A methodology for landscape characterisation based on GIS and spatially constrained multivariate analysis.

iLaria Marengo

Submitted to the School of Biological and Environmental Science

University of Stirling

August 2010

For the degree of Doctor of Philosophy

Supervisor: Prof. Ian Simpson

UNIVERSITY OF STIRLING

University of Stirling, FK9 4LA, Scotland

STATEMENT OF ORIGINALITY

I hereby confirm that this is an original piece of work conducted independently by the undersigned and all work contained herein has not been submitted for any other degree.
All research material has been duly acknowledged and cited.

Signature of Candidate

iLaria Marengo

Date: 25th August 2010








ABSTRACT

Landscape is about the relationship between people and place and can be defined and interpreted in several ways according to the objectivist or subjectivist point of view. The former considers the landscape as the area where ecological patterns and processes might be identified and therefore measured, while the latter supports the theory that landscape is in the beholders' eyes thus there are many landscapes according to human preferences. A possible compromise between these two different ideas of landscape might have been reached by the European Landscape Commission (ELC) that in 2000 stated that "landscape is an area as perceived by people whose character is the result of natural and human actions and interactions".

In the 70s the reason for studying the landscape was because of the necessity of attributing a value to it; nowadays the motivations behind managing, conserving and enhancing the landscape is because the landscape is the place where people belong to and, consciously or not, recognise themselves. In addition, people identify different landscapes on the basis of the particular combinations of the elements in the landscape. As a consequence a landscape can be distinguished from another on the basis of its character which, according to the Landscape Character Assessment (LCA) guidance for England and Scotland (C. Swanwick and Land Use Consultant, 2002), is defined as "a distinct, recognisable and consistent pattern of elements in the landscape that makes one landscape different from the other rather than better or worse".

This definition was the starting point of a PhD research project aimed at developing and implementing a methodology able to identify and quantify the character of the Scottish landscape through the application of GIS and statistics. The reason for doing this research was to provide the landscape architects, who are the main end-users, with a tool that could help them to overcome the main weaknesses of the current process of LCA and define the landscape character types in a more consistent, objective, and scientifically robust way.

The GIS approach considers the landscape as formed by physical and cultural/historical elements. No aesthetic and perceptual attributes of the landscape are taken into account, despite an attempt to quantify enclosure, diversity and dominance conducted for the Cairngorms National Park (CNP).

The research selected eleven areas of study distributed across Scotland so that various and different landscapes could be represented. The data collected for the analysis came in part from official sources: BGS (geology bedrock and superficial),

CEH (land-cover 2000), RCAHMS (historic land-use assessment), OS (DEM and Mastermap), and in part were derived from official data, for example landforms and settlement types were respectively obtained from DEM (calculation of the topographic positioning index) and OS Mastermap (calculation of the density of buildings).

Overall, the data about the eleven areas of study presented the same characteristics, and it was both multivariate and geographical; in other words the landscape of the sample areas was on average described by 50 variables and each one was related to a geographic reference (easting and northing coordinates).

One of the objectives of the analysis was to identify the spatial patterns formed by the landscape elements (the variables) by taking into account the influence of the spatial location. Space was recognised as important for the analysis on the basis of what is stated by the first law of geography: everything is related to everything else but near things are more related than the distant ones (W. Tobler 1970). Fundamentally the idea that supports this statement is that proximity facilitates the relationships between objects. However, similarity also contributes to the association between landscape elements, and this is taken into account with the assumption of the presence of spatial autocorrelation amongst the data. The null hypothesis is that the landscape elements are randomly distributed, while the alternative is that they are spatially correlated. Thus, if two geographic objects that are close to each other and similar (both show similar positive or negative values), then they are spatially autocorrelated and form detectable patterns.

As mentioned the data were also multivariate, thus the analysis required a method of calculation able to deal simultaneously with multivariate and spatial autocorrelation issues. After several attempts a spatially constrained multivariate technique was adopted and was revealed to be the most suitable tool for the data analysis. This technique allowed the identification of the spatial patterns (landscape characters) within the dataset by maximising the product between the variance and the spatial autocorrelation of the variables. Therefore, contrary to a standard principal component analysis (PCA), which maximises only the variance, the spatially constrained multivariate analysis (MULTISPATI-PCA) includes the effect of the spatial autocorrelation. In other words, the geographic location of the elements in the landscape, as well as their distribution, contributes to the character of the landscape.

Once designed, the GIS/MULTISPATI-based methodology for landscape characterisation was tried for each area of study, and the results showed that with the methodology it was possible to detect the spatial structure of the data and that each single pattern corresponded to a distinct landscape.

At the end of the process of landscape characterisation several tests simulating different situations, for example the effect of the boundaries and the change in scale, number of data input, and size of the units of analysis, were performed to verify the robustness of the methodology. The outcomes revealed that the identification of strong spatial patterns was not affected by the simulated situations.

Finally the outcomes and the methodology were presented to a selected group of end-users, largely comprised of landscape architects but including also members of the Cairngorms National Park, Historic Scotland and the RCAHMS, who provided useful comments and feedback. The maps of landscape character types were considered overall plausible and it was acknowledged that the methodology achieved a more quantitative and consistent landscape characterisation.

LIST OF ACRONYMS:

BGS: British Geological Survey
CEH: Centre for ecology and hydrology
CNP: Cairngorms National Park
ELC: European Landscape Convention
GIS: Geographic Information Systems
LCA: Landscape Character Assessment
OS: Ordnance Survey
PCA: Principal Component Analysis
RCAHMS: Royal Commission on the Ancient and Historical Monuments of Scotland
SNH: Scottish Natural Heritage

COPYRIGHT AKNOWLEDGMENT:

AKNOWLEDGEMENTS

This thesis has been completed thanks to the help and support of several kind people around me. I would like to express my gratitude to James Fenton and Caroline Read, who represented my sponsor the Scottish Natural Heritage, for the encouragement and helpful comments and suggestions they provided. Thank you for the time you dedicated at the numerous meetings and for sharing your ideas.

I am equally grateful and thankful to Stirling University for providing the structure and services I needed to complete the PhD. The main thanks go to my supervisor, Prof. Ian Simpson whose comments helped to finalise and refine the content of the thesis. My thanks go also to Dr. Sandy Winterbottom who gave me this opportunity and helped in the initial stages of the research. Poster session and presentation given at the School of Biological and Environmental Science offered the opportunity to show the stages of this research. I would like to thank the lecturers who made constructive comments in particular Luc Bussière and Mario Vallejo-Marin.

The contribution of MULTISPATI to the realisation of this research and to the development and implementation of the methodology was extraordinary and my total appreciation is addressed to Stéphane Dray, researcher at Biometry and Evolutionary Biology Lab of the University of Lyon 1, France. Stéphane demonstrated high degree of professionalism and kindness which I cannot compare to anyone. Sincere thanks to Ollier, Thibau, Chessel and Thioulouse, all researchers whose studies provided essential and significant references for this PhD.

Highly valuable was the support, the ideas and the constructive feedback provided by Matthew Hawkins, senior manager at the Cairngorms National Park and Piers Dixon operations manager at the RCAHMS. Thanks for listening and following this research throughout its three years. It was a positive and interesting experience to meet you since you provided an excellent example of what being open-minded and knowledgeable means.

I am very grateful to all the data providers, in particular I would like to thank people who make EDINA Digimap such a useful and excellent service. Many thanks for their assistance and for supplying the huge amount of maps by Ordnance Survey.

Thanks to Bill Jamieson for helping me with the print of the thesis and Scott Jackson for the impeccable IT support provided throughout the three years of this research.

A sincere thank goes to Prof. Carlo Socco for his ideas and his hard but first class teaching that I benefited from in the period spent in Turin.

I would like to thank two special people who were the closest to me in the last three years and made this “journey” a happier experience. Kate and Berenice, thank you so much for your friendship and the precious time spent in listening, helping and encouraging me throughout this time.

Kate, thanks for the opportunity you gave me to know West Moss-side, to meet so many interesting people, to learn that life is an endless series of projects and to grow…not only the vegetables in the garden. I spent three unforgettable and intense years.

Berenice, thanks for being such a good and patient friend who likes coming with me up to the hills or along the sea, exploring places in long treks or short breaks. You are never short of curiosity and I like when you want to see what is around the corner. Thanks for being tuned on “radio iLa” and for reminding me that Scotland is incredibly stunning not only for its natural heritage but also for its historical past.

Finally, grazie mille!! to my family who supported me with immeasurable love at any time of my life. To them, this thesis is dedicated.

TABLE OF CONTENT

1. The landscape and its character between philosophy and practical applications.

1.1 The importance of landscape and its character 2

1.2 The objectivist and subjectivist paradigm in the landscape analysis 5

1.3 Landscape as a whole 11

1.4 From landscape evaluation to landscape character assessment 15

1.5 The role of history as a contributor to the character of the landscape 23

1.6 European approaches to landscape character assessment 25

1.7 Examples of landscape character assessment in the rest of the world 26

1.8 Principal objectives of the PhD 30

2. Structure and concepts that support the Landscape Character Assessment.

2.1 The conceptual model of LCA 34

2.2 LCA programme: lights and shadows 62

3. The steps towards the definition of a quantitative and GIS-based methodology for landscape characterisation.

3.1 The need to develop a new, more consistent, approach to LCA 66

3.2 Aesthetic aspects of the landscape and GIS: an analytical challenge 75

3.3 On the characterisation and classification of diversity and dominance 77

3.4 Attempts towards a GIS-based methodology for landscape characterisation 81

4. From raw data to information.

4.1 The importance of the data in GIS-based analyses 86

4.2 Understanding the data 89

4.3 On the derivation and classification of landforms 93

4.4 On the definition of settlement types 105

4.5 On the interpretation and use of the Historic Land-use Assessment map 122

5. "Everything is related to everything else, but near things are more related than distant ones".

5.1 The role of statistics within the GIS-based methodology 132

5.2 An introduction to spatial statistics 135

5.3 The link between spatial autocorrelation and landscape characterisation 143

5.4 The areas of study and the sampling zones 147

6. MULTISPATI: a spatially constrained multivariate analysis for landscape characterisation.

6.1 A methodology for landscape characterisation 168

6.2 The approach with standard multivariate PCA analysis 172

6.3 The approach with MULTISPATI-PCA 182

6.6 Cluster analysis: the bridge between MULTISPATI and GIS 191

7. Understanding the results and testing the methodology for robustness.

7.1 Landscape characterisation through GIS and MULTISPATI "eyes" 200

7.2 How to read a GIS/MULTISPATI map? 222

7.2.1 North Uist and South Harris, an example of similarity between LCA and GIS/MULTISPATI landscape maps 222

7.2.2 The Southeast: an example of greater discrepancy between LCA and GIS/MULTISPATI landscape maps 233

7.3 Testing the GIS/MULTISPATI-based methodology for robustness 244

7.4 What happens if new polygons are generated? 245

7.4.1 The Cairngorms National Park 245

7.4.2 The Central area 251

7.5 What happens if two neighbouring areas are analysed? 260

7.6 What happens if the number of polygons changes? 272
(if the scale is modified)?

8. Listening to the end users and overall conclusions.

8.1 Learning from the feedback of the potential end users 286

8.2 The aims achieved after three years 292

8.3 Those weak points that can be improved 301

8.4 Looking towards the future 304

References 306

CHAPTER 1

The landscape and its character between philosophy and practical applications

The Blackwater reservoir with Buchaille Etive Mor and the Aonach ridge on the background (I.Marengo).

1.1 The importance of landscape and its character

Landscape is about the relationship between people and place. It provides the setting for the day to day lives. Landscape is an urban area and its surroundings, as much as a mountain range or an expanse of lowland plain, as showed in figure 1.1.

a)  b) 

c)

Figure 1.1. These three pictures depict a) Scottish highlands b) Stirling and Wallace monument and c) fields of barley in Aberdeenshire and they are all a form of landscape. Usually and mistakenly landscape is associated with wild and remote areas but in fact landscape is the place where people live everyday (© I.Marengo).

A landscape is made up of different components, both natural and cultural, and how these interact together and are perceived by people. On this basis it appears that people's perceptions turn land into the concept of landscape, and perception is not only visual but also related to the other four senses, the feelings, the memories and the associations that they evoke (Swanwick and Land Use Consultants, 2002).

Moreover behind its appearance a landscape is a source of livelihood. From a landscape we can understand much of the history and economic conditions of a society, as if landscape is a kind of genetic code of a geographic area. With this idea in mind landscape becomes important because it is a fundamental element of individuals and communities, and a well looked after landscape is an indication of civilisation (Berengo and Di Maio, 2008).

The idea of the importance of landscape and the need to recognise, protect, manage and renew it, because landscape is of great value for our life, has become more widespread in recent years. However, the majority of people are not aware of the landscape and its importance until a noticeable change, which disturbs or modifies the landscape, occurs. Thus in order to make society more aware of landscape 19 member states of the Council of Europe on the 20th October 2000 in Florence signed a convention[1] which represented a revolution because the landscape moved from being considered a pretty postcard to being officially "an area as perceived by people whose character is the result of natural and human actions and interactions" (Berengo and Di Maio, 2008).

Briefly each state that signed the European Landscape Convention (ELC) recognises that landscape is a basic element necessary for human life, similarly to the air that we breathe. The importance of landscape is in the concept of "sense of place", and namely in the fact that people, consciously or not, recognise themselves with an area rather than with another one. What people perceive as a landscape is basically what makes it different from the surrounding areas, and precisely the character, a unique and distinct

[1] A convention is an agreement between two or more states which decide to tackle a certain important issues or problems in the same way. The states that are part of a convention are then obliged to create laws based on the principles of the convention itself. 19 member states signed the convention in 2000, but the number grew throughout the years and at the 8/4/2010 the agreement has entered to force for 30 states.

combination of elements. However, making people more conscious of the landscape and of its importance as a natural resource and repository for cultural values and memories (Ahern et al., 2005) is a first step on which local authorities, landscape managers and conservationists should focus. Fundamentally by acknowledging the character of the landscape and by assessing it, local communities are empowered in activities such as planning zones, site reviews, establishing partnerships and having ideas for the landscape conservation (Ahern et al., 2005).

In addition, as clarified by the ELC, landscape is not only an area of special values, such as national parks or historical sites, but also an area of everyday life, the most "normal" and ordinary place (Berengo and Di Maio, 2008). On this basis all landscapes are equally important since their character reflects the individuals and the communities that live there. Thus the ELC adopted three principles to follow when tackling the landscape: protect what is a world heritage and of great value; manage the everyday landscape from a perspective of sustainable development; plan anew those areas where there is need for improvement and reorganisation (Berengo and Di Maio, 2008).

Those who are involved in making decisions about the landscape and in setting policy in planning and environmental management, restoration, conservation and enhancement have a useful tool, called Landscape Character Assessment (LCA), that allow them to know which character in the landscape is present, distinct/unique and threatened. Decisions should be taken so that changes do not revolutionise the local character in such a radical way that people no longer recognise themselves in their landscape. A substantial change in the original character, which occurs without consultation, is likely to make people lose part of their identity (Swanwick and Land Use Consultants, 2002). For instance the massive quantity of oil spill which has been spreading on some of the US coasts in the Gulf of Mexico for months is causing not only an environmental and economic disaster but also a psychological wound for the people living in that geographic area. The oil spill is now depriving people of their familiar sceneries and places that are changed and hardly recognisable.

The ELC convention and the LCA programme were not intended to halt the planning and development strategies of an area, rather to make it more sustainable and in favour of everyone. A landscape in its nature is dynamic and in evolution and will continue to be so, but there are “durable” landscapes that resist a change for long time and “weaker” landscapes that are easily threatened by human activities. The aim of both ELC and LCA is to make people aware of these realities and allow them to take the most suitable decisions about the way their natural and cultural resources should be managed, conserved and enhanced.

A question might rise at this point; on what basis does a local authority decides to protect, manage or plan a landscape? First of all it is essential to identify the number of landscapes that make up an area and their characteristics. Only the analysis of the association of the elements that form a landscape, namely geology, landforms, land-cover, historic remains and land-uses, settlement and field patterns, allows the local authority to have a better and comprehensive insight of the landscape of its area and subsequently understand what the best solution is for it. The landscape character assessment is the programme that in the UK and in other countries has been adopted in order to first identify the character of the landscape and second make judgements and decide about plans, strategies and policies that involve the landscape.

1.2 The objectivist and subjectivist paradigm in the landscape analysis

Scientific research usually starts with framing the subject, identifying the issue(s), stating the aim(s), listing the objectives, and formulating preliminary hypothesis. This study can be framed in the bullets points below:

- **TOPIC**: landscape characterisation.

- **ISSUE**: the current LCA lacks of scientific and quantitative basis.

- **AIM**: to develop and implement a methodology that enables landscape architects to carry a consistent, objective, easily updatable and flexible landscape characterisation.

- **OBJECTIVES**: to identify the spatial structures/patterns within the landscape; these will constitute and determine the landscape characters.

- **RESEARCH QUESTIONS**: “Can the GIS analytical capacities be applied successfully to identify the character of the landscape and provide a support tool for the landscape architects and practitioners?” “Can a computer based device read layers of maps, whose data is numeric, and detect different landscapes without seeing, feeling, perceiving them?”

From the definition of landscape provided by the ELC it is clear that landscape is the interaction between human activities and natural/physical elements as perceived by people. From the literature it emerges that landscape can be either analysed scientifically or experienced emotionally; in fact landscape researchers could be divided into two categories: those who defend a more reductionist and quantitative-objective approach to landscape characterisation and those who use artistic and psychological descriptors (see figure 1.2). The former look at the landscape with more scientific eyes the latter turn their attention to the aesthetics of the landscape (Pastor et al, 2007). The research focus of this thesis belongs to the first group.

Science and aesthetics recall the concepts of objectivity and subjectivity that gave birth to a long standing debate about their role in dealing with the landscape. According to Lothian (1999) the objectivist and subjectivist terms refer to the physical aspects of the landscape and to those that are product of the mind/eye of the beholder (Lothian, 1999).

OBJECTIVIST/PHYSICAL PARADIGM	SUBJECTIVIST/PHYCHOLOGICAL PARADIGM
- Landscape quality is an intrinsic physical attribute	- Landscape quality is in the eyes of the beholder
- Assessments are field based	- Surrogates (photos) are used for assessments
- Empirical	- Experimental
- Differentiate landscape quality on the basis of implicit assumptions	- Differentiate landscape quality on the basis of human preferences
- relatively easy and inexpensive to undertake	- Difficult and expensive to undertake

Figure 1.2. Summary of the main differences between the objectivist and subjectivist paradigms (source: Lothian, 1999).

Lothian observes that the objectivist paradigm, which can be summarised as viewing the landscape and its special quality in the physical scene in front of one's eyes, is not objective as it claims. Establishing classification based for example on high/medium/low or numerical scales or making assumptions, e.g. mountains and rivers have a higher quality than a plain, are subjective and personal preferences. Whereas, the subjective paradigm is paradoxically objective as it measures the preference of a community without the influence of the researcher's personal biases, although these may occur in framing the questionnaire and in the evaluation of the results (Lothian, 1999).

Similarly the LCA guidance points out that surveying the elements that make up a landscape, mapping and describing landscape types, which many might consider to be wholly objective, can still involve subjective judgements. For example a surveyor could record that there are 20 hectares of woodland in an area and this assessment would be quantitative and an objective fact. However, if the surveyor records woodland as a key characteristic of the area, necessarily he introduces an element of subjectivity into characterisation (Swanwick and Land Use Consultants, 2002).

The debate about the way landscape is conceptualised is probably endless but from the literature emerges that there is a clear division between those who conceive the landscape as an object and a "material thing" and those that consider the landscape as a "form of visualisation" of product of "our senses". The landscape ecologists are almost unanimously interested in the study of spatial variation in landscapes at a variety of scales and in general are concordant in defining landscape as an area within which ecological patterns and processes might be identified.

On the opposite, amongst those who consider the landscape in visual and qualitative terms, a greater variety of point of views is registered. Cosgrove (1984) suggests that landscape is not merely the world we see but it is a construction and a composition of that world. In other terms landscape is a way of seeing the worlds. Abrahamsson (1999) for example uses the term "inner landscape" to address the idea that landscape derives from our way of attributing a meaning to things and features around us by interpreting them and considering them as symbolic resources. Consequently, when people are asked to clarify and describe "what is the form of the landscape? Then the replies can be varied and dissimilar. People can understand what is meant by man-made and natural elements of the landscape such as "trees", "hill", "river", but they have trouble in quantifying them. For example, a "steep slope", "narrow valley" or "rolling hills" can have different interpretations. More complex definitions of concepts like "naturalness" or "wilderness" occur because people will tend to represent them mentally with different meanings and symbols. In other words, people "charge" the surrounding space with signs which reflect their cultural identities. Similarly, Palang et al. (2000) and Keisteri (1990) outlined the influence of the cultural background on the way people perceive the landscape and on the fact that a same physical object acquires a different meaning and a different symbol according to the cultural group. Ingold (2000) and Lorimer and Lund (2003) shift the attention on different senses and state that landscape is part of us; it is felt. According to these authors the landscape is felt through the action of walking and it is incorporated into the bodily experience.

The conclusion that can be drawn is that the uptake of the different meanings of landscape depends on disciplinary and professional affiliations as well as on research goals (Macpherson and Minca, 2006). Brabyn (2005) recognises that landscapes are both physical realities and social and cultural constructs. Therefore he is convinced that the separation between “material” and “cognitive” definition of landscape has led to numerous contrasts between disciplines and schools and has resulted in research that is largely polarised either towards the observer or towards the physical landscapes. The author indicates that people’s perception of the landscape should be taken into account while the characterisation and classification processes are still in phase to be defined. However, he agrees that “people’s minds” are significantly more difficult to capture, than the physical landscape components.

Understanding how people perceive the landscape and how they conceptualise terms such as “naturalness”, “wilderness” or “harmony” is a time consuming approach and it can lead to questions relating to the types of people who should be considered, the sample size and the nature of their perception, since it is known that cultural background affect the way of perceiving the landscape. The difficulty in providing a realistic representation of space that corresponds with human conceptualisations is widely recognised. Thomson and Bera’ (2007) show that knowledge can be modelled by an ontology which explicitly states how relevant concepts and their constituting objects relate to each other and manifest themselves in reality. The two authors developed a top-down model capable of representing the knowledge of the reality as acquired from people through a process of inner conceptualisation of the reality (see figure 1.3).

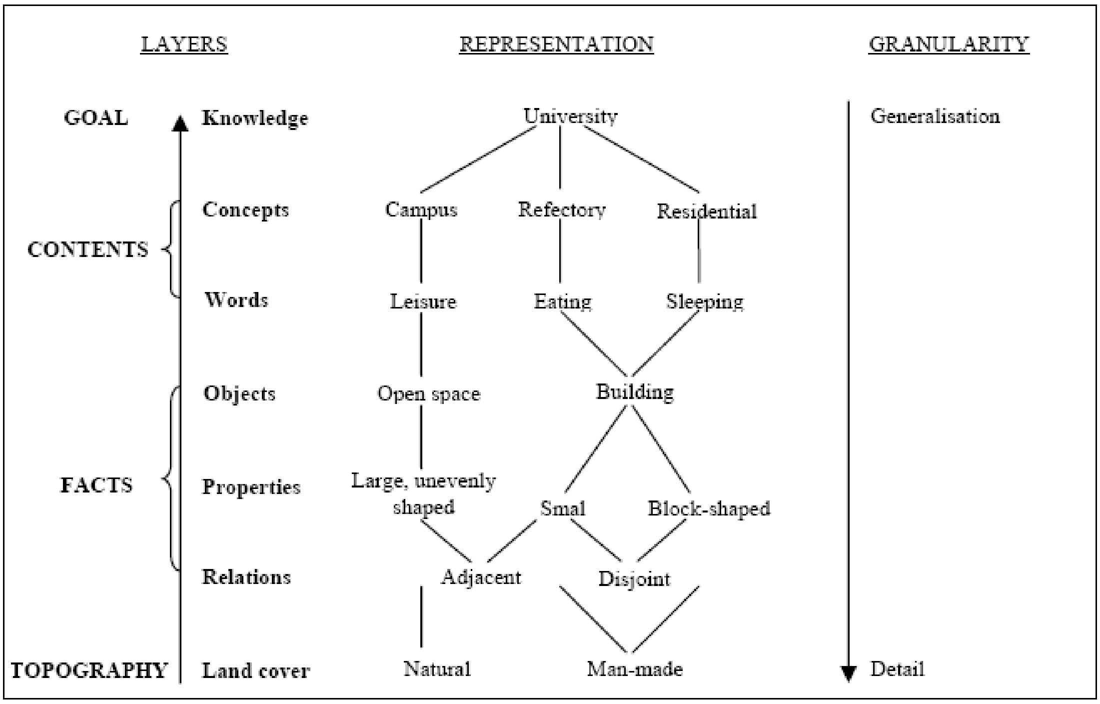


Figure 1.3. A top-down approach used by Thomson and Bera' to explain how a person may conceptualise the category of land-use, which in the specific example is represented by the geographic space occupied by the university (© Thomson and Bera').

As illustrated in the Thomson and Bera's model (see figure 1.3) layers, representation and granularity are the three key concepts. Layers represent the steps through which the knowledge of the reality is achieved; representation is the inner conceptualisation of the reality and granularity the different scale through which the knowledge about the reality takes place, and precisely from a general to a specific idea. The goal is to observe how people understand "university" as a category of land-use. The first layer is represented by the topography on top of which facts and contents are overlaid. With regard to representation the authors notice how people define the land-use category first of all in a generic terms, such as campus, refectory and residential, and then describe them further according to their purpose and role, spatial relations and physical properties. The authors conclude that the land-use category "university" is defined from the land-cover and by separating man-made and natural elements which have got specific properties and relations to each other.

On the basis of the papers reviewed it emerges that the landscape analysis is not only a difficult task due to the complexity of the landscape itself, but also it assumes the form of a philosophical debate. Thus in order to simplify the approach of landscape analysis, to make it feasible for GIS and to meet the aim of an objective and scientific based landscape characterisation it was decided not to consider public perception in landscape analysis, although this is an extremely important component for understanding the character of a landscape.

Following observations by many consultants and practitioners, in the characterisation phase of the LCA the influence of subjectivity could be generally avoided if the whole task is reduced to that of measuring attributes from maps and analysing the data quantitatively (Swanwick and Land Use Consultants, 2002). Hence, for the time being this research is addressed to the analysis of those measurable and objective landscape elements. Nevertheless it is suggested that in the future GIS-based landscape analysis could be enhanced by adding public perception amongst the explanatory variables of landscape character.

At this point, given that only physical and cultural elements of landscape in the GIS-based landscape characterisation are considered, someone might object that the GIS-based landscape analysis is partial and not holistic. Below it is suggested that in fact the GIS approach is far more comprehensive than other studies about landscape carried out until now.

1.3 Landscape as a whole

It is universally recognised that landscapes are complex and dynamic systems where natural and anthropic components are in continuous transition. These traits make them a difficult topic to deal with. Evidence that landscape analysis is more centred on the single elements and not on the whole is provided by the literature reviewed. The majority of the articles tackle the elements of the landscape singularly but offered interesting background and indications on how to carry quantitative analyses. Nagendra et al. (2004) integrated GIS, socio-economic and remote sensing techniques and from the analysis of images of land cover they explored the ways landscape pattern can be used to gain more information on land use process. The authors concluded that since landscape has heterogeneous and dynamic properties (characteristics), the integration of the temporal dimension would improve the results to their research (Nagendra et al., 2004).

Bocco et al. (2001), Giles (1997), Thompson et al. (2006) and Deng et al. (2005) focussed on the quantification of landforms, geomorphic signatures, topography and terrain attributes. Their studies were linked by a common thread since all provided examples of a similar method of analysing complex realities. Basically the approach is to simplify the calculation by facing the various components of the reality separately.

Instead of separating the landscape into its elements and analyse them separately other authors such as Antrop and Van Eetvelde (2000), Palang et al. (2000) reminded us that landscape should be viewed from a holistic point of view, as the result of the sum and interaction of its singular components and therefore geology and soils, landform and topography, field and settlements patterns should be linked to each other.

From the holistic point of view the identification of the landscape character occurs through the collection of the landscape elements and the analysis of their unique and distinctive association which is revealed in form of a spatial pattern. Hence, the objective of this research is to identify the spatial patterns within the landscape by using GIS and appropriate statistical techniques.

So far terms such as landscape characterisation and landscape character assessment have been widely employed but without any mention to their meaning, thus an exhaustive explanation of them is required and provided in the next paragraphs.

1.4 From landscape evaluation to landscape character assessment.

During the first half of the twentieth century, the need to incorporate landscape into the environmental decision making process was recognised, however, it wasn't until the 1970s that landscape became increasingly important as the emphasis on sustainability grew (see figure 1.4). Landscape Evaluation was born and the landscape started to be evaluated on the basis of those elements that make it better than another for particular purposes (C. Swanwick, 2002). These initial attempts at evaluation of landscape didn't succeed because many believed that it was inappropriate to reduce something complex, emotional and intertwined in human culture, to just a value expressed in number.

In the mid 1980s a different kind of tool emerged, known as Landscape Assessment. Contrary to the previous approach, this was based on an assessment which classified and described a landscape in two separated stages and on the basis of the elements that make a landscape different or distinct from another. As noticeable, the new approach did not distinguish the landscape in qualitative terms like "better" or "worse", but used words such as "unique" and "distinct".

Local authorities as well as practitioners in public and private sector became increasingly active. Both were encouraged in assessing the landscape in a wide range of applications and were assisted by a first model of guidance (C. Swanwick, 2002).

Since 1993 increasing emphasis has been placed on the role of the landscape character and the landscape assessment has been addressed to the characterisation of the landscape and the description of the forces for change on the basis of existing or derived data such as, land cover and land use, geological and topographical maps (Ho Kim and Pauleit, 2007). As a result Landscape Assessment turned into Landscape Character Assessment (LCA) and the Warwickshire landscape Project was the first application of LCA in the UK. From that moment the definition of character, as distinct and recognisable pattern of elements in the landscape that makes it different from another rather than better or worse, was set out explicitly (C. Swanwick, 2002).

Landscape Evaluation	Landscape Assessment	Landscape Character Assessment
• Focused on landscape value • Claimed to be an objective process • Compared value of one landscape with another • Relied on quantitative measurement of landscape elements	• Recognised role for both subjectivity and objectivity • Stressed differences between inventory, classification and evaluation of landscape • Provided scope for incorporating other people's perceptions of the landscape	• Focuses on landscape character • Divides process of characterisation from making judgements • Stresses potential for use at different scales • Links to Historic Landscape Characterisation • More recent emphasis on need for stakeholders to be involved
Early 1970s	**Mid 1980s**	**Mid 1990s**

Figure 1.4. The evolution and change in approach to the study of landscape throughout the last 40 years. Assessing landscape instead of evaluating was the most determining change because landscape wasn't considered "good or bad" but different according to the elements characterising it. Recently landscape character assessment is aiming at making stakeholders more actively involved in the phase of making judgements and taking decision. (© SNH and Countryside Agency – LCA. Guidance for England and Scotland)

In the UK the concept of LCA has been continuously developed and refined in the light of experience. On one hand the approach to LCA is common to England and Scotland where it has been carried out by the Countryside Agency and the Scottish Natural Heritage respectively, which in 2002 published jointly the "Guidance to the Landscape Character Assessment" that provides suggestions on how to carry out a LCA. This guide was accompanied by a second important reference entitled "Guidelines for landscapes and visual impact assessment" by the Landscape Institute and Institute of Environmental, Management and Assessment (2002).

On the other hand, Northern Ireland and Wales have undertaken a similar work but in a slight different way, especially Wales where the initiative is called LANDMAP (Swanwick and Land Use Consultants, 2002).

In Scotland, the information available for Scotland's landscapes was poor when compared to other aspects of the natural heritage, such as habitat or species. This imbalance became a problem when Scottish Natural Heritage, after its establishment in 1992, wanted to advise planning authorities on development control cases and strategic planning issues. It was evident that SNH needed to develop a methodology for making people aware of the landscape as a natural resource and heritage and for understanding the processes that constitute a landscape (Swanwick and Land Use Consultants, 2002).

With this premise, in 1994 SNH promoted a national Landscape Character Assessment involving 29 local authorities and other organisations such as the Forestry Commission, Scottish Enterprise, Historic Scotland, the Scottish executive Rural Affairs Division as well as small local groups. These various partners were involved either in the consultations or in the steering group.

Each study was carried out by landscape architects or landscape consultants who referred to the above mentioned "Guidance to the LCA" by C. Swanwick and Land Use Consultants. However, despite LCA is a systematic analysis of the landscape once completed it appeared widely dependant on the professional but personal judgement of the landscape architects. Additionally not only the different *modus operandi* of the landscape architects interfered with the outcomes of the 29 assessments but also the diversity of uses for which they were intended and the aspirations of the individual local authorities had a strong influence. Finally the scale adopted during the characterisation phase, which was mainly 1:50000 and in some cases 1:25000, contributed to make the 29 assessments fairly independent but each comprised of a description of the landscape character types and areas.

The descriptive work covers different contexts such as geology, landform, hydrology, land use and land cover, settlements and field patterns. Information on landscape experience and forces for change are added to the rest of the landscape description. GIS was used more for visualising the geographic data than for the analysis and identification of landscape characters.

When the national LCA programme was completed (with the exclusion of the Loch Lomond and the Trossachs National Park) the data from the 29 assessments were pulled together and since the number of landscape character areas totalled 3967, it was decided to create two more levels of landscape in order to aggregate areas having common features in types. As result of this, level 2 was created which is comprised of 366 landscape character types, and level 3, which is comprised of 106 more generalised categories of landscape. The full range of the three different levels represents the hierarchy of landscape character in Scotland based on a bottom-up grouping method (Swanwick and Land Use Consultants, 2002).

The geographic database that was prepared allowed the three levels of landscape character to be mapped for the whole country and it is currently available to the public on the SNH website; in this regard figure 1.5 illustrates the map of Scotland's landscape types at level 3 of aggregation (Swanwick and Land Use Consultants, 2002).

Figure 1.5

In Scotland, the relevance of LCA in decision making was recognised in various elements of planning legislation. For example the National Planning Policy Guideline (NPPG) 14 makes explicit reference to SNH's programme of LCAs and underlines the use of these assessments in the planning processes in providing guidance on the capacity of the landscape to accommodate development, and in informing policy development and development control processes. A year later, in 1999 another NPPG, number 18, stresses the relationship between historic buildings, cultural features and

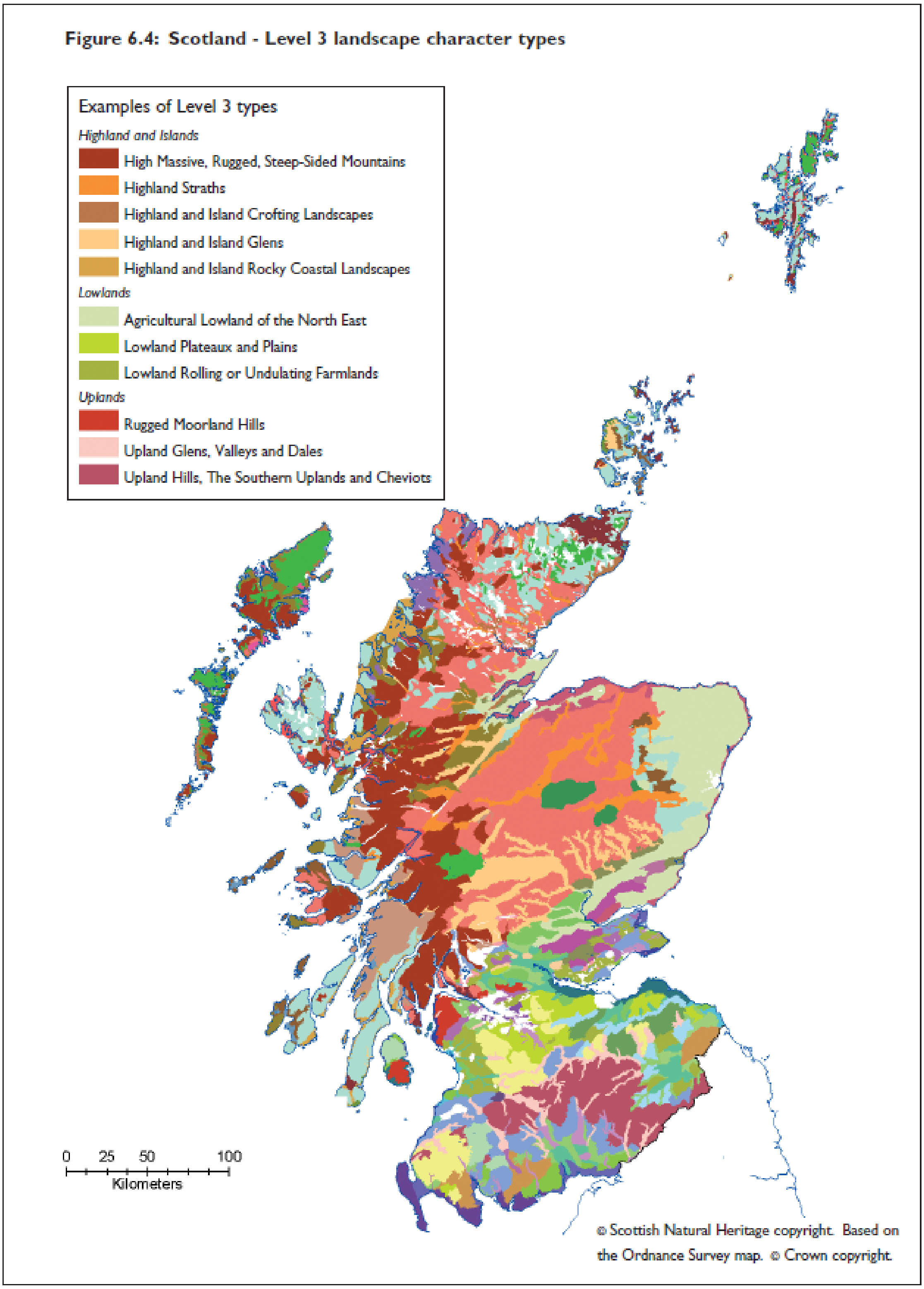


Figure 1.5. Level 3 of the LCA as carried out in Scotland. The landscape character types have been aggregated on the basis of common traits in a "bottom-up" classification system (© SNH and Countryside Agency – LCA. Guidance for England and Scotland).

the natural environment and highlights that this combination provides an area with its particular identity and character. In 2000, the Planning Advice Note 60 states that safeguarding and enhancing landscape character is an important planning objective (Swanwick and Land Use Consultants, 2002).

In England the Countryside Agency (formerly the Countryside Commission) began an innovative programme of work oriented to the analysis of the wider countryside in the early 1990s. A pilot study was carried out in the southwest regions in 1993/1994 which helped to develop a robust methodology that combined map analysis of the different variables that give the landscape its character, GIS data handling and computer classification methods.

In parallel English Nature launched its own Natural Areas programme to provide a similar framework for setting nature conservation objectives. Because both organisations were working on a similar subject they joint their forces to produce a single map underpinning both landscape and nature conservation measures in future. This was the Character of England map, illustrated in figure 1.6 (Swanwick and Land Use Consultants, 2002).

Figure 1.6

The map is accompanied by descriptions of the 159 character areas which include key characteristics, physical, historical and cultural influences, building, settlement and land cover patterns. The influences determining the character and the pressures for change were also noted. The methodology involved compiling a list of attributes of the different influences on the landscape and each square kilometre was coded with the relevant attribute, e.g. geology, soil type, settlement or field pattern. The coded database was then statistically analysed using a program called TWINSPAN which sought to organise typologies of landscape in a hierarchical structure.

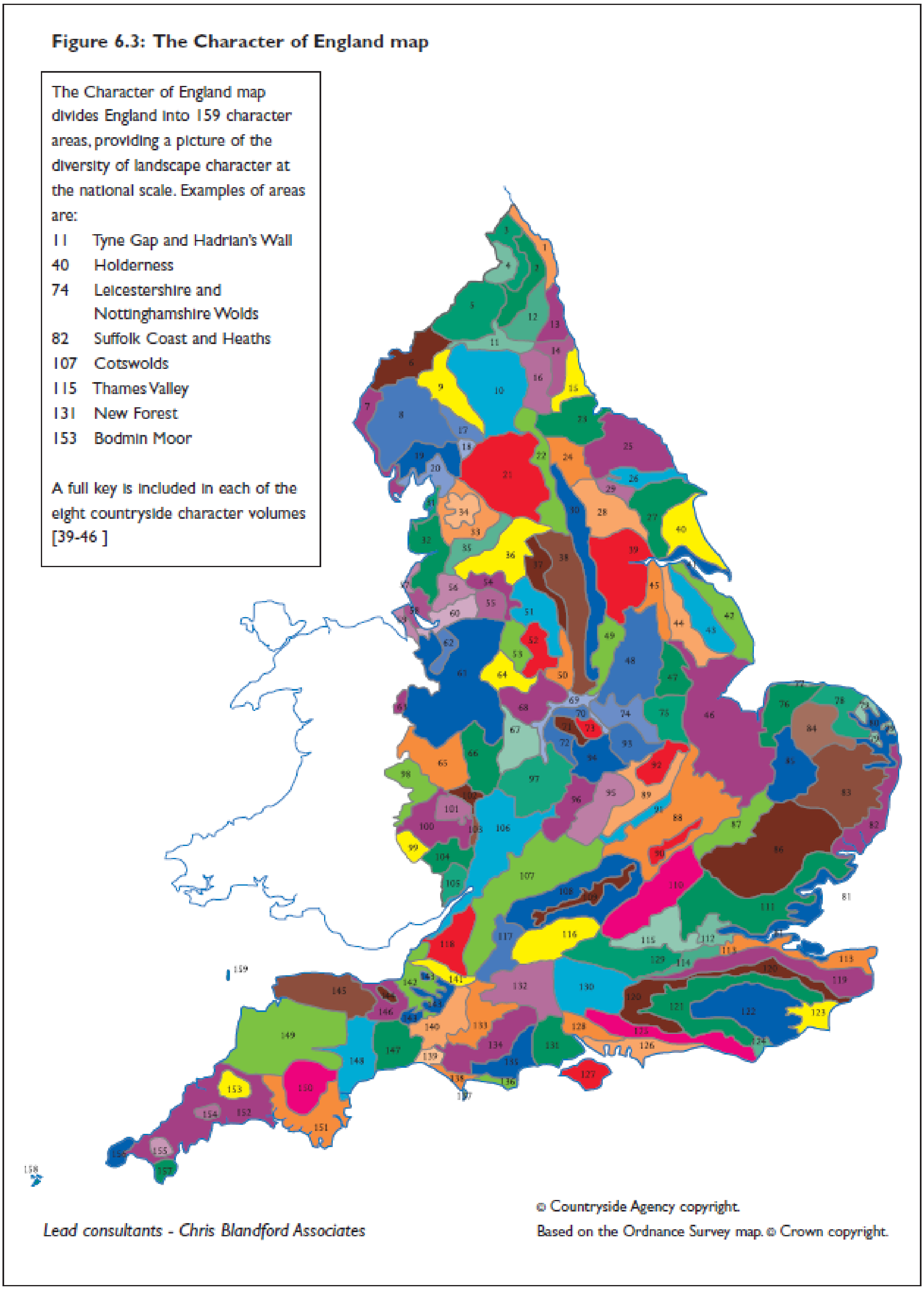


Figure 1.6. The Character of England map illustrates the 159 character areas identified with the LCA (© SNH and Countryside Agency – LCA. Guidance for England and Scotland).

After several iterations of the program were carried out using different thresholds of classification, a map was produced that broke the sample area down into a range of types at a regional scale (Bell, 2002).

However, the Character of England map dealt only with landscape character areas and lacked of a national landscape typology. Therefore, the Countryside Agency in collaboration with English Nature and English Heritage developed a new approach called National Mapping Project.

Typology has been derived by map analysis of the main physical, biological and cultural factors that determine the landscape character and GIS have been used for manipulating data. The resulting national typology is an intermediate level between that of the Character of England map and that of the local authority; hence it can inform work at both of these levels (Swanwick and Land Use Consultants, 2002).

In Wales, the tool for analysing and studying landscape is called LANDMAP, which started in 2003 and ended in 2009 and is the results of a partnership between the Countryside Council for Wales and the Unitary and National Park authorities of Wales. LANDMAP takes a whole landscape approach that covers all landscape designated and not designated. It includes the natural, rural, coastal, inland water and peri-urban areas (excluding Cardiff and Swansea). Many of the European Landscape Convention’s principles are reflected in the LANDMAP programme together with other landscape work by the Countryside Council for Wales (J. Bullen, LCN issue 34, 2010).

LANDMAP separates the landscape into five aspects and evaluates them: Geological landscape, Landscape Habitats, Visual and Sensory, the Historic Landscape and the Cultural Landscape. Each of the five spatial layers are divided into discrete geographical units (polygons) referred to as aspect areas which are distinctly defined by their recognisable landscape characteristics and qualities recorded during a field survey by specialists and local authority officers (J. Bullen, LCN issue 34, 2010). Data is recorded on a standard form in order to ensure transparency and consistency and allows comparison of the content and value of the landscape between areas. The main outcomes from LANDMAP are thematic landscape maps of Wales that can be produced to depict how landscapes are identified and classified for each of the five layers, illustrating landscape patterns and their distribution, diversity and representativeness. The maps and the survey information provide a consistent core of baseline information onto which any additional geographic detail can be added.

Of crucial importance to LANDMAP is the accessibility of the data through the LANDMAP website which can occur in two ways: the data can either viewed or downloaded and then used in people's own GIS system (J. Bullen, LCN issue 34, 2010).

In Northern Ireland the Environment and Heritage Survey started in 1997 and was completed in 2000. The LCA identifies 130 unique landscape character areas across the region. Broad descriptions of regional landscapes and more detailed descriptions of individual character areas are provided in a series of 26 LCA reports organised by local government district. These reports describe landscape character, analyse landscape qualities and features and provide guidance on accommodating development and other land use change. The classification and description is consistent across Northern Ireland because the assessment was undertaken as a single exercise and by a single firm of consultants. The level of detail is intermediate between that of the Character of England map and the Scottish LCAs. The results are available internally on GIS and externally on the web of the Environment and Heritage Service. Additional research included information on biodiversity and earth science for each landscape character area (Julie Martin Associates and Swanwick, 2003).

1.5 The role of history as a contributor to the character of the landscape

As the number of LCA initiatives increased throughout the years and across the UK other programmes, which were thought to reveal the historical and archaeological dimension of the landscape, emerged. Human actions in the past and their evidence in the landscape are without doubt highly significant to the present day landscape character. Thus, two free standing programmes were embarked upon by English Heritage and Historic Scotland: the Historic Landscape Characterisation (HLC) and the Historic Land-Use Assessment (HLA) (Swanwick and Land Use Consultants, 2002).

In contrast to the English HLC, in Scotland the HLA adopted immediately GIS technology from that outset which allowed a more dynamic and articulated analysis. Not only subtleties and changes within the landscape could be recorded more sensibly, but also data entry and metadata creation was more rigorous. As result the methodology of historic land-use assessment was adapted to respond better to GIS technology (Dyson-Bruce, 2003).

Conceptually, the HLA is meant to provide an understanding of the historic dimension of the modern landscape that should empower people to make more informed decisions about the future development of their landscape. Thus, as far as the aim is concerned the HLA does not differ from the LCA and it is interesting the way both programmes aspire to provide outcomes applicable to practical landscape management and conservation policies and decisions. HLA and LCA hope to disseminate information which can be translated into plans and strategies by decision makers, environmental conservationists, planners and developers, landscape managers and so on.

HLA methodology is a broad-brush and desktop based approach that assesses historical information related to the current land-use through the interpretation of aerial photographs, old maps, text, books and other documentary sources. The scale of the data capture is 1:25.000 while maps are digitised at 1:10.000. The format of the dataset is shape file which was then converted it into raster with a resolution (cell size) equal to 50 meters. The HLA is an on going programme and therefore is only partly completed. The attribute table of the HLA consists of the following main fields[2]:

- Historic land-use category, which is the historic land-use grouped into 14 broad classes;
- Historic land-use period, which reflects the current archaeological and historical understanding of that land-use;
- Historic land-use type, which is the historic land-use classified into finer classes according to the period, form and function of that land-use;
- Relicts (1,2,3), which identifies the number of historical evidences still visible/recognisable. Relicts have been classified into category, period and type as well.

In contrast to the LCA, the single ownership of the project emerged as an advantage of the HLA because it assured the use of a uniform method and a high standard of consistency in mapping and in interpreting data. However the differences between HLA and LCA became an obstacle. In fact despite everybody acknowledges that a landscape is both natural and historical heritage (figure 1.7 illustrates it clearly) operationally the LCA and HLC/HLA have such a different approach to characterisation which makes their integration difficult.

It is a personal opinion that the lack of collaboration between the two programmes is not beneficial. A better knowledge about the variable “time” in the process of landscape characterisation should help to answer questions like “Which areas are the oldest? Which areas are derived from prehistoric and medieval patterns? Which areas have changed the most, indicating high dynamism?”

[2] More detailed information on each field can be obtained from the Royal Commission on the Ancient and Historical Monuments of Scotland which is the body undertaking the HLA http://hla.rcahms.gov.uk/

Figure 1.7. Despite the wide agreement on considering the landscape as formed by natural and cultural elements, HLA and LCA proceed more in parallel than together. This research aims at incorporating information from HLA more efficiently, so that the landscape analysis result more holistic (© I.Marengo).

Whereas questions like “Which areas are the rarest or provide particularly significant evidence of history? Which areas have a greater diversity of types? Which areas show more evidence of the past and can tell us more about the history?” can be answered by investigating the relicts (another way of referring to evidence from the past) and the diversity in which they occur.

From the facts and experience gathered at the end of this PhD it would be desirable to see a more effective collaboration between landscape architects and historians since the synergy is likely to enhance the final results of landscape characterisation enormously.

1.6 European approaches to landscape character assessment

The differences in the approach to LCA showed by England, Scotland, Wales and Northern Ireland have their counterparts at European level. In 2005 a report was edited by the European Landscape Character Assessment Initiative (ELCAI) on the typology, cartography and indicators for the assessment of sustainable landscapes. The report captured 54 national LCA works within the European countries which adopted the LCA and are listed in table 1.1.

Austria	Czech Republic	Italy	Denmark
Belgium	Slovakia	Netherland	Hungary
Germany	Norway	Switzerland	Estonia
United Kingdom	Portugal	Ireland	Spain

Table 1.1. List of the European Countries that adopted the LCA (source: Wascher – European Landscape Character Areas)

From the reports it emerged that each country ended with a different LCA and the reason was the variety of objectives each country wanted to achieve through their LCA programme (Groom, 2005). The outputs differed mainly because of the scale used, the number of hierarchical level, and the identification of the landscape character into types or areas. Table 2 illustrates the factors that are considered relevant to different countries, providing an insight of the way LCA has been used and interpreted. Overall biophysical factors are those that almost every country has used for the LCA; however, cultural factors are also prevalent and mainly represented by land use, historical development, field and settlement pattern and heritage. As far as the methodology is concerned, the interpretation provided by experts is the most widely used method. Highly automated derivation of LCA is not as common as it might be expected given the widespread use of GIS. It is likely this computer device has been exploited more for its mapping facilities than for its powerful analytical functionalities (Groom, 2005).

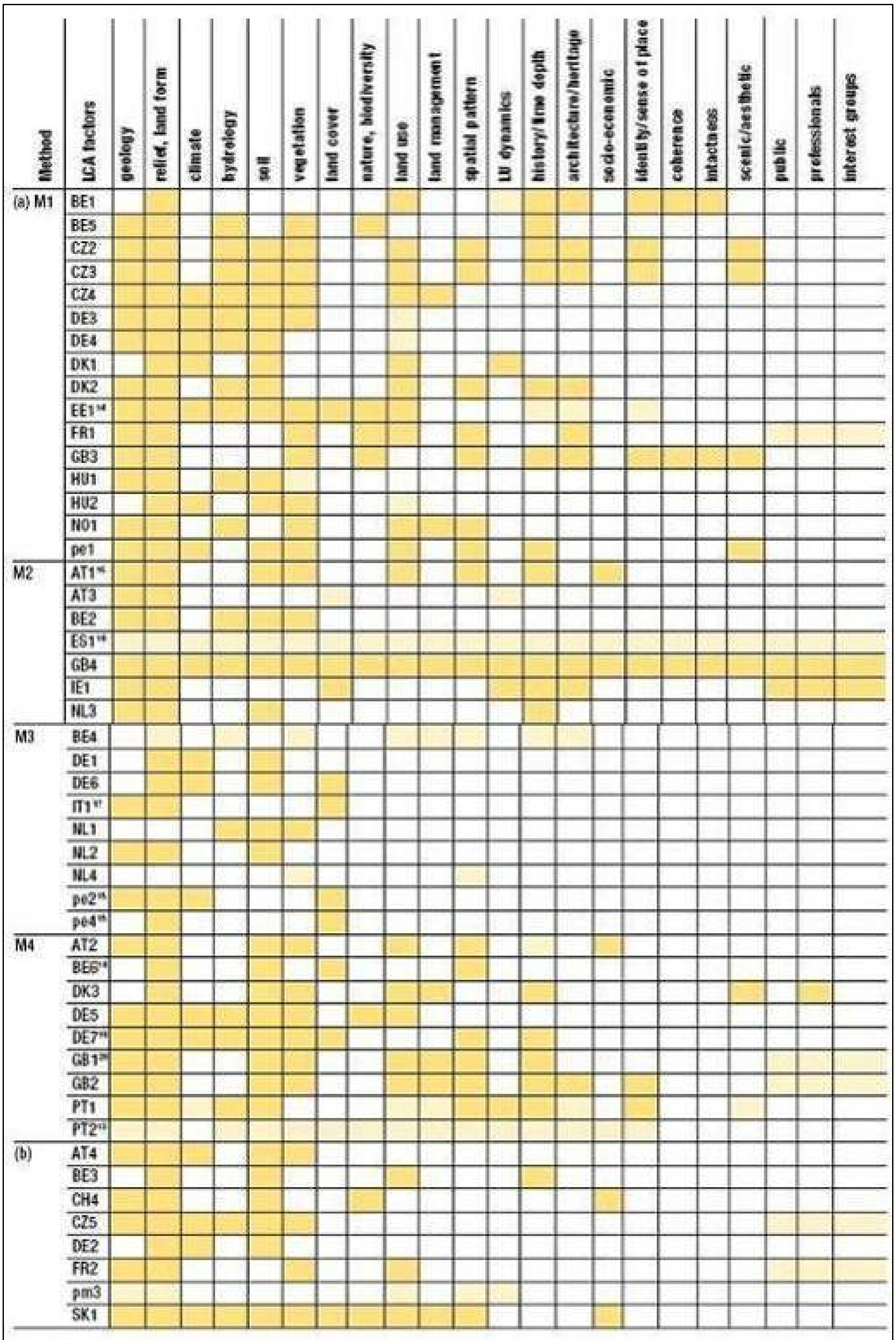
Table 1.2

The ELCAI report investigated also the period of time that the European countries undertook LCA initiatives and it emerged that few examples can be inscribed between 1950 and 1990 (Groom, 2005). The reason was likely due to the fact that the first half of the last century coincided with the transition from landscape evolution and landscape assessment and with the application of model-based and reductionist works to geographical issues. Therefore, geographers were engaged more in a “conceptual battle” than active in practical studies and works. On the contrary, from the 1990s onwards almost all the countries started LCA programmes, and were also supported by the development and availability of technology for handling, visualising and analysing large spatial datasets. At that time, detailed guidelines for the production of LCAs were published, contributing to the dissemination of LCA practice and providing a helpful support for regional and local applications (Groom, 2005).

1.7 Examples of LCA in the rest of the world

Not only in the UK or in Europe was LCA undertaken and in the 90s it rapidly became the common method of landscape analysis. Amongst the literature reviewed there are many examples of LCA initiatives carried out in different parts of the world and examples of case studies are provided below.

Ho Kim and Pauleit (2007) adapted and tested the UK approach to LCA in order to identify the landscape character types in a region of South Korea and to provide



Pale yellow = analysis is uncertain

14 The set of factors used is broad, but variable since the LCA approach is open.

15 Socio-economic criteria = land ownership patterns.

16 The actual factors used is not known.

17 Land cover = CORINE LC.

Table 1.2. Results from the review of the factors and methods used in each LCA work undertaken by some of the European countries. (© Wascher – European Landscape Character Areas)

quantitative information on the potential biodiversity occurring in each type (Ho Kim and Pauleit, 2007). Strong pressure on landscapes in South Korea due to extensive urbanisation is leading to severe environmental problems, changes in landscape character and the loss and fragmentation of woodland which threatens biodiversity. Using land cover data and aerial photographs, four landscape metrics were customised for the landscape analysis and the results provided the useful information required to support sustainable land use planning for nature conservation. In addition Ho Kim and Pauleit demonstrated that the UK approach for LCA was efficient when transferred to a completely different reality such as that of the Korean region (Ho Kim and Pauleit, 2007).

In Hong Kong, the planning department in 2001 commissioned private consultants to map the value of the landscape, which across the country reveals large undeveloped and natural tracts that support native plant species, a varied wildlife, a long history of human settlement and a variety of cultural relics associated with the settlement. Previous studies had been carried out but none had considered the landscape as a whole. The LCA carried out in 2001 sought to fill the gap due to the absence of comprehensive information on the existing conditions of landscape resources in Hong Kong. The information provided by the LCA helped to establish indicators for monitoring the landscape conditions and is available on the web and accessible to the public.
(http://www.pland.gov.hk/pland_en/p_study/prog_s/landscape/e_index.htm)

In the United States, Ahern et al. (2005) illustrated another application of LCA to the case of a landscape which was facing a rapid transformation. The American researchers conducted a landscape character study with the aim of making the local community of Cape Cod in Massachusetts aware of the radical transformation that was seriously threatening their landscape. In order to achieve their aim, the authors pointed out the links between place, character and people's memories and ended with stressing the point that for the local community the need to preserve the cultural landscape was as strong as the need to conserve the physical environment (Ahern et al., 2005).

Ahern et al. also demonstrated that the pressures on Cape Cod's landscape, due to the ongoing changes, influenced both the perception of the landscape and the lifestyle of the local people. Thus in order to turn their research into action Ahern et al. suggested three scenarios for the area of study and promoted their results in a conference involving the local community (Ahern et al., 2005).

In New Zealand, in the Coromandel Peninsula of the North Island a study was carried out by Brabyn (2005) who aimed to identify the naturalness of the area using three datasets: land cover, utility (roads, railways and transmission lines), density, and cadastre. GIS solutions along with fuzzy set theory were applied to the analysis of the three datasets; the results were compared to each other in order to understand which one provided the more satisfying and valid classification of naturalness (Brabyn, 2005). The findings showed that the methods of analysis applied to the three datasets identified the natural and developed areas quite consistently despite different representations, whereas, for the classes in between it was hard to provide a definition of naturalness. According to the author the reason was due to the fact that naturalness is a fuzzy concept as it involves human conceptualization. Therefore where forms of development or naturalness were not sharply defined and separated it was difficult to classify them and reach a consistent agreement. Brabyn's conclusions were that it was possible to use GIS for the landscape analysis but in order to achieve better results he recommended collaboration between GIS analysts and cognitive researchers, so that quantitative and more abstract concepts could be taken into account in the overall landscape analysis.

1.8 Principal objectives of the PhD

From the overview of the LCA initiatives taken at UK, European and international level a common thread emerges quite clearly, and namely that the LCA in first instance is addressed at finding the difference between one landscape to another, categorising them into types and/or areas and describing them. Maps are tools to distinguish essentially the different landscapes and indicate their distribution, extent and exposure to major influences (Mander, 2010).

Then, on the basis of the knowledge about the landscapes and the information on the maps, the LCA supports and guides planners and landscape practitioners in the judgement and decision making. More broadly the LCA can be seen as the bridge between scientific knowledge and socio-economic issues that are needed to meet the demands of sustainable planning and management.

Aims of this research are to make the landscape characterisation a more quantitatively and scientifically based process and, in order to achieve so, demonstrate whether or not GIS can be successfully applied as analytical support tool. The objectives are:

- to exploit the capacities of data storing and data analysis that GIS have showed to have throughout years and diverse fields of applications;
- to make the landscape characterisation more comprehensive by attributing more relevance to historical and cultural aspects of the landscape;
- to create a functional link between more powerful statistics and GIS.

The scheme illustrated in figure 1.8 exemplifies the way of proceeding followed by this research.

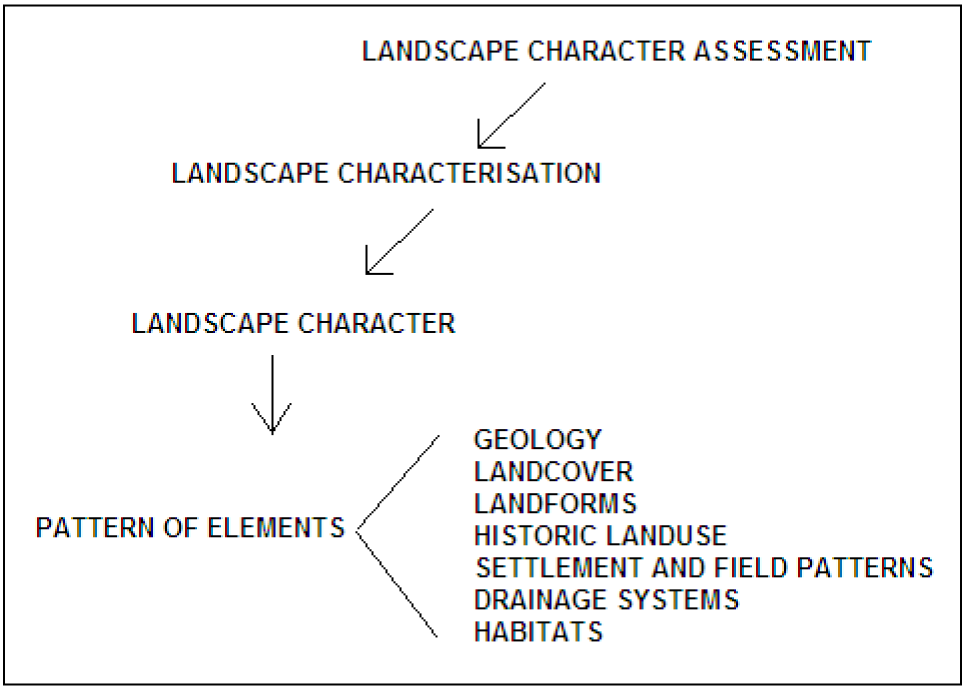


Figure 1.8. The way the PhD unfolded: started from the general concept of landscape character assessment and ended with a focus on the definition and identification of the character of the landscape.

The structure of the thesis can be summarised in brief. It started by understanding the structure and purposes of the LCA, secondly it concentrated on the characterisation phase and looked at the different steps of which it is comprised, third it moved on the key point of the investigation, and namely the character.

At that stage the methodology was firstly designed by considering the elements not in isolation but in association to each other. Then the first law of geography, stated by W.Tobler in 1970, and the theory of spatial autocorrelation were taken as main reference concepts on which to develop the methodology. Basically both ideas implied that the landscape elements can form geographic patterns which can be recognised as spatial structures within the data.

The identified patterns corresponded to landscape character types and the methodology was applied to eleven case studies selected all over Scotland in order to cover a large variety of landscapes and check its consistency and repeatability.

Finally the research concluded with a series of tests for verifying the robustness of the methodology under different situations.

Task of the following chapters is to describe clearly and comprehensively how the methodology was designed, developed and implemented. Thus the starting point is to understand the conceptual model of the LCA and its repartition into the characterisation and making judgment phases.

CHAPTER 2

Structure and concepts that support the Landscape Character Assessment

Lismore island, looking south towards Appin from the main road (I.Marengo).

2.1 **The conceptual model of LCA**

As explained so far, LCA is concerned with landscape character, rather than with landscape quality or value and understanding this point is vital because often the tendency is to associate character and quality while the difference between the two is conceptual and philosophical (Swanwick and Land Use Consultants, 2002).

According to the LCA guidance for Scotland and England (which is the main point of reference for this thesis and for LCA in general) LCA is primarily a tool to help understand what a landscape is nowadays, what it may come to be like and how it may change in the future (Swanwick and Land Use Consultants, 2002). With this in mind some of the common applications of LCA are found in practical fields such as planning and landscape conservation, management and enhancement. For instance LCA may be applied because it informs development plan policies at regional and local levels; contributes to landscape capacity studies; provides input to Environmental Assessment and it is basis for landscape management strategies; helps the identification of areas for designation and guides land use change in sustainable way and so on.

LCA is organised in two parts in the following order: first a relatively value-free *characterisation* and second the *judgements* made on the basis of the knowledge of landscape character (Swanwick and Land Use Consultants, 2002). Then each part is divided into a series of steps that come in a specific order and are summarised in figure 2.1.

Characterisation involves the identification of the character of the landscape into types and areas. The former are relatively homogeneous and generic types of landscapes which occur in different parts and geographic areas of the country. Landscape character types of the same class basically share a similar association of landscape elements, such as geology, land cover, land use, landforms settlement and field patterns and so on.

PART I - CHARACTERISATION

1. DEFINE SCOPE
2. DESK STUDY
3. FIELD SURVEY
4. CLASSIFICATION AND DESCRIPTION

PART II - MAKING JUDGEMENTS

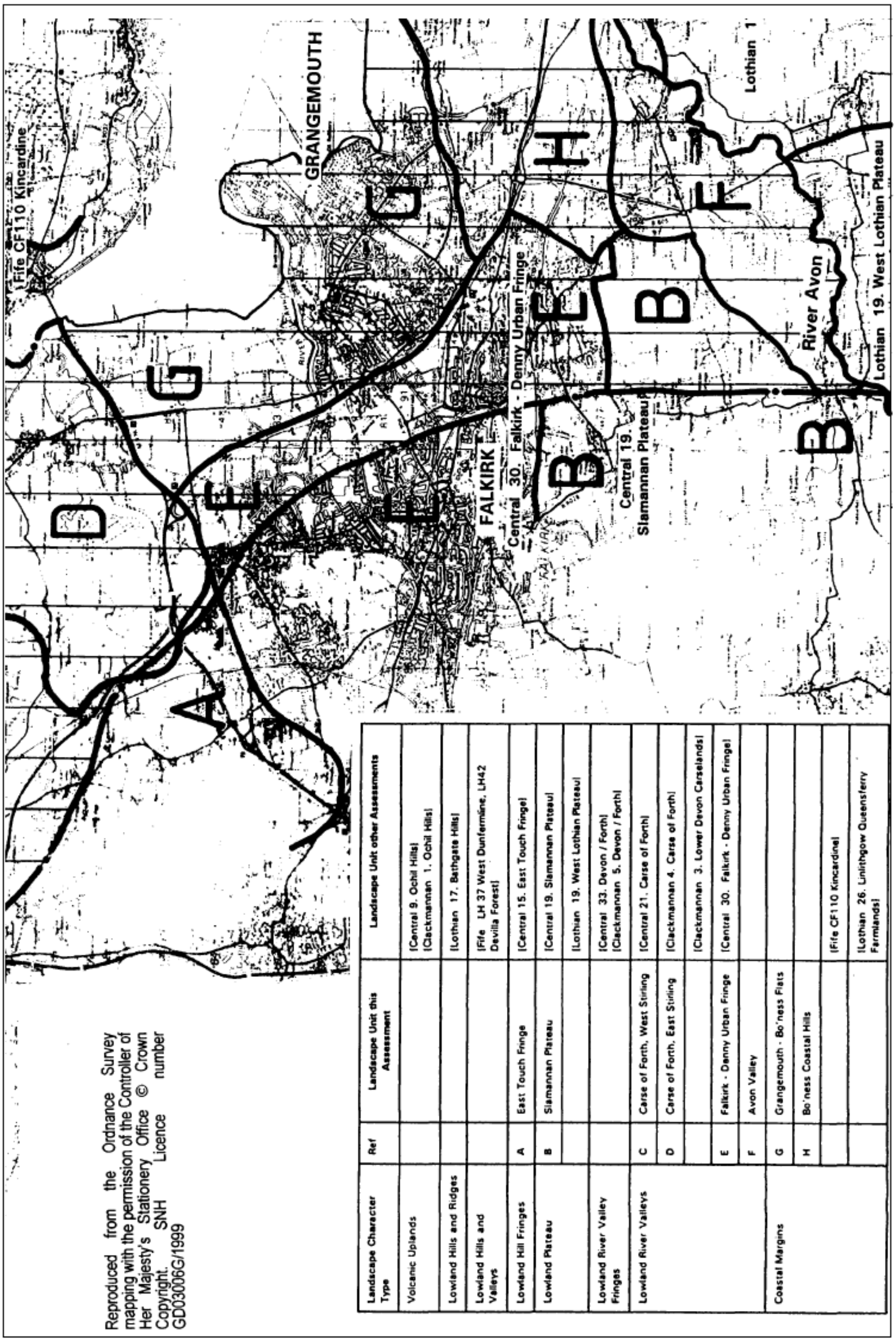
1. DECIDING THE APPROACH TO JUDGEMENTS
2. MAKING JUDGEMENTS

Figure 2.1. Characterisation and making judgments with their several steps form the LCA. The phases are chronological and each step needs to be taken in the illustrated order. However, at the start the purpose, the objectives and the sort of judgements are required from the LCA process, should be clarified (source: SNH and The Countryside Agency – The LCA. Guidance for England and Scotland).

Conversely landscape character areas are unique and refer to a specific geographic location, therefore they have their own identity and are named after the geography of the place. For instance the Stirling and Grangemouth LCA identified a landscape character type called “Lowland river valleys” to which the landscape character areas “the Carse of Forth West of Stirling” and “the Carse of Forth East of Stirling” were associated to (see figure 2.2).

Figure 2.2

As it emerges from the map, the landscape character types and areas after being identified are classified, mapped and described so that it is possible to locate them geographically and to gain more knowledge about them. For example the LCA for Stirling and Grangemouth describes the “Lowland river valleys” as “flat, open, large-scale *carseland* forming the floor and former floodplain of the River Forth…”



Landscape Character Type	Ref	Landscape Unit this Assessment	Landscape Unit other Assessments
Volcanic Uplands			[Central 9. Ochil Hills] [Clackmannan 1. Ochil Hills]
Lowland Hills and Ridges			[Lothian 17. Bathgate Hills]
Lowland Hills and Valleys			[Fife LH 37 West Dunfermline, LH42 Devilla Forest]
Lowland Hill Fringes	A	East Touch Fringe	[Central 15. East Touch Fringe]
Lowland Plateau	B	Slamannan Plateau	[Central 19. Slamannan Plateau]
			[Lothian 19. West Lothian Plateau]
Lowland River Valley Fringes			[Central 33. Devon / Forth] [Clackmannan 5. Devon / Forth]
Lowland River Valleys	C	Carse of Forth, West Stirling	[Central 21. Carse of Forth]
	D	Carse of Forth, East Stirling	[Clackmannan 4. Carse of Forth]
			[Clackmannan 3. Lower Devon Carselands]
	E	Falkirk - Denny Urban Fringe	[Central 30. Falkirk - Denny Urban Fringe]
	F	Avon Valley	
Coastal Margins	G	Grangemouth - Bo'ness Flats	
	H	Bo'ness Coastal Hills	
			[Fife CF110 Kincardine]
			[Lothian 26. Linlithgow Queensferry Farmlands]

Figure 2.2. Extract of the LCA map for Stirling and Grangemouth. In the table on the bottom left there is mention of the landscape character types and the landscape character areas (here referred as units), (source: LCA Stirling and Grangemouth – David Tyldesley and Associates 1999).

The LCA guidance invites the landscape architects or people involved in the preparation of the LCA to provide a clear and concise description of the landscape character types and areas. This description is important because not only it informs on what the landscape is but also it mentions which “forces for change” are likely to occur.

Here the term “forces for change” refers to the key development pressures and trends in land management that can affect and change a landscape character type or area (Swanwick and Land Use Consultants, 2002). In the above mentioned case of the “lowland rivers valleys” some of the identified forces for change are the urban expansion of Stirling, Falkirk and nearby centres; the lack of architectural composition or relationship with settings already present; flood prevention and other engineering works on rivers and burns; new or improved major and minor roads and so on (Julie Martins Associated and Swanwick, 2003).

It is evident from this brief introduction that the characterisation phase of the LCA aims at providing the necessary information for having a general but comprehensive picture of the landscape of a given area. Understanding the physical and cultural aspects of the landscape and the possible developments that can occur and modify it are meant to provide the basis for a sustainable and better landscape planning and management.

The making judgements part of the LCA formalises and makes operative the results obtained from the characterisation by informing different types of decisions to be taken regards the landscape. As previously mentioned, LCA comes to be a useful and key tool for local and national authorities when dealing with planning and landscape management and conservation issues. Indeed there is not one approach of making judgments but different ones according to the type of issue and objectives of the overall landscape assessment and as mentioned in the guidance the main approaches are:

- landscape strategies;
- landscape guidelines;

- landscape capacity and
- attaching status to landscape.

The outputs associated with the making judgements phase should ensure first of all that the judgements are clearly linked to the results of characterisation; secondly that there are specific reasons and criteria of adopting an approach instead of another; and finally that it is unambiguously stated who made the judgements and took the decision on the basis of the information from the characterisation.

This research is addressed to the development and implementation of a GIS-based methodology for landscape characterisation, therefore from now the attention and the emphasis will be drawn on the first part of LCA. The objective is to describe in more detail the current approach to landscape characterisation and analyse it briefly through its four steps depicted in figure 2.3.

Defining the scope of the LCA is the first action to take since it will influence the way of carrying out the assessment, the approach to judgements needed to inform decisions, and finally the use and dissemination of the results. Briefly, stating the ultimate purpose of the LCA allows people to clarify the scale and the level of detail of the assessment, the individuals or institutions that should be involved in the preparation of it, the skills, expertise, time and economic resources required, and finally the types of judgements needed to inform the decision.

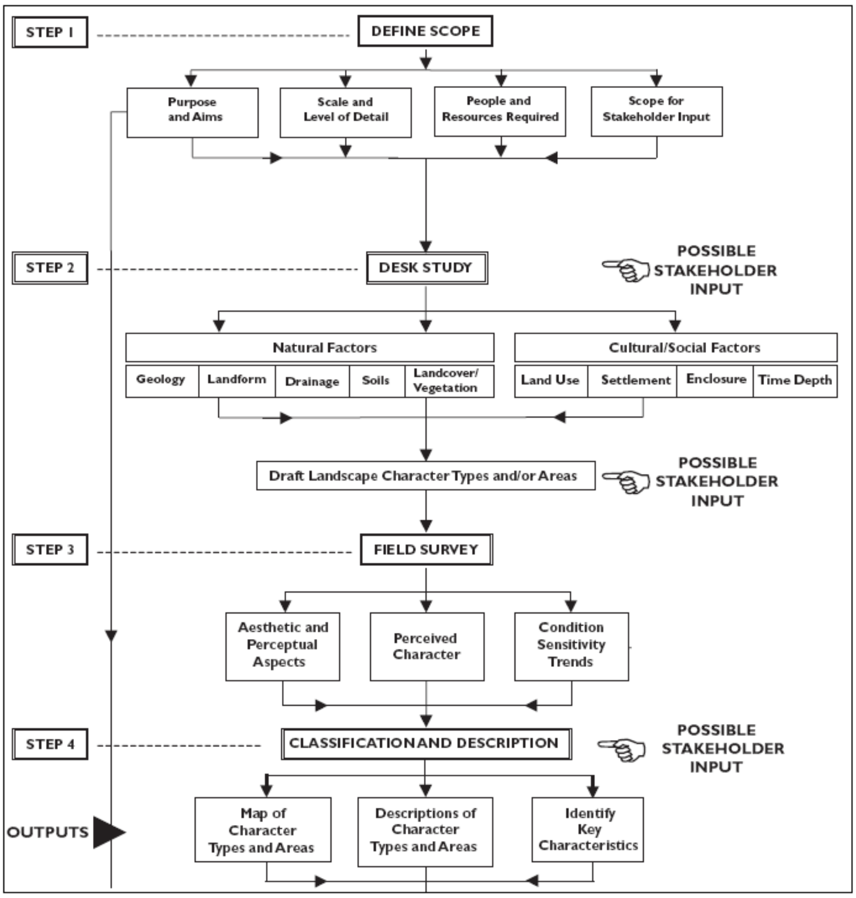


Figure 2.3. The four steps of the landscape characterisation phase in detail. As it appears each step requires further analysis and also a possible input from the stakeholders, which it is recommended but about which the local authority or the promoter of LCA can decide. (source: SNH and The Countryside Agency – The LCA. Guidance for England and Scotland).

Scale is an essential concept in both natural and social sciences and becomes a common issue when dealing with geographic data and landscape is not an exception. Scale can be expressed in spatial or temporal units and may be associated as the viewing window size of the observer or the broader framework in which diverse ecological phenomena take place (Wu and Qi, 2000).

The LCA process identifies three main scales: national, regional and local. In general an assessment at national and regional scale is carried out on the basis of 1:250000 scale. While at local authority scales it is preferable to adopt scales such as 1:50000 and 1:25000. For very small areas, for example country parish, estates or proposed development sites, a detailed landscape assessment is needed thus a 1:10000 scale should be preferred. Accompanied to scale is the concept of level of detail which translates into amount of information necessary to conduct an appropriate and exhaustive assessment. A GIS analyst and any other landscape assessor are well aware that the data collection is a very time consuming task, which often leads to the discovery that some important data is missing or only partially available. Nevertheless the advantage of a clear scope is that people can focus the investigation on the datasets that are needed for the overall assessment.

Strictly related to scale is the process of extrapolating or translating information from one scale to another. The opportunity of transferring information between scales is indispensable, particularly in a hierarchical process such as LCA where landscape character types at local level can be aggregated to constitute broader landscape type at regional or national level. This is the case of Scotland; as mentioned in the previous chapter landscape characterisation followed a bottom-up approach so that the “level three” results from the amalgamation of landscape types identified at finer scale.

In theory, both bottom-up and top-down approach seems as easy as assembling or disaggregating a Russian doll. However, when working with geographic data and above all with GIS, the scaling operation requires that the analyst bears in mind some issues with respect to the size of the study area defined as landscape, the level of spatial resolution and aggregation of the data. Extent, grain, detail of the data and area of study contribute to define the structure of a landscape in a way that a change would lead to different conclusions in the landscape analysis.

Figure 2.4 illustrates the same area but at different scale; it is noticeable how zooming in or out tends to reveal or conceal attributes of the map; as a consequence the aspect of the map varies. A similar situation occurs while the analysis of the landscape is carried at national, regional or local scale; for each level the outputs of the characterisation differ so that in order to be precise the scale to which a landscape character type/area belongs should be always mentioned.

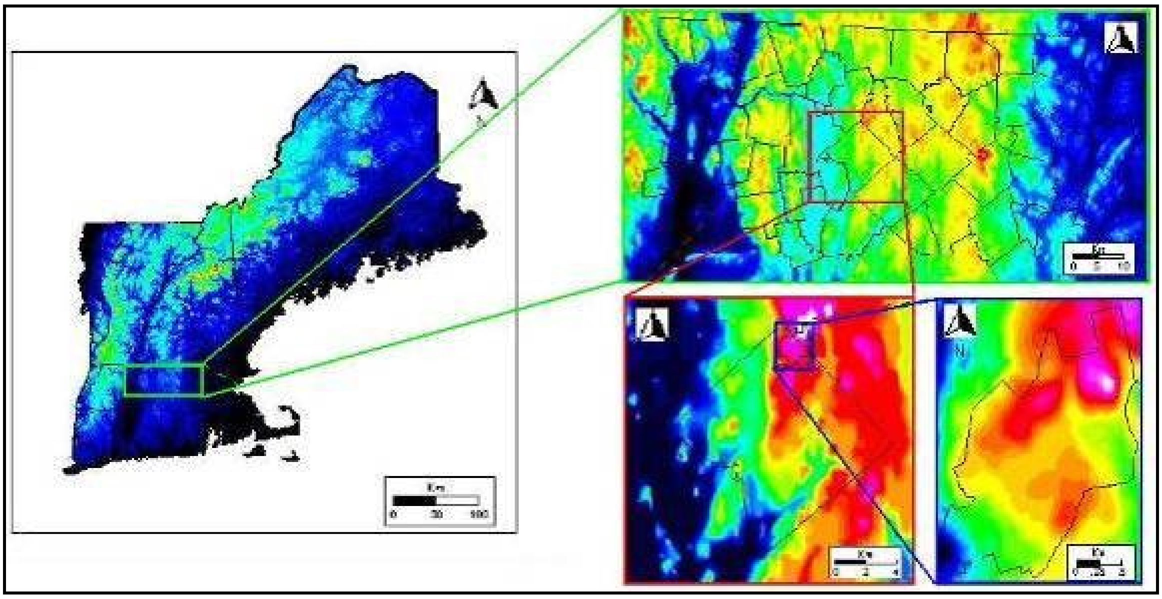


Figure 2.4. Information is more detailed every time it is zoomed in, moving from regional scale 1cm = 50km to sub regional scale 1 cm = 10 km to site scale 1 cm = 0.5 km. (© USGS – Land use History of the North America).

In the field of geography there are many examples on how researchers faced and dealt with scale as mapping and analytical issue. For example Weiers et al. (2004) applied a satellite data and raster GIS approach to assessing habitats at different scales. Patil et al. (2000) used a more complicated synoptic multivariate spatial data to provide a multiscale assessment of landscape and watershed. Gaucherel (2007) proposed a method to capture the local and scaling variations of the landscape. He started from the fact that the Shannon diversity index, which is widely used to quantify landscape diversity, provides different results according to scale. Thus the method by Gaucherel combines a multiscale diversity map, with a diversity profile averaging at each place and scale, to give the spatial information contained in the landscape. The results show a map of landscape diversity with directional trend and local information. Similarly Purtauf et al. (2005) tried to quantify the variation of patterns of land-use composition parameters through a multiscale analysis and multivariate statistics. Their conclusions highlighted that there is not an optimal scale for characterising and comparing landscape patterns but a range of sub-optimal ones. The results from

Purtauf et al. were interesting given the similarities with those of this current research, namely that each scale used for landscape characterisation leads to different outcomes.

Landscape however, doesn't have only a spatial dimension; in fact being dynamic and in continuous evolution it exhibits heterogeneity and changes **throughout the time** (see figure 2.5). In other words the landscape shows different aspects at both spatial and temporal scales. Even details like season and timing are not minor because a landscape responds differently to the cycles of the nature. As suggested by the LCA guidance, more than one season should be covered while assessing the character of a landscape (Swanwick and Land Use Consultants, 2002).

Figure 2.5

The literature provides again a series of examples of studies about landscape change (in particular land-use change) based on the use of satellite images or interpretation of aerial photo. Apan et al. (2002) used two sets of LANDSAT imagery, dated 1973 and 1997, and classified them independently in order to obtain land-cover and land-use maps. These were subsequently utilised as input to GIS-based landscape pattern calculation software to generate landscape structure indices. The use of two different temporal scales allowed the authors to gain information on the changing/developing state of the studied landscape. Awareness of the change in the landscape structure and its quantification, e.g. loss of 20% of woodland converted into pasture, provided land managers with the essential information to think of plans and strategies and make appropriate decisions (Apan et al., 2002).

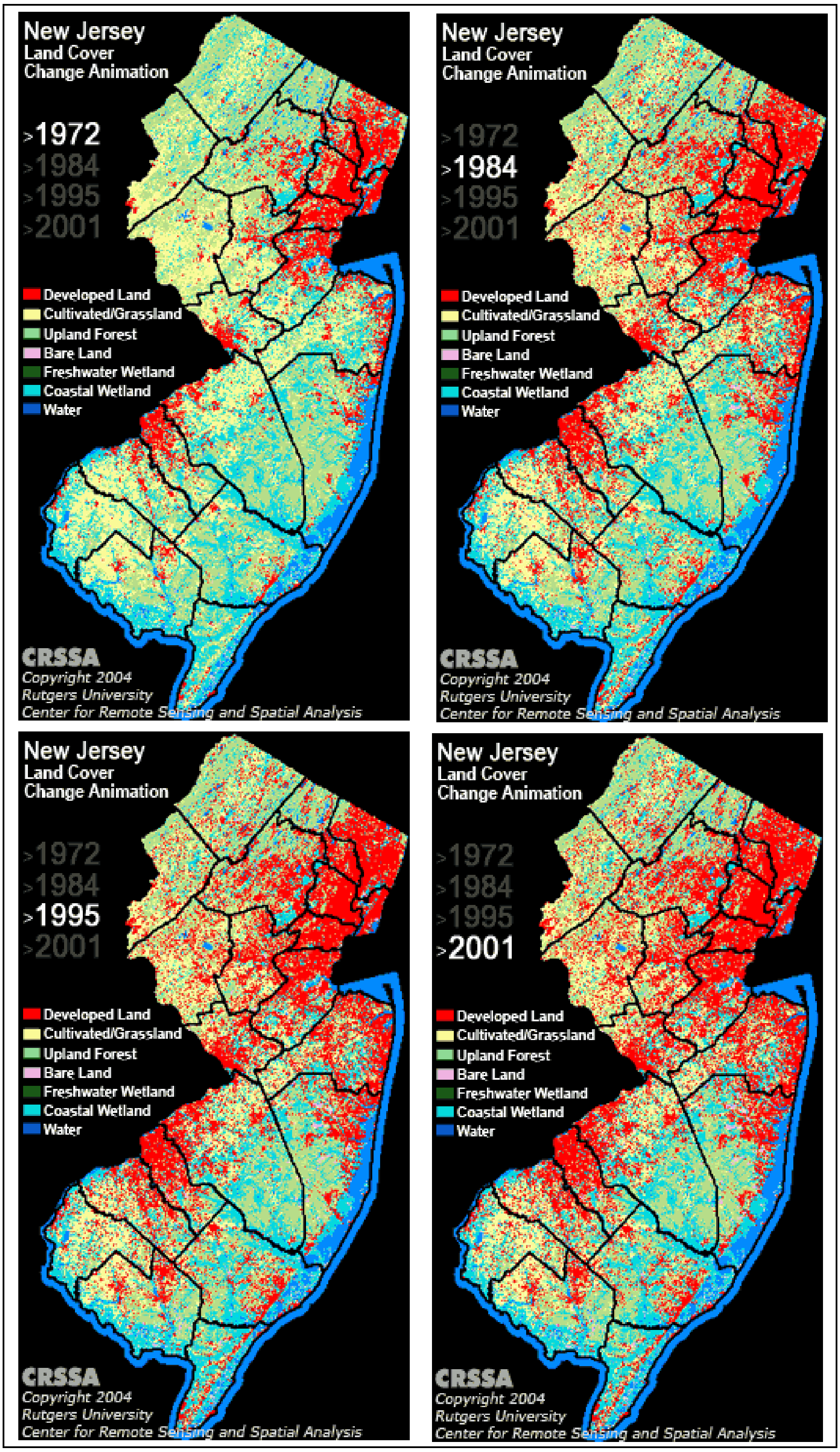


Figure 2.5. Spatial scale provides a different detail of information whereas as showed in the picture temporal scale helps understanding dynamics and changes occurred to the landscape. In this case the land cover change in New Jersey has been analysed from 1972 to 2001 using remote sensing and digital geospatial technologies (© Rutgers University – New Jersey).

Van Eetvelde and Antrop (2004) referred to aerial photographs covering a period from 1960 to 1999 with the intention of analysing changes at settlement level. The aerial images were scanned and then imported in GIS where could be analysed. The authors concluded that when it was possible to assess quantitatively the change in the overall landscape character, it was hardly feasible to measure the change in the different structural components of the settlements. This conclusion confirmed that the temporal scale helped revealing changes in the landscape but that the spatial scale (of the aerial photographs) was not appropriate to answer all the research questions posed by the authors. Ayad (2005) proposed a combined use of remote sensing and GIS in order to model visual landscape change in the north-western arid coast of Egypt between 1950 and 1990. The author wanted to provide the local land-use planners with valuable information on how the changes in the scenic characteristics of the landscape could be related to badly planned activities and could compromise the tourism of the area.

From the literature it emerges quite clearly that the temporal scale allows us to gain a better knowledge of the dynamics and interactions between people and resources. As mentioned in the first chapter, besides LCA two other independent programmes started in the mid 90s and were centred at the assessment of the historic landscape character. These are the English Historic Land-use Characterisation (HLC) and the Scottish Historic Land-use Assessment (HLA). The main concept at the basis of both assessments is time-depth which can be defined as the long sequence of events and actions (both natural and human) that have produced the present landscape (see topic paper 5 LCA guidance). In this research the past as explanatory variable of/for the present landscape character was taken into account in the analysis by using the HLA map. Therefore, the rich database of the HLA filled effectively the gap due to the lack of available satellite images and aerial photographs.

Another essential part during the definition of the scope of LCA is to judge people, skills and time needed for completing successfully the work. In general the people who undertake the LCA are professionals both landscape architects and landscape consultants; nevertheless it is good practice to include input from stakeholders. The European Landscape Convention stresses particularly the importance of people's perception in the definition of landscape and the introduction of new tools and working practices in public sector organisations (e.g. exchanging good practice, e-government, and local communities' action groups) has corresponded to a growth in importance of stakeholders within the process of LCA. Therefore as illustrated in figure 2.3 for each step of landscape characterisation there is space for a possible input from the stakeholders. Indeed, the public participation has to be planned while defining the scope, and it is necessary to clarify which range of stakeholders has to be involved, at which stage and with which method they will be active participants to the LCA.

If the scope of the LCA is stated clearly it is easier to recruit people with the necessary and appropriate skills, thus for instance the recently completed LCA for the Cairngorms National Park was carried out by a team of experts in geology, history and by landscape architects. The importance of having people with complementary skills also emerges and this to some extent is not a surprise since the landscape is a complex entity which requires different experts in order to be analysed. In addition, the last 20 years have been marked by an improvement in access to software such GIS, mainly employed as mapping tools since they are very helpful in handling different layers of mapped information. Although everybody recognises the importance and effectiveness of GIS, these modern and advanced tools do still look "unfriendly" and the amount of time required for training people to a correct GIS use and application is considerable. Furthermore, GIS so far are applied only as mapping tools and very little of their analytical capacity is exploited. This is a reason why the current PhD was conceived and funded by the Scottish Natural Heritage and the references amongst the literature reviewed were so extraordinarily scarce that the ideas for the investigation came mainly from landscape ecology, genetics or econometric rather than from previous works on landscape character assessment.

To summarise, the definition of the scope is the essential first step of landscape characterisation since it helps to identify two pivotal points for the assessment, such as scale and detail of the data, and to make a cost analysis in terms of human and technical resources and time needed. This preliminary step ends with a brief which should serve as a summary and reference text/note throughout the assessment (Swanwick and Land Use Consultants, 2002). In addition the guidance suggests a field visit to the area where the landscape is to be assessed; this sort of familiarisation to the area of study should help those involved in the assessment to learn more about the character of the landscape.

The second step of the characterisation is the **desk study** and this involves a review of data and literature available about the area of study which fundamentally should enable the team of landscape assessors to gain a better understanding of the landscape of the study area.

The analysis of relationships between different landscape elements, such as geology, land cover and settlement patterns, is central in order to identify, classify and describe the landscape character types/areas both with the standard and with the GIS approach (see figure 2.6). The landscape elements are basically spatial data represented by maps either in digital or paper format; indeed where the collected maps result in digital format then the process of storage, manipulation, analysis and presentation is facilitated by using GIS (Swanwick and Land Use Consultants, 2002). Archives and other documents such as historic photographs, literature on local architecture, approval of designation areas, statutory development plans, countryside and forestry strategies, provide another useful source of information.

LANDMAP Information

Main information types	Aspects and themes of information	
Contextual Layers	Landscape form	A description of the form of the landscape: e.g. "broadleaved woodland"
	Landscape function	A description of the primary function of the landscape e.g. "Amenity"
Evaluated Layers	Geological landscape	Description of the nature of the landscape within the hierarchical system. Each area is evaluated for quality, value and condition
	Landscape habitats	
	Visual and sensory	
	Historical landscape	
	Cultural landscape	
Socio-economic	People	National coverage of socio-economic information is provided within *LANDMAP*
	Economy	
	Wealth	
	Health	
Public Perception	"Top down" information	
	"Bottom up" information	

Figure 2.6. The desk study is mainly focused on the collection of data available. The figure illustrates an example of layers used by the Countryside Council for Wales to compose LANDMAP (© Countryside Council for Wales)

It becomes evident that the amount of data that should be collected and analysed for the landscape characterisation is quite remarkable. The guidance suggests that at the beginning in case the resources are limited and the time is short, it could be sufficient to retrieve data on geology, landform, land-cover and settlement distribution. These datasets should be then integrated at least with the historic land-use information as the resources and time become available (Swanwick and Land Use Consultants, 2002).

As previously mentioned history plays a fundamental role in the landscape characterisation because it offers a valuable background of information on the past land-uses. Nevertheless understanding the historic dimension of the landscape is not straightforward and it requires an expert analyst, possibly an historian or archaeologist. The HLA map comes with a glossary that is very useful and rich in explanations, in addition people who are in charge of the HLA, and namely a team of cartographers and surveyors at the RCAHMS, are available to provide clarification and more details about their work and the way of interpreting it.

Regardless the amount of data collected it is crucial, above all if dealing with GIS, to evaluate the quality of the data and the common criteria in use are: actuality, completeness, consistency, accessibility, accuracy, precision and source of errors caused by data entry and manipulation (Burrough and McDonnel cited in Antrop and Van Eetvelde, 2000). It is worthwhile remembering that checking data quality is essential since it might happen that the available data layers differ in several of the above mentioned criteria and some of the criteria might assume more importance than others. For example strongly dynamic and heterogeneous landscapes such as urban fringes require a high level of actuality and accuracy or, if using satellite images, high spatial resolution (Antrop and Van Eetvelde, 2000). Hence it is important to invest time in considering the characteristics of the data collected, and an inventory (a note) of the information available in digital or paper format is strongly recommended.

Moreover, the guidance suggests that it is useful to have a good aerial photograph of the area of study since it adds a "bird's eye" of the whole area of investigation. Nowadays aerial photographs can be retrieved from Google Earth and photograph of Britain for every grid square are available at the website www.geograph.org.uk. In the first case it is good practice to double check both the year and the quality (pixel resolution) of the images.

If the scope of the LCA requires the presence of stakeholders then it might be possible to add their input at this point of the characterisation, because their observations and opinions can provide a significant contribution in understanding which unique and distinct association of elements makes a landscape different and makes people feel the “sense of place”.

In the literature reviewed there are several interesting works that try to define the character of the landscape on the basis of what people feel and think. For instance, De la Fuentes de Val et al. (2006) focussed on how landscape preference is related to landscape spatial patterns using a sample of 98 people and 8 landscape photographs. For each photograph, 11 visual attributes were evaluated and 3 windows were defined to cover the different areas corresponding to foreground, mid-ground and background visual fields. Then the landscape spatial structure of the windows was analysed using spatial metrics and the correlation between each dimension and the spatial pattern indices of the landscape were calculated (De la Fuentes de Val et al, 2006). The results suggested that heterogeneity and a certain level of complexity (calculated as diversity and composition of elements in the landscape) influenced positively people’s perception of the landscape. On the contrary scenic beauty showed a limited correlation with landscape pattern indices because beauty is more a concept and it is hard to find an index that quantifies it. An attempt to define in spatial terms the concept of wild land in Scotland is provided by Habron (1998) who based its research study on people’s perceptions of the landscape. Habron used a photographic questionnaire in order to gather public perception of wild land both from people living in the areas where the pictures were taken and from potential visitors. The photographs represented a range of characteristic landscape attributes of the Cairngorms and Wester Ross which were rated for their wildness. The extent of visible landscape attributes was quantified using GIS.

Another example of the involvement of people in landscape analysis is given by Scott (2002) who worked on the development of a methodology to identify the public perception of landscape in Denbighshire (Wales). His results demonstrated that people’s perception can allow particular landscape types to be evaluated in qualitative and quantitative terms. In addition the strong attachment to managed rural landscapes, which was revealed by people, was interpreted by the author as very useful information that can help planners and policy makers to rethink their approaches to conventional landscape management strategies and planning. Ode et al. (2007) explored the relationships between landscape preference and three indicators of naturalness that were identified by computer generated visualisation of landscape containing pastures and broadleaved woodland. In total, 27 different visualisations were showed to 703 respondents and the results highlighted a high correlation between preference for number of woodland patches and level of succession, while a weak correlation was found between preference and the shape index of the edges.

Because people were not involved in this research, the developed GIS-based methodology characterises the landscape only on those elements that can be associated to digital maps. More precisely, GIS extract information from maps that metaphorically function like our eyes. In fact while we can observe a landscape and feel it through our senses, GIS can read the landscape through information that is stored digitally.

Consequently it is understandable that only data that can be measured or quantified with indices or other criteria can be mapped and used by GIS. This aspect of GIS can be both a disadvantage, because it limits the landscape characterisation to a smaller range of data, and at the same time an advantage since it ensures a quantitative analysis.

Once the digital information has been collected, the different layers can be overlaid; this operation facilitates the detection of the correlations between the various map elements and consequently the identification of the areas of common character. Nowadays maps overlay is achieved by using GIS, because most of the data is in digital format; however there might be cases that a manual or judged by eye analysis is required.

The level of information about the character of the landscape achieved at the end of the desk study depends entirely on the quantity and quality of the data collected and indeed on the brief written after the definition of the LCA scope. If the identification of areas of common character is a hard task and difficult to achieve, then the guidance suggests to prepare a map of areas for survey that are supposed to have a distinct character and verify the assumption with the field survey, which in any case is going to be the fundamental and unavoidable third step of the characterisation (Swanwick and Land Use Consultants, 2002).

Independently of the level of knowledge about the landscape reached with the desk study, the field survey is essential and plays a central role because is the best opportunity that the surveyors have to collect all the information necessary to:

- describe the character;
- identify aesthetic and perceptual qualities;
- assist in the final decision about the division into character types and areas;
- update and expanding the data gathered during the desk study;
- contribute to the process of making judgements about the future of the landscape (Swanwick and Land Use Consultants, 2002).

The field survey to some extent helps to validate the data collected during the desk study, particularly if satellite images or aerial photography had been used as reference.

According to the guidance for each area, that is assumed to have a distinct character, at least three points of observation of easy access and well representative of the landscape should be selected, and the information collected from the visual assessment should be annotated on a field record sheet which is exemplified in figure 2.7.

LANDSCAPE ASSESSMENT – FIELD SURVEY SHEET

VIEW POINT NO: **DATE:** **GRID REF:** **HEIGHT:** **PHOTO:**

LOCATION: **CHARACTER AREA:**

GEOLOGY: **SOILS:** **WEATHER:**

BRIEF DESCRIPTION OF VIEW POINT LOCATION:

OBJECTIVE ASSESSMENT:

Landform	Gorge Plateau Hummocky	Narrow Valley Sloping Scarp	Broad Valley Undulating Ridge Steep	Hollow Rolling Vertical	Flat Outcrops	Terrace Hills

Line

Horizon and X-section	Notes

Landcover	Peat Bog Improved	Marsh Meadow	Moor Cereal Crops	Scrubland Root Crops	Rough Grassland Forestry	Semi-improved

NOTES:

Woodland

DECIDUOUS		CONIFEROUS		MIXED		INDIVIDUAL TREES		NOTES
Plantation		Plantation		Plantation		Deciduous		
Woodland		Woodland		Woodland		Coniferous		
Shelterbelt		Shelterbelt		Shelterbelt		Along boundaries		
Clumps		Clumps		Clumps		Scattered within fields		
%		%		%		%		

Farming/agric	Arable Open	Pasture Enclosure	Intensive Livestock Common Land	Rough Grazing
Field Size	Small	Medium	Large	Very Large

Settlement

Settlement Type	Form Nucleated/Scattered/Linear	Name and Dominant Building Materials
Farm stead		
Individual residential		
Hamlet		
Village		
Town		

SUBJECTIVE ASSESSMENT:

Scale	Intimate	Small	Moderate	Large	Vast
Enclosure	Confined	Enclosed	Semi-Enclosed	Open	Exposed
Diversity	Uniform	Simple	Diverse	Complex	
Colour	Monochrome	Muted	Colourful	Garish	
Balance	Harmonious	Balanced	Discordant	Chaotic	
Accessibility	Wild	Remote	Easily Accessible		
Pattern	Random	Organised	Regular	Formal	
Visual Dynamic	Extensive	Dispersed	Channelled		
Management	Derelict	Neglected	Tended	Manicured	
Quality of Light	Bright	Shaded	Dark		

Tree Cover ____________________

Detractors	Offensive Smoke Pollution	Unpleasant Industrial Noise	Depressing Intrusive Lighting	Bland Intermittent Traffic Noise	Constant Traffic Noise

OTHER: ____________________

Stimulus	Pleasant	Interesting	Attractive	Beautiful	Invigorating	Stunning

OTHER: ____________________

Additional Comments:

Figure 2.7. Examples of field survey sheets used by the surveyors, generally two, during the observation carried in the field. It is noticeable the distinction made between the objective and subjective part of the assessment (source: East Riding of Yorkshire Council).

Amongst the LCA reports that have been consulted for this PhD it emerged that the field record sheets did not have exactly the same layout; nevertheless all of them showed approximately the same scheme and reserved the space for:

- a written description of the character observed at particular points or in certain areas, that should capture the overall impression of landscape character;
- an annotated sketch, that should draw how the different elements of the landscape interact together;
- a checklist that includes both landscape elements, their significance and aesthetics and perceptual factors. The list should act as aid memoir for the surveyors;
- a place for observation about the condition, sensitivity and management needs of the landscape.

In this way, as specified in the guidance, the surveyors were encouraged to make systematic observations of the landscape and to record them in a consistent way (Swanwick and Land Use Consultants, 2002). However, the consultation of several LCA reports revealed that the field survey could not be wholly objective and value free. Aesthetic and perceptual factors inevitably involve subjective judgments and evaluation, and in fact, as noticeable in figure 2.7b, they form the “subjective” assessment. In some occasions field record sheets were found asking the surveyors to quantify the physical and cultural elements of the landscape, for example in figure 2.7a the surveyors are asked to assess the percentage of woodland coverage, and in figure 2.8 there are terms such as “shallow valley”, “undulating”, “rolling” that are neither quantitative nor add more objectivity to the assessment.

Elements of objectivity can in fact be introduced only by incorporating criteria that measure mathematically and unambiguously the elements of the landscape. The GIS-based methodology aims at achieving this result and the fact that GIS work with maps which represent quantitatively the geographic objects is a promising starting point. In any case, the aesthetic and perceptual aspects of the landscape represent the most difficult challenge for GIS because defining indices or parameters that quantify concepts such as remoteness, unity, texture or form it is far to be straightforward. However, details about the way the GIS-based methodology has tried to tackle them are provided in chapter3.

DRAFT LANDSCAPE TYPE ____________________

DRAFT LANDSCAPE CHARATCTER AREA: ____________________

Keywords describing the landscape: ____________________

PHYSICAL FEATURES

ELEVATION

☐ Lowland (<50m)	☐ Transitional (50-75)	☐ Upland (>75m)

LANDFORM

☐ Flat	☐ Steep slopes	☐ Escarpment	☐ Broad valley
☐ Shelving	☐ Gentle slopes	☐ Knoll	☐ Narrow valley
☐ Rolling	☐ Floodplain	☐ Plateau	☐ Shallow valley
☐ Undulating	☐ Hills	☐ Coomb Valley	☐ Deep valley

WATER/HYDROLOGY

☐ River (S/M/L)	☐ Stream/tributary	☐ Flooded gravel pits	☐ Engineered/artificial
☐ Speed (F/M/S)	☐ Drainage channels	☐ Lake	☐ Locks/Weirs
☐ Clarity (C/M)	☐ Drainage ditches	☐ Ponds	☐ Other
☐ River Meanders?	☐ Spring	☐ Bog	☐ Other

LAND/VEGETATION COVER (INDICATE RELATIVE %)

☐ Arable	☐ Amenity grassland	☐ Small farm woods	☐ Heathland
☐ Perm. pasture	☐ Conif. plantation.	☐ Shelterbelts	☐ Scrub
☐ Pasture	☐ Christmas trees	☐ Copses/ clumps	☐ Wetland/Aquatics
☐ Ley/improved	☐ Decid. woodland	☐ Woodland belt	☐ Gardens
☐ Rough grazing	☐ Mixed woodland	☐ Hanging woodland	☐ Common
☐ Wet Meadow	☐ Parkland	☐ Scattered Trees	☐ Green
☐ Chalk grassland	☐ Avenues	☐ Hedgerow trees	☐ Paddocks
☐ Set-aside	☐ Orchards (type..)	☐ Hedgerows	☐ Other

Notes on ecological character ____________________

LAND USE

☐ Farmland	☐ Residential	☐ Commercial	☐ Natural
☐ Forestry/woodland	☐ Industrial	☐ Transportation	☐ Military
☐ Historic Parkland	☐ Leisure/Recreation	☐ Mineral Working	☐ Other

Figure 2.8. Example of field record sheet specific to the assessment of physical aspects of the landscape. Terms as rolling, undulating or small, medium and large(S/M/L) river are attempts made in order to quantify the landscape elements. However, these terms are disputable/arguable and do not make the assessment more objective (source: Land Use Consultants).

Photographs are other recommended recording tools available to the surveyors: these supplementary databases of information can result very useful in case the landscape is particularly subjected to forces for change and likely to be modified. In these specific situations it appears quite evident how photographs and the other outcomes from the field survey contribute to the making judgement phase. For example data on the condition of the landscape elements, the evidence and the causes of change can be relevant and inform subsequent decisions (Swanwick and Land Use Consultants, 2002).

Once the field survey is completed and an exhaustive set of information about the landscape has been put together then it is possible to move to the last step of the characterisation which comprises of classifying the landscape into character types/areas, mapping their extent and describing their content (Swanwick and Land Use Consultants, 2002).

For landscape researchers it is important to classify the landscape character in order to have a frame of reference for communicating and comparing their research. Generally classification is a difficult task because of the complex nature of landscapes and because it is meant to be explicit (Brabyn, 1996). On one hand the field survey helps the surveyors to retrieve information about the way the elements occur and are distributed in the landscape. On the other, the classification implies the identification of patterns in the landscape and the definition of a set of criteria upon which to base their recognition and recording. In this regard, the word "patterns" indicates the way the natural and cultural elements interact to each others and form the character of the landscape.

As mentioned at the beginning of the chapter, the character of the landscape is classified into types and areas; the former can be repeated in an area as the types show broadly similar patterns of geology, landform, soil and vegetation. In contrast, the latter have a unique and individual identity and appear only once in a geographic location.

On this basis it emerges that a classification method is required more for the landscape character types rather than the landscape character areas and it is a general rule that the former are identified using terms like for example “rocky moorland”, “upland glens”, “open farmland”, whereas the latter are named after geographic place names like “Strathspey”, “the Trossachs”. For instance figure 2.9 illustrates the way the early Cairngorms National Park LCA classified the landscape character into types and areas.

Figure 2.9

From the map we gather that there are three character types, namely plateaux, straths and upland and glens which comprise of several character areas. For example “Badenoch”, “Strathspey”, “Muir of Dinnet” are areas associated to the type strath.

Then, as far as the Scottish LCA is concerned, the description of each landscape character type covers context, geology, landforms, water, land-use, land-cover, settlement, other features, landscape experience and pressures for change. As suggested by the guidance the description should capture the essence of their character and recognise the forces for change such as key development pressures and trends in land management. Usually the description is written and accompanied by illustrations.

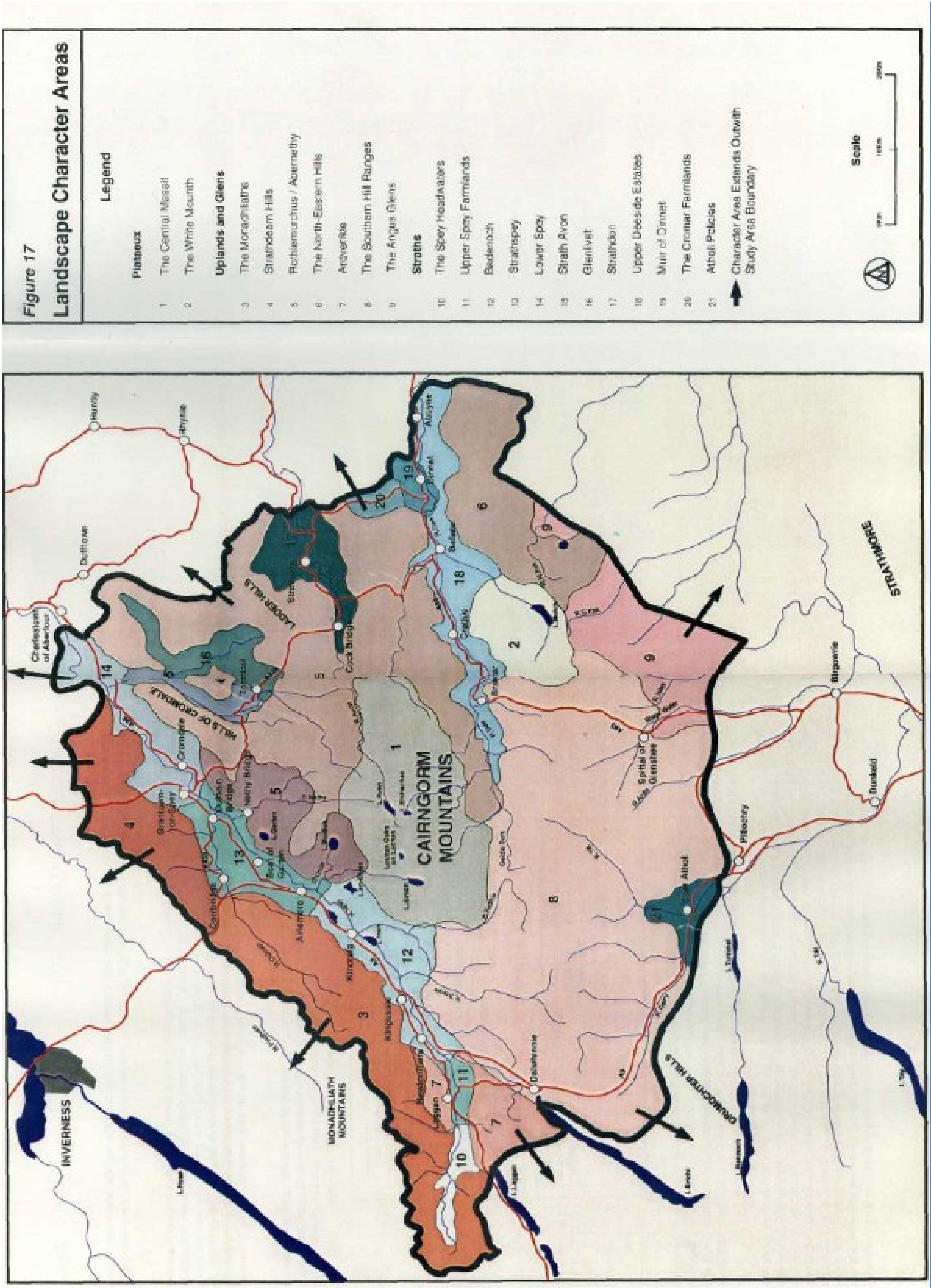


Figure 2.9. Maps representing the landscape character areas of the Cairngorms National Park (numbered from 1 to 21). The landscape character types can be derived from the legend and are written in bold (©LCA of the Cairngorms National Park - 1999 edition).

In order to clarify what a description looks like figure 2.10 is an example of landscape character type and area of Aberdeenshire local authority.

5 ***AGRICULTURAL HEARTLANDS LANDSCAPE CHARACTER TYPE***

5.1 ***DESCRIPTION OF LANDSCAPE CHARACTER TYPE***

The **Agricultural Heartlands** landscape character type contains ten landscape character areas. The predominant landscape of the study area is an agricultural one of generally intensive mixed farming on large fertile fields. Essentially it therefore represents a quite recent landscape character type whose character relies less on historic features and landuse practices than other landscapes.

There are three main centres of such farming within the study area. The coastal lowlands form the most extensive area which, despite their diverse lithology, have been eroded into an undulating plateau between Formartine in the north and the coastal slopes south of the River Dee. The second area fringes the northern edge of the wide rift valley of Strathmore, extending between the River North Esk and Stonehaven. The third comprises a belt of generally flat lowland that curves between Huntly and Oldmeldrum through Insch between surrounding high ridges of **Moorland Plateaux.**

Figure 2.10. Extract from the South and Central Aberdeenshire LCA report. Description and map provided for the landscape character type “Agricultural Heartlands”. As mentioned in the text 10 character areas are nested into the character type, and for all of them a specific description is provided (©LCA of South and Central Aberdeenshire - 1998).

Some of the recommendations provided by the LCA guidance are about avoiding repetitions and subjective judgements and suggest the writer be balanced between factual statements about the components which make up the landscape and more evocative statements about aesthetic quality and perception. If the stakeholders were involved in the LCA process then their opinions and quotes, which highlight what an area means to local people and/or visitors, should be added.

The descriptions should be followed by a list of the key characteristics of each landscape character type/areas. The key characteristics represent the combination of landscape elements that give an area its distinct sense of place. Therefore they are extremely important and should be the first to be considered when monitoring landscape change or identifying landscape indicators.

Finally the main advice given in the guidance is to write the description bearing in mind the scope of the LCA and that the target is to inform planning policies. In fact at the end of the characterisation phase the outputs should provide a good base to facilitate the judgements necessary to inform decisions on landscape and environmental management, conservation, restoration and enhancement.

Before moving to the introduction of how the GIS-based methodology was born, developed and tested on different areas of study it is opportune to conclude this chapter with a mention of the second part of the LCA, which is making judgements. Although this topic has not been explored by the PhD it is relevant to describe it concisely since it helps to have a complete overview and understanding of the reasons for conducting a LCA.

One of the duties of the scientists is to ensure that policy and decision makers are provided with the best available information. Strengthening capabilities for landscape management through raising awareness that landscape is a natural resource and part of our heritage is one of the aims of the LCA.

Most assessments, for this reason, will usually move beyond the characterisation stage to that of making judgements to inform particular decisions. Maintenance of existing character is one part of the decision to be made. The focus should be on ensuring that change and/or the development proposals are planned and designed to achieve an appropriate relationship with the surroundings, and wherever possible contribute to the enhancement of the landscape (Swanwick and Land Use Consultants, 2002).

The method of making judgements based on the results of the characterisation stage will vary depending on the particular issue that is being addressed (Swanwick and Land Use Consultants, 2002). This in turn will reflect the scope of the assessment. Therefore the judgements can be either specifically related to decision making based on landscape character or can be designed to contribute to wider environmental decision making tools, where landscape is only one of the topics addressed. Once the approach to the judgement has been made explicit and transparent then it can be translated into:

- *strategies,* based on what change is thought to be desirable for a particular landscape character type. The judgements require a certain degree of transparency and should be devised and tested through the stakeholder involvement. In the field of planning policy the strategies help providing basis for landscape and development policies, identifying areas for formulation of landscape status or designation;

- *guidelines* are usually designed to influence the way landscape are managed. Thus, it is desirable that many stakeholders involved in the day to day management participate actively in the process of LCA. This will help to ensure that the guidelines are not vague but based on the "real world", have a good understanding of land uses and land management practices, and can be practically implemented;

- *landscape capacity studies*, are about the ability of an area to accommodate change, either as a result of new development, or some other form of land use change. The judgement must be based on an understanding of the ability of the landscape to accommodate change without significant effect on its character. In planning policies or landscape strategies, criteria for what constitutes significant change need to be identified and will be informed by potential effects on the character of the landscape;

- *attaching status to landscapes*, occurs when a tract of landscape is selected for special recognition; therefore judgements need to be based on a range of different key considerations:

    - natural beauty, that encompasses flora, fauna, geological and geographical features;
    - recreational opportunities, that are those afforded for open air recreation, having regard both to landscape character and position in relation to centres of population;
    - natural beauty and amenity, that is a combination of terms covering the physical landscape and less tangible aspects such as remoteness, tranquillity and other aspects that appeal to senses such as sight, sound and smell.

From this general overview it emerges that the approach to making judgements is based not only on landscape character, but also on landscape quality, value, capacity and on the involvement of stakeholders. Despite the degree of subjectivity and value that are attributed to the judgement, there are examples of how GIS-based decision support tool can be successfully used and implemented in the decision making process and for participatory landscape planning. Moreira et al (2006), Tress and Tress (2003), Sisk et al. (2006), Mander et al. (2005), Weiers et al (2004) and Palmer (2004) are few examples.

2.2 LCA programme: lights and shadows.

This chapter has described so far the conceptual model of LCA. An explanation of the way the characterisation is tackled by the current approach to LCA was provided. It emerged that in order to identify the character of the landscape there is a series of intermediate steps which start with the definition of the scope of the LCA and end with the description and mapping of the landscape character types and areas. The chapter ended by showing that LCA is required in order to provide information to planners and policy makers and support them to make appropriate decisions in terms of landscape management and planning.

With the description provided above, the conceptual model of LCA seems compelling, credible and well structured. However, between the lines and throughout years of applications, some weaknesses and a series of issues have been detected and, for the general improvement of the programme, require now to be solved. Here are highlighted the main limitations of LCA and the way this research has been conceived in order to address them. It is worth noticing that the description of the LCA weaknesses is confined to the characterisation phase since this is the research focus.

The main limitations of the current LCA programme can be briefly summarised as it follows:

- in the field survey sheets, the terms that refer to the landscape elements are not universal and unambiguous; thus the principal consequence is that the results from the assessment are first of all different, according to the field survey sheet used, and secondly arguable, since a ”narrow valley” or a “smooth hill” can be interpreted differently by the assessors;
- the landscape character types and areas are described in a narrative form which on one hand implies subjectivity in the analysis and on the other had does not provide a strong support when decisions about planning issues have

to be made. The reason is because planners and developers are more accustomed to work with numbers and need to be able to quantify the impact of a proposed plan on the landscape;

- inconsistencies in terms of identification and description of landscape character are evident along the boundaries of the local authorities. This is a consequence of the lack of standard terminology in the field survey sheets, the different interpretation of the guidance given by the assessors, and their diverse methodology of work;
- ecological and historical aspects of the landscape are rarely assessed, therefore the identification of landscape character misses usually two important data;
- the LCA reports requires months of work and are not updatable annually. They lack of dynamicity and their style at time appears to be old-fashioned.

In order to face the above mentioned LCA limitations, this research aims at developing a new approach to landscape characterisation based on the use and application of GIS and statistics. These two are analytical tools and techniques that consider the landscape elements as measurable objects and ensure objective outcomes since the personal judgement and interpretation during the calculations is sufficiently contained.

In addition, GIS and statistics should also resolve the problem of inconsistency; in fact the target is to implement a methodology of analysis that can be repeated exactly in the same way throughout Scotland, regardless the differences between the areas of study. To be successful, the methodology should be able to detect the various landscape characters by following the same rigorous steps that will be defined only once in designing phase.

Another goal of the research is to attribute more importance to the historic aspects of the landscape by adding to the datasets the Historic Land-Use Assessment map whose rich database is meant to be used more extensively and appropriately.

Finally GIS characteristics to be interactive, flexible and easily updatable will be fully exploited in order to improve the outcomes and make them looking more dynamic and up-to-date.

The next chapter describes more in details the major limitations of the LCA programme, the importance and contribution that GIS may give to the enhancement of LCA. The chapter will show that the new methodology has links with the current LCA and is not conceived to replace the work carried by landscape architects or other practitioners. Rather, the use and application of GIS and spatial statistics are meant to become a support tool for both professional figures.

CHAPTER 3

The steps towards the definition of a quantitative and GIS-based methodology for landscape characterisation

The tors of Ben Avon from Beinn a Bhuird (I.Marengo).

3.1 The need to develop a new, more consistent, approach to LCA

As far as the Scotland national programme of LCA is concerned, it took approximately three years for SNH teams and external landscape consultants to undertake the 29 LCA which covered Scotland and over this time the methods between LCAs varied. Hence, as expressed by David Tyldesley and Associated who were involved in several LCA initiatives, "*it is not surprising to find that the LCA vary in many ways, both in respect of presentation and emphasis*" (Julie Martin associated and Swanwick, 2003). Nevertheless time needed for the completion of LCA and differences in LCA scopes cannot justify entirely some of the weaknesses and problems showed by the current LCA approach.

In fact landscape architects, practitioners and other professionals operating in the field of landscape assessment often were, and still are, confused by the inconsistent way the LCA is tackled and its results are presented and the question whether or not the approach taken for the LCA is right it is frequently posed (D. Carman, 2007).

In 2003 SNH commissions an overview of the LCA national programme which aims at highlighting the main strengths and weaknesses of the 29 reports, the way they had been used and at making recommendations for future works (Julie Martin Associates and Swanwick, 2003).

The comments about strengths and weaknesses are made on the basis of the degree to which the programme has met its objectives and on the comparison with good LCA practice in other countries (Julie Martin Associates and Swanwick, 2003).

According to who was in charge of the review, LCA shows several strengths, for example:

- it is the key tool for SNH staff to use in fulfilling SNH's landscape duties and remit, providing an inventory of the Scotland's landscapes;
- it has achieved a formal recognition in policy and advice from the central government;

- it has been widely used for a large number of different applications;
- its outputs are highly recognised and used by planners for development planning and control throughout Scotland.

In addition, LCA was the first full-coverage, detailed programme to be completed in Europe with the participation of local authorities and other partners across Scotland. Amongst the use of LCA in phase of planning, it seems that capacity work is very advanced and greatly regarded by all the local authorities that have been consulted during the process of LCA review. Moreover, as other good points of the LCA they highlighted:

- the fact that the programme raised awareness of landscape issues among other agencies, planners and developers;
- the legitimacy of landscape considerations above all in the context of planning decisions;
- the fact that the programme is cost-effective use of SNH research funding.

However, the LCA, despite meeting the majority of its objectives, presents a number of weaknesses. In fact it has not fully responded to a series of points relative to:

- increased awareness of Scotland's landscapes beyond the group of those most actively involved in the programme development;
- consistent and reliable identification of forces for change in Scotland's landscapes. LCA often provides a list of the forces for change that are wordy and strongly influenced by the perspective and the outlook of the surveyor. This makes them ineffective;
- informing national policy on issues relating to landscape interests;

- limited stakeholder input to the LCA preparation. The involvement of stakeholders was in fact restricted to those participating in the steering group for each assessment;
- limited range of LCA applications in landscape conservation and management;
- unclear distinction between the more objective characterisation stage of the assessment and the more subjective stage of making judgement.

In addition the contractors responsible for the review point out more technical weaknesses.

First of all the programme emphasised the landscape character types at the expense of the landscape character areas. A better balance would be beneficial as the difference between types and areas is more conceptual than "spatial": the latter are unique and explain better what is distinct in the landscape, communicating landscape character to non-specialists (Julie Martin Associates and Swanwick, 2003).

Secondly, a less strong national and regional characterisation has been developed in Scotland compared for example to England as a consequence of the preference for the "bottom-up" approach rather than the "top-down" in phase of mapping and classification. To some extent it appears that the national classification process evolved rather than being planned from the outset.

The classification adopted, moreover, needs to be improved looking at a more rigorous control and standardisation. At level 2 and 3 it is hard to distinguish clear pattern of highland, upland and lowland and the terminology is not very meaningful to most of the people (Julie Martin Associates and Swanwick, 2003).

Third, information on historic land-use and habitats are sometimes poorly covered and are not reflected within the LCA reports. However, historical and ecological aspects are part of the landscape and can contribute to our understanding and describing of the landscape character. In this regard data and digital maps from the Historic land use Assessment are highly recommended to be incorporated into the LCA.

Finally variation between LCAs in scale, detail, methodology, classification and description causes inconsistencies of approach and content especially across the boundaries. This is a major issue which has been raised by consultants using LCAs in preparation of planning applications and EIAs, by SNH staff and by other government departments and agencies. To this regard GIS are seen as the tool able to solve the omissions and inconsistencies and to bring accuracy and more usefulness to the outputs of LCA (Julie Martin Associates and Swanwick, 2003).

With an investment in technology LCA could produce a better end product that can be made available and understandable to a wider range of users in the future. LCA based on GIS will store digital information as databases that are a flexible, usable and updatable resource. In addition GIS are powerful tools for presenting and visualising data and linked to the internet, GIS can disseminate the LCA outputs in a quicker way thus reaching more people so that the targets of the LCA are fully achieved.

However, if many consultants agreed with the benefit of GIS, a few also highlighted the difficulties. First of all it requires technical skills, expensive software. Maintaining the database is expensive in terms of money, time and human resources. Where used, the GIS have been only been partially exploited; in fact they have been considered more as visual mapping tools than as analytical tools of spatial and geographical data.

Beside the costs and the skills required, the use and application of GIS for the LCA has become more complex. The few cases of work and research published provide evidence that implementing GIS that meets the LCA purposes is more of an experiment and a challenge than a straightforward operation. The literature about GIS applications for LCA is thin on ground whereas there are many articles dealing with issues of landscape ecology or urban planning. In the majority of situations, the landscape is not analysed as a whole or from a holistic perspective but in its individual features and aspects. The methodologies and approaches undertaken on individual elements, pattern and processes forming the landscape might be used as reference from which start and develop the basis for this research.

In conclusion, the key message emerging from the review is that in 2003 the need to develop a new approach to the LCA was already recognised since it could make it (LCA) more consistent, objective, flexible, updatable, less verbose and more accessible. The majority of people working on LCAs or using LCA outputs seemed to consider GIS as the tool able to solve/overcome the weaknesses of the LCA and enhance the overall process.

Perhaps on the basis of the results of the 2003 LCA review, in 2008 SNH commissioned another national overview of Scotland's landscapes; in this occasion the contractors were asked to examine the methods used and the experience gained in producing national classifications of landscape character across Europe in order to develop recommendations for classifying Scotland's landscape character at regional level (Land-Use Consultants, 2008).

In their report, the contractors describe an approach to regional character assessment in Scotland and amongst the elements of the methodology proposed they mention the use of GIS during the quantitative analysis of the datasets collected. In a change from the LCA guidance dated 2002, the role of GIS in the identification of areas where there are common patterns of landform, land-cover, settlement types and so on, is stated more clearly. The report specifies that the areas of different landscape character identified through GIS are provisional and the field survey is recommended in order to confirm or refine those (Land-use consultants, 2008). Therefore, despite GIS were applied for more analytical purposes, the contractors pointed out some risks that could derive from the GIS led approach and namely that the analysis is necessarily based on the datasets that cover the whole Scotland and incomplete data is likely to raise problem in achieving consistent analysis throughout the country.

Furthermore the analysis might be biased towards those aspects that can or have been measured; in fact it is acknowledged that not all the elements in the landscape can be quantified and mapped, thus there is a series of unmapped information, above all qualitative aspects of the landscape that is not included in the GIS analysis.

Finally, scale remains influential. In fact according to the level of detail of the data available GIS might or not be able to pick up relationships between landscape elements and consequently might or not be able to identify adequately different landscape in a given area.

The weak points highlighted by the contractors are substantially correct and well judged, and this research could confirm them. However, as highlighted in the report, it is worthwhile adopting/exploiting GIS in the quantitative analysis because their results

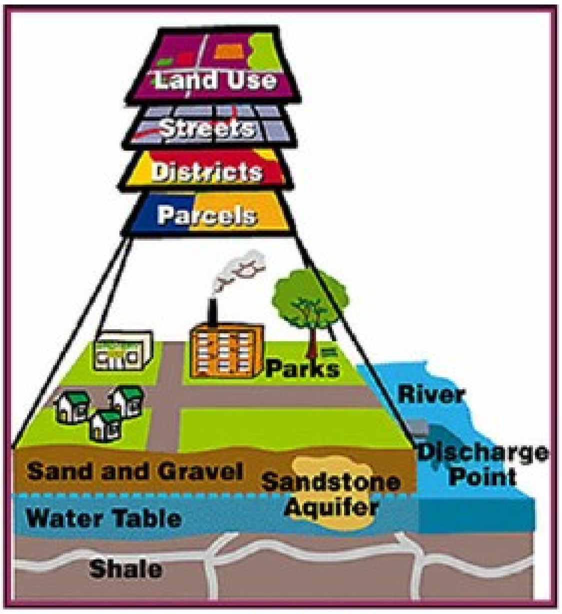
- could always been compared to the information from existing LCA in order to identify areas of disagreement/agreement;
- could be used to derive maps from raw data and fill gaps in information;
- could help to review/confirm or refine the areas identified by stakeholders in case their involvement took place during the characterisation phase.

To conclude, from 2003 to 2008 there is an increasing trend in considering GIS as more active tool for landscape characterisation with its pro and cons. This reasoning is based on the assumption that GIS have already proved to be an efficient support tool for assembling, analysing and distributing geographical information, then they might turn to be useful for recognising the intrinsic landscape formations and features that define the identity of each region (D. Carman, 2007). In addition, GIS generally convey the information resulting from the analysis in a more concise and direct way, which is a map, tables and/or graphs, easily accessible to people.

Nevertheless, although the idea of using and applying GIS to the quantitative analysis during the landscape characterisation phase of LCA is theoretically sound, in practice many doubts and uncertainty arises, above all amongst GIS analysts and as stated in the first chapter the main research questions are "how GIS can identify from digital maps unique and distinct patterns of elements in the landscape that make it different from another?" "How can a computer based device read layers of maps, whose data is numeric, and detect different landscapes without seeing, feeling, perceiving them?"

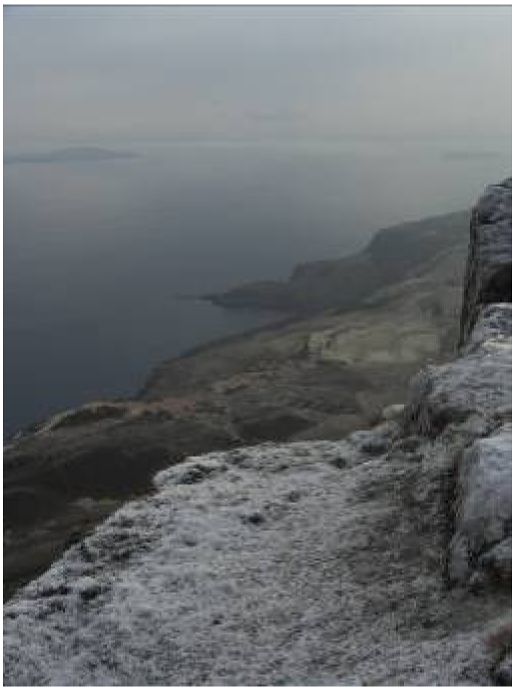
As it appears, the challenge is huge but not impossible and this PhD will demonstrate that there is an answer to the questions. However, the definition of the GIS-based methodology for landscape characterisation required time and a long journey, accompanied by several attempts undertaken in order to shape the methodology into its definitive form. This chapter will describe the steps that characterised the definition of the methodology, from its beginning to an important turning point. Interestingly it will emerge that there are several links between the GIS approach and the current LCA methodology. For instance the collection of landscape elements in a series of thematic maps is not new to LCA, nevertheless the selection and inventory of maps from the GIS point of view is a crucial aspect because the maps are the "eyes" of GIS and as noticed by Land-Use Consultants (2008) gaps in information and data coverage influence the overall GIS analysis.

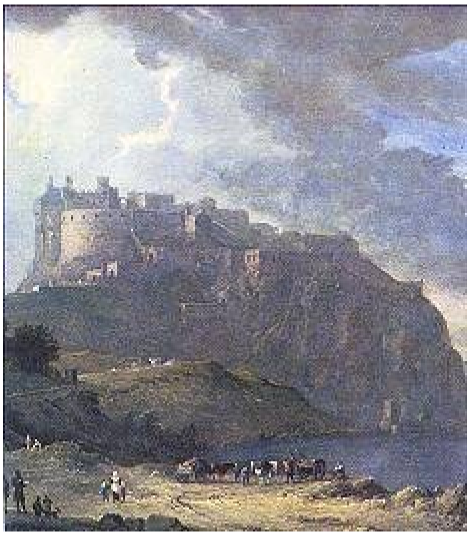
Possibly the representation of landscape in GIS terms is not as romantic as the image that a painter, a landscape architect or people in general have in their minds (figure 3.1). The GIS representation of a landscape doesn't have the appeal of the atmosphere created by a painter, the vivid colour of a photograph, and neither does it evoke any idea of beauty. However, it is thoroughly realistic and "logical" because it suggests that the landscape is derived by summing different elements together: for example land use, streets, hydrology, geology etc.



© Alexis Aguilar. Sonoma State University. 2006

a)





b) **c)**

Figure 3.1. The landscape from the GIS perspective (a) and the landscape as perceived by people (b) East coast of Raasay from Dun Can and (c) Edinburgh castle by A. Nasmyth 1780.

Indeed depicted in this way the landscape appears as complex as in the reality but, contrary to people, GIS treat the elements in the landscape as numbers and analyse them in more mathematical terms. Therefore, on the one hand it is possible to sum elements and obtain a landscape and on the other it is possible to operate inversely and decompose the whole (landscape) in elementary pieces (elements) and analyse them individually or in small groups.

The option of tackling the landscape in its individual parts is revealed to be an effective solution to the problem of dealing with complex realities/objects. Although GIS can handle and analyse data stored in several maps, it is true that a high number of layers corresponds to more complicated operations with a risk of obtaining confused and imprecise results. Therefore, the idea of breaking the landscape into smaller parts and understanding which elements contribute to the landscape is the starting point of the GIS approach to landscape characterisation.

It is worth making clear that the GIS approach to landscape characterisation is holistic despite it considers the landscape as divided into smaller parts. Contrary to the examples of GIS application to individual landscape elements, which were mentioned in the first chapter, the GIS-based characterisation starts from the identification of the elements that contribute to the character of the landscape and then continues with an integrated analysis of them. The approach adopted is to break the landscape into units, consider the elements in each unit and then assemble them back to obtain the picture of the landscape as a whole. As previously mentioned this research aims at carrying a quantitative and objective landscape characterisation, therefore it considers only the landscape elements that can be measured unambiguously. Nevertheless it is recognised that we differentiate landscapes through our feelings and perceptions and aesthetic and perceptual aspects undoubtedly contribute to make the landscape character. The main questions are “how to measure these subjective aspects of the landscape through GIS?” And “it would be possible to integrate these aspects within a quantitative landscape characterisation process?”

3.2 Aesthetic aspects of the landscape and GIS: an analytical challenge

Two works carried out by Habron (1998) and Carver et al. (2008) were identified as starting point from which tackle the analytical challenge that GIS have to face when asked to define subjective aspects of the landscape. Habron first and Carver et al. after, tried to define wild land and wildness respectively in Scotland and in the Cairngorms National Park. Despite the difference of 10 years in between, both works recognised that the problem with defining wild land and wildness lies in the subjectivity of these concepts and in the fact that there is not a universal definition of them but hundreds of interpretation since "Wilderness is what we think it is" (R.Nash, 1982, cited in Carver et al. 2008). The solution adopted by the authors were different: Habron tried to find out which landscape attributes influence people's perception of wild land by submitting to a sample of people a photographic questionnaire. Carver et al. started from thinking of wildness in terms of index, derived by combining attributes maps describing the components of wildness and weighted according to their order of priority, such that the relative value of wildness could be mapped for any areas (Carver et al., 2008). Carver et al. referred to the definition of wildness provided by SNH and looked at ways of measuring attributes maps such as absence of modern human artefacts, perceived naturalness of land cover, rugged nature of the terrain, remoteness from mechanised access. The analysis was carried out by applying GIS-based multi-criteria evaluation and fuzzy mapping methods.

Carver et al. study, as well as Palmer and Lankhorst (1998), who attempted to define spaciousness and enclosure in the Netherlands in terms of presence/absence of tree and buildings, inspired to look at how the LCA guidelines defines aesthetics and perceptual aspects of the landscape and see if from those definitions a series of indicators (or attributes) could be derive and calculated from the available maps using GIS.

The LCA guidance dedicates a part of its chapter 5 to the issue of dealing with the aesthetic and perceptual aspects of the landscape and points out that it is important to give equal attention to the experiential dimensions of the landscape character. The LCA guidance provides a brief explanation of the meaning of a series of aesthetic aspects and suggests a list of descriptive “adjectives” that could be used by landscape architects and practitioners in the field survey. Furthermore, the LCA guideline outlines the importance of indicating how specific landscape elements contribute to the aesthetic characteristics. For example, the guidance cites the case of “enclosure”, which might result from the presence of woodlands or from particular terrain morphology.

The guidance distinguishes the aesthetic aspects from the perceptual ones. The latter are defined as more subjective since people’s responses depend on the experience of the individuals and therefore are personal. The overall indication from the guidance is to incorporate both aesthetic and perceptual aspects into the survey in a transparent way by acknowledging the extent of subjectivity that is included (see figure 2.7 in chapter 2).

For this research which, as said, approaches the landscape characterisation from a quantitative point of view, the question was “how to measure the attributes of aesthetic/perceptual aspects that make the character of a landscape without a survey of people’s perception but only using GIS and the digital maps available?”

After consulting the literature and testing various methods, it was concluded that it was impossible to quantify the “perceived landscape” because landscape as such is an image and a construct of people’s minds and feelings (Arriaza et al., 2003). In contrast, it was thought useful to use the description of aesthetic aspects in the LCA guidance in order to understand which attributes and terms could be translated into indicators and be measured.

Two disadvantages of this approach could be identified. The first was that the definitions provided by the LCA guidance were hardly applicable within GIS. For example according to the guidance the relative quantities of different elements within the view affect balance and proportion. Criteria such as the rule of thirds (1/3 to 2/3 relationships) were suggested to assess how well balanced the landscape is in aesthetic terms. Proportions may change with the seasonal addition or loss of elements and temporal effects may influence these aspects as well and therefore they should be included in the analysis (Swanwick and Land Use Consultants, 2002). However, in this research temporal and seasonal effects could not be considered because of the lack of information and the rule of thirds, as intended in visual arts, was too complex to be applied to GIS.

The second disadvantage of working with aesthetic aspects was that the decision process about the variables that best explain the aesthetic aspects is more subjective, interpretative and the likely to be arguable. In order to not to make a decision based exclusively on personal judgement, Matthew Hawkins, senior ecology officer at the Cairngorms National Park and expert landscape architect, was consulted and he provided advice on how to consider diversity and dominance. The way GIS tackled these two aspects is discussed below.

3.3 On the characterisation and classification of diversity and dominance

Diversity was intended in terms of variety of unique and distinct classes of landscape elements occurring in a defined area while dominance was intended as the frequency of occurrence of a unique class of landscape elements compared to all the others within an area of analysis. It was decided to try and quantify these two definitions in GIS terms through the focal functions of variety (applied to calculate diversity) and majority (applied to measure dominance) which are illustrated in figure 3.2.

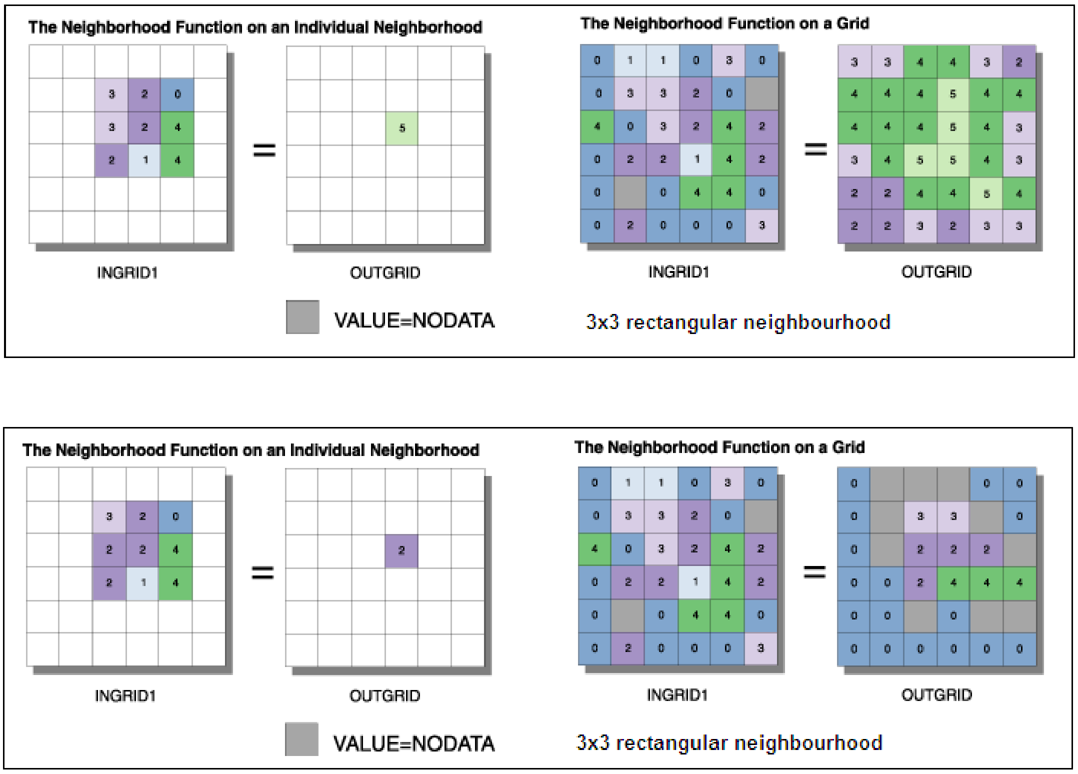


Figure 3.2 The way focal variety (top) and majority (bottom) functions work within GIS. On the left is an example of what happens to the value of the central cell in a single neighbourhood; while on the right is an illustration of what happens to the cells when the neighbourhood is calculated for the entire grid. ©ESRI

Fundamentally, the two focal statistics[3] perform a neighbourhood analysis. Thus, the neighbourhood, which can be rectangular, circular or annular, used to calculate the statistics, has to be defined first. Then in case of variety, figure 3.2 top, to the central cell of the neighbourhood area is attributed a new value indicating the number of unique classes identified within the neighbourhood area. This number is the measure of the variety of the area, hence the greater the number the higher the variety and the diversity.

In the case of dominance, figure 3.2 bottom, to the central cell of the neighbourhood area is attributed the value of the most frequent cell. For instance the value 2 is

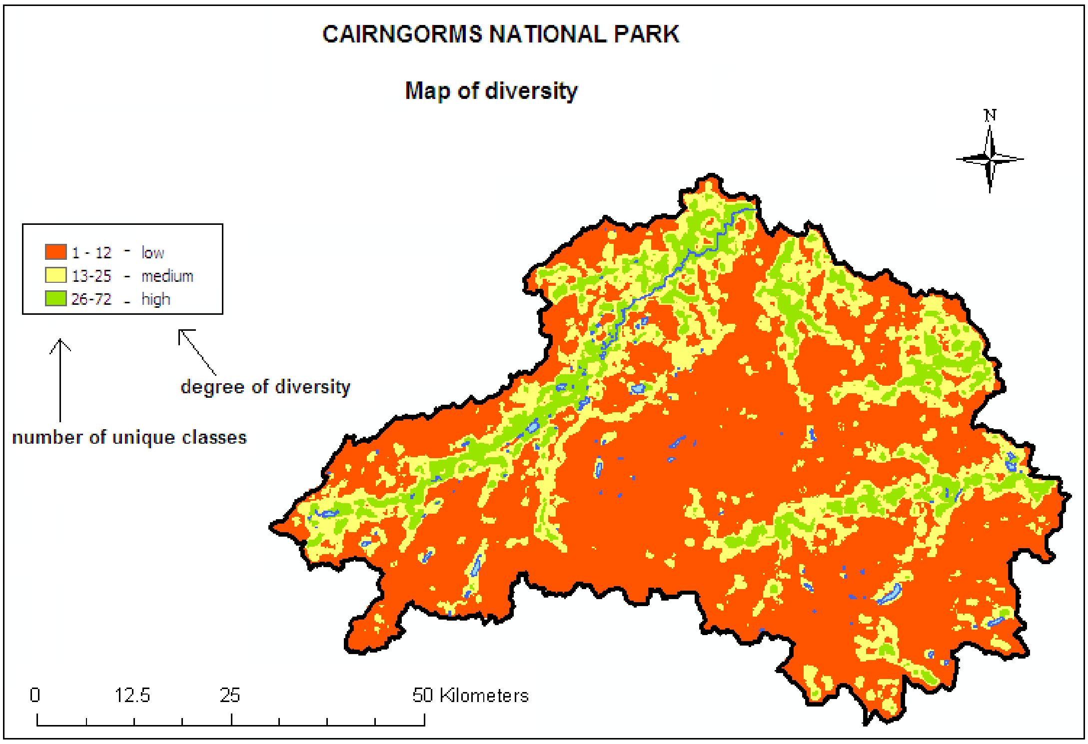
[3] In arcGIS 9.2 this tool is stored in arctool box\spatial analysis tools\neighbourhood.

assigned to the central cell since this number occurs most often within the neighbourhood area.

Although with the two focal statistics it was possible to quantify diversity and dominance it was still necessary to make personal judgement and decide for example on the dimension of the radius to use for the neighbourhood analysis and on which landscape elements calculate variety. With Matthew Hawkins it was agreed to base the analysis and quantification of the aesthetic aspects on the visible elements in the landscape since these are perceived and seen by people. Consequently the analysis included landforms, land-cover, and historic land-use. The geological maps of bedrock and superficial and the map of settlement types were excluded because the former were often hidden by other elements such as classes of land-cover or “behind” the landforms, the latter were counted as urban areas in the land cover map.

It is recognised that the analysis was based on a personal decision thus it doesn’t want to be “the way” of calculating diversity and dominance and its results (see figure 3.3 top and bottom) are arguable. Nevertheless the analysis was an attempt to demonstrate that a landscape characterisation through GIS can include landscape elements that are generally, and more straightforwardly, described through our senses and perceptions.

The maps of diversity and dominance resulting from the application of the focal functions of variety and majority looked overall plausible because revealed information which is largely recognised as it reflects the reality. Precisely the areas with greater diversity are those with the highest presence of human activities which contributed, throughout the centuries, to modify the landscape in a more remarkable way. Thus, the changes in the landscape are reflected by a greater diversity in historic land-use and land-cover and, in the case of the Cairngorms, this occurs along the main rivers in the flat areas.



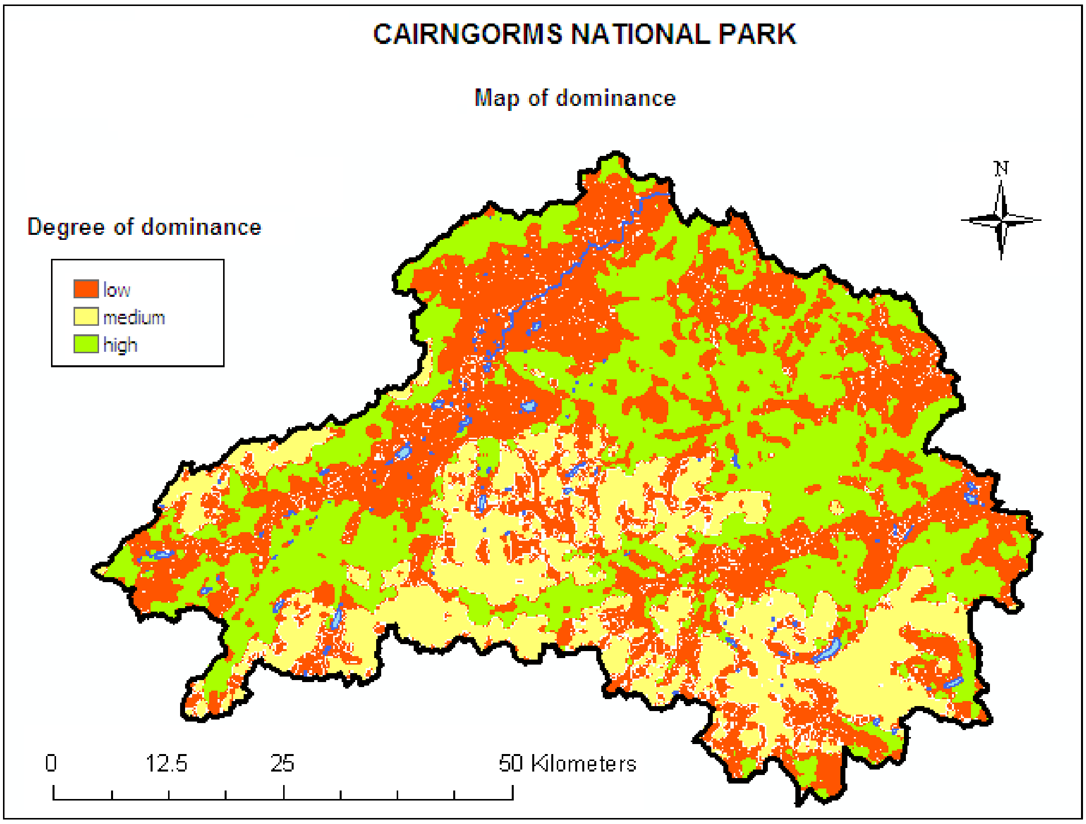


Figure 3.3. Final version of the map of diversity (top) and dominance (bottom) ready to be used for the landscape characterisation analysis.

In addition areas rated with high degree of diversity match with areas characterised by low dominance, and in fact in the Cairngorms there is a lower dominance in the flat areas and in the glens, compared to the top of the plateau and the surrounding moorlands and uplands which seem to be characterised more uniformly by an association of open slopes, shrubs/heaths, moorland and rough grazing.

It is worth remembering that these are comments made on a general overview of the resulting maps of diversity and dominance, however in order to validate the maps properly and with greater accuracy it is suggested to carry an operation of ground thruthing, which here could not be carried out because of a lack of time.

The way GIS were used to measure diversity and dominance demonstrated that it is not impossible to include aesthetic aspect into the process of an objective and quantitative based landscape characterisation but it is surely less easier than considering only those elements, such as geology, land-cover, land-use, landforms, that can be defined explicitly. Therefore for the rest of the research it is acknowledged that physical, historical, aesthetic and perceptual aspects of the landscape contribute to make its character but only the first two are included in the GIS-based landscape characterisation.

3.4 Attempts towards a GIS-based methodology for landscape characterisation

The first step to begin landscape characterisation with GIS is to have clear understanding of what character means. According to the English dictionary "character is the particular combination of qualities that makes someone a particular type of person". Although referred to a person, the definition suggests looking at the *combination of qualities* that make something (a) *unique and distinctive* (type) (Longman English Dictionary).

Surely more tailored to the landscape is the definition of character provided by the LCA guidance that says: ”character is a *distinct, recognisable and consistent pattern* of elements in the landscape that makes one landscape *different* from another rather than better or worse” (Swanwick and Land Use Consultants, 2002).

Finally, more “romantic” is the tone used in one ESRI Arc-Newsletter 2005/2006[4] from Hoesterev to describe the character “as the voice of the landscape, its *intrinsic asset* that define its scenic beauty and that contains its cultural heritage.” Therefore according to Hoesterev, who was responsible for the Greenprint Puget Sound project, listening to the landscape would allow the identification of those resources that define the character of each landscape and make people more aware of the landscape signatures and how much they are connected with their everyday life (ESRI Arc-newsletter).

Although very different, the three definitions showed similarities and expressed agreement in affirming that natural and cultural elements related to each other and repeated in distinct and different patterns denote a character. Consequently it is logical to look at the way the selected elements in the landscape distribute and group in relation to each other. If there are distinct groups of elements that combine together and make an area different from its neighbours, then these groups describe the different landscape characters of that area. Therefore, the characterisation of the landscape lies in the identification of elements that exhibit similar attributes, form recognisable patterns/clusters and reveal strong relationships amongst each others.

Slowly it becomes evident that characterisation implies an analysis of the structure of the data which can help to identify those unique and distinct patterns that correspond to the character of the landscape. The literature does offer several examples of studies carried with the intention of detecting patches and patterns in the data or the landscape structure. In the majority of the cases these analyses are carried out by landscape ecologists who tend to analyse the landscape from the perspective of species, e.g.

[4] http://www.esri.com/news/arcnews/winter0506articles/for-puget-sound.html

plants or animals. Although the difference in scale, which is very fine when dealing with species and coarser when analysing landscapes, and the focus on defining the causes/effects of an ecological process, the analyses conducted by landscape ecologists constitutes a valuable background in order to understand what a pattern analysis is about and how to deal with it.

Landscape differs structurally, in terms that its elements vary in size, shape, number, kind and configuration, but also functionally, according to the interaction between its components. In addition landscape changes and modifies both structure and function over time, which makes landscape a complex object and its quantification problematic, but not impossible (McGarigal and Marks, 1995). In fact nowadays a common way of measuring and quantifying the structure of the landscape exists and it is identified with the term of metrics. These are processed by a spatial pattern program for categorical maps called FRAGSTATS.

Several studies in landscape ecology (Kearns et al., 2005; Peng et al., 2010; Huang et al., 2006) provided examples of the way FRAGSTATS can be successfully applied to the identification of landscape classes/patches, however when the tool was experimented within this research and used for the definition of the character of the landscape it revealed not to be helpful. In fact several metrics need to be calculated for the identification of landscape classes thus, from one original map we ended to "*n*" maps, one for metric and the overall level of analysis increased in complexity, which was not the desired target. In addition the calculation of the metrics take place on a categorical map at time, hence the holistic perspective of the landscape analysis is not supported.

Williams and Wentz (2007) presented a method, TOSS (Type Orientation Shape and Size), which calculates attributes such as orientation, shape and size as a solution for the identification of landscape patterns. According to the authors these four attributes are the forming processes for polygon objects, thus by knowing their properties it is possible to identify the polygons. In fact the aim of TOSS is to determine whether or not geographic features with similar geometric properties and attribute types form a patter, which in turn may suggest that a similar process occurred in their creation (Williams and Wentz, 2007).

TOSS was experimented in order to see if could be helpful and meet the aim of the research. This method helped to cast light and quantify whether patterns in the landscape distribute in a clustered, random or dispersed way, however, similarly to FRAGSTAT, it did not perform and achieve a characterisation of the landscape in its context; in fact once again the calculation focussed more on single maps and not on the landscape as a whole and from an holistic perspective. Hence TOSS lacked the most important part of landscape characterisation, which is to consider the association amongst the elements in the landscape, which come from different maps and need to be calculated all together in the analysis.

At the end it was decided not to apply FRAGSTATS and TOSS for the landscape characterisation, because none of them considered the landscape as a whole and helped to display the patterns of landscape elements in terms of landscape character types and areas. Nevertheless, the experimentation of these two methods allowed few points and concepts to be highlighted.

First of all both methodologies taught that patterns are identified from quantifiable attributes of the landscape elements, such as their shape, size, orientation, contiguity, diversity, dispersion, proximity, total area and so on. Therefore TOSS and FRAGSTAT suggested that landscape characterisation requires GIS to be used more as analytical tools and not only as ordinary mapping tools like the majority of people is accustomed to. Thus, the lesson learnt is that there is a necessity of thinking and adopting indicators or criteria able to quantify the landscape elements.

Secondly the methodologies demonstrated that working with numeric and quantitative data, instead of qualitative ones, still involves some personal judgements; for instance we could question how many meaningful metrics should be calculated, how many clusters should be retained from the cluster analysis, how much different or similar the patterns are from the others and so on.

To summarise, the main conclusion from the application of TOSS and FRAGSTAT was that the GIS-based methodology for landscape characterisation in order to be innovative and successful should:

- decide on the parameters to be used in order to measure and take into account the elements in the landscape;
- investigate which statistical techniques are the most suitable for the identification of spatial patterns in the structure of the data;
- provide an effective and clear representation of the landscape characters on a map.

The following chapters describe the way the methodology was shaped and implemented in order to achieve these points.

CHAPTER 4

From raw data to information

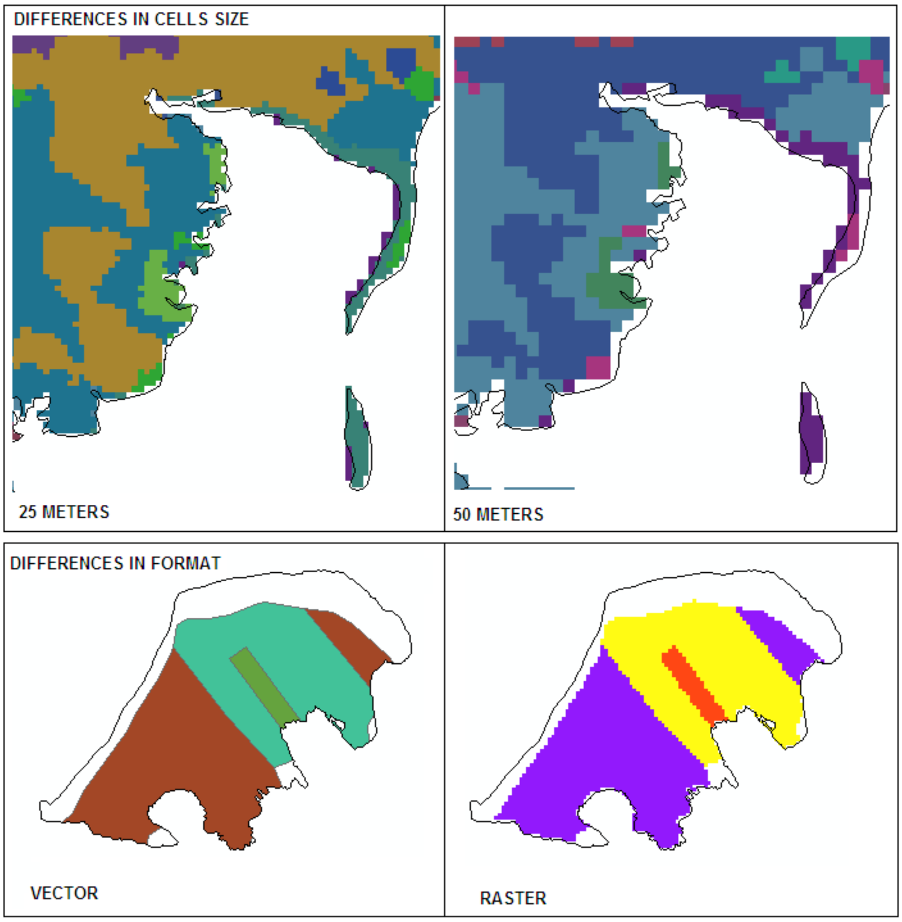
Loch Creran, Lismore and Mull from Beinn Sgulaird (I.Marengo).

4.1 The importance of the data in GIS-based analyses

The elements that form the landscape are numerous: lochs, hills, cliffs, stacks, woodlands, pastures, towns, villages, archaeological sites, roads, windfarms, pylons and so on. Each of them can be grouped into categories such as landforms, land-cover, geology which are illustrated and symbolised by thematic maps or images (aerial or satellite). Consequently, we know about the landscape either by observing and experiencing it directly or through the maps. Because of these two ways of becoming aware of the landscape and the opportunity to identify differences in the landscape we are privileged.

However, while approaching the analysis of the landscape we might wonder whether characterisation is an operation of identifying areas that share the same distinct, recognisable and consistent pattern of elements or is a process of separating geographical areas on the basis of the rarest combinations of elements. Seemingly there are two different perspectives on tackling landscape characterisation; on the one hand there are the “clumpers” who like to group objects and are keen to generalise and identify similarities. On the other, there are the “splitters” who are more concerned with details and prone to highlight differences. These two perspectives may depend on the method of analysis whether it is “bottom-up” (grouping) or “top-down” (splitting) (Brabyn, 2009). Contrary to people and their advantage to be observers and map readers, GIS work exclusively with digital maps which provide them with the information about the landscape of an area. Therefore while dealing with GIS the decision of being a “clumper” or a splitter depends on the purpose of the analysis and on the quantity and quality of data available.

In the context of this research, the methodology for the landscape characterisation had to face the fact that GIS could constrain the analysis in two ways: firstly GIS allow the analyst to perform calculations only on those landscape elements that can be measured and quantified by parameters. This means that if data is not available and cannot be measured then it is not considered and analysed through GIS. Secondly the properties of the data, such as scale, form and format influence the decision about the method of analysis and the tools and statistical techniques to be applied. For example as illustrated in figure 4.1, nominal or ratio data, a raster or vector format and a resolution of 25 or 50 meters (or a scale of 1:10000 or 1:1000000) translate into different maps and different information content which constrains the analyst to use certain tools and to think of the best way of structuring the analysis.



LCV_CLASSES	COUNT	XCENT_LCV
open shrub heath	561130	45.65
acidic grass	187084	15.22
bog	174524	14.2
water inland	69231	5.63
rough grass	66997	5.45

TYPES OF DATA

NOMINAL: lcv_classes
ORDINAL: count
RATIO: xcent_lcv

Figure 4.1. Differences between resolution (or scale), format and type of data. Before starting a GIS analysis it is important to know the characteristics of the data collected because it will affect the decision about the operations and calculations to carry out.

Due to the undoubted influence of the data in the GIS-based analyses the first step is to understand the data collected, which means to know the type of maps, the features and the attributes that they describe, and their date, if they are old or recently updated. This process is similar to make an inventory and it is useful also because it allows the gaps and missing data to be spotted. In case relevant information is not available, criteria can be set in order to retrieve it from the raw data.

4.2 Understanding the data

In this research the dataset collected is comprised partly of raster and partly of vector shape files at a variety of scales and coverage (see table 4.1).

Name of dataset	**Scale**	**Coverage**	**Source/owner**	**Data format**	**Year**
Land-Cover2000	25x25m cell size	Whole Scotland	SNH (CEH)	Raster (ESRI GRID)	2000
Geology (bedrock/superficial)	1:625,000	Whole Scotland	BGS	Vector (Shapefile)	-
National Vegetation Classification (NVC)	Sample quadrats 5 meters resolution	Partial	SNH (JNCC)	Vector	2007
Phase 1 Habitat	Sample quadrats 5 meters resolution	Partial	SNH (JNCC)	Vector	2007
Historic Land use Assessment map	1:25,000	Partial	RCAHMS	Vector	2007
Digital Terrain Model	50x50m cell size	Whole Scotland	EDINA DIGIMAP (OS)	Raster	-
Ancient woodland inventory	1:12,500	Whole Scotland	SNH	Vector	2000
Seminatural woodland inventory	1:12,500	Whole Scotland	SNH	Vector	2000
Land-Use	1:12,500	Whole Scotland	SNH (Macaulay Institute)	Vector	1993

Table 4.1. Characteristics of the data collected at the beginning of the research. Notice that only those highlighted in blue were used for the analysis.

Throughout this research the LCA guidance for England and Scotland was taken as main reference source and the data collection was planned according to the suggestions found in the guidance. The main target was to gather information on:

- geology
- landforms
- soils
- vegetation
- trees- woodland
- river and drainage systems
- land-use and field patterns
- settlement patterns

As noticeable from table 4.1 only some of the data suggested by the guidance was in fact available, partly because of the accessibility and costs of the data and partly because of the complete lack of information. For example it was fortunate to have the majority of the data provided either by the sponsor (SNH) or by Edina DIGIMAP, however, it was clear that maps such as landforms, settlement types, drainage systems needed to be derived in some how from existent dataset.

In addition, after the analysis of the content of the maps it was decided to use only some of them. The NVC and Phase 1 habitat maps were not used because their highly limited coverage (less than half of Scotland) and extremely detailed information. Equally, the land-use map was not considered, but mainly because of its obsolete information; in fact the map depicts Scotland's land-use in the late 80s which is a considerable time lag if the map is compared to the more recent land-cover 2000 and the historic land-use assessment (HLA) maps.

As far as the HLA is concerned its use was opportune despite the coverage being incomplete. There are two reasons for using HLA in landscape characterisation: first the landscape is comprised of physical and cultural elements and second HLA is the only data source that currently provides rich and specific information on historic categories, types and periods of land-use. Hence HLA information was too important to be dismissed and not comparable to the cases of NVC and Phase 1 habitat where the data was too detailed if compared to the rest of the maps.

A slightly different treatment was operated on the maps illustrating ancient and semi-natural woodland. These maps were used only during the attempt to derive the enclosure map, but not for the whole landscape characterisation. The reason has a simple explanation, it was noticed that the data about semi-natural and ancient woodland was in part covered by the land-cover map; therefore it was thought opportune not to carry redundant information.

To summarise, at the end of the first inspection of the content of the maps only five maps were retained for the analysis, namely:

- geology bedrock
- geology superficial
- land-cover
- historic land-use assessment
- digital terrain model (DTM)

Amongst the data suggested by the guidance, two maps, specifically landforms and settlement types, were missing but considered sufficiently important to try and fill the gap. The literature provided the basis to define the criteria from which derive the two maps; however before giving a detailed explanation about the way the criteria were applied, it is worth mentioning how the rest of the data was prepared for the analysis.

As the methodology took shape, it became evident that the data needed to be “fitted” to the identification of geographic patterns in the data. The GIS tools that suit best to this analysis are the spatial analysis tools that tend to prefer raster data instead of shape files and that in general produce outcomes in raster format. Consequently it was decided to convert the shape files into raster.

This conversion requires the analyst to decide the resolution (scale) of the output map. In this research a cell size of 50 metres was opted for two reasons: first smaller cell sizes result in larger datasets since the entire surface has to be covered by a greater number of cells which need more storage space and often make the processing time longer. Second, 50 metres was the resolution of the DTM and resampling this raster to have a smaller cell size does not improve the detail of its information.

After the conversion from vector to raster the second important operation carried on the data was to aggregate the information since the attribute tables of all the maps provided more information than needed. Generally the higher the number of categories classified in each map, the more complex and long the calculation; therefore the decision of grouping the data in few and more generic classes. Table 4.2 summarises the number of classes reduced from the original set.

Thematic maps	**FROM (number of classes)**	**TO (number of classes)**
geology bedrock	60	9
landcover	26	14
landforms	10	5
historic land-use	NA	NA

Table 4.2. Original and aggregated number of classes per thematic map. The classes of HLA, despite grouping, cannot be classified since the number of classes retained in each area of study was selected on a case by case basis.

The generalisation followed in some cases an official classification: the classes of geology bedrock were grouped according to the Geological map of Scotland (edited in

1976 by the former Institute of geological sciences, currently BGS); the historic land-use periods were merged into new classes on the basis of the advice given by Mr. Piers Dixon, operation manager at RCAHMS (see page for more details); the classes of land-cover were aggregated according to the Land Cover Map 2000 (LCM2000) level1. In this specific case the classes "bogs" and "fen, marsh and swamp" were grouped together to form the class "peatland and wetland" and the classes "littoral, supra-littoral rocks and saltmarsh" were grouped into a unique class. Notice that the geology superficial was not aggregated since the 9 classes were all distinct.

After this re-classification and aggregation the data was ready for the analysis which is explained step by step in chapter 6. The rest of this chapter is dedicated to the description of the way landforms and settlement types' maps were derived.

4.3 On the derivation and classification of landforms

Landforms are the product of both long and short-term processes that operate principally in response to climate, geology, surface water and ground water conditions, soil properties, vegetation and land use (Prima et al., 2006). From the literature emerged that the derivation and definition of landform elements computationally was a challenge that involved years of research. Amongst the studies reviewed, many classify landforms through the use of digital elevation model (DEM), and here the aim is to find the classification technique that is the most practical and feasible on the basis of the GIS capacity.

Over the last three decades DEMs have been developed to aid the terrain analysis by using computers and nowadays it is common practice to use DEMs to extract landform parameters such as slope, aspect, convexity and curvature. As illustrated in figure 4.2 these are fundamental terrain attributes that are calculated automatically and have become standard operations in most GIS packages.

Attribute	Significance
Slope (gradient) [percent, degrees]: Rate of change of elevation in the direction of steepest descent. Either calculated as steepest downhill slope to one of the eight neighbours or using a second-order finite difference method fitted to 4 or 8 closest neighbours	The distribution of soils, water content and erosion potential are directly dependent on the flow velocity of flows, which is controlled by slope.
Aspect (exposition) [degrees clockwise from north]: orientation of the line of steepest descent, usually measured in degrees clockwise from north. Aspect is undefined in flat areas, where slope is zero.	Aspect can be used or terrain visualization or in conjunction with slope to estimate solar radiation
Profile Curvature [radians per 100 meters]: rate of change of slope down a flowline as a measure of changes in flow velocity and sediment transport.	Identification of zones of enhanced soil erosion and deposition; differentiation between upper and lower slopes.
Plan Curvature [radians per 100 meters]: rate of change of aspect along a contour as a measure of topographic convergence and divergence.	Measure the propensity of water to converge as it flows through the landscape; identification of crests, ridges or depressions.

Figure 4.2. List of the four basic topographic attributes that are derived from the Digital Elevation Model (DEM) and explanation of their main uses. (source: B. Klingseisen, 2004.)

The accuracy and quality of the terrain attributes, which are also called surface derivatives, depend on both quality and resolution of the DEM (Klingseisen, 2004). The most influential factors are:

- Topographic complexity and roughness of the landscape
- Source of elevation data (ground survey) and DEM generation method;
- pixel size;
- precision;
- algorithms used to calculate the terrain (topographic) attributes.

Nowadays GIS calculate the surface derivatives automatically, but amongst the basic surface tools there is not one that deals with the definition of landform units and classification methods. Seemingly there are two justifications: first of all landforms may be modelled quite differently according to the geomorphic processes which are considered (Prima et al., 2006). Secondly the terms applied to geographic features and their definitions are generic and derive from the particular needs and applications of the organisations using them. Fundamentally the question: “What are the differences between mountain, hill and peak; lake and pond; river and creek?” cannot be properly answered. There is a distinct lack of a standard landform classification methods and it is possible to cite examples of the variety of the methods of landforms classification from the literature.

Pennock et al. (1987) for example defined several distinct landform elements by measures of gradient, plan and profile curvature as shown in figure 4.3.

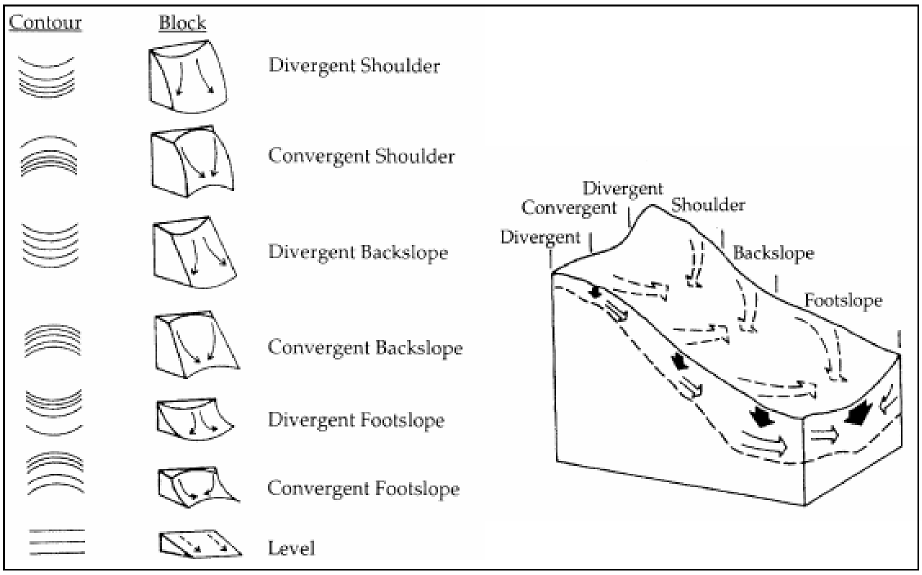


Figure 4.3. On the left are depicted the landform elements defined on the basis of contours and block diagrams while on the right it is illustrated a drawing with associated landform elements and the flow of water. (source: B. Klingseisen, 2004.)

Dikau (1989) subsequently described landforms though a hierarchical subdivision of the land surface into relief units with homogeneous slope, aspect and curvature. He traced out what Hammond elaborated in 1969 when he firstly identified different landform quantitatively, on the basis of the values of slopes, relative relief and relative proportion of flat and gently sloping terrain.

MacMillan et al. (2000) revisited the classification made by Pennock et al. (1987) and categorised the landform entities into types and elements. The former were defined as small areas of land surface that exhibit a relatively restricted range of morphological attributes, whereas the latter were considered as representative assemblages of characteristic patterns that repeat themselves (Prima et al., 2006). The authors ended with 15 landform elements as illustrated in figure 4.4.

Landform category	Landform element (also called landform facet)				Slope (%)	Slope curvature (deg/100 m)	
	No.	Name	Abbr.	Comments		Profile	Plan
Upper slope	1	Level crest	LCR	Level area in upper slope	0–2	+10 to −10	–
	2	Divergent shoulder	DSH	Convex upper, water shedding element	>2	> +10	–
	3	Upper depression	UDE	Depression in upper slope position	0–2	< −10	< −10
Mid-slope	4	Backslope	BSL	Rectilinear transition mid-slope segment	>2	+10 to −10	+10 to −10
	5	Divergent backslope	DBS	Sloping 'ridge'	>2	+10 to −10	> +10
	6	Convergent backslope	CBS	Sloping 'trough'	>2	+10 to −10	< −10
	7	Terrace	TER	Level mid-slope >2 m above base level	0–2	+10 to −10	na
	8	Saddle	SAD	Special case of a divergent footslope	na	< −10	>10
	9	Midslope depression	MDE	Depression in midslope position	0–2	< −10	< −10
Lower slope	10	Footslope	FSL	Concave, water receiving element	>2	< −10	na
	11	Toeslope	TSL	Rectilinear in lower slope >20% of low slope	>2	+10 to −10	+10 to −10
	12	Fan	FAN	Special case of a divergent toeslope	>2	+10 to 10	> +10
	13	Lower slope mound	LSM	Crown in lower slope <2 m above base level	>2	> +10	> +10?
	14	Level lower slope	LLS	Level in lower slope, >20% of low slope	0–2	+10 to −10	+10 to −10
	15	Depression	DEP	Concave element in lowest landform pos.	0–2	< −10	<0

Figure 4.4. Guideline for the 15 unit landform classification rule based applied by MacMillan et al. (2000). (source: B. Klingseisen, 2004)

As depicted in figure 4.5 MacMillan et al. (2000) also modelled a hierarchy of the landform entities which takes into account scale and resolution of the DEM.

Figure 4.5

In Australia, a decade earlier than MacMillan et al., Speight (1990 cited in Klingseisen, 2007) had developed a similar classification by evaluating the landform as a hierarchical mosaic of tiles where those of 300m radius formed landform patterns and those of 20m radius landform elements.

Appropriate Scale	DEM Resolution and Scale	Appropriate Scale
1:5 Million to 1:10 Million	10 x 10 km (ETOPO5)	Physiographic Province
1:1 Million to 1:5 Million	**1 x 1 km (GTOPO30)**	**Physiographic Region**
1:500,000 to 1:1 Million	500 x 500 m (DTED)	Physiographic District
1:125,000 to 1:500,000	**100 x 100 m (DTED)**	**Physiographic System**
1:50,000 to 1:125,000	**25 x 25 m (SRTM)**	**Landform Type**
1:10,000 to 1:50,000	10 x 10 m	Unnamed and undefined
1:5,000 to 1:10,000	**5 x 5 m**	**Landform Element**
1:1,000 to 1:5,000	1 x 1 m	Unnamed and undefined

Figure 4.5. Hierarchy of landform entities proposed by MacMillan et al. (2000). (source: B. Klingseisen, 2004)

Speight then identified 10 topographic positions in which landform elements fall in as shown in figure 4.6 (Klingseisen, 2004).

Name	Definitions of Speight (1990)
Crest	Area high in the landscape, having positive plan and/or profile curvature
Depression (open, closed)	Area low in the landscape, having negative plan and/or profile curvature, *closed*: local elevation minimum; *open*: extends at same or lower elevation
Flat	areas having a slope < 3%
Slope	Planar element with an average slope > 1%, sub classified by relative position:
Simple Slope	Adjacent below a crest or flat and adjacent above a flat or depression
Upper Slope	Adjacent below a crest or flat but not adjacent above a flat or depression
Mid-Slope	Not adjacent below a crest or flat and not adjacent above a flat or depression
Lower Slope	Not adjacent below a crest or flat but adjacent above a flat or depression
Hillock	Compound element where short slope elements meet at a narrow crest < 40m
Ridge	Compound element where short slope elements meet at a narrow crest > 40m

Figure 4.6. Classification of landforms on the basis of 10 topographic positions (Speight, 1990). The degree of adjacency between the terrain derivatives is the main determinant factor in this classification. (source: B. Klingseisen, 2004)

Nevertheless, all the cited models were used mostly in the field surveys and were missing a specific detail to be compatible to computers: an algorithm.

Coops et al. (1998 cited in Klingseisen, 2007) were the first authors who developed a compelling algorithm to predict topographic position through the definitions of Speight.

This is based on a mixture of terrain analysis and Boolean algebra and delineates the landform elements using thresholds on topographic attributes derived from a DEM.

The increasing use of terrain analysis in a wide range of applications was accompanied by the development of new software. Not only the generic GIS packages expanded their tools in order to calculate the most important topographic attributes (slope, aspect, curvature) but also numerous freely available software packages appeared offering a palette of tools for terrain analysis (Klingseisen, 2004).

Amongst the latter group of software packages, the topographic positioning index (TPI), created by Jenness (2006), and its marine counterpart the benthic positioning index (BPI), developed in U.S. by NOAA Coastal Service Centre and Oregon State University (OSU)[5], were selected and tested in this study.

TPI and BPI are measures of where a referenced location is relative to the locations surrounding it (NOAA and OSU, 2006) and are derived from an input DEM which can store elevation or bathymetric data.

In 2001 Andrew Weiss presented a poster at the ESRI International user conference in which he described the concept of topographic position index (TPI) and how it could be calculated (Jenness, 2006). Using the TPI at different scales and calculating the slope allowed the users to classify a landscape into slope position, for example ridge top, valley bottom, mid-slope, and into landform categories such as steep narrow canyons, gentle valleys, and open slopes.

The TPI became popular and is still widely used because is based on a clever and simple algorithm which calculates the difference between a cell elevation value, "the

---

[5] The Benthic Terrain Modeler (BTM) is a collection of ESRI® ArcGIS®-based tools that coastal and marine resource managers can use, with bathymetric data sets, to examine the deepwater benthic environment. The BTM was created as part of a cooperative agreement between Davey Jones' Locker Seafloor Mapping and Marine GIS Laboratory, Department of Geosciences at Oregon State University, and the National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center. (http://www.csc.noaa.gov/products/btm/)

cell target", and the average of the neighbourhood around the cell. The TPI of a cell is defined considering both the degree to which the cell value is higher or lower than its neighbours and the slope value of the cell (Paron and Vargas, 2007). Figure 4.7 shows the values assumed by TPI.

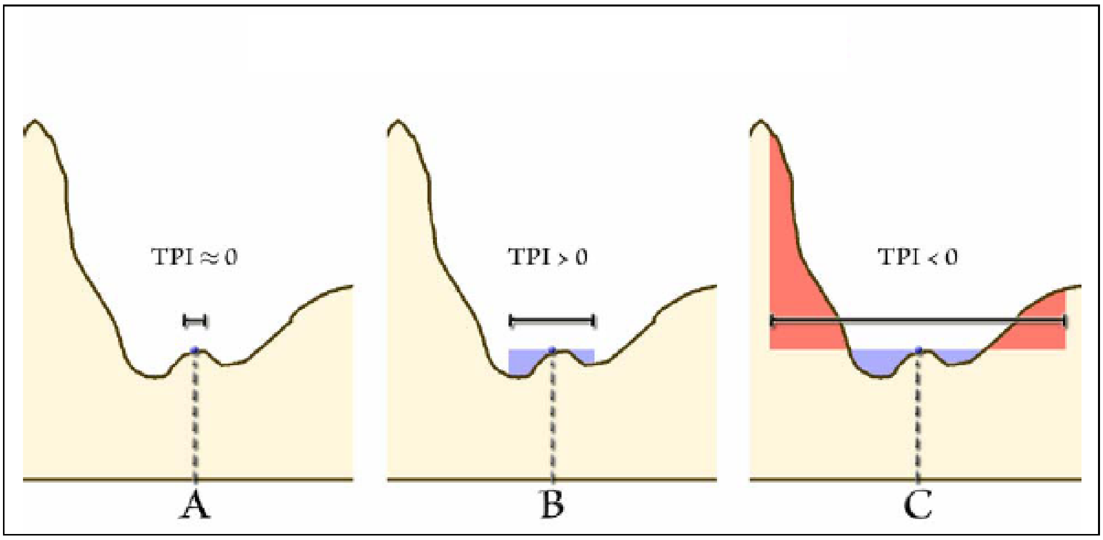


Figure 4.7. TPI at three scales. A: the scale is very small and the point is at about the same elevation of the entire region. Thus, TPI is near 0. B: the region of analysis is bigger and encompasses the entire hill plus the part of the bottom valley. Thus, the point is higher than the surrounding areas and TPI>0. C: the area of analysis is much wider and includes the valley and both its sides. The point is no longer the highest feature in the region. Thus TPI is <0. (source: J.Jennes, 2006)

Generalising this concept, if a cell has a negative TPI it means that it tends towards being a valley and at the bottom of the terrain, while if it has got a positive TPI then the cell tends towards the ridge tops and the hilltops. When the TPI is around zero, the cell is likely to be in a flat area or, according to the slope value, on a mid, upper, or lower slope.

A key point is that TPI is scale dependent. In fact the same point on a crest of a mountain range can be considered a ridge top from people standing on the bottom valley or a flat area from who lives on that mountain. Therefore users should be aware of the scale they want to use in the analysis since the classification produced is valid only for the landscape analysed at that extension and it is not universal (Paron and Vargas, 2007).

Scale depends on the resolution of the DEM and also on the neighbourhood which is defined by the number of cells considered around the cell value. For example if the kernel used to determine the neighbourhood has a radius of 500 meters then the TPI will reflect the difference in values between the cell and those that are in the considered radius. In this regard figure 4.8 provides an example of the way the neighbourhood size influences the resulting TPI maps. From the same DEM a larger and a smaller kernel are able to detect different features (Jenness, 2006).

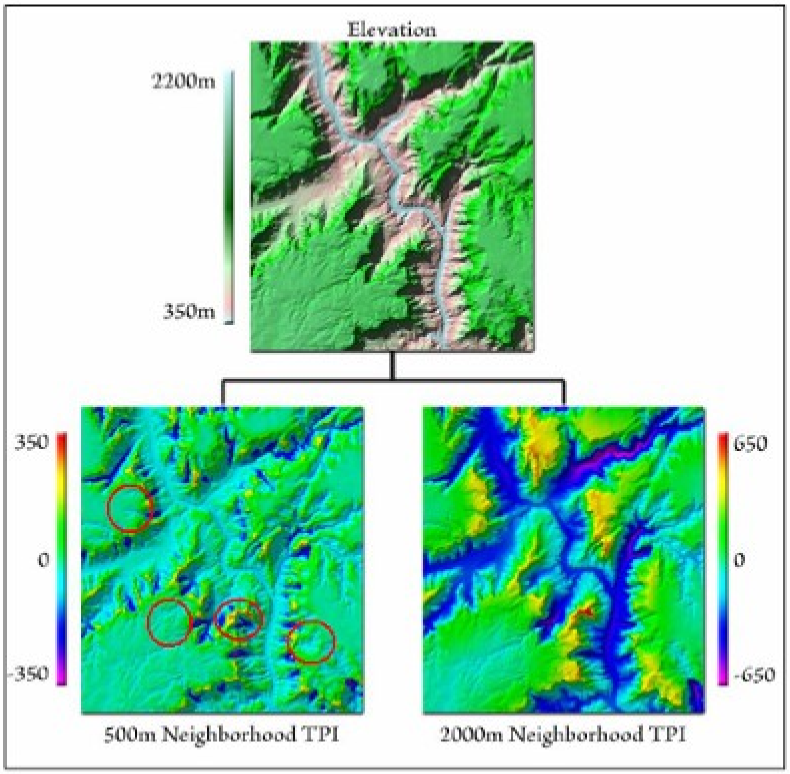


Figure 4.8. Influence of the kernel size in the identification of landforms. The results of the use of two kernels are compared to each other. In the bottom left map the red circles outline the extreme values that can be detected with a kernel of 500 meters radius. Because the scale is finer it is possible to pick a higher level of detail, while the broad scale of the 2000 meters kernel is more appropriate for an analysis at a larger context. (source: J.Jennes, 2006)

The main difference in using a small kernel and a larger one is that with the former it is possible to pick details e.g. the side drainages of the valleys (see red circles), while with the latter is possible to highlight the overall valley system.

TPI values increases with scale; a broad scale TPI analysis (1:50000) has lower values because a larger analysis neighbourhood is used and a consequence is that small variations of the terrain are averaged out. Whereas, at finer scales TPI analysis (1:25000) shows higher values because of the smaller neighbourhood that is used, therefore small and localised variations of terrain are detected.

Once the TPI datasets have been created at broad (large neighbourhood) and fine (small neighbourhood) scale, the next step of the terrain classification process is to standardise the values in order to make them comparable and classifiable at a common scale (NOAA and OSU, 2006). In fact the last step is to operate a classification of the resulting standardised broad and fine TPI datasets in landform units. The flowchart in figure 4.9 provides a summary of the operations involved in the derivation and classification of landform classes through the use of the TPI tool.

Figure 4.9

The most complex and likely subjective part of the process is the definition of the classes to apply during the classification because there is not a standard classification and consequently the values chosen for slope and elevation are specific to the area of study. Within TPI the landforms are classified in 10 classes, which can be further aggregated to each other, while BPI divides the landform units into two categories: zones and structures. The former are more generic representations of the terrain and are grouped into four classes: crests, depressions (valleys), slopes, and flats. The latter are more articulated and range, for example, from narrow depressions, broad flat, steep slopes, narrow crests. For this research, according to the way the rest of the datasets were treated, it was decided to classify the landform types more generically rather than specifically and four classes were identified, hill tops, plains, open slopes, and valleys.

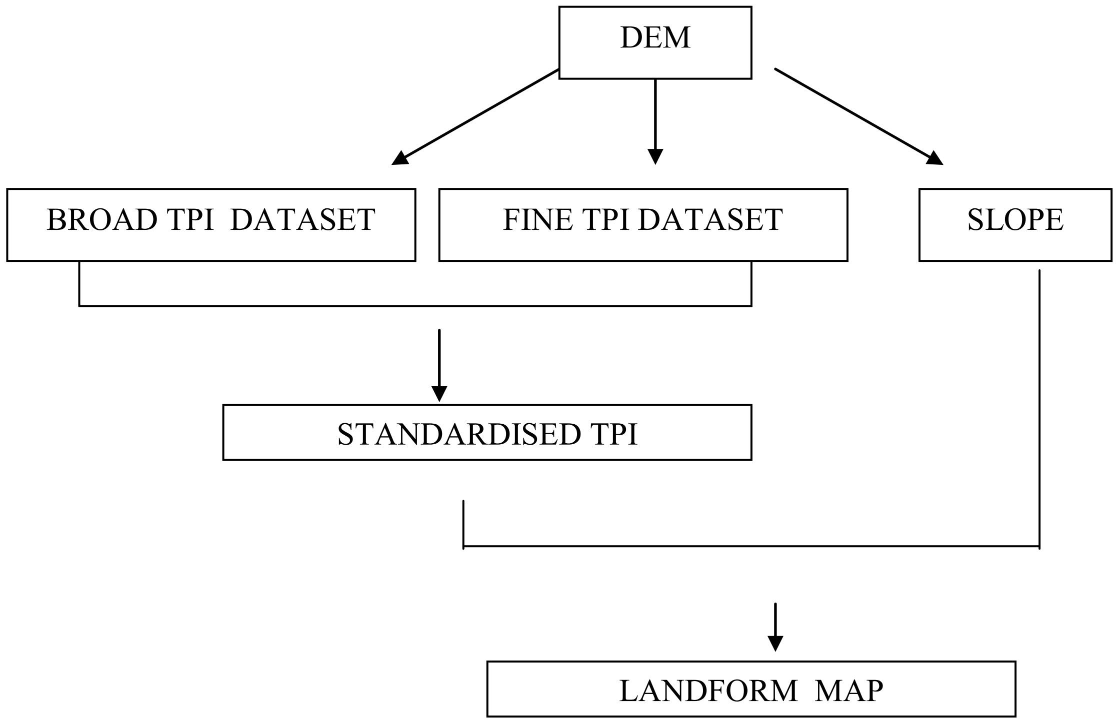


Figure 4.9. Flowchart of the steps required to derive a landform map from a DEM. As it appears slope and two sets of benthic positioning index (that is the equivalent of the terrain positioning index) are needed as background data for the classification of landforms in broad zones.

Another advantage of keeping the classification generic is that it allows other scientists to use them for qualitative and quantitative analysis without feeling constrained to a specific case and area of study (Lundblad and Wright, 2006).

From a closer analysis of the thresholds used by TPI and BPI it appeared that both indices operated in the same way and referred to the work conducted by A. Weiss (2001). He defined the threshold TPI values in terms of standard deviations from the elevation taking into account the variability of elevation values within the neighbourhood (Jenness, 2006). This means that grid cells with identical TPI value may be classified differently in different areas in relation to the variability in their respective neighbourhoods. Figure 4.10 summarises the thresholds applied for the classification of landforms into ten types, subsequently grouped into 4 classes, and namely hill tops, glens/valleys, flats/plains and open slopes.

landform types	small neighbourhood	large neighbourhood	slope	new_types
incised streams	TPI <= -1 STDEV	TPI <= -1 STDEV		glens/valleys
shallow valleys	TPI <= -1 STDEV	-1 < TPI < 1 STDEV		glens/valleys
headwaters	TPI <= -1 STDEV	TPI >= 1 STDEV		glens/valleys
u-shaped valleys	-1 < TPI < 1 STDEV	TPI <= -1 STDEV		glens/valleys
upper slopes	-1 < TPI < 1 STDEV	TPI >= 1 STDEV		small hills
hills in valley	TPI >= 1 STDEV	TPI <= -1 STDEV		small hills
small hills in plains	TPI >= 1 STDEV	-1 < TPI < 1 STDEV		small hills
plains	-1 < TPI < 1 STDEV	-1 < TPI < 1 STDEV	<=1	plains
open slopes	-1 < TPI < 1 STDEV	-1 < TPI < 1 STDEV	>1	open slopes
hill tops	TPI >= 1 STDEV	TPI >= 1 STDEV		hill tops

Figure 4.10. The classification of landform types followed that one suggested by A. Weiss who calculated the TPI values in standard deviation (STDEV) since this facilitated the definition of thresholds. The column "new_types" indicates the four landform types used for the reclassification of the original ten types.

Along with elevation, slope played a role in the classification of landforms, above all in the definition of flat and sloping areas. In order to have a better knowledge of the meaning of slope in the UK context, this research took into account the classification used by the Forestry Commission[6] (in the technical note 16/95) that adopted and modified a terrain classification used in Scandinavia. The description of slope is taken at right angle to the contour lines and is classified in terms of gradient or topographic form. Table 4.3 reports the five classes expressed in % or in degree as were recognised by the forestry commission.

CLASS	DESCRIPTION	%	Degree
1	Level	0-10	0-6
2	Gentle	10-20	6-11
3	Moderate	20-33	11-18
4	Steep	33-50	18-27
5	Very steep	>50	>27

Table 4.3. Classification of slope adopted by the forestry commission (source: technical report 16/95).

[6] The reason why Forestry Commission developed the terrain classification was because the definitions of ground conditions, ground roughness and slope have an influence on the type of machines and the systems to be used during forestry works

On this basis it was decided to consider "class 1" the threshold of slopes and then classify the areas less than 1 as a "flats" and those greater than 1 as "open slopes".

Once mapped the results were satisfying since the four landform types represented with a good approximation the location of hill tops, valleys, slopes and flats as they appear in reality and as figure 4.11 illustrates.

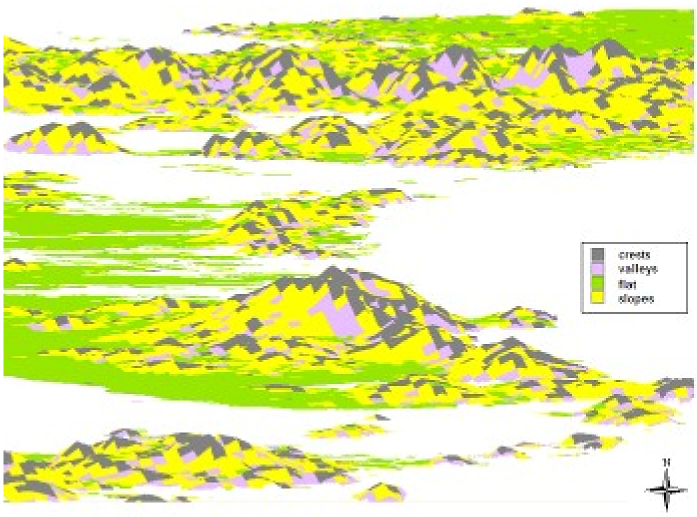


Figure 4.11. Detail of the Western Isles in 3D. The map of landforms, classified as crest, valley, flat and slope, was derived from the DEM and map of slopes. The red lines in the Google image indicate approximately the field of view of the 3D map.

To conclude, the definition of landform types was an important step in the process of landscape characterisation as it enabled to fill a critical gap in the data available for the analysis. In the application of TPI index it emerged that scale was an issue and a personal judgement was included in the definition of thresholds. Both elements/topics will occur in other parts of this research because scale is an intrinsic problem of the geographic data and personal judgement is perhaps unavoidable in any kind of analysis. The line followed by this research is to recognise both scale and personal judgement as a fact, be aware of their occurrence during the data analysis and point them out while reporting the results.

4.4 On the definition of settlement types

As outlined with the application of TPI for landform types classification, GIS-based landscape analysis requires a clear and precise definition of the criteria used to identify which elements constitute the character of the landscape.

There are landscape elements, for example those belonging to the cultural heritage, that require more interpretation. In fact a map depicting built-up areas doesn't help to understand the way settlements contribute to the character of the landscape because it does not define what a "settlement" is. Similarly the HLA map, which is a huge database of historic land-use types and periods, doesn't inform explicitly on the contribution of history to the landscape character. Consequently, for settlement types and historic land-use categories and periods it was necessary to think of criteria that could make them measurable.

The LCA guidance considers settlements and field patterns as contemporary markers of human activity, and suggests gathering information on them from the HLA map (or Historic Landscape Characterisation map in England). This current research highlighted that it was difficult to retrieve data about settlements and field patterns from HLA for two reasons. First, the patterns are described in great detail and are linked to historic periods, thus it is necessary to group them in larger classes in order to facilitate their analysis. Because of time constraint, this investigation was not carried out, but it is recommended for future works. The second reason is about the lack of the definitions describing the form and the status of the settlements. For example "elongated" and "circular" or "small village", "hamlet", and "sparse houses" were not available in any database. As a result, it was necessary to investigate whether or not an official classification and definition of "settlements" has already adopted in the UK.

In early 2002 five bodies, namely the Department for Environment, Food and Rural Affairs (Defra), the Office of the Deputy Prime Minister (ODPM), the office for National Statistics, the Welsh assembly government and the Countryside Agency formed a consortium to commission a new definition of urban and rural areas. The aim was to identify, define and derive populations for the small towns, villages, hamlets and isolated dwellings that make up the settlement pattern of rural areas (Bibby and Shepherd, 2004).

A population threshold of 10,000 or more inhabitants was taken to discriminate urban from rural areas. In their study Bibby and Shepherd covered England and Wales with a grid comprised of 35 million cells of 1ha. Individual residential addresses were captured by this grid and the household density pattern was calculated for each cell. These densities were subsequently averaged for each cell by using areas with varying radii around each cell. The result was the creation of a "density profile" that enabled, through a set of rules, a classification of settlement types. The classes identified were: small town and fringe, village and dispersed houses (Bibby and Shepherd, 2004).

The next stage taken by Bibby and Shepherd was to relate rural settlements to the Census Output Areas (COAs) in order to classify them by settlement types. The classification was based on the proportion of the population within each COA in settlements of various kinds. Then residential densities were averaged at a series of much larger geographic scales to give measurement of the context for those settlements reflecting a wider dispersion of the population. The final classification is depicted in figure 4.12.

In this current research, the Bibby and Shepherd's model was adopted and modified according to the data available and the geography of the areas of study. Here the Cairngorms National Park (CNP) is taken as example to explain how the method of analysis was developed. The CNP stretches 8,618 km² and hosts a population of 17,077 units (2001 Census).

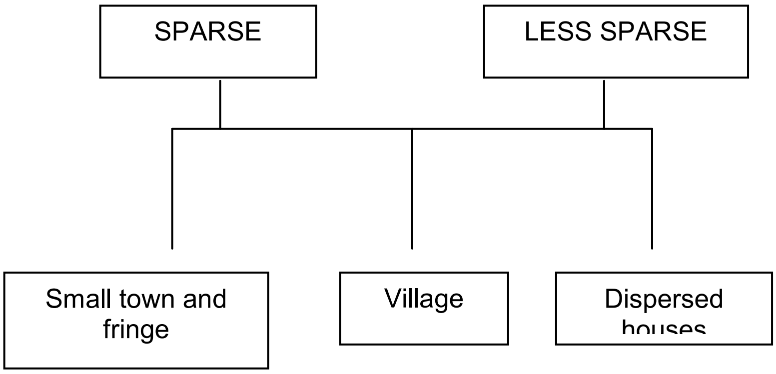


Figure 4.12 Classification of the rural settlements provided by Bibby and Shepherd. © Bibby and Shepherd (2004)

The settlements are mainly located in the highland straths that occupies only the 22% of the Park and are the most fertile lands as illustrated by figure 4.13.

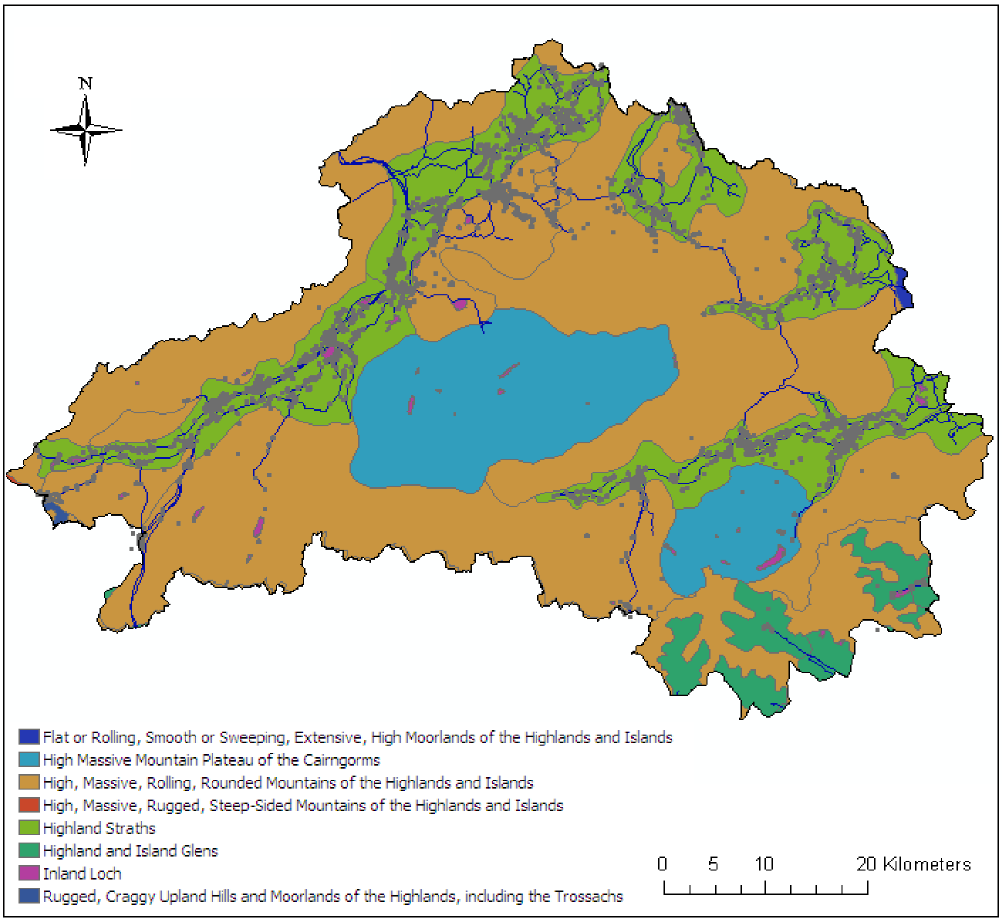


Figure 4.13. Distribution of built-up areas within the boundaries of CNPA and within the landscape character types obtained from the 1999 LCA map. There is a strong relationship between built-up areas and Highland straths.

The distinction between urban and rural area as defined by the ODPM and based on the threshold of 10,000 inhabitants was used to assess the presence or absence of urban areas in the CNPA. With reference to the data retrieved by the national statistics all the COAs in the Cairngorms ranged from a minimum of 53 to a maximum of 252 inhabitants, and the two main centres of the park, Aviemore and Grantown-on-Spey counted respectively 2,397 and 2,239 inhabitants. On this basis it was possible to claim that the CNP is characterised only by rural areas.

The next step was to look at the rural settlements inside the park and classify them into:

- rural towns;
- villages;
- small villages;
- dispersed houses and
- fringes.

Contrary to Bibby and Shepherd, who could use the population data derived by the dataset of the individual addresses, here we could rely on population per COAs whose limits, as showed in figure 4.14, sometimes crossed the boundaries of the park. As a consequence, the population recorded for the park might be overestimated and it was decided not to use it to calculate the density profile for the classification of settlement types.

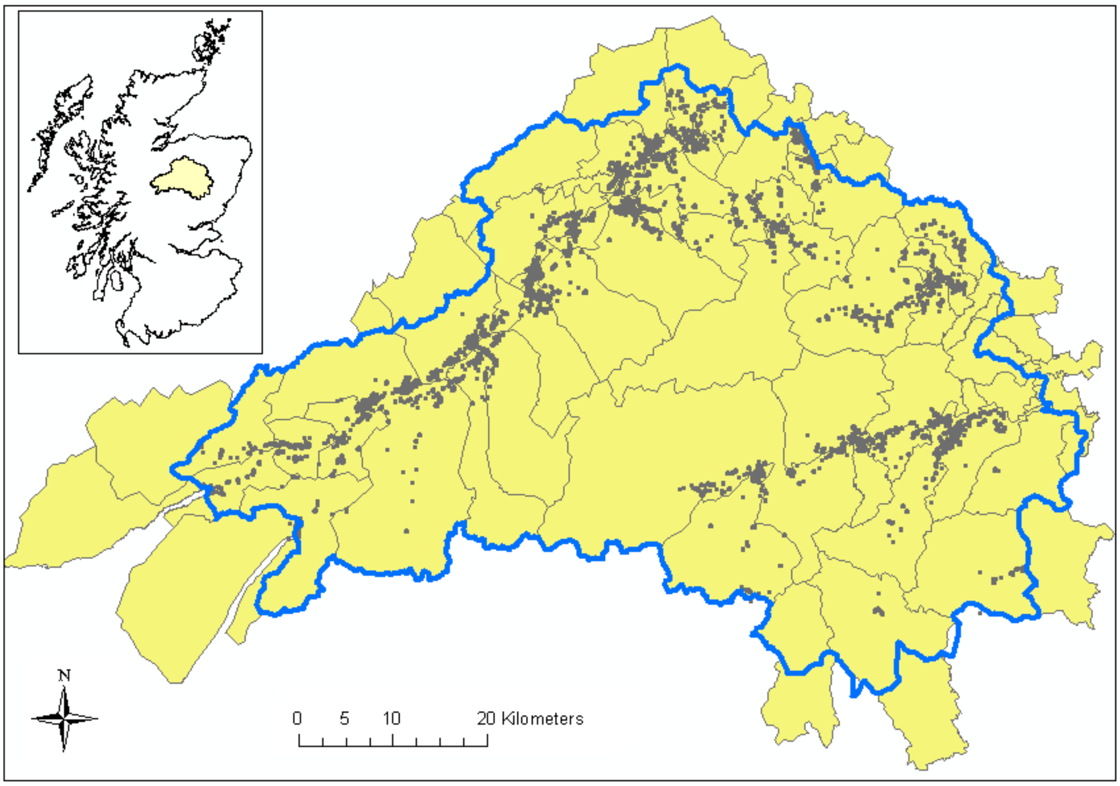


Figure 4.14. Census output areas, as in 2001, within the CNPA boundaries. In order to keep a high level of accuracy, population data wasn't used since some of the census areas consider population outside the park.

Hence the methodology adopted in this study was based on the assumption that settlements can be described in terms of the number of buildings, which correspond to the places where people live. Therefore the steps in order to classify the settlements into types were:

Retrieve a map of the buildings

Create the density profiles

Set a rule to identify thresholds for the settlements classes

Operate the classification

Through the service EDINA DIGIMAP, areas of 100 km² from OS Mastermap were downloaded and buildings were extracted and stored to form a new dataset. The scale of the OS Mastermap topography is 1:10000 over the moorland and mountain areas; 1:2500 over rural areas and 1:1250 over the urban areas. As highlighted by figure 4.15 overall the scale of the "buildings" map was largely finer than that in use for the rest of the analysis.

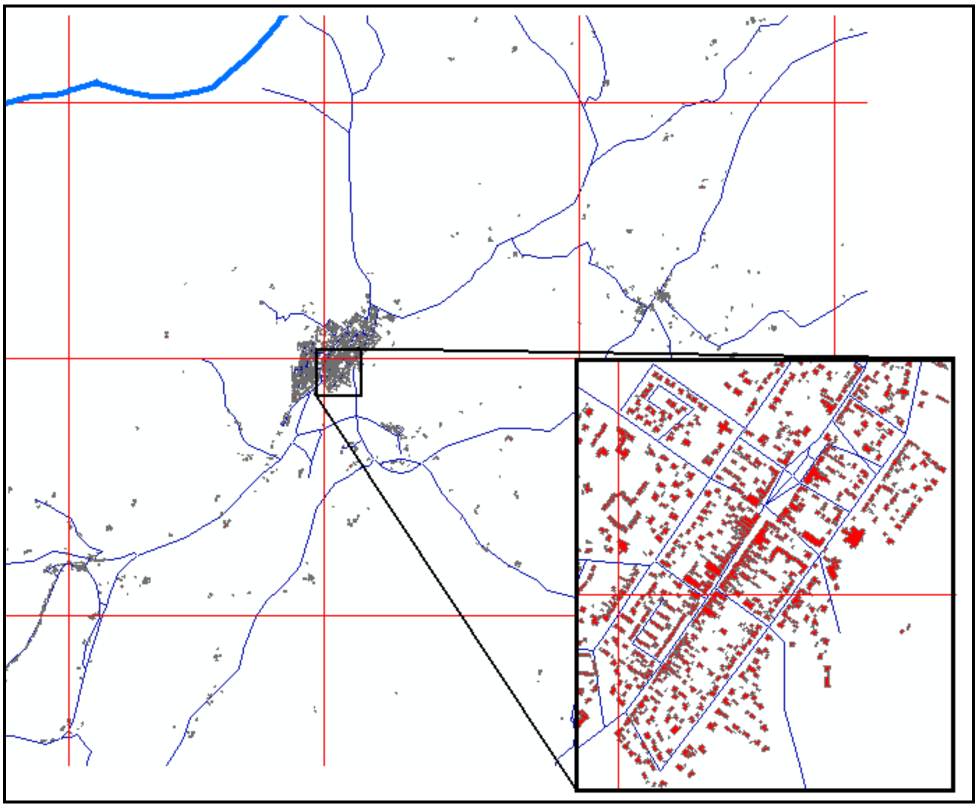


Figure 4.15. To understand the fine scale of the map of building a grid of cells having 3.06 m² was overlapped. The enlarged window shows how minute the buildings appear when compared to the grid cell size.

According to Bibby and Shepherd's methodology, the settlements types can be identified by calculating their density over a selected area. High density values indicate the presence of a large settlement, formed by numerous buildings such as a village or a small rural town; whereas low density values show less evident settlements which might be comprise of sparse houses or very contained clusters of buildings as in a small village.

Given the formula

Density= number of buildings/area

It is evident that as the area increases the density decreases, thus the scale of the area chosen for carrying out the density profile analysis was fundamental for the outcomes. The way Bibby and Shepherd faced this problem was to generate a density profile by using different areas set at four different distances from the buildings.

Within GIS there is a spatial analysis tool that performs density analysis and works with a defined neighbourhood. The density tool[7] takes known quantities of the phenomenon to be studied and totals the number of features that fall within the neighbourhood. Then it divides this number by the area of the neighbourhood, whose size becomes a factor; in fact a large radius considers a larger number of points within a wider open space as showed by figure 4.16. As consequence the output raster files are more generalised. Contrary, a small radius detects better local variations of the density values. Nevertheless, if the radius is too small broader patterns might be difficult to spot and therefore could be missed (Mitchell, 1999).

On this basis it was thought appropriate to analyse the density of the buildings by attempting different search radii with the intention to find out the density maps which could best be used to classify the settlements.

[7] The density tools are stored in spatial analysis\density (arctool box).

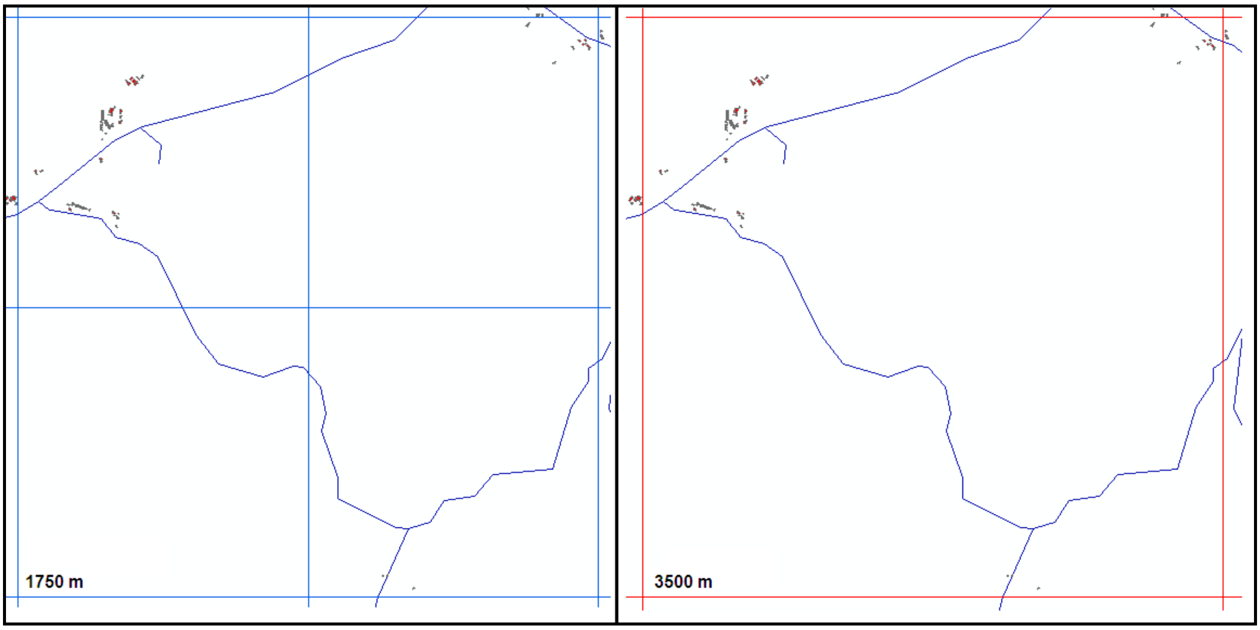


Figure 4.16. A change in the size of the neighbourhood leads to a change in the calculation of the neighbourhood area of analysis (from 4 cells to 1) and it generally affects other measurements such as density.

ArcGIS allows calculating density using a simple point or a kernel option. Both work on points, hence the buildings were converted from polygons to points, and both density tools were tested in order to see which one could offer the surface density map most suitable for the classification of the settlements.

Conceptually the simple point density option calculates the density of point features around each output raster cell. Basically it defines a neighbourhood around each raster cell centre and the number of points that fall within the neighbourhood is summed and divided by the area of the neighbourhood (ArcGIS Desktop help 9.2). ArcGIS requires a field called “population” in the attribute table of the point feature. This field stores the values that determine the number of times to count the point. For example a point with value “three” is counted as three points. To some extent the population field can be used to weight some points more heavily than others or to allow one point to represent several observations. In this research, contrary to the Bibby and Shepherd (2004) study, there is no population value associated to the building and the interest is only in classifying the settlements into types according to their density distribution. Thus the alternative field "NONE" is used and each point (each building) is calculated as individual (ArcGIS Desktop help 9.2).

The kernel density option works differently because it spreads the values associated to each point from the point location to the specified radius of the search area. The density is greatest at the point location and diminishes with increasing distance from the point until it reaches zero at the specified radius. Basically the kernel operates as if it fits a smoothly curved surface over each point and the volume under the surface equates to the "population" field value or to 1 if NONE is specified. The density at each output raster cell is calculated by adding the value of all the kernel surfaces where they overlay the raster cell centre. The kernel function is based on the quadratic kernel equation described in Silverman[8] (1986, p76 equation 4.5) (ArcGIS Desktop help 9.2).

In addition to this, the kernel option allows only a circular neighbourhood, consequently in order to make an accurate comparison between the options it was decided to use this shape also for the simple point density function.

As far as the definition of the radius is concerned the following six radii, measured in meters, were used to create the density profile:

- 750
- 1000
- 1500
- 3000
- 6000
- 12000

The decision for these measures of the radii was taken on the basis of the building distribution. It was noticed that many buildings were isolated and sparse in the

[8] Silverman, B.W. *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1986.

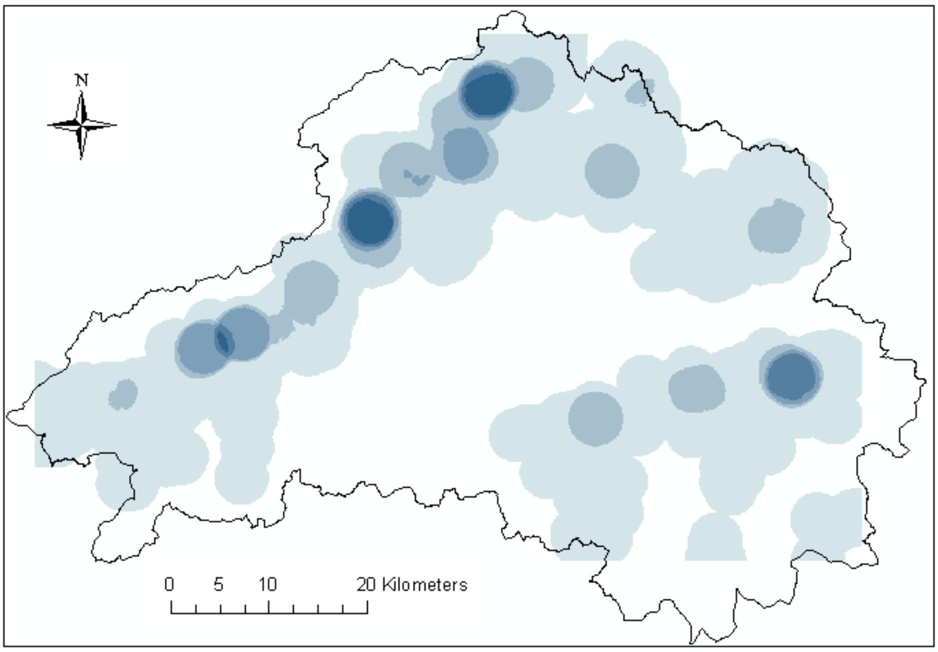
countryside, thus in order to capture them it was thought opportune to have radii up to 12 kilometres. The next step was to run the two density options, compare the resulting maps and decide which one was the most suitable for the classification of the settlements. The choice was in favour of the kernel option because the building densities were illustrated more neatly and clearly as showed in figure 4.17. The simple point option after 3km radius tended to overlap the density surfaces and generate areas of high density where in reality there weren't so many buildings, whereas the kernel option, regardless the radius size, was able to keep the density surfaces distinct and to match perfectly the areas of high density with the areas with more buildings. In other words the density maps resulting from the kernel option resulted in a more understandable and more helpful output for the classification.

Figure 4.17

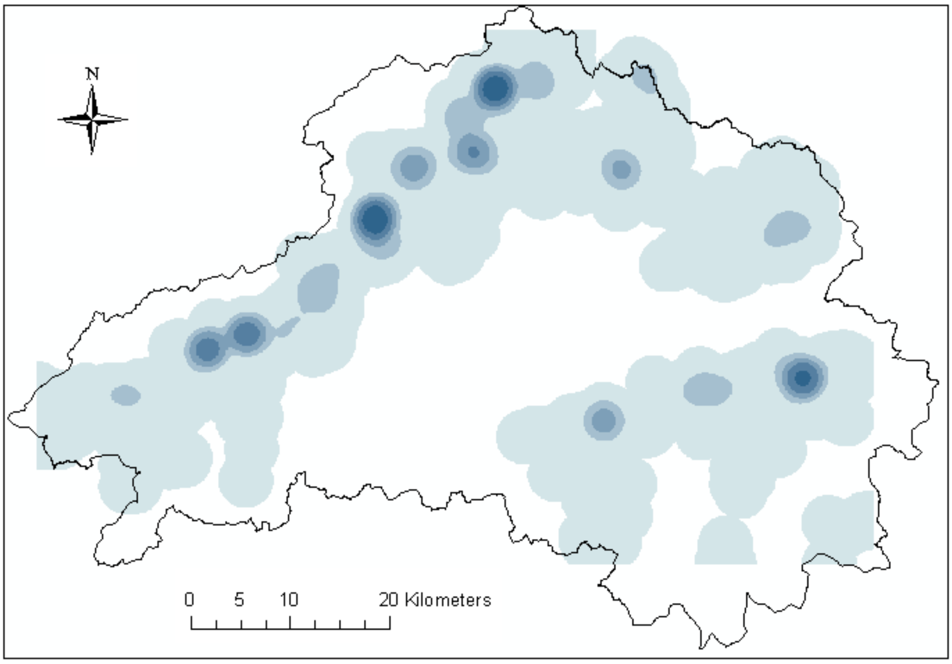
The measures attributed to the radii were also changed after noticing how their smoothing effect was amplified and visible on the outcomes. Table 4.4 stores the values of building density, which were classified into 5 classes with the natural Jenks classification system, obtained from the application of the six radii. The table helped to understand the profile of the distribution and variation of the building density values.

	Radii (m)					
classes	**750**	**1000**	**1500**	**3000**	**6000**	**12000**
1	1-64	1-53	1-34	1-15	1-6	1-3
2	64-244	53-191	34-115	15-42	6-17	3-7
3	244-557	191-423	115-252	42-82	17-31	7-13
4	557-981	423-728	252-439	82-139	31-47	13-18
5	981-1792	728-1279	439-700	139-221	47-67	18-24

Table 4.4. Measures of the radii and density values for each class according to natural break classification. As visible the greater the size of the radius the smaller the densities; this is the smoothing and generalisation effect typical of large neighbourhood areas.



a)



b)

Figure 4.17. Comparison between the output maps from two different density analysis methods. 5a) shows the surface density of the built-up areas calculated by using the single point method, while 5b) is the surface density map derived by applying a kernel method. The latter map makes the patterns of density easy to see thus it was adopted for the analysis.

As noticeable the density decreases with increasing the radii of the neighbourhood indicating that the density surface maps show a more generalised pattern (see figure 4.18)

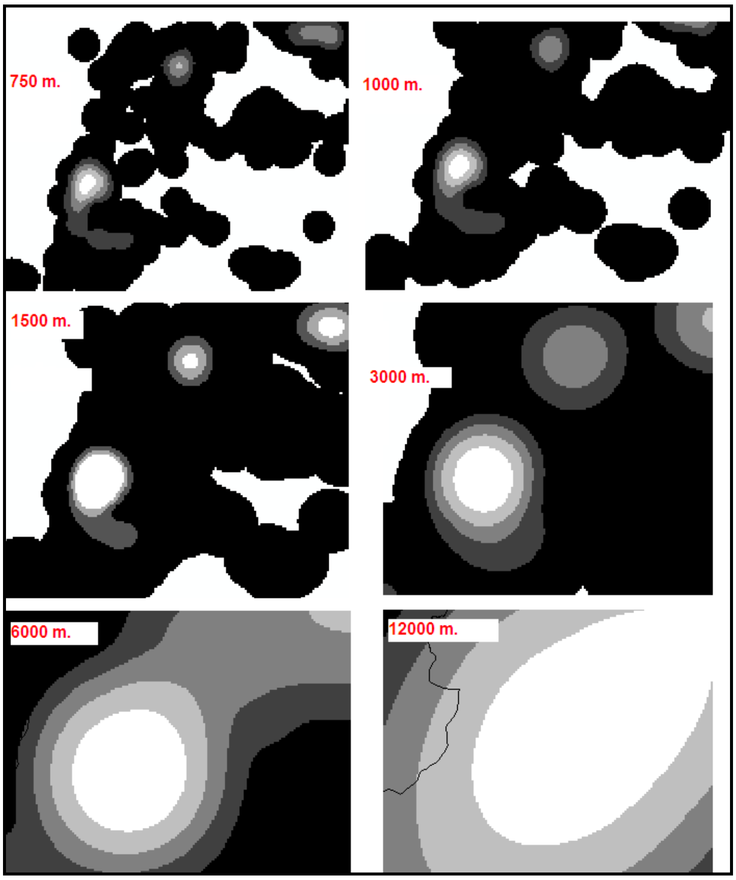


Figure 4.18. Differences in density surface patterns according to the size of the neighbourhood. It is crucial for detecting patterns of density successfully to decide for an appropriate size of the radius. If it is too big the risk is to incur in extreme generalisation while if it is too small broader patterns might not be detected.

On the basis of the exaggerate smoothing effect produced by large radii and with the consideration of the geography and the distribution of the buildings in the CNP and also in the other targeted areas of study (see chapter 7, figure7.1), the new measures that defined the radius of the kernel density tool were:

- 350 (metres)
- 750
- 1350 (this radius was used only in presence of urban areas)

Settlements all over Scotland were defined on the basis of these thresholds and classified into the following categories:

- dispersed houses;
- small villages;
- villages;
- fringes;
- rural town;
- urban town.

The classes were decided in consultation with Caroline Read and James Fenton, who both represented SNH, the sponsor of this research project, and the idea was to use terms that despite general could inform straightforwardly on the dimension of a settlement. For example dispersed houses would suggest isolated and scattered buildings, while villages would describe a more compact and structured group of buildings. Indeed the qualitative description of the settlements was supported by a quantitative measure of the density of the buildings whose calculation is briefly described below.

First of all the density surface maps were classified in 5 classes using the “natural break” classification system which emphasises with sufficient clarity the differences between the density values because it sets the breaks where there is a discontinuity (a variation) in the values (Mitchell, 1999). Afterwards the natural break classes were reclassified from 1 to 5 as showed in figure 4.20.

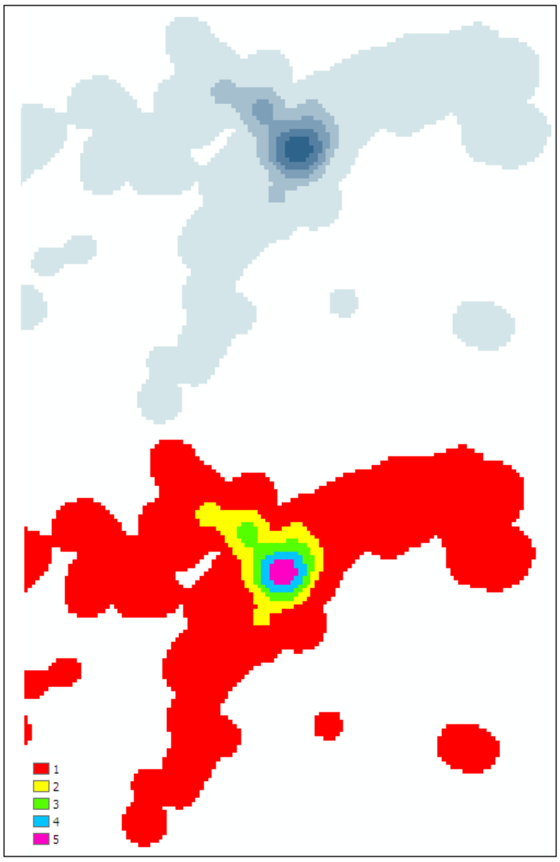


Figure 4.19. Example of reclassification of different levels of density surface. Each area was firstly classified according to the Natural Break classification, then each class was classified again in ascending order, from the lowest level of density (1) to the highest level of buildings density (5).

Secondly in ArcGIS the density maps were converted from raster to polygons[9] since shape files work better than raster when operations of selection (by attribute or by location) are performed. In addition, a field about the surface (hectares) of the density classes was added to the attribute table of the density map.

Afterwards, the buildings and the converted density maps were overlapped. This operation was meant to facilitate the classification of buildings into settlement types. The results were reached after a series of attempts aimed at grouping the density values into classes of settlements which could match the reality adequately and meaningfully. Notice that at this point a certain amount of knowledge about the area

[9] During the conversion from raster to polygon the option "simplify polygons" was kept unchecked.

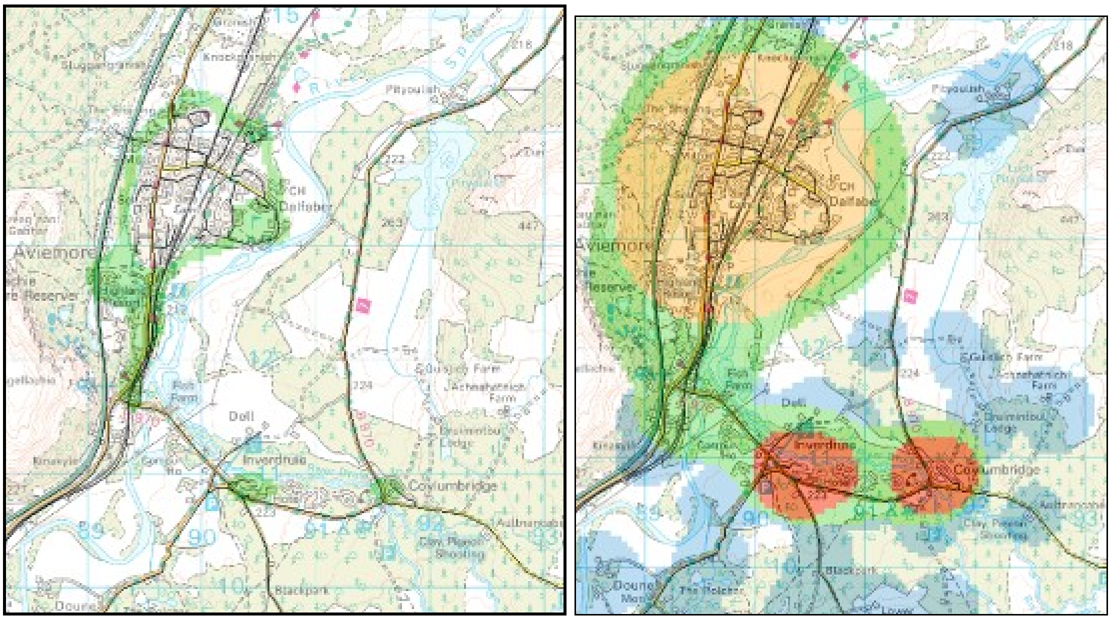
can help the classification and overall it is highly recommended a visit to the area of study before using the map of settlement types for the landscape characterisation analysis. Table 4.5 illustrates the cut off applied for the classification, while figure 4.20 depicts the resulting map of settlement types.

	density map radius		
settlement classes	350 m.	750 m.	1350 m.
Dispersed houses	class 1 and 2 (area<= 5ha)		
Small villages[10]	class 2 (area > 5ha) and class 3		
Villages		Class 3 and 4 (with buffer of 500 m.)	
Rural towns		class 5 (with buffer of 750 m.)	
Urban towns			Class 6 (with buffer of 1 Km.)

Table 4.5. The classes of densities were selected for the most suitable radius sizes in order to classify settlements into types.

As it emerges from the table, the classes of density were matched to the settlement types and in three occasions buffers had to be calculated to describe with more precision the settlement types. By proceeding in this association it became evident that some of the classes of density showed a hole in the middle and resembled more to a doughnut (see figure 4.20).

[10] It is worth pointing out that a change in the settlement classification occurred when it was applied to small islands (Jura and Islay for example), in order to be closer to those realities. The category small villages was referred as grouped houses, villages became small villages and small rural town changed into villages.



a) **b)**

Figure 4.20. During the process of classification of the settlement types it is possible to distinguish the fringes. The extract from the map of the class 3 of the 350 metres kernel shows clearly the difference between the density areas classified as fringes and the other classified either as small villages or villages. a) The green area similar to a doughnut is the fringe, while the full circles are the small villages. b) The same geographic area once the classification is finished.

These density areas were called “fringes” and were calculated both for the categories of small villages and villages. The identification of the fringes took place manually.

The third and last step was to convert the individual settlement types from vector to raster format and reclassify them as reported in table 4.6.

Settlement type (ATTRIBUTE)	**UNIQUE CODE**
Dispersed houses	1
Small villages	10
Villages	100
Fringes	20 (if from classes 2 and 3 – kernel 350m.) and 200 (if from classes 3 and 4 – kernel 750m.)
Rural towns	1000
Urban towns	2000

Table 4.6. Final reclassification applied to the raster map which will become the definitive map of settlement types.

Then, within the raster calculation tool the raster, which represented the individual settlement types, were added to each others and once again reclassified to generate the final map of settlement types as illustrated in figure 4.21.

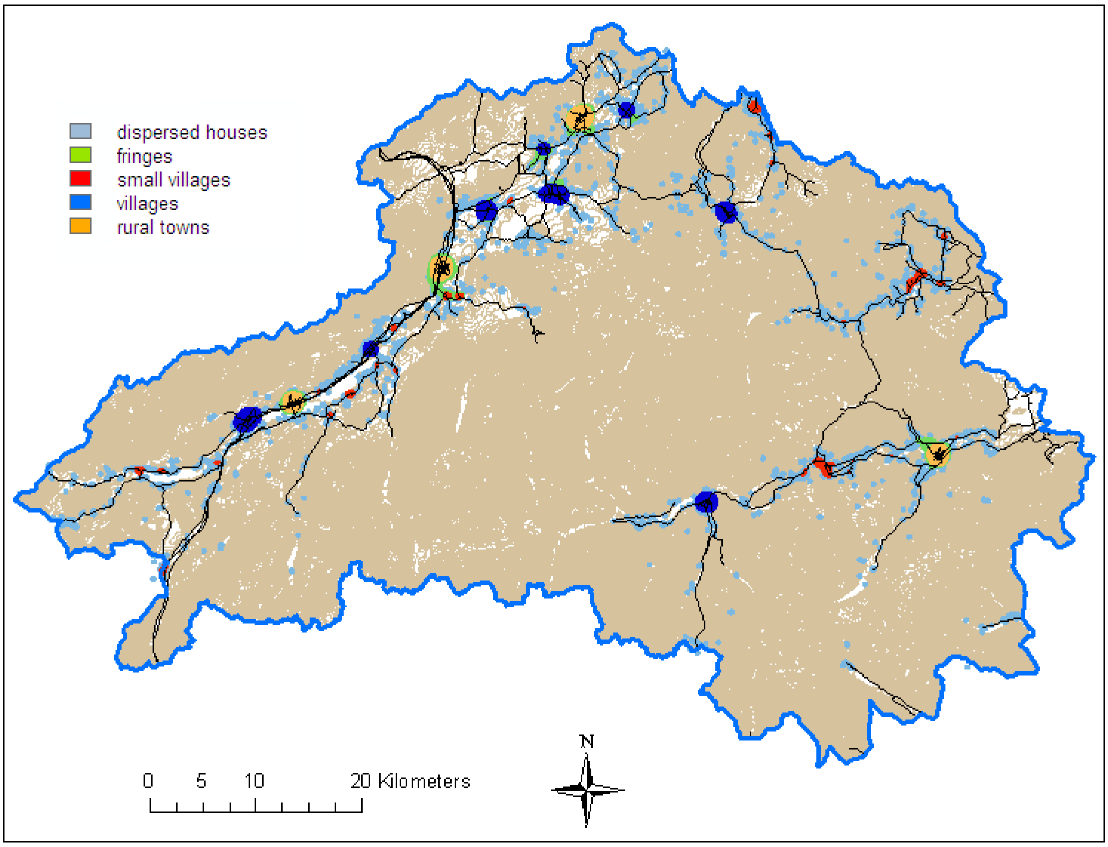


Figure 4.21. The way of representing settlements as areas is symbolic but effective from a point of view of the analysis.

To conclude, the process of identifying settlement types is very experimental and many decisions had to be taken about the computational method, the thresholds, and even the names for the categories of settlements. Thus the process of derivation of the settlement types' map can be disputable, but the results proved to be less arguable. To test its validity, the method for the identification and classification of settlement types was applied to the other areas of study, which show differences in geography and building distribution. In all the cases no changes had to be made in order to accommodate the method to the characteristic of the area and the same thresholds could be applied in each area. The results obtained for the Cairngorms National park were assessed and validated with the help of Matthew Hawkins, while the rest of the areas were verified on the basis of personal knowledge.

The work carried out on the HLA map was slightly different from the previous two because in this case a map already existed. The aim was to generalise and summarise part of its content and at the same time condense three kinds of information, which were thought more relevant to landscape characterisation, into one. The way the analysis developed is explained below.

4.5 On the interpretation and use of the Historic Land-use Assessment map

For the effective use of the HLA map in this research and to understand the way the historic elements contribute to the character of the landscape, it is crucial to be aware of what is considered to be historically essential and to distinguish why it matters. To this regard the work of Fairclough (1998), who defined some of the attributes of the historic characterisation, was particularly interesting. The approach of historians and archaeologists is to consider the whole landscape as historic and they attribute historic values to the present-day landscape. Fairclough draws a distinction between "landscape archaeology", that is the study of the past of all periods through archaeological methods, and "historic landscape characterisation", that characterises the present-day landscape in terms of historic origin, process and change (Fairclough, 1998). On this basis time, diversity of types and change are important keys in the assessment.

In addition, historic characterisation is aimed at extending the interpretation and understanding of the present-day landscapes recognising that these have been inherited from the past, both remote (prehistoric, roman and medieval period) or recent (from 18th century to date). Remains from earlier periods can help to add more evidence of the historic component of the landscape and make people aware of it. In fact commonly very few people realise the time-depth of the landscapes, the speed and the frequency in which they have changed and are changing (Dyson-Bruce, 2003).

Because of this lack of consciousness the historic aspect is more difficult to be detected and its importance, as contributor to the landscape character, can be underestimated or missed. The archaeological perspectives of the landscape are "vertical" in terms that they seek the chronological depth underneath the visible surface of the landscape. The "vertical" view of the landscape distinguishes the archaeologists from the architects who predominantly base the landscape assessment on a horizontal, surface-based and aesthetic view (Fairclough, 1998). Thus, disciplines such as archaeology and history can help to understand the landscape along with geography and a better dialog between landscape practitioners and archaeologists would help to enhance landscape analyses and research. This collaboration to some extent was already outlined in the LCA guidance where expert analysis was suggested in order to understand the time-depth aspects of the landscape.

Time was the new and important dimension introduced in the landscape by looking at it from the historical point of view. However, time is not only a variable to be measured during the GIS analysis; it also suggests the idea of movement and dynamism. Both are reflected/ found in the landscape which can be represented in a constant state of flux due to a combination of natural and anthropogenic forces occurring throughout the periods. As a consequence, the landscape transforms and changes continuously and inevitably through time, both the capacity of a landscape to absorb the changes and the strength manifested by the changes themselves have an impact on what can be recorded today as historic land-use.

In addition to time-depth, historians and archaeologists are also interested in capturing the surviving and visible historic components of our present-day landscapes that could inform the process of landscape development. As explained above, from the point of view of an archaeologist there are no areas unaffected by past or continuing human activity and there are no areas in a state of unaltered nature (Herring, 2009). Even in very modern or recently modified landscapes there are traces of the more distant past that can be contributors to the landscape character, sense of place and local distinctiveness. Nevertheless, due to the unavoidable processes of alteration, evidence from the past become valuable records and their rarity/paucity is counted as a special character of the landscape and indeed the more remote the past, the more historically relevant is the evidence.

Therefore, together with period (time depth), two other variables are useful for the characterisation of the historic land-use. The first is the description of the types of land-use which is determined by identifying the predominant current land-use evident from field survey and other data sources. The types of land-use are assessed from an historical perspective and then grouped into main categories. The second variable is represented by the survival of relict features in the modern landscape. These are classified as "Relict land-uses".

Historic period, categories and types of land-use and relicts[11] are all recorded in a very detailed way so that the attribute table of the HLA contains information on historic periods, category and type of the land-use, and similarly historic period, category and type of three different relicts, indicated as "Rel1", "Rel2" and "Rel3" as illustrated in figure 4.22.

The information conveyed by the attribute table, despite being enormous, is not suitable for the landscape characterisation analysis because of the way it is stored. The aim of the research is to understand how the historic and cultural aspects contribute to the character of the landscape; therefore here the intention is to extract the necessary and relevant information able to meet the aim.

[11] For more precise information on these variable please refer to the following website http://hla.rcahms.gov.uk/ where maps and glossary can be downloaded.

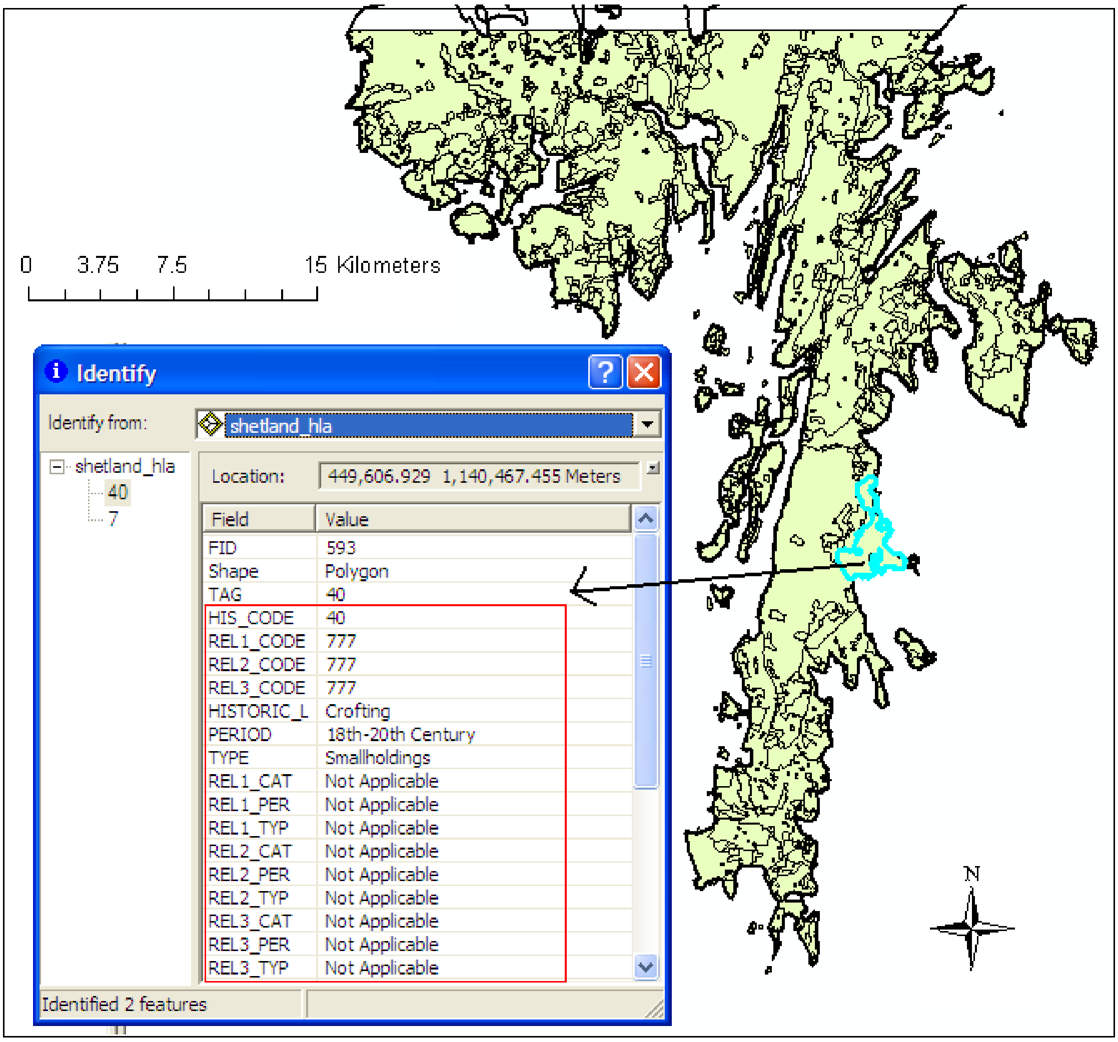


Figure 4.22. The attribute table of the HLA map as it appears in its original form. The red square highlights the information on the variables recorded by the RCAHMS surveyors. The details about the relicts (REL 1 to 3) are not necessary to the purpose of the landscape characterisation analysis and this information needed to be aggregated and made more usable.

During the analysis it became clear that in first instance the data needed to be reduced to a smaller number of classes and possibly amalgamated in fewer fields. Indeed the disadvantage of any process of generalisation is the loss of information but it was thought opportune to continue and ask the experts at the RCAHMS for advice.

In the end only the information on historic periods and relicts could be generalised. The former were grouped in 5 classes, as illustrated in figure 4.23, and stored in a new field, called “new_value”, in the attribute table of the HLA map, while the latter were summarised in terms of presence/absence.

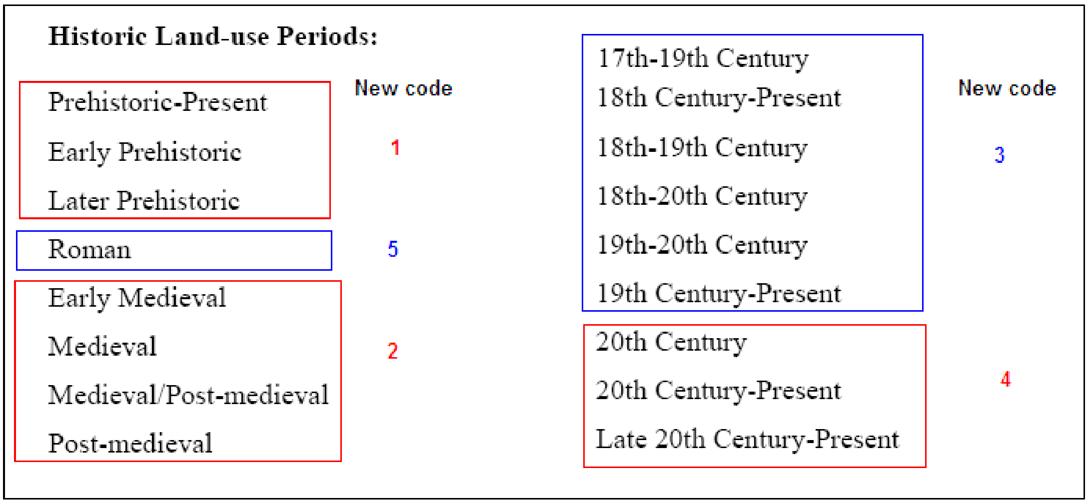


Figure 4.23. Original classes of historic land-use periods and the new codes attributed to them during the process of generalisation (© HLA glossary)

The decision on keeping information on presence/absence of relicts instead of period, types and categories was made on the consideration that for the purpose of landscape characterisation it was sufficient to highlight whether or not a relict land-use was recorded for a surveyed area.

The presence/absence of relicts was calculated in different steps. Firstly three new fields were created, each for relict. Secondly, the code “0” was associated to absence of relicts, this information could be retrieved by selecting the original code 777 (see attribute table in figure 4.22), and the code “1” was attributed to presence of relict by switching the previous selection. Thirdly a new field called “count_rel” was added and basically it contained the sum of the three fields above mentioned, hence a value of 1 corresponded to one relict, 2 indicated two relicts and finally 3 equated to three relicts. By operating in this way it was possible to know how many relicts were recorded for each surveyed area but no information on period, type and category was available anymore.

As far as the types and categories are concerned, it was opted to keep the latter instead of the former since the discrepancies of their number was remarkable:

precisely 14 categories against 59 types. The categories were also re-coded, as depicted in figure 4.24, and stored in a new field called "new_categ".

New code	Historic Land-use Categories	New code	Historic Land-use Categories
1	Built-up Area	8	Moorland and Rough Grazing
2	Crofts and Smallholdings	9	Planned Village
3	Defensive Establishment	10	Recreation Area
4	Designed Landscape	11	Ritual Area
5	Energy Establishment	12	Transport
6	Fields and Farming	0	Water Body
7	Mineral, Waste and Peat Industries	13	Woodland and Forestry

Figure 4.24. Original classes of historic land-use periods and the new codes attributed to them to facilitate the process of data amalgamation. (© HLA glossary)

In the end the new codes for categories, historic periods and the count of relicts were summed together in a new field, and to make its content clearer a descriptive field was added. Therefore the table of the HLA map at the end of all these operations appeared as illustrated in figure 4.25. For the subsequent analysis[12] the descriptive code was used.

[12] The subsequent analysis corresponds to the application of the tabulate area tool of arcGIS which in chapter 5 is explained in much greater detail, thus please refer to it.

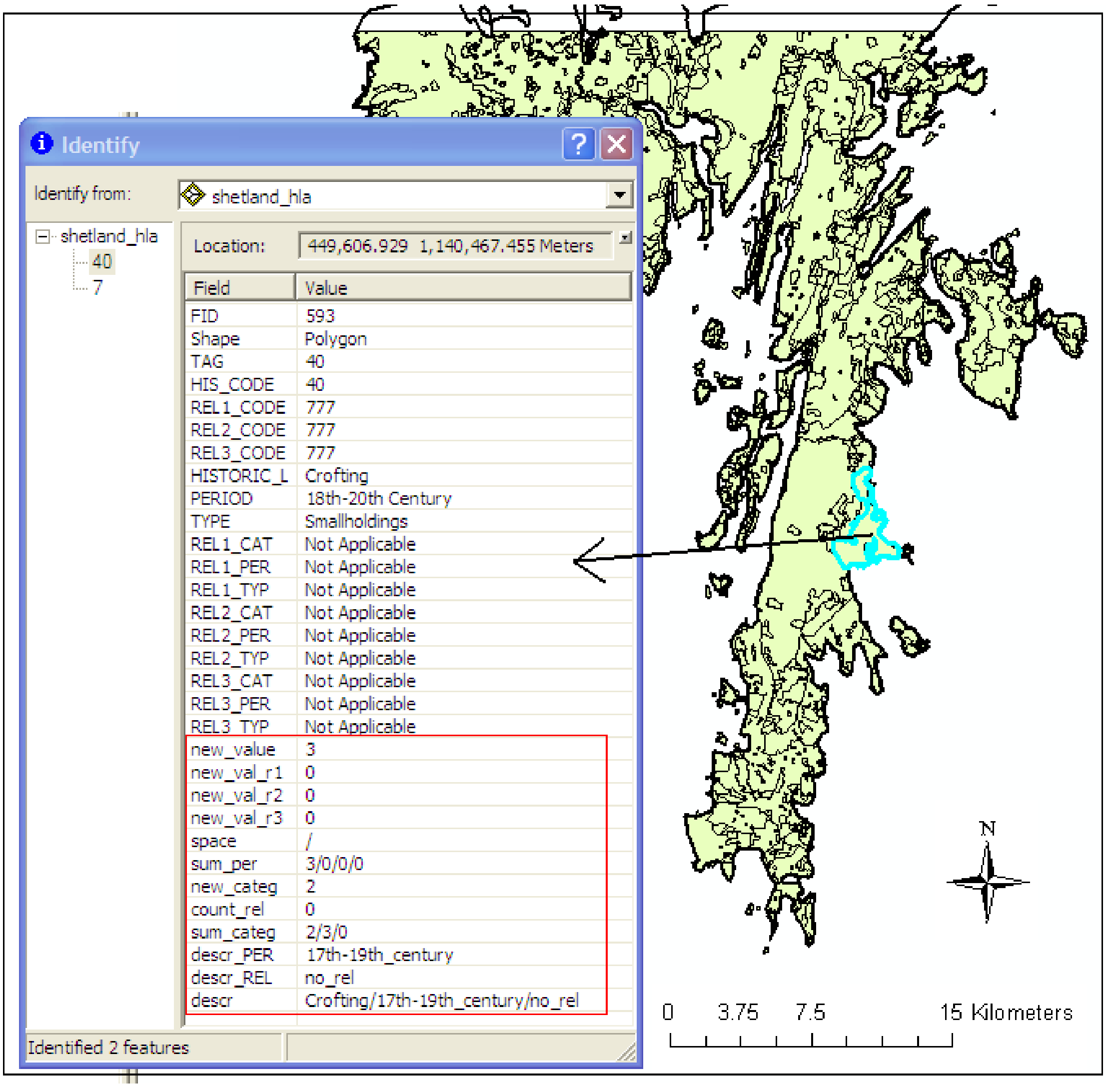
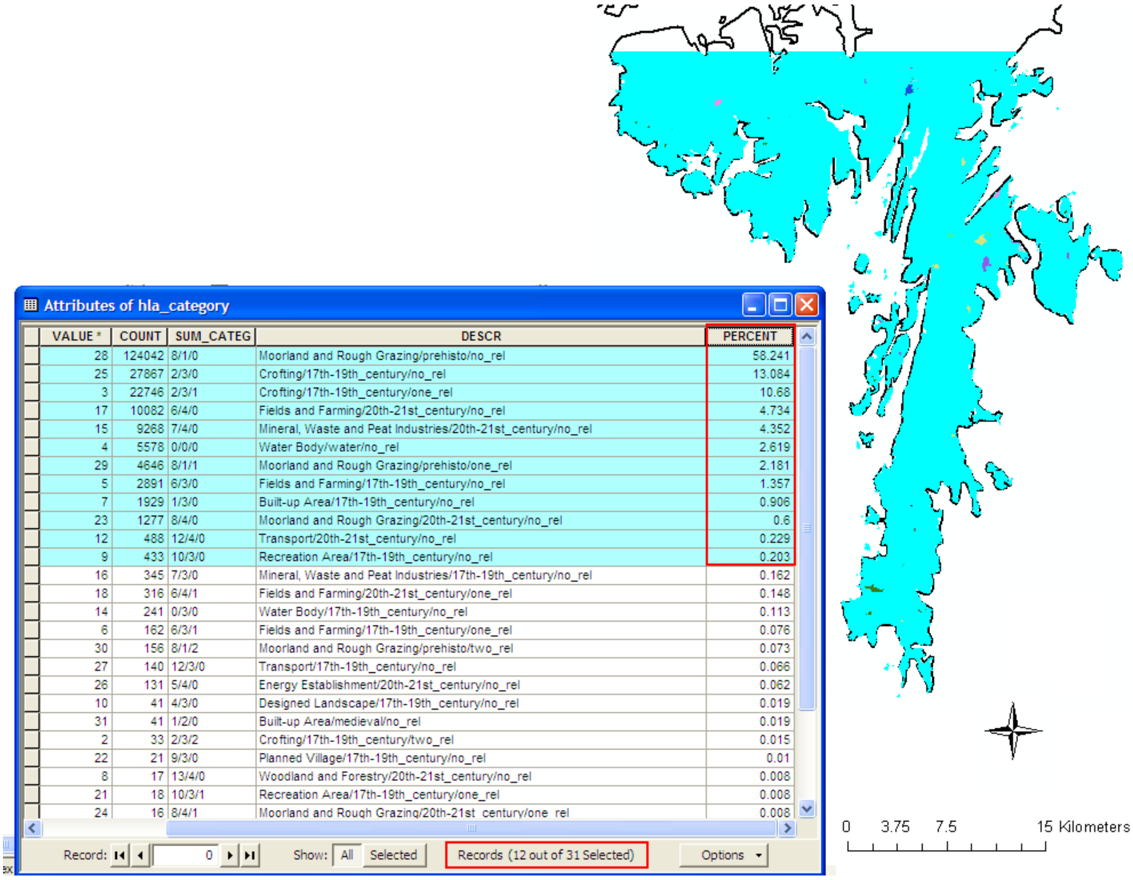


Figure 4.25. The attribute table of the HLA map (in this case for South Shetland) as it appears after the reduction of the number of variables. The new fields added during the calculations are in the red square, notice that only the last field, which condenses three variables (land-use category, period and presence/absence of relicts), is used for the landscape characterisation analysis.

The reclassification did not reduce the number of classes but it operated in a way that few classes could describe the majority of the area of study, as illustrated in figure 4.26.



Attributes of hla_category

VALUE *	COUNT	SUM_CATEG	DESCR	PERCENT
28	124042	8/1/0	Moorland and Rough Grazing/prehisto/no_rel	58.241
25	27867	2/3/0	Crofting/17th-19th_century/no_rel	13.084
3	22746	2/3/1	Crofting/17th-19th_century/one_rel	10.68
17	10082	6/4/0	Fields and Farming/20th-21st_century/no_rel	4.734
15	9268	7/4/0	Mineral, Waste and Peat Industries/20th-21st_century/no_rel	4.352
4	5578	0/0/0	Water Body/water/no_rel	2.619
29	4646	8/1/1	Moorland and Rough Grazing/prehisto/one_rel	2.181
5	2891	6/3/0	Fields and Farming/17th-19th_century/no_rel	1.357
7	1929	1/3/0	Built-up Area/17th-19th_century/no_rel	0.906
23	1277	8/4/0	Moorland and Rough Grazing/20th-21st_century/no_rel	0.6
12	488	12/4/0	Transport/20th-21st_century/no_rel	0.229
9	433	10/3/0	Recreation Area/17th-19th_century/no_rel	0.203
16	345	7/3/0	Mineral, Waste and Peat Industries/17th-19th_century/no_rel	0.162
18	316	6/4/1	Fields and Farming/20th-21st_century/one_rel	0.148
14	241	0/3/0	Water Body/17th-19th_century/no_rel	0.113
6	162	6/3/1	Fields and Farming/17th-19th_century/one_rel	0.076
30	156	8/1/2	Moorland and Rough Grazing/prehisto/two_rel	0.073
27	140	12/3/0	Transport/17th-19th_century/no_rel	0.066
26	131	5/4/0	Energy Establishment/20th-21st_century/no_rel	0.062
10	41	4/3/0	Designed Landscape/17th-19th_century/no_rel	0.019
31	41	1/2/0	Built-up Area/medieval/no_rel	0.019
2	33	2/3/2	Crofting/17th-19th_century/two_rel	0.015
22	21	9/3/0	Planned Village/17th-19th_century/no_rel	0.01
8	17	13/4/0	Woodland and Forestry/20th-21st_century/no_rel	0.008
21	18	10/3/1	Recreation Area/17th-19th_century/one_rel	0.008
24	16	8/4/1	Moorland and Rough Grazing/20th-21st_century/one_rel	0.008

Figure 4.26. With 12 classes (out of 31) that summarise the information on historic period, category of land-use it is possible to describe the whole South Shetland. The column with the percentage values quantifies the coverage of the classes.

To conclude, as far as this current research is concerned, the rich but unmanageable documentation provided by the HLA map needed to be generalised and reduced. Consequently, the contribution of the historic and cultural elements to the character of the landscape was summarised by the reference to historic period, categories of land-use and presence/absence of relicts. The process of generalisation was supervised by operation manager Piers Dixon at RCAHMS who kindly arranged two meetings, one for discussing the way the generalisation should have carried out and the other for checking the outcomes. These ones were presented to a larger group of experts in history and archaeology who commented on the results positively. The advantage of the new approach to the use of HLA data for landscape characterisation is that for the first time the historical elements were considered as relevant as the physical elements.

Indeed this research made a first attempt to include history in a more effective way in the process of landscape characterisation. Further investigation and applications are recommended in order to verify whether or not any kind of loss of information can be reduced and the outcomes can be enhanced.

As explained throughout the chapter, GIS depend on data quality and quantity. To start the process of identification of landscape character a series of datasets was needed and time was spent on the collection of data. It occurred that some datasets were available and accessible and other were missing, nevertheless it was possible to derive information from raw data and subsequently recover the gap within the datasets. In addition new maps could be obtained from other official ones but only when the information was quantifiable through criteria. It was demonstrated that what is measurable is also “mappable”: for example the landforms and settlement types maps were derived respectively from the DEM, by calculating the TPI index, and from a map depicting only buildings, by applying a density kernel.

With these maps the whole framework of data to be used in the analysis for landscape characterisation was completed. Thus, the final and definitive list of datasets used for this research comprises of:

- geological maps of bedrock and superficial;
- landforms;
- land-cover;
- historic land-use;
- settlement types.

The next chapter describes the first part of the statistics applied to the landscape analysis and the results obtained.

CHAPTER 5

"Everything is related to everything else, but near things are more related than distant ones".

Beaches of South Harris from Chaipaval (I.Marengo).

5.1 The role of statistics within the GIS-based methodology

It has been established that characterisation means identification of areas that share the same distinct, recognisable and consistent pattern of elements that makes a landscape different from another. Thus pattern and its identification are placed at the centre of the investigation and are the goals that GIS-based methodology should achieve.

In general there are two ways of identifying patterns: one is by displaying the geographic data on a map, the other is by calculating statistics in order to measure the extent to which features or values are dispersed, clustered or random (Mitchell, 2005). The first approach is likely to lead to a subjective interpretation of the map, while the second approach is certainly more objective and quantitative but, to be functional, requires suitable statistical techniques, which GIS may or not directly provide.

Patterns are the result of the way data distribute and associated to each other (see figure 5.1), consequently the data is the building blocks of the patterns and it is important to understand the characteristics of the data before tackling the analysis of patterns.

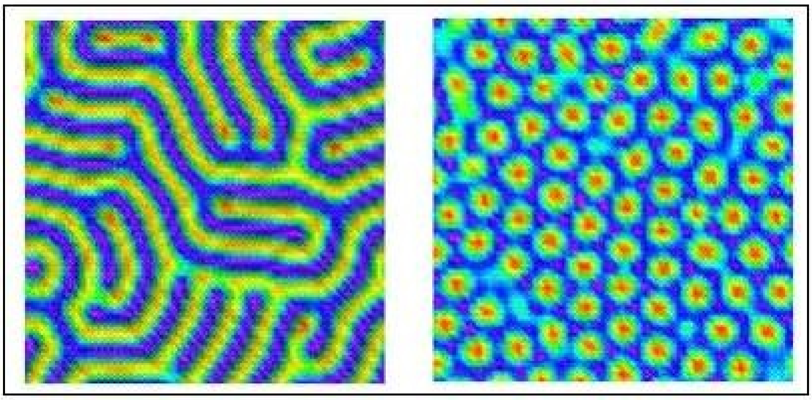


Figure 5.1. The stripes and squares are the data that repeat themselves and form the building blocks that generate the two patterns.

To a GIS analyst geographic data assumes two relevant properties: first of all it is spatially referenced; that means it has a location on the map reflecting its real world position. Secondly it is accompanied by a table that stores its attributes. As a result, the geographic data is not only a coloured point or area on the map but becomes a value that can be measured, analysed and that is related to the space. This way of understanding geographic data is fundamental and at the basis of any GIS analysis, in fact it denotes that if geographic data is not just a symbol then its description through a GIS-based analysis shifts from being purely qualitative to be quantitative.

Statistics provide the tools used in order to analyse data that is expressed numerically because they help to reveal the characteristics of the data collected and under observation, how data is distributed, whether or not the data generate spatial patterns, if data forms trends and if there are outliers (extremely high or low values). In addition, statistics, if properly used, give valuable support in the extraction of additional information from the data that is not obviously detectable by simply looking at a map or plotting raw data.

Nevertheless statistics can range from basic to very complicated concepts and theories and often they do not raise enthusiasm amongst landscape practitioners and officers. In the everyday life the majority of GIS users are normally satisfied with mapping where things are and how they change, finding what is nearby or inside an area, and identifying the largest and smallest values in the field of investigation. In other words the majority of people apply simple descriptive statistics in order to answer to their questions (Krivorunchko and Gotway, 2002).

On the contrary, in the field of scientific disciplines and social sciences, GIS analysts and more demanding users require much more sophisticated methods for spatial analysis and modelling, thus the use of statistics is more a necessity than an option. This current research finds its place amongst the studies that need more than ordinary statistics for the data analysis because it attributes relevance to the space and the spatial relationships between the geographic objects. Therefore, it develops a methodology of analysis around two points:

1. it thinks of landscape character “spatially” and
2. it wants to support the analysis with new statistical techniques called spatial statistics.

“Thinking spatially” means that the spatial reference of the data, for instance easting/northing or latitude /longitude, is as relevant as the value of the data, and this idea translates into the mathematical concept of spatial autocorrelation which says that observations (data) are correlated over some distance in space (Storfer et al, 2007).

Spatial autocorrelation is a spatial statistics that is formulated specifically in order to consider the spatial location of the observed geographic data and is a technique able to capture how geographical data occurs and distribute in relation to space. Consequently spatial autocorrelation recognises the role that concepts such as distance, location, proximity, neighbourhood and region play in the data distribution.

To summarise, thinking spatially considers a geographic objects as a values (variables) taken at specific locations (site) in a defined geographic space (area of study) and considered not in isolation but in relation to its neighbours (Boots, 2003).

The interest in using spatial statistics is that they can be helpful in the identification of spatial patterns. In general three types of patterns can be recognised according to the way the geographic objects associate to each other (see figure 5.2).

Clustered patterns occur when the data shows a distinct and recognisable distribution over the area of study; dispersed patterns generate if data is more uniformly spread across the area, and random patterns happen where data doesn’t show any particular structure and they are unlikely to be controlled by specific processes or mechanisms. This division is more abstract than concrete; in fact usually in the real world it is difficult to observe a well defined boundary between these three types of patterns and the approach commonly adopted aims at finding how close a spatial pattern is to one of the three main categories (Lee and Wong, 2001).

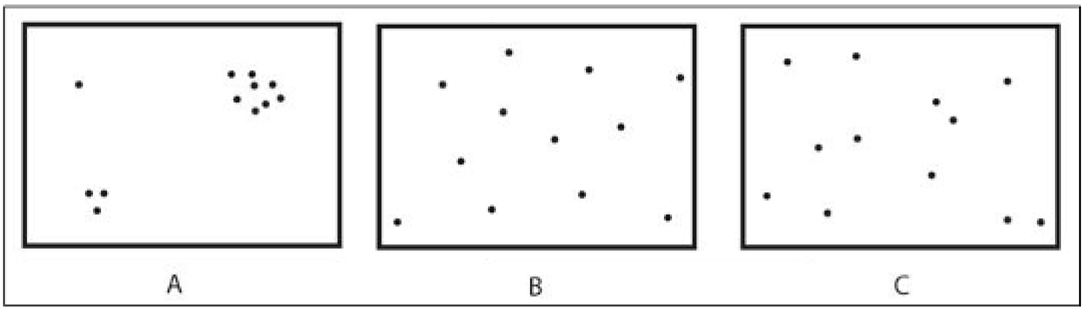


Figure 5.2. Three types of patterns recognised in statistics: A clustered, B dispersed, C random (source: *de Smith, Goodchild, Longley, Geospatial Analysis - a comprehensive guide. 3rd edition).*

Before introducing the main spatial statistics techniques it is relevant to clarify that the current GIS-based methodology is not about explaining the reasons underlying the spatial patterns generated from the unique association of landscape elements. Here we recognise that there are spatial processes that drive and determine spatial patterns; however these processes are not at the centre of the investigation. Instead the aim is to achieve the identification of the spatial patterns since they correspond to the different characters of the landscape.

5.2 An introduction to spatial statistics

After decades of application, GIS have proved to be efficient at spatial data manipulation through a set of analytical tools such as buffering, overlay, map algebra, queries, reclassification, and surface calculations. Thus GIS provide a range of spatial statistics that focuses on how compact or dispersed the features are, whether they are oriented in a particular direction, and whether they form a clustered, uniform or random pattern across the region. Specifically the patterns can be calculated by GIS in terms of distance (how far apart features are), level of clustering (if high or low) or spatial autocorrelation and within the software arcGIS 9.2, the spatial analyst tool comprises of the following set of spatial statistics: the average nearest neighbour distance, the high/low clustering, the spatial autocorrelation (global Moran's *I*) and the multi distance spatial cluster analysis.

For completeness we should mention that GIS also perform geostatistical analyses that use sample points taken at different locations in a landscape and create, by the application of interpolation techniques, a continuous surface. The sample points are used to predict, for any other location in the landscape, the continuous values of particular attribute, such as elevation, temperature, and concentration (ESRI help on line).

Despite geostatistical and spatial analysis are provided by GIS, their use is still fairly restricted to GIS analysts and users with a high confidence and knowledge about the functionalities and limits of these statistical tools. Spatial statistics are often complicated, "unfriendly" and not accessible to everyone because of their difficult task; in addition there are many disciplinary barriers that prevent a fruitful interaction between people from different traditions and disciplines (Krivorunchko and Bivand, 2003).

The difficulties described so far were also encountered in this current research, and in fact the development of the GIS/spatial statistics-based methodology for landscape characterisation took time and effort. Both were spent in learning and understanding the theory behind spatial statistics techniques and which, amongst these, was the most suitable for meeting the purposes of the research. At the end it was discovered that more powerful statistical tools are available from other software and they are generally compatible to the main functions of arcGIS 9.2. This confirmed what Wise et al. (2000) already highlighted, namely that despite the request to expand the range of spatial analysis tools, GIS software vendors are reluctant to include such tools in standard software packages. Hence the current methodology was implemented by using R[13] and some of its free packages as statistical support units, and GIS helped to map and visualise the results.

There are three types of spatial analysis that might be of interest to those working with GIS and are summarised in figure 5.3.

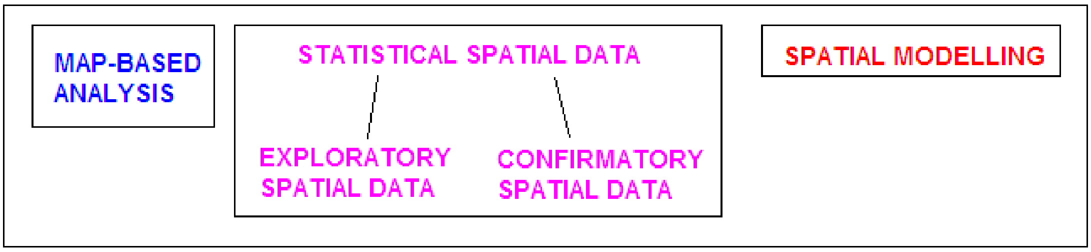


Figure 5.3. The three main spatial analysis that can be carried also jointly with GIS.

The commonest and most accessible is the map-based analysis, which mainly corresponds to a set of tools that help in visualising data distribution and dispersion (Wise et al, 2001). This map-based analysis can also be used for showing interesting patterns; however care should be taken in the interpretation of the output maps, since the representation of the patterns depends on the accuracy used in the classification methods and risks being misleading. For example figure 5.4 illustrates how the distribution of the land cover class “shrubs” in Orkney present a different pattern according to the technique of classification adopted.

[13] R is a free open source environment and it is devoted to the computation of statistics and their advancement.

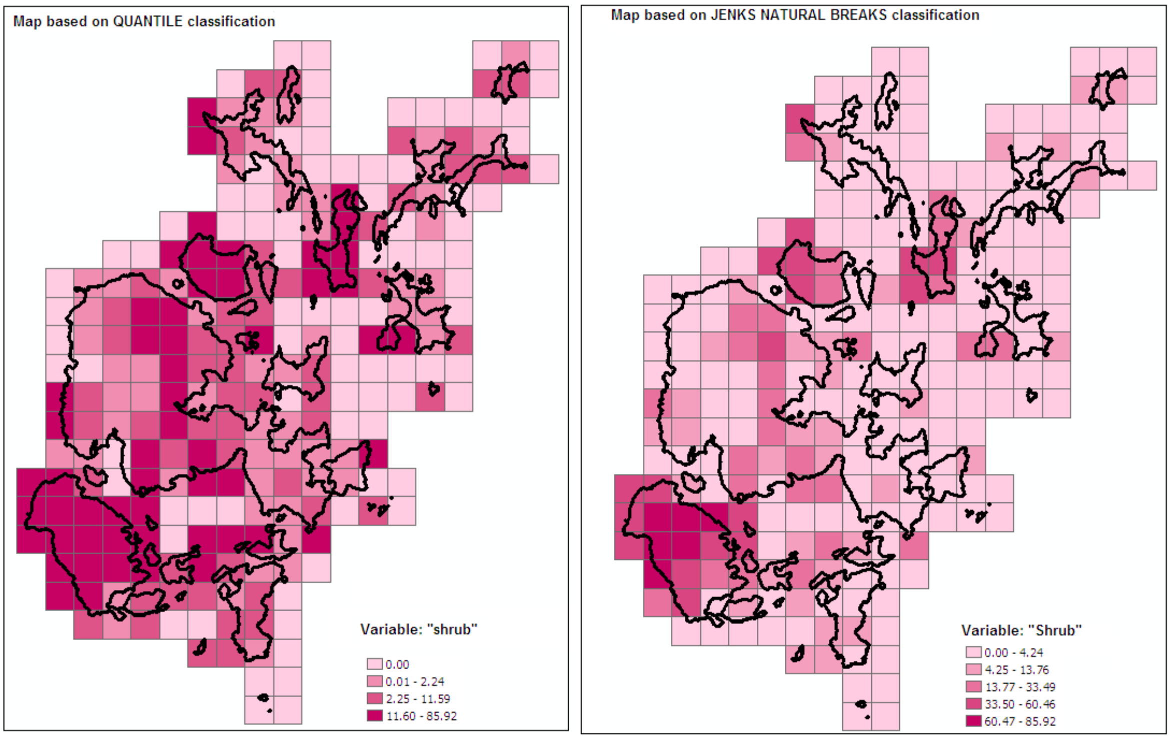


Figure 5.4. Examples of the way thematic maps can be misleading in a pattern recognition based on visual inspection of the map. The classification methods applied (Quantile on the left and Jenks natural breaks on the right) in the maps differ and influence the overall data distribution.

Additionally and equally important, the choropleth map in figure 5.4 does not indicate straightforwardly the typology of a pattern and let us wonder whether the pattern depicted is random, clustered or dispersed.

The second type of spatial analysis is statistical spatial data analysis (SSDA). This can be described as the analysis of empirical spatial data using statistical methods. Two types of SSDA can be identified. The first is exploratory spatial data analysis (ESDA) that is concerned with detecting spatial patterns in data, identifying unusual or interesting spatial features of the data (for example spatial outliers), formulating hypotheses and validating spatial models. The second type of SSDA is called confirmatory spatial data analysis (CSDA) which involves model building, the estimation of parameters and their errors, and hypothesis testing (Wise et al, 2001).

Finally the third type of spatial analysis is spatial modelling and regression which aim at explaining interesting patterns (Wise et al, 2001). The focus of the present research is on the utility of ESDA to uncover interesting patterns of a set of elements in a given landscape taking into account spatial autocorrelation

.

Suppose observing the value of *p* statistical variables on *n* areal units of a study region. Such data is referred as lattice (Scrucca, 2005) and may be represented by either regular, as on a grid, or irregular regions such as zip codes, administrative boundary or census areas (Sain and Cressie, 2007). These kinds of data are often associated to spatial econometric techniques (Scrucca, 2005) and commonly spatial autocorrelation is the method applied for the data analysis. This technique makes a step forward in the detection of spatial patterns because it takes into account not only the location of the data but also its attributes. Spatial autocorrelation recognises the fact that locations may not have similar characteristics; therefore for the determination of spatial patterns the value assumed by the geographic data at each location becomes relevant (Lee and Wong, 2001). From the spatial autocorrelation perspective, the identification of spatial pattern is concerned with the degree to which areas at one location show attributes that are similar to areas in the neighbouring locations. As clearly illustrated in figure 5.5 a positive spatial autocorrelation corresponds to a clustered spatial pattern, while a negative spatial autocorrelation identifies a dispersed pattern. If the value of spatial autocorrelation is close to zero then the spatial pattern is classified as random.

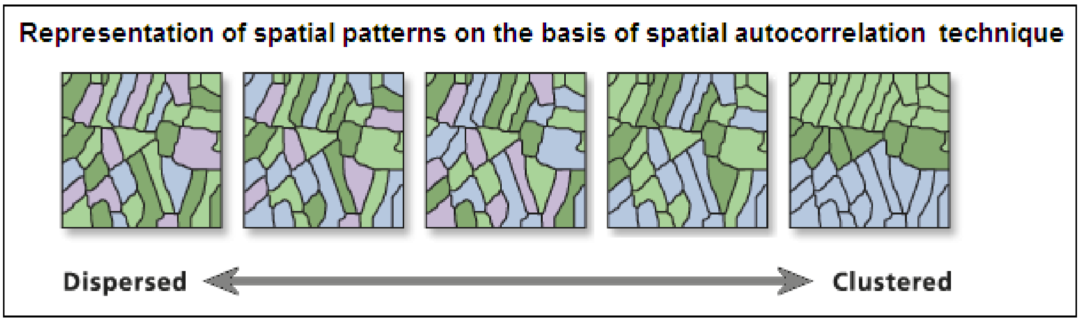


Figure 5.5. In case of areas or regions the spatial statistics technique is spatial autocorrelation which identify a clustered pattern when adjacent areas show the same properties and attributes, otherwise the pattern is classified as dispersed (source arcGIS9.2 help).

Expressed in these terms, spatial autocorrelation translates operationally the first law of geography stated by W. Tobler (1970), which says "everything is related to everything else, but near things are more related than distant things". This sentence apparently looks obvious but is actually quiet profound. On this basis it is hardly surprising that most geographic patterns of interest involve groupings of similar values in clusters (Unwin, 1996) and in fact the basic property of spatially autocorrelated data is the assumption that values are not randomly distributed in space, but spatially related to each other. As a consequence the significance test for spatial autocorrelation compares the observed value to randomly distributed values.

In practice spatial autocorrelation consider both location and attribute values of the data by measuring both the *proximity* of the locations of points/areas and the *similarity* of the attributes of points/areas at their locations. Proximity is calculated in terms of distance between points/areas, while similarity of the attributes is measured as the difference in the values of spatial adjacent points/areas (Lee and Wong, 2001).

There are two indices that measure spatial autocorrelation: Geary's *c* and Moran's *I*. The indices differ in their formula but the concept is mainly the same and namely for measuring spatial autocorrelation both combine the two measures of proximity and similarity into a single index as depicted in figure 5.6.

$$\Gamma = \sum_{i}^{n} \sum_{j}^{n} a_{ij} \; w_{ij}$$

w_{ij} represents the proximity of the point/area *i's* and the neighbouring point/areas *j's*

a_{ij} represents the similarity of points/areas *i's* and points/areas *j's* attributes

Figure 5.6. Index of spatial autocorrelation: Γ is the function that aggregate proximity and similarity in Moran's *I* and Geary's *c*.

The existence of spatial autocorrelation between neighbouring locations can be assessed globally or locally. In the former case the index is calculated across the whole domain of observations and allows the identification of spatial patterns, while in the latter case the index focuses on individual features and their relationship to nearby features allowing the detection of clusters. However in both cases the aim is to determine whether and how value similarities are linked to location similarities.

A map depicting the results from spatial autocorrelation assigns each area a value quantifying how similar it is to its neighbours and where clustered patterns are located. Nevertheless the visualisation of the data is important prior the analysis; nowadays there are software, GeoDA[14] and GGobi just to name a few, that offer a series of graphs, for instance parallel coordinates plot, connectivity histograms and operations, like brushing and linking, that help to examine better the distribution of the data and the association between variables.

[14] GeoDa was designed by Luc Anselin, who is one of the main reference authors in the field of the econometrics studies in order to implement exclusively techniques for ESDA on lattice data; it combines maps with statistical graphics and intends to provide a user friendly and graphical interface to methods such as spatial auto-correlation and indicators of spatial outliers (L.Anselin, 2003). GeoDA was used in this research in order to uinderstand how spatial autocorrelation works and can be calculated. Since geoDA performs only a uni and bi-variate analysis, it could not be applied as the final tool for the identification of the landscape character.

While employing spatial autocorrelation for the pattern analysis two concepts should be kept in mind. The first one is that the indices of spatial autocorrelation, such as Moran's *I*, assume stationarity and that the underlying processes, which determine the spatial patterns, have approximately the same parameters values, specifically mean and variance, for the entire study area (Wagner and Fortin, 2005). Nevertheless, every day we observe an incredible variety of forms, shapes, elements, and landscapes that it is almost unreasonable to describe the real world as an "average" place or take a subset of the Earth's surface and use it as a representative sample of the whole planet (de Smith et al, 2006). The evident variety of forms that we perceive is called spatial heterogeneity and implies non-stationarity, which indicates that mean, variance, covariance of a variable vary across a study area (Wagner and Fortin, 2005). Thus the application of spatial autocorrelation indices is in conflict with the reality, and several authors suggested applying detrending techniques and non parametric methods to address the problem of non stationarity (Haining 1997, Kabos and Csillag 2002 in Wagner Fortin, 2005).

Despite what said about spatial heterogeneity, it is also true that a proper observation of the Earth's surface reveals regions that exhibit internal similarity (de Smith et al., 2006). For example there are areas characterised by abundant rainfall and dense vegetation, or areas dominated by high peaks, glaciers and scarce vegetation, or densely populated areas around major rivers. Indeed there are exceptions to this broad scale pattern and, for example, there are places where the conditions change very rapidly over short distances. For instance between the Carse of Stirling and the Highland boundary fault there is a neat and clear division between the alluvial plain of the Forth river and the hills.

The similarity and homogeneity of forms, landscapes, shapes that we can notice in the real world at a large or fine scale is explained through the concept of spatial dependence, which is implied by spatial autocorrelation. Recalling W. Tobler's law it emerges clearly that without the dependence due to the location of the geographic data everything would be chaotically and randomly distributed, and there would not be the patterns or the similarities that we notice in the real world (de Smith et al., 2006). The implication of spatial dependence in spatial regression and correlation analyses, which try to explain the spatial process underlying the spatial patterns, is significant and in fact spatial dependence violates the assumption of independence amongst the observations. Thus spatial dependence might or not be an issue according to the analyses that are conducted; consequently the analysts are in front of the choice of removing spatial dependence or dealing with it while explaining the causes that generate a pattern.

5.3 The link between spatial autocorrelation and landscape characterisation

In the context of this research the reasons for choosing statistics such as exploratory data analysis which considers spatial autocorrelation are several. First of all, both in its global and local form, the technique identifies patterns, analysing the location proximity and the similarity of the attribute values of the data. This analytical property of spatial autocorrelation corresponds exactly to that described for the lattice data analysis and it was fundamental for the purpose of the research, since the input data could be recognised and distinguished according to their different categories (classes) and location.

Secondly within spatial autocorrelation data is not analysed in isolation rather in the context of its neighbours. This supports the holistic perspective of the landscape, which should be considered as a whole and not as its individual elements.

Thirdly, since the research wanted to identify spatial patterns and not investigate the underlying spatial processes, the presence of spatial dependence was not an issue and was recognised and measured through spatial autocorrelation. At the same time the research ignored the concept of spatial heterogeneity on the grounds that this issue was not such important as for other landscape disciplines.

Nevertheless the main problem in adopting spatial autocorrelation is that it requires lattice data or regions. Hence this research needed to define boundaries within which the landscape elements could be counted. Instinctively, and as suggested by econometrics and other social sciences studies, the first idea was to define the zones on the basis of the official boundaries of the local authorities (see figure 5.7), however two problems were immediately identified and both were related to the size of the areas. First of all, the local authorities extend over surfaces that are too large for the purpose of the analysis. For instance the Highland comprises of so many and different landscapes that smaller zones were needed in order to be able to detect changes in landscape types.

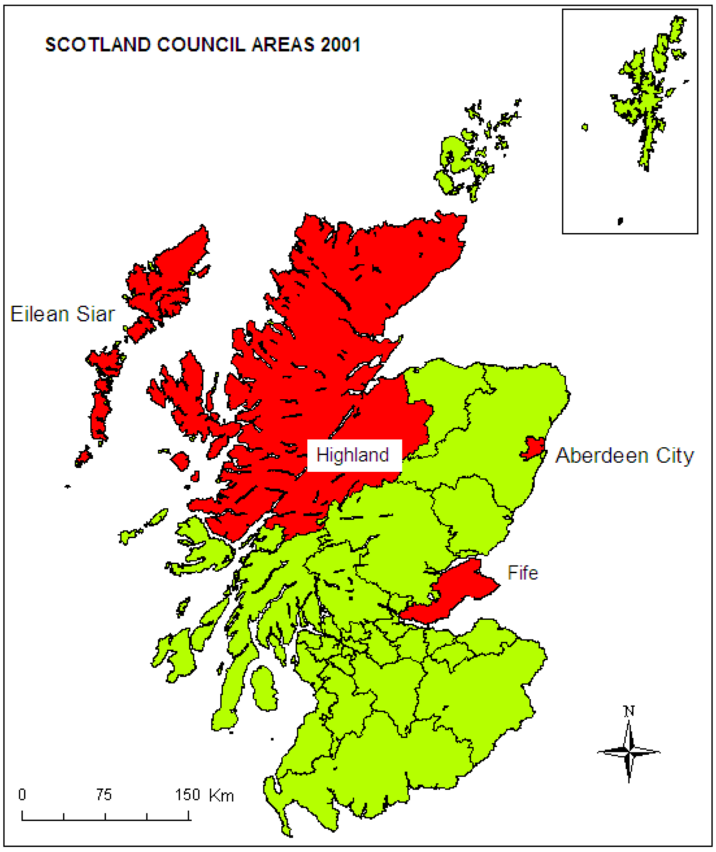


Figure 5.7. Scotland, as it appeared in 2001, divided into its council areas. (Source Census data)

The second evident problem was linked to the difference in size showed by the local authorities. Areas such as Aberdeen city, for example, were dwarfed when compared to Highland, Outer Hebrides (Eilean Siar) or Fife. These discrepancies (see figure 5.8) were also noticed using different administrative units, such as the Census output areas (COAs), the post codes, or the constituencies. In the example of the COAs, was also clear that the small zones corresponded to densely populated areas while the large zone represented mainly the countryside and the less populated areas. Thus, the main concern was that demographic and socio/economic constraints would influence the analysis and could seriously bias the outcomes from the landscape characterisation.

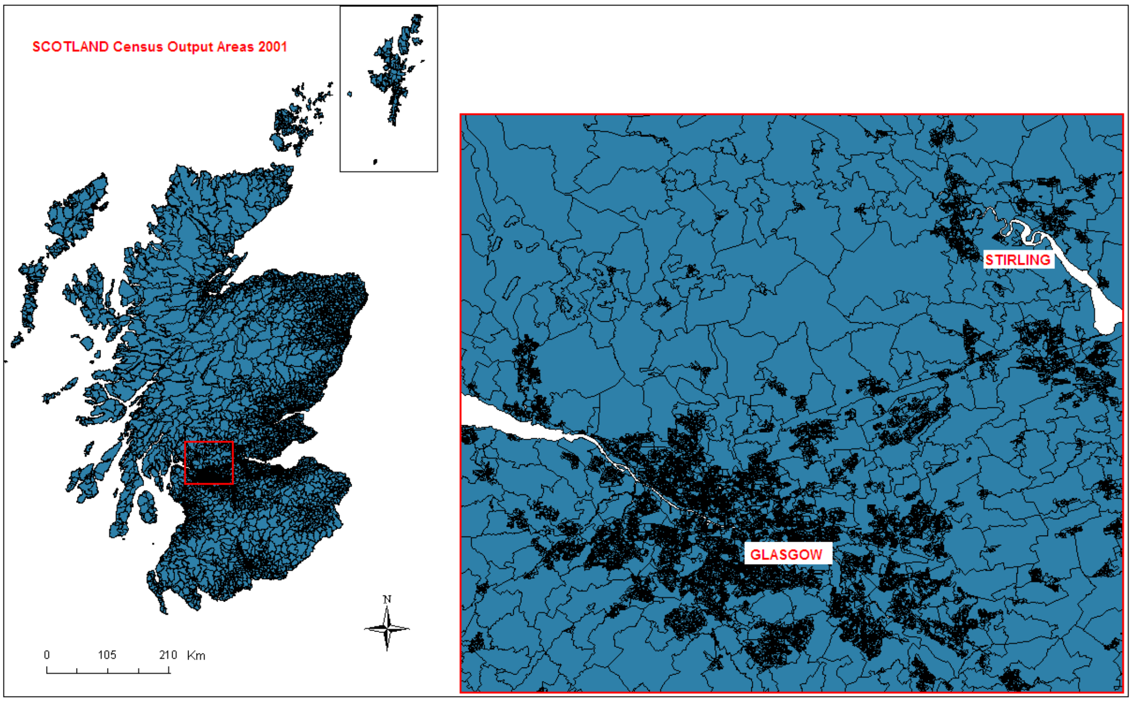


Figure 5.8. Scotland's Census Output Areas (2001). On the right the zoom for the Glasgow and Stirling areas reveal how much the size of the COAs changes when moving from urban areas (small) to the countryside (large).

In addition the use of official boundaries raised other worries; there was awareness that such artificial zones, in particular post codes and COAs, tend to change throughout the years with the result that a comparison between landscape characterisation carried at different years becomes meaningless. In this regard, it is worth remembering that the intention of the research was to develop a methodology able to identify the landscape character types in Scotland regardless of (independently from) their location, therefore tying the landscape characterisation to artificial boundaries was considered not opportune (adequate).

Finally, thoughts of the Modifiable Areal Unit Problem or MAUP, which is defined as "a problem resulting from the imposition of artificial units of spatial reporting on continuous geographical phenomenon resulting in the generation of artificial spatial patterns (Heywood et al., 2006)", were unavoidable. Areal data, such as those that frequently occur in the social sciences, cannot be measured at a single point but has to be contained within a boundary. For example the percentage of unemployed or the rate of births has to be related and calculated over an area. Therefore it is common that individual data are assigned to areal units that form the base layer of spatial units for the study of some phenomenon (Lock and Leigh Molyneaux, 2006). The two problems connected to MAUP are scale and zonation effects (see figure 5.9). The former relates to the variation that occurs when the same set of areal units is grouped into larger ones, while the latter is the variability in the outcomes that can occur when the set of areal units is recombined in a different way (Armheim, 1998 in Kang-Rae and Banister, 2006).

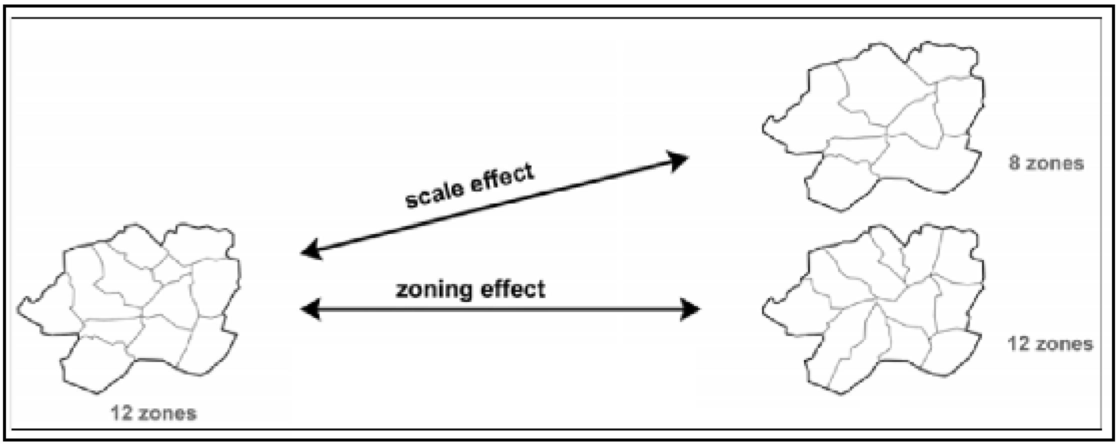


Figure 5.9. The scale and zoning effects related to MAUP. (Source: Kang-Rae and Banister D., 2006)

As explained, this current research had to use areal units in order to calculate the spatial autocorrelation and the results may be affected by MAUP. Consequently tests[15] were carried out in order to verify whether or not the MAUP would be a problem for the analysis and for the methodology implemented. Chapter 7 provides the details and the results of the tests.

To conclude, on the basis of what described above, it was decided not to use official boundaries but instead it was chosen to refer to an artificial grid mad up by randomly generated polygons. A more detailed explanation is provided below.

5.4 The areas of study and the sampling zones

The first step consisted in selecting eleven areas across Scotland which could represent different landscapes (see figure 5.10). The choice of the areas was mainly based on the personal knowledge of them which was thought being advantageous due to their distance from Stirling.

Nine areas were identified by a square of 2500 km², one (Islay and Jura) was framed by a polygon of 2390 km², and one had “real” boundaries since it corresponded to Cairngorms National Park, precisely 3816 km². The dimension of the areas of study was considered adequate for the analysis since it was able to assured heterogeneity in terms of classes of landscape elements and manageability of data; in fact the raster maps extracted for each area comprised of 1000 columns and 1000 rows, which totalled a million of cells of 2500 m² each.

[15] Due to the initial stage of the analysis we were not sure about the influence of MAUP on the outcomes. Zonation might have an effect, but scale should not be an issue in this research since the units were not aggregated. The test in chapter 7 will provide more details.

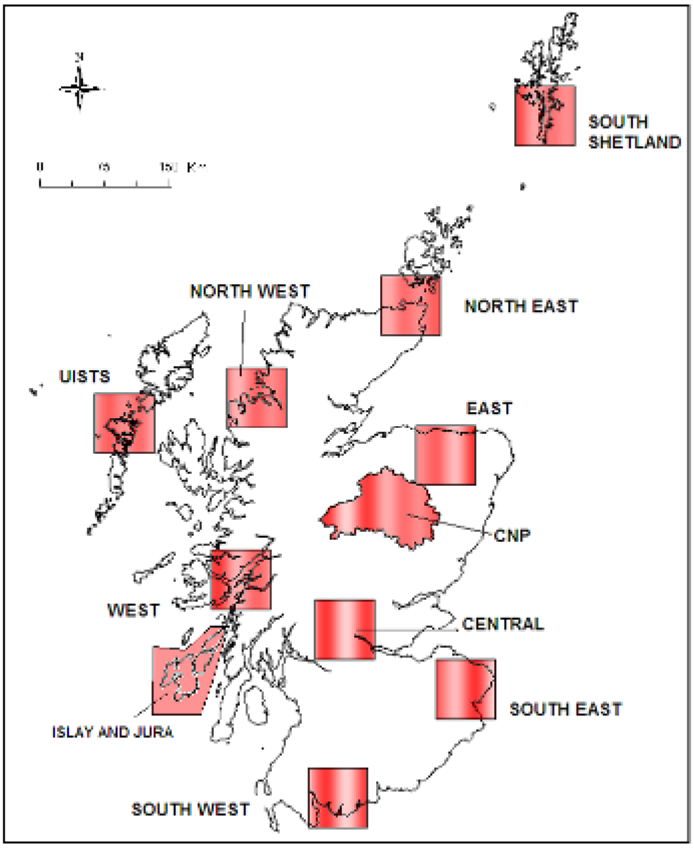


Figure 5.10. Areas of study selected across Scotland in order to cover different landscapes. Except for the Cairngorms National Park, which represents the only official boundary, the rest of the areas are defined by squares of 2500 km² that may cut across different local authorities. For instance the area "East" covers part of Moray and Aberdeenshire.

The second step was to divide the areas in smaller units that would correspond to the sampling zones within which the classes of landscape elements were counted. Quantifying an individual by counting its presence is a technique used in landscape genetics[16] and landscape ecology when for example the purpose is to count the number of species at observed sites (Storfer et al., 2007). Because landscape ecologists and geneticists often look at the landscape as a factor that can influence spatial patterns within their data, it was judged appropriate to try and identify a sampling method similar to that used by these researchers. The parallelism between

[16] In the last five years, landscape genetics has emerged as a new research area that integrates population genetics, landscape ecology and spatial statistics. As a result the literature provides reference of several attempts of understanding the landscape effects on gene flow, genetic discontinuities and genetic population structures (Storfer et al, 2007).

the sampling methodology used by landscape geneticists and ecologists and this current research is explained in figure 5.11.

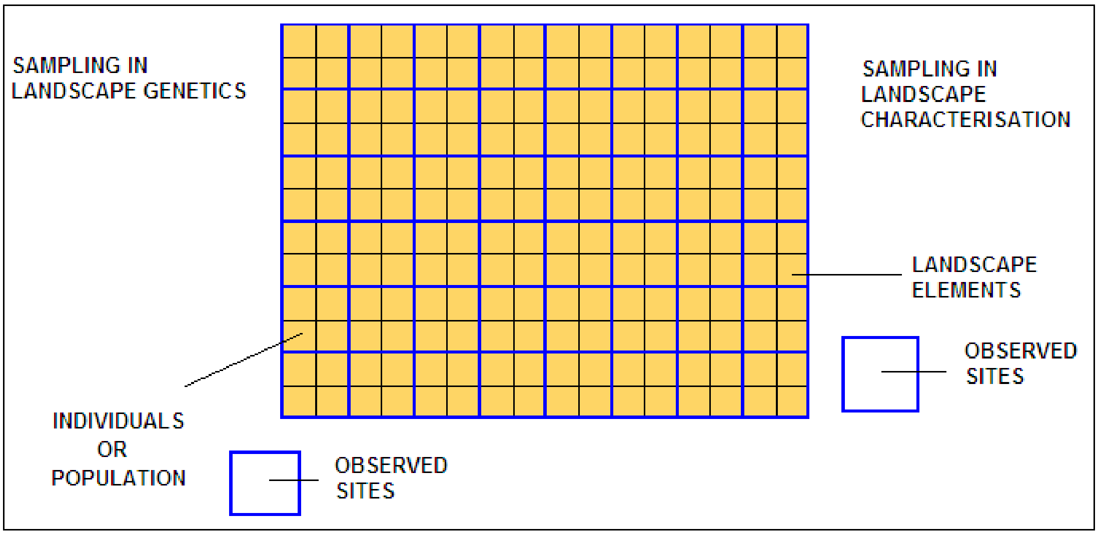


Figure 5.11. Representation of the similarities caught between the way of sampling in landscape genetics and that used for the analysis of landscape characterisation

As depicted, the individuals or population were matched with the landscape elements counted at observed sites (or sampling zones) within the study area. There are two ways of defining the sampling zones: one is through the creation of a uniform grid and the other is through randomly generated polygons (see figure 5.12).

The uniform grid is a systematic method of generating areal units; it consists of defining the size of the cells of the grid and then replicating the cells *n* times across the area of study. This method was discarded because it might happen that by coincidence the grid pattern has the same frequency of a pattern of landscape elements. If this perfect overlap occurs, the grid would not detect the landscape element pattern (Storfer et al, 2007); in contrast, this risk cannot occur when a series of contiguous polygons is generated randomly, and consequently the choice was made in favour of this method.

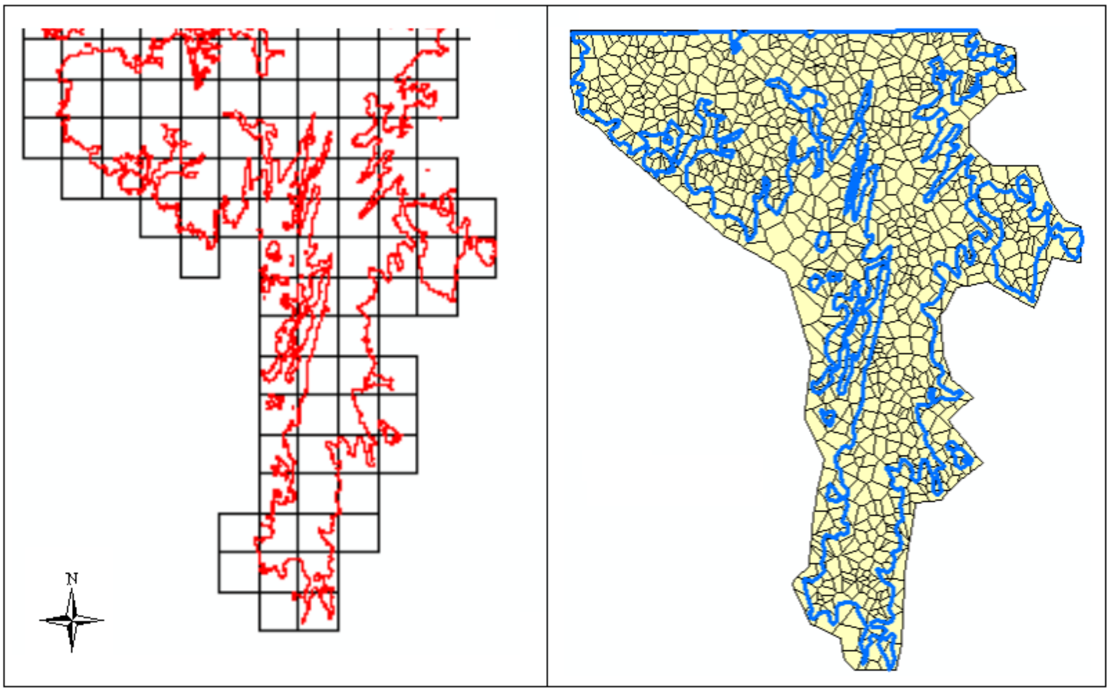


Figure 5.12. Two ways for defining the sampling zones: on the left a regular grid and on the right a randomly generated set of contiguous polygons. Area of study: South Shetland.

At this point it is important to remember that for a correct calculation of spatial autocorrelation no gaps are allowed within the areal units. Thus all the polygons randomly generated have to be continuous and the so called "islands effect" has to be avoided. For instance figure 5.13 shows a correct way of generating random polygons across a geographic area comprised of real islands. As noticeable the small islands between North Uist and South Harris are all connected through the polygons, hence each of them will have neighbours.

Afterwards a set of random points[17] was generated independently for each area of study in order to cover the whole surface; nevertheless because the interest was in identifying patterns of landscape elements on land, some of the points that fell into the sea were selected and moved onto the land. This data manipulation did not influence the final result of the characterisation analysis because the sea was always identified

[17] The random points can be easily generated within arcGIS 9.2 with a tool "create random points" that is stored in data management tools/feature class (arc toolbox)

as a distinct and unique pattern, well differentiated from the others. The main visible consequence from moving the random points from the sea to the land was that in the sea the points were fewer and more distant from each other, thus the size of the polygons generated from these points was slightly larger than those on the land (see figure 5.13).

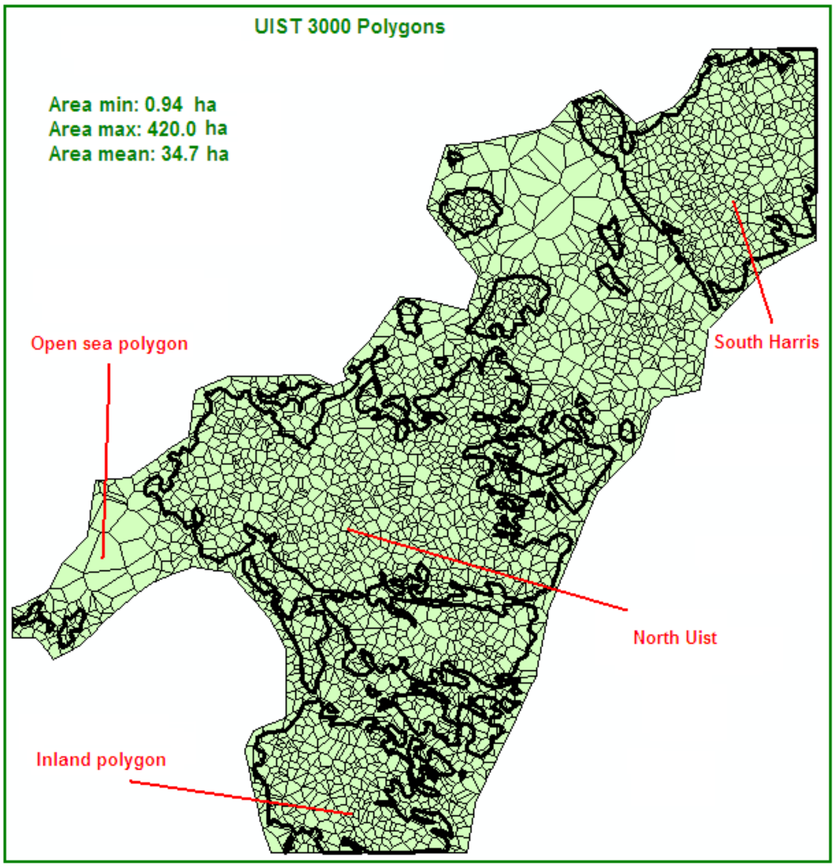


Figure 5.13. The 3000 randomly generated polygons cover entirely and contiguously the area of study of North Uist and South Harris which is comprised of many islands. Notice the difference in size of the polygons drew in the open sea and inland.

Due to the importance of location, the geographic coordinates, easting and northing of the British National Grid, were calculated for each random point, and then the Voronoi or Thiessen polygons were derived from the points as illustrated in figure 5.14.

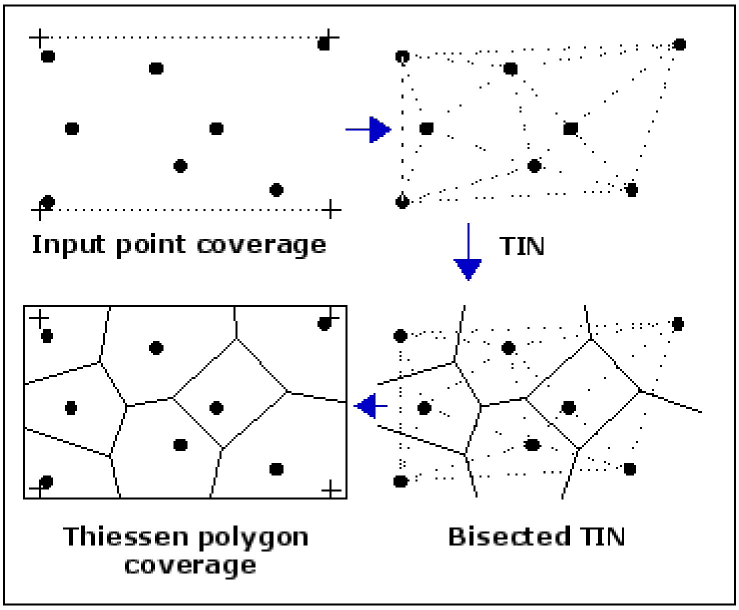


Figure 5.14. Given a set of random points a TIN (triangulated irregular network) are generated then perpendicular bisectors of the lines connecting the points are traced in order to derive the Thiessen polygons (Source: ArcGIS help-on-line).

There are two key points about the sampling technique adopted that need to be kept in mind. First of all the number of points, and consequently polygons, determines the scale[18] of the analysis, in a way that the greater the number of points, the smaller the area of the polygons and the finer the scale (see figure 5.15). Due to the well-known influence exerted by scale in geographical analyses it was thought opportune to quantify how much scale would contribute to the spatial variation of the landscape character types by generating two sets of random points and precisely 1000 and 3000 points[19]. Tests to verify the differences in the maps of landscape character types at 1000 and 3000 polygons (scale) are described in chapter 7.

[18] There is a difference between the scale as intended in MAUP and here. In the first case the scale effect is caused by the aggregation of the spatial units, here scale is determined by the change of the number of random polygons. Being randomly generated, 1000 and 3000 show different boundaries, hence 1000 polygons shape file is a new file and not an aggregation of the 3000 polygons.
[19] In the case of North Uist and South Harris area of study, the minimum area calculated in the map with 1000 polygons is 5.59 hectares which correspond to the size of 12.42 football pitches. While the minimum area measured in the map at scale 3000 polygons is 0.94 hectares, and namely 2.08 football pitches. The area of a cell in a raster map was 2500 m^2. The area of a football pitch is approx. 4500 m^2. Thus the two scale adopted seemed appropriate for the analysis.

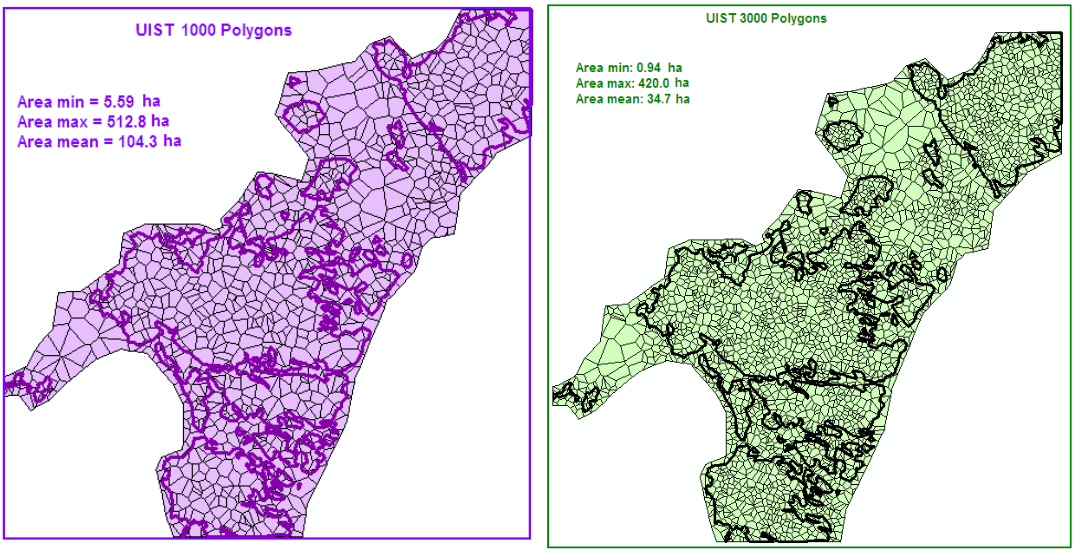


Figure 5.15. Comparison of cell size values between sets of randomly generated polygons. It is clear how 3000 polygons are more densely distributed and smaller than the respective 1000 ones. These characteristics translate into a finer scale of investigation.

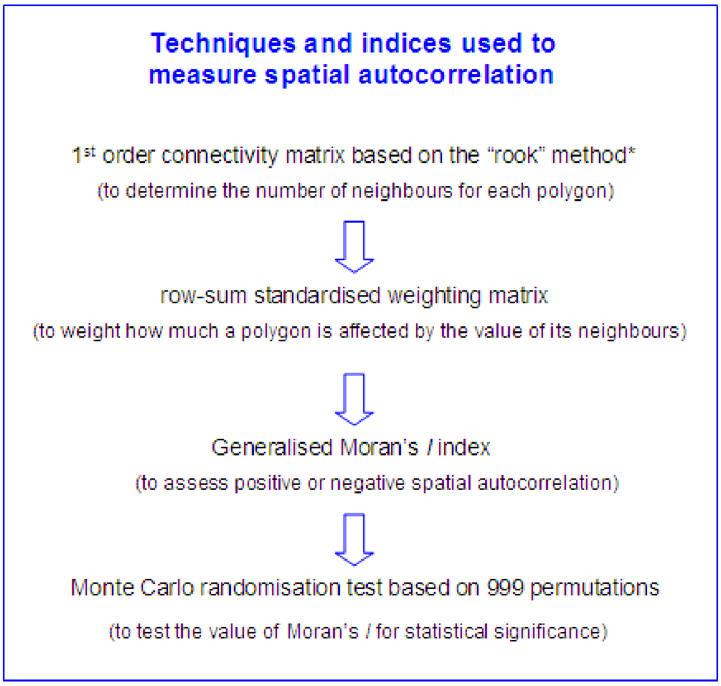
In addition to scale, the choice about the number of random points was constrained by the computational capacity of both the tabulate area tool and the cluster analysis. ArcGIS 9.2 crashed several times when tabulating a shape file of 6000 polygons and the raster maps (which, as previously mentioned, have 1000 rows and 1000 columns). Similarly Minitab, software used for the cluster analysis, gave up the calculations when more than 3000 polygons were processed.

The decision on the number of points and polygons highlights another time that a personal judgement is almost unavoidable, therefore it is acknowledged that the methodology is quantitative but not totally objective.

At this point the information for the calculation of spatial autocorrelation was ready to be processed in R by the package called *spdep*.

5.5 Calculation of the neighbourhood network and spatial autocorrelation

Spatial autocorrelation implies a series of operational steps that are summarised in the scheme below.



As noticeable the starting point is to determine the number of neighbours for each Thiessen polygon. There are two main ways of defining a neighbour relationship: one is based on taking into account adjacency and the other distance. The literature reviewed is relatively limited about the use of method and criteria for defining the neighbourhood relationships. However, Dray (2008) recommended for areal units, like polygons, the use of the adjacency criterion and therefore this was chosen for the current research.

Adjacency uses either the rook or the queen methods and as illustrated in figure 5.16 if the queen method is applied then 8 neighbours are identified for the unit number 6. In contrast, if the rook method is chosen then the number of neighbours for the unit 6 is halved. This discrepancy happens because in the first case the units surrounding the unit number 6 have to share only a point in order to be classified neighbours. In the second case only the units that share a boundary, with length greater than zero, are considered neighbours of the unit number 6.

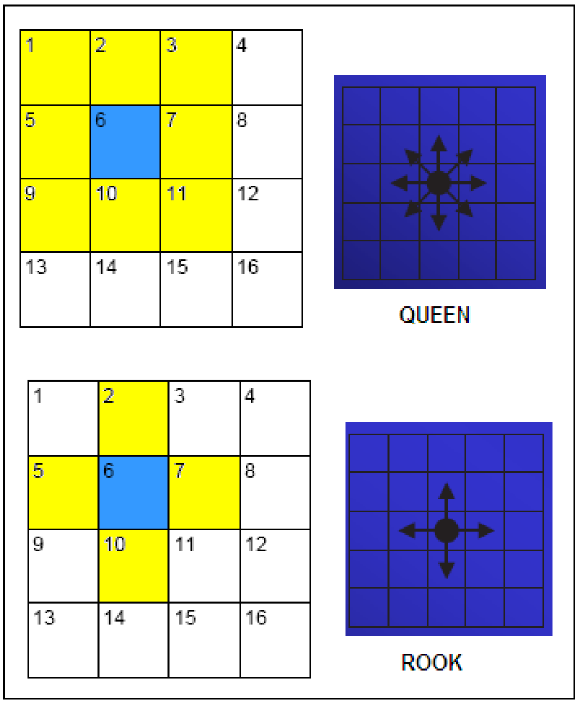


Figure 5.16. The most simple/easy example of lattice data is represented by a grid (left). The two methods that identify the neighbours on the basis of first order contiguity rules are the rook and queen (right). The target unit (number 6) is in blue and the neighbours are in yellow. The difference between rook and queen lies on the criterion adopted by the two methods (source: Voss and Ramsay, 2006).

In its works Dray (2008 and Dray et al. 2008) did not add further specification about the rook and the queen method. In order to clarify which method would be the most suitable for the Thiessen polygons a further investigation was carried out. The area of study “East”, in figure 5.17, is taken here as example. The results were very interesting and added more knowledge about the way the contiguity method works when applied to polygons.

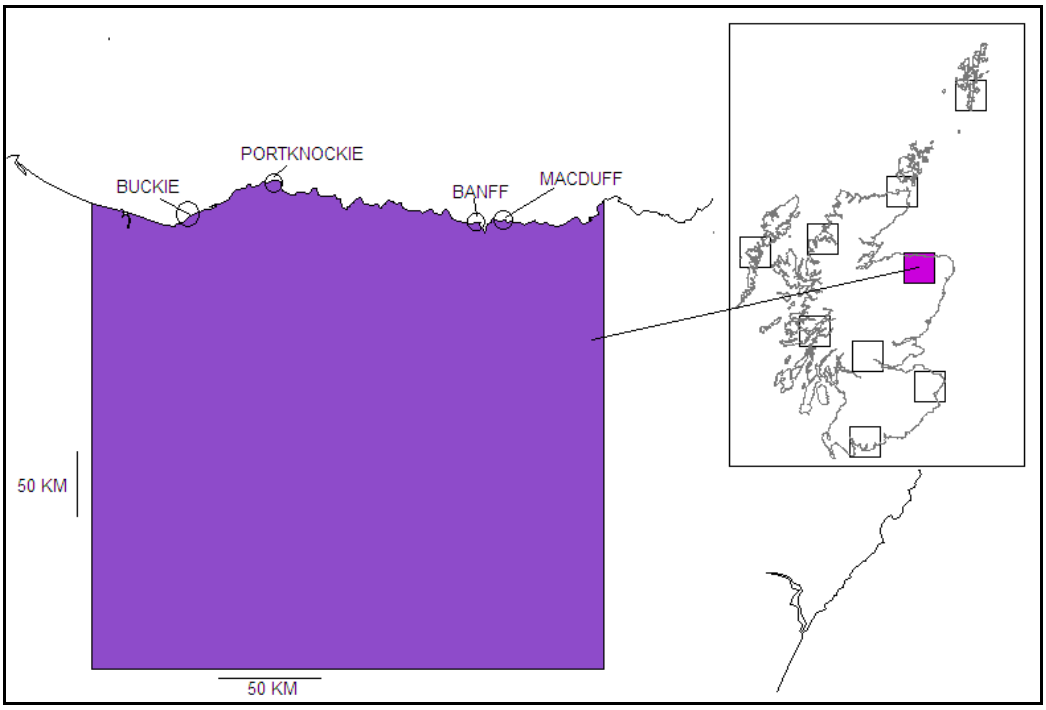


Figure 5.17. Location of the area of study “East”

As depicted in figure 5.16, for the polygon considered (in blue) it appears that the queen and rook methods lead respectively 8 and 4 neighbours. This is true in a regular grid, nevertheless it was observed that for Thiessen polygons, derived from randomly generated points, the queen and the rook methods produce exactly the same number of neighbours as illustrated in figure 5.18. The reason is because the points are connected through lines that follow identical directions (see previous figure 5.14).

Figure 5.18

Before carrying on with the calculation it is essential to check that all the polygons have at least a neighbour, which means that all of them are contiguous, otherwise there is a risk to incur in the so called “island effect”. For an island in fact it is impossible to define neighbours, thus isolated polygons similar to islands have to be avoided.

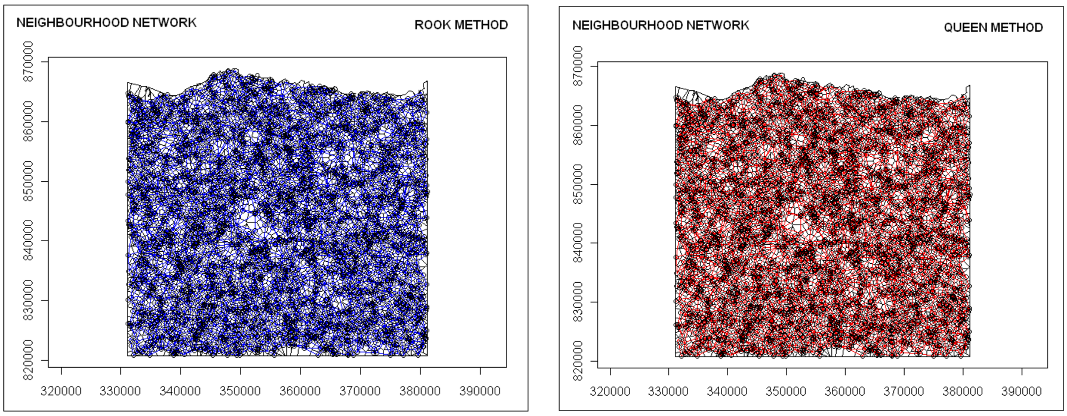
```
> east.nb <- poly2nb(eastpoly, queen = F)
> plot(east.nb, xy_east, col="blue", add=T)
> summary(east.nb)
Neighbour list object:
Number of regions: 3000
Number of nonzero links: 17592
Percentage nonzero weights: 0.1954667
Average number of links: 5.864
Link number distribution:

  2   3   4   5   6   7   8   9  10  11
  8  72 393 792 846 504 275  81  24   5
8 least connected regions:
1473 1511 1647 1653 2308 2920 2943 2981 with 2 links
5 most connected regions:
575 2067 2393 2442 2613 with 11 links

> queen_east.nb <- poly2nb(eastpoly, queen = T)
> plot(queen_east.nb, xy_east, col="red", add=T)
> summary(queen_east.nb)
Neighbour list object:
Number of regions: 3000
Number of nonzero links: 17592
Percentage nonzero weights: 0.1954667
Average number of links: 5.864
Link number distribution:

  2   3   4   5   6   7   8   9  10  11
  8  72 393 792 846 504 275  81  24   5
8 least connected regions:
1473 1511 1647 1653 2308 2920 2943 2981 with 2 links
5 most connected regions:
575 2067 2393 2442 2613 with 11 links
```

a)



b)

Figure 5.18. Contrary to a regular grid, the queen and rook methods applied to Thiessen polygons provide the same results in terms of number of neighbours. a) Shows the summary of the links (neighbours) and weights calculated by using *spdep* package in R. As noticeable the values are exactly the same and it is not a surprise that the graphs (b) displaying the neighbourhood network coincide.

As illustrated in figure 5.19 in R it is relatively easy to retrieve the information about the polygons and their neighbours. If "zero" is absent in the line that indicates the *neighbours by number of polygons* it means that all the polygons have at least a neighbour. In the example eight polygons have the minimum number of neighbours which for East is two.

```
> east_ne.nb <- poly2nb(east_nepoly, queen = F)
> plot(east_ne.nb, xy_east_ne, col="blue", add=T)
> summary(east_ne.nb)
Neighbour list object:
Number of regions: 1839
Number of nonzero links: 10740
Percentage nonzero weights: 0.317571
Average number of links: 5.84013
Link number distribution:

  2   3   4   5   6   7   8   9  10  11
  8  52 221 496 514 337 154  45   8   4
8 least connected regions:
94 177 352 690 718 1299 1425 1740 with 2 links
4 most connected regions:
493 660 1296 1821 with 11 links
```

Total polygons
Total neighbours
Neighbours by number of polygons
Polygons by number of neighbours

Figure 5.19. The summary of the rook contiguity method is very informative and allows the number of neighbours to be easily and quickly checked. In R the polygons and the neighbours are referred respectively as regions and links. In this case there are not island polygons and as indicated, the range of neighbours goes from 2 to 11. Respectively, 496 and 514 polygons have 5 or 6 neighbours.

The spatial relationships amongst the polygons are conventionally stored and organised in a matrix and are represented individually by a row and a column (Lee and Wong, 2001). In this research *spdep* produced a binary connectivity matrix because it attributed value equals to 1 to adjacent (neighbouring) cells otherwise the cell were equal to 0. A binary connectivity matrix reveals several interesting characteristics: the elements along the major diagonal score zero because it is assumed that no polygon is a neighbour of itself; the sum of the cell values by row provides the number of neighbours identified for a unit; the matrix is symmetric, so that from the major diagonal, which divides the matrix in two triangles, the upper triangle of cells is the mirror of the lower triangle (Lee and Wong, 2001).

Overall this is not a good point for the matrix, because it means that redundant information is stored, thus in order to keep a more manageable amount of information it should be necessary to isolate only the polygons with neighbours and then group them in a new matrix and use it for measuring the spatial autocorrelation.

Reducing the size of the binary connectivity matrix and avoiding redundancy in the information stored is optional and it depends essentially on the capacity of the personal computer in use. Instead, the enhancement of the connectivity matrix through an operation of standardisation is compulsory. In fact although the matrix tells us whether or not the polygons have neighbours, it doesn't indicate the strength of the relationship between the spatial units (Dray, 2008). The solution is a standardised (weighted) matrix, although the choice of the weight is a critical decision because it defines the limits of the autocorrelation indices[20] (Dray, 2008).

When the influence of each neighbour is considered the same, the standardised (weighted) matrix is called row-sum standardised matrix. In brief, the weights are derived by calculating the ratios of each polygon with respect to their total influence which is equal to 1 since each neighbour exerts the same amount of influence on the central polygon (Dray, 2008 and Chessel et al, 2003).

On this basis, the weight is simply the average of the neighbours and can be expressed as

weight = 1/number of neighbours

The resulting matrix is no more symmetric but still presents a major diagonal with cells equal to zero (see figure 5.20).

The row sum standardise weighing matrix is in general preferred because it is not dependant on the number of neighbours, thus it allows a correct analysis in case of an irregular lattice dataset (Dray et al., 2008). In fact in case of a regular lattice data, for example a uniform grid, the number of neighbours is constant and a binary connectivity matrix can be used for the calculation of Moran's *I* without being weighted.

[20] For instance a row-sum standardised weighting matrix the maximum of Moran's *I* is 1. While for non standardised weighting matrix the maximum depends on the values in the matrix (Dray et al, 2008)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Σ
1	0	1			1	1											3
2	1	0	1		1	1	1										5
3		1	0	1		1	1	1									5
4			1	0			1	1									3
5	1	1			0	1			1	1							5
6	1	1	1		1	0	1		1	1	1						8
7		1	1	1		1	0	1		1	1	1					8
8			1	1			1	0			1	1					5
9					1	1			0	1			1	1			5
10					1	1	1		1	0	1		1	1	1		8
11						1	1	1		1	0	1		1	1	1	8
12							1	1			1	0			1	1	5
13									1	1			0	1			3
14									1	1	1		1	0	1		5
15										1	1	1		1	0	1	5
16											1	1			1	0	3

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Σ
1	0	1/3			1/3	1/3											1
2	1/5	0	1/5		1/5	1/5	1/5										1
3		1/5	0	1/5		1/5	1/5	1/5									1
4			1/3	0			1/3	1/3									1
5	1/5	1/5			0	1/5			1/5	1/5							1
6	1/8	1/8	1/8		1/8		1/8		1/8	1/8	1/8						1
7		1/8	1/8	1/8		1/8	0	1/8		1/8	1/8	1/8					1
8			1/5	1/5			1/5	0			1/5	1/5					1
9					1/5	1/5			0	1/5			1/5	1/5			1
10					1/8	1/8	1/8		1/8	0	1/8		1/8	1/8	1/8		1
11						1/8	1/8	1/8		1/8	0	1/8		1/8	1/8	1/8	1
12							1/5	1/5			1/5	0			1/5	1/5	1
13									1/3	1/3			0	1/3			1
14									1/5	1/5	1/5		1/5	0	1/5		1
15										1/5	1/5	1/5		1/5	0	1/5	1
16											1/3	1/3			1/3	0	1

Figure 5.20. Top Binary contiguity matrix; the value 1 is associated to the cells that are adjacent and therefore neighbours. The last column sums the values by row and indicates the number of neighbours. **Bottom** Row-sum standardised matrix. The value of the cell corresponds to the weight (1/number of neighbours). As noticeable the last column, indicated with Σ, is always equal to 1 and is the sum of the influence exerted by each polygon (source: Voss and Ramsay, 2006). For clarity the value of 0 has been associated only to the diagonal at the centre of the matrices; however all the blank cells have the value of 0.

On the contrary, in presence of an irregular lattice data, the number of neighbours is not constant and has an effect on the calculation of Moran's *I*. The row sum standardisation method controls/overcomes the effect due to the number of neighbours by making each neighbour to exert the same influence (Dray et al, 2008).

On the basis of the considerations explained above and because the data in this current research assumes the form of irregular lattice (randomly generated Thiessen polygons) it was decided to adopt a row-sum standardised weighting matrix, based on the first order contiguity rook method. To be precise the package *spdep* firstly calculates the binary connectivity matrix (first order rook method) and secondly derives the row-sum standardised matrix as illustrated in figure 5.21.

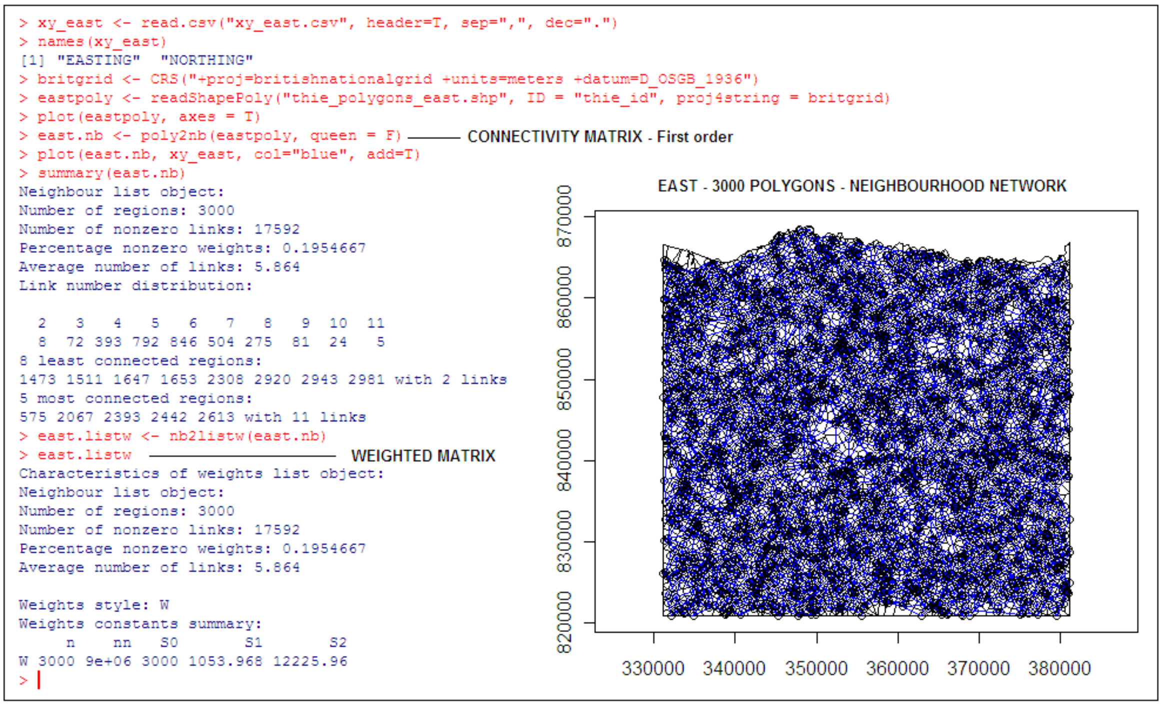


Figure 5.21. Binary connectivity matrix and row sum standardised weighted matrix calculated with *spdep* package for the East area of study. Notice that the graphical representation of the neighbourhood network is also displayed.

From the explanation provided so far it is clear that the role played by the neighbours in the measurement of the spatial autocorrelation is pivotal. This evidence stresses once more that in the pattern analysis the elements count more in their whole and not as individuals; furthermore the link to the ideas of studying the landscape from a holistic perspective is reinforced.

After the neighbourhood network is defined and a standardised connectivity matrix is derived, all the conditions are present to measure the generalised Moran's I index of spatial autocorrelation which is stated in equation 1.

(1)

$$I(\mathbf{x}) = \frac{n\sum_{(2)} c_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{(2)} c_{ij} \sum_{i=1}^{n} (x_i - \bar{x})^2} \quad \textbf{and if} \quad [x_i - \bar{x}] = [z_i] \quad \textbf{then} \quad I(\mathbf{x}) = \frac{n\sum_{(2)} c_{ij} z_i z_j}{\sum_{(2)} c_{ij} \sum_{i=1}^{n} z_i^2}$$

where n is the number of spatial units indexed by i and j; x is the variable of interest; $\bar{x}$ is the mean of x; and c_{ij} is the connectivity matrix.

If c_{ij} equates to W it is transformed in a row-sum standardised weighting matrix as below

(2)

$$\mathbf{W} = \left[c_{ij} / \sum_{j=1}^{n} c_{ij} \right]$$

When c_{ij} is applied to the index, Lee (2001, cited in Dray et al., 2008) proposes a decomposition of Moran's *I* into two parts using the concept of spatial lag (Anselin, 1996 cited in Dray et al., 2008). The lag vector equates to the averages of the neighbours weighted by the connectivity matrix and it is computed as

(3)

$$\tilde{\mathbf{x}} = \mathbf{W}\mathbf{x} \text{ (i.e. } \tilde{x}_i = \sum_{j=1}^{n} w_{ij} x_j \text{)}$$

Thus the use of a row-sum standardised weighting matrix transforms equation (*1*) to

(4)

$$I(\mathbf{x}) = \frac{\sum_{(2)} c_{ij} z_i z_j}{\sum_{i=1}^{n} z_i^2} = \frac{\sum_{i=1}^{n} z_i \tilde{z}_i}{\sum_{i=1}^{n} z_i^2} = \frac{\mathbf{z}^t \tilde{\mathbf{z}}}{\mathbf{z}^t \mathbf{z}}$$

which indicates that Moran's *I* is reduced to a ratio of quadratic forms and provides the lag vector ($\tilde{z}=\tilde{x}=Wx$) as a smoother operator since it represents the weighted average of the neighbouring values.

Moran's *I* is based on the hypothesis that spatial autocorrelation affects the spatial structure of the data, in contrasts with the null hypothesis of no spatial autocorrelation which implies a random distribution of the data (Lee and Wong, 2001). In case of null hypothesis the expected value of Moran's *I* equates to:

(5)

$$E(I) = \frac{-1}{N-1}$$

With this in mind, the stage after calculating Moran's index is to verify whether or not the hypothesis of spatial autocorrelation is valid. However, since the simple difference between *observed* Moran's *I* and *expected* Moran's *I* does not inform if the calculated value of the index is statistically significant, it is necessary to carry out a tailored test for significance, usually performed through a Monte Carlo randomisation test. The number of permutations or randomisation is defined by the analyst and usually is 99 or 999 (Lee and Wong, 2001). The differences in the number of permutations are explained in terms of significance level, for instance if 999 permutations are used then the correspondent significance level is 0.001, whereas in case of 99 permutations, the significance level equates to 0.01 (Anselin, 2005). Therefore it is up to the analyst to determine the level of accuracy when performing a randomisation test; the smaller the value of the significant level (p-value) the higher the confidence that the data is statistically significant.

Figure 5.22 illustrates the graph commonly used to plot the results of the randomisation test. This graph is a simpler way of understanding how the randomisation test works.

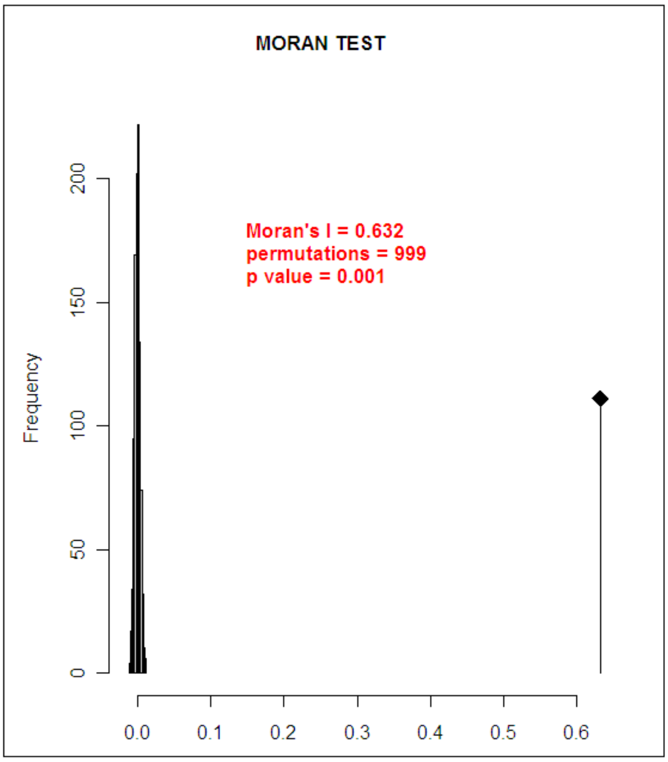


Figure 5.22. Example of graphic representation of the test based on 999 permutations carried out to test the statistical significance of the observed Moran's I.

The black lines on the left refer to the randomly simulated values while the single line, which represents the original data, is on the right. If this line is far apart from the black ones then the graph indicates that Moran's *I* is statistically significant, the spatial autocorrelation is confirmed and the null hypothesis can be rejected. In contrast, if the single line is within the range of the black ones, the data is randomly distributed and the null hypothesis has to be accepted.

As explained, Moran's *I* informs whether or not the spatial autocorrelation within pair of observations across the whole dataset is positive or negative. Nevertheless in order to know more about the strength of spatial autocorrelation in the relationship between neighbouring polygons it is necessary to refer to the Moran scatter plot (see figure 5.23), which was proposed by Anselin in 1996 (Dray et al, 2008).

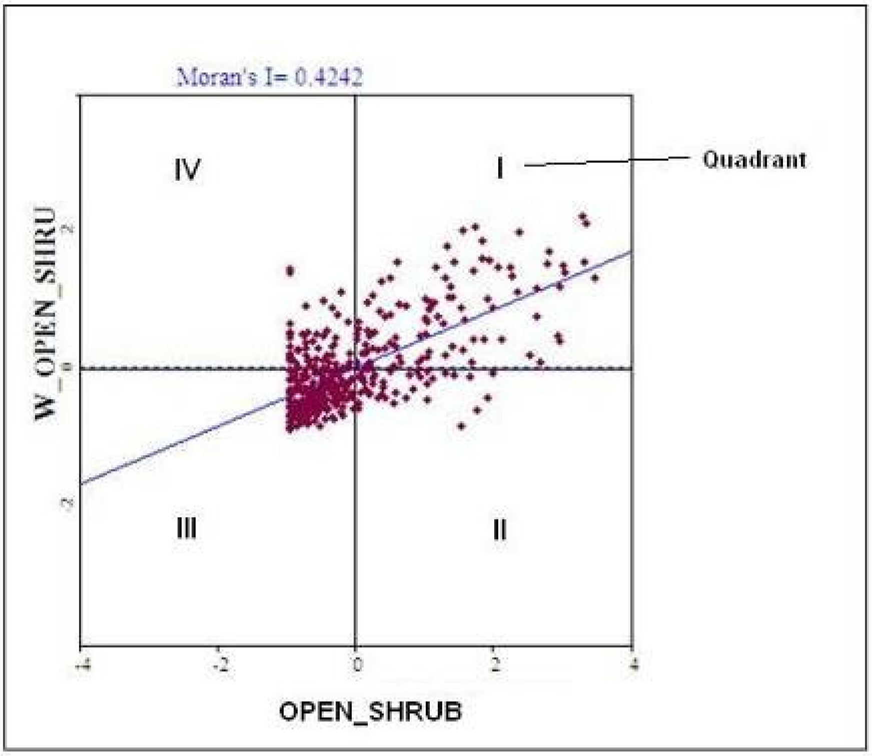


Figure 5.23. Example of a Moran scatter plot in a univariate analysis. The variable "open shrub" is compared to the lagged variable "W_open shrub". The strength of spatial autocorrelation can be read along the line which crosses the first and third quadrat. The points along this line indicate a strong positive spatial autocorrelation. The overall Moran's I index is calculates and equates to 0.42.

The graph determines the strength of spatial autocorrelation by plotting the original variable *n* against the spatial lag[21] *Wn* (the weighted average of the neighbouring values) of the variable *n*. The scatter plot shows four quadrants, the first and the third indicate positive spatial autocorrelation. In these quadrants, the variable and the lagged variable are similar and reveal a strong relationship. Data plotted in the second and fourth quadrants indicates a negative spatial autocorrelation and hence both variable and lagged variable are dissimilar and show a weak relationship.

The determination of the causes and effects of a spatial relationships is not in the interest of this research, nevertheless once again it must be remembered that Moran's *I* does not indicate what is causing the autocorrelation but it informs on the strength of spatial autocorrelation amongst the data and warns who investigates the spatial processes underlying patterns about the absence of independence amongst the observations.

Overall through the chapter it emerged that GIS need the support of statistics in order to conduct geographical analyses on complex issues (topics). Therefore, from now it is more appropriate to refer to a GIS/statistics-based methodology for landscape characterisation.

The next chapter is dedicated to the description the spatial statistics applied for the identification of the spatial patterns within an area of study, namely the character types of the landscape. The statistics perform multivariate analysis by taking into account spatial autocorrelation.

[21] A spatially lagged variable is an essential part of the analysis of spatial autocorrelation. It corresponds to the sum of spatial weights multiplied with the values for observations at neighbouring locations. The lagged variable is indicated with a "W" (Anselin, 2005).

CHAPTER 6

MULTISPATI: a spatially constrained multivariate analysis for landscape characterisation

Glen Strathfarrar and loch Beannacharan (I.Marengo).

6.1 A methodology for landscape characterisation

After the spatial connection between the Thiessen polygons was created the tabulate area tool[22] of ArcGIS could be run for each thematic map in order to quantify the landscape elements composition within each Thiessen polygon. The outcomes of the tabulate area tool were tables *n* x *p* where *n* are the rows indicating the polygons and *p* are the columns referring to the class of landscape elements (from now are called variables). The cells of the table contain the area (square meters) occupied by the elements in each polygon. Since the minimum mapping unit is 2500 square meters (50m x 50m cell size) it was easier to transform the surface area into the "count" of cell as figure 6.1 illustrates.

BROADLEAVE	CONIFERS	ARABLE
0	797500	0
42500	25000	582500
115000	0	1470000
277500	567500	157500
530000	1002500	430000
45000	812500	5000
22500	0	775000
15000	247500	1405000
0	0	0

→

BROADLEAVE	CONIFERS	ARABLE
0	319	0
17	10	233
46	0	588
111	227	63
212	401	172
18	325	2
9	0	310
6	99	562
0	0	0

Figure 6.1. In order to simplify the tables the landscape elements were quantified in terms of number of cells (table on the right) instead of area (m²) occupied (table on the left).

Afterwards, the tables were joined to the Thiessen polygon shape file so that for each polygon it was possible to obtain a summary of the landscape elements composition (see figure 6.2). The length of the table, on average 45-50 variables, was not a surprise; however it was interesting to see how the landscape elements occurred and distributed within the polygons. By querying the polygons as showed in figure 6.2 it was very difficult to define the way the landscape elements were associated to each other and it was almost impossible to try and describe the character of the landscape.

[22] The tabulate are tool is stored in the spatial analyst tool/zonal statistics (arc toolbox)

Indeed, the information was in the table but needed to be extracted with another statistical technique able to take into account the spatial autocorrelation amongst the polygons.

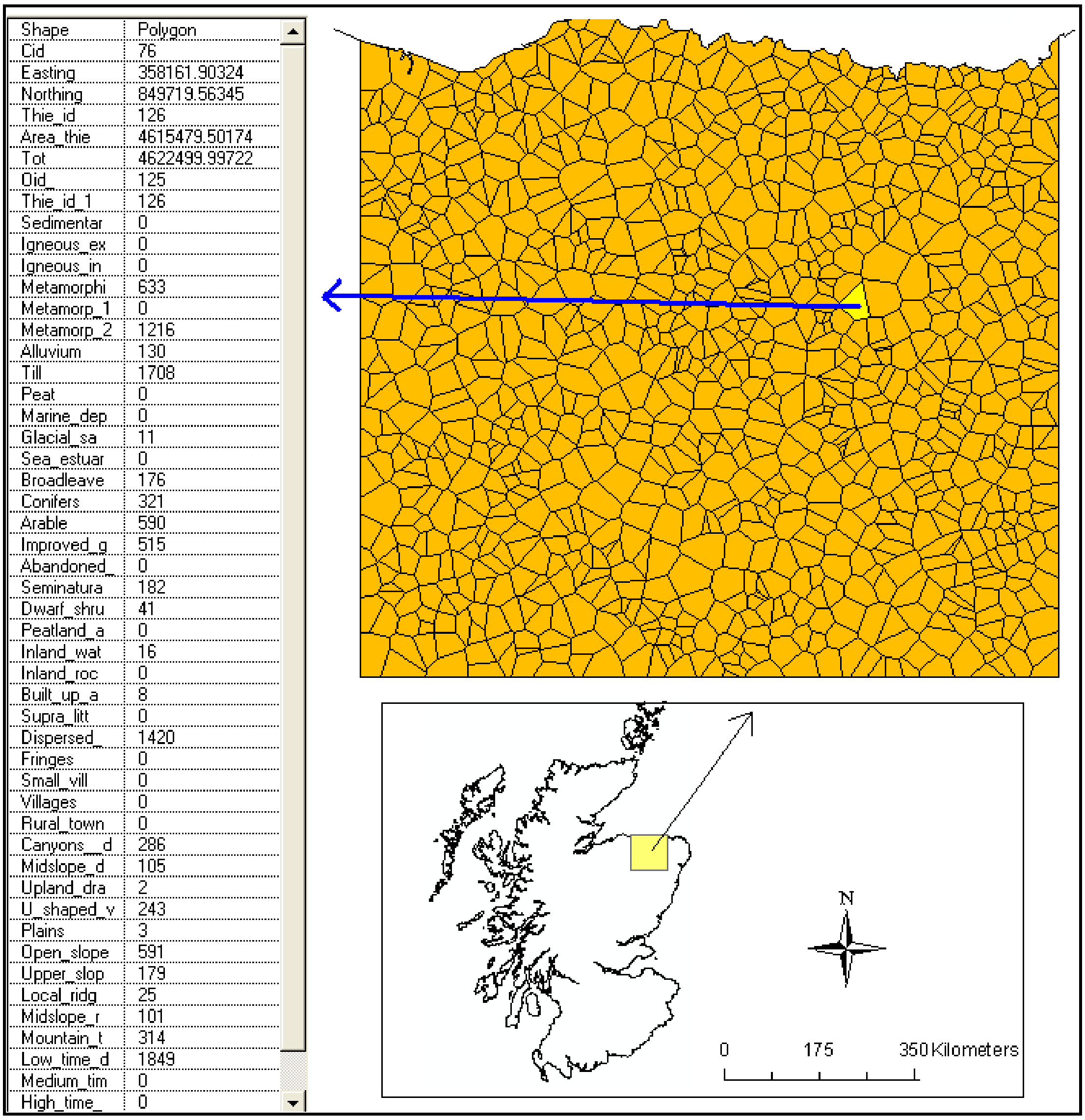


Figure 6.2. By selecting each polygon it is possible to derive the information about the associated landscape elements composition. Here the example of East area of study.

The information collected so far for the identification of the character of the landscape can be summarised as it follows:

- a row-sum standardised weighted matrix, obtained with the package *spdep,* ready to be used for the calculation of the generalised Moran's *I*;
- a table with the quantification of the landscape elements for each Thiessen polygons, retrieved with the tabulate area of ArcGIS 9.2.

The next phase was to choose the spatial statistics technique that could use both tables and define quantitatively and spatially those spatial structures within the data that would correspond to the character types of the landscape.

As described in chapter 5 the statistical program R was chosen to integrate the statistical tools provided by arcGIS 9.2 for several reasons; R it is free of charge, entirely accessible from the internet and identifies the latest and powerful support for statistical analyses. ArcGIS 9.2 offers a series of mapping and analytical tools but the level and complexity of statistics that it can perform is relatively low. The previous chapter described how the package *spdep* helped in the calculation of the neighbourhood network, in this chapter the focus is on other two packages, precisely *Ade4* and *adegenet,* which were chosen for this research due to their successful application of spatially constrained multivariate analysis (Drat et al., 2008, Saby et al, 2009, Jombart, 2009, Dray and Jombart, 2010).

MULTISPATI, which stands for multivariate spatially constrained analysis, belongs to *Ade4* and, after the adoption of the first law of geography, marks the second fundamental turning point of this research. In fact it allows the analysis of all the landscape elements at the same time and simultaneously it takes into account the effect of spatial autocorrelation.

With the three mentioned R packages, this research has finally got the spatial statistics able to identify the character of the landscape. On this basis the methodology of analysis was drawn and its steps are summarised in figure 6.3.

The first step consists in dividing the area of study into sampling zones, which for this research correspond to the randomly generated Thiessen polygons. In ArcGIS 9.2

these polygons are overlapped to the maps of landscape elements and through the tabulate area tool, the count of the classes of landscape elements within each polygon is performed.

From the same Thiessen polygons, in R the package *spdep* defines the neighbourhood network which allows the completion of the second step of the methodology and provides the row sum standardised matrix which is used for the calculation of the Moran's *I* index.

The third step is about the application of MULTISPATI which carries simultaneously a multivariate analysis (Principal Component Analysis or PCA) and spatial autocorrelation analysis by calculating Moran's *I* on the basis of the weighed matrix. The results from MULTISPATI are then processed in a cluster analysis intended to extract the spatial patterns in a way that they could be imported and visualised in arcGIS 9.2.

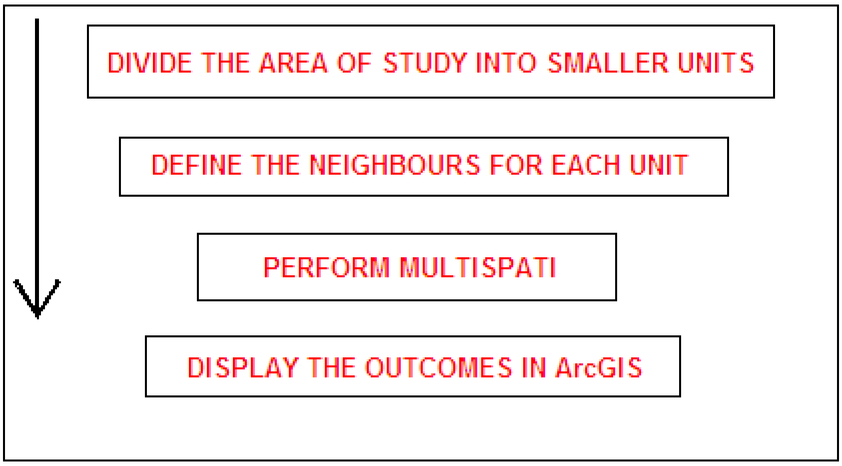


Figure 6.3. Main steps of the methodology of analysis for landscape characterisation based on GIS and MULTISPATI.

Thus, GIS and R worked in synergy and the rest of the chapter describes how the GIS/MULTISPATI-based methodology for landscape characterisation performs.

6.2 The approach with standard multivariate PCA analysis

Multivariate statistics is concerned about the analysis of more than one statistical variable and can be conducted through several models and different types of analysis. This current research wants to understand which variables (elements in the landscape) are associated in a unique way so that they contribute to the character of a landscape. From the tables retrieved by the tabulate area tool it emerged that some elements were either completely not recorded (the count was equal to zero) or they occurred rarely (the count was less than a third of the total possible cells). Therefore, the question was whether or not it was more effective to try and extract from all the variables analysed only those that contributed to explain significantly the landscape elements composition. On this basis it was decided to use the Principal Component Analysis (PCA) as statistical type of multivariate statistics. The main advantage of using PCA is that it transforms a number of possibly correlated variables in a smaller number of uncorrelated variables called principal components; hence PCA separates the variables according to the way they are associated and distributed in relation to each other.

PCA was calculated using the package *Ade4*, which allows the calculation of both a standard and a spatially constrained (MULTISPATI) PCA. The former was necessary in order to summarize the variability of the landscape elements. The latter was used to reveal spatial patterns due to the effects of spatial autocorrelation.

The PCA analysis in *Ade4* uses a data table as a statistical triplet (**X,Q,D**) where **X** is a *n x p* data table with *n* row (polygons) and *p* columns (variables); **Q** is a *p x p* positive symmetric matrix used to measure the relationships between the variables and **D** is a *n x n* symmetric matrix used to measure the differences between the polygons (Dray et al, 2008, Dray and Dufour 2007). The transformation of the data into a statistical triplet is a consequence of the application of the duality diagram (*dudi*) theory[23] to the data analysis. Two are the key principles of *dudi*:

1) computing the differences between observed sites and
2) identifying the associations between the variables

and both are related in that the associations between variables are used to explain the differences between observed sites. On this basis *dudi* allows the identification of spatial structures within the data and the detection of patterns.

To obtain optimal performance from PCA it is necessary to know more about the data and its characteristics, since it may be necessary to adjust it prior to analysis. The basic statistics and the test for normality were carried out first and the results showed a non normal distribution in the observed variables along with significant differences in variances and means. Some of them were very small and others were much larger. As a result, it was thought opportune to try and approximate the data distribution to a Gaussian line and to bring the mean and variance of the variables roughly to similar values. The last operation was particularly important because, as seen in chapter 5, one of the assumptions of spatial autocorrelation is stationarity, that occurs when the observed values have approximately similar means and variances across the whole area of study. A range of methods can be used to transform the data and for this study, the mean centring and the logarithmic transformation were chosen in order to achieve the condition of stationarity assumed by Moran's *I*.

Mean centring consists of calculating the mean for each variable then subtracting it from the variable values for each observation. As a result the recalculated mean equates to zero. Centring adjusts for differences in the offset between high and low

[23] The duality diagram theory belongs to the French School of Analyse des données that stands for data analysis. For further information, refer to Dray and Dufour (2007) who provide a more detailed description of the duality diagram theory.

values assumed by the variables and it investigates variance from the mean data rather than the data as a whole (Van den Berg et al, 2006).

The second type of transformation focuses on the variance from the mean and not from the whole data. The logarithmic transformation is potentially appropriate because it works efficiently for data that are positively skewed (like that in this research) and whose distribution requires to be more symmetrical. The transformation converts the original data in logarithms by compressing the data and reducing the variation in the mean and variance values; precisely the large values in the dataset are reduced more than the small values (Van den Berg et al, 2006). The results from the data transformation are illustrated in figure 6.4.

ORIGINAL DATASET

Variable	Mean	Variance
SEDIMENTAR	116.46	86998.73
IGNEOUS_EX	1.613	499.120
IGNEOUS_IN	143.0	119949.6
METAMORPHI	39.61	18772.76
METAMORP_1	0.421	96.160
METAMORP_2	589.6	271302.9
ALLUVIUM	42.59	14454.39
TILL	664.8	195897.3
PEAT	26.13	18428.52
MARINE_DEP	3.37	1594.55
GLACIAL_SA	57.18	25100.48
SMALL_HILL	89.90	7728.66
GLENS_VALL	128.70	16848.36
HILL_TOPS_	53.84	4763.88
OPEN_SLOPE	533.0	106190.4
STRATHS_PL	65.06	16288.29
SEA_ESTUAR	0.01500	0.06284
BROADLEAVE	48.59	4169.15
CONIFERS	126.72	49425.33
ARABLE	332.97	93298.43
IMPROVED_G	206.19	33047.25
ABANDONED_	31.00	4714.71
SEMINATURA	37.64	3481.03
DWARF_SHRU	93.09	42038.54
PEATLAND_A	6.25	1753.61
INLAND_WAT	0.781	15.378
INLAND_ROC	4.725	360.546
BUILT_UP_A	12.03	2003.26
SUPRA_LITT	0.425	15.992
DISPERSED_	476.3	125213.9
FRINGES	6.182	888.293
SMALL_VILL	5.442	885.026
VILLAGES	15.09	7449.92
RURAL_TOWN	17.90	10430.70

LOG_TRANSFORMATION

Variable	Mean	Variance
SEDIMENTAR	1.2196	6.0258
IGNEOUS_EX	0.0441	0.2121
IGNEOUS_IN	1.4803	6.7295
METAMORPHI	0.7045	3.3250
METAMORP_1	0.01528	0.06628
METAMORP_2	5.0416	7.0604
ALLUVIUM	0.9717	4.0453
TILL	5.9999	2.4365
PEAT	0.4140	2.0915
MARINE_DEP	0.0639	0.3447
GLACIAL_SA	1.1018	4.6400
SMALL_HILL	3.9043	1.7956
GLENS_VALL	4.1847	2.0889
HILL_TOPS_	2.9609	3.1075
OPEN_SLOPE	6.0462	0.6362
STRATHS_PL	2.8856	3.3982
SEA_ESTUAR	0.00610	0.00952
BROADLEAVE	2.8829	2.9769
CONIFERS	3.0501	5.2028
ARABLE	4.7591	4.7868
IMPROVED_G	4.5959	2.9651
ABANDONED_	1.7376	3.6658
SEMINATURA	2.3860	3.3618
DWARF_SHRU	2.7701	4.1675
PEATLAND_A	0.2512	0.9852
INLAND_WAT	0.1517	0.3581
INLAND_ROC	0.4667	1.2717
BUILT_UP_A	1.0352	2.0925
SUPRA_LITT	0.0643	0.1728
DISPERSED_	5.5423	2.8761
FRINGES	0.2598	1.1109
SMALL_VILL	0.2489	1.0018
VILLAGES	0.2863	1.4295
RURAL_TOWN	0.2705	1.4638

MEAN CENTRING

Variable	Mean	Variance
SEDIMENTAR	-0.0000	6.0258
IGNEOUS_EX	0.0000	0.2121
IGNEOUS_IN	0.0000	6.7295
METAMORPHI	0.0000	3.3250
METAMORP_1	0.00000	0.06628
METAMORP_2	-0.0000	7.0604
ALLUVIUM	0.0000	4.0453
TILL	0.0000	2.4365
PEAT	-0.0000	2.0915
MARINE_DEP	0.0000	0.3447
GLACIAL_SA	0.0000	4.6400
SMALL_HILL	0.0000	1.7956
GLENS_VALL	-0.0000	2.0889
HILL_TOPS_	0.0000	3.1075
OPEN_SLOPE	0.0000	0.6362
STRATHS_PL	-0.0000	3.3982
SEA_ESTUAR	0.00000	0.00952
BROADLEAVE	-0.0000	2.9769
CONIFERS	0.0000	5.2028
ARABLE	0.0000	4.7868
IMPROVED_G	0.0000	2.9651
ABANDONED_	0.0000	3.6658
SEMINATURA	-0.0000	3.3618
DWARF_SHRU	0.0000	4.1675
PEATLAND_A	0.0000	0.9852
INLAND_WAT	0.0000	0.3581
INLAND_ROC	0.0000	1.2717
BUILT_UP_A	-0.0000	2.0925
SUPRA_LITT	0.0000	0.1728
DISPERSED_	0.0000	2.8761
FRINGES	0.0000	1.1109
SMALL_VILL	-0.0000	1.0018
VILLAGES	0.0000	1.4295
RURAL_TOWN	-0.0000	1.4638

Figure 6.4. Summary of the values of mean and variance assumed by the data of “East” area of study before and after mean centrings and logarithmic transformation. The differences from the original to the transformed data stand out clearly.

The transformation in general succeeded and the data could be processed in *dudi.pca*[24] which creates a new smaller set of variables that captures as much variance in the original dataset as possible. These new variables, called components, are sorted into descending order in a plot illustrated in figure 6.5.

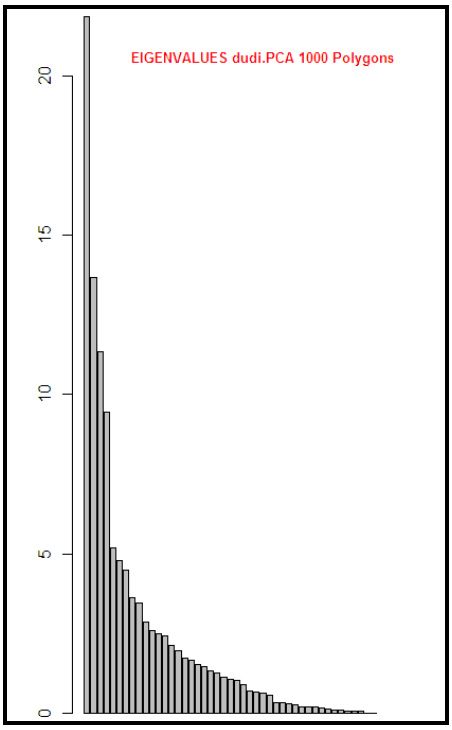


Figure 6.5. Bar plot (or scree plot) obtained by performing the standard PCA on the transformed data of “East” area of study. The vertical axis shows the eigenvalues, while the bars represent the components, or uncorrelated variables which summarise the variance of the original data.

The y axis of the bar plot in figure 6.15 refers to the total variance of the original data explained by each component and this is called eigenvalues. The components, which are the variables in the dataset, are represented by the bars. For instance here the first five eigenvalues equates to 21.85, 13.66, 11.34, 9.44 and 5.20. An abrupt decrease in eigenvalues in general indicates the boundary between true patterns and not interpretable structures.

[24] *dudi.pca* is the code used in R to perform a PCA on the basis of the duality diagram theory.

Commonly, the decision about the number of components to keep for the rest of the analysis is made on the basis of the results from the bar plot and the calculated values of the inertia (variance) and cumulative inertia, which are the proportion/percentage of the total variance accounted by each component. Figure 6.6 shows how the components are sorted into descending order according to the amount of inertia that they account for in the original dataset. Cumulatively, all the variables account for 100% of the variation and each of them accounts as much of the remaining total variance of the original data as possible.

```
> (variance <- 100 * eastdata1_pca$eig/sum(eastdata1_pca$eig))
 [1] 19.757725612 12.352229179 10.254656834  8.539341123  4.707001915
 [6]  4.345420198  4.075351874  3.278395003  3.135300907  2.594840482
[11]  2.361542705  2.268409443  2.201254304  1.934621352  1.787343349
[16]  1.552282511  1.491764796  1.383055582  1.324305483  1.204779546
[21]  1.149566089  1.019289747  0.965907464  0.923380927  0.804425588
[26]  0.645249259  0.596161408  0.581772393  0.523562397  0.294998334
[31]  0.288555821  0.264630915  0.239760448  0.192554378  0.176148519
[36]  0.169611586  0.154569883  0.119098396  0.091994083  0.085327977
[41]  0.054803579  0.051729062  0.049637889  0.006467375  0.001174282
> cumsum(variance)
 [1]  19.75773  32.10995  42.36461  50.90395  55.61095  59.95637  64.03173
 [8]  67.31012  70.44542  73.04026  75.40181  77.67022  79.87147  81.80609
[15]  83.59343  85.14572  86.63748  88.02054  89.34484  90.54962  91.69919
[22]  92.71848  93.68439  94.60777  95.41219  96.05744  96.65360  97.23538
[29]  97.75894  98.05394  98.34249  98.60712  98.84688  99.03944  99.21559
[36]  99.38520  99.53977  99.65887  99.75086  99.83619  99.89099  99.94272
[43]  99.99236  99.99883 100.00000
```

Figure 6.6. Inertia and cumulative inertia expressed by each component. The percentage values decrease from the first to the last component indicating that the relevance of each component in the overall explanation of the data decreases. The cumulative inertia shows that the first 4 components account for the 51% of the total variance of the original data and so should be chosen for interpreting the structure of the data.

The first and second components explain 19.7% and 12.3% respectively of the variance of the original dataset while cumulatively both express the 32.10% of the variance. Fundamentally the main expectation from conducting PCA is that the correlations among original variables are large enough so that the first few principal components account for most of the variance and therefore are able to explain the data. If this happens, then it is possible to derive a clear structure of the relationships amongst the data with a reduced number of variables and without losing information. In the example of East, on the basis of the eigenvalues, inertia and cumulative inertia, it was decided to retain the first four components.

An interesting representation of the variables and the way they occur in each component is given by the loadings plot which is derived by using the package *adegenet*. The loadings represent the relative weight (or importance) of each variable in each component rescaled with the amount of variance expressed by the component. For example figure 6.7 depicts the loadings of the first component which accounts for 19.7% of the variance of the original variables; the plot revealed that 9 variables amongst the 44 analysed stood out above the threshold of the 3rd quantile[25] and they were indicative as clearly relevant variables for the first component. In other words the representation of the loadings gives further clues about the composition of the components and it is a way of gaining an insight into the most representative variables (landscape elements) that contribute to the character of the landscape.

The investigation of the loadings was carried out for all the 4 components retained in the analysis and the variables that were the most likely to contribute to the character of the landscape of the study area were:

- Geology bedrock: metamorphic dalradian (metamorph_2 in the plot);
- Landforms: straths, hill tops;
- Landcover: arable, conifers, improved grassland, dwarf shrubs;
- Settlement types: dispersed houses;
- Historic period/time depth: prehistoric period, 17th-19th and 20th- 21st centuries

[25] The threshold of the 3rd quantile is set by default in *adegenet*, however it can be changed. Here the decision was to keep it since it allows the variables that contributed most to the spatial structure within the data to be identified with sufficient accuracy.

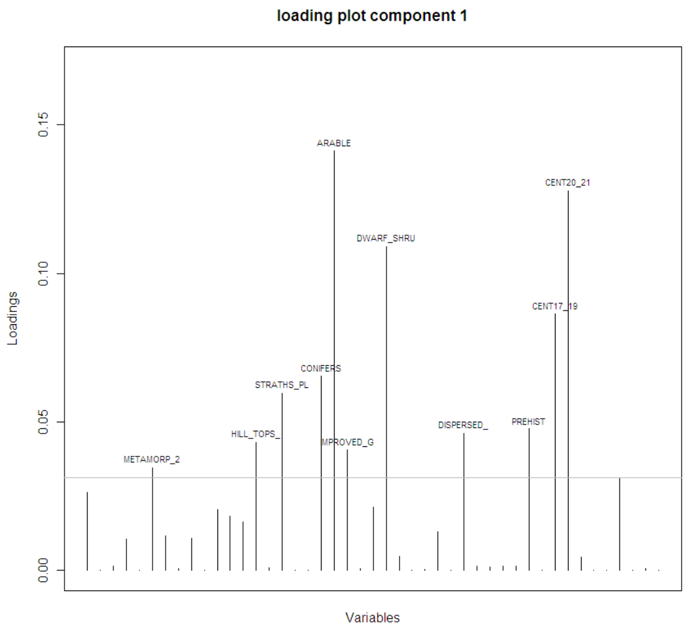


Figure 6.7. The loading plot allows the visualisation of the contribution of each variable, expressed as squared loadings, within each principal component. In this research the analysis of the loadings is carried out only for the components that are retained and considered relevant to the identification of patterns in the data. The horizontal line in the graph in light grey is the 3rd quantile threshold.

The biplot is another helpful plot to understand the operation carried out by PCA because it illustrates the distribution of the variables along the first two PCA axes (see figure 6.8). The variables are represented through vectors, whose length is an indication of the influence of the variable: e.g. short vectors represent variables with low importance. Moreover, if the vectors are close to each other and point towards the same direction, the variables are positively correlated to each other, whereas the variables show a negative correlation when their vectors diverge to opposite points in the plot.

In the example of East, the first axis clearly separates areas characterised by arable and farming fields, dispersed houses, broadleaves and improved grassland, from areas identified by moorland, plantations, dwarf shrubs, hills and glens. In contrast, the second axis provides a neat differentiation only on the basis of geological bedrock. In fact areas with sedimentary and igneous rocks are well distinct from those characterised by metamorphic dalradian rocks. However, the second axis highlights an interesting detail; the sedimentary and igneous rocks are typical of the flat areas where a deposit of gravel and sand has accumulated (notice in the plot the positive correlation amongst these variables), while the dalradian metamorphic rocks are positively correlated to a more undulating terrain, represented by open slopes, hill tops and glens.

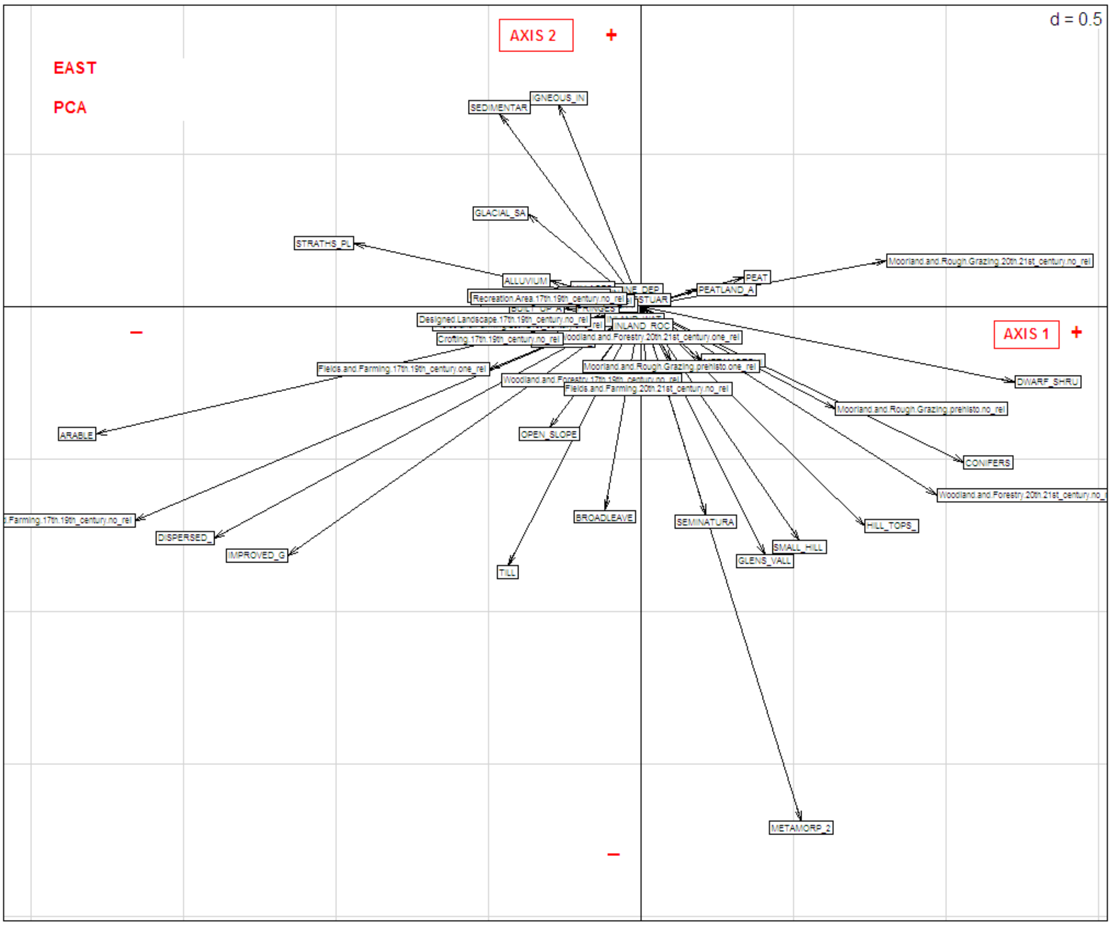


Figure 6.8. Principal component analysis (PCA) biplot of the 49 variables representing the landscape elements of East, are of study. The variables are distributed along the first two axes of PCA according to their degree of relevance (length of the lines) and their degree of correlation (positive if the lines are pointing in the same direction and closer to each other, whereas negative if the lines diverge).

The information conveyed by the biplot assumes great relevance because it helps to understand and interpret correctly the map of scores (see figure 6.9). The scores correspond to a matrix in which the columns are the principal axes of PCA and the rows are the individual data records (the polygons). The value in the cells of the matrix is the linear function of the original variables for which the variance has been maximised by PCA.

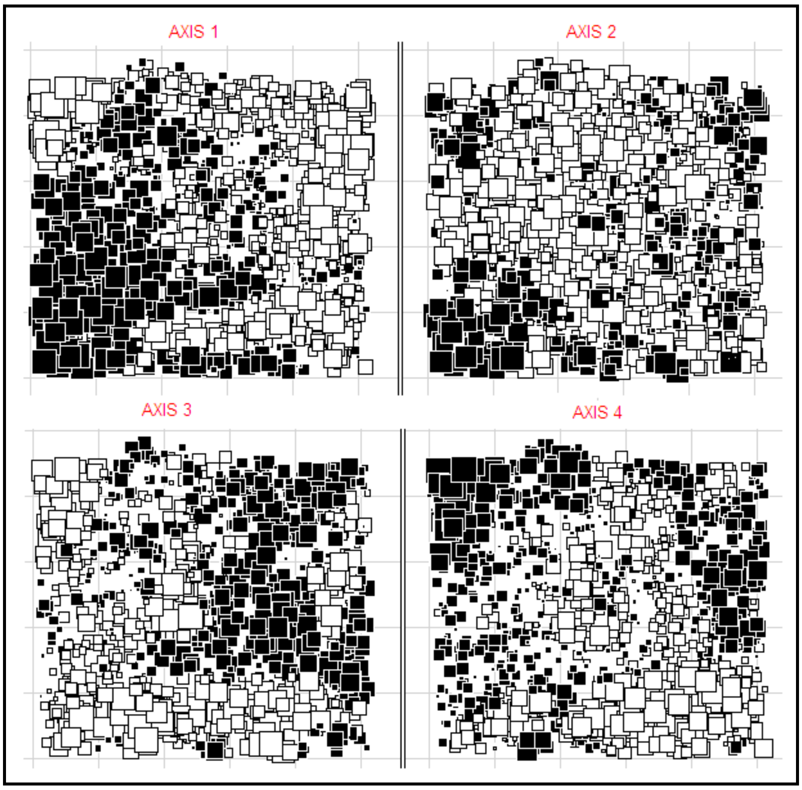


Figure 6.9. Geographical representation of the structures detected with the standard PCA analysis.

The squares indicate the scores (that on the input map are the polygons) and their size depends on the value of score. Instead the colour, black and white, marks the positive or negative sign of the value of the scores. The maps, still calculated in R, allow the analyst to have a first idea of where patterns of landscape elements occur in the study area.

It is worth noticing that these patterns reflect the distribution of the variables along the PCA axes (see figure 6.18). As far as the example of East is concerned, in the map of scores of the first PCA axis it was possible to identify the areas with arable and farming fields (the white squares) against the areas with moorland and conifers (the black squares). A less clear pattern could be retrieved from the map of scores of the second PCA axis, but it was not a surprise since the axis separated the areas only on the basis of the geological bedrock. Sedimentary and igneous rocks correspond to the black squares, while dalradian rocks are depicted with the white squares. Similar analysis can be carried out to interpret the remaining maps.

The main conclusion was that PCA was able to identify spatial patterns in the data but only by taking into account their variability. In fact the components retained for the analysis were chosen on the basis of their maximised variance and not because of the influence of the geographic location. A standard PCA does not carry any measure of spatial autocorrelation in its calculation; nevertheless the maps do suggest the presence of a spatial pattern. As a consequence Moran's *I* was measured in order to help to understand whether or not the structure of the data highlighted in the maps of scores could be driven by spatial autocorrelation. A high value of Moran's *I* would suggest evidence of spatial autocorrelation in the organisation of the data. The results of Moran's *I* calculated for the 4 axes were then tested for statistical significance with a Monte Carlo randomisation test and the positive results are illustrated in figure 6.10.

Figure 6.10

The higher values of Moran's *I* in axis 1, 3 and 4 compared to that in axis 2 confirmed what the maps of the scores depicted, and namely that the patterns in axis 2 were less specified and more difficult to interpret.

A standard PCA, like *dudi.pca,* detects structures in the data that are associated with the strongest variance, however, if the aim is to highlight the influence of the spatial location as contributor of spatial patterns within the data then it is appropriate to use MULTISPATI-PCA because it integrates Moran's *I* in the multivariate analysis. Hence, the next stage in the analysis is to apply MULTISPATI-PCA for the identification of spatial patterns on the basis of the influence of spatial autocorrelation.

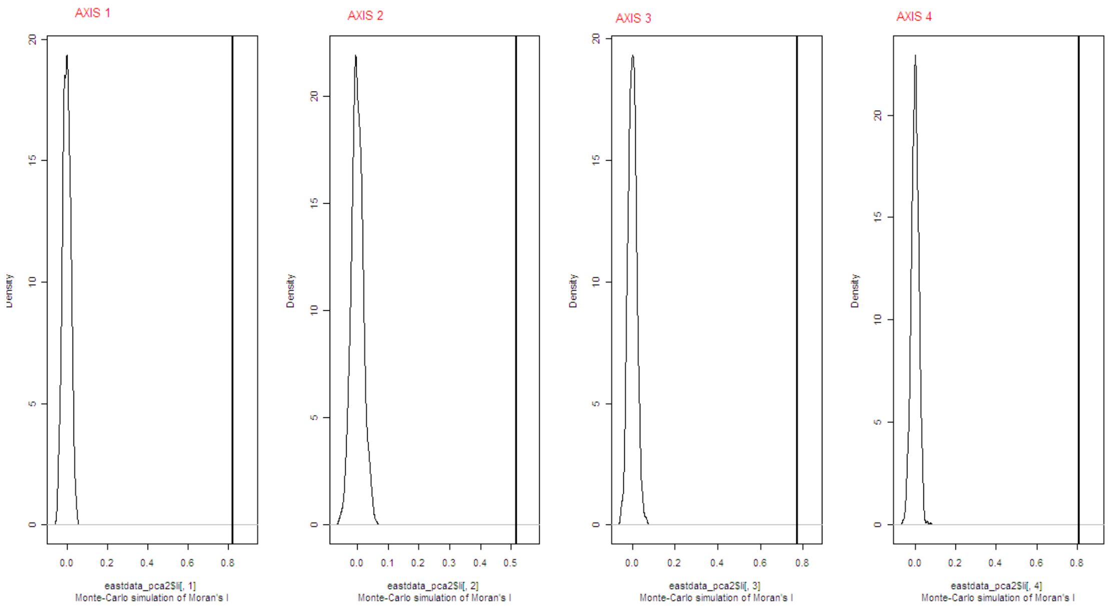


Figure 6.10. Moran's *I* calculated for each component in order to verify if the strength of the spatial autocorrelation in the data. In this case all the components show a high value of spatial autocorrelation which suggests that the patterns detected by standard PCA could be due to spatial autocorrelation.

The outcomes from MULTISPATI-PCA that are expected to change are the maps of scores (above all that one referring to the second axis) which should show clearer and more defined spatial structures. In this way MULTISPATI-PCA would demonstrate that spatial location affects the way variables associate with each other and that it enables a better identification of the spatial patterns.

6.3 The approach with MULTISPATI-PCA

MULTISPATI-PCA generalises the measure of autocorrelation of a variable through the application of Moran's *I* in the context of a multivariate analysis. In other words, it determines the relationships amongst many variables and their spatial structures by introducing the row-sum standardised weighted matrix W in the statistical triplet X,Q,D which was previously calculated in *dudi.pca* (Dray et al, 2008).

The calculations in the spatially constrained PCA proceeded exactly as in *dudi.pca* with the fundamental difference that the spatial weighted matrix was introduced in the identification of the principal components see figure 6.11.

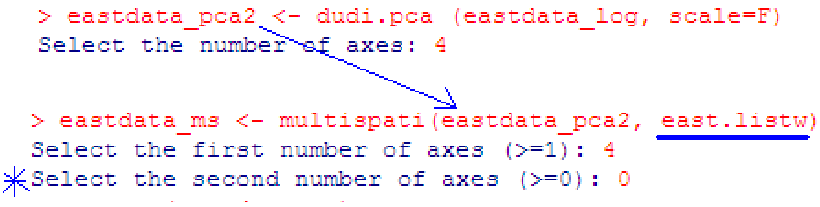


Figure 6.11. The difference in the calculation of the standard and spatially constrained PCA are visible in the code used in R. The first red row indicates the way of calculating standard PCA, which is based on the data transformed (eastdata_log). While the second red row shows how MULTISPATI integrates the spatial weight matrix (east.listw) to the analysis of the principal components (eastdata2_pca). In MULTISPATI it is required to select the first and second number of axes: the second number of axis refers to the negative eigenvalues which MULTISPATI can calculate since spatial autocorrelation can also assume negative values. (Jombart, 2009)

Similarly to *dudi.pca* it was necessary to select the most relevant components by looking at the scree plot of the eigenvalues and again the first 4 axes seemed the most informative (see figure 6.12).

Figure 6.12

Unlike standard multivariate analysis, the eigenvalues of MULTISPATI are composite because they measure both variance and spatial autocorrelation (Jombart, 2009). The first differences from *dudi.pca* could be read in the summary of the calculations performed in MULTISPATI and illustrated in figure 6.13.

Figure 6.13

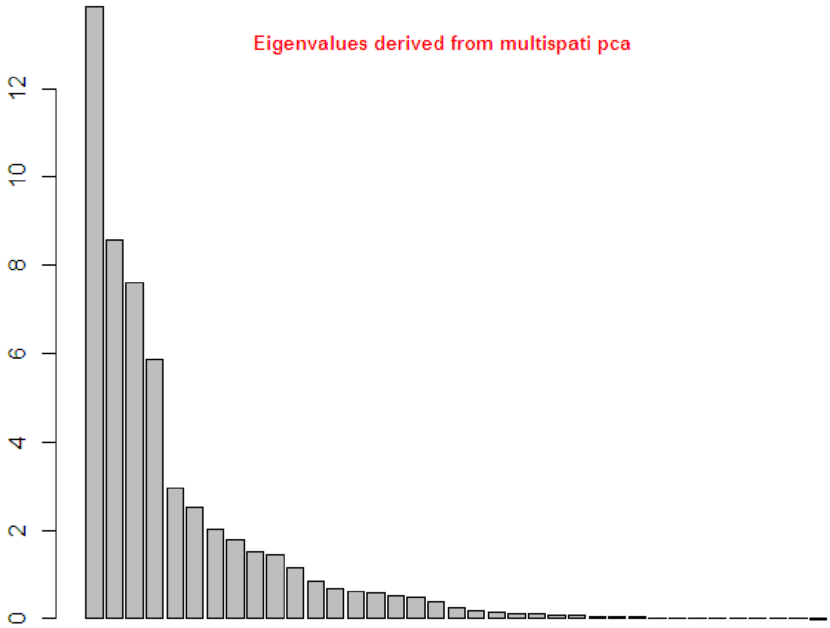


Figure 6.12. Scree plot derived after calculating MULTISPATI-PCA. Contrary to the standard PCA the eigenvalues indicate both variance and spatial autocorrelation. Information on both these aspects can be extracted by decomposing the eigenvalues by using a different graph (Jombart, 2009).

```
> summary(eastdata_ms)

Multivariate Spatial Analysis
Call: multispati(dudi = eastdata_pca2, listw = east.listw)

Scores from the initial duality diagramm:
         var      cum     ratio     moran
RS1 16.64777 16.64777 0.1761251 0.8246155
RS2 12.26318 28.91095 0.3058634 0.5161139
RS3 10.35211 39.26307 0.4153836 0.7733675
RS4  9.18970 48.45277 0.5126060 0.8115687

Multispati eigenvalues decomposition:
          eig       var     moran
CS1 13.850486 16.544464 0.8371674
CS2  8.564600 10.677438 0.8021213
CS3  7.618240  9.328297 0.8166807
CS4  5.876511 11.263240 0.5217425
```

Figure 6.13. The summary of the calculations performed in MULTISPATI is highly informative. The comparison between the values of variance and Moran's *I* obtained with *dudi.pca* and MULTISPATI helps to understand the influence of space as driving factor in the presence and distribution of spatial patterns. "Var" (variance) decreases and Moran's I increases as indication of the contribution of spatial autocorrelation to the formation of patterns in the original data.

The summary is relevant and informative because it explains what MULTISPATI-PCA achieves. Two are the key information to understand how MULTISPATI works: the variance (“var”) and Moran’s *I*. All the scores in MULTISPATI maximise the spatial autocorrelation between the sites, whereas those in *dudi.pca* maximise the variance, hence the analyst expects a loss in variance (“var”) between the compared values of “var” in *dudi.pca* and in MULTISPATI. Here the loss was equal to 16.64 versus 16.54 for component 1; 12.26 versus 10.6 for component 2 and 10.35 versus 7.6 for component 3.

Operationally MULTISPATI extends the concept of lag vector, that it was mentioned in chapter 5, to construct a lag matrix **WX**. The original data table **X**, which corresponds to the triplet (**X,Q,D**) is fully matched with the lag matrix **W**, so that the two tables contain the same variables for the same sites. MULTISPATI analyses this pair of tables by a coinertia analysis that maximises the scalar product between the linear combination of the original variables (here the principal components calculated in PCA) and the linear combination of lagged variables. The result is a linear combination of variables which maximises the product of the autocorrelation (a generalised version of Moran’s) by the variance calculated with a standard PCA (Dray et al, 2008, Ollier, 2005).

The maximisation of the spatial autocorrelation corresponds at the same time to a loss in the value of variance and to a gain in the value of Moran. In the example in figure 6.23 the increment of the index was consistent for the components 2 and 3, where Moran rose from 0.51 to 0.80 and from 0.77 to 0.81 respectively.

A plot that can help to clarify the correlation between *dudi.pca* and the MULTISPATI-PCA components is the correlation circle which plots the principal components selected in *dudi.pca* against the first two components of MULTISPATI. If the correlation circle records a null or weak correlation between the *dudi.pca* and the MULTISPATI-PCA components, there is indication that a change occurred in the spatial structure of the data (see figure 6.14).

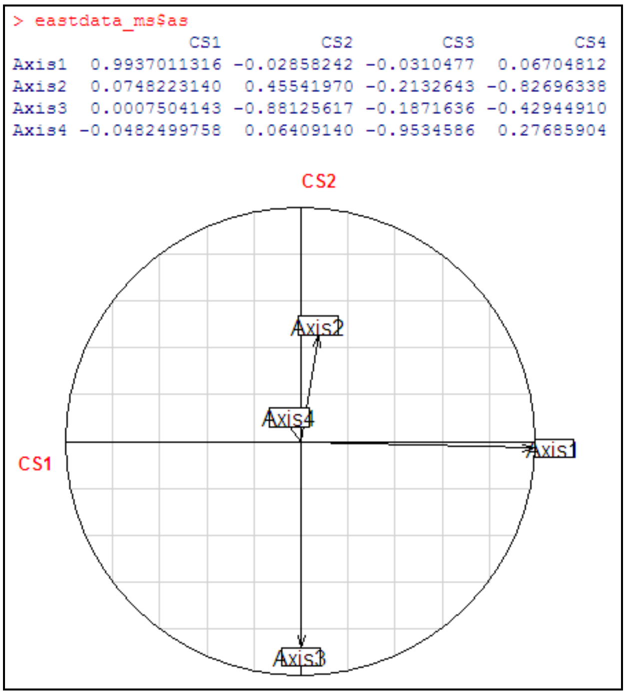


Figure 6.14. The analysis of the projection of the first axes selected in the standard PCA onto the two first components of MULTISPATI-PCA is carried out in order to understand whether or not relationships between the two patterns identified by the two analyses are registered.

In the example of “East”, figure 6.14 reveals that axis 1 of *dudi.pca* is strongly correlated to first component of MULTISPATI (CS1 = 0.99), while axis 2, 3 and 4 of *dudi.pca* show very weak correlation to its correspondent MULTISPATI components (CS2 = 0.45; CS3 = -0.18; CS4 = 0.27). The differences between the last three PCA axes and the MULTISPATI-PCA components can be retraced in MULTISPATI map of scores which should depict different spatial patterns.

In contrast to standard PCA, MULTISPATI plots both scores and lagged scores. The last ones are added because tend to project a clearer picture of the spatial distribution of the data. In fact the lagged scores are the sum of spatial weights multiplied with values for observation at neighbouring locations (Anselin, 2003) and act as a smoothing operator (see figure 6.15).

The map of scores for axis 1 showed similarities with the map obtained with a standard PCA (see previous figure 6.9) and it confirmed what already illustrated by the correlation circle. In contrast, the map of scores of the other axes revealed the most noticeable difference, if compared to its peer in figure 6.9. The MULTISPATI-PCA maps of scores/lagged scores depict more distinct and recognisable spatial patterns than the PCA maps, and the reason is because of the gain in the values of Moran's *I* index.

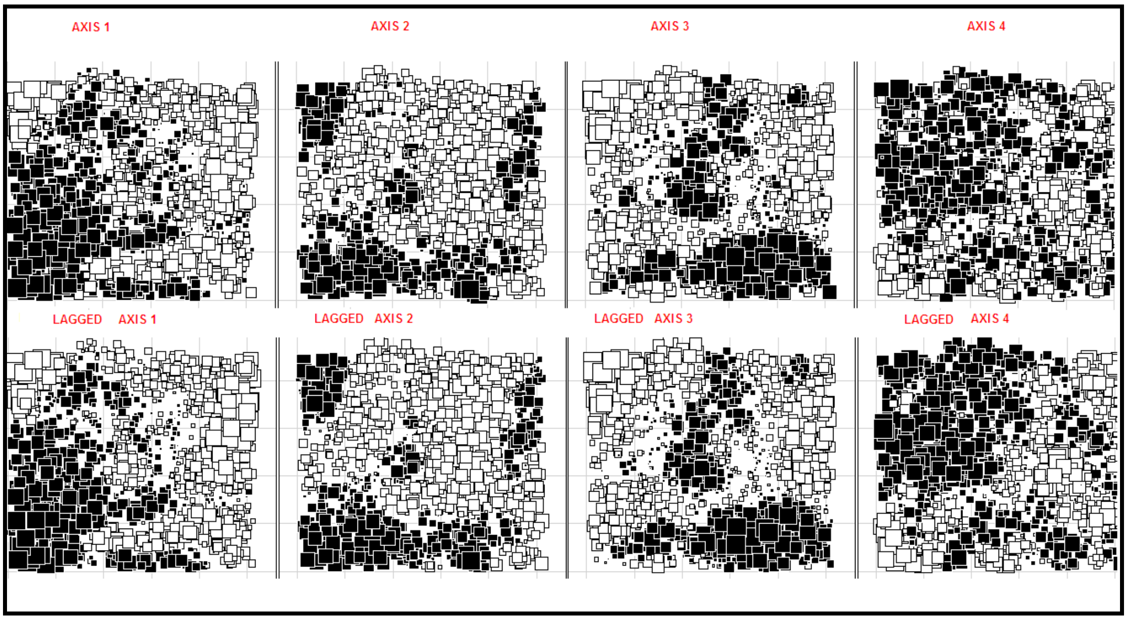


Figure 6.15. The maps of scores and lagged scores plotted for the four principal axes. There are cases, for instance the fourth axis, when the maps of the lagged variables display the patterns in a clearer way. This is due to the fact that the lagged variables are the average of the neighbours' values, thus they tend to smooth the patterns.

Similarly to *dudi.pca* it is possible to retrieve information on the variables and the way they contribute to the spatial patterns by looking at the biplot (figure 6.16).

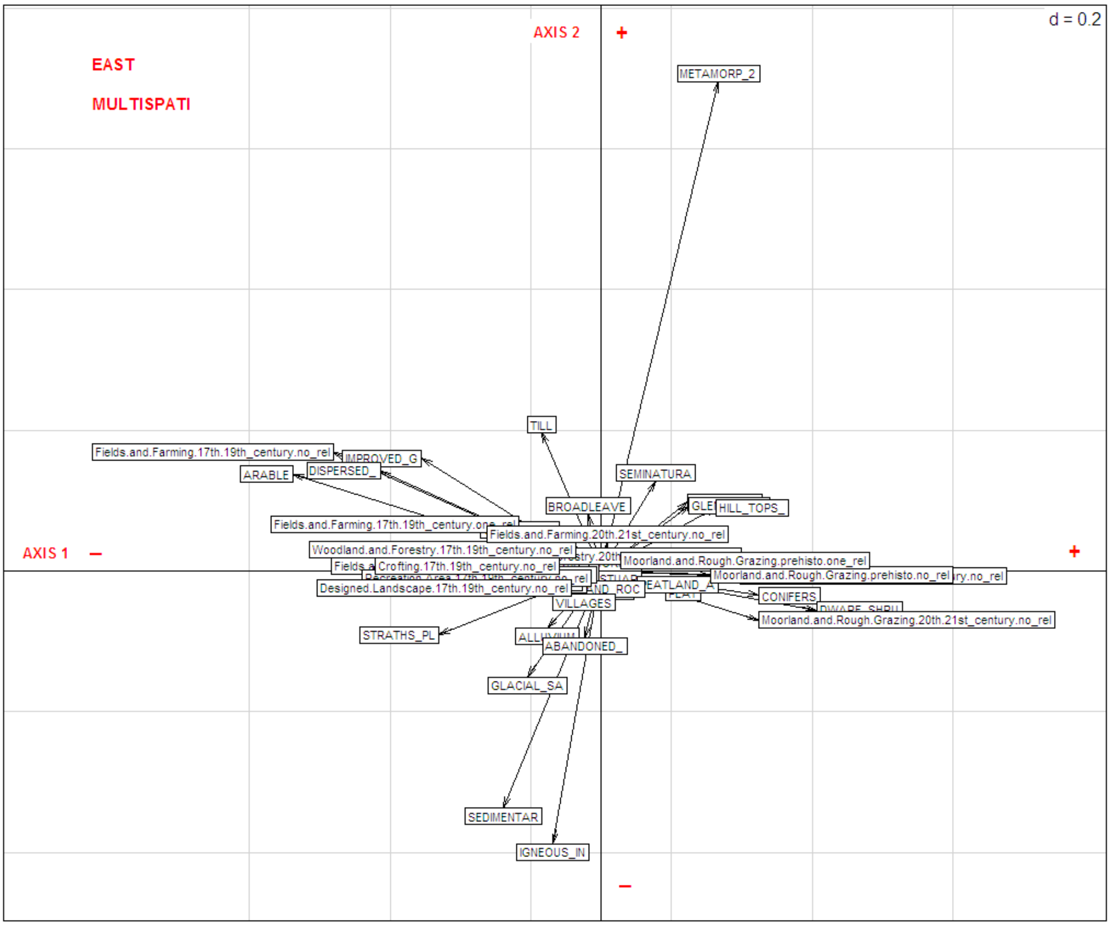


Figure 6.16. MULTISPATI-PCA biplot of East area of study. This plot can be compared to that one in figure 6.18 in order to understand how the spatial patterns are influenced by spatial autocorrelation

Along the first axis the similarity with the biplot obtained by running PCA is evident; whereas differences in the location and composition of the variables along the second axis are more noticeable. The importance of metamorphic dalradian rocks is more pronounced (longer line than in PCA biplot) at indication of an increased influence in the identification of spatial patterns. Metamorphic dalradian rocks are opposite to sedimentary and igneous rocks which are correlated to flat areas comprised of glacial sand and gravel, alluvium and abandoned grassland. The last two variables are new if compared to the biplot of PCA and contribute to form the spatial pattern that on the map of scores is represented by white squares (see figure 6.15, map of scores, axis 2).

Finally, as with the standard PCA, the plot of the loadings can be used as a tool to understand the variables that explain the highest variability in each component. The variables that appear in figure 6.17 are the most recurrent and possibly dominant in the landscape.

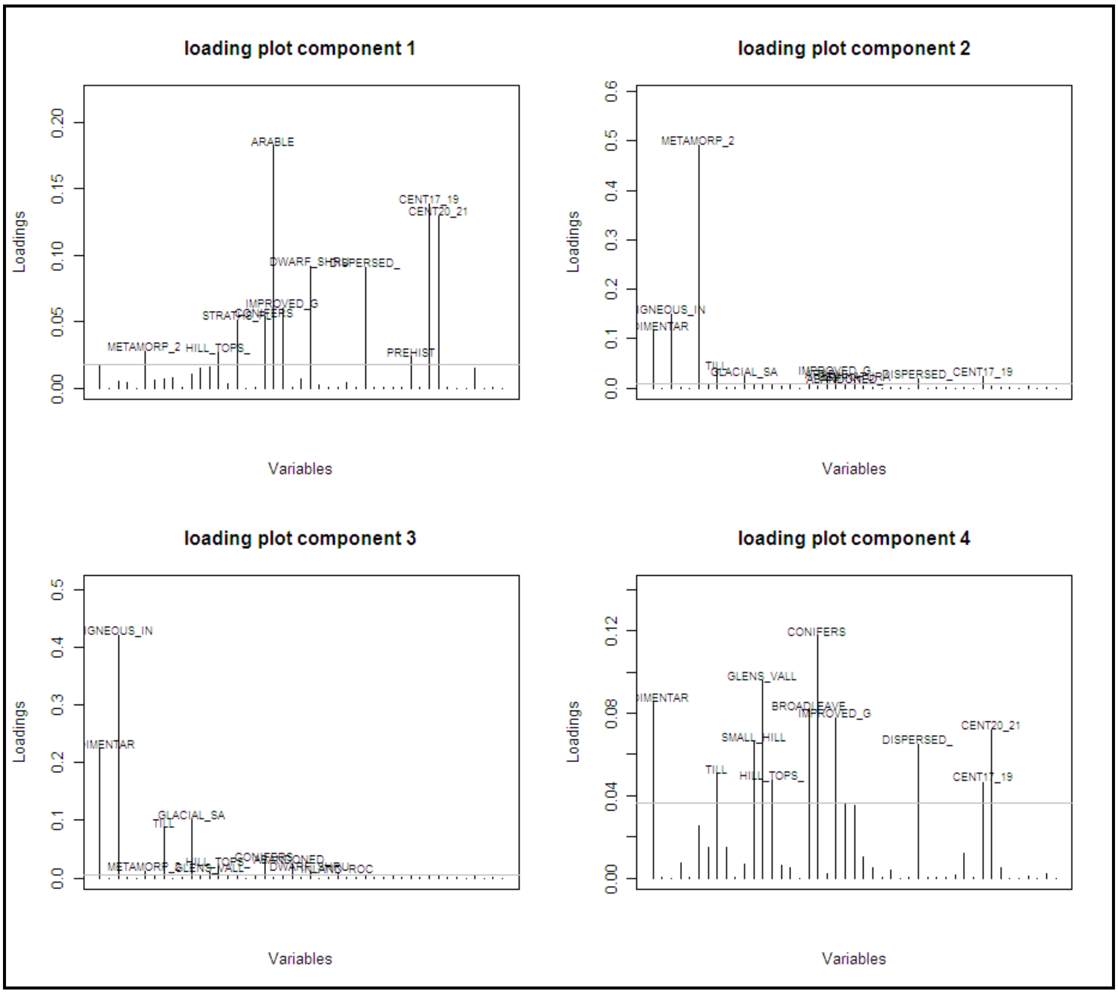


Figure 6.17. Loading plots for the first four components retained during MULTISPATI-PCA analysis. It is interesting to see the different ways the variables account for each component and the different association between them. These observations will reflect different patterns once displayed a map.

The results from MULTISPATI-PCA inform that the variables generate clustered patterns; nevertheless it is necessary to perform a statistical test in order to define whether the patterns occurred by chance or were determined by spatial autocorrelation. A Monte Carlo permutation test is run with 999 permutations of the rows of table **X**, the total inertia of the analysis is computed and its value increases with the intensity of the link between the matrix **X** of the original variables and the matrix **WX** of the lagged variables (Ollier, 2005).

In the case of "East" figure 6.18 illustrates that the observed Moran's *I* in the graph is far apart from the simulated values as indication that none of the simulated values were larger than the observed ones. Hence, the test revealed that the spatial autocorrelation measured in MULTISPATI was statistically significant.

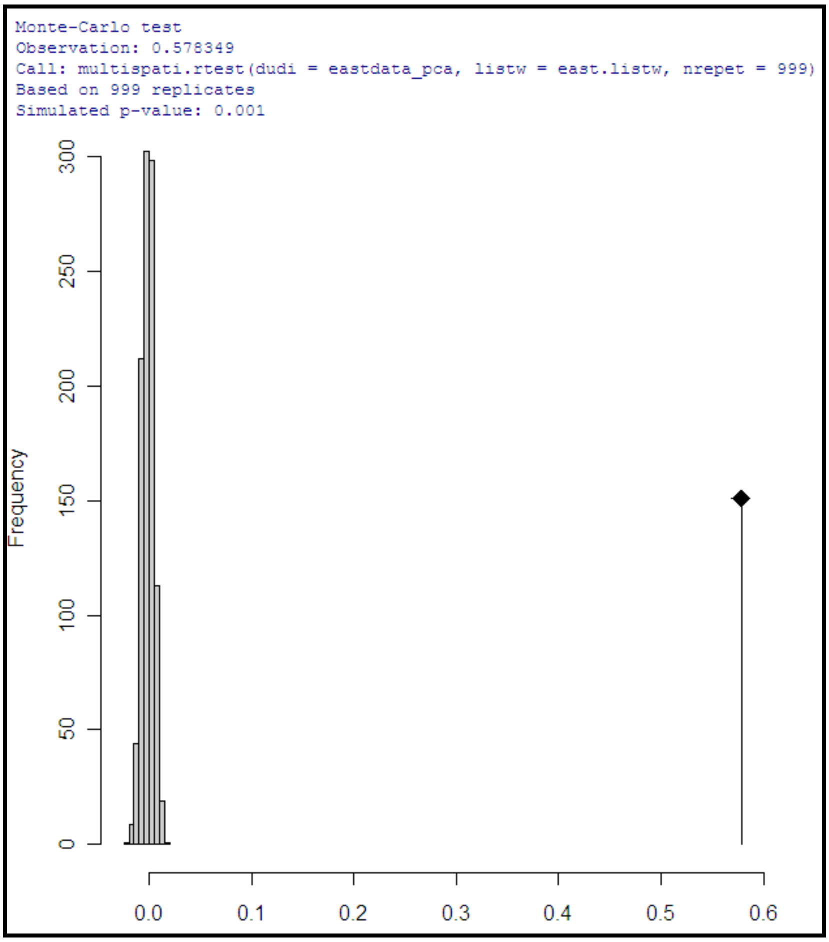


Figure 6.18. The test for significance depicts that Moran's *I* was statistically valid. The line on the right side of the graph indicates the observed value of Moran's *I* and it is well away from the bars on the left that refer to the simulated values generated through a randomisation process.

From the explanation and description of the results obtained from the application of MULTISPATI-PCA it emerges that this statistical technique is effective in the identification of spatial patterns within the data. However, two points need to be absolutely remembered:

- MULTISPATI is a purely descriptive method based on linear algebra and on geographical and geometrical properties of the data. It does not rely on any model of fitting and it can be applied on any type of variables such as binary, counts, or qualitative ones (Saby et al, 2009);
- MULTISPATI should be the preferred method of analysis if the aim is to study the spatial structure of one dataset. If descriptors (e.g. environmental variables) other than space are included in order to perform variation partitioning, then MULTISPATI should not be used. The reason is due to the fact that primarily MULTISPATI works with spatial autocorrelation and not with variances. Other techniques such as redundancy analysis with Moran's eigenvector maps and principal coordinates of neighbour matrices should be preferred (Dray et al, 2008).

6.6 Cluster analysis: the bridge between MULTISPATI and GIS

As seen above the maps of the scores and lagged scores, plotted axis by axis, illustrate the spatial patterns identified with MULTISPATI-PCA, however the representation of the polygons in terms of big or small squares is not useful in practical terms. In fact it is difficult to understand from those maps which polygon corresponds to a spatial pattern. Therefore in order to have a clearer display of the different landscape characters it is necessary to think of a way to export the results from MULTISPATI-PCA and import them in arcGIS where clearer maps can be achieved.

According to a previous work carried by S. Ollier (2005) a cluster analysis of the scores and lagged scores seems to be the bridge between MULTISPATI and GIS. In its general definition a cluster analysis aims at the designation of groups of similar items within a dataset. Commonly the cluster analysis is split into two different methods: partitioning and hierarchical; the former divides the dataset into a number of groups pre-designated by the user, while the latter produces a hierarchy of clusters from the small ones characterised by very similar values to the large ones that include more dissimilar values (Holland, 2006). Hierarchical clusters can be further split using two methods: the first is called divisive, which divides a large cluster into two smaller ones and it repeats this process until all the clusters are divided. The second is called agglomerative and works exactly in the opposite direction to the previous method (Holland, 2006).

The agglomerative hierarchical clusters method is commonly used in the natural sciences and allows displaying similarities of values across a wide range of scales (Holland, 2006). For this reason, they are the most appropriate for landscape analysis and were applied in this study. The approach used for cluster analysis is once again a critical moment because a series of decisions was required about:

- the linkage method for considering the values,
- the method for considering the distance between the values,
- the option of measuring the cluster on the basis of similarities or dissimilarities
- and finally the optimal number of clusters for the observed dataset of values.

In this research the cluster analysis was run using the software Minitab, which was preferred to R because we wanted to try and explore different software as part of the learning process that characterises the PhD. The use of Minitab highlighted that in terms of performance R is a much better tool and for the future it is suggested the use of the package *hclust* in order to carry out the cluster analysis.

After several tests and with reference to Ollier (2005), the linkage method “Ward” and the Euclidean distance[26] provided the most convincing results, hence both were adopted for the cluster analysis. The determination of the number of clusters[27] is based on personal judgement, although this is backed by the analysis of the graph and the table generated from the cluster analysis.

The dendrogram[28] is a useful graph because it allows the analyst to have a better idea of the way the data is grouped and it shows clearly how the clusters are related to each other. Figure 6.19 illustrates two dendrograms obtained by clustering the scores and lagged scores of the area of study “East” at two different levels, respectively 10 and 5 number of clusters.

Figure 6.19

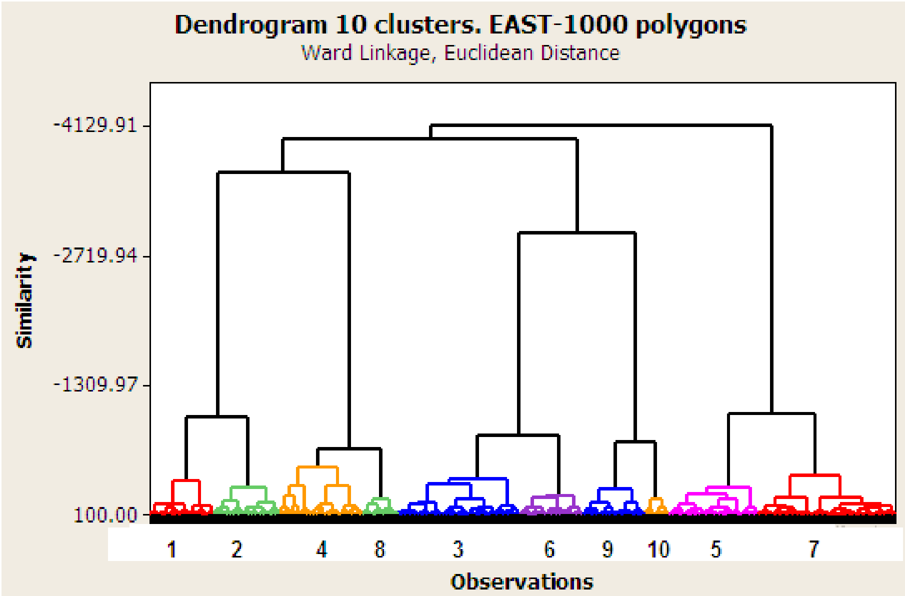
The strength of clustering is indicated by the level of similarity at which the elements join a cluster. This level is read on the vertical axis of the dendrogram and is graphically showed by the length of the stems connecting all the data observed; therefore the shorter the stem the higher the similarity (Holland, 2006). In addition to the dendrogram the values of similarities and distance between clusters are provided in two tables displayed by Minitab. These are illustrated in figure 6.20 and refer to the two dendrograms showed above.

Figure 6.20

[26] Specifically the decision of using Euclidean distance was base on the personal comment received by Dray who explained that “Euclidean” was the more logical distance since MULTISPATI is based on it.

[27] The patterns identified on the maps of scores/lagged scores correspond to the clusters of the cluster analysis.

[28] A dendrogram is the graphical output of the agglomerative hierarchical cluster analysis that shows the grouping of the data. It looks as a tree and it is comprises of lines, referred as stems, which link data on the basis of distance and similarity. A dendrogram that clearly differentiates data in groups has small distances in the furthest branches of the tree and big distances in the nearest branches.



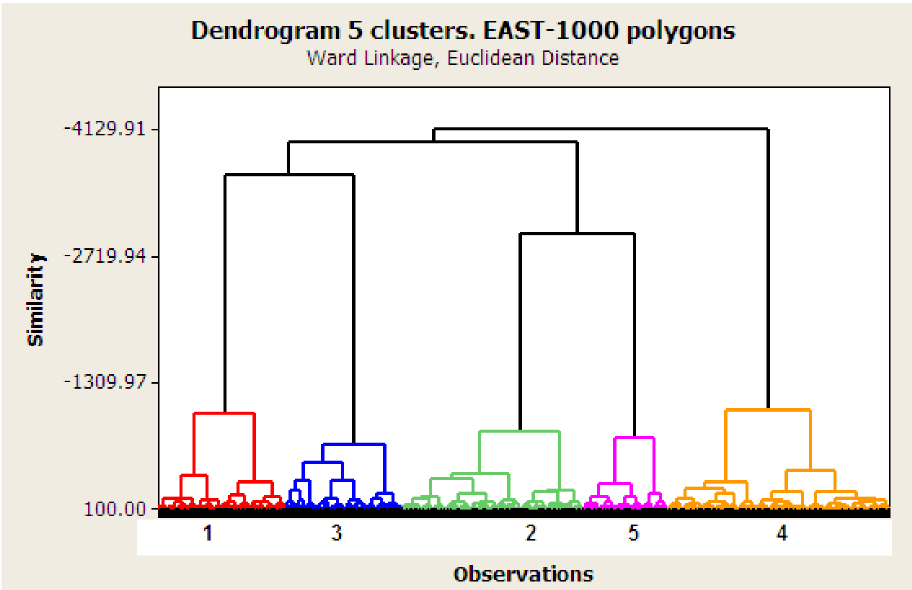


Figure 6.19. Two dendrograms of 10 and 5 clusters calculated for the “East” area of study with 1000 polygons.

steps A	number of clusters B	level of similarity C	Euclidean distance D
988	12	-259.49	27.565
989	11	-390.16	37.584
990	10	-454.87	42.545
991	9	-619.82	55.193
992	8	-819.09	70.472
993	7	-1001.71	84.474
994	6	-1095.48	91.665
995	5	-1359.42	111.902
996	4	-1977.35	159.283
997	3	-4452.97	349.104
998	2	-5127.30	400.809
999	1	-6224.42	484.932

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	237	907.834	1.85662	3.76557
Cluster2	189	175.837	0.90978	2.13678
Cluster3	346	459.522	1.03059	3.40527
Cluster4	128	322.172	1.47931	3.07798
Cluster5	100	108.224	0.96184	2.26150

	Number of observations	Within cluster sum of squares	Average distance from centroid	Maximum distance from centroid
Cluster1	60	35.545	0.72196	1.55417
Cluster2	89	291.131	1.75378	3.40978
Cluster3	189	175.837	0.90978	2.13678
Cluster4	142	97.120	0.78726	1.56395
Cluster5	90	105.648	1.03517	1.87076
Cluster6	158	68.168	0.61102	1.72557
Cluster7	38	64.950	1.19224	2.57566
Cluster8	100	108.224	0.96184	2.26150
Cluster9	46	59.697	1.09040	1.79093
Cluster10	88	104.622	1.03244	2.01547

Figure 6.20. Tables related to the dendrograms depicted in the previous figure. The most important information can be read in column C and D where it is possible to understand the values assumed by similarity and Euclidean distance. It stands out that the smaller the distance between the variables forming a cluster, the higher their level of similarity.

In deciding the optimal number of clusters, it is necessary to be cautious about the level of similarity to be maintained. If the similarity is kept very high, many clusters are likely to be generated and their large number does not simplify the analysis. However, if the similarity is low, the analyst will deal with a smaller number of clusters which are likely to be comprised of very different variables (Holland, 2006).

At the end the choice about the number of clusters to retain was taken on the basis of the information conveyed by the dendrograms and by the GIS maps (see figure 6.21) due to their greater readability and clarity in delivering the information.

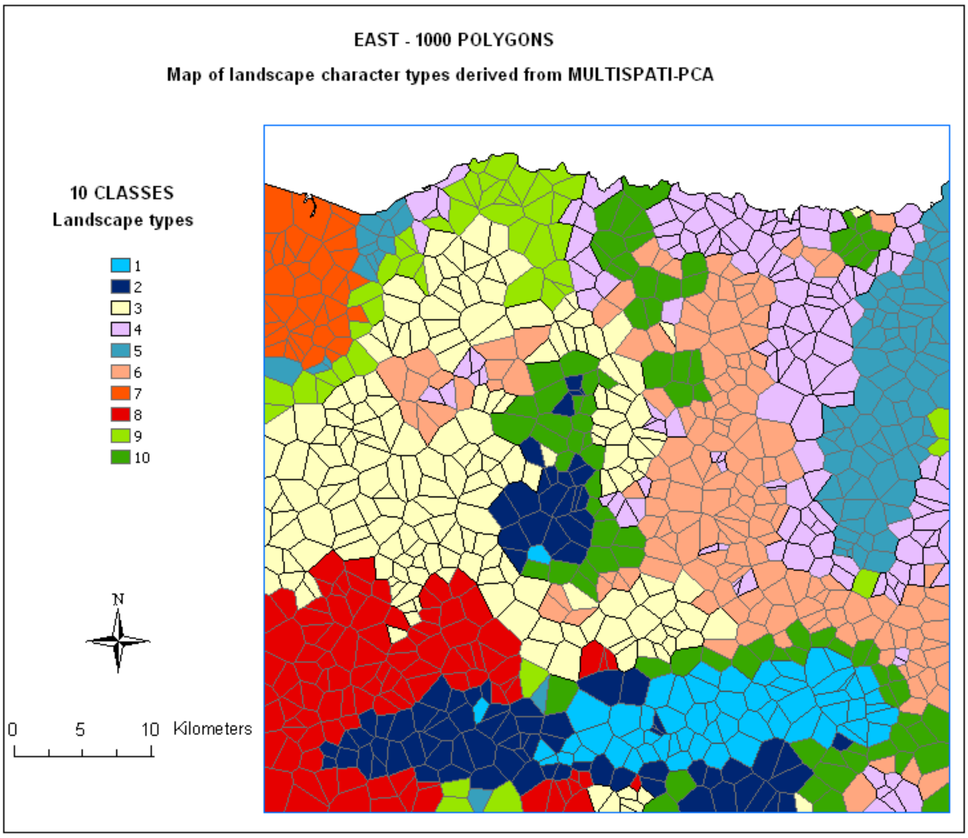
Figure 6.21

As previously mentioned, a greater number of clusters makes the map of landscape character types more complex and perhaps more difficult to read. However, if the need is for a landscape characterisation that looks at the character types in more detail, then a large number of clusters should be maintained. Otherwise, for a general overview of the landscape character types, a small number of clusters should be preferred.

Knowledge about the landscape elements composition of each cluster can help to refine the decision about the best number of clusters to retain in order to represent the landscape character types of a given area of study. From the attribute table associated to each cluster it is possible to extract the landscape elements composition and verify whether or not redundant information about the landscape character types is conveyed. This operation is explained in the next chapter, which also provides more examples of application of MULTISPATI-PCA.

This chapter has illustrated the different stages of the second part of statistical techniques applied to define the landscape character types of a study area.

The choice of PCA is taken with the intention of simplifying a highly complex geographical system such as the landscape, while the decision to constrain the analysis to the space with MULTISPATI-PCA follows the theory of spatial autocorrelation, namely that the elements are not randomly distributed but they combine to each other according to their proximity and similarity in their values.



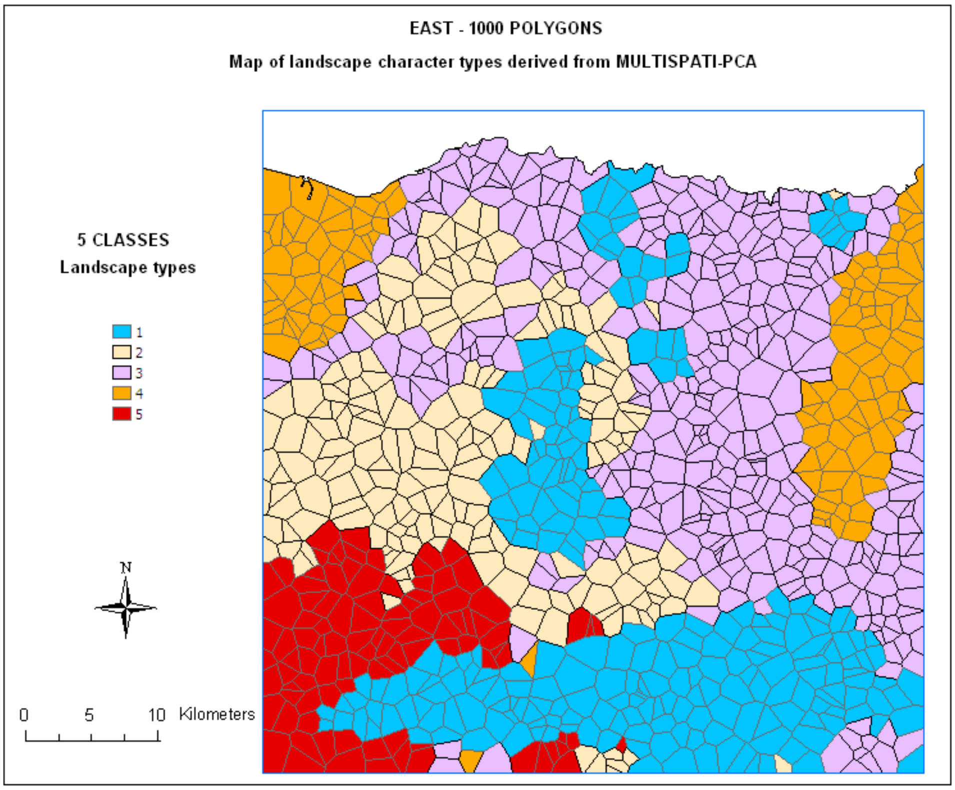


Figure 6.21. Maps of the landscape character types of the area of study “East” at scale 1000 polygons if 10 and 5 clusters, which are calculated on the scores and lagged scores of MULTISPATI-PCA, are retained.

In brief MULTISPATI-PCA summarizes and extracts the landscape elements that occur more frequently and that contribute significantly to describe the whole landscape by taking into account the influence of the space (geographical location) occupied by the elements. In order to be visualised in GIS, the spatial structures revealed from MULTISPATI-PCA are analysed through a cluster analysis which forms them again by grouping the scores and lagged scores (the value of single polygons and the average of the values of the neighbouring polygons) into classes with similar characteristics.

In brief the GIS-MULTISPATI based methodology for landscape characterisation is based on calculating:

- MULTISPATI-PCA, which constrains a standard PCA to spatial autocorrelation rules by applying a row-sum standardised weighted matrix W to the data (the statistical triplet X, Q, D of the duality diagram *dudi* at pages 172-173) previously calculated in standard PCA;
- Hierarchical agglomerative cluster analysis, which is applied to the scores and lagged scores calculated by MULTISPATI-PCA and re-organises the spatial patterns by making them readable and “mappable” through GIS.

The following chapter illustrates the outcomes from the application of the GIS/MULTISPATI-based methodology for landscape characterisation to the entire set of case studies, and describes the several tests of robustness that have been run on the methodology.

CHAPTER 7

Understanding the results and testing the methodology for robustness

Loch Nevis from Sgurr na Ciche (I.Marengo).

7.1 Landscape characterisation through GIS and MULTISPATI "eyes"

The way MULTISPATI is applied to reveal how the landscape elements are spatially structured and correlated to each others has been so far described generally but it has been demonstrated to have the great advantage of being objective, scientifically robust and logically designed. Evidence of spatial relationships amongst the landscape elements are operationally translated into the presence of spatial clusters which are identified and referred in terms of landscape character types.

In the first part of this chapter attention is given to the outcomes from the application of MULTISPATI to the selected areas of study (see figure 7.1) and their comparison to LCA maps; in addition, an explanation of how the GIS/MULTISPATI maps should be read and interpreted is provided.

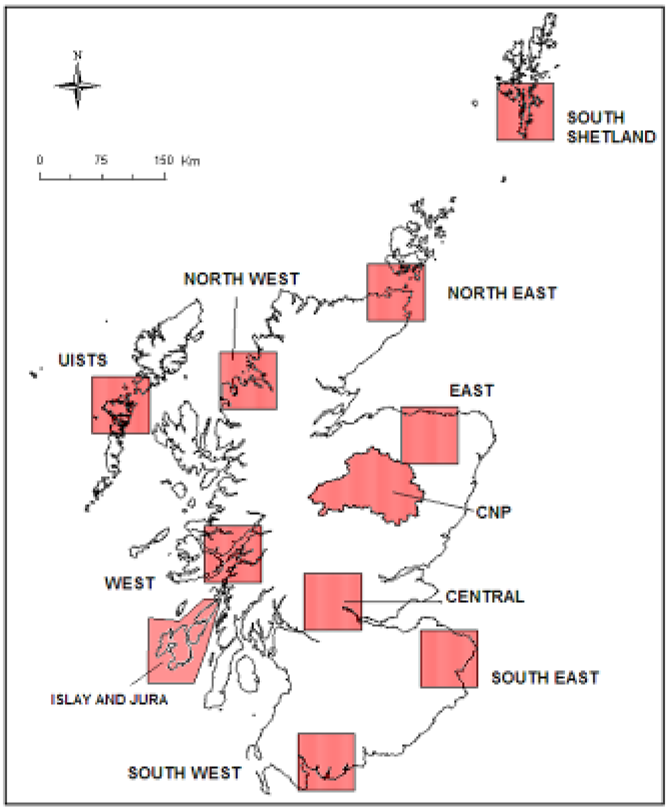


Figure 7.1. Location of the 11 areas of study across Scotland with contrasting landscapes. On these areas the GIS methodology was applied and tested. Each area covers 2,500 km² except for the Cairngorms National Park which stretches for 3,800 km² and the isles of Islay and Jura which cover 2,300 km².

In the second half of the chapter, how the GIS/MULTISPATI-based methodology would work when particular situations occur is elucidated. Precisely, simulation of what happens if a different number of polygons is used, new polygons (but same number) are generated, if two neighbouring areas are analysed and finally if different datasets for the same area are considered.

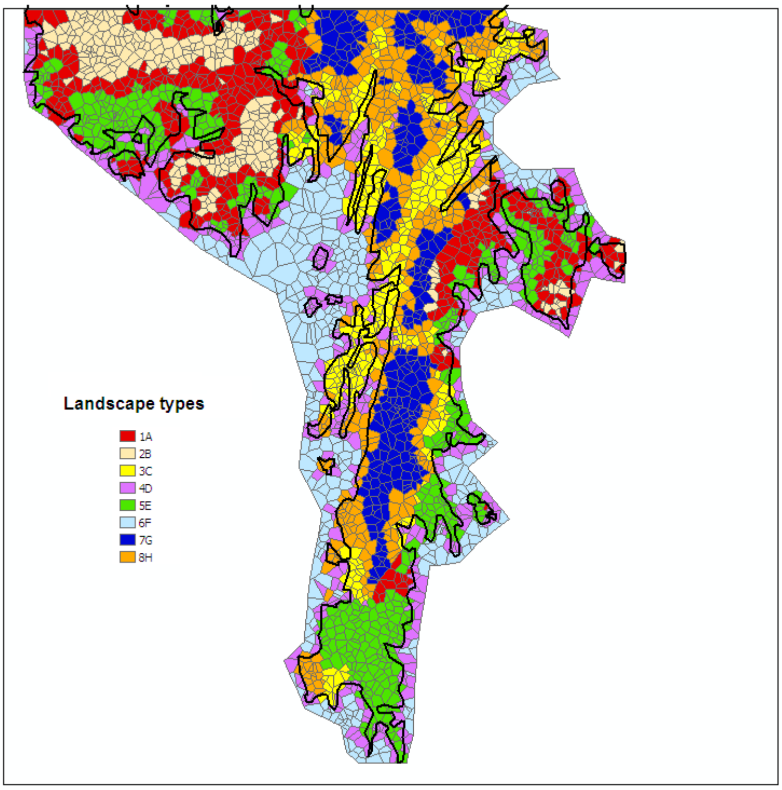
The application of MULTISPATI provided the following maps (see list in figure 7.2) of landscape character types for each area of study.

Figure 7.2

At first sight some of the maps in figure 7.2, and specifically those of North Uist, South Shetland, Cairngorms, and Northeast, showed a surprising similarity with the correspondent LCA maps. For these areas the GIS/MULTISPATI-based landscape analysis seemed to produce a compelling landscape characterisation. On the contrary, the rest of the areas revealed a lower degree of similarity which suggests uncertainty in the accuracy and success of the GIS/MULTISPATI-based landscape characterisation. Nevertheless, a visual comparison of the map is not a scientifically robust method to determine the degree of similarity/discrepancy between GIS and LCA maps; in fact the judgement and the impression about the look of the GIS maps and the way their appearance matches with that of the LCA maps is based on a personal and therefore subjective interpretation. This can be sometimes correct but other times misleading thus two more objective approaches were adopted in order to be confident about the calculations operated and to verify the validity of the GIS/MULTISPATI-based landscape analyses. A first approach was based on statistics and the second on the factual evidence collected from the way the LCA was conducted.

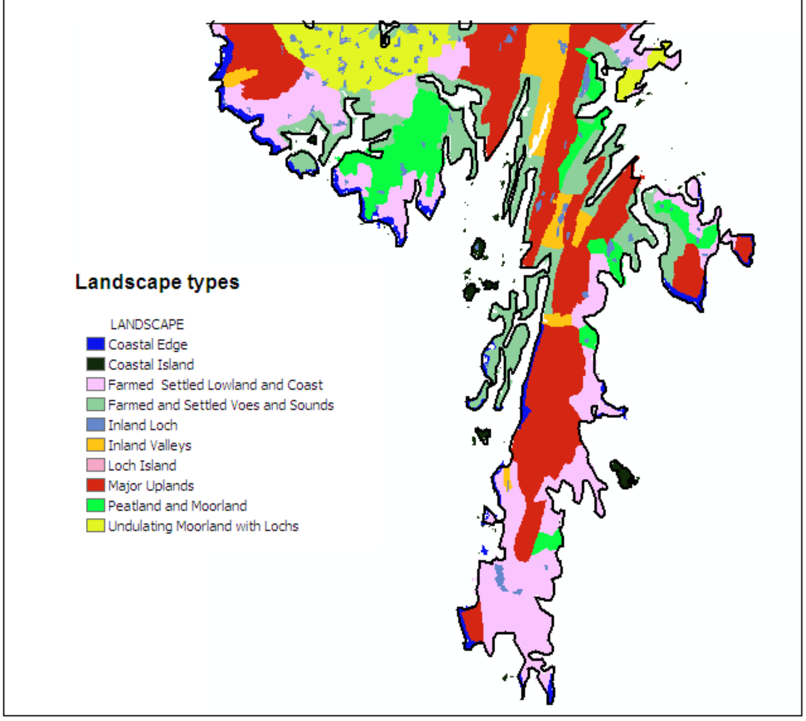
SOUTH SHETLAND

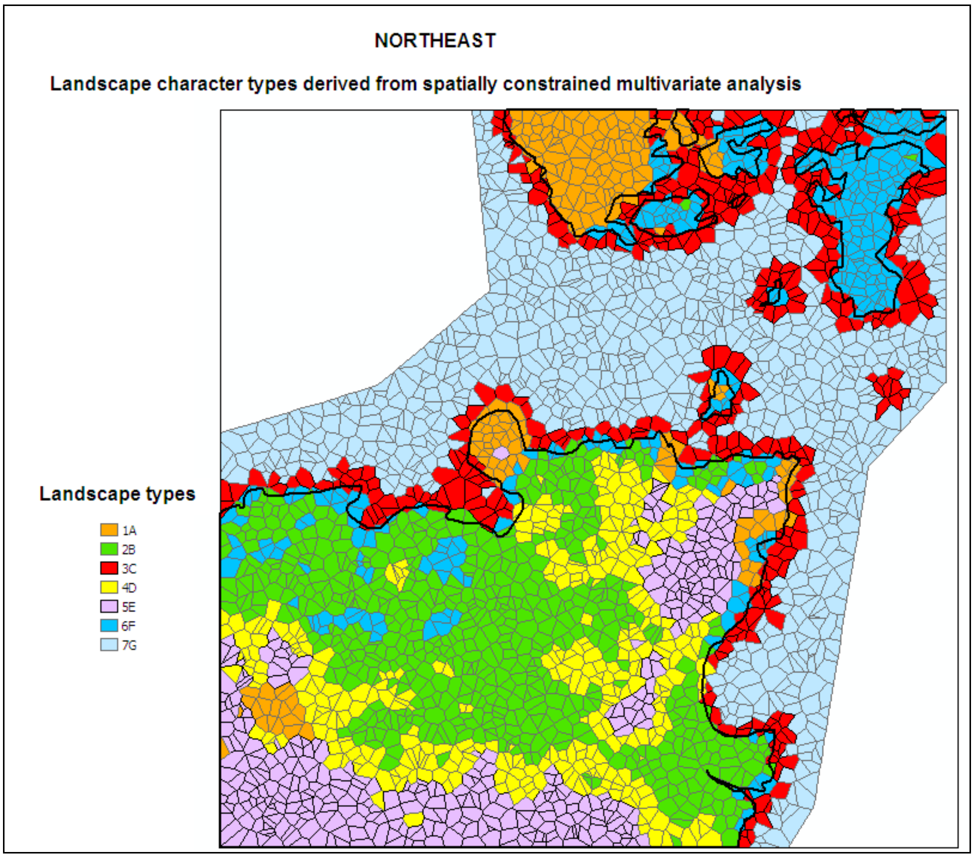
Landscape character types derived from spatially constrained multivariate analysis

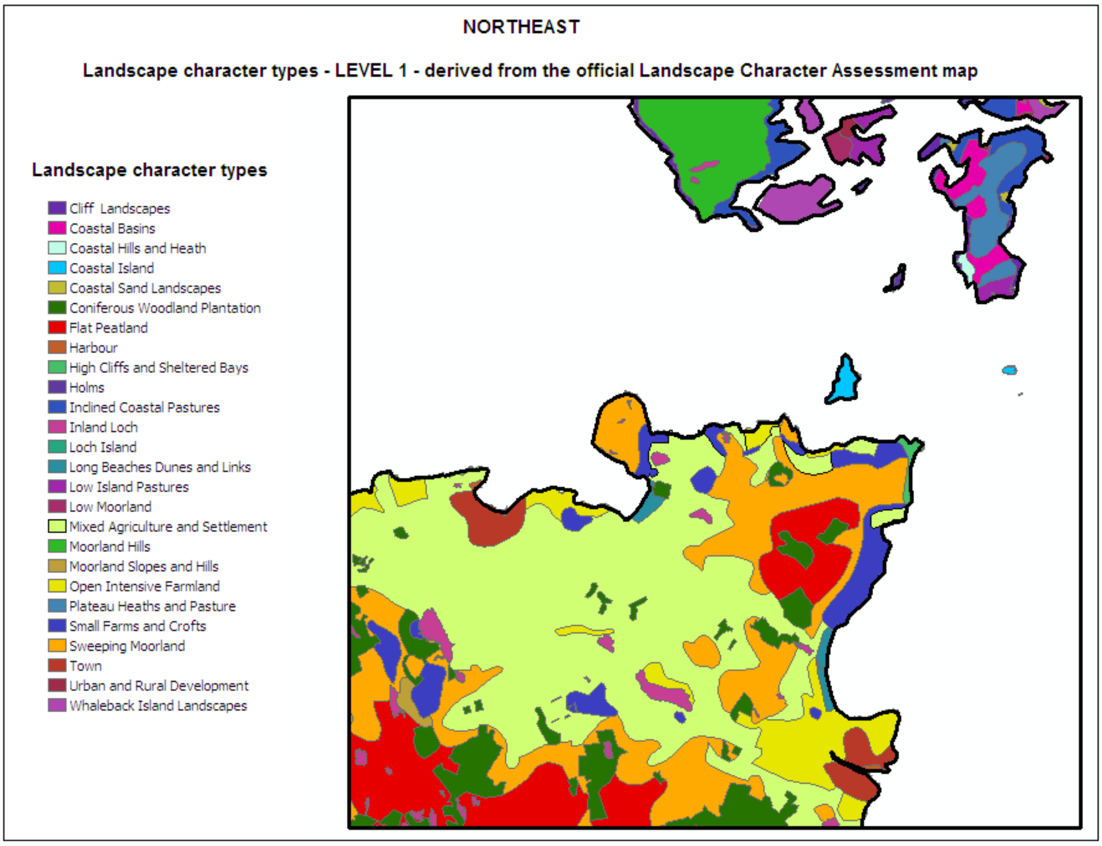


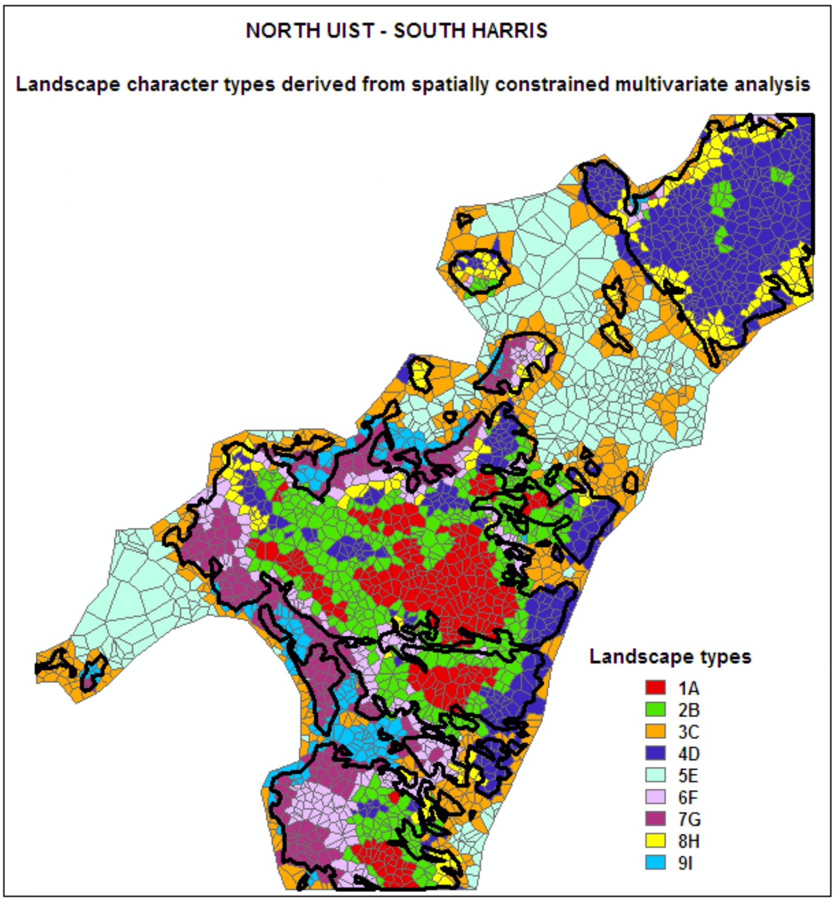
SOUTH SHETLAND

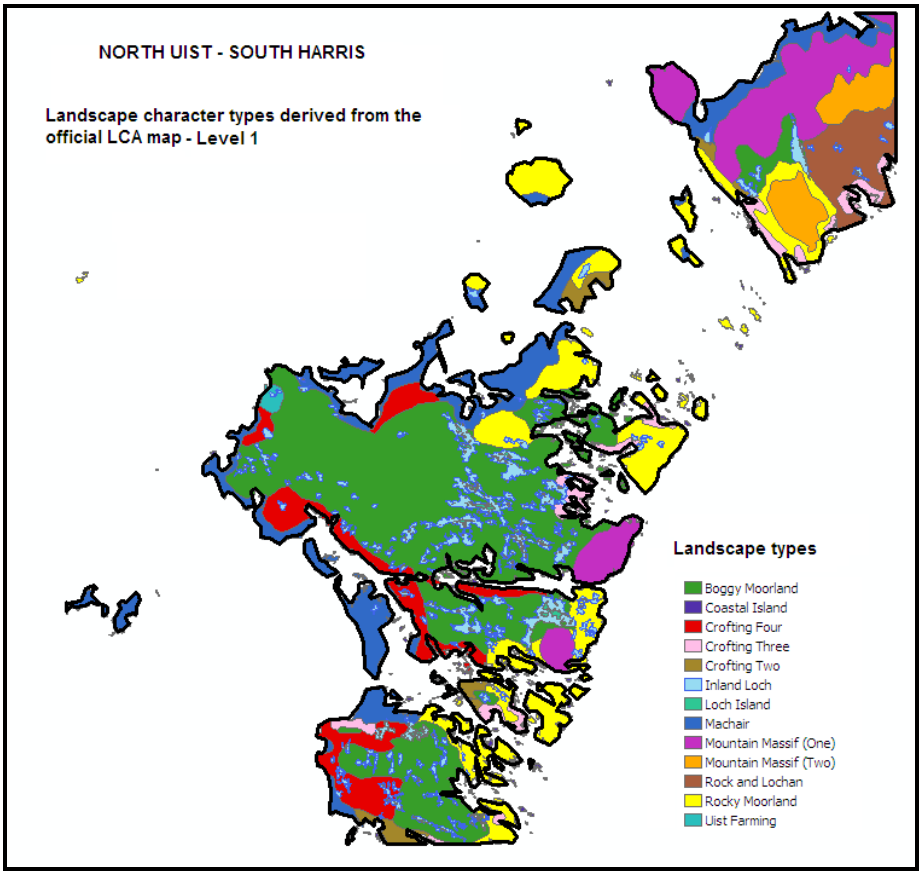
Landscape character types derived from official LCA - Level 1

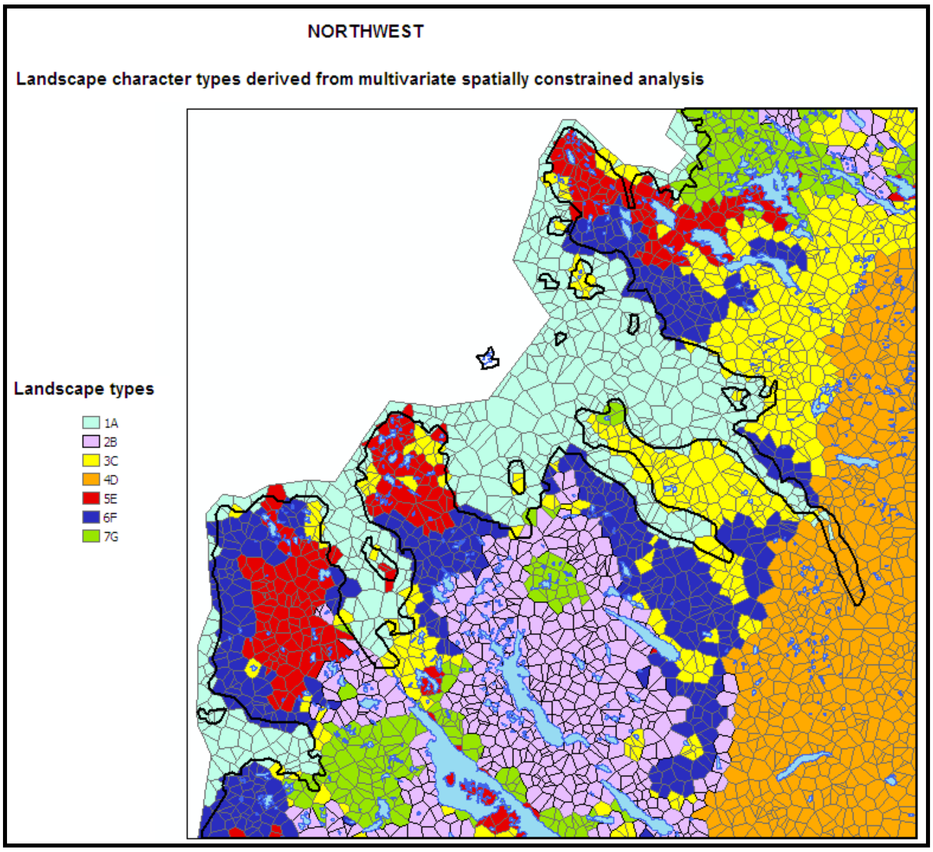


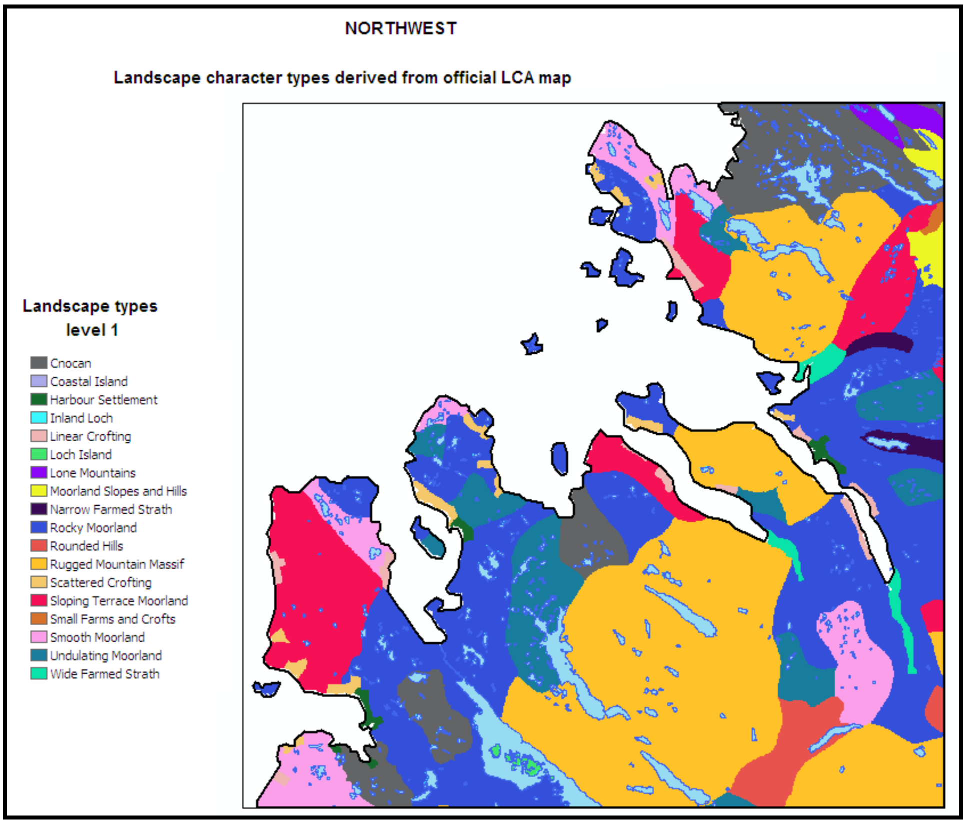

NORTHEAST
Landscape character types derived from spatially constrained multivariate analysis
Landscape types
1A
2B
3C
4D
5E
6F
7G

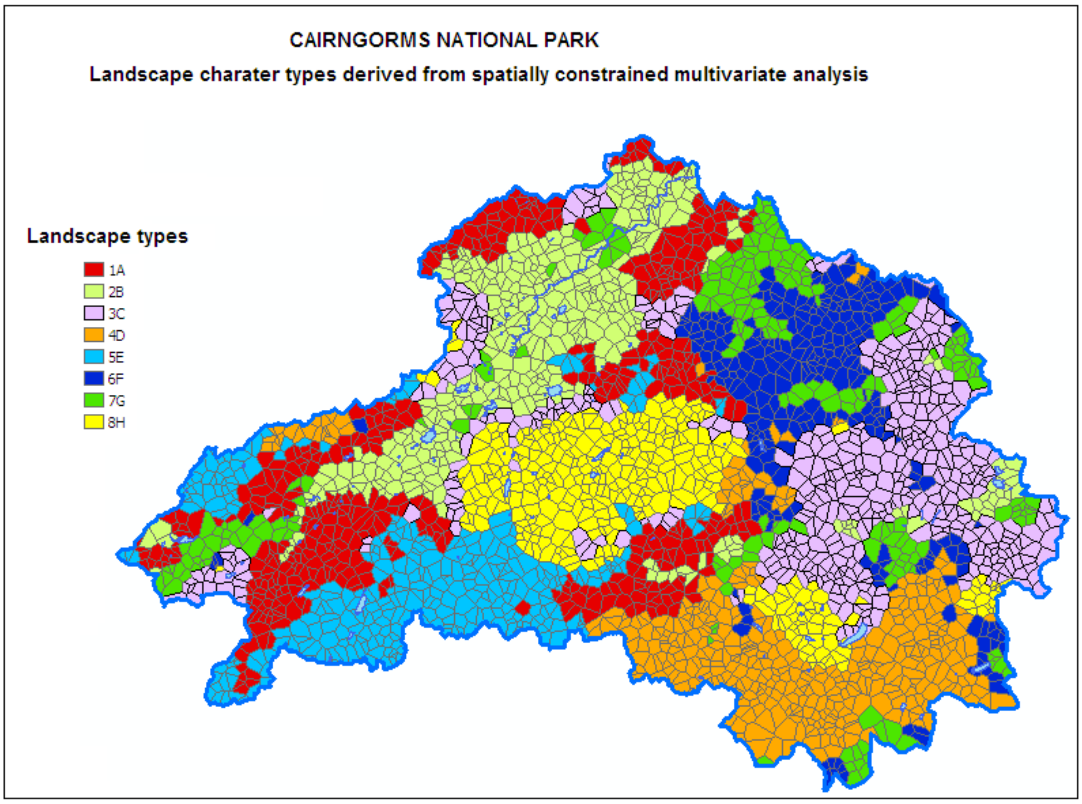

NORTHEAST
Landscape character types - LEVEL 1 - derived from the official Landscape Character Assessment map
Landscape character types
Cliff Landscapes
Coastal Basins
Coastal Hills and Heath
Coastal Island
Coastal Sand Landscapes
Coniferous Woodland Plantation
Flat Peatland
Harbour
High Cliffs and Sheltered Bays
Holms
Inclined Coastal Pastures
Inland Loch
Loch Island
Long Beaches Dunes and Links
Low Island Pastures
Low Moorland
Mixed Agriculture and Settlement
Moorland Hills
Moorland Slopes and Hills
Open Intensive Farmland
Plateau Heaths and Pasture
Small Farms and Crofts
Sweeping Moorland
Town
Urban and Rural Development
Whaleback Island Landscapes

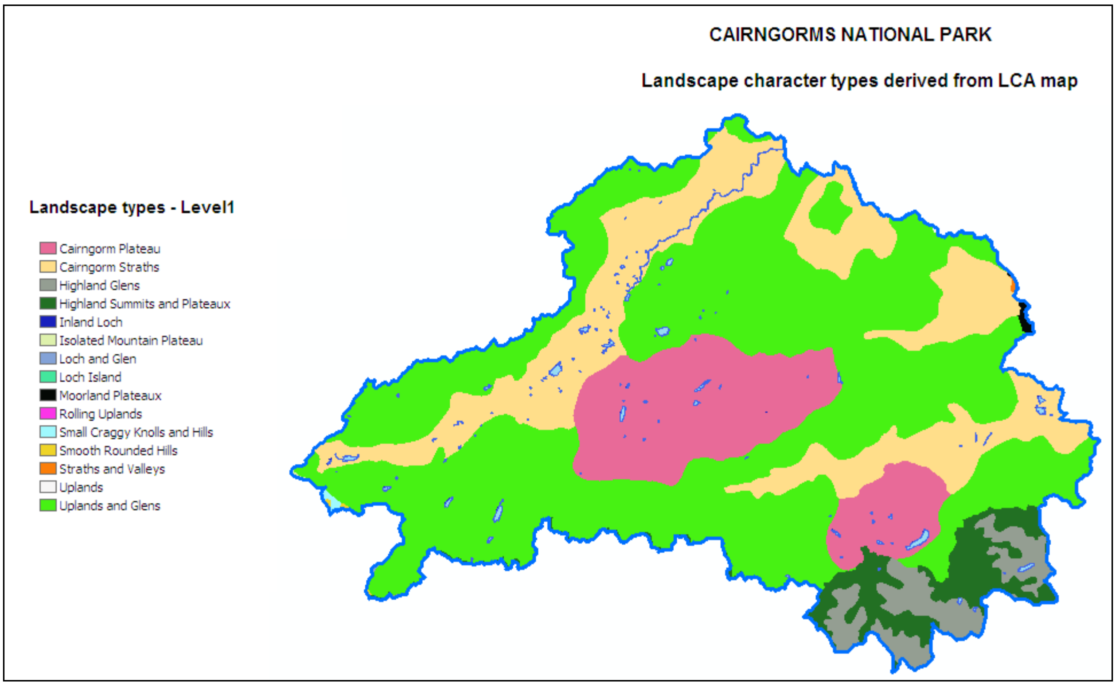

NORTH UIST - SOUTH HARRIS
Landscape character types derived from spatially constrained multivariate analysis
Landscape types
1A
2B
3C
4D
5E
6F
7G
8H
9I

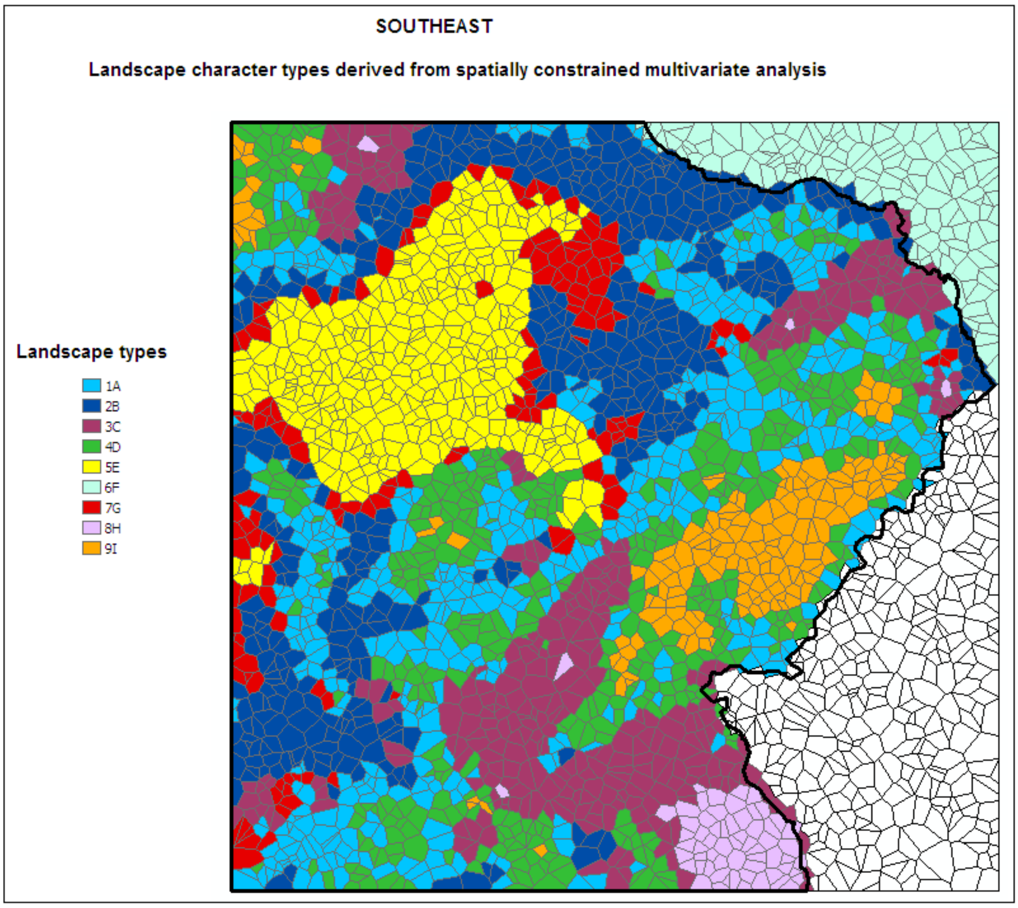

NORTH UIST - SOUTH HARRIS
Landscape character types derived from the official LCA map - Level 1
Landscape types
Boggy Moorland
Coastal Island
Crofting Four
Crofting Three
Crofting Two
Inland Loch
Loch Island
Machair
Mountain Massif (One)
Mountain Massif (Two)
Rock and Lochan
Rocky Moorland
Uist Farming

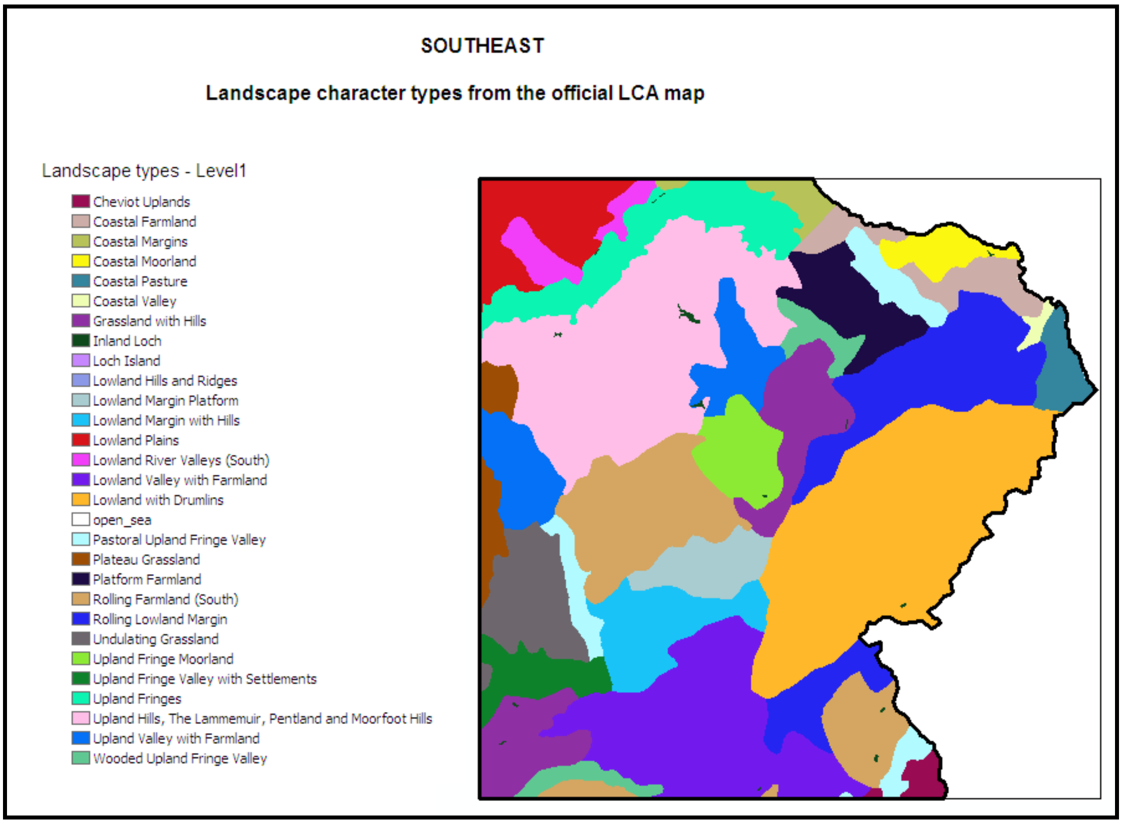

NORTHWEST
Landscape character types derived from multivariate spatially constrained analysis
Landscape types
1A
2B
3C
4D
5E
6F
7G

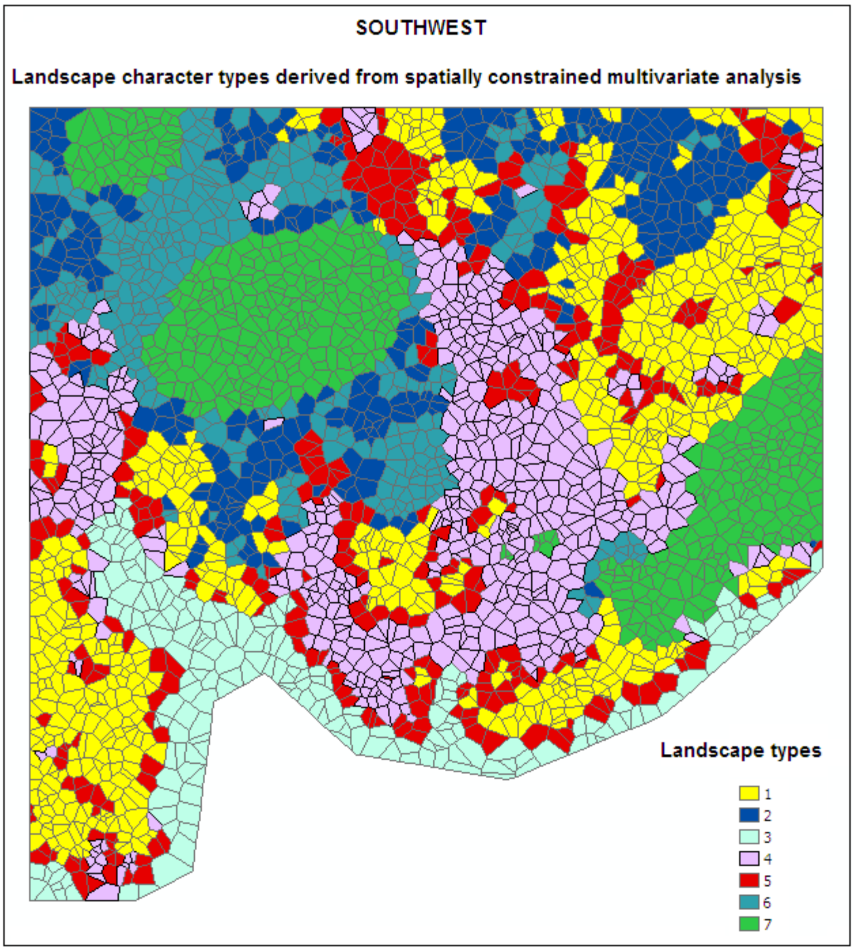

NORTHWEST
Landscape character types derived from official LCA map
Landscape types
level 1
Cnocan
Coastal Island
Harbour Settlement
Inland Loch
Linear Crofting
Loch Island
Lone Mountains
Moorland Slopes and Hills
Narrow Farmed Strath
Rocky Moorland
Rounded Hills
Rugged Mountain Massif
Scattered Crofting
Sloping Terrace Moorland
Small Farms and Crofts
Smooth Moorland
Undulating Moorland
Wide Farmed Strath

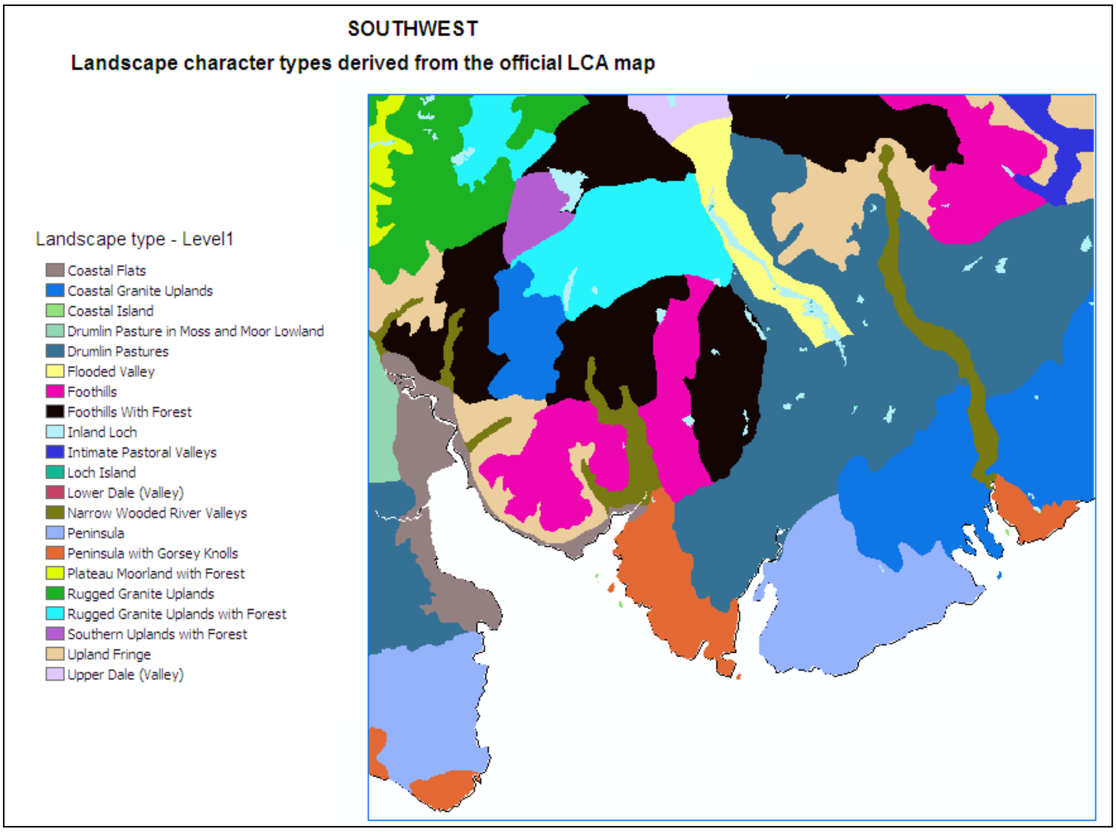

CAIRNGORMS NATIONAL PARK
Landscape charater types derived from spatially constrained multivariate analysis
Landscape types
1A
2B
3C
4D
5E
6F
7G
8H

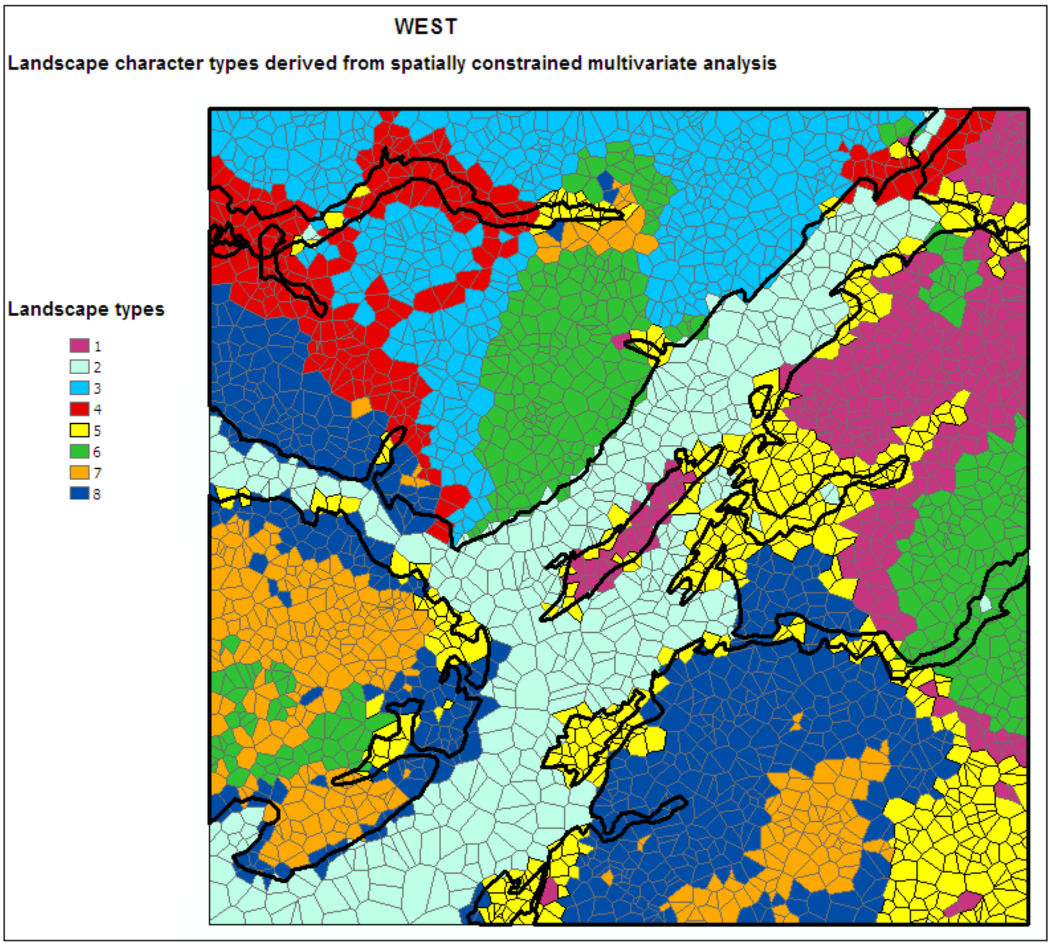

CAIRNGORMS NATIONAL PARK
Landscape character types derived from LCA map
Landscape types - Level1
Cairngorm Plateau
Cairngorm Straths
Highland Glens
Highland Summits and Plateaux
Inland Loch
Isolated Mountain Plateau
Loch and Glen
Loch Island
Moorland Plateaux
Rolling Uplands
Small Craggy Knolls and Hills
Smooth Rounded Hills
Straths and Valleys
Uplands
Uplands and Glens

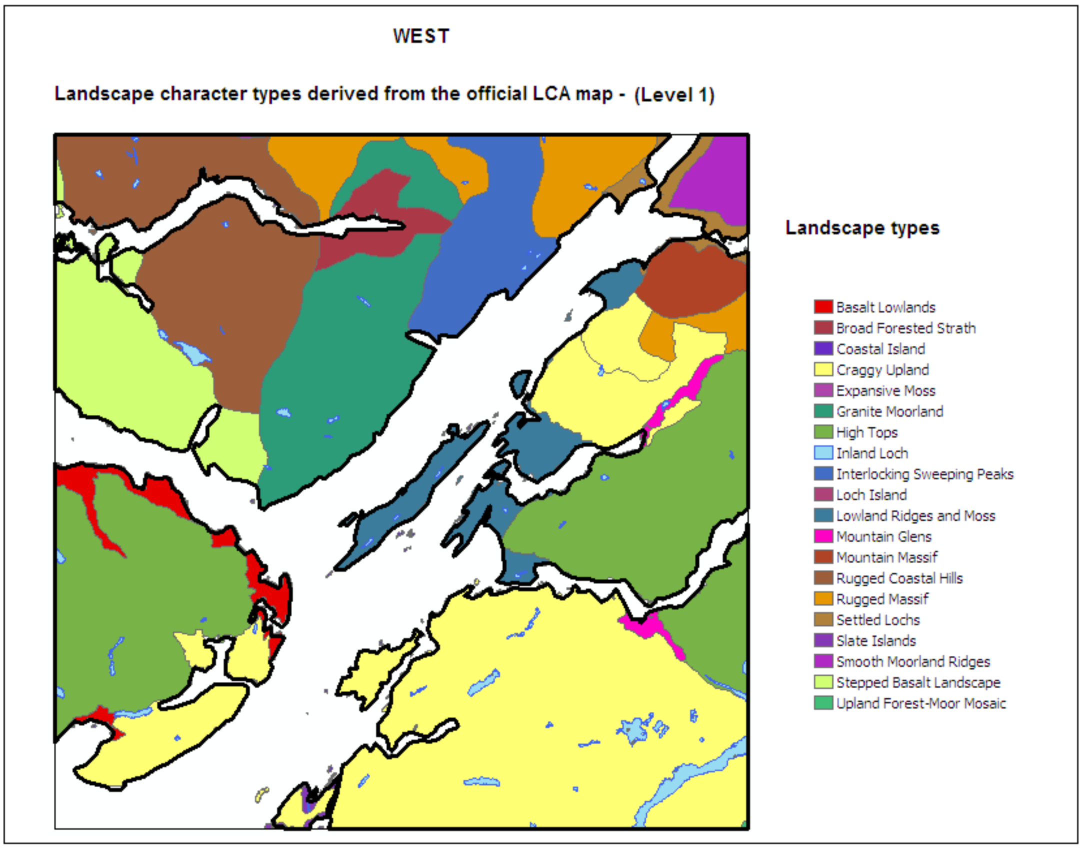

SOUTHEAST
Landscape character types derived from spatially constrained multivariate analysis
Landscape types
1A
2B
3C
4D
5E
6F
7G
8H
9I

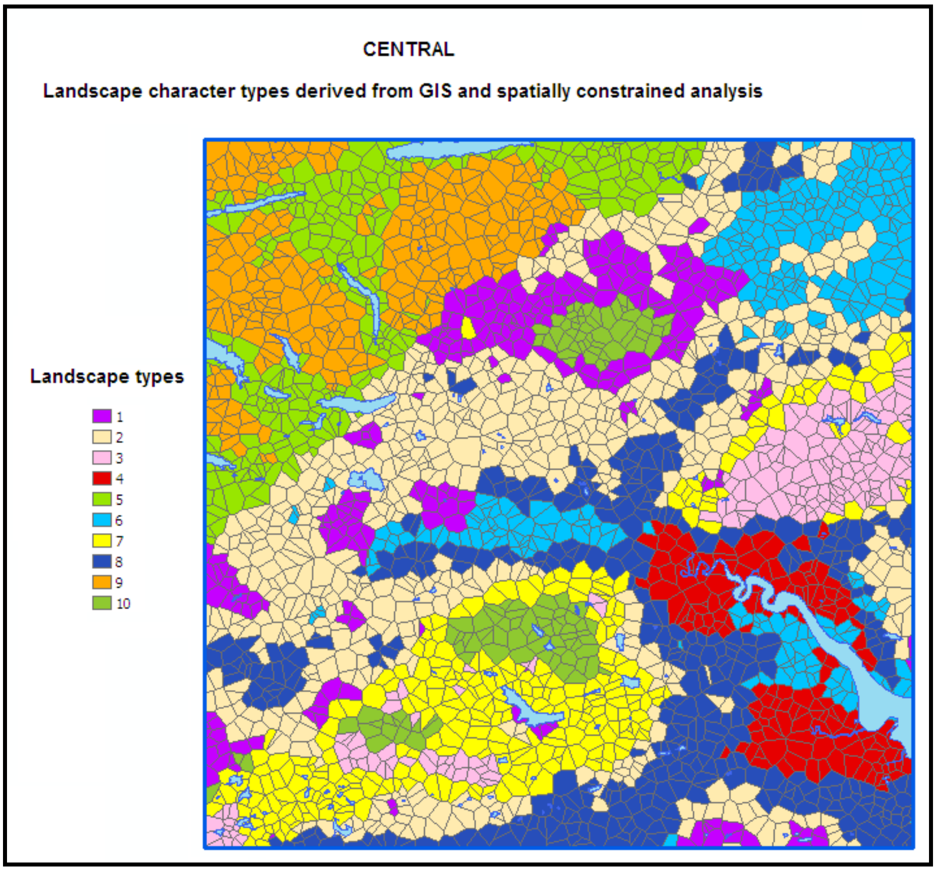

SOUTHEAST
Landscape character types from the official LCA map
Landscape types - Level1
Cheviot Uplands
Coastal Farmland
Coastal Margins
Coastal Moorland
Coastal Pasture
Coastal Valley
Grassland with Hills
Inland Loch
Loch Island
Lowland Hills and Ridges
Lowland Margin Platform
Lowland Margin with Hills
Lowland Plains
Lowland River Valleys (South)
Lowland Valley with Farmland
Lowland with Drumlins
open_sea
Pastoral Upland Fringe Valley
Plateau Grassland
Platform Farmland
Rolling Farmland (South)
Rolling Lowland Margin
Undulating Grassland
Upland Fringe Moorland
Upland Fringe Valley with Settlements
Upland Fringes
Upland Hills, The Lammemuir, Pentland and Moorfoot Hills
Upland Valley with Farmland
Wooded Upland Fringe Valley


SOUTHWEST
Landscape character types derived from spatially constrained multivariate analysis
Landscape types
1
2
3
4
5
6
7

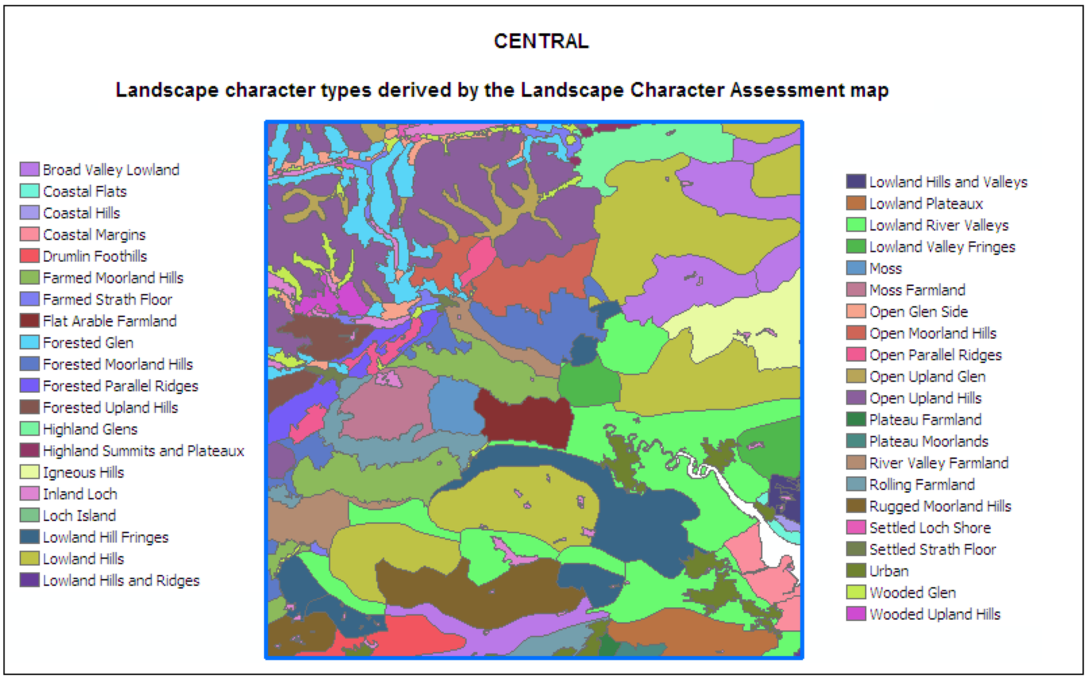

SOUTHWEST
Landscape character types derived from the official LCA map
Landscape type - Level1
Coastal Flats
Coastal Granite Uplands
Coastal Island
Drumlin Pasture in Moss and Moor Lowland
Drumlin Pastures
Flooded Valley
Foothills
Foothills With Forest
Inland Loch
Intimate Pastoral Valleys
Loch Island
Lower Dale (Valley)
Narrow Wooded River Valleys
Peninsula
Peninsula with Gorsey Knolls
Plateau Moorland with Forest
Rugged Granite Uplands
Rugged Granite Uplands with Forest
Southern Uplands with Forest
Upland Fringe
Upper Dale (Valley)


WEST
Landscape character types derived from spatially constrained multivariate analysis
Landscape types
1
2
3
4
5
6
7
8

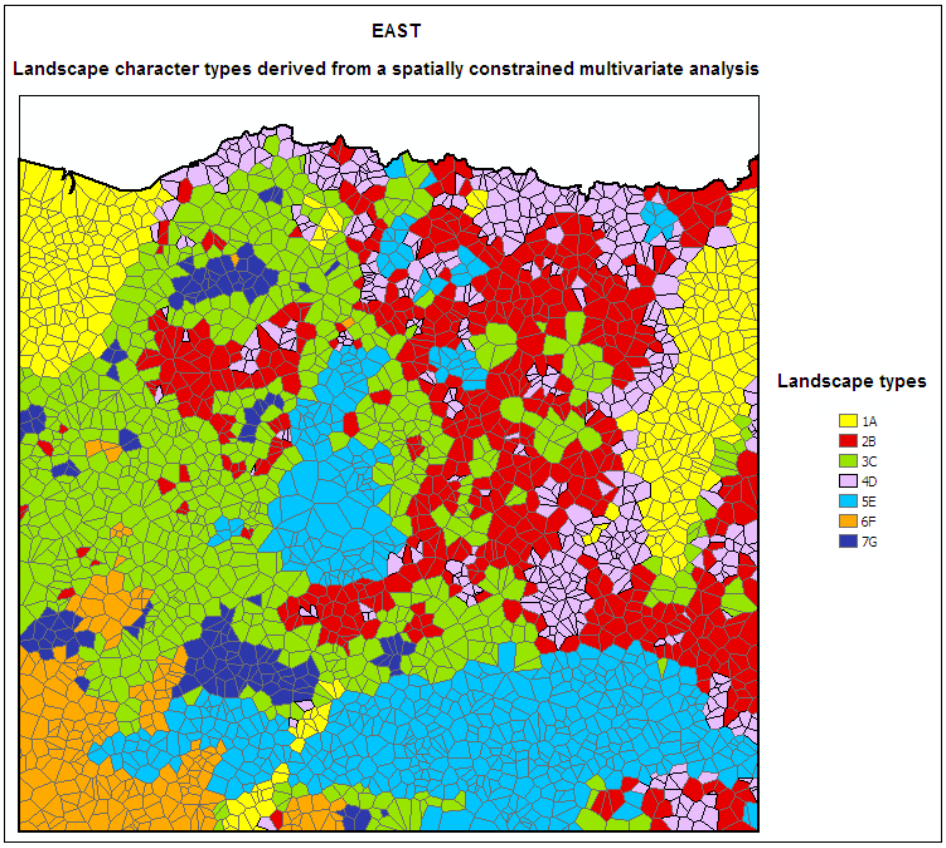

WEST
Landscape character types derived from the official LCA map - (Level 1)
Landscape types
Basalt Lowlands
Broad Forested Strath
Coastal Island
Craggy Upland
Expansive Moss
Granite Moorland
High Tops
Inland Loch
Interlocking Sweeping Peaks
Loch Island
Lowland Ridges and Moss
Mountain Glens
Mountain Massif
Rugged Coastal Hills
Rugged Massif
Settled Lochs
Slate Islands
Smooth Moorland Ridges
Stepped Basalt Landscape
Upland Forest-Moor Mosaic

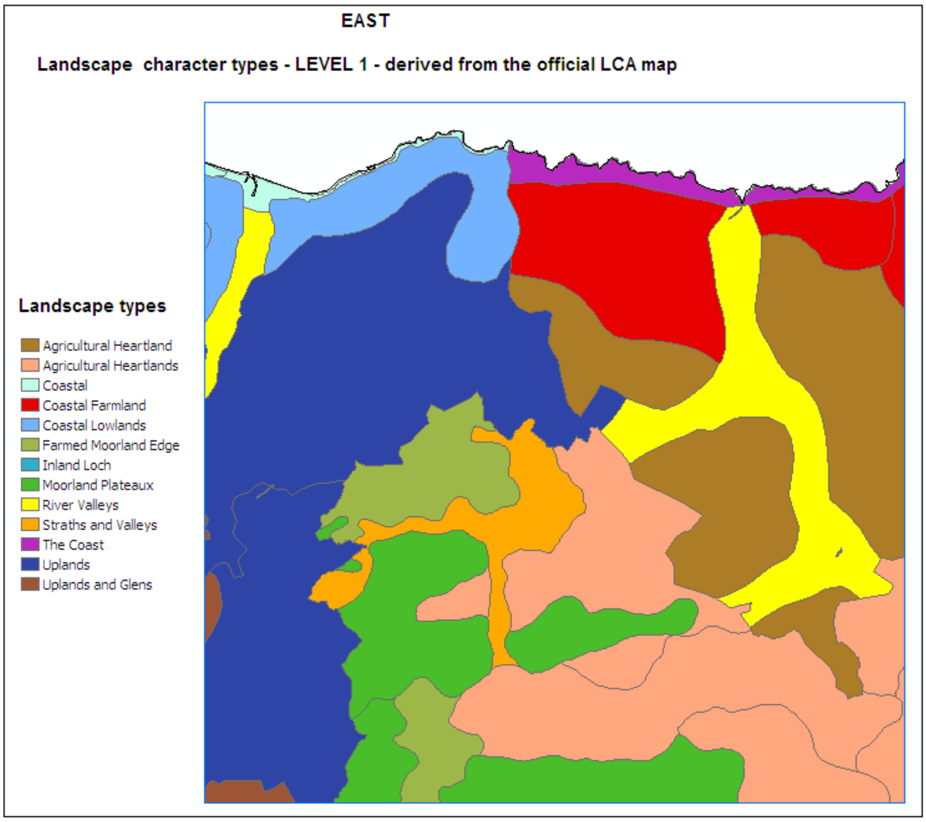


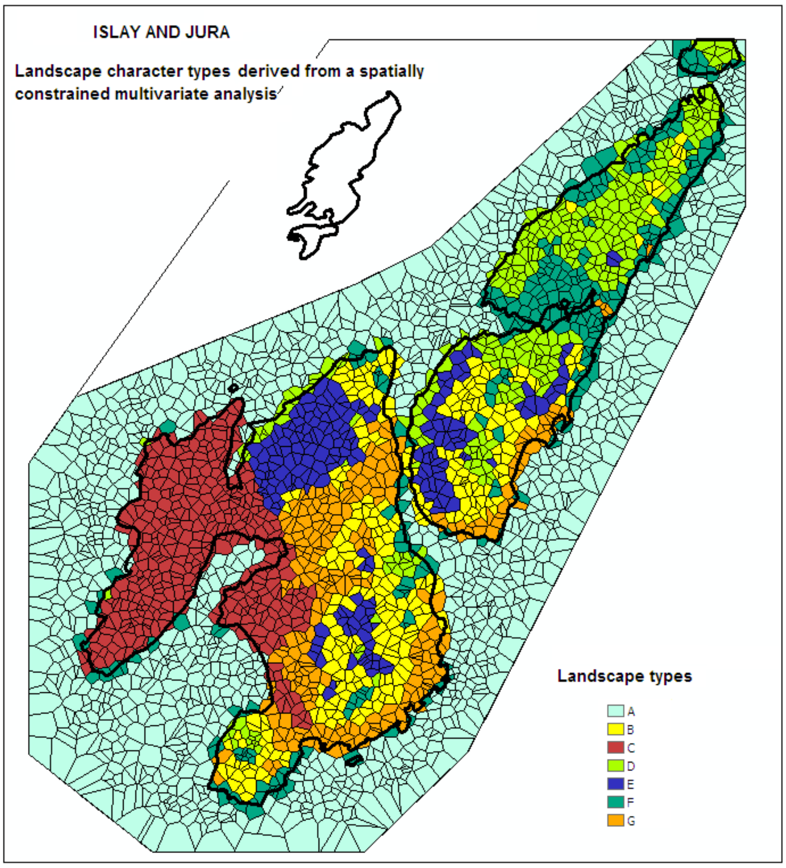
CENTRAL
Landscape character types derived from GIS and spatially constrained analysis
Landscape types
1
2
3
4
5
6
7
8
9
10



CENTRAL
Landscape character types derived by the Landscape Character Assessment map
Broad Valley Lowland
Coastal Flats
Coastal Hills
Coastal Margins
Drumlin Foothills
Farmed Moorland Hills
Farmed Strath Floor
Flat Arable Farmland
Forested Glen
Forested Moorland Hills
Forested Parallel Ridges
Forested Upland Hills
Highland Glens
Highland Summits and Plateaux
Igneous Hills
Inland Loch
Loch Island
Lowland Hill Fringes
Lowland Hills
Lowland Hills and Ridges
Lowland Hills and Valleys
Lowland Plateaux
Lowland River Valleys
Lowland Valley Fringes
Moss
Moss Farmland
Open Glen Side
Open Moorland Hills
Open Parallel Ridges
Open Upland Glen
Open Upland Hills
Plateau Farmland
Plateau Moorlands
River Valley Farmland
Rolling Farmland
Rugged Moorland Hills
Settled Loch Shore
Settled Strath Floor
Urban
Wooded Glen
Wooded Upland Hills


EAST
Landscape character types derived from a spatially constrained multivariate analysis
Landscape types
1A
2B
3C
4D
5E
6F
7G


EAST
Landscape character types - LEVEL 1 - derived from the official LCA map
Landscape types
Agricultural Heartland
Agricultural Heartlands
Coastal
Coastal Farmland
Coastal Lowlands
Farmed Moorland Edge
Inland Loch
Moorland Plateaux
River Valleys
Straths and Valleys
The Coast
Uplands
Uplands and Glens



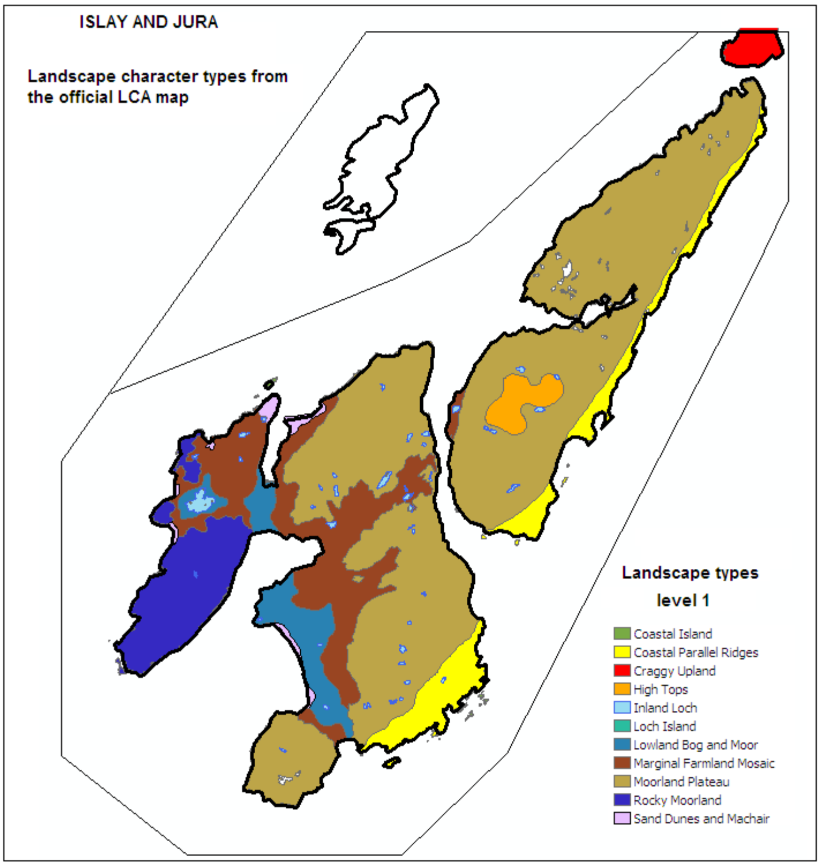


Figure 7.2. Pairs of maps of landscape character types obtained from GIS/MULTISPATI-based methodology and official LCA maps for each of the ten areas of study.

The statistical technique adopted was suggested by recent research carried out by S.Dray and T.Jombart (2010) who revisited an old study on moral statistics by Andre' Michel Guerry (1833, cited in Dray and Jombart, 2010) through the application of different multivariate analyses to the dataset. The idea at the basis of Dray and Jombart' s work wanted to demonstrate the efficiency of statistics able to perform spatially constrained multivariate analyses and the way they help to reveal the spatial structure of multivariate dataset considering simultaneously the influence of the geographical position of the data on the relationships amongst the data. For their case study the authors illustrated several techniques and included MULTISPATI and a between classes analysis (BCA)[29].

If MULTISPATI is best applied in order to identify landscape character types by detecting spatial clusters from a multivariate dataset, BCA performs a similar spatially constrained multivariate analysis by starting from a known spatial partition. In this case the spatial partition was represented by the types of landscape belonging to the original LCA maps and the task of the BCA was to group the 3000 polygons according to the known LCA types.

In brief, both MULTISPATI and BCA were performed and compared to each other and since both techniques offered the same typology of graphs it was relatively easy to analyse them in parallel. Figure 7.3 provides an example of graphs obtained by calculating BCA on South Shetland: the input data is identical to that used for MULTISPATI, but BCA uses the official landscape character types of LCA as spatial constrain, while MULTISPATI derives spatial patterns on the basis of spatial autocorrelation (quantified through the spatial weighting matrix, see chapter 6).

[29] Conceptually BCA is similar to a discriminant analysis. There is an exception though between a linear discriminant analysis and BCA, and namely that the latter is not influenced by collinearity among variables and it does not require the number of variables to be smaller than the number of observations (Dray and Jombart, 2009).

Therefore the aim is to compare the outcomes and calculate the degree of correspondence between the two analyses.

All the graphs of BCA are important and informative: the eigenvalues barplot (A) indicates that two components account for the overall variance of the data; the biplot entitled “canonical weights” (B) illustrates the distribution and relationships of the variables along the two main axes; the factorial map (C) shows the repartition of the polygons (these are the black dots from which the vectors start) into the official LCA character types; the second factorial map (E) represents the distribution of the LCA types along the first two axes and can be compared to the distribution of the variables in the biplot.

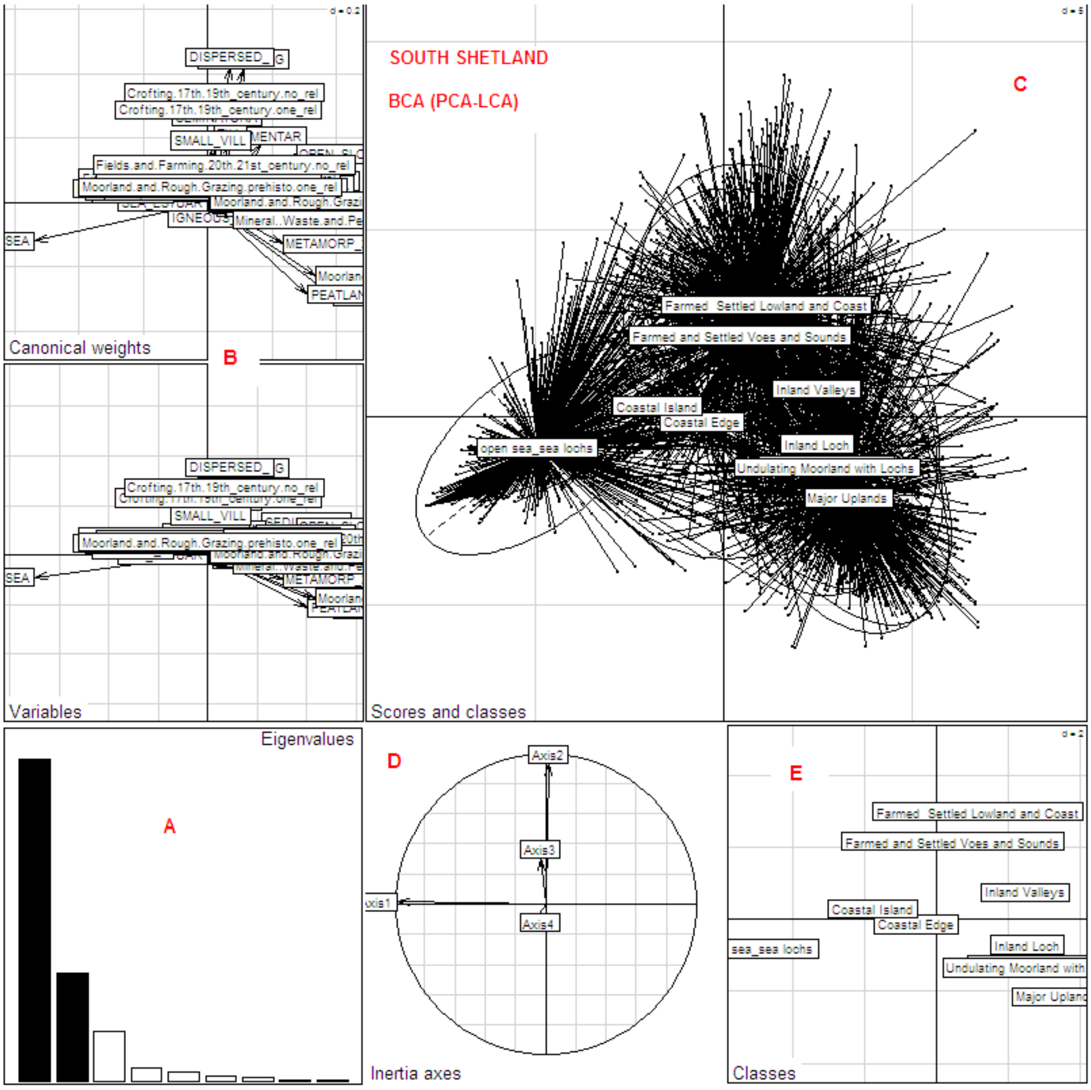


Figure 7.3. Summary of all the results obtained by running a between classes analysis.

As stated previously MULTISPATI and BCA can be easily compared since they present similar graphs; thus the first way of assessing the similarity between the LCA and GIS/MULTISPATI maps was to look at the biplots (figure 7.3 B). It emerged that they generally appear similar to each other, in terms that the variables (the landscape elements) are combined in the same way, despite their different location along the axes. For instance figures 7.4 (top and bottom), illustrate the biplots of South Shetland obtained with BCA and MULTISPATI analysis. It is clearly evident that the first two axes of the plot explain similar correlations amongst the variables despite these being rotated with a different angle. In both plots sea is separated from land, and this result was excellent and evidence that MULTISPATI correctly recognised Shetland as islands. Then, straths and plains are combined with improved and semi-natural grassland, dispersed houses, till and crofting (historic land-use); hill tops, glens and valleys and dwarf shrubs are associated together as well as peat, peatlands and wetlands and metamorphic Dalradian rocks.

Figure 7.4

A second way to measure the level of concordance between BCA and MULTISPATI analyses is to compare the factorial maps (graphs C in figure 7.3). The scores of the polygons on the first two axes of BCA and MULTISPATI are compared by using *procrustes* statistics. A value of 1 would indicate a perfect match between the factorial maps. The results from the *procrustes* statistics are then tested for significance through a randomisation test which indicates that the statistics are always significant with $p = 0.001$.

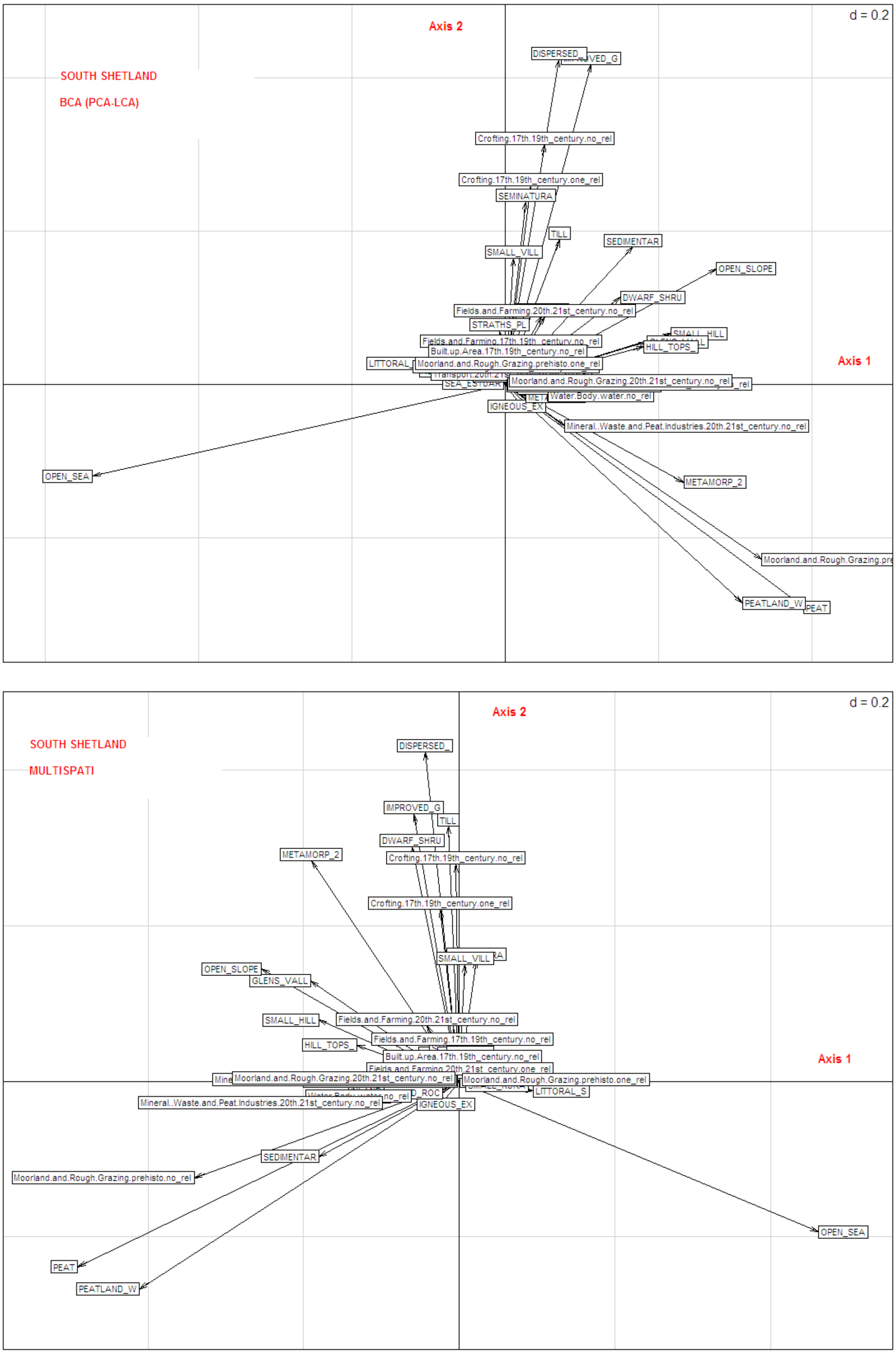


Figure 7.4. Biplots of South Shetland derived from the between classes analysis (BCA) (top) and MULTISPATI-PCA (bottom)

Table 7.1 represents the values of concordance between BCA and MULTISPATI analysis calculated with *procrustes* analysis.

areas of study	BCA vs MULTISPATI
N. UIST/S. HARRIS	0.993
NORTHEAST	0.998
CENTRAL	0.84
WEST	0.995
SOUTHEAST	0.955
SOUTHWEST	0.935
EAST	0.872
NORTHWEST	0.952
CNPA	0.887
SOUTH SHETLAND	0.994
ISLAY-JURA	0.96

Table 7.1. Results from the Procrustes analysis based on the comparison of the scores calculated with the between classes analysis and MULTISPATI.

From table 7.1 it is evident that the match between the results from BCA and MULTISPATI is very high overall. The reason is due to the fact that each area of study revealed data with strong spatial structures and high level of spatial autocorrelation already detectable with the standard PCA. Fundamentally after performing a standard PCA, BCA and MULTISPATI-PCA it emerges that, because of the significant influence of spatial autocorrelation, the results from the three analyses are similar. As reported in table 7.2, the *procrustes* statistics show that the values of concordance between the results from both standard PCA and BCA, and from standard PCA and MULTISPATI-PCA are also very high.

Due to the similarity between the results obtained from standard PCA and MULTISPATI-PCA, it is worth remembering why the latter statistics are applied in this research and are considered more advantageous. The main reason is that spatially constrained multivariate statistics focus on the spatial aspect of the data and are a more integrated and flexible approach which provides a new way of tackling complex datasets (Dray and Jombart, 2010).

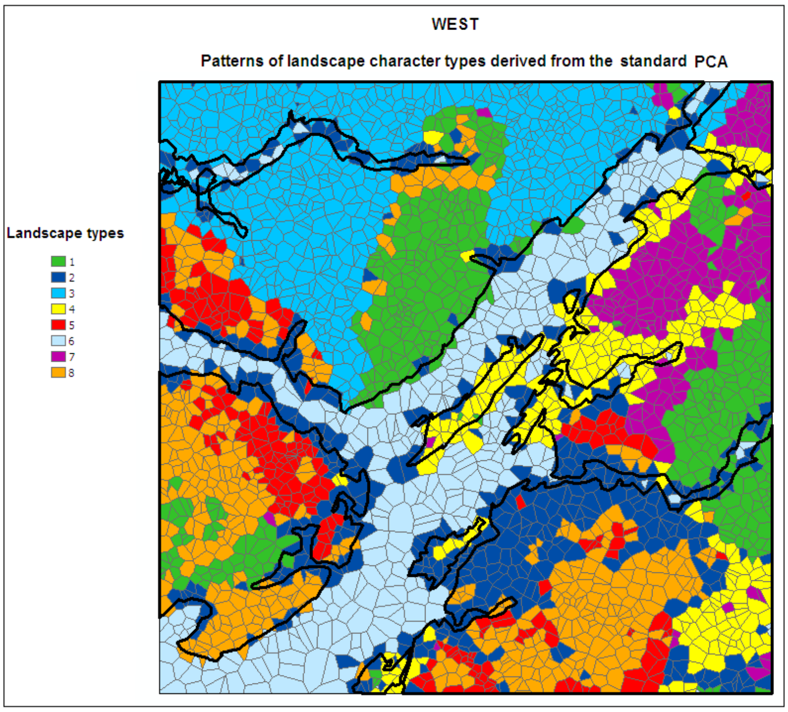
areas of study	PCA vs BCA	PCA vs MULTISPATI
N. UIST/S. HARRIS	0.992	0.997
NORTHEAST	0.997	0.999
CENTRAL	0.76	0.905
WEST	0.991	0.995
SOUTHEAST	0.952	0.999
SOUTHWEST	0.932	0.997
EAST	0.87	0.995
NORTHWEST	0.956	0.996
CNPA	0.828	0.935
SOUTH SHETLAND	0.938	0.942
ISLAY-JURA	0.948	0.996

Table 7.2. Degree of similarity between the values of standard PCA, BCA and MULTISPATI-PCA calculated by using *procrustes* statistics.

In contrast, a standard PCA is designed to maximise the variance of the variables analysed and to identify structure that may or not be spatially influenced. Thus, if spatial autocorrelation is weak, the spatial structure might be missed by using a standard PCA (Dray and Jombart, 2010). Consequently, in presence of large datasets and when the objective of the analysis is the detection of spatial patterns it would be preferable to use a spatially constrained multivariate statistics since they allow more accurate, clearer and smoother results (see figure 7.5).

Figure 7.5

A closer look at tables 7.1 and 7.2 reveals that some of the areas showed values slightly lower than the others, for examples Central, East and the Cairngorms. The last one was initially a surprise since the GIS and LCA maps visually seemed very similar to each other, but a closer examination at both maps allowed us to establish that in fact the GIS analysis was more detailed than the LCA since it highlighted differences in the type “Glens and uplands” that were entirely missed by the LCA. Indeed the GIS landscape character types need to be validated officially but informally they were considered sufficiently plausible by Matthew Hawkins (personal comment), the senior heritage manager of the Cairngorms National Park.



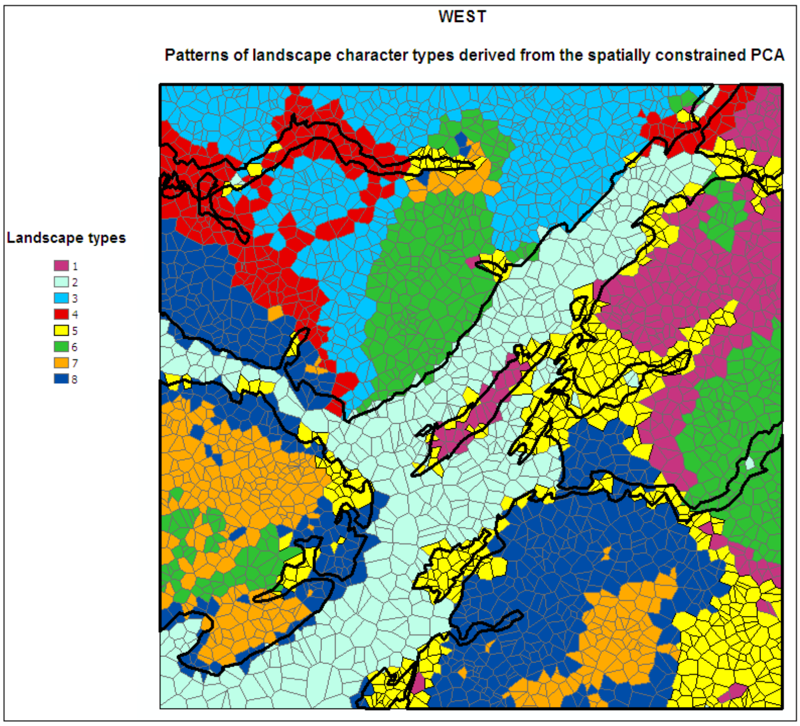


Figure 7.5. Comparison between the outcomes of a landscape characterisation carried out with a standard PCA and MULTISPATI-PCA In the example of West area of study, there are differences in the size and shape of some landscape character types, e.g. those in red, blue and orange. Overall the landscape character types obtained with MULTISPATI-PCA show a more compact and tidy form than those derived by standard PCA.

A legitimate/reasonable question can rise at this point. Why, despite the significance and the evidence of the statistics[30], some GIS maps appeared to be more similar to LCA maps than others?

Several answers can be found and they vary according to each area of study, however in general it is true that the level of similarity between the maps depends on the way the LCA was carried out by the landscape architects and on the complexity/simplicity of the landscape itself. North Uist/South Harris, South Shetland and the Cairngorms for example are areas where it is relatively easy to differentiate the landscapes because their differences are remarkable. If we wander in North Uist and South Harris for example the contrast between the machair, the moorland punctuated by island lochs and rocky hill tops is clear. Similarly, the flat fertile lands and the voes of South Shetland contrast with the undulating peaty hills, and in the Cairngorms, the Plateau is a distinct natural environment, with its corries and the arctic-like vegetation that make it unique if compared with the inhabited and forested straths.

Another point to bear in mind while looking at the LCA and GIS maps is to think of the way the landscape assessment was undertaken within the current LCA. As mentioned in chapter 2, the visual assessment which took place during the field survey has a great impact in the definition of the character types. The landscape architects in fact tend to interpret personally the official LCA guidance and to focus on highlighting the key characteristics[31] of a landscape. In contrast, within this methodology, GIS assume that all the elements in the landscape have the same importance, and analyse them according to the way they combine and associate to each other. However, GIS depend on the quality and quantity of information in input; therefore if some relevant data is missing then GIS carry a "partial" or skewed landscape characterisation.

Finally, since the LCA starts with a clear and defined reason, this is likely to affect the degree of detail in drawing the boundary of the landscape character types.

---

[30] The *procrustes* statistics prove that BCA and MULTISPATI describe the same spatial patterns.

[31] Key characteristics are defined as specific associations of elements that were relevant and important for the definition of the character of the landscape.

Alternatively, the level of detail achievable with GIS is provided by the number of polygons generated at the beginning of the analysis and the resolution/scale of the maps.

To sum up, two main results can be gathered from a first inspection of the GIS landscape character type's maps and their comparison to the official LCA ones. First of all the objectivity of the statistics should be preferred to the subjectivity of the visual interpretation of the maps, and the results from the statistics confirmed that despite the appearance there is a numeric correspondence between GIS and LCA maps. Secondly the reasons why LCA maps differ from GIS maps are numerous and above all vary from area to area according to the assessment carried out by the architects.

On the basis of the statistics it is possible to claim that in general there is correspondence between GIS and LCA maps, which is a positive result because it indicates that overall the GIS methodology works and is able to identify landscape character types with credibility. In parallel, the research showed that there are cases where a further investigation on the datasets and the scale to be used during the analysis is recommended and should be taken into account in order to improve the GIS/MULTISPATI-based landscape characterisation and the final maps.

The observations and comments carried out so far allowed us to explain and gain awareness of the meaning of similarity/discrepancy between the GIS and the LCA maps. However, no mention has been made in order to let people understanding how to retrieve information about the landscape elements composition of the landscape character types depicted in the GIS maps. Hence the way of interpreting and reading the information conveyed in the maps is described below and different areas of study are taken each time as explanatory examples.

7.2 How to read a GIS/MULTISPATI map.

7.2.1 North Uist and South Harris: an example of similarity between LCA and GIS/MULTISPATI landscape maps.

Commonly, and without an explicit map of the landscape character types, the way of obtaining information on the landscape of a given area of study is to collect maps depicting the physical and cultural aspects of the landscape and try to summarize their information. In their simplest application, GIS allow us to retrieve a general idea of the landscape of a given region through two basic operations on digital maps: quantification and overlapping. The former is achieved by calculating the area covered by each landscape element in terms of percentages whereas the latter is obtained by superimposing, relating and querying the maps so that their information is accessed jointly. However, the summary of the elements in each map provide a table that is informative and a valid support for a general description of the landscape, but does not show where the patterns occur (see figure 7.6).

UIST LANDCOVER	%	% NO SEA	UIST LANDFORMS	%
open sea	72.18		straths/plains	37.35
dwarf shrub heath	11.37	40.86	open slopes	23.92
supra littoral rock sediments	3.44	12.35	glens/valleys	20.25
seminatural natural grassland	2.50	8.98	small hills	9.29
abandoned and derelict grassland	2.38	8.57	hill tops/high ridges	9.19
inland water	1.89	6.80		
sea/estuary	1.78	6.42		
improved grassland	1.50	5.39	UIST BEDROCK	%
inland rock	1.48	5.34	metamorphic lewisian complex	96.43
wetlands peatlands	1.27	4.56	water	3.33
conifer wood	0.17	0.60	igneous intrusive	0.25
broadleave wood	0.02	0.07		
built-up areas	0.02	0.06		
UIST SETTLEMENT TYPES	%		UIST SUPERFICIAL	%
dispersed_houses	45.30		peat	65.83
villages	26.00		blown sand	13.06
fringes	13.70		till	11.13
small_villages	11.74		unknown lithology	8.69
small_rural_towns	3.26		alluvium	1.29

Figure 7.6. Summary of the landscape elements that form North Uist and South Harris. Each map has few dominant variables, for instance the metamorphic Lewisian complex constitute 93% of the geological bedrock of the islands while peat represent 65.83% of the superficial

deposits. The historic land-use data is not included since at the time of investigation the HLA map was not completed. Currently the data is available.

While the overlapped maps help to understand where and which elements are associated to each other but do not explicitly describe a pattern (see figure 7.7). In addition the analysis and readability of more than two maps becomes complicated due to the numerous ways the elements combine to each other.

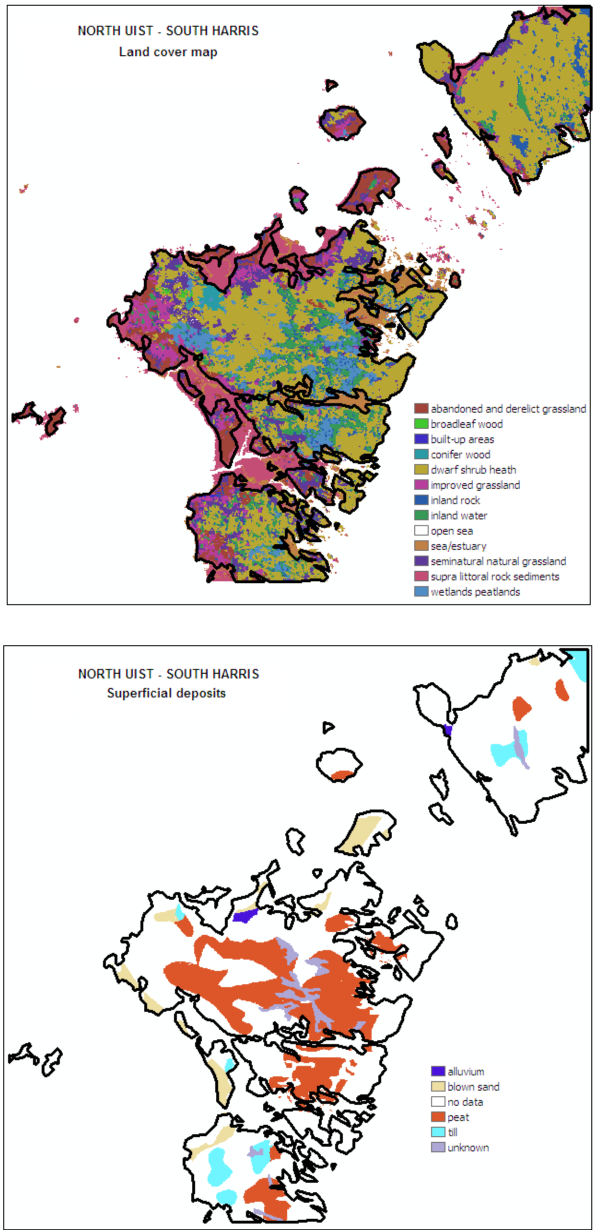


Figure 7.7. Land-cover (top) and superficial deposits (bottom) maps used during the landscape analysis. By overlapping and relating each feature of both maps it is possible to select an area

and retrieve integrated information from both maps. For example the combination of dwarf shrub heath and inland water match largely with peat in the centre of north Uist and in South Harris while seminatural and natural grassland are more related to blown sand and this combination is found particularly along the coastline.

The novelty introduced by the synergy between GIS and MULTISPATI is that the structure and patterns within the data are clearly and explicitly summarised, retrieved and displayed as illustrated in figure 7.8.

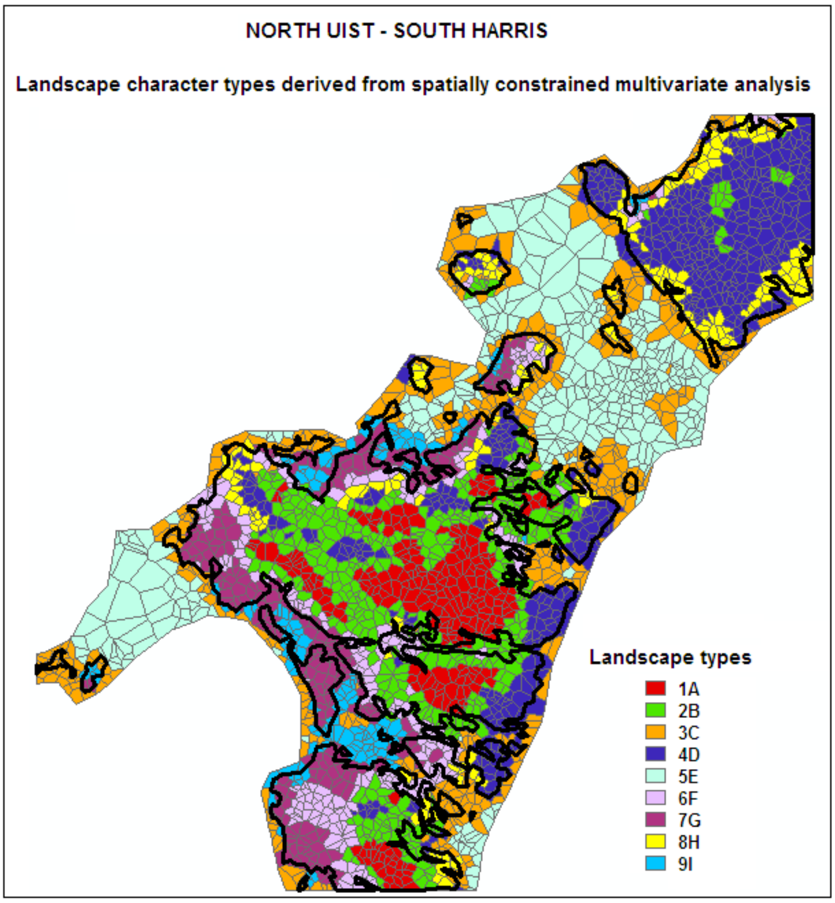


Figure 7.8. Map of the landscape character types of North Uist and South Harris as obtained from the application of GIS/MULTISPATI analysis.

There are two ways to understand the meaning of the nine classes of landscape character types. The first one is to read the pie charts obtained by calculating the percentage of area covered by each landscape element within a landscape type. For instance the pie charts of the landscape types coded 5E and 7G, both depicted in Figure 7.9, inform us that the former landscape is mainly represented by "open sea", while the latter is characterised by a combination of straths, metamorphic Lewisian rocks, abandoned, improved and semi-natural grassland, blown sand and inhabited areas.

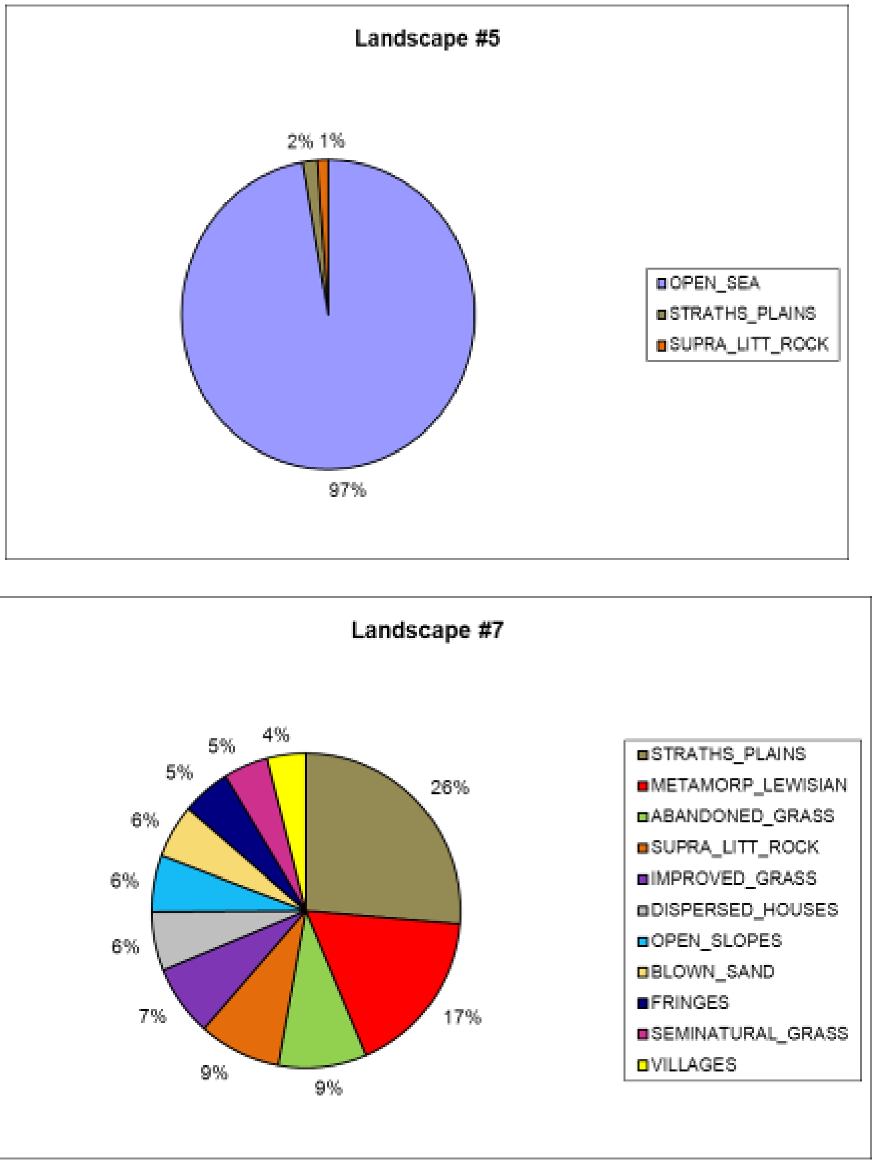


Figure 7.9. Pie charts generated by calculating the percentage of area covered by each landscape element within the boundary of a landscape type. This calculation was performed through the spatial analyst tool of arcGIS called "cross tabulate area"**.**

Moreover it might be interesting to compare two neighbouring landscape types in order to highlight/reveal which elements contribute to differentiate them. For example here the landscape type 7G is compared to the 6F and it stands out a change amongst the land-cover types (see figure 7.10): in fact in the landscape 6F dwarf shrubs and inland water replace blown sand. The same type looks also more undulating, since the presence of open slopes increases from 6% to 12%. The conclusion is that the landscape type 6F, compared to 7G, is hillier, covered by heather and interspersed by lochs.

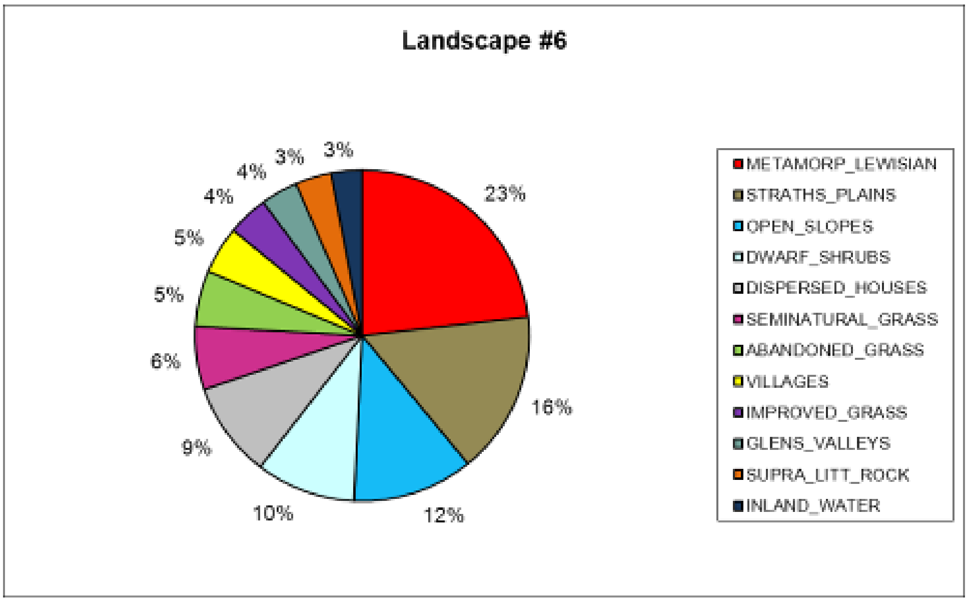


Figure 7.10. Pie chart of the landscape type 6F

A method to double check that the differences observed between types 6F and 7G occurr in reality is through direct ground truthing, however if the area of study is not easily accessible it is possible to rely on satellite/aerial images which can be downloaded from the web. In this case a picture from Google Earth of the area in the red square (see figure 7.11) is downloaded and used in comparison with the GIS map. To the naked eye it is already possible to spot the difference in land-cover between the two landscape types and the aerial picture effectively helps to validate the accuracy of the pie charts.

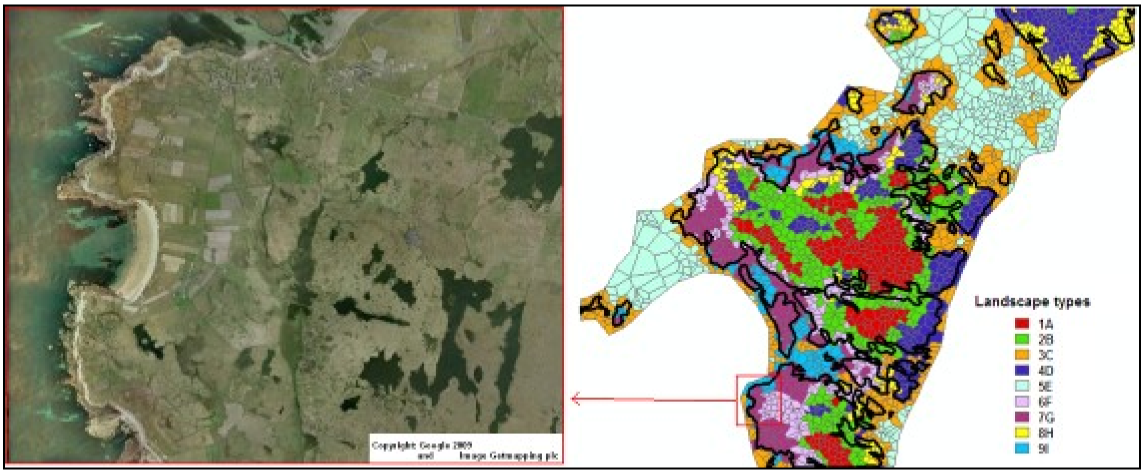


Figure 7.11. Aerial picture of the area in the red square in North Uist and South Harris landscape character types map**.** The aim was to verify that the information from the pie charts on differences in land-cover between the types 6F (dark pink) and 7G (light pink) was accurate. The line on the aerial image helps to identify the boundary between the two landscape types.

The LCA provides a detailed description and classification of landscape character types which can be achieved also through GIS. As demonstrated, the pie charts offer the opportunity to read the composition of the landscape types so that for example the landscape 7G can be defined as an area largely flat, with 26% of straths and plains and 6% of open slopes, where the geology is dominated by the metamorphic lewisian complex. The presence of settlements is also relevant since dispersed houses, fringes and villages occur in percentages close to 6%, 5% and 4%. The land cover is mainly characterised by grassland (abandoned, improved and semi-natural) and blown sand that, with littoral and supra littoral rocks, suggests being in proximity to the sea.

After describing the landscape types it may be opportune to classify them so that the meaningless code (number + alphabetical letter) can be redefined into a more explicit class. Because the LCA already provides classes of landscape character types it is possible to borrow them and associate them to the most suitable/opportune GIS types as illustrated in table 7.3.

Code GIS landscape type	Description
1A	Boggy moorland with lochans
2B	Boggy moorland
3C	Coastal flats and rocks
4D	Mountain massif, rocky moorland and lochans
5E	Sea
6F	Crofting and farming
7G	Machair
8H	Crofting "3"
9I	Sea estuary and supra littoral rocks

Table 7.3. The table provides an example of how the official classification of landscape types can be helpful in order to classify the GIS types too. In the case of South Harris and North Uist finding the most suitable classes for the GIS types was a relatively easy task. There might be cases that require variation to the original LCA classes and adaptation to them in order to match properly to the GIS type.

Along with the pie charts and the support from Google Earth there is another robust way to read the results from the GIS map, which is referring to the biplots (described in the previous chapter) and the factorial maps obtained from MULTISPATI. Figure 7.12 shows the biplot obtained from MULTISPATI analysis for North Uist and South Harris. The two axes in the plot indicate the first and second component of the PCA and correspond to the largest fraction of the overall variability of the data; therefore the variables that are located along these components are highly significant and explain the structure and composition of the data.

In figure 7.12 for example it is evident that the first component opposes clearly open sea to the rest of the variables while the second component separated straths and plains, abandoned and improved grassland, the settlements types from hill tops, small hills, glens and valleys and dwarf shrubs. Finally peat, inland water, wetlands and peatlands stand out as another correlated group of variables. It is important to mention that the variables close to the centre of the graph account very little for the variance expressed by the first and second components.

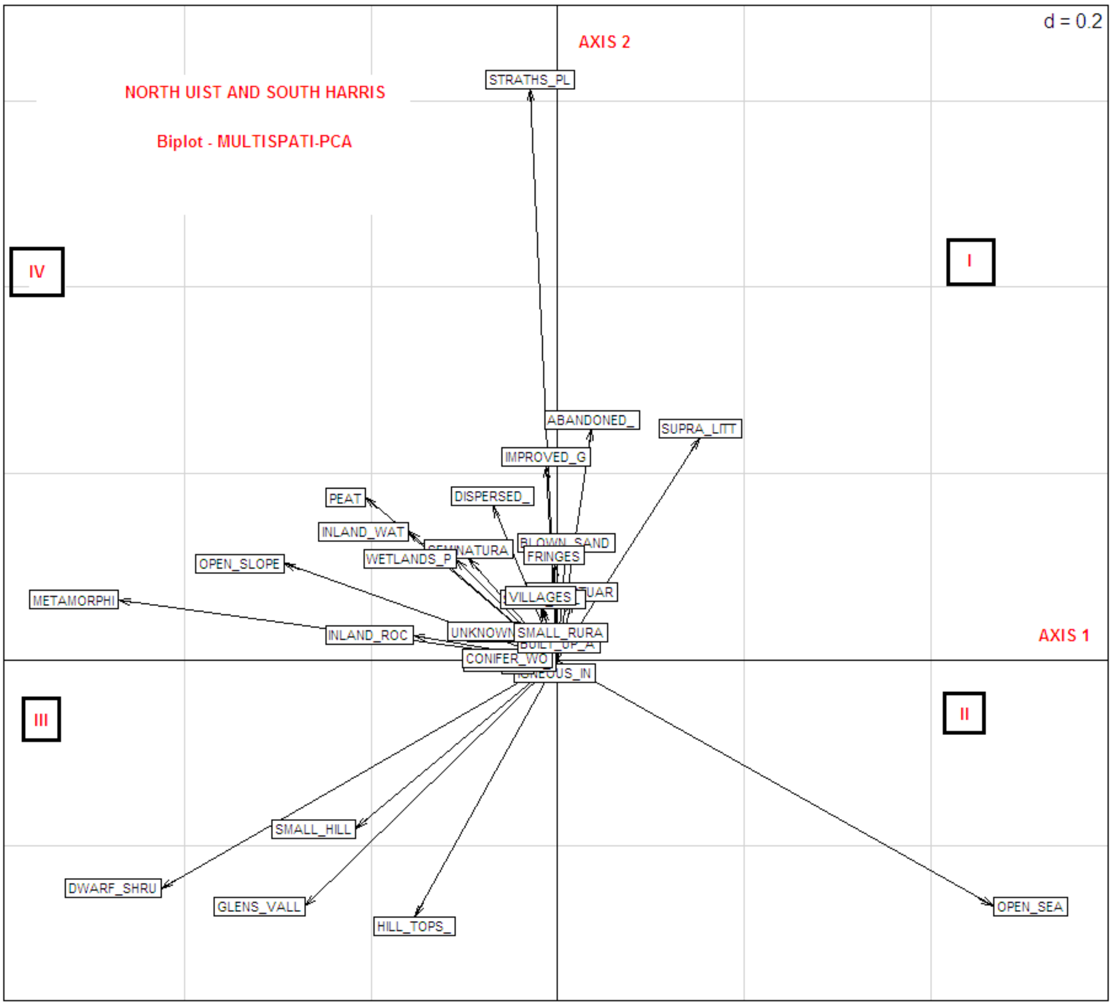


Figure 7.12. Biplot of the variables along the first and second axes of spatially constrained PCA. The black squares contain numbers identifying the quadrants of the plot.

In addition the biplot informs on the degree of correlation between the variables, for instance here open sea stands alone and is not correlated to the other variables, while hill tops and straths are clearly opposite to each other and show correlations with two different set of variables (see quadrants IV and III in figure 7.12).

Overall the location of the variables and their relationship are quite neat and defined and let us suppose that North Uist and South Harris are characterised by a strong spatial structure. Five large clusters could be identified: open sea; hill tops/glens/small hills and dwarf shrubs constituted by metamorphic Lewisian rocks; straths and open slopes formed by peatlands/wetlands; improved grassland and straths characterised by dispersed houses and blown sand; sea estuary formed by littoral rocks and abandoned grassland.

From these five broad groups, MULTISPATI was able to identify nine more specific landscape character types and the factorial map in figure 7.13 illustrates the way the 3000 polygons distributed along the first two axes of PCA. A series of ellipses, which identify the landscape character types, helps to see more clearly the clustering performed by MULTISPATI.

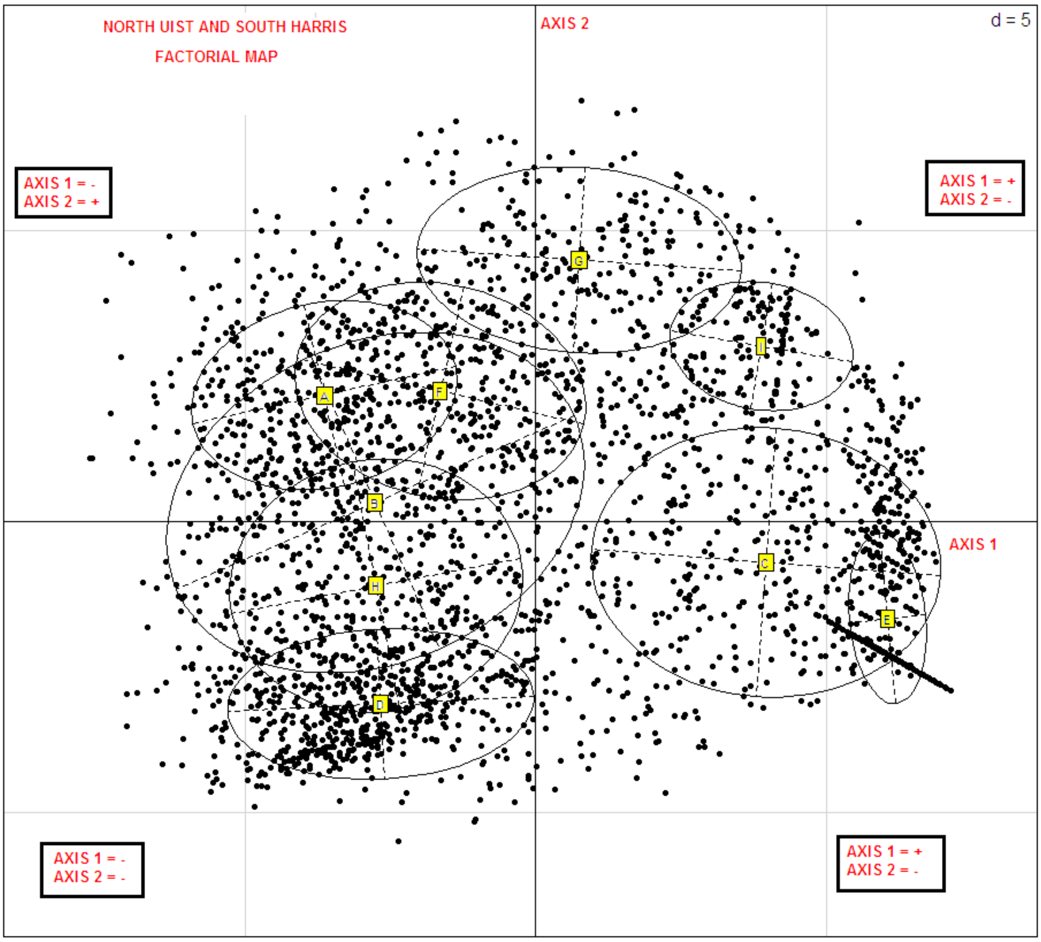


Figure 7.13. Factorial map obtained with MULTISPATI. The black dots correspond to the polygons observed during the analysis, while the ellipses indicate the landscape character types, whose ID code is highlighted in yellow. The factorial map is helpful in order to have an idea of the way the polygons are distributed along the axes of the PCA and how they are grouped into character types. In the black squares there is a reminder of the signs associated to the PCA's axes.

The ellipses compute the mean, variance and covariance of each group of black dots (the 3000 polygons in the map) on both axes of PCA. The centre of the ellipse corresponds to the average of the values assumed by the observed 3000 polygons, the two axes of the ellipse represent the two first principal components for each group, the slope of the ellipse's axes (which is the ratio between the major axis and the minor axis) is given by the value of covariance. With this in mind it is clear that the shape and the size of the ellipses reflect the values assumed by the mean, variance and covariance and inform us about the spatial structure of the data. For example if an ellipse presents an elongated (stretched) shape then it indicates that the polygons in the ellipse are formed by variables that are correlated to each other and show similar characteristics. On the contrary if an ellipse tends to resemble a sphere then the polygons are constituted by variables with weak correlation and different characteristics. In brief, compact ellipses reveal strong and defined landscape types, while large ellipses are indication of landscape types with a less distinctive character.

The location of the ellipses in one of the four quadrants of the factorial map is linked to the variance expressed by the groups along the first two PCA axes, as showed in table 7.4.

	Axis1	Axis2
A	-3.63	2.17
B	-2.75	0.32
C	3.96	-0.70
D	-2.66	-3.13
E	**6.05**	**-1.66**
F	-1.63	2.24
G	0.74	4.50
H	-2.74	-1.09
I	3.88	3.02

Table 7.4. Values of variance expressed by the 9 landscape character types calculated with MULTISPATI- PCA. The location of the landscape types in figure 13 is given by the numbers in column. Axis 1 corresponds to values on x axis and axis 2 refers to the values on y axis.

With reference to figure 7.13 where the positive or negative signs of the axes are marked in red, it is possible to understand better the content of table 7.4. For instance the landscape character type “E” is located in the second quadrant since the correspondent values of variance are 6.0 for axis 1 and -1.6 for axis 2.

The factorial map can be compared to the biplot map in figure 7.12 and by looking at both graphs it is possible to describe, in general terms, the landscape element composition for each ellipse (thus for each landscape type). In the case of South Harris and North Uist it is possible to perform the following matches:

1. landscapes types E (sea), I (sea estuary) and C (coastal flats and rocks) were represented by open sea, supra littoral rocks and sea estuary;
2. landscape type D (mountain massif and rocky hills) matched with the variables hill tops, dwarf shrubs, glens, small hills and metamorphic rocks;
3. landscape type A (boggy moorland with lochans) corresponded to peat, inland water, open slopes, peatland/wetland, inland rocks;
4. landscape types F and G were described by straths and plains and then respectively F (crofting and farming) was characterised by peat, inland water, seminatural grassland, dispersed houses while G (machair) looked to be better represented by abandoned and improved grassland, blown sand, villages and dispersed houses;
5. landscape types B (boggy moorland) and H (crofting 3) showed very large ellipses at indication of a high variability of values. In this case B and H represented transitional landscapes.

Reading the factorial map together with the biplot allows double checking the information contained in the pie charts and confirm that the descriptive classes borrowed from the LCA in order to classify the GIS types are overall appropriate.

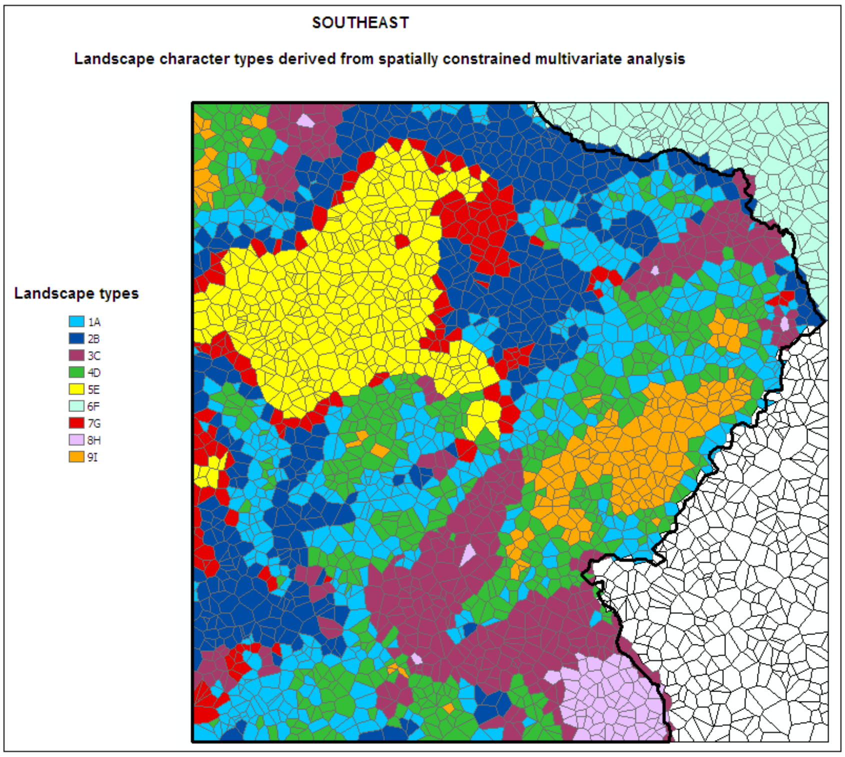
7.2.2 The Southeast: an example of greater discrepancy between LCA and GIS/MULTISPATI landscape maps.

North Uist and South Harris showed a simple landscape and the GIS/MULTISPATI map was very similar to the official LCA map so that it was possible to borrow the classes of LCA types in order to classify the GIS types. Nevertheless during the analyses there were situations, like the Northwest, Southeast, Central, Southwest and West areas, where the discrepancies between the official LCA maps and the GIS based maps were more evident and the number of landscape types more numerous than that retrieved with the GIS methodology. Hence for these areas there was uncertainty in deciding which LCA types corresponded to the GIS types and could picture it properly. For example figure 7.14 illustrate the GIS and LCA maps of the Southeast area.

Figure 7.14

Here, the difference in the number of landscape types is remarkable and before thinking of a classification of the GIS types it is reasonable to wonder whether or not the nine GIS types were accurate enough to describe the landscape of the area of study.

The decision about the number of GIS types takes place during the cluster analysis, which is a delicate phase of the statistical computations because of the degree of arbitrariness introduced into the analysis. As mentioned in chapter 6, the task of the GIS analyst is to decide about the number of clusters to retain on the basis of the results displayed in the dendrogram.



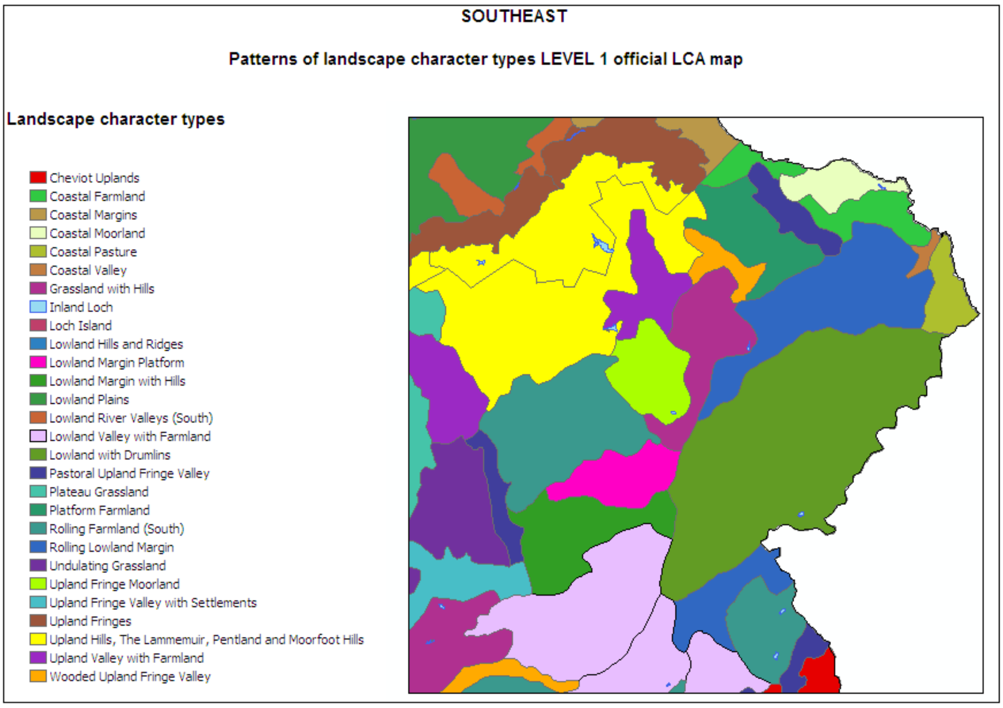


Figure 7.14. Maps of landscape character types derived from GIS/MULTISPATI analysis (top) and traditional LCA approach (bottom).

For instance figure 7.15 illustrates the dendrogram of the Southeast area. As noticeable the clusters, which are the landscape character types on the map, are well differentiated from each other and the length of the stems of the dendrograms (the black lines joining the groups) indicated a very large dissimilarity between the clusters.

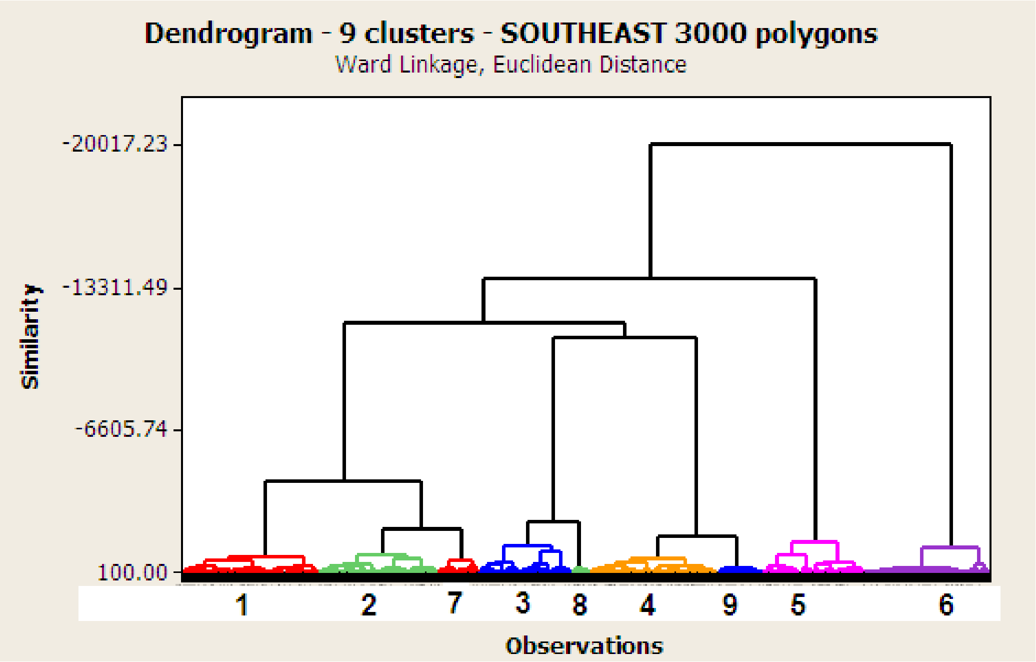
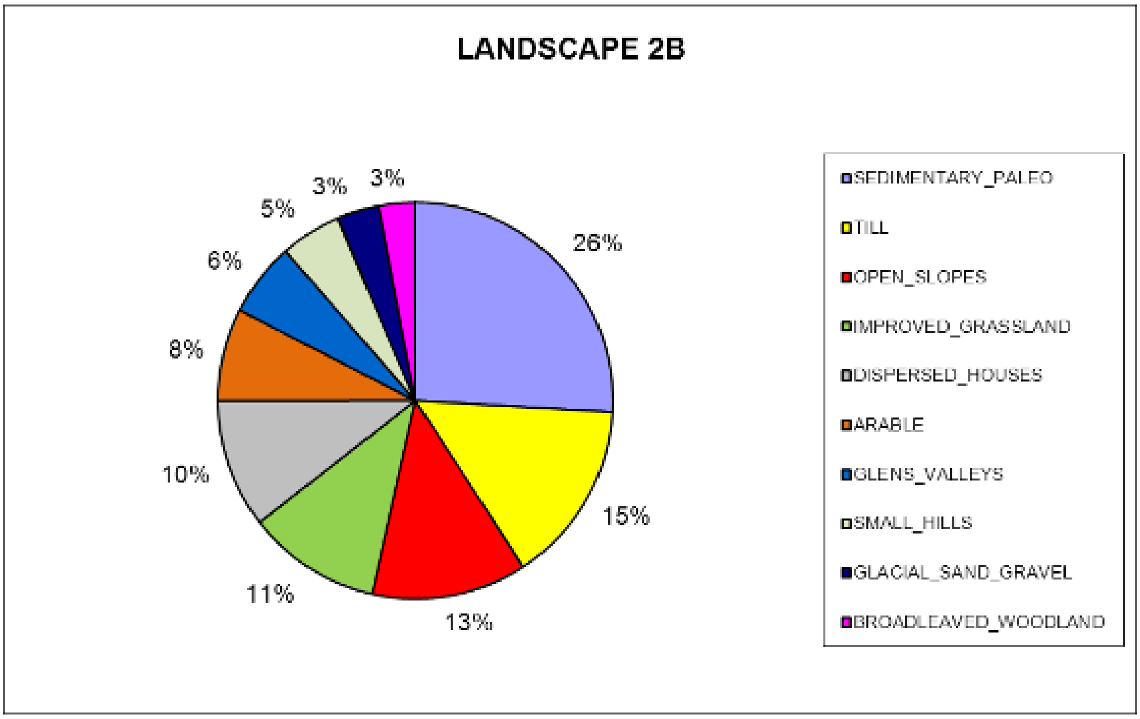
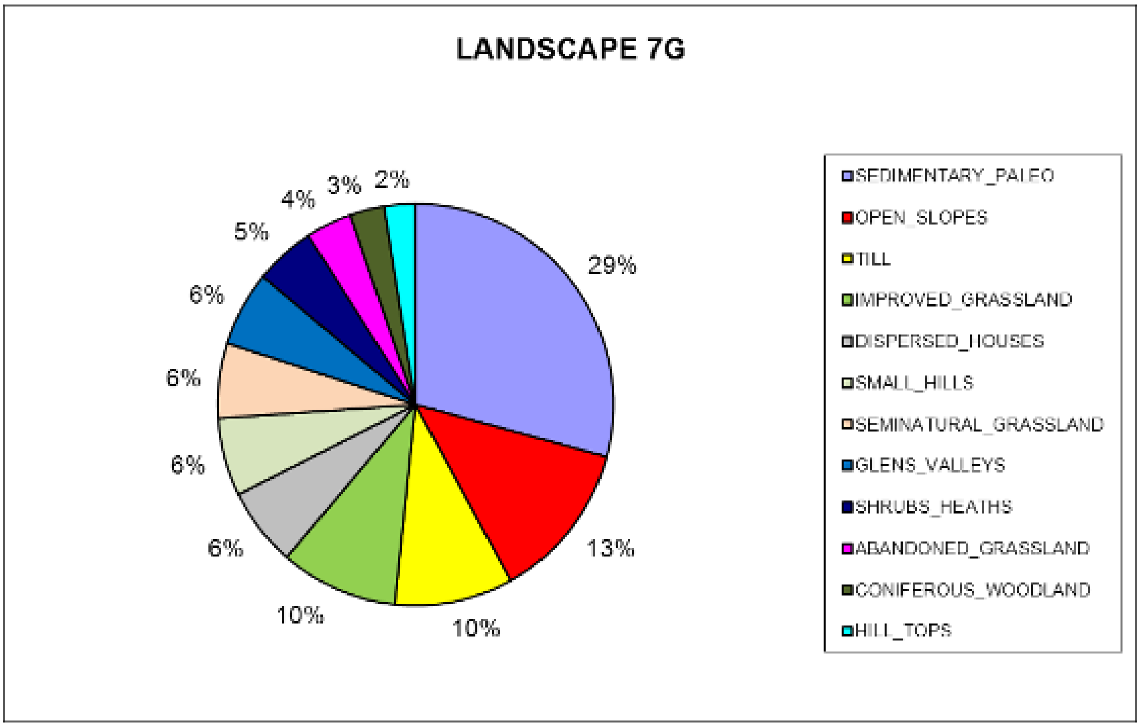
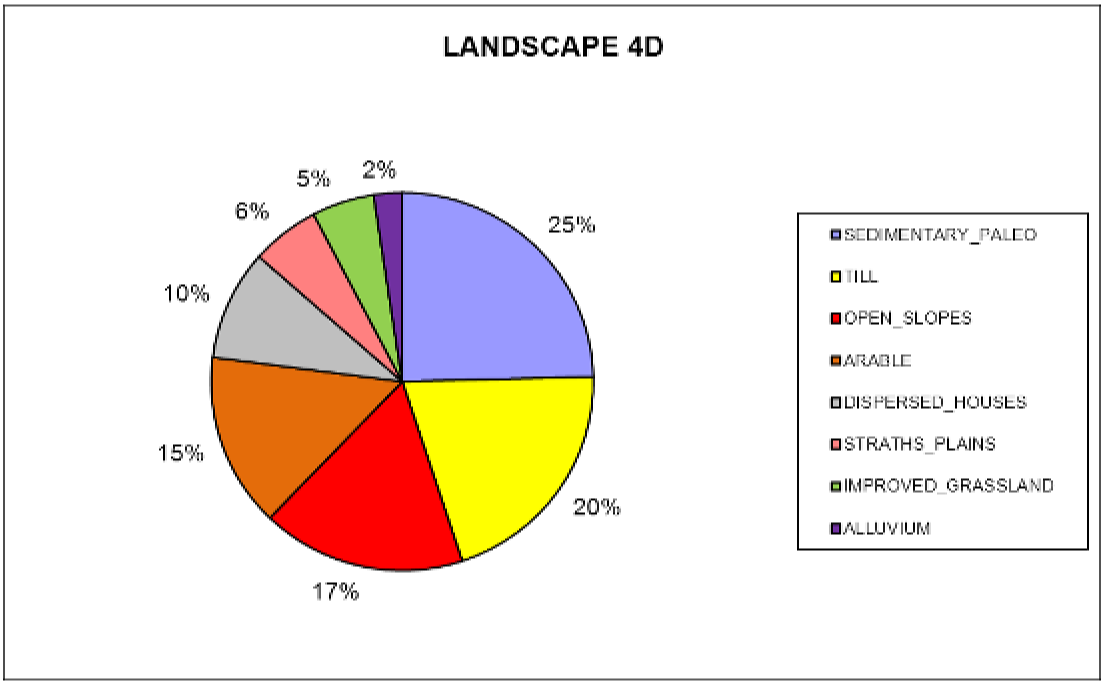


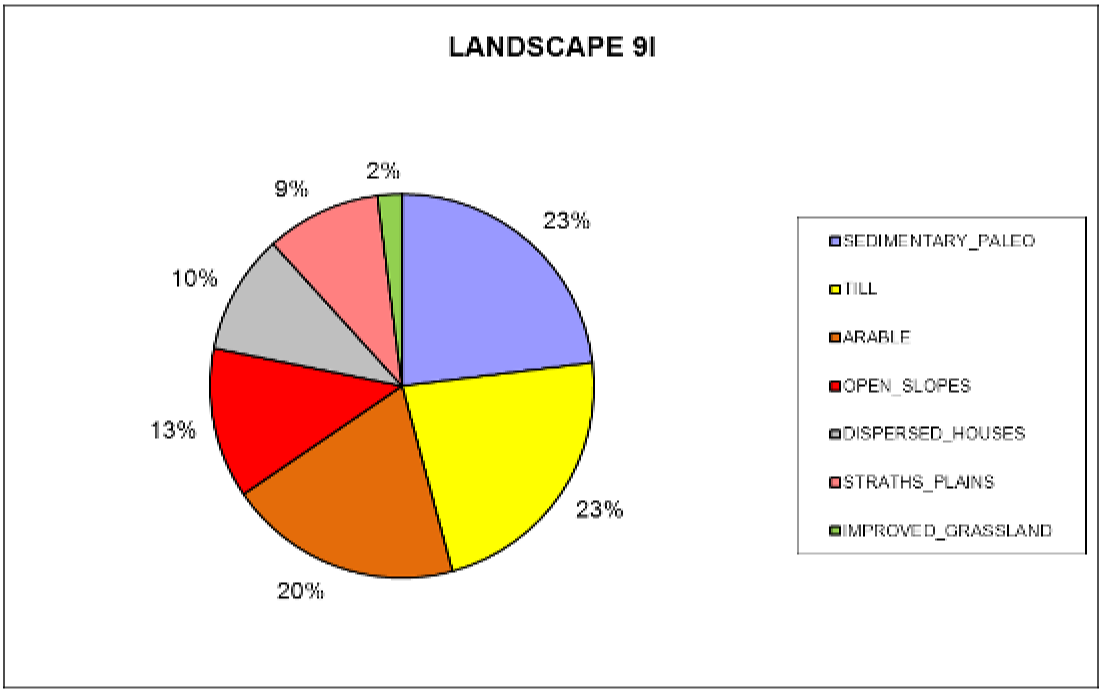
Figure 7.15. Dendrogram obtained after the cluster analysis performed on the results of MULTISPATI. The numbers from 1 to 9 identify the clusters which correspond to the landscape types on the map in figure 7.14a.

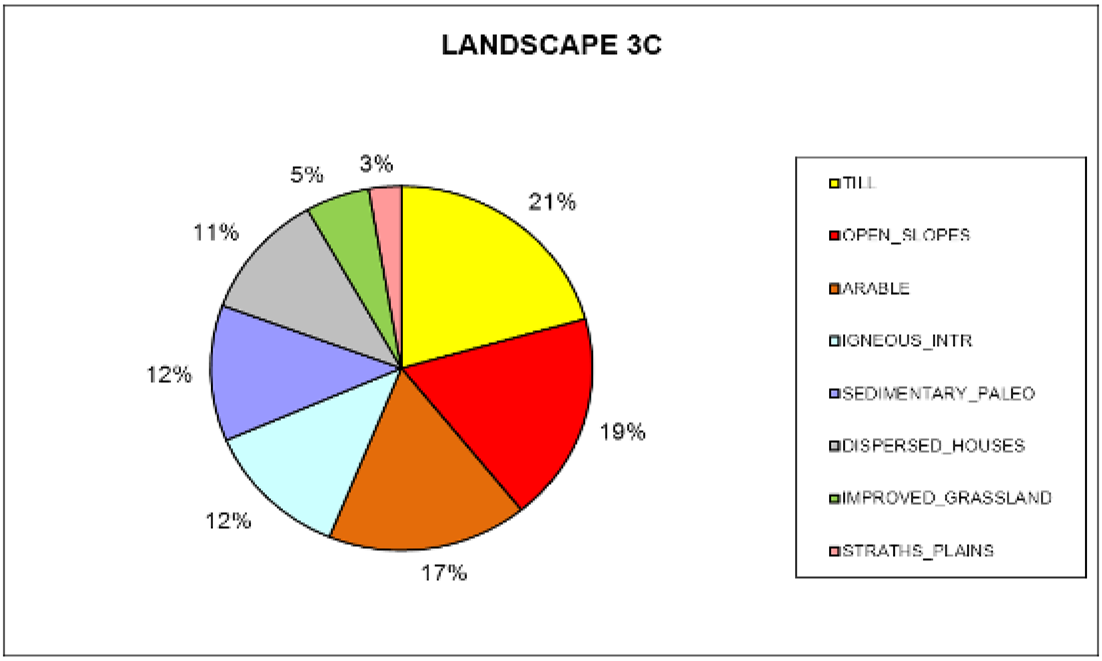
The dendrogram suggests that the 9 clusters are sufficient to describe the landscape character of the area of study and possibly a certain degree of similarity may be found between clusters 2 and 7, 3 and 8 and 4 and 9 since they are linked by shorter stems. In order to be sure that a further clustering would not cause a loss of information but would group landscape character types that effectively are similar, it is necessary to look and compare the pie charts of the GIS types 2B and 7G, 3C and 8H and 4D and 9I depicted in figure 7.16.


LANDSCAPE 2B
26%
15%
13%
11%
10%
8%
6%
5%
3%
3%
SEDIMENTARY_PALEO
TILL
OPEN_SLOPES
IMPROVED_GRASSLAND
DISPERSED_HOUSES
ARABLE
GLENS_VALLEYS
SMALL_HILLS
GLACIAL_SAND_GRAVEL
BROADLEAVED_WOODLAND


LANDSCAPE 7G
29%
13%
10%
10%
6%
6%
6%
6%
5%
4%
3%
2%
SEDIMENTARY_PALEO
OPEN_SLOPES
TILL
IMPROVED_GRASSLAND
DISPERSED_HOUSES
SMALL_HILLS
SEMINATURAL_GRASSLAND
GLENS_VALLEYS
SHRUBS_HEATHS
ABANDONED_GRASSLAND
CONIFEROUS_WOODLAND
HILL_TOPS


LANDSCAPE 4D
25%
20%
17%
15%
10%
6%
5%
2%
SEDIMENTARY_PALEO
TILL
OPEN_SLOPES
ARABLE
DISPERSED_HOUSES
STRATHS_PLAINS
IMPROVED_GRASSLAND
ALLUVIUM


LANDSCAPE 9I
23%
23%
20%
13%
10%
9%
2%
SEDIMENTARY_PALEO
TILL
ARABLE
OPEN_SLOPES
DISPERSED_HOUSES
STRATHS_PLAINS
IMPROVED_GRASSLAND



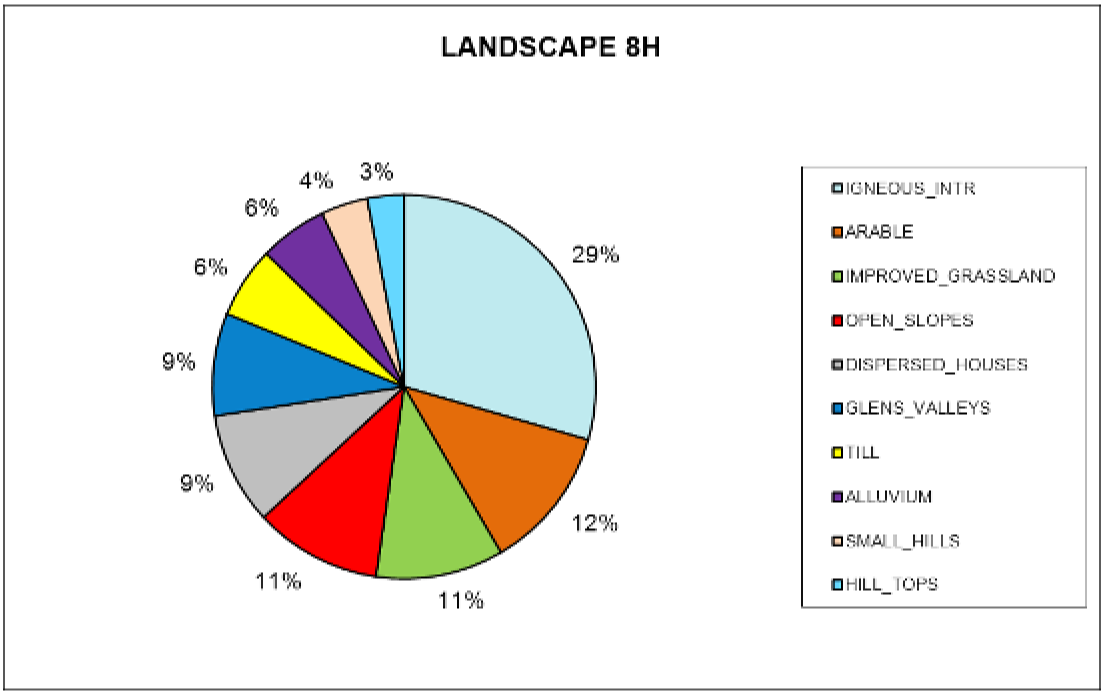


Figure 7.16. Pie charts illustrating the landscape elements composition of the landscape types 2B-7G, 4D-9I, 3C-8H that the dendrogram suggests having a certain degree of similarity. The pie charts are consulted in order to verify whether or not this similarity is real.

The comparison is extremely interesting: as far as the landscape elements are concerned, several differences are recorded between the three pairs of pie charts but, interestingly, these follow a different pattern. In fact types 2B and 7G show a complete correspondence of the first five landscape elements which accounted for 70% of the landscape, but the remaining 30% in type 2B comprises of arable fields, glacial sand and gravel and broadleaved woodland, while in type 7G it is formed by semi-natural and abandoned grassland, shrubs/heaths and coniferous woodland. A similar pattern occurs between types 4D and 9I where 98% of the landscape is described by the same elements, and only 2% of alluvium (in type 4D) marks the difference between the two types.

In contrast, the pie charts of type 3C and 8H are quite different and the geology, a mix of sedimentary paleozoic and igneous rocks in type 3 compared to the only igneous rocks in type 8, marks the boundaries between the two types. As visible from the geological map in figure 7.17, the whole Southeast area is mainly dominated by the sedimentary paleozoic rocks and regions of igneous intrusive and extrusive rocks are found particularly in the southeast corner of the area of study, where the type 8H is identified.

Figure 7.17

The main conclusion from the pie charts analysis is that types 2 and 7, and types 4 and 9 could be further aggregated, while types 3 and 8 should be kept separated since GIS/MULTISPATI landscape analysis is able to differentiate them on the basis of the geological bedrock.

As previously seen, the statistics can also help to understand the landscape type composition and the data structure of the landscape types. In the case of Southeast the combined analysis of the factorial map and the biplot in figures 7.18 and 7.19 confirmed that a further clustering could be attempted. It was noted that in general the information from the factorial map matched well with that retrieved from the dendrogram: for example the types 6F and 5E were distant and well separated from the others as they appeared in the dendrogram (figure 7.15) and similarly the types 4D and 9I were close to each other as indication of high degree of similarity. The position of types 3C and 8H also confirmed what depicted in the dendrogram and measured with the pie charts. The two types are closer but different because of the geological bedrock.

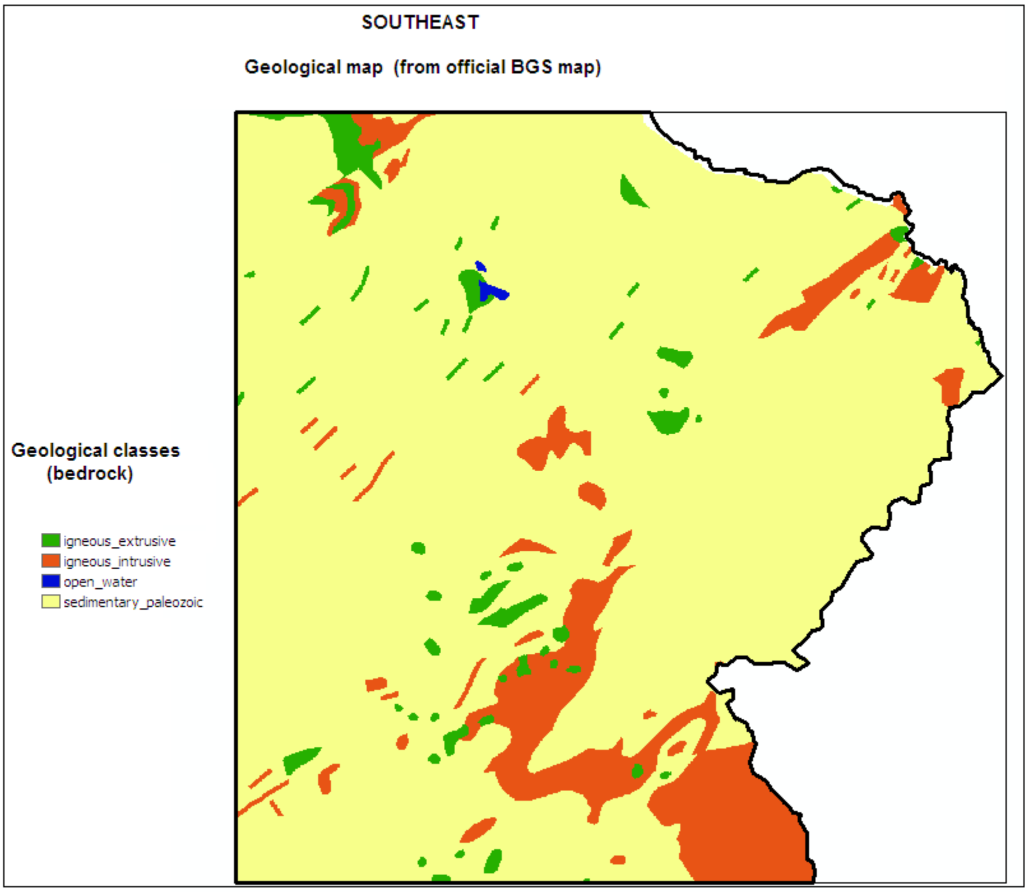


Figure 7.17. Map illustrating the geological bedrock of the Southeast area of study.

Finally, it was interesting to see the positions occupied by types 2B and 7G which according to the dendrogram were expected to be closer to each other, while in the factorial map were more distant. On the basis of the factorial map the 9 clusters could be grouped in 4 major areas (highlighted in red squares) and only the type 7G was excluded since it was considered more a transitional landscape.

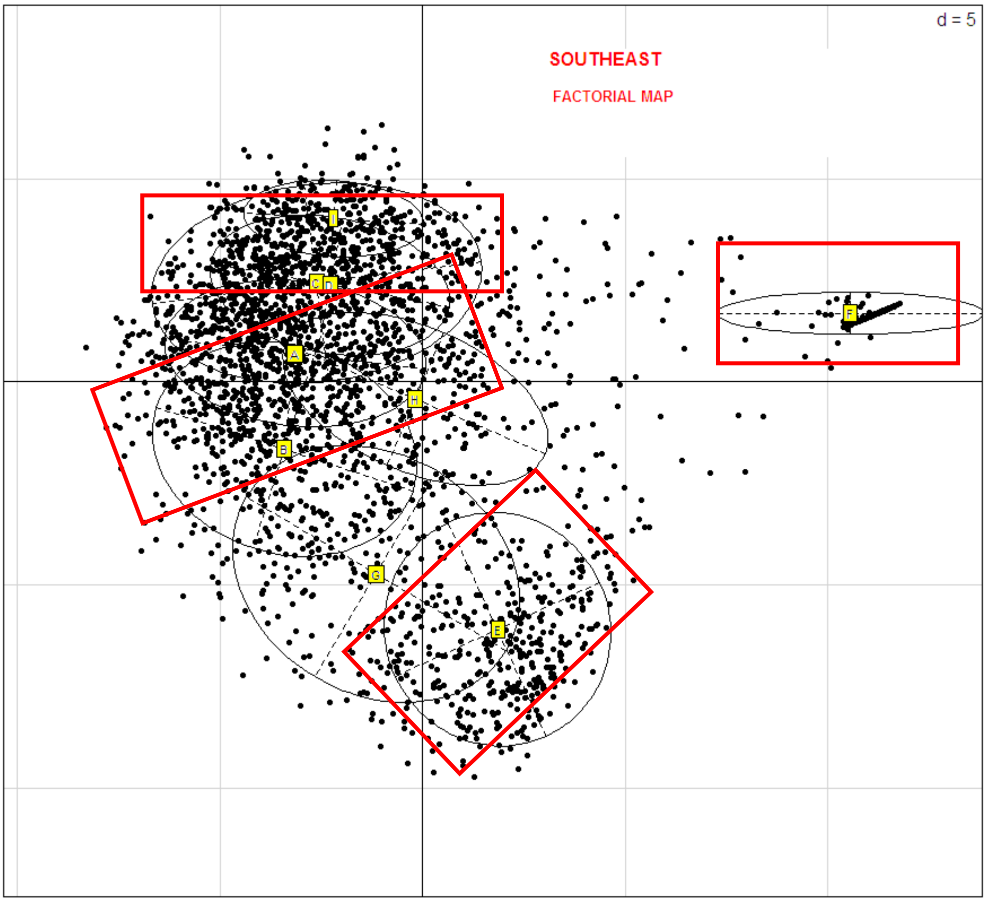


Figure 7.18. Factorial map derived from MULTISPATI analysis. It is interesting to see the way the polygons group together in the ellipses. The position of the type F (open sea) and E (the upland hills), located far apart, is an indication that these two types differ considerably from all the others.

The biplot in figure 7.19 was helpful since it explained the way the variables were correlated to each other and distributed along the first two axes of MULTISPATI-PCA. As previously mentioned the biplot can be compared to the factorial map and can be used to understand the most dominant landscape elements for each landscape character type. If the analysis is accurate, the general information of the biplot should match with that conveyed by the pie charts, which is more detailed. In the example of Southeast, the correspondence between biplot and pie charts was achieved.

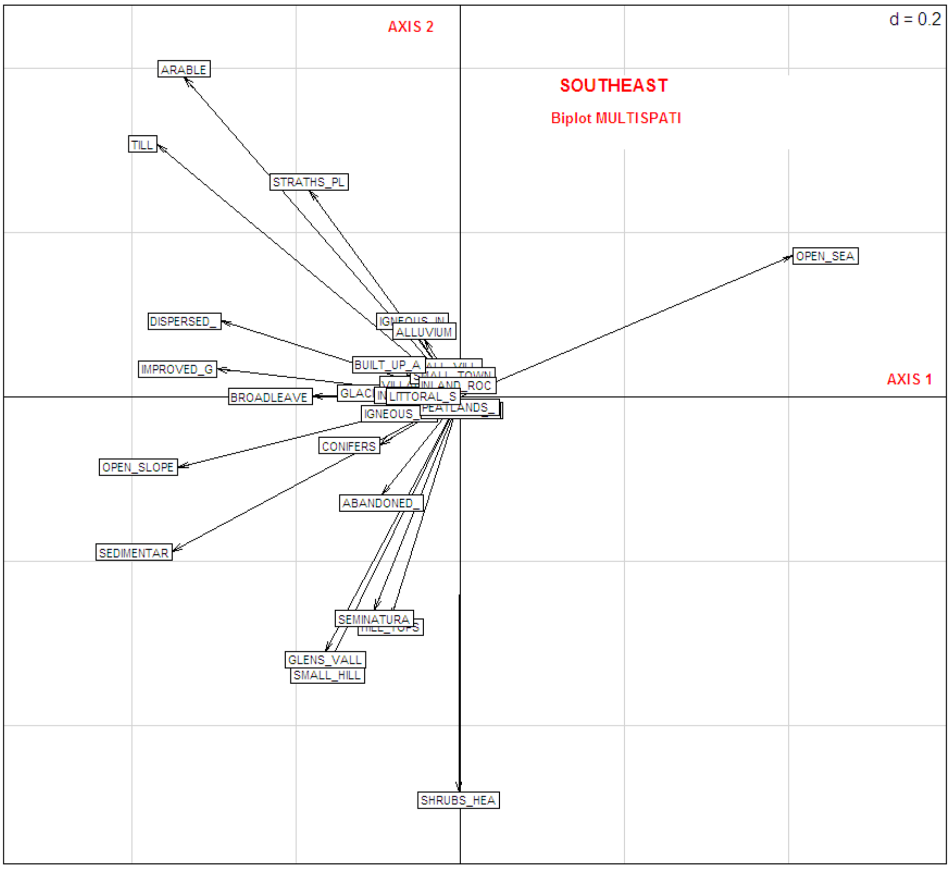


Figure 7.19. The biplot illustrates how the variables correlated to each others and comprise the first and second component of the PCA.

With the results obtained from the analysis of the pie charts and the statistics it was decided to keep the types 2B and 7G, and 3C and 8H separated, therefore a second cluster analysis was performed to reduce the number of types from 9 to 8. Figure 7.20 displays the resulting map.

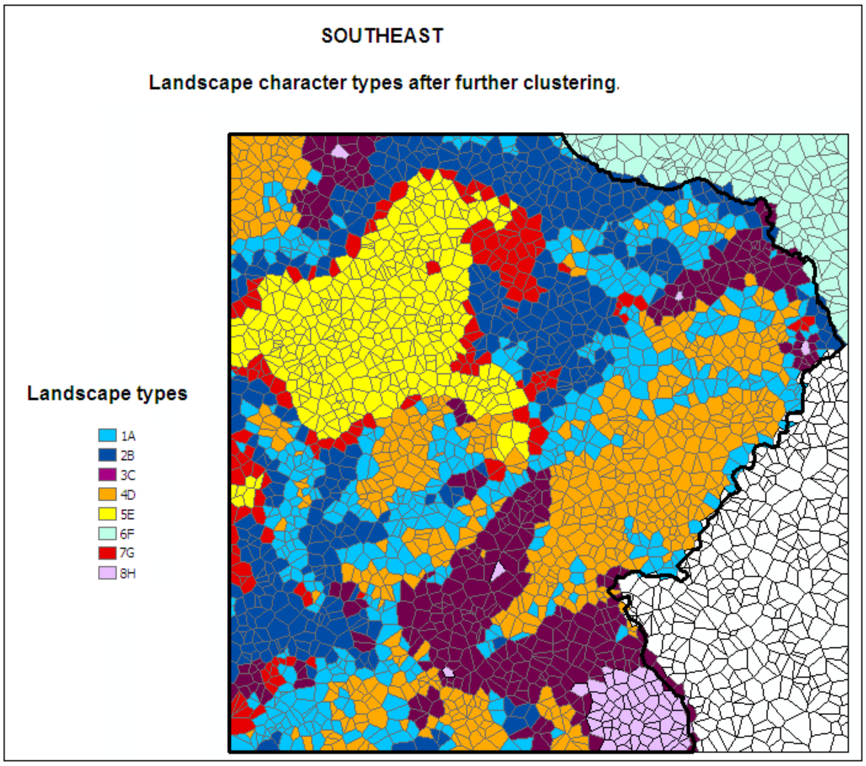


Figure 7.20. Map of the landscape character types for the southeast area after a further clustering. It is worthwhile noticing the type 7G which in the factorial map in figure 19 appeared to be a transitional landscape between the types 5E and 2B.

At this point, after having verified that the number of GIS types is sufficient to describe adequately the character of the Southeast area, it was possible to try and classify the landscape types with a more explicit description. Due to the discrepancies with the official LCA only some of the classes can be borrowed and table 7.5 illustrates the classification attempt.

As it has emerged so far, reading the maps of landscape character types obtained from the combined GIS/MULTISPATI analysis requires time and knowledge about the statistics used thus it is not a straightforward operation. Nevertheless the combined analysis of dendrograms, pie charts, factorial maps and biplots is the procedure that allows us to verify the accuracy of the map and gain a better understanding of what is depicted on the map.

Code GIS landscape types	Description
1A	Rolling farmlands with dispersed houses, improved grassland, grazing and broadleaved woodlands
2B	Upland fringes and valley, with dispersed houses, improved grassland and mixed woodlands.
3C	Lowland valleys and margins with farmland, dispersed houses and broadleaved woodland
4D	Lowland with drumlins and undulating agricultural and woodland (broadleaves) areas.
5E	Upland hills
6F	Open sea
7G	Upland fringes with arable grassland, dispersed houses and woodland plantations (conifers)
8H	Undulating farmlands, hills and valleys with villages and seminatural and abandoned grassland

Table 7.5. GIS landscape character types and an attempt to provide them with a more explicit classification.

Overall, the first part of the chapter has demonstrated that the GIS/MULTISPATI methodology provides the opportunity to carry out an objective and quantitative assessment and comment about the landscape elements composition of each landscape type. The classification of the GIS types into explicit classes and not meaningless code (number plus alphabetical letter) is not yet automatic but, perhaps, this can be the aim of a new research for the future.

7.3 Testing the GIS/MULTISPATI-based methodology for robustness

Once the characterisation of all the areas of study was completed, a series of tests to double check the robustness of the GIS/MULTISPATI-based methodology was conducted. Specifically it was asked “What does the GIS/MULTISPATI analyses do if:

A) New thiessen polygons are generated (but the amount of polygons is kept the same)?

B) Two neighbouring areas are analysed?

C) A different number of Thiessen polygons is used for the analysis. This corresponds to a change in the scale of analysis.

Each question is answered separately here below.

7.4 What happens if new polygons are generated?

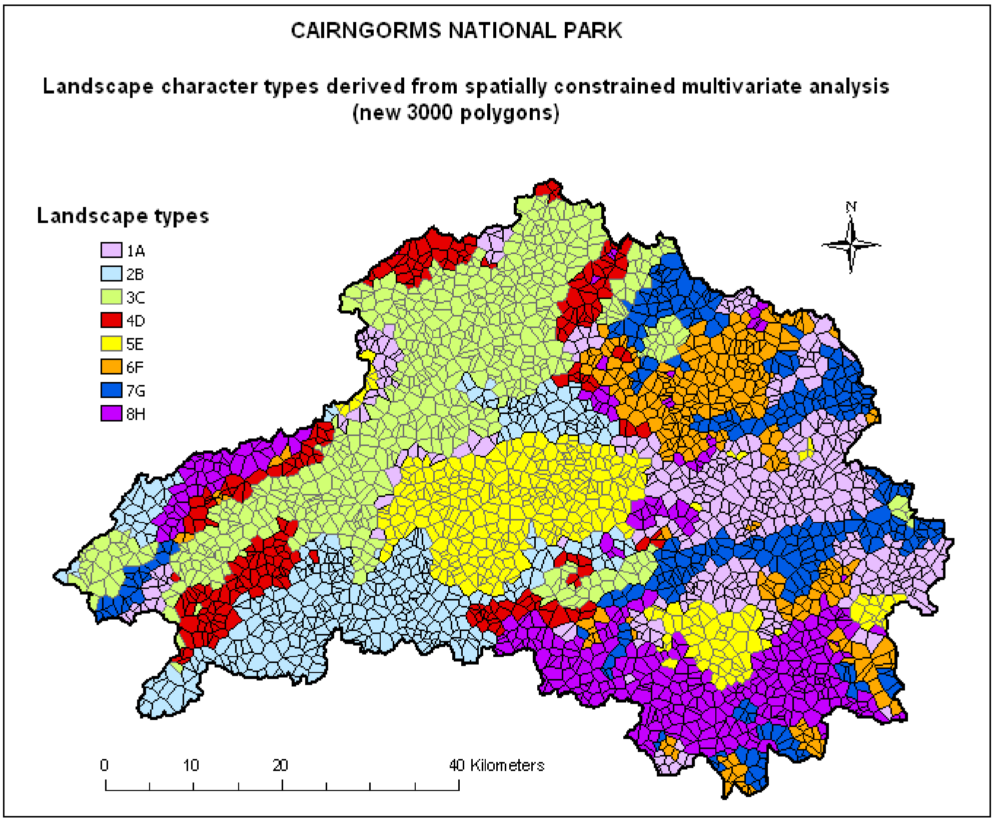
As explained in chapter 6, one of the first steps before starting the GIS/MULTISPATI-based landscape characterisation is to break the area of study into smaller units in order to make possible and facilitate the GIS calculations. Fundamentally, 3000 random points are generated and from them the same amount of Thiessen polygons is created. As mentioned in chapter 6, the polygons are generated randomly instead of being regularly spaced in order to avoid the risk of missing regular patterns of the elements in the landscape.

During the analysis it was asked what would happen if new random points and therefore new polygons were generated. Would the outcomes from the landscape characterisation be the same? The answer to the questions was sought by creating 3000 new polygons and by running the GIS/MULTISPATI analysis again from scratch. The CNP and the Central areas of study were selected in order to perform the test, and the results are presented below.

7.4.1 The Cairngorms National Park

Figure 7.26 illustrates the outcomes from the landscape characterisation of CNP based respectively on the new and the original set of polygons.

.



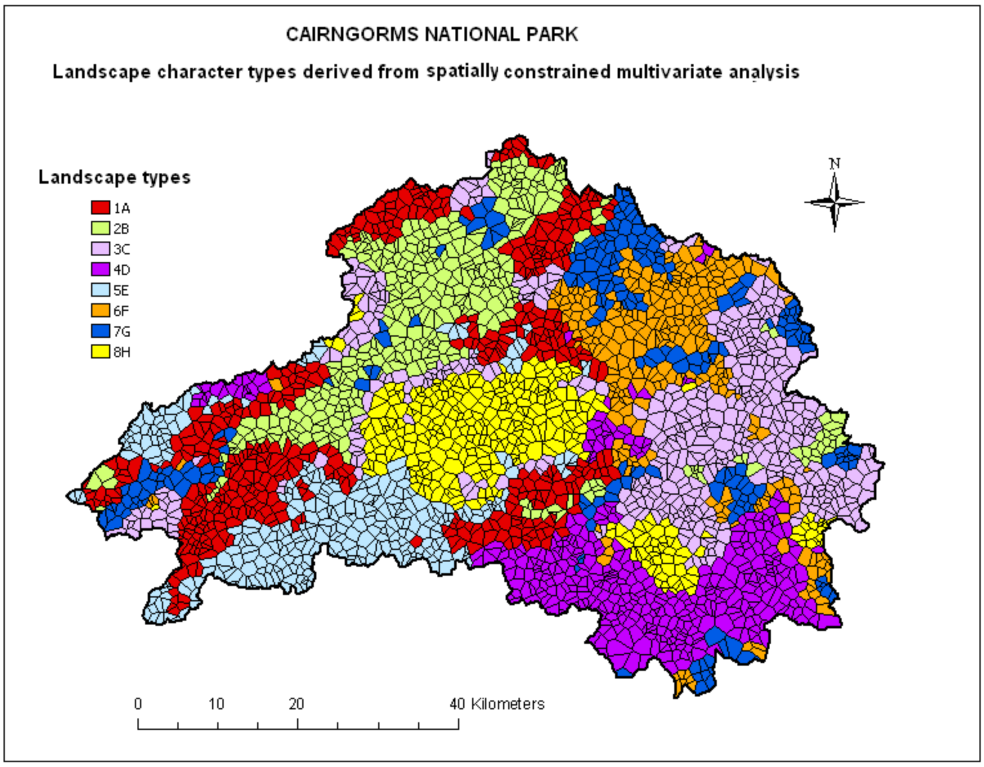


Figure 7.26. Map of the landscape character types of the Cairngorms National Park retrieved from two different set of polygons. The type 7G (strath with coniferous woodlands) identified by using new 3000 polygons (top) is the most different character type if compared to those based

on the original set of polygons (bottom). Notice that the landscape types are matched by colour and not by the name of the class because of the way the cluster analysis operates.

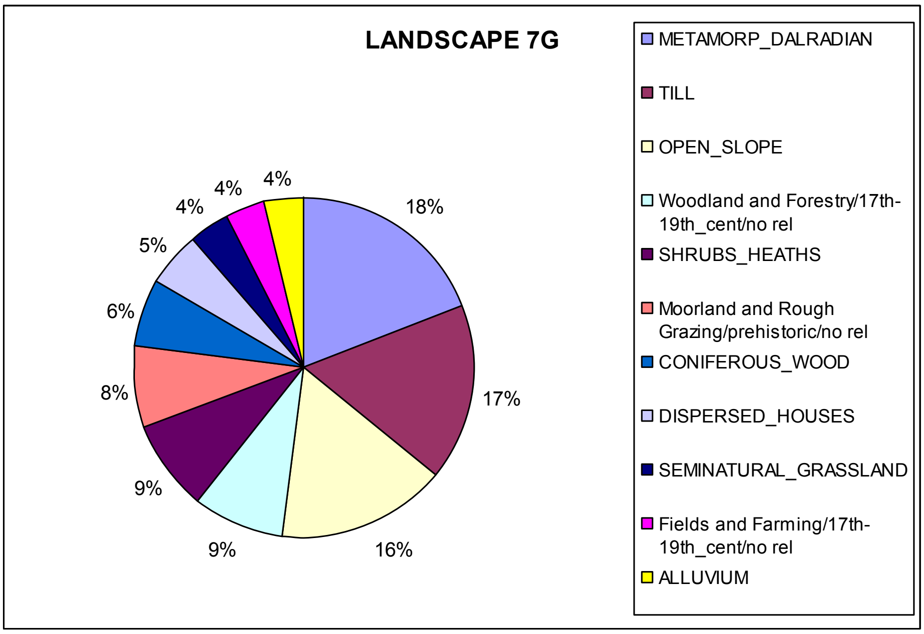
At first sight the landscape character maps of the Cairngorms looked slightly different, again in the south east corner of the Park. However, the visual comparison was followed by the quantitative analysis of the pie charts and the statistics that gave particular relevance to the types 7G since on both maps they showed the most remarkable discrepancies (see figures 7.27a and 7.27b).

The analysis of the pie charts revealed that, contrary to the visual assessment, the difference in the landscape elements composition between the two types was minimal and due to the presence of igneous intrusive rocks which covered 6% of the area in the type 7G obtained from the new set of polygons.

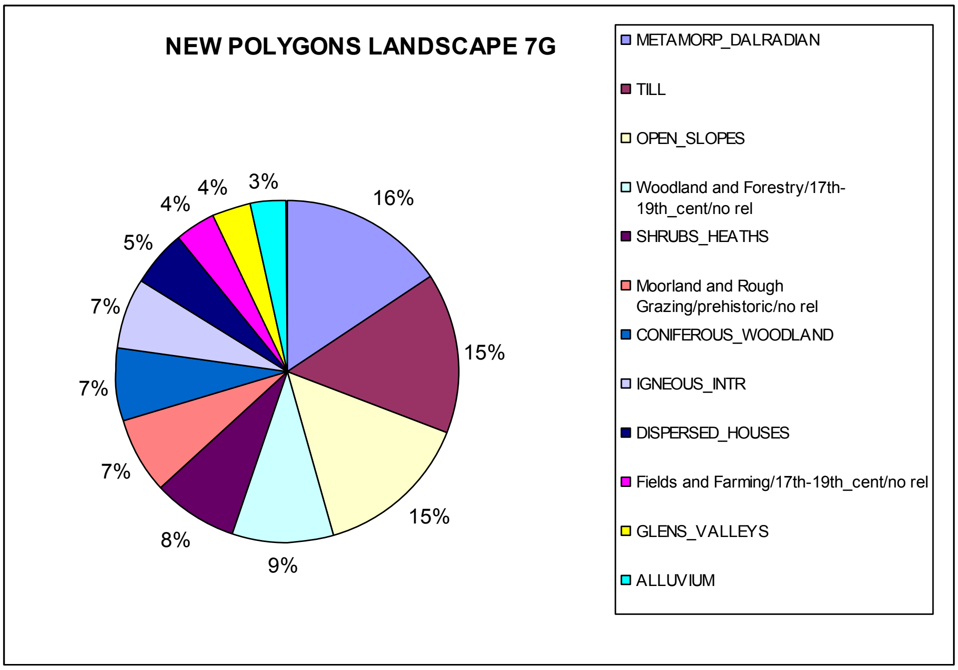
Figures 7.27a and 7.27b

The combined analysis of the biplots and factorial maps helped to confirm the results from the pie charts and again underlined how, on its own, the visual interpretation of the maps can be misleading. Figures 7.28 and 7.29 illustrate a clear correspondence and similarity between the statistics from the two sets of random points.

Figures 7.28 and 7.29



a)



b)

Figure 7.27. The comparison of the pie charts of the types 7G (a = original and b = new polygons) indicate that the differences between the types are minimal, despite graphically on the maps they look more emphasised.

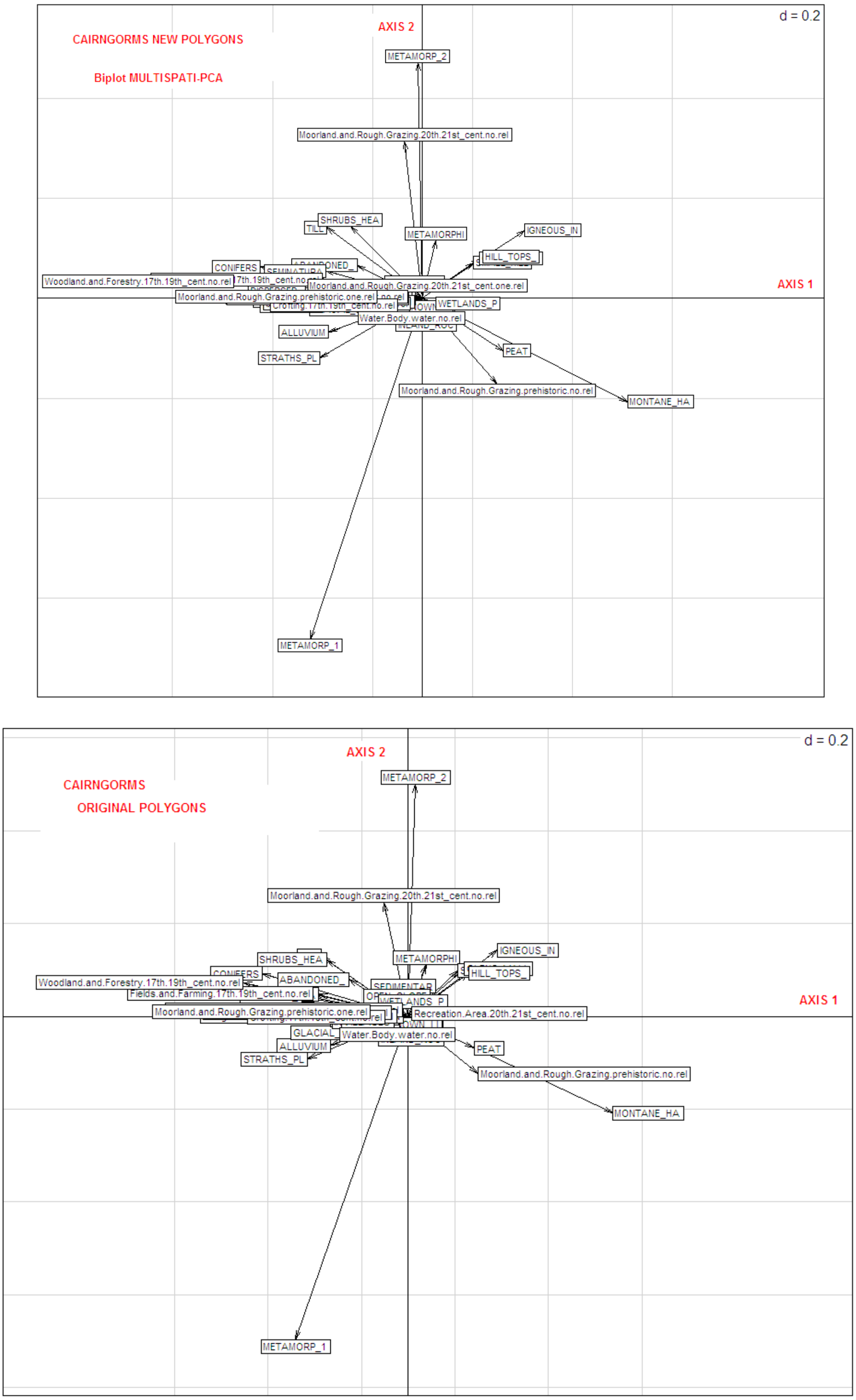
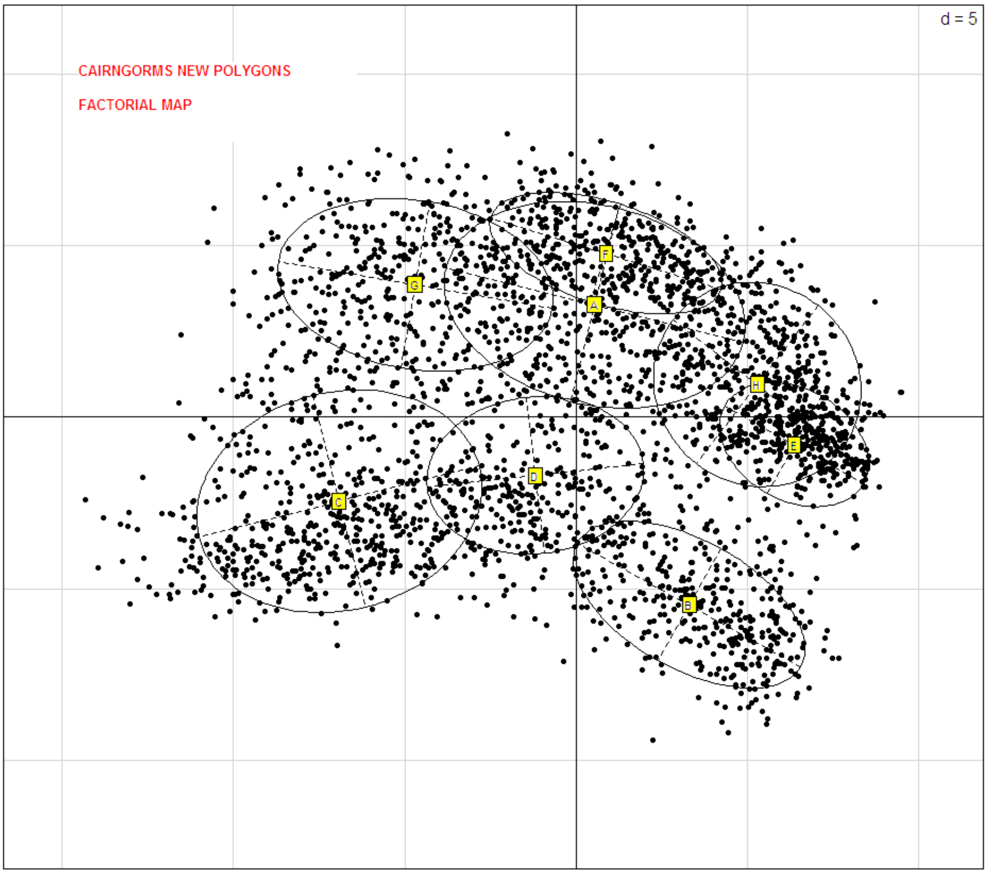


Figure 7.28. The biplots illustrate unambiguously that the correlation between the variable does not vary when moving from the analysis on the new set of polygons (top) to that one on the original polygons (bottom).



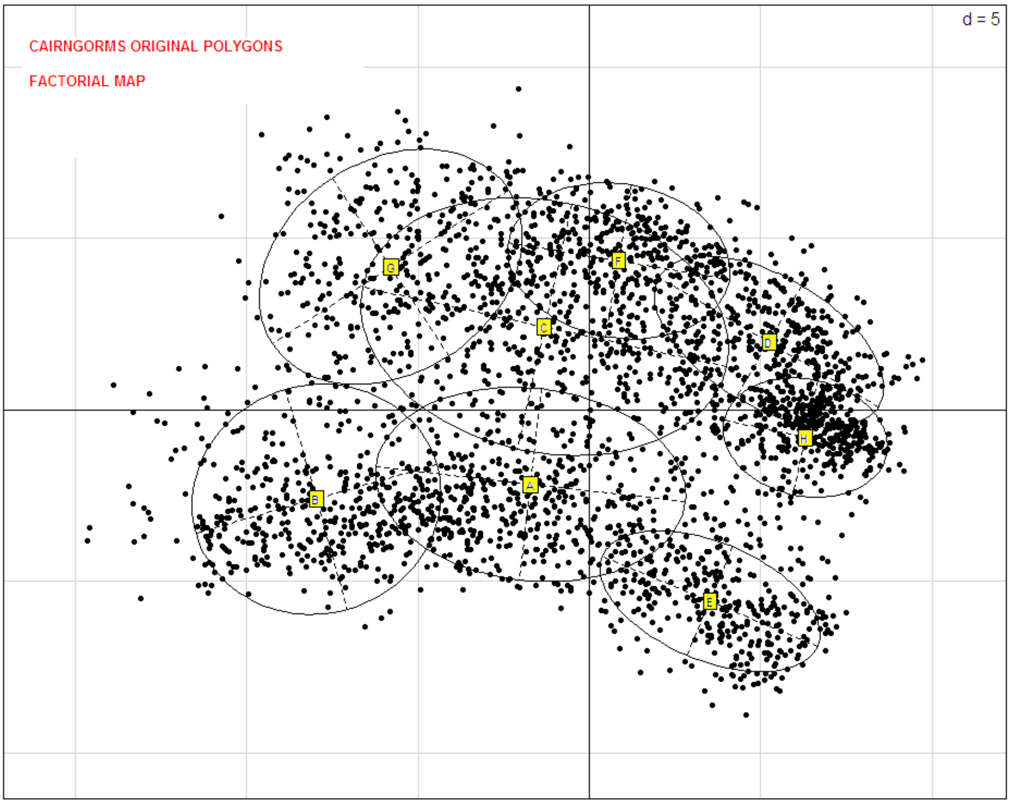


Figure 7.29. The factorial maps from the new set of polygons (top) and the original one (bottom) confirm the results in the biplot; the organisation of the landscape types in both analyses is strongly equivalent.

7.4.2 The Central area

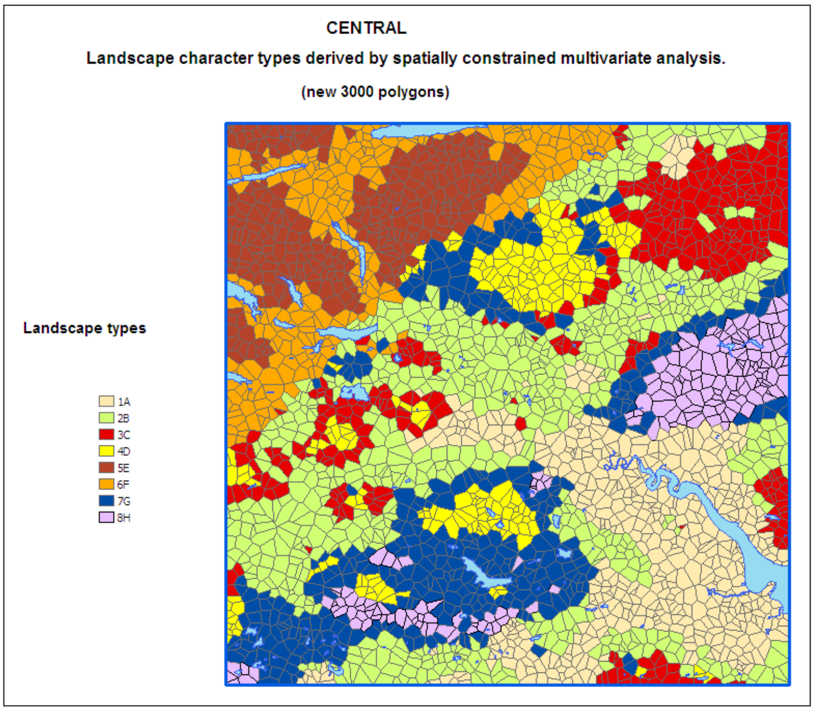
In order to verify that the results from the test were not specific to the structure of the landscape elements in the Cairngorms National Park, it was thought opportune to repeat the test over another area. Thus, the Central area of study, with its more complex landscape, was chosen and figure 7.30 depicts the two maps obtained from the original and new sets of 3000 Thiessen polygons.

Once again both pie charts and statistics were used and analysed to determine whether or not the differences noticed from the visual assessment of the maps matched to differences in landscape elements composition. Particular attention was given to landscape types 1A, 3C and 7G (see figure 7.30 top) and 1A, 4D, 6F and 8H (see figure 7.30 bottom)

Figure 7.30

The statistics in figure 7.31 revealed that there was a correspondence between the biplots of both sets of polygons; in fact the variables looked correlated in the same way and this suggested that the structure of the data in both maps was overall similar.

Figure 7.31



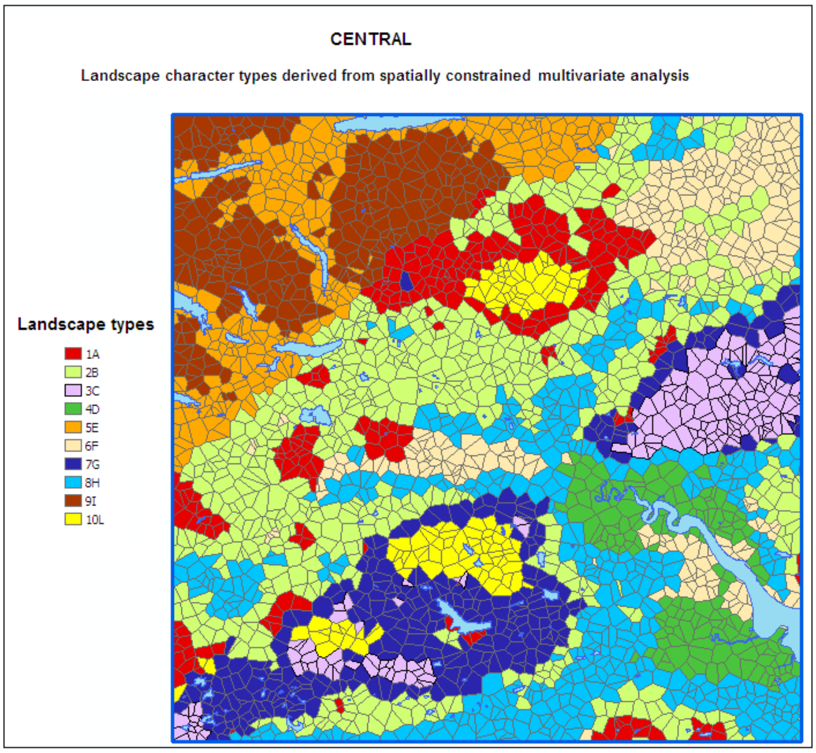
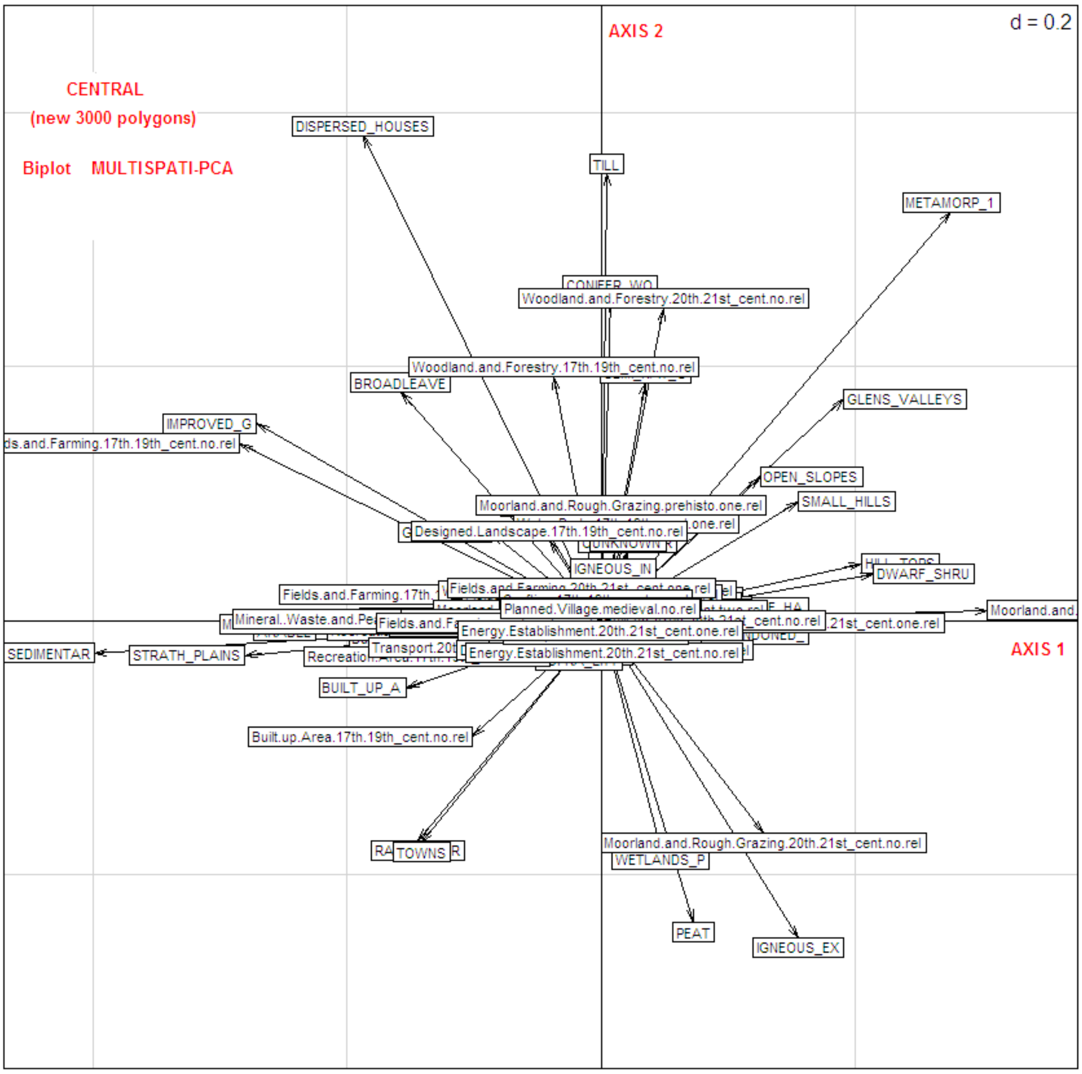
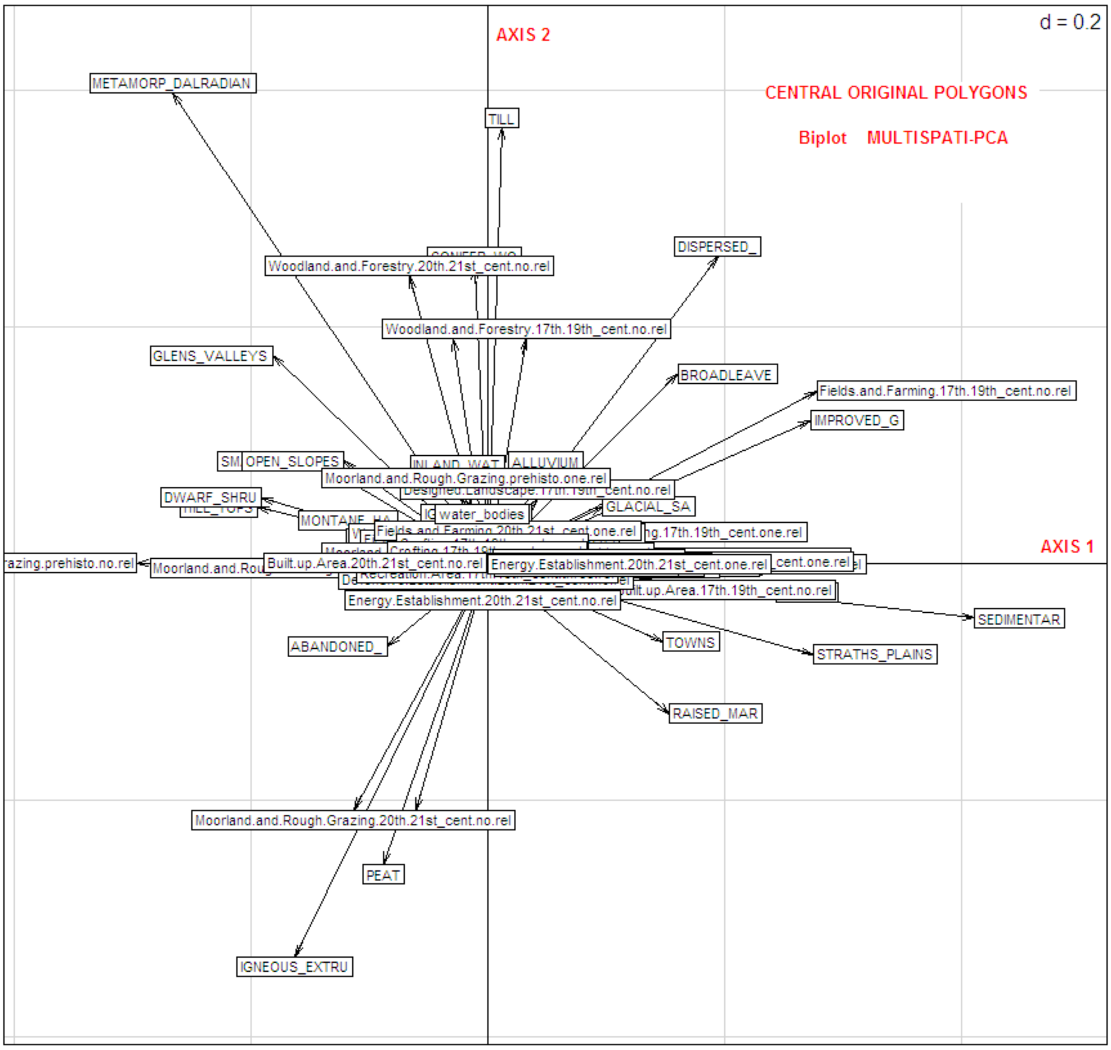


Figure 7.30. Landscape character types of Central area of study retrieved from the analysis of a different set of polygons



d = 0.2
AXIS 2
AXIS 1
CENTRAL
(new 3000 polygons)
Biplot MULTISPATI-PCA
DISPERSED_HOUSES
TILL
METAMORP_1
Woodland.and.Forestry.20th.21st_cent.no.rel
Woodland.and.Forestry.17th.19th_cent.no.rel
BROADLEAVE
GLENS_VALLEYS
IMPROVED_G
OPEN_SLOPES
SMALL_HILLS
Moorland.and.Rough.Grazing.prehisto.one.rel
Designed.Landscape.17th.19th_cent.no.rel
IGNEOUS_IN
DWARF_SHRU
Planned.Village.medieval.no.rel
Energy.Establishment.20th.21st_cent.one.rel
Energy.Establishment.20th.21st_cent.no.rel
SEDIMENTAR
STRATH_PLAINS
BUILT_UP_A
Built.up.Area.17th.19th_cent.no.rel
TOWNS
Moorland.and.Rough.Grazing.20th.21st_cent.no.rel
WETLANDS_P
PEAT
IGNEOUS_EX

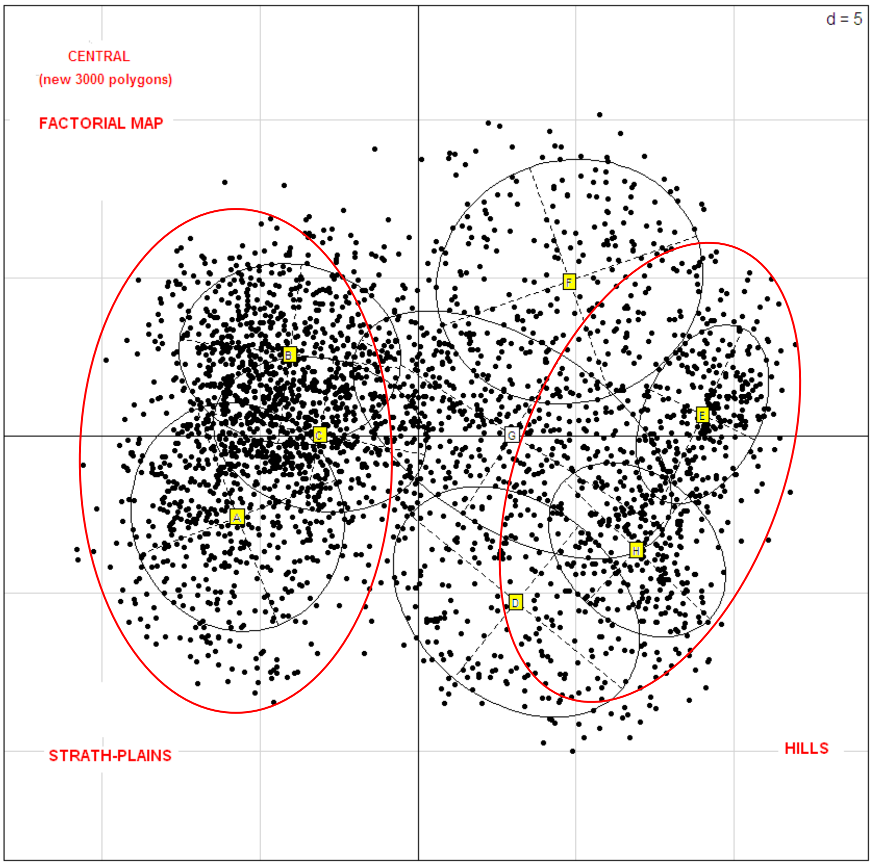
a)



b)

Figure 7.31. The biplots of Central, similarly to what happened in the Cairngorms, do not display remarkable changes in the way the variables are correlated to each others.

The factorial map in figures 7.32 (top and bottom) allowed the consideration of the way the polygons were distributed and grouped into landscape types. It was noticed that, for example, type 1A (figure 7.32 top) corresponded to the types 4D and 8H (of figure 7.32 bottom) since all the types were explained by the combination of straths and plains, built up areas, sedimentary rocks, arable fields and farming. Moreover, types E and H (figure 7.32 top) matched well with I and C (figure 7.32 bottom) since both pairs represented hills, moorland rough grazing and metamorphic Moine rock.



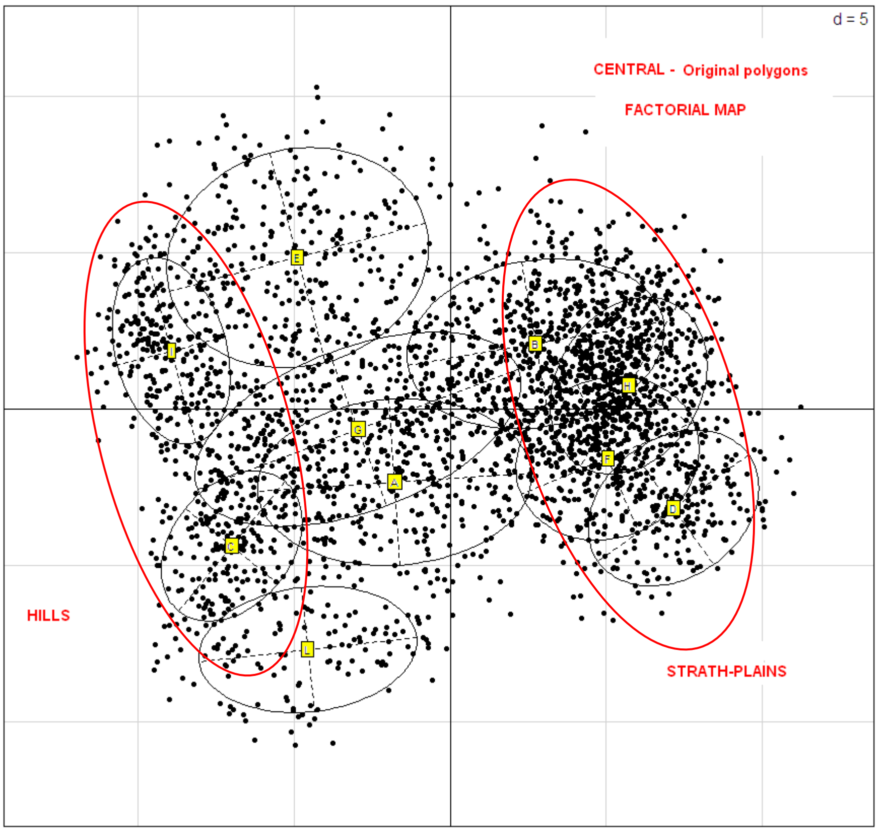
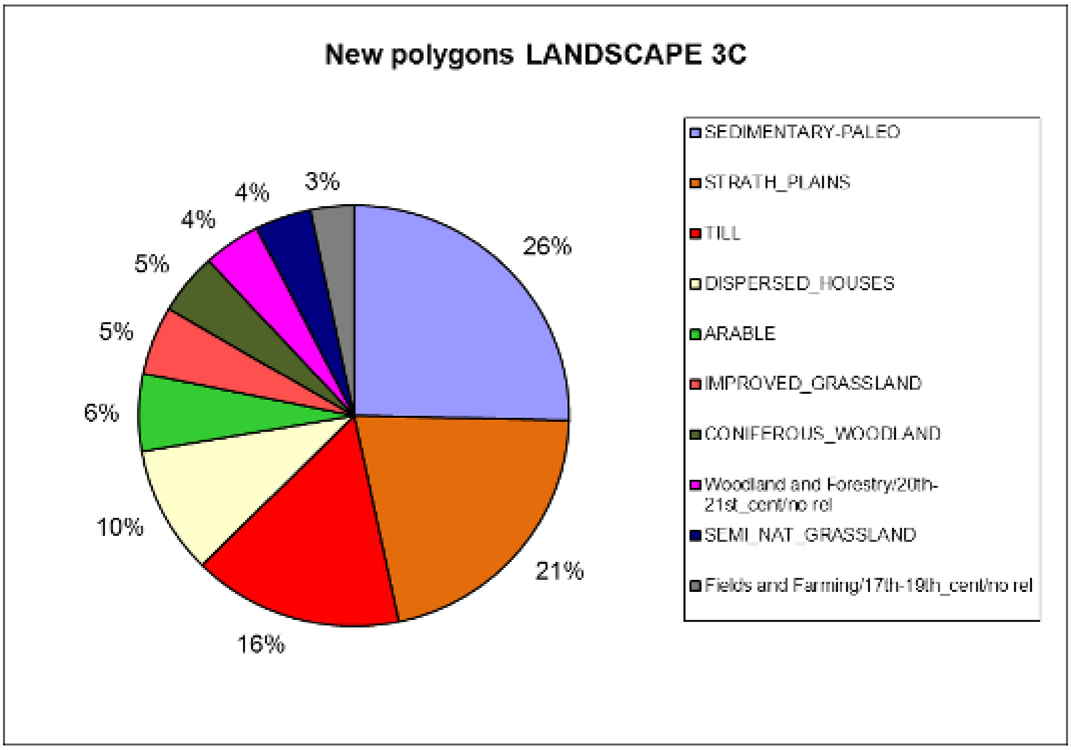
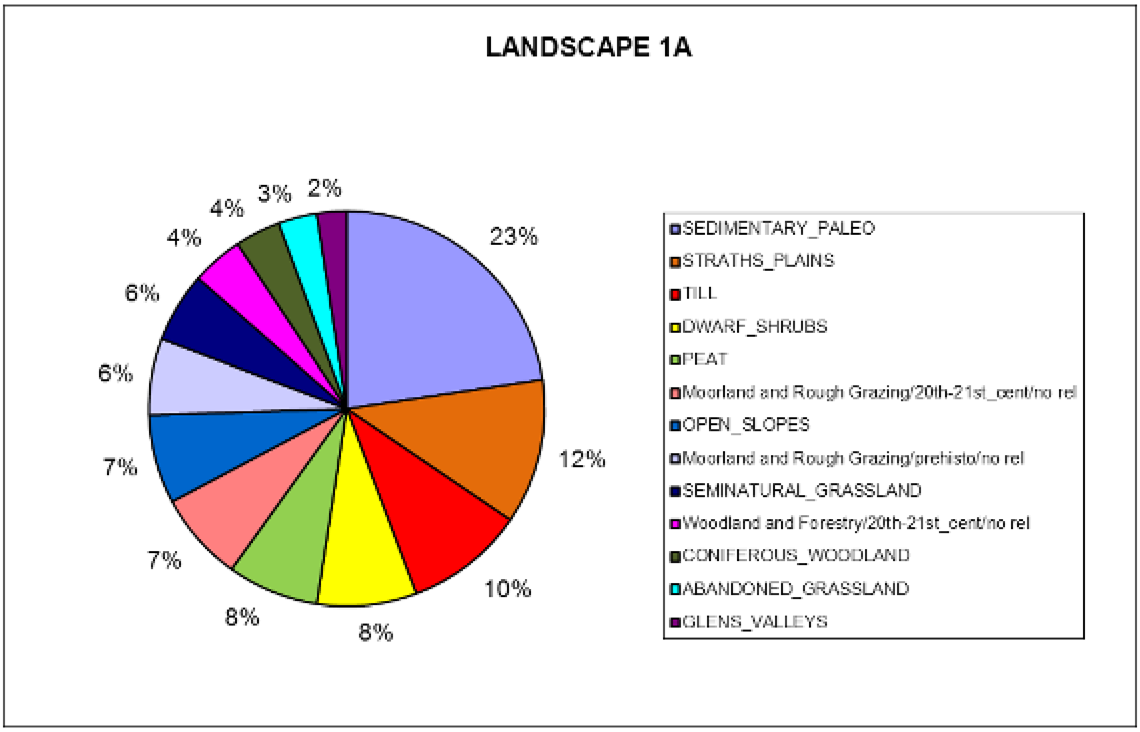


Figure 7.32. The factorial maps reflect the distribution of the landscape types seen in the biplots (fig. 7.31). In the red circles are the character types very different from each others representing the straths and plains and the hills. In between a transitional zone can be identified.

The type 3C (in figure 7.32 top) matched partly with the type 6F and partly with the type 1A (the last two in figure 7.32 bottom) and the correspondent pie charts were taken into account in order to understand the association between these landscape character types (see figure 7.33). The pie charts were helpful since they clarified unambiguously the landscape elements composition of the character types in the new and old set of 3000 polygons.





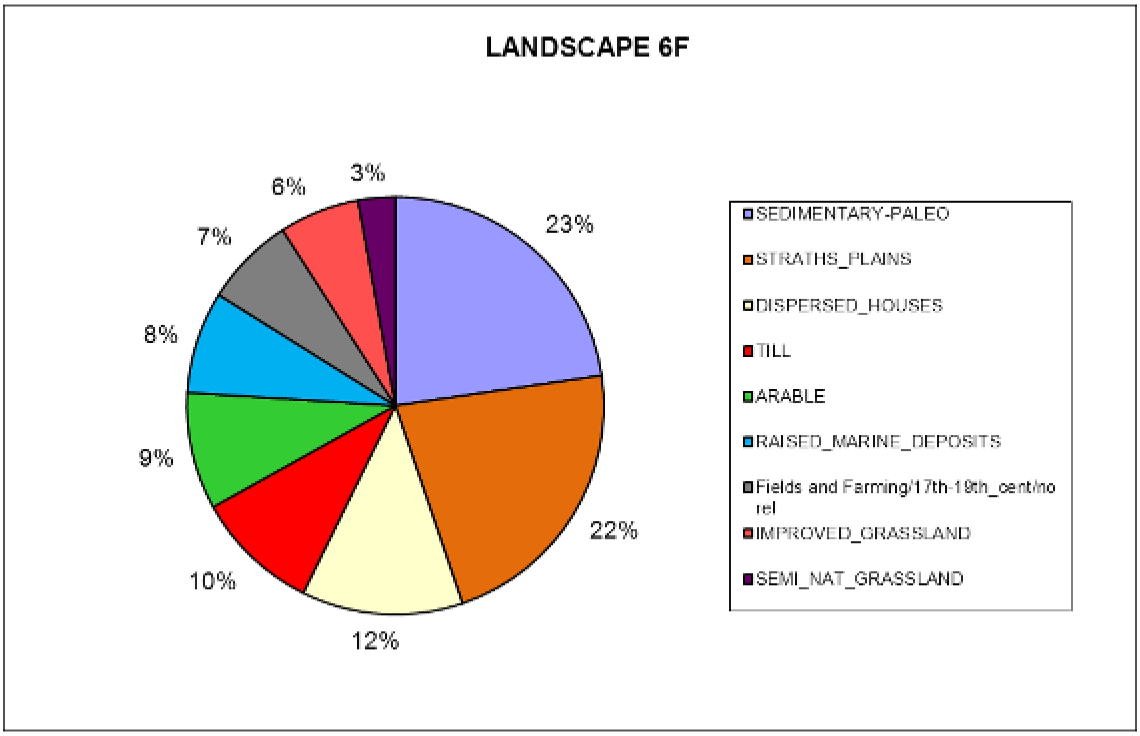


Figure 7.33. The type 3C seems to correspond well with the type 6F and only in part with 1A.

In contrast to the case of the Cairngorms National Park, the new set of polygons tested in the Central area of study demonstrated that the rearrangement of the polygons had a more significant influence on the GIS/MULTISPATI-based landscape characterisation despite being confined to a small number of character types.

Interestingly, for the Central area of study, the biplots revealed a great degree of similarity but the factorial maps and the pie charts showed greater differences. For instance in the factorial map the similarity between landscape types was recognised only with two clusters (in the red circle in figure 7.32) while some of the pie charts conveyed different information. The character type 3C of the new set of polygons appears to be a mix of the character 1A and 6F; in fact a part sharing the three main elements, namely sedimentary paleozoic rocks, straths and dispersed houses, it comprises of arable, seminatural grassland and the historic period of field and farming 17^{th}-19^{th} century which belongs to the type 6F. At the same time, the type 3C is composed of woodland, peat and dwarf shrubs which are more characteristics of the type 1A.

The main conclusion is that in the new 3000 polygons the type 3C contains a mix of landscape elements that in the original polygons could be attributed to two distinct landscape character types. To a certain extent it is possible to claim that the change of the arrangement and shape of the polygons affects in part the landscape characterisation.

Finally it was noticed that similar data emerged from the analysis of the new set of polygons both in the CNP and in Central area of study; in both areas the landscape character types of straths and plains were well separated from those identifying hills or plateaux (in the case of the Cairngorms). Figure 7.34 refers to Central area of study and highlights the landscape types identified as hills and open moorland.

The ability to separate flat areas from hilly grounds was a strong point in favour of the GIS/MULTISPATI landscape characterisation that demonstrates its efficiency of reading the landscape and detecting its main characteristics.

Figure 7.34

Overall, the main conclusion after the two tests on the CNP and Central was that in general the change in the set of polygons was partly influential in the identification of landscape character types and an interesting analogy with the previous test about the introduction of new data input was spotted. It clearly emerged that regardless the set of polygons used during the analysis the landscape character types with a strong spatial structure, with elements organised in spatial clusters and strongly autocorrelated were always detectable.

Figure 7.34. The Central area of study as it appears from the satellite. Satellite images, downloaded from Google earth, can be very useful to make a first and approximate validation of GIS/MULTISPATI results; in fact they can cover large areas and are a solution in case the area is remote and difficult to be reached (although this was not the case for the Central). Here the green polygons highlight the landscape types characterised by hill tops, dwarf shrubs, moorland and grazing (copyright image: Google 2010).

For instance in the Cairngorms National Park this was evident for the types identifying the Cairngorms Plateau and strath Spey, while in the Central area of study the separation involved the well defined character types of the hills covered by moorland, rough grazing and the arable and farmland plains.

7.5 What happens if two neighbouring areas are analysed?

So far the examples of landscape characterisation have referred to single areas of study located in geographic areas far apart from each other. What results might be expected from the GIS/MULTISPATI approach if two neighbouring areas are analysed?

For this test it was thought opportune to take one of the 10 areas of study as an example and divide it into two parts. The aim was to verify that a landscape character falling along the boundary line between the two areas was detected regardless if belonging to one area or the other.

It was decided to attempt the test on the Cairngorms National Park because it showed a distinct and recognisable landscape character type, the “Cairngorms Plateau”, located in the middle of the park. Both the geographic location and the strong spatial autocorrelation amongst the landscape elements of the Cairngorms Plateau provided an ideal case study in order to verify whether or not the GIS methodology was able to identify this character type. Consequently the map was divided in two parts and each one was analysed separately. Notice that in figure 7.35 the Cairngorms Plateau is coded as 8H.

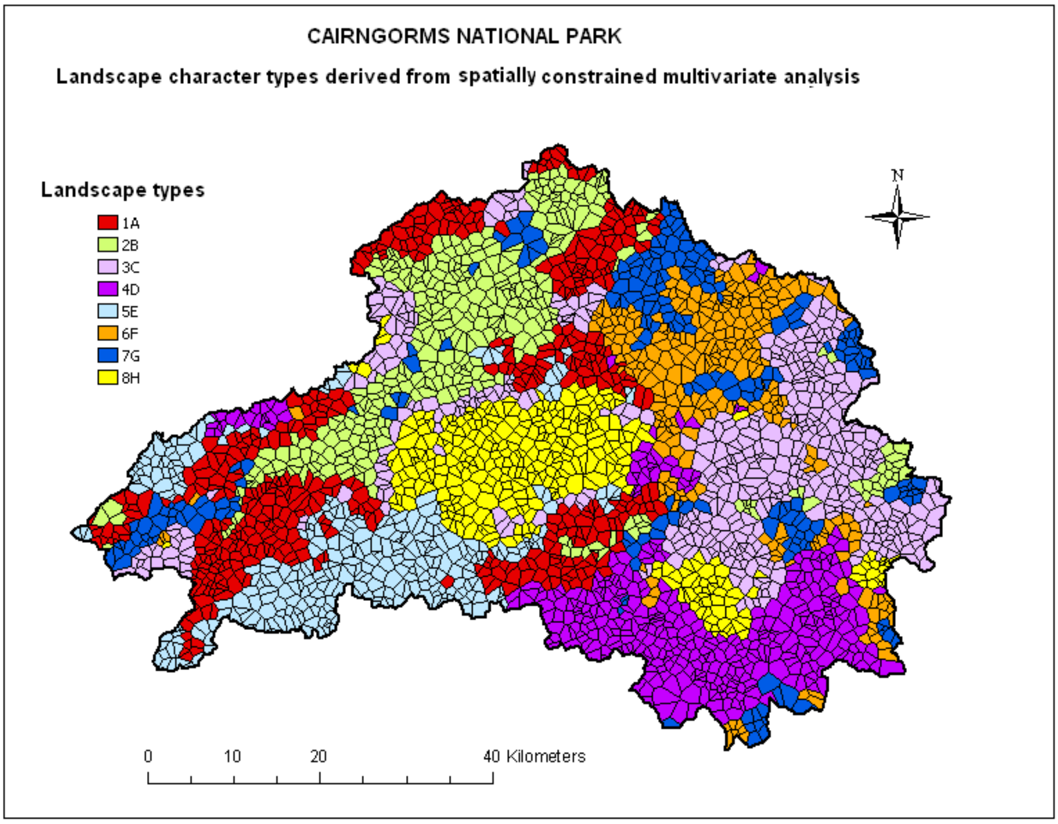


Figure 7.35. Map of the landscape character types for the Cairngorms National Park.

The first step was to draw the hypothetical boundary line approximately in the middle of the park and to identify the two parts as "West" and "East" (see figure 7.36).

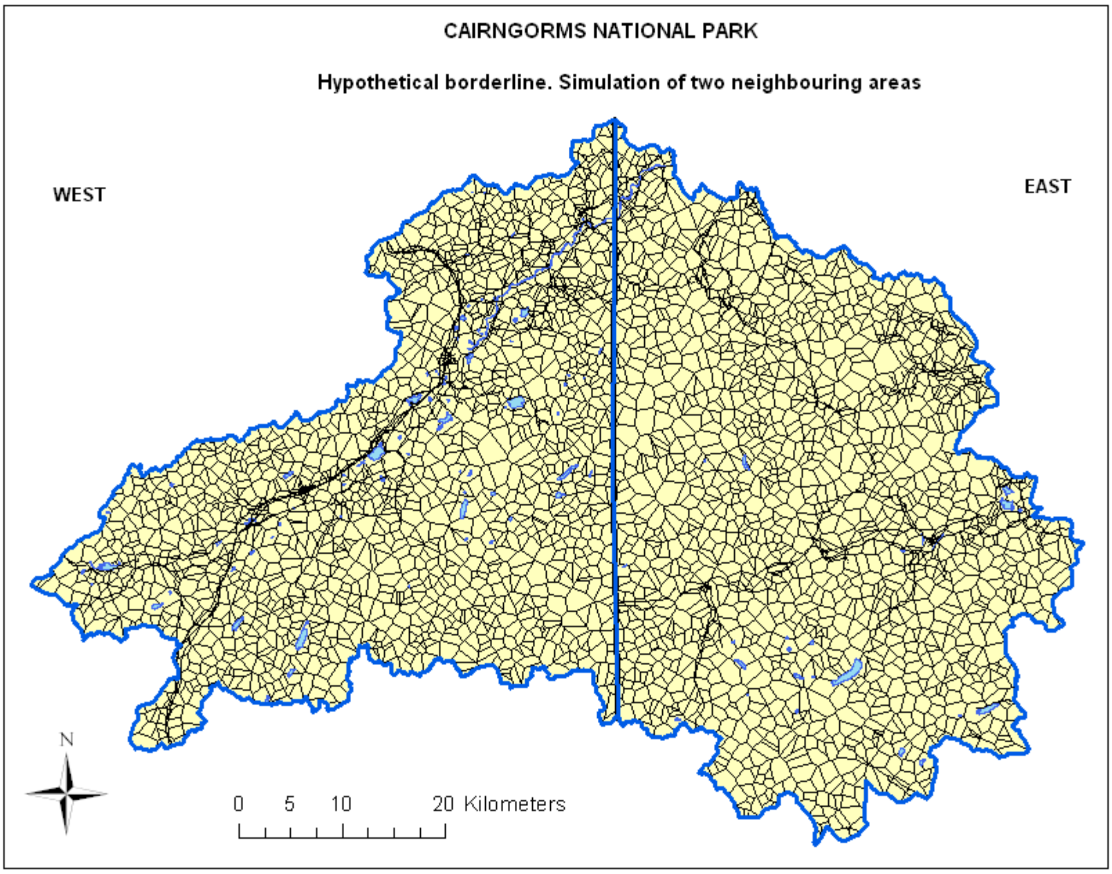


Figure 7.36. Simulated boundary line in the Cairngorms National Park. West and East were the two new areas where the landscape characterisation was carried out.

Then a new set of 3000 Thiessen polygons[32] was generated for each part since the analysis began again from scratch. The third important step was to create around the boundary line a buffer with a width of 2 kilometres which were thought to be sufficient for this test (see figure 7.37). The buffer played an important role in order to inform the GIS that polygons on both sides of the boundary line had to be taken into account when the two separated analyses of the western and eastern part of the Cairngorms were run.

[32] The 3000 polygons were identified with an ID number, their easting and northing coordinates and their belonging to the "West" or the "East" part.

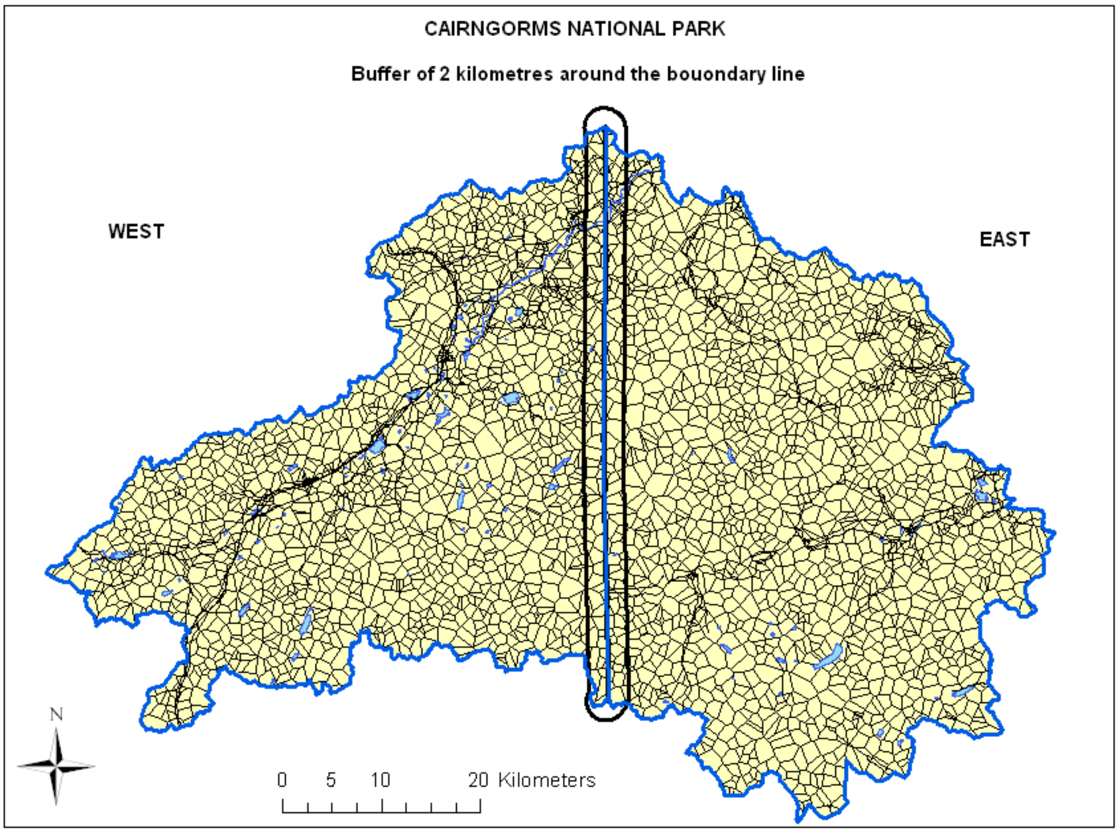


Figure 7.37. The buffer of 2 kilometres around the boundary line had an important role in the analysis since it allowed the GIS to consider the polygons on both sides of the boundary line.

After the polygons were selected for both areas, multivariate spatial analysis was carried out and the outcomes were imported in arcGIS where the maps of landscape character types for West and East of the Cairngorms were plotted (see figure 7.38).

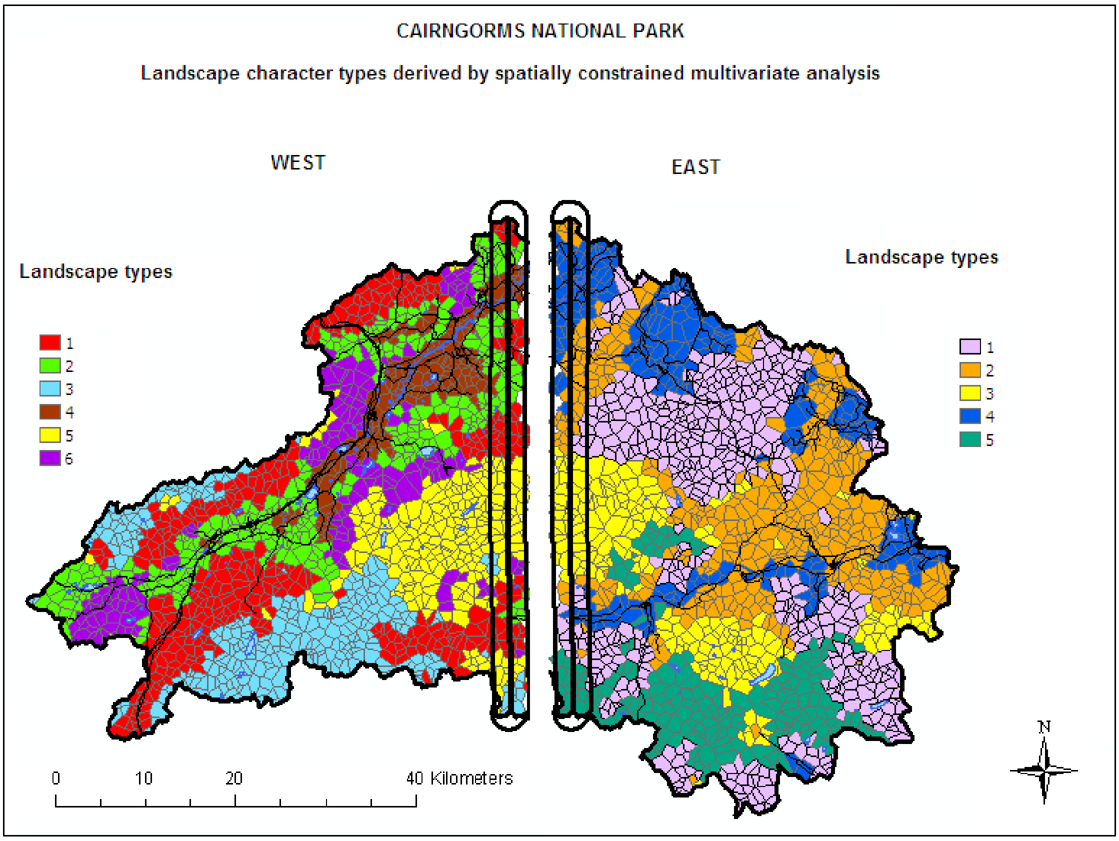
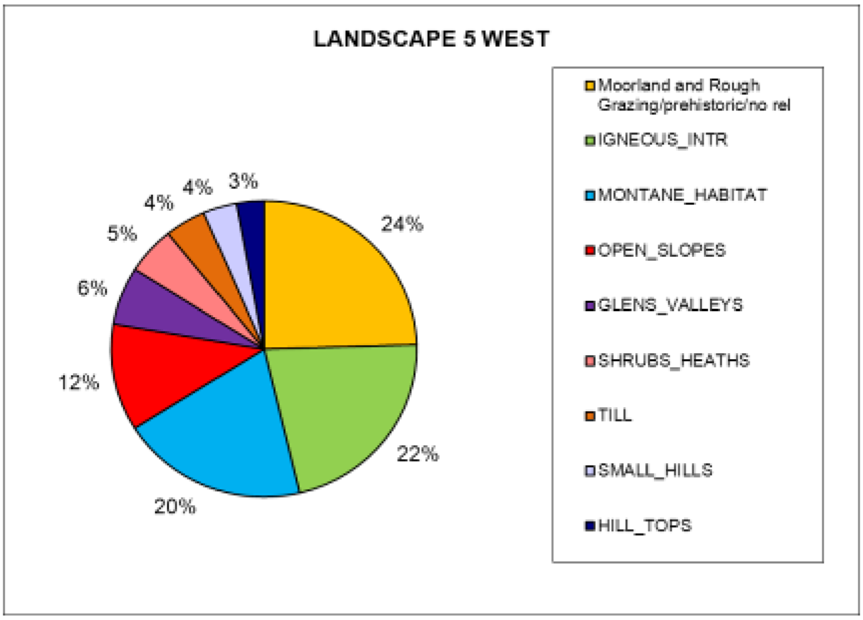


Figure 7.38. Maps of landscape character types of West and East areas of the Cairngorms. It was decided not to overlap them but show separately in order to have a better understanding on the way the GIS/MULTISPATI analysis characterised the polygons within the buffer.

The analysis of the pie charts associated to the Cairngorms Plateau, and specifically type 5 in West and type 3 in East, were the first to be analysed and a high level of similarity between the two types was recorded (see figure 37.9). Basically the match in terms of landscape elements composition equated to 98% with small variation in the percentage of individual elements: for example igneous intrusive rocks in type 5 (West) described 21% of the overall landscape elements, while in type 3 (East) they accounted for 23%.



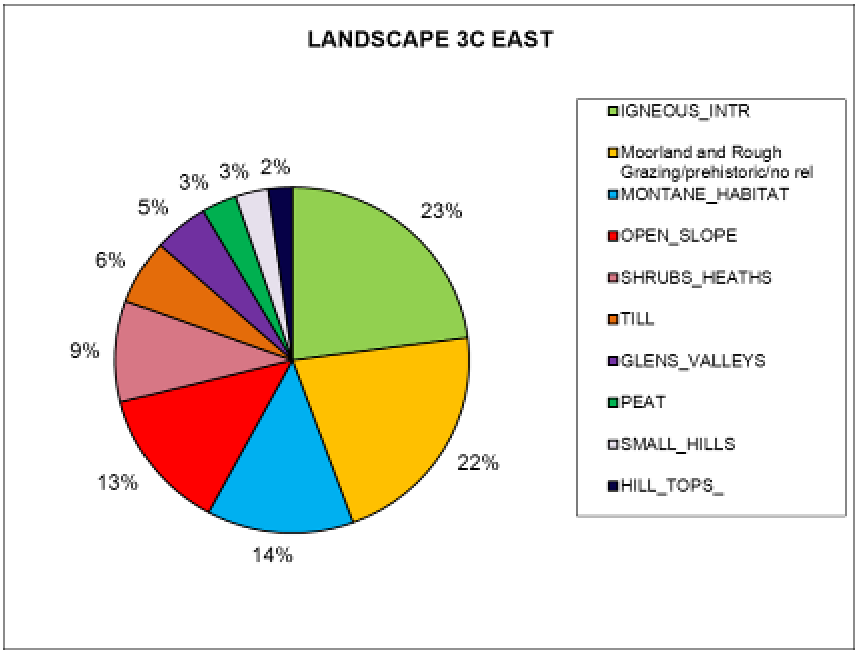


Figure 7.39. Pie charts of the types 5 and 3 that represent the Cairngorms Plateau in the western and eastern part of the Park. The degree of similarity of the pie charts in terms of percentage of landscape elements is an indication that both types describe the same landscape character.

In order to verify that the landscape character type identified with types 5E and 3C matched with that one of type 8H in the original map (see figure 4.35) a comparison with the correspondent pie chart was carried out and the results were positive. As illustrated in figure 7.40 the pie chart of type 8H reported exactly the same landscape elements composition, which indicated that the GIS/MULTISPATI analysis was able to identify that specific character of the CNP's landscape regardless the presence of the boundary line.

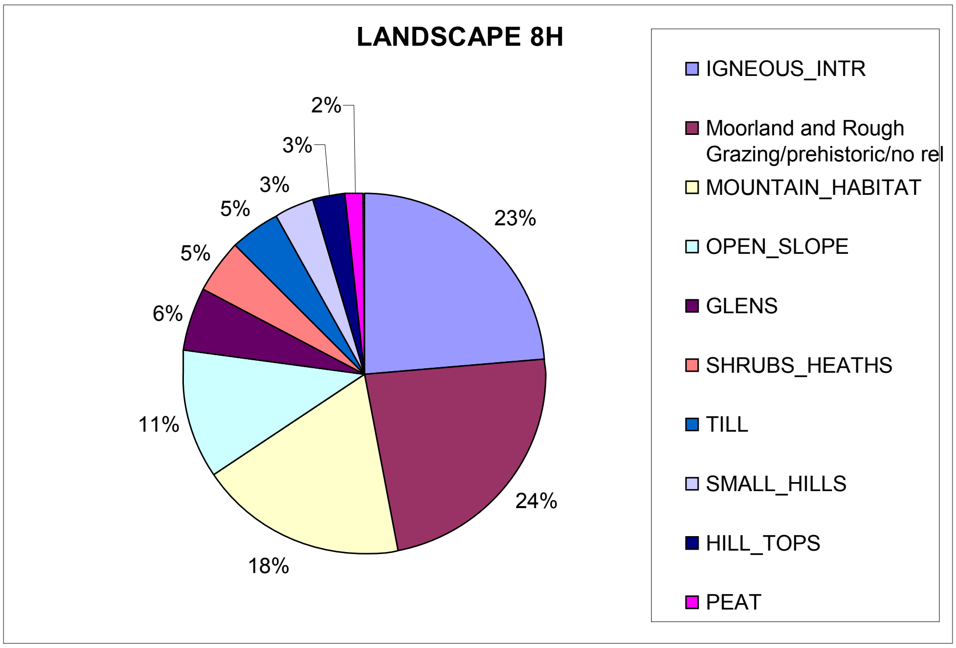
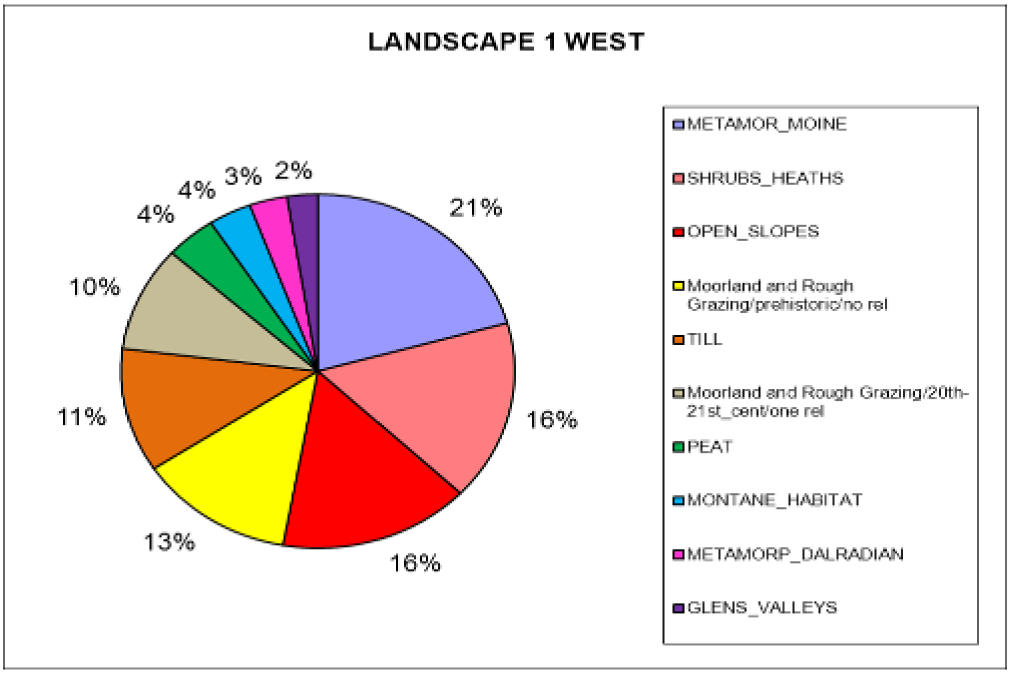
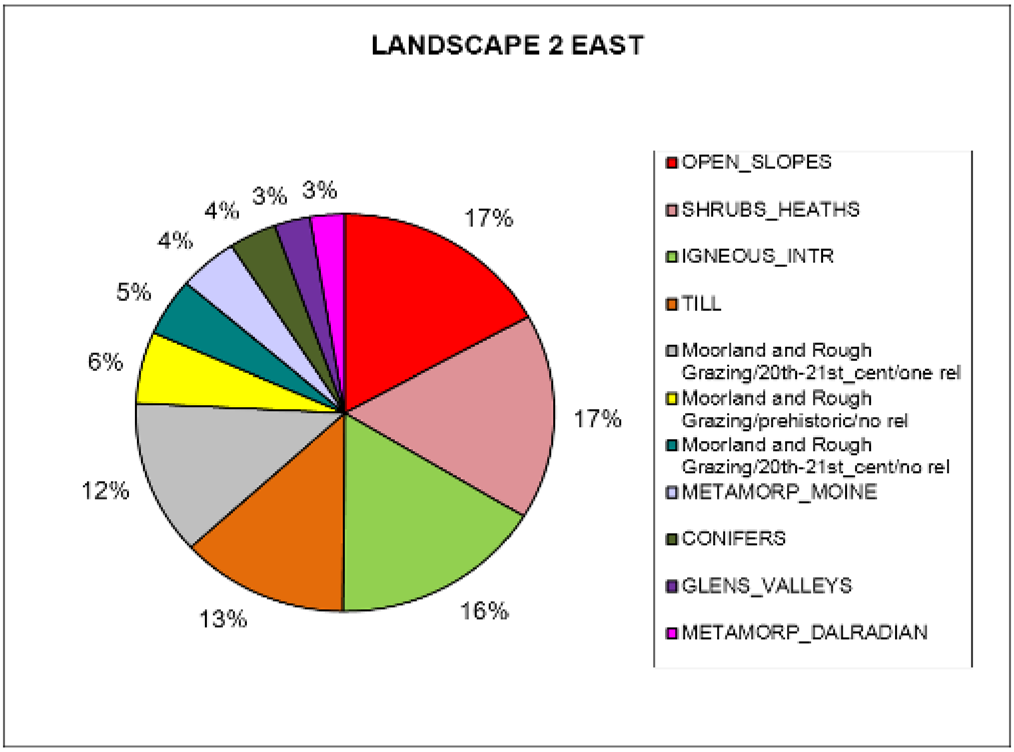


Figure 7.40. Pie chart of the Cairngorms Plateau character type from the original map (before the artificial.

The positive outcome nevertheless looked confined to the Cairngorms Plateau character type. In fact from the original map in figure 7.35, other two types were located on both sides of boundary, precisely types 1A and 2B, but they were identified differently. For example, type 1A was classified as type 1 in West and type 2 in East, while type 2B corresponded to types 2 and 4 in West and type 4 in East. The analysis of the pie charts helped to understand whether or not the types were similar to each other and represented the same character type depicted in the original map.

Figure 7.41 shows the pie charts for types 1 West, 2 East and 1A of the original map.


LANDSCAPE 1 WEST
21%
16%
16%
13%
11%
10%
4%
4%
3%
2%
METAMOR_MOINE
SHRUBS_HEATHS
OPEN_SLOPES
Moorland and Rough Grazing/prehistoric/no rel
TILL
Moorland and Rough Grazing/20th-21st_cent/one rel
PEAT
MONTANE_HABITAT
METAMORP_DALRADIAN
GLENS_VALLEYS


LANDSCAPE 2 EAST
17%
17%
16%
13%
12%
6%
5%
4%
4%
3%
3%
OPEN_SLOPES
SHRUBS_HEATHS
IGNEOUS_INTR
TILL
Moorland and Rough Grazing/20th-21st_cent/one rel
Moorland and Rough Grazing/prehistoric/no rel
Moorland and Rough Grazing/20th-21st_cent/no rel
METAMORP_MOINE
CONIFERS
GLENS_VALLEYS
METAMORP_DALRADIAN

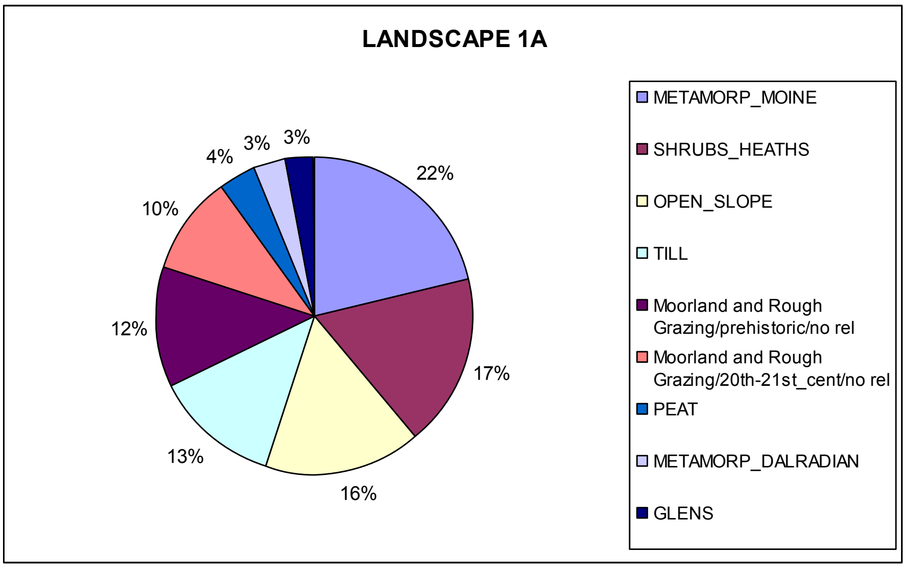


Figure 7.41. Comparison between pie charts of landscape character types which belong to the two areas, West and East, of the CNP and to the original map (landscape type 1A). The effect of the boundary is noticeable, in fact landscape types 1 West and 2 East correspond to the type 1A only f considered together.

The comparison highlighted clearly that type 1 in West corresponded to type 1A of the original map while it differentiated from type 2 in East mainly because of the geology. In fact 20% of type 1 is characterised by metamorphic moine rocks whereas 16% of igneous intrusive rocks form type 2. In addition conifers account for 4% of type 2, while the same percentage of montane habitat occurs in type 1.

The main conclusion was that contrary to the Cairngorm Plateau, which showed similar geology, landforms, land-cover, and historic land-use over a large area on both side of the boundary, the pattern determining the landscape character type 1 was confined more in the West than in the East and in fact it matched perfectly with type 1A that is mainly located on the western part of the National Park (see figure 7.35). A similar explanation was given to describe the differences between types 2 and 4 in West, 4 in East and 2B in the original map. Overall it was noticed that the geology of the National Park influenced the identification of the landscape character types more than the rest of the landscape elements,.

The impact of geology was explained by looking at the geological maps, in particular that one depicting the bedrock, which revealed a quite neat and distinct distribution of metamorphic rocks between West and East (see figure 7.42).

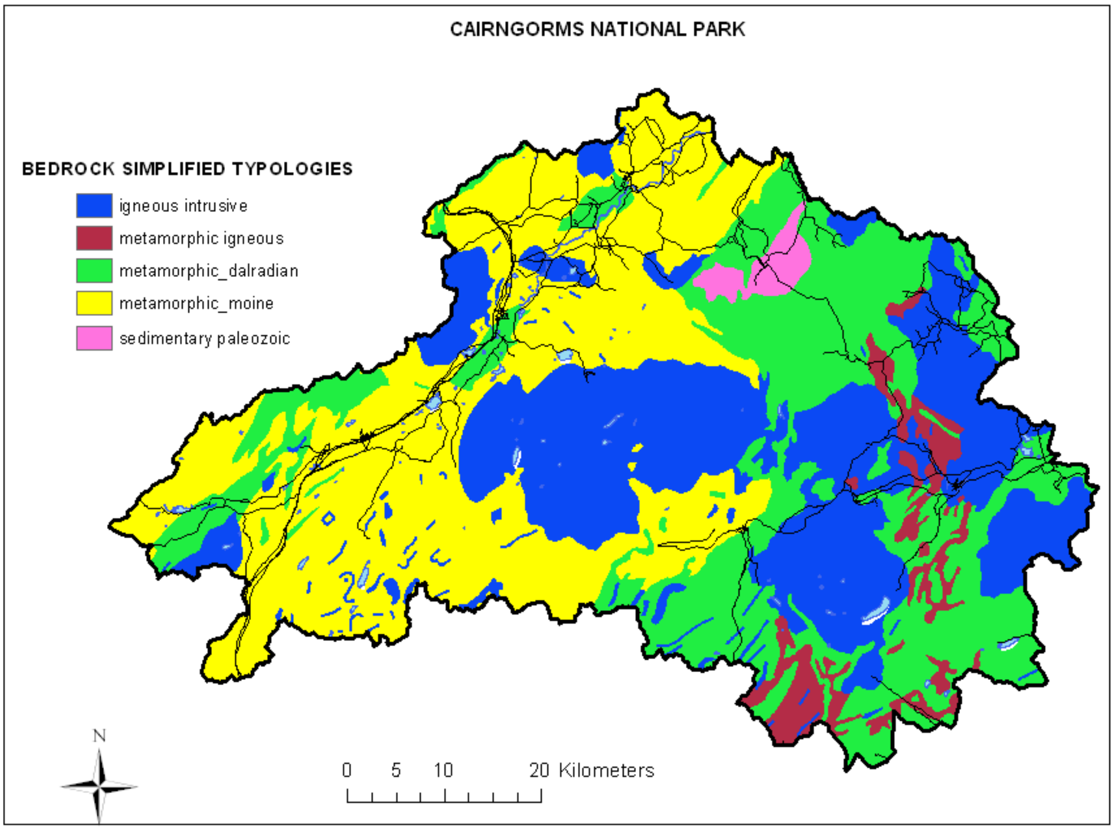


Figure 7.42. Map of the geological bedrock of the Cairngorms National Park. It is very interesting to note the separation between metamorphic Moine rocks, in the West, to metamorphic Dalradian rocks, in the East. In addition, the Eastern part of the Park presents overall a greater complexity in terms of geology compared to the neighbouring Western part.

The evident change in the geological bedrock, from West (metamorphic Moine rocks) to East (metamorphic Dalradian rocks), determined a clear cut in the spatial structure of the data which was confirmed by the biplots (see figure 7.43). From both figures it is possible to derive two observations: the geological elements are the most dominant since account for the highest variability in the data (their vectors are the longest), then they split unequivocally in the West igneous rocks are opposite to metamorphic Moine rocks, while in the East igneous rocks contrast with metamorphic Dalradian rocks. In both cases this clear-cut characterises the spatial structure of the data.

Figure 7.43

Only the Cairngorms Plateau is not affected by the change in geology since, as mentioned before, the association of landscape elements does not vary. In fact in both biplots, igneous rocks are associated to the same group of landscape elements: specifically montane habitat, moorland and rough grazing, hill_tops, small hills and glens.

To summarise, a test was carried out in order to verify whether or not a landscape character type falling along the boundary of two neighbouring areas was detected consistently and unambiguously by the GIS/MULTISPATI-based landscape characterisation. The Cairngorms National Park was taken as an example since it presented a well defined and extended landscape character type at its centre: the Cairngorms Plateau.

The results showed that the GIS/MULTISPATI-based methodology was able to recognise unambiguously only the Cairngorms Plateau character type and less accurately the other landscape types along the boundary line. This confirmed something already noticed in the two previous tests, and specifically that a strong spatial structure is always recognised and detected by the GIS/MULTISPATI analysis. In this case the Cairngorm Plateau was the strongest pattern along the boundary line between West and East, because it was not affected by a change in geological elements and in both parts of the Park showed the same association of landscape elements.

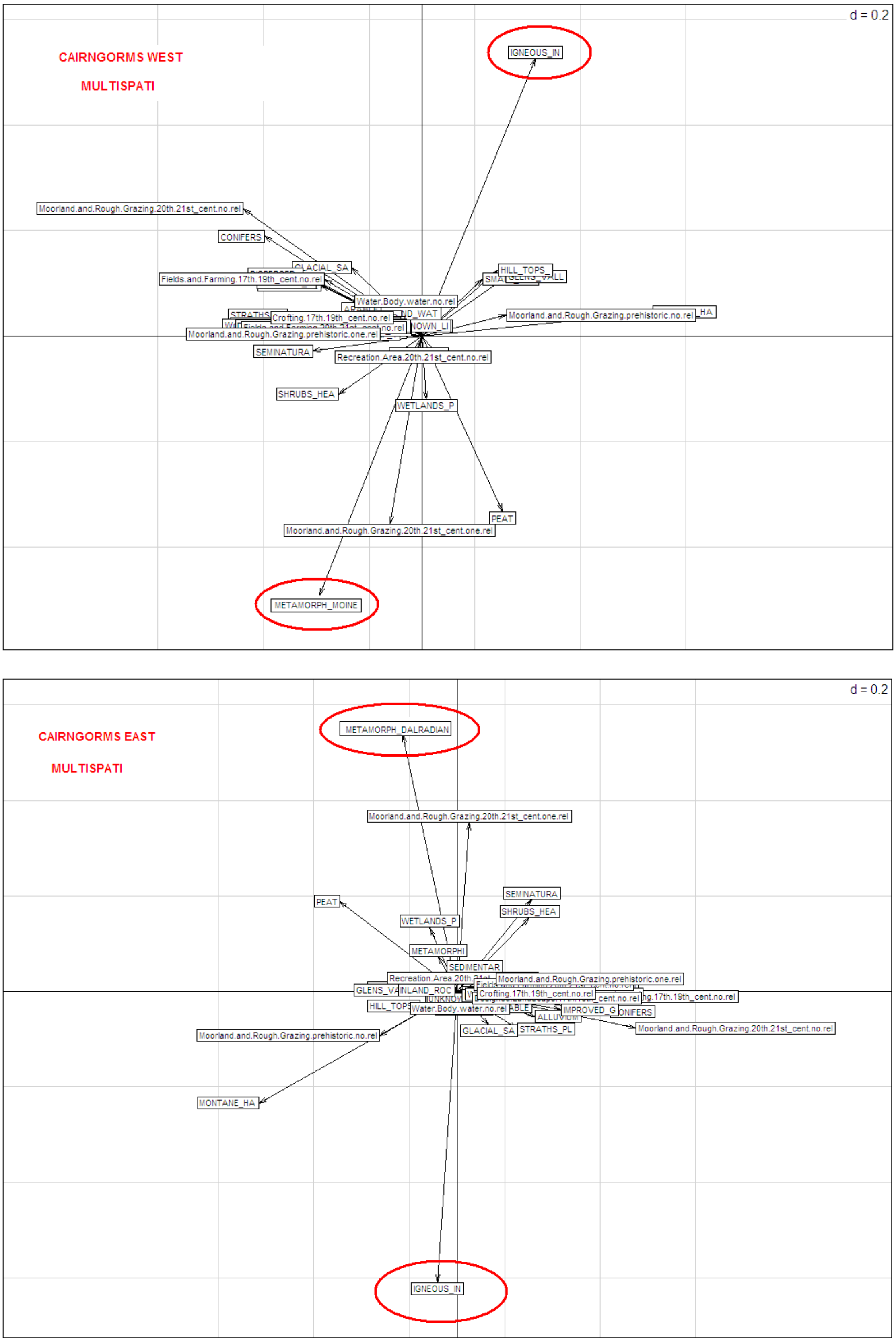


Figure 7.43 Comparison between the biplots of the two areas of the CNP. In the red circles are the geological classes that differentiate the character of the landscape in the two areas: Metamorphic Moine (East) and metamorphic Dalradian (West). Interestingly both metamorphic rocks are negatively correlated to igneous intrusive rocks which are common to East and West areas.

The test provided a second important outcome since it highlighted the role played by the geology in the overall landscape characterisation; in fact a difference in terms of landscape types between West and East was identified on the basis of a change from metamorphic Moine to metamorphic Dalradian rocks.

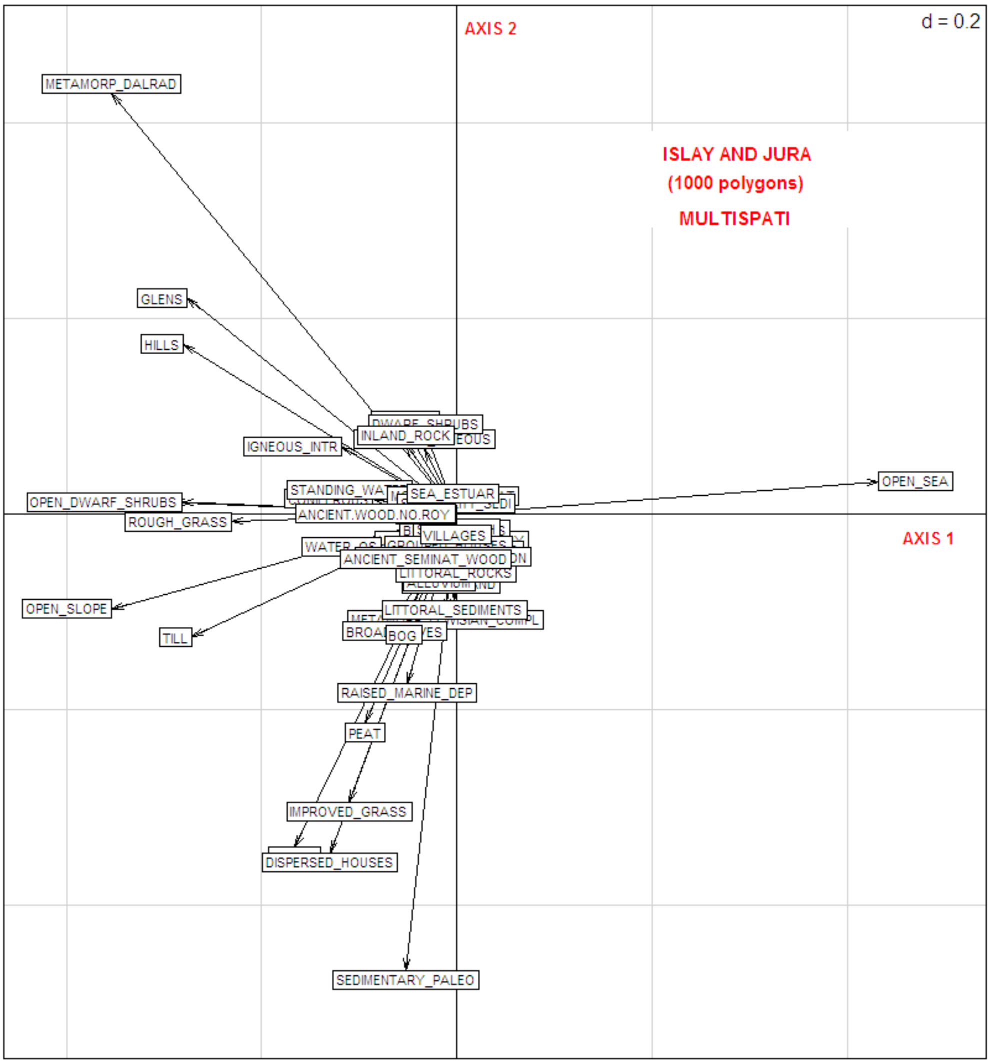
Finally also this test highlighted that the official LCA classified the landscape types of the Park generically, in particular the “upland and glen”. For the fourth time the GIS/MULTISPATI analysis identified a larger variety of landscape types within this LCA type.

7.6 What happens if the number of polygons changes (if the scale is modified)?

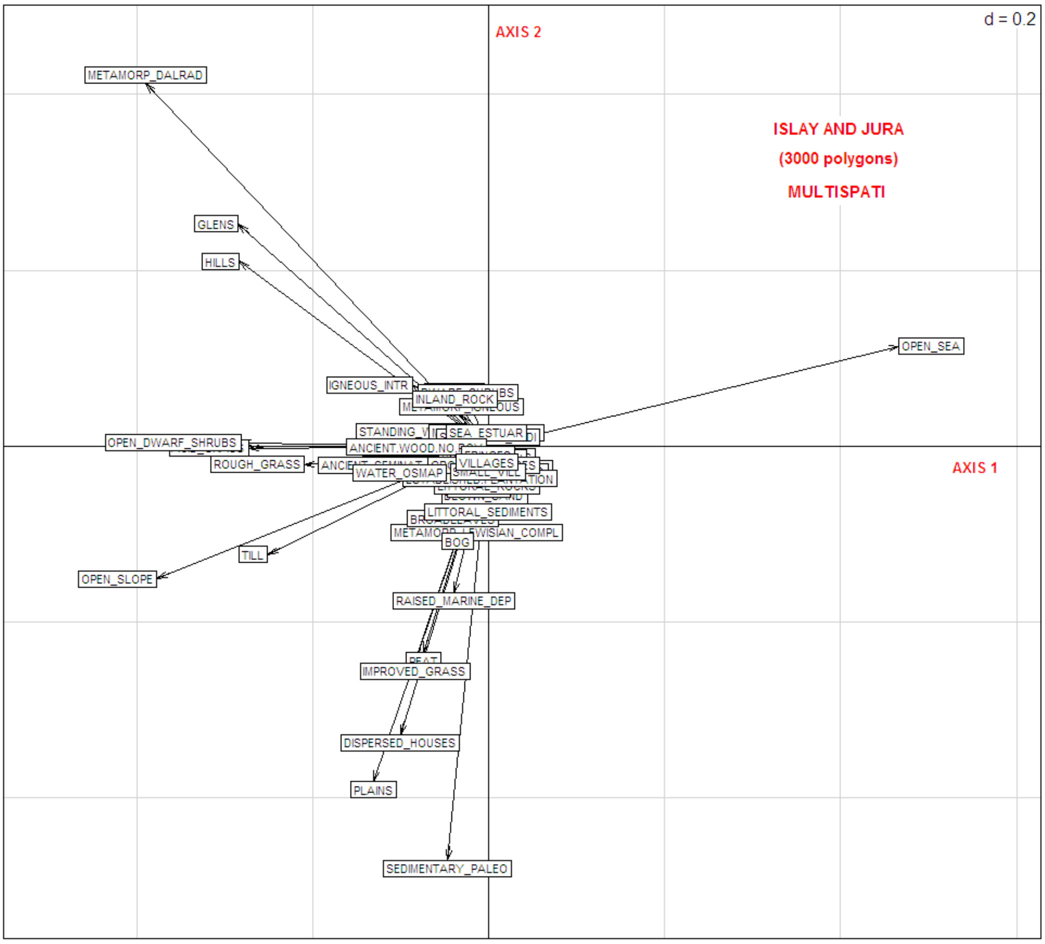
With the intention of exploring the way the GIS/MULTISPATI-based methodology reacts to a change in the scale of analysis, the number of Thiessen polygons was modified. On this basis, the results obtained from the landscape characterisation conducted at the fine scale of 3000 polygons were compared against the outcomes retrieved by using the coarse scale of 1000 polygons. The effect of scale on the landscape characterisation based on the integrated use of GIS and MULTISPATI was tested on all the areas of study, and here the isles of Islay and Jura are selected as explanatory example.

The two standard and spatially constrained PCA were carried out in R as explained in chapter 6 and the same operations were performed at 1000 and 3000 polygons. The attention is focussed on the graphs that help to better understand whether or not the scale has an influence on the results of the landscape characterisation. Note that since the interest is in the spatially constrained multivariate analysis, all the figures refer to the application of MULTISPATI-PCA 1000 and 3000 polygons.

The starting point is to determine how the landscape elements at both scales distribute along the first and second axes of MULTISPATI-PCA (see figure 7.44a and 7.44b).



a)



b)

Figure 7.44. The biplots show the distribution of the landscape elements on the first two axes of MULTISPATI-PCA. The difference between 1000 (45a) and 3000 (45b) polygons are minimal. In other words the spatial structure of the data looks very similar regardless the scale.

The biplots do not show relevant differences: in both scales the division between land and open sea is clearly identified by the first axis, while from reading the information carried on the second axis it emerges that the land is distinguished because of the geology and landforms. In fact the second axis separates two geological bedrocks, Dalradian rocks versus sedimentary paleozoic rocks, and a landscape characterised by hills and glens against one more undulating and flat (plains and open slopes). In terms of land-cover and settlement types it is more difficult to find a neat separation. Certainly there is an evident difference between dwarf shrubs and improved grassland, since the former are associated to the Dalradian rocks and the latter occurs in presence of paleozoic sedimentary rocks. Due to their location in the middle of the second axis, open dwarf shrubs and rough grassland characterise both landscapes of metamorphic and sedimentary rocks. In contrast the settlement types are identified mainly by dispersed houses which show a greater correlation with the sedimentary rocks rather than with the Dalradian rocks.

As a result, the biplot graph offers a first informative insight on the way data are correlated to each other. It is possible to know more about the distribution of the spatial patterns by examining the maps of scores depicted in figures 7.45 (top and bottom).

Figures 7.45

When comparing the first two columns of both set of maps (1000 and 3000 polygons) no major discrepancies are noticeable, whereas the third column, which represents the third axis of MULTISPATI-PCA, illustrates a diverse pattern. The reasons of these differences can be clarified by consulting the biplot again. Figures 7.46 (top and bottom) represent the biplots for Islay and Jura at 1000 and 3000 polygons. In both plots the x axis corresponds to the first axis of MULTISPATI-PCA, while the y axis represents the third axis of MULTISPATI-PCA.

Figures 7.46

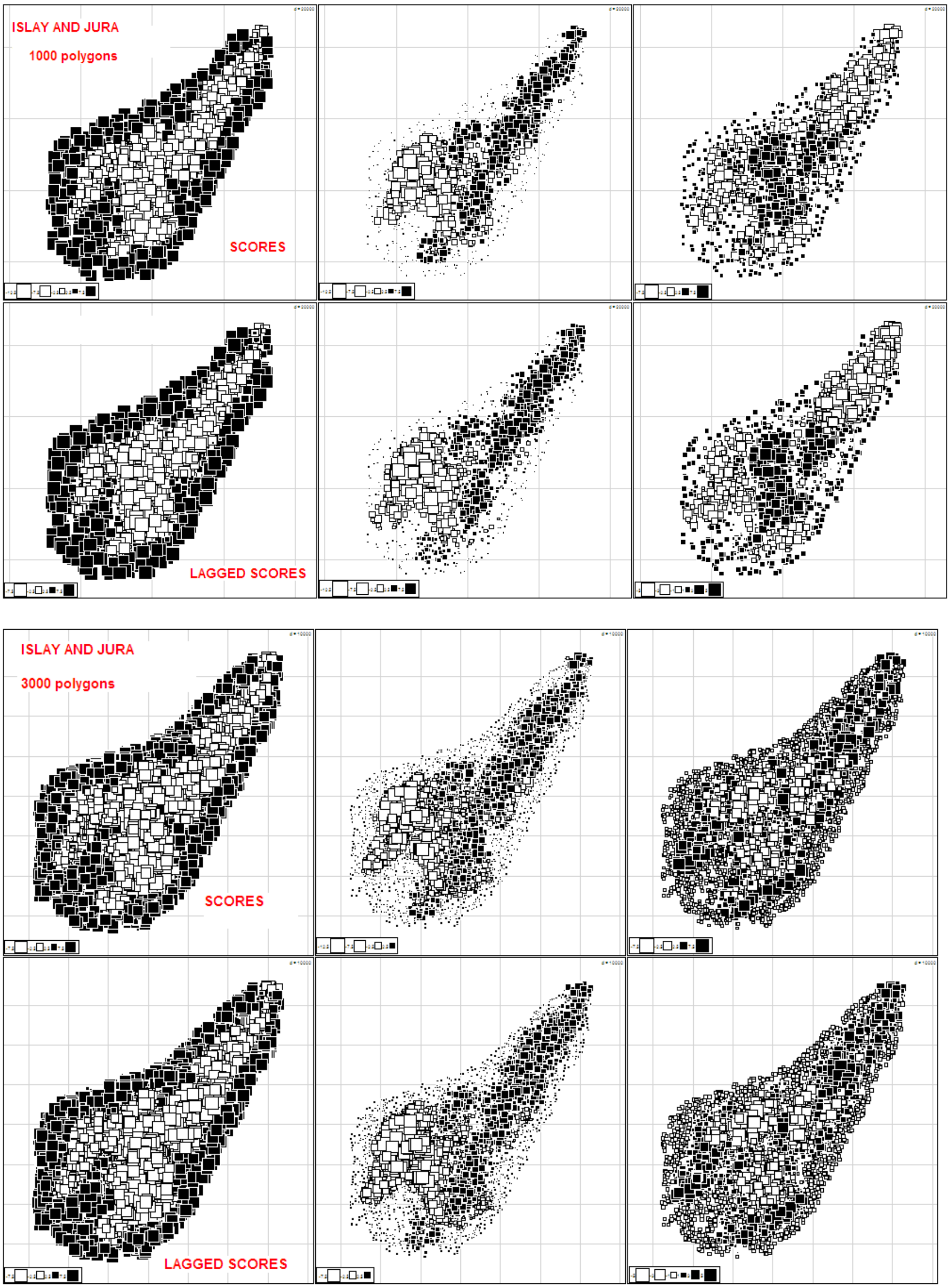


Figure 7.45. The factorial maps depicting the positive and negative values of the scores and lagged scores provide a first impression of the spatial patterns within the data. In the first column for both scale, 1000 and 3000 polygons, it is recognisable the difference between land (white squares) and sea (black squares). This reflects what is illustrated in figure 7.44 by the first axis of MULTISPATI-PCA. The second column shows the difference between sedimentary rock (white squares) and Dalradian rocks (black squares). Again, this was already evident in the biplot in figure 7.44. The third column is the most interesting because it shows the influence of scale in the landscape analysis. The differences in spatial structure can be explained by the way the variables and the loadings associate to each other.

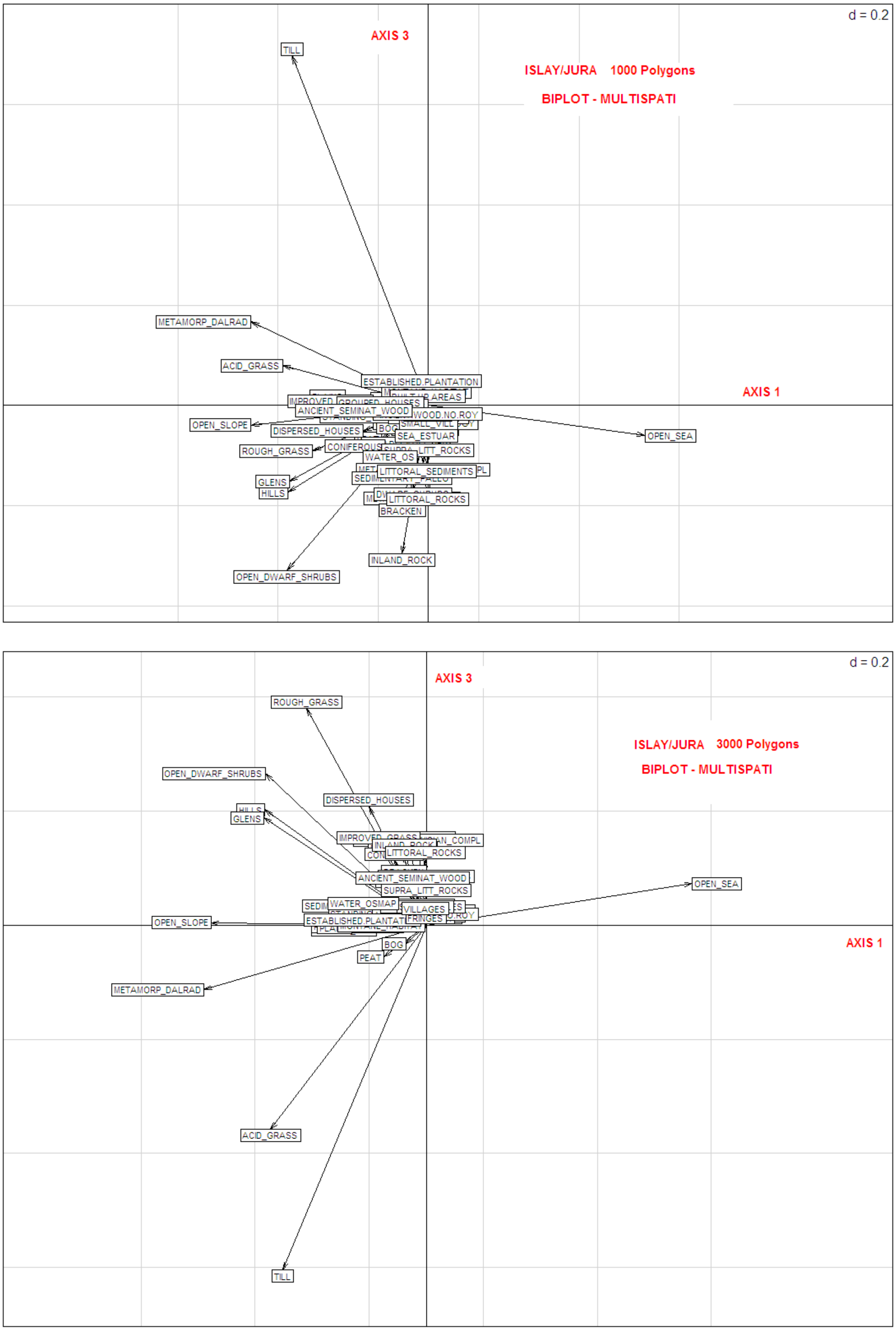


Figure 7.46. Biplots of the Islay and Jura at 1000 (top) and 3000 (bottom) polygon scale. These show the way the variables are distributed along the first and third axes of MULTISPATI –PCA.

It is noticeable that in both plots the first axis keeps on identifying and separating the sea from the land, while the third axis associate the variables differently: at 1000 polygons the group of till, acidic grass and metamorphic Dalradian rocks is distinguished from the more numerous group formed by open dwarf shrubs, inland rock, braken, rough grass, glen, hill-tops and open slopes. In contrast at 3000 polygons, the group of till, acidic grass, metamorphic Dalradian rocks, peat and bog is separated from that comprised of rough grass, open dwarf shrubs, dispersed houses, glen, and hill-tops. On this basis it is concluded that the change in the association amongst the variables, despite not being outstanding, determined the difference in the spatial structure of the data, which was registered in the third column of the map of scores (see figure 7.45).

The cluster analysis follows and translates the scores and lagged scores of MULTISPATI-PCA into information (landscape character types) that can be plotted in arcGIS. As explained in chapter 6, the landscape character types are derived on the basis of the highest similarity and lowest distance between the values observed.

The results from the cluster analysis are displayed in the GIS maps illustrated in figure 7.47. The first impression is that visually the landscape character types identified in the two maps are not remarkably different; in fact it is possible to associate the landscape type B, of the 1000 polygons map, with the type C, in the 3000 polygons map. Then the types E, B and G identified at fine scale seem to share the same location of the types C, F and G at coarse scale. The main difference is recorded in the northern part of Jura which, according to the analysis at 1000 polygons, is characterised only to the type D while on the basis of the analysis at 3000 polygons is characterised by the types D and F.

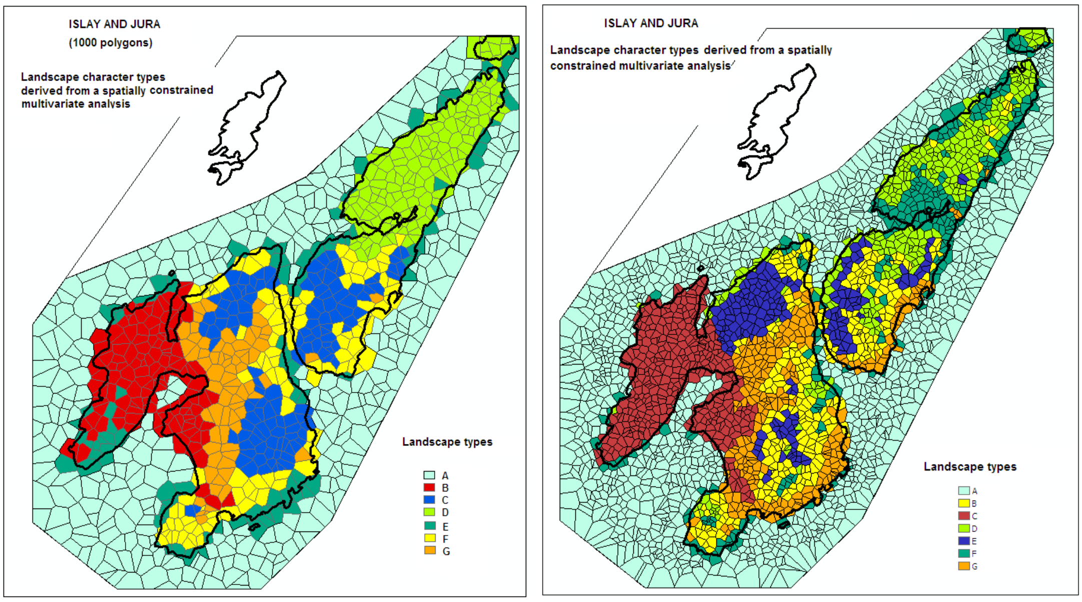
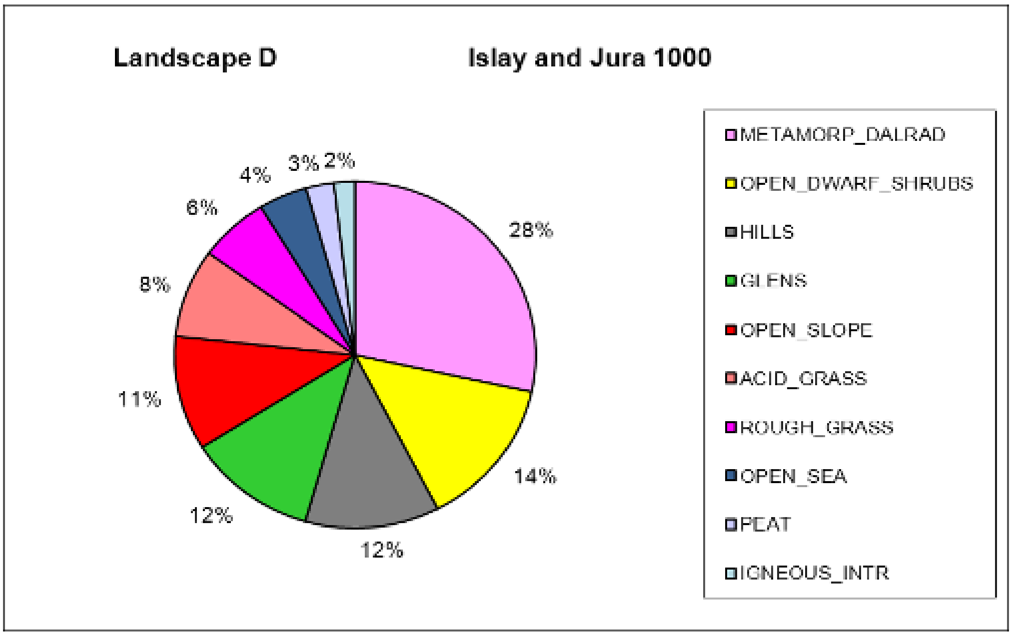


Figure 7.47. Comparison between the distribution of the spatial patterns at 1000 (left) and 3000 (right) polygons as identified in Islay and Jura with MULTISPATI-PCA. As it is evident, the main changes in pattern, due to scale, is more obvious on the isle of Jura.

As mentioned several times, the visual analysis of the maps is not scientific and can be misleading, therefore it is recommended to always quantify the landscape character types in terms of landscape elements composition and afterwards carry out a further comparison. For instance, the identified similarity between the types B and C at 1000 and 3000 polygons is verified by the pie charts in figure 7.48.

The percentage of the landscape elements and their occurrence in both types gives close results and clearly confirm that the same pattern is detected, hence it is possible to conclude that for this specific case, the change in the scale of analysis does not influence the landscape characterisation.



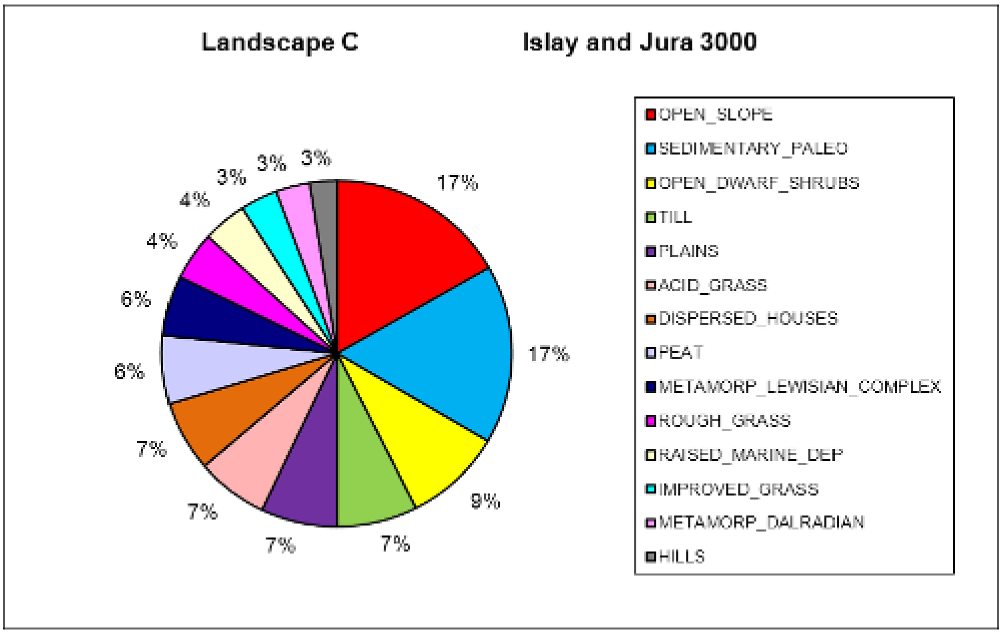
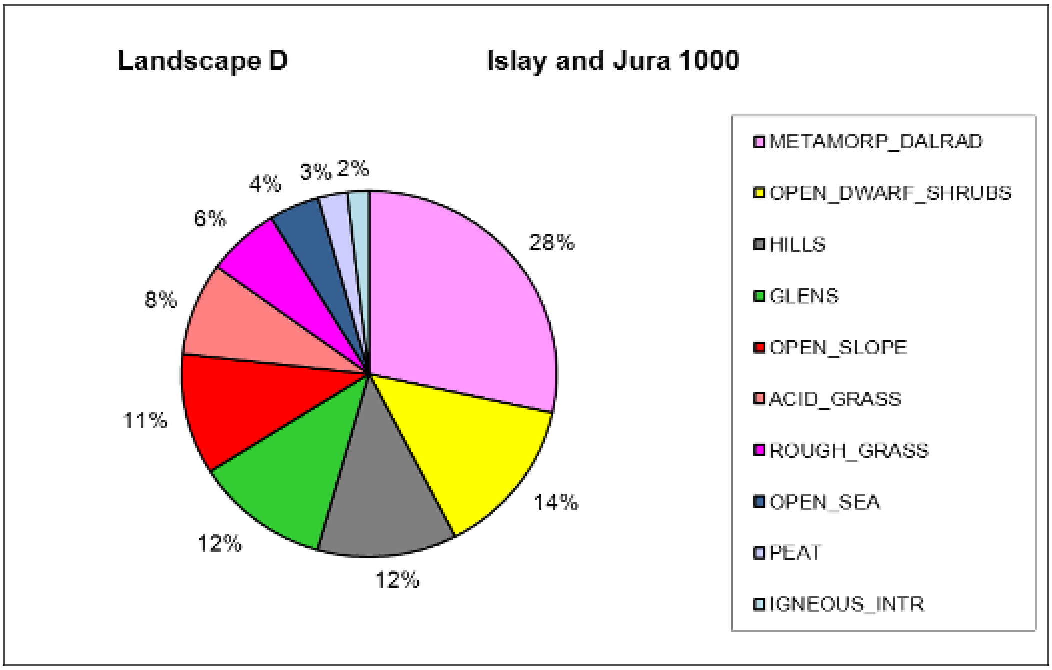
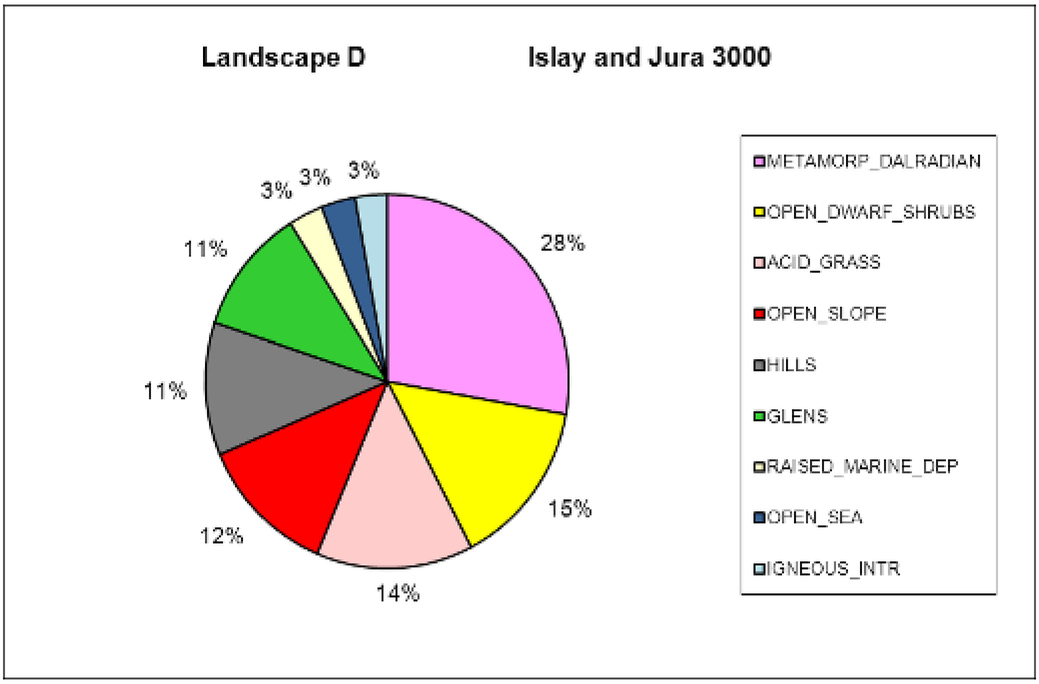


Figure 7.48. Pie charts illustrating the landscape element composition of the types B and C as identified at 1000 and 3000 polygons. The similarity is evident and highlights the fact that this specific character type is so well defined and strong that it is not affected by the change in scale.

A second example is provided by figure 7.49 that represents the landscape elements composition of the northern part of Jura. As pointed out before, for this part of the area of study the number of character types identified by the two analyses changes from one character (1000 polygons) to two character types (3000). The goal is to investigate whether differences in landscape elements composition occur.





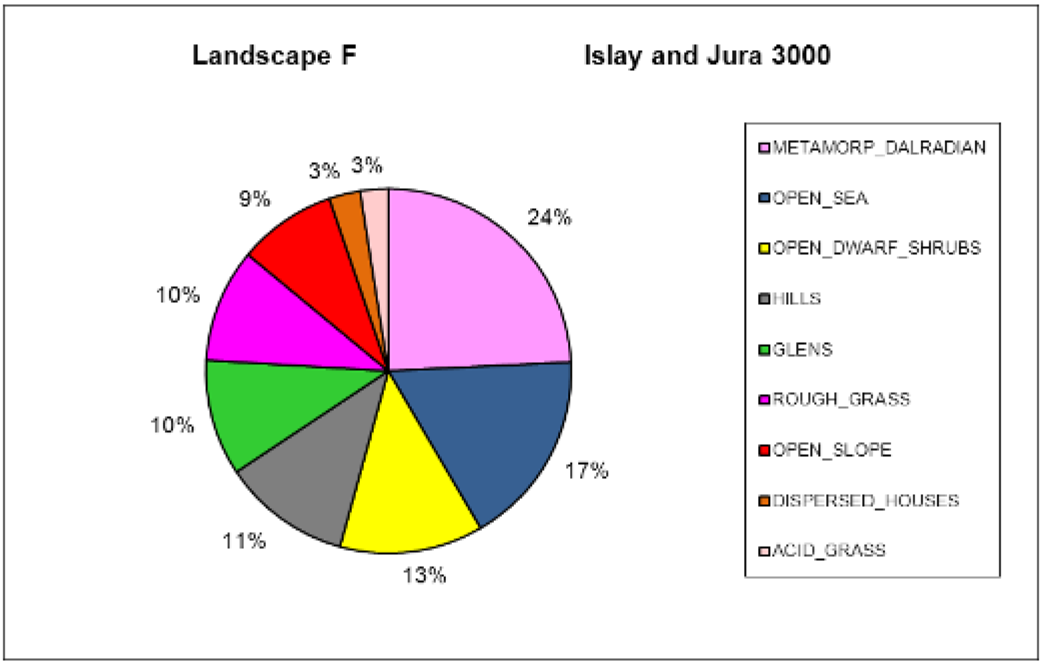


Figure 7.49. The pie chart illustrating the landscape element composition of the types D at 1000 polygons is compared to that of the types D and F identified at 3000 polygons and in the same geographical location (north of Jura).

The pie charts provided a positive answer. At the scale of 3000 polygons the analysis is able to differentiate a character type, D from the character type F. In the former there are 6 main elements: Dalradian rocks (25%), open dwarf shrubs (14%), acidic grass (13%), open slopes (12%), hills (11%) and glens (10%). In the latter, the landscape is comprised by 7 main elements: Dalradian rocks (23%), open sea (17%), open dwarf shrubs (12%), hills (11%) and glens (10%), rough grass (10%) and open slopes (8%). Apart from the difference in the percentage values, which overall varies in one or two units, the comparison between types D and F shows that open sea and rough grass are new variables and mark the difference in terms of character type.

Type D obtained with the landscape analysis at 1000 polygons shows a very interesting landscape composition because it is the sum of the main landscape elements of types D and F; in fact it comprises a majority of Dalradian rocks (27%) followed by open dwarf shrubs (13%), hills (11%), glens (11%), open slopes (8%), acidic grass (8%), rough grass (6%) and open sea (4%). Therefore it is possible to conclude that at coarser scale the size of the polygons is too big to detect details and consequently the subtleties between landscape character types are missed. This translates into an aggregation of landscape character types, here for instance acidic grass, rough grass and open sea are all together in type D while at 3000 polygons these elements are separated to form the type D and type F.

This trend was recorded for all the areas of study and it was also noticed that well defined and strong spatial structures were generally detected regardless the scale. This last outcome supported the idea that at coarse scale there is a tendency of obtaining character types that summarize the element composition within each type, at finer scale the differences are highlighted and contribute to the definition of new character types.

To conclude, it is possible to summarise in a few key points the results achieved by all the tests. First of all, the GIS/MULTISPATI-based methodology is demonstrated to have a robust and scientific basis and keeps consistency throughout the analysis of the data. The variation of the shape and size of the Thiessen polygons did not affect the final results of the characterisation, whereas change in scale highlighted that according to the distribution of the landscape elements and their degree of correlation, scale can be influential in the definition of the character of the landscape.

With regard to the introduction of more data to the original dataset, it was observed that their influence is significant and translates into new character types only when the new data are strongly correlated to other elements and also largely spread over the area of study. Similarly, the test on two neighbouring areas revealed that strong patterns could be identified easily along the boundaries of two contiguous areas of study, while weaker patterns are more difficult to identify. Once again the definition and identification of the character of the landscape depends on the way the elements are located in space and correlated to each other. In conclusion the tests clarified one fundamental point: although MULTISPATI-PCA enhances the effect of spatial autocorrelation, the characteristics of the original data play a main role. In fact if the variables are correlated and spatially dependent (high values of Moran's *I*), then MULTISPATI-PCA is able to detect with accuracy the spatial patterns within the data and identify the character of the landscape. Otherwise if these two conditions are missing then there is low confidence that the landscape character types identified by MULTISPATI-PCA are accurate and correspondent to the reality.

The last chapter of the thesis is entirely dedicated to summarise the main achievement reached by the research and to highlight the role played by the possible end-users.

CHAPTER 8

Listening to the end-users and overall conclusions

Fields of barley. East coast near St. Cyrus (I.Marengo).

8.1 Learning from the feedback of the potential end users

From the outset this study was considered to be applied research intended to deliver outcomes to end users including landscape architects, landscape practitioners and private landscape consultants working for SNH or more generally for local authorities. The GIS/MULTISPATI-based methodology was not designed to replace the work of the landscape architects, but rather it was implemented with the intention to be a useful tool to support them in the identification of landscape character types. However, how can we be sure that the methodology adequately meets this aim, and it is an efficient and helpful tool for the landscape architects and other practitioners? The best way to find an answer is to ask directly the end users. For logistic reasons it was not feasible to work jointly with them during the development and implementation of the methodology, nevertheless their opinions and comments were recorded and contributed significantly to conclude the study. Thus a group of possible end-users, comprised of SNH landscape architects and policy and advice officers, operations managers from RCAHMS and Historic Scotland, and the senior heritage manager of the Cairngorms National Park, were contacted and engaged in the final part of this research.

When the end users were approached, the primary aim was to understand whether or not GIS provided a clear, accessible and useful method of landscape analysis, if the outcomes looked plausible and if they could detect the weaknesses and strengths of the new approach to landscape characterisation. In order to reach this aim several meetings were arranged and a presentation of the research, at its different stages, was delivered. Each participant contributed effectively with constructive discussions and, most important, completed a questionnaire which was structured as follows:

1) Was the GIS approach to landscape characterisation interesting and clear?
2) Was the GIS approach user-friendly and would you consider using it yourself?

3) If the GIS approach is considered complex and difficult, in your opinion what would make it more user-friendly and more likely to be used?
4) During the presentation did you identify strengths and weak points in the GIS approach to the landscape characterisation? If so, what are they?
5) From your experience could you tell whether or not the GIS methodology can be useful and applicable to your job?
6) Could you please specify examples of where you would like to apply the GIS-based landscape characterisation?
7) Other comments?

The description and observations of the feedback and comments returned from the end users is described here below.

In general the GIS approach to landscape characterisation was found interesting and clearly delivered when it was presented. Some of the end users commented that they caught the general picture of how GIS operated, others pointed out the need for further explanation about the way the landscape elements were introduced in the analysis and contributed to the identification of the landscape character types. Finally, a few admitted to struggle with everything involved numbers and GIS in general. On this basis it emerged that the GIS/MULTISPATI methodology as currently implemented is not directly intuitive but results interesting thanks to some promising results. The end-users confirmed that more work is necessary in order to make the methodology more user-friendly.

With regard to simplifying the method of running the analysis, the end users had clear and specific ideas about the way of addressing this issue. First of all they proposed the preparation of a booklet as guidance which could explain how to input the data in the system; how the spatial autocorrelation and the multivariate analysis operate and how the selection of the number of landscape character types from the cluster analysis takes place. The end-users' suggestion was not a surprise since this study already assumed that not everybody has got skills in GIS and knowledge about the statistics applied. Therefore, where the research is applied to a real case study, then the preparation and release of an explanatory booklet with instructions and advice will occur.

A second suggestion was to identify the landscape character types with a more explicit indication such as a short title which would describe the types and provide a straightforward understanding of them. Currently the landscape character types are coded with alphabetical letters and numbers because these are the easiest reference to record data with GIS and MULTISPATI. Hence, it is only through the analysis of the pie charts associated to the landscape character types that it is possible to know about the landscape elements composition of each type and consequently describe them with a text.

The lack of a clearer and unambiguous descriptive identifier for the landscape character types was already noticed during the implementation of the methodology. To date, a solution has not been found yet, and it is hard to see how a character type can be automatically illustrated at the end of the GIS/MULTISPATI analysis. Thus, as described in chapter 7, at the moment the way of operating is to consult both the pie charts and the official LCA classification. In case there is an evident similarity between the GIS types and the LCA types it is recommended to use the official classes of character types. Otherwise, if the outcomes from the two methodologies are discrepant, it is necessary to find an appropriate title for the character type.

The end users were very attentive to the aesthetic of the GIS maps and a third request was made in order to improve the boundaries of the polygons, which provide an "artificial" and untidy picture of the landscape. This it has been attempted by smoothing the polygons and figure 8.1 illustrates the results obtained.

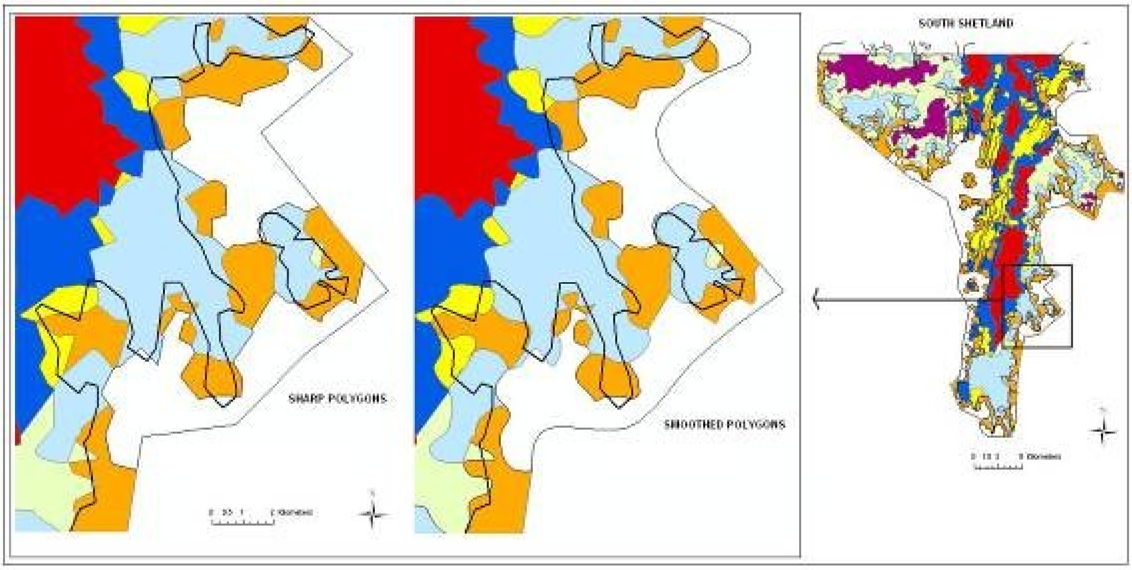


Figure 8.1. First attempt to ameliorate the aesthetic of the GIS map by smoothing the boundaries of the polygons. The two areas of South Shetland are the detail of the landscape types in the green square. The smoothing effect seems minimal and requires to be verified by the end users.

As it emerges from the picture, the map looks slightly improved. The problem of having polygons with sharp boundaries is not only an aesthetical issue but also conceptual. In fact sharp lines on thematic maps conventionally represent abrupt changes of phenomena. Here, since there is uncertainty about the change in the landscape character types, in terms that it may not occur exactly where the boundaries are, it would be more appropriate to use fuzzy techniques for representing the boundaries. Basically the fuzzy logic works in a way that the membership of a landscape character type is not absolute (yes or no), but ambiguous (possibly yes or possibly no). For instance if the final maps, which are shape files, are converted into raster it would be possible to attribute the membership of each pixel to the correspondent landscape character type with values that range from 0 to 1 and not with a binary code 0-1. Thus, for future applications, fuzzy techniques are recommended.

The end users showed some perplexity about the discrepancies between the landscape character types identified in the official LCA maps and those described by the GIS maps. As mentioned in chapter 7 the degree of similarity/dissimilarity between the two maps, and the two methodologies applied, varied from area to area. Some of the end users wondered whether the use of different scales and datasets could help in drawing landscape character types more accurately. This point was correct and highlighted one of the reasons[33] why the maps from GIS landscape analysis show diverse degrees of similarity with the official LCA map. As mentioned in chapter 7, according to the area of study and the defined scope of the LCA, each landscape element has a different relevance during the analysis. In other words, there might be areas where the landforms and the geology are particularly important and others where the size and the pattern of the fields and the orientation and typology of the settlements count more. It is clear that if the maps with this kind of relevant information are missing, then GIS and MULTISPATI identify the boundaries of landscape types on a partial analysis of the reality and the chances to have a map similar to the official one decrease.

A possible solution to this problem may be the definition of what elements are relevant for a correct and complete landscape characterisation of a given area prior to the start of the GIS/MULTISPATI analysis. Sometimes data can be derived from other already in use, and in this research the maps about landforms, historic land-uses and periods, and settlement types provided an example. Nevertheless, as pointed out by the end users it is necessary to double check the outcomes of the derivation process in order to be sure that the derived data keeps on being meaningful for practical purposes and respectful of the content of the original raw data. In the event that the data is unavailable it is a matter of acknowledging this gap when delivering the results of the analysis to the end users.

In addition to the GIS dependence on data quality and availability and the costs related to the data, the end users flagged up the lack of expertise and skills in GIS which make people afraid of using GIS for their work. As mentioned at the beginning,

[33] Another reason that explains the different degrees of equivalence between GIS and LCA map is the necessity of having landscape elements with simultaneously a strong spatial structure and marked occurrence/variability.

this remark was accepted and again it confirmed one of the already known disadvantages of using GIS. If there is little familiarity with these tools, it is unavoidable that they look complicated, difficult and unfriendly. Therefore besides having a booklet that explains how to input the data and how the statistics work and process the data, it is important to train people and teach them the essential tools and the philosophy of GIS. Supporting documentation and training courses should be a solution to this issue along with the idea of persuading the landscape architects to work with GIS analysts and achieve synergy at work.

The end users had also positive comments on the results retrieved by using GIS/MULTISPATI-based methodology. They pointed out that GIS offer an objective and consistent analysis of the landscape and minimise the impact that the judgments and the interpretation due to professional expertise, familiarity and knowledge about a landscape have on the outputs from the landscape assessment.

In addition GIS received appreciation because they incorporate data from different sources and specifically the end users with expertise in history were quite pleased to see the cultural heritage considered as important as the natural heritage and a relevant element for the landscape characterisation. The recognition that the GIS methodology carries out the landscape analysis from a holistic point of view was also highly valued. Moreover the end users acknowledged the capacity of GIS to be easily updatable and to perform the analyses quite quickly if compared to the current way of performing the LCA.

The final suggestions and comments collected from the questionnaire were about the possible uses and applications of the methodology to real projects. The suggestions from the end users in this case were varied. Some of them clearly stated that an application of GIS to their current job was far from being feasible because of the complexity of the tool; however they mentioned other areas and sectors, such as development planning and development management, where the GIS-based methodology could find adequate and advantageous application. Others advised on trying and seeing the methodology applied either to landscape characterisation within the local authority boundaries and for specific topics or to a broader scale, such as the whole Scotland, for more general themes. In both cases it was evident the suggestion of exploring further how the GIS methodology copes with different scales. Finally some end users proposed to apply the GIS-based landscape characterisation as a tool for informing plans for designation areas, for land-use and landscapes policies, and for selecting landscape objectives in line with the European Landscape Convention.

In general the end users provided constructive feedback and several hopes that this research can be useful and applied to real case studies in future. The record of the end-users' thoughts was the starting point for a series of questions that inevitably come at the end of any kind of research. Looking back at the last three years we can ask whether or not the work carried out reached its targets, if the time dedicated to the investigation contributed positively to the scientific and general community, if the methodology implemented and developed is innovative and if the results achieved are meaningful and useful not only within the context of this research but also for a wider range of people and possible end users. The following chapter is meant to find an answer to all the questions by trying and summarising the work carried out, by describing the way it evolved from its beginning to its end, by highlighting strengths of the methodology and by understanding how to improve the weak point in order to ensure that this research has got a future.

8.2 The aims achieved after three years

In order to clarify whether or not the research has succeeded in its intent, it can be helpful to start with a reminder of the aims and reasons of doing this research. As explained in the first two chapters, landscape is important and matters because it is not only a natural and cultural resource but also the place where people recognise themselves and associate their memories and the everyday life. Hence the landscape is a physical and abstract entity where the natural elements come together with colours, sounds, feelings and perceptions. This idea of landscape was formalised in 2000 when 19 European states signed the European Landscape Convention (ELC), a formal act that intended to promote the protection, management and planning of the European landscapes.

Previous to the official definition of landscape by the ELC, the landscape was already taken into account and at the centre of the attention of planners and policy makers. As illustrated in figure 8.2, it is possible to draw a sort of chronology about the approaches to the study of the landscape as they evolved and developed throughout the last century. The 70's are identified with the period of landscape evaluation during which the landscape was analysed according to its value and the idea was to consider the landscape on the basis of its quality, and therefore decide if one landscape was better or worse than another. However, after more or less a decade, it became evident that this approach was not fruitful and noticeably restrictive since only few landscapes could be effectively protected and conserved. Therefore the step forward was to consider the landscapes on the basis of the unique and distinct traits that differentiate them from each other. As a consequence the focus of the landscape analyses shifted from the value of the landscape to its character.

Figure 8.2

In the UK, England and Scotland launched a programme in the mid of the 90's called Landscape Character Assessment (LCA) which was undertaken by the local authorities and the national parks. The LCA was born with the intention of being a valid support tool for the planners, developers, decision and policy makers. The LCA comprises of two stages: the first is addressed to get people aware of the landscape and of what makes it different from the surrounding ones. The second stage is centred on the judgements that are required in order to prescribe landscape management recommendations which will guide the future of the landscape towards the conservation and enhancement of its character.

Landscape EVALUATION

(1970)

"One landscape is better and then more valuable than another"

Landscape ASSESSMENT

(1980)

"Landscapes are different from each other but not better or worse"

Landscape CHARACTER ASSESSMENT

(1990)

"The character of the landscape is a distinct, unique and recognisable pattern of elements in the landscape"

EUROPEAN LANDSCAPE CONVENTION

(2000)

"The landscape is an area, as perceveid by people, whose character is the result of the action and interaction of natural and/or human factors"

Figure 8.2. The changes in the way policy makers, planners and landscape practitioners intended and approached the landscape throughout the 20th century.

After being used for more than a decade by local authorities, landscape architects or private landscape consultants, concerns have risen about the LCA and specifically about the way the characterisation phase is structured. The most evident weak points are the lack of consistency and robustness and the fact that the results are not quantitative. From the LCA reports it is evident that the landscape character types and areas are basically described to people in a narrative form and inevitably the descriptions reflect the personal perception that the assessors had about the landscape. Therefore, the content of the reports appears arguable, subjective and not sufficiently scientific to face debates or public enquiries about planning issues.

The main obstacle is that planners and developers deal generally with numbers and to a certain extent they require to see the landscape from a numeric and quantitative point of view, so that the impact of proposed plans can be not only described but also visualised and measured in terms of changes in the character of the landscape.

In addition, this research had to face other conceptual and technical problems connected not only to finding a solution for the LCA weaknesses but also to the way of approaching the analysis of the landscape. The main difficulty was due to the fact that the landscape is such a generic theme that plays a relevant and influential role in many and diverse disciplines which, from their side, tackle the landscape analysis from opposite points of view. From the literature reviewed many dichotomies emerged, for example between ecological and social sciences, between the holistic and reductionist theories, between the objectivism and subjectivism supporters, to name a few.

The presence of such different multi-approaches to the same topic generated confusion and made the investigation on the character of the landscape more difficult and complex than what was expected at the beginning. Moreover this research had the challenge to verify whether or not the use and application of GIS to the landscape analysis was feasible and fruitful. GIS were thought to be the solution to the LCA weaknesses because of the numerous examples of successful applications recorded from the 90's, when the use of GIS as mapping and analytical tools started to increase exponentially.

With the confidence that GIS analytical capacities could be exploited in support of the LCA a series of points were annotated and worked as starter for the definition of the aim of the research, that is the development and implementation of a methodology for landscape characterisation based on the analysis of all the elements that comprise the landscape through GIS (see figure 8.3).

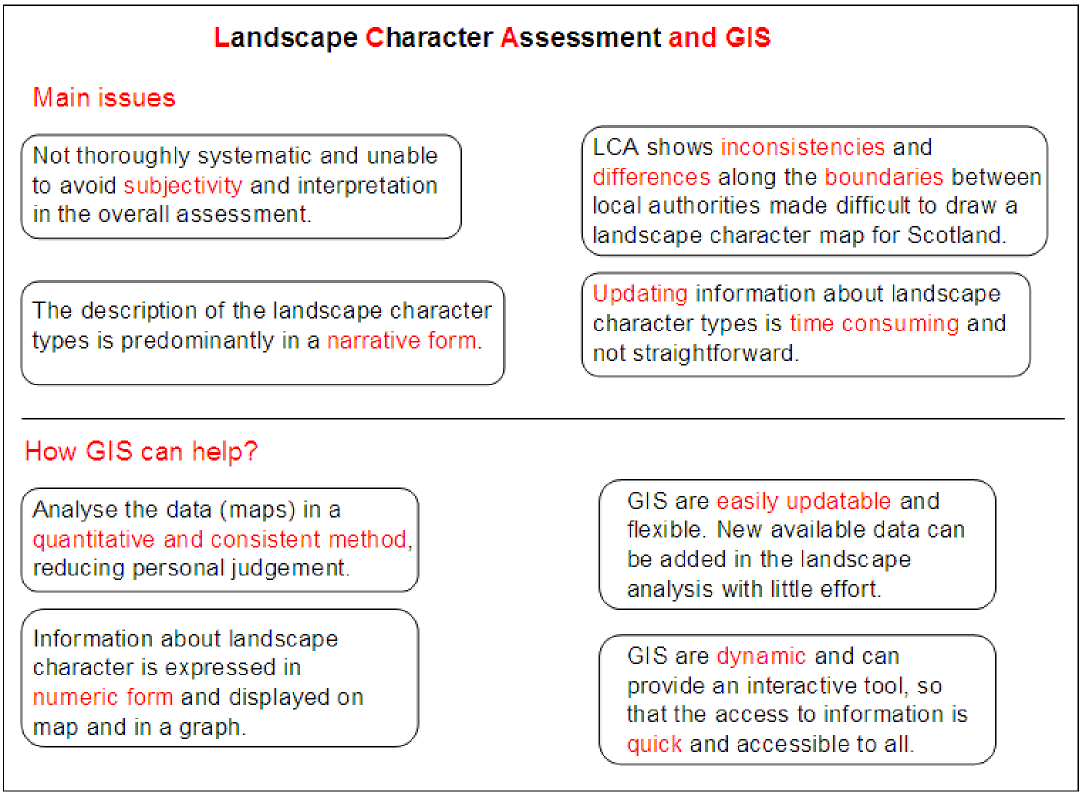


Figure 8.3. Some ideas about the way the GIS can help to solve the most evident weaknesses of the LCA.

The ways GIS can be profitably employed to overcome the main LCA issues provided the basis on which design the methodology for landscape characterisation. This was shaped to consider the landscape as a whole, marrying the holistic theory and to be quantitative and repeatable in order to meet the targets of objectivity and consistency.

It has to be remembered that during the early 90's there had been attempts of using GIS as support tools for the LCA; nevertheless, in every occasion, GIS were employed either as mapping tools, to show and display the maps of landscape character types and areas, or as editing tools, to draw the boundaries of the landscape character types and areas that the landscape architects had designed on paper during or after the field survey. Therefore, to a certain extent people were aware of GIS and their potential contribution to LCA and the idea of introducing more GIS into the LCA process was already widespread, but the scarce knowledge and familiarity with GIS amongst the local authority officers and landscape architects, prevented the use of GIS as analytical tool.

This research intended to demonstrate whether or not GIS could be effectively used as analytical support tools (thus not only for mapping or editing) for the LCA. As a consequence here GIS are not required to produce only maps of landscape character types but primarily they are asked to use official maps in order to detect the character of the landscape. Chapters 3, 4 and 5 extensively described that the research reached the definition of an innovative GIS-based methodology for landscape characterisation after a long process characterised by the use of various techniques of analysis of spatial data.

Throughout the investigation it became apparent that GIS on their own could not give a robust analytical support on which develop the methodology but they needed to be integrated and enhanced by more sophisticated and complex statistics since these could deal better with the analysis of the landscape elements. The choice of the most suitable statistical technique was not straightforward and followed an approach characterised by several attempts whose failures and successes helped to make the final decision. A summary of the experimentation of the statistics and method of analysis is summarised in table 8.1

Techniques	Positive points	Negative points	Final decision
ArcGIS spatial statistics tools	Straightforward application and clear summary of the outcomes.	It calculates only a global Moran's *I* and consider a variable at time.	rejected
FRAGSTAT	It provides a large number of metrics from which detect the spatial patterns.	Both analyse the maps individually, without considering that landscape is an association of variables from different maps.	rejected
TOSS	It bases the identification of spatial patterns on measurable criteria.		
GeoDa	It has a user-friendly interface to quantify global and local spatial autocorrelation.	It performs only uni and bi-variate analyses.	rejected
MULTISPATI-PCA	It calculates spatial correlation and multivariate analysis together. It is free.	It requires a stronger statistical background and expertise in R.	Approved and applied to the methodology

Table 8.1 Summary of the analytical techniques of landscape characterisation experimented throughout the PhD. According to the positive and negative aspects, concisely described, MULTISPATI-PCA was selected as the most suitable technique to meet the aims of this study.

Initially FRAGSTAT and TOSS were tried in order to identify the spatial patterns which, in statistical terms, correspond to the landscape character types. In both situations, the techniques were unable to identify the character of the landscape from all the maps collected. Instead the analysis was possible only on a map at a time which was not the purpose of the research.

Subsequently the software GeoDa was tested, and this was particularly helpful because it introduced the concept of spatial autocorrelation but it also showed great limitations because it was impossible to compare more than two variables at time.

At the end, FRAGSTAT, TOSS and GeoDa remained only experiments and were not employed for the rest of the research; nevertheless they played an important role because improved the knowledge about spatial statistics and above all paved the way to the statistical technique that was finally adopted, which is MULTISPATI-PCA.

The research challenged the analytical capacities of GIS and discovered the importance of the statistics as powerful tool able to enhance the analytical capacities of the GIS. Since the application of spatially constrained multivariate statistics was the first attempt ever tried in the field of landscape characterisation, it is opportune to clarify that the methodology developed at the end should not be considered as “the” only possible solution to the application of GIS to landscape characterisation. Perhaps there are other ways to apply GIS and statistics to landscape characterisation. Certainly, this methodology is innovative and introduces effective concepts and a functional approach to the analysis and identification of the character of the landscape. Therefore, compared to three years ago, it is undeniable that this research made evident progresses which are entirely supported by the plausible results obtained. As described in chapters 6 and 7 the statistics assured the research about the significance and accuracy of the calculations while the direct assessment of the landscape added to the positive comments expressed by several experts validated the final maps and allowed the research to be considered a success.

The question posed at the beginning, “is the use and application of GIS feasible and fruitful for the landscape characterisation?” has then received a positive answer. The methodology applied over eleven contrasting Scottish landscapes is illustrated in figure 8.5 and showed that GIS with the integration of spatially constrained multivariate statistics can be very useful tools because they offer a quantitative, unbiased, and repeatable analysis of the landscape elements and allow the character of the landscape to be defined in numeric and descriptive terms. By leaving little space for subjectivity and personal interpretation about the real landscape the GIS methodology made the landscape characterisation more objective, dynamic and flexible.

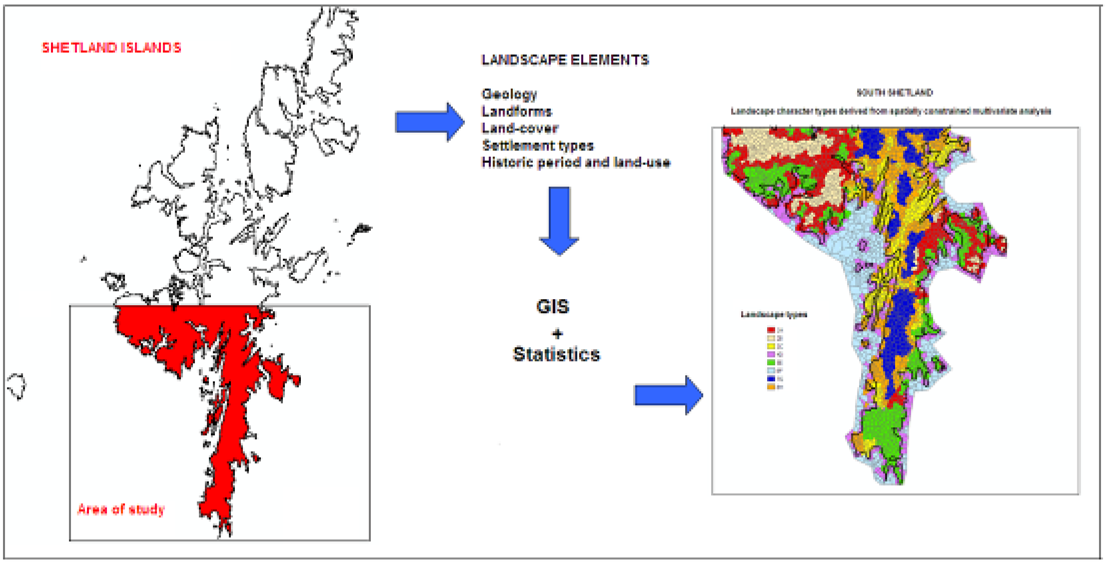


Figure 8.5. Example of the implementation of the methodology for landscape characterisation, which starts with the definition of the area of study, continues through the GIS/spatial statistics analysis, and ends with the map of landscape character types.

An additional strength of the implemented methodology is that it does not alter the final purpose of the LCA but possibly it enhances it. The idea at the basis of LCA is to try and ensure that all landscapes are managed sustainably for the future generations regardless of their location, designation and perceived quality. The LCA is not a tool against development but, on the contrary, it is a support for a better one which gives awareness of the character and sensitivity of the landscape. Planning the development of an area by taking into account the outcomes of the LCA process means essentially to consider the character of the landscape and that it should not be changed, but rather strengthened or maintained.

When deciding in favour or against any strategic plan or development policy, numbers or statistically/scientifically proved figures add robustness and credibility to the assertions and definitions about the landscape, its character and its importance as natural and cultural resource. By adopting this methodology it is possible to provide the planners and policy makers with both descriptions and measures about the landscape character types. In addition, updating the datasets through GIS is a much quicker process and the access to information is available to everyone along with the possibility to add or delete data from the analysis, which is always feasible and not too lengthy.

8.3 Those weak points that can be improved

Certainly, at the moment the methodology is structured in a way that end users have to know how GIS and the applied spatial statistics work. As previously mentioned, the end users were the first to express some concern about the complexity of the techniques used to develop the GIS/MULTISPATI-based methodology. The lack of GIS skills and knowledge about the statistics can be a major problem and might hinder the application of GIS in the short term. It is recognised that in order to make the methodology more accessible and familiar to everyone it would be necessary to automate many procedures and make them more user-friendly by realising an interface within the software arcGIS and by creating a link with R which would run the code written for the spatial statistics.

However, even if the GIS-based methodology for landscape characterisation is automated and can be run through a simple "click" with the mouse, it would be always necessary to train and teach people about the basic tools of GIS, the theoretical framework of the methodology and the meaning of the statistics applied. By knowing how GIS and MULTISPATI perform and what they analyse is a way of reducing the amount of errors and inaccuracies in the final data.

In the first chapters it was outlined that GIS are powerful tools able to deal with a large quantity of information and that are able to read many maps by overlapping them and crossing their geographic data. In contrast to this great advantage is the fact that the results of GIS analyses depend on the quality and the quantity of the data input which, in case of gaps or poor level of accuracy, become limiting factors. Nevertheless as mentioned in chapter 4, the lack of maps may not always be an obstacle to the analysis; in fact if criteria, which help to measure and quantify the missed information from other data sources, are defined then it is feasible to derive maps.

Undoubtedly the process of deriving maps from raw data or other datasets is more complicated if the aim is to represent qualitative data, such as the aesthetic and perceptual aspects of the landscape. Thus, contrary to the current LCA and the advantage of carrying a visual landscape assessment, the GIS-based analysis was applied only to measurable landscape elements that are ensured to be “mappable”. Whereas, it was possible to make only attempts in order to map aesthetic aspects of the landscape, and here the main difficulty encountered was to find the set of criteria able to describe quantitatively concepts such as openness, texture, balance and proportion (the appendix provides examples of the attempts made). Since there were many doubts about the validity and credibility of the maps derived, the outcomes from the attempts were not used for landscape characterisation, but were a valuable exercise and evidence of how difficult it can be dealing with landscape characterisation in the absence of an important means such as our eyes.

If compared to the official LCA maps, the GIS maps looked either similar or dissimilar according to different areas of study. As explained to the end-users, and described in chapter 7, the *procrustes* statistical technique revealed that in fact there was a high correspondence between the maps, although visually some of them looked remarkably different. The attempt to quantify the degree of similarity/dissimilarity through statistics is more scientific than the visual comparison and therefore should be trusted.

In this regard it is worth noting that the visual comparison between LCA and GIS maps can be misleading because it does not recognise that at the basis of the two maps there are two different approaches and ways of investigation. First of all different groups of consultants and landscape practitioners were involved in the LCA programme and each one followed its own techniques of assessment to meet the requests made by the local authorities. Hence according to the scope of the LCA either only few landscape elements, identified as the key characteristics, contributed to the identification of the character of the landscape or the entire set of elements was considered relevant to the definition of character. In contrast, the GIS/MULTISPATI-based methodology always considers all the elements on the maps, without making a selection, but indeed the quantity of elements analysed depends on the availability of the maps.

From the GIS/MULTISPATI analysis it clearly emerges that the landscape character types derived from a well defined spatial structure are always detected and consistently identified. Thus it appears that the way the input data are correlated to each other and distributed over the area of study influences the identification of spatial structures within the data analysed. In fact the results demonstrated that when the landscape elements formed recognisable and strong spatial clusters, then external effects such as scale, polygons distribution and shape, new data input and boundaries across the character type had overall a negligible effect on the identification of the landscape character. Alternatively, in areas of study where the landscape elements occurred in large variety and showed a weak correlation to each other, the resulting spatial patterns were less defined and more influenced by the variation in scale, data entry, and presence of boundaries and polygons distribution.

Therefore, while comparing the LCA and GIS maps, these relevant details should be taken into account along with the fact that the landscape character types correspond to the clusters defined during the cluster analysis where a personal judgement is involved in the analytical process. As noticed in chapters 6 and 7 the choice about the number of clusters relies on reading jointly the dendrogram and the associated table of the similarity and proximity values.

More than once it has been stressed that one of the strength of the GIS methodology is to support landscape characterisation with an analysis which is quantitatively based and always repeats the same structure; hence objectivity and consistency are both achieved. This is entirely true, but throughout the chapters it emerged that the personal judgement was sometimes unavoidable and the choice about the number of clusters to retain is an example. Another case is the definition of the thresholds for the classification of landforms and settlement types that can be arguable and still modified according to geographic areas. For instance England or Wales, due to the different morphology of the terrain, may require other thresholds.

A slightly different situation was described in chapter 5, where the decision about the rook/queen contiguity model for the neighbourhood analysis and the division of the areas of study in Thiessen polygons and not regular grid were backed by more scientific and indisputable reasons.

In conclusion, the methodology shows limits that, if not structural, can be overcome. For example, it is discouraged to carry out a visual comparison between the LCA and the GIS maps to assess the quality and credibility of the landscape character types identified by using GIS and MULTISPATI. Thus, it is suggested either to rely on the statistics or to validate the map through ground truth and the help of experts. It is recognised that presently the methodology is not able to analyse the aesthetic and perceptual elements of the landscape, however the challenge of including these qualitative aspects in the GIS/MULTISPATI-based methodology is not impossible but postponed and it may be the topic for another research. Finally, training courses and an increased availability and quality of the datasets should be achievable in the future and both should make the methodology more user-friendly and facilitate the access to information.

8.4 Looking towards the future

On the basis of what has been explained so far, it emerges that the new methodology introduces technical and conceptual innovations to the entire process of LCA. For the first time the first law of geography and the theory of spatial autocorrelation are applied successfully to landscape characterisation, and GIS are used both as analytical and mapping tools.

For the future the starting point is the feedback provided by the end users; in fact from the comments and suggestions received it is possible to visualise the next steps that should help the methodology to improve and create the conditions for its use in practical case studies.

First of all the whole process needs to be simplified and made more accessible and user-friendly. Moreover care should be taken to the way the final maps are presented and submitted. Secondly it would be beneficial to work closely with the end-users and shape the methodology according to their needs. For instance it seems important to know if the key characteristics, instead of all the elements in the landscape, assume a major role during the landscape characterisation. In fact with this indication it might be possible to think of criteria able to measure and quantify those elements that count most in the landscape. Finally attention should be paid to other landscape researches and see whether this methodology can be integrated to other techniques and improved. For example it might be interesting to investigate more in the use of voxel model applied by Pyysalo et al. (2009) and Washtell et al. (2009) in viewshed analysis for the visualisation and classification of landscapes using respectively airborne laser scanner data and geomorphometrics.

This research succeeded for two reasons: first of all it was a rewarding learning process and experience marked by personal improvement in the knowledge about GIS, statistics and landscape analysis. Secondly it managed to deal with the initial challenge without losing its track. It concluded by formulating a possible answer: with the results achieved the research demonstrates that GIS coupled with multivariate statistics can be used for landscape characterisation.

REFERENCES

Abrahamsson K.V., 1999, Landscape lost and gained: on change in semiotic resources, *Human Ecology Review*, 6:2, 51-61

Ahern J., Carr E., Hamin E., Glassberg D., 2005, People and Places on the Outer Cape: A landscape character study, *EDRA/Places awards Research,*28-31.

Anselin L., 2003. *GeoDA™ 0.9 User's Guide.* Spatial Analysis Laboratory. Department of Agricultural and Consumer Economics. University of Illinois, Urbana-Champaign. Urbana, IL61801 and Centre for Spatially integrated Social Science. Copy revisited June, 2003. Luc Anselin all right reserved.

Anselin L., 2005, Exploring spatial data with *GeoDA™:* a Workbook. Spatial Analysis Laboratory. Department of Agricultural and Consumer Economics. University of Illinois, Urbana-Champaign. Urbana, IL61801 and Centre for Spatially integrated Social Science. Copy revisited March 06, 2005. Luc Anselin all right reserved.

Antrop M. and Van Eetvelde V., 2000, Holistic aspects of suburban landscapes: visual image interpretation and landscape metrics, *Landscape and Urban Planning*, 50, 43-58.

Apan A.A., Raine S.R., and Paterson M.S., 2002, Image analysis techniques for assessing landscape structural change: A case study of the Lockyer valley catchment, Queensland, Australia, *Landscape and Urban Planning*, 59, 43-57.

ArcGIS desktop 9.2 help. Accessed in 2009
http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Density_calculations

Arriaza M., Cañas-Madueño J.F., Cañas-Madueño J.A., Ruiz-Aviles P., 2004, Assessing the visual quality of rural landscapes, *Landscape and Urban Planning*, 69, 115-125.

Ayad Y., 2005, Remote sensing and GIS in modelling visual landscape change: a case study of the north-western arid coast of Egypt, *landscape and Urban Planning,* 73, 307-325.

Bell S., 2002, Reading the landscape. Understanding the landscape, the process that formed it and the pressure for change.

Berengo C. and Di Maio S., 2008, *We are the landscape*, Giunti progetti educative, Prato – Firenze.

Bibby P., Shepherd J., 2004, *Developing a new classification of urban and rural areas for policy purposes-the Methodology*.
http://www.statistics.gov.uk/geography/downloads/Methodology_Report.pdf

Bocco G., Mendoza M., and Velazquez A., 2001, Remote sensing and GIS-based regional geomorphological mapping – a tool for land use planning in developing countries, *Geomorphology,* 39, 211-219.

Boots B., 2003, Developing local measures of spatial association for categorical data, *Geographical Systems*, 5, 139-160

Brabyn L, 1996, Landscape classification using GIS and national digital databases, *Landscape research*, 21, 3, 277-300.

Brabyn L., 2005, Solutions for characterising natural landscapes in New Zealand using geographical information systems, *Journal of Environmental Management,* 76, 23-34.

Brabyn L., 2009, Classifyin landscape character, *Landscape Research*, 34(3), 299-321.

Bullen J., 2010, *LANDMAP: recording and mapping the landscape of Wales,* Landscape Character Network. Issue 34 – Winter 2010, 11-13.

Carman D., 2007. *Isn't it time fro an approach to landscape, townscape and coastal seascape assessment on which everyone agrees?* Landscape Character Network News. Issue 25. Spring 2007. pg 8

Carver S., Comber L., Fritz S., McMorran R., Taylor S., Washtell J., 2008, *Wildness study in the Cairngorms National Park*, Work commissioned by the Cairngorms National Park Authority and Scottish Natural Heritage. University of Leeds.

Chessel D., Ollier S., Dray S., 2003, Ordination sous constrainte spatiale, http://pbil.univ-lyon1.fr/R/stage/stage8.pdf

Cosgrove D., 1984, *Social formation and symbolic landscape*. Barnes and Noble Books, New Jersey.

De la Fuente de Val G., Atauri J.A., and De Lucio J.V., 2006, Relationship between landscape visual attributes and spatial pattern indices: A test study in Mediterranean-climate landscapes, *Landscape and Urban Planning,* 77, 393-407.

De Smith M.J., Goodchild M.F., Longley P.A., 2006-2009, *Geospatial Analysis. A comprehensive Guide to principles, techniques and software tools. 3rd edition.,* Matador, Leicester.

Deng Y.X., Wilson J.P., and Bauer B.O., 2005, DEM resolution dependencies of Terrain Attributes across Landscape Typologies, *International Journal of Geographic Information Science,* 80-118.

Dikau R., Brabb E.E., Mark R.M., 1989, *Landform classification of New Mexico by computer*, U.S. Department of Interior, U.S. Geological Survey. Open-File report 91-634.

Dray S. and Dufour B., 2007, The ade4 Package: implementing the Duality Diagram for Ecologists, *Journal of Statistical Software*, Volume 22, Issue 4.

Dray S., 2008, Moran's eigenvectors of spatial weighting matrices in R, appendix to Dray S and Legendre P. and Peres-Neto P.R., 2006, Spatial modelling : a comprehensive framework for principal coordinate analyisis of neighbour matrices (PCNM), *Ecological Modelling*, 196, 483-493

Dray S., Said S., Debias F., 2008, Spatial ordination of vegetation data using a generalisation of Wartenberg's multivariate spatial correlation, *Journal of Vegetation Science*, 19, 45-56.

Dray S. and Jombart T., 2010, Revisiting Guerry's data: introducing spatial constraints in multivariate analysis, *Annals of applied Statistics (in press).*

Dyson-Bruce I., 2003, Historic Landscape Assessment: the East of England experience. Paper product to GIS delivery, *Journal of GIS in Archaeology*, volume1, 63-72.

Fairclough G., 2003, *The "long chain": archaeology, historical landscape characterisation and time depth in the landscape,* chapter 16, 295-318. Landscape interfaces H. Palang and G.Frey (eds) Kluwer Academic Press, Netherlands.

Forestry Commission Technical note 16/95, 1996
http://www.forestry.gov.uk/pdf/technote16_95.pdf/$file/technote16_95.pdf

Gaucherel C., 2007, Multiscale heterogeneity map and associated scaling profile for landscape analysis, *Landscape and Urban Planning,* 82, 95-102.

Giles P.T., 1998, Geomorphological signatures: Classification of aggregated slope unit objects from digital elevation and remote sensing data, *Earth Surface Processes and Landform,* 23, 581-594.

Groom G., 2005, European Landscape Character Areas. Typology, cartography and indicators for the assessment of sustainable landscapes, FP5-EU Accompanying measure, ELCAI-EVK2-CT-2002-80021, Dirk Wascher.

Gustafson E.J., 1998, Quantifying landscape spatial patterns: what is the state of the art?, *Ecosystems*, 1, 143-156.

Habron A.D., 1998, *Defining Wild land in Scotland through GIS based wilderness perception mapping*, Thesis submitted to the University of Stirling for the degree of Doctor of Philosophy.

Heywood I., Cornelius S., Carver S., 2006, *An Introduction to Geographical Information Systems.* Pearson Education Limited.

Herring P.C., 2009, Framing perceptions of the historic landscape: Historic Landscape Characterisation (HLC) and Historic Land-use Assessment (HLA), *Scottish Geographical Journal*, 125:1, 61-77.

Hoesterev R., 2005, in
ESRI arc-newsletter http://www.esri.com/news/arcnews/winter0506articles/for-puget-sound.html

Holland S.M., 2006, *Cluster Analysis*,
http://www.uga.edu/strata/software/pdf/clusterTutorial.pdf

Huang C., Geiger E.L., and Kupfer J.A., 2006, Sensitivity of landscape metrics to classification scheme, *International Journal of Remote Sensing,* Vol. 27, No. 14, 2927-2948.

Ingold T., 2000, *The perception of the environment: essays on livelihood, dwelling and skill*. London. Routledge.

Jenness J., 2006. Topographic Position Index (tpi_jen.avx) extension for ArcView 3.x, v 1.3a Jennes enterprise. Available at http://www.jennessent.com/arcview/tpi.htm

Jombart T., 2009, Multivariate analysis of genetic markers as a tool to explore the genetic diversity: some examples. Practical course using the R software. http://adegenet.r-forge.r-project.org/files/globaldiv-PC.1.3.pdf

Julie Martin Associates and Swanwick C., 2003, Overview of Scotland National Programme of Landscape Character Assessment, *SNH commissioned report F03 AA307*.

Kang-Rae and Banister D., 2006, Excess Commuting: A Critical Review, *Transport Reviews*, vol.26, numb.6, 749 – 767.

Kearns F.R., Maggi Kelly N., Carter J.L., Resh V.H., 2005, A method for the use of landscape metrics in freshwater research and management, *Landscape Ecology,* 20, 113-125.

Keisteri T., 1990, The study of change in cultural landscapes. *Fennia*, 168:1, 31-115.

Kim H. K. and Pauleit S., 2007, Landscape character, biodiversity and land use planning: The case of Kwangju City Region, South Korea, *Land Use Policy,* 24, 264-274.

Klingseisen B., 2004. *GIS based generation of topographic attributes for landform classification*. Diploma thesis. Villach 2004.

Klingseisen B., 2007. Geomorphometric landscape analysis using a semi-automated GIS approach. *Environmental Modelling and software,* xx, 1-13.

Krivorunchko K. and Bivand R., 2003.*GIS, users, developers and spatial statistics: on monarchs and their clothing.* Paper presented at StatGIS, International workshop on interfacing geostatistics, GIS and spatial databases, Sept 29 – Oct 1, 2003 Pörtschach, Austria.

Krivorunchko K. and Gotway C.A., 2002. *Expanding the "S" in GIS: incorporating spatial statistics in GIS.* Paper presented at the CSISS specialist meeting on spatial data analysis software tools. Santa Barbara. May 2002.

Land Use Consultants, 2008, National overview of Scotland's landscapes, Scottish Natural Heritage technical report, March 2008, Perth, UK.

Lausch A., Herzog F., 2002, Applicability of landscape metrics for the monitoring of landscape change: issued of scale, resolution and interpretability, *Ecological Indicators*, 2, 3-15.

Lee J. and Wong D.W.S., 2001, *Statistical Analysis with Arcview GIS*, John Wile and Son, Printed in the United States of America.

Lock G and Leigh Molyneaux B., 2006, *Confronting scale in Archaeology. Issues of Theory and Practice*, Springer Science and Business Media, New York.

Lorimer H. and Lund K., 2003, *Performing facts: finding a way over Scotland's mountains*, The Editorial Board of the Sociological Review, Oxford, Blackwell.

Lothian A., 1999, Landscape and the philosophy of aesthetics: is the landscape quality inherent in the landscape or in the eye of the beholder?, *Landscape and urban planning,* 44, 177-198.

Lundblad E.R., Wright D.J., Miller J., Larkin E.M., Rinehart R., Naar D.F., Danahue B.T., Anderson S.M. and Battista T., 2006. A Benthic Classification Scheme for American Samoa. *Marine Geodesy,* 29, 89-111.

MacMillan R.A., McNabb D.H., Jones R.K., 2000, *Automated landform classification using DEMs: a conceptual framework for a multilevel, hierarchy of hydrologically and morphologically oriented physiographic mapping units*, 4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4). Banff, Alberta – Canada, September 2-8, 2000.

Macpherson H. and Minca C., *Landscape embodiment and visual impairment: an exploration of the limits of landscape knowledge*, Forum UNESCO University and Heritage, 10th International Seminar, "cultural Landscapes in the 21st Century". Newcastle-upon-Tyne, 11-16 April 2005.

Mander Ü., Muller F., and Wrbka T., 2005, Functional and structural landscape indicators: Upscaling and downscaling problems, *Ecological Indicators,*5, 267-272.

Mander Ü., 2010, Landscape assessment for sustainable planning - Editorial, *Ecological Indicators,* 10, 1-3.

McGarigal K. and Marks B., 1995, *FRAGSTATS: spatial pattern analysis program for quantifying landscape structures*, United States Department for agriculture and Pacific Northwest Research Station, General Technical Report PNW-GTR-351.

Mitchell A., 1999. *The ESRI guide to GIS analysis. Volume 1: geographic patterns and relationships.* Redlands. California. ESRI.

Mitchell A., 2005. *The ESRI guide to GIS analysis. Volume 2: spatial measurement and statistics.* Redlands. California. ESRI.

Moreira F., Queiroz A.I., and Aronson J., 2006, Restoration principles applied to cultural landscapes, *Journal for Nature Conservation,* 14, 217-224.

Nagendra H., Munroe D.K., and Southworth J., 2004, From pattern to process: Landscape fragmentation and the analysis of land use land cover change, *Agriculture Ecosystem and environment,* 101, 111-115.

NOAA Coastal Services Center – Linking people, information and technology. Oregon State Unniversity OSU.
http://www.csc.noaa.gov/products/btm/

Ode Å, Fry G., Tveit M.S., Messager O., Miller D., 2007, Indicators of perceived naturalness as drivers of landscape preference, *Journal of Environmental Management*, 1-9.

Ollier S., 2005, *Des outils pour l'intégration des contraintes spatiales, temporelles et évolutives en analyse de données écologiques*, Thèse pour obtenir le diplôme de doctorat, Université Claude Bernard- Lyon1.

Palang H., Alumäe H., Mander U., 2000, Holistic aspects in landscape development: a scenario approach, *Landscape and Urban Planning,* 50, 85-94.

Palmer J.F., 2004, Using spatial metrics to predict scenic perception in a changing landscape: Dennis, Massachusetts, *Landscape and Urban Planning,* 69, 201-218.

Palmer J.F. and Lankhorst J.R.-K., 1998, Evaluating visible spatial diversity in the landscape, *Landscape and Urban Planning*, 43, 65–78

Paron P. and Vargas R., 2007. *Landform selected study areas in Somaliland and Southern Somalia. Integrated Landform mapping approach at semi-detailed scale using remote sensing and GIS techniques*. FAO-SWALIM. Project Report L-02. Nairobi, Kenya.

Pastor O.I., Casermeiro Martinez M.A., Ezquerra Canalejoa A., Esparcia Marino P., 2007, Landscape evaluation: Comparison of evaluation methods in a region of Spain, *Journal of Environmental Management,* 85, 204-214.

Patil G.P., Myers W.L., Luo Z., Jhonson G.D., Taille C., 2000, Multiscale assessment of landscapes nad watersheds with synoptic multivariate spatial data in environmental and ecological statistics, *Mathematical and Compputer Modelling,* 32, 257-272.

Peng J., Wang Y., Zhang Y, Wu J., Li W., Li Y., 2010, Evaluating the effectiveness of landscape metrics in quantifying spatial patterns, *Ecological Indicators*, 10, 217-223.

Pennock D.J., Zerbath B.J., DeJong E., 1987, Landform classification and soil distribution in hummocky terrain, Saskatchewan, Canada, *Geoderma*, 40, 297-315.

Pyysalo U., Oksanse J., Sarjakoski T., 2009, Viewshed analysis and visualisation of landdscsape Voxel model. 24th International Cartographic Conference. 15th-21st November 2009. Chile

Prima O. D. A., Echigo A., Yokoyama R. and Yoshida T., 2006. Supervised landform classification of Norteast Honshu from DEM-derived maps. *Geomorphology,* 78, 373-386.

Purtauf T., Thies C., Ekschmitt K., Wolters V., Dauber J., 2005, Scaling properties of multivariate landscape structure, *Ecological Indicators,* 5, 295-304.

Saby N.P.A., Thioulouse J., Jolivet C.C., Ratie' C., Boulonne L., Bispo A., Arrouays D., 2009, Multivariate analysis of the spatial patterns of 8 trace elements using the French soil monitoring network data., *Science of the Total Environment*,407 , 5644-5652.

Sain S.R. and Cressie N., 2007, A spatial model for multivariate lattice data, *Journal of Econometrics*, 140, 226-259.

Scott A., 2002, Assessing Public Perception of Landscape: the LANDMAP experience, *Landscape Research*, 27(3), 271-295.

Scrucca L., 2005, Clustering multivariate spatial data based on local measures of spatial autocorrelation. An application to the labour market of Umbria. *Quaderni del Dipartimento di Economia, Finanza e Sattistica*, Universita' degli studi di Perugia, Number 20, October 2005.

Sisk T., Prather J.W., Hampton H.M., Aumack E.N., Xu Y., Dickson B.G., 2006, Participatory landscape analysis to guide restoration of ponderosa pine ecosystems in the American Southwest, *Landscape and Urban Planning,* 78, 300-310.

Storfer A., Murphy M.A., Evans J.S., Goldberg C.s., Robinson S., Spear S.F., Dezzani R., Delmelle E., Vierling L., Waits L.P., 2007, Putting the "landscape" in landscape genetics, *Heredity*, 98, 128-142

Swanwick C. and Land Use Consultants, 2002, Landscape Character Assessment. Guidance for England and Scotland, Countryside Agency and Scottish Natural Heritage publication.

Thompson J.A., Pena-Yewtukhiw E.M., and Grove J., 2006, Soil-landscape modelling across a physiographic region: topographic patterns and model transportability, *Geoderma,* 133, 57-70.

Thomson, M.K., Béra, R., 2007. Relating Land Use to the Landscape Character: Toward an Ontological Inference Tool. Proc. GISRUK 2007, NUI Maynooth, Ireland, 11-13 April 2007

Tress B. and Tress G., 2003, Scenario visualisation for participatory landscape planning: a study from Denmark, *Landscape and Urban Planning,* 64, 161-178.

Unwin D.J., 1996, GIS, spatial analysis and spatial statistics, *Progress in Human Geography*, 20, 4, 540-551.

Van den Berg R.A., Hoefsloot H.CJ., Westerhuis J.A., Smilde A.K., Van der Werf M.J., 2006, Centring, scaling and transformations: improving the biological information content of metamolomics data, *BMC Genomics*, 7:142.

Van Eetvelde V. and Antrop M., 2004, Analyzing structural and functional changes of traditional landscapes – two examples from Southern France, *Landscape and Urban Planning*, 67, 79-85.

Voss P. and Ramsay S., 2006. Five day course on spatial regression analysis. University of Manchester. Documentation of pdf files is available at the following link: http://www.wun.ac.uk/ggisa/elearning.html

Wagner H.H. and Fortin M-J., 2005, Spatial analysis of landscapes: concepts and statistics, Ecology, 86(8), 1975-1987.

Washtell J., Carver S., Arrel K., 2009, A viewshed based classification of landscapes using geomorphometrics, Proceedings of Geomofphometry 2009. Zurich, Switzerland, 31 August – 2 September, 2009

Weiers S., Bock M., Wissen M., Rossner G., 2004, Mapping and indicator approaches for the assessment of habitats at different scales using remote sensing and GIS methods, *Landscape and Urban Planning,* 67, 43-65.

Weiss A., 2001, *Topographic Position and landforms analysis*, Poster presentation, ESRI international user conference, San Diego, CA, 2001

Williams E.A. and Wentz E.A., 2008, Pattern analysis based on Type, Orientation, Size and Shape, *Geographical Analysis,* 40, 97-122.

Wise S., Haining R, Ma J., 2001, Providing spatial statistical data analysis functionality for the GIS user: the SAGE project, International Journal Geographical Information Science, 15, 3, 239-254.

Wu J. and Qi Y., 2000, Dealing with scale in landscape analysis: an overview, *Geographic Information Science,* 6, 1-5.