

https://dspace.org/

Investigation of Inter-analyst Variability in Stone-walled Structure Classification

Hunt, Tamsin

2013-07-01
http://hdl.handle.net/10539/12795

Downloaded from a DSpace repository. https://dspace.org/

Investigation of Inter-analyst Variability in Stone-walled Structure Classification



School of Geography, Archaeology and Environmental Studies

HONOURS RESEARCH PROJECT

Tamsin Helen Hunt

365237

Supervised by Professor Karim Sadr

Declaration

I declare that this research project is my own, unaided work. It is being submitted in partial fulfilment for the Degree of Bachelor of Science (Honours), at the University of Witwatersrand, Johannesburg. It has not been submitted for any degree or examination in any other university.

_____

(Signature of candidate)

Date of Submission: 24 May 2013

Abstract

For many years stone-walled structures have been classified into different groups using aerial photography. The development of technology such as Google Earth and Geographic Information Systems has resulted in an increase in such classifications. What must be considered, however, is to what extent the results differ depending on the person analysing the structures. A study has been conducted by three different analysts, classifying the stone-walled structures south of the Suikerbosrand Game Reserve. A statistical and visual comparison of the three sets of analyses using Google Earth, QGIS and CrimeStat. These methods showed that variability is obvious between the sets of classifications. It is then important to consider what causes the variability in the classifications and how it can be remedied. This is important as variability in the classifications of stone-walled structures will have an effect on the larger interpretations of the sites and the people affiliated with them.

Acknowledgements

Firstly I would like to thank my supervisor Professor Karim Sadr for his support and encouragement throughout the course of this study. His consistent cheerfulness and patience made the experience of writing my first dissertation a pleasant and enjoyable one.

I would also like to thank Prof. Sadr for spending the time it took to teach me how to use programs such as QGIS, DNR Garmin and CrimeStat.

Thank you to the other two analysts 'K' and 'M' who spent the time locating, outlining and classifying stone-walled structures on Google Earth.

Thank you to my mother, Professor Maureen Coetzee and my friend, Jen Fitchett for their help in editing all my different chapters; particularly in my Results section where Jen's help with some of the statistics was invaluable.

And lastly, thank you to all my friends and family who have supported me and kept me smiling through a very tough year.

Contents

Page number:

Declaration..... ii

Abstract..... iii

Acknowledgements..... iv

List of Figures..... vi

List of Tables..... vii

Chapter 1: Introduction..... 1

1.1 Aims and Objectives..... 1

1.2 Location and Environment..... 2

1.3 Theoretical Background..... 3

Chapter 2: Literature Review..... 5

Chapter 3: Methods and Techniques..... 10

Chapter 4: Results..... 14

4.1 Descriptive Statistics..... 14

4.1.1 Number of Structures..... 14

4.1.2 Levels of Agreement between Analysts..... 15

4.1.3 Area Averages..... 16

4.1.4 Nearest Neighbour Analysis..... 20

4.1.5 Elevation..... 23

4.2 Spatial Statistics..... 24

4.2.1 Mean Centre..... 24

4.2.2 Standard Deviation Ellipse..... 25

4.2.3 Convex Hull..... 29

4.3 Visual Comparisons..... 32

4.3.1 Outline Overlap..... 32

4.3.2 Classification versus Shape..... 33

4.3.3 Large versus Small Structures..... 35

Chapter 5: Discussion and Conclusion..... 37

5.1 Group 1 Structures..... 37

5.2 Group 2 Structures..... 38

5.3 Group 3 Structures..... 39

5.4 Group 4 Structures..... 40

5.5 Unknown..... 41

5.6 General Observations..... 41
5.6.1 Google Earth Imagery..... 41
5.6.2 Training..... 43
5.6.3 Typology..... 43
5.7 Conclusion..... 44
Reference List..... 45
Turnitin Report..... 47

List of Figures

Page number:

Figure 1.1 Google Earth image of the study area, Pam 2..... 2
Figure 3.1 Group 1 according the Taylor's 1979 Masters thesis (see Sadr & Rodier 2012)..... 10
Figure 3.2 Group 2 and Group 3 according to Taylor's 1979 Masters thesis (see Sadr & Rodier 2012)..... 11
Figure 4.1 Graph showing the numbers of structures identified by each analyst..... 14
Figure 4.2 Graph showing the mean areas for each group and analyst..... 19
Figure 4.3 Graph showing the mean altitudes for each group and analyst..... 24
Figure 4.4 QGIS image showing the One Standard Deviation Ellipses for Group 1..... 26
Figure 4.5 QGIS image showing the Standard Deviation Ellipses of Group 2..... 26
Figure 4.6 QGIS image showing the Standard Deviation Ellipses for Group 3 sites..... 27
Figure 4.7 QGIS image showing the Standard Deviation Ellipses for Group 4..... 28
Figure 4.8 QGIS image showing the Standard Deviation Ellipses for the Unknown Group..... 29
Figure 4.9 QGIS image showing each analysts CHULL for Group 1..... 29
Figure 4.10 QGIS image showing each analysts CHull for Group 2..... 30
Figure 4.11 QGIS image showing each analysts CHull for Group 3..... 30
Figure 4.12 QGIS image showing each analysts CHULL for Group 4..... 31
Figure 4.13 QGIS image showing each analysts CHULL for the Unknown Group..... 32
Figure 4.14 QGIS image of each analysts outlines of a Group 1 structure..... 33
Figure 4.15 QGIS image of each analysts pathways..... 34

Figure 4.16 QGIS image of the differing pathways of each analyst…….. 34
Figure 4.17 QGIS image showing the overlaid pathways of all three analysts…….. 35
Figure 5.1 QGIS image showing how each analyst's outlines have shifted due to the use of different GE imagery…….. 42

List of Tables

Page number:

Table 4.1 The number of sites on which the analysts agreed with each other…… 15
Table 4.2 Area data for Group 1…… 17
Table 4.3 Area data for Group 2…… 17
Table 4.4 Area data for Group 3…… 18
Table 4.5 Area data for Group 4…… 19
Table 4.6 Area data for the Unknown Group…….. 20
Table 4.7 Nearest neighbour analyses for Group 1…….. 21
Table 4.8 Nearest neighbour analyses for Group 2…… 21
Table 4.9 Nearest neighbour analyses for Group 3…… 22
Table 4.10 Nearest neighbour analyses for Group 4…….. 22
Table 4.11 Nearest neighbour analyses for the Unknown Group…… 23
Table 4.12 The distance between the Mean Centre points for each analyst…….. 25

Chapter 1

Introduction

1.1. Aims and objectives

Aerial photographs have been used since the 1960's to map stone-wall structures found in southern Africa and it is from these that the different architectural styles were first classified (Mason 1968; Maggs 1976; Sadr & Rodier 2012). These classification groups are believed to represent the different cultures that have inhabited the area at different times through history. The problem being addressed in this study is that of inter-analyst variability in classifying such structures from aerial and satellite images and how much of an effect it has on the results of the interpretation of past settlement patterns

The study consisted of several objectives:

- To survey the study area using Google Earth (GE) imagery to locate and outline the individual stone-walled structures.
- To classify each pre-colonial structure in the area by placing each structure into either group 1, 2, 3 or 4, with a separate group for those that are impossible to classify or are simply unknown. These groups are based on the classification system set out by Taylor (1979) in his Masters thesis that was originally based on stone-wall structures found in the Vredefort Dome. With the exception of group 4
- To carry out statistical analysis of the data after they had been transferred into GIS. This included (amongst others) comparisons of altitude, size and clustering. These are similar to analyses by Sadr and Rodier (2012) who used Google Earth and Geographic Information System software to find and digitally outline stone-wall structures in a nearby polygon similar to my study area south of the Suikerbosrand Nature Reserve near Johannesburg.
- To compare my data with those collected from the same polygon and analysed by two other analysts. This comparison will show the extent of inter-analyst variability and how our results differ.

This study is important as it determines the extent to which researcher bias plays a role in the classification of stone-walled sites that is significant in the interpretation of archaeological data. It is important to take into account how the initial step of data analysis may vary between analysts.

This study will also show whether a study like this can be carried out by a number of analysts after a short training period, or if the analysts need more intensive training. It is here that this study will make a contribution to future capacity building in the field of archaeology.

1.2. Location and Environment

The study area, Pam 2, is located near the Suikerbosrand Game Reserve to the south of Johannesburg and very close to Vereeniging (Figure 1). It is bordered by main roads such as the R54, the R549, the R42 and the R557. This area is about 400 square kilometres that includes +/- 200 stone-walled structures. Pam 2 is situated just north of the Vaal dam and one of its tributaries, the Suikerbosrand River, runs through it creating a river valley surrounded by hilly terrain. Because of this slight valley there is some variation in altitude throughout Pam 2, with the lowest areas situated at around 1450 m above sea level and the highest levels around 1700 m.

Much of the region is characterised by open grassland, with little in the way of wooded terrain which would explain the heavy reliance on stone as a building material rather than wood (Maggs 1976). This kind of terrain also makes the stone-walled structures easy to identify using satellite imagery (Sadr & Rodier 2012). The current average rainfall that the area receives varies between 650 and 700 mm annually (Birss & Green 2007; AGIS 2007). These grasslands therefore represent much agricultural and grazing potential (AGIS 2007).

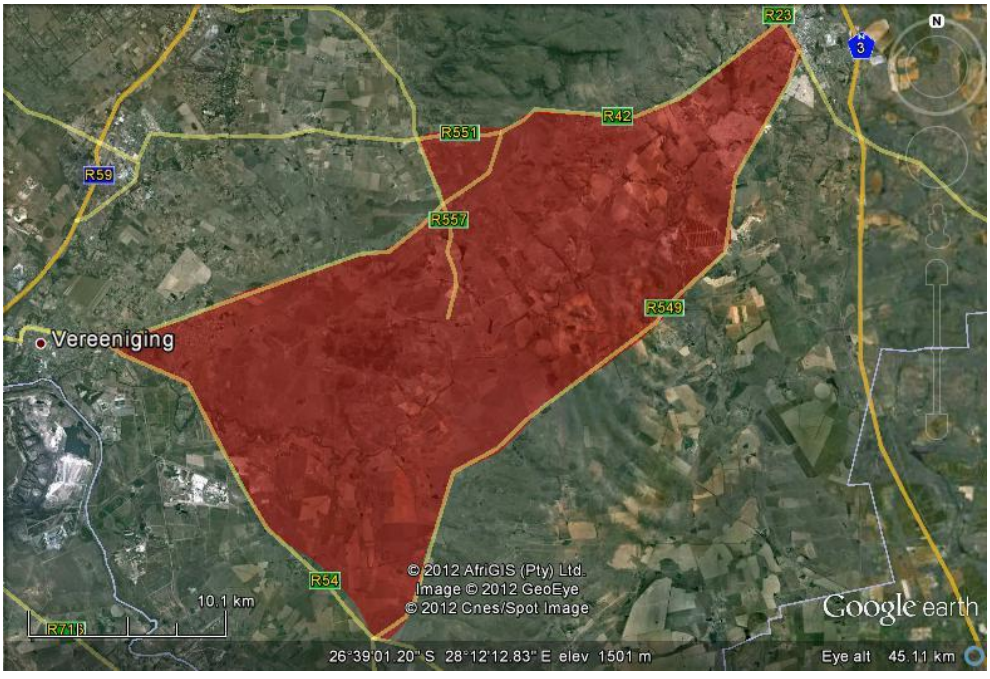


Figure 1.1 Google Earth image of the study area, Pam 2.

1.3. Theoretical Background

Middle-range theory came to the forefront of the social sciences during the rise of 'new archaeology'. It was during this time that it was decided that archaeology's claim to science stood or fell not on statistical methods or systematic thinking but rather on the ability of archaeologists to link the static record (the archaeological material) with the cultural dynamics of past societies (Thomas 2004: 72). During this time, theory was concerned predominantly with theorising how this link was made, and this resulted in the development of middle-range theory.

The assumptions that are made between the observation of the static record and the interpretation, are the 'middle-range' assumptions. It was argued that these assumptions must be clearly stated and justified, and whether they are right or wrong, they need to be made explicit and assessed. It was made apparent that in order to assess these assumptions, one required the use of analogy and ethnography (Raab & Goodyear 1984). For Lewis Binford, one of the leading theorists in this field, middle-range theory was intended to provide logico-empirical bridges between the material remains found in the archaeological record and the behavioural dynamics that are supposed to have brought the material remains into being (Raab & Goodyear, 1984). Binford (1977) describes middle-range theory as being one level in the archaeological investigation. He states that the observations made on the archaeological record are contemporary facts which are inherently static in their nature and therefore the problems facing archaeologists include translating these facts into statements about the past, and transforming these static facts into statements of dynamics (Binford, 1977). In other words, the static facts, the material remains that make up the archaeological record, can be described as the lower range while the higher range is the "big" theories on what happened in the past and how things worked. The idea behind middle-range theory was therefore to build a bridge between the two, to make interpretations of the material record that make statements about the past. Middle-range theory is "really just a way of bridging the gap between what archaeologists find and the behaviour that created the stuff" (Praetzellis, 2011).

Lewis Binford spent much of his career seeking to bridge the gulf between static remains in the present and dynamic processes in the past by attempting to establish unambiguous connections between material signatures and the behaviour that created them (Thomas, 2004). To reduce the ambiguity of the archaeological record, conventions were set up that stated that if a certain thing was found, it meant something very specific in terms of the past (Thomas 2004). However, if one is uncertain about a particular aspect of the material remains, and this cannot be resolved,

Binford suggests that it should be set aside because our present knowledge is too limited to resolve the ambiguity (Thomas, 2004). It is this reasoning that led Binford to be suspicious of interpretive arguments that link material evidence to events or processes in the past (Thomas, 2004). It is a criticism of Binford's work that the dynamics to which he aimed to connect the archaeological record do not include larger questions about the causes of change or stability in cultural systems (Raab & Goodyear 1984). When set against the larger goal of explaining cultural dynamics, middle-range theory is primarily methodological as it allows archaeologists to deal with material records but not with problems of cultural dynamism (Raab & Goodyear, 1984).

The present project on inter-analyst variability is not concerned with the larger questions about the past, nor even with past events or dynamics. It is primarily a methodological project and therefore, the weaknesses of middle-range theory concerning interpretive archaeology do not apply. My project is about investigating the strength of stone-wall analysis, and how much the results vary between analysts. If middle-range theory is about bridging the gap between today's archaeological record and the dynamic processes of the past (Raab & Goodyear, 1984), then my project is a study of how well that gap is bridged with respect to the classification of stone-walled structures in southern Gauteng. Placing each stone-wall structure into a certain group is not purely determined by the architectural styles but is also based on the interpretation of the Principle Identifying Attributes (PIA's) and interpretations of PIA's are likely to differ between analysts.

Chapter 2

Literature Review

The Highveld area of Gauteng has been studied extensively by many researchers over the past 45 years (Mason 1968; Maggs 1976; Taylor 1979; Huffman 2007). The great wealth of stone walled structures in these areas has stimulated much research on their classification using architectural details and general layout. The different types of structures have often been assigned to different ethnic or linguistic groups, making typologies of this kind crucial to the interpretation of archaeological data. While some researchers have focused predominantly on the typologies, others have focused more on the functions of such structures.

Huffman (2007) and Hall (1998: 235) are two such researchers who have focused predominantly on the functions of the stone walled structures rather than just the typologies. They assert that settlement structure is closely associated with a specific worldview (Huffman 1986; 2007). Different worldviews will produce different settlement organisations, especially in pre-industrial societies such as those found in the iron age of southern Africa (Huffman 2007: 23). It was with this as a basis that Huffman then formed his 'Central Cattle Pattern' (the CCP) theory. This is characterised by a male domain in the middle of the structure, where the cattle were kept, where important people were buried, and where men held court. All of this was then surrounded by a residential area and women's domain (Huffman 2007: 25). It has been shown that cattle held significant value in iron age settlements far beyond their economic value. This can be seen by the position they were given in the structure of the settlements (Maggs 1976: 23). It is usually situated in the centre of the settlement, and the men's meeting place, such as the *kgotla* of the Sotho-Tswana, which was also the administrative centre, was located in or right next to the cattle enclosure (Maggs 1976: 23).

Within the study of stone-walled structures, the use of aerial photography has been widespread. This is due to the fact that such photography gives the researcher a much better view of the layout of the structures than that which could be obtained from the ground. The first use of aerial photography in studying Iron Age sites in southern Africa was at Great Zimbabwe by Caton Thompson in the early 1930's (Maggs 1976). However, it was not until the 1950's that it was demonstrated just how useful such photographs were through the work of Walton (see Mason 1968) and the first systematic survey using aerial photography was done by Mason in 1964

(Mason 1968; Maggs 1976). Seddon (1968) was another important researcher to discover the uses of aerial photography in mapping the spatial layouts of settlements to further understand how the structures were used. Through the use of aerial photographs that covered up to 9km², Seddon (1968) identified up to 75% of the settlements in the area. Mason (1968) was to also have an impact through the use of aerial photography in his surveys of the region; and it was with these photographs that the identified structures were classified into different groups using the variation in architectural styles. From these studies they discerned 5 main classes of settlements and defined each one by its architectural characteristics (Mason 1968: 169). Mason (1968) also recognised the possibility of variation of results between analysts and took care to record how much disagreement occurred in the classification of the stone-walled structures. For example, in classifying class 1 structures, 4 of the 7 analysts identified between 12 and 17 structures within a given area while the other 3 identified between 23 and 27. This variability in interpretation will inevitably have an effect on the interpretation of the functions and of the builders of the stone-walled structures.

There have been many studies conducted into the problem of inter-analyst variability, not only in archaeology but in numerous other fields of study as well. There have been studies carried out on inter-analyst variability in fields such as accounting and auditing (Milne & Adler 1999), health promotion research (Oldenburg *et al* 1999) and social sciences (Kalton & Stowell 1979). A good example is the field of zooarchaeology where faunal analysis plays an important role, as does inter-analyst variability. The analysis of cut marks on bone is often done through the use of hand-drawn diagrams created by different analysts (Abe *et al*. 2002). While diagrammatic data provides a useful overview of the placement of the cutmarks, it is almost impossible to use such data in any comparative analysis due to the high frequency of inter-analyst variability (Abe *et al*. 2002). The work of the analyst in this field is further complicated by various processes of bone destruction after the bone had been discarded, such as carnivore scavenging and fragmentation through time (Abe *et al.* 2002). One way in which researchers aim to reduce the level of inter-analyst variability and thus increase the accuracy of comparative studies is through the use of Geographic Information Systems (GIS). The use of GIS has standardised the analysis of cutmarks (Abe *et al*. 2002) thus making the study of cutmarks far more reliable.

A further study into inter-analyst variability in archaeology has been carried out by Lyman and VanPool (2009). They argue that archaeology is essentially a 'comparative science', meaning that measurements and classifications are arrived at through the procedure of assigning descriptive

labels to an artefact according to a widely agreed upon set of units by following a certain set of rules (Lyman & VanPool 2009). For example, in this present study I will be assigning the descriptive labels of Group 1, 2, 3 or 4 to the stone-walled structures found in my given study area by following the rules set out by Taylor in his 1979 Masters thesis and the article written by Sadr and Rodier (2012). Lyman and VanPool (2009) argue that archaeologists use such measurements and classifications to such a large extent that it justifies carrying out studies into inter-analyst variability. The descriptions of certain attributes of an artefact are used to distinguish types, identify functional differences, track social interaction and change, as well as identify craft specialization (Lyman & VanPool 2009), showing that such measurements are in fact crucial to the study of archaeology.

Lyman and VanPool (2009) go on to discuss the ways in which the accuracy of such measurements can be tested. One suggestion is that the study be replicated by different people using different tools, if the results obtained by the different analysts are consistent then one can argue that the individual results are accurate. However, this will not be the case if there is a systematic bias throughout the analysts and their tools (Lyman & VanPool 2009). Lyman and VanPool (2009) have identified three different categories of error: blunders, bias and random errors.

- Blunders are simply problems of accuracy and comparability due to the misreading of data or the misplacing of decimal points (Lyman & VanPool 2009).
- Bias can be produced in various different ways, from either the analyst or the tools. For example, a specific tool may give out consistently lower or higher results relative to an accurate measurement, which may not always be immediately detectable (Lyman & VanPool 2009). Specific to this study, bias may have been created through the possibility of some researchers having access to higher resolution Google Earth (GE) imagery than the others. Bias amongst the analysts may be produced through differences in training or previous experience in the field of study.
- Random errors are also significant to this study as they involve situations where the analyst is forced to estimate a value based on what they can see (Lyman & VanPool 2009). Within my study this can occur when a section of stone walling is not visible on GE imagery and we have to estimate the most likely place that it would be, and this would inevitably end in variation between the analysts, both in terms of mapping out the sites, and later in their classifications.

The classification set out by Taylor in his 1979 Masters dissertation is as follows: Group 1 can be identified by "an elliptical wall enclosing a group of smaller enclosures in the centre" (Taylor 1979: 10); group 2 is a series of discontinuous, semi-circular walls that replace the distinct perimeter wall and face inwards towards an inner group of smaller enclosures (Taylor 1979: 31); and group 3 is identified as being a collection of circular enclosures, with an outer wall consisting of differing lengths of curved walls and small circular enclosures (Taylor 1979: 79).

While Mason (1968) uses the term Class 1, Huffman (2007) and Maggs (1976) refer to Group 1 sites as Type N but the classifications are otherwise the same. Type N is defined as a few cattle kraals in the centre linked together by other walls and then surrounded by a continuous outer wall (Maggs 1976: 33; Huffman 2007: 33). The walling identified as being Group 1 has been dated back to the 15th and 17th centuries in the Free State during which time these communities also spread across the Vaal region in to what is now known as Gauteng (Huffman 2007: 33). According to Huffman (2007: 38) Type N then developed into the Klipriviersberg walling which is classified as Group 3 by Taylor (1979). These settlements are more complex as the outer wall is often scalloped, there are more kraals, and there are straight walls that separate households in the residential zone. These settlements date back to the 18th and 19th centuries and are associated with the Fokeng people, as is Type N/Group 3 (Huffman 2007: 38). On the other hand, Group 2 is associated with the western Sotho-Tswana communities, such as the Hurutshe and the Kwena, and the walling in this group has been termed 'Molokwane' by Huffman (2007: 38). This is also the same as Mason's Class 6 and 9 (Mason 1968; Huffman 2007: 38). From aerial photographs, these settlements are represented by a kind of 'sunflower' shape, with the outer wall forming the back courtyards of households while the inner enclosures form the kraals where the sheep and cattle were kept (Huffman 2007). According to Boeyens (2003) it is customary to divide the Sotho-Tswana into two different phases. When identifying the difference between the two phases Boeyens (2003) points to an increase in stone-wall building and a change in settlement location, from flat lands to spurs and hilltops. This change in location and building style is attributed to a severe dry spell that occurred in the latter half of the 17th century (Boeyens 2003).

In the Suikerbosrand River basin there were also found to be small stone walled structures that did not fit this typology. Some of these small stone circles are isolated and may possibly be satellites of Group 1, 2 or 3. But there are instances where these stone circles are densely clustered so seem to form a group of their own. These have thus been classified as Group 4 sites by Sadr and Rodier (2012).

Sadr and Rodier (2012) carried out a survey on a section of the Suikerbosrand Nature Reserve. Google Earth (GE) imagery was used to survey a given area and place each stone-wall structure into the groups set out by Taylor (1979). A series of spatial analyses were then carried out on the data using GIS, including altitude, size, clusters, ranking, chronology, and an analysis of the distance to arable land and the inner enclosures (Sadr & Rodier 2012). The aim of their study was to investigate new technologies and see how they can aid in archaeological studies such as the classification of stone-walled structures.

The use of new technologies such as satellite imaging, GE and GIS, has had a profound impact on archaeological research. These tools effectively reduce the margins of error in accuracy by making the data and methods of data collection more advanced and systematized (Sadr & Rodier 2012). These tools also make it possible to easily return to survey areas at a later date when more advanced methods and tools are created to obtain even more accurate findings, effectively making studies such as this one easily repeatable. An example of a study concerned with the efficacy of these research tools is one that was carried out by MacQuilkan and Sadr (2010) in which aerial photography, Google Earth and ground-truthing was used to determine how best to survey stone-walled structures. It was found that Google Earth had the advantage of having the function of allowing one to see historical imagery of any tract of land, meaning that one could view the same area in different seasons and therefore see different details as the vegetation changed (MacQuilkan & Sadr 2010). Aerial photography may have the advantage of covering several decades, and the resolution is often uniformly good, but the images are often not easy to get hold of and they are often in black and white (MacQuilkan & Sadr 2010). Colour photography is often more useful as it is easier to differentiate between features such as stone walls, vegetation and open ground (MacQuilkan & Sadr 2010). This study found that about 78% of the stone-walled structures could be seen in GE imagery (MacQuilkan & Sadr 2012) as opposed to the 60-75% that could be seen in aerial photography (Seddon 1968).

The study of stone-walled structures has evidently played a significant role in archaeological research for a long time now, and the development of new technologies has made a huge impact on the accuracy of such studies and made them easily repeatable. However, inter-analyst variability is still an important issue which certainly has an impact on the results, but to an unknown extent. There are many factors that may have an impact on the degree of inter-analyst variability, including training, previous experience, and differing access to high resolution GE imagery.

Chapter 3

Methods and Techniques

Sadr and Rodier (2012) sets out the details of this study and provides a breakdown of the classification methods. All three analysts received this article and took part in a short course on how to use Google Earth and Quantum Geographic Information Systems (QGIS). This was the extent of our training. We then received two GE files from Professor Sadr: one containing the outline of the area that I was to study (named Pam 2); and the second file containing yellow markers where one or several of the 200+ stone-walled structures had previously been located. Using these files, I then systematically scoured the area for stone-wall structures which were then outlined using the 'path' tool on Google Earth. This meant that I often had to estimate where the outer walls lay due to parts of the wall not being visible any longer.

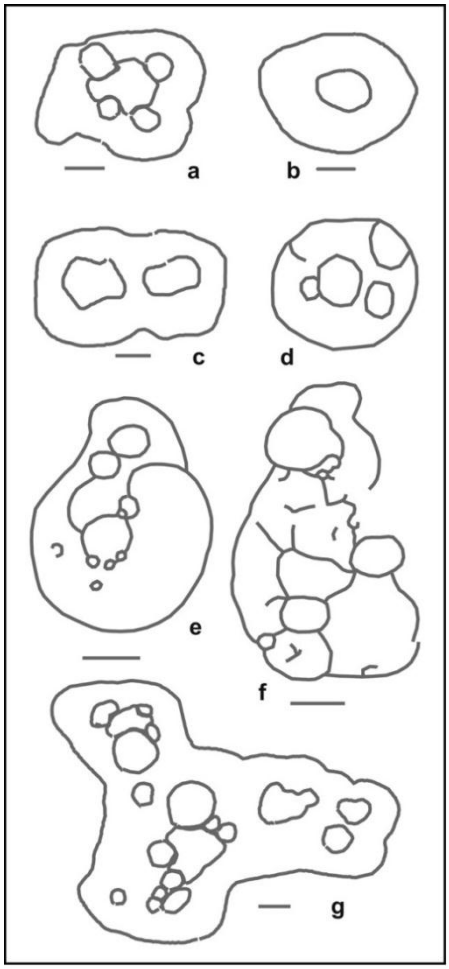


Figure 3.1 Group 1 according the Taylor's 1979 Masters thesis (see Sadr & Rodier 2012).

Each structure then had to be classified according to the classification method set out by Taylor (1979) and Sadr and Rodier (2012).

This involved placing each structure into one of four groups:

- Group 1 can be described as having an outer elliptical wall surrounding a group of smaller enclosures, it is often circular or oval, sometimes irregular, and there are sometimes small primary circles attached (fig. 3.1).
- Group 2 however, has a discontinuous series of c-shaped walls serving as an outer wall, a central group of inner enclosures which are less regular but are usually central and depending on the season, ash middens are sometimes visible (fig. 3.1).
- Group 3 has a continuous outer wall characterised by differing lengths of curved and scalloped walls, the inner enclosures are confusing, rarely central, and sometimes bisect the homestead. There are also more than in Group 1, but less than Group 2 (fig. 3.2).

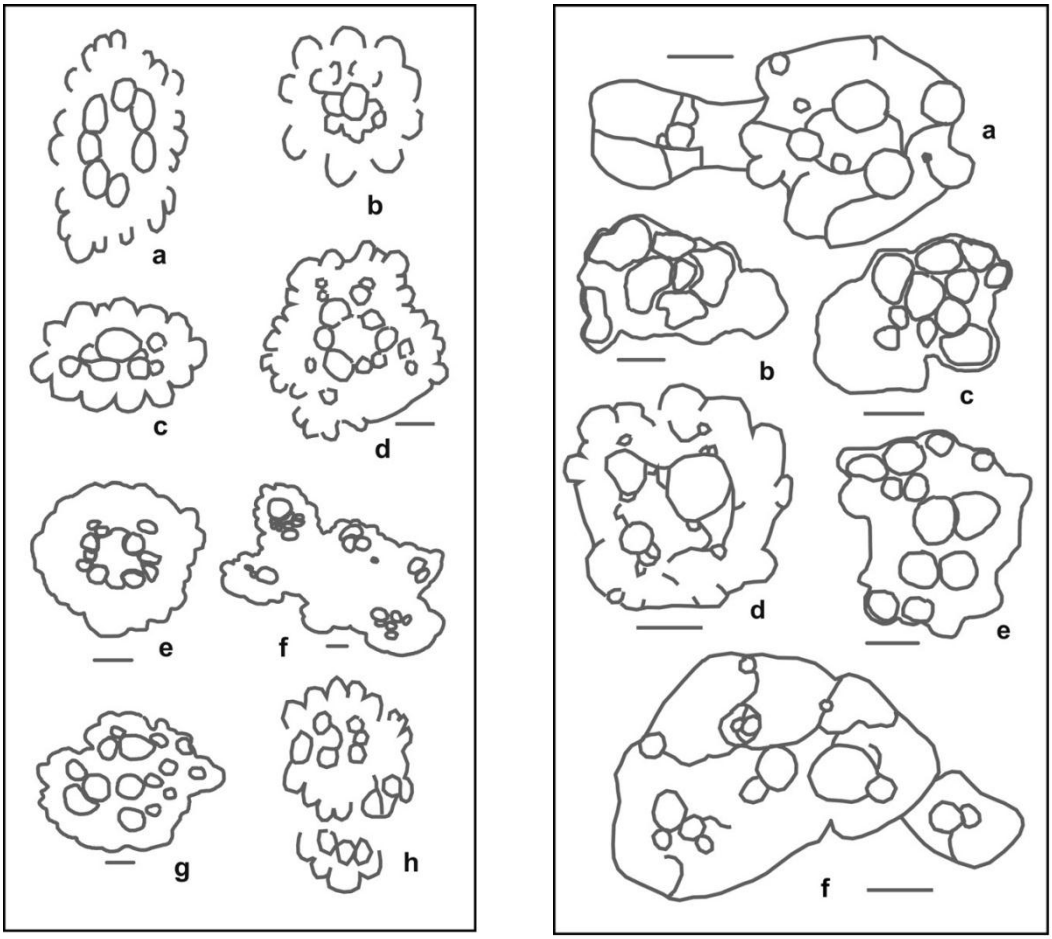


Figure 3.2 Group 2 (left) and Group 3 (right) according to Taylor's 1979 Masters thesis (see Sadr & Rodier 2012).

- The last group is Group 4 which is characterised by isolated primary circles with no outer walls. These could be satellites of groups 1, 2 or 3 but sometimes they are seen to be densely clustered which implies that they represent a distinct type of stone-wall structures.

Each group was then assigned a different colour marker, group 1 had white markers, group 2 was green, group 3 was pink, group 4 was blue, and all stone walls that I was unable to classify were marked with a question mark.

The 'Historical Imagery' tool in Google Earth was used to find the clearest image of each site. This tool allows the analyst to view the same site in different years and different seasons. This is extremely useful as the vegetation changes drastically according to the season. For example, in summer the vegetation will be thick and green due to the rainy season which is helpful in some situations. This is because the growth of bushes and thick vegetation along the stone walls will sometimes make them stand out against the short grass. On the other hand, sometimes the

vegetation is so thick that you cannot see any stone walls at all. In these cases it is necessary to view the imagery taken in dryer seasons where the vegetation is much sparser, brown and sometimes burnt, and will therefore reveal more of the stone walls. For every structure I found it necessary to see the site in several different seasons to find the one in which the walls are the clearest. However, going too far back into the historical imagery had the disadvantage that the resolution was not as good as the most recent imagery.

Once all stone-wall structures in the designated area had been outlined and classified, the data were imported into Quantum GIS 1.8.0, the Wroclaw Geographic Information Systems software. Using this progam each file was converted into ESRI shapefiles and projected using UTM 35S. Each path was converted into separate polygons, and a series of basic statistical analyses were done on the data. The altitude of the structures was found using the Point Sampling Tool in Quantum GIS software, version 1.8.0. The elevation of the geographical mean centre of each polygon was obtained from the Aster 30 m digital elevation model of the survey area. These statistics also allowed me to derive information about other aspects of the stone-wall structures such as average area and nearest neighbour of each group using software features such as 'Analysis Tools and 'Data Management Tools'.

I then received the results obtained by other students and I carried out the same statistical methods on their data. Their results were then compared to my findings using the descriptive statistics that I had pulled from Quantum GIS.

Each set of results was then imported into a spatial statistics program called CrimeStat which has the tools for doing spatial statistics on point locations (Smith & Bruce 2008). These statistics include:

- The Mean Centre for each group which identifies the mean centre of any group of points (Smith & Bruce 2008).
- The Standard Deviation Ellipse which "creates an ellipse representing one standard deviation around the mean centre for a group of points" (Smith & Bruce 2008).
- The Convex Hull which creates a polygon using the external points as a boundary (Smith & Bruce 2008).

To do this, the shape files from QGIS first had to be imported into the program DNR Garmin which has the function of providing each separate point with longitude and latitude values.

Once each site had longitude and latitude values it could be imported into CrimeStat which provided me with shapefiles containing the different statistics. These shapefiles could then be imported back into QGIS where the statistics were illustrated in the form of polygons and points. This made it easier to compare my results, group by group, with the results of the other analysts and determine the ways in which they differ.

The last step was to create a database in which I assigned each stone-walled structure a number. Each number was then loaded into Excel where I noted what classification each stone-walled structure had been given by each analyst. This made it easier for me to compare each analyst's classifications for individual stone-walled structures. The individual numbering system also enabled me to identify particular structures that were identified completely differently by the three analysts and to locate them again on Google Earth to find out what each analyst was seeing.

Chapter 4

Results

The analysis of my results involved three different stages. The first focused on descriptive statistics including average areas, nearest neighbour analysis, average elevations, and an analysis of the numbers of structures found by each analyst as well as how much the analysts agreed on individual structures. The second stage of my analysis involved a spatial analysis that included a comparison of the statistics drawn from CrimeStat such as the mean centres, Standard Deviation Ellipses and the Convex Hulls. Thirdly, a visual comparison of how each analyst was outlining and classifying the stone-walled structures is included.

4.1 Descriptive Statistics

4.1.1 Number of structures

The first thing that became evident when comparing my results to the other analysts was that we all identified a different total number of structures. We all received the same GE file with 114 points showing us where stone-walled structures were clustered. Using these points as a guide, analyst 'M' identified a total of 161 sites, analyst 'T' 234 sites and analyst 'K' 341 sites. It is already evident that there is a high rate of inter-analyst variability between the three people with 'M' identifying less than half the number of sites compared to 'K'. Figure 4.1 shows that 'T' and 'M' agreed most on the number of Group 1

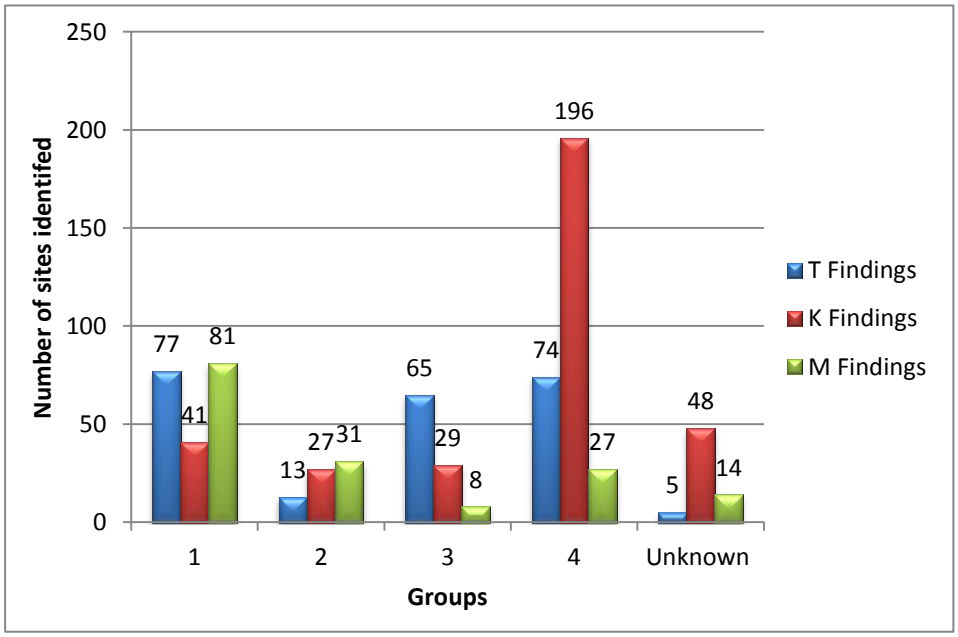


Figure 4.1 Graph showing the numbers of structures identified by each analyst.

structures in Pam 2, identifying 77 and 81 respectively, while 'K' identified only 41, much less than the others. In terms of Group 2 structures, all three seemed to agree that there were very few to be found in Pam 2 with 'T' identifying 13, 'K' 27 and 'M' 31 in total. Groups 3 and 4 also showed a large amount of variation between analysts (Fig. 4.1) with 8.1 and 7.3 fold difference between highest and lowest numbers respectively. Each analyst then made a fifth group named 'Unknown' in which they placed those structures that could not be classified as any of the four groups. In this there was also much variation in deciding what could be classified and what could not.

4.1.2. Levels of agreement between analysts

One thing descriptive statistics like fig. 4.1 do not show us is how many individual sites were identically identified by each analyst. Table 4.1 shows how many sites were classified as each group by all the analysts. For example, out of all the sites in Pam 2, 119 were classified as Group 1 by at least one of the analysts. However, only 19 of these sites were classified as Group 1 by all three (Table 4.1), showing only a 16% level of three-way agreement. In a two-way comparison the results were a bit better, 'T' and 'K' agreed with each other on 19% of the sites, 'T' and 'M' agreed with each other the most with 34%, and 'K' and 'M' agreed with each other on 27% of the Group 1 sites.

Table 4.1. The number of sites on which the analysts agreed with each other. Several sites have been counted more than once as they have been classified as different groups by each analyst.

	Group 1	Group 2	Group 3	Group 4	Unknown
Total sites	119	38	74	246	61
Agreed by all	19	10	3	14	0
T and K	23	10	18	30	0
T and M	40	13	5	15	0
K and M	32	23	5	25	3

Group 2 saw much less variation between the analysts, with a 26% level of three-way agreement. 'K' and 'M' had the least amount of variation in their results, as they agree on 60% of their Group 2 sites. 'T' and 'M' came next with 34%, and 'T' and 'K' agreed the least with 26%. At first it may seem as though Group 3 had high levels of disagreement, however, these numbers will be skewed because 'M' had so few Group 3 sites and 'T' and 'K' had so many. Therefore, the fact that 'M' agreed with 'T' and 'K' on 5 of their 8 sites shows a small level of agreement on the part of 'M'. 'T' and 'K' only saw eye to eye on 24% of their Group 3 sites.

The numbers on Table 4.1 for Group 4 showed that 'K' and 'M' agreed on 25 sites which is actually very good considering that 'M' only identified 27 sites overall. 'T' and 'M' agreed on 15, and 'T' and 'K' agreed on 30 which is also not too bad since 'T' identified only 74 Group 4 sites. No individual site was classified as Unknown by all three analysts and only 3 were classified as Unknown by both 'K' and 'M' (Table 4.1).

Of a grand total of 382 sites identified by at least one of the three analysts, only 46 were classified as the same group by all three analysts, therefore showing that only 12% of the sites were agreed upon in Pam 2. 'T' and 'K' agreed with each other 81 times out of the 382, showing a 21% rate of agreement. 'T' and 'M' agreed with each other 73 times, therefore having a 19% rate of agreement; and 'K' and 'M' agreed with each other 88 times, (23%).

4.1.3 Area Averages

When comparing the average areas given by each analyst, the mean value was often the most valuable as it was the average area of all structures classified in a particular group by each analyst. However, a problem arises if there are outliers, in other words, structures that are much smaller or larger than the average, and these outliers have an effect on the mean. The median for each group tells us the middle value for each dataset and therefore when the median is very different to the mean for a particular analyst it shows us that the dataset is skewed and therefore will have more outliers. On the other hand, a median that is close to the mean shows that there are few outliers in the dataset. For example, from Table 4.2 we can see that 'T' and 'M' both have mean areas that are much larger than that of 'K', but their max areas are also much larger, more than double in the case of 'M', this maximum value is then a statistical outlier and it would have skewed the mean.

The medians for 'T' and 'M' are very different from their means and one can therefore show that they have far more outliers than 'K', whose median is very close to the mean. This is further illustrated by the Coefficient Variance (CV) and standard deviation which basically shows how close all the values are to the mean, but in this case the CV will be more useful than the Standard deviation as we are dealing with sets of data that differ in size, and the Standard Deviation differs according to the size of the dataset. Therefore, a high CV (one that is close to or higher than 1) means that there is a greater spread of data, which includes extreme values as well as values between the mean and the extreme. From Table 4.2, we can see that the CV values for 'T' and 'M'

are higher than that of ‘K’ showing that there are many more extreme values in their dataset than in that of ‘K’s and is therefore the reason for the elevated mean areas of ‘T’ and ‘M’.

Table 4.2 Area data for Group 1

	Group 1 by T	Group 1 by K	Group 1 by M
Mean (m^2)	3345.6	2537.2	3420.1
Median (m^2)	2580.2	2484.8	2929.7
Min Area (m^2)	107	721.5	167.2
Max Area (m^2)	10641	5405.8	11630.4
Standard deviation (m^2)	2342.77	1161.32	2593.2
CV	0.7	0.5	0.8

Table 4.3 Area data for Group 2.

	Group 2 by T	Group 2 by K	Group 2 by M
Mean	4845.8	7734.9	4796
Median	4857.9	7005.3	4130.5
Min Area	1450.7	676.2	1081.1
Max Area	7886.8	23024.2	14769.1
Standard deviation	2019.7	5676	2955.1
CV	0.4	0.7	0.6

Group 2 showed similar results to Group 1 in terms of average areas, as ‘T’ and ‘M’ seem to have a very similar mean area. However, the max area for each shows that ‘M’ is almost double that of ‘T’ (Table 4.3), and the median also varies from the mean, indicating that ‘M’ has many more outliers than ‘T’. ‘K’ has a larger CV value than both ‘M’ and ‘T’ and a max value that is almost triple that of ‘T’. Correspondingly, the mean area for ‘K’ is much larger than that of ‘T’ and ‘M’. This shows that ‘K’s sites are more dispersed around the mean, are much larger and vary a lot in size compared to the other analysts.

Table 4.4 Area data for Group 3.

	Group 3 by T	Group 3 by K	Group 3 by M
Mean	6918.7	6107.87	9417.4
Median	4698.3	4832.1	9619.5
Min Area	163.9	717.8	3432.9
Max Area	45739.7	17221.9	17744.6
Standard deviation	7714	4334.7	4607.2
CV	1.1	0.7	0.5

Group 3 showed a lot of variation in terms of classification and size. By looking at Table 4.4 it may first appear that there is a measure of agreement between 'T' and 'K' as their mean areas are very similar. However, the max area for 'T' is more than double that of 'K' and both the medians for 'T' and 'K' vary a lot from the mean showing that there are many outliers. This shows that 'T' had several sites that were much larger, but as shown above (Fig. 4.1), 'T' also identified many more Group 3 sites than the others. While 'M' has a low CV, showing a measure of consistency in size, it was stated above that they only found 8 Group 3 sites, and the fact that the mean area for 'M' is much higher than that of 'T' or 'K' shows that their Group 3 sites were consistently larger than the other analysts.

Figure 4.2 shows that there is a more similarity between 'T' and 'M' than either to 'K' in that both analysts classified all the largest structures as Group 3. This is evident as the mean area for Group 3 is higher than that of any other group. 'K' is different in this respect as their mean area for Group 2 is significantly higher than it is for Group 3. One can also see from Fig. 4.2 that all three analysts are mostly in agreement in terms of Group 4, as it was consistently found to have the smallest area and it was therefore agreed that Group 4 was the smallest group (Table 4.5). 'K' is quite a lot higher than 'T' and 'M' but it is still smaller than any other group classified by 'K'.

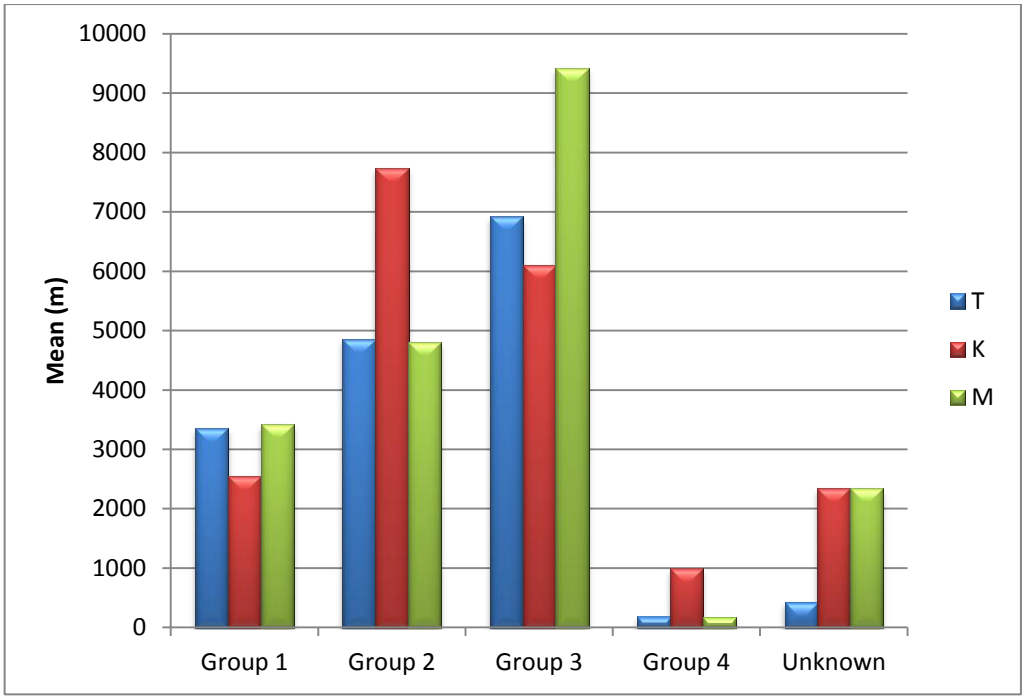


Figure 4.2 Graph showing the mean areas for each group and analyst.

Table 4.5 Area data for Group 4.

	Group 4 by T	Group 4 by K	Group 4 by M
Mean	181.5	989.9	176.7
Median	94.2	192.5	172.6
Min Area	19.7	8.4	114.9
Max Area	835.8	17064.3	257.1
Standard deviation	192	2361.4	37
CV	1.1	2.4	0.2

The most striking aspect for Group 4 is the evident variation in size. 'K' has by far the greatest spread of values with a CV of 2.4 while 'M' seems to be the most consistent with a CV of only 0.2 (Table 4.5). This is also evident in the max areas in which 'K' is more than 20 times larger than 'T' and more than 66 times larger than 'M', this is further illustrated by 'K's median which is much smaller than their mean showing that there were extreme outliers in their dataset. However, it must be taken into consideration that 'M' only found 27 Group 4 sites while 'T' and 'K' were working with a much larger dataset, with their Group 4's numbering at 74 and 196 respectively. With a larger sample size there was a greater probability that 'T' and 'K' would record an outlier, however their mean was also taken from a greater number of samples and is moderated by this.

We already know that the Unknown Group has a high rate of inter-analyst variability due to the differing sizes of each dataset as illustrated in Fig. 4.1. In terms of mean areas, it seems that 'K' and 'M' are almost identical while 'T' is much smaller than the others (Table 4.6), but this can be because 'T' had much fewer Unknown sites than either 'K' or 'M'. In terms of variation of size, 'K' has the highest amount due to the high CV of 1.1 (Table 4.6) while 'T' is the most consistent with a CV of only 0.4.

Table 4.6 Area data for the Unknown Group.

	Unknown by T	Unknown by K	Unknown by M
Mean	432.7	2337.4	2340.8
Median	366.1	1204.9	2056.8
Min Area	251.7	136	314.4
Max Area	747	14337.4	5692.3
Standard deviation	191	2653.4	1815.5
CV	0.4	1.1	0.8

These area averages tell us that while each analyst's outlines do differ significantly, there are certain patterns throughout the data. All three analysts agreed that Groups 1 and 4 were on average the smallest in size while Groups 2 and 3 were the largest. There was some disagreement as to which was the largest as 'K' found Group 2 to be larger than Group 3, while 'T' and 'M' agreed that Group 3 was the largest group of all. It is also evident that there is more often two-way agreement than there is three-way. Groups 1, 2 and 4 shows two-way agreement between 'T' and 'M' while Group 3 shows more agreement between 'T' and 'K' than either with 'M'.

4.1.4. Nearest Neighbour Analysis

A nearest neighbour analysis was done to determine the distance of one type of stone-walled structure to its nearest neighbour of the same type. The nearest neighbour index (NNI), often known as R (Clark & Evans 1954) shows the degree to which the structures are clustered or dispersed. The NNI is calculated by dividing the observed mean distance by the expected mean distance (Clark & Evans 1954). The observed mean distance is the mean of the actual measurements between the centres of the structures while the expected mean distance is the mean distance to the nearest neighbour in an infinite, hypothetical distribution (Clark & Evans 1954). Therefore, if the NNI is closer to zero it means that the sites of that particular group are

more clustered, and if the NNI is close to or higher than one it means that the sites are more dispersed.

All analysts found Group 1 structures to be quite clustered (Table 4.7), with NNI's of 0.5 for 'T' and 'M' and 0.6 for 'K'. Even though the sample sizes for each analyst were so different, it is important to note that all found Group 1 to have around the same level of fairly dense clustering.

Table 4.7 Nearest neighbour analyses for Group 1.

	Group 1 by T	Group 1 by K	Group 1 by M
Observed Mean Distance (m)	414.1	711.3	363.8
Expected Mean Distance (m)	851.5	1121.3	713.4
Nearest Neighbour Index	0.5	0.6	0.5
Number of Sites	77	41	81

Group 2 showed more varied results between the analysts. 'K' and 'M' had very similar NNI's of 0.3 and 0.4 respectively (Table 4.8) indicating clustering of Group 2 sites even though the observed mean distance for 'M' is much smaller than for 'K'. 'T' has the highest observed mean distance and NNI so it can be shown that while there is still a large amount of clustering, the Group 2 sites for 'T' are more dispersed than for 'K' or 'M'.

Table 4.8 Nearest neighbour analyses for Group 2.

	Group 2 by T	Group 2 by K	Group 2 by M
Observed Mean Distance (m)	262.4	213.6	192.7
Expected Mean Distance (m)	493.5	813.6	506.6
Nearest Neighbour Index	0.5	0.3	0.4
Number of Sites	13	27	31

Group 3 also shows a lot of variation in NNI between the analysts. The most obvious difference is 'M' with an NNI of 1.1 (Table 4.9) that indicates that the sites are very dispersed, but this is understandable when one takes into account that there are only eight Group 3 sites in this analysts dataset which makes it an unrepresentative sample. With such a small dataset the chance of randomness in spatial distribution is greater, which means that you could have a highly dispersed group, or a highly clustered one. Conversely, a large dataset is more representative and

easier to draw conclusions from. Analysts 'T' and 'K' have similar results, both with evidence of clustering in their datasets (Table 4.9).

Table 4.9 Nearest neighbour analyses for Group 3

	Group 3 by T	Group 3 by K	Group 3 by M
Observed Mean Distance (m)	319.9	549.3	485.2
Expected Mean Distance (m)	1004.5	1004.7	458
Nearest Neighbour Index	0.3	0.6	1.1
Number of Sites	65	29	8

Group 4 showed the most consistent rate of clustering between the analysts out of all the groups. 'T' and 'K' once again have similar results showing extremely dense clustering of sites, while 'M's sites are slightly more dispersed yet still quite clustered (Table 4.10). The difference is in the observed mean distance as 'M' shows a mean distance that is almost 6 times larger than that of either 'T' or 'K' (Table 4.10). The number of sites is also significant as 'M' identified far fewer than the other analysts. This shows that while 'M's Group 4 sites are clustered, there are far fewer of them and they are much further apart than those in 'T' or 'K's datasets.

Table 4.10 Nearest neighbour analyses for Group 4.

	Group 4 by T	Group 4 by K	Group 4 by M
Observed Mean Distance (m)	79.7	80.7	479.9
Expected Mean Distance (m)	649.3	516	950.4
Nearest Neighbour Index	0.1	0.2	0.5
Number of Sites	74	196	27

The Unknown Group is also a source of variation between the analysts. 'T' shows the most dispersion with an NNI of 1.7 (Table 4.11) and an observed mean distance of over 3500m showing that their Unknown structures are very few and far between. 'M' has slightly more Unknown structures which are relatively clustered with an NNI of 0.8 and 'K' had the most clustering, with the shortest mean distance and that smallest NNI (Table 4.11).

Table 4.11 Nearest neighbour analyses for the Unknown Group.

	Unknown by T	Unknown by K	Unknown by M
Observed Mean Distance (m)	3531.6	591.4	1317.2
Expected Mean Distance (m)	2038.2	1202	1755.5
Nearest Neighbour Index	1.7	0.5	0.8
Number of Sites	5	48	14

The nearest neighbour analysis tells us exactly how clustered each group is and how much each analysts group differs in that way. From the analysis above we can see that in terms of clustering the analysts agree that Group 1 and 2 sites are fairly clustered together. There was also agreement on the clustering of Group 4 but more so between 'T' and 'K' than with 'M'. Group 3 shows a high level of disagreement with 'T' showing a high level of clustering while 'M's Group 3 sites are very dispersed. 'K' seems to fall in the middle of the two analysts. Overall the analysts seem to agree that the different groups are highly clustered, Group 3 and the Unknown Group are the only groups that show sites that are not very clustered, and then its only those sites from one of the analysts.

4.1.5. Elevation

A study of the mean elevation of each group from each analyst is important as it may reveal trends concerning a specific group of structures being built at a higher altitude to the others. The altitude of the stone-walled structures in Pam 2 ranged from 1446 to 1709 metres above sea level. The average altitudes for 'T' showed little variation between the five groups except for Group 2 which seemed to be at a much lower altitude than the others at 1520 metres above sea level (Fig. 4.3). The results for 'K' showed that the altitudes for Groups 1, 4 and Unknown averaged between 1569 and 1573 metres above sea level, Group 2 was comparatively much lower with an average altitude of 1534 metres, and Group 3 at a higher altitude than the rest with an average of 1614 metres (Fig. 4.3). 'M' shows the most variation, with Group 2 at 1515 metres above sea level, Group 3 at 1633 metres and the rest varied in between (Fig. 4.3). All three analysts seem to have very similar altitudes for Groups 1 and 2, with Group 2 being located at a predominantly lower altitude than all the other groups (Fig. 4.3). Groups 3 and 4 have the most variation between analysts with 'T' showing consistently lower altitudes while 'M' shows the highest (Fig. 4.3).

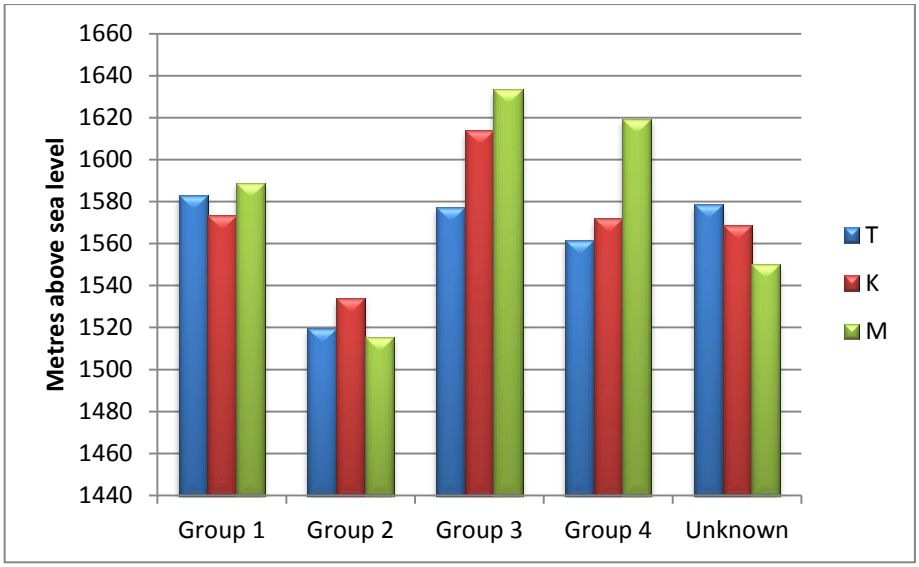


Figure 4.3 Graph showing the mean altitudes for each group and analyst.

4.2 Spatial Statistics

This section concerns spatial statistics calculated using CrimeStat that, when imported into QGIS, gave more information on how the classifications of each analyst differed. The statistics include the Mean Centre, the Standard Deviation Ellipse (SDE) and the Convex Hull (CHull).

4.2.1 Mean Centre

The mean centre identifies a point that is the average centre of a group of points (Smith & Bruce 2008) and by comparing the distances between the mean centres for each analysts group we are able to tell how much the spatial distribution of each analysts group differs. When comparing the mean centres for each analyst it stands to reason that a shorter distance between the points means that the spatial distribution for those analysts are more similar than if the distance is larger. Therefore, in Table 4.12 we can see that the mean centre for 'T' and 'K' in terms of Group 1 is only 1.5km while 'T' and 'M' is 2.7km. In this case 'K' and 'M' differ the most with a distance of 4km between their mean centre points. This shows that in terms of the spatial distribution of Group 1 sites, 'T' and 'K' agree with each other while 'M' agrees more with 'T' than with 'K'.

Out of all the groups, Group 2 shows the least variation between analysts with the distance between the mean centre points at a minimum (Table 4.12). This does not necessarily mean that there is agreement on all Group 2 sites, in fact as discussed above there is very little agreement on any of the Group 2 sites. It does mean however that all analysts agree that Group 2 sites are widely spread across Pam 2. Group 3 on the other hand, shows very little agreement with the

distances between mean centre points varying between 3.8km and 8.7km. This shows that 'K' and 'M' have a more similar distribution of points than either one has with 'T' (Table 4.12).

Table 4.12 The distance between the Mean Centre points for each analyst.

	Group 1	Group 2	Group 3	Group 4	Unknown Group
T and K (km)	1.5	2.2	5.3	1.2	4.1
T and M (km)	2.7	1.1	8.7	8.5	1.3
K and M (km)	4	1.6	3.8	8.1	4.4

Considering the statistics discussed above, it is not surprising that 'T' and 'K' have such a similar spatial distribution for Group 4. It is also not surprising that the mean centre point for 'M's Group 4 is so far away from either 'T' or 'K' as their Group 4 was much smaller than that of 'T' and 'K' and they did not agree on many of the sites. The Unknown Group shows a short distance of 1.3km between the mean centres of 'T' and 'M', while the distances between the mean centres of 'T' and 'K', and 'K' and 'M' are very similar with 4.1 and 4.4km respectively (Table 4.12).

4.2.2 Standard Deviation Ellipse

The SDE represents the standard deviation around the mean centre for each group (Smith & Bruce 2008). Much like the CV, the standard deviation shows how close all the values are to the mean and therefore a comparison of these ellipses will show us how varied each group is from the other. The spatial standard deviation ellipse is different to the numerical one seen in the area analysis in the way that the ellipse also indicates the size but also the shape of the area in which the dataset lies. For example, Fig. 4.5 shows the SDE's for Group 2 and you can see immediately that the dataset for 'K' lies in an entirely different area to that of 'T' or 'M'. For each of the following figures, the results for each analyst are colour coded in such a way as to make them easily identifiable.

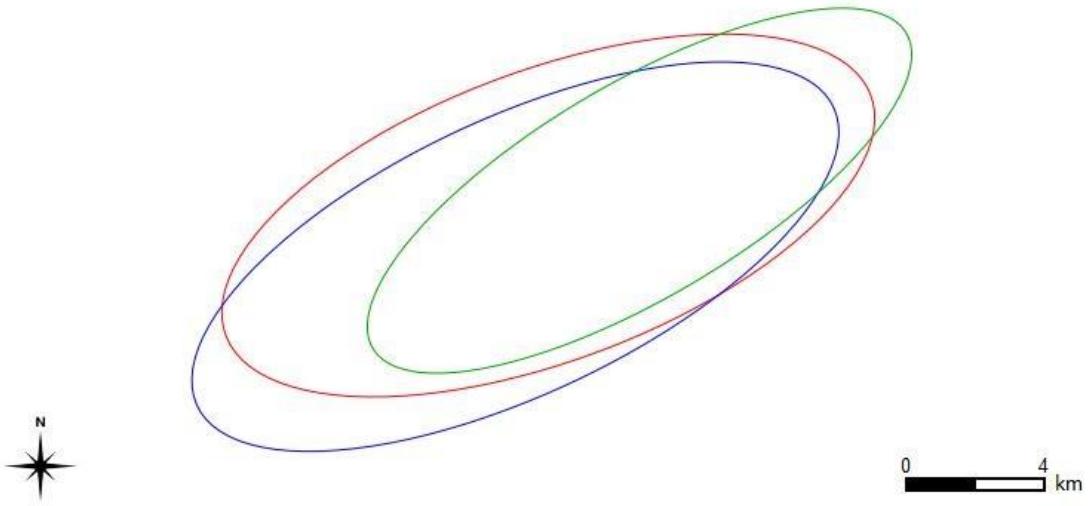


Figure 4.4 QGIS image showing the One Standard Deviation Ellipses for Group 1. ‘T’ is red, ‘K’ is blue and ‘M’ is green.

Fig. 4.4 shows that the values for Group 1 for all the analysts are widely spread throughout Pam 2. We can tell this because the scale is quite high showing that the SDE’s cover a great distance, meaning that on average, the Group 1 sites are spread considerably away from their mean centres compared to some of the other groups. We can also see that the results of ‘T’ and ‘K’ are similar while ‘M’ is slightly different to either of them. However, it is also evident that though they differ slightly, the standard deviations of the spatial distribution for Group 1 are very similar.

Figure 4.5 shows quite clearly that there is much more variation in the results for Group 2 than

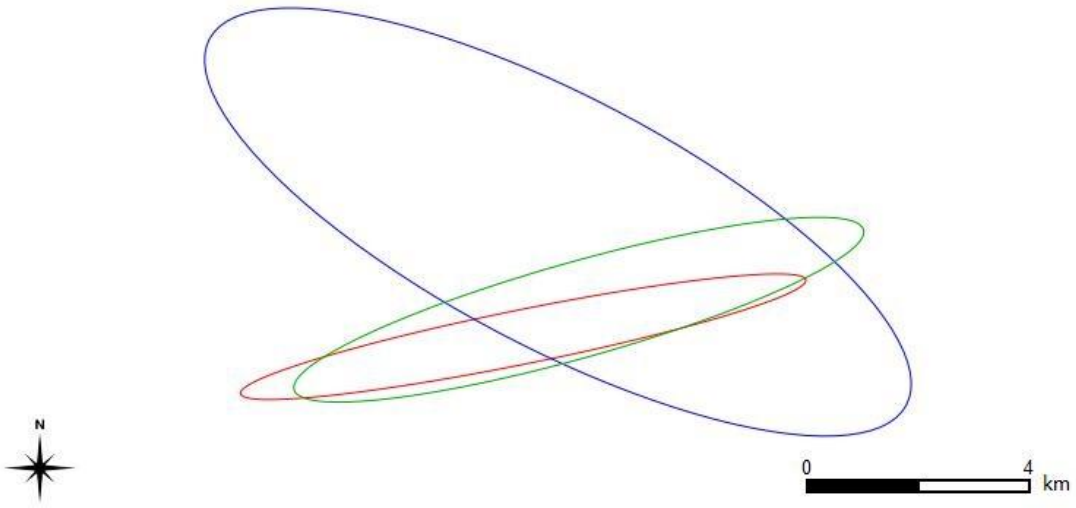


Figure 4.5 QGIS image showing the Standard Deviation Ellipses of Group 2. 'T' is red, 'K' is blue and 'M' is green.

was evident in Table 4.12. This is unusual considering that in many of the other analyses Group 2 showed the least variance. The SDE for 'K' is much larger and in a completely different direction to that of 'T' or 'M' showing that the Group 2 values for 'K' are not as close to the mean as those of 'T' and 'M'. It also shows that 'K' has completely different Group 2 sites than 'T' or 'M' which has resulted in their SDE covering a completely different area to those of the other analysts. On the other hand, 'T' and 'M' seem to agree quite closely with each other on the spatial distribution of Group 2 sites (Fig. 4.5).

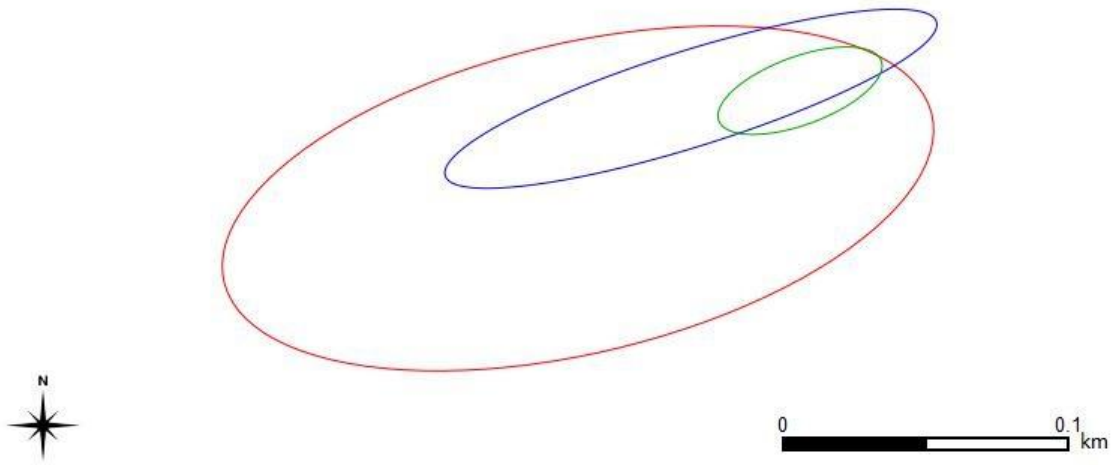


Figure 4.6 QGIS image showing the Standard Deviation Ellipses for Group 3 sites. 'T' is red, 'K' is blue and 'M' is green.

Figure 4.6 shows that there is small overlap on the spatial distribution of Group 3 sites. 'T' has the largest SDE almost encircling the SDE's of the other analysts. This shows that their Group 3 points are very far away from their mean centre and are therefore widely distributed. The SDE for 'K' shows that their results are closer to their mean centre but are distributed in a completely different way to those of either 'T' or 'M'. The SDE for 'M' is much smaller than either of the others which shows that their results are extremely close to their mean centre.

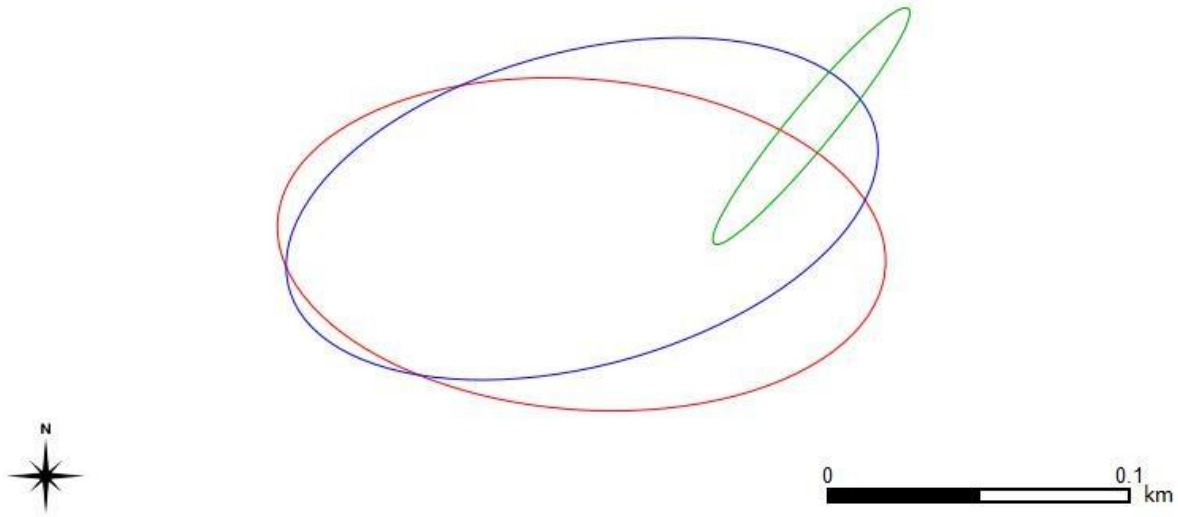


Figure 4.7 QGIS image showing the Standard Deviation Ellipses for Group 4. 'T' is red, 'K' is blue and 'M' is green.

The SDE's depicted in Fig. 4.7 support the statistics that have been discussed above such as the average areas and nearest neighbour analysis. It is evident that 'T' and 'K' agree strongly on the spatial distribution of Group 4 sites with very similar SDE's. Both are quite large ellipses which shows that the sites are widely scattered throughout Pam 2. 'M' seems to disagree with the other analysts with the SDE showing a spatial distribution that is very different from that of 'T' or 'K'. The SDE for 'M' shows that their Group 4 sites are fairly close to their mean centre and are therefore not as widely distributed as those of the other analysts. It is also a different shape (Fig. 4.7) showing that their Group 4 sites have a very different spatial distribution to those classified by 'T' and 'M'.

The SDE's for the Unknown Groups for all three analysts are overall quite similar (Fig. 4.8). All of the SDE's for this group are very large showing that the sites are very far away from their mean centres and are therefore widely distributed. 'T' has the largest SDE out of all three and therefore has sites that are more widely distributed than the other analysts.

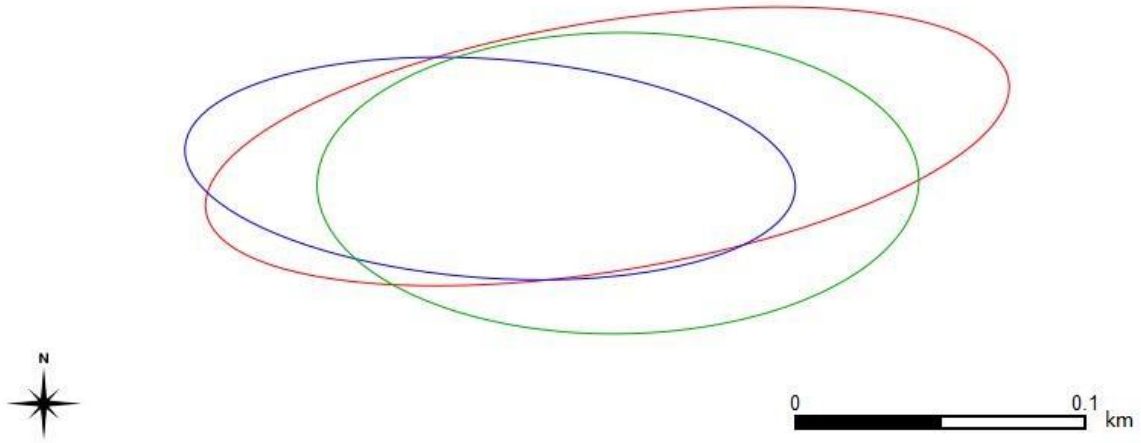


Figure 4.8 QGIS image showing the Standard Deviation Ellipses for the Unknown Group. 'T' is red, 'K' is blue and 'M' is green.

4.2.3 Convex Hull

The CHull is a polygon using the external points of a particular group as its boundary (Smith & Bruce 2008), this is useful as the differences in shape will tell us how much the spatial distribution of the groups differ between analysts. For example, Fig. 4.9 shows the CHull's for all three analysts for Group 1. You can immediately see that this corresponds with the spatial statistics described above, where all analysts mostly agree on the spatial distribution of their Group 1 sites. The first thing that is noticeable when looking at Fig. 4.10 is that the CHull for 'K' is much larger

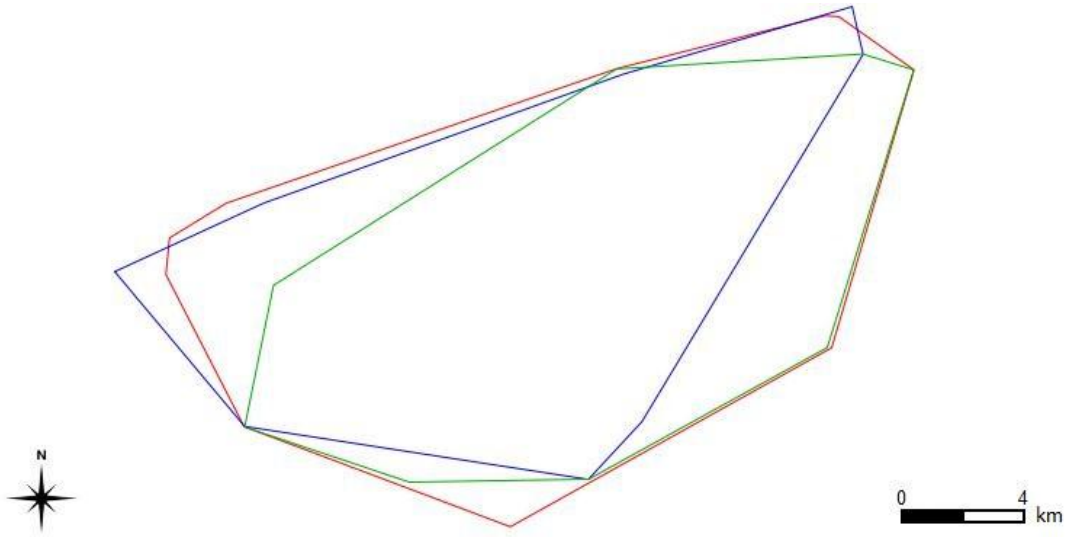


Figure 4.9 QGIS image showing each analysts CHULL for Group 1. 'T' is red, 'K' is blue and 'M' is green.

than that of ‘T’ or ‘M’. This means that the Group 2 sites for ‘K’ are much more dispersed than those of ‘T’ and ‘M’. ‘T’ has the smallest CHull which shows that their Group 2 sites are very close together. You can see from Fig. 4.10 that there is not very much overlap at all showing that there are very few sites that were classified as Group 2 by all three analysts.

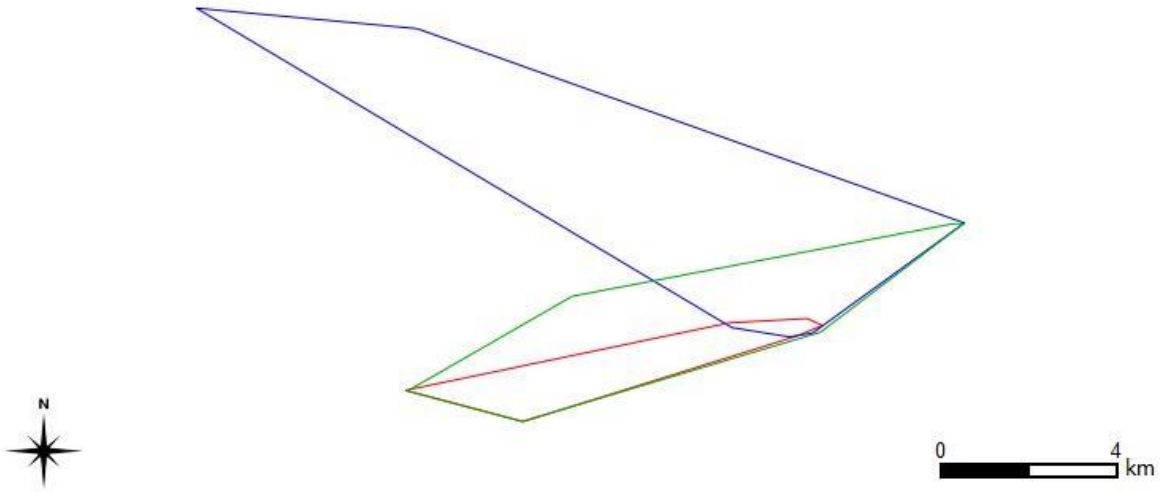


Figure 4.10 QGIS image showing each analysts CHull for Group 2. ‘T’ is red, ‘K’ is blue and ‘M’ is green.

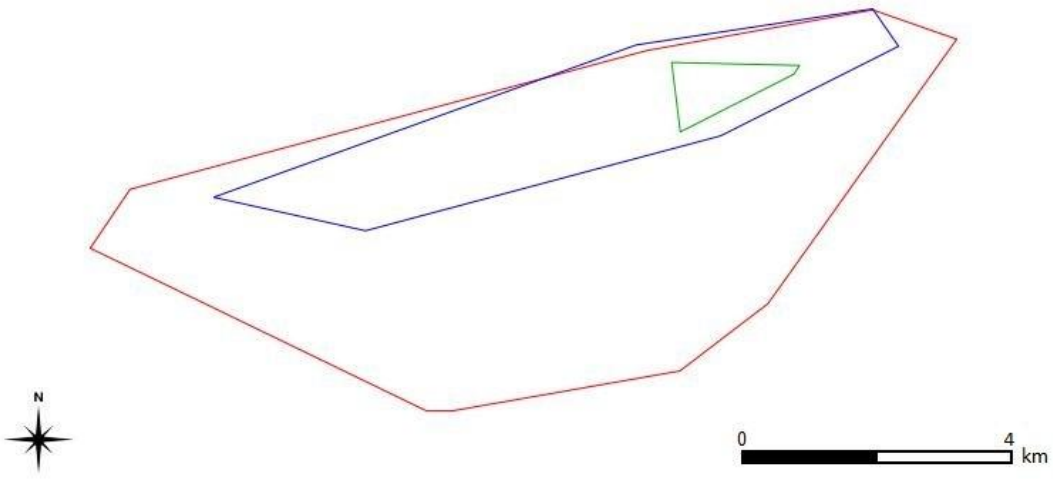


Figure 4.11 QGIS image showing each analysts CHull for Group 3. ‘T’ is red, ‘K’ is blue and ‘M’ is green.

For Group 3 (Fig. 4.11), ‘T’ has the largest CHull and therefore their sites are more dispersed than those of the other analysts. The fact that the CHull for ‘M’ is so small is not surprising since they had so few sites but it does show that what few sites they had were very close together. The fact that the CHull’s for each analyst are all inside one another (Fig. 4.11) shows us that there must have been some overlap in the classification of some of the sites.

The green shaded section in Fig. 4.12 for Group 4 is the CHull for ‘M’. It was shaded to make the diagram more clear as ‘M’ shares the same border as ‘K’. Compared with ‘T’ and ‘K’, ‘M’ has a very strange spatial distribution. Where ‘T’ and ‘K’ Group 4 sites cover a large, general area, ‘M’ Group 4 sites are located in a very particular strip running down the length of Pam 2.

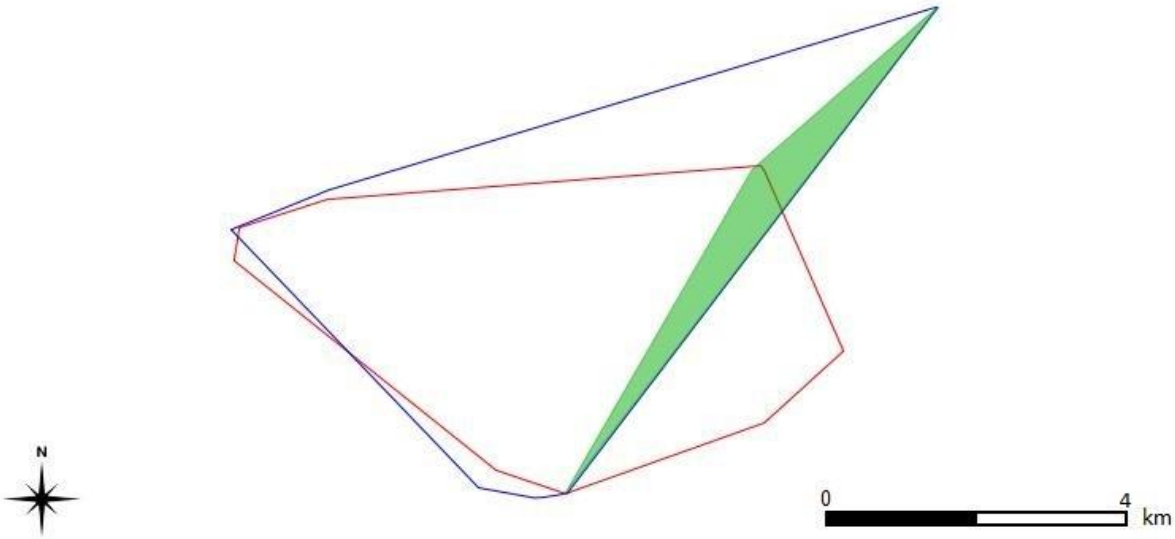


Figure 4.12 QGIS image showing each analysts CHull for Group 4. ‘T’ is red, ‘K’ is blue and ‘M’ is green.

The CHull’s for the Unknown Group show a lot of variation between the three analysts, where ‘K’ has the largest CHull and therefore their Unknown sites were much more dispersed than those of ‘T’ or ‘M’. ‘T’ had the least amount of Unknown sites and this is illustrated in Fig. 4.13 where their CHull is much smaller than those of the other analysts.

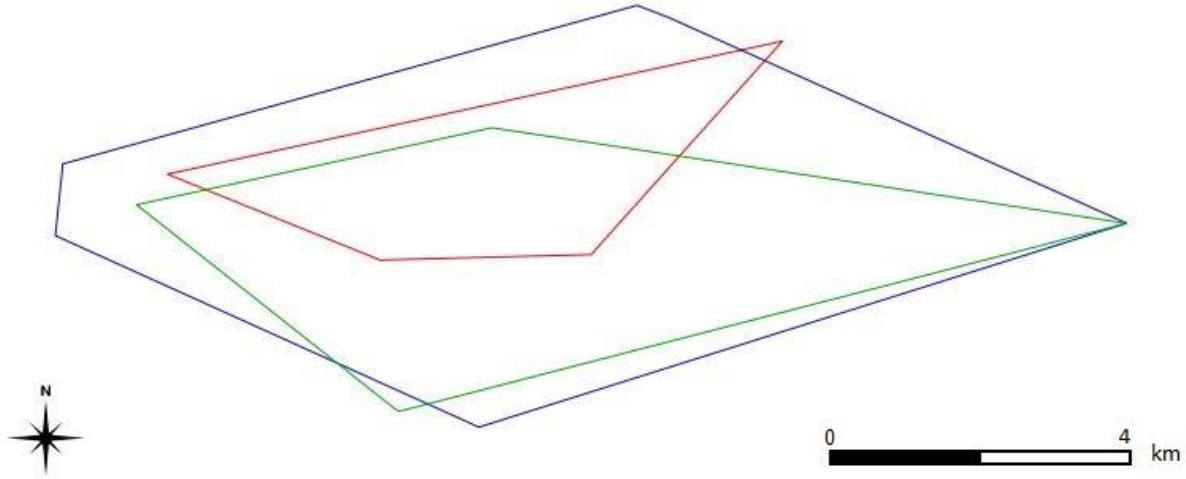


Figure 4.13 QGIS image showing each analyst's CHull for the Unknown Group. 'T' is red, 'K' is blue and 'M' is green.

The spatial statistics all show mostly the same results. They show that in terms of spatial distribution, Group 1 shows the most similarity between the analysts. Group 2 shows two-way agreement between 'T' and 'M' while 'K's Group 2 sites lie in a different area to those of the other analysts (fig. 4.10). Group 3 shows little agreement between any of the analysts, and Group 4 shows more agreement between 'T' and 'K' than either with 'M'.

4.3 Visual Comparisons

The significant inter-analyst variability meant that I had to go back to Google Earth to find out why the results differed so much. A study of the outlines of the compounds by each analyst was done to determine how each analyst was outlining the structures and would therefore make it clearer as to why they were classifying each structure as they were. For clarity each analyst's outlines of certain stone-walled structures were overlaid and colour coded and a QGIS image was generated in each case to make the differences clear.

4.3.1. Outline Overlap

Overall there was a large amount of variation when it came to drawing outlines around the outer walls of the stone-walled structures. However, the group that had the most amount of overlap, where all three analysts agreed on the position of the outer wall, was predominantly Group 1. Figure 4.14 is a QGIS image of one of the sites that had been classified as Group 1 by all three

analysts. This image clearly shows that each analyst was seeing the same thing when outlining this structure, and this was, for the most part, the trend with the Group 1 structures that were classified the same by all the analysts.

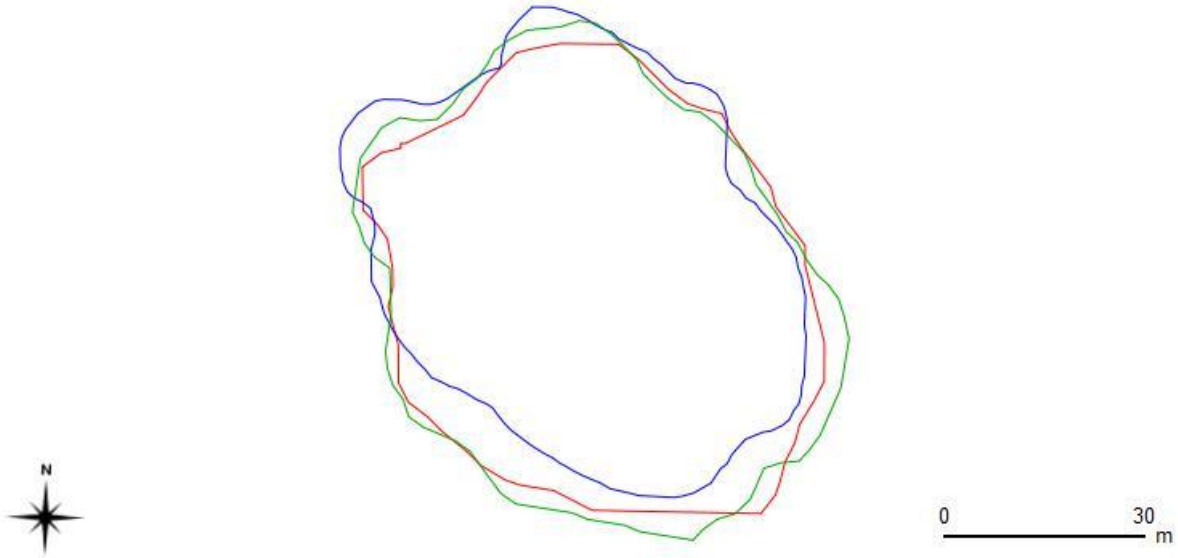


Figure 4.14 QGIS image of each analysts outlines of a Group 1 structure, red is 'T' Group 1, blue is 'K' Group 1 and green is 'M' Group 1.

4.3.2. Classification versus Shape

Several patterns emerged concerning the ways in which each analyst was likely to outline and classify structures. One pattern that became obvious when comparing the outlines of all the analysts was that very often all three would classify the structure as the same group but outlined different shapes. For example, Fig. 4.15 shows a stone-walled structure that was classified as Group 1 by all three analysts and yet the outlines are different. The opposite was also evident, where a site would be classified as different groups and yet the outlines were similar. But the pattern most frequently found was that all three analysts saw different shapes and therefore classified them into different groups. Figure 4.16 shows the outlines of one such structure: 'T' outlined the a large Group 1 site that enclosed all the other circles, 'K' identified three different structures, one of which was classified as Group 2, one as Group 4 and the other as Unknown, and 'M' identified three Unknown sites (Fig. 4.16). An unusual aspect to this site is that one structure was outlined exactly the same by 'K' and 'M'. It is difficult to pick up on this imagery but the site classified as Group 2 by 'K' is in fact overlaid directly on top of one of 'M's Unknown sites (Fig. 4.16). Figure 4.17 shows the same phenomenon where, out of the four structures identified by 'K'

and ‘M’ the third from the left is outlined exactly the same by both analysts, however, in this case both agreed on the site being Group 1.

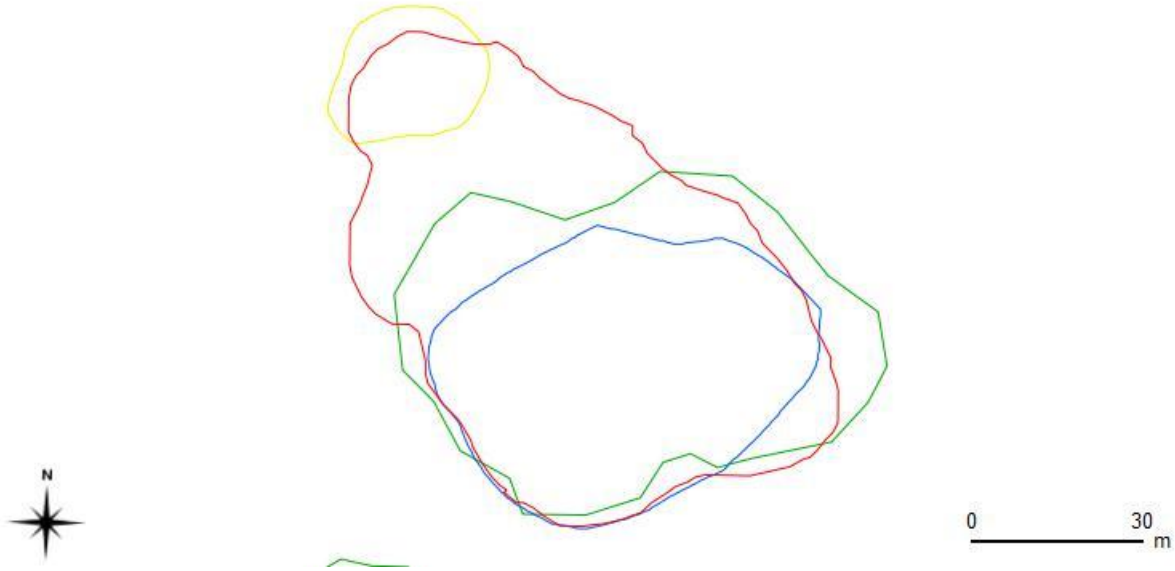


Figure 4.15 QGIS image of each analysts pathways. Red is ‘T’ Group 1, blue is ‘K’ Group 1, green is ‘M’ Group 1 and yellow is ‘K’ Unknown.

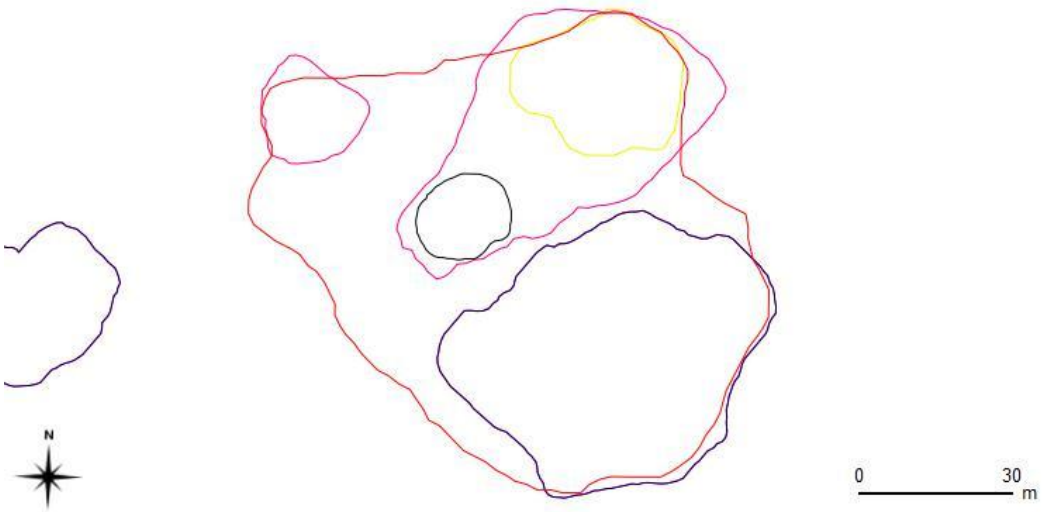


Figure 4.16 QGIS image of the differing pathways of each analyst. Red is ‘T’ Group 1, dark blue is ‘K’ Group 2, black is ‘K’ Group 4, yellow is ‘K’ Unknown and pink is ‘M’ Unknown.

4.3.3. Large versus Small Structures

A common difference between the three analysts pathways involved the division of large stone-walled structures. I found that when confronted with a large site, 'T' was most likely to trace a large outer wall encircling all the stone walling. 'M' was most likely to divide the site up into two or three sections and name them Group 1 or Unknown, while 'K' was most likely to either do the same as 'M' or outline all of the inner enclosures and name each one Group 4. The difference in classification is understandable in this case because when you're looking at a larger site one would naturally classify it as a more complex structure rather than a simpler one. This also accounts for the fact that 'K' identified so many more Group 4 sites than the other analysts, as well as finding many more sites in general. 'K' did not, by any means, do this with every site, but it did become obvious that 'M' and 'K' were much more likely to divide a site up into several different sites while 'T' outlined a large outer wall around all of them and named it as one site (Fig. 4.17).

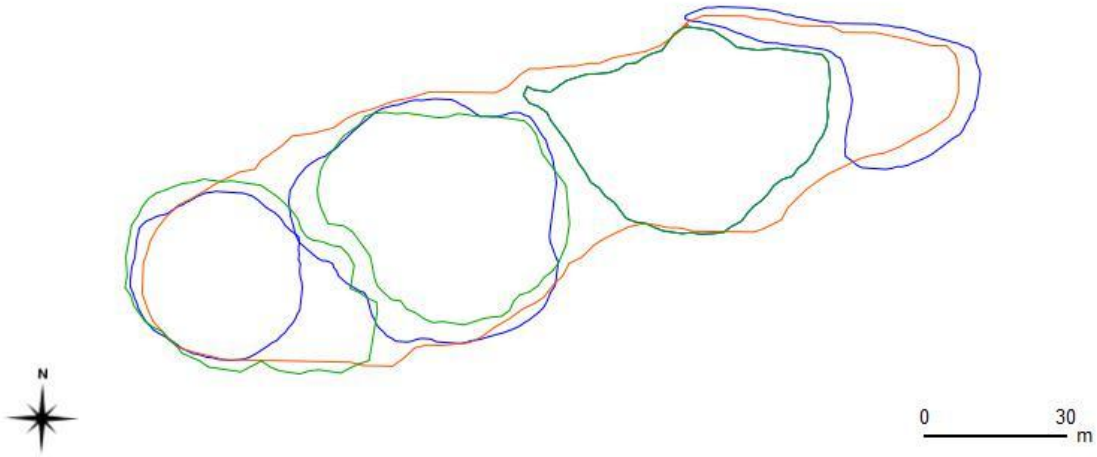


Figure 4.17 QGIS image showing the overlaid pathways of all three analysts. Orange is 'T' Group 3, blue is 'K' Group 1 and green is 'M' Group 1.

Figure 4.17 illustrates this pattern as it shows the way in which the three analysts outlined a particular structure. The orange outline going all the way around was done by 'T' who classified this structure as Group 3, while blue and green circles were done by 'K' and 'M' respectively. 'K' classified all four of their outlines as Group 1's, while 'M' did the same but only outlined three. As shown above in Table 4.1, there was not very much agreement on classification overall, and nor was there agreement on the outlines of the structures which explains why there are such huge discrepancies in the number of stone-walled structures identified by each analyst (Fig. 4.1).

It has become evident that there are statistics that point the same way and those that point in different directions concerning how each group differs between the analysts. These will be discussed in the next chapter and reasons for such variance will be explained.

Chapter 5

Discussion and Conclusion

It was evident from the analysis in Chapter 4 that there was a high rate of inter-analyst variability between the three analysts. The next step was to identify the reasons for such high variability and to find ways in which it can be reduced. There are certain patterns, common to each individual analyst, which would have an effect on the variability between the analysts. Each stone-walled structure group presented different problems and patterns between the analysts. In addition, there is the interesting situation where the statistics don't fully agree with one another. Each group will therefore be discussed individually below.

5.1 Group 1 Structures

Group 1 had the least amount of variability between the analysts, especially concerning the spatial statistics. However, the descriptive statistics show a very different picture. Both 'T' and 'M' found that Group 1 structures were the most common structures in Pam 2 while 'K' found far fewer than the other analysts. This is not because 'K' was not finding the sites but rather because they were classifying them as a different group, this will be discussed further in section 5.6. 'T' and 'M' found very similar numbers of Group 1 sites but when one looks at how many individual sites they actually agreed upon we find that it was only about half of the sites that they each found. There was much more two-way agreement in this group than there was three-way as we can see by the fact that only 19 sites were classified as Group 1 by all three analysts. Therefore, in terms of the classification of individual sites, there was a huge amount of variability. However, there are similarities between the analyst's Group 1 results.

Even though the sample sizes are so different, it became evident that each analyst found roughly the same amount of clustering for this group. All the analysts found Group 1 sites to have the same kind of distribution, as can be seen in the section on spatial statistics, which is important as it shows that even though their sample sizes varied dramatically, all the analysts were finding that Group 1 sites were widely distributed but were also often clustered together. The Convex Hulls for the three analysts for Group 1 showed that the analysts were mostly in agreement in terms of the spatial distribution of the sites.

Overall, when there was agreement on a Group 1 site, the analysts would agree not only on the classification but also on the placement of the stone-walling. This was illustrated by the area

averages which were fairly consistent for all the analysts, and fig. 4.2 shows that to a large extent there was agreement between the analysts in terms of the sizes of Group 1 structures. The visual comparisons of the outlines in QGIS supports this conclusion as it shows that when the analysts agreed on a Group 1 structure they were all seeing the same features in the stone-walling. A reason for this is that out of all the structure types, Group 1 is the simplest and the easiest to identify with a round outer wall and a few inner enclosures (Taylor 1979). It stands to reason that a simpler stone-walled structure with distinct features would be easier to identify and classify than a more complicated one and therefore it is understandable that when there was agreement on these sites, the analysts were all seeing the same site.

5.2 Group 2 Structures

Group 2 sites showed a high amount of agreement between the analysts in some respects and a high amount of variability in others. In terms of the classification, there was in fact quite a high level of agreement. All analysts found that Group 2 was the rarest group in Pam 2 with all three finding the least number of sites in their respective datasets. Out of the 13 Group 2 sites found by 'T', 10 of them were agreed upon by either 'K' or 'M' which shows quite a high level of agreement. There is also little variability between 'K' and 'M' as they agreed on 23 sites out of a total of 27 and 31 sites respectively. Out of all the Group 2 sites identified, between a quarter and a third of them were agreed on by all three analysts. This group showed more three-way agreement than any other group.

However, when it comes to the spatial analysis of Group 2 it was found that the results differed more than was previously evident, and that there was more two-way agreement between 'T' and 'M' than there was with 'K'. Their mean centres were the closest together for this group, their SDE's were similar and their Convex Hulls were far closer together than they were with 'K'. In terms of spatial distribution, 'K' found sites in completely different areas to that of 'T' and 'M' as evidenced by the Convex Hulls and the SDE's. This means that even though there were a large number of sites on which 'K' agreed upon with 'T' and 'M', there were many that were not only classified differently but were located in a completely different area of Pam 2.

The high levels of agreement can be explained by the complexity of the Group 2 typology. The complexity makes for a very distinct group, one that would be very difficult to confuse with another group. The discontinuous c-shaped outer walls and the central inner enclosures are distinct and that is why there was a fair amount of general agreement on the identification and

classification of Group 2 sites. The sites on which the analysts disagreed may have been due to other factors that are discussed in section 5.6 below.

One other aspect of agreement between the analysts was that Group 2 sites were to be found at consistently lower elevations to the other groups. Figure 4.3 shows that Group 2 was the only group to shows significantly different elevation values compared to the other groups. This shows that Group 2 sites were consistently being built on lower ground, probably in river valleys, where they would be closer to a source of water and more fertile land could be found. All the other groups show average elevation values that are generally much higher than those of Group 2.

5.3 Group 3 Structures

This group, along with Group 4, had the most amount of disagreement between the analysts. This can be seen immediately by the number of Group 3 sites identified by each analyst. 'T' classified by far the most with 65 sites, 'K' with 29 and 'M' with only 8. Most of the Group 3 sites identified by 'K' were also identified by 'T', and the same can be said for 'M' as they agree on 5 of the 8 sites. Therefore, the analysts seemed to agree on the general idea of what a Group 3 site ought to look like and yet 'T' found far more than either 'K' or 'M'. What became clear when analysing the database was that most of the sites that 'T' classified as Group 3 had been broken up into several sites by 'K' and 'M' and were often classified as Group 1 or 4 sites.

Overall, the analysts were seeing completely different things when looking at these sites but, nevertheless, common patterns arose. It became evident that 'T' was the most likely to outline a large, complex structure and classify it as Group 3, 'K' was most likely to outline the inner enclosures and classify them as individual Group 4 structures, while 'M' was more likely to divide the larger structures into two or three separate structures and classify them as Group 1. This is largely understandable. If you find a large complex structure you will be more likely to classify it as the complex group, Group 3 (Taylor 1979). However, if you do not identify the outer wall, and all you see are the inner enclosures, it is logical to classify them as Group 4. The same goes for 'M' who was obviously seeing several simple enclosures rather than one complex one and would therefore have classified them all as Group 1. The area data showed that all the analysts were on average classifying Group 3 as having the largest stone-walled structures and this supports my theory that while the largest sites are classified as Group 3, when an analyst sees only a portion of the site they will be less likely to classify it as Group 3.

The spatial statistics showed further variation between the analysts. The NNI's and the SDE's show that 'T's sites are very clustered and are found over a large area while 'K' and 'M' are less clustered and cover a much smaller stretch of land. This is further illustrated by the Convex Hulls that showed that the spatial distribution for Group 3 sites were very different for each analyst. This is not surprising considering the variation in the numbers of sites identified by the analysts with 'T' bound to have a larger distribution as they had by far the most Group 3 sites.

5.4 Group 4 Structures

Group 4 is a group that was not originally a part of Taylors 1979 typology and is characterised by isolated primary circles that may have been satellites of Group 1, 2 or 3 structures (Sadr & Rodier 2012). There was a high rate of variability between the analysts when it came to Group 4 and it would seem that the analysts all thought very differently about what a Group 4 site ought to be. 'T' stuck mostly to the description given by Sadr and Rodier (2012) and classified a Group 4 site only when there was no evidence of an outer wall. They were also consistently much smaller than any other group, were widely distributed and were often clustered together with other Group 4 sites, as can be seen by the area averages Convex Hull and NNI.

'K' identified far more Group 4 sites than either of the other analysts, as stated above, this was mostly because this analyst was in the habit of identifying only the inner enclosures and would often completely miss the outer walls of stone-walled structures that should have, in fact, been classified as a different group. As can been seen from the area averages for 'K', their Group 4 sites were also much larger than those of 'T' or 'M', and some were in fact classified as Group 1 sites by either 'T' or 'M' or both. This shows that not only did 'K' often miss the outer walls of structures, they would also occasionally not identify the inner enclosures of small Group 1 sites and then classify it as a large Group 1 site. This resulted in a dataset of Group 4 sites that were much more numerous and physically much larger than they should be.

'M' had the smallest dataset of Group 4 sites with a total of only 27. But out of that 27, 15 were also classified as Group 4 by 'T', and 25 were classified as such by 'K'. This shows that while 'M' has a similar idea about what a Group 4 site should look like to the other analysts, for some reason they were just not finding them. Most of the Group 4 sites classified by the other analysts were simply not identified as anything by 'M'.

All three analysts found Group 4 sites to be very clustered together as their NNI's were all consistently low, and overall the area averages were the smallest compared to the other groups. Therefore there was a degree of agreement on size and spatial distribution of Group 4 structures.

5.5 Unknown Structures

There is not much of an analysis that can be carried out on this group. Each analyst found different structures difficult to classify and for different reasons. Out of the 61 sites classified as Unknown by any one analyst, only three were classified as such by more than one analyst. One pattern that I identified was that quite a few of the sites classified as Unknown by any one of the analysts were classified as Group 4 by one of the others, and often not found at all by the other. Which analyst took which role depended heavily on the site, but it was often either 'T' or 'K' who identified the site as either Group 4 or Unknown, and 'M' who would miss the site entirely.

Quite often sites that had been classified as Unknown by any one analyst would be classified as completely different groups by the others, showing that the sites were at least being found. However, the fact that analysts were classifying sites as Unknown meant that there must have been factors making the structures unclear or difficult to identify, this is supported by the fact that quite often one of the analysts wouldn't find the sites at all. This could mean that in these cases one or more of the analysts was guessing at what group the structure would fall into rather than making an informed decision according to the typology.

5.6 General Observations

There are several reasons that can explain some of the inter-analyst variability that is evident in data belonging to the three analysts. These include access to good quality images in Google Earth, the use of different seasons in Google Earth, the amount of training received by the three analysts, and the typology used to classify the structures.

5.6.1 Google Earth Imagery

Google Earth has a tool which allows a person to view the same area in different seasons and in different years. This allowed the analysts to flick through different imagery of the same stone-walled structure to find the image in which the walls are clearest. One may think that the most obvious choice would be to use the most recent imagery as it would have the best resolution, but I found that the clarity of the stone-walled structure depended more on the season in which it was photographed. I also found that different seasons suited different structures the best,

sometimes a dry season would show the walls more clearly as the vegetation would be low and brown. However, in a rocky area the walls would blend in with the other rocks and would therefore be difficult to identify. In these cases I found that the walls were easier to see in a greener season as the vegetation grows with a darker colour along the walls compared with the surrounding grass, making them easily identifiable. The general idea was to flick through all the different images to find the one that made the walls as clear as possible. However, this opens up the possibility of the analysts using different imagery to view the same sites and possibly seeing different things.

When an analyst changes the year of GE imagery, the georeferencing changes as well. This results in outlines that identify the same structure and yet they do not line up with each other, showing without a doubt that the analysts were using different imagery to view the structures. Figure 5.1 shows an example of this where a group of structures were outlined by all the analysts but confusion reigns to the extent that it is difficult to identify which outlines are supposed to identify the same stone walls. In most cases the analysts used the same GE imagery as their outlines almost always line up with each other. However, when the outlines don't line up it is most often 'K' who used different imagery to the other analysts. It is evident that while 'T' and 'M' would flick through the different seasons for each site to find the optimum clarity, 'K' would often just settle for the most recent imagery with the belief that the best resolution would provide the clearest image. Therefore, if ways in which the analysts were using Google Earth differed, it is not surprising that their subsequent classifications of structures differed so much.

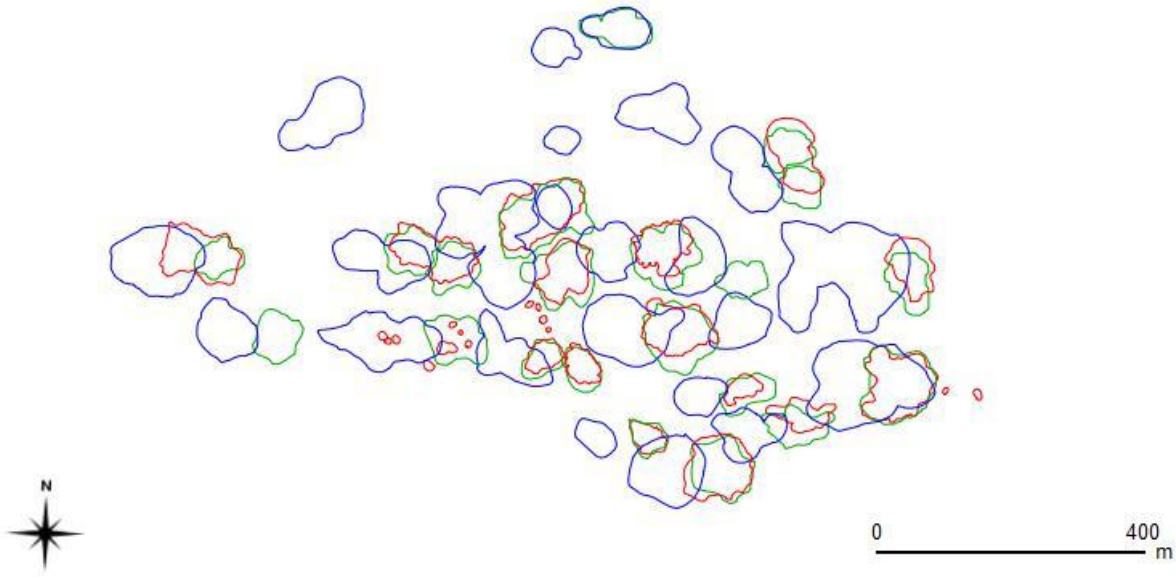


Figure 5.1 QGIS image showing how each analyst's outlines have shifted due to the use of different GE imagery. 'T' is red, 'K' is blue and 'M' is green.

5.6.2 Training

The training that each analyst received was not extensive but was consistent for all three. The training consisted of being given the article written by Sadr and Rodier (2012) which described the area that we would be focusing on as well as the typology that we would be using. We were then given a short tutorial on how to use the different tools in Google Earth and QGIS which completed the training. It is possible that with a more extensive training course on the classification of stone-walled structures, there would have been more agreement on the classification of the structures. Perhaps a workshop with all the analysts together would be sufficient, where they go through the typology together and classify different sites together, thus ensuring that they all have the same idea as to what each group looks like. This would hopefully result in less inter-analyst variability with analysts doing the classifications at the same time and having access to imagery of the same resolution.

One more thing to consider with respect to the analysts is the matter of motives. For example, when I was doing the digital surveying and classifications it was with the final aim of my honours project in mind, and I was therefore motivated to be as thorough and careful as possible. However, the other analysts may have had other motives, for example they may have been paid research assistants doing the work in between their other studies. This may have created some of the variation in our analyses and may provide some explanation as to why 'M' found far fewer sites than the other analysts. Perhaps it is as simple as 'M' not looking as carefully as the others.

5.6.3 Typology

It is also possible that the typology is not as clear as it could be. The fact that there is the least variability in Group 2 shows that the typology for Group 2 is sufficiently clear to make the structures easily identifiable. However, if you compare Groups 1 and 3 you will see that there is a lot of overlap between the groups and many of the structures in Pam 2 could have been identified as either Group 1 or 3. This would go a long way to explaining why there is so much variation in those two groups. The fact that Group 4 is not included in the original typology may have resulted in a certain amount of confusion for some of the analysts and this could be a reason for the variability in that group. Is it a problem with the typology or with the training? If the typology was clearer would the need for more training fall away or visa versa? These are valuable questions that need further investigation.

5.7 Conclusion

The classification of stone-walled structures is an important aspect of the field of archaeology in southern Africa. It tells us much about the people who lived there and the way they structured their lives (Huffman 2007: 23). More than that, the classification of the structures into different groups enables researchers to make certain interpretations about the ethnic origins and movements of the people associated with the different groups. Unfortunately, high levels of inter-analyst variability in this field of study is a problem that affects any further interpretations made from the data. Archaeology is not the only field in which inter-analyst variability is a concern; it is an important factor that influences any research that is based on information gathered by several different analysts. There are many aspects that will have an impact on the levels of inter-analyst variability including the amount of training received by each analyst and the past experience of each analyst in the field in which they are working. Concerning the present study, the quality of the GE imagery used by each analyst would have had an impact on each analyst's classifications as well as the clarity of the typology which contained some overlap between the different groups. It can also be argued that three analysts are not sufficient for a study such as this and that it is possible that with a few more analysts the results will show, on average, less inter-analyst variability. If the analysts take part in more extensive training sessions and they are more careful in the outlining and classification of the sites while making use of the technology available to them, then it is possible that the high levels of inter-analyst variability as seen in this study may be reduced.

Reference List

Abe, Y., Marean, C.W., Nilssen, P.J., Assefa, Z. & Stone, E.C. 2002. The Analysis of Cutmarks on Archaeofauna: A Review and Critique of Quantification Procedures, and a New Image-Analysis GIS Approach. *American Antiquity*. 67(4): 643-663.

AGIS. 2007. Agricultural Geo-Referenced Information System accessed from www.agis.agric.za in January 2013.

Binford, L.R. 1977. Introduction. In *For Theory Building in Archaeology*. Edited by Lewis R. Binford, pp. 1-10. Academic Press, New York.

Birss, C. & Green, N. 2007. Suikerbosrand Nature Reserve and Extension Ecological Management Plan. Gauteng Provincial Government Department of Agriculture. Conservation and Environment, Johannesburg.

Boeyens, J.C.A. 2003. The Late Iron Age sequence in the Marico and early Tswana history. *The South African Archaeological Bulletin* 58 (178), 63-78.

Clark, P.J. & Evans, F.C. 1954. Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations. *Ecology*. 35 (4): 445-453.

Hall, S. 1998. A Consideration of Gender Relations in the Late Iron Age 'Sotho' Sequence of the Western Highveld, South Africa. In: Kent, S. (ed.) *Gender in Prehistory*: 235-258. London: Altamira Press.

Huffman, T.N. 1986. Archaeological Evidence and Conventional Explanations of Southern Bantu Settlement Patterns. *Journal of the International African Institute*. 56(3): 280-298.

Huffman, T.N. 2007. Handbook of the Iron Age: The Archaeology of Pre-Colonial Farming Societies in Southern Africa. KZN Press, Pietermaritzburg.

Kalton, G. & Stowell, R. 1979. A study of Coder Variability. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*. 28(3): 276-289.

Lyman, R.L. & VanPool, T.L. 2009. Metric Data in Archaeology: A Study of Intra-Analyst and Inter-Analyst Variation. *American Antiquity*. 74(3): 485-504.

MacQuilkan, P. & Sadr, K. 2010. Remote Sensing of Archaeological Ruins: A Comparison. *PositionIT* Nov/Dec 2010: 50-51.

Maggs, T.O.'C. 1976. Iron Age Communities of the Southern Highveld. Council of the Natal Museum, Pietermaritzburg.

Mason, R.J. 1968. Transvaal and Natal Iron Age settlement revealed by aerial photography and excavation. African Studies 27 (4), 167-180.

Milne, M.J. & Adler, R.W. 1999. Exploring the Reliability of the social and Environmental Disclosures Content Analysis. *Accounting, Auditing and Accountability Journal*. 12(2): 237-256.

Oldenburg, B.F., Sallis, J.F., Ffrench, M.L. & Owen, N. 1999. Health Promotion Research and the Diffusion and Institutionalization of Interventions. *Health Education Research*. 14(1): 121-130.

Praetzellis, A. 2011. *Death by Theory: A Tale of Mystery and Archaeological Theory*. pp. 43. AltaMira Press, United Kingdom.

Raab, L.M. & Goodyear, A.C. 1984. Middle-Range Theory in Archaeology: A Critical Review of Origins and Applications. *American Antiquity*. 49(2): 255-268.

Sadr, K. & Rodier, X. 2012. Google Earth, GIS and stone-walled structures in southern Gauteng, South Africa. Journal of Archaeological Science. 39: 1034-1042.

Seddon, J.D. 1968. An Aerial Survey of Settlement and Living Patterns in the Transvaal Iron Age: Preliminary Report. *African Studies*. 27(4): 189-194.

Smith, S.C. & Bruce, C.W. 2008. *CrimeStat III: User Workbook*. The National Institute of Justice: Washington, DC.

Taylor, M. 1979. Late Iron Age Settlements on the Northern Edge of the Vrededort Dome. M.A. Dissertation, University of the Witwatersrand, Johannesburg.

Thomas, J. 2004. *Archaeology and Modernity*. pp 72-76. Routledge, London.

Turnitin Originality Report

Alex Schoeman 09300863 Honours Research Project.docx by Null Null

From quicksubmit_09300863_2013 (quicksubmit_09300863_2013_WITSE_2013)

Similarity Index	Similarity by Source	
2%	Internet Sources:	1%
	Publications:	N/A
	Student Papers:	1%

Processed on 23-May-2013 11:34 AM
SAST
ID: 332191836
Word Count: 15800

sources:

1 1% match (student papers from 18-Oct-2012)
Class: Project
Assignment:
Paper ID: 276797009

2 < 1% match (Internet from 15-Jan-2013)
http://wiredspace.wits.ac.za/bitstream/handle/10539/7061/Testing%20and%20validation%20of%20coal%20combustion%20prediction%20indices.pdf?sequence=1

3 < 1% match (Internet from 14-Jan-2013)
http://etd.lib.nsysu.edu.tw/ETD-db/ETD-search-c/view_etd?URN=etd-0727111-214121

4 < 1% match (Internet from 18-Aug-2003)
http://www.me.binghamton.edu/miles/theses/ethan%20thesis.pdf

5 < 1% match (Internet from 18-Jan-2013)
http://www.oregonscca.com/content/view/310/258/

6 < 1% match (student papers from 30-Oct-2011)
Submitted to University of Queensland on 2011-10-30

7 < 1% match (student papers from 03-Nov-2008)
Submitted to University of Pretoria on 2008-11-03

8 < 1% match (Internet from 20-Jan-2013)
http://www.coursehero.com/file/2286919/weikelrossr200508phd/

9 < 1% match (Internet from 24-Feb-2010)
http://proteus.brown.edu/gisandremotesensing/admin/download.html?attachid=6323391&view=1

10 < 1% match (Internet from 15-Jun-2009)
http://www.john-libbey-eurotext.fr/e-docs/00/04/00/E3/vers_alt/VersionPDF.pdf

11 < 1% match (Internet from 14-Sep-2012)
http://www.cupw.ca/multimedia/website/publication/English/PDF/2010/persp-may-2010_en.pdf

12 < 1% match (Internet from 16-Dec-2007)
http://waas.stanford.edu/~wwu/papers/gps/PDF/Thesis/SharonHouckThesis01.pdf

13 < 1% match (Internet from 30-Apr-2008)
http://www.tnrsp.com/empv11.pdf

14 < 1% match (Internet from 08-May-2012)

http://www.berr.gov.uk/files/file28947.pdf

15 < 1% match (Internet from 29-Dec-2007)
http://www.swanvalleyforest.ca/pdfs/1999-2000%20FML3%20Annual%20Report.pdf

16 < 1% match (Internet from 14-Sep-2010)
http://www.otters.co.za/VDOME%20CONSERVATION%20&%20MGMNT%20PLAN%202004FINAL.pdf

17 < 1% match ()
http://www.beckmancoulter.com/literature/ClinDiag/9282%20dat%209%20METQ%201002.pdf