased sample of the archaeological material present on the surface, biases introduced by differences in land use, survey conditions, collection methods, and differences between the individuals taking part in the survey may conspire to create a highly non-representative data sample instead. My aim in this article is to outline

* Much work on this article was done while I was working with the Wroxeter Hinterland Project at the Birmingham (UK) University Field Archaeology Unit. My thanks go to Vince Gaffney, for allowing me to develop my interest in GIS-implemented bias modeling, and to Roger White and his volunteers for providing me with some of the most systematically collected survey data imaginable. Similar thanks should go to the surveyors – too numerous to mention here - for the various RPC project campaigns in south-central Italy, and to Peter Attema, for sharing forthcoming survey data. Luke Dalla Bona's kind invitation to attend his Ontario workshop on predictive modelling in spring 1997 provided a second and very welcome push to work on the problem of biases; and as always, members of the GISARCH discussion list were extremely helpful with their comments on various draft versions.
approaches toward dealing with the significant research biases which do occur in our data sets, and to substantiate this with several case studies.

The study of processes transforming the archaeological record is a natural consequence of the theoretical stance of the New Archaeology. Originally, interest centred on depositional and taphonomic processes because these provided the link between archaeological assemblages and their anthropological/ethnographic parallels. As processual archaeology took hold in the later 1970s, a more general interest in formation processes emerged because it was hoped that if these could be modelled in sufficient detail, the ‘original’ record might be reconstructed (‘behavioral archaeology’; Schiffer 1976). This extreme processualist stance led to the emergence of the post-processual movement in the early 1980s.

Unfortunately the latter movement has given rise to a general rejection of ‘rule-finding’ approaches in archaeology, thus throwing away the baby with the bath water. This has been particularly clear in current criticism of GIS-based landscape studies as being ‘environmentally deterministic’ (see discussion in Gaffney & Van Leusen 1995). In view of this, I should stress that this article is not about the modeling of past or current taphonomic processes, but rather about past and current research processes - that is, the processes by which we study the archaeological record1. These have long been recognised as a separate set of distorting processes (eg, Daniels 1972:202) but have had most of their impact north of the Mediterranean region (see esp. Hamond 1978, 1980).

A subsidiary point I should make here is that biases operate at all spatial and temporal scales and resolutions, but that different biases may become significant at different scales and resolutions. For example, although land use and pedological conditions for the fields surveyed in the WHP were relatively uniform, the distribution of these conditions within the larger WHP survey transect and 30 by 40 km study area was decidedly non-random (Van Leusen & Gaffney 1996:303); good examples of the occurrence of biases at larger scales (smaller distances) are provided by the RPC surveys: historic earth movement in the Fogliano area, and the ploughing-up of pebbles from underlying conglomerates in the Francavilla area. I will refer to the issue of scale and resolution throughout this article, especially during the discussion of GIS-based landscape modeling.

In the remaining part of this introductory section, I will review the background and history of the treatment of biases in Mediterranean landscape archaeology (concentrating on Italy), in general methodological studies, and in GIS-based landscape studies. I will then briefly introduce the concept of bias modelling, which will be treated in detail in section 2. This will be followed by a concluding discussion (section 3) which also serves to establish a context for the case studies. These were conducted to demonstrate, firstly, the relevance of bias factors to the interpretation of survey data and of landscape archaeological data in general; and secondly, approaches to the inclusion of bias factors in geographic models of archaeological landscapes. The first of the case studies addresses biases occurring at a regional or supra-regional scale where, usually, ‘sites’ are the analytical unit, concentrating on the effects of differential land use / land cover history in the hinterland of the Roman civitas capital at Wroxeter (Shropshire, England). The second case study addresses research biases occurring at a local or regional scale where land parcels are the analytical unit, and concentrates on visibility-related biases in RPC project surveys in central and south Italy. The third case study again concerns the history of land use, this time in the context of 20th century land improvement of the Fogliano area of the Pontine region (chapter 17, published as Feiken & Van Leusen 2001).

---

1 I prefer this to the alternative name ‘discovery processes’ which stresses accidental effects.
1.2 HISTORY & TREATMENT OF BIASES IN MEDITERRANEAN LANDSCAPE
ARCHAEOLOGY

RESEARCH TRADITIONS

The treatment of biases is intimately tied up with more general traditions of archaeological research. Terrenato (1996) began his discussion of landscape archaeology in the Italian regions of Etruria and Umbria by remarking that ‘it did not succeed in standardising itself in the period 1960 – 1980’ (nor later; see also Cambi & Terrenato 1994), and that, in order to compare the settlement histories of these regions, we must first understand what methods were used in each. He defined two methodological traditions: the *Forma Italiae* approach originating in the early 20th century (including such non mapsheet-based projects as the South Etruria Survey), and the systematic sampling designs using transects that began in the late 1970s (such as the Ager Cosanus, Ager Tarraconensis, and Agro Pontino projects; Terrenato 1996:217-220). The latter all shared one important design trait - their stratified sampling designs were a-spatial, in the sense that they aimed to answer questions about properties of the sampled population as a whole, rather than to record spatial variables. Their aim was to establish differences and similarities between the strata which could be interpreted as differential use of landscape units (see Orton 2000 for a detailed review of methodology). In the late 1980s it became clear that neither approach produced satisfactory results, as ongoing research began to show that vegetation cover and recent geopedological phenomena played an important role in the process of finding sites in Central Italy (e.g., Malone & Stoddart 1994:5).

In the early 1980s the general recognition of the existence and importance of biases led to a phase of experimentation (continuing at a slower pace today) aimed at determining the precise effect of the biases. An important publication because of its effect on survey work carried out by Dutch teams in Italy was Shennans East Hampshire study, which aimed to solve “many problems concerning the significance of information derived from surface survey”, in particular observer effects, differential visibility effects, and the effects of sub-recent landuse (Shennan 1985:2).

From the early 1990s onwards British and Dutch archaeologists in Italy also began to record off-site densities (for example, in the Pontine Region, Gubbio Basin, and Rieti projects) while Italians stuck to the site-based approach developed for the *Forma Italiae* series. The increasing research intensity required by off-site approaches, with its emphasis on high resolution data collection and quantification based on land parcels, has forced many researchers to recognise the importance of coming to terms with biases in the archaeological record. This point was reiterated recently in the influential (but sadly delayed) POPULUS project edited volume on ploughsoil assemblages (Francovich & Patterson 2000). It was recognised at the time (circa 1995) that a practical standardised approach to bias correction would have to be developed (Millett 2000:92-4). On the other hand, the category of ‘site’ remains a mainstay of archaeological interpretation in both research traditions, allowing the issue of bias to be ignored to a greater or lesser extent. Since both of the research traditions mentioned above have contributed data to modern GIS-based regional archaeological records created for landscape approaches to archaeological analysis and interpretation, the challenge now is to conceive of approaches which will be able to deal with biases in wide area, site-based data as well as in local area, parcel-based data.

GIS-BASED ANALYSIS OF ARCHAEOLOGICAL LANDSCAPES

The emergence, in the early 1990s, of GIS as data management and analytical tools in landscape archaeology has had a marked influence on both data collection and analytical methods, including those that have a bearing on biases. The need to collect and standardise data on a regional basis has highlighted gaps, biases, errors and uncertainties. Not surprisingly, low archaeological data quality has been widely identified as a fundamental problem for the application of cartographic modelling techniques in archaeology. Solutions have been sought in two directions: firstly, to accept low data quality and find a

2 Most of these projects still await full publication.
lowest common denominator in point-based archaeological site records; and secondly, to study ways in which data quality can be improved.

Published cartographic models of archaeological landscapes (including my own) have, without exception, been based on point-based records rather than parcel-based distribution maps. However, as I have argued elsewhere (Van Leusen 1996) the set of point observations (‘sites’) resulting from a typical desktop-based survey of an archaeological landscape is equally, though differently, biased by the non-random effects of recent post-depositional research processes. This point will be substantiated in a case study on biases caused by land use, land cover, and research methods in the Wroxeter Hinterland. GIS analyses of point-based archaeological records have also suffered from two deficiencies which have prevented them from dealing adequately with biases, namely the absence of ‘non-site’ observations, and the fact that cartographic models cannot be tested with the aid of independent control samples. The absence of ‘non-site’ observations means that the landscape characteristics of sites cannot be compared to those of non-sites; the lack of independent control samples means that we cannot test cartographic models for biases in our input data.

In order to explain why these two are of such great significance when it comes to avoiding biases, I must briefly digress to explain some principles of locational (location-allocation) modelling in archaeology. Archaeological locational modeling in GIS is the modeling of the probability of the presence of archaeological site locations based on the properties of each topological unit (polygon, cell) in the study area; or in other words, the prediction of the chances of finding a site in any particular location on the basis of models of the ancient landscape. Many such applications have derived landscape ‘predictors’ by studying the known record of archaeological sites and by assuming that this input (the ‘test sample’) is more or less representative of all archaeological sites (the ‘population’). Exactly what that ‘population’ is, almost always remains undefined, but it is implicitly assumed to be either a) original distributions of archaeological sites, b) distributions of sites modified through taphonomic processes, or c) distributions of sites modified by both taphonomic and discovery processes (Hamond 1980, Orton 2000).

In order to find which landscape characteristics are predictive of site location, and which are not, a statistical analysis of site characteristics is performed. However, most such methods require that ‘site’ characteristics be tested against ‘non-site’ characteristics, that is, the test sample should include data about site absence as well as site presence. But records of such negative observations were never kept, and thus the preconditions for such statistical tests cannot be met. To get around this, it has been argued that random observations, given the scarcity of sites, are likely to be non-site observations; and therefore a set of such random observations may be used as a stand-in for the missing non-site records. This method implicitly assumes that the intensity of research has not varied across the study area, but we know in fact that this is not true – both the quality and the intensity of research display huge variations as soon as the scale of the study area transcends that of a single project! Thus, the absence of ‘non-site’ observations in regional archaeological records is one cause of the tendency to ignore spatial variation in taphonomic and discovery processes as a factor in archaeological landscape analysis.

A second way lies in the manner in which models are tested. There are two criteria by which to measure the quality of any prediction: specificity (yield) and accuracy. Models with a low specificity are uninteresting; models with a low accuracy are wrong. The specificity of an archaeological locational model is usually expressed as the ratio of the percentage of input sites and the percentage of total study area. For example: if the model can describe 75% of its input sites by characteristics occurring in only 10% of the landscape, its yield is said to be 75/10 or 7.5. Once a model has been made, its specificity can be calculated; but its accuracy can only be assessed by testing the model - applying it to a set of independent observations which were not used to construct the model in the first place. Because such independent observations cannot be collected in a reasonable amount of time and money by doing new fieldwork, this ‘control sample’ nearly always derives from the same data set of known archaeological sites from which the test sample was taken. Since, ipso facto, it has been subject to the same biases as the test sample of sites, it cannot be used to remove those biases. Therefore, the typical locational ‘predictive’ model is in fact only a descriptive model, and what it describes are the characteristics of the known archaeological sites.
The second major approach to problems with archaeological data quality in landscape studies has been to improve matters either by collecting ‘better’ data or by correcting existing data. Both approaches entail greater attention to the understanding, recording and correction or avoidance of post-depositional and research biases, usually within the context of intensive ‘non-site’ or ‘off-site’ surveys. In contrast to the site-based analyses discussed above, GIS has been applied to the spatially continuous distributions of archaeological data collected in such surveys at a much more local scale (usually not more than a few square km). Since most of these studies – among them the analyses of the WHP- and RPC project surveys - are not yet published, there has not yet been a thoroughgoing discussion of approaches to dealing with biases in such data. The recent publication of Bintliff’s reanalysis of the Boeotia survey data and the discussion that followed in the pages of the Journal of Mediterranean Archaeology (Bintliff et al 1999; Bintliff & Howard 2000; JMA 13) may be regarded as the opening shots in this debate. GIS in general are eminently suited for the management and analysis of archaeological data collected in the form of quantitative measurements per land parcel, but the higher spatial resolution required by modern intensive surveys does present the problem that increasingly low densities of artifact are being recorded, which have a lower diversity, a lower diagnosticity, and are relatively more sensitive to stochastic effects.

The increased emphasis on high resolution quantitative processing has led researchers to devise increasingly formalised and standardised methods for collecting and recording survey data. To avoid slowing down the fieldwork too much, these processes are now being automated through the use of digital field equipment. The potential of portable GPS attached to handheld field computers in particular is now being explored, and recent experiments suggest that future surveys will operate new, powerful and versatile navigation, recording and communication tools which are likely to replace all paper-based recording practises (Van Leusen & Ryan forthcoming, see chapter 7). The gain may be put to good use for increased attention to biases and methods for dealing with them.

TYPES OF BIASES

As indicated in section 1.1, I define biases not as distortions of some idealised archaeological record, but distortions in the way we go about obtaining knowledge about that record, and in that sense all biases are ‘research’ biases. The point I wish to make in the current section is, that research biases can (and do) occur at any stage from the definition of research aims through the design and execution of the fieldwork, and into the analytical and interpretative stages. In order to limit confusion in what follows, I will distinguish the following bias types occurring during these research stages: Conceptual bias, Visibility bias, and Observer bias.

Conceptual biases are biases caused by the classification of data under preconceived concepts. These play an important role at all scales and stages of archaeological enquiry and include the tendency to study only some geographical, typological, and chronological parts of the available archaeological record.

Paraphrasing the quote made famous by Cherry, Mediterranean archaeologists have been like frogs around a pond, with very little study taking place in inland and highland areas. Until very recently archaeological interest in southern Italy for example, was mainly targeted at the Hellenistic colonies there. Following Livy (XXV: I, 1), indigenous settlement was regarded as uninteresting. A comprehensive and systematic overview of indigenous settlement in southern Italy was therefore unavailable until recently, when the University of Lecce began building a regional archaeological site database containing settlements from the 9th (EIA) to the 3rd century BC. This highlighted both the previously ignored presence of a dense network of indigenous settlements contemporary to the Hellenistic colonies, and that of major geographical hiatuses in research, for example in central Puglia and Basilicata (D’Andria 1999). It also highlighted the typological concentration of research (especially for the Archaic, Classical, and Hellenistic period) onto necropoleis, wall circuits, and cult places. Similar conceptual biases occur at the local scale as well – in the ‘judgmental’ choice of the study area for example, based on landscape physiography and on the character of the known archaeological record. Many surveys are targeted at exploring the ‘catchment’ of some known highly monumental site – a town, hillfort, or large villa – and therefore the results of such surveys are biased in favour of central places and inadvertently stress the the role of settlement continuity and hierarchisation in the landscape.
Another form of conceptual bias occurs during classification – the process of sorting our observations into sets of mutually exclusive classes both in the field and during finds processing. Among the most important high-level classes we use in the field are those of 'site' and 'non-site'. An extensive literature exists on the theory of the siteless survey and why we should not employ the concept of 'site' in the field (Dunnell & Dancey 1983; Ebert 1992); in a very practical sense the continued definition of 'sites' in the field (as opposed to during the interpretation) may cause more problems than it solves. If we increase the 'intensity' of our survey once we have defined a 'site', not only does the density (numbers per unit area) of our samples increase, but so does their diversity (number of distinct material categories). Whereas the quantitative effect has been noted by many survey practitioners in the past, the more subtle effect on diversity has escaped general notice. For example, Small (1998:349) noted that the most significant concentrations of paleolithic flint in his Basentello surveys occur at 'gridded' sites, where the likelihood of recovery would be much higher; he did not note whether the diversity of flint types within those concentrations increased as well. It is to be hoped that the recent publication of re-analyses of the Boeotia survey data will kindle more interest in the subject (Gillings & Sbonias 1999:45; Bintliff & Howard 2000; see also below in section 2.1).

A second important set of conceptual biases comes into play during the collection, recording, and processing of individual finds. The distinction of 'diagnostic' and 'indeterminate' finds gives rise to a tendency to concentrate collection and recording practices on those parts of the surface archaeological record which are amenable to functional or chronological interpretation. Many past and current survey methodologies have relied to a greater or lesser extent on the walkers' selection, and the ceramicists' subsequent preferential processing, of 'diagnostic' sherds. Besides the admittedly practical reasons – a wish to limit the volume of finds to be processed - even the value of collecting non-diagnostic types of material is sometimes doubted. Thus Yntema, in his introduction to the diagnostic pottery types of the Brindisi area, writes “A field survey with a very limited quantity of diagnostic wares or without expert knowledge of regionally current ceramics of the periods on which the survey centres offers no sound basis for regional studies and is to be considered a waste of time and money” (Yntema 1993:29).

In the RPC survey areas in central and southern Italy the concentration on diagnostic wares and forms leads to biases in favour of the better preserved and better studied pottery types, which are chronologically tied to dated sequences outside the regions themselves (sometimes even outside Italy). This introduces the further danger of assigning too strict a dating range to particular types of ceramics (see Bintliff & Howard 2000). In this, survey ceramic specialists are persevere dependent on similar practises in excavation. Malone and Stoddart (2000:95-6) note that the tendency to date excavated features from lower frequency diagnostic elements (which are rare in surface deposits) is especially marked in the later prehistoric metal producing periods. Since the progressive degradation of prehistoric pottery leads to a relative dearth of diagnostic sherds anyway (Malone & Stoddart 2000: fig 11.1), and the recognisability of ceramics varies wildly over time with low recognisability occurring at phases of important reorganisation within societies (Malone & Stoddart 2000: fig 11.2), our reliance on diagnostic material indeed introduces strong biases.

All non-total collection strategies seem to me to contain built-in conceptual biases of this type because they require the field team, or even the individual walker, to make decisions about which finds are 'diagnostic' and 'indeterminate'. Under typical survey conditions, not even a team consisting of period specialists could reliably make such decisions.

Visibility biases are research biases caused by regional and local variations in the visibility of the archaeological record. The term 'visibility' has been used to describe a number of things, but in the current context I will equate it to the retrieval rate - the probability that an artefact lying within a walker's transect will be recorded. On the widest scales this type of bias is obviously mediated by land use and land cover, which circumscribe the outcomes of different types of regional research. Examples of this are the peculiar sets of sites generated by aerial reconnaissance in the Wroxeter Hinterland (LULC bias, see chapter 14) and by 'topographic' surveying in the Pontine Region, Etruria, and more generally in the Mediterranean (see chapter 13). On the scale of an individual survey, visibility bias is mediated by local variations in a large number of factors, of which traditionally the amount of vegetation covering the
soil surface has been seen as the most important one. As I will argue in more detail below, this is a simplistic view and the understanding of local visibility biases will require much more research effort.

Observer biases, as distinct from visibility biases, are concerned with the ability of the observer to record information which is available in principle. Again, examples can be adduced for both the regional and the local scale of enquiry. In taking aerial photographs, and again in mapping from aerial photography, researchers have consistently selected interpretable shapes such as circles, squares, and linear features. In surveys, many observers have noted the different abilities of individual walkers to ‘see’ different artefact types such as flint, but opinions differ as to how significant such biases are. My own opinion, which I feel confident would be backed-up by a survey of the relevant psychometric literature, is that the differences between observers are much larger than most students of the problem realise, and extend to such areas as general ability to recognise shapes and colours, and to concentrate for any length of time.

Besides individual differences, there are also generalised observer effects for which many examples can be adduced from the RPC surveys. The height of the observer is one such effect: the closer one comes to the surface, the more detail (including finds) one sees. ‘Hands and knees’ inspection of limited areas yields finds where normal upright survey does not; it follows that small people (all other things being equal) will find more stuff than tall people. Another effect is distraction, where one or more finds categories may become ‘invisible’ because the eye is distracted by more prominent features of the soil surface. For example, even though prehistoric ceramics may have lain on the surface in parts of the SIBA2000 survey, the presence of overwhelming numbers of similar-coloured and -shaped fragments of ploughed-up conglomerate rock made it impossible to pick up and inspect many fragments in detail; in other survey areas a similar role may be played by large numbers of more ‘attractive’ Hellenistic/Roman or recent ceramics.

Observer biases are treated here as a separate category despite their links with both conceptual and visibility biases, because they highlight weaknesses in one particular link in the whole research process – our ability to make reliable observations – which has been largely ignored so far. It is precisely in high intensity surveys of off-site areas that even minor observer biases can have major consequences on results and, eventually, interpretations.

Summarising all of the above, archaeological research can be biased by conceptual, visibility, and observer biases. Some of these we cannot avoid; others we may not even wish to avoid. But we must study these biases if we want to be aware of the systematic distortions they create in our archaeological record; and if the distortions can be quantified we can attempt to correct for them (see section 2).

2 BIAS MODELLING

Bias modelling is the cartographic modelling, usually with the help of a GIS, of the presence and value of factors influencing the discovery rates of parts of the surface archaeological record, and of the nature of that influence, with the aim of correcting the outcome of primary research results. The following sections detail the stages involved in this (with examples), ending with a discussion of current issues.

2.1 DEALING WITH BIASES

Archaeologists’ responses to the question of how to deal with biases have been three-fold: firstly, to ignore it because the problem is thought to be insignificant or insoluble; secondly, to avoid it; and thirdly, to model and attempt to correct for it. Before detailing my own approach to this issue I will need to say a few brief words about the first two types of response.

In spite of the arguments expounded in section 1 above, some will disagree with either the need or the feasibility of dealing with biases. As noted in section 1.2, while many researchers involved in
surveying in the Mediterranean will agree that there is a vast underestimation of the impact of distorting effects on the quantification of surface material, some believe that the effects of these biases are too complex to be modeled and corrected (Fentress and Ammerman, cited in Cambi & Terrenato 1994:168-71; Terrenato & Ammerman 1996:93-5; Fentress 2000). If taken literally, from this belief it would follow that the current practice of recording bias factors is largely a waste of time. It is difficult to judge the value of this belief without the benefit of having detailed published arguments available, but it may be pointed out we are not seeking to model all the distorting processes (which would indeed be impossible) but only their effects.

Bias modelling is an essential step in the production of locational models in any study that uses biased data sets – and there aren’t any others in archaeology. It is similar in this respect to other historical sciences, which employ source criticism to trace and correct for distortions in their sources. Yet despite the occasional research into methods needed to correct survey data for known biases, current everyday practice seems very little affected.

One can think of both general and specific reasons for this: archaeologists (like other people) are generally overconfident when they estimate the reliability of their results; and having to deal with biases in our data adds to an already high workload. Archaeologists would rather spend their fieldwork time collecting fresh archaeological data than on what they perceive to be mere methodological niceties.

The second response type – to avoid biases altogether – is clearly a valuable approach if it can be realised; but so far I have been able to find only a single example of this. Van de Velde (1996, 2001) claims to have avoided surface visibility bias during the Riu Mannu survey in Sardinia by conducting, alongside a traditional extensive survey for diagnostic materials, a systematic gridded point sample survey, in which two square meters of ground surface were cleared at each sample point ‘to provide quantitative control without visibility effects’. It is not clear that such a method can be extended to any other factors causing visibility bias or other types of bias; and Van de Velde himself acknowledged that geopedological effects cannot be controlled this way. In my concluding section I will return briefly to this issue.

Turning now to the third response type, methodologies for dealing with biases can be described as following three consecutive steps – recognition, recording, and correction. Each of these is described in more detail below, and then discussed in the following sections.

RECOGNISING BIASES

A bias factor must first be recognised as such, and its significance must be assessed in a preliminary manner in order to determine whether it is necessary to continue with the next step. Since the significance and even presence of bias factors varies over geographical space, it is probable that no regionally or supra-regionally valid ‘standard’ set of bias factors can exist. Instead, significant bias factors must be recognised locally and for each material category separately. Excellent examples can be cited from RPC fieldwork experience. Land improvement schemes in the Pontine region had resulted in the localised burial or removal of ancient land surfaces depending on the terrain morphology and soil type; since the Fogliano survey area straddled several such terrain units this had an obvious effect on the outcome of the survey (chapter 17). In the SIBA2000 survey, by contrast, the localised ploughing-up of conglomerate bedrock proved to be the most significant bias factor in many parts of the landscape unit under investigation, and the discovery of dull and undiagnostic prehistoric pottery was most significantly affected (chapter 12: 11). At this stage, surveys may be designed to avoid some biases.

RECORDING BIASES

Once it has been established which bias factors are likely to have a significant effect on the outcome of a regional or local survey, they must be recorded and assessed. These are two separate things though they are often confused. To record the bias factor means to measure or estimate the degree of its presence; to assess it means to measure or estimate the effect of that presence on the retrieval rates of all material categories.
For example, the stoniness of a field may be recorded as a bias factor by estimation using the MOLAS field guide (Spence 1990: chart 11), or by measurement in the field or from photogrammetry, but the actual amount of bias due to this factor is not yet known at that stage - 20% stoniness does not equal 20% less finds. Yet the two are often equated, especially in the case of vegetation cover which is the single most recorded bias factor in Mediterranean surveys. In other words, a linear relation is assumed to exist between the bias factor and the bias itself, whereas in fact this relation has to be assessed either by estimation from field observations (the current, though rarely explicit, practice at almost all current surveys) or by statistical analysis of the correlation between bias factors and retrieval rates, or has to be measured by conducting field experiments.

Since in most cases there will be multiple significant bias factors present, it may be more relevant to evaluate their combined effect on the retrieval of each material category (for example by factor analysis or logistic regression techniques) than to do this for each factor individually. Typically, the end result of this step is one or more biases expressed on a percentage scale from 0 (no bias; perfect retrieval) to 100 (total bias; no retrieval).

The distinction between the bias factor and the bias itself, seemingly of academic interest only, may be seen to be important in the handling of differences between material categories. If the bias is equated to the bias factor, it becomes an objectified, measurable, ‘environmental’ variable rather than one which is relative to the characteristics of each material category - implying that the bias affects all material categories to an equal degree, which is patently untrue.

CORRECTING BIASES

Once a bias factor has been recorded and its effect on the retrieval rates of all material categories assessed, the next logical step is to correct the ‘raw’ retrieval rates for this effect. Nance (1983:350) already advocated the formulation of correction factors in the context of probabilistic sampling. In current practice, corrections to raw survey data are typically applied by multiplying the counts or weights of retrieved material categories by the inverse of the relevant bias percentage. A good example of this is Gaffney’s (Gaffney et al. 1991) correction of raw survey data for surface visibility percentage in the Hvar project. However, there are problems with this method at both the ‘high’ and the ‘low’ ends of the distribution of densities per collection unit, albeit for different reasons. At low retrieval rates, quantitative correction for biases ignores the problem of statistical diversity and enhances the effects of statistical noise: multiplying up low retrieval rates will not increase the number of distinct types within the assemblage, and insignificant density variations can become significant by multiplication. High retrieval rates typically occur at ‘sites’, where the collection strategy is likely to have been different in various respects from the one normally used, so that the collected assemblages are no longer representative of what lay in the field. Multiplication by bias correction factors would then result in completely unrealistic densities for some material categories.

In other words, current methods will not correct the fact that rare categories will be underrepresented in areas with a low retrieval rate, while changes in the types of samples taken from areas with high retrieval rates will tend to lead to overrepresentation of rare and ‘diagnostic’ categories, and underrepresentation of ‘indeterminate’ or ‘uninteresting’ categories. As will be suggested below, the problem of low diversity may be countered, at the expense of spatial resolution, by merging neighbouring collection units with low retrieval rates until sufficiently large assemblages have been created. The ‘high end’ problem can be avoided only by a rigorous separation of the ‘standard’ samples taken from every collection unit, from the ‘special’ samples taken once a ‘site’ has been defined.

2.2 IDENTIFICATION AND ASSESSMENT

IDENTIFICATION

One of the effects of the ‘New Archaeology’ has been that, beginning in the late 1970s, many authors have identified and discussed the impact of factors that bias our knowledge of the archaeological
record. This is not the place to reproduce long lists of such factors; I will concentrate here on those factors that are most pertinent to the discussion.

The three factors which, since 1980, have attracted most attention both in the surveying literature and in the literature on GIS and spatial analysis, are geomorphological processes such as erosion and deposition (eg, Vermeulen & De Dapper 2000; Allen 1991), land use and land cover (eg, Van Leusen 1996), and sampling and surveying technique itself (Shennan 1985, Ammerman & Terrenato 1995, Verhoeven 1986).

Geo(pedo-)logical research is increasingly seen to be a requisite part of regional project designs, not just to map one of the most important factors in determining past land use, but also to map bias in the survey results caused by natural geomorphological and human processes. While geomorphological processes work over relatively long time-scales, RPC fieldwork experience in the Pontine and Salento Isthmus regions confirms that many landscapes have been seriously affected by recent or sub-recent anthropogenic soil movement and restructuring (see chapters 11 and 17).

The fact that the effects of bias factors vary by region, by period and by material category has been widely recognised in the last decades, but has so far not led to a systematic approach to the treatment of biases in Mediterranean archaeology. A major reason for this may be the difference in climates and sedimentary regimes as compared with north-western Europe. Whereas Mediterranean archaeologists have an extremely rich and, in the arid climate, well-preserved surface archaeological record to study, the situation in many parts of north-western Europe is radically different, with sedimentation burying many sections of the archaeological records and wetter climates giving rise to extensive areas of grasslands.

It is no wonder then, that north-western archaeologists were forced relatively early on to consider bias factors and how they affect different site types, periods, and regions. Among the published studies of regional landscape archaeology in northern and western European archaeology, Fokkens’ (1991, 1998) thesis on the settlement history of the north-western Netherlands stands out for its clear and systematic approach. His regional, site-based study of ‘map formation processes’ stands out as a model still unrivalled by later GIS-based studies. Fokkens, using a somewhat different bias typology from mine, recognises post-depositional bias factors and research bias factors. Within the scope of his study, soil types, land use, and historical land reclamation activity are identified as significant for the former, while the localised activity of amateur archaeologists is identified as the main significant factor for the latter.

The fact that the Mediterranean surface archaeological record appeared to be generally much more complete than its north-western European counterpart has caused Mediterranean archaeologists to take a different approach to biases, stressing one factor in particular that has a direct, observable, effect on retrieval rates: land use / land cover (LULC). Cambi and Terrenato (1994:151-2) report that, following the then recent recognition of the significance of biases by Italian archaeology, new volumes in the Forma Italiae series of regional site-based surveys have begun to include mapping of non-visible areas. As an advocate of the ‘Anglo-Saxon’ approach, Terrenato himself, in the Cecina survey, considered zones of geological deposition and zones where land cover largely impeded survey to be the most significant bias factors (Terrenato & Ammerman 1996). On the local scale, because surveys are targeted towards areas of optimal land use/land cover to begin with, many practitioners continue to equate ‘visibility’ with the percentage of ground surface not covered by vegetation.

As I have argued above, the tendency to think of the surface archaeological record as something which is there to be observed, and will in fact be reliably observed by any qualified observer unless blocked by intervening sediment or vegetation, has one further consequence for the identification and assessment of bias factors in Mediterranean archaeology - the low retrieval rates of material categories such as flint are regarded as a consequence of choices made during survey design (in this example, whether a lithic specialist would be included in the field teams, or not), rather than as a permanent, built-in feature of any type of archaeological fieldwork. While it is to be hoped that a specialist will attain a higher retrieval rate for his or her particular material category, it does not follow that his or her assemblages are unbiased.
Within Italy, the assessment of the significance of bias factors relative to material categories attracted early interest through the evaluation of the South Etruria project data. At the time, Di Gennaro and Stoddart (1982), already noticed the low visibility of prehistoric and medieval ceramics; later on, Malone and Stoddart (2000, echoed in Kuna 2000:34-6) went on to provide several quantitative arguments to explain this phenomenon in terms of formation, taphonomic, and recovery processes.

These authors begin their argument by pointing out that the relative abundances of prehistoric ceramics of different periods may be culture-specific, that is, related to a cultural tendency to dig pits (thus preserving archaeological materials sufficiently long for their discovery in the present). Kuna calculates that a prehistoric residential farm in his Bohemian study area may yield only a few ceramic fragments in a survey at 10% coverage. Although such calculations provide support for an assumption I made for the interpretation of the Fogliano survey data (Attema et al. 2001), they can only explain why prehistoric material is underrepresented in the ploughsoil (formation processes) rather than why archaeologists have trouble recovering a representative sample of the material which is in the ploughsoil. However, Kuna goes on to demonstrate that the probability of occurrence of a distinctive (diagnostic) fragment within such a sample varies by period, and decreases with elapsed time. And as we have seen in the discussion of conceptual biases above (section 1.2.3) the preferential treatment accorded to diagnostic finds implies that both the variable ratio of diagnostic fragments and its general decline through time constitute significant bias factors for some material categories.

RPC field survey experience suggests that even this does not come close to identifying all the relevant biases in regional and local archaeological data sets, as the following brief discussion will show.

At the regional and supra-regional scale, the site-based records compiled by desktop studies or in ‘topographic’ survey can be shown to be influenced by differences in accessibility of the terrain. Before about 1970, the probability of archaeological material being observed, recognised, and reported was largely due to chance and the activities of local amateur archaeologists. Field observations could only be made in accessible areas, and the more accessible an area the more likely that it would be visited by a person able to recognise archaeological features; ‘chance’ finds are therefore more likely to occur nearby modern infrastructure than away from it. Surveys conducted in the ‘topographic’ tradition (which requires a very large area to be surveyed by one or a few persons at most) continue this trend, as is shown clearly by the Forma Italiae volumes for the Pontine region which were produced between 1920 and 1970 (cf. the discussion in chapter 13).

At both the regional and local scales, the significance of the influence of subsequent human occupation and land use on earlier remains appears to be severely understudied by Mediterranean archaeologists. Authors such as Shennan (1985), Verhoeven (1991), Van de Velde (1996) and Burgers and Yntema (1996) do not mention this factor at all, although it is in operation on any but the most short-lived site and can cause bias both because earlier periods tend to be ‘lost’ among remains of a later date, and because the greater research intensity at sites of a later date can lead to the discovery of low density remains of an earlier period. Both effects have a bearing on the apparent amount of settlement continuity in a study region, as is shown by Attema’s (1993, forthcoming; see also Van Leusen 1998) study of a possible Roman Republican ‘villa colonisation’ in the foothills of the Lepine mountains.

My third and final example concerns a previously unidentified bias in the recording of the surface area of collection units and the subsequent calculation of finds densities on that basis. Typically, the raw counts and weights per collection unit resulting from a survey will be ‘normalised’ to account for any differences in the size and coverage rate of the collection units. It was found during analysis of the Ostuni survey data (Attema et al 2001) that the digitised areas of many collection units were approximately 10% smaller than those mapped in the field on topographic maps at scale 1:10,000. The difference is explained by the fact that topographic maps generalise and omit features such as paths, wayside berms and scrub, and buildings; however these features were included in the detailed hand-drawn field maps made of each

---

1 Fokkens (1998:59) does mention this factor.
collection unit which formed the basis for the digital map. In the relatively well-mapped Ostuni survey zone the effect remains fairly small, but in the topographically complex and poorly mapped terrain of the SIBA2000 survey the areas of some agricultural fields were overestimated by as much as 50%, because the boundaries of ploughed fields were ‘stretched’ to include the surrounding unmapped scrubland as well. The variable percentage by which the areas of collection units are overestimated indicates how large the potential errors can be when ‘normalisation’ of survey results takes place. This is therefore a potentially significant bias, especially where it is intended to detect and interpret relatively small variations in finds densities across the landscape.

Assessing a bias factor means asking the question: Does this factor have a significant effect on the types or amounts of information I can retrieve from the archaeological record? A simple way to begin answering this question is by plotting the two against each other in a graph. Examples of such graphs were produced as far back as 1978, when Fred Plog and his colleagues plotted the number of person/days per square km surveyed against the number of sites discovered for 12 surveys in the American southwest (see figure 2; Plog et al. 1978, 389-94 and fig 10.1). Cherry (1983, fig 1) did the same for a large number of survey projects in Greece, adding a historical perspective by grouping the projects into broad types as practised since about 1950. In figure 1 I have extended his plot for Italy, including both extensive regional and intensive local surveys for the Pontine Region. Although Cherry used surveyed area, rather than the more precise effort in person/days, as his measure of research intensity, both studies clearly show that more intensive research leads to a higher retrieval rate (in this case, of sites). While Cherry’s historical perspective uncovers the incredible increase in site density achieved since the earliest surveys (10-fold even in the last 30 years), Plog’s diagram highlights the fact that there is an almost linear relation between intensity and result – for every duplication of research intensity, an approximate 75% rise in the number of sites may be expected.

Figure 1 – Site productivity in relation to surveyed area for different survey types in Greece and Italy (adapted from Cherry 1983, fig 1). Note that the scale on both axes is logarithmic.
On a regional and supra-regional scale, differences in research intensity have been shown to play an important part in creating spatial variation in site density. As noted already by Nance (1983:311), ‘much of the variation apparent in data from survey to survey may be explained by differences in survey intensity’. Moving to more local scales, examples of the assessment of bias factors have been published since about 1990 and have concentrated on the relation of ‘visibility’ (however defined) and site density. Again, Cherry, in his analysis of the results of the Keos survey (Cherry et al 1991, fig 3.6), appears to have been the first to publish a plot of visibility versus site density (my figure 3), apparently confirming that vegetation cover is inversely related to site discovery rate. However, such plots presuppose the detailed recording of the bias factor, which is the subject of the next section, so I will discuss them there.

Four aspects of the examples discussed above are noteworthy: firstly, the relation between research intensity and site density cannot remain linear - there must be a point of diminishing returns, where increasing the research intensity will no longer result in the discovery of more sites. It is likely that modern high-intensity surveys, with crew spacings of less than 10m, have reached this point. Secondly, measuring research intensity is not trivial. Crew spacing and walking speed are the two most obvious factors involved, but in almost all cases only the former is recorded. A more reliable proxy measure of research intensity is therefore likely to be the ratio of the area covered and number of person/days spent, expressed as either person-days per unit area or area per person/day. Thirdly, measuring the ‘result’ of a survey by counting the number of sites per surveyed unit area becomes increasingly meaningless as the definition of ‘site’ has begun to shift and blur, including smaller and less dense scatters, and is even ceasing to be used as a unit of discovery. Fourth and lastly, increased research intensity leads not only to the discovery of more sites, as the examples above have shown, but also to the discovery of different types of sites: the ones that are smaller and less visible. In my intra-regional comparison of survey data sets of the Pontine region (chapter 13 section 2) all of these aspects are shown to play a role in confounding successful quantitative comparisons.

In view of the above, our assessment of research intensity as a bias factor at the regional and supra-regional scale must be, that differences in research intensity will not only result in significant spatial variation in site densities, but also in significant spatial variation in distributions of site types. This assessment can help us avoid misinterpreting variations caused by the bias factor, with obvious consequences for settlement chronology and landscape history; but only if it is properly recorded and if we succeed in quantifying its effects.
2.3 RECORDING AND EVALUATION

Whilst bias factors such as vegetation cover can be recorded, their effect on the retrieval rate (i.e., the bias itself) cannot; it has to be evaluated. This section deals with methods for doing both, again with examples ranging from the supra-regional to the local scale.

RECORDING

Current approaches to the recording of bias factors in the field are based on, and in fact almost identical to, methods first developed for Mediterranean surveys in the early 1980s. In the Pontine Region, approaches to the recording of visibility and observer factors developed for the Agro Pontino Survey project (Voorrips et al. 1991: 82 ff.) were adopted by subsequent Dutch survey projects directed by Attema, and in turn formed the starting point for the recording experiments carried out during the RPC project surveys in the Pontine Region and elsewhere.

The most basic method imaginable for recording any bias factor is the ‘binary’ approach, in which each factor can assume only one of two values – present / absent. An example of this is the recording by Terrenato and Ammerman (1996) of visibility factors during their survey of the Cecina valley, during which conditions of geopedology and vegetation were classed as favorable or unfavorable (see figure 4a). While such an approach can be valid for some factors, it does tend to simplify the reality of field situations, and most other researchers have therefore adopted various simple ordinal rating scales to record bias factors. Where multiple bias factors have been identified, each is usually recorded separately on an ordinal scale. In the RPC surveys, five factors are recorded in this manner: vegetation cover, stoniness, sun/shade, soil weathering, and ploughing conditions. Overall visibility is independently recorded as a sixth variable. In the Fogliano survey (Attema et al. 2001) overall visibility was recorded as two related ordinal variables: Field visibility and Block visibility. The former, which rated the overall visibility of an agricultural field from low through normal and high to optimal, had to be corrected using the relative visibility rating recorded for each collection unit as low, normal, or high. Low relative visibility, for example, resulted in ‘demoting’ a particular collection unit to a lower overall visibility class.

The rating of bias factors by estimation along an ordinal scale of measurement is to some (large but unknown?) extent subjective, so that it is not possible to compare these across projects. It may be possible to improve ratings using a method called CARS (criterion anchored rating scale) which anchors such ratings to specific criteria (Suenson-Taylor et al 1999). A rating will then be arrived at by assessing these criteria much like grain size is arrived at by comparison with standard grain sizes during the recording of soil properties. This method is claimed to give a proper numeric scale along which to...
measure the factor, but has not been applied to field archaeology yet as far as I am aware.

**EVALUATION**

The effect that a specific bias factor has on the retrieval rate of a specific material category can be evaluated by estimation, measurement, or statistical analysis. Measurement requires the setting up of controlled field experiments in which known densities of the material are surveyed under different conditions. Statistical analysis of the influence of bias factors on retrieval rates has only been pursued by relatively few researchers, among whom Shennan (1985) still takes pride of place. Shennan used techniques of logistical regression in order to study the impact of a large number of factors on the results of his experimental survey area in Hampshire; identical methods were later applied to the results of the Agro Pontino Survey project by Verhoeven (1991).

However, the large majority of past and current practitioners of field survey employs simple, ‘intuitive’ methods for evaluating the effect of bias factors. Because there is no formal method underlying this step, it receives little or no discussion in the literature. Examples from the RPC field surveys include the estimation of the overall percentage visibility effect for each of the visibility classes distinguished in the Fogliano and Ostuni surveys (see chapter 11, page 3). Often the rating for the bias factor is taken to be the rating for the effect itself, as in the Boeotia survey where the ground surface visibility score is derived directly from the 10 point rating scale used to rate land use/land cover (Gillings & Sbonias 1999:36). At a larger scale Fokkens (1998:64-5), too, based his calculations of archaeological visibility indices on estimations of the bias caused by each of his distorting factors, with 100% indicating optimal recovery conditions for each.

The statistical approach to evaluating bias factors initially seemed to hold promise of a much more precise and replicable method. Shennan (1985:38-9) concluded that, although not entirely negligible, distorting factors did not seem to have a major impact on the retrieval rates of his three most frequently occurring material categories (explaining some 17-18% of variation in them). Certainly their effect was less than that of environmental variables, which he calculated to be on the order of 40-60%. Verhoeven (1991:87) found that the effects of weather and field conditions on the results of the Agro Pontino Project surveys had a similarly limited effect but that ploughsoil conditions (dust/rain) do have a strong influence on the recovery of flint and obsidian. Going on to consider less frequent material categories, Shennan found that light conditions did play an important role (34% of variation explained) in the recording of Romano-British pottery.

Going one step further and using his detailed records of who surveyed where, Shennan (1985:40-44) statistically removed the visibility biases he identified, and went on to look for observer bias in the residuals. His analysis confirmed that observer bias, though fairly minor, was definitely present. For chipped stone, his analysis showed a small (3%) but highly significant observer bias; for burnt flint and post-medieval pottery the variability of the walkers was somewhat greater (explaining 9-10% of the variation in retrieval rates). Shennan also noted that the differences between walkers in picking up various materials were quite large.

It is unfortunate that, after these initial attempts at formal evaluation of bias effects, no further research was published, and practitioners reverted to estimation methods. Two new twists to this approach were added in the 1990s by, firstly, plotting the effect of bias factors against retrieval rates and, secondly, by employing simulation studies. Figures 3 and 4 present two examples of site density plotted against

---

4 This confirms my observation during the Fogliano survey – observer bias and differential visibility differ by finds category. The fact that Shennan chose to study his most frequently occurring find categories already entails less distortion!

5 Van de Velde (2001) notes that the method used by Shennan and Verhoeven to calculate ‘field effects’ underestimates the true effect because it disregards ‘zero’ observations.
visibility percentage for the Keos and Cecina surveys. Terrenato (forthcoming) also investigated the effects of surface visibility (geopedology and vegetation) on the recovery of archaeological sites through simulated survey coverages.

To conclude: what seems to be lacking in the ‘estimation’ approach to recording and evaluation of bias factors, is the presence of a testing phase in which the accuracy and replicability of both the recording method and of the evaluation ‘rule’ can be assessed.

2.4 CORRECTION

Although bias factors are being recorded, the apparent failure of formal evaluation methods has meant that many researchers used this information only for informal adjustment of the raw survey results. Even simple map overlays of the bias factors onto the dataset being studied can be very effective in this respect, and neither Fokkens (1998) nor Attema (pers. comm.) went on to apply formal correction methods to their data. Three quantitative correction factors, however, are currently being applied in many Mediterranean surveys because of their computational simplicity – unit area, coverage, and overall visibility; these will be discussed in the following sections.

What constitutes ‘correction’? Following Shennan, Van de Velde (1996:23) and Verhoeven (1991), believe that observer effects such as the well-known ‘flint expert’ phenomenon can be corrected for by rotating individual walkers during a field survey. I believe this to be incorrect; while rotating the walkers would certainly randomise such effects, it does not eliminate them. For example, where-ever a colour blind observer is put in a field walking line-up, the transects walked by that person will turn up less small sites.

UNIT AREA

Correction for unit area is needed because (ceteris paribus) a larger area will yield proportionally more finds. This requires the unit area to be measured with a specific precision and accuracy, relating to the desired degree of confidence that a specific density difference is indeed a significant archaeological variation rather than a result of random errors in the measured unit area. As we have seen above, such measurements can indeed be surprisingly inaccurate, and require the adoption of more stringent field methods (see chapter 7). The correction itself transforms finds counts per unit into finds densities, expressed as number of finds per hectare within the RPC project.

COVERAGE

The single variable which has been successfully corrected for in surveys (in regional site-based records it was never recorded in the past), is the coverage (percentage of the collection unit actually observed). This correction method has even been applied in the field, in cases where the distance between walkers was adjusted whenever changes in vegetation cover were encountered. Van de Velde (2001) applied this relatively simple correction based on the % coverage of his survey to arrive at a corrected total number of sites for one of his survey transects, and reports that this outcome agrees reasonably well with the number of sites independently discovered by local amateurs within the same area. Such observations are the first step toward proving that retrieval rates corrected by formal methods are less biased, and form a better basis for interpretation.

However, coverage is itself composed of two variables: the average distance between walkers and the swath width. While the former is generally recorded, the latter is set to a notional value (usually 2 m, but values from 1 to 5 m have also circulated). Seemingly small differences in the swath width (eg, 1.5 vs 2 m) represent a large percentage change (25-33%) which, if used in a correction formula, will make comparison between surveys difficult. The correction usually takes place by multiplying the original finds count per unit by the inverse of the percent coverage for that unit, so that the new finds count represents what would have been recovered had the unit been fully covered.
OVERALL VISIBILITY

The rationale for correcting for visibility is, as we have seen, that the probability of recovering a surface artefact decrease as the surface itself becomes less visible. Correction takes place by multiplying the original finds count with the inverse of the visibility score. Gillings and Sbonias (1999:36), for example, used the 10-point ground surface visibility score recorded by the Boeotia survey to correct raw artefact counts using the formula

\[ \text{Corrected} = \text{recorded} \times \left( \frac{10}{\text{visibility score}} \right) \]

This, in effect, represents a conversion of the visibility score to a visibility coefficient, which is then used to multiply raw artefact counts in areas of low visibility. Similar methods were used in the analysis of the RPC surveys near Ostuni and Francavilla Marittima (chapters 11 and 12).

THE PROBLEM OF LOW DENSITIES

Modern field surveys, with their increasingly detailed and intensive study of ‘off-site’ areas, tend to result in large numbers of collection units containing very few finds. In addition, these finds tend to be less different from each other – in statistical terms, the diversity of the assemblage is lower. This gives rise to several interpretative problems. Firstly, the multiplication of finds counts in the course of bias correction, as suggested in the previous paragraph, will increase the numbers of finds but not their diversity, thereby causing a relative drop in variability. A method for avoiding this, namely the aggregating of collection units until a given minimum number of finds is obtained, was suggested as early as 1973 (Dunnell & Dancey 1983:272), but it remains to be seen how the loss of spatial resolution affects subsequent analysis.

A second implication is that low finds densities can no longer automatically be dismissed as insignificant ‘noise’ or ‘off-site material’. Instead, the survey intensity and visibility biases should be taken into account. Higher survey intensity equates to a higher diversity of the recovered finds assemblage; for example, Gillings & Sbonias’ (1999:46-53) detailed chronological discussion of a single site assemblage of the Boeotia survey ignores the fact that the site itself was much more intensively surveyed than its surroundings, and low density unobtrusive find types present on the site could just as well have been present outside the site grid. Likewise, re-surveying experiments at Fogliano have confirmed that a single sherd recovered under circumstances of very low visibility can develop into a scatter if the survey takes place under more favourable circumstances. As it is recognised that even intensive modern surveys cannot hope to collect representative samples of low density, low visibility categories such as the (possibly) prehistoric impasto occurring in the Sibaritide foothills (chapter 12), dedicated specialist (re-)surveys are required for such materials.

Beyond the purely quantitative aspects of interpreting low density ceramics, there are also some less easily categorised aspects. Low densities can, for example, be an artefact of the finds classification process, in that material from some periods can only be recognised if diagnostic forms or decorations are present. If these are rare or absent, finds will be classified into broad undiagnostic categories, or even as ‘indeterminate’. Finally, the belief that material can be transported very far from its origin by slope processes or ploughing, prevalent among archaeologists of the Roman school and invoked regularly to explain away the occurrence of low-density ceramic scatters near urban sites, seems unwarranted in many cases. Bintliff’s recent work on re-interpretation of the extra-urban ‘blankets’ of surface finds from the Boeotia survey is confronting this issue head-on, and is attempting to distinguish between several alternative causative processes for low-density distributions.⁶

---

⁶ Rather than posing a problem, Bintliff and Howard (2000) note that low densities, in combination with certain diagnostic types, may indicate the presence of a Hellenistic cemetery.
CONCLUDING DISCUSSION

Our ability to record surface archaeological material is not perfect; it is biased by visibility and research biases. An example of the former is current and historical land use / land cover (LULC; see especially chapters 14 and 17); examples of the latter are the recording and classification methods used (chapters 13 and 16). Of the two courses of action that have been proposed to remedy the problem, that of avoiding biases seems to have little if any potential beyond Van de Velde’s experiments with surface cleaning. Recording and attempting to correct for biases is the only alternative, but before this can be done with any degree of confidence in the results, dedicated studies will need to underpin the methodology. Data on research and visibility biases have so far not been systematically collected at regional scales, although suitable methods have been developed. Terrenato (1996: 227 –228) underlined the importance of recording bias factors if we are to attempt ‘the correction, at least partially, of incomplete distributions’, and advocated ‘a series of methodological experiments dealing with the various aspects of how to document surface scatters’, e.g. by replication studies, all within a local (regional) context. This research was then, and still is now, ‘a high priority for survey practitioners in Central Italy’. In fact such experiments should probably have precedence over the collection of bias data, because we do not yet understand how best to do the latter.

The issue of bias is almost universally agreed to be an important one, but neither the intensive interest and study conducted in the early 1980s, nor the constant popularity of surveying or, latterly, the use of GIS, have so far led to anything resembling a concerted effort to develop a methodology which is valid across projects. The plethora of biases discussed in this article requires that we return briefly to the fundamental question raised most recently by Fentress: Could it be that we will not be able to disentangle the mess, and correct the distortions? I hope I have been successful in arguing that many lines of research remain to be explored before giving in to such a counsel of despair.

CASE STUDIES

The case studies presented in chapters 13, 14, and 17 were conducted to demonstrate a) the relevance of bias factors to the interpretation of survey data and of landscape archaeological data in general; and b) methods by which bias factors can be included in geographic models of archaeological landscapes. At the regional scale, studies of the data collected by the Wroxeter Hinterland Project and the Agro Pontino Project (Voorrips et al. 1991) demonstrate this for systematically surveyed data and general archaeological records; at the scale of a ‘local’ survey such as the Ninfa and Fogliano surveys conducted in 1998-9, case studies demonstrate this for specific visibility and research biases. In the WHP surveys, conducted in 1994-6, the choice of fields was limited by modern land use and land cover (LULC), in particular the availability of freshly ploughed surfaces. Since these are not randomly distributed over the landscape – relief, distance to Severn, soil type and hydrology all play a role – they result in the taking of a biased archaeological sample. In the surveys conducted by the Agro Pontino project, paleosurfaces dating to the paleolithic period had been covered in some parts of the Pontine plain by more recent alluvial and colluvial deposits, and similar though less clearly evident biases must have been present for material dating to later periods.

REFERENCES

Alcock, SE, JF Cherry & JL Davis 1994

Allen, MJ 1991

Ammerman, AJ 1985

Attema, PAJ (ed) 1995

Attema, P, M van Leusen & E van Joolen in press

Attema, P, G-J Burgers & M van Leusen forthcoming

Baxter, MJ 1994

Bintliff, J 1992

Bintliff, J & Ph Howard 2000

Bintliff, J & A Snodgrass 1988

Bintliff, J, Ph Howard & A Snodgrass 1999

Blankholm, HP 1991

Bradley, R, T Durden & N Spencer 1994

Burrough, PA and RA McDonnell 1998
*Principles of geographical information systems*. Oxford: Oxford University Press.

Cherry, JF, JL Davis & E Mantzourani (eds) 1991
*Landscape Archaeology As Long-Term History. Northern Keos in the Cycladic Islands*. Los Angeles:??.

Cressie, NAC 1993


Daniels, SGH 1972

Di Gennaro, F & S Stoddart 1982

Dunnell, RC & WS Dancey 1983

Ebert, JI 1992
Distributional Archaeology. Albuquerque: University of New Mexico Press.

Feiken, H & M Van Leusen in press

Fentress, E 2000
What are we counting for? In Francovich & Patterson (eds) 2000:44-52.

Fokkens, H 1998

Francovich, R & H Patterson (eds) 2000

Gaffney, VL, J Bintliff & B Slapsak 1991
Gaffney, V & PM van Leusen 1995

Gillings, M & K Sbonias 1999

Goovaerts, P 1997

Hamond, FW 1978

Hamond, FW 1980

Isaaks, EH & RM Srivastava 1989

Kuna, M 2000

Leusen, PM van 1993

Leusen, PM van 1996
Unbiasing the Archaeological Record, Archeologia e Calcolatori 6:129-136.

Leusen, PM van 1998

Leusen, PM van & VL Gaffney 1996

Livy

Malone, C & S Stoddart (eds) 1994
Territory, time and state. The archaeological development of the Gubbio Basin. Cambridge: Cambridge UP.

Malone, C & S Stoddart 2000
The current state of prehistoric ceramic studies in Mediterranean survey, in Francovich, R & Patterson (eds) 2000: 95-120.

Nance, JD 1983

O’Brien, L 1992

Orton, C 2000

Plog, S & M Hegmon 1993
The sample size-richness relation: the relevance of research questions, sampling strategies and behavioral variation, American Antiquity 58(3):489-96.

Plog, S, F Plog & W Wait 1978

Rossignol, J & L Wandsworth (eds) 1992

Shennan, S 1985
Experiments in the Collection and Analysis of Archaeological Survey Data: The East Hampshire Survey. Sheffield: Department of Archaeology and Prehistory.

Shennan, S 1988

Siegel, S & NJ Castellan 1988

Small, A et al. 1998

Spence, C 1990

Suenson-Taylor, K, D Sully & C Orton 1999

Terrenato, N & AJ Ammerman 1995

Terrenato, N 2000

Terrenato, N 1996

Velde, P van de 1996

Velde, P van de 2001

Verhoeven, AAA 1991

Vermeulen, F & M De Dapper (eds) 2000

Yntema, D 1993