QUANTIFYING THE EFFECTS OF VARIABLE SELECTION, SPATIAL SCALE AND SPATIAL DATA QUALITY IN MARINE BENTHIC HABITAT MAPPING

by

© Vincent Lecours

A Thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Geography

Memorial University of Newfoundland

September 2016

St. John's

Newfoundland and Labrador

Abstract

Mapping benthic habitats has become critical in many contexts like conservation and management. While marine habitat mapping methods strongly rely on tools and methods from geography and geomatics, habitat mapping practitioners with a background outside of these specialized areas do not always have a full understanding of the spatial concepts behind these tools and methods. This phenomenon is amplified when marine geomorphometry, the science used to quantify seafloor terrain characteristics, is integrated into the marine habitat mapping workflow. This dissertation reviews the use of spatial concepts in the field of marine benthic habitat mapping; many concepts are poorly understood or poorly implemented in the habitat mapping workflow, among which spatial scale and spatial data quality stand out as being of particular importance.

While geomorphometry is commonly used in marine benthic habitat mapping, no framework existed to test which terrain attributes should be used as surrogates of species distribution, leading to an inability to compare results from different studies. This dissertation explores different options for terrain attribute selection and proposes an optimal combination that can be used as standard in all habitat mapping studies. This selection is then tested using two approaches to benthic habitat mapping and is shown to perform better than others.

Bathymetric data, the primary input for marine geomorphometry analyses and one of the main data inputs for habitat mapping, are commonly impacted by data acquisition artefacts. Very little work has been done on trying to understand how these artefacts propagate throughout the habitat mapping workflow. The impact of artefacts on the bathymetry and its derived terrain attributes is described, and it is shown that artefacts modify the spatial and statistical distributions of depth and terrain attribute values. However, when these affected data are used in habitat mapping, their impact is not always predictable. Some artefacts were found to sometimes inflate measures of accuracy and performance and sometimes decrease them.

Overall, habitat maps were shown to be very sensitive to the effects of variable selection, spatial scale and data quality, and as such have serious implications when they are used to inform decision-making, for instance in marine conservation and management. This dissertation raises awareness about these issues and highlights the need for careful integration of spatial data in habitat mapping practices.

Co-Authorship Statement

This dissertation is written in manuscript style with the goal of publishing Chapters 2 to 6. The author of this dissertation has been the primary researcher behind the work performed in each of the chapters, including developing the ideas, reviewing the literature, designing the experiments, analyzing the data, interpreting the results, and preparing the manuscripts. The co-authors and committee members have contributed ideas, assisted in the development of the different methods and the interpretation of results, and reviewed the manuscripts. In terms of the manuscripts, the author of this dissertation is the primary author on all the manuscripts included in this dissertation. The co-authors contributed by critically reading and providing feedback on the different manuscripts.

Chapter 2, "Spatial Scale and Geographic Context in Benthic Habitat Mapping: Review and Future Directions" was published in 2015 in the journal *Marine Ecology Progress Series* (vol. 535, p. 259-284). This manuscript was written in close coordination with co-supervisor Dr. Rodolphe Devillers (Department of Geography, Memorial University), and with the help of Dr. David Schneider (Department of Ocean Sciences, Memorial University). Committee members Dr. Vanessa Lucieer (Institute for Marine and Antarctic Studies, University of Tasmania) and Dr. Craig Brown (Nova Scotia Community College), and co-supervisor Dr. Evan Edinger (Department of Geography, Memorial University) provided feedback on versions of this manuscript.

Chapter 3, "Towards a Framework for Terrain Attribute Selection in Environmental Studies", has been submitted to the journal *Environmental Modelling & Software* in June

2016 and is currently under review. The analysis performed in this manuscript was supervised by Dr. Devillers and Dr. Alvin Simms (Department of Geography, Memorial University). Dr. Devillers, Dr. Simms, and committee members Dr. Lucieer and Dr. Brown provided feedback on versions of this manuscript.

Chapter 4, “Comparing Selections of Environmental Variables for Ecological Studies: A Focus on Terrain Attributes”, has been submitted to the journal *PLoS One* in June 2016 and is currently under review. Dr. Brown and Dr. Devillers supervised the development of the methods, and provided feedback on versions of the manuscript, together with Drs. Lucieer and Edinger.

Chapters 5 and 6 were written for submission to a journal after completion of this dissertation. Drs. Devillers, Lucieer and Brown provided feedback on versions of Chapter 5, and Drs. Devillers and Edinger provided feedback on versions of Chapter 6.

Other publications led by the author are not presented in this dissertation but were part of the author’s PhD work. A short book chapter and a full-length review article of the field of marine geomorphometry, an important component of this dissertation, were published. Preliminary results from Chapter 3, gained using slightly different methods from those in Chapter 3, were also published as a short book chapter. In terms of spatial data quality, an evaluation of uncertainty propagation to bathymetric data collected from a remotely operated vehicle in the deep sea was published in conference proceedings. A technical report that explored benthic habitat mapping methods to study cold-water coral habitats at multiple scales using geomorphometry was also presented to the Department of Fisheries and Oceans Canada. Finally, a short paper looking at the application of a new algorithm for photo-mosaicking images of the seafloor to habitat mapping was published in *Sea Technology*. The relevant references are listed below.

Lecours, V., Dolan, M.F.J., Micallef, A., & Lucieer, V. (2016) A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrology and Earth System Sciences*, 20:3207-3244.

Lecours, V., Lucieer, V.L., Dolan, M.F.J., & Micallef, A. (2015) An ocean of possibilities: applications and challenges of marine geomorphometry. In: Jasiewicz, J., Zwolinski, Z., Mitasova, H., & Hengl, T. (Eds) *Geomorphometry for geosciences*. Bogucki Wydawnictwo Naukowe, Poznan, Poland, p. 23–26.

Lecours, V., Simms, A., Devillers, R., Lucieer, V. & Edinger, E. (2015) Finding the best combinations of terrain attributes and GIS software for meaningful analysis. In: Jasiewicz, J., Zwolinski, Z., Mitasova, H., & Hengl, T. (Eds) *Geomorphometry for geosciences*. Bogucki Wydawnictwo Naukowe, Poznan, Poland, p. 133-136.

Lecours, V., & Devillers, R. (2015) Assessing the spatial data quality paradox in the deep-sea. In: Sieber, R.E. (Ed.) *Proceedings of Spatial Knowledge and Information Canada*, 2015:1–8.

Lecours, V., Miles, L.L., Devillers, R., & Edinger, E.N. (2013) Data analysis towards the multiscale characterization of cold-water coral and sponge habitats in Canadian waters. *Technical Report REQ. No. F6160-120010*. Department of Fisheries and Oceans Canada, Newfoundland and Labrador Region, St. John’s, 260 p.

Bagheri, H., & Lecours, V. (2012) Benthic habitat mapping using high-resolution image mosaicking. *Sea Technology*, 53:15-20.

Acknowledgements

The journey to completing this dissertation was shaped by all the individuals who I encountered in the last few years and who have given me time, expertise, financial and moral support. I have grown professionally and personally in the last five years, and these people have definitely contributed to that growth.

First and foremost, I would like to acknowledge my supervisors Rodolphe Devillers and Evan Edinger, who have been there for me every time I needed, whether it was work-related or otherwise. They provided guidance, expertise and financial support, but I am most grateful for all the opportunities they gave me to learn and experience throughout my degree. They understand that a PhD is not just a piece of paper they give you at the end of a long and arduous journey, but it is actually a training that involves interactions with others as much as analyses. Whether it was for conferences, workshops, professional meetings, teaching opportunities, or allowing me to go get lost on a boat in the middle of the Pacific Ocean with other scientists, they were always supportive, and that means a lot to me. I am very grateful for their trust in me and the independence they gave me.

Many thanks to my research committee, Craig Brown and Vanessa Lucieer, for their insightful comments and discussions on my research and for dealing with my tendency to always do more (or too much…). Thanks are also due to David Schneider, Alvin Simms, Marie-Josée Fortin, Margaret Dolan, and Aaron Micallef who contributed directly to my work, and to all the others that I contacted throughout the years for advice or help on topics related to my research. Also, thanks to Trevor Bell, Chris McGonigle and Eric Vander Wal for reviewing my dissertation and providing useful comments on it.

My gratitude also goes to the Geography Department at Memorial University. The faculty, students, and staff have made my time in the department unforgettable. I am thankful to have been part of such a great group in the Marine Geomatics Research Lab and the Marine Habitat Mapping Research Group. To all alumni, current students, postdocs and visiting researchers with whom I shared more than just a working space, thanks for tolerating my occasional grumpiness. Arnaud and Daniel deserve special mention for the freshly-made coffee every morning.

I am grateful to all the organizations that provided funding for my research or for conferences: the Natural Sciences and Engineering Research Council of Canada (NSERC), the School of Graduate Studies of Memorial University, the Department of Fisheries and Oceans Canada, the Office of the Vice-President (Research) of Memorial University, the Canadian Foundation for Innovation (CFI), the Atlantic Division of the Canadian Association of Geographers, the Department of Geography of Memorial University, GeoHab and Esri Canada.

Finally, thanks to my friends, in Québec, Newfoundland and elsewhere, who supported, encouraged and most importantly entertained me throughout this whole thing. Thanks to Emma, who has been more than patient with me, listening and providing me with daily, unconditional support (and food!). Thanks to my parents, who encouraged me to jump into this adventure, my sister who showed me the example, and to my grand-parents and aunts who are always happy to hear about what I am doing, even if they do not always understand much!

Dedicated to the memory of Jean-Denis Lecours.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AUC	Area Under the Curve
AUV	Autonomous Underwater Vehicle
CC	Coefficient of Congruence
CHS	Canadian Hydrographic Service
COS	Change-Of-Support
DBM	Digital Bathymetric Model
DEM	Digital Elevation Model
DFO	Department of Fisheries and Oceans Canada
DTM	Digital Terrain Model
GEBCO	General Bathymetric Chart of the Oceans
GIS	Geographic Information System
GPS	Global Positioning System
GWR	Geographically Weighted Regression
ICES	International Council for the Exploration of the Sea
IID	Independent and Identically Distributed
IMU	Inertial Measurement Unit
LiDAR	Light Detection and Ranging
MAP	Minimum Average Partial Correlation
MAUP	Modifiable Areal Unit Problem
MaxEnt	Maximum Entropy
MBES	Multibeam Echosounder System

MESH	Mapping European Seabed Habitats
MI	Minimum Information
MIM	Minimum Information Matrix
MMU	Minimum Mapping Unit
PA	Parallel Analysis
PCA	Principal Components Analysis
RDMV	Relative difference to Mean Value
ROC	Receiver Operating Characteristic
ROV	Remotely Operated Vehicle
SAC	Spatial Autocorrelation
SC	Similarity Coefficient
SCUBA	Self-Contained Underwater Breathing Apparatus
SDM	Species Distribution Model
SWATH	Small Waterplane Area Twin Hull
TASSE	Terrain Attribute Selection for Spatial Ecology
TPI	Topographic Position Index
VIF	Variable Inflation Factor
VRM	Variable Ruggedness Measure

List of Appendices

1. Introduction and Overview

1.1 Introduction

While the oceans are estimated to comprise up to 90% of the inhabitable area for life on Earth (Tittensor *et al.*, 2009), and despite over 150 years of exploration, the scientific community still knows very little about the marine environment compared to its terrestrial counterpart (Roberts, 2002). Historically, knowledge about the marine environment was often gained and driven by human use of the oceans, partly through the exploitation of natural resources. For instance, the presence of cold-water corals has been documented by fishermen since the mid-eighteenth century (Roberts, 2002), and more recently coral observations from fisheries bycatch have helped describe their presence, abundance and geographic distribution (Cogswell *et al.*, 2009; Murillo *et al.*, 2011). The realization in the twentieth century that the oceans, especially in the deep sea, are not muddy and lifeless triggered interest in documenting more than just species occurrences by mapping marine habitats. Marine habitat mapping then became a scientific endeavour, often an applied one designed to answer specific scientific and management questions. For instance, marine habitat mapping has been used to inform conservation efforts (*e.g.* Laffoley & Hiscock, 1993; Light, 1998), to study juvenile mortality in benthic invertebrates (*e.g.* Gosselin & Qian, 1997), to study the disturbance of seabed habitats from fishing gear (*e.g.* Friedlander *et al.*, 1999), and to evaluate ecological associations between different species (*e.g.* Olsgard *et al.*, 2003).

In the last 25 years, the importance of anthropogenic pressure on seafloor environments and the growing realization of the significance of these ecosystems, for instance in terms of ecosystem services (Galparsoro *et al.*, 2014), have steered many nations towards increasing efforts to better manage and protect marine resources (Borja, 2014). Such efforts led scientists to define standards for marine habitat mapping, and many definitions of benthic habitats were proposed (*e.g.* Kostylev *et al.*, 2001; Harris & Baker, 2012). In its simplest form, a benthic habitat is a distinct area of the seafloor characterized by a combination of specific chemical, physical and/or biological characteristics. Mapping the seafloor based on species' habitat requirements has become critical in many contexts, and is often the first step in implementing scientific management, monitoring environmental change, and assessing the impacts of anthropogenic disturbance on benthic ecosystems (Roff *et al.* 2003; Cogan & Noji 2007). Seafloor mapping also provides the data necessary for the identification and monitoring of marine protected areas (Le Pape *et al.*, 2014). Habitat maps enable the interpretation of the nature, distribution, and extent of distinct physical environments, and allow predictions of species or communities distribution based on their associations with the environment (Harris and Baker, 2012).

The field of marine benthic habitat mapping has evolved rapidly during these 25 years. Technological and methodological developments in marine benthic habitat mapping were partly driven by innovations in geomatics, more specifically in Geographic Information Systems (GIS), spatial analysis methods and remote sensing technologies (Wright & Heyman, 2008; Brown *et al.*, 2011). In their "Review of Standards and Protocols for Seabed Habitat Mapping", Coggan *et al.* (2007) defined a very general approach to marine benthic habitat mapping as the spatial integration of different datasets, usually within a geospatial environment. Within a GIS environment, spatial analytical techniques are combined with high-resolution geoscientific and environmental data with *in situ* observations to enable accurate quantification and representation of habitats. This provides a framework for mapping the distribution of benthic species and interpreting spatial patterns in biodiversity (Whitmire *et al.*, 2007; Brown *et al.*, 2011; Harris and Baker, 2012). For a long time, habitat mapping was done using available, often broad-scale data. For instance, while bycatch data and bathymetric data from satellite radar altimetry provide some understanding of species biogeography at a regional scale (*e.g.* Bryan & Metaxas, 2007), they present challenges when trying to understand habitat characteristics at a more local scale that is often more meaningful for purposes such as conservation and management (Etnoyer & Morgan, 2007). However, the development of acoustic remote sensing and bathymetric LiDAR techniques in recent decades has helped fill the gap in high-resolution spatial environmental data necessary to map benthic habitats at scales relevant to such purposes. Data provided by this type of remote sensing, specifically bathymetry and backscatter data, have revolutionized benthic habitat mapping, both in terms of the methods used and our ability to map habitats efficiently and with relative ease. Bathymetric and backscatter data are used for seafloor characterization, habitat mapping, and the derivation of surrogates to predict species distribution (Butler *et al.*, 2006; Anderson *et al.*, 2008; Dolan *et al.*, 2008). These data have proven their value for habitat mapping and their potential to help the scientific community advance its understanding of seafloor ecosystems (Anderson *et al.*, 2008; Brown *et al.*, 2011).

Like benthic habitat mapping, geomorphometry – the science used to derive quantitative measurements of terrain characteristics from digital terrain models (DTM) – has been fueled by advances in remote sensing and GIS in recent decades (Florinsky, 2012) and now strongly rely on methods and techniques from geomatics. Geomorphometry has traditionally focused on the investigation of terrestrial landscapes, but the dramatic increase in the availability of digital bathymetric data and the increasing ease by which geomorphometry can be analyzed using GIS has prompted interest in employing geomorphometric techniques to investigate the marine environment (*e.g.* Lundblad *et al.*, 2006; Wilson *et al.*, 2007). Over the last decade or so, a multitude of geomorphometric techniques have been applied to characterize the seafloor, and marine benthic habitat mapping is a major area where the use of marine geomorphometry has grown in recent years (reviewed in Lecours *et al.*, 2016). Linked to the increasing use of multibeam and bathymetric LiDAR data for benthic habitat mapping (Brown *et al.*, 2011; Smith and McConnaughey, 2016), the vast majority of habitat mapping studies with access to bathymetric data are now using, or at least testing, some form of terrain attributes (*e.g.* slope, orientation, rugosity) in their workflow. Since most marine habitats are difficult to access, observe and sample (Solan *et al.*, 2003; Robinson *et al.*, 2011), bathymetric data are often the only reliable dataset available to characterize benthic habitats; by enabling the extraction of quantitative information from bathymetric data, geomorphometry provides an invaluable source of additional and relevant information for benthic habitat mapping.

In the context of marine geomorphometry, the roles of several spatial concepts are still not studied (*e.g.* spatial data quality) or fully understood (*e.g.* spatial scale), and these spatial concepts are often overlooked when marine geomorphometric techniques are used in marine benthic habitat mapping. By its spatial and data-driven natures and the near ubiquitous use of GIS, remote sensing and spatial analysis in its workflow, marine habitat mapping and its practices are also directly influenced by spatial concepts such as spatial scale and spatial autocorrelation. However, the scientists framing the questions related to the study of benthic habitats and the managers applying the answers are frequently not trained in the geomatics tools that underpin much of the new techniques in marine habitat mapping. While geographers and other spatial scientists have been studying these fundamental concepts for a long time, the understanding of their role and their integration in the habitat mapping workflow – including during geomorphometric analyses – have received scant attention in the past. Since marine habitat maps have become a critical tool in decision-making, especially in marine conservation and management (Reiss *et al.*, 2014; Buhl-Mortensen *et al.*, 2015; Davies *et al.*, 2015; Rolet *et al.*, 2015; Howell *et al.*, 2016), there is an urgent need to improve our understanding of how these concepts influence the representation of benthic ecosystems and habitats. This dissertation focuses on marine benthic habitat mapping methods, and particularly on how a better integration of spatial concepts like spatial scale can improve the marine habitat mapping workflow. It is aimed at developing best practices in the application of geomatics-based marine habitat mapping to ecological and management questions. A particular focus is given to the concepts of variable selection, spatial scale and spatial data quality.

1.2 Research Problem and Research Gap

The general problem that this dissertation addresses is that of a widely used but poorly understood spatial framework for marine benthic habitat mapping research. Some work has been done (*e.g.* Dolan *et al.*, 2008; Brown *et al.*, 2011; Rengstorf *et al.*, 2012; Dolan & Lucieer, 2014; Rattray *et al.*, 2014) to improve our understanding of the influence of certain spatial concepts (*e.g.* spatial scale, spatial data quality) on the way we represent and understand benthic habitat. However, much work remains to be done to enable the full integration or consideration of spatial concepts within the habitat mapping workflow. This problem is rooted in the disconnection between marine habitat mapping – a field that is spatial in nature – and the field of geomatics, which provides the spatial concepts that partly support many habitat mapping practices. Geomatics is intrinsically linked to geography and is defined as the "discipline dedicated to the management of spatially referenced data, and thus relies on the scientific concepts and technologies implied in the acquisition, storage, analysis, and distribution of the data" (Caron *et al.*, 2008, p. 295).

Through the development of GIS, which offer tools to analyze and represent spatial data, geomatics has become very accessible to a wide range of scientists involved in marine benthic habitat mapping (*e.g.* geologists, ecologists, biologists). The effectiveness of geomatics concepts and methods – and of GIS tools to assist in the exploration of questions from other disciplines – has given rise to the multidisciplinary and interdisciplinary nature of the field (Wright *et al.*, 1997; Mark, 2000, 2003; Blaschke *et al.*, 2011, 2012). First, these concepts and tools enable a better understanding of complex phenomena by integrating information in a spatial context, and incorporating both qualitative and quantitative spatial reasoning with concepts from other disciplines (*e.g.* geology or ecology) into a common framework. Then, geomatics provides a set of core concepts that improve communication and mutual understanding about spatial data and information among researchers with different backgrounds (Kuhn, 2012). As a consequence, geomatics has been widely adopted by many disciplines in the last 25 years (Raper, 2009; Blaschke & Merschdorf, 2014), including geomorphometry (Zhou & Zhu, 2013) and marine habitat mapping (Wright & Heyman, 2008).

Geomatics is an integrative science. As Lam & Kemp (2012, p. 2194) stated: “Integrative science is the cornerstone of this field – both helping others integrate and integrating other sciences into ours to see where the technology and science are lacking.” Geomatics is thus considered a multiparadigmatic science in which approaches from other disciplines are commonly applied within the field, and vice-versa (Blaschke and Merschdof, 2014). However, the integrative nature of geomatics also has downsides. The development of easily accessible tools brings hidden dangers by facilitating non-critical use by end-users (*e.g.* computer scientists, ecologists, geologists) who may have limited appreciation of spatial concepts, for instance spatial scale, spatial representation and spatial data quality. Because the tools are made to be intuitive, they often do not require end-users to fully understand the characteristics of the underlying processes and parameters implemented in the tools, and of the spatial data that form the basis for analysis. The scientific concepts and foundations are thus often hidden behind increasingly “black-box”, user-friendly tools. In addition, the fast pace at which tools and techniques are developing may commonly prevent end-users from remaining apprised of new developments in these scientific concepts and foundations. This leads to the danger of inappropriate use of geospatial data and tools and the lack of appropriate consideration of important spatial concepts, from which misinformed and potentially erroneous interpretations or inferences could be made.

In summary, there is a lack of consideration, likely caused by a lack of understanding, of spatial concepts (*e.g.* spatial scale, spatial data quality, spatial autocorrelation) within marine benthic habitat mapping practices. There is a need to reunite the community of end-users – *e.g.* the geologists, ecologists, geographers, or decision-makers involved in the production of benthic habitat maps – with the spatial foundations that underpin the data, tools and methods they use. The specific research gap that this dissertation addresses is the following: there are currently no best practices defined to better integrate spatial concepts like spatial scale and spatial data quality in the marine benthic habitat mapping workflow, particularly when marine geomorphometric analyses are part of this workflow.

1.3 Research Questions

The research problem can be addressed by increasing geographic literacy in disciplines like marine habitat mapping and marine geomorphometry that commonly use spatial data and GIS tools (Blaschke & Strobl, 2010).This dissertation attempts to answer the following questions to increase the body of knowledge related to spatial concepts in marine benthic habitat mapping, and more particularly on spatial scale and spatial data quality. Since marine geomorphometry is an important part of the marine habitat mapping workflow, these questions will be answered through the integration of geomorphometric analyses in the habitat mapping practices.

1. Which particular spatial concepts are poorly integrated in the marine benthic habitat mapping workflow, and is it possible to identify specific ways to improve the integration of spatial concepts in marine benthic habitat mapping?
2. Which combination of terrain attributes best capture seafloor characteristics while minimizing spatial covariation? Can these terrain attributes form a standard protocol for using marine geomorphometry in different approaches to habitat mapping?
3. How sensitive are these terrain attributes to different types of data acquisition artefacts? Does the sensitivity of terrain attributes to data acquisition artefacts vary with spatial scale?
4. Do data acquisition artefacts propagate to habitat maps and species distribution models? Are the impacts of data acquisition artefacts on habitat maps and species distribution models scale-dependent?

1.4 Research Hypotheses

The following five hypotheses are examined in this dissertation:

1. There is a lack of understanding of the role of different spatial concepts, for instance like spatial scale and spatial data quality, in marine benthic habitat mapping practices.
2. Different optimal combinations of terrain attributes exist and their composition varies depending on seafloor characteristics (*e.g.* roughness).

3. The different optimal combinations of terrain attributes are generalizable, *i.e.* they can be integrated in the marine habitat mapping workflow regardless of the approach used to map habitats.
4. Terrain attributes, and habitat maps and species distribution models produced from these terrain attributes, are sensitive to data acquisition artefacts.
5. The sensitivity of terrain attributes, habitat maps and species distribution models to data acquisition artefacts is scale-dependent; finer-scale data, and maps and models built from these data, are more sensitive to data acquisition artefacts than broader-scale data and maps and models produced with these data.

1.5 Research Objectives

The overarching objective of this dissertation is to identify spatial concepts that currently lack a proper consideration in marine benthic habitat mapping practices, and to propose solutions towards a better (re)integration of these concepts into the marine benthic habitat mapping workflow. A particular focus is given to spatial scale, spatial covariation, and spatial data quality.

Specific objectives are:

1. Review existing knowledge on spatial concepts in benthic habitats and their mapping, including the related practices of surrogacy assessment and species distribution modelling.
2. Identify ways to improve marine benthic habitat mapping practices through a better integration of spatial concepts.

3. Identify combinations of environmental variables that minimize spatial covariation and optimize the information extracted from these data.
4. Demonstrate the importance of using spatial data with an ecological meaning for marine benthic habitat mapping by describing the effects of subjectively selecting spatial data on the production of habitat maps and species distribution models.
5. Describe the impacts of a poor spatial data quality on the marine benthic habitat mapping workflow.
6. Evaluate the relationship between spatial scale and spatial data quality in a marine benthic habitat mapping context.

1.6 Methods

Through its exploration of the role of spatial concepts in marine benthic habitat mapping, this dissertation focuses on one specific type of spatial data: bathymetry. Since it often is the only reliable continuous dataset available to characterize benthic habitats, particularly in deeper waters, bathymetry has become the most important dataset in marine benthic habitat mapping. By definition, bathymetry is the representation of the seafloor that defines the benthic component of benthic habitats. Bathymetry is also the primary input to geomorphometric analyses. Based on this, this dissertation focuses on the integration of spatial concepts in marine benthic habitat mapping when marine geomorphometry practices are also integrated in the workflow. While the dissertation addresses many spatial concepts, including spatial autocorrelation and spatial heterogeneity, it focuses on issues related to spatial scale, spatial data quality, and spatial covariation among environmental variables like terrain attributes that are commonly used in marine habitat mapping.

This dissertation is based on the evaluation of different practices and methods related to the production of benthic habitat maps. Consequently, it often necessitated a control dataset on which hypotheses could be tested. Two main datasets were used in this dissertation. The first one includes nine artificial terrain surfaces that provided a controlled environment for defining a framework for the use of marine geomorphometry in marine benthic habitat mapping (*cf.* Chapter 3). The second dataset is that of the German Bank, an area off Nova Scotia that has been extensively studied before (*e.g.* DFO, 2006; Todd *et al.*, 2012; Brown *et al.*, 2012). This dataset include bathymetry, backscatter data, and two types of ground-truth data (photographs of the seafloor and biological observations). This dataset is an excellent example of complete dataset that can be found in the literature, and enabled controlling variables for testing hypotheses in Chapter 4, Chapter 5, and Chapter 6. Its previous uses also provide opportunities for comparison of methods and results.

1.7 Significance of Research

By demonstrating the impacts of an inappropriate integration of spatial concepts in the marine benthic habitat mapping workflow, this dissertation will increase awareness amongst habitat mapping practitioners of the importance to consider these concepts. This should lead to an increased geographic literacy within marine habitat mapping. It should also result in more tools being developed and made accessible through GIS to a wide range of marine scientists, which should facilitate the integration of spatial concepts in the workflow and eventually improve standards and protocols. Ultimately, an increased realization amongst the community of the importance of spatial concepts should lead to a change in practices, which will enable the production of knowledge on benthic habitats grounded on a sound, spatially-explicit inferential basis. Consequently, such knowledge will better support decisions made from habitat maps in contexts such as conservation and management. Finally, this dissertation also defines standards for the use of geomorphometry in marine habitat mapping, which now provide an operational framework for habitat mapping practitioners willing to integrate geomorphometric analyses in their workflow. The adoption of that framework by the community will now enable valid comparisons between studies.

1.8 Organization of the Dissertation

This dissertation is organized into seven sections: the introduction, five manuscripts, and a conclusion chapter.

Chapter 2 reviews the marine benthic habitat mapping literature and highlights the importance of incorporating ecological scaling and geographical theories in this field. Recommendations are provided on more effective practices for marine benthic habitat mapping.

Chapter 3 focuses on the use of terrain attributes derived from DTMs. An optimal combination of terrain attributes for use in marine benthic habitat mapping was proposed after testing 230 tools and algorithms on nine artificial surfaces.

Chapter 4 uses the recommended selection of terrain attributes from Chapter 3 – defined in a theoretical context – in a practical context of marine benthic habitat mapping.

It demonstrates the selection's superiority over others in the production of marine habitat maps and species distribution models.

Chapter 5 addresses spatial data quality at multiple spatial scales. It describes how different types of data acquisition artefacts that are common in multibeam bathymetric data impact the derivation of terrain attributes from bathymetry.

Chapter 6 pushes forward the analyses from Chapter 5 by exploring how the artefacts errors propagate to habitat maps and species distribution models when bathymetry and terrain attributes are used in their production.

In *Chapter 7* the relevance and implications of this research is discussed within the context of marine benthic habitat mapping but also in a wider context of ecological and marine research. Future research areas that can build upon this research but are beyond the scope of this dissertation are also discussed.

1.9 Literature Cited

Anderson, J.T., Holliday, D.V., Kloser, R., Reid, D.G., & Simard Y. (2008) Acoustic seabed classification: current practice and future directions. *ICES Journal of Marine Science*, 65:1004-1011.

Blaschke, T., & Merschdorf, H. (2014) Geographic information science as a multidisciplinary and multiparadigmatic field. *Cartography and Geographic Information Science*, 41:196-213.

Blaschke, T., & Strobl, J. (2010) Geographic Information Science Developments. *GIS.Science Zeitschrift für Geoinformatik*, 23:9-15.

Blaschke, T., Strobl, J., & Donert, K. (2011) Geographic information science: building a doctoral programme integrating interdisciplinary concepts and methods. *Procedia – Social and Behavioral Sciences*, 21:139-146.

Blaschke, T., Strobl, J., Schrott, L., Marschallinger, R., Neubauer, F., Koch, A., Beinat, E., Heistracher, T., Reich, S., Leitner, M., & Donert, K. (2012) Geographic information science as a common cause for interdisciplinary research. *In*: Gensel, J., Josselin, D., & Vandenbroucke, D. (Eds) *Bridging the Geographic Information Sciences, Springer Lecture Notes in Geoinformation and Cartography*. Berlin: Springer, pp. 411-427.

Borja, A. (2014) Grand challenges in marine ecosystems ecology. *Frontiers in Marine Science*, 1:1-6.

Brown, C.J., Sameoto, J.A., & Smith, S.J. (2012) Multiple methods, maps, and management applications: purpose made seafloor maps in support of ocean management. *Journal of Sea Research*, 72:1-13.

Brown, C.J., Smith, S.J., Lawton, P., & Anderson, J.T. (2011) Benthic habitat mapping: a review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuarine, Coastal and Shelf Science*, 92:502-520.

Bryan, T.L., & Metaxas, A. (2007) Predicting suitable habitat for deep-water gorgonian corals on the Altantic and Pacific continental margins of North America. *Marine Ecology Progress Series*, 330:113-126.

Buhl-Mortensen, L., Buhl-Mortensen, P., Dolan, M.F.J., & Gonzalez-Mirelis, G. (2015) Habitat mapping as a tool for conservation and sustainable use of marine resources: some perspectives from the MAREANO Programme, Norway. *Journal of Sea Research*, 100:46-61.

Butler, J., Neuman, M., Pinkard, D., Kvitek, R., & Cochrane, G. (2006) The use of multibeam sonar mapping techniques to refine population estimates of the endangered white abalone (*Haliotis sorenseni*). *Fishery Bulletin*, 104:521-532.

Caron, C., Roche, S., Goyer, D., & Jaton, A. (2008) GIScience journal ranking and evaluation: an international Delphi study. *Transaction in GIS*, 12:293-321.

Cogan, C.B., & Noji, T.T. (2007) Marine classification, mapping, and biodiversity analysis. In: Todd, B.J., & Greene, H.G. (Eds) Mapping the seafloor for habitat characterization, *Geological Association of Canada*, Special Paper 47, pp. 129-139.

Coggan, R., Populis, J., White, J., Sheehan, K., Fitzpatrick, F., & Piel, S. (2007) Review of standards and protocols for seabed habitat mapping, 2nd edition. MESH (Mapping European Seabed Habitats). www.emodnet-seabedhabitats.eu/default.aspx?page=1442.

Cogswell, A.T., Kenchington, E.L.R., Lirette, C.G., MacIsaac, K., Best, M.M., Beazley, L.I., & Vickers, J. (2009) The current state of knowledge concerning the distribution of coral in the Maritimes Provinces. *DFO Canadian Technical Report of Fisheries and Aquatic Sciences*, 2885, 66 p.

Davies, J.S., Stewart, H.A., Narayanaswamy, B.E., Jacobs, C., Spicer, J., Golding, N., & Howell, K.L. (2015) Benthic assemblages of the Anton Dohrn Seamount (NE Atlantic): Defining deep-sea biotopes to support habitat mapping and management efforts with a focus on vulnerable marine ecosystems. *PLoS ONE*, 10:e0124815.

DFO (2006) Presentation and review of Southwest Nova Scotia benthic mapping project. *DFO Canadian Science Advisory Secretariat*, Proceedings Series 2006/047.

Dolan, M.F.J., Grehan, A.J., Guinan, J.C., & Brown, C. (2008) Modelling the local distribution of cold-water corals in relation to bathymetric variables: adding spatial context to deep-sea video data. *Deep-Sea Research I*, 55:1564-1579.

Dolan, M.F.J., & Lucieer, V.L. (2014) Variation and uncertainty in bathymetric slope calculations using geographic information systems. *Marine Geodesy*, 37:187-219.

Etnoyer, P., & Morgan, L.E. (2007) Predictive habitat model for deep gorgonian needs better resolution: comment on Bryan & Metaxas (2007). *Marine Ecology Progress Series*, 339:311-312.

Florinsky, I.V. (2012) Digital terrain analysis in soil science and geology. Elsevier/Academic Press, The Netherlands, 379 p.

Friedlander, A.M., Boehlert, G.W., Field, M.E., Mason, J.E., Gardner, J.V., & Dartnell, P. (1999) Sidescan-sonar mapping of benthic trawl marks on the shelf and slope off Eureka, California. *Fishery Bulletin*, 97:786-801.

Galparsoro, I., Borja, A., & Uyarra, M.C. (2014) Mapping ecosystem services provided by benthic habitats in the European North Atlantic Ocean. *Frontiers in Marine Science*, 1:23-36.

Gosselin, L.A., & Qian, P.Y. (1997) Juvenile mortality in benthic marine invertebrates. *Marine Ecology Progress Series*, 146:265-282.

Harris, P.T., & Baker, E.K. (2012) *Seafloor geomorphology as benthic habitat: GeoHab atlas of seafloor geomorphic features and benthic habitats*. Elsevier, Amsterdam.

Howell, K.-L., Piechaud, N., Downie, A.-L., & Kenny, A. (2016) The distribution of deep-sea sponge aggregations in the North Atlantic and implications for their effective spatial management. *Deep-Sea Research Part I*, 115:203-220.

Kostylev, V.E., Todd, B.J., Fader, G.B.J., Courtney, R.C., Cameron, G.D.M., & Pickrill, R.A. (2001) Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs. *Marine Ecology Progress Series*, 219:121-137.

Kuhn, W. (2012) Core concepts of spatial information for transdisciplinary research. *International Journal of Geographical Information Systems*. 26:2267-2276.

Laffoley, D., & Hiscock, K. (1993) The classification of benthic estuarine communities for nature conservation in Great Britain. *Netherland Journal of Aquatic Ecology*, 27:181-187.

Lam, N.S., & Kemp, K.K. (2012) Reflections on Geographic Information Science: special issue in honor of Michael Goodchild. *International Journal of Geographical Information Science*, 26:2193-2196.

Lecours, V., Dolan, M.F.J., Micallef, A., & Lucieer, V. (2016) A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrology and Earth System Sciences*, 20:3207-3244.

Le Pape, O., Delavenne, J., & Vaz, S. (2014) Quantitative mapping of fish habitat: a useful tool to design spatialised management measures and marine protected area with fishery objectives. *Ocean & Coastal Management*, 87:8-19.

Light, J.M. (1998) Marine molluscan conservation: the value of mapping as a conservation tool. *Journal of Conchology*, special issue, 147-154.

Lundblad, E., Wright, D.J., Miller, J., Larkin, E.M., Rinehart, R., Naar, D.F., Donahue, B.T., Anderson, S.M., & Battista, T. (2006) A benthic classification scheme for American Samoa. *Marine Geodesy*, 29:89-111.

Mark, D.M. (2000) Geographic information science: critical issues in an emerging cross-disciplinary research domain. *Journal of the Urban and Regional Information Systems Association*, 12:45-54.

Mark, D.M. (2003) Geographic information science: defining the field. *In*: Duckham, M., Goodchild, M.F. & Worboys, M.F. (Eds) *Foundations of Geographic Information Science*. London: Taylor & Francis, pp. 3-18.

Murillo, F.J., Munoz, P.D., Altuna, A., & Serrano, A. (2011) Distribution of deep-water corals of the Flemish Cap, Flemish Pass, and the Grand Banks of Newfoundland (Northwest Atlantic Ocean): interaction with fishing activities. *ICES Journal of Marine Science*, 68:319-332.

Olsgard, F., Brattegard, T., & Holthe, T. (2003) Polychaetes as surrogates for marine biodiversity: lower taxonomic resolution and indicator groups. *Biodiversity & Conservation*, 12:1033-1049.

Raper, J. (2009) Geographical information science. In: Cronin, B. (Ed.) *Annual Review of Information and Science and Technology*.

Rattray, A., Ierodiaconou, D., Monk, J., Laurenson, L.J.B., & Kennedy, P. (2014) Quantification of spatial and thematic uncertainty in the application of underwater video for benthic habitat mapping. *Marine Geodesy*, 37:315-336.

Reiss, H., Birchenough, S. Borja, A., Buhl-Mortensen, L., Craeymeersch, J., Dannheim, J., Darr, A., Galparsoro, I., Gogina, M., Neumann, H., Populus, J., Rengstorf, A.M., Valle, M., Hoey, G.V., Zettler, M.L., & Degraer, S. (2014) Benthos distribution modeling and its relevance for marine ecosystem management. *ICES Journal of Marine Science*, 72:297-315.

Rengstorf, A.M., Grehan, A., Yesson, C., & Brown, C. (2012) Towards high-resolution habitat suitability modeling of vulnerable marine ecosystems in the deep-sea: resolving terrain attribute dependencies. *Marine Geodesy*, 35:343-361.

Roberts, C.M. (2002) Deep impact: the rising toll of fishing in the deep sea. *TRENDS in Ecology & Evolution*, 17:242-245.

Robinson, L.M., Elith, J., Hobday, A.J., Pearson, R.G., Kendall, B.E., Possingham, H.P., & Richardson, A.J. (2011) Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. *Global Ecology and Biogeography*, 20:789-802.

Roff, J.C., Taylor, M.E., & Laughren, J. (2003) Geophysical approaches to the classification, delineation and monitoring of marine habitats and their communities. *Aquatic Conservation*, 13:77-90.

Rolet, C., Spilmont, N., Dewarumez, J.-M., & Luczak, C. (2015) Linking macrobenthic communities structure and zonation patterns on sandy shores: mapping tool toward management and conservation perspectives in Northern France. *Continental Shelf Research*, 99:12-25.

Smith, T.A., & McConnaughey, R.A. (2016) The applicability of sonars for habitat mapping: a bibliography. U.S. Department of Commerce, NOAA Technical Memo NMFS-AFSC-317, 129 p.

Solan, M., Germano, J.D., Rhoads, D.C., Smith, C., Michaud, E., Parry, D., Wenzhöfer, F., Kennedy, B., Henriques, C., Battle, E., Carey, D., Iocco, L., Valente, R., Watson, J., & Rosenberg, R. (2003) Towards a greater understanding of pattern, scale and process in marine benthic systems: a picture is worth a thousand worms. *Journal of Experimental Marine Biology and Ecology*, 285-286:313-338.

Tittensor, D.P. Baco, A.R., Brewin, P.E., Clark, M.R., Consalvey, M., Hall-Spencer, J., Rowden, A.A., Schlacher, T., Stocks, K.I., & Rogers, A.D. (2009) Predicting global habitat suitability for stony corals on seamounts. *Journal of Biogeography*, 36:1111-1128.

Todd, B.J, Kostylev, V.E., & Smith, S.J. (2012) Seabed habitat of a glaciated shelf, German Bank, Atlantic Canada. In: Harris, P.T., & Baker, E.K. (Eds) *Seafloor geomorphology as benthic habitat*. Amsterdam: Elsevier, pp. 555-568.

Whitmire, C.E., Embley, R.W., Wakefield, W.W., Merle, S.G., & Tissot, B.N. (2007) A quantitative approach for using multibeam sonar data to map benthic habitats. In: Todd, B.J., & Greene, H.G. (Eds) Mapping the seafloor for habitat characterization. *Geological Association of Canada Special Paper* 47:111-126.

Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., & Grehan, A.J. (2007) Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30:3-35.

Wright, D.J., Goodchild, M.F., & Proctor, J.D. (1997) GIS: tool or science? Demystifying the persistent ambiguity of GIS as "tool" versus "science". *Annals of the Association of American Geographers*, 87:346-362.

Wright, D.J., & Heyman, W.D. (2008) Introduction to the special issue: marine and coastal GIS for geomorphology, habitat mapping, and marine reserves. *Marine Geodesy*, 31:223-230.

Zhou, Q., & Zhu, A.-X. (2013) The recent advancement in digital terrain analysis and modeling. *International Journal of Geographic Information Science*, 27:1269-1271.

2. Spatial Scale and Geographic Context in Benthic Habitat Mapping: Review and Future Directions

2.1 Introduction

The volume of space that can host life on Earth is at least 150 times greater in the oceans than on land (Gjerde, 2006). However, scientific knowledge about marine environments is still sparse compared to terrestrial environments due to difficulties to access, observe, and sample most places in the marine realm (Solan *et al.*, 2003; Robinson *et al.*, 2011). The oceans, which cover 70% of our planet's surface, are estimated to be 90% unexplored (Gjerde, 2006). Ocean research led by several international initiatives and groups (*e.g.* Census of Marine Life and the International Council for the Exploration of the Sea [ICES]) has increased significantly over the last decade (Heyman & Wright 2011; Borja, 2014), driven by efforts by many nations to better manage and protect marine resources. In the ocean realm, benthic ecosystems provide important services (Thurber *et al.*, 2013; Galparsoro *et al.*, 2014) but are also increasingly impacted by human activities (*e.g.* bottom-contact fishing, oil and gas extraction) (Halpern *et al.*, 2008; Williams *et al.*, 2010; Harris, 2012). Research on near-bottom environments and their associated biota has become essential to support effective monitoring and management strategies (Thrush & Dayton, 2002; Ramirez-Llodra *et al.*, 2011). Anthropogenic impacts on the seafloor alter benthic biodiversity (Cook *et al.*, 2013; Grabowski *et al.*, 2014), habitats (Jones, 1992; Puig *et al.*, 2012), and modify ecosystem structures and functions (Koslow *et al.*, 2000; Olsgard *et al.*, 2008). Ramirez-

Llodra *et al.* (2011) noted that exploration, scientific research, monitoring, and conservation measures are essential to ensure that exploitation of resources does not lead to massive destruction of ecosystems. To protect benthic species from such threats, distribution patterns and ecological dynamics must be better understood (Ramirez-Llodra *et al.*, 2011; Mengerink *et al.*, 2014). Managers need accurate, quantitative and spatially explicit information, at scales relevant to their objectives, in order to support protection and management plans (Anderson *et al.*, 2008; Davies & Guinotte, 2011). Marine habitat mapping has become mandatory in some countries and contexts, such as the 1996 amendment to the United States Magnuson-Stevens Fishery Conservation and Management Act regarding the description and identification of essential fish habitats (Benaka, 1999). To ensure that these efforts are as representative as possible, species distributions should be mapped at multiple scales (Lourie & Vincent, 2004; Smith & Brennan, 2012; Shucksmith & Kelly, 2014). Mapping seafloor based on species' habitat requirements is essential and is the first step in implementing scientific management, monitoring environmental change, and assessing the impacts of anthropogenic disturbance on benthic habitats (Roff *et al.*, 2003; Cogan & Noji, 2007; Harris & Baker, 2012a).

Habitats can be defined as physical spaces characterized by a combination of variables of different types in which species can survive (Whittaker *et al.*, 1973). Several definitions of benthic habitats have been proposed. Harris & Baker (2012a, p. 8) define them as being "physically distinct areas of seabed that are associated with the occurrence of a particular species". A more comprehensive definition of benthic habitats could include the chemical environment and water properties known to influence benthic faunal distribution (Kostylev *et al.*, 2001; Cogan & Noji, 2007; Brown *et al.*, 2011a). A benthic habitat can hence be defined as an area of the seabed that is distinct from its surrounding in terms of physical, biological, and chemical variables. Brown *et al.* (2011a) provide a comprehensive review of types of benthic habitat maps, techniques of data collection, and methods that can be used to create habitat maps. Habitat-based approaches to estimate organism response to landscape heterogeneity have been used for decades in landscape ecology (Turner *et al.*, 2001; Robinson *et al.*, 2011). Because species have a range of environmental preferences and requirements (Hutchinson & MacArthur, 1959), many of these approaches focus on the structure and quantity of potential habitats, either instead of, or in addition to, the distribution of biological populations at the time of sampling.

Habitat maps must be placed in context with the appropriate spatial, temporal, and thematic scales (Cogan & Noji, 2007). Scale is considered to be "one of the most critical aspects in habitat mapping, as well as one of the most misunderstood" (Greene *et al.*, 2007, p. 145). As Boyce (2006, p. 274) stated: "Ecologists are still at a fairly naïve pattern-documentation phase in understanding the importance of scale." Despite the well-known importance of spatial scale in benthic habitat mapping (Brown *et al.*, 2011a), the topic is only briefly mentioned in texts (*e.g.* Todd & Greene, 2007; Harris & Baker, 2012b) and only a few publications address the implications of scale for benthic habitat mapping. Brown *et al.* (2011a) includes a complete section on spatial scale in benthic habitat mapping. Other publications have addressed spatial resolution (*e.g.* Anderson *et* *al.*, 2008), the impact of scale in management and surrogacy assessment (*e.g.* McArthur *et al.*, 2009, 2010), and its impact in shallow water monitoring (*e.g.* Van Rein *et al.*, 2009).

Scale is only briefly acknowledged in the extensive literature on benthic habitat mapping, often with little or no treatment of the role of spatial scale in the production of benthic maps and the interpretation of research results. This lack of treatment likely indicates little awareness and understanding of the importance and role that spatial scale plays in benthic habitat mapping. Figure 2.1 illustrates the increase in publications on benthic habitat mapping for the period 1995-2014, and the number of cases that address scale. Approximately a third of the articles and reviews used the term "scale" in the title, abstract or keywords, with 22% for "spatial scale", less than 5% for "multiple scales" and 1% for "multiscale"; these numbers are much lower than in landscape ecology-related publications, where scale is still considered as being insufficiently described (Lechner *et al.*, 2012a).

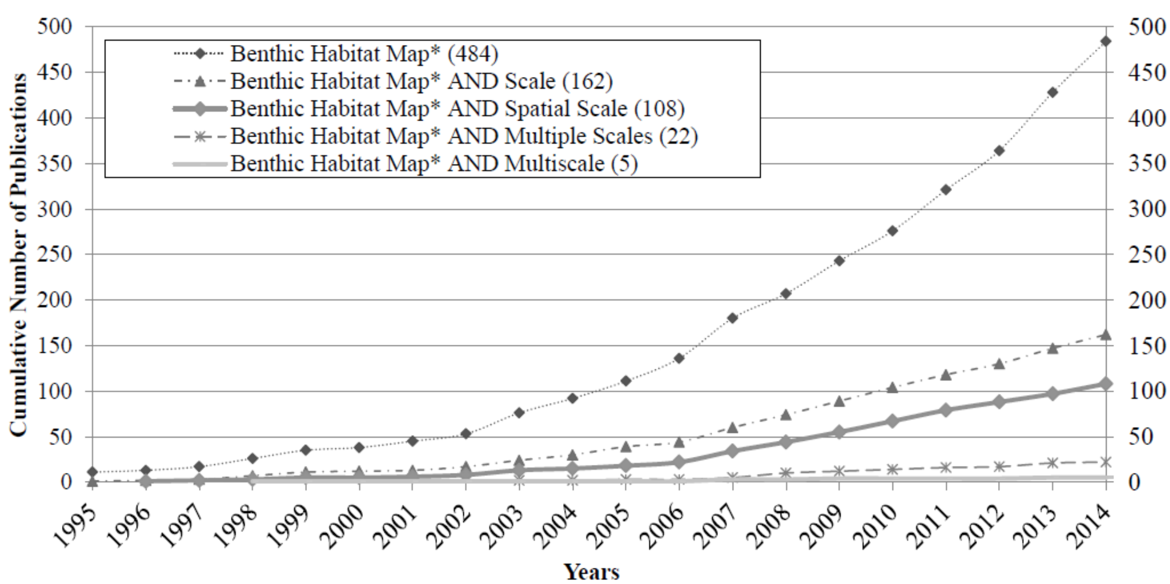

Figure 2.1: Cumulative number of publications (articles or reviews) listed in the Scopus database mentioning specific keywords (see key) in their title, abstract or keywords, by the end of 2014.

The aims of this contribution are (1) to review existing knowledge on spatial scale in benthic habitats and their mapping, including the related practices of surrogacy assessment and species distribution modelling, and (2) identify ways to improve benthic habitat mapping practices. The paper is organized as follows. We first review knowledge of scale in ecology, including the difference between scales of phenomenon, observation and analysis. We then introduce the concepts of benthic habitat mapping, including the natural characteristics that can influence marine species distribution, the basis of their representation as spatial data and of their analysis, and the importance of characterizing habitat at multiple scales. Thirdly, we emphasize the need to consider the spatial nature of data in analyzing species' relationships with their environment. Fourth, we discuss current needs and future directions in habitat mapping, and propose a new standard for defining benthic habitat that includes the explicit statement of scale. Finally, we make recommendations regarding the integration of ecological scaling and geographical theories in habitat mapping.

2.2 Scale in Ecology

Three types of scale are typically recognized in the ecological literature: spatial, temporal, and thematic. Several definitions of spatial scale have been given depending on the contexts (Schneider, 1994, 2001a; Dungan *et al.*, 2002; Lechner *et al.*, 2012b). Spatial scale commonly refers to the spatial characteristic of an object or process, including both its spatial resolution (*i.e.* level of detail) and geographic extent (Schneider, 1994; Gustafson, 1998). Like spatial scale, temporal scale is characterized by both resolution (*e.g.* days vs. minutes) and extent (*i.e.* range of time) (Schneider, 1994). Space and time are intrinsically linked and often depicted in joint space-time diagrams (Stommel, 1963; Steele, 1978; Delcourt *et al.*, 1983). Thematic scale, also called level of organization, organizational scale, or ecological organization, is linked to the level at which objects of study are described, for instance taxonomic resolution (Levin, 1992; Larsen & Rahbek, 2005). Thematic scale is important because the observed relationships of any 2 variables can vary across thematic scales (Pearson, 2002; Larsen & Rahbek, 2005; Brown *et al.*, 2011a, 2012). For instance, grouping species with different habitat requirements can result in conclusions that differ from when species are studied individually (*e.g.* Grober-Dunsmore *et al.*, 2007; Lecours *et al.*, 2013). Understanding the effects of spatial, temporal, and thematic scale is challenging but essential, as many important ecological processes are scale-dependent (Turner *et al.*, 2001; Schneider, 2009; De Knegt *et al.*, 2010). Changes in pattern with changes in scale have been recognized in ecology since the 1950s (Greig-Smith, 1952), but the importance of scale only became widely acknowledged in the 1980s (Meentemeyer, 1989; Schneider, 2001a). Research on scale, on methods to scale-up and scale-down across scales, and on the problem of relating phenomena across scales is fundamental and remains an important focus in many sciences (Wiens, 1989; Levin, 1992; Schoch & Dethier, 1996). According to Turner *et al.* (2001, p. 330): “The effects of scale are now well recognized, but the need for improved quantitative understanding remains critical.”

Lechner *et al.* (2012a, b) distinguish the scale at which a pattern or process occurs from the scale of observation and the scale of analysis. The scale at which a pattern or process occurs is often referred to as intrinsic, operational, or ecological scale. The observational scale relates to the data that are used to describe natural phenomena (*e.g.* the pixel size, or spatial resolution, on gridded bathymetric data), while the analysis scale relates to the method used to analyze these data (*e.g.* the size of the analytic window used to perform focal statistics in spatial analysis).

Issues can arise when there is a mismatch between ecological, observational, and analytical scales: appropriate detection of species-habitat relationships and ecological patterns is dependent on the chosen observational and analysis scales (García & Ortiz-Pulido, 2004; Gambi & Danovaro, 2006). For instance, using a 1 km resolution bathymetric dataset would likely not allow the understanding of how bathymetry relates to species distribution in a coral reef, as knowledge of smaller changes in depth would be required. Observational and analytic scales are often arbitrarily chosen in ecological studies (Levin, 1992), due to financial, technical or time constraints (Meentemeyer, 1989) and are typically not reported in sufficient details (Pittman & McAlpine, 2003). Wheatley & Johnson (2009) reviewed the use of multiple scales in terrestrial wildlife-habitat studies, finding that 70% of the articles used arbitrarily chosen scales, with no consideration of the scales relevant to wildlife or to environmental variables. They mentioned that when such choices are made, "published results may reflect scale artefacts" and scale-dependent processes may be missed "by examining irrelevant or redundant scales of observation" (Wheatley & Johnson, 2009, p. 151). Scale artefacts are observations that seem to explain the studied pattern or process, but may not be causally linked or cannot be validated due to the choice of observational scale (Wheatley & Johnson, 2009; Lechner *et al.*, 2012b). In order to avoid scale artefacts and missing important patterns or processes, data and analysis need to capture the essential elements of the habitat, meaning that the observational and analytic scales should encompass the ecological scales of the biological or environmental phenomenon being studied (Hobbs, 2003; Mayor *et al.*, 2009; Goodchild, 2011). Habitat structure must then be measured at spatial scales relevant to the organism of interest (Pearson, 2002; Gallucci *et al.*, 2009; De Knegt *et al.*, 2010). For instance, the habitat of a wide-ranging shark would not be measured at the same scales as the habitat of a small cavity-dwelling reef fish, even if they are found within the same geographic area.

No single scale, be it spatial, temporal or thematic, is appropriate for the study of all ecological problems, and all scales do not have similar explanatory powers (Clark, 1985; Wiens, 1989; Levin, 1992; Willis & Whittaker, 2002). For instance, coarse-scale data can help understand regional patterns of terrestrial and marine species biogeography (*e.g.* Rahbek & Graves, 2001; Davies *et al.*, 2008) but may be insufficient for identifying specific conservation areas (Davies & Guinotte, 2011). Models created with coarse-scale data to predict a species' geographic distribution can be improved using better knowledge of its habitat requirements gained from finer-scale information (Bryan & Metaxas, 2007; Etnoyer & Morgan, 2007; Davies & Guinotte, 2011; Ross & Howell, 2013). However, saying that no single scale is appropriate does not mean that all scales serve a purpose equally well or that scaling laws or patterns cannot be defined (Levin, 1992).

Spatial scale is an important consideration when studying organism and habitat structure interactions (McCoy *et al.*, 1991; Pearson, 2002). Habitat selection by a particular species can occur and be measured at some scales and not necessarily at others

(Owen, 1972; Boyce, 2006). For instance, Anderson *et al.* (2005) found that elks select their habitat based on broad-scale spatial distribution of wolves in conjunction with fine-scale selection of forage areas. In a marine context, the associations of infaunal (De Leo *et al.*, 2014), sessile (Schneider *et al.*, 1987), and mobile epibenthic species (Grober-Dunsmore *et al.*, 2007; Kendall *et al.*, 2011) with their environment were all found to vary with spatial scale. The concept of habitat is both scale-dependent (Pearson, 2002) and species-specific (Pandit *et al.*, 2009). For instance, habitat specialists, such as coral reef gobies (Munday *et al.*, 1997), live in a very specific habitat characterized by a narrow range of environmental conditions and respond to more fine-scale processes. On the other hand, habitat generalists, such as the copepod *Nitocra spinipes* (Pandit *et al.*, 2009), can tolerate a broad range of environmental conditions and respond to more broad-scale processes. More generally, Schneider *et al.* (1987) found that mobile species often show decoupling from the environment at finer scales, and habitat association at coarser scales, compared to finer-scale coupling with habitat by sessile species. Meyer & Thuiller (2006) reported that the majority of species respond to habitat characteristics at more than one scale at the same time. Despite that, a response measured at one particular scale cannot always be used to predict habitat use at another scale (VanderWerf, 1993; Apps *et al.*, 2001).

2.3 Scale in Benthic Habitat Mapping

2.3.1 Review of Concepts and Methods

2.3.1.1 Habitat Mapping

The complex interactions between biological, physical, chemical, and behavioural elements of the marine environment can make benthic habitats difficult to map (Zajac, 2008; Rigby *et al.*, 2010). The integration of data representing these elements at multiple scales is especially challenging (Brown *et al.*, 2011a). Traditional data-acquisition techniques can be limited by varying factors, including depth (as with optical remote sensing that only captures data in shallow waters), visibility (as with cameras), and time (as with SCUBA diving) (Dunn & Halpin, 2009; Costa *et al.*, 2014). Whilst some techniques can help delineate benthic habitats at some specific scales, they present challenges when trying to delineate benthic habitats at other scales. For instance, seafloor acoustic mapping from the surface and sparse ground-truthing in deeper waters provide information at a scale that Davies *et al.* (2008) considered regional, but lack the capacity to characterize finer-scale patterns and processes (Stone, 2006; Davies *et al.*, 2008; Tittensor *et al.*, 2009). On the other hand, SCUBA diving allows the collection of fine-scale data in shallow waters but cannot generate a broader characterization of ecosystem pattern (Costa *et al.*, 2014). The use of bathymetric LiDAR (in shallow waters) and acoustic remote sensing (in deeper waters) can help reduce these sampling gaps, by providing continuous and high-resolution data necessary for mapping over greater areas, and thus at scales that may be more relevant for understanding pattern and process in these habitats (Kenny *et al.*, 2003). Despite their strengths, these techniques have their own limitations, as they do not necessarily provide data of sufficient resolution to understand very fine ecological processes. However, combining acoustic or LiDAR data with *in situ* observations, high-resolution geoscientific and environmental information, and spatial analytical techniques does allow for more accurate quantitative characterization of habitat at multiple scales, in addition to providing a framework for mapping the distribution of benthic species and interpreting spatial patterns in biodiversity (Whitmire *et al.*, 2007; Wedding *et al.*, 2008; Brown *et al.*, 2011a; Harris & Baker, 2012a).

Brown *et al.* (2011a) identified 3 of the most common approaches to benthic habitat mapping: abiotic surrogate mapping that does not consider biological data, and unsupervised (top-down approach) and supervised (bottom-up approach) classifications that integrate biological data in different ways (see Figure 4 in Brown *et al.*, 2011a). These methods correspond to what the "Review of Standards and Protocols for Seabed Habitat Mapping" published by MESH (Mapping European Seabed Habitats) identified as the general approach to benthic habitat mapping: the spatial integration of different datasets, usually within a geospatial environment (Coggan *et al.*, 2007). While a number of studies (*e.g.* Brock *et al.*, 2004; Wedding & Friedlander, 2008) mapped benthic habitats in shallow environments using bathymetric LiDAR, optical remote sensing, or SCUBA diving, this approach often focuses on the use of acoustic remote sensing (*e.g.* multibeam echosounders, sidescan sonars) to collect spatial information on the characteristics of the seafloor (Brown *et al.*, 2011a); most of the 57 case studies presented in Harris & Baker (2012b) used either backscatter or bathymetric data, or both. For example, Copeland *et al.* (2012) combined information extracted from bathymetric and backscatter data with biota in a sub-Arctic fjord to determine 6 types of benthic habitats and to identify patterns of biodiversity. All the techniques used to map both shallow and deeper waters influence or determine the scale of data collection and analysis. For instance, the spatial resolution and extent of acoustic bathymetric data depends on the sensor-to-seafloor distance (*e.g.* Lecours & Devillers, 2015) and the systems used (Kenny *et al.*, 2003): the shorter the distance, the higher the resolution and the lower the extent.

In parallel, approaches from terrestrial ecology are increasingly used in marine ecology to represent environmental heterogeneity as habitat maps. Seascape ecology draws on techniques from landscape ecology, using spatial pattern metrics to quantify the seascape structure and delineate patch-based models of habitat type (see Boström *et al.*, 2011; Pittman *et al.*, 2011; Wedding *et al.*, 2011). The size of habitat patches can be an indicator of the spatial scale at which species use an environment when linked to species distribution and behaviour (Pittman & McAlpine, 2003; Pittman *et al.*, 2007). For instance, Hitt *et al.* (2011) tracked fish movements, linking them to seascape structures to study habitat use in relation to patch types and connectivity, which allowed quantifying the extent of the environment that the fish were using. The literature on seascape ecology is however still scarce (Pittman *et al.*, 2011). Applications are mostly in coastal shallow environments, using optical remote sensing (*i.e.* aerial photography or satellite remote sensing) (*e.g.* Kendall & Miller, 2010) or bathymetric LiDAR data (*e.g.* Purkis & Kohler, 2008), and are often applied to reef fishes (*e.g.* Kendall *et al.*, 2011). Despite its potential to explain marine ecological patterns and processes at multiple scales (Schoch & Dethier, 1996), seascape ecology has yet to be implemented in deeper water using acoustic bathymetric data. Habitat maps developed in a seascape ecology context also involve the consideration of spatial scale, as the spatial pattern metrics are dependent on the resolution and extent of the input data that influence the minimum mapping unit (MMU) (Saura, 2002; Fassnacht *et al.*, 2006; Kendall *et al.*, 2011). MMU is the size of the smallest area to be mapped as a discrete unit, and its selection determines the scale at which patches are defined in a seascape: as the MMU increases, rare and smaller features tend to not be considered by the analysis, which can lead to erroneous interpretation (see Kendall & Miller, 2008).

Significant progress has been made in the understanding of benthic habitats in the last decade (see Todd & Greene, 2007; Harris & Baker, 2012b) despite the difficulties associated with their mapping, modelling, and management (Diaz *et al.*, 2004). Much work remains to be done to gain an adequate understanding of these complex ecosystems at relevant scales. For instance, very little work has been done on infaunal benthos (see De Leo *et al.*, 2014). Not only is most benthic diversity infaunal, but the rate of release of nutrients into the water column, a key benthic variable, is driven mostly by infaunal activity. Mapping benthic diversity to the species level is not possible in the absence of continuously mappable surrogates (see next subsection) for any one species. However, with sufficiently fine-scale data, it would be possible to map evidence of biogenic flux, such as castings or burrow diameters, through the sediment surface. For instance, acoustic reflectivity (backscatter) can capture fine-scale information of the sediment surface, which can then be combined with *in situ* ground-truthing in a benthic modelling approach (*e.g.* Brown *et al.*, 2011b; Freitas *et al.*, 2011; Copeland *et al.*, 2012). Another issue that constrains complete understanding of benthic ecosystems is the species-specific relation to habitat as a function of scale, which in turn complicates the study of species assemblages (Grober-Dunsmore *et al.*, 2007; Howell *et al.*, 2010, 2011; see also Brennan *et al.*, 2002; Brown *et al.*, 2011a). For example, Schneider *et al.* (1987) found that the scale-dependent association between population density and substrate differed between mobile and sedentary fauna.

2.3.1.2 Surrogacy

As in terrestrial ecology, the challenges associated with sampling marine organisms in relation to their environment has led to an increasing use of surrogates, also known as "proxies" (McArthur *et al.*, 2009, 2010; Anderson *et al.*, 2011). A surrogate can be defined as "a measurable entity that will represent, or substitute for, a more complex element of biodiversity that is more difficult to define or measure" (Harris & Baker, 2012b, p. 899). Surrogates can be any measurable characteristic of the environment, sampled either *in situ* at specific locations (*e.g.* sediment pH), or provided as continuous or near-continuous coverage, such as bathymetry derivatives (*e.g.* seabed roughness or slope). Before mapping habitats, surrogate variables for a particular species first need to be identified, together with the strength of covariation in the study, and the establishment of a biological basis for the covariation. For instance, the selection of surrogates to be tested and the scale at which they should be tested may be based on knowledge gained from previous observations, experimental work or some evidence of causal connection

(Brennan *et al.*, 2002). Surrogates may be relevant only at particular scales (Urban *et al.*, 1987; Gambi & Danovaro, 2006). For instance, Tong *et al.* (2013) found aspect (the geographic orientation of the slope) to be a good surrogate of the cold-water coral *Paragorgia arborea*'s presence over areas of 30 × 30 m and 90 × 90 m, but not at a broader scale. They linked this result to the presence of finer-scale bottom currents in the study area that bring food to the corals, which is not the case for broader-scale currents (Tong *et al.*, 2013). Once defined, surrogates can be used to map habitats, predict or estimate species distribution, and build habitat suitability models (*e.g.* Lucieer *et al.*, 2013; Hill *et al.*, 2014). The use of surrogates for these purposes cannot be trusted at spatial scales other than the scale at which the surrogate was defined.

2.3.1.3 Species Distribution Modelling

Combining georeferenced species occurrence data with environmental variables to develop habitat suitability and predictive distribution models is an important approach increasingly used in the marine environment (Heyman & Wright, 2011; Robinson *et al.*, 2011; Brown *et al.*, 2012; Hill *et al.*, 2014; Vierod *et al.*, 2014), especially for protection and management purposes (Ross & Howell, 2013). These models build on existing knowledge of species-environment relationships, either directly or via surrogates, to predict the location and extent of potential habitat in areas where only environmental information is available (see Elith & Leathwick, 2009; Zimmermann *et al.*, 2010 for general reviews; and Robinson *et al.*, 2011 and Vierod *et al.*, 2014 for specific reviews for the marine environment). The criteria (Brennan *et al.*, 2002; Franklin, 2009) to consider in the selection of a model for a particular application are (1) species characteristics, (2) data availability, (3) the observational and analysis scales, (4) stability in time (*e.g.* bathymetry compared to temperature), and (5) the biological and physical underpinnings (if any) of the model. As in terrestrial ecology, few marine studies address the issues of choosing an appropriate range of spatial scales at which to identify surrogates of species habitat or identify the appropriate scales at which to develop predictive models (Franklin, 2009). A coarser scale model may underrepresent the area of suitable habitat since the finer-scale habitat features that drive species distribution are not captured by the data (Seo *et al.*, 2009; Vierod *et al.*, 2014) (see also Figure 2.2). The scale (extent) of the study area also has a direct impact on the quality of the models (VanDerWal *et al.*, 2009; Hijmans, 2012), and Meyer & Thuiller's (2006) meta-analysis of species distribution modelling studies found that the use of environmental variables at more than one scale tends to give more accurate predictions. In the deep sea, the implementation of effective habitat suitability models is limited by the resolution and extent of environmental data (Vierod *et al.*, 2014), and will only be possible if high-resolution data become globally available (Davies *et al.*, 2008). Some data may not be available for an area, or may be available at an inappropriate scale. Often, certain variables (*e.g.* temperature, bottom current speed) are only available at a coarser resolution than other variables (*e.g.* slope and rugosity, measured using acoustic remote sensing techniques). Down-scaling or improved spatial measurement of the former to a level in line with the latter is needed to free models from errors in cross-scaling and to put knowledge of species distribution relative to habitat on a sound basis.

2.3.2 Ecological Scale: Benthic Species and Their Environment

2.3.2.1 Environmental and Biological Surrogates

Several environmental variables were found useful in characterizing marine habitats, with differing degrees of importance depending on species (*e.g.* Freeman & Rogers, 2003), locations (*e.g.* Georgian *et al.*, 2014), settings (*e.g.* submarine canyons) (*e.g.* De Leo *et al.*, 2014), and spatial scales (*e.g.* Gambi & Danovaro, 2006; Henry *et al.*, 2013). This diversity in use of environmental variables highlights the difficulties in quantifying the distribution of benthic organisms in relation to habitat. Reviews of potential surrogates of marine benthic biodiversity can be found (*e.g.* McArthur *et al.*, 2009; Howell, 2010; Harris & Baker, 2012b) but only McArthur *et al.* (2010) discuss the usefulness of surrogates in relation to spatial scale.

In addition to physical and chemical factors, biological factors and ecological interactions likely explain the distribution of benthic organisms at different scales (Robinson *et al.*, 2011). For example, reproduction strategies can influence species distribution, following spatial patterns in which organisms expect to disperse gametes over greater areas to reduce aggregation (Gage & Tyler, 1999). Ecological interactions can also be used as surrogates; if predation or commensalism is observed between 2 species (*e.g.* between structure-forming species and fishes), the presence of one could predict the other (Ward *et al.*, 1999; Tissot *et al.*, 2006; Mumby *et al.*, 2008; Baillon *et al.*, 2012). Both intraspecific and interspecific interactions vary with scale on land (Wiens *et al.*, 1986b; Sherry & Holmes, 1988) and in the ocean (Haury *et al.*, 1978). Mellin *et al.*'s (2011) meta-analysis of the effectiveness of biological surrogates in marine studies showed that biological surrogates tend to be more effective at finer spatial scales (*i.e.* smaller spatial extent). According to Leaper *et al.* (2012, p. 858), "the need for effective biological surrogates is especially critical in the marine realm, where a large number of species remain undescribed". Yet biological surrogates are rarely used in habitat mapping as they are difficult to assess at meaningful (often fine) spatial scales (Muotka *et al.*, 1998; Mellin *et al.*, 2011; Snickars *et al.*, 2014). The addition of biological surrogates to species distribution models can potentially improve predictions (Austin, 2002; Robinson *et al.*, 2011).

2.3.2.2 Combined Environmental Influence and Multicollinearity

Environmental variables identified as surrogates can act together to influence species distribution. For instance, the combination of topography and currents influences the levels of connectivity among populations for reproduction at different scales (Adams & Flieri, 2010; Rex & Etter, 2010). On seamounts, millimetre-scale colonization patterns are affected by coarser-scale flow patterns, however the motion of the fine-scale benthic boundary layer is also a determining factor (Gage & Tyler, 1999; Young, 2009). On continental slopes, rough seafloors interact with meso-scale currents to create complex circulation patterns that could potentially lead to the isolation of populations (Rex & Etter, 2010). Such relationships complicate data analysis because multicollinearity among variables occurs within and across scales (Rengstorf *et al.*, 2012; Laffan *et al.*, 2014). Multicollinearity occurs when 2 or more explanatory variables (*e.g.* water depth and temperature) are highly correlated (see Tabachnick & Fidell, 2013), obscuring the influence of each variable (Hengl & MacMillan, 2009; Tabachnick & Fidell, 2013).

Multicollinear explanatory variables are common in marine ecology but rarely considered in analyses (Wedding *et al.*, 2011): multicollinearity should systematically be tested (Pittman *et al.*, 2009). Statistical methods to address the problem can be found in Dormann *et al.* (2013) and Tabachnick & Fidell (2013).

That explanatory variables covary raises a question: how many and which variables are necessary to best characterize a habitat? In the past, a single surrogate was often used, but it is now widely accepted that biogeographic patterns are best explained by a combination of multiple variables (Hagberg *et al.*, 2003; McArthur *et al.*, 2009). Too few covariates can result in an overly general habitat characterization (Barry & Elith, 2006; VanDerWal *et al.*, 2009). The opposite, too many variables, can result in model over-fitting (Peterson & Nakazawa, 2008). According to Peterson *et al.* (2011), the number of variables will depend on the studied species, the complexity of the habitat, the availability of data, and the observational and analysis scales. Mateo Sánchez *et al.* (2014) argue that it is as important to identify the relevant environmental factors as to identify the scales at which these drive species distributions. Selecting relevant variables and at relevant scales is essential to the quality of habitat maps and the performance of predictive models (Austin, 2002; Williams *et al.*, 2012). The choices of variables and observational and analysis scales need to be based on their ecological relevance as these choices can impact the measurements of relationships between fauna and environmental variables (Araújo & Guisan, 2006; Synes & Osborne, 2011). Austin & Van Niel (2011) however report that assumptions made in the literature about the ecological relevance of variables vary among publications, are sometimes inconsistent, and so need to be revisited with a consideration of spatial scale.

2.3.3 Observational Scale: Representing Nature with Spatial Data

2.3.3.1 Adequacy of Spatial Data

When expressed as ecogeographical data (*i.e.* ecological variables with a geographic component), surrogates have a spatial dimension defined by their latitude, longitude, and depth (or altitude for terrestrial applications). A measure derived from these 3 spatial variables is geographical distance, an important predictor of fish species distributions in coral reefs (Pittman & Brown, 2011) and hard-bottom habitats (Dunn & Halpin, 2009). These spatial variables define the spatial scale (resolution and extent) of ecogeographical data and can themselves be used as surrogates (McArthur *et al.*, 2009). For instance, small changes in depth can better explain changes in populations than larger changes in latitude and longitude (Rex & Etter, 2010). However, their ecological meaning is arguable (Pittman & Brown, 2011). Depth for instance may itself be a surrogate of a causal variable such as light or temperature. A strong relation between a biological variable and non-causal surrogate (*e.g.* depth) can therefore obscure the relation to an underlying causal variable (*e.g.* light or temperature), reducing the predictive power of important covarying environmental variables (*e.g.* Araújo & Williams, 2000; Clarke & Lidgard, 2000; Hothorn *et al.*, 2011).

All data are not equally good at capturing the relevant information. Figure 2.2 illustrates this idea with the example of bathymetry: if one finds that only broad-scale bathymetric features, such as a large seamount (dotted ellipses), drive species distribution, then finer-scale data are not needed. On the other hand, if intermediate-scale features (*e.g.* smaller pinnacles or banks; dashed ellipses in Figure 2.2) influence species biogeography, finer-scale data would be required. If the detailed topography (*e.g.* single boulder; solid ellipses in Figure 2.2) represent ecologically important habitats, even finer-scale data would then be essential to capture the important information.

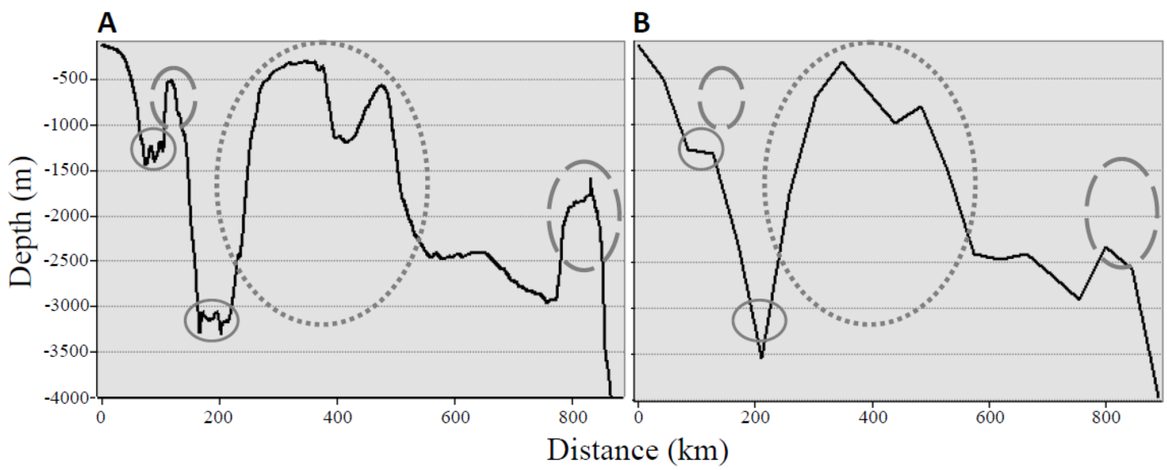

Figure 2.2: Seabed profiles (black lines) showing fine-scale (solid gray ellipses), intermediate-scale (dashed gray ellipses) and broad-scale (dotted gray ellipses) topographic features delineated using (A) finer-scale and (B) coarser-scale bathymetric data. By using only a coarse observational scale, information on potentially ecologically important finer-scale features is not captured. (Conceptual figure shows bathymetric profiles derived from the General Bathymetric Chart of the Oceans [GEBCO] dataset; www.gebco.net/).

The role of spatial scale has never been formally assessed in marine habitat mapping, despite repeated calls for an improved scientific understanding of benthic habitats at finer scales to allow better prediction of the geographic distribution of benthic species (Etnoyer & Morgan, 2007; Davies *et al.*, 2008; Davies & Guinotte, 2011; Rengstorf *et al.*, 2013).

This lack of assessment makes it difficult to define which observational scales are "fine enough" and which ones represent the upper limit of usefulness (Wilson *et al.*, 2007). In terrestrial environments, local biological interactions often complicate the observation of the relationships between species and abiotic variables; the opposite occurs at coarser scales (Levin, 1989; Sarkar *et al.*, 2005), which makes fine-scale studies more appropriate to investigate details of biological mechanisms and broad-scale studies for generalizations (Wiens, 1989). Wiens (1989) suggested that these patterns were likely to be the same in the marine realm but Steele (1991) showed that biological and physical phenomena do not scale in the same way in the ocean as on land. Planktonic life stages, the ability of some pelagic larvae to remain in an undeveloped stage until they find a suitable location to settle, and ocean fluid dynamics allow broad-scale dispersal into fine-scale suitable environments, which is not comparable to the finer-scale dispersal of many terrestrial species (Gray, 1966; Carr *et al.*, 2003; Kinlan & Gaines, 2003). In benthic habitat mapping, it is possible that an intermediate observational scale finer than the current coarse-scale studies (although not too fine) could provide more useful information (*cf.* dashed ellipses in Figure 2.2). For instance, Roberts *et al.* (2008) investigated communities at a local scale and concluded that intermediate-scale mapping might be useful to improve their results.

2.3.3.2 Data Quality and Spatial Scale

Several factors, including multicollinearity, autocorrelation (see "Adding geographic context: Spatial autocorrelation"), and spatial and thematic scales, can influence the accuracy of habitat maps (see Figure 4 in Wedding *et al.*, 2011). Despite recommendations to investigate and map variable uncertainty and error propagation when mapping habitats and species distribution (Rocchini *et al.*, 2011; Beale & Lennon, 2012; Vierod *et al.*, 2014), uncertainty and quality issues associated with spatial and non-spatial data are rarely addressed in habitat mapping (Lechner *et al.*, 2012a). Spatial data quality directly impacts the reliability of habitat maps, predictive models, and statistical description of species-habitat relationships (Menke *et al.*, 2009; Moudrý & Šímová, 2012). Data quality is conceptually related to spatial scale (Zhang *et al.*, 2014; Lecours & Devillers, 2015; Pogson & Smith, 2015). For instance, the finer the data resolution, the more that uncertainty and poor positional accuracy influence relationships between variables (Hanberry, 2013). Spatial matching between ecogeographical variables is particularly important: the positional error on biological data should always be smaller than the spatial resolution of the environmental data (Moudrý & Šímová, 2012; Lecours & Devillers, 2015) to avoid the emergence of false relationships between species and the environment, or the overestimation of a variable's range of values associated with a species (Guisan & Thuiller, 2005; Guisan *et al.*, 2007).

In habitat mapping and predictive modelling, a trade-off between data quality (*i.e.* accuracy and precision), sample size, and spatial scale (*i.e.* resolution and extent) must be considered (Brennan *et al.*, 2002; Lecours & Devillers, 2015). Despite attempts to address this challenge (*e.g.* Braunisch & Suchant, 2010), it is still unclear which characteristics should be given a higher priority in sampling strategy. Fine resolution data arguably yields better predictive models if the data quality is adequate, even if the sample size of biological data is smaller (Huston, 2002; Engler *et al.*, 2004; Kaliontzopoulou *et al.*, 2008; Reside *et al.*, 2011; Williams *et al.*, 2012), but this conclusion is not unanimous (Braunisch & Suchant, 2010). Some authors suggest using uncertainty to weight the variables in modelling and statistical analyses; information with less positional error can increase precision and thus improve models (Beale & Lennon, 2012; Moudrý & Šímová, 2012).

2.3.4 Analysis Scale: Influence on Analyzing Ecogeographical Data

Statistical relationships depend on the scale of analysis and results can vary as a function of it (Greig-Smith, 1952; Rahbek & Graves, 2001; Dungan *et al.*, 2002). An example is given in this sub-section using surrogate variables derived from bathymetry, which are among the most sensitive to the scale of analysis. Bathymetric data have proven their potential to advance understanding of seafloor ecosystems and their value for habitat mapping (Anderson *et al.*, 2008; Brown *et al.*, 2011a), and can be used in geomorphometry (*i.e.* terrain analysis) to quantify seafloor topography and complexity (Lecours *et al.*, 2015). In the last decade, a range of terrain attributes (*e.g.* slope, curvature) were found to have a relationship to marine biodiversity (McArthur *et al.*, 2009), thus inducing an increase in the application of geomorphometric techniques in marine habitat mapping (*e.g.* Wedding *et al.*, 2008; Zieger *et al.*, 2009; Rengstorf *et al.*, 2012; Tong *et al.*, 2013; Dolan & Lucieer, 2014). The relationship between spatial scale and terrain attributes has become an important research focus in geomorphometry (*e.g.* Florinsky & Kuryakova, 2000; Schmidt & Andrew, 2005; Deng *et al.*, 2007; Li, 2008), but a good understanding of scaling methods is still missing from geomorphometric analysis (Drăgut *et al.*, 2009). Terrain attributes vary with scale (Evans, 1972) and so their computation does not result in only one true, real fixed value, but in a range of possible values that depend on the resolution of the data and the extent of the analysis window (Shary *et al.*, 2002; Hengl, 2006). In the marine environment, coarse-scale geomorphometric analyses may not be adequate to resolve smaller features important for benthic biodiversity (Rengstorf *et al.*, 2012; Lecours *et al.*, 2013). The effects of the spatial resolution of bathymetry and terrain attributes on habitat suitability models are discussed in more detail by Rengstorf *et al.* (2012). Issues related to scale in geomorphometric analysis are similar to those in ecology and habitat mapping: it is widely accepted that a single scale (fixed resolution and window size) cannot completely describe a surface and capture all features of interest in an area (*cf.* Figure 2.2) (MacMillan & Shary, 2009; Goodchild, 2011). Yet many applications use a single scale, with an arbitrary choice of spatial resolution for the input surface and a single neighbourhood size (MacMillan & Shary, 2009). This limits analysis to those features that are observable at a single scale, which can have a significant impact on habitat maps and consequently on the resulting conclusions on species-habitat relationships. Similar scaling issues arise in the analysis of environmental data other than bathymetry.

2.3.5 Multiscale and Multi-Design Approaches

2.3.5.1 Multiple Scales and the MAUP

It has long been argued that ecology and geography would benefit from the adoption of a multiscale perspective in research, applications, and management (*e.g.* Stone, 1972; Legendre & Demers, 1984; Wiens *et al.*, 1986a; Addicott *et al.*, 1987; Meentemeyer, 1989; Conroy & Noon, 1996; Brennan *et al.*, 2002; Pittman & McAlpine, 2003). According to Wiens (1989, p. 394), "studies conducted at several scales or in which grain and extent are systematically varied independently of one another will provide a better resolution of domains, of patterns and their determinants, and of the interrelationships among scales." The implementation of multiscale analysis is, however, challenging and thus remains sporadic (Wheatley & Johnson, 2009) due to various difficulties including objective choice of sampling scales, simultaneous sampling of multiple scales (Addicott *et al.*, 1987; Brennan *et al.*, 2002), and the modifiable areal unit problem (MAUP) (Gehlke & Biehl, 1934; Openshaw, 1984; Marceau, 1999), also known as change-of-support (COS) in spatial statistics (Cressie, 1993; Cressie & Wikle, 2011). MAUP is defined by Harvey (2008, p. 284) as "the assumption that a relationship observed at one level of aggregation holds at another" and by Heywood *et al.* (2006, p. 416) as a "problem arising from the imposition of artificial units of spatial reporting on continuous geographic phenomena resulting in the generation of artificial spatial patterns." Combining data from 2 observational scales (*e.g.* when developing a habitat map) is invalid due to MAUP, and results from 2 different analytic scales (*e.g.* results of the quantification of species-habitat relationships at different scales) are not comparable: the aggregation of information taking place across changing spatial resolution or extent modifies the statistical properties (*e.g.* means, variances, and covariances) of the data, possibly resulting in distorted relationships between variables. Thematic scales can be very sensitive to MAUP. MAUP is related to Goodchild's (2011) concept of cross-scale inference, which occurs when inferences made at a coarser scale are transferred to a finer scale. Cross-scale inference is directly related to the concepts of ecological and atomistic fallacies (Robinson, 1950; Cressie & Wikle, 2011; see Lloyd, 2014). The action of inferring across scales without checking for MAUP or cross-scale inference may lead to misinterpretation of results (Openshaw & Taylor, 1979; Meentemeyer, 1989) and unfounded conclusions. However, methods exist to deal with MAUP: Zhang *et al.* (2014, p. 147) elaborate on multivariate geostatistics "to facilitate multisource and multiscale data integration", a relevant method for habitat mapping where data are often collected at different scales and with different sensors.

2.3.5.2 Multiscale and Multi-Design Frameworks

Wheatley & Johnson (2009) distinguish multiscale from multi-design sampling. The former is characterized by 2 elements: (1) the same environmental variables must be analyzed across scales and (2) there needs to be a change in only one of the 2 elements of spatial scale (*i.e.* resolution or extent). When both the spatial extent and resolution are changed, a study is multi-designed rather than multiscale (see Figure 2 in Wheatley & Johnson, 2009). Figure 2.3 illustrates the difference between the 2 approaches. In a number of studies, the term "multiscale" is inappropriately used to characterize the independent use of multiple scales, thus corresponding to a "multi-design" approach (*e.g.* Brennan *et al.*, 2002; Anderson & Yoklavich, 2007; Georgian *et al.*, 2014). The distinction is important as multi-designed studies cannot allow generalization and comparison of results between the different scales due to MAUP (Jelinski & Wu, 1996; Wu *et al.*, 1997; Nelson, 2001; see Lechner *et al.*, 2012b for MAUP in multiscale studies). Despite potential errors of interpretation caused by MAUP, comparisons between scales are often performed in the literature without exploring its effects.

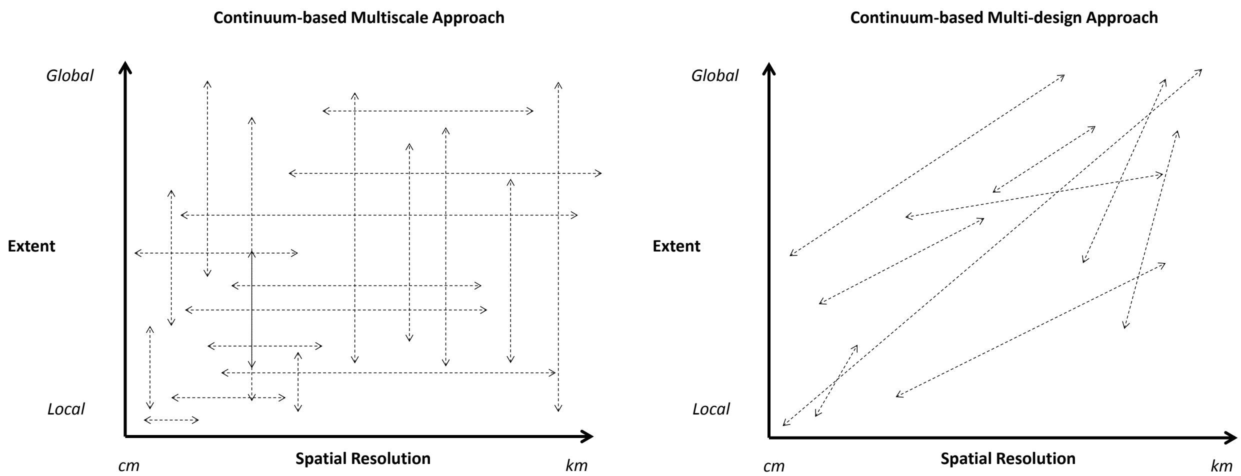

Figure 2.3: (A) Multiscale and (B) multi-design continuum-based approaches. Both extent and resolution vary in a multi-design approach, while only one of these 2 scale characteristics is modified in a multiscale survey; each dotted line illustrates an example of how a single study could be framed.

2.3.5.3 Studying Benthic Habitats at Multiple Scales

Benthic habitat studies at multiple scales were first performed along transects (*e.g.* Schneider *et al.*, 1987; Schneider & Haedrich, 1991). Extending knowledge gained from this type of study to 2-dimensional mapping is challenging in terms of logistics, data volume, and analytic complexity. Recent work has begun to meet these challenges by looking at the differences between local and regional settings, and showing the importance of observing and mapping seafloor habitats at more than one scale (Wilson *et al.*, 2007; Davies *et al.*, 2008; Wedding *et al.*, 2008; Zieger *et al.*, 2009; Tong *et al.*, 2013). In species distribution modelling, combining data from different scales has improved model reliability and performance (Wu & Smeins, 2000; Store & Jokimäki, 2003; Mateo Sánchez *et al.*, 2014).

Benthic habitat studies at multiple spatial scales have generated several insights. For instance, some variables (substrates, food supply) were found to best explain species distribution at relatively fine scales (Davies & Guinotte, 2011; Edinger *et al.*, 2011). Conversely, other variables (*e.g.* productivity) were found to have a stronger influence at relatively coarse scales (Davies *et al.*, 2008). Still other variables (*e.g.* depth) were found to be important at both finer and coarser scales. However, these conclusions are constrained by the observational and analysis scales used in these studies, which did not cover a broad continuum of spatial scales. For instance, fine-scale ocean chemistry could also be found to be locally important if studied within an appropriate range of fine scales.

2.4 Adding Geographic Context by Considering the Spatial Nature of Data

When mapping habitats, it is important to consider the spatial attributes of measurements. Beyond the questions of spatial scale, considering spatial properties of the data is vital in understanding ecological complexity in benthic habitats (Brown *et al.*, 2011a) and in supporting management decisions about these habitats (Katsanevakis *et al.*, 2011; Galparsoro *et al.*, 2014). Spatial heterogeneity (spatial non-stationarity) and spatial autocorrelation (spatial dependence) are properties of most ecogeographical data: spatial heterogeneity refers to the level of variation of a property across space, *i.e.* if an observed variable varies locally or globally (Miller, 2012), while spatial autocorrelation (SAC) is "the correlation of a variable with itself" (Lloyd, 2014, p. 13) and quantifies the observation that spatially closer objects tend to be more similar than spatially distant objects (Tobler, 1970). These 2 properties can strongly affect observed relationships and predictive models (Foody, 2004; Hothorn *et al.*, 2011; Hijmans, 2012). Finley (2011) compared predictive statistical models that account for spatial heterogeneity and SAC to regular regression models. This comparison showed that models accounting for both properties performed better than non-spatial models or models accounting for SAC alone. Spatial heterogeneity and SAC are also strongly scale-dependent, varying with both resolution and extent (Meentemeyer, 1989; Legendre, 1993; Dutilleul & Legendre, 1993; Lloyd, 2014). Zhang *et al.* (2014, p. 67) stated that "the interactions between spatial dependence and spatial heterogeneity have been shown previously to alter local definitions of scales." Consequently, standard statistics based on the assumptions of independent and identically distributed (IID) variables, while used in many ecological studies, should not be used if they violate these statistical assumptions (Meentemeyer & Box, 1987; Marceau & Hay, 1999; Brennan *et al.*, 2002; Goodchild, 2004; Beale *et al.*, 2010; Windle *et al.*, 2010). Demšar *et al.* (2013) identify the need to promote "spatially aware" statistical methods, and other authors advocate for "the need to move beyond potentially misleading global regression models which can obscure the space-varying nature of relationships between the outcome variable of interest and covariates" (Finley, 2011, p. 149, based on Foody, 2004). Nevertheless, standard IID statistics are still often used (Austin, 2002; Brennan *et al.*, 2002; Fortin *et al.*, 2005).

2.4.1 Spatial Autocorrelation

While rarely considered in marine habitat mapping studies, SAC is a well-known scale-dependent phenomenon in geography and ecology (Legendre & Fortin, 1989;

Legendre, 1993) that should always be assessed before conducting spatial analysis (Dormann *et al.*, 2007; Moudrý & Šímová, 2012; Laffan *et al.*, 2014; Vierod *et al.*, 2014). SAC can be present even when samples are collected using random sampling schemes (Lecours *et al.*, 2013). Samples presenting SAC are not statistically independent, which can influence standard statistical tests (Moran, 1948; Cressie, 1993), introducing redundancy into the analyses, and often can induce cross-scale correlation among the variables (Battin & Lawler, 2006; Kristan, 2006; Rigby *et al.*, 2010). In species distribution models, SAC of environmental covariates can increase the influence of positional uncertainty in species occurrence data (Moudrý & Šímová, 2012), and artificially increase the performance of models (Veloz, 2009; Hijmans, 2012). Segurado *et al.* (2006) demonstrated that SAC inflated the significance estimates of their species distribution models up to 90-fold.

Several tools can be used to measure and handle SAC (see Zhang *et al.*, 2014). The spatial scale at which SAC occurs needs to be identified to deal with SAC effects. Techniques to identify this scale include spectral analysis (*e.g.* Legendre & Demers, 1984), study of the 3-term local quadrat variance metric (*e.g.* Boyce, 2006), and neutral landscape models (*e.g.* With & King, 1997). SAC has a strong potential to help resolve ecological complexities. Legendre (1993) indicates that it should be considered as one of the structural attributes of the landscape that needs to be understood, and not considered only as nuisance. SAC can be an indicator of spatial variability, and can be used to study patchiness as a function of scale across a landscape or seascape (Sokal & Oden, 1978; Sokal, 1979). The exploration of the structure of SAC in occurrence data can help improve predictive models by presenting information on the dispersal potential of the organisms (Smith, 1994; Araújo & Williams, 2000; Keitt *et al.*, 2002), even more when this is done at multiple scales (Václavík *et al.*, 2012). De Oliveira *et al.* (2014) showed that accounting for SAC in environmental variables prevents over-fitting of models whilst improving accuracy. Despite its importance, Dormann (2007) found that less than 20% of species distribution modelling studies accounted for SAC, and most of them focused on trying to remove it, something that cannot be done (Mizon, 1995, see discussion in Fortin & Dale, 2009). According to Vierod *et al.* (2014), none of the species distribution modelling work performed in the deep sea has explicitly considered SAC (*e.g.* Ross & Howell, 2013). Failure to account for SAC can result in the selection of predictors with the greatest level of autocorrelation (Lennon 2000), the selection of broad-scale predictors over finer-scale ones (Diniz-Filho *et al.*, 2003), and selection of models with too many predictors (Hoeting *et al.*, 2006; Latimer *et al.*, 2006). Beale *et al.* (2007) showed that precision tends to rapidly decrease when SAC increases when using standard non-spatial models. Dormann *et al.* (2007), Miller *et al.* (2007), Veloz (2009) and Miller (2012) review SAC in a context of species distribution modelling.

The spatial structure of species distribution is influenced by the autocorrelation among environmental variables (exogenous autocorrelation) and by the autocorrelation among biological variables (endogenous autocorrelation) (Miller, 2012). Failing to consider SAC in the analysis and interpretation of data can lead to misinterpretation and incorrect conclusions about spatial structure and the variables that influence it (Lennon, 2000; Keitt *et al.*, 2002; Segurado *et al.*, 2006). Incorporating SAC into modelling effort allows additional knowledge to be gained from the analysis, allowing for habitat characterizations that are closer to reality (Hothorn *et al.*, 2011; De Oliveira *et al.*, 2014). Physical and biological processes can be used to generate testable hypotheses concerning change in SAC in benthic habitat structure and benthic fauna (Schneider & Haedrich, 1991). Developments in geostatistical theory now allow prediction of changes in SAC and adaptation of standard statistics for use with spatial data, without violating any IID assumptions. These adaptations often result in better performance than standard statistics when compared on the same datasets (*e.g.* Brunsdon *et al.*, 1996; Fotheringham *et al.*, 2002; Jombart *et al.*, 2008).

2.4.2 Using Spatial Statistics to Account for Spatial Heterogeneity

The interpretation of species-environment relationships and predictive models can be influenced by the choice of statistics used to perform the analysis (Dormann *et al.*, 2007; Finley, 2011). Most habitat mapping studies have relied on simple statistics to test species-environment relationships (*e.g.* Pearson's correlation) before the application of multivariate statistics (Brown *et al.*, 2011a). Multivariate techniques such as linear discriminant function (*e.g.* McLeod *et al.*, 2007) or principal components analysis (PCA) (*e.g.* Anderson *et al.*, 2011) allow the inclusion of correlation structure in models and are now more common. With these techniques, the independent variables correspond to the values of environmental covariates at certain point locations corresponding to species occurrences. Often, geographical effects are not considered when statistical analyses are performed on these points and their associated environmental values, and results are represented non-spatially in tables (*e.g.* Antunes *et al.*, 2008; Preston, 2009). Other works use raster-based statistical analyses where each pixel is considered a sample point (*e.g.* Maina *et al.*, 2008; Verfaillie *et al.*, 2009). However, despite the fact that pixels are georeferenced, the geographical effects are not taken into consideration in the calculations, but only in the representation of the output maps (Demšar *et al.*, 2013).

Current developments in statistical sciences extend traditional methods to include the spatial component. Locally and geographically weighted statistical methods that account for spatial heterogeneity are becoming increasingly common (Lloyd 2014), particularly in social sciences (*e.g.* Lloyd, 2010a, b), helped by the development of tools for implementation (*e.g.* the R package GWmodel) (Lu *et al.*, 2014b). Rare examples of their use in marine ecology come from Windle *et al.* (2010, 2012), who demonstrated that the use of Geographically Weighted Regression (GWR) could improve the detection of interspecies relationships (cod and invertebrates) and species-environment relationships, with identification of the scale(s) at which these relationships were relatively strong. In addition to these methods that consider spatial effects, future developments in geostatistics will likely improve capacity to detect patterns of variations across spatial scales (see Atkinson & Tate, 2000; Zhang *et al.*, 2014). For instance, Pardo-Igúzquiza & Dowd (2002) introduced a geostatistical technique (namely a factorial cokriging) to identify how cross-correlation between variables varies with scale.

2.5 Future Directions – Integrating Spatial Concepts in Habitat Mapping

2.5.1 Past, Current and Future Trends in Benthic Habitat Mapping

Studies of species-environment relationships often use a limited number of surrogates at either one scale or at multiple arbitrarily chosen scales (Lechner *et al.*, 2012b). Studies of habitats at multiple scales tend to be multi-designed rather than multiscale. While such studies can contribute to our knowledge of marine ecosystems, they may produce results that are not comparable among scales and studies (*e.g.* because of MAUP) (Mayor *et al.*, 2009; Lechner *et al.*, 2012b). Also, the lower and upper limits of "useful" scales at which to study benthic habitats are unknown: while there is a belief in the benthic habitat mapping community that finer-scale data will improve the understanding of benthic ecosystems, such an assumption is not necessarily correct as fine-scale data do not always reveal associations present at coarser spatial scales (Schneider *et al.*, 1987).

As highlighted in this review, a multiscale perspective needs to be adopted in benthic habitat mapping (Nash *et al.*, 2014), using objective and non-arbitrary methods to select observational and analytic scales (Wiens, 1989; Lechner *et al.*, 2012b). Data collection should be planned to characterize as much as possible of the physical, chemical, and biological environment, with emphasis on those variables relevant to the purpose of the survey. Over the past 10 years, bathymetric LiDAR, acoustic remote sensing, and underwater vehicles have revolutionized how the seafloor environment can be mapped and studied. There are, however, some fundamental technical limitations, such as the footprint size (the size of the area of seafloor surveyed at a particular moment), that will dictate the scale at which the data are available (Kenny *et al.*, 2003; Diaz *et al.*, 2004).

These considerations should be integrated in the scale assessment of given studies even though they are often neglected or ignored once the data enter the realm of geographic information systems (GIS) for analysis and map production (Brown *et al.*, 2011a).

The importance of identifying changes in spatial pattern on a continuum has long been recognized in physical and biological oceanography (*e.g.* Stommel, 1963; Steele, 1978), and in ecology (Wiens, 1989; Levin, 1992; Brennan *et al.*, 2002). In terrestrial ecology, Mayor *et al.* (2009) recommended using a spatial, continuum-based approach to identify the ranges of scales over which organisms associate with their habitat. Advances in spatial statistics (Cressie, 1993) put continuum-based analysis on a sound mathematical basis. A continuum-based approach, using coarse-graining, has been applied to benthic transect data (Schneider *et al.*, 1987), but has yet to be implemented in 2-dimensional benthic habitat mapping due to the lack of available data covering a substantial range of scales. The computational power needed to analyze and store such data (Vierod *et al.*, 2014) further limits the application of such approach to quantify the strength of association with habitat as a function of scale. Hierarchical data models could eventually be used to map habitats at multiple scales and implemented in GIS environments so that one habitat map can be represented in different ways depending on the intended application or question.

The idea of identifying the “best” or “right” scale to study habitat association and habitat selection has proven elusive. A logical candidate for “best” scale is that at which variance in either density or a habitat variable reaches a maximum. However, spectral analyses show no peaks in variance in physical and biological variables in either the pelagic (Horne & Schneider, 1997) or benthic realms (Schneider *et al.*, 1987). Similarly, peaks in the scale at which organisms are associated with habitat are another logical candidate for "best scale". Peaks in covariance were not found for any epibenthic species in a study on the outer continental shelf of Newfoundland (Schneider *et al.*, 1987) and have yet to be reported in subsequent studies. Competing with the idea of "right" scale, Wiens (1989) introduced the concept of scale domains, which he defined as ranges of continuous scales for which there is no change (or a constant change) in the observed pattern or process and separated by "chaotic" transitions (see Figure 4 in Wiens, 1989). He argued that these domains were key to understanding ecological systems and could define the limits of generalizations (*i.e.* the bounds within which it is possible to scale-up or scale-down). Scale domains, as defined graphically by Wiens (1989), have not yet been confirmed by empirical data. Graphic representations of patterns and processes as a function of resolution scale in a benthic context (Schneider *et al.*, 1987; Schneider & Haedrich, 1991) show a variety of patterns, with no evidence of transitions as depicted by Wiens (1989). The term "scale domain" has however been used by other authors to characterize levels in hierarchical theory and modelling frameworks (*e.g.* Wu, 1999; Pearson & Dawson, 2003; Muñoz-Reinoso, 2009). The concept of "scale-dependent pattern and process" is arguably of more utility in habitat mapping than attempts to "detect the right scale" or identify "scale domains". Scaling manoeuvres (Schneider, 2001b), in either the distance domain (*e.g.* lagging) or frequency domain (*e.g.* coarse-graining) are available for characterizing the association of benthic biota with habitat, and quantifying habitat association as a function of scale. Figure 2.4 illustrates how such techniques can be implemented by quantifying the association between a species and several characteristics of its environment at multiple scales.

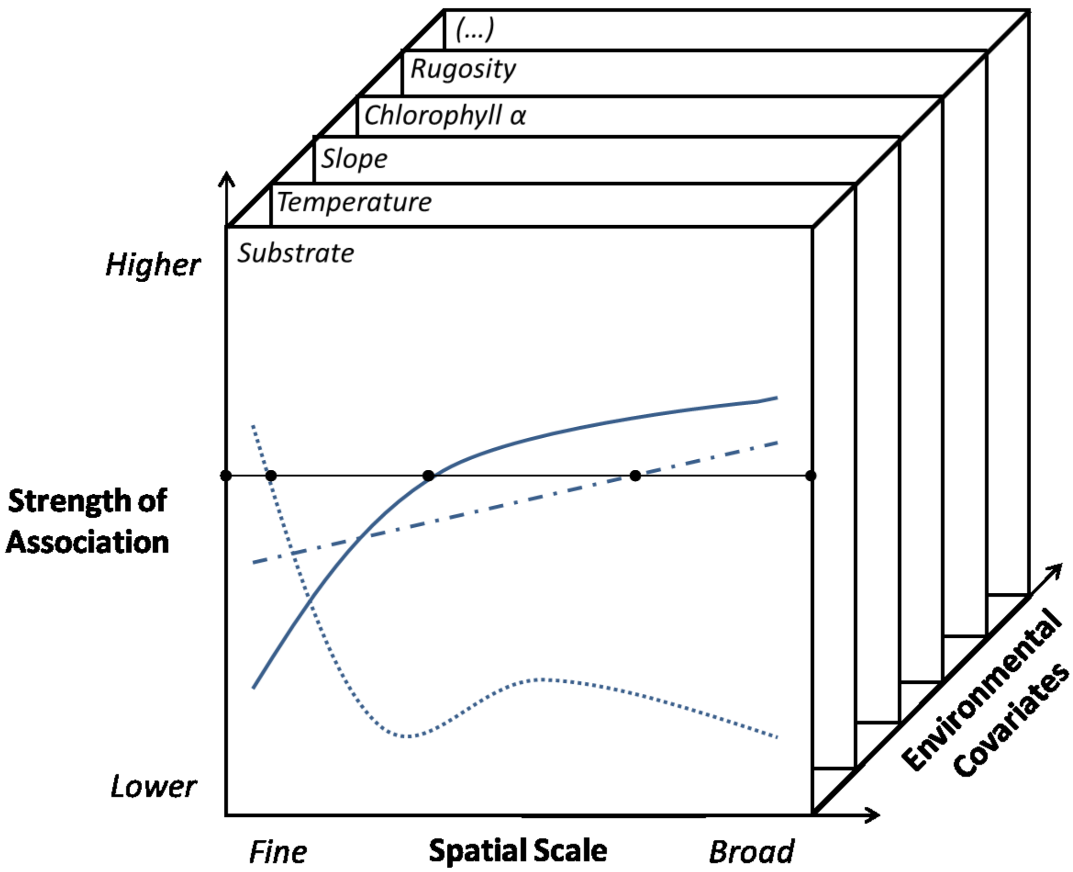

Figure 2.4: Conceptual representation of the implementation of a continuum-based multiscale approach to explore scale-dependency of species-environment relationships. By sampling several environmental characteristics (z axis) at multiple spatial scales (x axis), it is possible to quantify the strength of association (y axis) between a species and its habitat as a function of scale (blue curves). The black horizontal line represents a given significance threshold. Note that if a coefficient of correlation was to be used to measure significance, there would be 2 significance thresholds: one for strongly positive correlations and one for strongly negative correlations. Curves are hypothetical and inspired by results from Horne & Schneider (1997) (pelagic species), and Schneider *et al.* (1987) and Kendall *et al.* (2011) (benthic and epibenthic species).

Because benthic habitats are being altered or destroyed at a faster pace than we discover and understand them (Ramirez-Llodra *et al.*, 2011), it becomes urgent to make effective use of resources to map benthic habitats. Identifying useful surrogates will become possible as this field shifts from studies at multiple scales that only tell part of the story, to continuum-based multiscale approaches. When studying species-habitat relationships, it is as important to identify the scales at which environmental factors drive species distributions as to identify the relevant environmental factors (Williams *et al.*, 2012; Mateo Sánchez *et al.*, 2014). Sampling should be planned with a full combination of efforts to survey as many characteristics of the environment as possible and at as many scales as possible. Because all species cannot be studied, species assemblages (*e.g.* Howell *et al.*, 2010) or those species that interact strongly with other species (*e.g.* Buhl-Mortensen *et al.*, 2010; Baker *et al.*, 2012), or that modify/create habitats (engineer species) (*e.g.* Howell *et al.*, 2011), or that serve as umbrella species in a conservation context (Larsen & Rahbek, 2005), should be targeted. Techniques such as bivariate scaling (*e.g.* Mateo Sánchez *et al.*, 2014), spectral analysis (*e.g.* Schneider *et al.*, 1987), or scalewise variance (*e.g.* Detto & Muller-Landau, 2013) could then be used to identify the strength of association of a particular species with habitat variables at multiple scales. Muotka *et al.* (1998) demonstrated how geostatistics can be efficiently used to characterize the spatial associations between lotic fish and macroinvertebrate species and their habitat at multiple scales while avoiding MAUP effects. Geostatistics and spatial analysis also include methods to deal with the concept of fuzzy boundaries, which are characteristic of many habitats (Dale & Fortin, 2014). The problem is rarely acknowledged in the practice of habitat mapping, which typically imposes sharp boundary delineation.

Benthic ecosystem research often lacks sufficiently extensive datasets at several scales, particularly in the deep sea where sampling is limited and sporadic (Benn *et al.*, 2010). Current data acquisition techniques often cannot capture biological and environmental patterns and processes at a fine resolution over extensive areas (Wilson *et al.*, 2007; Huang *et al.*, 2012), resulting in the need to identify tools to fill the gap. Ongoing improvements in bathymetric LiDAR and multibeam echosounders data analysis are generating some of the most extensive and accurate seafloor data available (Costa *et al.*, 2009; Schimel *et al.*, 2010). Development of remotely operated vehicles (ROV) and autonomous underwater vehicles (AUV) has increased both the range and extent of seafloor data (Wright, 1999; Heyman & Wright, 2011) at ever decreasing costs per megabyte. ROV-and AUV-mounted sensors have the capacity to sample the chemical, physical, and biological environment at fine spatial scales. These new technologies allow biological, geological, chemical, and physical observations to be situated in an accurate multiscale and geospatial context, allowing identification of surrogate variables (*e.g.* Costa *et al.*, 2014; see Van Rein *et al.*, 2009). Metadata are essential to improve the use of geospatial data and to build what Devillers *et al.* (2007) call a “quality-aware” community: all collected datasets will need to be associated with complete metadata files reporting scale information, error and uncertainty quantification, the species or environmental variables that were targeted, the other species that were observed, and other information relevant to further use of the datasets.

Technological developments will continue to drive progress in benthic habitat mapping. Of interest are developments in automatic species detection and analysis on video data (*e.g.* Purser *et al.*, 2009; Lüdtke *et al.*, 2012; Seiler *et al.*, 2012; Tanner *et al.*, 2015), in methods for generating photomosaics of the seafloor for accurate georeferencing (*e.g.* Prados *et al.*, 2012; Kwasnitschka *et al.*, 2013; Marsh *et al.*, 2013), in spatial statistics (*e.g.* Harris *et al.*, 2011; Lu *et al.*, 2014a,b), in computationally fast algorithms capable of processing high-dimensional datasets (*e.g.* Mumby, 2006; Filzmoser *et al.*, 2008; Bermejo *et al.*, 2011; Oyana *et al.*, 2012), in species distribution models that consider spatial autocorrelation, non-stationarity, and scale (*e.g.* Miller & Hanham, 2011; Robinson *et al.*, 2011; Beale *et al.*, 2014; Vierod *et al.*, 2014), and in geomorphometry (Gessler *et al.*, 2009; Guth, 2013). Analyses at multiple scales with many datasets require substantial computational time and effort, and tools that can iterate analyses at multiple scales will become necessary. Surveying multiple characteristics of an area at multiple scales generates immense amounts of data. As in satellite remote sensing (Turner *et al.*, 2015), adequate software and institutional arrangements are needed to realize the potential for these data to be used for purposes other than habitat mapping, to become a valued repository (Borja, 2014), and to notify stakeholders of their existence. This resource-sharing philosophy is important to implement (Turner *et al.*, 2015) if marine scientists are to make effective use of the data and to understand benthic ecosystems before they become substantially altered (Vierod *et al.*, 2014). In some cases, data have been stored for decades waiting for the development of appropriate analytical tools (Knobles *et al.*, 2008). Conversely some researchers might have developed tools applicable to more than their own application, but lack the platform to share these tools with the relevant communities.

2.5.2 Improving Standards for Defining Benthic Habitats

In the previous sections we review the ways that scale and the spatial nature of data influence the way we perceive, measure, analyze, and interpret the environment and species-habitat relationships in benthic habitats. We found that information on scale is not always clearly reported in published works, that a quantitative understanding of habitats and scale is needed, and that results depend on the geographic context of habitat mapping. We thus propose a better standard for defining benthic habitat, one that builds upon the habitat definition of Harris & Baker (2012a). With these standards benthic habitats can be defined as "areas of seabed that are (geo)statistically significantly different from their surroundings in terms of physical, chemical and biological characteristics, when observed at particular spatial and temporal scales". This revised definition of benthic habitat addresses some of the critiques discussed in the previous sections. First, it addresses the growing realization that habitats must be quantitatively delineated and that what constitutes the description of a habitat is dictated by the scale of the techniques employed (Diaz *et al.*, 2004). Then, it addresses the argument for considering the chemical environment in the characterization of benthic habitats (Kostylev *et al.*, 2001; Brown *et al.*, 2011a). Finally, it addresses the case made by Cogan & Noji (2007) that habitats be placed in context with the appropriate spatial, temporal and thematic scales when being mapped. The reference to geostatistics encompasses the consideration of the spatial nature of data and the concepts of fuzzy boundary delineation, while the biological characteristics relate to thematic scale and allow the study of species assemblages as much as individual species, and the mention of spatial and temporal scales makes habitats explicit about scale. Being explicit about temporal scale is important when studying migratory species that do not inhabit the same space through time.

2.5.3 Recommendations

The previous section on trends in benthic habitat mapping highlighted some of the main issues currently encountered in benthic habitat mapping, proposed some solutions and gave an insight on what the future developments might bring to the field. Based on this discussion, it is possible to identify 3 elements in the habitat mapping process that can be improved: project planning and data collection, data analysis and interpretation, and communication/dissemination of research results and data. Project planning and data collection can be improved from a biological, environmental and/or approach point of view. For the biology, we recommend focusing on the study of ecosystem engineer or umbrella species that would indirectly allow collecting data on other species. For the environment, we recommend sampling as many environmental variables as possible to aim for a comprehensive understanding of the environment and its dynamics. In terms of approach, we recommend adopting continuum-based multiscale methods, which involves sampling the environment over an extensive range of spatial scales. To improve data analysis and interpretation, we recommend using spatial statistical analyses that consider spatial heterogeneity and autocorrelation of data, rather than standard statistics based on the assumptions of IID, to establish results on a sound inferential basis. We also suggest always quantifying errors and spatial uncertainty. Finally, to improve communication and dissemination of research and data, we recommend making available metadata in which the results from the quantification of errors would be reported together with the spatial scales at which the data was collected (observation scale), at which the research was intended to be conducted (ecological scale), and at which the analysis was performed (analysis scale). In terms of dissemination, we suggest developing and automating tools (*e.g.* GIS, statistical, ecological) for processing or analyzing data and make them available, together with datasets and complete metadata, to maximize research and application potential.

2.6 Conclusions

Organisms inhabit a space that suits their needs. Understanding what controls benthic species distribution requires understanding the physico-chemical properties and dynamics within the water column, and at the seafloor interface (Clark *et al.*, 2012; Vierod *et al.*, 2014). The structure and spatial arrangement of habitats constrain, and can potentially become predictors of, species distribution, abundance, and richness. The cost and difficulties associated with sampling the marine environment highlight the need for better predictions of species distributions and improvement in sampling strategies. This will become possible with a better understanding of ecological patterns and processes as a function of scale, and should bring an overall improvement to benthic research efficiency. Using appropriate surrogates at appropriate scales is likely to be more effective than the use of opportunistic or arbitrarily chosen variables and scales. Generating habitat maps is a complex process that requires multidisciplinary efforts (Heyman & Wright, 2011). Technological advances will help marine scientists address the current challenges of their field and develop new approaches to understand and so protect benthic habitat structure and function (Ramirez-Llodra *et al.*, 2011). Geospatial data and techniques from geomatics and geostatistics show potential to tackle core issues in spatial ecology (Skidmore *et al.*, 2011; Laffan *et al.*, 2012) and in the marine sciences (Wright & Goodchild, 1997; Heyman & Wright 2011).

The need for fundamental ecological and conservation theory, including explicit treatment of spatial scale has been noted repeatedly (*e.g.* Guisan & Thuiller, 2005; Levin & Dayton, 2009). Spatial scale is central to understanding habitat use, to selecting a sampling method, and to statistical analysis. Despite being recognized as a central issue, scales are often arbitrarily chosen, and studies regularly fail to report the scale(s) investigated and how the results depend on spatial scale. As stated by Dungan *et al.* (2002, p. 632): "If ecologists are explicit about all of the components and dimensions of scale so that the spatial characteristics of the quantities measured can be correctly interpreted, there will be new opportunities to gain experience and improve understanding of the effects of observations and analysis scale changes." Evidence-based scaling functions, which link pattern to process as a function of scale, are needed to identify reliable surrogates of species distribution, to scale-up and scale-down relevant information, and for improved quantitative understanding of benthic habitats.

Based on this review, we provide 8 recommendations that could lead to more efficient practices in benthic habitat mapping: (1) umbrella species' habitats should be prioritized for mapping and prediction; (2) sampling should be conducted to obtain data covering an extensive range of spatial scales and as many environmental variables as possible; (3) continuum-based habitat characterization approaches should be adopted; (4) statistical methods that consider the spatial nature of data should systematically be used; (5) errors and spatial uncertainty should be quantified at every step of habitat mapping (*i.e.* data collection, surrogacy testing, predictive modelling); (6) existing tools should be automated and new tools (*e.g.* GIS, statistical, ecological) should be developed for processing data and defining surrogates of species distribution and habitat at multiple scales; (7) data, complete metadata, and tools should be made available to maximize research and applications potential; and (8) the spatial extent and resolution (scale) at which the research was intended to be conducted, at which the data was collected, and at which predictive or monitoring aims were directed should always be clearly reported. We further recommend that benthic habitat be defined to the following standards: (1) explicit statement of observational scale (*i.e.* spatial resolution and extent); (2) inclusion of chemical variables along with physical and biological variables; and (3) placement in context with the appropriate spatial, temporal and thematic scales when being mapped.

2.7 Literature Cited

Adams, D.K., & Flieri, G.R. (2010) Modeled interactions of mesoscale eddies with the East Pacific Rise: implications for larval dispersal. *Deep-Sea Research Part I*, 57:1163-1176.

Addicott, J.F., Aho, J.M., Antolin, M.F., Padilla, M.F., Richardson, J.S., & Soluk, D.A. (1987) Ecological neighborhoods: scaling environmental patterns. *Oikos*, 49:340-346.

Anderson, T.J., & Yoklavich, M.M. (2007) Multiscale habitat associations of deepwater demersal fishes off central California. *Fishery Bulletin*, 105:168-179.

Anderson, D.P., Turner, M.G., Forester, J.D., Zhu, J., Boyce, M.S., Beyer, H., & Stowell, L. (2005) Scale-dependent summer resource selection by reintroduced elk in Wisconsin, USA. *Journal of Wildlife Management*, 69:298-310.

Anderson, J.T., Holliday, D.V., Kloser, R., Reid, D.G., & Simard Y. (2008) Acoustic seabed classification: current practice and future directions. *ICES Journal of Marine Science*, 65:1004-1011.

Anderson, T.J., Nichol, S.L., Syms, C., Przeslawski, R., & Harris, P.T. (2011) Deep-sea bio-physical variables as surrogates for biological assemblages, and example from the Lord Howe Rise. *Deep-Sea Research Part II*, 58:979-991.

Antunes, S.C., Pereira, E., Sousa, J.P., Santos, M.C., & Gonçales, F. (2008) Spatial and temporal distribution of litter arthropods in different vegetation covers of Porto Santo Island. *European Journal of Soil Biology*, 44:45-56.

Apps, C.D., McLellan, B.N., Kinley, T., & Flaa, J.P. (2001) Scale-dependent habitat selection by mountain caribou, Columbia Mountains, British Columbia. *Journal of Wildlife Management*, 65: 65-77.

Araújo, M.B., & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33:1677-1688.

Araújo, M.B., & Williams, P.H. (2000) Selecting areas for species persistence using occurrence data. *Biological Conservation*, 96: 331-345.

Atkinson, P.M., & Tate, N.J. (2000) Spatial scale problems and geostatistical solutions: a review. *The Professional Geographer*, 52:607-623.

Austin, M.P. (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, 157:101-118.

Austin, M.P., & Van Niel, K.P. (2011) Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, 38:1-8.

Baillon, S., Hamel, J.F., Wareham, V.E., & Mercier A. (2012) Deep cold-water corals as nurseries for fish larvae. *Frontiers in Ecology and the Environment*, 10:351-356.

Baker, K.D., Haedrich, R.L., Fifield, D.A., & Gilkinson, K.D. (2012) Grenadier abundance examined at varying spatial scales in deep waters off Newfoundland, Canada, with special focus on the influence of corals. *Journal of Ichthyology*, 52:678-689.

Barry, S., & Elith, J. (2006) Error and uncertainty in habitat models. *Journal of Applied Ecology*, 43:413-423.

Battin, J., & Lawler, J.J. (2006) Cross-scale correlations and the design and analysis of avian habitat selection studies. *Condor*, 108:59-70.

Beale, C.M., & Lennon, J.J. (2012) Incorporating uncertainty in predictive species distribution modelling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367:247-258.

Beale, C.M., Lennon, J.J., Elston, D.A., Brewer, M.J., & Yearsley, J.M. (2007) Red herrings remain in geographical ecology: a reply to Hawkin *et al.* (2007). *Ecography*, 30:845-847.

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J., & Elston D.A. (2010) Regression analysis of spatial data. *Ecology Letters*, 13: 246-264.

Beale, C.M., Brewer, M.J., & Lennon, J.J. (2014) A new statistical framework for the quantification of covariate associations with species distributions. *Methods in Ecology and Evolution*, 5: 421-432.

Benaka, L. (1999) Fish habitat: essential fish habitat and rehabilitation. *American Fisheries Society*, Bethesda, MD.

Benn, A.R., Weaver, P.P., Billett, D.S.M., Van den Hove, S., Murdock, A.P., Doneghan, G.B., & Le Bas, T. (2010) Human activities on the deep seafloor in the North East Atlantic: an assessment of spatial extent. *PLoS ONE* 5:e12730.

Bermejo, P., Gámez, J.A., Puerta, J.M. (2011) A GRASP algorithm for fast hybrid (filter-wrapper) feature subset selection in high-dimensional datasets. *Pattern Recognition Letters*, 32: 701-711.

Borja, A. (2014) Grand challenges in marine ecosystems ecology. *Frontiers in Marine Science*, 1:1-6.

Boström, C., Pittman, S.J., Simenstad, C., & Kneib, R.T. (2011) Seascape ecology of coastal biogenic habitats: advances, gaps, and challenges. *Marine Ecology Progress Series*, 427:191-217.

Boyce, M.S. (2006) Scale for resource selection functions. *Diversity and Distribution*, 12:269-276.

Braunisch, V., & Suchant, R. (2010) Predicting species distributions based on incomplete survey data: the trade-off between precision and scale. *Ecography*, 33:826-840.

Brennan, J.M., Bender, D.J., Contreras, T.A., & Fahrig, L. (2002) Focal patch landscape studies for wildlife management: optimizing sampling effort across scales. In: Liu, J., & Taylor, W.W. (Eds) *Integrating landscape ecology into natural resource management*. Cambridge University Press, Cambridge, p. 68-91.

Brock, J.C., Wright, C.W., Clayton, T.D., & Nayegandhi, A. (2004) LIDAR optical rugosity of coral reefs in Biscayne National Park, Florida. *Coral Reefs*, 23:48-59.

Brown, C.J., Todd, B.J., Kostylev, V.E., & Pickrill, R.A. (2011b) Image-based classification of multibeam sonar backscatter data for objective surficial sediment mapping of Georges Bank, Canada. *Continental Shelf Research*, 31:S110-S119.

Brown, C.J., Smith, S.J., Lawton, P., & Anderson, J.T. (2011a) Benthic habitat mapping: a review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuarine, Coastal and Shelf Science*, 92:502-520.

Brown, C.J., Sameoto, J.A., & Smith, S.J. (2012) Multiple methods, maps, and management applications: purpose made seafloor maps in support of ocean management. *Journal of Sea Research*, 72:1-13.

Brunsdon, C., Fotheringham, A.S., & Charlton, M.E. (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis*, 35:58-82.

Bryan, T.L., & Metaxas, A. (2007) Predicting suitable habitat for deep-water gorgonian corals on the Atlantic and Pacific Continental Margins of North America. *Marine Ecology Progress Series*, 330:113-126.

Buhl-Mortensen, L., Vanreusel, A., Gooday, A.J., Levin, L.A., Priede, I.G., Buhl-Mortensen, P., Gheerardyn, H., King, N.J., & Raes, M. (2010) Biological structures as a source of habitat heterogeneity and biodiversity on the deep ocean margins. *Marine Ecology*, 31:21-50.

Carr, M.H., Neigel, J.E., Estes, J.A., Andelman, S.J., Warner, R.R., & Largier, J.L. (2003) Comparing marine and terrestrial ecosystems: implications for the design of coastal marine reserves. *Ecological Applications*, 13:90-107.

Clark, W.C. (1985) Scales of climatic impacts. *Climatic Change*, 7: 5-27.

Clark, M.R., Schlacher, T.A., Rowden, A.A., Stocks, K.I., & Consalvey, M. (2012) Science priorities for seamounts: research links to conservation and management. *PLoS ONE*, 7:e29232.

Clarke, A., & Lidgard, S. (2000) Spatial patterns of diversity in the sea: bryozoan species richness in the North Atlantic. *Journal of Animal Ecology*, 69:799-814.

Cogan, C.B., & Noji, T.T. (2007) Marine classification, mapping, and biodiversity analysis. In: Todd, B.J., & Greene, H.G. (Eds) Mapping the seafloor for habitat characterization. *Geological Association of Canada Special Paper*, 47:129-139.

Coggan, R., Populis, J., White, J., Sheehan, K., Fitzpatrick, F., & Piel, S. (2007) Review of standards and protocols for seabed habitat mapping, 2nd edition. MESH (Mapping European Seabed Habitats). www.emodnet-seabedhabitats.eu/default.aspx?page=1442.

Conroy, M.J., & Noon, B.R. (1996) Mapping of species richness for conservation of biological diversity: conceptual and methodological issues. *Ecological Applications*, 6:763-773.

Cook, R., Fariñas-Franco, J.M., Gell, F.R., Holt, R.H.F., Holt, T., Lindenbaum, C., Porter, J.S., Seed, R., Skates, L.R., Stringell, T.B., & Sanderson, W.G. (2013) The substantial first impact of bottom fishing on rare biodiversity hotspots: a dilemma for evidence-based conservation. *PLoS ONE*, 8:e69904.

Copeland, A., Edinger, E., Bell, T., LeBlanc, P., Wroblewski, J., & Devillers, R. (2012) Geomorphic features and benthic habitats of a sub-Arctic fjord: Gilbert Bay, Southern Labrador, Canada. In: Harris, P.T., & Baker, E.K. (Eds) *Seafloor geomorphology as benthic habitats: GeoHab atlas of seafloor geomorphic features and benthic habitats*. Elsevier, Amsterdam, p. 309-327.

Costa, B.M., Battista, T.A., & Pittman, S.J. (2009) Comparative evaluation of airborne LiDAR and ship-based multibeam SoNAR bathymetry and intensity for mapping coral reef ecosystems. *Remote Sensing of Environment*, 113:1082-1100.

Costa, B., Taylor, J.C., Kracker, L., Battista, T., & Pittman, S. (2014) Mapping reef fish and the seascape: using acoustics and spatial modeling to guide coastal management. *PLoS ONE*, 9:e85555.

Cressie, N. (1993) Statistics for spatial data. Wiley, New York, NY.

Cressie, N., & Wikle, C.K. (2011) Statistics for spatio-temporal data. Wiley, Hoboken, NJ.

Dale, M.R.T., & Fortin, M.J.F. (2014) Spatial analysis: a guide for ecologists (2nd edition). Cambridge University Press, Cambridge.

Davies, A.J., & Guinotte, J.M. (2011) Global habitat suitability for framework-forming cold-water corals. *PLoS ONE*, 6: e18483.

Davies, A.J., Wisshak, M., Orr, J.C., & Murray Roberts, J. (2008) Predicting suitable habitats for the cold-water coral *Lophelia pertusa* (Scleractinia). *Deep-Sea Research Part I*, 55:1048-1062.

De Knegt, H.J., Van Langevelde, F., Coughenour, M.B., Skidmore, A.K., de Boer, W.F., Heitkönig, I.M.A., Know, N.M., Slotow, C., van der Waal, C., & Prins, H.H.T. (2010) Spatial autocorrelation and the scaling of species-environment relationships. *Ecology*, 91:2455-2465.

De Leo, F.C., Vetter, E.W., Smith, C.R., Rowden, A.A., & McGranaghan, M. (2014) Spatial scale-dependent habitat heterogeneity influences submarine canyon macrofaunal abundance and diversity off the Main and Northwest Hawaiian Islands. *Deep-Sea Research Part II*, 104:267-290.

De Oliveira, G., Rangel, T.F., Lima-Ribeiro, M.S., Terribile, L.C., & Diniz-Filho, J.A.F. (2014) Evaluating, partitioning, and mapping the spatial autocorrelation component in ecological niche modeling: a new approach based on environmentally equidistant records. *Ecography*, 37:637-647.

Delcourt, H.R., Delcourt, P.A., & Webb, T. (1983) Dynamic plant ecology: the spectrum of vegetational change in space and time. *Quaternary Science Reviews*, 1:153-175.

Demšar, U., Harris, P., Brunsdon, C., Fotheringham, A.S., & McLoone, S. (2013) Principal component analysis on spatial data: an overview. *Annals of the Association of American Geographers*, 103: 106-128.

Deng, Y., Wilson, J.P., & Bauer, B.O. (2007) DEM resolution dependencies of terrain attributes across a landscape. *International Journal of Geographical Information Science*, 21:187-213.

Detto, M., & Muller-Landau, H.C. (2013) Fitting ecological process models to spatial patterns using scalewise variances and moment equations. *American Naturalist*, 181:E68-E82.

Devillers, R., Bédard, Y., Gervais, M., Pinet, F., Schneider, M., Bejaoui, L., Lévesque, M.-A., Salehi, M., & Zargar, A. (2007) How to improve geospatial data usability: from metadata to quality-aware GIS community. In: Wachowicz, M., & Bodum, L. (Eds) Spatial data usability, an AGILE pre-conference workshop. *Proceedings of the 10th*

AGILE International Conference on Geographic Information Science, Aalborg, May 8. AGILE (Association of Geographic Information Laboratories in Europe), p. 1–8.

Diaz, R.J., Solan, M., & Valente, R.M. (2004) A review of approaches for classifying benthic habitats and evaluating habitat quality. *Journal of Environmental Management*, 73:165-181.

Diniz-Filho, J.A.F., Bini, L.M., & Hawkins, B.A. (2003) Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, 12:53-64.

Dolan, M.F.J., & Lucieer, V.L. (2014) Variation and uncertainty in bathymetric slope calculations using geographic information systems. *Marine Geodesy*, 37:187-219.

Dormann, *CF.* (2007) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, 16:129-138.

Dormann, *CF.*, McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M., & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30:609-628.

Dormann, *CF.*, Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., García Marquéz, J.R., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., & Lautenbach, S. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36:27-46.

Drăgut, L., Eisank, C., Strasser, T., & Blaschke, T. (2009) A comparison of methods to incorporate scale in geomorphometry. In: Purves, R., Gruber, S., Straumann, R., & Hengl, T. (Eds) Geomorphometry 2009 Conference Proceedings, University of Zürich, Switzerland, p. 133-139.

Dungan, J.L., Perry, J.N., Dale, M.R.T., Legendre, P., Citron-Prousty, S., Fortin, M.-J., Jakomulska, A., Miriti, M., & Rosenberg, M.S. (2002) A balanced view of scale in spatial statistical analysis. *Ecography*, 25:626-640.

Dunn, D.C., & Halpin, P.N. (2009) Rugosity-based regional modeling of hard-bottom habitat. *Marine Ecology Progress Series*, 377:1-11.

Dutilleul, P., & Legendre, P. (1993) Spatial heterogeneity against heteroscedasticity: an ecological paradigm versus a statistical concept. *Oikos*, 66:152-171.

Edinger, E.N., Sherwood, O.A., Piper, D.J.W., Wareham, V.E., Baker, K.D., Gilkinson, K.D., & Scott, D.B. (2011) Geological features supporting deep-sea coral habitat in Atlantic Canada. *Continental Shelf Research*, 31:S69-S84.

Elith, J., & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40:677-697.

Engler, R., Guisan, A., & Rechsteiner, L. (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41:263-274.

Etnoyer, P., & Morgan, L.E. (2007) Predictive habitat model for deep gorgonian needs better resolution: Comment on Bryan & Metaxas (2007). *Marine Ecology Progress Series*, 339:311–312.

Evans, I.S. (1972) General geomorphometry, derivatives of altitude, and descriptive statistics. In: Chorley, R.J. (Ed.) *Spatial analysis in geomorphology*. Harper & Row, London, p. 17-90.

Fassnacht, K.S., Cohen, W.B., & Spies, T.A. (2006) Key issues in making and using satellite-based maps in ecology: a primer. *Forest Ecology and Management*, 222:167-181.

Filzmoser, P., Maronna, R., & Werner, M. (2008) Outlier identification in high dimensions. *Computational Statistics and Data Analysis*, 52: 1694-1711.

Finley, A.O. (2011) Comparing spatially-varying coefficients models for analysis of ecological data with non-stationarity and anisotropic residual dependence. *Methods in Ecology and Evolution*, 2:143-154.

Florinsky, I.V., & Kuryakova, G.A. (2000) Determination of grid size for digital terrain modelling in landscape investigations exemplified by soil moisture distribution scale. *International Journal of Geographical Information Science*, 14:815-832.

Foody, G.M. (2004) Spatial nonstationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. *Global Ecology and Biogeography*, 13:315-320.

Fortin, M.J., & Dale, M.R. (2009) Spatial autocorrelation in ecological studies: a legacy of solutions and myths. *Geographical Analysis*, 41:392-397.

Fortin, M.J., Keith, T.H., Maurer, B.A., Taper, M.L., Kaufman, D.M., & Blackburn, T.M. (2005) Species' geographic ranges and distributional limits: pattern analysis and statistical issues. *Oikos*, 108:7-17.

Fotheringham, A.S., Brunsdon, C., & Charlton, M. (2002) Geographically weighted regression: the analysis of spatially varying relationships. Wiley, Chichester.

Franklin, J. (2009) Mapping species distributions: spatial inference and prediction. Cambridge University Press, Cambridge.

Freeman, S.M., & Rogers, S.I. (2003) A new analytical approach to characterisation of macro-epibenthic habitats: linking species to the environment. *Estuarine, Coastal and Shelf Science*, 56:749-764.

Freitas, R., Ricardo, F., Pereira, F., Sampaio, L., Carvalho, S., Gaspar, M., Quintino, V., & Rodrigues, A.M. (2011) Benthic habitat mapping: concerns using a combined approach (acoustic, sediment and biological data). *Estuarine, Coastal and Shelf Science*, 92:598-606.

Gage, J.D., & Tyler, P.A. (1999) Deep-sea biology: a natural history of organisms at the deep-sea floor. Cambridge University Press, Cambridge.

Gallucci, F., Moens, T., & Fonesca, G. (2009) Small-scale patterns of meibenthos in the Arctic deep sea. *Marine Biodiversity*, 39: 9-25.

Galparsoro, I., Borja, A., & Uyarra, M.C. (2014) Mapping ecosystem services provided by benthic habitats in the European North Atlantic Ocean. *Frontiers in Marine Science*, 1:23-36.

Gambi, C., Danovaro, R. (2006) A multiple-scale analysis of metazoan meiofaunal distribution in the deep Mediterranean Sea. *Deep-Sea Research*, 53:1117-1134.

García, D., & Ortiz-Pulido, R. (2004) Patterns of resource tracking by avian frugivores at multiple spatial scales: two case studies on discordance among scales. *Ecography*, 27: 187-196.

Gehlke, C.E., & Biehl, K. (1934) Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, 29:169-170.

Georgian, S.E., Shedd, W., & Cordes, E.E. (2014) High-resolution ecological niche modelling of the cold-water coral *Lophelia pertusa* in the Gulf of Mexico. *Marine Ecology Progress Series*, 506:145-161.

Gessler, P., Pike, R., MacMillan, R.A., Hengl, T., & Reuter, H.I. (2009) The future of geomorphometry. In: Hengl, T., & Reuter, H.I. (Eds) Geomorphometry: concepts, software, applications. (*Developments in soil science 33*). Elsevier Science, Amsterdam, p. 637-652.

Gjerde, K.M. (2006) Ecosystems and biodiversity in deep waters and high seas. UNEP Regional Seas Reports and Studies 178, UNEP/IUCN, Gland.

Goodchild, M.F. (2004) The validity and usefulness of laws in geographic information science and geography. *Annals of the Association of American Geographers*, 94:300-303.

Goodchild, M.F. (2011) Scale in GIS: an overview. *Geomorphology*, 130:5-9.

Grabowski, J.H., Bachman, M., Demarest, C., Eayrs, S., Harris, B.P., Malkoski, V., Packer, D., & Stevenson, D. (2014) Assessing the vulnerability of marine benthos to fishing gear impacts. *Reviews in Fisheries Science and Aquaculture*, 22:142-155.

Gray, J.S. (1966) Selection of sands by *Protodrilus symbioticus* (Giard). *Veroeff Inst Meeresforsch Bremerhav*, 2:106-116.

Greene, H.G., Bizzarro, J.J., O'Connell, V.M., & Brylinsky, C.K. (2007) Construction of digital potential marine benthic habitat maps using a coded classification scheme and its application. In: Todd, B.J., & Greene, H.G. (Eds) Mapping the seafloor for habitat characterization. *Geological Association of Canada Special Paper* 47:141-155.

Greig-Smith, P. (1952) The use of random and contiguous quadrants in the study of the structure and plant communities. *Annals of Botany*, 16:293-316.

Grober-Dunsmore, R., Frazer, T.K., Lindberg, W.J., & Beets, J. (2007) Reef fish and habitat relationships in a Caribbean seascape: the importance of reef context. *Coral Reefs*, 26:201-216.

Guisan, A., & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8: 993-1009.

Guisan, A., Graham, C.H., Elith, J., Huettmann, F., & NCEAS Species Distribution Modelling Group (2007) Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distribution*, 13:332-340.

Gustafson, E. (1998) Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems*, 1:143-156.

Guth, P.L. (2013) The giga revolution in geomorphometry: gigabytes of RAM, gigabyte-sized data sets, and gigabit internet access. In: Guoan, T., Qiming, Z., Mitasova, H., & Hengl, T. (Eds) *Geomorphometry 2013 Conference Proceedings*, Nanjing Normal University, Nanjing, p. K41-K44.

Hagberg, J., Jonzén, N., Lundberg, P., & Ripa, J. (2003) Uncertain biotic and abiotic interactions in benthic communities. *Oikos*, 100:353-361.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C.V., Micheli, F., D'Agrosa, C., Bruno, J.F., Casey, K.S., Ebert, C., Fox, H.E., Fujita, R., Heinemann, D., Lenihan, H.S., Madin, E.M.P., Perry, M.T., Selig, E.R., Spalding, M., Steneck, R., & Watson, R. (2008) A global map of human impact on marine ecosystems. *Science* 319:948-952.

Hanberry, B.B. (2013) Finer grain size increases effects of error and changes influence of environmental predictors on species distribution models. *Ecological Informatics* 15:8-13.

Harris, P.T. (2012) Anthropogenic threats to benthic habitats. In: Harris, P.T., & Baker, E.K. (Eds) *Seafloor geomorphology as benthic habitats: GeoHab atlas of seafloor geomorphic features and benthic habitats*. Elsevier, Amsterdam, p. 39-60.

Harris, P.T., & Baker, E.K. (2012a) Why map benthic habitats? In: Harris, P.T., & Baker, E.K. (Eds) *Seafloor geomorphology as benthic habitats: GeoHab atlas of seafloor geomorphic features and benthic habitats*. Elsevier, Amsterdam, p. 3-22.

Harris, P.T., & Baker, E.K. (Eds) (2012b) *Seafloor geomorphology as benthic habitats: GeoHab atlas of seafloor geomorphic features and benthic habitats*. Elsevier, Amsterdam.

Harris, P., Brunsdon, C., & Charlton, M. (2011) Geographically weighted principal components analysis. *International Journal of Geographical Information Science*, 25:1717-1736.

Harvey, F. (2008) A primer of GIS: fundamental geographic and cartographic concepts. The Guilford Press, New York, NY.

Haury, L.R., McGowan, J.A., & Wiebe, P.H. (1978) Patterns and processes in the time-space scales of plankton distributions. In: Steele, J.H. (Ed.) *Spatial pattern in plankton communities*. Plenum Press, New York, NY, p. 277-327.

Hengl, T. (2006) Finding the right pixel size. *Computers and Geosciences*, 32:1283-1298.

Hengl, T., & MacMillan, R.A. (2009) Geomorphometry – a key to landscape mapping and modelling. In: Hengl, T., & Reuter, H.I. (Eds) Geomorphometry: concepts, software, applications. (*Developments in soil science 33*). Elsevier Science, Amsterdam, p. 433–460.

Henry, L.A., Moreno Navas, J., & Roberts, J.M. (2013) Multi-scale interactions between local hydrography, seabed topography, and community assembly on cold-water coral reefs. *Biogeosciences*, 10:2737-2746.

Heyman, W.D., & Wright, D.J. (2011) Focus: marine geomorphology as a determinant for essential life habitat and marine protected area design. *The Professional Geographer*, 63:429-442.

Heywood, I.D., Cornelius, S., & Carver, S. (2006) *An introduction to geographic information systems, 3rd edition*. Pearson Education Limited, Harlow.

Hijmans, R.J. (2012) Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology*, 93:679-688.

Hill, N.A., Lucieer, V., Barrett, N.S., Anderson, T.J., & Williams, S.B. (2014) Filling the gaps: predicting the distribution of temperate reef biota using high resolution biological and acoustic data. *Estuarine, Coastal and Shelf Science*, 147:137-147.

Hitt, S., Pittman, S.J., & Nemeth, R.S. (2011) Diel movements of fishes linked to benthic seascape structure in a Caribbean coral reef ecosystem. *Marine Ecology Progress Series*, 427: 275-291.

Hobbs, N.T. (2003) Challenges and opportunities in integrating ecological knowledge across scales. *Forest Ecology and Management*, 181:223-238.

Hoeting, J.A., Davis, R.A., Merton, A.A., & Thompson, S.E. (2006) Model selection for geostatistical models. *Ecological Applications*, 16: 87-98.

Horne, J.K., & Schneider, D.C. (1997) Spatial variance of mobile marine organisms: capelin and cod in Newfoundland coastal waters. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, 352: 633-642.

Hothorn, T., Müller, J., Schröder, B., Kneib, T., & Brandl, R. (2011) Decomposing environmental, spatial, and spatiotemporal components of species distributions. *Ecological Monographs*, 81: 329-347.

Howell, K.L. (2010) A benthic classification system to aid in the implementation of marine protected area networks in the deep/high seas of the NE Atlantic. *Biological Conservation*, 143: 1041-1056.

Howell, K.L., Davies, J.S., & Narayanaswamy, B.E. (2010) Identifying deep-sea megafaunal epibenthic assemblages for use in habitat mapping and marine protected area net-work design. *Journal of the Marine Biological Association of the UK*, 90:33-68.

Howell, K.L., Holt, R., Endrino, I.P., & Stewart, H. (2011) When the species is also a habitat: comparing the predictively modelled distribution of *Lophelia pertusa* and the reef habitat it forms. *Biological Conservation*, 144:2656-2665.

Huang, Z., Nichol, S.L., Siwabessy, J.P.W., Daniell, J., & Brooke, B.P. (2012) Predictive modelling of seabed sediment parameters using multibeam acoustic data: a case study on the Carnavon Shelf, Western Australia. *International Journal of Geographical Information Science*, 26:283-307.

Huston, M.A. (2002) Introductory essay: critical issues for improving predictions. In: Scott, J.M., Heglund, P.J., Samson, F., Haufler, J., Morrison, M., Raphael, M., & Wall, B. (Eds) *Predicting species occurrences: issues of accuracy and scale*. Island Press, Washington DC, p. 7-21.

Hutchinson, G.E., & MacArthur, R.H. (1959) A theoretical ecological model of size distributions among species of animals. *American Naturalist*, 93:117-125.

Jelinski, D., & Wu, J. (1996) The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, 11: 129-140.

Jombart, T., Devillard, S., Dufour, A.B., & Pontier, D. (2008) Revealing cryptic patterns in genetic variability by a new multivariate method. *Heredity*, 101:92-103.

Jones, J.B. (1992) Environmental impact of trawling on the seabed: a review. *New Zealand Journal of Marine and Freshwater Research*, 26:59-67.

Kaliontzopoulou, A., Brito, J.C., & Carretero, M.A. (2008) Modelling the partially unknown distribution of wall lizards (*Podarcis*) in North Africa: ecological affinities, potential areas of occurrence, and methodological constraints. *Canadian Journal of Zoology*, 86:992-1001.

Katsanevakis, S., Stelzenmüller, V., South, A., Sorensen, T.K., Jones, P.J.S., Kerr, S., Badalamenti, F., Anagnostou, C., Breen, P., Chust, G., D'Anna, G., Duijn, M., Filatova, T., Fiorentino, F., Hulsman, H., Johnson, K., Karageorgis, A.P., Kröncke, I., Mirto, S., Pipitone, C., Portelli, S., Qiu, W., Reiss, H., Sakellariou, D., Salomidi, M., van Hoof, L., Vassilopoulou, V., Fernández, T.V., Vöge, S., Weber, A., Zenetos, A., & Hofstede, R.T. (2011) Ecoystem-based marine spatial management: review of concepts, policies, tools, and critical issues. *Ocean and Coastal Management*, 54:807-820.

Keitt, T.H., Bjørnstad, O.N., Dixon, P.M., & Citron-Pousty, S. (2002) Accounting for spatial pattern when modeling organism-environment interactions. *Ecography*, 25:616-625.

Kendall, M.S., & Miller, T.J. (2008) The influence of thematic and spatial resolution on maps of a coral reef ecosystem. *Marine Geodesy*, 31:75-102.

Kendall, M.S., & Miller, T.J. (2010) Relationships among map resolution, fish assemblages, and habitat variables in a coral reef ecosystem. *Hydrobiologia*, 637:101-119.

Kendall, M.S., Miller, T.J., & Pittman, S.J. (2011) Patterns of scale-dependency and the influence of map resolution on the seascape ecology of reef fish. *Marine Ecology Progress Series*, 427: 259-274.

Kenny, A.J., Cato, I., Desprez, M., Fader, G., Schüttenhelm, R.T.E., & Side, J. (2003) An overview of seabed-mapping technologies in the context of marine habitat classification. *ICES Journal of Marine Science*, 60:411-418.

Kinlan, B.P., & Gaines, S.D. (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. *Ecology*, 84:2007-2020.

Knobles, D.P., Wilson, P.S., Goff, J.A., & Cho, S.E. (2008) Seabed acoustics of a sand ridge on the New Jersey continental shelf. *The Journal of the Acoustical Society of America*, 124:EL151–EL156.

.

Koslow, J.A., Boehlert, G.W., Gordon, J.D.M., Haedrich, R.L., Lorance, P., & Parin, N. (2000) Continental slope and deep-sea fisheries: implications for a fragile ecosystem. *ICES Journal of Marine Science*, 57:548-557.

Kostylev, V.E., Todd, B.J., Fader, G.B.J., Courtney, R.C., Cameron, G.D.M., & Pickrill, R.A. (2001) Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs. *Marine Ecology Progress Series*, 219:121-137.

Kristan, W.B.I. (2006) Sources and expectations for hierarchical structure in bird-habitat associations. *Condor*, 108:5-12.

Kwasnitschka, T., Hansteen, T.H., Devey, C.W., & Kutterolf, S. (2013) Doing fieldwork on the seafloor: photogrammetric techniques to yield 3D visual models from ROV video. *Computers and Geosciences*, 52:218-226.

Laffan, S.W., Skidmore, A.K., & Franklin, J. (2012) Geospatial analysis of species, biodiversity and landscapes: introduction to the second special issue on spatial ecology. *International Journal of Geographical Information Science*, 26:2003-2007.

Laffan, S.W., Skidmore, A.K., & Franklin, J. (2014) Species distribution and diversity, habitat selection and connectivity: introduction to the third special issue on spatial ecology. *International Journal of Geographical Information Science*, 28:1527-1530.

Larsen, F.W., & Rahbek, C. (2005) The influence of spatial grain size on the suitability of the higher-taxon approach in continental priority-setting. *Animal Conservation* 8:389-396.

Latimer, A.M., Wu, S.S., Gelfand, A.E., & Silander, J.A. (2006) Building statistical models to analyze species distributions. *Ecological Applications*, 16:33-50.

Leaper, R., Dunstan, P.K., Foster, S.D., Barrett, N.J., & Edgar, G.J. (2012) Comparing large-scale bioregions and fine-scale community-level biodiversity predictions from subtidal rocky reefs across south-eastern Australia. *Journal of Applied Ecology*, 49:851-860.

Lechner, A.M., Langford, W.T., Bekessy, S.A., & Jones, S.D. (2012a) Are landscape ecologists addressing uncertainty in their remote sensing data? *Landscape Ecology*, 27:1249-1261.

Lechner, A.M., Langford, W.T., Jones, S.D., Bekessy, S.A., & Gordon, A. (2012b) Investigating species-environment relationships at multiple scales: differentiating between intrinsic scale and the modifiable areal unit problem. *Ecological Complexity*, 11:91-102.

Lecours, V., & Devillers, R. (2015) Assessing the spatial data quality paradox in the deep-sea. In: Sieber, R.E. (Ed.) *Proceedings of Spatial Knowledge and Information Canada (SKI-Canada) 2015*, February 27–28, Banff, Alberta, p. 1–8.

Lecours, V., Miles, L.L., Devillers, R., & Edinger, E.N. (2013) Data analysis towards the multiscale characterization of cold-water coral and sponge habitats in Canadian waters. *Technical Report REQ. No. F6160-120010*. Department of Fisheries and Oceans Canada, Newfoundland and Labrador Region, St. John's.

Lecours, V., Lucieer, V.L., Dolan, M.F.J., & Micallef, A. (2015) An ocean of possibilities: applications and challenges of marine geomorphometry. In: Jasiewicz, J., Zwolinski, Z., Mitasova, H., & Hengl, T. (Eds) *Geomorphometry for geosciences*. Bogucki Wydawnictwo Naukowe, Poznan, p. 23–26.

Legendre, P. (1993) Spatial autocorrelation: problem or new paradigm. *Ecology*, 74:1659-1673.

Legendre, L., & Demers, S. (1984) Towards dynamic biological oceanography and limnology. *Canadian Journal of Fisheries and Aquatic Sciences*, 41: 2-19.

Legendre, P., & Fortin, M.J. (1989) Spatial pattern and ecological analysis. *Vegetatio*, 80:107-138.

Lennon, J.J. (2000) Red shifts and red herrings in geographical ecology. *Ecography*, 23:101-113.

Levin, S.A. (1989) Challenges in the development of a theory of ecosystem structure and function. In: Roughgarden, J., May, R.M., & Levin, S.A. (Eds) *Perspectives in ecological theory*. Princeton University Press, Princeton, NJ, p. 242-255.

Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, 73:1943-1983.

Levin, L.A., & Dayton, P.K. (2009) Ecological theory and continental margins: where shallow meets deep. *Trends in Ecology and Evolution*, 24:606-617.

Li, Z. (2008) Multi-scale terrain modelling and analysis. In: Zhou, Q., Lees, B., & Tang, G.A. (Eds) *Advances in digital terrain analysis*. Springer, Berlin, p. 59-83.

Lloyd, C.D. (2010a) Analysing population characteristics using geographically weighted principal component analysis: a case study of Northern Ireland in 2001. *Computers, Environment and Urban Systems*, 34:389-399.

Lloyd, C.D. (2010b) Exploring population spatial concentrations in Northern Ireland by community background and other characteristics: an application of geographically weighted spatial statistics. *International Journal of Geographical Information Science*, 24: 1193-1221.

Lloyd, C.D. (2014) *Exploring spatial scale in geography*. Wiley-Blackwell, Chichester.

Lourie, S.A., & Vincent, A.C.J. (2004) Using biogeography to help set priorities in marine conservation. *Conservation Biology*, 18: 1004-1020.

Lu, B., Charlton, M., Harris, P., & Fotheringham, A.S. (2014a) Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. *International Journal of Geographical Information Science*, 28:660-681.

Lu, B., Harris, P., Gollini, I., Charlton, M., & Brunsdon, C. (2014b) The GWmodel R package: further topics for exploring spatial heterogeneity using geographically weighted models. *Geo-spatial Information Science*, 17:85-101.

Lucieer, V., Hill, N.A., Barrett, N.S., & Nichol, S. (2013) Do marine substrates 'look' and 'sound' the same? Supervised classification of multibeam acoustic data using autonomous underwater vehicle images. *Estuarine, Coastal and Shelf Science*, 117: 94-106.

Lüdtke, A., Jerosch, K., Herzog, O., & Schlüter, M. (2012) Development of a machine learning technique for automatic analysis of seafloor image data: Case example, *Pogonophora* coverage at mud volcanoes. *Computers and Geosciences*, 39:120-128.

MacMillan, R.A., & Shary, P.A. (2009) Landforms and landform elements in geomorphometry. In: Hengl, T., & Reuter, H.I. (Eds) Geomorphometry: concepts, software, applications. (*Developments in soil science 33*). Elsevier Science, Amsterdam, p. 227–254.

Maina, J., Venus, V., McClanahan, T.R., & Ateweberhan, M. (2008) Modelling susceptibility of coral reefs to environmental stress using remote sensing data and GIS models. *Ecological Modelling*, 212:180-199.

Marceau, D.J. (1999) The scale issue in social and natural sciences. *Canadian Journal of Remote Sensing*, 25:347-356.

Marceau, D.J., & Hay, G.J. (1999) Remote sensing contributions to the scale issue. *Canadian Journal of Remote Sensing*, 25:357-366.

Marsh, L., Copley, J.T., Huvenne, V.A.I., Tyler, P.A., & Isis ROV Facility (2013) Getting the bigger picture: using precision Remotely Operated Vehicle (ROV) videography to acquire high-definition mosaic images of newly discovered hydrothermal vents in the Southern Ocean. *Deep-Sea Research Part II*, 92:124-135.

Mateo Sánchez, M.C., Cushman, S.A., & Saura, S. (2014) Scale dependence in habitat selection: the case of the endangered brown bear (*Ursus arctos*) in the Cantabrian Range (NW Spain). *International Journal of Geographical Information Science*, 28:1531-1546.

Mayor, S.J., Schneider, D.C., Schaefer, J.A., & Mahoney, S.P. (2009) Habitat selection at multiple scales. *Ecoscience*, 16: 238-247.

McArthur, M.A., Brooke, B., Przeslawski, R., Ryan, D.A., Lucieer, V.L., Nichol, S., McCallum, A.W., Mellin, C., Cresswell, I.D., & Radke, L.C. (2009) A review of surrogates for marine benthic biodiversity. *Geoscience Australia Record 2009/42*, Geoscience Australia, Canberra.

McArthur, M.A., Brooke, B., Przeslawski, R., Ryan, D.A., Lucieer, V.L., Nichol, S., McCallum, A.W., Mellin, C., Cresswell, I.D., & Radke, L.C. (2010) On the use of abiotic surrogates to describe marine benthic biodiversity. *Estuarine, Coastal and Shelf Science*, 88:21-32.

McCoy, E.D., Bell, S.S., & Mushinsky, H.R. (1991) Habitat structure: synthesis and perspectives. In: Bell, S.S., McCoy, E.D., & Mushinsky, H.R. (Eds) *Habitat structure: the physical arrangement of objects in space*. Chapman & Hall, London, p. 427-430.

McLeod, I., Skewes, T.D., Gordon, S.R., & Pitcher, C.R. (2007) A method for seabed habitat mapping: integrating acoustic information with biogeophysical observations; case study – Scott Reef. In: Todd, B.J., & Greene, H.G. (Eds) Mapping the seafloor for habitat characterization. *Geological Association of Canada Special Paper* 47:309-317.

Meentemeyer, V. (1989) Geographical perspectives of space, time, and scale. *Landscape Ecology*, 3:163-173.

Meentemeyer, V., & Box, E.O. (1987) Scale effects in landscape studies. In: Turner, M.G. (Ed.) Landscape heterogeneity and disturbance. *Ecological Studies*, 64:15-34.

Mellin, C., Delean, S., Caley, J., Edgar, G., Meekan, M., Pitcher, R., Przeslawski, R., Williams, A., & Bradshaw, C. (2011) Effectiveness of biological surrogates for predicting patterns of marine biodiversity: a global meta-analysis. *PLoS ONE* 6:e20141.

Mengerink, K.J., Van Dover, C.L., Ardron, J., Baker, M., Escobar-Briones, E., Gjerde, K., Koslow, J.A., Ramirez-Llodra, E., Lara-Lopez, A., Squires, D., Sutton, T., Sweetman, A.K., & Levin, L.A. (2014) A call for deep-ocean stewardship. *Science*, 344: 696-698.

Menke, S.B., Holway, D.A., Fisher, R.N., & Jetz, W. (2009) Characterizing and predicting species distributions across environments and scales: Argentine ant occurrences in the eye of the beholder. *Global Ecology and Biogeography*, 18:50-63.

Meyer, C.B., & Thuiller, W. (2006) Accuracy of resource selection functions across spatial scales. *Diversity and Distribution*, 12:288-297.

Miller, J.A. (2012) Species distribution models: spatial auto-correlation and non-stationarity. *Progress in Physical Geography*, 36: 681-692.

Miller, J.A., & Hanham, R.Q. (2011) Spatial nonstationarity and the scale of species-environment relationships in the Mojave Desert, California, USA. *International Journal of Geographical Information Science*, 25: 423-438.

Miller, J., Franklin, J., & Aspinall, R. (2007) Incorporating spatial dependence in predictive vegetation models. *Ecological Modelling*, 202:225-242.

Mizon, G.E. (1995) A simple message for autocorrelation correctors: don't. *Journal of Econometrics*, 69:267-289.

Moran, P. (1948) The interpretation of statistical maps. *Journal of the Royal Statistical Society B*, 10:243-251.

Moudrý, V., & Šímová, P. (2012) Influence of positional accuracy, sample size and scale on modelling species distributions: a review. *International Journal of Geographical Information Science*, 26:2083-2095.

Mumby, P.J. (2006) Connectivity of reef fish between mangroves and coral reefs: algorithms for the design of marine reserves at seascapes scales. *Biological Conservation*, 128:215-222.

Mumby, P.J., Broad, K., Brumbaugh, D.R., Dahlgren, C.P., Harborne, A.R., Hastings, A., Holmes, K.E., Kappel, C.V., Micheli, F., & Sanchirico, J.N. (2008) Coral reef habitats as surrogates of species, ecological functions, and ecosystem services. *Conservation Biology*, 22:941-951.

Munday, P.L., Jones, G.P., & Caley, M.J. (1997) Habitat specialisation and the distribution and abundance of coral-dwelling gobies. *Marine Ecology Progress Series*, 152:227-239.

Muñoz-Reinoso, J.C. (2009) Boundaries and scales in shrublands of the Doñana biological reserve, Southwest Spain. *Landscape Ecology*, 24:509-518.

Muotka, T., Mäki-Petäys, A., Kreivi, P., & Högmander, H. (1998) Spatial associations between lotic fish, macroinvertebrate prey and the stream habitat: a multi-scale approach. *Boreal Environment Research*, 3:371-380.

Nash, K.L., Allen, C.R., Angeler, D.G., Barichievy, C., Eason, T., Garmestani, A.S., Graham, N.A.J., Granholm, D., Knutson, M., Nelson, R.J., Nyström, M., Stow, G.A., & Sundstrom, S.M. (2014) Discontinuities, cross-scale patterns, and the organization of ecosystems. *Ecology*, 95:654-667.

Nelson, A. (2001) Analysing data across geographic scales in Honduras: detecting levels of organisation within systems. *Agriculture, Ecosystems and Environment*, 85:107-131.

Olsgard, F., Schaanning, M.T., Widdicombe, S., Kendall, M.A., & Austen, M.C. (2008) Effects of bottom trawling on ecosystem functioning. *Journal of Experimental Marine Biology and Ecology*, 366:123-133.

Openshaw, S. (1984) The modifiable areal unit problem. *CATMOG – Concepts and Techniques in Modern Geography*, Vol. 38, Geo Books, Norwich.

Openshaw, S., & Taylor, P.J. (1979) A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In: Wrigley, N. (Ed.) *Statistical applications in the spatial sciences*. Pion, London, p. 127-144.

Owen, M. (1972) Some factors affecting food intake and selection in white-fronted geese. *Journal of Animal Ecology*, 41:79-92.

Oyana, T.J., Achenie, L.E.K., & Heo, J. (2012) The new and computationally efficient MIL-SOM algorithm: potential benefits for visualization and analysis of a large-scale high-dimensional clinically acquired geographic data. *Computational and Mathematical Methods in Medicine*, 2012:683265.

Pandit, S.N., Kolasa, J., & Cottenie, K. (2009) Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology*, 90: 2253-2262.

Pardo-Igúzquiza, E., & Dowd, P.A. (2002) FACTOR2D: A computer program for factorial cokriging. *Computers and Geosciences*, 28: 857-875.

Pearson, S.M. (2002) Interpreting landscape patterns from organism-based perspectives. In: Gergel, S.E., & Turner, M.G. (Eds) *Learning landscape ecology*. Springer, New York, NY, p. 187-198.

Pearson, R.G., & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, 12:361-371.

Peterson, A.T., & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, 17:135-144.

Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Nakamura, M., & Araújo, M.B. (2011) Ecological niches and geographic distributions. *Monographs in Population Biology*, 49.

Pittman, S.J., & Brown, K.A. (2011) Multi-scale approach for predicting fish species distributions across coral reef seascapes. *PLoS ONE* 6:e20583.

Pittman, S.J., & McAlpine, C.A. (2003) Movements of marine fish and decapod crustaceans: process, theory and application. *Advances in Marine Biology*, 44:205-294.

Pittman, S.J., Caldow, C., Hile, S.D., & Monaco, M.E. (2007) Using seascape types to explain the spatial patterns of fish in the mangroves of SW Puerto Rico. *Marine Ecology Progress Series*, 348: 273-284.

Pittman, S.J., Costa, B.M., & Battista, T.A. (2009) Using LiDAR bathymetric and boosted regression trees to predict the diversity and abundance of fish and corals. *Journal of Coastal Research*, 53:27-38.

Pittman, S.J., Kneib, R.T., & Simenstad, C.A. (2011) Practicing coastal seascape ecology. *Marine Ecology Progress Series*, 427:187-190.

Pogson, M., & Smith, P. (2015) Effect of spatial data resolution on uncertainty. *Environmental Modelling and Software*, 63:87-96.

Prados, R., Garcia, R., Gracias, N., Escartín, J., & Neumann, L. (2012) A novel blending technique for underwater gigamosaicing. *IEEE Journal of Oceanic Engineering*, 37:626-644.

Preston, J. (2009) Automated acoustic seabed classification of multibeam images of Stanton Banks. *Applied Acoustics*, 70: 1277-1287.

Puig, P., Canals, M., Company, J.B., Martín, J., Amblas, D., Lastras, G., Palanques, A., & Calafat, A.M. (2012) Ploughing the deep sea floor. *Nature*, 489:286-289.

Purkis, S.J., & Kohler, K.E. (2008) The role of topography in promoting fractal patchiness in a carbonate shelf landscape. *Coral Reefs*, 27:977-989.

Purser, A., Bergmann, M., Lundälv, T., Ontrup, J., & Nattkemper, T.W. (2009) Use of machine-learning algorithms for the automated detection of cold-water coral habitats: a pilot study. *Marine Ecology Progress Series*, 397:241-251.

Rahbek, C., & Graves, G.R. (2001) Multiscale assessment of patterns of avian species richness. *Proceedings of the National Academy of Science USA*, 98:4534-4539.

Ramirez-Llodra, E., Tyler, P.A., Baker, M.C., Aksel Bergstad, O., Clark, M.R., Escobar, E., Levin, L.A., Menot, L., Rowden, A.A., Smith, C.R., & Van Dover, C.L. (2011) Man and the last great wilderness: Human impact on the deep sea. *PLoS ONE*, 6:e22588.

Rengstorf, A.M., Grehan, A., Yesson, C., & Brown, C. (2012) Towards high-resolution habitat suitability modeling of vulnerable marine ecosystems in the deep-sea: resolving terrain attribute dependencies. *Marine Geodesy*, 35:343-361.

Rengstorf, A.M., Yesson, C., Brown, C., & Grehan, A.J. (2013) High-resolution habitat suitability modelling can improve conservation of vulnerable marine ecosystems in the deep sea. *Journal of Biogeography*, 40:1702-1714.

Reside, A.E., Watson, I., & VanDerWal, J. (2011) Incorporating low-resolution historic species location data decreases performance of distribution models. *Ecological Modelling*, 222: 3444-3448.

Rex, M.A., & Etter, R.J. (2010) *Deep-sea biodiversity: pattern and scale*. Harvard University Press, Cambridge, MA.

Rigby, P., Pizarro, O., & Williams, S.B. (2010) Toward adaptive benthic habitat mapping using Gaussian process classification. *Journal of Field Robotics*, 27:741-758.

Roberts, J.M., Henry, L.A., Long, D., & Hartley, J.P. (2008) Cold-water coral reef frameworks, megafaunal communities and evidence for coral carbonate mounds on the Hatton Bank, north east Atlantic. *Facies*, 54:297-316.

Robinson, W.S. (1950) Ecological correlation and the behavior of individuals. *American Sociological Review*, 15:351-357.

Robinson, L.M., Elith, J., Hobday, A.J., Pearson, R.G., Kendall, B.E., Possingham, H.P., & Richardson, A.J. (2011) Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. *Global Ecology and Biogeography*, 20:789-802.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jiménez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. (2011) Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Progress in Physical Geography*, 35, 211-226.

Roff, J.C., Taylor, M.E., & Laughren, J. (2003) Geophysical approaches to the classification, delineation and monitoring of marine habitats and their communities. *Aquatic Conservation*, 13:77-90.

Ross, R.E., & Howell, K.L. (2013) Use of predictive habitat modelling to assess the distribution and extent of the current protection of 'listed' deep-sea habitats. *Diversity and Distribution*, 19: 433-445.

Sarkar, S., Justus, J., Fuller, T., Kelley, C., Garson, J., & Mayfield, M. (2005) Effectiveness of environmental surrogates for the selection of conservation area networks. *Conservation Biology*, 19: 815-825.

Saura, S. (2002) Effects of minimum mapping unit on land cover data spatial configuration and composition. *International Journal of Remote Sensing*, 23:4853-4880.

Schimel, A.C.G., Healy, T.R., Johnson, D., & Immenga, D. (2010) Quantitative experimental comparison of single-beam, sidescan, and multibeam benthic habitat maps. *ICES Journal of Marine Science*, 67:1766-1779.

Schmidt, J., & Andrew, R. (2005) Multi-scale landform characterization. *Area*, 37:341-350.

Schneider, D.C. (2001a) The rise of the concept of scale in ecology. *Bioscience*, 51:545-553.

Schneider, D.C. (2001b) Concepts and effects of scale. In: Levin, S.A. (Ed.) *Encyclopedia of Biodiversity*. Academic Press, San Diego, CA, p. 245-254.

Schneider, D.C. (1994) *Quantitative ecology: spatial and temporal scaling*. Academic Press, San Diego, CA.

Schneider, D.C. (2009) *Quantitative ecology, 2nd edition: measurements, models and scaling*. Academic Press, San Diego, CA.

Schneider, D.C., & Haedrich, R.L. (1991) Post-mortem erosion of fine-scale spatial structure of epibenthic megafauna on the outer Grand Bank of Newfoundland. *Continental Shelf Research*, 11:1223-1236.

Schneider, D.C., Gagnon, J.M., & Gilkinson, K.D. (1987) Patchiness of epibenthic megafauna on the outer Grand Banks of Newfoundland. *Marine Ecology Progress Series*, 39:1-13.

Schoch, G.C., & Dethier, M.N. (1996) Scaling up: the statistical linkage between organismal abundance and geomorphology on rocky intertidal shorelines. *Journal of Experimental Marine Biology and Ecology*, 201:37-72.

Segurado, P, Araújo, M.B., & Kunin, W.E. (2006) Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology*, 43:433-444.

Seiler, J., Friedman, A., Steinberg, D., Barrett, N., Williams, A., & Holbrook, N.J. (2012) Image-based continental shelf habitat mapping using novel automated data extraction techniques. *Continental Shelf Research*, 45:87-97.

Seo, C., Thorne, J.H., Hannah, L., & Thuiller, W. (2009) Scale effects in species distribution models: implication for conservation planning under climate change. *Biology Letters*, 5:39-43.

Shary, P.A., Sharava, L.S., & Mitusov, A.V. (2002) Fundamental quantitative methods of land surface analysis. *Geoderma*, 107:1-32.

Sherry, T.W., & Holmes, R.T. (1988) Habitat selection by breeding American redstarts in response to a dominant competitor, the least flycatcher. *Auk*, 105:350-364.

Shucksmith, R.J., & Kelly, C. (2014) Data collection and mapping – principles, processes and application in marine spatial planning. *Marine Policy*, 50:27-33.

Skidmore, A.K., Franklin, J., Dawson, T.P., & Pilesjö, P. (2011) Geospatial tools address emerging issues in spatial ecology: a review and commentary on the special issue. *International Journal of Geographical Information Science*, 25:337-365.

Smith, P.A. (1994) Autocorrelation in logistic regression modelling of species' distribution. *Global Ecology and Biogeography*, 4: 47-61.

Smith, G., & Brennan, R.E. (2012) Losing our way with mapping: Thinking critically about marine spatial planning in Scotland. *Ocean and Coastal Management*, 69:210-216.

Snickars, M., Gullström, M., Sundblad, G., Bergström, U., Downie, A.L., Lindegarth, M., & Mattila, J. (2014) Species-environment relationships and potential for distribution modelling in coastal waters. *Journal of Sea Research*, 85:116-125.

Sokal, R.R. (1979) Ecological parameters inferred from spatial correlograms. In: Patil, G.P., & Rosenzwig, M.L. (Eds) *Contemporary quantitative ecology and related ecometrics*. International Co-operative Publishing House, Fairland, MD, p. 167-196.

Sokal, R.R., & Oden, N.L. (1978) Spatial autocorrelation in biology: some biological implications and four applications of evolutionary and ecological interest. *Biological Journal of the Linnean Society*, 10:229-249.

Solan, M., Germano, J.D., Rhoads, D.C., Smith, C., Michaud, E., Parry, D., Wenzhöfer, F., Kennedy, B., Henriques, C., Battle, E., Carey, D., Iocco, L., Valente, R., Watson, J., & Rosenberg, R. (2003) Towards a greater understanding of pattern, scale and process in marine benthic systems: a picture is worth a thousand worms. *Journal of Experimental Marine Biology and Ecology*, 285-286:313-338.

Steele, J.H. (1978) Some comments on plankton patches. In: Steele, J.H. (Ed.) *Spatial pattern in plankton communities*. Plenum Press, New York, NY, p. 1-20.

Steele, J.H. (1991) Can ecological theory cross the land-sea boundary? *Journal of Theoretical Biology*, 153:425-436.

Stommel, H. (1963) The varieties of oceanographic experience. *Science*, 139:572-576.

Stone, K.H. (1972) A geographer's strength: the multiple scales approach. *Journal of Geography*, 71:354-362.

Stone, R.P. (2006) Coral habitat in the Aleutian Islands of Alaska: depth distribution, fine-scale species associations, and fisheries interactions. *Coral Reefs*, 25:229-238.

Store, R., & Jokimäki, J. (2003) A GIS-based multi-scale approach to habitat suitability modeling. *Ecological Modelling*, 169:1-15.

Synes, N.W., & Osborne, P.E. (2011) Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Global Ecology and Biogeography*, 20:904-914.

Tabachnick, B.G., & Fidell, L.S. (2013) *Using multivariate statistics, 6th edition*. Pearson Education Limited, Harlow.

Tanner, J.E., Mellin, C., Parrott, L., & Bradshaw, C.J.A. (2015) Fine-scale biodiversity patterns inferred from image processing. *Ecological Complexity*, 22:76-85.

Thrush, S.F., & Dayton, P.K. (2002) Disturbance to marine benthic habitats by trawling and dredging: implications for marine biodiversity. *Annual Review of Ecology and Systematics*, 33:449-473.

Thurber, A.R., Sweetman, A.K., Narayanaswamy, B.E., Jones, D.O.B., Ingels, J., & Hansman, R.L. (2013) Ecosystem function and services provided by the deep sea. *Biogeosciences Discussions*, 10:18193-18240.

Tissot, B.N., Yoklavich, M.M., Love, M.S., York, K., & Amend, M. (2006) Benthic invertebrates that form habitat on deep banks off southern California, with special reference to deep sea coral. *Fishery Bulletin*, 104:167-181.

Tittensor, D.P., Baco, A.R., Brewin, P.E., Clark, M.R., Consalvey, M., Hall-Spencer, J., Rowden, A.A., Schlacher, T., Stocks, K.I., & Rogers, A.D. (2009) Predicting global habitat suitability for stony corals on seamounts. *Journal of Biogeography*, 36:1111-1128.

Tobler, W.R. (1970) A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46:234-240.

Todd, B.J., & Greene, H.G. (Eds) (2007) Mapping the seafloor for habitat characterization. *Geological Association of Canada Special Paper* 47.

Tong, R., Purser, A., Guinan, J., & Unnithan, V. (2013) Modeling the habitat suitability of deep-water gorgonian corals based on terrain variables. *Ecological Informatics*, 13:123-132.

Turner, M.G., Gardner, R.H., & O'Neill, R.V. (2001) *Landscape ecology, in theory and practice: pattern and process*. Springer-Verlag, New York, NY.

Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A.K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., Koh, L.P., Leimgruber, P., Taubenboeck, H., Wegmann, M., Wikelski, M., & Woodcock, C. (2015) Free and open-access satellite data are key to biodiversity conservation. *Biological Conservation*, 182:173-176.

Urban, D.L., O'Neill, R.V., & Shugart Jr., H.H. (1987) Landscape ecology. *Bioscience*, 37:119-127.

Václavík, T., Kupfer, J.A., & Meentemeyer, R.K. (2012) Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography*, 39:42-55.

Van Rein, H.B., Brown, C.J., & Quinn, R. (2009) A review of sub-littoral monitoring methods in temperate waters: a focus on scale. *Underwater Technology*, 28:99-113.

VanDerWal, J., Shoo, L.P., Graham, C., & Williams, S.E. (2009) Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling*, 220:589-594.

VanderWerf, E.A. (1993) Scales of habitat selection by foraging 'elepaio in undisturbed and human-altered forests in Hawaii. *Condor*, 95:980-989.

Veloz, S.D. (2009) Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, 36:2290-2299.

Verfaillie, E., Du Four, I., Van Meirvenne, M., & Van Lancker, V. (2009) Geostatistical modeling of sedimentological parameters using multi-scale terrain variables: application along the Belgian part of the North Sea. *International Journal of Geographical Information Science*, 23:135-150.

Vierod, A.D.T., Guinotte, J.M., & Davies, A.J. (2014) Predicting the distribution of vulnerable marine ecosystems in the deep sea using presence-background models. *Deep-Sea Research Part II*, 99:6-18.

Ward, T.J., Vanderklift, M.A., Nicholls, A.O., & Kenchington, R.A. (1999) Selecting marine reserves using habitats and species assemblages as surrogates for biological diversity. *Ecological Applications*, 9:691-698.

Wedding, L.M., & Friedlander, A.M. (2008) Determining the influence of seascape structure on coral reef fishes in Hawaii using a geospatial approach. *Marine Geodesy*, 31:246-266.

Wedding, L.M., Friedlander, A.M., McGranaghan, M., Yost, R.S., & Monaco, M.E. (2008) Using bathymetric LiDAR to define nearshore benthic habitat complexity: implications for management of reef fish assemblages in Hawaii. *Remote Sensing of Environment*, 112:4159-4165.

Wedding, L.M., Lepczyk, C.A., Pittman, S.J., Friedlander, A.M., & Jorgensen, S. (2011) Quantifying seascape structure: extending terrestrial spatial pattern metrics to the marine realm. *Marine Ecology Progress Series*, 427:219-232.

Wheatley, M., & Johnson, C. (2009) Factors limiting our understanding of ecological scale. *Ecological Complexity*, 6:150-159.

Whitmire, C.E., Embley, R.W., Wakefield, W.W., Merle, S.G., & Tissot, B.N. (2007) A quantitative approach for using multibeam sonar data to map benthic habitats. In: Todd, B.J., & Greene, H.G. (Eds) Mapping the seafloor for habitat characterization. *Geological Association of Canada Special Paper* 47:111-126.

Whittaker, R.H., Levin, S.A., & Root, R.B. (1973) Niche, habitat, and ecotope. *American Naturalist*, 107:321-338.

Wiens, J.A. (1989) Spatial scaling in ecology. *Functional Ecology*, 3: 385-397.

Wiens, J.A., Addicott, J.F., Case, T.J., & Diamond, J. (1986a) Overview: the importance of spatial and temporal scale in ecological investigations. In: Diamond, J., & Case, T.J. (Eds) *Community ecology*. Harper & Row, New York, NY, p. 145-153.

Wiens, J.A., Rotenberry, J.T., & Van Horne, B. (1986b) A lesson in the limitations of field experiments: shrubsteppe birds and habitat alteration. *Ecology*, 67:365-376.

Williams, A., Schlacher, T.A., Rowden, A.A., Althaus, F., Clark, M.R., Bowden, D.A., Stewart, R., Bax, N.J., Consalvey, M., & Kloser, R.J. (2010) Seamount megabenthic assemblages fail to recover from trawling impacts. *Marine Ecology*, 31:183-199.

Williams, K.J., Belbin, L., Austin, M.P., Stein, J.L., & Ferrier, S. (2012) Which environmental variables should I use in my biodiversity model? *International Journal of Geographical Information Science*, 26:2009-2047.

Willis, K.J., & Whittaker, R.J. (2002) Species diversity – scale matters. *Science* 295:1245-1248.

Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., & Grehan, A.J. (2007) Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30:3-35.

Windle, M.J.S., Rose, G.A., Devillers, R., & Fortin, M.J. (2010) Exploring spatial non-stationarity of fisheries survey data using geographically weighted regression (GWR): an example from the Northwest Atlantic. *ICES Journal of Marine Science*, 67:145-154.

Windle, M.J.S., Rose, G.A., Devillers, R., & Fortin, M.J. (2012) Spatio-temporal variations in invertebrate-cod-environment relationships on the Newfoundland-Labrador Shelf, 1995-2009. *Marine Ecology Progress Series*, 469:263-278.

With, K.A., & King, A.W. (1997) The use and misuse of neutral landscape models in ecology. *Oikos*, 79:219-229.

Wright, D.J. (1999) Getting to the bottom of it: tools, techniques, and discoveries of deep ocean geography. *The Professional Geographer*, 51:426-439.

Wright, D., & Goodchild, M.F. (1997) Data from the deep: implications for the GIS community. *International Journal of Geographical Information Science*, 11: 523-528.

Wu, J. (1999) Hierarchy and scaling: extrapolating information along a scaling ladder. *Canadian Journal of Remote Sensing*, 25:367-380

Wu, J., Gao, W., & Tueller, P.T. (1997) Effects of changing spatial scale on the results of statistical analysis with landscape data: a case study. *Geographic Information Sciences*, 3:30-41.

Wu, X.B., & Smeins, F.E. (2000) Multiple-scale habitat modeling approach for rare plant conservation. *Landscape and Urban Planning*, 51:11-28.

Young, C.M. (2009) Communities on deep-sea hard bottoms. In: Wahl, M. (Ed.) *Marine hard bottom communities: patterns, dynamics, diversity, and change*. Springer, Berlin, p. 39-60.

Zajac, R. (2008) Challenges in marine, soft-sediment benthoscape ecology. *Landscape Ecology*, 23:7-18.

Zhang, J., Atkinson, P.M., & Goodchild, M.F. (2014) *Scale in spatial information and analysis*. CRC Press, Boca Raton, FL.

Zieger, S., Stieglitz, T., & Kininmonth, S. (2009) Mapping reef features from multibeam sonar data using multiscale morphometric analysis. *Marine Geology*, 264:209-217.

Zimmermann, N.E., Edwards, T.C., Graham, C.G., Pearman, P.B., & Svenning, J.C. (2010) New trends in species distribution modelling. *Ecography*, 33:985-989.

3. Towards a Framework for Terrain Attribute Selection in Environmental Studies

3.1 Introduction

Combining georeferenced species data with environmental datasets has become common practice in environmental studies both in the terrestrial (Elith & Leathwick, 2009) and marine realms (Brown *et al.*, 2011). Exploring species-environment relationships is important for habitat mapping, biogeographical classification, conservation, and management (Harris & Baker, 2012). Research in these fields has been fueled by progresses in remote sensing and Geographic Information Systems (GIS), along with the increase in data availability and computing power (Wiersma *et al.*, 2011; Vierod *et al.*, 2014). In parallel, these elements have motivated the development of geomorphometry (Bishop *et al.*, 2012; Zhou & Zhu, 2013), the field that helps quantitatively describe digital terrain models (DTM) using terrain attributes such as slope, orientation or rugosity (Pike, 1995). Terrain attributes have been found to be linked with the distribution of many terrestrial and marine species in different types of environments (*e.g.* forests, agroecosystems, deep-sea, continental shelf) and are now routinely integrated in environmental studies (Bouchet *et al.*, 2015). Other environmental disciplines that make use of terrain attributes include hydrology, soil mapping, vegetation mapping, geomorphology, meteorology and agriculture (Florinsky *et al.*, 2002; Lacroix *et al.*, 2002; Schwanghart & Heckmann, 2012; Hengl & Reuter, 2009).

Led by the increasing availability of different types of intuitive GIS tools that "automatically" derive terrain attributes from DTMs (Bishop *et al.*, 2012; *e.g.* Klingseisen *et al.*, 2008; Han *et al.*, 2012; Rigol-Sanchez *et al.*, 2015) - either digital elevation (DEM) or bathymetric (DBM) models - ecologists and other GIS users often select a small subset of terrain attributes to perform their analyses. Non-expert GIS users do not always understand the underpinnings of the numerous options available (Bishop & Shroder, 2004; Bouchet *et al.*, 2015), and a lack of guidance can lead them to select an arbitrary and sub-optimal set of terrain attributes. Such selections are often based on the availability and simplicity of the GIS tools rather than on statistical grounds or ecological, biological, or geomorphological relevance. An inappropriate selection of terrain attributes can however produce results that do not accurately represent the observed phenomenon, fail to capture the key properties of the terrain relevant to the question or problem, and influence subsequent analysis (*e.g.* species-environment relationship measurements).

A same terrain attribute derived using different algorithms can also produce significantly different outcomes. For instance, Dolan & Lucieer (2014) demonstrated that five different slope algorithms derived from DBMs resulted in different slope surfaces, confirming previous work performed on DEMs (Jones, 1998a) and artificial surfaces (Jones, 1998b). Since the algorithms used by GIS tools are not always made explicit within the software, users are often left with little choice on which one to use and sometimes are not free to decide the details of particular parameters such as the neighbourhood size. These elements, combined with the lack of explicit statements in the ecological literature of algorithms and parameters used for deriving terrain attributes

(Dolan & Lucieer, 2014), may lead to misleading and incorrect comparisons of results from different studies. To add to the confusion, geomorphometry is a field recognized for its ambiguous terminology (Bishop *et al.*, 2012), where terrain attributes measuring a same terrain characteristic can be named differently depending on the source or software.

Finally, a poor selection of terrain attributes may cause covariation between variables. Being all derivatives of the same DTM, terrain attributes are likely to covary and induce redundancy in the analysis (Pittman *et al.* 2009), violating the basic assumptions of many statistical analysis methods used. For instance, Rooper & Zimmermann (2007) calculated a correlation of 0.90 between their measures of slope and rugosity. Assessing covariation between variables is however rarely performed (Graham, 2003), despite being recognized to obscure the influence of individual drivers on a response variable, and to impact statistical models, species distribution models and regression analyses (Hijmans, 2012; Dormann *et al.*, 2013).

Selecting a suitable set of independent variables, including terrain attributes, is essential to ensure robust analyses and increase reliability of results in environmental studies (King & Jackson, 1999). A theoretical and operational framework to geomorphometric analysis is still to be defined (Pike, 1995), and "the use of quantitative geomorphological knowledge must be revisited in an analytical framework" (Bishop *et al.*, 2012, p.6). This paper bridges geomorphometry and environmental studies by proposing an operational framework that addresses the common issue of terrain attribute selection in environmental applications like ecology. It aims to identify combinations of available terrain attributes that minimize covariation between attributes and optimize the information given on the characteristics of a terrain. The specific objectives are to 1) explore existing GIS software to compute available terrain attributes, 2) identify groups of terrain attributes that represent unique morphological terrain characteristics, 3) and explore the relationship between the importance of these groups and terrain complexity.

3.2 Materials and Methods

A summary of the methods is presented in Figure 3.1. First, DTMs were generated from which terrain attributes were derived. Then, three iterative statistical methods were used to explore independence and both linear and non-linear relationships amongst terrain attributes: Principal Component Analysis (PCA), Variable Inflation Factor (VIF), and Mutual Information (MI). Because of the intricacy of the methods used in this study, more details on how and why they were used are provided in Appendix A to allow potential replication and generalization.

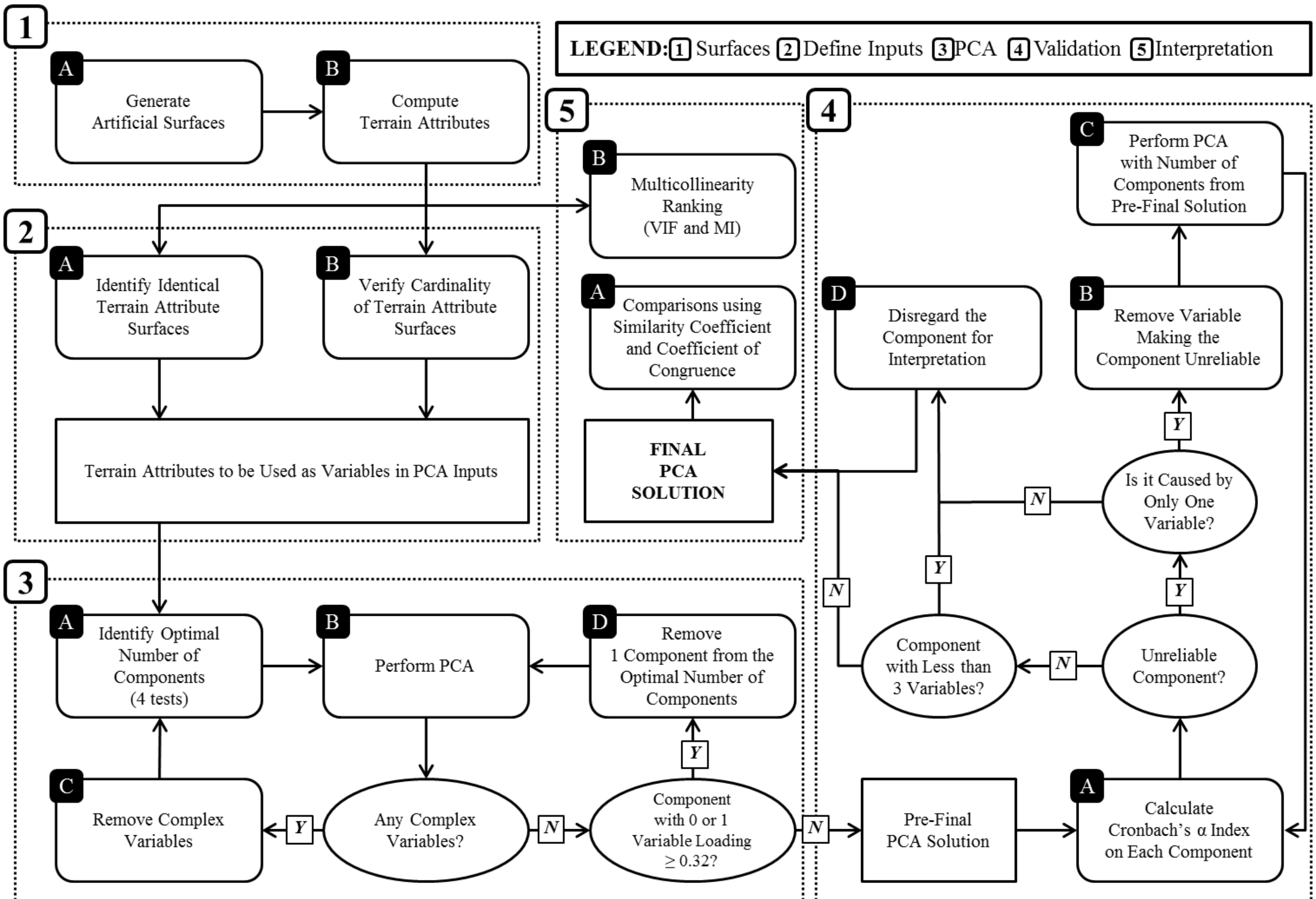

Figure 3.1: Conceptual model of the analysis performed on each artificial surface.

3.2.1 Surfaces and Terrain Attributes

3.2.1.1 Artificial Surfaces

Artificial surfaces have proven to be valuable tools in ecology (With & King, 1997) and geomorphometry (Jones, 1998b). Since terrain attributes are sensitive to DTM errors and uncertainty (Raaflaub & Collins, 2006; Kinsey-Henderson & Wilkinson, 2013), artificial surfaces provide a controlled environment in which to test hypotheses (Halley *et al.*, 2004). Nine artificial surfaces of 106 m x 106 m, with a 1 m spatial resolution, and presenting different complexity levels were created using spectral synthesis in LandSerf 2.3 (step 1A in Figure 3.1; Figure 3.2). The spectral synthesis technique, developed by Peitgen & Saupe (1988), is one of the methods that provide the most realistic artificial landscapes (Chipperfield *et al.*, 2011). It generates surfaces so that their correlative characteristics, which are studied in this paper, are present at all spatial scales (Keitt, 2000). LandSerf requires a user-specified fractal dimension to create surfaces. Fractal dimension expresses and encodes the degree of complexity of objects (Dimri, 2005) and is a "useful simple summary of roughness distribution" that can be used to compare DTM characteristics (Lloyd, 2014, p.152). Fractal dimension was chosen to differentiate surfaces because of its scale-invariance properties (Pfeifer, 1984). These properties and the preservation of the correlative characteristics of the surfaces make results from this study generalizable to other surfaces with the same fractal dimension at different scales (see Appendix A). The fractal dimension of the computed surfaces ranges from 2.06 to 2.79 (Figure 3.2), covering most complexity levels found in real terrain (2.20 to 2.60, Hofierka *et al.* (2009)) and other natural elements (2.28 to 2.61 in coral reefs, Zawada & Brock (2009)).

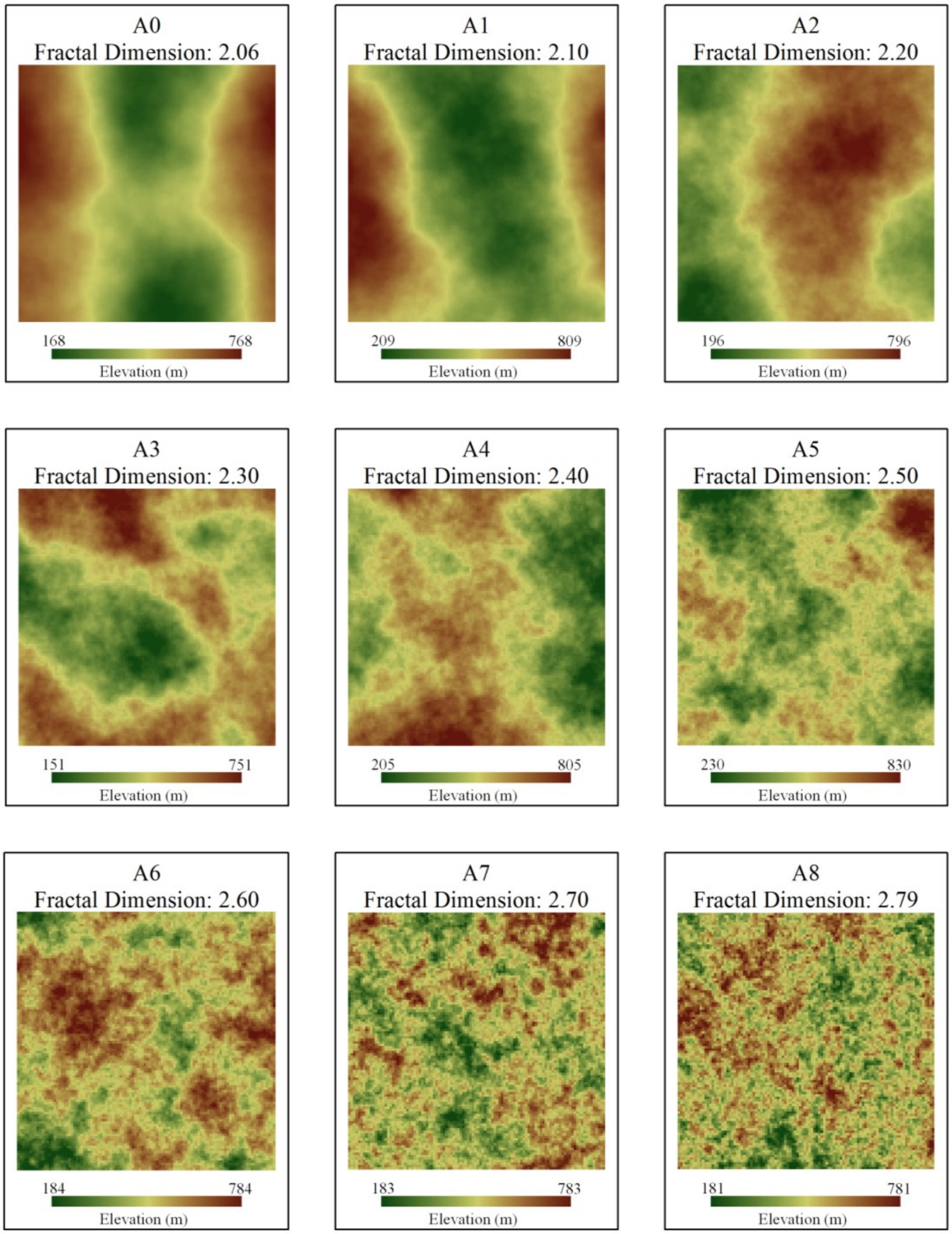

Figure 3.2: Artificial surfaces computed and analyzed in this study, from least complex (A0) to most complex (A8).

3.2.1.2 Terrain Attributes and GIS Tools

For each of the nine surfaces, 230 terrain attribute surfaces were derived (step 1B in Figure 3.1; Appendix B) using 11 different commercial and open-source software (Table 3.1). A 3 x 3 analysis window was used as several software packages use this as default and do not allow modifying it. As commonly performed in ecology, measures of aspect (orientation) were transformed into northerness and easterness to remove circularity in the data (Olaya, 2009). To eliminate edge contamination, the outer 3 m were removed from all resulting attribute surfaces, reducing the extent to 100 m x 100 m surfaces.

Table 3.1: List of software used, number of terrain attributes that were computed using each of them, and percentage of these terrain attributes that reached the final PCA solution for each surface.

Software and Versions	Number of Attributes Computed	Percentage of Terrain Attributes in the Final Solution								
		A0	A1	A2	A3	A4	A5	A6	A7	A8
ArcGIS 10.2.2 with Python 2.7.8	22	86	91	82	77	73	77	77	82	82
ArcGIS 10.2.2 with DEM Surface Tools (v.2.1.399)	17	82	76	76	76	71	65	65	71	65
ArcGIS 10.2.2 with Benthic Terrain Modeler 3.0 rc3	12	83	92	83	83	83	83	83	83	67
Diva-GIS 7.5.0	7	86	86	100	100	100	100	100	100	100
Idrisi Selva 17.0	7	86	86	71	71	86	86	86	86	86
Landserf 2.3	12	75	58	58	58	58	50	50	42	42
Quantum GIS 2.4.0 Chugiak	13	100	100	100	100	100	100	100	100	85
SAGA GIS 2.0.8	96	76	74	80	80	78	58	71	68	66
TNTmips Free 2014 (MicroImages)	25	100	100	80	84	96	56	92	84	80
uDig 1.4.0b	9	89	100	78	67	78	78	78	56	56
Whitebox GAT 3.2.1 Iguazu	10	90	90	100	90	90	70	80	80	70
Total Number of Attributes:	230	83	83	81	80	81	67	77	74	70

3.2.1.3 Preparation of the Data

To identify terrain attributes computed using the same algorithm, each attribute was tested against the others to detect pairs giving identical results (step 2A, Figure 3.1). For each set of duplicates, only one of the two attributes was kept for further analysis to account for each algorithm only once. To keep the analysis objective in terms of software used, the attributes removed were added back before interpreting the results, assuming that their behaviour through the analyses would have been the same as their duplicate.

Cardinality (*i.e.* the number of different values for a variable) was assessed for each terrain attribute (step 2B, Figure 3.1). Terrain attributes with low cardinality were identified and not used as input in PCA analyses (Section 2.2): such variables are known to complicate PCA solution as they account for only a negligible amount of the total variance (Tabachnick & Fidell, 2014). They were however carried over to the VIF and MI analyses (Section 3.2.3 below).

3.2.2 Principal Component Analysis

The first statistical analysis performed on the datasets of terrain attributes for each surface was an iterative PCA (step 3, Figure 3.1). PCA was chosen to address the second objective as it helps removing collinearity and redundancy, and regrouping variables into uncorrelated groups of correlated variables (*i.e.* components) while minimizing information loss (Tabachnick & Fidell, 2014). PCA is often used in ecology to explore patterns in large multidimensional datasets as it can handle variable dependency where other multivariate techniques cannot (King & Jackson, 1999; McGarigal *et al.*, 2000). PCA were performed using IBM SPSS Statistics software v.22, correlation matrices with standardized data, and a Varimax orthogonal rotation with Kaiser Normalization (Kaiser, 1958) (step 3B in Figure 3.1; Appendix A). The orthogonal rotation of the solution allows for a "theoretically more meaningful" component structure that is easier to interpret

(McGarigal *et al.*, 2000, p.59), and the Varimax method is the most widely used and often performs better than others (Bhattacharyya, 1981; Henson & Roberts, 2006).

3.2.2.1 Identifying the Optimal Number of Components

An over-extraction or under-extraction of components in ecological studies can lead to inferential issues (Franklin *et al.*, 1995; Wood *et al.*, 1996) (Appendix A). Several methods exist to estimate the appropriate number of PCA components to use (Zwick & Velicer, 1986). Since these methods are known to often give different results (Zwick & Velicer, 1986; O'Connor, 2000), Thompson & Daniel (1996, p.200) state that it is "appropriate and often desirable" to simultaneously use multiple methods. Using O'Connor's (2000) SPSS programs, four methods were combined to determine the statistically optimal number of components to retain (step 3A, Figure 3.1): Minimum Average Partial Correlation (MAP) (Velicer, 1976), Parallel Analysis (PA) (Horn, 1965), and modifications of MAP and PA respectively proposed by Velicer *et al.* (2000) and O'Connor (2015). The mode of the four results was chosen as the appropriate number of components to retain. Since this number depends on the number of PCA input variables, this step was performed before each PCA computation.

3.2.2.2 Complexity of Variables

Our third objective required comparing components from different solutions, which is only possible if solutions reach a simple structure. A simple structure has an invariance property (Kaiser, 1958) that allows generalization (Rummel, 1970). When a simple structure is reached, PCA solutions from different computations will always find the same components regardless of the insertion of other variables in the dataset (Rummel, 1970). A simple structure is deemed to be reached when most components have "marker" variables, *i.e.* those that load strongly on only one component. Variables that load strongly on more than one component are called "complex" variables. They are redundant, do not contribute to the model, and have to be removed for the solution to reach a simple structure (Tabachnick & Fidell, 2014). After removing complex variables (step 3C, Figure 3.1), steps 3A and 3B were performed again, and more terrain attributes were removed if new complex variables appeared. Iterations stopped once the PCA ceased to isolate complex variables.

3.2.2.3 Towards a Final Solution

After removing complex variables, the previously obtained optimal number of components may need to be adjusted. If the last component had no or only one marker variable loading on it, the PCA was re-run with one less component (step 3D, Figure 3.1). Once there were no more complex variable and no more components with less than two variables, the solutions had reached a simple structure and were ready for validation.

3.2.2.4 Validation

Components' internal consistency was assessed in SPSS using Cronbach's α coefficient of reliability (Cronbach, 1951) (step 4A, Figure 3.1; Appendix A). Unreliable components (α index below 0.6) were made reliable when possible (*i.e.* if unreliability was caused by a single variable, Appendix A) by removing the variable, and PCA was re-run (steps 4B and 4C, Figure 3.1). When unreliable components could not be made reliable, they were not considered in the interpretation of results.

Components with less than three variables are considered weak and unstable (Costello & Osborne, 2005) and do not meet the criteria of replicability (Gorsuch, 1983). Variables in these components have a higher probability to be grouped by chance (Gorsuch, 1983). Components demonstrating this characteristic were thus not carried over to the interpretation (step 4D).

3.2.2.5 Comparisons

In step 5A (Figure 3.1), a quantitative comparison of the general configuration of the solutions (*i.e.* one per surface) was performed using two measures that together represent a "geometrically important and meaningful" way of comparing components (Rummel, 1970, p.462): a similarity coefficient (SC) (Harman, 1967) and a coefficient of congruence (CC) (Tucker, 1951) (see Appendix A). These coefficients measure different underlying characteristics of the PCA solution. SC quantifies similarity of components in terms of magnitude and direction of the components, while CC looks at pattern and magnitude of the components through the shared variance of the components (Cattell, 1978).

3.2.2.6 Interpretation

Rummel (1970) identified five elements to look at when interpreting and comparing PCA solutions: number of components, variance, complexity, communality, and configuration. The number of components is an indication of the convergence level of the dataset to a certain dimensionality. The variance indicates the relative importance of each component and helps investigate if a component's importance is specific to a particular surface or is generalizable (Gorsuch, 1983). Communality is a measure of the variance of a variable that is accounted for by the sum of the components. It can help identifying which variables are unique: if the communality of a particular variable is very low, this variable does not match very well the solution and needs to be further explored (see Section 3.2.3). The configuration corresponds to the pattern and magnitude of loadings of variables on a solution. Finally, the complexity represents the behaviour of a variable, *i.e.* if it shifts from a component to another in a different solution. The PCA results are presented below and interpreted according to these five elements.

3.2.3 Covariation assessment: variable inflation factor and mutual information

Two drawbacks of PCA are that it does not account for non-linear relationships and that if a variable does not covary with any others (potentially being unique), it will not load on any component. To address these issues, two independent stepwise measures of covariation were computed in the statistical software R v. 3.1.1: the Variable Inflation Factor (VIF) and Mutual Information (MI) (step 5B in Figure 3.1; Appendix A). The VIF is one of the most commonly used methods to detect covariation due to its simplicity (Belsey *et al.*, 2004) and is recommended over other methods by Dormann *et al.* (2013) to assess covariation in datasets. MI is a measure of co-expression widely used in information theory that has the capacity of detecting non-linear relationships and is often used to generalize the correlation structure of a dataset through a Mutual Information Matrix (MIM) (Steuer *et al.*, 2002; Song *et al.*, 2012). For each measure, the variables were ranked from least covarying to most covarying and an average of the two rankings was made (Appendix A). This average ranking was used to confirm PCA results and explore the uniqueness of the variables that did not load on any component, had a very low amount of variance accounted for by the sum of the components (*i.e.* low communality), or were not considered in the PCA because of their low cardinality.

3.3 Results

3.3.1 Terrain Attributes and Software

Of the 230 computed terrain attributes, 48 were found to be identical to another one and were removed from further analysis (Appendix B). Eight other terrain attributes were not considered for the PCA analyses because of their low cardinality (Appendix C). SAGA GIS offered the most terrain attributes (n=96), while Diva-GIS and Idrisi Selva offered the least (n=7) (Table 3.1). However, more terrain attributes computed from a software package did not necessarily result in more of them being useful in describing surface variability. For instance, the 13 attributes computed using Quantum GIS were all found in the final PCA solutions of the eight least complex surfaces (A0 to A7). LandSerf was the software that had the lowest percentage of terrain attributes reaching the final solutions, with an average of 55% (Appendix C).

3.3.2 Principal Component Analysis

3.3.2.1 Number of Components, Variance and Communality

A general pattern did not emerge from the number of components extracted with regards to surface complexity (Appendix C). A slight decreasing trend of total variance accounted for could be observed with increasing surface complexity (Appendix C). The solution corresponding to the lowest complexity surface accounted for 91.67% of the topographic structure, while the solution of the highest complexity surface accounted for 86.56%. A minimum was reached with surface A6 (85.35%). The average percentage of variance was 88.91%. When only considering reliable components, the percentage of variance ranged from 80.88% (A7) to 87.37% (A0), with an average of 84.34%.

Information on the communalities of each variable for each PCA solution can be found in Appendix C. 72% of the terrain attributes showed a negative trend where communality decreased with increasing surface complexity, which is also shown by the decreasing proportion of high communality terrain attributes as complexity increases (Appendix C). Over 90% of the variables of the least complex surfaces (A0 to A3) have high communalities. Surfaces A4 and A8 have the highest percentage of variables with low communalities, with 2.7% and 2.4% respectively.

3.3.2.2 Configuration

Several terrain attributes were found to be consistently grouped on the same components across solutions. These groups are presented in Figure 3.3 and their category, or interpretation, was based on the variables with the highest loadings in each of them

(McGarigal *et al.*, 2000). The locations of these groups on the different solutions are represented in the configuration summary presented in Figure 3.4.

A visual interpretation of Figure 3.4 allows identifying the groups that loaded on the same component for more than one surface, thus indicating similarity between the solutions. This similarity was confirmed quantitatively by the SC and CC values, measured on reliable components of consecutive solutions: the pairs of solutions A0-A1, A2-A3, A4-A5, and A5-A6 are all considered identical, with a one to one match of their components. Three other pairs of solutions were considered very similar, with only one inversion in components: A1 and A2 are the same except for the inversion of the second and fifth components, A6 and A7 show the same configuration but the inversion of the statistical parameters and slope attributes (fifth and sixth components), and A7 and A8 would be considered identical if it were not for the inversion of the fourth and fifth components.

MAJOR GROUPS	Group 1	Group 2	Group 3A	Group 3B	Group 4	Group 5A	Group 5B
Category	Curvatures and Relative Position	Rugosity Indices	Orientation (Aspect)		Local Statistical Attributes	Slope	
Terrain Attributes	Topographic Position Index, General Curvature, Deviation from Mean Value, Plan Curvature, Convergence Index, Minimum Curvature, Maximum Curvature, Mean Curvature, Morphometric Position Index	Terrain Ruggedness Index, Surface Ratio, Standard Deviation, Melton Ruggedness Index, Roughness, Range	Easterness	Northerness	Mean, Median, Minimum, Maximum	Slopes (Horn's Method)	Slopes (4-Cell Method, Zevenbergen & Thorne Method)

MINOR GROUPS	Group 5C	Group 5D	Group 5E	Group 6A	Group 6B	Group 7	Group 8
Category	Slope			Vector Ruggedness Measures		Rugosity Indices	
Terrain Attributes	Slopes (Maximum Slope Algorithm and Maximum Triangle Slope Algorithm)	Slopes (Quadratic Surfaces, Least-Square Fit Algorithm)	Other Slope Algorithms	Vector Ruggedness Measures (Sappington's Method)	Vector Ruggedness Measures (SAGA GIS Algorithms)	Planimetric-to-Surface Ratio, Ruggedness Index	Representativeness, Residual at Centre

MINOR GROUPS	Group 9A	Group 9B	Group 9C	Group 9D	Group 9E	Group 9F	Group 9G
Category	Curvatures						
Terrain Attributes	Profile Curvatures (TNT mips Algorithms)	Plan Curvatures (TNT mips Algorithms)	Profile & Plan Curvatures (SAGA GIS Algorithms)	Curvatures (DEM Surface Tools' Algorithms)	Curvatures (WhiteBox GIS Algorithms)	Curvatures (Haralick's Algorithm)	Curvatures (uDig Algorithms)

Figure 3.3: Main groups of multicollinear terrain attributes that consistently loaded together on the different components.

<table>
<tr><th rowspan="2">COMPONENTS</th><th colspan="9">SURFACES' PCA SOLUTIONS</th><th rowspan="2">COMPONENTS</th></tr>
<tr><th>A0</th><th>A1</th><th>A2</th><th>A3</th><th>A4</th><th>A5</th><th>A6</th><th>A7</th><th>A8</th></tr>
<tr><td rowspan="4">1</td><td colspan="9">Group 1</td><td rowspan="4">1</td></tr>
<tr><td colspan="4">Cross-sectional Curvature</td><td colspan="5"></td></tr>
<tr><td></td><td colspan="8">Profile Curvature</td></tr>
<tr><td colspan="2">Tangential Curvature</td><td colspan="7"></td></tr>
<tr><td rowspan="2">2</td><td colspan="2">Groups 5A, 5B, 5D, 5E & 6A, Fractal Dimension, Surface Roughness Index</td><td colspan="2" rowspan="2">Groups 2 & 7, Variance, Coefficient of Variation</td><td colspan="5" rowspan="2">Group 3A</td><td rowspan="2">2</td></tr>
<tr><td>Group 5C</td><td></td></tr>
<tr><td>3</td><td colspan="4">Group 3A</td><td colspan="5">Group 3B</td><td>3</td></tr>
<tr><td rowspan="2">4</td><td colspan="4" rowspan="2">Group 3B</td><td colspan="4">Groups 2 & 7, Coefficient of Variation</td><td rowspan="2">Group 4</td><td rowspan="2">4</td></tr>
<tr><td colspan="2">Variance</td><td colspan="2"></td></tr>
<tr><td rowspan="2">5</td><td colspan="2" rowspan="2">Groups 2 & 7, Variance</td><td colspan="5">Groups 5A & 5E</td><td rowspan="2">Group 4</td><td rowspan="2">Group 2, Coefficient of Variation</td><td rowspan="2">5</td></tr>
<tr><td>Group 5B</td><td colspan="3"></td><td>Group 5D</td></tr>
<tr><td>6</td><td colspan="7">Group 4</td><td colspan="2">Group 5A</td><td>6</td></tr>
<tr><td>7</td><td colspan="4">Group 9A</td><td colspan="5">Group 5B, Fractal Dimension</td><td>7</td></tr>
<tr><td>8</td><td colspan="4">Group 9B</td><td colspan="5">Group 9A</td><td>8</td></tr>
<tr><td>9</td><td colspan="3">Groups 9E & 9G</td><td>Group 5B</td><td>Group 9B</td><td>Groups 9E & 9G</td><td colspan="3">Group 9B</td><td>9</td></tr>
<tr><td>10</td><td colspan="4">Group 6B</td><td>Groups 9E & 9G</td><td></td><td colspan="3">Groups 6A & 6B</td><td>10</td></tr>
<tr><td>11</td><td colspan="3">Total Curvature, Slope Variability</td><td>Groups 9E & 9G</td><td>Group 6B</td><td></td><td><u>Group 5C</u></td><td><u>Group 5C</u></td><td>Group 9F</td><td>11</td></tr>
<tr><td>12</td><td colspan="2"></td><td>Group 9C</td><td>Total Curvature</td><td>Group 9C</td><td></td><td>Groups 9E & 9G</td><td>Group 5D</td><td>Groups 9E & 9G</td><td>12</td></tr>
<tr><td>13</td><td colspan="2"></td><td>Group 8</td><td>Group 9C</td><td><u>Group 5C</u></td><td colspan="2"></td><td><u>Group 9F</u></td><td>Group 5D</td><td>13</td></tr>
<tr><td>14</td><td colspan="3"></td><td>Group 8</td><td>Group 5D</td><td colspan="2"></td><td><u>Group 6A</u></td><td><u>Group 9D</u></td><td>14</td></tr>
<tr><td>15</td><td colspan="4"></td><td>Group 8</td><td colspan="2"></td><td><u>Group 9D</u></td><td></td><td>15</td></tr>
<tr><td>16</td><td colspan="7"></td><td><u>Group 9E</u></td><td></td><td>16</td></tr>
</table>

Figure 3.4: Summary of the PCA solutions' configuration. The different groups are detailed in Figure 3.3. Characters in italic represent the unreliable components as assessed by Cronbach's α, and underlined characters indicate components with less than three variables.

Only one pair of consecutive solutions, A3-A4, is considered significantly different. Only three out of ten pairs of components were found to be matching: the first, fifth and sixth components. The overall mismatch is caused by the decreasing relative importance of the two groups of rugosity indices (groups 2 and 7), and the increased relative importance of the second group of slope algorithms (group 5B).

3.3.2.3 Complexity of Variables

Figure 3.4 shows how the complexity of the major groups of terrain attributes presented in Figure 3.3 vary with surface complexity. Only Group 1 (interpreted as curvatures and relative position) does not vary in complexity across solutions. In general, the aspect (Groups 3A-3B) and local statistical attributes (Group 4) shift to a higher component as surface complexity increases. On the opposite, slopes (Groups 5A-5B) shift to a lower component as surface complexity increases. Group 2 does not demonstrate any particular pattern.

Table 3.2 presents a summary of the solutions according to variable's behaviour (*e.g.* complex or marker variables), and Appendix C helps assessing changes in behaviour with changing surface complexity. Some of the variables that are not included in the major groups of terrain attributes previously identified show particular patterns. For instance, a small number of northerness measures (ID104, ID105, ID107, ID108, ID109) are marker variables in solutions from low complexity surfaces, and become complex variables as surface complexity increases. A similar pattern is observed for measures of total curvature. On the other hand, some terrain attributes, such as a few measures of profile curvature (ID135, ID141, ID142, ID144) are complex variables for low complexity surfaces and become marker variables as surface complexity increases.

Table 3.2: Percentage of the 230 variables that formed each solution or were removed during the iterative PCA. Complex variables that loaded equally on more than one component were removed, while complex variables that loaded more strongly on one component were kept (Appendix A).

	In the Final Solution		Removed during Iterations		
	Marker Variables	Complex Variables		Reliability	Cardinality
A0	66.1	17.4	13.0	0.0	3.5
A1	63.0	19.6	13.9	0.0	3.5
A2	62.2	19.1	14.8	0.4	3.5
A3	63.5	17.0	16.1	0.0	3.5
A4	65.2	15.7	15.2	0.4	3.5
A5	63.9	3.0	29.1	0.4	3.5
A6	70.4	7.0	18.3	0.9	3.5
A7	63.9	10.9	20.9	0.9	3.5
A8	62.2	8.3	24.8	1.3	3.5

3.3.3 Variable Inflation Factor and Mutual Information

The average rank of each variable according to VIF and MI is presented in Appendix C. 114 variables became more multicollinear (*i.e.* ranked lower) with increasing surface complexity, while 116 variables became less multicollinear with increasing surface complexity. Variables from Group 1 were generally losing some positions in the ranking, while variables from Group 2 and 7 were gaining ranks.

Of the eight variables whose cardinality values were too low to enter the PCA, three of them consistently ranked in the top 50 least collinear variables: center versus neighbours variability (ID2), and one measure of easterness (ID42) and northerness (ID101). Of the variables that had low communality for any surface, mean of residuals (ID70), representativeness (ID158), standard deviation of slope (ID195), and one measure of profile curvature (ID146), plan curvature (ID120), tangential curvature (ID204), easterness (ID42) and northerness (ID104) were consistently falling in the top 50 least collinear variables.

3.4 Discussion

3.4.1 Surface Complexity and Importance of Terrain Attributes

Results indicate that the relative importance of terrain attributes in capturing relevant information on terrain characteristics varies with surface complexity. For instance, the relative importance of slope diminishes as surfaces become more complex, while local statistical attributes importance increases. Also, the importance of differentiating the algorithms used to compute a terrain attribute seems to vary with surface complexity. For instance, most slope algorithms are grouped into one component for the least complex surfaces, indicating that they are all highly correlated, but get separated across several components as surface complexity increases, indicating that some of them are not correlated anymore. An opposite trend can be observed for vector ruggedness measures (VRM) algorithms. For low complexity surfaces, VRMs computed with Sappington's (2007) method are relatively important (they load on the second component), while the VRMs computed with SAGA GIS' algorithms load on the tenth component. VRMs however all load together on the tenth component of the most complex surfaces, with no difference between the algorithms used.

Also, both the quantitative (SC and CC) and qualitative (Figure 3.4) comparative assessments of the solutions' configuration allowed finding an important break when surface complexity reaches a fractal dimension of 2.40 (surface A4). This is likely to correspond to a threshold where a significant number of terrain attributes' behaviour changes in relation to terrain characteristics. This threshold is located somewhere between a fractal dimension of 2.30 and 2.40, representing approximately the middle of the range of complexities tested.

3.4.2 Terrain Attributes, Algorithms and Software

The diversity and amount of variables loading on the same components demonstrate high level of covariation amongst terrain attributes. For low complexity surfaces for instance, measures of slope covary with local fractal dimension, measures of roughness, and VRM, attributes often used together in ecological studies (Harris & Baker, 2012). Group 1 also shows high variable diversity (Figure 3.3). About one third of the attributes were found to correlate with several components (*i.e.* the complex variables), thus being very collinear and holding low potential for ecological applications. For instance, measures of longitudinal curvature were complex variables for all surfaces, and their high level of covariation was confirmed by their VIF-MI ranking. Of the 230 computed attributes, 79% were found to be unique (with no identical attribute surfaces), confirming and extending the conclusions from Dolan & Lucieer (2014) and Jones (1998a; 1998b) that same terrain attributes derived using different algorithms or software can give different results.

The extent to which different slope algorithms are divided into sub-groups as surface complexity increases and how they rank on the different components is particularly interesting. Skidmore (1989) argued that Zevenbergen & Thorne's (1987) method was better than Horn's (1981) even if results for the latter were not very different from those of the former. Our results demonstrate that while this is the case for low complexity surfaces, Horn's method captures more variance in more complex surfaces. Skidmore's (1989) conclusions may then be limited by, and dependent on, the surfaces that were used. On the other hand, Hodgson (1995) argued that Zevenbergen & Thorne's (1987) method worked best on rough surfaces, while Horn's (1981) worked best on smooth surfaces, without further distinction on what are rough and smooth surfaces. Our results suggest that Horn's (1981) method (Group 5A) captures more variance than others for rougher surfaces (fractal dimension ≥ 2.30) and that there are no differences for low complexity, or smoother, surfaces (fractal dimension ≤ 2.20). Jones (1998b) found that the performance of slope algorithms depended on the relative importance of random noise versus mean smooth elevation difference: in general the 4-cell method and Horn's method performed better than others, but when a surface was very noisy, Sharpnack & Akin's (1969) method, from which Zevenbergen & Thorne (1987) method is derived, performs better.

Some ambiguity in the names and types of terrain attributes was found during the analysis, confirming previous observations (Bishop *et al.*, 2012). For instance, topographic position index (TPI) measures from SAGA GIS (ID214, ID217) were found identical to "difference from mean values" (ID23, ID26) attributes. Some attributes identified simply as "curvature" (ID14, ID16) were found to be duplicates of measures of profile (ID143, ID145) and plan curvature (ID123, ID125); profile and plan curvatures are however supposed to describe different terrain characteristics. Solely using "curvature" to describe a measure could mislead a user into believing that it is a measure of general or total curvature. Also, different types of curvatures were often found to be inter-correlated when computed with the same software, but uncorrelated to those generated from other software. For instance, Groups 9C, 9D, 9E, 9F and 9G all combine correlated measures of curvature that should theoretically be uncorrelated, but that are generated from the same software. Finally, several terrain attributes from SAGA GIS computed from three different methods (no distance weighting, inverse of the distance, and squared inverse of the distance) often gave identical results. Since these measures are directly dependent on the window size used (distance-based algorithms), it is possible that these attributes would give different outcomes if a bigger window size would be used. Further work will be necessary to assess the behavior of terrain attributes with changing surface complexity using different window sizes.

3.4.3 Suitable Subset of Terrain Attributes

The five major groups of terrain attributes presented in Figure 3.3 were always the ones accounting for the most variance, regardless of terrain complexity. These groups alone accounted for an average of 75% of the overall topographic structure (Appendix C). Based on psychometric methods, a suitable subset of variables in a PCA consists of selecting one variable from each component (Gorsuch, 1983). Since these five groups always loaded on different but the highest components for all levels of topographic complexity, a suitable subset of terrain attributes would include one variable from each of the groups 1, 2, 3A, 3B, 4, and 5A.

While any attributes from each of these groups could be a good choice, we recommend a specific combination that, in addition to reducing covariation and redundancy, also reduces ambiguity: relative difference to mean value (Group 1), local standard deviation (Group 2), easterness and northerness (Groups 3A-3B), local mean (Group 4), and a measure of slope preferably computed with Horn's method (Group 5A, Table B.2 in Appendix B). Relative difference to mean value is a measure of relative position, local standard deviation is a measure of rugosity, and easterness and northerness are measures of orientation derived from aspect. Local mean may however not be required if users include elevation or depth in the analysis, as the local mean will be highly correlated with the input DTM. However, local mean could potentially be more reliable than the initial DTM if a surface is noisy, as using the mean may filter out some of the noise. This recommendation of six attributes increases replicability and generality as it selects terrain attributes that are easily computed from any software over terrain attributes that are only available in some of them. For instance, we recommend using local standard deviation in Group 2 over terrain ruggedness index or roughness as there is no ambiguity on how to compute standard deviation and all software have focal statistics tools that can compute it. To help potential users of this method, we provide with this paper a toolbox developed for ArcGIS, named TASSE (Terrain Attribute Selection for Spatial Ecology), that automatically generates the six proposed terrain attributes (Lecours, 2015).

While using the proposed terrain attributes helps optimize the information extracted from the terrain, it does not necessarily mean that all of the proposed attributes will be useful for a given ecological application (Lecours *et al.*, submitted). For instance, using them in species distribution modelling exercises will not necessarily result in these six terrain attributes being drivers or surrogates of species distribution. Rather, it means that most of the terrain properties, or topographic structure, and the variation in these properties will be accounted for when performing the analysis and modelling. An analyst that includes such selection can be assured that most of the topographic structure is considered in the analysis, and can then focus on the integration of the selection with other environmental data (*e.g.* remotely sensed data, climate data, oceanographic data, vegetation data). Also, if slope and rugosity are spatially found in the same regions of a study area, the slope and standard deviation terrain attributes would be correlated, despite extracting different information from the surface.

It is important to note that this combination of attributes and the alternatives (if selecting one different attribute from each group) are mainly based on the PCA analysis. However, the exploration of communality combined with the covariation assessment using VIF and MI allowed identifying unique variables that do not correlate highly with any other but could potentially give information on unique terrain characteristics. Some of these variables may improve the proposed selection of terrain attributes, and we recommend further exploration of these unique attributes in future work.

3.5 Conclusion

If the use of terrain attributes in environmental studies is now common, selecting an appropriate set of terrain attributes is a challenge rarely approached carefully in these studies. This can have potential consequences that can invalidate the analyses and not make the best use of terrain data. This study conducted an extensive analysis to suggest a suitable selection of terrain attributes that optimizes the information derived from DTMs.

Three methods were used to iteratively assess the relationships between 230 terrain attribute surfaces computed on nine artificial surfaces of different complexity using 11 different software packages. Results confirm that (1) different algorithms computing a same terrain attribute (*e.g.* slope, curvature) can produce different outcomes, that (2) terrain attributes are highly multicollinear, that (3) there is some ambiguity in the denomination of terrain attributes, and that (4) their selection for any application needs to be carefully performed. Our conclusions highlight the importance to explicitly report the software, algorithms, and parameters used to generate terrain attributes in ecological studies to allow careful interpretation of the results and comparison between studies. We also encourage software and tools developers to be explicit in their documentation or metadata about the methods or algorithms used by their products.

Based on our analysis, we recommend the use of six terrain attributes that optimize the topographic structure accounted for when performing environmental studies: (1) relative difference to mean value (a measure of relative position), (2) local standard deviation (a measure of rugosity), (3) easterness and (4) northerness (measures of orientation), (5) local mean, and (6) slope (preferably computed with Horn's method). The proposed selection reduces redundancy, covariation and ambiguity, and improves generality and replicability, and can be applied across a wide range of terrain complexity. While the six proposed attributes can easily be computed using any GIS package, an ArcGIS toolbox was provided with this paper to easily generate these attributes (Lecours, 2015). Our work has also identified unique terrain attributes that were not considered by the PCA analysis but that showed potential to represent different characteristics of a terrain and that needs to be further explored.

Our recommendations provide an operational framework to any users willing to incorporate topography or bathymetry and their derivatives (*i.e.* terrain attributes) in environmental models and analyses. While the six proposed attributes may not necessarily be useful for all applications and should be tested separately, for instance as effective surrogates, their combination ensures that the amount of topographic structure accounted for is optimized in the analysis. The proposed operational framework can help users make more robust analyses and bridge geomorphometry with disciplines like ecology, biogeography, habitat mapping and distribution modelling. An application of our recommended terrain attributes and their comparison to other subsets is presented in a real ecological application, namely a marine benthic habitat mapping exercise, in Lecours *et al.* (submitted).

3.6 Literature Cited

Bouchet, P.J., Meeuwig, J.J., Kent, C.P.S., Letessier, T.B., & Jenner, C.K. (2015) Topographic determinants of mobile vertebrate predator hotspots: current knowledge and future directions. *Biological Reviews*, 90:699-728.

Belsey, D.A., Kuh, E., & Welsch, R.E. (2004) *Regression diagnostics: identifying influential data and sources of multicollinearity*. Wiley, Hoboken.

Bhattacharyya, H., (1981) Theory and methods of factor analysis and principal components. In: Capen, D.E. (Ed.) *The use of multivariate statistics in studies on wildlife habitat*, U.S. Forest Service General Technical Report RM-87, pp. 72-79.

Bishop, M.P., & Shroder Jr., J.F. (2004) GIScience and mountain geomorphology: overview, feedbacks, and research directions. In: Bishop, M.P., & Shroder, J.F. (Eds) *Geographic Information Science and Mountain Geomorphology*, Springer-Praxis, Chichester, pp.1-31.

Bishop, M.P., James, L.A., Shroder Jr., J.F., & Walsh, S.J. (2012) Geospatial technologies and digital geomorphological mapping: concepts, issues and research. *Geomorphology*, 137:5-26.

Brown, C.J., Smith, S.J., Lawton, P., & Anderson, J.T. (2011) Benthic habitat mapping: a review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuarine, Coastal and Shelf Science*, 92:502-520.

Cattell, R.B. (1978) *The scientific use of factor analysis in behavioral and life sciences*. Plenum Press, New York.

Chipperfield, J.D., Dytham, C., & Hovestadt, T. (2011) An updated algorithm for the generation of neutral landscapes by spectral synthesis. *PLos ONE*, 6:e17040.

Costello, A.B., & Osborne, J. (2005) Best practices in explanatory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10:1-9.

Cronbach, L.J. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika*, 16:297-334.

Dimri, V.P. (2005) Fractals in geophysics and seismology – an introduction. In: Dimri, V.P. (Ed.) *Fractal Behaviour of the Earth System*, Springer-Verlag, Berlin, pp. 1-22.

Dolan, M.F.J., & Lucieer, V.L. (2014) Variation and uncertainty in bathymetric slope calculations using geographic information systems. *Marine Geodesy*, 37:187-219.

Dormann, *CF.*, Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., García Marquéz, J.R., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., & Lautenbach, S. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36:27-46.

Elith, J., & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40:677-697.

Florinsky, I.V., Eilers, R.G., Manning, G.R., & Fuller, L.G. (2002) Prediction of soil properties by digital terrain modelling. *Environmental Modelling and Software*, 17:295-311.

Franklin, S.B., Gibson, D.J., Robertson, P.A., Pohlmann, J.T., & Fralish, J.S. (1995) Parallel analysis: a method for determining significant principal components. *Journal of Vegetation Science*, 6:99-106.

Gorsuch, R.L. (1983) *Factor analysis (second edition)*. Erlbaum, Hillsdale.

Graham, M.H. (2003) Confronting multicollinearity in ecological multiple regression. *Ecology*, 84:2809-2815.

Halley, J.M., Hartley, S., Kallimanis, A.S., Kunin, W.E., Lennon, J.J., & Sgardelis, S.P. (2004) Uses and abuses of fractal methodology in ecology. *Ecology Letters*, 7:254-271.

Han, W., Di, L., Zhao, P., & Shao, Y. (2012) DEM Explorer: an online interoperable DEM data sharing and analysis system. *Environmental Modelling and Software*, 38:101-107.

Harman, H.H. (1967) *Modern factor analysis (revised edition)*. University of Chicago Press, Chicago.

Harris, P.T., & Baker, E.K. (2012) *Seafloor geomorphology as benthic habitat: GeoHab atlas of seafloor geomorphic features and benthic habitats*. Elsevier, Amsterdam.

Hengl, T., & Reuter, H.I. (2009) *Geomorphometry: concepts, software, applications*. Developments in Soil Science, 33, Elsevier, Amsterdam.

Henson, R.K., & Roberts, J.K. (2006) Use of explanatory factor analysis in published research: common errors and some comment on improved practice. *Organizational Research Methods*, 7:191-205.

Hijmans, R.J. (2012) Cross-validation of species distribution models: removing spatial bias and calibration with a null model. *Ecology*, 36:27-46.

Hodgson, M.E. (1995) What cell size does the computed slope/aspect angle represent? *Photogrammetric Engineering and Remote Sensing*, 61:513-517.

Hofierka, J., Mitášová, H., & Neteler, M. (2009) Geomorphometry in GRASS GIS. In Hengl, T., & Reuter, H.I. (Eds) *Geomorphometry: concepts, software, applications*, Elsevier Science, Amsterdam, pp. 387-410.

Horn, B.K.P. (1981) Hill shading and the reflectance map. *Proceedings of the IEEE*, 69:14-47.

Horn, J.L. (1965) A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30:179-185.

Jones, K.H. (1998a). A comparison of algorithms used to compute hill slope as a property of the DEM. *Computers and Geosciences*, 24:315-323.

Jones, K.H. (1998b) A comparison of two approaches to ranking algorithms used to compute hill slopes. *GeoInformatica*, 2:235-256.

Kaiser, H.F. (1958) The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23:187-200.

Keitt, T.H. (2000) Spectral representation of neutral landscapes. *Landscape Ecology*, 15:479-493.

King, J.R., & Jackson, D.A. (1999) Variable selection in large environmental data sets using principal components analysis. *Environmetrics*, 10:67-77.

Kinsey-Henderson, A.E., & Wilkinson, S.N. (2013) Evaluating Shuttle radar and interpolated DEMs for slope gradient and soil erosion estimation in low relief terrain. *Environmental Modelling and Software*, 40:128-139.

Klingseisen, B., Metternicht, G., & Paulus G. (2008) Geomorphometric landscape analysis using a semi-automated GIS-approach. *Environmental Modelling and Software*, 23:109-121.

Lacroix, M.P., Martz, L.W., Kite, G.W., & Grabrecht, J. (2002) Using digital terrain analysis modeling techniques for the parametrization of a hydrologic model. *Environmental Modelling and Software*, 17:127-136.

Lecours, V. (2015) *Terrain Attribute Selection for Spatial Ecology (TASSE), v. 1.0*. URL www.marinegis.com.

Lecours, V., Brown, C.J., Devillers, R., Lucieer, V.L., & Edinger, E.N. (submitted). Comparing selections of environmental variables for ecological studies: a focus on terrain attributes.

Lloyd, C.D. (2014) *Exploring spatial scale in geography*. Wiley, United States.

McGarigal, K., Cushman, S.A., & Stafford, S.G. (2000) *Multivariate statistics for wildlife and ecology research*. Springer-Verlag, New York.

O'Connor, B.P. (2000) Using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32:396-402.

O'Connor, B.P. (2015) *SPSS, SAS, MATLAB, and R programs for determining the number of components and factors using parallel analysis and Velicer's MAP test*. URL https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html [accessed 21 January 2015]

Olaya, V. (2009) Basic land-surface parameters. In: Hengl, T., & Reuter, H.I. (Eds) *Geomorphometry: concepts, software, applications*, Elsevier Science, Amsterdam, pp. 141-169.

Peitgen, H.O., & Saupe, D. (1988) *The science of fractal images*. Springer-Verlag, New York.

Pfeifer, P. (1984) Fractal dimension as working tool for surface-roughness problems. *Applications of Surface Science*, 18:146-164.

Pike, R.J. (1995) Geomorphometry: progress, practice, and prospect. *Zeitschrift für Geomorphologie*, 101:221-238.

Pittman, S.J., Costa, B.M., & Battista, T.A. (2009) Using lidar bathymetric and boosted regression trees to predict the diversity and abundance of fish and corals. *Journal of Coastal Research*, 53:27-38.

Raaflaub, L.D., & Collins, M.J. (2006) The effect of error in gridded digital elevation models on the estimation of topographic parameters. *Environmental Modelling and Software*, 21:710-732,

Rigol-Sanchez, J.P., Stuart, N., & Pulido-Bosch, A. (2015) ArcGeomorphometry: A toolbox for geomorphometric characterisation of DEMs in the ArcGIS environment. *Computers & Geosciences*, 85:155-163.

Rooper, C.N., & Zimmermann, M. (2007) A bottom-up methodology for integrating underwater video and acoustic mapping for seafloor substrate classification. *Continental Shelf Research*, 27:947-957.

Rummel, R.J. (1970) *Applied factor analysis*. Northwestern University Press, Evanston.

Sappington, J.M., Longshore, K.M., & Thompson, D.B. (2007) Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave desert. *Journal of Wildlife management*, 71:1419-1426.

Schwanghart, W., & Heckmann, T. (2012) Fuzzy delineation of drainage basins through probabilistic interpretation of diverging flow algorithms. *Environmental Modelling and Software*, 33:106-113.

Sharpnack, D.A., & Akin, G. (1969) An algorithm for computing slope and aspect from elevations. *Photogrammetric Engineering*, 35:247-248.

Skidmore, A.K. (1989) A comparison of techniques for calculating gradient and aspect from a gridded digital elevation model. *International Journal of Geographical Information Systems*, 3-4:323-334.

Song, L., Langfelder, P., & Horvath, S. (2012) Comparison of co-expression measures: mutual information, correlation, and model based indices. *BMC Bioinformatics*, 13:328.

Steuer, R., Kurths, J., Daub, C.O., Weise, J., & Selbig, J. (2002) The mutual information: detecting and evaluating dependencies between variables. *BMC Bioinformatics*, 18:S231-S240.

Tabachnick, B.G., & Fidell, L.S. (2014) *Using multivariate statistics, sixth edition*. Pearson Education Limited.

Thompson, B., & Daniel, L.G. (1996) Factor analytic evidence for the construct validity of scores: a historical overview and some guidelines. *Educational and Psychological Measurement*, 56:197-208.

Tucker, L.R. (1951) *A method for synthesis of factor analysis studies*. Personnel Research Section Report 984, Department of the Army, Washington D.C.

Velicer, W.F. (1976) Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41:321-327.

Velicer, W.F., Eaton, C.A. & Fava, J.L. (2000) Construct explication through factor or component analysis: a review and evaluation of alternative procedures for determining the number of factors or components. In: Goffin, R.D., & Helmes, E. (Eds) *Problems and solution in human assessment*, Kluwer, Boston, pp. 41-71.

Vierod, A.D.T., Guinotte, J.M., & Davies, A.J. (2014) Predicting the distribution of vulnerable marine ecosystems in the deep sea using presence-background models. *Deep-Sea Research II*, 99:6-18.

Wiersma, Y.F., Huettmann, F. & Drew, C.A. (2011) Introduction. Landscape modeling of species and their habitats: history, uncertainty, and complexity. In: Drew, C.A., Wiersma, Y.F., & Huettmann, F. (Eds) *Predictive Species and Habitat Modeling in Landscape Ecology*, Springer, New York, pp. 1-6.

With, K.A., & King, A.W. (1997) The use and misuse of neutral landscape models in ecology. *Oikos*, 79:219-229.

Wood, J.M., Tataryn, D.J., & Gorsuch, R.L. (1996) Effects of under- and overextraction on principal axis factor analysis with Varimax rotation. *Psychological Methods*, 1:354-365.

Zawada, D.G., & Brock, J.C. (2009) A multiscale analysis of coral reef topographic complexity using lidar-derived bathymetry. *Journal of Coastal Research*, 53:6-15.

Zevenbergen, L.W., & Thorne, C.R. (1987) Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms*, 12:47-56.

Zhou, Q., & Zhu, A.X. (2013) The recent advancement in digital terrain analysis and modeling. *International Journal of Geographical Information Science*, 27:1269-1271.

Zwick, W.R., & Velicer, W.F. (1986) Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99:432-442.

4. Comparing Selections of Environmental Variables for Ecological Studies: a Focus on Terrain Attributes

4.1 Introduction

Due to the difficulty in sampling ecological data at sufficient spatial and temporal resolutions, many ecological studies rely on the use of surrogates to understand species distribution and ecological processes. Amongst commonly used surrogates, terrain attributes (*e.g.* slope, rugosity, aspect) derived from digital elevation (DEM) or bathymetric (DBM) models have proven their value in a broad range of terrestrial and marine ecological studies (Bouchet *et al.*, 2015). Such attributes can now be derived easily using tools available in most Geographic Information Systems (GIS). While tools are increasingly user-friendly, a lack of transparency in most software on the actual algorithms used (Dolan & Lucieer, 2014) can prevent users from making an informed decision on which tools to use. Also, terrain attributes sharing the same name but generated using different algorithms (*e.g.* slope) have been shown to produce different derivative surfaces (Jones, 1998; Dolan & Lucieer, 2014). Software developers and authors of published work are often not explicit on the methods they use to derive terrain attributes (*e.g.* algorithm or tool). This lack of information can possibly influence the analysis and interpretation of the resulting terrain attribute surfaces, and consequently the ecological application for which they are being used.

Choosing an appropriate selection of terrain attributes for specific ecological applications can be challenging, and users will often simply use the terrain attributes made available by the software they have access to or are familiar with, without further questioning if those attributes are the most appropriate ones for their study. In a related study, Lecours *et al.* (submitted) showed that many terrain attributes covary, which may cause potential problems for many statistical analyses. In an attempt to identify an optimal combination of terrain attributes to use in ecology, the authors recommended using six easily computable attributes for ecological studies that consider topography or bathymetry: (1) relative difference to mean value, which is a measure of relative position that can identify local peaks and valleys, (2) standard deviation, which is a measure of rugosity, (3) local mean, (4) slope, and (5-6) easterness and northerness, which together provide information on the orientation of the slope (*i.e.* aspect).

This article aims to describe the effects of subjectively selecting input variables for ecological applications. The specific objectives are (1) to compare the performance of Lecours *et al.* (submitted) recommended selection of terrain attributes to other selections in a real ecological context, (2) to demonstrate the relative importance of terrain morphology in aiding our understanding of ecological questions compared to other environmental variables, and (3) to report on the consequences of selecting different input variables on both the accuracy of habitat maps and the spatial distribution of the outputs.

4.2 Materials and Methods

Benthic habitat mapping is the act of mapping significantly distinct areas of the seafloor based on their physical, chemical and biological characteristics at particular spatial and temporal scales (Lecours *et al.*, 2015). The marine environment presents particular challenges in observing and sampling seafloor characteristics. However, developments in acoustic remote sensing technologies, specifically multibeam echosounders (MBES), now allow the collection of high-resolution remotely sensed data of the seafloor. Bathymetric measurements from MBES can be used to generate DBMs, from which terrain attributes can be derived (Lecours *et al.*, 2016). Additionally, MBES systems can also record acoustic reflectance (backscatter) data that provide information on seafloor properties (*e.g.* surficial geology, porosity). In combination, these attributes are commonly used for the production of benthic habitat maps.

4.2.1 Data

Datasets from Brown *et al.* (2012), covering 3,650 km^2 of German Bank, an area of the Canadian continental shelf off Nova Scotia (Figure 4.1), were used to address the objectives of this study. These data comprised a 50 m resolution DBM, 3,190 geo-referenced underwater images of the seabed visually classified into five bottom types (glacial till, silt and mud, rippled silt, rippled sand, reef), 4,816 geo-referenced sea scallop observations, and three backscatter data derivatives (Q1, Q2, Q3; Figure 4.1). Details on how the data were collected and processed can be found in Brown *et al.* (2012). According to Brown *et al.* (2012), the 50 m resolution was chosen to reduce the computation time and limitations for the classifications (*cf.* next sub-sections), and because that scale met a range of ocean management needs in the context of their study. For comparison with surfaces used in Lecours *et al.* (submitted), the fractal dimension, which is a quantitative representation of surface complexity, was measured over 10,000 m^2 areas of German Bank. Values ranged from 2.09 to 2.93, thus including regions of low (towards 2.00), moderate and high complexities (towards 3.00).

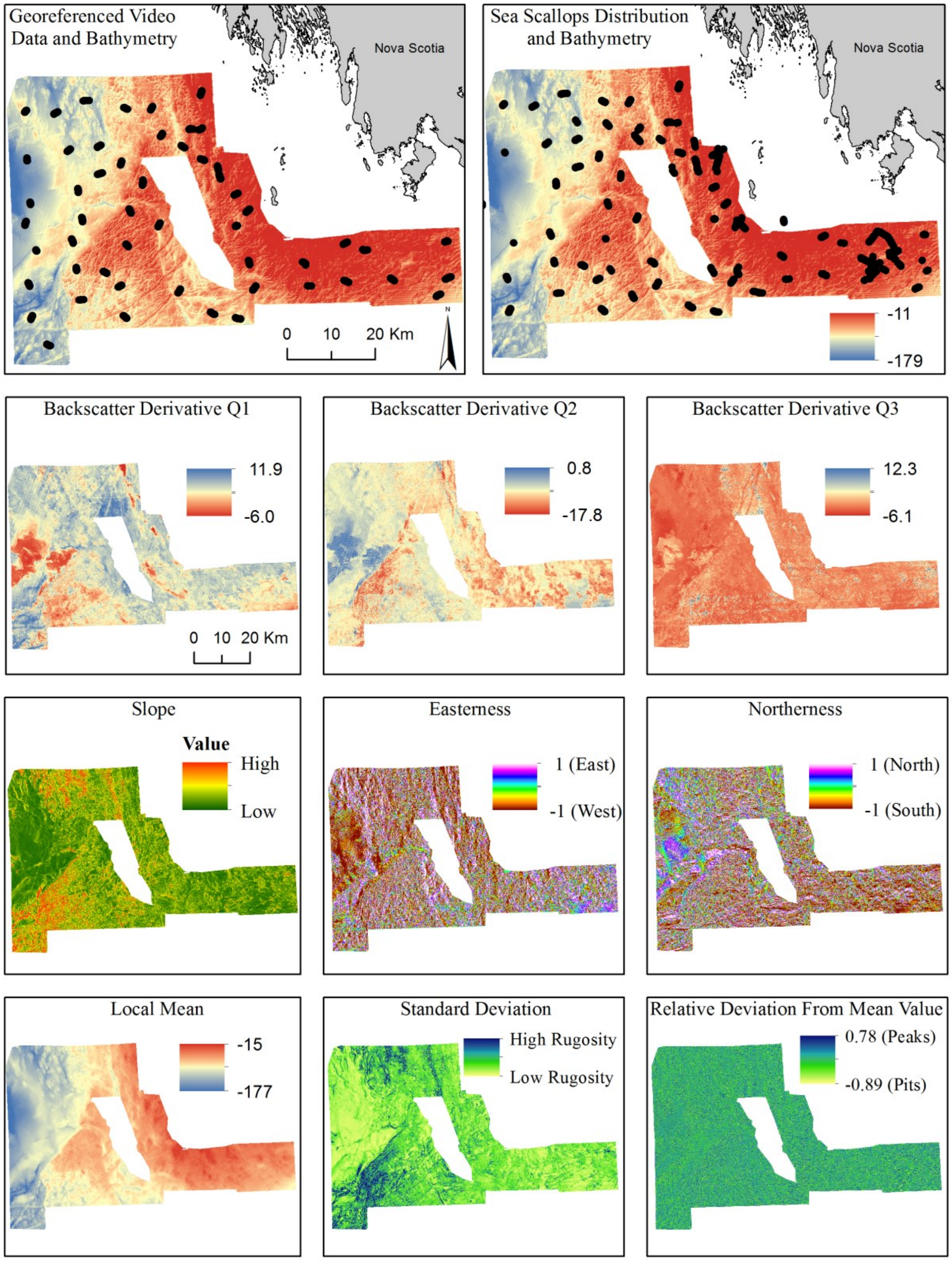

Figure 4.1: German Bank study area with some of the input variables used in this study: the ground-truth data for the bottom types, the sea scallops observations, the bathymetry, the three backscatter derivatives and the six terrain attributes from Selection 1.

Two types of habitat maps were generated from this dataset. First, biophysical classifications of the area were performed to create benthoscape maps (Zajac, 2008). This top-down, unsupervised approach segments the MBES derived data layers into a statistically optimum number of units which are then compared and subsequently recombined based on best match against independently classified *in situ* photographic data, classified into broad biophysical benthoscape classes. Using this approach, biophysical features can be delineated at a broader scale over the study area to generate a benthoscape map. Second, a bottom-up supervised approach was performed in which the *in situ* data were used to segment the environmental data to predict sea scallop (*Placopecten magellanicus*) habitat on German Bank. In both cases, different selections of terrain attributes were used in combination with other environmental data (seafloor bathymetry and backscatter derivatives) within the classification methodologies to test which combination of variables performed the best.

A total of 24 different terrain attributes were derived from the DBM and grouped into seven selections of six terrain attributes each (Table 4.1). The terrain attributes were selected from groups of variables that exhibited various behaviours during the statistical analyses performed in Lecours *et al.* (submitted) (see caption of Table 4.1). Selection 1 corresponds to our recommended selection of six terrain attributes. These terrain attributes were computed using the TASSE (Terrain Attribute Selection for Spatial Ecology) toolbox for ArcGIS (Lecours, 2015). Selections 2 to 7 were built to maximize variability and resemblance to Selection 1 (*e.g.*to avoid having two measures of slope or curvatures within one selection). Particular focus was also given to terrain attributes that were identified as potentially important by Lecours *et al.* (submitted).

Table 4.1: Selections of terrain attributes used to build the habitat maps and models. The ID numbers refer to Lecours *et al.* (submitted) and allow finding the software and parameters with which the attributes were generated. Marker variables correspond to important variables; whether they were found on strong components (Sel. 1) or weak components (Sel. 4) is linked to the amount of topographic structure they accounted for. Variables with low cardinality (Sel. 2) did not have many different values, thus limiting their ability to explain slight variations in terrain morphology. Complex variables (Sel. 3) correspond to redundant variables.

Selection 1 *Marker Variables on Strong Components*		**Selection 2** *Variables with Low Cardinality*		**Selection 3** *Complex Variables*	
ID31	Easterness	ID1	Bathymetric Position Index	ID70	Mean of Residuals*
ID67	Local Mean	ID2	Center vs Neighbor Variability*	ID116	Plan Curvature
ID90	Northerness	ID42	Easterness*	ID136	Profile Curvature
ID157	Relative Deviation from Mean Elevation	ID101	Northerness*	ID178	Slope
ID166	Slope	ID111	Percentile	ID201	Surface Roughness
ID190	Standard Deviation	ID143	Profile Curvature	ID221	Value Range
Selection 4 *Marker Variables on Weak Components*		**Selection 5** *Mix of Selections 1 and 2*		**Selection 6** *Mix of Selections 1 and 3*	
ID132	Plan Curvature	ID1	Bathymetric Position Index	ID70	Mean of Residuals*
ID153	Profile Curvature	ID2	Center vs. Neighbor variability*	ID178	Slope
ID158	Representativeness*	ID42	Easterness*	ID221	Value Range
ID188	Slope Variability	ID67	Local Mean	ID31	Easterness
ID219	Total Curvature	ID90	Northerness	ID90	Northerness
ID227	Vector Ruggedness Index	ID166	Slope	ID190	Standard Deviation
Selection 7 *Mix of Selections 1 and 4*		*Previously Identified as Potentially Important [4]			
ID158	Representativeness*				
ID188	Slope Variability	Scenario A: Each Selection used Alone (6 layers)			
ID227	Vector Ruggedness Index	Scenario B: Each Selection used with Depth (7 layers)			
ID67	Local Mean	Scenario C: Each Selection used with the Three Backscatter Derivatives (9 layers)			
ID31	Easterness	Scenario D: Each Selection used with Depth and the Three Backscatter Derivatives (10 layers)			
ID90	Northerness				

4.2.2 Unsupervised Classifications of Potential Habitat Types

A total of 29 benthoscape maps were built using the Modified k-Means unsupervised classification tool in Whitebox GAT v.3.2 “Iguazu”. This modified k-means algorithm is similar to the more common k-means or isocluster algorithms, but in this case the user does not have to subjectively select a number of desired classes. The algorithm starts by overestimating the number of classes, and then iteratively merges classes that have cluster centres close to each other, as defined by a user-defined threshold that relates to the minimum mapping unit. The resulting number of classes is thus statistically optimal and objectively achieved. To assess the relative importance of the different environmental variables and the consequences of using different input variables in habitat mapping, four scenarios were tested with each of the seven selections, resulting in 28 habitat maps. Maps were first created using each selection alone (six input layers), then adding the bathymetry (seven layers), the three backscatter derivatives (nine layers), and finally both the bathymetry and the backscatter derivatives (ten layers). In order to quantify the relative influence of terrain morphology in potential habitat characterization of German Bank, an additional habitat map was produced using only the backscatter derivatives and the bathymetry (four layers, not accounting for terrain morphology).

Following the method outlined in Brown *et al.* (2012), the resulting clusters for each classification were spatially compared to the 3,190 photographs of the seabed. Clusters corresponding to the same habitat types were grouped together and mapped as the corresponding habitat types. Confusion matrices, summarizing agreement and disagreement between the ground-truth data and the results from the classified bottom types (Jensen, 2005), were built to compute the overall accuracy and kappa coefficient of agreement of each habitat map. The two measures are commonly used in ecology (Boyce *et al.*, 2002) and in remote sensing (Congalton, 1991; Jensen, 2005). The success of the discrimination of each individual bottom type by the 29 classifications was assessed using the producer's accuracy (Jensen, 2005), and a spatial comparison of the outputs was made to assess the amplitude of change caused by selecting different variables. This was quantified using the percentage of pixels that were classified as the same bottom type by different classifications.

4.2.3 Supervised Classifications of Sea Scallop Habitats

Maximum entropy (MaxEnt) (Jaynes, 1957; Phillips *et al.*, 2004), which was shown to perform better than other species distribution models (SDM) in both terrestrial (Phillips *et al.*, 2006) and marine realms (Monk *et al.*, 2010), was used to perform supervised classification of scallops habitat. Following the method of Brown *et al.* (2012), the classifier was run in the MaxEnt software v.3.3.3k with the default settings, except that the number of background points was increased to 50,000 to account for background conditions in full measure in such a large area. The 3,813 scallop observations selected by Brown *et al.* (2012) were used to train the model, while the remaining 1,003 observations were kept for validation. A total of 29 MaxEnt models were run: for each of the seven selections, four models were run according to the scenarios previously mentioned resulting in 28 models, and one model was run without terrain attributes.

The MaxEnt software was also used to perform jackknife tests and to calculate the area under the curve (AUC) derived from threshold independent receiver operating characteristic (ROC) curves; the former quantify the percentage contribution of each input variable to the models while the latter serves to assess the performance of SDMs (Phillips *et al.*, 2006). We acknowledge that there is currently a debate in the literature surrounding the use of AUC as a measure of model evaluation (*e.g.* Jiménez-Valverde, 2012); some authors argue that AUC can be inappropriate when different modelling techniques are used (Peterson *et al.*, 2008) or if two different species or areas are compared (Lobo *et al.*, 2008). However, AUC often performs better than other measures (McPherson *et al.*, 2004; Vaughan & Ormerod, 2005) and is appropriate when the species, study area, and the training and test samples are the same across the compared models (Elith *et al.*, 2011; Merow *et al.*, 2013), like in the current study.

Model outputs were evaluated in terms of their statistical fit to the validation data (AUC_{Test}) (Fitzpatrick *et al.*, 2013). A non-parametric 95% confidence interval around the AUC_{Test} values was used to identify the significant differences in performances (Zweig & Campbell, 1993). The goodness-of-fit of the models to the training data (AUC_{Train}) was used to assess models' generalizability (*i.e.* transportability, transferability). Generalizability is described by Vaughan & Ormerod (2005, p.720) as "a basic requirement for predictive models" that describes the ability of a model to produce accurate predictions with data other than the training dataset. Generalizability was measured using the difference (AUC_{Diff}) between AUC_{Train} and AUC_{Test} (Warren & Seifert, 2011). A model that over-fits the training data will have a high AUC_{Train} but a low AUC_{Test} as it performs poorly on the test dataset, thus resulting in a high AUC_{Diff}. Such a model is too specific to the training data and less generalizable. A diagnostic of the input variables contribution to the different models was also performed based on the results from the jackknife procedure, in order to identify the loss or gain in explanatory power as each variable is removed from the models or used alone (Khatchikian *et al.*, 2011). Finally, a spatial comparison of the models was performed to evaluate the consequences of variable selection on the model outputs.

4.3 Results

4.3.1 Unsupervised Classifications

4.3.1.1 Performance of Classifications

The overall accuracies and kappa coefficients of the 29 habitat maps are presented in Figure 4.2. Selection 1 (*i.e.* the proposed attribute selection) outperformed the others with the highest overall accuracy and kappa coefficient in three of the four scenarios. The highest kappa coefficient was obtained when combining Selection 1 with bathymetry and the backscatter derivatives. The highest overall accuracy, 68.3%, was reached when combining Selection 5 with bathymetry (Figure 4.2B). Selection 1 combined with bathymetry had the second highest overall accuracy (67.1%). Selections with only three attributes from Selection 1 (*i.e.* Selections 5, 6 and 7) usually outperformed their related selection with none of the proposed attributes (*i.e.* Selections 2, 3 and 4). Selection 4 resulted in poor classifications, and Selections 2, 3 and 6 performed generally poorly except when bathymetry was added. Compared to the classification that only used bathymetry and the backscatter derivatives (*i.e.* no topography), eight classifications had a higher overall accuracy: the four classifications that used Selection 1 as input, Selections

5 and 6 combined with bathymetry, and Selections 5 and 6 combined with both bathymetry and the backscatter derivatives. In terms of kappa coefficients, only four classifications performed better than the one with no topography: Selection 1 with the backscatter derivatives, Selection 1 with both bathymetry and the backscatter derivatives, Selection 5 with bathymetry, and Selection 6 with both bathymetry and the backscatter derivatives.

Differences up to 45.5% were observed between the overall accuracy values and the kappa coefficients for a same selection and scenario. Differences were substantial with an average of 28.5% and a standard deviation of 11.9%. The average difference between the two measures of accuracy for the four maps using Selection 1 was the lowest, followed by the average difference for the four maps of Selections 5, 7, 6, 2, 3 and 4.

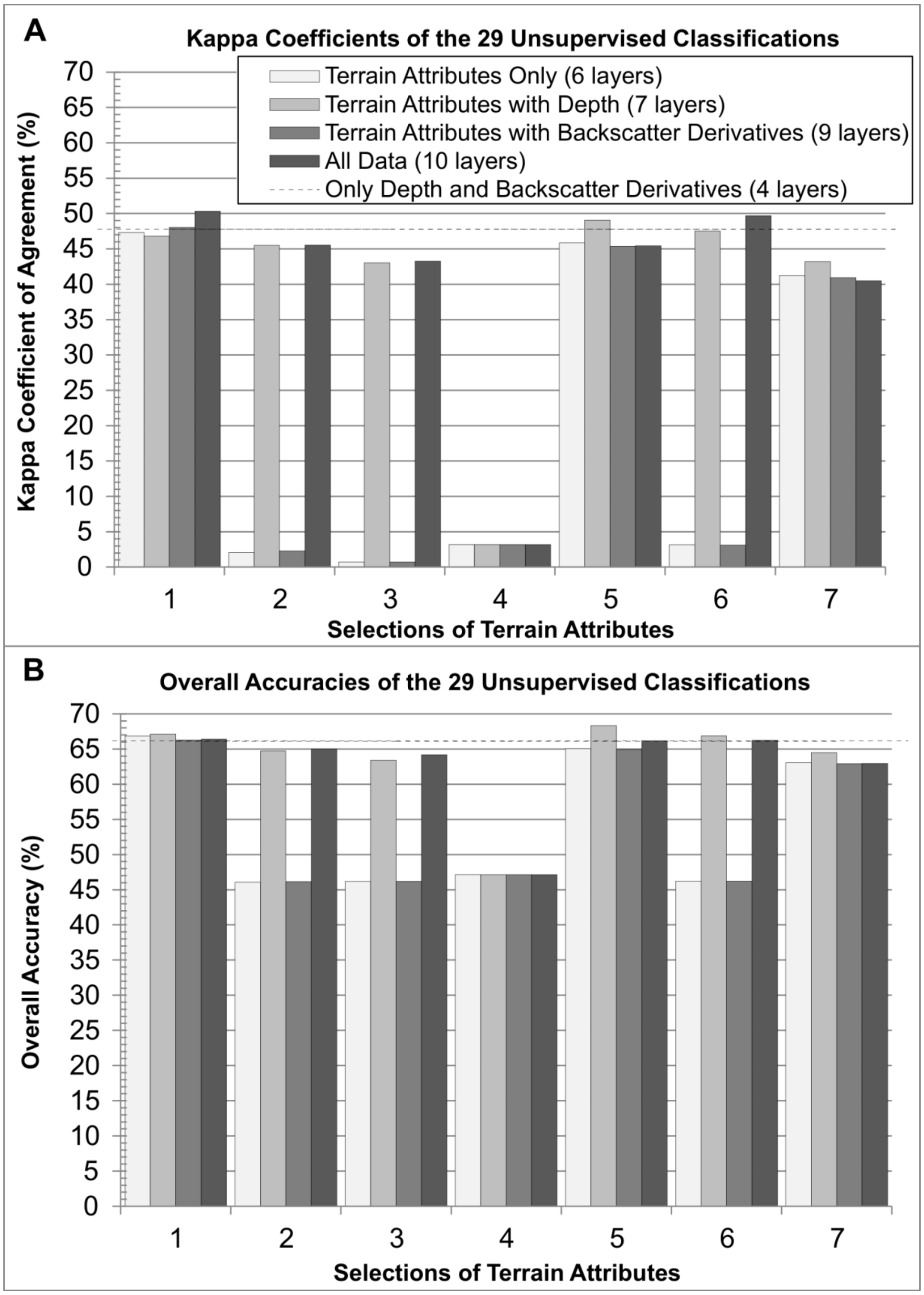

Figure 4.2: Map accuracies measured with (A) a kappa coefficient of agreement and (B) the overall accuracy.

4.3.1.2 Discrimination of Benthoscape Classes

Selection 1 performed on average better than the others when discriminating between the five bottom types (see Figure D.1 of Appendix D). When looking at the individual habitat types, 25 of the 28 other classifications discriminated glacial till better than the classification with only bathymetry and the backscatter derivatives (producer's accuracy of 77.1%), indicating that terrain morphology is not a good surrogate of the presence of glacial till. The "silt and mud" class seemed driven primarily by bathymetry and sediment properties (*i.e.* backscatter derivatives), with a producer's accuracy of 87.4% for the classification that did not account for terrain morphology. Only two of the 28 remaining classifications discriminated that habitat type better, although several other classifications were very close to achieving that accuracy. Reefs were generally poorly discriminated. The classification with no topography reached a producer's accuracy of 19.8%, and only six of the remaining classifications performed better, including three of the classifications using Selection 1. Rippled silt seemed to be better explained by the bathymetry and the backscatter derivatives, with the corresponding classification reaching an accuracy of 57.6%. Only four other classifications did better, including two classifications that included Selection 1. Finally, rippled sand was very poorly discriminated by all the classifications, which may be due to its small sample size (only 49 photographs).

In terms of mean producer's accuracy for the five habitat types, only three classifications did better than the one with no topography (48.4%): Selection 1 with the backscatter derivatives (48.5%), Selection 1 with bathymetry and the backscatter derivatives (51.6%), and Selection 6 with bathymetry and the backscatter derivatives

(51.3%). When averaging the mean producer's accuracies from the four scenarios for each selection, Selection 1 ranked first, followed by Selections 5, 7, 6, 3, 2 and 4.

4.3.1.3 Spatial Variations of Outputs from Different Selections

The most accurate map according to the kappa coefficients of agreement was made from Selection 1 combined with bathymetry and the backscatter derivatives. The spatial similarity indices of that map with the other habitat maps built with ten layers are presented in Table 4.2. Compared to Selection 1, Selection 6 produced the most spatially similar map with 90.0% similarity. Selection 4 is the least similar with only about 41.7% identically classified pixels. The other maps were between 73.3% and 79.4% similar to the map with Selection 1, except for the map with no topography (*i.e.* only bathymetry and the backscatter derivatives) with 82%.

Table 4.2: Spatial similarity of the habitat maps and SDMs generated from Selections 2 to 7, compared to the map and model built from Selection 1. A similarity of 90% indicates that 90% of the pixels were classified as the same habitat type in the two compared maps, or that 90% of the pixels were within ±5% of probability distribution in the two compared models.

	Spatial Similarity with Selection 1 (%) Scenario with 10 Layers	
	Unsupervised Classifications	*Supervised Classifications (within ±5% probability)*
Selection 2	73.3	72.9
Selection 3	77.0	64.6
Selection 4	41.7	65.3
Selection 5	79.4	81.4
Selection 6	90.0	71.5
Selection 7	78.4	70.9
No topography	82.1	66.9

4.3.2 Supervised Classifications

4.3.2.1 Predictive Capacity and Robustness

Figure 4.3 shows the performance of the 29 MaxEnt models. All models performed significantly better than random (*i.e.* AUC_{Test} ± 95% confidence interval > 0.500). Models with higher AUC_{Test} and lower standard deviations are more robust and present the highest predictive capacity (Palialexis *et al.*, 2011). In general, adding bathymetry, the backscatter derivatives, or all of them to the terrain attributes improved the models predictive capacity. However, Selection 1 and other selections that include terrain attributes from Selection 1 did not always follow that trend. For instance, Selection 1 used alone (only six terrain attributes; black diamond in Figure 4.3) performed better than other selections combined with bathymetry or the backscatter derivatives (*e.g.* blue and green squares and triangles in Figure 4.3). Selection 1 combined with the backscatter derivatives (black triangles in Figure 4.3) performed better than other selections that were combined with both bathymetry and the backscatter derivatives (*i.e.* most circles in Figure 4.3).

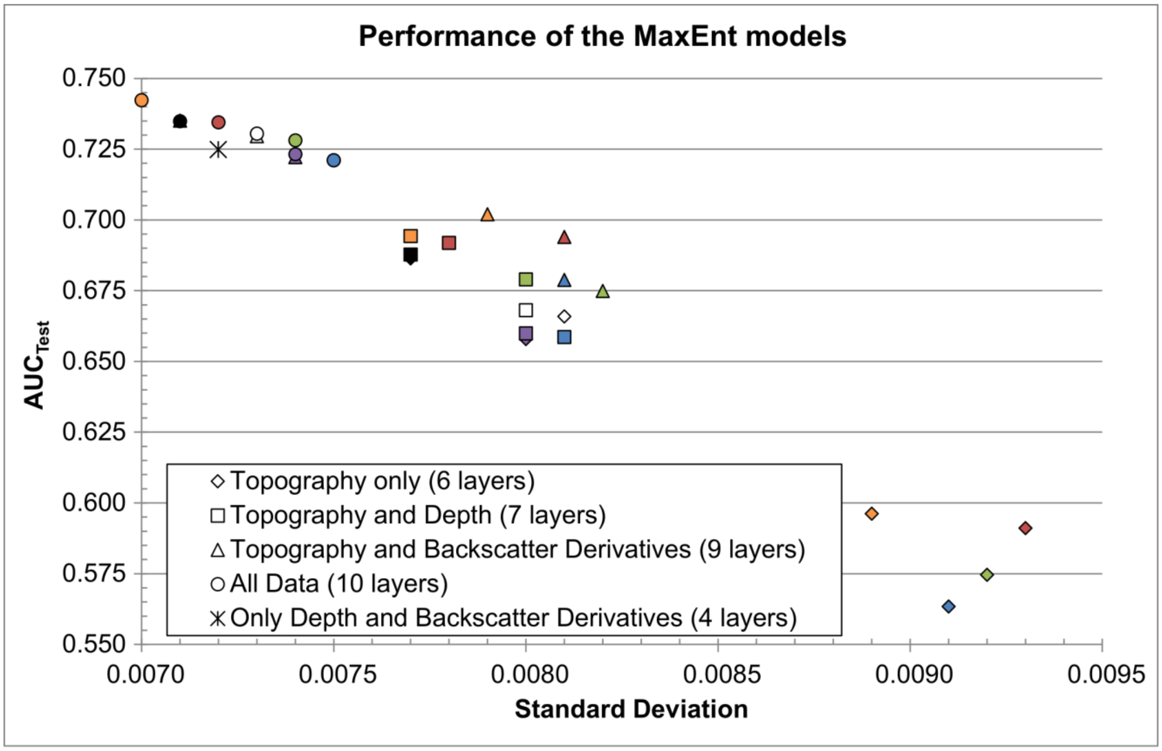

Figure 4.3: Performance and robustness of the 29 MaxEnt models. Models in the top-left corner of the graph performed better and are more robust. Colour legend: Selection 1 (black), Selection 2 (blue), Selection 3 (red), Selection 4 (green), Selection 5 (purple), Selection 6 (orange), Selection 7 (white).

In the scenario where only terrain attributes are used (diamonds in Figure 4.3), Selection 1 performed the best, followed by the three selections that include three terrain attributes from Selection 1 (Selections 5, 6, and 7). The same pattern was observed when combining the selections with the three backscatter derivatives (triangles in Figure 4.3). A different pattern arose when adding bathymetry to the selections, one in which Selection 1 performed second best behind Selection 6. However, the 95% confidence intervals measured around the AUC values show that the difference in performances between

Selection 1 and 6 are not significant for the two scenarios where Selection 6 performed better than Selection 1.

4.3.2.2 Generalizability

Figure 4.4 shows the generalizability of the 29 SDMs. Models with higher AUC_{Train} fitted better the training data while models with lower AUC_{Diff} predicted more efficiently the validation data. Models with high AUC_{Train} and low AUC_{Diff} are therefore the most generalizable, as they do not over-fit the training data (Palialexis *et al.*, 2011). Figure 4.4 shows that the models that included bathymetry (scenarios with seven and ten layers; squares and circles in Figure 4.4) are more similar than the other models, especially for the models that combined ten input layers.

Models that used only terrain attributes or combined them with backscatter derivatives (diamonds or triangles in Figure 4.4) showed similar patterns, where the best models in terms of AUC_{Train} also had a higher AUC_{Diff}, an indication that the best models were also the ones that over-fitted the data the most. In those two scenarios, Selection 1 clearly stands out as a good trade-off between predictive ability and over-fitting of data, making it the most likely to be generalizable and to perform well. When considering bathymetry (squares and circles in Figure 4.4), a similar pattern emerged whether or not the backscatter derivatives were added: Selections 1, 3 and 6 stand out as being more generalizable. Selections 7 and 4 have the highest AUC_{Train}, but also the highest AUC_{Diff}, therefore having a tendency to over-fit the training data.

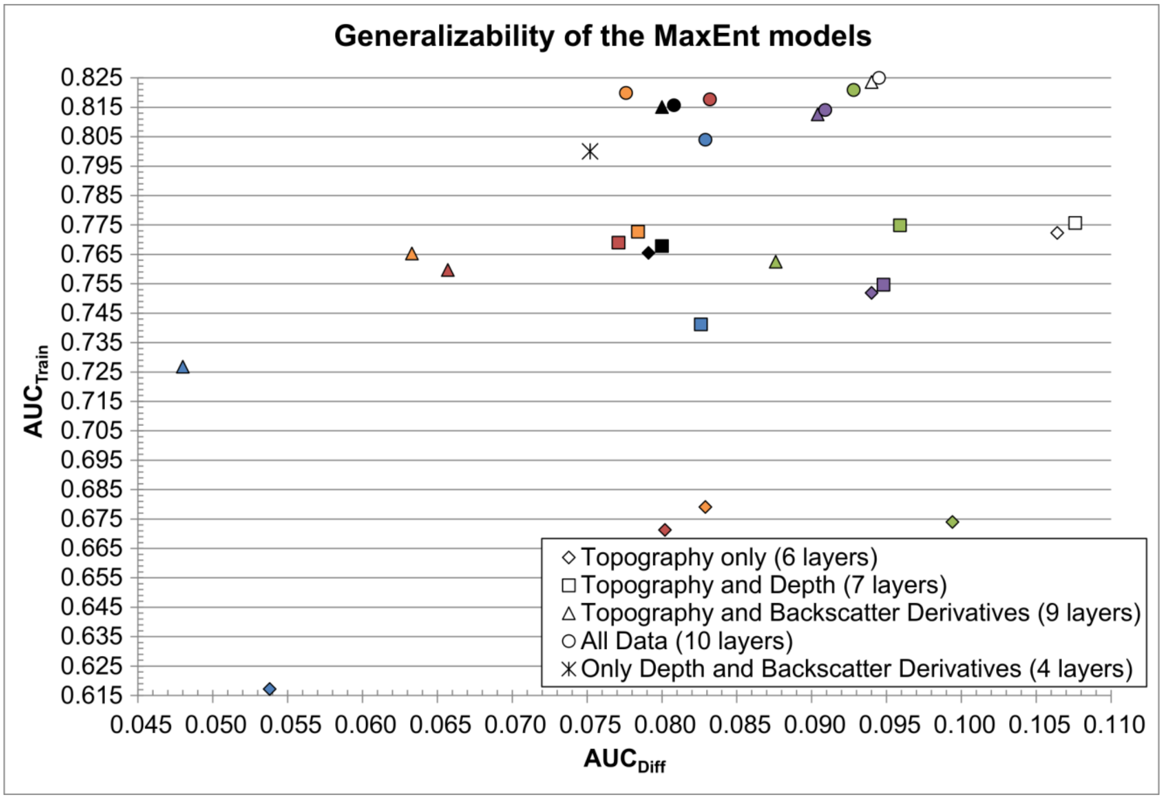

Figure 4.4: Generalizability of the 29 MaxEnt models. Models closer to the top-left corner are more generalizable as they performed well on the training data and replicated well to the validation data. See Figure 4.3 for colour legend.

4.3.2.3 Variables Contribution

The percentage of contribution of each variable used as input in the 29 models can be found in Figure D.2 of Appendix D. When used, bathymetry and two of the backscatter derivatives (Q1 and Q2) contributed the most to the models, with a respective average of 39.2%, 25.4% and 19.6% for the 15 models that used them. Bathymetry contributed less to the models that include local mean as input, resulting from the high collinearity between these two variables; when two variables are correlated, MaxEnt is known to assign a more important percentage contribution to one of the two and a lower one to the other (Khatchikian *et al.*, 2011). Consequently, local mean is a surrogate of bathymetry and appears as an important variable, with an average contribution of 51.2% for the 12 models that include it. In general, measures of rugosity like standard deviation and vector ruggedness measure also contributed to the models.

The analysis of changes in model gain based on the jackknife procedure described the impact on model gain of removing each variable from the models, in addition to provide what would be the model gain if each variable would be used alone. This analysis provided additional information on the variables contribution and the performance of models. In MaxEnt, a variable with a high gain when used alone in a model contributes useful information to the model (Khatchikian *et al.*, 2011). On the other hand, a variable that contributes unique information to a model makes the gain decrease when it is excluded from the model (Khatchikian *et al.*, 2011). In this study, all variables in all models provided unique information in training the models, except for the four models that included Selection 2. In terms of transferability of this uniqueness to the training data (*i.e.* if the variables still provide unique information when applied to the validation data), Selection 1 performed better than the others in three of the four scenarios. It only failed to outperform the other selections when ten layers were used, likely due to spatial correlations between local mean and bathymetry, and slope and local standard deviation. Regarding usefulness, Selection 1 generally did not provide as many useful variables to the models being trained as the other selections. However, these useful variables were generally also useful for the validation data, thus transferable, which was not the case for the other selections. For instance, Selection 1 in combination with bathymetry and the backscatter derivatives had six variables providing useful information to the trained model, and these six variables were all useful for the validation data, an indication of robustness and generalizability. Finally, only two models would not have reached a higher AUC_{Test} if any one of their inputs were removed: Selection 1 combined with the backscatter derivatives and Selection 1 combined with bathymetry.

4.3.2.4 Spatial Variations of Predictions from Different Selections

The model computed from the combination of Selection 1 with bathymetry and the backscatter derivatives showed the best trade-off between robustness, uniqueness and generalizability. It was therefore used as a reference to spatially compare the outputs of comparable models, *i.e.* those computed with ten layers (Table 4.2). The most similar model to the reference one, based on a ±5% margin in probability distribution, was the model computed with Selection 5 (81.4% similar). The lowest similarity was 64.6% (Selection 3). In average, the six other models were 71.1% similar to the one made from Selection 1. The map produced without terrain morphology had a spatial similarity index of 66.9% with the map from Selection 1 combined with bathymetry and the backscatter derivatives.

4.4 Discussion

4.4.1 Framework for Terrain Attribute Selection

The operational framework proposed in Lecours *et al.* (submitted) was based on two literature-grounded assumptions: fractal-based surfaces created with spectral synthesis are appropriate representations of natural surfaces (Peitgen & Saupe, 1988; Chipperfield *et* *al.*, 2011), and the scale-invariance property of fractals allows results to be generalized to other spatial scales (*i.e.* different resolution and/or extent) (Pfeifer, 1984; Keitt. 2000). Artificial surfaces proved their value in ecology (With & King, 1997) and geomorphometry (Jones, 1998). DTMs of real terrains are actually geographic "representations" of real terrains, thus in theory no different than DTMs representing artificial terrains with characteristics found in real terrains. However, a number of authors argue that fractal-based surfaces should be limited to the development of null hypotheses (With & King, 1997). The debate is still unsettled; while some claim that "it is heuristically clear that seafloor or landscape topography is best described by fractal geometry" (Herzfeld & Overbeck, 1999, p. 981), others prefer to argue that despite demonstrating fractal-like properties (Milne, 1992), real terrains are not perfectly fractal (Halley *et al.*, 2004). Without necessarily contributing to this debate, the current study confirmed that results gained from the artificial fractal surfaces in Lecours *et al.* (submitted) hold when using a DTM representation of a real terrain (*i.e.* German Bank) at another spatial scale (*i.e.* an extent of 3,650 km^2 represented at 50 m resolution). Consequently, it confirmed the appropriateness of the proposed framework for selecting terrain attributes and its application to any terrestrial and marine ecological application, regardless of the scale of the environmental data.

4.4.2 Selections of Terrain Attributes

The findings from this study, utilizing MBES-derived surfaces from German Bank, support many of the findings presented by Lecours *et al.* (submitted) based on terrain attributes generated from artificial surfaces. First, the proposed selection of terrain attributes performed better than the other selections tested, both in the application of top-down and bottom-up approaches to habitat mapping: they generally (1) produced more accurate habitat maps, (2) better discriminated individual habitat types, (3) produced SDMs with higher AUC values, (4) produced more robust and generalizable SDMs, (5) provided SDMs with the most variables carrying unique information, and (6) had the highest number of variables carrying useful information that replicated well to the validation data. Using real data, these results confirm that the six recommended terrain attributes best describe the topographic structure of the terrain by capturing different and unique characteristics of the terrain. Results also indicate robustness and generalizability of the proposed framework. Many aspects of this study highlighted better performances of Selection 1 compared to Selections 2, 3 and 4, thus confirming the limited ability of these three selections to adequately and fully describe terrain geomorphology.

4.4.3 Terrain Morphology as an Environmental Factor

Results of both types of classifications indicate that bathymetry and substrate characteristics (for which the backscatter derivatives were a proxy) had a positive, and sometimes more important impact on the performance of the classifications than terrain morphology (quantified through terrain attributes); adding bathymetry and the backscatter derivatives to terrain attribute variables often increased map accuracy for both benthoscape and sea scallop suitability distributions on German Bank. For the unsupervised classifications, only two of the five bottom types (glacial till and reefs) seemed to be driven to a certain level by local geomorphology. In addition, reefs and rippled silt were poorly discriminated by a majority of classifications, likely because their distribution is influenced by other environmental factors or that the variables tested were measured and analyzed at a scale that did not match the scale of the relevant geomorphological features (Lecours *et al.*, 2015). In agreement with results from Brown *et al.* (2012), the MaxEnt analysis showed that bathymetry, sediment properties and rugosity are important variables in predicting sea scallops distribution, but that aspect, slope and relative position are not. Only four SDMs out of 28 performed better than the model with no topography (only bathymetry and the three backscatter derivatives).

Other variables (*e.g.* physical, oceanographic, ecological) may drive particular species or assemblage distributions more than terrain geomorphology. However, they were not used in this study as they were not available at the same spatial scale as the MBES data. When including more variables, users need to keep in mind that covariation may influence models like MaxEnt. If an oceanographic variable is correlated with a terrain characteristic, the user needs to keep only one of them. This is also true of the proposed selection of terrain attributes; as demonstrated in Lecours *et al.* (submitted) and confirmed in the current study, each of the six proposed terrain attributes captures a unique characteristic of the terrain, but some of these characteristics may be spatially correlated in a certain area.

The framework for selecting terrain attributes for ecological studies proposed by Lecours *et al.* (submitted), and supported by the findings of this study, aims at helping the end-users select a robust combination of terrain attributes that best captures the different characteristics of terrain geomorphology. The recommended selection of six terrain attributes serves as a guide as to which set of attributes should be tested in order to achieve the best outcome. It provides end-users with an optimal set of attributes, from which a subset combined with other environmental variables can result in a high accuracy map or model output. The best results will not necessarily come from the use of all six terrain attributes, but may only come from some of them. For instance, if particular terrain characteristics have no ecological meaning in an application, using the terrain attributes that capture these characteristics will not yield the best outcome. It is therefore highly site and case specific as to which variables should be included (Lecours *et al.*, 2015). Nonetheless, the recommended approach provides the optimal starting point from which terrain attributes can be selected.

4.4.4 Consequences of Variable Selection

Results highlight the importance of appropriately selecting input variables in both unsupervised and supervised classifications, and consequently the inappropriateness of making such selection arbitrarily. For instance, the benthoscape map generated from the combination of Selection 1 with bathymetry and the backscatter derivatives yielded an overall accuracy and a kappa coefficient of agreement that are respectively only 0.2% and 0.6% different than the map built from Selection 6, bathymetry and the backscatter derivatives. It would be quite intuitive to interpret the difference in map outputs as insignificant based only on these measures of accuracy. However, 10.0% of the study area was classified differently by these two classifications, an area corresponding to about 362 km^2. In addition, the differences occurred in all regions of the study area and across all the habitat types. In the worst case scenario (*i.e.* the difference between Selection 4 and Selection 1, Table 4.2), the total area that was mapped differently covers over 2,115 km^2.

The results of this study indicate that a subjective selection of terrain attributes could potentially provide a map that is in average 26.7% different in terms of the location and boundaries of benthoscape classes, which has serious implications for ecological applications that use these maps and models for decision-making.

4.4.5 Comparisons with Other Studies: Terrestrial and Marine

Many different terrain attribute selections have been used in terrestrial and marine ecology (Bouchet *et al.*, 2016; Lecours *et al.*, 2016; references therein). In a meta-analysis of ecological studies using geomorphometry, Bouchet *et al.* (2015) found that about a third of the studies only used one terrain attribute and that very few authors used more than four. While focusing on forest ecosystems, Sharaya & Sharyi (2011) wrote that in general, one to three basic terrain attributes are used to study landscape phenomena and that the "insufficient representativeness" (*ibid*, p. 2) of terrain attributes makes for an inefficient use of topography as a variable in ecology. In a management context and using the same dataset as in the current study, Brown *et al.* (2012) selected six terrain attributes based on previous use in marine ecology studies and "iterative testing of a large number of different layers by the authors" (*ibid*, p. 3). This relatively subjective way of selecting terrain attributes is the most common one in ecology. However, it provides many significant and valid insights for many applications; most of the common terrain attributes found in the ecological literature (*e.g.* local mean, slope, aspect) (Bouchet *et al.*, 2015) are part of our proposed selection, or are related to one of the proposed attributes. For instance, different types of curvature are commonly used, which Lecours *et al.* (submitted) found to be correlated to the recommended relative difference to mean value, although more ambiguously defined and thus not included in the recommended selection. Despite using a subjective selection of terrain attributes, Brown *et al.* (2012) yielded valid results. Their MaxEnt model had a high predictive capacity, although it had some level of over-fitting and was less robust than some of the best models of the current study. If implemented in the current study using the same method, an unsupervised classification made from their selection of variables would rank amongst the best benthoscape maps and be 90.8% similar to the map built with Selection 1 and the four other environmental variables. This demonstrates that despite potentially resulting in huge differences (*cf.* "Consequences of Variable Selection" above), subjective selection of terrain attributes can sometimes produce relevant and valid results.

Finally, the observed differences between the overall accuracy measures and the kappa coefficients of agreement confirm that the overall accuracy might be a poor guide of the value of a classification, something that has been already argued in the literature (Felix & Binney, 1989; Fielding & Bell, 1997). Based on our results, we recommend the kappa coefficient as a more appropriate measure than the overall accuracy for ecological mapping.

4.5 Conclusion

Selecting the most appropriate environmental variables to use in a specific study can be very challenging. This study demonstrated the importance of carefully selecting variables for ecological work; maps and models that perform similarly can still produce very different spatial outcomes, which can have important implications when these maps and models are used in decision-making for conservation and management. Using two different approaches to habitat mapping, this paper also confirmed that the selection of terrain attributes recommended in Lecours *et al.* (submitted) performs better than other selections, thus serving as a guide to make better use of geomorphometry in ecology. Results also showed that while this selection of terrain attributes ensures that most of the local topographic structure is captured when performing terrestrial or marine ecological studies, and while terrain morphology can help improve maps and models, it is not always the most important environmental factor for all ecological applications. The relationship between terrain morphology and ecological phenomena is species, area and scale-dependent (Lecours *et al.*, 2015). The use of the proposed selection of terrain attributes, in combination with other environmental variables (*e.g.* precipitations, climate, currents), will help ecologists produce more robust analyses and generate maps and models with a higher degree of confidence. In order to get the best representation of the environment as possible and to best inform policy, conservation and management efforts, we recommend (1) that stakeholders prepare more than a single map using different combinations of environmental variables, and (2) that they select the best outcome based on map accuracy or model performance quantification.

4.6 Literature Cited

Bouchet, P.J., Meeuwig, J.J., Kent, C.P.S., Letessier, T.B., & Jenner, C.K. (2015) Topographic determinants of mobile vertebrate predator hotspots: current knowledge and future directions. *Biological Reviews*, 90:699-728.

Boyce, M.S., Vernier, P.R., Nielsen, S.E., & Schmiegelow, F.K.A. (2002) Evaluating resource selection functions. *Ecological Modelling*, 157:281-300.

Brown, C.J., Sameoto, J.A., & Smith, S.J. (2012) Multiple methods, maps, and management applications: purpose made seafloor maps in support of ocean management. *Journal of Sea Research*, 72:1-13.

Chipperfield, J.D., Dytham, C., & Hovestadt, T. (2011) An updated algorithm for the generation of neutral landscapes by spectral synthesis. *PLoS ONE*, 6:e17040.

Congalton, R.G. (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37:35-46.

Dolan, M.F.J., & Lucieer, V.L. (2014) Variation and uncertainty in bathymetric slope calculations using geographic information systems. *Marine Geodesy*, 37:187-219.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distribution*, 17:43-57.

Evans, I.S., & McClean, C.J. (1995) The land surface is not unifractal: variograms, cirque scale and allometry. *Zeitschrift für Geomorphologie*, 101:127-147.

Felix, N.A., & Binney, D.L. (1989) Accuracy assessment of a Landsat-assisted vegetation map of the coastal plain of the Arctic National Wildlife Refuge. *Photogrammetric Engineering and Remote Sensing*, 55:475-478.

Fielding, A.H., & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24:38-49.

Fitzpatrick, M.C., Gotelli, N.J., & Ellison, A.M. (2013) MaxEnt versus MaxLike: empirical comparisons with ant species distributions. *Ecosphere*, 4:1-15.

Halley, J.M., Hartley, S., Kallimanis, A.S., Kunin, W.E., Lennon, J.J., & Sgardelis, S.P. (2004) Uses and abuses of fractal methodology in ecology. *Ecology Letters*, 7:254-271.

Herzfeld, U.C., & Overbeck, C. (1999) Analysis and simulation of scale-dependent fractal surfaces with application to seafloor morphology. *Computers and Geosciences*, 25:979-1007.

Jaynes, E.T. (1957) Information theory and statistical mechanics. *Physical Review*, 106:620-630.

Jensen, J.R. (2005) *Introductory digital image processing, a remote sensing perspective, 3rd edition*. Upper Saddle River: Pearson Prentice Hall.

Jiménez-Valverde, A. (2012) Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, 21:498-507.

Jones, K.H. (1998) A comparison of two approaches to ranking algorithms used to compute hill slopes. *GeoInformatica*, 2:235-256.

Keitt, T.H. (2000) Spectral representation of neutral landscapes. *Landscape Ecology*, 15:479-493

Khatchikian, C., Sangermano, F., Kendell, D., & Livdahl, T. (2011) Evaluation of species distribution model algorithms for fine-scale container breeding mosquito risk prediction. *Medical and Veterinary Entomology*, 25:268-275.

Lecours, V. (2015) *Terrain Attribute Selection for Spatial Ecology (TASSE), v. 1.0*. URL www.marinegis.com.

Lecours, V., Devillers, R., Simms, A.E., Lucieer, V.L., & Brown, C.J. (submitted) Towards a framework for terrain attribute selection in environmental studies.

Lecours, V., Dolan, M.F.J., Micallef, A., & Lucieer, V.L. (2016) A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrology and Earth System Sciences*, 20:3207-3244.

Lobo, J.M., Jiménez-Valverde, A., & Real, R. (2008) AUC : a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17:145-151.

McPherson, J.M., Jetz, W., & Rogers, D.J. (2004) The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of Applied Ecology*, 41:811-823.

Merow, C., Smith, M.J., & Silander Jr., J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*. 36:1058-1069.

Milne, B.T. (1992) Spatial aggregation and neutral models in fractal landscapes. *American Naturalist*, 139:32-57.

Monk, J., Ierodiaconou, D., Versace, V.L., Bellgrove, A., Harvey, E., Rattray, A., Laurenson, L., & Quinn, G.P. (2010) Habitat suitability for marine fishes using presence-only modelling and multibeam sonar. *Marine Ecology Progress Series*, 420:157-174.

Palialexis, A., Georgakarakos, S., Karakassis, I., Lika, K., & Valavanis, V.D. (2011) Prediction of marine species distribution from presence-absence acoustic data: comparing the fitting efficiency and predictive capacity of conventional and novel distribution models. *Hydrobiologia*, 670:241-266.

Peitgen, H.O., & Saupe, D. (1988) *The science of fractal images*. New York: Springer-Verlag.

Peterson, A.T., Papes, M., & Soberón, J. (2008) Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, 213:63-72.

Pfeifer, P. (1984) Fractal dimension as working tool for surface-roughness problems. *Applied Surface Science*, 18:146-164.

Phillips, S.J., Dudík, M., & Schapire, R.E. (2004) A maximum entropy approach to species distribution modeling. *Proceedings of the 21st International Conference on Machine Learning*, New York: ACM Press, pp. 655-662.

Phillips, S.J., Anderson, R.P., & Schapire, R.E. (2006) Maximum entropy modeling of species distributions. *Ecological Modelling*, 190:231-259.

Sharaya, L.S., & Sharyi, P.A. (2011) Geomorphometric study of the spatial organization of forest ecosystems. *Russian Journal of Ecology*, 42:1-8.

Tate, N.J. (1998) Estimating the fractal dimension of synthetic topographic surfaces. *Computers and Geosciences*, 24:325-334.

Vaughan, I.P., & Ormerod, S.J. (2005) The continuing challenges of testing species distribution models. *Journal of Applied Ecology*, 42:720-730.

Warren, D.L., & Seifert, S.N. (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21:335-342.

With, K.A., & King, A.W. (1997) The use and misuse of neutral landscape models in ecology. *Oikos,* 79:219-229.

Zajac, R. (2008) Challenges in marine, soft-sediment benthoscape ecology. *Landscape Ecology*, 23:7-18.

Zweig, M.H., & Campbell, G. (1993) Receiver-operating characteristics (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry*, 39:561-577.

5. A Multiscale Analysis of the Impact of Artefacts in Multibeam Bathymetric Data on Terrain Attributes

5.1 Introduction

In recent years, the growing availability of remotely sensed data and the diversity of geospatial tools, such as geographic information systems (GIS), have revolutionized the science of geomorphometry that studies the quantitative measurement of terrain morphology. Geomorphometric analyses are commonly conducted in five steps (Pike *et al.*, 2009): (1) sampling the terrain, (2) generating a digital terrain model (DTM) from the samples, (3) processing the DTM to prepare it for the next step, (4) generating terrain attributes (*e.g.* slope, aspect, rugosity) or extracting terrain features (*e.g.* ridges, channels, peaks), and (5) employing these surfaces to answer a specific question or to use in further statistical analyses. The third step, the processing of the DTM, usually involves the detection and possibly correction of artefacts, noise and systematic errors (Wise, 2000; Pike *et al.*, 2009). Artefacts are "distinct erratic features", often but not always systematic in nature, caused by unrealistic, erroneous values (Reuter *et al.*, 2009, p.91). They can potentially be found in all types of DTMs, regardless of the data collection technique and the interpolation method used to generate the surface (Gessler *et al.*, 2009). In terrestrial DTMs, or Digital Elevation Models (DEM), artefacts can be obvious, very subtle, or invisible (Regan *et al.*, 2002; Albani & Klinkenberg, 2003). Artefacts were shown to impact data quality more than random noise (Rousseaux, 2003; Van Niel *et al.*, 2004), and to propagate and sometimes amplify in derived terrain attributes like curvature

(Temme *et al.*, 2009; Sofia *et al.*, 2013). Despite this, DTMs and other remotely sensed data are often analyzed by end-users as if they were error-free and a real surface, rather than a representation or model of a real surface. This was observed both on land (*e.g.* Evans, 1997; Oksanen & Sarjakoski, 2005) and underwater (*e.g.* Dolan & Lucieer, 2014).

Developments in geomorphometry have traditionally focused on the exploration of terrestrial and extra-terrestrial environments, but recent efforts have been made to highlight the need for a dedicated science of marine geomorphometry (Lecours *et al.*, 2015a; 2016); the nature of the marine environment and the techniques used to sample depth have implications for the subsequent geomorphometric analyses that are different than in terrestrial applications. The increased availability of multibeam echosounder systems (MBES) that enable the collection of reliable continuous underwater terrain data (*i.e.* bathymetry) (Brown *et al.*, 2011; Erikstad *et al.*, 2013) has revolutionized several fields of research and applications like marine habitat mapping (Smith & McConnaughey, 2016) and marine geomorphology (Hughes-Clarke *et al.*, 1996). MBES bathymetric data, or Digital Bathymetric Models (DBM), are now commonly used in GIS to derive terrain attributes that can be used as surrogates for other phenomena in different disciplines (Lecours *et al.*, 2016). For instance, measures of aspect, which informs on the orientation of the slope, can act as a proxy of currents in hydrodynamics modelling (Gille *et al.*, 2004) or of food supply in habitat mapping (*e.g.* Tong *et al.*, 2013). Rugosity has also been shown to be a good surrogate of biodiversity: complex habitats are known to shelter higher levels of biodiversity than less complex areas (*e.g.* Kostylev *et al.*, 2005; Dunn & Halpin, 2009).

DBM are more prone to errors and artefacts than DEM for several reasons (also see discussion in Lecours *et al.*, 2016). First, the motion of the supporting platform, which can be a ship, a remotely operated vehicle (ROV), or an autonomous underwater vehicle (AUV), is more influenced by environmental conditions (*e.g.* wind, currents, waves, tides) than are airborne platforms or satellites. Second, similarly to the influence of atmospheric conditions on electromagnetic radiations, the properties of the water column (*i.e.* temperature, salinity and pressure) strongly influence sound waves propagation (Jianhu & Jingnan, 2003). However, water properties are less predictable than atmospheric conditions as they demonstrate finer spatial and temporal variability and are more difficult to measure at an appropriate scale (Cushman-Roisin & Beckers, 2011). Consequently, it is sometimes difficult to account for the effect of these properties on sound waves, which may result in what is called "refraction artefacts" in the DBM (see Lurton, 2010). Third, when using underwater platforms such as ROV and AUV, the quality of the DBM is limited by the precision and accuracy of ancillary data; because of the inability of GPS to work underwater, additional instruments (*e.g.* ultra-short baseline, Doppler velocity log) need to be integrated to the system, which can result in issues of positional accuracy, timing, and data logging (Lecours & Devillers, 2015). Finally, the limitations of instruments like inertial measurement units (IMU) and GPS receivers to measure truly continuous phenomena like the movement of the platform caused by wave actions prevent the appropriate (continuous) correction of data (Hughes-Clarke *et al.*, 1996; Zhao *et al.*, 2007). As a result of the combination of these elements, most DBMs generated from MBES data present a certain level of artefacts (Hughes-Clarke, 2003a;

Roman & Singh, 2006) that can be directly visible in published works (*e.g.* Lucieer *et al.*, 2012; Georgian *et al.*, 2014). Since artefacts are usually within hydrographic error standards (Hughes-Clarke, 2003a), they are typically considered as acceptable in hydrography. However, problems may arise when such data are used by end-users (*e.g.* habitat mappers, ecologists, geologists) with limited knowledge about the characteristics of the data and how they were collected: the presence of artefacts can lead to misinterpretation of seafloor features, patterns and processes in many disciplines (de Moustier & Kleinrock, 1986; Hughes-Clarke *et al.*, 1996).

While each of the five steps of geomorphometry has been extensively studied in terrestrial settings (Hengl & Reuters, 2009), marine geomorphometry studies have mainly focused on the two last steps. There have been studies that looked at errors and uncertainty in DBMs, particularly in the hydrographic surveying literature: some work has been done on quantifying uncertainty propagation at the data collection level (*e.g.* Calder & Mayer, 2003), identifying the cause of artefacts (Hughes-Clarke, 2003a), and reducing refraction and motion artefacts in post-processing (Yang *et al.*, 2007; Landmark *et al.*, 2015). However, despite their prevalence, there has been no studies of how artefacts in bathymetry can impact the derivation of terrain attributes, resulting in little understanding of artefacts influence on the geomorphometric workflow, even though "we are more often than not interested in how these errors and artifacts influence our analysis rather than elevation [or depth] per se" (Wilson, 2012, p. 108).

As reviewed in Lecours *et al.* (2016), most marine geomorphometry applications tend to overlook the presence of artefacts that are not removed in post-processing, using judgement to exclude them for practical purposes. Some methods for dealing with artefacts have been proposed in both the terrestrial (*e.g.* Lindsay & Creed, 2006) and marine (*e.g.* Hughes-Clarke, 2003a; Landmark *et al.*, 2015) literature, but are not widely used in a more applied context. There could be many reasons for this. First, it could be because of the lack of implementation of proper tools in software that are easily accessible to end-users, like open-source GIS (Bonin & Rousseaux, 2005). It could also be because spatial analysts cannot always access the hydrographic software to improve the output. Finally, the data may simply be of poor quality. Principles of error propagation make artefacts in DBM very likely to impact any subsequent analyses regardless of context (Heuvelink, 1998; Bangen *et al.*, 2014). A goal of this paper is to start answering the call made by Wilson (2012, p.117) to improve "knowledge of the presence of and propagation of errors" in remote sensing data sources, namely in MBES data, as previously done by Fisher & Tate (2006) for other technologies. The specific objectives are (1) to assess how different types of artefacts in DBMs impact derived terrain attributes for marine geomorphometric applications, (2) to assess if those impacts vary with spatial scale, and (3) to explore the role of sampling density in attenuating or increasing the influence of artefacts.

5.2 Material and Methods

5.2.1 Data

An area covering 3,650 km^2 of German Bank, off Nova Scotia, Canada, was mapped by the Canadian Hydrographic Service (CHS) (Figure 5.1) using a Simrad Subsea

EM1000 MBES attached to the hydrographic Canadian Coast Guard Ship "Frederick G. Creed". This vessel is a Small Waterplane Area Twin Hull (SWATH) designed to limit the impact of sea surface conditions on the ship, making it a highly suitable platform for bathymetric data collection: Collins *et al.* (2005) showed that in the same conditions, the range of motion of a SWATH vessel was about half the range recorded by a more traditional single-hull vessel. Details of the German Bank surveys are described in DFO (2006) and Brown *et al.* (2012). Raw data (*i.e.* soundings) were imported in the bathymetric processing software CARIS HIPS and SIPS, were corrected for tide, motion, and sound velocity, and erroneous soundings were removed. This standard post-processing was performed by the CHS. Corrected soundings were used to generate reference DBMs at five different spatial resolutions (10 m, 25 m, 50 m, 75 m, and 100 m), using the interpolation algorithm "swath angle" in CARIS HIPS and SIPS. While no terrain models can really be free of artefacts, these five surfaces were assumed artefact-free and considered as reference surfaces for further analysis, as done in other studies (*e.g.* Sofia *et al.*, 2013). While some artefacts were present – although very subtle (*cf.* Figure 5.1) – in the data associated with the deeper areas of German Bank, their amplitude was small compared to that of artefacts introduced in this study (see below).

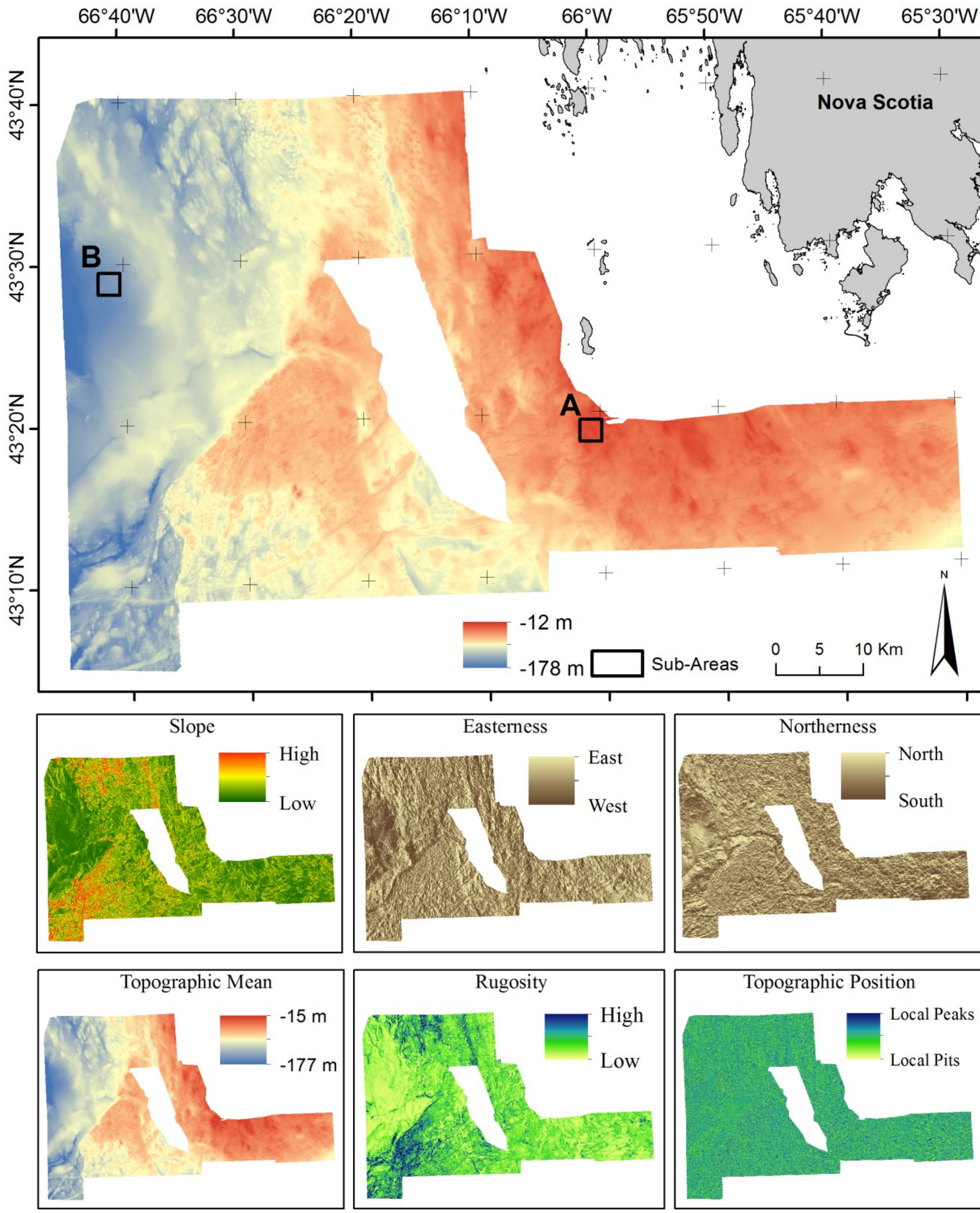

Figure 5.1: 50 m resolution DBM of German Bank (top) with its six terrain attributes derived (bottom). Two specific sub-areas, one in shallower waters (A) and one in deeper waters (B), were used in the analyses and are indicated by a black square in the main map.

Four types of artefacts that are typical sources of error in bathymetric surveys were chosen for this study based on their impact on the bathymetric data: heave, pitch, roll and time artefacts. The first three often result from misalignments among the different systems during calibration or an inappropriate correction of motion, whilst time can be caused by an improper synchronisation of the different components used in data collection. Pitch induces a horizontal displacement of soundings – either ahead or behind the platform – while heave induces a vertical shift of the soundings. Roll impacts the outer beams in the vertical plane and thus affects areas that overlap between different survey lines. Finally, time synchronisation error can shift adjacent survey lines on the horizontal plane. Animated visual representations of the effect of these artefacts on bathymetric data can be found in Hughes-Clarke (1997, 2002), and details on how these artefacts occur, their characteristics and how they alter DBMs can be found in Hughes-Clarke (1996, 2003a, 2003b) and Lurton (2010). We also note that these four types of artefacts form only a subset of the different possible types of artefacts and were chosen based on their prevalence in other studies. However, some of the results gained in the current study can potentially be extrapolated to other types of artefacts: different authors have mentioned that artefacts caused by roll are very similar to those caused by refraction, that imperfect tide correction can have a similar impact to an improper calibration of heave, that heading artefacts are thought to behave similarly than time artefacts, and that in some cases a heading misalignment can also introduce roll and pitch errors (Hughes-Clarke, 2003a; Roman & Singh, 2006; Yang *et al.*, 2007).

The recorded ship motion at the time of the surveys was used to compute statistics on the vessel's range of motion (Table 5.1). CARIS HIPS and SIPS was then used to introduce artificial motion artefacts in the original data: within the vessel configuration file, calibration values of roll, pitch and heave were altered at 10 different levels based on the standard deviation of the recorded motion (Table 5.1). Time artefacts were simulated by inducing time delays between sensors. The time delays were arbitrarily chosen to encompass the range of calibration values used by the CHS: five calibration values ranging from -0.70 to 0.26 seconds (Table 5.1). These alterations applied artificial systematic corrections to the raw soundings that were then used to generate new DBMs with artefacts at the five scales of study, totalizing 200 altered surfaces (*i.e.* 10 levels of artefacts for four types of artefacts at five spatial scales). We acknowledge that systematic errors are not the only type of errors found in bathymetric data, but they provide controlled conditions that enable direct comparisons of results. Such approach is commonly adopted in geomorphometry to evaluate terrain attributes sensitivity to properties of the input DBMs (Florinsky, 1998; Zhou & Liu, 2004; Reuter *et al.*, 2009). We also acknowledge that while many surveying systems have internal quality control filters and would potentially flag some of these artefacts when surveying from a surface vessel, this is not necessarily the case when surveying from a submersible platform in deeper waters, where the uncertainty associated with different components of the system is known to prevent the appropriate correction of motion and to contribute to the presence of artefacts (Lecours *et al.*, 2013; Lecours & Devillers, 2015).

Table 5.1: Statistics of the range of motion recorded during the surveys, and levels of artefact induced to the five reference DBMs. Statistics for time values are based on the five calibration values used by the CHS. The sign convention used is positive pitch with the bow up and positive roll with the port side up.

Total Records	Mean	Standard Deviation	Minimum	Maximum		Level of Induced Artefact									
						-5σ	-4σ	-3σ	-2σ	-σ	σ	2σ	3σ	4σ	5σ
4094390	-0.16	0.33	-2.72	21.27	***Heave (m)***	-1.65	-1.32	-0.99	-0.66	-0.33	0.33	0.66	0.99	1.32	1.65
4094382	-0.76	1.65	-9.88	8.94	***Pitch (°)***	-8.25	-6.60	-4.95	-3.30	-1.65	1.65	3.30	4.95	6.60	8.25
4077894	-0.24	1.01	-11.95	10.04	***Roll (°)***	-5.05	-4.04	-3.03	-2.02	-1.01	1.01	2.02	3.03	4.04	5.05
5	-0.07	0.37	-0.70	0.26	***Time (s)***	-1.25	-1.00	-0.75	-0.50	-0.25	0.25	0.50	0.75	1.00	1.25

As recommended in Chapter 3, a combination of six terrain attributes that together best capture the topographic variability of an area were derived from the reference and altered DBMs: relative difference from mean value (RDMV) (*i.e.* a measure of topographic position), local standard deviation (*i.e.* a measure of rugosity), easterness and northerness (*i.e.* measures of orientation), local mean, and slope using Horn's (1981) algorithm. A total of 1,230 terrain attribute surfaces were generated (30 from the reference surfaces and 1,200 computed from the altered DBMs) using the TASSE toolbox for ArcGIS (Lecours, 2015). Examples computed from the 50 m resolution DBM are presented in Figure 5.1. For the remainder of this paper, in order to avoid confusion with statistical terms, RDMV will be referred as topographic position, local standard deviation as rugosity, and local mean as topographic mean.

Analyses were performed on the full extent of the study area and on two sub-areas (Figure 5.1) to evaluate differences in the effect of artefacts based on sampling density (*cf.* third objective). The two sub-areas were selected for having similar depth distribution characteristics and similar complexity (quantified with fractal dimension), but different density and total amount of soundings (*i.e.* data density) (Table 5.2).

Table 5.2: Comparison of the characteristics of the shallower (A in Figure 5.1) and the deeper (B in Figure 5.1) sub-areas.

		Shallower Sub-Area (A)	Deeper Sub-Area (B)
Density of Soundings per Pixel (10m Resolution)	Minimum	1	1
	Maximum	121	24
	Mean	33.4	6.9
	Standard Deviation	16.5	4.2
Total Number of Soundings in the Area		208,129	42,896
Statistics of Depth Values	Mean	-21.94	-134.12
	Range	14.48	14.47
	Standard Deviation	2.14	2.71
	Skewness	-0.10	-0.14
	Kurtosis	2.71	2.42
	Moran's I	0.97	0.98
Fractal Dimension		2.42	2.44

5.2.2 Comparisons

Four elements were studied for bathymetric and terrain attributes surfaces: (1) how the reference surfaces' spatial and statistical distributions vary with spatial scale, (2) the spatial similarity of the altered surfaces to the reference surfaces, (3) the error induced by the artefacts in the altered surfaces, and (4) the impacts of artefacts on the values of the different surfaces (*e.g.* depth or slope values). Spatial similarity between surfaces was quantified using a correlation coefficient (r). The impact of scale and artefacts on the values of the different surfaces was quantified using different descriptive statistics: range of values, mean, standard deviation, distribution (quantified through skewness and kurtosis), spatial autocorrelation (Moran's I), and fractal dimension (*i.e.* representation of terrain complexity; only for bathymetry and topographic mean). The error induced was considered as the absolute difference between the altered bathymetric and terrain attribute surfaces and their respective reference surfaces (Zhang *et al.*, 2014). Absolute differences were calculated using ArcGIS 10.2.2, and were used to avoid having mean errors of zero (Wise, 2011a). Visual representations of error for each type of artefacts are presented in Figure 5.2 and Figure 5.3 using the two sub-areas. As done in other studies (Fisher, 1998; Holmes *et al.*, 2000; Van Niel *et al.*, 2004), the same statistics as for the previous analysis were calculated to characterize the error, except for fractal dimension that has no meaning in this case. These statistics were computed using a combination of tools from the GIS software ArcGIS 10.2.2, WhiteBox GAT 3.3, and LandSerf 2.3. Spatial similarity among the different levels of error was also calculated to estimate if the error affects the same locations based on the intensity of the artefact.

To enable comparisons of the impact of artefacts on bathymetric data and terrain attributes that have different units and ranges of possible values, statistical changes (*e.g.* in mean or correlation) were modelled against the amplitude of artefacts using quadratic regressions, thus enabling the comparison of rates of change (*i.e.* the first term of the equations). The relationships between the statistics of errors and the amplitude of artefact were also modelled with quadratic regressions, while the relationships between scale and the different surfaces were modelled using linear regressions. The quadratic regressions provided the best fit for the studied relationships. The proportion of the variance explained by the models was assessed with coefficients of determination (r^2), model significance was assessed using the F statistic of overall significance based on model residuals, and the significance of each term in the regression equations was assessed using Student's t-test. All significance tests were performed using $\rho = 0.05$.

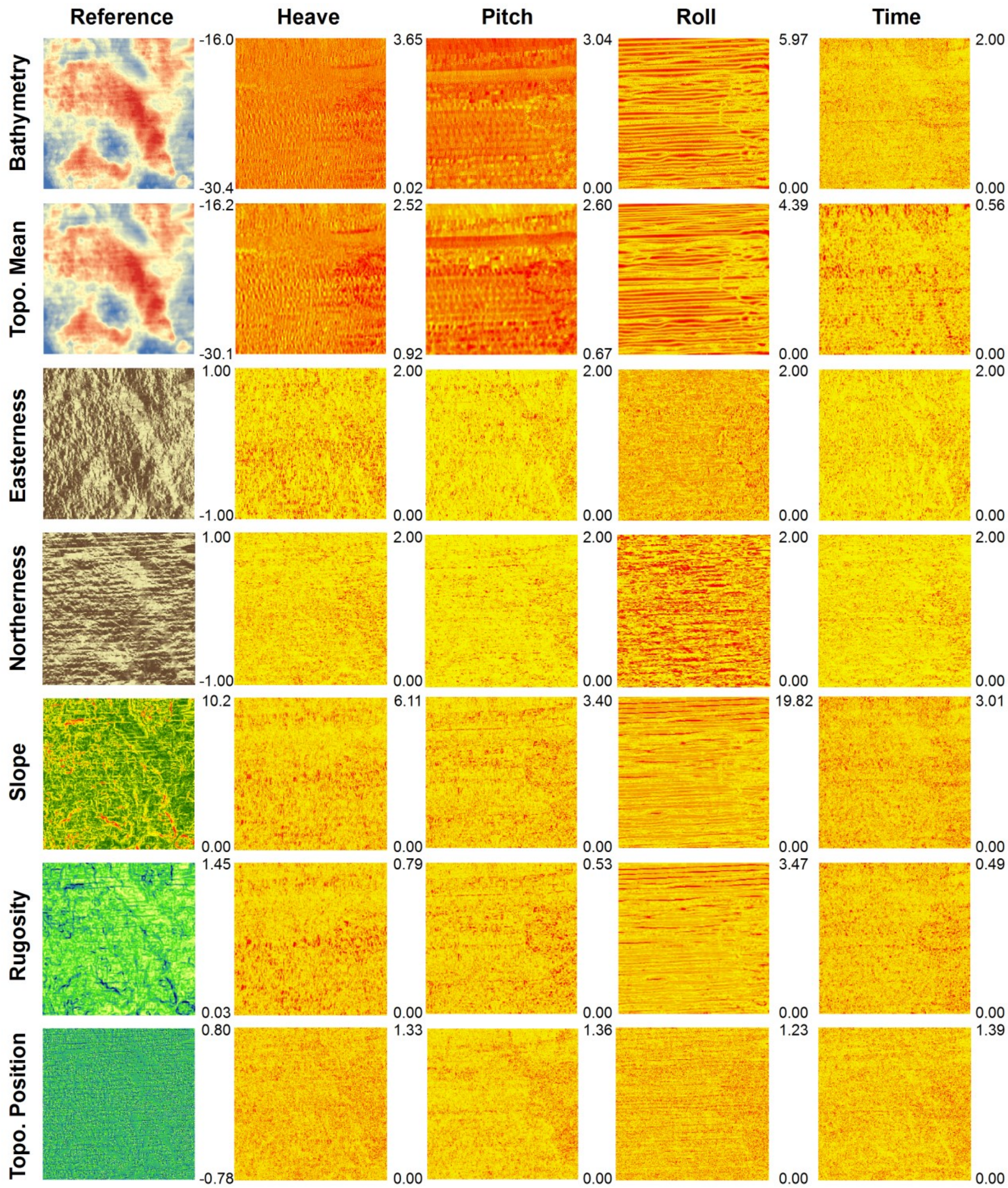

Figure 5.2: The reference panels to the left show the bathymetry and terrain attributes of the shallower sub-area, A, at 10 m resolution (2.5 km by 2.5 km; same colour scheme as Figure 5.1). The panels to the right show examples of error (absolute difference between the reference and altered surfaces) for that sub-area. The level of error represented is the highest one (5σ, Table 5.1). Numbers in the right margins indicate error ranges (note differences in scales). Colours range from yellow (lowest) to red (highest).

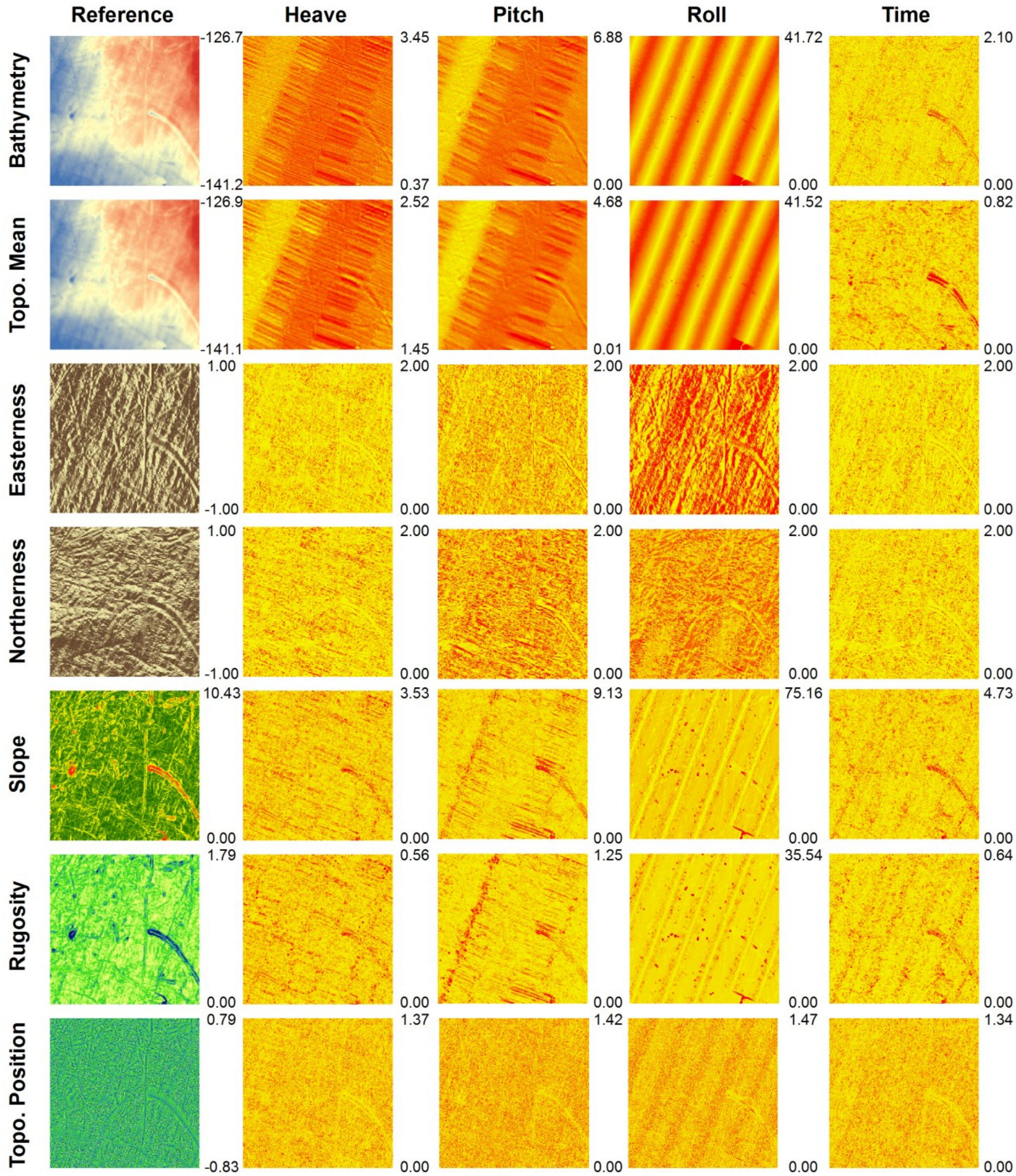

Figure 5.3: The reference panels to the left show the bathymetry and terrain attributes of the deeper sub-area, B, at 10 m resoltuion (2.5 km by 2.5 km; same colour scheme as Figure 5.1). The panels to the right show examples of error (absolute difference between the reference and altered surfaces) for that sub-area. The level of error represented is the highest one (5σ, Table 5.1). Numbers in the right margins indicate error ranges (note differences in scales). Colours range from yellow (lowest) to red (highest).

5.3 Results and Interpretation

5.3.1 Reference Surfaces (No Artefact)

Figure 5.4 shows correlations between the reference bathymetric and terrain attribute surfaces at 10 m resolution and their corresponding surfaces at the four other scales. While no correlation was perfect, bathymetry and topographic mean had an r value very close to 1 for the full extent (r = 0.996 for bathymetry and r = 0.995 for topographic mean at 100 m resolution) and the deeper sub-area (r = 0.999 for both types of surface at 100 m resolution). Overall, topographic mean was always the least impacted by change in spatial resolution, followed by bathymetry. The most impacted was always topographic position, having the sharpest drop in correlation strength between 10 m and 25 m resolution, and consistently the lowest overall correlations. In general, changes in correlation were smaller for the full extent than for the sub-areas, except for bathymetry and topographic mean in the deeper sub-area that showed the smallest changes in correlation. When comparing the two sub-areas, bathymetry, topographic mean and rugosity changed more with spatial resolution in the shallower sub-area than in the deeper one. An opposite trend was observed for measures of easterness, which can be caused by the orientation of the seabed features found in the different areas. No consistent trend was found for northerness, slope and topographic position. Overall, these results show that different characteristics of the terrain are captured when the spatial resolution at which the terrain attributes are derived changes.

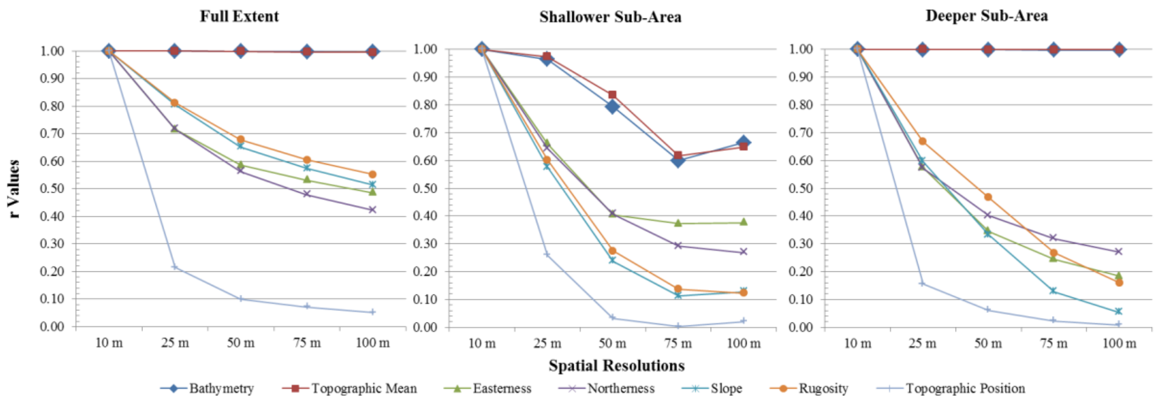

Figure 5.4: Correlations between the 10 m resolution surfaces and the same surfaces computed at different scales. For instance, the 50 m topographic position surface of the full study area is correlated at about 0.10 with the 10 m topographic position surface of the full area.

Table 5.3 summarizes changes in statistics as spatial resolution changed, quantified using linear regression models, which provided the best fit to the data. When looking at the full extent of the study area, significant models adequately explained the observed relationships, having r^2 values ranging from 0.793 to 0.998 (average of 0.935). In general, the range of values of the surfaces decreased as the spatial resolution of the surfaces coarsened, except for easterness and northerness. The mean and standard deviation of slope and northerness decreased with coarser spatial resolution and increased for rugosity. The mean of easterness values decreased with coarser resolutions while their standard deviation increased. In terms of statistical distribution of the values (*i.e.* skewness and kurtosis), bathymetry, topographic position and topographic mean tended towards a normal distribution, easterness and northerness were symmetrical but slightly platykurtic, and rugosity and slope were skewed right and leptokurtic. The distribution of bathymetry, topographic mean, slope and rugosity values did not change with scale. As scale coarsened, the distribution of topographic position became more left-skewed. The skewness and kurtosis values of northerness increased with spatial resolution, while only skewness increased significantly for easterness. Spatial autocorrelation was positive at all scales for all surfaces except for topographic position values that were not autocorrelated. In all cases but northerness, models showed that spatial autocorrelation decreased with coarsening spatial resolution. Finally, the fractal dimension of bathymetry increased with coarser scales, and fractal dimension values of topographic mean were always smaller than those of bathymetry.

Table 5.3: Change in statistical distribution of values with coarsening scales, based on linear regressions. Grey cells indicate non-significant relationships (assessed with the F statistic). White cells indicate significant relationship; the sign in them indicates whether the equations had a positive or negative slope.

		Range	Mean	Standard Deviation	Skewness	Kurtosis	Spatial Autocorrelation	Fractal Dimension
Bathymetry	Full Extent	-					-	+
	Shallower Area						-	+
	Deeper Area	-	-				-	-
Topographic Mean	Full Extent	-					-	
	Shallower Area			-			-	-
	Deeper Area	-	-				-	-
Easterness	Full Extent		-	+	+		-	
	Shallower Area	-		+		-	-	
	Deeper Area		-	-	+	+		
Northerness	Full Extent		-	-	+	+		
	Shallower Area		-	-		+		
	Deeper Area		-	-	+	+		
Slope	Full Extent	-	-	-			-	
	Shallower Area		-	-	-	-	-	
	Deeper Area		-	-	-	-		
Rugosity	Full Extent	-	+	+			-	
	Shallower Area	+	+	+	-	-	-	
	Deeper Area	-	+	+	-	-	-	
Topographic Position	Full Extent	-			-		-	
	Shallower Area		-				+	
	Deeper Area							

Significant models for the sub-areas also had high r^2 values, ranging from 0.774 to 0.997 (average of 0.921). Some differences in the patterns of change in distribution statistics as a function of scale were observed for the sub-areas compared to the full extent. First, the fractal dimension of bathymetry decreased with coarsening spatial resolution in the deeper sub-area, similarly to the fractal dimension of topographic mean in the two sub-areas. Standard deviation of topographic mean values also decreased with coarser resolutions in the shallower sub-area, similarly to the range of easterness values. The standard deviation of easterness values in the deeper sub-area decreased with coarsening scale. The range of rugosity values increased in the shallower sub-area as the resolution became coarser, so did the spatial autocorrelation of topographic position values.

5.3.2 Altered Surfaces

A description of the key results is presented in this section but details are described in Appendix E. Three elements are presented in Appendix E for both the full extent and the two sub-areas. First, the analysis of spatial similarity between the altered surfaces and the reference surfaces, for which 420 regressions were computed, is described. Then, the analysis of error modelling is presented. Overall, 2,436 regressions were computed to describe changes in errors with artefacts. Finally, results of analysis of changes in bathymetric and terrain attributes surfaces caused by the introduction of artefacts are presented. A total of 2,548 regressions were computed to describe changes in surfaces with artefacts.

5.3.2.1 Artefacts

Results indicate that artefacts impact bathymetry and propagate significantly to its derived terrain attributes (*cf.* Appendix E).

The assessment of spatial similarity demonstrated that heave artefacts do not alter the relative spatial distribution of bathymetry and topographic mean but do alter the spatial distribution of values of other terrain attributes. While the mean error, its standard deviation and spatial autocorrelation increased with greater heave, it did not affect the statistical distribution of any of the bathymetric and terrain attributes surfaces. This is explained by the fact that since the heave artefact was induced in a systematic way, it shifted all depth values from the reference DBMs by a given height, thus altering the absolute values of bathymetry but not their relative values (see Franklin *et al.*, 2000; Wise, 2000). Consequently, it did not create actual changes in relief and therefore did not alter values of aspect (*i.e.* easterness and northerness), slope, rugosity and topographic position. In the particular case of heave, the necessity to have controlled conditions (Reuter *et al.*, 2009) to evaluate the impact of artefacts on DBMs and terrain attributes prevented an appropriate representation of what typically occur during surveys. Instead of the constant vertical shift induced in this study, heave artefacts are usually recognized as irregular vertical shifts that can result in a bathymetric surface characterized by waves patterns, as represented in Schmidt *et al.* (2010). Conclusions gained from the analysis of heave are however still transposable to real surveys: Hughes-Clarke (2003b) highlighted the difference between relative and absolute alignment of the different sensors. The former consist of aligning the different sensors in reference to each other, while the latter consist of positioning the different sensors in the platform's frame of reference. Hughes-Clarke (2003b) indicated that the relative alignment is always performed through the patch test, but that the absolute alignment is not consistently addressed in calibration procedure, which is of concern and can result in systematic position biases (Hughes-Clarke, 2003a,b), just like the one introduced by the current study design. Such systematic bias, often seen in other types of remote sensing like LiDAR (Filin, 2003; Lichti & Skaloud, 2010), can be significant for subsequent analyses as they are not obvious (Brown & Bara, 1994). The magnitude and direction of the induced bias need to be known to correct for them; however these are often challenging to determine (Regan *et al.*, 2002; Wilson, 2012).

The study design provided many insights on how systematic pitch, roll and time artefacts influence bathymetric surfaces and their derived terrain attributes. The analysis of spatial similarity between altered and reference surfaces showed that pitch generally had a negative impact on the relative spatial distribution of depth and terrain attributes values. The distribution of errors caused by pitch became less centered on the mean as pitch amplitude increased. Pitch also transformed the studied surfaces by altering their statistical distribution, particularly their mean, standard deviation, and spatial autocorrelation. Pitch decreased spatial autocorrelation of values of most surfaces except topographic mean. The decrease in kurtosis and increase in skewness and standard deviation of northerness and topographic position suggest an increase in extreme values. Error characteristics suggest that pitch impacted the deeper sub-area more than the shallower one, which was also confirmed by the statistical distribution of the values in each surface. This could be caused by the higher density of soundings in shallow water that could attenuate pitch impact, or could result from the increased effect of angular differences with deeper water that shift soundings positions further along-track (*i.e.* fore and aft of the transducer) with increasing water depth. Roll impacted bathymetry and terrain attributes even more than pitch in all three studied elements (spatial similarity, statistical characteristics of errors, and statistical characteristics of depth and terrain attributes values). However, the shallower sub-area appeared to be more impacted by roll than the deeper sub-area, suggesting that a higher density of soundings amplifies the error. This particular result goes against what would be expected from such angular error: in theory, angular errors like those caused by pitch and roll are amplified when the sensor-to-seafloor distance increases. We would thus expect from these errors to have greater effects in the deeper waters, like what was observed for pitch in this study. However, because roll has a greater impact on outer beams, the overlapping areas between survey lines would be most impacted by roll; these observations are thus likely explained by a higher degree of overlap between survey lines in shallower waters compared to in deeper waters. As expected since they induce a shift in the horizontal plane, time artefacts changed the relative spatial distribution of values for all types of surfaces. In terms of statistical distribution of error, bathymetry and topographic mean behaved differently than the other surfaces by becoming more spatially autocorrelated with increasing time artefacts. Time artefacts did very little to the statistical distribution of the different surfaces, except that it increased spatial autocorrelation in easterness, northerness and topographic position values. Although the impacts of time artefacts were subtle compared to those of pitch and roll, the deeper sub-area was more impacted than the shallower sub-area, which could be a consequence of using different ping rates; it is common to use slower ping rates in deeper waters compared to shallower waters as sound takes more time to reach the seafloor and come back towards the sensor.

Our results provided quantitative confirmation of more anecdotal observations previously made in the literature. For instance, while evaluating disparities in bathymetric measurements made from repeated surveys, Roman & Singh (2006) observed that pitch and roll artefacts did induce important errors in bathymetry. In trying to reduce the influence of different errors in bathymetry, Singh *et al.* (2000) noted that roll was often the most significant source of error. Hughes-Clarke *et al.* (1996) mentioned that roll artefacts were more common in deeper waters. Our results showed that while it is true that errors caused by roll in DBMs were greater in deeper waters because of their angular nature, the impacts on the statistical distribution of bathymetric and terrain attributes values were greater in shallower waters, likely because of the overlap between survey lines that amplifies the effect of this type of error.

5.3.2.2 Bathymetry and Terrain Attributes

As expected for environmental data (Legendre, 1993), bathymetry and most terrain attributes were spatially autocorrelated. This was however not true of topographic position, a very spatially heterogeneous terrain attributes (*cf.* Figure 5.1). Likely because of this spatial heterogeneity, topographic position was very often the terrain attribute most impacted by artefacts. There was however one exception to this observation with roll artefacts, for which slope was the most impacted in terms of loss of spatial similarity with the reference slope surfaces, followed by rugosity and topographic position.

Topographic mean, which is directly related to bathymetry (see Chapter 3), was generally the least impacted terrain attribute, and was the only terrain attribute that was less impacted than bathymetric surfaces. This confirms observations made in terrestrial contexts that errors in DTMs are amplified in terrain attributes (Temme *et al.*, 2009; Sofia *et al.*, 2013). For bathymetry and topographic mean, pitch decreased their range and standard deviation of values while roll increased these two statistics. This raises questions about whether or not the combined effect of these two artefacts (*i.e.* if they are both impacting the same DBM) could be smaller than their individual effect. Often there will be multiple problems influencing data, and it is likely that these problems will have a combined effect on the resulting surface. Such effect should be documented in future work. Both pitch and roll artefacts artificially increased measures of fractal dimension, making surface representations (*i.e.* DBMs) appear more complex than they actually are. Roll also decreased spatial autocorrelation and increased standard deviation of depth values, likely because of the introduction of local outlier values.

5.3.2.3 Spatial Scale

Results yielded several insights in terms of scale. First, the description of the reference surfaces confirmed that spatial autocorrelation is scale-dependent (Meentemeyer, 1989; Legendre, 1993), changing with both resolution and extent (more autocorrelated at finer resolutions and greater extent). In terms of spatial similarity, results showed that when the same level of artefact is affecting the data, finer-scale altered surfaces are more dissimilar to the reference surfaces than broader-scale altered surfaces. The spatial similarity between bathymetric and topographic mean surfaces of different scales was particularly impacted in the shallower sub-area. As expected, the descriptive statistics of bathymetric and terrain attributes surfaces, combined with the multiscale assessment of spatial similarity, demonstrated that the range of values decreases as scale broadens, which results in the attenuation of the observed terrain characteristics. This phenomenon has been described many times before, particularly in a context of coarse-graining, both in terrestrial (*e.g.* Chow & Hodgson, 2009; Grohmann, 2015) and marine DTMs (*e.g.* Wilson *et al.*, 2007; Rengstorf *et al.*, 2012), although less for terrain attributes other than slope. In one of the earliest study, Evans (1980) reported a decrease in mean and standard deviation of slope with coarser scales, which was confirmed by our results. Northerness also behaved that way, but interestingly easterness' standard deviation increased with coarser scales. This is likely caused by the fact that northerness and easterness are the only two measures that cover their entire range of possible values (-1 to 1), and that coarsening the scale increases the frequency of the minimum and maximum values, which represent for easterness slopes fully oriented towards the East of the West. Because of the anisotropy in the dataset, for instance caused by the decrease in depth values as one moves west through the study area and which causes more slope values to be oriented west, may lead to more extreme measures of easterness and explain the increase in standard deviation. Also, the orientation of the survey line may cause the artefacts to appear in a specific direction, thus inducing local anisotropy that would impact values of easterness and northerness. Measures of rugosity also behaved differently: both mean and standard deviation increased with coarser scales, which is in line with the increase in fractal dimension of bathymetric surfaces for both the full extent and the shallower sub-area as spatial resolution coarsens. Roll also had a greater impact on the complexity of coarser-scale DBMs, as shown by measures of fractal dimension. In the shallower sub-area, the mean of topographic position also increased with coarser scales. The deeper sub-area sometimes presented a different trend from observations made from the shallower sub-area. For instance, when roll artefacts were propagating to aspect, slope and rugosity, measures of spatial similarity were lower at coarser scales than at finer scales. In terms of error, error values were more spatially autocorrelated in finer-scale surfaces than in coarser-scale surfaces.

5.4 Discussion

5.4.1 Comparisons with Other Studies: Terrestrial and Marine

As expected, results confirm that higher DBMs error increases the error in derived terrain attributes (Oksanen & Sarjakoski, 2005), although the relationship was not as clear for heave artefacts when analyzing smaller spatial extents. As commonly performed in error propagation analyses (Wise, 2011b), the study design implemented in the current paper assumed that the error had the same characteristics across the entire study area, *i.e.* it was systematically introduced. However, it is much likely that in a real context, seafloor characteristics (*e.g.* complexity, orientation of natural features), which vary spatially, have an influence on how the error is revealed. This was also suggested by the observed differences between the shallower and deeper sub-areas. Wise (2011a, b) discussed cases where the consideration of the nature of the terrain yielded better error model. However, these results may not be directly applicable in the marine environment, as errors are different in nature and distribution, *e.g.* in terms of artefacts having a greater influence than random noise (Hughes-Clarke, 2003a). These differences are partly due by the different technologies used to collect terrestrial and marine DTMs, for instance in terms of surveying geometry, dependency to ancillary data, and characteristics of sound propagation in water.

In terrestrial settings, it is well known that DTM errors, including artefacts and random noise, are spatially autocorrelated (Holmes *et al.*, 2000; Temme *et al.*, 2009; Leon *et al.*, 2014). Our results confirm that pitch and roll errors were autocorrelated in the bathymetry and that the error from all four artefacts was autocorrelated in the topographic mean, which is a generalization of bathymetry (*cf.* Chapter 3). Errors in terrain attributes were however not spatially autocorrelated, except for some exceptions. In general, error values were becoming more autocorrelated and closer to a normal distribution as the level of artefact was increasing. While looking at a different type of artefacts, namely those caused by interpolation, Wise (2011a) found that when data density was low, DTM error was more spatially autocorrelated than when data density was higher. This was confirmed by our study, for which spatial autocorrelation values were in average higher among errors in the deeper sub-area than in the shallower sub-area for the four types of artefacts, except for topographic position. Wise (2011a) also mentioned that the distribution of errors in lower density data was closer to a normal distribution than for higher density data that had more of a leptokurtic distribution. Our results agreed with these conclusions only for heave errors, and to a much lesser extent for roll errors. Pitch errors did not show a significant pattern, and time errors in bathymetry were much more leptokurtic in the deeper sub-area (low density) than in the shallower one (high density). This could be explained by the fact that time affects the boundary of different survey lines, and consequently the area that overlaps. In shallower waters, there was more overlap between survey lines than in the deeper waters, which could indicate that it is more likely that the shallower sub-area get higher error values that would result in an error distribution less centered on the mean than in the deeper sub-area. Wise (2011a) results also showed that error distribution in DTMs was centered on zero. Our results confirm those of Wise (2011a), except for time errors that were highly right-skewed, likely because of the use of the absolute value of the differences. Results are however in line with other studies that gave DTM error distribution longer tails (*e.g.* Kyriakidis *et al.*, 1999; Bonin & Rousseaux, 2005; Oksanen & Sarjakoski, 2006). Wise (2011a) discussed the disagreements within the literature regarding the statistical distribution of DTM errors, and concluded that the situation may be more complex than what is discussed in that body of literature. The author also found that the distribution of errors for slope, aspect, and curvature followed a leptokurtic distribution, which was true for our results for all types of artefacts but roll, for which easterness and northerness in the deeper sub-area and the full extent had a kurtosis similar to that of a normal distribution. Slope in the shallower sub-area also had such distribution. Topographic position, that is related to curvature (*cf.* Chapter 3), had indeed a leptokurtic distribution. Roll also impacted measures of aspect and topographic position more in shallower waters.

Regarding specific terrain attributes, previous research has mainly focused on slope and curvature. Csillik *et al.* (2015) argued that the statistical distribution of slope and curvature are often long-tailed; our results confirmed it for slope, and showed that it was also the case for rugosity. However, the values of other terrain attributes tested in our study were normally distributed. Curvature was previously found to be very sensitive to interpolation artefacts (Wise, 2007), and results from the current study confirmed that topographic position (related to curvature), was often the most impacted of terrain attributes. Interesting parallels can be made between the current study and results from Sofia *et al.* (2013), who studied the presence of outliers and striping artefacts in LiDAR DTM and their propagation to minimum curvature. First, they showed that the distribution of curvature changed with scale in their reference surfaces. Our results also showed that change in scale affect the distribution of terrain attributes (Figure 5.4), although it does not affect bathymetry as much. Among other similarities with our study, Sofia *et al.* (2013) concluded that the modelling of errors in relation to spatial scale was crucial in obtaining reliable curvature measurements. They also showed that errors had a greater effect on curvature at finer scales. Finally, they showed that “when DTMs include errors, curvature distributions become controlled by these errors, whose propagation depends on error distribution, error spatial correlation, and the scale of analysis” (Sofia *et al.*, 2013, p. 1116).

5.4.2 Implications for Marine Geomorphometry

Artefacts in DBMs are problematic in marine geomorphometry studies for several reasons. First, results suggest in many ways that artefacts will potentially impact subsequent analyses using the bathymetry and terrain attributes. For instance, spatial autocorrelation is known to affect predictive models and observed relationships between marine species and their environment (Foody, 2004; Hotorn *et al.*, 2011), and our results showed that spatial autocorrelation levels vary with the presence of artefacts. Second, while future developments in MBES technology and processing software may help improve data quality, the potential for error in any survey will never be fully eliminated. Artefacts are usually within error specifications (Hughes-Clarke, 2003a) and appear even when appropriate calibrations were completed and motions compensated for (Erikstad *et al.*, 2013). Similarly to some errors in terrestrial DTMs (Wilson, 2012), artefacts in marine DTMs cannot always be removed. Our results also showed that the errors induced by different levels of artefacts are not correlated, indicating that the distribution of error changes depending on the amplitude of the artefact. This prevents proper predictions of the spatial distribution of errors that could have helped correcting or accounting for them. Third, underwater artefacts may be more challenging to identify and correct than their terrestrial counterparts: they are more likely to resemble natural features such as sand waves than the strips, terraces or bands typically found in terrestrial DTM (Fisher & Tate, 2006). When visual validation (*e.g.* from video data) is not available to confirm if such pattern in the data are natural, it may be erroneous to apply filters or other techniques to remove these features, especially as some of these techniques may themselves alter the initial surface and consequently its derived terrain attributes (Franklin *et al.*, 2000). Fourth, the options available to analyse error propagation are more limited in marine DTMs than in terrestrial DTMs. Analysis in terrestrial settings are often performed by comparing the DTM to reference data (*e.g.* GPS ground data) that are considered true or more accurate than the DTM (Shortridge, 2001), which is logistically very complicated underwater and thus rarely, if not ever implemented (Roman & Singh, 2006). Podobnikar (2009) argued that visual methods of evaluation of spatial data are underused in terrestrial environments, but they are often the only available methods in the marine environments (Erikstad *et al.*, 2013). Finally, marine geomorphometry studies often strongly rely on expensive, remotely sensed MBES bathymetric data. In waters too deep for bathymetric LiDAR or optical remote sensing, MBES are among the only systems available that can provide high enough resolution data for many applications (Brown *et al.*, 2011; Lecours *et al.*, 2016). There is often no alternative to these datasets because most places in the marine environment are difficult to access, observe, and sample (Solan *et al.*, 2003; Robinson *et al.*, 2011).

These arguments highlight that despite the presence of artefacts, most research and applications are dependent on such DBMs. For this reason it is critical to understand the impacts of artefacts on the derivation of terrain attributes to enable their proper consideration in the interpretation of results. It is becoming crucial to integrate error modelling to the marine geomorphometry workflow, at least as an informative tool to make cognizant decisions regarding fitness for use (Fisher & Tate, 2006). Fitness for use is the concept of determining whether or not a dataset is of sufficient quality for a particular purpose (Devillers & Jeansoulin, 2006). Agumya & Hunter (2002) estimated that for a long time end-users would consider the assessment of fitness for use unnecessary when they would not have any alternative data available like it is often the case for MBES data. However, since artefacts are likely to propagate into analyses past the generation of terrain attributes, it is fundamental to assess the quality of the DTM and its derived terrain attributes before they are included in further analyses (Wechsler & Kroll, 2006; Van Niel & Austin, 2007). Such assessment will help evaluate fitness for use and situate analyses and interpretation of results on a sound inferential basis. This is particularly important when there is a decision-making process involved at the end of the workflow, for instance when terrain attributes are used to build habitat maps to assist in conservation, or to make geomorphological maps used for navigation or dredging. In the end, the assessment of fitness for use requires an intended purpose and specific application goals or questions (Devillers *et al.*, 2002; Rocchini *et al.*, 2011; Wilson, 2012; Passalacqua *et al.*, 2015). In the context of using terrain attributes impacted by artefacts in marine habitat mapping exercises, Vierod *et al.* (2014, p. 14) stated that "a missing predictor may be less detrimental to a models prediction than a distorted predictor containing false information". Erikstad *et al.* (2013) demonstrated that poor quality terrain attributes were not fit for their use as they prevented the proper identification of terrain types associated with the geology and marine habitats of their study area.

An important characteristic of the assessment of fitness for use is the consideration of spatial scale. While Grohmann (2015) recommended using DTMs with the highest resolution available to derive terrain attributes, many others claim that in practice, the highest resolution DTM may not be the most suitable based on the successful identification, or not, of the features of interest in a particular area and for a particular purpose (Cavazzi *et al.*, 2013; Lecours *et al.*, 2015b). Our results confirm the latter opinion: the spatial similarity analysis clearly demonstrated that many terrain attribute surfaces were not highly correlated with their corresponding surfaces computed at different scales, thus capturing and describing different seafloor features and characteristics. Our results also demonstrate that in marine geomorphometry, attempting to match the observational scale with the geomorphological scale of the features of interest should not be the only factor when deciding at which resolution a DTM should be generated. Since the influence of DTM error was stronger on higher resolution bathymetric surfaces and terrain attributes, there is a trade-off between higher-resolution information that has for instance the potential to improve models prediction, and spatial data quality that could potentially worsen predictions. Such trade-off should be part of defining fitness for use. Podobnikar (2009) mentioned that the higher the spatial resolution of a DTM, the more difficult is the evaluation of its quality. We also note that another characteristic of spatial scale that was not considered in this study is the analytical scale, which in this case refer to the size of the window of analysis used to derive terrain attributes. Albani *et al.* (2004) showed that using a greater window of analysis diminished the influence of errors in terrestrial DTMs, which was also shown in the current study by the lower impact of artefacts on topographic mean compared to bathymetry. This element should also be considered in the evaluation of the trade-off between scale and quality (Wilson *et al.*, 2007). Finally, results showed that errors caused by artefacts in bathymetry are sometimes amplified in the terrain attributes, which highlights the idea that if a DBM is deemed fit for use, it does not mean that the terrain attributes derived from it will also be fit for use.

5.5 Conclusions

Florinsky (1998) listed several factors on which depends the quality of terrain attributes: the roughness of the terrain, sampling density, spatial resolution, interpolator, vertical resolution and the type of geomorphometric analysis performed (*i.e.* the types of terrain attributes extracted and the algorithms used). Our study looked at the impact of artefacts on terrain attributes by considering three of the factors listed by Florinsky (1998): sampling density, spatial resolution, and the types of terrain attributes. The interpolator was kept consistent across the study design, and the assessment of data quality as a function of variations in seafloor characteristic – like roughness – is identified as a future area of research later in this section. Results showed that artefacts do impact bathymetry and terrain attributes, and that the level of impact depends on the three studied factors. Time and pitch artefacts had a greater impact on surfaces having a lower sampling density, while roll artefacts had a greater impact on surfaces created with a higher sampling density. In terms of spatial resolution, finer-scale data were generally more impacted by artefacts than broader-scale data. For terrain attributes, topographic mean was most often the least impacted terrain attributes while topographic position was often the most impacted. Results confirmed that errors in bathymetry do not only propagate to terrain attributes but are amplified in DBMs derivatives.

Based on the results, it is difficult to make clear recommendations on which terrain attributes users should select when they are aware of the presence of artefacts in their data. However, as topographic mean was less impacted by artefacts than bathymetry, we recommend using the former over the latter when artefacts are prevalent, which confirms a hypothesis made in Chapter 3 that topographic mean may be more reliable than bathymetry if the DTM is noisy or of poor quality. The selection of terrain attributes can be better informed if the cause of the artefacts can be identified. For instance, measures of aspect (*i.e.* easterness and northerness) were most sensitive to pitch artefacts, slope and rugosity to roll and pitch artefacts, and topographic position to time artefacts. In all cases, users need to carefully evaluate the impacts of artefacts in their DBM on the derivation of terrain attributes in order to make more informed decisions on whether or not terrain attributes are suitable for their application.

Future work should carry forward the analysis performed in this paper by (1) studying the impact of other types of artefacts on bathymetry and terrain attributes, (2) looking at inducing artefacts of random amplitude across the study area, (3) evaluate if the combined effect of different artefacts is additive, multiplicative, or if they cancel or attenuate each other's influence, and (4) expand the analysis of sub-areas to quantify the effects of artefact with varying terrain characteristics (*e.g.* roughness, depth). With more knowledge on the impacts of artefacts in DTMs on terrain attributes, we can then move further into improving our understanding of artefact propagation into the workflow of specific applications of marine geomorphometry, like habitat mapping or hydrodynamics modelling. This would enable a better quantification of uncertainty and error associated with models produced using bathymetric data with artefacts. Only from that point onward will we be able to establish application-related practices, protocols and standards for fitness for use evaluation. In the meantime, better practices should be implemented, including the assessment and acknowledgement of artefacts and other types of errors in both the bathymetry and its derived terrain attributes, and the potential implications of these errors for the analysis (Rocchini *et al.*, 2011; Lecours *et al.*, 2016).

In conclusion, understanding data limitations is crucial when integrating them into a workflow (Brown *et al.*, 2005) like in geomorphometry. Since MBES bathymetric data are often among the only available data for studying seafloor environments and are frequently impacted by artefacts, there is a need to find ways to efficiently report spatially explicit error and uncertainty assessments for both DBMs and derived terrain attributes. Such assessments will become important informative tools to make cognizant decisions regarding fitness for use and enable a better understanding of potential implications for analyses. This will ultimately encourage responsible use of data in research and applications.

5.6 Literature Cited

Agumya, A., & Hunter, G.J. (2002) Responding to the consequences of uncertainty in geographical data. *International Journal of Geographical Information Systems*, 16:405-417.

Albani, M., & Klinkenberg, B. (2003). A spatial filter for the removal of striping artifacts in digital elevation models. *Photogrammetric Engineering & Remote Sensing*, 69:755-765.

Albani, M., Klinkenberg, B., Andison, D.W., & Kimmins, J.P. (2004). The choice of window size in approximating topographic surfaces from Digital Elevation Models. *International Journal of Geographical Information Systems*, 18:577-593.

Bangen, S.G., Wheaton, J.M., Bouwes, N., Bouwes, B., & Jordan C. (2014) A methodological intercomparison of topographic survey techniques for characterizing wadeable streams and rivers. *Geomorphology*, 206:343-361.

Bonin, O., & Rousseaux, F. (2005) Digital terrain model computation from contour lines: how to derive quality information from artifact analysis. *GeoInformatica*, 9:253-268.

Brown, C. J., Sameoto, J. A., & Smith, S. J. (2012). Multiple methods, maps, and management applications: purpose made seafloor maps in support of ocean management. *Journal of Sea Research*, 72:1-13.

Brown, C. J., Smith, S. J., Lawton, P., & Anderson, J. T. (2011). Benthic habitat mapping: a review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuarine, Coastal and Shelf Science*, 92:502-520.

Brown, D.G., & Bara, T.J. (1994) Recognition and reduction of systematic error in elevation and derivative surfaces from 71/2-minute DEMs. *Photogrammetric Engineering & Remote Sensing*, 60:189-194.

Brown, J.D., Heuvelink, G.B., & Refsgaard, J.C. (2005) An integrated methodology for recording uncertainties about environmental data. *Water Science and Technology*, 52:153-160.

Calder, B.R., & Mayer, L.A. (2003) Automatic processing of high-rate, high-density multibeam echosounder data. *Geochemistry, Geophysics, Geosystems*, 4:1048-1070.

Cavazzi, S., Corstanje, R., Mayr, T., Hannam, J., & Fealy, R. (2013) Are fine resolution digital elevation models always the best choice in digital soil mapping? *Geoderma*, 195-196:111-121.

Chow, T.E., & Hodgson, M.E. (2009) Effects of lidar post-spacing and DEM resolution to mean slope estimation. *International Journal of Geographical Information Systems*, 23:1277-1295.

Collins, C.A., Clynch, J.R., & Rago, T.A. (2005) Comparison of SWATH and monohull vessel motion for regional class research vessels. *Office of Naval Research*, Technical Report #NPS-OC-05-001, 61 p.

Csillik, O., Evans, I.S., & Drăguţ, L. (2015) Transformation (normalization) of slope gradient and surface curvatures, automated for statistical analyses from DEMs. *Geomorphology*, 232:65-77.

Cushman-Roisin, B., & Beckers, J.-M. (2011) *Introduction to geophysical fluid dynamics, 2nd edition*. Academic Press, 875 p.

De Moustier, C., & Kleinrock, M.C. (1986) Bathymetric artefacts in Sea Beam data: how to recognize them and what causes them. *Journal of Geophysical Research*, 91:3407-3424.

Devillers, R., Gervais, M., Bédard, Y., & Jeansoulin, R. (2002) Spatial data quality: from metadata to quality indicators and contextual end-user manual. *OEEPE/ISPRS Joint Workshop on Spatial Data Quality Management*, March 21-22, Istanbul, p. 45-55.

Devillers, R., & Jeansoulin, R. (2006) *Fundamentals of spatial data quality*, ISTE, London, U.K., 312 p.

DFO (2006). Presentation and review of Southwest Nova Scotia benthic mapping project. *DFO Canadian Science Advisory Secretariat*, Proceedings Series 2006/047.

Dolan, M.F.J., & Lucieer, V.L. (2014) Variation and uncertainty in bathymetric slope calculations using geographic information systems. *Marine Geodesy*, 37:187-219.

Dunn, D.C., & Halpin, P.N. (2009) Rugosity-based regional modeling of hard-bottom habitat. *Marine Ecology Progress Series*, 377:1-11.

Erikstad, L., Bakkestuen, V., Bekkby, T., & Halvorsen, R. (2013). Impact of scale and quality of digital terrain models on predictability of seabed terrain types. *Marine Geodesy*, 36:2-21.

Evans, B.J. (1997) Dynamic display of spatial data-reliability: does it benefit the map user? *Computers & Geosciences*, 23:409-422.

Evans, I.S. (1980) An integrated system of terrain analysis and slope mapping. *Zeitschrift für Geomorphologie Supp. Bd.*, 36:274-295.

Filin, S. (2003) Recovery of systematic biases in laser altimetry data using natural surfaces. *Photogrammetric Engineering and Remote Sensing*, 57:1235-1242.

Fisher, P. (1998) Improved modelling of elevation error with geostatistics. *GeoInformatica*, 2-3:215-233.

Fisher, P. F., & Tate, N. J. (2006). Causes and consequences of error in digital elevation models. *Progress in Physical Geography*, 30:467-489.

Florinsky, I.V. (1998) Accuracy of local topographic variables derived from digital elevation models. *International Journal of Geographical Information Science*, 12:47-61.

Foody, G.M. (2004) Spatial nonstationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. *Global Ecology and Biogeography*, 13:315-320.

Franklin, J., McCullough, P., & Gray, C. (2000) Terrain variables used for predictive mapping of vegetation communities in Southern California. In: Wilson, J.P., & Gallant, J.C. (Eds), *Terrain Analysis: Principles and Applications* (pp. 331-353). New York: John Wiley & Sons.

Georgian, S., Shedd, W., & Cordes, E.E. (2014) High-resolution ecological niche modelling of the cold-water coral Lophelia pertusa in the Gulf of Mexico. *Marine Ecology Progress Series*, 506:145-161.

Gessler, P., Pike, R., MacMillan, R. A., Hengl, T., & Reuter, H. I. (2009). The future of geomorphometry. In: Hengl, T., & Reuter, H. I. (Eds), *Geomorphometry: Concepts, Software, Applications* (pp. 637-652). Developments in Soil Science, 33. Amsterdam: Elsevier.

Gille, S.T., Metzger, E.J., & Tokmakian, R. (2004) Seafloor topography and ocean circulation. *Oceanography*, 17:47-54.

Grohmann, C.H. (2015) Effects of spatial resolution on slope and aspect derivation for regional-scale analysis. *Computers & Geosciences*, 77:111-117.

Hengl, T., & Reuter H.I. (2009) Geomorphometry: concepts, software, applications. *Developments in Soil Science*, 33, Elsevier, Amsterdam, 765 p.

Heuvelink, G. B. M. (1998). *Error propagation in environmental modelling with GIS*, London: Taylor and Francis, 127 p.

Holmes, K.W., Chadwick, O.A., & Kyriakidis (2000) Error in a USGS 30-meter digital elevation model and its impact on terrain modelling. *Journal of Hydrology*, 233:154-173.

Horn, B.K.P. (1981) Hill shading and the reflectance map. *Proceedings of the IEEE*, 69:14-47.

Hotorn, T., Müller, J., Schröder, B., Kneib, T., & Brandl, R. (2011) Decomposing environmental, spatial, and spatiotemporal components of species distributions. *Ecological Monographs*, 81:329-347.

Hughes-Clarke, J.E. (1997) Are you really getting “full bottom coverage”? http://www.omg.unb.ca/~jhc/coverage_paper.html

Hughes-Clarke, J.E. (2002) The challenge of technology: improving sea-floor mapping methodologies. http://www.omg.unb.ca/AAAS/UNB_Seafloor_Mapping.html

Hughes-Clarke, J.E. (2003a) Dynamic motion residuals in swath sonar data: ironing out the creases. *International Hydrographic Review*, 4:6-23.

Hughes-Clarke, J.E. (2003b) A reassessment of vessel coordinate systems: what is it that we are really aligning? *US Hydrographic Conference 2003*, Biloxi, MS, 1-12.

Hughes-Clarke, J.E., Mayer, L.A., & Wells, D.E. (1996) Shallow-water imaging multibeam sonars: a new tool for investigating seafloor processes in the coastal zone and on the continental shelf. *Marine Geophysical Researches*, 18:607-629.

Jianhu, Z., & Jingnan, L. (2003) Development of method in precise multibeam acoustic bathymetry. *Geo-spatial Information Science*, 6:71-74.

Kostylev, V.E., Erlandsson, J., Ming, M.Y., & Williams, G.A. (2005) The relative importance of habitat complexity and surface area in assessing biodiversity: fractal application on rocky shores. *Ecological Complexity*, 2:272-286.

Kyriakidis, P.C., Shortridge, A.M., & Goodchild, M.F. (1999) Geostatistics for conflation and accuracy assessment of digital elevation models. *International Journal of Geographical Information Science*, 13:677-707.

Landmark, K., Solberg, A.H.S., Albregtsen, F., Austeng, A., & Hansen, R.E. (2015) A radon-transformed-based image noise filter – with applications to multibeam bathymetry. *IEEE Transactions on Geoscience and Remote Sensing*, 53:6252-6273.

Lecours, V. (2015) Terrain Attribute Selection for Spatial Ecology (TASSE), v.1.0, URL www.marinegis.com.

Lecours, V., & Devillers, R. (2015). Assessing the spatial data quality paradox in the deep-sea. In: Sieber, R. E. (Ed.), *Proceedings of Spatial Knowledge and Information – Canada 2015*, pp. 1-8.

Lecours, V., Devillers, R., Schneider, D.C., Lucieer, V.L., Brown, C.J., & Edinger, E.N. (2015b) Spatial scale and geographic context in benthic habitat mapping: review and future directions. *Marine Ecology Progress Series*, 535:259-284.

Lecours, V., Dolan, M.F.J., Micallef, A., & Lucieer, V.L. (2016) A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrology and Earth System Science*, 20:3207-3244.

Lecours, V., Lucieer, V.L., Dolan, M.F.J., & Micallef, A. (2015a) An ocean of possibilities: applications and challenges of marine geomorphometry. In Jasiewicz, J., Zwolinski, Z., Mitasova, H., & Hengl, T. (Eds), *Geomorphometry for Geosciences* (pp. 23-26), Adam Mickiewicz University in Poznan – Institute of Geoecology and Geoinformation, Poland.

Lecours, V., Miles, L.L., Devillers, R., & Edinger, E.N. (2013) Data analysis towards the multiscale characterization of cold-water coral and sponge habitats in Canadian waters. *Technical Report REQ. No. F6160-120010*. Department of Fisheries and Oceans Canada, Newfoundland and Labrador Region, St. John's, 260 p.

Legendre, P. (1993) Spatial autocorrelation: problem or new paradigm. *Ecology*, 74:1659-1673.

Leon, J.X., Heuvelink, G.B.M., & Phinn, S.R. (2014) Incorporating DEM uncertainty in coastal inundation mapping. *PLoS ONE*, 9:e108727.

Lichti, D.D., & Skaloud, J. (2010), Registration and calibration. In: Vosselman, G., & Maas, H.-G. (Eds), *Airborne and terrestrial laser scanning* (pp. 83-133). CRC Press Inc.

Lindsay, J., & Creed, I.F. (2006) Distinguishing actual and artefact depressions in digital elevation data. *Computers & Geosciences*, 32:1192-1204.

Lucieer, V., Barrett, N., Hill, N., Nichol, S.L. (2012) Characterization of shallow inshore coastal reefs on the Tasman Peninsula, Southeastern Tasmania, Australia. In: Harris, P.T., & Baker, E.K. (Eds), *Seafloor geomorphology as benthic habitat* (pp. 481-492). Amsterdam: Elsevier.

Lurton, X. (2010) *An introduction to underwater acoustics: principles and applications, 2nd edition*. Springer/Praxis Publishing, Berlin, Germany, 704 p.

Meentemeyer, V. (1989) Geographical perspective of space, time, and scale. *Landscape Ecology*, 3:163-173.

Oksanen, J., & Sarjakoski, T. (2005). Error propagation of DEM-based surface derivatives. *Computers & Geosciences*, 31:1015-1027.

Oksanen, J., & Sarjakoski, T. (2006). Uncovering the statistical and spatial characteristics of fine toposcale DEM error. *International Journal of Geographical Information Science*, 20:345-369.

Passalacqua, P., Belmont, P., Staley, D.M., Simley, J.D., Arrowsmith, J.R., Bode, C.A., Crosby, C., DeLong, S.B., Glenn, N.F., Kelly, S.A., Lague, D., Sangireddy, H., Schaffrath, K., Tarboton, D.G., Wasklewicz, T., & Wheaton, J.M. (2015) Analyzing high resolution topography for advancing the understanding of mass and energy transfer through landscapes: a review. *Earth-Science Reviews*, 148:174-193.

Pike, R.J., Evans, I.S., & Hengl, T. (2009) Geomorphometry: a brief guide. In: Hengl, T., & Reuter, H. I. (Eds), *Geomorphometry: Concepts, Software, Applications* (pp. 3-30). Developments in Soil Science, 33. Amsterdam: Elsevier.

Podobnikar, T. (2009) Methods for visual quality assessment of a digital terrain model. *S.A.P.I.E.N.S*, 2:1-10.

Regan, H. M., Colyvan, M., & Burgman, M. A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications*, 12:618-628.

Rengstorf, A.M., Grehan, A., Yesson, C., & Brown, C. (2012) Towards high-resolution habitat suitability modeling of vulnerable marine ecosystems in the deep-sea: resolving terrain attribute dependencies. *Marine Geodesy*, 35:343-361.

Reuter, H.I., Hengl, T., Gessler, P., & Soille, P. (2009) Preparation of DEMs for geomorphometric analysis. In: Hengl, T., & Reuter, H. I. (Eds), *Geomorphometry: Concepts, Software, Applications* (pp. 87-120). Developments in Soil Science, 33. Amsterdam: Elsevier.

Robinson, L. M., Elith, J., Hobday, A. J., Pearson, R. G., Kendall, B. E., Possingham, H. P., & Richardson, A. J. (2011). Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. *Global Ecology and Biogeography*, 20:789-802.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J.M., Jinénez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. (2011) Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Progress in Physical Geography*, 35:211-226.

Roman, C., & Singh, H. (2006) Consistency based error evaluation for deep sea bathymetric mapping with robotic vehicles. *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, Orlando, Florida, 3568-3574.

Rousseaux, F. (2003). Étude de l’impact de la représentation du relief sur les applications. *Revue Internationale de Géomatique*, 13:493-509.

Schmidt, V., Raineault, N.A., Skarke, A., Trembanis, A., & Mayer, L.A (2010) Correction of bathymetric survey artifacts resulting from apparent wave-induced vertical position of an AUV. *Center for Coastal and Ocean Mapping, University of New Hampshire*, paper 793, 1-13.

Shortridge, A. (2001). Characterizing uncertainty in digital elevation models. In: Hunsaker, C. T., Goodchild, M. F., Friedl, M. A., & Case, T. J. (Eds), *Spatial uncertainty in ecology: Implications for remote sensing and GIS applications* (pp. 238-257). New York: Springer-Verlag

Singh, H., Whitcomb, L., Yoerger, D., & Pizarro, O. (2000) Microbathymetric mapping from underwater vehicles in the deep ocean. *Computer Vision and Image Understanding*, 79:143-161.

Smith, T. A., & McConnaughey, R. A. (2016). The applicability of sonars for habitat mapping: a bibliography. *U.S. Dep. Commer., NOAA Tech. Memo*. NMFS-AFSC-317, 129 p.

Sofia, G., Pirotti, F., & Tarolli, P. (2013). Variations in multiscale curvature distribution and signatures of LiDAR DTM errors. *Earth Surface Processes and Landforms*, 38:1116-1134.

Solan, M., Germano, J. D., Rhoads, D. C., Smith, C., Michaud, E., Parry, D., Wenzhöfer, F., Kennedy, B., Henriques, C., Battle, E., Carey, D., Iocco, L., Valente, R., Watson, J., & Rosenberg, R. (2003). Towards a greater understanding of pattern, scale and process in marine benthic systems: a picture is worth a thousand worms. *Journal of Experimental Marine Biology and Ecology*, 285-286:313-338.

Temme, A.J.A.M. , Heuvelink, G.B.M., Schoorl, J.M., & Claessens, L. (2009). Geostatistical simulation and error propagation in geomorphometry. In: Hengl, T., & Reuter, H. I. (Eds), *Geomorphometry: Concepts, Software, Applications* (pp. 121-140). Developments in Soil Science, 33. Amsterdam: Elsevier.

Tong, R., Purser, A., Guinan, J., & Unnithan, V. (2013) Modelling the habitat suitability of deep-water gorgonian corals based on terrain variables. *Ecological Informatics*, 13:123-132.

Van Niel, K.P., & Austin, M.P. (2007) Predictive vegetation modeling for conservation: impact of error propagation from digital elevation data. *Ecological Applications*, 17:266-280.

Van Niel, K.P., Laffan, S.W., & Lees, B.G. (2004) Effect of error in the DEM on environmental variables for predictive vegetation modelling. *Journal of Vegetation Science*, 15:747-756.

Vierod, A.D.T., Guinotte, J.M., & Davies, A.J. (2014) Predicting the distribution of vulnerable marine ecosystems in the deep sea using presence-background models. *Deep-Sea Research II*, 99:6-18.

Wechsler, S.P., & Kroll, C.N. (2006) Quantifying DEM uncertainty and its effect on topographic parameters. *Photogrammetric Engineering and Remote Sensing*, 9:1081-1090.

Wilson, J.P. (2012) Digital terrain modelling. *Geomorphology*, 137:107-121.

Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., & Grehan, A.J. (2007) Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30:3-35.

Wise, S. (2000) Assessing the quality for hydrological applications of digital elevation models derived from contours. *Hydrological Processes*, 14:1909-1929.

Wise, S. (2007) Effect of differing DEM creation methods on the results from a hydrological model. *Computers & Geosciences*, 33:1351-1365.

Wise, S. (2011a) Cross-validation as a means of investigating DEM interpolation error. *Computers & Geosciences*, 37:978-991.

Wise, S. (2011b) Geomorphometry and DEM error. In: Hengl, T., Evans, I.S., Wilson, J.P., & Gould, M. (Eds), *Geomorphometry 2011* (pp. 41-44), Redlands, CA.

Yang, F., Li, J., Wu, Z., Jin, X., Chu, D., & Kang, Z. (2007) A post-processing method for the removal of refraction artifacts in multibeam bathymetry data. *Marine Geodesy*, 30:235-247.

Zhao, J., Zhang, H., & Hughes-Clarke, J. (2007) Determination of precise instantaneous height at multibeam transducer. *Geo-spatial Information Science*, 10:250-254.

Zhang, J., Atkinson, P.M., & Goodchild, M. (2014) *Scale in spatial information and analysis*. CRC Press, Boca Raton, FL, 353 p.

Zhou, Q., & Liu, X. (2004) Analysis of errors of derived slope and aspect related to DEM data properties. *Computers & Geosciences*, 30:369-378.

6. Influence of Artefacts in Digital Terrain Models on Habitat Maps and Species Distribution Models: A Multiscale Assessment

6.1 Introduction

In the last decades, remote sensing has become the main method used for collecting elevation data used in the production of Digital Terrain Models (DTM); methods used to collect elevation data include both passive, optical techniques (*e.g.* photogrammetry, stereoscopy) and active techniques (*e.g.* interferometric synthetic aperture Radar, LiDAR). All DTMs carry a certain level of error (Gessler *et al.*, 2009) caused by random noise, systematic errors and artefacts (Wise, 2000). Artefacts were characterized by Reuter *et al.* (2009, p. 91) as "distinct erratic features" that are made of improbable and incorrect values. Artefacts can be found in DTMs collected from any remote sensing systems (Fisher & Tate, 2006; Sofia *et al.*, 2013) and at all scales. While it is often assumed that higher resolution data (*e.g.* collected with airborne LiDAR) result in higher quality DTMs (Nelson *et al.*, 2009), high resolution DTMs are more sensitive to survey conditions (Su & Bork, 2006) and more prone to errors like artefacts (Podobnikar, 2009; Zandbergen, 2011). Artefacts can be induced by the interpolation method used to create the DTM (Sofia *et al.*, 2013), the motion and location of the acquisition platform (Harrison *et al.*, 2009), timing or log frequency issue in the surveying system (Lecours & Devillers, 2015) or a lack of or an inappropriate correction of ionospheric and atmospheric conditions (Li & Goldstein, 1990). Overall, artefacts can be problematic as they influence data quality more than other types of errors (Rousseaux, 2003) and can be very subtle in the DTM (Filin, 2003), making them “the most significant errors in a spatial or statistical analysis because they are not easily detected yet introduce significant bias” (Brown & Bara, 1994, p.189).

DTMs are now commonly used in Geographic Information Systems (GIS) to derive terrain attributes (*e.g.* slope, orientation, rugosity) that can be used as surrogates for other phenomena in fields such as ecology (Bolstad *et al.*, 1998), biogeography (Franklin, 2013), digital soil mapping (Behrens *et al.*, 2010), hydrology (Nikolakopoulos *et al.*, 2015), and biology (Kozak *et al.*, 2008). Artefacts in DTMs were shown to sometimes propagate to the derived terrain attributes (*e.g.* Sofia *et al.*, 2013; Chapter 5) and are likely to impact subsequent analyses (Arbia *et al.*, 1998; Heuvelink, 1998). Mapping and quantifying error propagation throughout analysis have received significant attention in the geospatial literature (*e.g.* Fisher & Tate, 2006; Wilson, 2012) but are rarely performed by DTM users from other disciplines (*e.g.* ecology, biogeography) (van Niel & Austin, 2007). In recent years, repeated calls for the appropriate consideration of error propagation in environmental modelling and mapping have been made (*e.g.* Guisan *et al.*, 2006; Rocchini *et al.*, 2011; Lecours *et al.*, 2015), with a strong focus on uncertainty propagation (*e.g.* van Horssen *et al.*, 2002; Jager & King, 2004; Beale & Lennon, 2012; Lechner *et al.*, 2012a).

Quantifying error propagation from DTM is especially relevant for the production of species distribution models (SDM) and habitat maps (van Niel *et al.*, 2004; Peters *et al.*, 2009) that often combine terrain attributes with other environmental data (Franklin, 1995; Guisan & Zimmermann, 2000; Williams *et al.*, 2012; Leempoel *et al.*, 2015). These maps and models are often used to support decision-making in conservation and management (Miller, 2010; Guisan *et al.*, 2013). However, a lack of understanding of errors, their propagation and spatial distribution in maps may result in inaccurate maps and models that could lead to inappropriate decisions (Beale & Lennon, 2012), and negative impacts on biodiversity or stakeholders (Beven, 2000; Regan *et al.*, 2005; Etnoyer & Morgan, 2007). The awareness of data quality in SDM and habitat mapping is increasing (Barry & Elith, 2006; Lek, 2007), with studies that for instance investigated the influence of species distribution data accuracy (Moudrý & Šímová, 2012), the choice of algorithm or model (Pearson *et al.*, 2006), the choice of predictor variables (Synes & Osborne, 2011), and issues associated with collinearity (Watling *et al.*, 2015). With some exceptions (*e.g.* van Niel *et al.*, 2004; Livne & Svoray 2011), issues related to spatial data error are often overlooked. To our knowledge, the influence of data acquisition artefact errors in a DTM has never been assessed on maps resulting from SDM or habitat mapping exercises.

The objective of this study was to describe the impact of some common remotely sensed data acquisition artefacts on habitat maps and SDMs. Our specific objectives were to 1) quantify the impact of artefacts on habitat maps accuracy and SDMs performance, to 2) assess if impacts are dependent on spatial scale, and to (3) assess if impacts can be attenuated when combining the affected data with other environmental data of better quality. Our hypotheses were that artefacts do negatively impact habitat maps and SDMs, that the impacts are greater at finer scales, and that the addition of better quality data reduces the impacts of artefacts on maps and models.

6.2 Material and Methods

6.2.1 Case Study and Data

This paper explored the impact of DTM artefacts on habitat maps using a case study from the marine environment. The marine realm provides an ideal case as it has been suggested that underwater DTMs, or Digital Bathymetric Models (DBM), may be more prone to errors and artefacts than terrestrial DTMs (Hughes-Clarke *et al.*, 1996; Passalacqua *et al.*, 2015; Lecours *et al.*, 2016). Except for shallow waters that can be mapped using LiDAR (Brock & Purkis, 2009) or optical data (*e.g.* Eugenio *et al.*, 2015), acoustic remote sensing remains the main technology to collect reliable continuous terrain data in deeper waters (Brown *et al.*, 2011). DBMs are often the only available datasets used to characterize deep-water environments due to difficulties to observe and sample other environmental characteristics (Solan *et al.*, 2003; Robinson *et al.*, 2011). If multibeam echosounders are currently the best technology allowing the collection of large DBMs (Kenny *et al.*, 2003), most bathymetric surfaces generated from these systems still contain some artefacts (Hughes-Clarke, 2003a; Roman & Singh, 2006). Since these artefacts are often within hydrographic error standards (Hughes-Clarke, 2003a) and appear even when appropriate calibration and corrections are made (Erikstad *et al.*, 2013), they are often considered inherent to the data and tend to be overlooked by DBM end-users. No mention of DBM and multibeam echosounders were found in recent texts reviewing spatial data quality in DTM (*e.g.* Hunsaker *et al.*, 2001; Wu *et al.*, 2006; Shi, 2010), making them considerably less explored and understood than digital elevation models (DEM) generated from LiDAR, Radar or optical remote sensing.

This paper used bathymetric data for German Bank, off Nova Scotia (Canada), in the eastern Gulf of Maine (Figure 6.1). The surveyed area covers 3,650 km^2 of the Scotian Shelf and has been extensively studied in previous works (*e.g.* DFO, 2006; Todd *et al.*, 2012; Brown *et al.*, 2012). Bathymetric data were collected by the Canadian Hydrographic Service (CHS) and were corrected in post-processing for tide, motion, and sound velocity. The corrected soundings were used to generate reference DBMs at five different spatial resolutions: 10 m, 25 m, 50 m, 75 m and 100 m, using the interpolation algorithm “swath angle” implemented in the bathymetric processing software CARIS HIPS and SIPS v.9.0. These five reference DBMs were assumed to be free of artefacts, and following methods described in Chapter 5, ten different amplitudes of heave, pitch, roll and time artefacts were artificially introduced in them (Table 6.1) using CARIS HIPS and SIPS. As described in Chapter 5, these common artefacts were selected based on their different theoretical impact on bathymetric data; pitch impacts bathymetric data in a horizontal plane, heave impacts them in a vertical plane, roll affects soundings that are further away from the nadir in a vertical plane – consequently affecting areas that overlap between different survey lines – and time causes a relative shift of adjacent lines in the horizontal plane (see Hughes-Clarke, 1996, 1997, 2002, 2003a, 2003b and Lurton, 2010). Similar artefacts can also be found in other types of remote sensing like LiDAR (Brown & Bara, 1994; Filin, 2003; Lichti & Skaloud, 2010). The artefacts were systematically introduced to provide controlled conditions that enable comparisons of results, as often performed in evaluations of the impact of error on analyses (Reuter *et al.*, 2009).

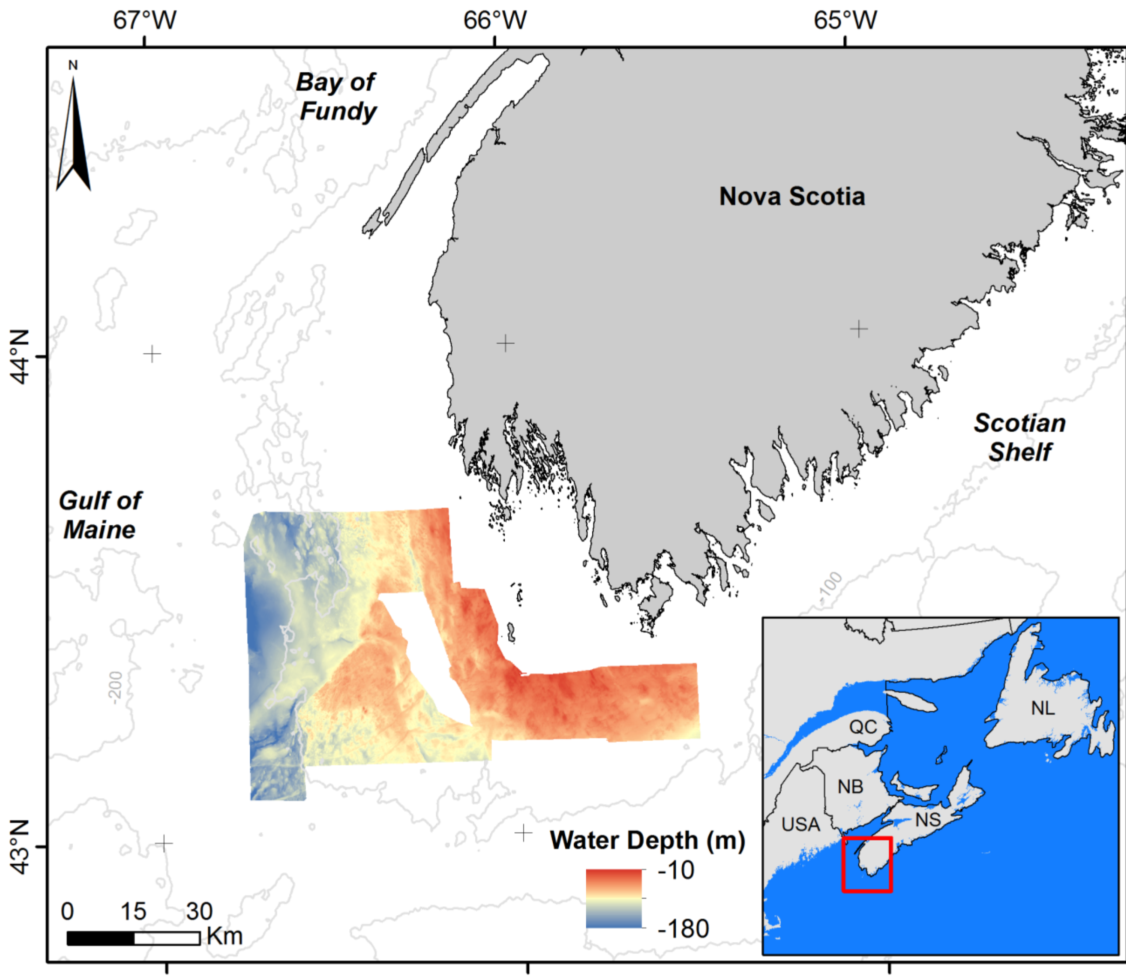

Figure 6.1: Digital Bathymetric Model of the German Bank study area.

Table 6.1: Levels of artefacts introduced in the five reference DBMs. Standard deviations (σ) were derived from the recorded motion at time of survey. A positive pitch indicates that the bow is up and a positive roll means that the port side is up.

	Level of Induced Artefact									
	-5σ	-4σ	-3σ	-2σ	-σ	σ	2σ	3σ	4σ	5σ
Heave (m)	-1.65	-1.32	-0.99	-0.66	-0.33	0.33	0.66	0.99	1.32	1.65
Pitch (°)	-8.25	-6.60	-4.95	-3.30	-1.65	1.65	3.30	4.95	6.60	8.25
Roll (°)	-5.05	-4.04	-3.03	-2.02	-1.01	1.01	2.02	3.03	4.04	5.05
Time (s)	-1.25	-1.00	-0.75	-0.50	-0.25	0.25	0.50	0.75	1.00	1.25

Six terrain attributes that together summarize topographic variability of an area (*cf.* Chapter 3) were derived from each of the five reference and 200 altered DBMs using the TASSE toolbox for ArcGIS (Lecours, 2015). The attributes are slope, two measures of aspect (easterness and northerness), topographic mean, rugosity and topographic position.

Backscatter data (*i.e.* acoustic reflectance) were simultaneously recorded with the bathymetric data. The backscatter data were processed and transformed by Brown *et al.* (2012) into three derivative layers that inform on seafloor properties (*e.g.* surficial geology, porosity): Q1, Q2 and Q3. Since the raw backscatter data could not be accessed, the original 50 m resolution backscatter derivatives were resampled using ArcGIS at the four other studied scales to match the resolutions of the bathymetric and terrain attributes surfaces. Resampling is a common practice in ecological studies that combine remote sensing data with other environmental data of different spatial resolutions (Costa *et al.*, 2009; Chust *et al.*, 2010; Davies & Guinotte, 2011; Erikstad *et al.*, 2013). Finally, two sets of ground-truth data from Brown *et al.* (2012) were used: (1) 3,190 geo-referenced photographs of the seafloor classified into five habitat types (reef, glacial till, silt and mud, silt with sediment bed forms, sand with sediment bed forms and highly abundant sand dollars (*Echinarachnius parma*)), and (2) 4,816 geo-referenced sea scallop observations (*Placopecten magellanicus*) (*cf.* Figure 4.1).

6.2.2 Habitat Maps and SDMs

Using the 205 sets of bathymetric and terrain attribute surfaces (*i.e.* five reference and 50 altered sets per artefact type), habitat maps and SDMs were produced for three scenarios. First, maps and models were generated using only the bathymetry and the six terrain attribute surfaces, thus accounting only for terrain morphology (*i.e.* hereafter referred to as "7 layers" scenario). Then, maps and models were produced using all the available data, *i.e.* bathymetry, six terrain attributes and three backscatter derivatives (*i.e.* "10 layers" scenario). Finally, maps and models were built using only non-correlated variables (*i.e.* "8 layers" scenario). On German Bank, the steepest areas are also the ones with the highest rugosity, resulting in a high correlation between the slope and rugosity data layers. Also, bathymetry is highly correlated with topographic mean as they are closely related (see Chapter 3). Rugosity and topographic mean were therefore not used for the last sets of maps and models. The data for the 8 layers scenario were thus bathymetry, slope, easterness, northerness, topographic position, Q1, Q2, and Q3. Overall, 615 habitat maps and 615 SDMs were produced and analyzed.

The method used to generate habitat maps is based on the concept of benthoscape (Zajac, 2008), which is a representation of the biophysical characteristics of an area. Benthoscape maps are generated by adopting a landscape style approach similar to when maps of landscape features are generated from terrestrial datasets. Such approach was used by Brown *et al.* (2012) to map features on the seafloor that could be resolved within the acoustic remotely sensed data, thus not attempting to delineate seafloor attributes beyond what the remote sensing techniques were capable of resolving. This top-down, unsupervised approach to habitat mapping segments the different data layers into a statistically optimum number of classes. These classes are then spatially compared to the geo-referenced photographs and recombined based on best match with the different habitat types (see Brown *et al.*, 2012; Chapter 4). The Modified k-Means unsupervised classification tool of Whitebox GAT v.3.2 was used to produce these maps, and confusion matrices were built to calculate the overall accuracy and kappa coefficient of agreement of each map (*e.g.* Boyce *et al.*, 2002).

SDMs were generated using a bottom-up, supervised approach to habitat mapping in which the sea scallop observations were used to segment the environmental data based on maximum entropy (MaxEnt), a common and effective method used to build SDMs (Phillips *et al.*, 2006; Monk *et al.*, 2010). The MaxEnt software v.3.3.3k was used to compute the models with the same settings as used in Brown *et al.* (2012) and in Chapter 4: most default settings were used except for that the number of background points was increased to 50,000. A set of 3,813 scallop observations were used to train the model while 1,003 were kept for validation. The MaxEnt software was also used to calculate the area under the curve (AUC) derived from threshold independent receiver operating (ROC) curves to quantify the performance of the models and enable comparisons (Phillips *et al.*, 2006). Two types of AUC values were calculated: AUC_{Train} that measures the goodness-of-fit of models to the training data, and AUC_{Test} that evaluates the ability of models to perform well on an independent dataset – in this case the validation samples (Fitzpatrick *et al.*, 2013). Combining these two measures enabled comparisons between models' performance, robustness, and generalizability (*i.e.* transportability, transferability) (Vaughan & Ormerod, 2005; Warren & Seifert, 2011). Generalizability is quantified using AUC_{Diff}, which is the difference between AUC_{Train} and AUC_{Test}. A high value of AUC_{Diff} is an indication that a model over-fitted the training data and does not replicate well to a different dataset. Finally, correlations between model outputs were calculated to evaluate spatial similarity of predictions.

6.3 Results

6.3.1 Habitat Maps

6.3.1.1 Reference Habitat Maps

Kappa coefficients of agreement and overall accuracies of the reference habitat maps show that maps with no correlated input data (8 layers) performed better than the two other types of maps at finer scales (10 and 25 m) (Figure 6.2). At coarser scales (50 m to 100 m), the maps using all available data (10 layers) performed best. At all spatial scales, the maps accounting only for terrain morphology and depth (7 layers) had the lowest coefficients of agreement and overall accuracies. In general, coarser-scale data produced more accurate maps than finer-scale data, except for maps built from uncorrelated data that were more consistent across scales.

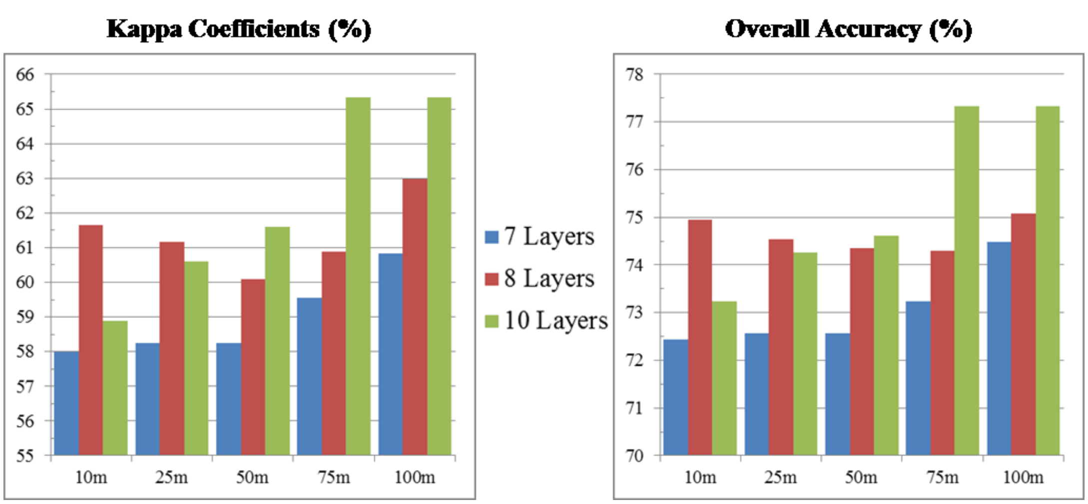

Figure 6.2: Kappa coefficients of agreement and overall accuracies of the 15 reference habitat maps.

6.3.1.2 Habitat Maps from Altered Data

The average kappa coefficients of agreement of all ten maps made from altered data for each scenario and their standard deviation are presented in Table 6.2, while each individual kappa of the 615 habitat maps is presented in Figure 6.3. Since overall accuracy has previously been found to be a poor indicator of classification performance (Felix & Binney, 1989; Fielding & Bell, 1997; Chapter 4), the associated results are not discussed here. Results are however provided in Appendix F for comparison with previous habitat mapping studies from German Bank that reported overall accuracies (*e.g.* Brown *et al.*, 2012). In general, habitat maps produced using 10 layers provided the best classifications, followed by those using 8 layers and those built from only 7 layers. For all artefact types, the average ranges of measured kappa coefficients across the ten levels of introduced artefacts were lower for the scenario with 8 layers. The second lowest average ranges belonged to the scenario with 7 layers when maps were impacted by pitch and heave, and to the scenario with 10 layers for roll and time. The lowest average range was 2.7% (heave, 8 layers) and the highest one was 11.4% (roll, 7 layers). In average (Table 6.2), only three sets of maps showed a scale-dependent pattern for which maps produced from finer-scale data were more impacted by artefacts than maps made from broader-scale data. These sets all belong to the scenario with 10 layers and were maps impacted by pitch, roll and time. A deeper exploration of the results showed that the presence of artefacts in data used to produce habitat maps has a noticeable influence on the spatial distribution of the habitats (*cf.* Figure 6.4 and Figure 6.5). When matching the total area misclassified because of artefacts – *i.e.* when comparing the classifications made from altered data to one made from reference data – to the difference in kappa coefficient between maps made from altered data and reference maps, results show that a little difference in kappa coefficient can translate into large differences in spatial output (*cf.* Figure 6.5).

Table 6.2: Mean and standard deviation of kappa coefficients of agreement of the ten maps made from altered data for each type of artefact, each scenario and each scale.

		7 Layers		8 Layers		10 Layers	
	Resolution	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Heave	10 m	60.6	0.6	61.6	0.6	63.9	0.8
	25 m	58.0	0.9	62.1	1.5	60.2	0.7
	50 m	59.3	0.9	60.3	0.9	61.0	1.0
	75 m	59.8	0.5	60.5	0.9	62.5	1.3
	100 m	62.2	1.1	63.1	0.6	65.7	1.1
Pitch	10 m	57.0	0.9	61.9	0.9	60.5	1.3
	25 m	58.3	0.9	60.3	1.1	60.7	1.2
	50 m	58.1	0.6	60.1	0.8	60.9	0.7
	75 m	60.1	1.3	61.8	0.6	62.0	1.3
	100 m	61.6	1.6	62.6	1.2	64.3	1.1
Roll	10 m	49.7	3.5	55.4	3.6	52.9	3.4
	25 m	49.0	4.2	54.8	3.7	53.4	3.7
	50 m	50.8	3.4	55.5	2.6	55.4	3.1
	75 m	53.1	3.0	56.1	2.8	56.6	3.9
	100 m	55.2	3.0	57.0	3.4	58.8	3.7
Time	10 m	57.9	1.2	62.3	1.2	60.1	1.3
	25 m	57.6	1.0	60.6	1.2	60.4	0.7
	50 m	58.2	1.5	60.4	1.3	61.3	1.0
	75 m	60.0	1.0	61.1	0.8	62.4	1.0
	100 m	61.8	1.3	62.8	0.9	64.4	1.0

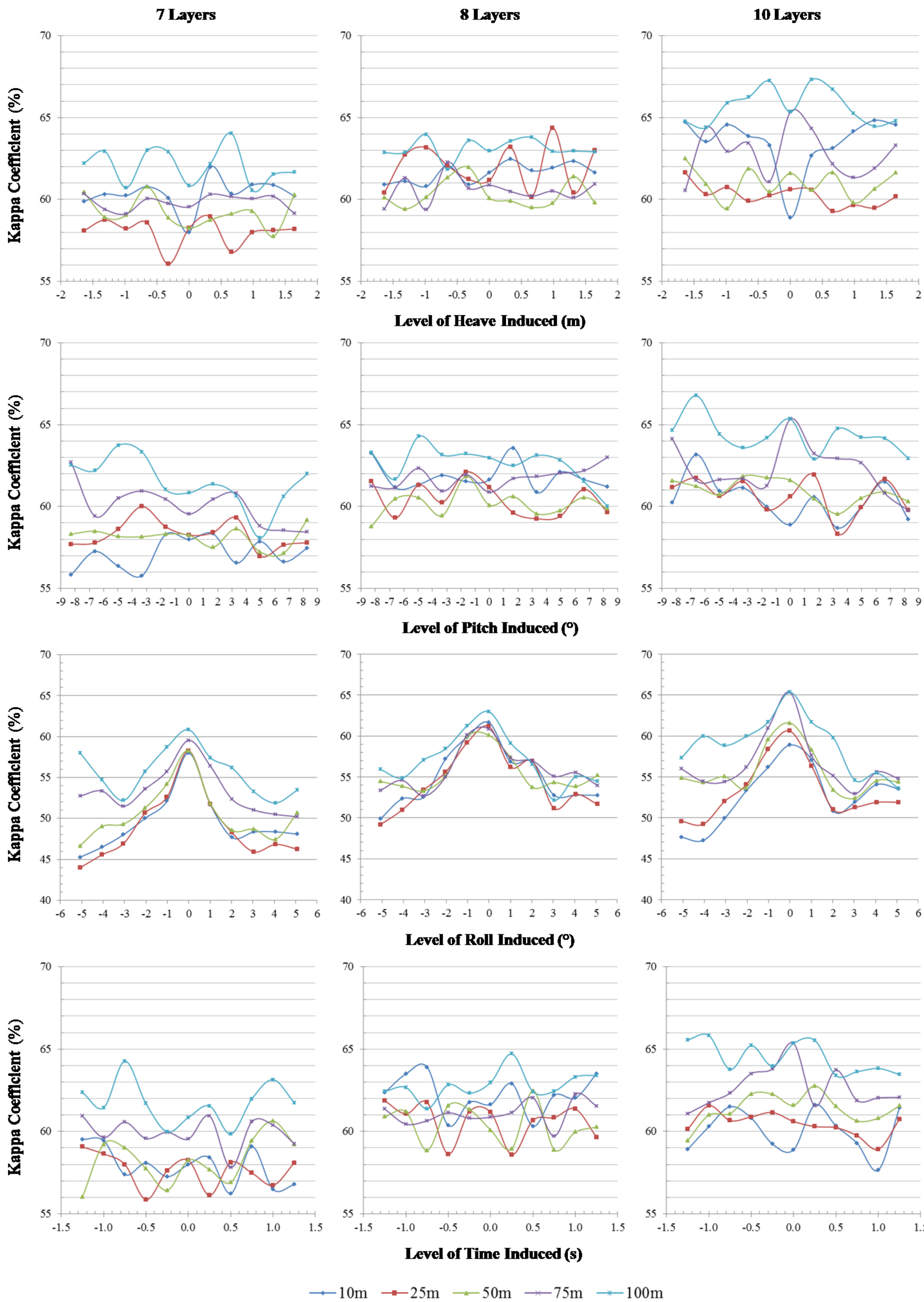

Figure 6.3: Kappa coefficients of agreement of the 615 habitat maps.

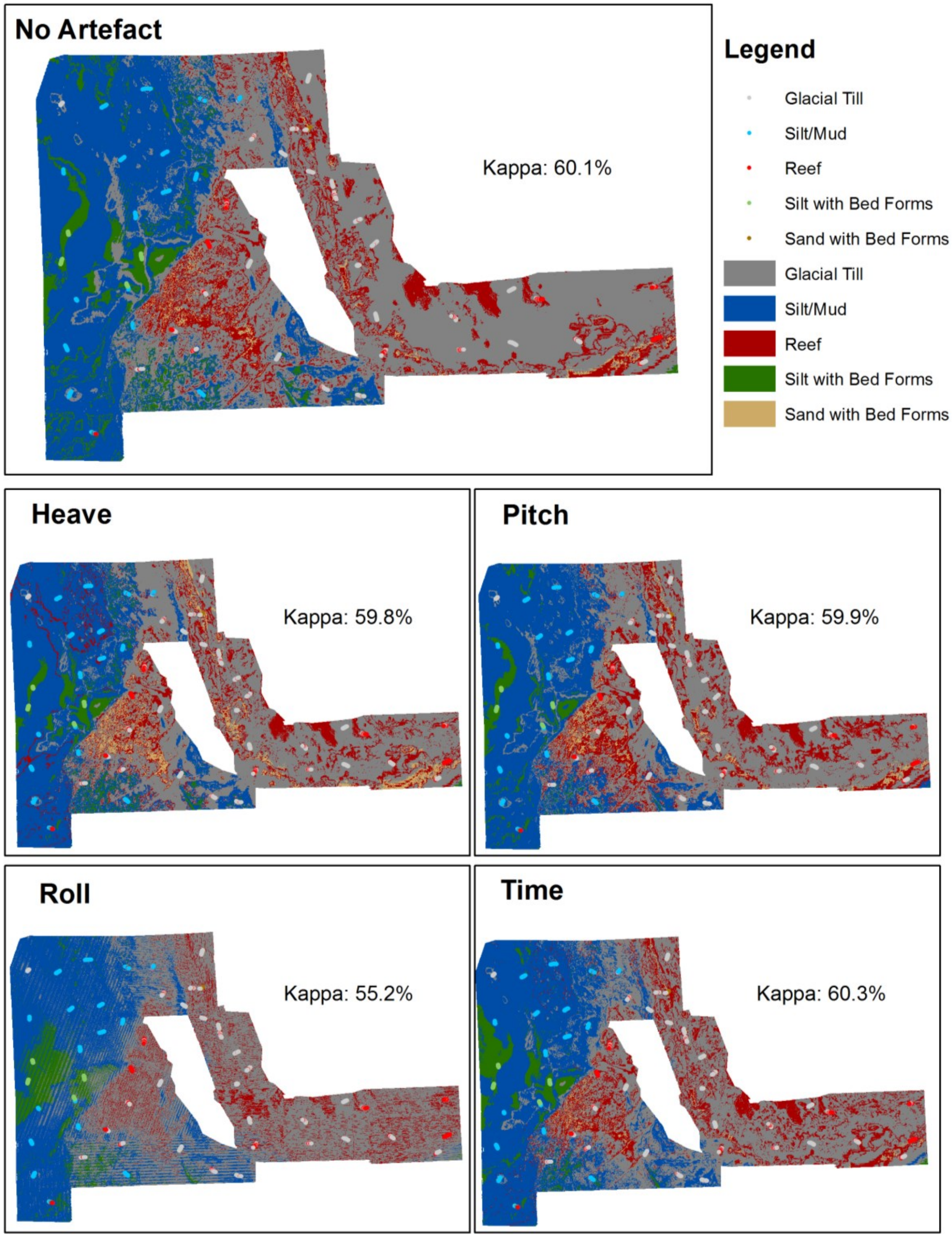

Figure 6.4: Examples of habitat maps produced with 8 layers at 50 m resolution, overlaid by the ground-truth data. The level of error represented in the lower maps is the highest one (5σ, Table 6.1). The colour of the ground-truth data should spatially match the classification's colour when appropriately classified.

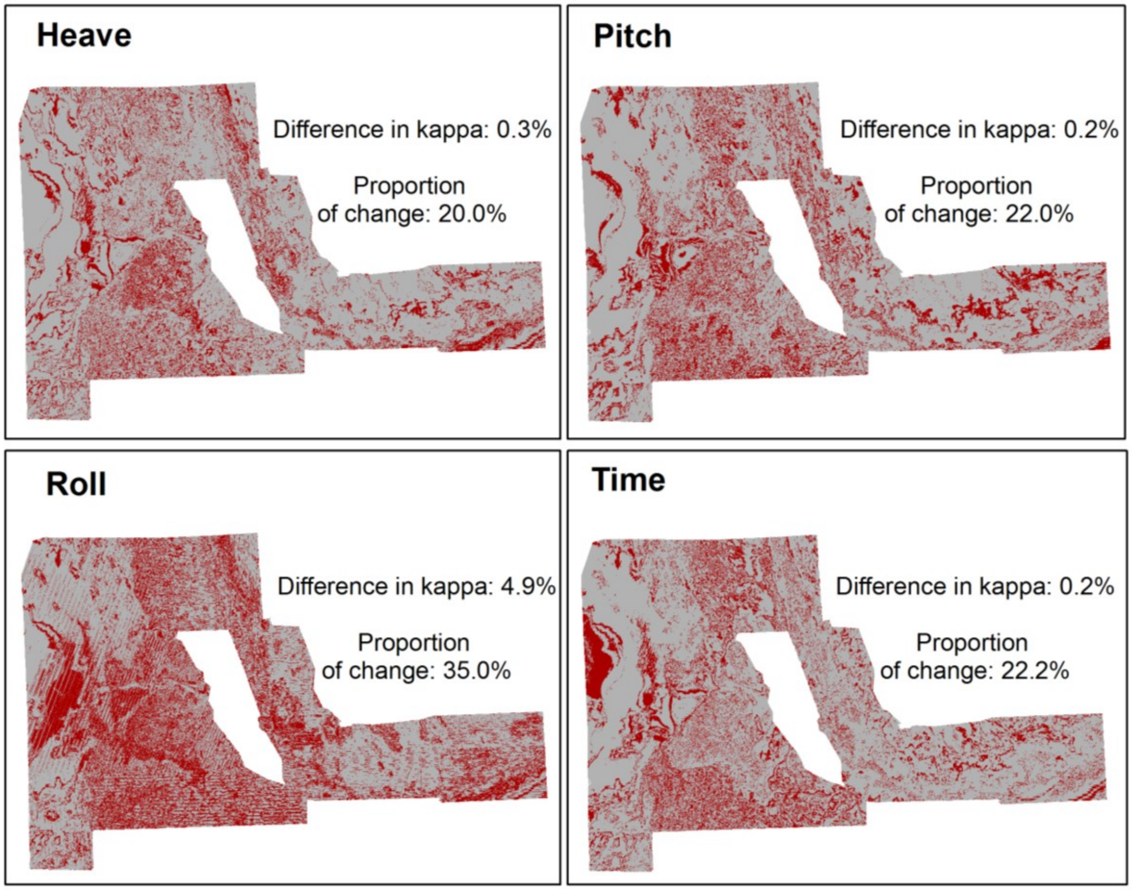

Figure 6.5: Spatial distribution of the change in habitat map classification between the maps presented in the bottom of Figure 6.4 and the reference map shown on top of Figure 6.4. Red pixels indicate change while grey pixels mean that these pixels were classified as the same habitat type in the two compared classifications.

Except for maps affected by roll artefacts, the reference maps did not always produce the best outcome in terms of accuracy: while all habitat maps impacted by any level of roll artefact performed worse than the reference maps, Figure 6.6 shows that an important number of maps made from altered data had a higher kappa coefficient than their corresponding reference maps. This was observed regardless of scale and scenario. Overall, 47% of the habitat maps altered by pitch had a higher kappa coefficient than their corresponding reference map, and that percentage was higher for time (50%) and heave (55%). No particular scale-dependent patterns were observed, except for habitat maps made from seven layers and impacted by pitch, for which a greater amount of maps performed better than the reference maps at broader scales. The absence of maps that performed better than the reference map at 75 m resolution for the 10 layers scenario is noteworthy.

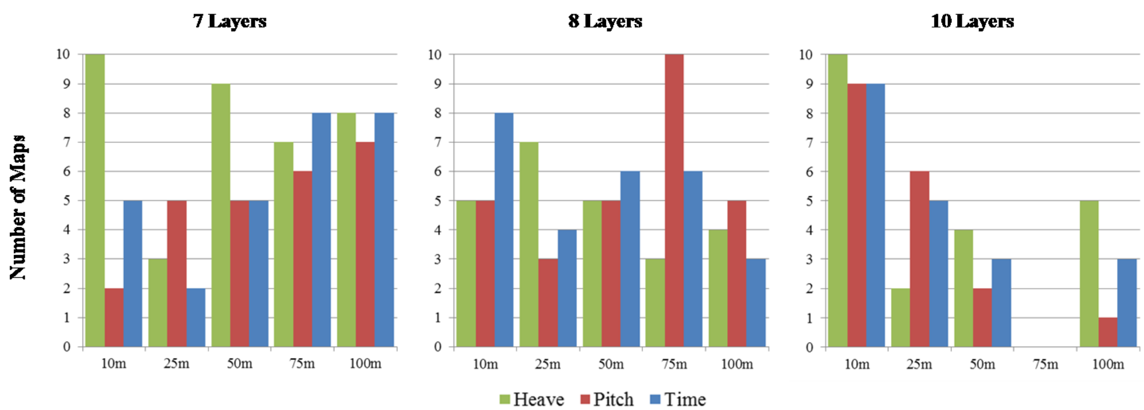

Figure 6.6: Number of habitat maps made from altered bathymetric and terrain attribute data that had a higher kappa coefficient of agreement than the reference habitat maps. No habitat maps impacted by roll performed better than the reference maps.

6.3.2 Species Distribution Models

6.3.2.1 Reference SDMs

Figure 6.7 shows the performance, robustness and generalizability of the MaxEnt models generated from the reference data. In general, models with 7 layers performed poorly compared to others. Models with10 layers performed better and were more robust than models using 8 layers at 10 m and 100 m scales, while models from 8 layers were the best at 25 m, 50 m and 75 m resolution. In these two scenarios (8 and 10 layers), models at 50 m resolution showed the best trade-off between performance and robustness, followed by those at 25 m, 10 m, 100 m and 75 m.

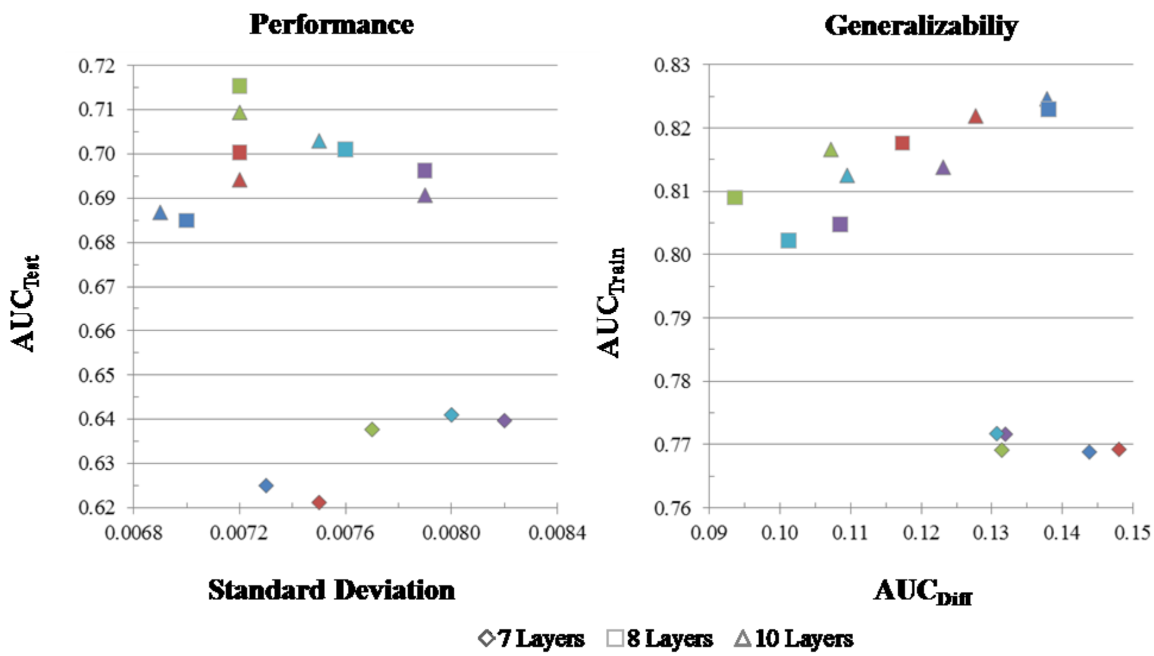

Figure 6.7: Performance and robustness (left) and generalizability (right) of the 15 reference MaxEnt models. Models are colour-coded following Figure 6.3: dark blue (10 m resolution), red (25 m), green (50 m), purple (75 m) and light blue (100 m). High AUC_{Test} values indicate that models performed well on validation data and low standard deviations indicate robust models. High AUC_{Train} values indicate that models performed well on training data and low AUC_{Diff} indicate that they replicated well on validation data. In both graph, the best models are thus located in the top left quadrant.

Models accounting only for topography (7 layers) were not highly generalizable as they showed low AUC_{Train} combined with high AUC_{Diff} relative to other models. Models with 10 layers usually had a higher AUC_{Train} than models with 8 layers, but the former also had a higher AUC_{Diff} than the latter. Better models were thus less replicable. The model presenting the highest index of generalizability (the ratio of AUC_{Train} on AUC_{Diff})

was the one built from uncorrelated data at 50 m (*cf.* green square in Figure 6.7). It was followed by the model at 100 m (8 layers), and those at 50 and 100 m (10 layers), 75 and 25 m (8 layers and then 10 layers), and models at 10 m resolution (10 layers followed by 8 layers). All models generated from 7 layers (*cf.* diamonds in Figure 6.7) were among the least generalizable. Regarding scale, a similar relationship to that of robustness was observed between AUC_{Train} and AUC_{Diff}: higher performance on training data also involved lower replicability on test data.

In terms of spatial outputs, Figure 6.8A shows the quantification of how predictions generated by the MaxEnt models at scales ranging from 25 m to 100 m vary in space compared to the same model generated from 10 m resolution data. In all scenarios, the correlation between predictions decreased as spatial scale coarsened. Models with 8 and 10 layers followed a very similar trend (*cf.* red and green lines), and Figure 6.8B shows that their outputs are highly correlated at all scales (*cf.* orange line). Models from different scenarios were generally more similar at finer scales, except for the reference models with 7 and 10 layers that became more similar at coarser scales (*cf.* light blue line in Figure 6.8B).

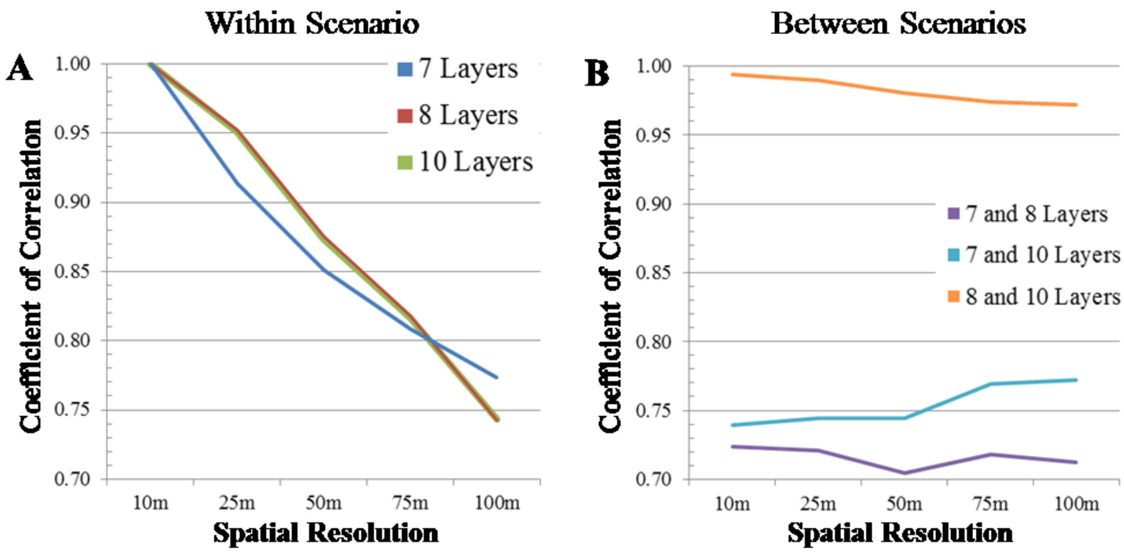

Figure 6.8: A) Spatial correlation between the reference MaxEnt models for the three scenarios and their corresponding reference models at other scales (*e.g.* the models computed with 7 layers at 50 m and 10 m have a r value of 0.85). B) Spatial correlation between the reference models of different scenarios at each scale (*e.g.* models computed at 50 m resolution with 7 and 8 layers have a r value slightly above 0.70).

6.3.2.2 SDMs from Altered Data

Figure 6.9 shows changes in MaxEnt SDM performance and robustness as artefacts are introduced in the input DBMs, and Figure 6.10 shows examples of discrepancies in probability distribution that were introduced by artefacts. For all types of artefacts, no pattern could be observed regarding whether some scales were more impacted than others, or whether a greater level of artefact resulted in higher or lower performance or robustness.

In general, results show that introducing heave artefacts decreased models' performance and tend to also decrease models' robustness (*i.e.* higher standard deviation).

For instance, the 8 layers models at 50 and 75 m resolution were as robust as their respective reference model, although the reference models performed better (*i.e.* higher AUC_{Test}). One major exception was observed to these patterns: at 75 m resolution, 7 and 10 layers models performed better than the reference models. The two reference models in these cases had a higher standard deviation and a lower AUC_{Test}. Models with pitch artefacts sometimes performed better than the reference models: 39% of the models built from data showing this type of artefacts had a higher AUC_{Test} than their respective reference model, and 36% of them were more robust than the reference models. The reference models were the best in terms of AUC_{Test} only at 10 m with 7 layers, and 50 m with 8 and 10 layers. Only the reference model at 100 m built from uncorrelated data (8 layers) was the most robust compared to its respective models made from altered data. Roll artefacts also increased model performance, as 87% of models built from altered data had higher AUC_{Test} measures than the reference models. In terms of robustness, 26% of models from altered data had a lower standard deviation than the reference models. In terms of time artefacts, 29% of the models that were built from altered data performed better than the reference models and 11% were more robust.

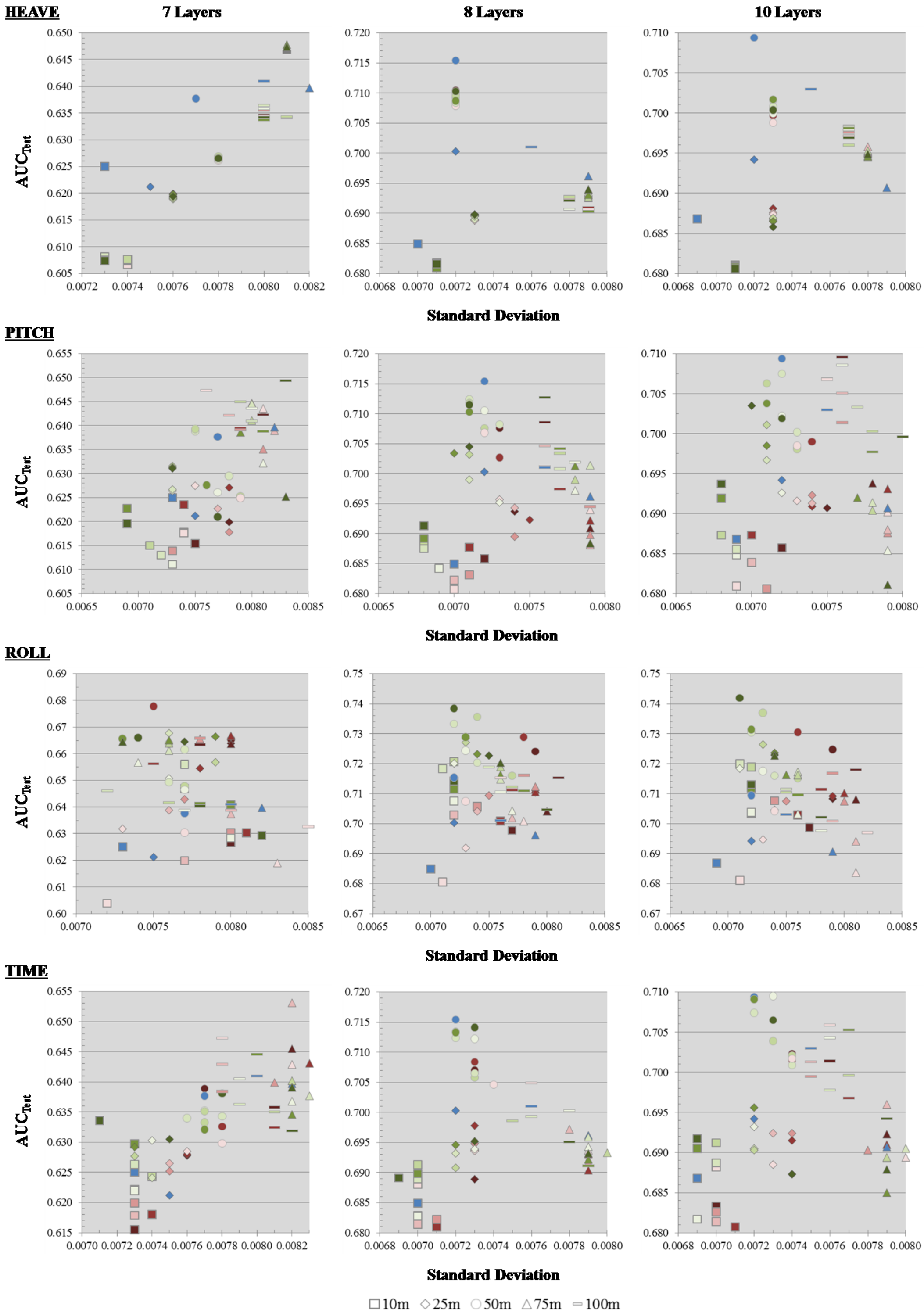

Figure 6.9: Change in performance and robustness as the level of artefacts in the data changes. The blue symbols are the reference models. The darker a symbol, the greater the level of artefact is.

Green indicates positive artefact alterations (*e.g.* positive pitch) while red indicate negative alterations (Table 6.1). High AUC_{Test} indicate that models performed well and low standard deviations indicate robust models. Good and robust models are located in the top left quadrant.

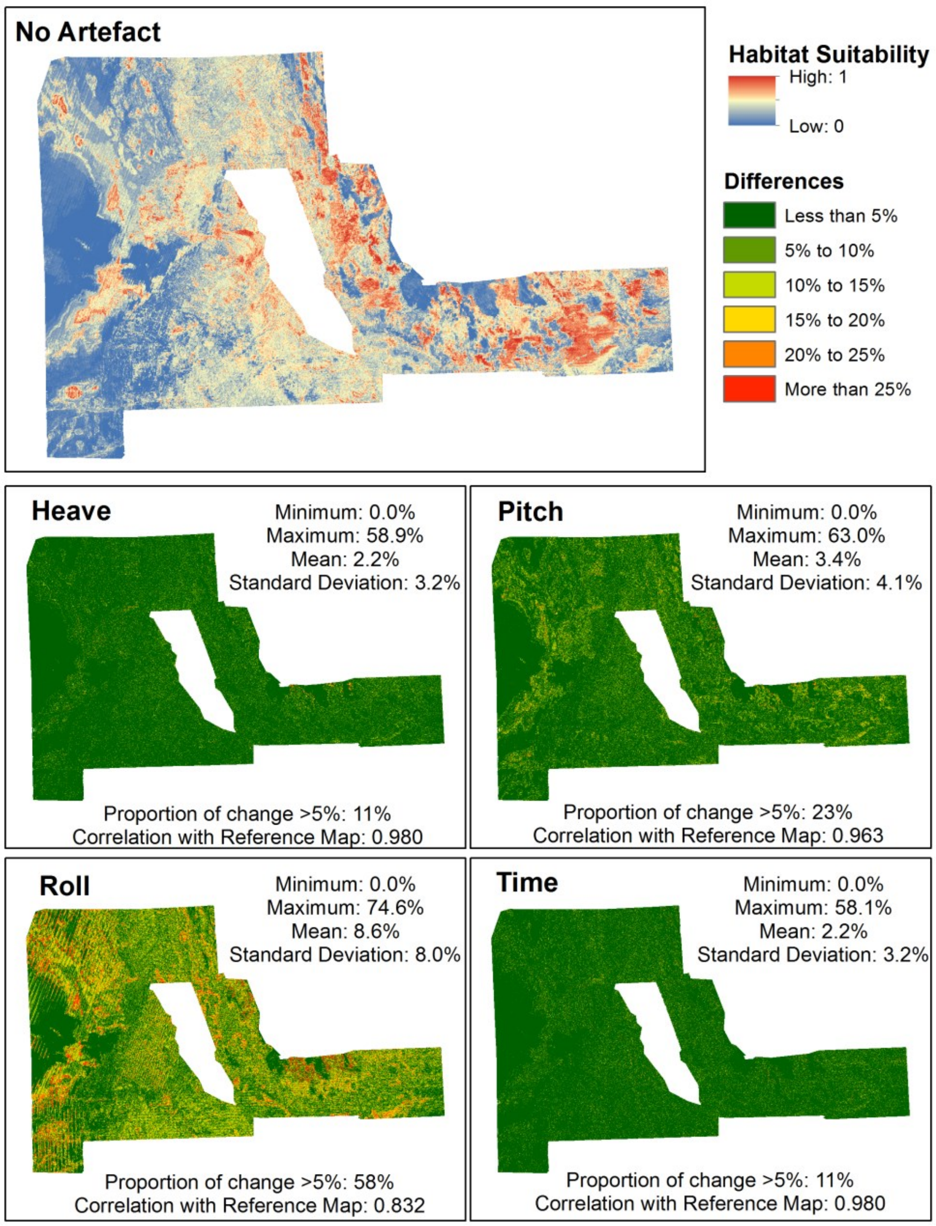

Figure 6.10: Differences in probability distribution between models affected by artefacts and a reference model (top). The scenario represented is the one with 8 layers at 50 m resolution. The level of error represented is the highest one (5σ, Table 6.1).

Heave artefacts decreased the generalizability of all models at 10, 50 and 100 m resolution, regardless of the number of input layers. At 25 m resolution however, heave decreased the generalizability of models built from 8 and 10 layers but increased the generalizability of models with 7 layers. At 75 m resolution, generalizability was also increased by the presence of heave artefacts in models built with 7 and 10 layers, but decreased when using uncorrelated dataset. Pitch artefacts however did not demonstrate such a specific pattern. At 10, 75 and 100 m, some models made from altered data were more generalizable than the reference models, while others were not. This was also true of models with 7 layers at 50 m, and models with 8 or 10 layers at 25 m resolution. Models from altered data accounting only for topography at 25 m were however all more generalizable than the reference model, and the two reference models at 50 m built from uncorrelated data and all available data were the most generalizable. In terms of roll artefacts, most models built from altered data became more generalizable than the reference models. Finally, similarly to what was seen for pitch, time artefacts did not affect generalizability in any particular patterns at 10, 75 and 100 m resolution, with some models being more and some being less generalizable than the reference models. At 25 and 50 m however, most models affected by time artefacts were less generalizable than the reference models when 8 or 10 layers were used. At 25 m, models made from altered data and accounting only for topography were all more generalizable than the reference models, while no particular pattern was found at 50 m resolution.

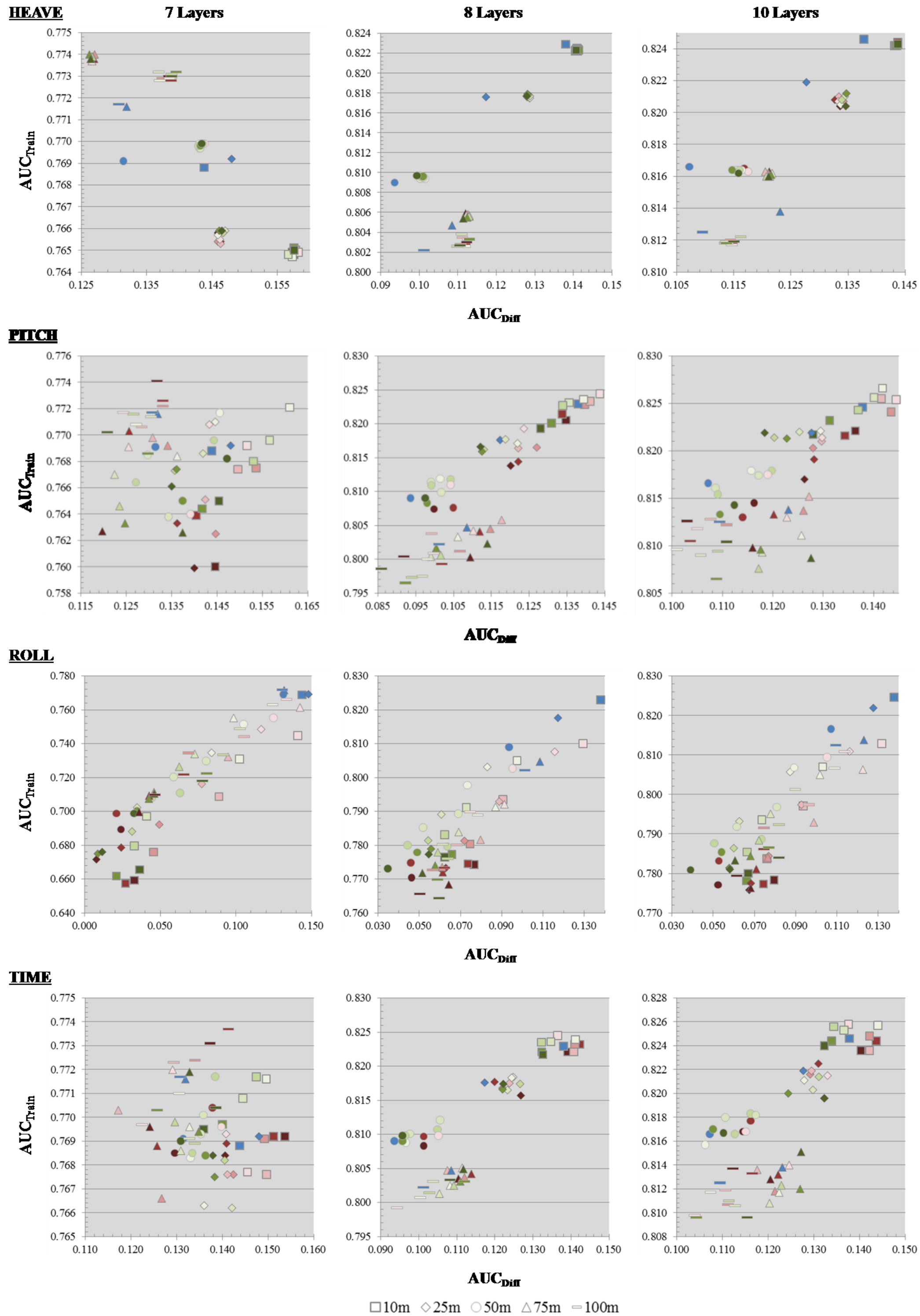

Figure 6.11: Generalizability of all SDMs as the level of artefacts in the data changes. High AUC_{Train} indicate that models performed well on training data and low AUC_{Diff} indicate that they replicated

well on validation data. More generalizable models are thus located in the top left quadrant. See Figure 6.9 for legend.

In terms of spatial outputs, models impacted by heave artefacts were different from the reference models. However, that alteration was not very variable depending on the level of artefact introduced. Models built from uncorrelated data (8 layers) were the least impacted at scales ranging from 25 m to 75 m. Models accounting only for topography and depth (7 layers) were consistently the most different to the reference models. Models affected by pitch artefacts also produced different outputs compared to the reference models. Models with 7 layers were the most variable and also the least similar to reference models. For the two other scenarios, the ranges of correlation coefficients were similar but the one with uncorrelated data seemed to be slightly less impacted by artefacts. Roll produced similar results to pitch, although the recorded correlation values were much lower. Finally, time artefacts also produced similar results to pitch. Overall, roll seemed to have the most impact on the spatial distribution of predictions of sea scallop probabilities of occurrence. Models built from only 7 layers were the most impacted, followed by those with 8 layers and those made from 10 layers. No clear pattern was observed in terms of scale, although roll artefacts seemed to produce models that were more similar to the reference ones at coarser scales, and the extreme scales (10 and 100 m) seemed to be a bit more impacted by heave, time and pitch than the intermediate scales (25 to 75 m).

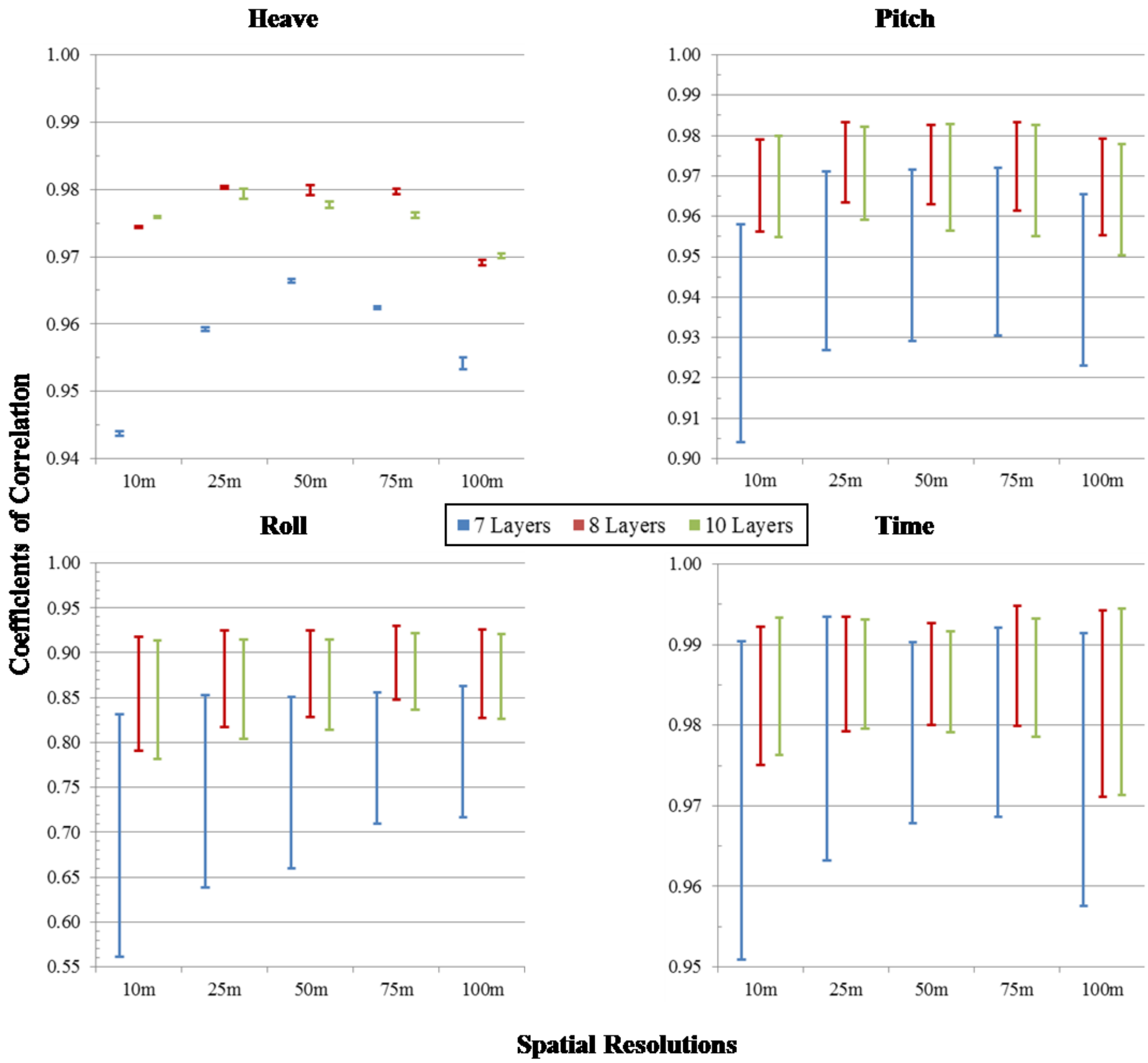

Figure 6.12: Spatial variation in predictions of sea scallops distribution as quantified by the range in correlation coefficients between models built from altered data and the reference models.

6.4 Discussion

6.4.1 Reference Habitat Maps and SDMs

In general, results expanded on those from Chapter 4 that focused on a single scale of 50 m resolution. Results clearly indicate that adding backscatter data to the terrain variables (*i.e.* 8 and 10 layers scenario) improved classification accuracy of habitat maps and SDMs performance, robustness, and generalizability. This result is consistent with the results of Chapter 4, and confirms the importance of carefully selecting the variables used to capture the phenomenon of interest. Since many of the benthoscape classes were based on sediment characteristics (*e.g.* silt and mud), it is expected that backscatter data, which capture this information, improved the classifications. Results also show that the use of either correlated data or a greater amount of data layers (*cf.* 10 layers scenario), or both, can increase accuracy measurements of habitat maps, particularly at coarser spatial scales, and SDMs performance and robustness (*cf.* models at 10 and 100 m). Correlated data are known to impact SDMs and statistical models (Hijmans, 2012; Dormann *et al.*, 2013), and the use of too many variables has been shown to result in model over-fitting (Peterson & Nakazawa, 2008). The latter is confirmed by the lower generalizability of reference models made from 10 layers compared to those built from 8 layers (*cf.* Figure 6.7). Results gained from maps and models built from uncorrelated data may thus be more reliable than those from maps and models built from all available data; even if maps are less accurate or models have a lower performance, they could be more replicable and more reliable.

In terms of spatial scale, results indicate that physical habitat types on German Bank are best captured at coarser spatial scales. These results are logical considering that benthoscape maps are used to delineate broader-scale biophysical features (Zajac, 2008; Brown *et al.*, 2012). This also confirms claims that finer-scale data may not always be the most appropriate to capture the relevant environmental characteristics for a particular habitat or purpose (*e.g.* Cavazzi *et al.*, 2013; Lecours *et al.*, 2015). However, these conclusions need to be approached with care as they could partly be caused by coarse-graining, a phenomenon well documented in the geospatial literature and in terrestrial (*e.g.* Chow & Hodgson, 2009) and marine studies (*e.g.* Wilson *et al.*, 2007; Rengstorf *et al.*, 2012). When the action of coarsening data resolution reduces the range of occurring values and transforms their statistical distribution by centering them on the mean, it removes fine-scale heterogeneity in habitat spatial distribution and possibly artificially increases measures of classification accuracy. In Chapter 5, we demonstrated that the terrain attributes used in the current study were showing the effects of coarse-graining. It is thus likely that these effects translated into coarse-grained habitat maps. We note however that the habitat maps generated from uncorrelated data were less sensitive to changes in scale, which could be an indication that what was observed was not a result of coarse-graining, but simply the appropriate matching of the observational scale with the environmental scale at which the habitats were occurring (Lecours *et al.*, 2015).

Results from the MaxEnt analysis showed that models using 50 m resolution usually performed better, were more robust and more generalizable, an indication that the environmental preferences of sea scallops in terms of seafloor characteristics may be best captured at this scale, confirming once again that finer-scale data may not always be the most appropriate. Another factor to consider regarding the higher performance of the 50 m resolution data is the relative importance of backscatter data in this context. Since these data were initially provided at 50 m resolution, their performance at that particular scale may have been enhanced. Results might have been different if the raw backscatter data could have been accessed and used in this study; if the raw backscatter data had been processed at 10 m resolution, it would likely have added fine-scale heterogeneity in sediment characteristics. It is however impossible to know for sure if a higher level of spatial heterogeneity would improve or decrease finer-scale models' performance. If access to the raw backscatter data had been possible, it would have been better to generate the backscatter derivatives at the five different scales directly from the raw data, as done with the bathymetric data.

Finally, the examination of spatial similarity across scales and scenarios highlight the importance of decisions regarding which variables should be included in analyses and at which scale, in line with previous calls made in the literature (*e.g.* Araujo & Guisan, 2006; Synes & Osborne, 2011; Williams *et al.*, 2012; Mateo Sánchez *et al.*, 2014; Lecours *et al.*, 2015). While the measured correlations between model outputs may seem high (>0.70), Figure 6.10 show that two model outputs with a correlation coefficient of 0.963 between prediction probabilities can be different (*i.e.* differences greater than 5%) for 23% of the area. A coefficient of 0.832 translated into differences for 58% of the study area. These results are in line with those from Chapter 4. The differences in correlation coefficients among scales and the three scenarios that were calculated in the current study (*cf.* Figure 6.8) are thus indications of the likelihood of finding very different model outputs when different variables and scales are used. The assessment of spatial differences in probability predictions is thus a more adequate measure than the correlation coefficient to evaluate differences in model outputs.

6.4.2 Impacts of Artefacts on Habitat Maps and SDMs

Our first hypothesis was that artefacts in bathymetry that propagate to terrain attributes would impact habitat maps and SDMs in a negative way. Results show that this is not always the case. While we were expecting map accuracy to decrease as a function of level of artefacts, only maps impacted by roll demonstrated such relationship. Results show that the other types of artefacts sometimes artificially increased map accuracy, although not in a predictable way. A higher level of artefact did not necessarily result in a better or worse map or model than a lower level of artefact. About half of the habitat maps produced with data altered by heave, pitch and time artefacts performed better than the reference maps. These results may however have been influenced by the approach used to quantify map accuracy. Since the habitats are represented on maps as clusters of pixels showing similar characteristics, they share some characteristics with areal data. Artefacts might thus influence the boundaries of these “zones” more than the area inside them. Because the ground-truth data are points that are more likely to fall within the middle of a zone than at its boundary, the kappa coefficients of agreement and overall accuracies may not capture the change in boundary. A spatial assessment of the differences between the different habitat maps, as performed in Chapter 4 and in Figure 6.5, could help better capture the influence of artefacts on the delineation of the different habitat zones. Considering the amount of maps produced in this study, this would be computationally intensive but such an approach should be considered in future work. These results however yield an important conclusion regarding the methods commonly used in the literature to quantify classification and habitat map accuracy: measures using punctual data to validate classifications of zones may be biased by not capturing the variability of the classifications along zone boundaries.

The analysis of SDMs yielded similar conclusions to the analysis of habitat maps but from different types of artefacts. Heave artefacts had generally a negative impact on the performance of models, with some exceptions (*e.g.* 75 m resolution models). Many models impacted by pitch and time performed better than the reference models. Models impacted by roll artefacts clearly contradicted our hypothesis: the performance of most of these models was artificially increased by the presence of roll artefacts. This could be explained by the fact that sea scallops distribution is driven by rugosity (Brown *et al.*, 2012; Chapter 4), and artefacts like roll and pitch artificially increase the rugosity of an area. Models produced with data altered by these artefacts would thus artificially increased the distribution probability of sea scallops across the entire area, resulting in a higher prediction success when validated against the test data. The increase in distribution probability was confirmed by visual comparison (*cf.* Figure 6.10) but also by the high differences in spatial correlation recorded for pitch and particularly for roll (*cf.* Figure 6.12).

Our second hypothesis stated that the impacts of artefacts in bathymetry and terrain attributes should be greater at finer scales. This hypothesis was based on the fact that the propagation of artefacts from DTM to terrain attributes was previously found to be scale-dependent (*cf.* Chapter 5). Results from both unsupervised and supervised classifications did not confirm, neither did they refute, this hypothesis as no particular scale-dependent patterns could be identified. The difference between the scale-dependent propagation of

DTM artefacts in terrain attributes and the scale-independent propagation of these artefacts in habitat maps and SDMs may be explained by the integration of a biological/ecological context. The presence of artefacts in finer-scale data may not result in a poor habitat classification if these data and the scales at which they were collected and analyzed do not have an ecological meaning or do not match the ecological scale of the phenomenon being studied, thus being unsuitable regardless of their quality.

Finally, our third hypothesis was that the addition of better quality data would reduce the impacts of artefacts on maps and models. Results suggest that this hypothesis is true for the habitat maps, as maps built with the relatively good quality backscatter data were generally more accurate. It however remains unclear whether this improvement was caused by the quality of the data or their nature (*i.e.* backscatter), which in this particular case was known to be ecologically relevant. The latter option is the most likely, considering results from Chapter 4 that showed that maps produced only with backscatter data and depth performed very well. Further work is thus required to validate or invalidate this hypothesis with more certainty. In addition, results showed that the range in measures of accuracy was more stable when uncorrelated data were used, which could be an indication that a better choice in input variable has the potential to stabilize and attenuate the impacts of artefacts. Results also showed that SDMs with artefacts behave differently depending on how many or which variables were included. It was however challenging to find a predictable pattern in this behaviour.

6.4.3 Consideration of Spatial Errors in Ecology

In the general literature, much focus has been given to uncertainty and random error or noise (Li *et al.*, 2012). In the ecological literature, research has also been oriented towards measurement uncertainty and the work performed on errors has largely focused on the positional accuracy of species observations (*e.g.* Moudrý & Šímová, 2012). The impact of DTM artefacts has been studied before in a geomorphology and geomorphometry context (*e.g.* Bonin & Rousseaux, 2005) but rarely in ecology. Of note is the work by Van Niel *et al.* (2004) that studied the effect of error in DTM on terrain attributes and other geomorphometric variables. Despite being set in a context of predictive vegetation modelling, the authors did not study the impact of these errors on the actual predictive models. This issue was however addressed in a related study (Van Niel & Austin, 2007), where sets of random errors were distributed across DTMs and predictive vegetation models were built using generalized additive and generalized linear models. Despite different approaches and types of error studied, this study and the current one yielded similar conclusions regarding the fact that errors do propagate throughout analyses, and impact distribution models although not in an easily predictable way. In another ecological study that looked at uncertainty and error propagation, Livne & Svoray (2011) identified the need to focus on assessing the behaviour of ecological models to spatial errors at different spatial resolutions. While this was addressed in the current study, results did not indicate any scale-dependent pattern.

In the marine environment, researchers are aware of artefacts as they are often, although not always (*e.g.* Lucieer *et al.*, 2012), acknowledged (*e.g.* Blondel & Gómez

Sichi, 2009). When acknowledged, their implications for the ecological analysis being performed are often not discussed (*e.g.* Kostylev *et al.*, 2001). The presence of artefacts in MBES data sometimes prevents their use or the use of their derived terrain attributes in ecological applications (*e.g.* Clements *et al.*, 2010). When such data are still used, artefacts have been linked to habitat misclassifications (*e.g.* Costa & Battista, 2013; Micallef *et al.*, 2012), to noise in results from unsupervised classifications (*e.g.* Galparsoro *et al.*, 2015), and to difficulties associated with identification of seabed features (*e.g.*, Dolan & Lucieer, 2014), among other consequences. In the context of the MAREANO program, the Norwegian Hydrographic Service indicated that "seabed features shall not be camouflaged by artefacts and artefacts must not appear as seabed features" (NHS, 2013, p. 5), and that artefacts in the processed bathymetry "shall be kept at an insignificant level not disturbing the seabed image" (NHS, 2013, p. 14). However, no procedures are indicated to deal with artefacts when they cannot be removed. This overview of the literature reflects the lack of understanding of how artefacts, particularly in MBES bathymetry, impact ecological analyses and interpretations, and the lack of knowledge on how to respond to the presence of artefacts. The work by Zieger *et al.* (2009) is however noteworthy as they used terrain attributes and seafloor classification to identify artefacts in flat areas before correcting for the misclassifications caused by artefacts. Such methods could become a suitable option to deal with artefacts in habitat mapping and SDMs, although they would need to be tested on areas that are not flat and may still require expert knowledge to distinguish which bathymetric patterns are artefacts and which are actual natural features.

While this study has focused on artefacts in multibeam bathymetric data, backscatter data are also often impacted by artefacts (*e.g.* Collier & Brown, 2005; Che Hasan *et al.*, 2012). Like for bathymetric data, some of these artefacts can be removed in post-processing (*e.g.* De Falco *et al.*, 2010; Lamarche *et al.*, 2011) but a complete removal is not always achieved. Backscatter data with artefacts have been widely used (*e.g.* Rattray *et al.*, 2009; Roberts *et al.*, 2009) as they may still yield useful observations. Other times however, they are judged unusable for the mapping or modelling exercise (*e.g.* Holmes *et al.*, 2008). It has been recognized that there is a broad misunderstanding of backscatter within the end user community (Lurton & Lamarche, 2015). In this study, backscatter data were used to evaluate the impact of adding good quality data to poor quality data within the same analysis. As done with bathymetry in this study, future work should evaluate the impacts of artefacts in backscatter data on habitat maps and SDMs. It is to be expected that like for bathymetry and terrain attributes, artefacts in backscatter data will have a greater impact if sediment properties are ecologically relevant to the species, area or problem studied. For instance, Copeland *et al.* (2013) noted that artefacts in the backscatter data resulted in an apparent striping pattern in their habitat classifications.

6.4.4 Implications for Ecological Applications

The results of this study have critical implications for ecological studies that use DTMs and their derived terrain attributes in their applications, which is a very common practice (Bouchet *et al.*, 2015; Lecours *et al.*, 2016). The use of environmental variables such as terrain attributes has been shown to improve predictions accuracy in SDMs (Dobrowski *et al.*, 2008). However, this study showed that when systematic artefact errors are present in DTMs, there is a trade-off between the improved prediction that would be gained from including the DTM and its derived terrain attributes and the risk to produce inaccurate predictions. Results showed that such predictions are not necessarily revealed as lower or absence of predictions, but can be important inflation in predictions. For instance, artefacts may alter the quantification of species-environment relationships by artificially increasing the importance of rugosity in habitat characterization. When rugosity is known to be a surrogate of a particular species distribution, this leads to an overestimation of the suitable habitat for that species.

Conservation and management, especially in the marine environment, often rely on habitat maps and SDMs to inform decisions (Le Pape *et al.*, 2014). Those maps and SDMs are often produced with arbitrarily chosen observational and analytical scales and a limited number of data based on availability (Levin, 1992; Wheatley & Johnson, 2009; Lechner *et al.*, 2012b). The situation in regards to spatial scale is complex. While the observational and analytical scales should match the ecological scale to capture what is relevant (Lecours *et al.*, 2015), the observational scale is defined by the survey methodology (Lecours *et al.*, 2016). In addition to this limitation, the scales that are found to best capture the ecological patterns and processes may not be deemed appropriate by the end users. For instance, Brown *et al.* (2012) selected a 50 m resolution observational scale because (1) a finer resolution was not considered particularly useful in the management process by the end users (*e.g.* fisheries managers, scientists) and a 50 m resolution coincided well with the needs of the managers, (2) such resolution was in line with the footprint of the sea scallops fishery, and (3) it corresponded well with the scale of the features on German Bank for broad-scale characterization. The relationships among observational and analytical scales, ecological scale, and the end use of the maps and models make decisions regarding scales very difficult when habitat mapping is implemented in the process of conservation and management planning.

In addition, studies that include any assessment of data quality are rare (Van Niel & Austin, 2007). The current study highlighted the sensitivity and volatility of maps and models to the choice of variables, the observational scale, and spatial errors like artefacts. Many calls have been made in the literature for the quantification of uncertainty and error propagation throughout ecological analyses (*e.g.* Guisan *et al.*, 2006; Lecours *et al.*, 2015), and tools have been proposed to deal with uncertainty (*e.g.* the Data Uncertainty Engine by Brown & Heuvelink, 2007) but not with errors. The ecological community that makes use of GIS tools and remote sensing techniques is usually aware of this need but such protocol are not yet implemented in any workflow. As stated by Li *et al.* (2012, p. 2277): "there are user communities who may be aware of spatial data quality issues but may not have at their disposal techniques and tools for data quality assurance." Such tools, associated with proper standards, protocols and metadata, are becoming crucial to enable a proper incorporation of error modelling in the different applications workflow. This will eventually lead to results and interpretation that are grounded on solid foundations, and more informed decisions. While it is impossible to avoid error and uncertainty in ecological analyses, it is also important that practitioners stop avoiding it. An acknowledgement of errors like artefacts and a discussion on their potential impact on analyses will increase the chances to make more informed decisions when these data and analyses are used in contexts like conservation planning.

6.5 Conclusions

DTM and terrain attributes are now commonly used in ecological studies. Despite an awareness of the presence of errors like artefacts in these spatial data, their quality is rarely assessed, acknowledged or discussed. The goal of this study was to develop evidence linking the presence of artefacts in DTM with the accuracy of analyses performed in ecological applications. In a context of marine habitat mapping and species distribution modelling, results demonstrated that artefacts do impact habitat maps and SDMs, although not in a predictable way. Roll artefacts showed the most predictable influence, decreasing the accuracy of habitat maps and artificially increasing the performance and generalizability of SDMs. Other types of artefacts sometimes increased map accuracy and model performance and generalizability or decreased them. Results showed that the importance of the impacts of artefacts on ecological applications strongly depend on whether or not the methods are grounded in ecological relevance, particularly in terms of the choice of variables and the spatial scale of the data. While the influence of errors on an analysis depends on the type and requirements of the analysis (Friedl *et al.*, 2001), results gained in this study are transposable to other applications that use remotely sensed data like LiDAR-derived DTMs and encounter similar artefacts. This study also highlighted requirements for error quantification tools to become widely available to scientists and practitioners with a wide range of background and expertise. This will improve standards and protocols and lead to more quality-aware decisions in contexts such as conservation and management.

6.6 Literature Cited

Araújo, M.B., & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33:1677-1688.

Arbia, G., Griffith, D., & Haining, R. (1998). Error propagation and modelling in raster GIS: overlay operations. *International Journal of Geographical Information Science*, 12:145-167.

Barry, S., & Elith, J. (2006). Error and uncertainty in habitat models. *Journal of Applied Ecology*, 43:413-423.

Beale, C. M., & Lennon, J. J. (2012). Incorporating uncertainty in predictive species distribution modelling. *Philosophical Transactions of the Royal Society B*, 367:247-258.

Behrens, T., Zhu, A., Schmidt, K., & Scholten, T. (2010). Multi-scale digital terrain analysis and features selection for digital soil mapping. *Geoderma*, 155:175-185.

Beven, K. (2000). On model uncertainty, risk and decision making. *Hydrological Processes*, 14:2605-2606.

Blondel, P., & Gómez Sichi, O. (2009) Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf. *Applied Acoustics*, 70:1288-1297.

Bolstad, P. V., Swank, W., & Vose, J. (1998). Predicting Southern Appalachian overstory vegetation with digital terrain data. *Landscape Ecology*, 13:271-283.

Bonin, O., & Rousseaux, F. (2005) Digital terrain model computation from contour lines: how to derive quality information from artifact analysis. *GeoInformatica*, 9:253-268.

Bouchet, P.J., Meeuwig, J.J., Salgado Kent, C.P., Letessier, T.B., & Jenner, C.K. (2015) Topographic determinants of mobile vertebrate predator hotspots: current knowledge and future directions. *Biological Reviews*, 90:699-728.

Boyce, M.S., Vernier, P.R., Nielsen, S.E., & Schmiegelow, F.K.A. (2002) Evaluating resource selection functions. *Ecological Modelling*, 157:281-300.

Brock, J. C., & Purkis, S. J. (2009). The emerging role of lidar remote sensing in coastal research and resource management. *Journal of Coastal Research*, 53:1-5.

Brown, C. J., Sameoto, J. A., & Smith, S. J. (2012). Multiple methods, maps, and management applications: purpose made seafloor maps in support of ocean management. *Journal of Sea Research*, 72:1-13.

Brown, C. J., Smith, S. J., Lawton, P., & Anderson, J. T. (2011). Benthic habitat mapping: a review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuarine, Coastal and Shelf Science*, 92:502-520.

Brown, D. G., & Bara, T. J. (1994). Recognition and reduction of systematic error in elevation and derivative surfaces from 71/2-minute DEMs. *Photogrammetric Engineering & Remote Sensing*, 60:189-194.

Brown, J.D., & Heuvelink, G.B.M. (2007) The Data Uncertainty Engine (DUE): a software tool for assessing and simulating uncertain environmental variables. *Computers & Geosciences*, 33:172-190.

Cavazzi, S., Corstanje, R., Mayr, T., Hannam, J., & Fealy, R. (2013) Are fine resolution digital elevation models always the best choice in digital soil mapping? *Geoderma*, 195-196:111-121.

Che Hasan, R., Ierodiaconou, D., & Laurenson, L. (2012) Combining angular response classification and backscatter imagery segmentation for benthic biological habitat mapping. *Estuarine, Coastal and Shelf Science*, 97:1-9.

Chow, T.E., & Hodgson, M.E. (2009) Effects of lidar post-spacing and DEM resolution to mean slope estimation. *International Journal of Geographical Information Systems*, 23:1277-1295.

Chust, G., Grande, M., Galparsoro, I., Uriarte, A., & Borja, A. (2010) Capabilities of the bathymetric Hawk Eye LiDAR for coastal habitat mapping: a case study within a Basque estuary. *Estuarine, Coastal and Shelf Science*, 89:200-213.

Clements, A.J., Strong, J.A., Flanagan, C., & Service, M. (2010) Objective stratification and sampling-effort allocation of ground-truthing in benthic-mapping surveys. *ICES Journal of Marine Science*, 67:628-637.

Collier, J.S., & Brown, C.J (2005) Correlation of sidescan backscatter with grain size distribution of surficial seabed sediments. *Marine Geology*, 214:431-449.

Copeland, A., Edinger, E., Devillers, R., Bell, T., LeBlanc, P., & Wroblewski, J. (2013) Marine habitat mapping in support of Marine Protected Area management in a subarctic fjord: Gilbert Bay, Labrador, Canada. *Journal of Coastal Conservation*, 17:225-237.

Costa, B.M., & Battista, T.A. (2013) The semi-automated classification of acoustic imagery for characterizing coral reef ecosystems. *International Journal of Remote Sensing*, 34:6389-6422.

Costa, B.M., Battista, T.A., & Pittman, S.J. (2009) Comparative evaluation of airborne LiDAR and ship-based multibeam SoNAR bathymetry and intensity for mapping coral reef ecosystems. *Remote Sensing of Environment*, 113:1082-1100.

Davies, A.J., & Guinotte, J.M. (2011) Global habitat suitability for framework-forming cold-water corals. *PLoS ONE*, 6:e18483.

De Falco, G., Tonielli, R., Di Martino, G., Innangi, S., Simeone, S., & Parnum, I.M. (2010) Relationships between multibeam backscatter, sediment grain size and Posidonia oceanica seagrass distribution. *Continental Shelf Research*, 30:1941-1950.

DFO (2006). Presentation and review of Southwest Nova Scotia benthic mapping project. *DFO Canadian Science Advisory Secretariat*, Proceedings Series 2006/047.

Dolan, M.F.J., & Lucieer, V.L. (2014) Variation and uncertainty in bathymetric slope calculations using geographic information systems. *Marine Geodesy*, 37:187-219.

Erikstad, L., Bakkestuen, V., Bekkby, T., & Halvorsen, R. (2013). Impact of scale and quality of digital terrain models on predictability of seabed terrain types. *Marine Geodesy*, 36:2-21.

Etnoyer, P., & Morgan, L. E. (2007). Predictive habitat model for deep gorgonians needs better resolution: comment on Bryan & Metaxas (2007). *Marine Ecology Progress Series*, 339:311-312.

Eugenio, F., Marcello, J., & Martin, J. (2015). High-resolution maps of bathymetry and benthic habitats in shallow-water environments using multispectral remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 53:3539-3549.

Felix, N.A., & Binney, D.L. (1989) Accuracy assessment of a Landsat-assisted vegetation map of the coastal plain of the Arctic National Wildlife Refuge. *Photogrammetric Engineering & Remote Sensing*, 55:475-478.

Fielding, A.H., & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24:38-49.

Filin, S. (2003). Recovery of systematic biases in laser altimetry data using natural surfaces. *Photogrammetric Engineering & Remote Sensing*, 69:1235-1242.

Fisher, P. F., & Tate, N. J. (2006). Causes and consequences of error in digital elevation models. *Progress in Physical Geography*, 30:467-489.

Franklin, J. (1995). Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*, 19:474-499.

Franklin, J. (2013). Species distribution models in conservation biogeography: developments and challenges. *Diversity and Distributions*, 19:1217-1223.

Galparsoro, I., Agrafojo, X., Roche, M., & Degrendele, K. (2015) Comparison of supervised and unsupervised automatic classification methods for sediment types mapping using multibeam echosounder and grab sampling. *Italian Journal of Geosciences*, 134:41-49.

Gessler, P., Pike, R., MacMillan, R. A., Hengl, T., & Reuter, H. I. (2009). The future of geomorphometry. In: Hengl, T., & Reuter, H. I. (Eds), *Geomorphometry: Concepts, Software, Applications* (pp. 637-652). Developments in Soil Science, 33. Amsterdam: Elsevier.

Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J. M. C., Aspinall, R., & Hastie, T. (2006). Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology*, 43:386-392.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M. R., Possingham, H. P., & Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, 16:1424-1435.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135:147-186.

Halpern, B.S., Regan, H.M., Possingham, H.P. & McCarthy, M.A. (2006) Accounting for uncertainty in marine reserve design. *Ecology Letters*, 9:2-11.

Harrison, J. W., Ferrie, F. P., Hefford, S. W., Samson, C., Kusevic, K., Mrstik, P., & Iles, P. J. W. (2009). Finding anomalies in high-density LiDAR point clouds. *Geomatica*, 63:397-405.

Heuvelink, G. B. M. (1998). *Error propagation in environmental modelling with GIS*, London: Taylor and Francis, 127 p.

Holmes, K.W. Van Niel, K.P., Radford, B., Kendrick, G.A., & Grove, S.L. (2008) Modelling distribution of marine benthos from hydroacoustics and underwater video. *Continental Shelf Research*, 28:1800-1810.

Hughes-Clarke, J.E. (1997) Are you really getting "full bottom coverage"? http://www.omg.unb.ca/~jhc/coverage_paper.html

Hughes-Clarke, J.E. (2002) The challenge of technology: improving sea-floor mapping methodologies. http://www.omg.unb.ca/AAAS/UNB_Seafloor_Mapping.html

Hughes-Clarke, J.E. (2003a) Dynamic motion residuals in swath sonar data: ironing out the creases. *International Hydrographic Review*, 4:6-23.

Hughes-Clarke, J.E. (2003b) A reassessment of vessel coordinate systems: what is it that we are really aligning? *US Hydrographic Conference 2003*, Biloxi, MS, 1-12.

Hughes-Clarke, J.E., Mayer, L.A., & Wells, D.E. (1996) Shallow-water imaging multibeam sonars: a new tool for investigating seafloor processes in the coastal zone and on the continental shelf. *Marine Geophysical Researches*, 18:607-629.

Hunsaker, C. T., Goodchild, M. F., Friedl, M. A., & Case, T. J. (2001). *Spatial uncertainty in ecology: implications for remote sensing and GIS applications* (pp. 1-402). New York: Springer-Verlag.

Jager, H. I., & King, A. W. (2004). Spatial uncertainty and ecological models. *Ecosystems*, 7:841-847.

Kenny, A. J., Cato, I., Desprez, M., Fader, G., Schüttenhelm, R. T. E., & Side, J. (2003). An overview of seabed-mapping technologies in the context of marine habitat classification. *ICES Journal of Marine Science*, 60:411-418.

Kostylev, V.E., Todd, B.J., Fader, G.B.J., Courtney, R.C., Cameron, G.D.M., & Pickrill, R.A. (2001) Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs. *Marine Ecology Progress Series*, 219:121-137.

Kozak, K. H., Graham, C. H., & Wiens, J. J. (2008). Integrating GIS-based environmental data into evolutionary biology. *Trends in Ecology and Evolution*, 23:141-148.

Lamarche, G., Lurton, X., Verider, A.-L., & Augustin, J.-M. (2011) Quantitative characterisation of seafloor substrate and bedforms using advanced processing of multibeam backscatter – application to Cook Strait, New Zealand. *Continental Shelf Research*, 31:S93-S109.

Lechner, A. M., Langford, W. T., Bekessy, S. A., & Jones, S. D. (2012a). Are landscape ecologist addressing uncertainty in their remote sensing data? *Landscape Ecology*, 27:1249-1261.

Lechner, A.M., Langford, W.T., Jones, S.D., Bekessy, S.A., & Gordon, A. (2012b) Investigating species-environment relationships at multiple scales: differentiating between intrinsic scale and the modifiable areal unit problem. *Ecological Complexity*, 11:91-102.

Lecours, V. (2015) Terrain Attribute Selection for Spatial Ecology (TASSE), v.1.0, URL www.marinegis.com.

Lecours, V., & Devillers, R. (2015). Assessing the spatial data quality paradox in the deep-sea. In: Sieber, R. E. (Ed.), *Proceedings of Spatial Knowledge and Information – Canada 2015*, pp. 1-8.

Lecours, V., Devillers, R., Schneider, D.C., Lucieer, V.L., Brown, C.J., & Edinger, E.N. (2015) Spatial scale and geographic context in benthic habitat mapping: review and future directions. *Marine Ecology Progress Series*, 535, 259-284.

Lecours, V., Dolan, M.F.J., Micallef, A., & Lucieer, V.L. (2016) A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrology and Earth System Science*, 20:3207-3244.

Leempoel, K., Parisod, C., Geiser, C., Daprà, L., Vittoz, P., & Joost, S. (2015). Very high-resolution digital elevation models: are multi-scale derived variables ecologically relevant? *Methods in Ecology and Evolution*, 6:1373-1383.

Lek, S. (2007). Uncertainty in ecological models. *Ecological Modelling*, 207:1-2.

Le Pape, O., Delavenne, J., & Vaz, S. (2014) Quantitative mapping of fish habitat: a useful tool to design spatialised management measures and marine protected area with fishery objectives. *Ocean & Coastal Management*, 87:8-19.

Levin, S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, 73:1943-1983.

Li, D., Zhang, J., & Wu, H. (2012) Spatial data quality and beyond. *International Journal of Geographical Information Science*, 26:2277-2290.

Li, F. K., & Goldstein, R. M. (1990). Studies of multibaseline spaceborne interferometric synthetic aperture radars. *IEEE Transactions on Geoscience and Remote Sensing*, 28:88-97.

Lichti, D.D., & Skaloud, J. (2010), Registration and calibration. In: Vosselman, G., & Maas, H.-G. (Eds), *Airborne and terrestrial laser scanning* (pp. 83-133). CRC Press Inc.

Livne, E., & Svoray, T. (2011). Components of uncertainty in primary production model: the study of DEM, classification and location error. *International Journal of Geographical Information Science*, 25:473-488.

Lucieer, V., Barrett, N., Hill, N., Nichol, S.L. (2012) Characterization of shallow inshore coastal reefs on the Tasman Peninsula, Southeastern Tasmania, Australia. In: Harris, P.T., & Baker, E.K. (Eds), *Seafloor geomorphology as benthic habitat*, Amsterdam: Elsevier, p. 481-492.

Lurton, X. (2010) *An introduction to underwater acoustics: principles and applications, 2nd edition*. Springer/Praxis Publishing, Berlin, Germany, 704 p.

Lurton, X., & Lamarche, G. (2015) *Backscatter measurements by seafloor mapping sonars, guidelines and recommendations*. GeoHAB, 200 p.

Mateo Sánchez, M.C., Cushman, S.A., & Saura, S. (2014) Scale dependence in habitat selection: the case of the endangered brown bear (Ursus arctos) in the Cantabrian Range (NW Spain*). International Journal of Geographical Information Science*, 28:1531-1546.

Micallef, A., Le Bas, T.P., Huvenne, V.A.I., Blondel, P., Hühnerback, V., & Deidun, A. (2012) A multi-method approach for benthic habitat mapping of shallow coastal areas with high-resolution multibeam data. *Continental Shelf Research*, 39-40:14-26.

Miller, J. (2010). Species distribution modeling. *Geography Compass*, 4:490-509.

Monk, J., Ierodiaconou, D., Versace, V.L., Bellgrove, A., Harvey, E., Rattray, A., Laurenson, L., & Quinn, G.P. (2010) Habitat suitability for marine fishes using presence-only modelling and multibeam sonar. *Marine Ecology Progress Series*, 420:157-174.

Moudrý, V., & Šímová, P. (2012). Influence of positional accuracy, sample size and scale on modelling species distributions: a review. *International Journal of Geographical Information Scien*ce, 26:2083-2095.

Nelson, A., Reuter, H. I., & Gessler, P (2009). DEM production methods and sources. In: Hengl, T., & Reuter, H. I. (Eds), *Geomorphometry: Concepts, Software, Applications* (pp. 65-85). Developments in Soil Science, 33. Amsterdam: Elsevier.

NHS (Norwegian Mapping Authority, Hydrographic Service) (2013) *Specifications for seabed mapping within the MAREANO program*. 17 p.

Nikolakopoulos, K. G., Choussiafis, C., & Karathanassi, V. (2015). Assessing the quality of DSM from ALOS optical and radar data for automatic drainage extraction. *Earth Science Informatics*, 8:293-307.

Pearson, R. G., Thuiller, W., Araújo, M. B., Martinez-Meyer, E., Brotons, L., McClean, C., Miles, L., Segurado, P., Dawson, T. P., & Lees, D. C. (2006). Model-based uncertainty in species range prediction. *Journal of Biogeography*, 33:1704-1711.

Peters, J., Verhoest, N. E. C., Samson, R., van Meirvenne, M., Cockx, L., & De Baets, B. (2009). Uncertainty propagation in vegetation distribution models based on ensemble classifiers. *Ecological Modelling*, 220:791-804.

Phillips, S.J., Anderson, R.P., & Schapire, R.E. (2006) Maximum entropy modeling of species distributions. *Ecological Modelling*, 190:231-259.

Podobnikar, T. (2009). Methods for visual quality assessment of a digital terrain model. *Survey and Perspectives Integrating Environment and Society (S.A.P.I.E.N.S.)*, 2:1-10.

Rattray, A., Ierodiaconou, D., Laurenson, L., Burq, S., & Reston, M. (2009) Hydro-acoustic remote sensing of benthic biological communities on the shallow South East Australian continental shelf. *Estuarine, Coastal and Shelf Science*, 84:237-245.

Regan, H. M., Ben-Haim, Y., Langford, B., Wilson, W. G., Lundberg, P., Andelman, S. J., & Burgman, M. A. (2005). Robust decision-making under severe uncertainty for conservation management. *Ecological Applications*, 15:1471-1477.

Rengstorf, A.M., Grehan, A., Yesson, C., & Brown, C. (2012) Towards high-resolution habitat suitability modeling of vulnerable marine ecosystems in the deep-sea: resolving terrain attribute dependencies. *Marine Geodesy*, 35:343-361.

Reuter, H.I., Hengl, T., Gessler, P., & Soille, P. (2009) Preparation of DEMs for geomorphometric analysis. In: Hengl, T., & Reuter, H. I. (Eds), *Geomorphometry: Concepts, Software, Applications* (pp. 87-120). Developments in Soil Science, 33. Amsterdam: Elsevier.

Roberts, J.M., Davies, A.J., Henry, L.A., Dodds, L.A., Duineveld, G.C.A., Lavaleye, M.S.S., Maier, C., van Soest, R.W.M., Bergman, M.J.N., Hühnerbach, V., Huvenne, V.A.I., Sinclair, D.J., Watmough, T., Long, D., Green, S.L., & van Haren, H. (2009) Mingulay reef complex: an interdisciplinary study of cold-water coral habitat, hydrography and biodiversity. *Marine Ecology Progress Series*, 397:139-151.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J. M., Jiménez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. (2011). Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Progress in Physical Geography*, 35:211-226.

Roman, C., & Singh, H. (2006) Consistency based error evaluation for deep sea bathymetric mapping with robotic vehicles. *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, pp. 3568-3574.

Rousseaux, F. (2003). Étude de l'impact de la représentation du relief sur les applications. *Revue Internationale de Géomatique*, 13:493-509.

Shi, W. (2010). *Principles of modeling uncertainties in spatial data and spatial analyses*, Boca Raton, FL: CRC Press, 412 p.

Sofia, G., Pirotti, F., & Tarolli, P. (2013). Variations in multiscale curvature distribution and signatures of LiDAR DTM errors. *Earth Surface Processes and Landforms*, 38:1116-1134.

Solan, M., Germano, J. D., Rhoads, D. C., Smith, C., Michaud, E., Parry, D., Wenzhöfer, F., Kennedy, B., Henriques, C., Battle, E., Carey, D., Iocco, L., Valente, R., Watson, J., & Rosenberg, R. (2003). Towards a greater understanding of pattern, scale and process in marine benthic systems: a picture is worth a thousand worms. *Journal of Experimental Marine Biology and Ecology*, 285-286:313-338.

Su, J., & Bork, E. (2006). Influence of vegetation, slope, and lidar sampling angle on DEM accuracy. *Photogrammetric Engineering & Remote Sensing*, 72:1265-1274.

Synes, N. W., & Osborne, P. E. (2011). Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Global Ecology and Biogeography*, 20:904-914.

Todd, B.J, Kostylev, V.E., & Smith, S.J. (2012) Seabed habitat of a glaciated shelf, German Bank, Atlantic Canada. In: Harris, P.T., & Baker, E.K. (Eds), *Seafloor geomorphology as benthic habitat* (pp. 555-568). Amsterdam: Elsevier.

van Horssen, P. W., Pebesma, E. J., & Schot, P. P. (2002). Uncertainties in spatially aggregated predictions from a logistic regression model. *Ecological Modelling*, 154:93-101.

van Niel, K. P., & Austin, M. P. (2007). Predictive vegetation modeling for conservation: impact of error propagation from digital elevation data. *Ecological Applications*, 17:266-280.

van Niel, K. P., Laffan, S. W., & Lee, B. G. (2004). Effect of error in the DEM on environmental variables for predictive vegetation modelling. *Journal of Vegetation Science*, 15:747-756.

Vaughan, I.P., & Ormerod, S.J. (2005) The continuing challenges of testing species distribution models. *Journal of Applied Ecology*, 42:720-730.

Warren, D.L., & Seifert, S.N. (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, 21:335-342.

Watling, J. I., Brandt, L. A., Bucklin, D. N., Fujisaki, I., Mazzotti, F. J., Romañach, S. S., & Speroterra, C. (2015). Performance metrics and variance partitioning reveal sources of uncertainty in species distribution models. *Ecological Modelling*, 309-310:48-59.

Wheatley, M., & Johnson, C. (2009) Factors limiting our understanding of ecological scale. *Ecological Complexity*, 6:150-159.

Williams, K. J., Belbin, L., Austin, M. P., Stein, J. L., & Ferrier, S. (2012). Which environmental variables should I use in my biodiversity model? *International Journal of Geographical Information Science*, 26:2009-2047.

Wilson, J. P. (2012). Digital terrain modeling. *Geomorphology*, 137:107-121.

Wilson, M.F.J., O'Connell, B., Brown, C., Guinan, J.C., & Grehan, A.J. (2007) Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy*, 30:3-35.

Wise, S. (2000) Assessing the quality for hydrological applications of digital elevation models derived from contours. *Hydrological Processes*, 14:1909-1929.

Wu, J., Jones, K. B., Li, H., & Loucks, O. L. (2006). *Scaling and uncertainty analysis in ecology: methods and applications*, Dordrecht, The Netherlands: Springer, 351 p..

Zandbergen, P. A. (2011). Characterizing the error distribution of lidar elevation data for North Carolina. *International Journal of Remote Sensing*, 32:409-430.

Zieger, S., Stieglitz, T., & Kininmonth, S. (2009) Mapping reef features from multibeam sonar data using multiscale morphometric analysis. *Marine Geology*, 264:209-217.

7. Conclusion

Mapping marine benthic habitats has become an important practice supporting marine conservation and resources management. The different approaches to map benthic habitats typically involve the integration of various types of data into a same geographic framework. In the context of an increasing availability and user-friendliness of GIS tools, the producers of those maps are not always aware of theoretical foundations of these tools, which have roots in geography and geomatics. The issue is amplified when techniques from marine geomorphometry, another field with strong connections in geography and geomatics, are integrated within the habitat mapping workflow. Since marine habitat mapping is strongly data-driven, and that these data are most often of spatial nature, marine habitat mapping practices are directly affected by these theoretical foundations, or spatial concepts (*e.g.* spatial scale, spatial autocorrelation). A lack of understanding of how these spatial concepts impact our representation of benthic ecosystems and ultimately our understanding of these environments and their dynamics can potentially lead to misinformed and inappropriate conservation and management decisions. This dissertation reviewed how marine habitat mapping practices lack the proper consideration of some core spatial concepts and proposed best practices to (re)integrate these concepts in the marine habitat mapping workflow. A particular focus was given to issues of spatial scale, spatial covariation and spatial data quality when marine geomorphometry is integrated in the habitat mapping workflow.

7.1 Summary of Findings

7.1.1 Findings

This dissertation aimed at answering four research questions. The first question was: "Which particular spatial concepts are poorly integrated in the marine benthic habitat mapping workflow, and is it possible to identify specific ways to improve the integration of spatial concepts in marine benthic habitat mapping?" To answer this question, a review of the marine benthic habitat mapping literature was presented in Chapter 2. In this review, the importance of incorporating ecological scaling and geographic theories to this field was highlighted. It was found that spatial scale is one of the core spatial concepts in both ecology and geography, and that its role in benthic habitats is not well understood from an ecological perspective and a geospatial representation perspective. This review also showed that despite repeated calls for an improved consideration of spatial scale in ecology and the adoption of multiscale approaches in habitat characterization, these two elements have yet to be fully implemented and adopted by the community. Other spatial concepts that were found to be poorly implemented into the benthic habitat mapping workflow include spatial covariation, spatial data quality, spatial autocorrelation, and spatial heterogeneity. Chapter 2 also demonstrated the intrinsic link between these spatial concepts and spatial scale.

The second question of this dissertation was linked to the use of marine geomorphometry in marine benthic habitat mapping. The question was: "Which combinations of terrain attributes best capture seafloor characteristics while minimizing spatial covariation, and can these terrain attributes form a standard protocol for using marine geomorphometry in different approaches to habitat mapping?" The first part of this question was answered in Chapter 3. Independent groups of correlated terrain attributes were identified, and an optimal combination of terrain attributes was extracted from these groups. This chapter also enabled the description of different algorithms and tools available to derive terrain attributes, and it was found that different algorithms computing a same terrain attribute could produce different outcomes. It was also shown that terrain attributes show a high level of spatial covariation and are sometimes ambiguously defined within software documentation, thus confirming that tools are often "black-box".

The second part of this second question was answered in Chapter 4, in which the appropriateness of the optimal combination of terrain attributes was confirmed when it performed better than other combinations in both top-down and bottom-up approaches to habitat mapping. Chapter 4 also highlighted the extent of the sensitivity of habitat maps and species distribution models to the choice of input variables, and how measures of map accuracy may not be representative of the magnitude of these variations in terms of the spatial extent and distribution of habitats. The analysis performed in this chapter also provided insights on the biophysical characteristics of German Bank and the distribution of sea scallops. On German Bank, the benthoscape habitats were not mostly driven by terrain morphology but were strongly delineated by sediment characteristics. Sea scallop distribution was found to be driven by depth, sediment properties and seafloor rugosity. Finally, a methodological finding of this chapter was the demonstration of the inappropriateness of overall accuracy compared to kappa coefficient of agreement when quantifying classification performance.

The third question asked in this dissertation was: "How sensitive are terrain attributes to different types of data acquisition artefacts, and does that sensitivity vary with spatial scale?" Results showed that artefacts affected terrain attributes significantly, and finer-scale data were generally more impacted by artefacts than broader-scale data. The findings show that roll and pitch artefacts have a more significant negative impact on bathymetry and terrain attributes than time and heave artefacts. Additional findings demonstrated that the same terrain attributes of the same areas are not capturing the same information when they are computed at different spatial scales. Analyses also showed that data affected by artefacts can be significantly different than data without artefacts, thus providing different – and false – information on terrain characteristics. This observation was more acute at finer scales. An unexpected result was when the representation of terrain complexity, quantified using fractal dimension, was greater at broader scales than at finer scales (*i.e.* the terrain was represented as being more complex at broader scales).

The last question asked was "Do data acquisition artefacts propagate to habitat maps and species distribution models, and if yes, are the impacts scale-dependent?" This question related to artefact propagation from the marine geomorphometry workflow (*cf.* previous question and Chapter 5) to the marine benthic habitat mapping workflow. Results showed that artefacts do impact habitat maps and species distribution models, although not in predictable ways – with the exception of roll artefacts. Maps impacted by roll followed the expected pattern that the higher the level of artefact, the lower is habitat map accuracy. Maps impacted by other types of artefacts did show alterations but not in a predictable way: the classification accuracy of some maps impacted by artefacts was higher than that of the reference maps, while the accuracy of other maps was lower than that of the reference maps. Similar conclusions were made from results from the species distribution models analysis. However, in this case roll artefacts were artificially increasing model performance, likely because roll artificially increases the rugosity of an area and sea scallop distribution has been shown to be driven by rugosity. While the benthoscape classes of German Bank were better delineated at coarser scales, sea scallop distribution was best predicted at 50 m resolution. Findings also confirmed some of those from Chapter 4, including that the choice of input variables must be carefully made as it changes significantly the map and model outputs. Results also show that sediment properties are important delineators of benthoscape classes on German Bank.

7.1.2 Research Hypotheses

The literature review outlined in Chapter 2 confirmed that my first hypothesis, which was that there is a poor understanding of the role of different spatial concepts in marine benthic habitat mapping, was true. The concepts of spatial heterogeneity, spatial dependency, spatial covariation, spatial scale, spatial data quality, spatial representation and spatial data selection were all identified as being poorly integrated in the benthic habitat mapping workflow.

The study presented in Chapter 3 showed that my second hypothesis, which was that different optimal combinations of terrain attributes would be found and vary based on seafloor characteristics, was wrong. While optimal combinations of terrain attributes were found, results show that they do not vary with seafloor characteristics like roughness. The importance of individual terrain attributes do vary with terrain complexity, but their combination is valid for all types of terrain.

The third research hypothesis was that the different optimal combinations of terrain attributes would be generalizable, *i.e.* they could be integrated in the marine habitat mapping workflow regardless of the approach used to map habitats. This hypothesis was validated by results from Chapter 4, which showed that the optimal combination of terrain attributes defined in Chapter 3 is applicable to real benthic habitat mapping exercises. The proposed combination was found to work best compared to other combinations in both top-down and bottom-up approaches.

Part of the fourth hypothesis was tested in Chapter 5 while another part was tested in Chapter 6. My hypothesis was that terrain attributes, and habitat maps and species distribution models produced from these terrain attributes, would be sensitive to data acquisition artefacts. That hypothesis was validated: results from Chapter 5 demonstrated that artefacts affect terrain attributes in a predictable negative way, and results from Chapter 6 showed that artefacts also affect habitat maps and species distribution models but in an often unpredictable way (except for roll artefacts).

Finally, the last hypothesis was about the scale-dependence of the sensitivity of terrain attributes, habitat maps and species distribution models to data acquisition artefacts. Finer-scale data and maps/models built from these data were expected to be more sensitive to data acquisition artefacts than broader-scale data and maps/models produced with these data. Results from Chapter 5 showed that the impact of artefacts on terrain attributes is scale-dependent, with finer-scale data being more impacted than broader-scale data. However, results from Chapter 6 showed that while habitat maps and species distribution models are affected by artefacts, this effect is not scale-dependent.

7.2 Research Contributions and Highlights

This dissertation has contributed to the growing body of literature (*e.g.* Rengstorf *et al.*, 2012; Rattray *et al.*, 2014) that looked at improving our understanding of the influence of spatial concepts like spatial scale on the way benthic habitats are represented and understood. Consequently, it raised awareness on the importance of these spatial concepts in marine benthic habitat mapping, and should improve geographic literacy within the community of practitioners. It also provided new standards to define benthic habitats that are explicit about scale, and consider the spatial nature of data and the chemical environment as a potential delineator of habitats.

Spatial scale in marine habitat mapping was addressed in many chapters. Its study enabled finding how representations of environmental phenomena can be altered by the spatial scale at which they are observed. It also showed how important it is to identify the scale(s) that will capture the relevant phenomena, which confirms the need to move towards continuum-based multiscale approaches.

This dissertation also contributed to the definition of standards for using geomorphometry in marine habitat mapping. By statistically finding an optimal combination of terrain attributes to describe all terrains in any contexts, this dissertation provided an operational framework for best using geomorphometry in applications like marine benthic habitat mapping. This will hopefully lead to a standardization of practices, which will then enable proper and valid comparisons among studies making use of geomorphometry. To facilitate the implementation of these new standards, a free toolbox for ArcGIS was provided to end-users, associated with appropriate metadata.

Chapters 4, 5 and 6 yielded one of the most important contributions of this dissertation: while habitat maps are an invaluable tool in contexts such as conservation and management of marine resources to inform and support decision-making, they should always be produced, and their results should always be interpreted, critically and carefully. This research showed that habitat maps are highly sensitive to the variables selected as input, the spatial scale at which these variables are defined, and the quality of these data. That sensitivity was shown both in evaluation measures (*i.e.* overall accuracy, kappa coefficient of agreement, AUCs), but also and most importantly in the spatial distribution of habitats or distribution probabilities.

While this dissertation was focused on marine benthic habitat mapping practices, it also contributed knowledge to the relatively recent field of marine geomorphometry (reviewed in Lecours *et al.*, 2016). The combination of terrain attributes that was proposed in Chapter 3 is applicable to any geomorphometry applications, whether it is archeology, geomorphology, or hydrodynamics modelling. Chapter 5 also assessed the impact of common data acquisition artefacts in bathymetric data that are often used in marine geomorphometry. This was the first study of this kind in the marine geomorphometry literature.

7.3 Future Directions

7.3.1 Limitations and Future Opportunities

The methods used in this dissertation all have some kinds of limitations. In many chapters, geomorphometric analyses were performed using a single analytical scale/window of analysis of 3x3 pixels, which is the default setting in many GIS software. Keeping the analytical scale constant enabled performing multiscale analyses (rather than multi-design) as only the observation scale (*i.e.* the spatial resolution of the data) was changed. However, future work should try to assess if the different observations made in this dissertation vary with different analytical scales. This could be further explored by looking at the five different methods to generate terrain attributes at multiple scales, as presented in Dolan (2012). The five methods include (1) resampling the bathymetry before deriving terrain attributes, (2) averaging the bathymetry over different analytical windows and then deriving terrain attributes, (3) deriving terrain attributes from bathymetry before averaging results over different analytical windows, (4) deriving terrain attributes using different analytical windows, and (5) using the multiscale method developed by Wood (1996) that calculates terrain attributes across a series of analytical windows and reports the mean value and standard deviation of terrain attributes values across analytical scales. Future work should try to evaluate differences in map outputs when the different methods for multiscale geomorphometric analysis are used in a habitat mapping context.

Another limitation of this research is the lack of consideration of data other than topographic and sediment variables. Such data (*e.g.* oceanographic data) were not available at the same spatial scale as the multibeam data, but could have improved maps and models and yielded additional information on the role of terrain morphology in the studied habitats and in determining sea scallop distribution. Depth has been shown to be a good driver of sea scallop distribution, but it may be an indirect surrogate that is actually acting as a proxy of some other environmental gradient, *e.g.* in the chemical characteristics of the environment.

In terms of the impacts of artefacts on terrain attributes and habitat maps, the conclusions reached in this dissertation are limited to the four types of artefacts that were studied. Future work should look at different types of artefacts and study the impact of artefacts of random amplitude, which would be more characteristic of bathymetric surveys. In addition, future studies should asses how the impacts from the combination of different types of artefacts differ from those of individual artefacts. The impacts on habitat maps of errors like artefacts in other types of data, for instance in backscatter data, should also be described as they are also likely to significantly impact maps and models.

Finally, while this dissertation demonstrated the impact of spatial scale and spatial data quality on habitat maps, raising awareness on these issues, it did not provide concrete solutions for end-users (apart from the proposed selection of terrain attributes and the TASSE toolbox for ArcGIS). Future work will need to focus on providing tools, standards and protocols for instance for error modelling or fitness for use assessment. Now that the problems and some solutions/best practices have been identified, research should focus on carrying them forward in a way that they can be easily implemented and widely adopted by the marine geomorphometry and marine habitat mapping communities.

7.3.2 Emerging Questions

The conclusions from this dissertation yielded new questions that could drive future research. These questions include:

1. Do the five methods to implement multiscale analyses in geomorphometry, suggested by Dolan (2012), produce different outputs when applied to the same DTM? If they do, how does it impact habitat maps and species distribution models, and can we define best practices or standards based on that assessment?
2. While this dissertation addressed methods from general geomorphometry (*i.e.* that deals with continuous surfaces and provide terrain attributes like slope), can we define standards for the use of specific geomorphometry (*i.e.* that deals with the extraction of specific landforms like moraines) to ensure that geomorphometry is used and developed to its full potential in the marine benthic habitat mapping workflow?
3. Can we use methods from both types of geomorphometry (*i.e.* general and specific) to automatically identify artefacts in bathymetry and extract them based on a set of predefined rules? Which methods would be adequate to interpolate the areas left out by the removal of the artefacts?
4. Do data acquisition artefacts in backscatter data propagate to habitat maps and species distribution models? If yes, what are the combined impacts of data

acquisition artefacts in both bathymetric and backscatter data on habitat maps and species distribution models?

5. Related to the increase in data uncertainty with depth when surveying with submersible platforms demonstrated in Lecours & Devillers (2015), can we define survey protocols by evaluating the different trade-offs (*e.g.* in spatial scale, spatial data quality, equipment, financial resources) involved in data acquisition so that efforts are not spent collecting uncertain and poor quality data in the deep sea?
6. Can we provide a tool to enable proper quantification of uncertainty (*e.g.* the Data Uncertainty Engine by Brown & Heuvelink, 2007) and spatial error at the end of an analysis workflow? If it is based on a predefined set of rules, would it be able to inform the end-user of the finest spatial scale at which the data should be used in order to minimize the impact of uncertainty and errors on the results and interpretation?
7. Which geovisualization tools could be implemented to enable a clear spatial representation of uncertainty and errors associated with habitat maps and models, which would be suitable for decision-makers and would not be associated with negative perceptions of uncertainty and errors from the end-users?

7.3.3 Recommendations for Marine Habitat Mapping Practices

The findings in this dissertation enabled making many recommendations to improve the integration of spatial concepts in the marine benthic habitat mapping workflow.

In Chapter 2, recommendations were made to improve the marine benthic habitat mapping workflow, based on a review of past and current trends in this discipline regarding the integration of spatial concepts. First, I recommended prioritizing ecosystem engineers or umbrella species' habitats for mapping and prediction, and sampling the environment in a way that covers an extensive range of scales and environmental characteristics. Then, I suggested moving away from single scale habitat characterization and adopting continuum-based habitat characterization approaches. This was followed by a recommendation to use statistical methods and analytical approaches that consider the spatial nature of data. From a technical point of view, I recommended quantifying spatial errors and uncertainty at every step of the habitat mapping workflow, automating existing tools, and developing new tools (*e.g.* GIS, statistical, ecological) for processing data and defining surrogates of species distribution and habitat at multiple scales. A key element related to these suggestions is related to the recommendation to make data, metadata, and tools available with appropriate documentation to maximize research and applications potential. Finally, I recommended explicitly reporting the spatial extent and resolution at which the research was intended to be conducted, the data were collected, and at which the goals of the habitat mapping exercise were directed. I also proposed new standards for defining benthic habitats, which are "areas of seabed that are (geo)statistically significantly different from their surroundings in terms of physical, chemical and biological characteristics, when observed at particular spatial and temporal scales". This definition is different from previous ones because of the addition of chemical characteristics as components of the environment, the explicit consideration of different types of scales, and the consideration of the spatial nature of data.

In Chapter 3, the study of available terrain attributes yielded recommendations regarding an optimal combination of terrain attributes to use in ecological studies like habitat mapping. This combination includes slope, easterness and northerness (*i.e.* two measures of orientation, or aspect), local standard deviation (*i.e.* a measure of local rugosity), local mean, and relative difference to mean value (*i.e.* a measure of relative position). I recommended using that combination because it reduces redundancy, covariation, and ambiguity, while improving generality and replicability. Based on results from this analysis that demonstrate a high level of covariation among terrain attributes, I also recommended explicitly reporting the software, algorithms, and parameters used to generate terrain attributes in any studies to enable transparency and comparisons among studies. Finally, I encouraged software and tools developers to be explicit in their documentation and metadata about the methods or algorithms used by their products.

Based on insights gained in Chapter 4, I recommend selecting input variables with an ecological meaning and that are not covarying when mapping habitats. Because of the sensitivity of habitat maps to variable selection, I recommended that stakeholders prepare more than a single map using different combinations of environmental variables, and that they select the best outcome based on a combination of expert knowledge of the area and species or habitat, and map accuracy or model performance quantification. When using measures of accuracy like kappa coefficients of agreement and overall accuracies, I also recommended using the former over the latter as it was more consistent and more representative of the overall classification accuracy.

In Chapters 5 and 6, I recommended using local mean over the original bathymetry when important artefacts are present in the bathymetry. I also recommended that users always acknowledge the presence of artefacts in their work and discuss their potential implications for their particular analyses. Similarly to recommendations made in other chapters, I encourage users to make use of data with an ecological meaning and at scales that are relevant to the species or habitat being studied. One of the most important recommendations from these chapters concerns the integration of error modelling in marine geomorphometry and marine habitat mapping to enable proper decisions regarding fitness for use of data for particular applications.

7.4 Conclusions

In conclusion, the field of marine benthic habitat mapping will keep evolving as new tools and methods are developed. Since habitat maps are often used in decision-making processes, it is crucial to make sure that we provide the best possible information to decision-makers. Such information will come from the adoption of multiscale approaches and methods that consider the spatial nature of data, and a better understanding and quantification of error and uncertainty in our data and analyses. Tools that enable an easy but transparent implementation of these concepts will be necessary to assist habitat mapping practitioners in making habitat maps that are grounded on a sound, spatially-explicit basis. By improving standards and protocols and implementing practices like the assessment of fitness for use in the habitat mapping workflow, the efficiency of habitat maps to represent benthic habitats and provide the relevant information to decision-makers will increase, together with the trust that is put in these maps. That way, these maps will become an even more powerful communication tools than they currently are.

7.5 Literature Cited

Anderson, J.T., Holliday, D.V., Kloser, R., Reid, D.G., & Simard Y. (2008) Acoustic seabed classification: current practice and future directions. *ICES Journal of Marine Science*, 65:1004-1011.

Brown, J.D., & Heuvelink, G.B.M. (2007) The Data Uncertainty Engine (DUE): a software tool for assessing and simulating uncertain environmental variables. *Computers & Geosciences*, 33:172-190.

Dolan, M.F.J. (2012) Calculation of slope angle from bathymetry data using GIS – effects of computation algorithms, data resolution and analysis scale. *Geological Survey of Norway*, report no. 2012.041, 44 p.

Lecours, V., & Devillers, R. (2015) Assessing the spatial data quality paradox in the deep-sea. In: Sieber, R.E. (Ed.) *Proceedings of Spatial Knowledge and Information Canada*, 2015:1–8.

Lecours, V., Dolan, M.F.J., Micallef, A., & Lucieer, V.L. (2016) A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrology and Earth System Sciences*, 20:3207-3244.

Wood, J. (1996) The geomorphological characterisation of digital elevation models. PhD Thesis, University of Leicester, U.K.

Appendix A: Detailed Material and Methods (Chapter 3)

A.1 Surfaces and Terrain Attributes

A.1.1 Artificial Surfaces

Despite not being perfectly fractal (Halley *et al.*, 2004), real terrains often demonstrate fractal-like properties (Milne, 1992; With & King, 1997) and several authors indicate that such fractal-based surfaces are appropriate to develop “null hypotheses” (Evans & McClean, 1995; Tate, 1998; Halley *et al.*, 2004). Because of the scale-dependency of terrain attributes and topography (Tate & Wood, 2001), a scale-invariant measure was essential to characterize the complexity of the representation of the surface (*i.e.* the Digital Terrain Model (DTM)), rather than the complexity of the terrain itself. For instance, if a rough terrain is represented using a broad-resolution DTM, it may appear smooth: since terrain attributes are dependent on the DTM and not the real terrain, a suitable subset of terrain attributes to characterize this particular DTM would be one that is appropriate for smooth surfaces. This information would be captured by the fractal dimension of the DTM. If the DTM had a higher spatial resolution, it would represent better the roughness of the terrain, the fractal dimension would be higher, and the appropriate subset of terrain attributes would be chosen accordingly. A surface’s fractal dimension can theoretically range from 2.0 (very smooth) to 2.9 (very complex) (Peterson, 1984).

A.2 Principal Component Analysis

A.2.1 Identifying the Optimal Number of Components

Minimum Average Partial Correlation (MAP) (Velicer, 1976) and Parallel Analysis (PA) (Horn, 1965) are commonly used and recommended by statisticians as they tend to give better results than other methods (Zwick & Velicer, 1986; Henson & Roberts, 2006). MAP is however recognized to sometimes extract too few components (O'Connor, 2000) while PA is recognized to sometimes extract too many components (Buja & Eyuboglu, 1992). The revised MAP (Velicer *et al.*, 2000) raises the partial correlations used in the calculation to the fourth power rather than squared. The modified PA (O'Connor, 2015) uses the raw dataset (*i.e.* the actual values of terrain attributes) rather than theoretical values to determine the number of components. This method is considered very accurate and relevant for datasets that are not normally distributed, which is often the case in environmental datasets (Austin, 1987; O'Connor, 2015).

A.2.2 Complexity of Variables

The complex variables that loaded significantly more on one component than the others were kept in the analysis as they may help exploring the behaviour of variables during interpretation.

A.2.3 Validation

A component was found unreliable when its Cronbach's α index (Cronbach, 1951) was lower than 0.6 (Nunnally, 1978). Although, during the computation of α, SPSS also computes the potential values of α if each individual variable is removed from the component. This allowed checking if only one specific variable was making the component unreliable. When this was the case, this specific variable was removed from the component (step 4B, Figure 3.1) and the PCA was re-run (step 4C, Figure 3.1).

A.3 Covariation Assessment: Variable Inflation Factor and Mutual Information

These covariation measurements lack meaningful thresholds to separate the variables that demonstrate covariation to those that do not (Belsey *et al.*, 2004), but allow ranking the attributes from least covarying to most covarying. Such ranking is often performed in machine-learning as a pre-processing step (Kohavi & John, 1997), and has proven to be computationally efficient and statistically robust (Guyon & Elisseeff, 2003).

Stepwise calculations of VIF and MI were necessary because of the changing levels of co-association with variables being removed from the datasets. The two stepwise algorithms were computed in the statistical software R v. 3.1.1 and work the same way: they (1) rank terrain attributes based on the calculation of the VIF or MI measure of each of them, (2) remove the most redundant or least informative terrain attribute, (3) save it in a separate list, and (4) repeat the process until all the attributes are ranked in the list.

A.4 Adequacy of Methods: Replicability, Reliability and Generalization

The Global Moran's I values of the artificial surfaces (Figure A.1) show high spatial autocorrelation of the elevation values, which is expected in environmental data (Legendre, 1993). We thus believe that the artificial surfaces are good surrogates of natural terrain. In addition, the fractal dimension tested cover most of what can be find in natural environment (Hofierka *et al.*, 2009; Zawada & Brock, 2009).

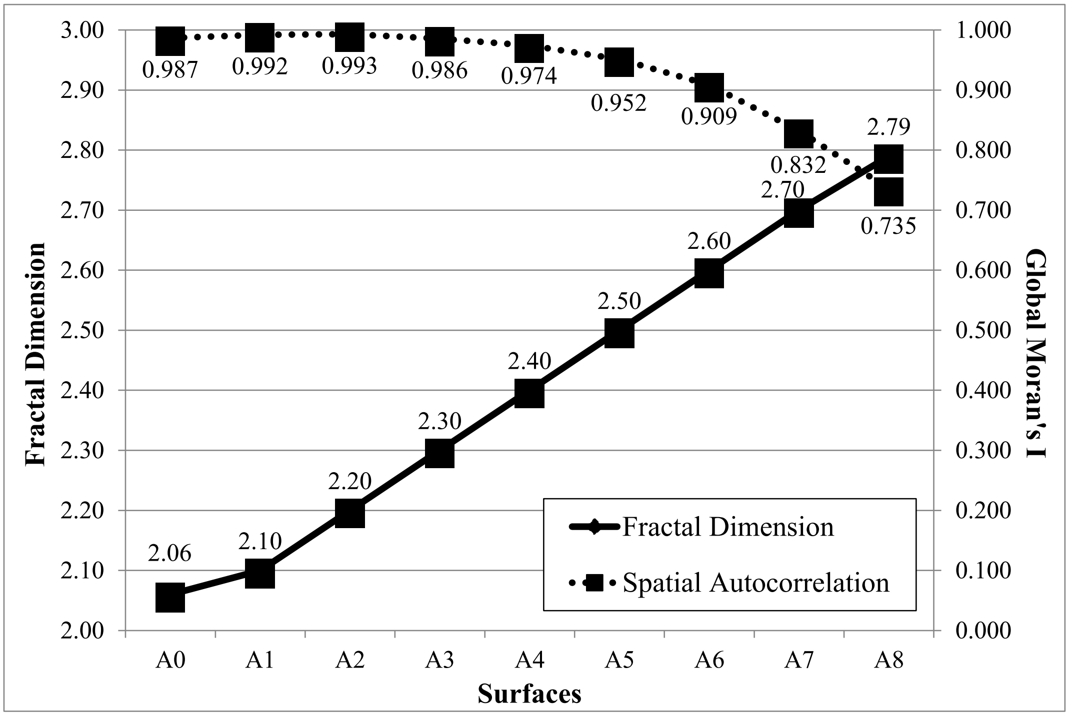

Figure A.1: Fractal dimensions and Moran's I values of the nine artificial surfaces.

The comparability of the study design applied on each surface (Figure 3.1) allowed us to compare PCA solutions and draw conclusions from the comparisons (Rummel, 1970). The 10,000 pixels included in the statistical analyses are considered enough to obtain generalizable and replicable results and produce more accurate solutions (Barrett & Kline, 1981; Costello & Osborne, 2005); Comrey & Lee (1992) advise that more than 1,000 samples is excellent, but that 10 observations per variables is also good. 10,000 samples is thus significantly more than the 1,740 samples that would have been necessary (230 variables – 8 low cardinality variables – 48 duplicates = 174 variables in the PCA).

According to Gorsuch (1983), communality is important for replicability as it assesses the appropriateness of the PCA model and consequently serves to validate the method. In the current study, the high average communalities of each solution are an indication that the iterative PCA were appropriate, stable, and replicable (Cliff & Pennell, 1967).

The relatively constant variance of the first component across the solutions is an indication of the invariance property of Group 1 of terrain attributes (Kaiser, 1958). An invariant component is highly reliable and replicable as is indicates that the importance of the variables loading on it do not exclusively belong to these solutions, but are inherent to any similar datasets (Gorsuch, 1983).

The results of VIF and MI were compared to make sure that one of the two methods was not consistently ranking variables higher or lower than the other method. For each surface, about half of the variables were ranked higher by one of the two methods, thus indicating that none of them influenced the average more than the other. Since they both measure covariation from different factors, we believe that their average is a good indication of the overall covariation behaviour of the variables.

Finally, the high loadings on each component (Appendix C), high correlations between variables loading on the components (Appendix C), and the meaningfulness of each component (Figure 3.3) are indications of the appropriateness of the method (Wood *et al.*, 1996). These are characteristics of a simple structure solution that allows for generalization of results (Kaiser, 1958): the same clusters of variables are consistently found (Figure 3.3 and Figure 3.4).

A.5 Literature Cited

Austin, M.P. (1987) Models for the analysis of species' response to environmental gradients. *Vegetatio*, 69:35-45.

Barrett, P.T., & Kline, P. (1981) The observation of variable ratio in factor analysis. *Personality Study and Group Behavior*, 1:23.

Belsey, D.A., Kuh, E., & Welsch, R.E. (2004) *Regression diagnostics: identifying influential data and sources of multicollinearity*. Wiley, Hoboken.

Buja, A., & Eyuboglu, N. (1992) Remarks on parallel analysis. *Multivariate Behavioral Research*, 27:509-540.

Cliff, N., & Pennell, R. (1967) The influence of communality, factor strength, and loading size on the sampling characteristics of factor loadings. *Psychometrika*, 32:309-326.

Comrey, A.L., & Lee, H.B. (1992) *A first course in factor analysis*. Erlbaum, Hillsdale.

Costello, A.B., & Osborne, J. (2005) Best practices in explanatory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10:1-9.

Cronbach, L.J. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika*, 16:297-334.

Evans, I.S., & McClean, C.J. (1995) The land surface is not unifractal: variograms, cirque scale and allometry. *Zeitschrift für Geomorphologie*, 101:127-147.

Gorsuch, R.L. (1983) *Factor analysis (second edition)*. Erlbaum, Hillsdale.

Guyon, I., & Elisseeff, A. (2003) An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157-1182.

Halley, J.M., Hartley, S., Kallimanis, A.S., Kunin, W.E., Lennon, J.J., & Sgardelis, S.P. (2004) Uses and abuses of fractal methodology in ecology. *Ecology Letters*, 7:254-271.

Henson, R.K., & Roberts, J.K. (2006) Use of explanatory factor analysis in published research: common errors and some comment on improved practice. *Organizational Research Methods*, 7:191-205.

Hofierka, J., Mitášová, H., & Neteler, M. (2009) Geomorphometry in GRASS GIS. In: Hengl, T., & Reuter, H.I. (Eds), *Geomorphometry: concepts, software, applications*, Elsevier Science, Amsterdam, pp. 387-410.

Horn, J.L. (1965) A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30:179-185.

Kaiser, H.F. (1958) The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23:187-200.

Kohavi, R., & John, G. (1997) Wrappers for feature selection. *Artificial Intelligence*, 97:273-324.

Legendre, P. (1993) Spatial autocorrelation: trouble of new paradigm? *Ecology*, 74:1659-1673.

Milne, B.T. (1992) Spatial aggregation and neutral models in fractal landscapes. *American Naturalist*, 139:32-57.

Nunnally, J.C. (1978) *Psychometric Theory (second edition)*. McGraw-Hill, New York.

O'Connor, B.P. (2000) Using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32:396-402.

O'Connor, B.P. (2015) *SPSS, SAS, MATLAB, and R programs for determining the number of components and factors using parallel analysis and Velicer's MAP test*. URL https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html [accessed 21 January 2015]

Peterson, I. (1984) Ants in the labyrinth and other fractal excursions. *Science News*, 125:42-43.

Rummel, R.J. (1970) *Applied factor analysis*. Northwestern University Press, Evanston.

Tate, N.J. (1998) Estimating the fractal dimension of synthetic topographic surfaces. *Computers & Geosciences*, 24:325-334.

Tate, N.J. & Wood, J. (2001) Fractals and scale dependencies in topography. In: Tate, N., & Atkinson, P. (Eds) *Scale in Geographical Information*, Wiley, Chichester, pp. 35-51.

Velicer, W.F. (1976) Determining the number of components from the matrix of partial correlations. *Psychometrika*, 41:321-327.

Velicer, W.F., Eaton, C.A., & Fava, J.L. (2000) Construct explication through factor or component analysis: a review and evaluation of alternative procedures for determining the number of factors or components. In: Goffin, R.D., & Helmes, E. (Eds) *Problems and solution in human assessment*, Kluwer, Boston, pp. 41-71.

With, K.A., & King, A.W. (1997) The use and misuse of neutral landscape models in ecology. *Oikos*, 79:219-229.

Wood, J.M., Tataryn, D.J., & Gorsuch, R.L. (1996) Effects of under- and overextraction on principal axis factor analysis with Varimax rotation. *Psychological Methods*, 1:354-365.

Zawada, D.G., & Brock, J.C. (2009) A multiscale analysis of coral reef topographic complexity using lidar-derived bathymetry. *Journal of Coastal Research*, 53:6-15.

Zwick, W.R., & Velicer, W.F. (1986) Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99:432-442.

Appendix B: Artificial Surfaces and List of Derived Terrain Attributes, with Software, Algorithms, References, and Duplicates (Chapter 3)

Table B.1: List of terrain attributes derived from each surface with references and duplicates.

ID	Attributes Names' in Software	Software	Algorithms/Methods/References	Identical to…
1	Bathymetric Position Index	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Wright et al. (2012)	
2	Center versus Neighbors Variability	Idrisi Selva 17.0	Not Specified	
3	Coefficient of variation	Diva-GIS 7.5.0	Not Specified	
4	Convergence Index	SAGA GIS 2.0.8	Using Aspect	
5	Convergence Index	SAGA GIS 2.0.8	Using Gradient	
6	Cross-sectional Curvature	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Evans (1979)	
7	Cross-sectional Curvature	Landserf 2.3	Not Specified	
8	Curvature	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Not Specified	9
9	Curvature	ArcGis 10.2.2 and Python 2.7.8	Zevenbergen and Thorne (1987)	8
10	Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Bauer et al., 1985)	
11	Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Haralick, 1983)	
12	Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Heerdegen and Beran, 1982)	
13	Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Zevenbergen and Thorne, 1987)	
14	Curvature	SAGA GIS 2.0.8	Least Squares Fitted Plane (Horn, 1981, Costa-Cabral and Burgess, 1996)	16, 123, 125, 143, 145
15	Curvature	SAGA GIS 2.0.8	Maximum Slope (Travis et al., 1975)	
16	Curvature	SAGA GIS 2.0.8	Maximum Triangle Slope (Tarboton, 1997)	14, 123, 125, 143, 145
17	Deviation from Mean Elevation	Whitebox GAT 3.2.1 Iguazu	Not Specified	
18	Deviation from Mean Value	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	21
19	Deviation from Mean Value	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	
20	Deviation from Mean Value	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	
21	Deviation from Mean Value	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	18
22	Difference from Mean Elevation	Whitebox GAT 3.2.1 Iguazu	Not Specified	
23	Difference from Mean Value	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	26, 214, 217
24	Difference from Mean Value	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	
25	Difference from Mean Value	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	216

ID	Attributes Names' in Software	Software	Algorithms/Methods/References	Identical to…
26	Difference from Mean Value	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	23, 214, 217
27	Easterness	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Not Specified	
28	Easterness	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	4-Cell Method (Fleming and Hoffer, 1979)	
29	Easterness	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Horn (1981)	51
30	Easterness	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Sharpnack and Akin (1969)	
31	Easterness	ArcGis 10.2.2 and Python 2.7.8	Zevenbergen and Thorne (1987)	
32	Easterness	Idrisi Selva 17.0	Not Specified	
33	Easterness	Landserf 2.3	Not Specified	
34	Easterness	Quantum GIS 2.4.0 Chugiak	Horn (1981)	
35	Easterness	Quantum GIS 2.4.0 Chugiak	Zevenbergen and Thorne (1987)	
36	Easterness	Quantum GIS 2.4.0 Chugiak	Not Specified	
37	Easterness	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Bauer et al., 1985)	39
38	Easterness	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Haralick, 1983)	
39	Easterness	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Heerdegen and Beran, 1982)	37
40	Easterness	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Zevenbergen and Thorne, 1987)	
41	Easterness	SAGA GIS 2.0.8	Least Squares Fitted Plane (Horn, 1981, Costa-Cabral and Burgess, 1996)	
42	Easterness	SAGA GIS 2.0.8	Maximum Slope (Travis et al., 1975)	
43	Easterness	SAGA GIS 2.0.8	Maximum Triangle Slope (Tarboton, 1997)	
44	Easterness	SAGA GIS 2.0.8	Not Specified	
45	Easterness	TNTmips Free 2014 (MicroImages)	Exact fit to 4 nearest neighbors and center cell	
46	Easterness	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit	47
47	Easterness	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, match central cell	46
48	Easterness	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by 1/distance	
49	Easterness	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by 1/distance^2	
50	Easterness	uDig 1.4.0b with Spatial Toolbox	Not Specified	

ID	Attributes Names' in Software	Software	Algorithms/Methods/References	Identical to…
51	Easterness	Whitebox GAT 3.2.1 Iguazu	Not Specified	29
52	Fractal Dimension	Idrisi Selva 17.0	Not Specified	
53	General Curvature	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Evans (1979)	
54	Gradient	uDig 1.4.0b with Spatial Toolbox	Evans (1979)	
55	Gradient	uDig 1.4.0b with Spatial Toolbox	Finite Differences	
56	Gradient	uDig 1.4.0b with Spatial Toolbox	Horn (1981)	
57	Longitudinal Curvature	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Evans (1979)	
58	Longitudinal Curvature	Landserf 2.3	Not Specified	
59	Maximum	ArcGis 10.2.2 and Python 2.7.8	Not Specified	60*
60	Maximum	Diva-GIS 7.5.0	Not Specified	59*
61	Maximum Curvature	Idrisi Selva 17.0	Not Specified	
62	Maximum Curvature	Landserf 2.3	Not Specified	
63	Maximum Value	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	64, 65, 66
64	Maximum Value	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	63, 65, 66
65	Maximum Value	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	63, 64, 66
66	Maximum Value	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	63, 64, 65
67	Mean	ArcGis 10.2.2 and Python 2.7.8	Not Specified	68*
68	Mean	Diva-GIS 7.5.0	Not Specified	67*
69	Mean Curvature	Landserf 2.3	Not Specified	
70	Mean of Residuals	Landserf 2.3	Not Specified	
71	Mean Value	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	74
72	Mean Value	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	
73	Mean Value	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	
74	Mean Value	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	71
75	Median	Diva-GIS 7.5.0	Not Specified	

ID	Attributes Names' in Software	Software	Algorithms/Methods/References	Identical to…
76	Minimum	ArcGis 10.2.2 and Python 2.7.8	Not Specified	77
77	Minimum	Diva-GIS 7.5.0	Not Specified	76
78	Minimum Curvature	Idrisi Selva 17.0	Not Specified	
79	Minimum Curvature	Landserf 2.3	Not Specified	
80	Minimum Value	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	81, 82, 83
81	Minimum Value	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	80, 82, 83
82	Minimum Value	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	80, 81, 83
83	Minimum Value	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	80, 81, 82
84	Modified/local Melton Ruggedness Index	ArcGis 10.2.2 and Python 2.7.8	Melton (1965)	
85	Morphometric Protection Index	SAGA GIS 2.0.8	Not Specified	
86	Northerness	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Not Specified	
87	Northerness	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	4-Cell Method (Fleming and Hoffer, 1979)	
88	Northerness	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Horn (1981)	110
89	Northerness	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Sharpnack and Akin (1969)	
90	Northerness	ArcGis 10.2.2 and Python 2.7.8	Zevenbergen and Thorne (1987)	
91	Northerness	Idrisi Selva 17.0	Not Specified	
92	Northerness	Landserf 2.3	Not Specified	
93	Northerness	Quantum GIS 2.4.0 Chugiak	Horn (1981)	
94	Northerness	Quantum GIS 2.4.0 Chugiak	Zevenbergen and Thorne (1987)	
95	Northerness	Quantum GIS 2.4.0 Chugiak	Not Specified	
96	Northerness	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Bauer et al., 1985)	98
97	Northerness	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Haralick, 1983)	
98	Northerness	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Heerdegen and Beran, 1982)	96
99	Northerness	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Zevenbergen and Thorne, 1987)	

ID	Attributes Names' in Software	Software	Algorithms/Methods/References	Identical to…
100	Northerness	SAGA GIS 2.0.8	Least Squares Fitted Plane (Horn, 1981, Costa-Cabral and Burgess, 1996)	
101	Northerness	SAGA GIS 2.0.8	Maximum Slope (Travis et al., 1975)	
102	Northerness	SAGA GIS 2.0.8	Maximum Triangle Slope (Tarboton, 1997)	
103	Northerness	SAGA GIS 2.0.8	Not Specified	
104	Northerness	TNTmips Free 2014 (MicroImages)	Exact fit to 4 nearest neighbors and center cell	
105	Northerness	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit	106
106	Northerness	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, match central cell	105
107	Northerness	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by 1/distance	
108	Northerness	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by $1/distance^2$	
109	Northerness	uDig 1.4.0b with Spatial Toolbox	Not Specified	
110	Northerness	Whitebox GAT 3.2.1 Iguazu	Not Specified	88
111	Percentile	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	112, 113, 114
112	Percentile	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	111, 113, 114
113	Percentile	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	111, 112, 114
114	Percentile	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	111, 112, 113
115	Plan Curvature	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Not Specified	117
116	Plan Curvature	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Evans (1979)	
117	Plan Curvature	ArcGis 10.2.2 and Python 2.7.8	Zevenbergen and Thorne (1987)	115
118	Plan Curvature	Landserf 2.3	Not Specified	
119	Plan Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Bauer et al., 1985)	
120	Plan Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Haralick, 1983)	
121	Plan Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Heerdegen and Beran, 1982)	
122	Plan Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Zevenbergen and Thorne, 1987)	
123	Plan Curvature	SAGA GIS 2.0.8	Least Squares Fitted Plane (Horn, 1981, Costa-Cabral and Burgess, 1996)	14, 16, 125, 143, 145
124	Plan Curvature	SAGA GIS 2.0.8	Maximum Slope (Travis et al., 1975)	

ID	Attributes Names' in Software	Software	Algorithms/Methods/References	Identical to…
125	Plan Curvature	SAGA GIS 2.0.8	Maximum Triangle Slope (Tarboton, 1997)	14, 16, 123, 143, 145
126	Plan Curvature	SAGA GIS 2.0.8	Not Specified	
127	Plan Curvature	TNTmips Free 2014 (MicroImages)	Exact fit to 4 nearest neighbors and center cell	
128	Plan Curvature	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit	
129	Plan Curvature	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, match central cell	
130	Plan Curvature	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by 1/distance	
131	Plan Curvature	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by $1/\text{distance}^2$	
132	Plan Curvature	Whitebox GAT 3.2.1 Iguazu	Not Specified	
133	Planar Curvature	uDig 1.4.0b with Spatial Toolbox	Not Specified	
134	Planimetric-to-surface ratio	ArcGis 10.2.2 and Python 2.7.8	Cooley (2014), Rashid (2010), Berry (2007)	
135	Profile Curvature	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Not Specified	137
136	Profile Curvature	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Evans (1979)	
137	Profile Curvature	ArcGis 10.2.2 and Python 2.7.8	Zevenbergen and Thorne (1987)	135
138	Profile Curvature	Landserf 2.3	Not Specified	
139	Profile Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Bauer et al., 1985)	
140	Profile Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Haralick, 1983)	
141	Profile Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Heerdegen and Beran, 1982)	
142	Profile Curvature	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Zevenbergen and Thorne, 1987)	
143	Profile Curvature	SAGA GIS 2.0.8	Least Squares Fitted Plane (Horn, 1981, Costa-Cabral and Burgess, 1996)	14, 16, 123, 125, 145
144	Profile Curvature	SAGA GIS 2.0.8	Maximum Slope (Travis et al., 1975)	
145	Profile Curvature	SAGA GIS 2.0.8	Maximum Triangle Slope (Tarboton, 1997)	14, 16, 123, 125, 143
146	Profile Curvature	SAGA GIS 2.0.8	Not Specified	
147	Profile Curvature	TNTmips Free 2014 (MicroImages)	Exact fit to 4 nearest neighbors and center cell	
148	Profile Curvature	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit	
149	Profile Curvature	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, match central cell	

ID	Attributes Names' in Software	Software	Algorithms/Methods/References	Identical to…
150	Profile Curvature	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by 1/distance	
151	Profile Curvature	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by 1/distance^2	
152	Profile Curvature	uDig 1.4.0b with Spatial Toolbox	Not Specified	
153	Profile Curvature	Whitebox GAT 3.2.1 Iguazu	Not Specified	
154	Range	ArcGis 10.2.2 and Python 2.7.8	Not Specified	155*
155	Range	Diva-GIS 7.5.0	Not Specified	154*
156	Real Area	SAGA GIS 2.0.8	Not Specified	
157	Relative deviation from mean	ArcGis 10.2.2 and Python 2.7.8	Not Specified	
158	Representativeness	SAGA GIS 2.0.8	Boehner et al. (1997)	
159	Residual at centre	Landserf 2.3	Not Specified	
160	Roughness	Quantum GIS 2.4.0 Chugiak	Not Specified	
161	Ruggedness Index	Quantum GIS 2.4.0 Chugiak	Not Specified	
162	Slope	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Not Specified	164*, 166, 187*
163	Slope	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	4-Cell Method (Fleming and Hoffer, 1979)	
164	Slope	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Horn (1981)	162*, 166*, 187
165	Slope	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Sharpnack and Akin (1969)	
166	Slope	ArcGis 10.2.2 and Python 2.7.8	Horn (1981)	162, 164*, 187*
167	Slope	Diva-GIS 7.5.0	Not Specified	
168	Slope	Idrisi Selva 17.0	Not Specified	
169	Slope	Landserf 2.3	Not Specified	
170	Slope	Quantum GIS 2.4.0 Chugiak	Horn (1981)	
171	Slope	Quantum GIS 2.4.0 Chugiak	Zevenbergen and Thorne (1987)	
172	Slope	Quantum GIS 2.4.0 Chugiak	Not Specified	
173	Slope	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Bauer et al., 1985)	175
174	Slope	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Haralick, 1983)	

ID	Attributes Names' in Software	Software	Algorithms/Methods/References	Identical to…
175	Slope	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Heerdegen and Beran, 1982)	173
176	Slope	SAGA GIS 2.0.8	Fit 2 Degree Polynom (Zevenbergen and Thorne, 1987)	
177	Slope	SAGA GIS 2.0.8	Least Squares Fitted Plane (Horn, 1981, Costa-Cabral and Burgess, 1996)	
178	Slope	SAGA GIS 2.0.8	Maximum Slope (Travis et al., 1975)	
179	Slope	SAGA GIS 2.0.8	Maximum Triangle Slope (Tarboton, 1997)	
180	Slope	SAGA GIS 2.0.8	Not Specified	
181	Slope	TNTmips Free 2014 (MicroImages)	Exact fit to 4 nearest neighbors and center cell	
182	Slope	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit	183
183	Slope	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, match central cell	182
184	Slope	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by 1/distance	
185	Slope	TNTmips Free 2014 (MicroImages)	Quadratic surface, least-squares fit, weighted by 1/distance^2	
186	Slope	uDig 1.4.0b with Spatial Toolbox	Not Specified	
187	Slope	Whitebox GAT 3.2.1 Iguazu	Not Specified	162*, 164, 166*
188	Slope Variability	ArcGis 10.2.2 and Python 2.7.8	Ruszkiczay-Rudiger et al. (2009)	
189	Standard Deviation	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Not Specified	190
190	Standard Deviation	ArcGis 10.2.2 and Python 2.7.8	Ascione et al. (2008)	189
191	Standard Deviation	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	194
192	Standard Deviation	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	
193	Standard Deviation	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	
194	Standard Deviation	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	191
195	Standard Deviation of Slope	ArcGis 10.2.2 and Python 2.7.8	Naruse and Oguchi (2013)	
196	Surface Area	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Not Specified	198
197	Surface Area	ArcGis 10.2.2 and Python 2.7.8	Cooley (2014), Rashid (2010), Berry (2007)	200
198	Surface Area to Planar Surface	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Not Specified	196
199	Surface Ratio	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Jenness (2013)	

ID	Attributes Names' in Software	Software	Algorithms/Methods/References	Identical to…
200	Surface Ratio	ArcGis 10.2.2 and Python 2.7.8	Cooley (2014), Rashid (2010), Berry (2007)	197
201	Surface Roughness Index	ArcGis 10.2.2 and Python 2.7.8	Hobson (1972)	
202	Tangential Curvature	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Evans (1979)	
203	Tangential Curvature	uDig 1.4.0b with Spatial Toolbox	Not Specified	
204	Tangential Curvature	Whitebox GAT 3.2.1 Iguazu	Not Specified	
205	Terrain Ruggedness (VRM)	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Wright et al. (2012)	
206	Terrain Ruggedness Index	ArcGis 10.2.2 and Python 2.7.8	Riley et al. (1999)	
207	Terrain Ruggedness Index	Quantum GIS 2.4.0 Chugiak	Not Specified	
208	Terrain Ruggedness Index	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	211
209	Terrain Ruggedness Index	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	
210	Terrain Ruggedness Index	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	
211	Terrain Ruggedness Index	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	208
212	Topographic Position Index	ArcGis 10.2.2 and Python 2.7.8	Cooley (2014)	
213	Topographic Position Index	Quantum GIS 2.4.0 Chugiak	Not Specified	
214	Topographic Position Index	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	23, 26, 217
215	Topographic Position Index	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	
216	Topographic Position Index	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	25
217	Topographic Position Index	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	23, 26, 214
218	Topographic Ruggedness Index	Whitebox GAT 3.2.1 Iguazu	Not Specified	
219	Total Curvature	ArcGis 10.2.2 and DEM Surface Tools for ArcGIS 10 (v.2.1.399)	Evans (1979)	
220	Total Curvature	Whitebox GAT 3.2.1 Iguazu	Not Specified	
221	Value Range	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	222, 223, 224
222	Value Range	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	221, 223, 224
223	Value Range	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	221, 222, 224
224	Value Range	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	221, 222, 223
225	Variance	ArcGIS 10.2.2 and Benthic Terrain Modeler 3.0 Release Candidate 3	Wright et al. (2012)	
226	Vector Ruggedness Measure	ArcGis 10.2.2 and Python 2.7.8	Sappington et al. (2007); Hobson (1972)	
227	Vector Ruggedness Measure	SAGA GIS 2.0.8	Exponential (Wilson and Gallant, 2000)	230
228	Vector Ruggedness Measure	SAGA GIS 2.0.8	Gaussian weighting (Wilson and Gallant, 2000)	
229	Vector Ruggedness Measure	SAGA GIS 2.0.8	Inverse distance to a power (Wilson and Gallant, 2000)	
230	Vector Ruggedness Measure	SAGA GIS 2.0.8	No distance weighting (Wilson and Gallant, 2000)	227

***Except for surface A8**

Table B.2: Algorithms used to measure slope by each studied software packages. The definition of the groups can be found in Figure 3.3.

	Group 5A	Group 5B	Group 5C	Group 5D	Group 5E
ArcGIS 10.2.2 with Python 2.7.8	√				
ArcGIS 10.2.2 with DEM Surface Tools (v.2.1.399)	√	√			√
ArcGIS 10.2.2 with Benthic Terrain Modeler 3.0 rc3	√				
Diva-GIS 7.5.0	√				
Idrisi Selva 17.0		√			
Landserf 2.3					√
Quantum GIS 2.4.0 Chugiak	√	√			
SAGA GIS 2.0.8	√	√	√		√
TNTmips Free 2014 (MicroImages)				√	
uDig 1.4.0b	√	√			√
Whitebox GAT 3.2.1 Iguazu	√				

References of algorithms/methods

Ascione, A., Cinque, A., Miccadei, E., Villani, F., & Berti, C. (2008) The Plio-Quaternary uplift of the Apennine chain: new data from the analysis of topography and river valleys in Central Italy. *Geomorphology*, 102:105-118.

Bauer, J., Rohdenburg, H., & Bork, H.R. (1985) Ein digitales reliefmodell als Vorraussetzung fuer ein deterministisches Modell der Wasser – und Stoff-Fluesse. In: Bork, H.R. & Rohdenburg, H. (Eds) *Landschaftsgenese und Landschaftsoekologie, Parameteraufbereitung fuer deterministische Gebiets-Wassermodelle*, *Grundlagenarbeiten zu Analyse von Agrar-Oekosystemen*, 1-15.

Berry, J.K. (2007) Beyond Mapping III – Map analysis: understanding spatial patterns and relationships. *GeoTec Media Publisher*.

Boehner, J., Koethe, R., & Trachinow, C. (1997) Weiterentwicklung der automatischen Reliefanalyse auf der Basis von Digitalen Gelandemodellen. *Gottinger Geographische Abhandlungen*, 100:3-21.

Cooley, S., 2014. *GIS 4 Geomorphology: Terrain Roughness – 13 ways*. http://gis4geomorphology.com/roughness-topographic-position/

Costa-Cabral, M., & Burgess, S.J. (1996) Digital Elevation Model Networks (DEMON): a model of flow over hillslopes for computation of contributing and dispersal areas. *Water Resources Research*, 30:1681-1692.

Evans, I.S. (1979) An integrated system of terrain analysis and slope mapping. *Final report on grant DA-ERO-591-73-G0040*, University of Durham, England.

Fleming, M.D., & Hoffer, R.M. (1979) Machine processing of Landsat MSS data and DMA topographic data for forest cover type mapping. *Laboratory for Applications of Remote Sensing, Purdue University*, *West Lafayette, Indiana*, LARS Technical Report 062879.

Haralick, R.M. (1983) Ridge and valley detection on digital images. *Computer Vision, Graphics and Image Processing*, 22:28-38.

Heerdegen, R.G., & Beran, M.A. (1982) Quantifying source areas through land surface curvature. *Journal of Hydrology*, 57:359-373.

Hobson, R.D. (1972) Surface roughness in topography: quantitative approach. In: Chorley, R.J. (Ed.) *Spatial Analysis in Geomorphology*, Harper and Row, New York, United States.

Horn, B.K.P. (1981) Hill shading and the reflectance map. *Proceedings of the IEEE*, 69:14-47.

Jenness. J.S. (2013) DEM Surface Tools. *Jenness Enterprises*, Flagstaff, Arizona, United States, 98p.

Melton, M.A. (1965) The geomorphic and paleoclimatic significance of alluvial deposits in southern Arizona. *Journal of Geology*, 73:1-38.

Naruse, K., & Oguchi, T. (2013) Classification and formation environment of glacial valleys based on morphometric analyses. *Conference Proceedings of Geomorphometry 2013*, Nanjing, China, p.41-44.

Rashid, H. (2010) 3-D surface-area computation of the state of Jammu & Kashmir using Shuttle Radar Topographic Mission (SRTM) data in Geographical Information System (GIS). *Journal of Geomatics*, 4:77-82.

Riley, S.J., DeGloria, S.D., & Elliot, R. (1999) A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences*, 5:23-27.

Ruszkiczay-Rüdiger, Z., Fodor, L., & Horváth, E. (2009) Discrimination of fluvial, eolian and neotectonic features in a low hilly landscape: a DEM-based morphotectonic analysis in the Central Pannonian Basin, Hungary. *Geomorphology*, 104:203-217.

Sappington, J.M., Longshore, K.M., & Thompson, D.B. (2007) Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave desert. *Journal of Wildlife management*, 71:1419-1426.

Sharpnack, D.A., & Akin, G. (1969) An algorithm for computing slope and aspect from elevations. *Photogrammetric Engineering*, 35:247-248.

Tarboton, D.G., (1997) A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33:309-319.

Travis, M.R., Elsner, G.H., Iverson, W.D., & Johnson, C.G. (1975) VIEWIT: computation of seen areas, slope, and aspect for land-use planning. *USDA F.S. General Technical Report* PSW-11/1975, 70p.

Wilson, J.P., & Gallant, J.C. (2000) *Terrain analysis: principles and applications.* Wiley 520p.

Wright, D.J., Pendleton, M., Boulware, J., Walbridge, S., Gerlt, B., Eslinger, D., Sampson, D., & Huntley, E. (2012) ArcGIS Benthic Terrain Modeler (BTM), v. 3.0, *Environmental Systems Research Institute, NOAA Coastal Services Center, Massachusetts Office of Coastal Zone Management*. Available online at http://esriurl.com/5754.

Zevenbergen, L.W., & Thorne, C.R. (1987) Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms*, 12:47-56.

Appendix C: Extended Results of the Iterative Principal Component Analysis (PCA), Variable Inflation Factor (VIF), and Mutual Information (MI), for all Surfaces and Derived Terrain Attributes (Chapter 3)

Table C.1: Summary of PCA solutions. Legend: Car = Removed for low cardinality. Mx = Marker variable on component x. Coy = Complex variable removed at iteration y. NL = No loadings. CFa.b.c = Complex variable that reached the final solution as it does not load relatively equal on more than one component, but does load on more than one. a is the component on which it loads the most, followed by b and c, when applicable. Negative loadings are indicated with the minus sign. The categories are based on Comrey & Lee (1992) scale of measurement for loadings cutoff (Comrey, AL., and H.B. Lee, 1992, A first course in factor analysis, Hillsdale, New Jersey: Erlbaum). Regular characters: excellent loadings (>0.71). Characters in italic: very good loadings ([0.63-0.71[). Underlined characters: good loadings ([0.55-0.63[). Underlined characters in italic: fair loadings ([0.45-0.55[). Bold characters: poor loadings ([0.32-0.45[).

ID	TERRAIN ATTRIBUTES	POSITION AND STRENGTH IN FINAL PCA SOLUTIONS								
		A0	A1	A2	A3	A4	A5	A6	A7	A8
1	Bathymetric Position Index	Car	Car	Car	Car	Car	Car	Car	Car	Car
2	Center versus Neighbors Variability	Car	Car	Car	Car	Car	Car	Car	Car	Car
3	Coefficient of variation	Co1	Co1	CF2.6	CF2.6	CF4.6	CF4.6	CF4.6	CF4.5	CF5.4
4	Convergence Index	M1	Co5	CF1.12	CF1.12	CF1.12	Co2	M1	Co1	Co1
5	Convergence Index	M1	M1	M1	M1	M1	M1	M1	M1	M1
6	Cross-sectional Curvature	M1	CF1.8	CF1.8	CF1.8	Co1	Co2	Co1	Co1	Co1
7	Cross-sectional Curvature	M1	CF1.8	Co1	Co1	Co1	Co2	Co1	Co1	Co1
8	Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
9	Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
10	Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
11	Curvature	Co1	Co1	Co1	Co1	Co1	Co3	Co1	CF13.1	CF11.1
12	Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
13	Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
14	Curvature	Car	Car	Car	Car	Car	Car	Car	Car	Car
15	Curvature	M1	M1	M1	M1	M1	M1	M1	CF1.12	M1
16	Curvature	Car	Car	Car	Car	Car	Car	Car	Car	Car
17	Deviation from Mean Elevation	CF1.9	CF1.9	M1	M1	M1	M1	M1	M1	M1
18	Deviation from Mean Value	M1	M1	CF1.12	M1	M1	M1	M1	M1	M1
19	Deviation from Mean Value	M1	M1	M1	M1	M1	M1	M1	M1	M1
20	Deviation from Mean Value	M1	M1	M1	M1	M1	M1	M1	M1	M1
21	Deviation from Mean Value	M1	M1	CF1.12	M1	M1	M1	M1	M1	M1
22	Difference from Mean Elevation	*CF3.1*	Co1	CF1.3.1 2	CF1.12	CF1.12	M1	M1	CF1.13	CF1.11
23	Difference from Mean Value	M1	M1	M1	M1	M1	M1	M1	M1	CF1.11
24	Difference from Mean Value	M1	M1	M1	M1	M1	M1	M1	M1	M1
25	Difference from Mean Value	M1	M1	M1	M1	M1	M1	M1	M1	M1
26	Difference from Mean Value	M1	M1	M1	M1	M1	M1	M1	M1	CF1.11
27	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
28	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
29	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
30	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
31	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2

32	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
33	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
34	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
35	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
36	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
37	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
38	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
39	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
40	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
41	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
42	Easterness	Car	Car	Car	Car	Car	Car	Car	Car	Car
43	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
44	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
45	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
46	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
47	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
48	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
49	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
50	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
51	Easterness	M3	M3	M3	M3	M2	M2	M2	M2	M2
52	Fractal Dimension	CF2.5	CF2.5	Co1	Co1	CF7.5.4	CF7.4.5	CF7.5.4	M7	M7
53	General Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
54	Gradient	M2	M2	CF5.2	CF5.2	CF5.4	M5	M5	Co2	Co3
55	Gradient	M2	M2	CF5.2	CF9.5.2	CF7.5	M7	M7	M7	M7
56	Gradient	M2	M2	CF5.2	CF5.2	M5	M5	M5	M6	M6
57	Longitudinal Curvature	Co1	Co1	Re	Co2	Co2	Co2	Re	Co1	Co1
58	Longitudinal Curvature	Co1	Co1	Co2	Co2	Co2	Co2	Co1	Co1	Co1
59	Maximum	M6	M6	M6	M6	M6	M6	M6	M5	M4
60	Maximum	M6	M6	M6	M6	M6	M6	M6	M5	M4
61	Maximum Curvature	-M1	-M1	-M1	-M1	-M1	-M1	-M1	-M1	-M1
62	Maximum Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
63	Maximum Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
64	Maximum Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
65	Maximum Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
66	Maximum Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
67	Mean	M6	M6	M6	M6	M6	M6	M6	M5	M4
68	Mean	M6	M6	M6	M6	M6	M6	M6	M5	M4
69	Mean Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
70	Mean of Residuals	Co2-NL	Co2-NL	**M13**	**M14**	*M15*	Co5-NL	Co6-NL	Co1	Co1
71	Mean Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
72	Mean Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
73	Mean Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
74	Mean Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
75	Median	M6	M6	M6	M6	M6	M6	M6	M5	M4
76	Minimum	M6	M6	M6	M6	M6	M6	M6	M5	M4
77	Minimum	M6	M6	M6	M6	M6	M6	M6	M5	M4
78	Minimum Curvature	-M1	-M1	-M1	-M1	-M1	-M1	-M1	-M1	-M1
79	Minimum Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
80	Minimum Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
81	Minimum Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
82	Minimum Value	M6	M6	M6	M6	M6	M6	M6	M5	M4
83	Minimum Value	M6	M6	M6	M6	M6	M6	M6	M5	M4

84	Modified/local Melton Ruggedness Index	CF5.2	CF5.2	M2	M2	M4	M4	M4	M4	M5
85	Morphometric Protection Index	-M1	-M1	-M1	-M1	-M1	-M1	-M1	-M1	-M1
86	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
87	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
88	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
89	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
90	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
91	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
92	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
93	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
94	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
95	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
96	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
97	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
98	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
99	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
100	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
101	Northerness	Car	Car	Car	Car	Car	Car	Car	Car	Car
102	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
103	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
104	Northerness	M4	M4	M4	M4	M3	M3	M3	*M3*	Co1
105	Northerness	M4	M4	M4	M4	M3	Co4	Co2	Co1	Co1
106	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
107	Northerness	M4	M4	M4	M4	M3	M3	M3	Co1	Co1
108	Northerness	M4	M4	M4	M4	M3	M3	M3	Co1	Co1
109	Northerness	M4	M4	M4	M4	M3	*M3*	*M3*	<u>M3</u>	Co1
110	Northerness	M4	M4	M4	M4	M3	M3	M3	M3	M3
111	Percentile	Car	Car	Car	Car	Car	Car	Car	Car	Car
112	Percentile	Car	Car	Car	Car	Car	Car	Car	Car	Car
113	Percentile	Car	Car	Car	Car	Car	Car	Car	Car	Car
114	Percentile	Car	Car	Car	Car	Car	Car	Car	Car	Car
115	Plan Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
116	Plan Curvature	Co4	Co1	Co1	Co2	**M1**	Co4-NL	Co9	CF15.1	CF14.1
117	Plan Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
118	Plan Curvature	Co3	Co3	Co1	Co1	Co1		Co1	Co1	Co1
119	Plan Curvature	M1	M1	CF1.8	CF1.8	M1	Co2	Co1	Co1	Co1
120	Plan Curvature	Co1	Co1	Co1	Co1	Co1	Co3	Co1	*CF13.1*	<u>CF11.1</u>
121	Plan Curvature	M1	CF1.8	CF1.8	CF1.8	CF1.9	M1	CF1.9	M1	CF1.15
122	Plan Curvature	M1	M1	M1	M1	M1	M1	M1	M1	M1
123	Plan Curvature	Car	Car	Car	Car	Car	Car	Car	Car	Car
124	Plan Curvature	M1	M1	M1	M1	CF1.13	M1	CF1.11	CF1.12	M1
125	Plan Curvature	Car	Car	Car	Car	Car	Car	Car	Car	Car
126	Plan Curvature	Co2	Co6-NL	<u>M12</u>	<u>*M13*</u>	<u>*M12*</u>	Co3	Co5	Co3	Co5
127	Plan Curvature	CF8.1	M8	M8	M8	M9	Co1	M9	CF9.8	CF9.8
128	Plan Curvature	CF8.1	M8	M8	M8	M9	Co1	M9	M9	M9
129	Plan Curvature	CF8.1	M8	M8	M8	M9	Co1	M9	M9	M9
130	Plan Curvature	CF8.1	M8	M8	M8	M9	Co1	M9	M9	M9
131	Plan Curvature	CF8.1	M8	M8	M8	M9	Co1	M9	M9	M9
132	Plan Curvature	Co1	CF9.1	CF9.1	Co3	CF10.1	Co1	Co1	Co4	Co2
133	Planar Curvature	Co1	CF9.1	CF9.1	Co3	CF10.1	Co1	Co1	Co4	Co2
134	Planimetric-to-surface ratio	Co2-NL	Co2-NL	Co2-NL	Co1-NL	Co2-NL	Co3-NL	Co1-NL	Co1-NL	Co2-NL

135	Profile Curvature	Co2	-CF1.7	-CF1.7	-M1	-M1	-M1	-M1	-M1	-M1
136	Profile Curvature	Co1	Co1	Co1	Co1	Co1	Co2	Co1	Co1	Co3
137	Profile Curvature	Co2	-CF1.7	-CF1.7	-M1	Re	Re	Re	Re	Re
138	Profile Curvature	-M9	Co1	Co1	Co2	Co2-NL	Co1	Co3	Co3	Co4
139	Profile Curvature	Co1	CF1.7	CF1.7	CF1.7	CF1.8	Co2	CF1.8	Co1	Co1
140	Profile Curvature	Co1	Co1	Co1	Co1	Co1	Co1	Co1	Co2	*CF11.1*
141	Profile Curvature	Co1	Co6	CF1.7	CF1.7	CF1.8	M1	M1	M1	M1
142	Profile Curvature	Co2	CF1.7	CF1.7	M1	M1	M1	M1	M1	M1
143	Profile Curvature	Car	Car	Car	Car	Car	Car	Car	Car	Car
144	Profile Curvature	Co1	Co1	Co1	Co1	Co1	Co1	M1	M1	*M1*
145	Profile Curvature	Car	Car	Car	Car	Car	Car	Car	Car	Car
146	Profile Curvature	Co2	Co4	M12	*M13*	*M12*	Co1	Co4	Co1	**M4**
147	Profile Curvature	M7	M7	M7	M7	M8	M8	M8	CF8.9	M8
148	Profile Curvature	M7	M7	M7	M7	M8	M8	M8	M8	M8
149	Profile Curvature	M7	M7	M7	M7	M8	M8	M8	M8	M8
150	Profile Curvature	M7	M7	M7	M7	M8	M8	M8	M8	M8
151	Profile Curvature	M7	M7	M7	M7	M8	M8	M8	M8	M8
152	Profile Curvature	M9	M9	M9	M11	M10	M9	M12	Re	M12
153	Profile Curvature	M9	M9	M9	M11	M10	M9	M12	M16	M12
154	Range	CF5.2	CF5.2	M2	M2	M4	M4	M4	M4	M5
155	Range	CF5.2	CF5.2	M2	M2	M4	M4	M4	M4	M5
156	Real Area	CF5.2	CF5.2	CF2.5	M2	CF4.7	CF4.7	*CF4.7*	Co6	Co1
157	Relative deviation from mean	-CF1.9	-M1	-M1	-M1	-M1	-M1	-M1	-M1	-M1
158	Representativeness	Co3-NL	Co7-NL	-CF13.5	-M14	-CF15	Co1	Co7	Co5	-M5
159	Residual at centre	*M1*	Co3	Co1	Co1	Co1	Co1	Co1	Co2	Co1
160	Roughness	CF5.2	CF5.2	M2	M2	M4	M4	M4	M4	M5
161	Ruggedness Index	CF5.2	CF5.2	M2	M2	M4	M4	M4	CF4.12	Co1
162	Slope	M2	M2	CF5.2	CF5.2	M5	M5	M5	M6	M6
163	Slope	M2	M2	CF5.2	CF9.5.2	CF7.5	M7	M7	M7	M7
164	Slope	M2	M2	CF5.2	CF5.2	M5	M5	M5	M6	M6
165	Slope	M2	M2	CF5.2	CF5.2	CF5.4	M5	M5	Co2	Co3
166	Slope	M2	M2	CF5.2	CF5.2	M5	M5	M5	M6	M6
167	Slope	M2	M2	CF5.2	CF5.2	M5	M5	M5	M6	M6
168	Slope	M2	M2	CF5.2	CF9.5.2	CF7.5	M7	M7	M7	M7
169	Slope	M2	M2	CF5.2	CF5.2	CF5.4	M5	M5	Co2	Co3
170	Slope	M2	M2	CF5.2	CF5.2	M5	M5	M5	M6	M6
171	Slope	M2	M2	CF5.2	CF9.5.2	CF7.5	M7	M7	M7	M7
172	Slope	M2	M2	CF5.2	CF5.2	M5	M5	M5	M6	M6
173	Slope	M2	M2	CF5.2	CF5.2	CF5.4	M5	M5	Co2	Co3
174	Slope	Co2	Co1	CF2.5	Co2	Co1	Co2	Co1	Co1	Co1
175	Slope	M2	M2	CF5.2	CF5.2	CF5.4	M5	M5	Co2	Co3
176	Slope	M2	M2	CF5.2	CF9.5.2	CF7.5	M7	M7	M7	M7
177	Slope	M2	M2	CF5.2	CF5.2	M5	M5	M5	M6	M6
178	Slope	CF2.9	Co1	Co1	Co4	CF13.1	Co1	CF11.1	CF11.1	Co1
179	Slope	CF2.9	Co1	Co1	Co4	CF13.1	Co1	CF11.1	CF11.1	Co1
180	Slope	M2	Co2	Co2	M6	**M7**	Co4	*M6*	*M5*	*M4*
181	Slope	M2	M2	Co1	Co1	Co1	Co1	Co1	Co1	Co2
182	Slope	M2	M2	Co1	Co1	CF14.5.4	Co1	M5	M12	M13
183	Slope	M2	M2	Co1	Co1	CF14.5.4	Co1	M5	M14	M14
184	Slope	M2	M2	Co1	CF5.2	CF14.5.	Co1	M5	CF12.6	CF13.6

						4				
185	Slope	M2	M2	Co1	Co1	CF14.5.4	Co1	M5	CF12.6	CF13.6
186	Slope	CF5.2	*CF5.6.2*	Co1	Co2	Co1	*CF1.4*	*M1*	*CF1.5*	CF1.4
187	Slope	M2	M2	CF5.2	CF5.2	M5	M5	M5	M6	M6
188	Slope Variability	*M11*	M11	*M11*	Co1	Co1	Co1	Co1	Co1	Co1
189	Standard Deviation	CF5.2	CF5.2	M2	M2	M4	M4	M4	M4	M5
190	Standard Deviation	CF5.2	CF5.2	M2	M2	M4	M4	M4	M4	M5
191	Standard Deviation	CF5.2	CF5.2	M2	CF2.14	CF4.15	Co2	Co1	Co1	Co1
192	Standard Deviation	CF5.2	CF5.2	M2	M2	M4	M4	CF4.10	M4	M5
193	Standard Deviation	CF5.2	CF5.2	M2	M2	CF4.15	Co2	*CF4.10*	*M4*	Co1
194	Standard Deviation	CF5.2	CF5.2	M2	CF2.14	CF4.15	Co2	Co1	Co1	Co1
195	Standard Deviation of Slope	-M2	*-CF2.5*	Co1	Co1	Co1	**-M5**	Co1	**-M6**	**-M6**
196	Surface Area	CF5.2	CF5.2	M2	M2	M4	M4	M4	M4	Co6
197	Surface Area	CF5.2	CF5.2	CF2.5	CF2.5	CF4.5	CF4.5	CF4.5	CF4.10	CF5.6
198	Surface Area to Planar Surface	CF5.2	CF5.2	M2	M2	M4	M4	M4	M4	Co6
199	Surface Ratio	CF5.2	CF5.2	M2	M2	M4	M4	M4	M4	Co6
200	Surface Ratio	CF5.2	CF5.2	CF2.5	CF2.5	CF4.5	CF4.5	CF4.5	CF4.10	CF5.6
201	Surface Roughness Index	M2	*CF2.11*	Co1	Co1	Co1	Co1	-M10-Re	M10-Re	Re
202	Tangential Curvature	M1	Co3	Co1	Co1	Co1	Co2	Co8	CF15.1	CF14.1
203	Tangential Curvature	-M1	-CF1.9	Co1	Co2	Co1	Co1	Co1	Co1	Co1
204	Tangential Curvature	*M9*	M9	M9	M11	M10	Co1	M12	*M16*	M12
205	Terrain Ruggedness (VRM)	-M2	*-CF2.11*	Co1	Co1	Co1	Co1	M10	CF14.10	M10
206	Terrain Ruggedness Index	Co1	Co1	CF2.5.6	CF2.6	CF4.6	CF4.6	CF4.6	CF4.5	CF5.4
207	Terrain Ruggedness Index	CF5.2	CF5.2	M2	M2	M4	M4	CF4.11	CF4.12	Co1
208	Terrain Ruggedness Index	CF5.2	CF5.2	M2	CF2.14	CF4.15	Co1	Co1	Co1	Co1
209	Terrain Ruggedness Index	CF5.2	CF5.2	M2	M2	M4	M4	CF4.11	CF4.12	Co1
210	Terrain Ruggedness Index	CF5.2	CF5.2	M2	M2	M4	Co1	CF4.11	CF4.12	Co1
211	Terrain Ruggedness Index	CF5.2	CF5.2	M2	CF2.14	CF4.15	Co1	Co1	Co1	Co1
212	Topographic Position Index	M1	M1	M1	M1	M1	M1	M1	M1	M1
213	Topographic Position Index	M1	M1	M1	M1	M1	M1	M1	M1	M1
214	Topographic Position Index	M1	M1	M1	M1	M1	M1	M1	M1	CF1.11
215	Topographic Position Index	M1	M1	M1	M1	M1	M1	M1	M1	M1
216	Topographic Position Index	M1	M1	M1	M1	M1	M1	M1	M1	M1
217	Topographic Position Index	M1	M1	M1	M1	M1	M1	M1	M1	CF1.11
218	Topographic Ruggedness Index	CF5.2	CF5.2	M2	M2	M4	M4	M4	CF4.12	Co1
219	Total Curvature	M11	M11	M11	M12	Co3	Co1	Co3	Co1	Co1
220	Total Curvature	M11	M11	M11	M12	Co1	Co1	Co1	Co1	Re
221	Value Range	CF5.2	CF5.2	CF2.13	CF2.14	Co2	Co2	Co1	Co1	Co4
222	Value Range	CF5.2	CF5.2	CF2.13	CF2.14	Co2	Co2	Co1	Co1	Co4
223	Value Range	CF5.2	CF5.2	CF2.13	CF2.14	Co2	Co2	Co1	Co1	Co4
224	Value Range	CF5.2	CF5.2	CF2.13	CF2.14	Co2	Co2	Co1	Co1	Co4
225	Variance	CF5.2	CF5.2	M2	M2	M4	M4	M4-Re	M4-Re	M5-Re
226	Vector Ruggedness Measure	-M2	*-CF2.11*	Co1	Co1	Co1	Co1	M10	CF14.10	M10
227	Vector Ruggedness Measure	M10	M10	M10	M10	M11	Co1	M10	M10	M10
228	Vector Ruggedness Measure	M10	M10	M10	M10	M11	Co1	M10	M10	M10
229	Vector Ruggedness Measure	M10	M10	M10	M10	M11	Co1	M10	M10	M10
230	Vector Ruggedness Measure	M10	M10	M10	M10	M11	Co1	M10	M10	M10

Table C.2: Percentage of explained variance and total variance explained by each component for each surface. Legend: The numbers in italic represent the unreliable components as assessed by Cronbach's α, and underlined numbers correspond to components with less than three variables.

	% Explained Variance									Total Variance								
Component	**A0**	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8**	**A0**	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8**
1	18.29	18.67	19.86	19.78	19.36	21.71	20.17	20.28	20.54	28.16	28.20	29.59	29.07	29.04	26.48	28.64	27.78	27.31
2	17.09	14.64	13.86	13.18	12.67	15.28	12.86	13.16	13.30	26.31	22.11	20.65	19.37	19.01	18.65	18.27	18.03	17.69
3	13.89	13.63	13.47	13.16	12.13	13.98	11.80	11.09	10.67	21.40	20.58	20.08	19.34	18.20	17.06	16.76	15.19	14.19
4	13.20	12.92	12.76	12.82	11.78	11.68	9.31	9.53	8.26	20.33	19.50	19.01	18.85	17.68	14.24	13.22	13.05	10.99
5	10.56	11.16	8.55	7.51	6.91	7.86	7.61	6.92	6.57	16.26	16.86	12.74	11.04	10.36	9.58	10.81	9.47	8.74
6	5.92	5.96	6.40	6.75	6.30	7.56	6.62	4.70	5.68	9.11	8.99	9.53	9.92	9.45	9.23	9.41	6.44	7.56
7	3.59	3.73	3.43	3.40	3.63	4.49	4.14	4.41	4.57	5.53	5.63	5.12	5.00	5.45	5.48	5.88	6.05	6.07
8	2.99	3.24	3.29	3.38	3.22	3.72	3.28	3.37	3.50	4.61	4.89	4.90	4.97	4.83	4.54	4.66	4.62	4.66
9	*2.63*	*2.79*	*2.40*	2.18	3.12	*1.64*	3.21	3.35	3.39	*4.05*	*4.21*	*3.58*	3.20	4.68	*2.00*	4.56	4.59	4.51
10	1.84	1.97	1.93	1.96	*2.48*		2.70	2.30	2.66	2.83	2.98	2.87	2.87	*3.71*		3.83	3.15	3.53
11	*1.68*	*1.74*	*1.53*	1.79	1.97		1.86	1.93	2.22	*2.58*	*2.62*	*2.28*	2.62	2.95		2.64	2.64	2.95
12			*1.44*	*1.52*	*1.45*		*1.77*	1.78	1.94			*2.14*	*2.24*	*2.17*		*2.52*	2.44	2.58
13			*1.32*	*1.40*	1.43			1.75	1.73			*1.97*	*2.06*	2.14			2.40	2.30
14				*1.34*	1.34			1.64	1.54				*1.97*	2.01			2.24	2.05
15					*1.34*			1.40						*2.01*			1.92	
16								1.12									1.53	
Total	91.67	90.45	90.24	90.16	89.13	87.92	85.35	88.71	86.56	141.17	136.57	134.46	132.54	133.69	107.26	121.19	121.53	115.12

Table C.3 Communalities from PCA

ID	TERRAIN ATTRIBUTES	COMMUNALITIES OF EACH VARIABLE FOR EACH SURFACE								
		A0	A1	A2	A3	A4	A5	A6	A7	A8
1	Bathymetric Position Index	Low Cardinality								
2	Center versus Neighbors Variability	Low Cardinality								
3	Coefficient of variation			0.916	0.924	0.937	0.943	0.943	0.947	0.953
4	Convergence Index	0.684		0.815	0.774	0.799		0.581		
5	Convergence Index	0.905	0.881	0.903	0.890	0.912	0.900	0.902	0.911	0.911
6	Cross-sectional Curvature	0.876	0.843	0.812	0.814					
7	Cross-sectional Curvature	0.877	0.836							
8	Curvature	0.912	0.901	0.962	0.970	0.962	0.904	0.895	0.977	0.976
9	Curvature	0.912	0.901	0.962	0.970	0.962	0.904	0.895	0.977	0.976
10	Curvature	0.911	0.911	0.911	0.908	0.896	0.838	0.834	0.883	0.878
11	Curvature								0.870	0.913
12	Curvature	0.977	0.982	0.986	0.983	0.980	0.979	0.978	0.977	0.978
13	Curvature	0.912	0.901	0.962	0.970	0.962	0.904	0.895	0.977	0.976
14	Curvature	0.916	0.893	0.889	0.899	0.914	0.924	0.907	0.910	0.894
15	Curvature	0.864	0.835	0.833	0.812	0.882	0.773	0.855	0.878	0.758
16	Curvature	0.916	0.893	0.889	0.899	0.914	0.924	0.907	0.910	0.894
17	Deviation from Mean Elevation	0.898	0.879	0.901	0.893	0.902	0.842	0.845	0.908	0.902
18	Deviation from Mean Value	0.847	0.803	0.925	0.920	0.931	0.874	0.893	0.949	0.959
19	Deviation from Mean Value	0.922	0.889	0.946	0.945	0.956	0.942	0.956	0.966	0.967
20	Deviation from Mean Value	0.874	0.835	0.934	0.931	0.942	0.903	0.920	0.957	0.964
21	Deviation from Mean Value	0.847	0.803	0.925	0.920	0.931	0.874	0.893	0.949	0.959
22	Difference from Mean Elevation	0.854		0.883	0.882	0.901	0.824	0.846	0.940	0.962
23	Difference from Mean Value	0.878	0.881	0.952	0.942	0.948	0.894	0.907	0.973	0.982
24	Difference from Mean Value	0.945	0.951	0.973	0.970	0.973	0.964	0.969	0.985	0.988
25	Difference from Mean Value	0.902	0.906	0.960	0.952	0.957	0.921	0.932	0.978	0.984
26	Difference from Mean Value	0.878	0.881	0.952	0.942	0.948	0.894	0.907	0.973	0.982
27	Easterness	0.999	0.996	0.989	0.985	0.981	0.978	0.974	0.972	0.968
28	Easterness	0.996	0.985	0.955	0.934	0.917	0.895	0.880	0.870	0.843
29	Easterness	0.999	0.996	0.989	0.985	0.981	0.978	0.974	0.972	0.968
30	Easterness	0.998	0.992	0.977	0.967	0.952	0.943	0.932	0.927	0.913
31	Easterness	0.999	0.996	0.989	0.985	0.981	0.978	0.974	0.972	0.968
32	Easterness	0.995	0.984	0.954	0.935	0.917	0.896	0.880	0.870	0.843
33	Easterness	0.998	0.992	0.977	0.967	0.952	0.943	0.932	0.927	0.913
34	Easterness	0.999	0.996	0.989	0.985	0.981	0.978	0.974	0.972	0.968
35	Easterness	0.996	0.985	0.955	0.934	0.917	0.895	0.880	0.870	0.843
36	Easterness	0.999	0.996	0.989	0.985	0.981	0.978	0.974	0.972	0.968
37	Easterness	0.998	0.992	0.977	0.967	0.952	0.943	0.932	0.927	0.913
38	Easterness	0.996	0.983	0.949	0.924	0.892	0.861	0.847	0.822	0.798
39	Easterness	0.998	0.992	0.977	0.967	0.952	0.943	0.932	0.927	0.913
40	Easterness	0.996	0.985	0.955	0.934	0.917	0.895	0.880	0.870	0.843
41	Easterness	0.999	0.996	0.989	0.985	0.981	0.978	0.974	0.972	0.968
42	Easterness	Low Cardinality								
43	Easterness	0.968	0.918	0.806	0.731	0.689	0.658	0.618	0.599	0.568
44	Easterness	0.994	0.919	0.938	0.823	0.854	0.810	0.756	0.718	0.719
45	Easterness	0.986	0.949	0.868	0.816	0.763	0.734	0.688	0.666	0.630
46	Easterness	0.991	0.967	0.917	0.881	0.839	0.828	0.805	0.796	0.778
47	Easterness	0.991	0.967	0.917	0.881	0.839	0.828	0.805	0.796	0.778
48	Easterness	0.991	0.967	0.920	0.884	0.846	0.838	0.813	0.807	0.790

49	Easterness	0.991	0.968	0.920	0.885	0.846	0.837	0.813	0.806	0.790
50	Easterness	0.957	0.938	0.876	0.819	0.807	0.781	0.756	0.741	0.722
51	Easterness	0.999	0.996	0.989	0.985	0.981	0.978	0.974	0.972	0.968
52	Fractal Dimension	0.943	0.917			0.952	0.956	0.951	0.962	0.943
53	General Curvature	0.911	0.911	0.911	0.908	0.896	0.838	0.834	0.883	0.878
54	Gradient	0.971	0.938	0.943	0.950	0.925	0.889	0.818		
55	Gradient	0.937	0.874	0.866	0.995	0.989	0.989	0.985	0.996	0.994
56	Gradient	0.982	0.972	0.982	0.972	0.962	0.944	0.899	0.997	0.995
57	Longitudinal Curvature	Complex Variable								
58	Longitudinal Curvature	Complex Variable								
59	Maximum	0.998	0.999	0.999	0.996	0.998	0.996	0.990	0.982	0.979
60	Maximum									0.979
61	Maximum Curvature	0.798	0.819	0.884	0.894	0.868	0.825	0.818	0.913	0.898
62	Maximum Curvature	0.802	0.790	0.791	0.770	0.726	0.667	0.678	0.774	0.723
63	Maximum Value	0.997	0.998	0.997	0.991	0.992	0.961	0.928	0.916	0.860
64	Maximum Value	0.997	0.998	0.997	0.991	0.992	0.961	0.928	0.916	0.860
65	Maximum Value	0.997	0.998	0.997	0.991	0.992	0.961	0.928	0.916	0.860
66	Maximum Value	0.997	0.998	0.997	0.991	0.992	0.961	0.928	0.916	0.860
67	Mean	0.998	0.999	0.999	0.995	0.997	0.994	0.989	0.980	0.977
68	Mean									0.993
69	Mean Curvature	0.911	0.911	0.911	0.908	0.896	0.838	0.834	0.883	0.878
70	Mean of Residuals			0.151	0.192	0.244				
71	Mean Value	0.998	0.999	0.999	0.995	0.998	0.993	0.982	0.989	0.984
72	Mean Value	0.998	0.999	0.999	0.996	0.998	0.998	0.993	0.993	0.991
73	Mean Value	0.998	0.999	0.999	0.996	0.998	0.995	0.987	0.991	0.988
74	Mean Value	0.998	0.999	0.999	0.995	0.998	0.993	0.982	0.989	0.984
75	Median	0.998	0.999	0.999	0.995	0.997	0.995	0.989	0.980	0.968
76	Minimum	0.998	0.999	0.999	0.996	0.998	0.998	0.995	0.993	0.993
77	Minimum	0.998	0.999	0.999	0.995	0.997	0.994	0.989	0.980	0.977
78	Minimum Curvature	0.817	0.819	0.874	0.876	0.864	0.815	0.826	0.913	0.894
79	Minimum Curvature	0.806	0.801	0.796	0.783	0.733	0.675	0.692	0.776	0.721
80	Minimum Value	0.997	0.997	0.998	0.991	0.993	0.962	0.941	0.910	0.867
81	Minimum Value	0.997	0.997	0.998	0.991	0.993	0.962	0.941	0.910	0.867
82	Minimum Value	0.997	0.997	0.998	0.991	0.993	0.962	0.941	0.910	0.867
83	Minimum Value	0.997	0.997	0.998	0.991	0.993	0.962	0.941	0.910	0.867
84	Modified/local Melton Ruggedness Index	0.964	0.961	0.948	0.949	0.941	0.910	0.924	0.931	0.975
85	Morphometric Protection Index	0.823	0.802	0.859	0.855	0.874	0.859	0.869	0.879	0.878
86	Northerness	0.995	0.991	0.985	0.984	0.978	0.977	0.974	0.976	0.980
87	Northerness	0.986	0.966	0.945	0.939	0.923	0.909	0.893	0.890	0.861
88	Northerness	0.995	0.991	0.985	0.984	0.978	0.977	0.974	0.976	0.980
89	Northerness	0.992	0.982	0.967	0.963	0.948	0.936	0.928	0.922	0.917
90	Northerness	0.995	0.991	0.985	0.984	0.978	0.977	0.974	0.976	0.980
91	Northerness	0.981	0.961	0.944	0.938	0.923	0.908	0.893	0.890	0.861
92	Northerness	0.992	0.982	0.967	0.963	0.948	0.936	0.928	0.922	0.917
93	Northerness	0.995	0.991	0.985	0.984	0.978	0.977	0.974	0.976	0.980
94	Northerness	0.986	0.966	0.945	0.939	0.923	0.909	0.893	0.890	0.861
95	Northerness	0.995	0.991	0.985	0.984	0.978	0.977	0.974	0.976	0.980
96	Northerness	0.992	0.982	0.967	0.963	0.948	0.936	0.928	0.922	0.917
97	Northerness	0.981	0.960	0.927	0.925	0.883	0.863	0.839	0.827	0.790
98	Northerness	0.992	0.982	0.967	0.963	0.948	0.936	0.928	0.922	0.917
99	Northerness	0.986	0.966	0.945	0.939	0.923	0.909	0.893	0.890	0.861
100	Northerness	0.995	0.991	0.985	0.984	0.978	0.977	0.974	0.976	0.980

101	Northerness	Low Cardinality								
102	Northerness	0.863	0.783	0.733	0.739	0.671	0.654	0.636	0.603	0.571
103	Northerness	0.967	0.808	0.917	0.865	0.865	0.815	0.769	0.735	0.712
104	Northerness	0.930	0.862	0.788	0.793	0.735	0.660	0.634	0.552	
105	Northerness	0.941	0.886	0.826	0.830	0.780				
106	Northerness	0.941	0.886	0.826	0.830	0.780				
107	Northerness	0.942	0.890	0.834	0.838	0.795	0.705	0.690		
108	Northerness	0.942	0.889	0.831	0.836	0.792	0.696	0.680		
109	Northerness	0.739	0.690	0.662	0.572	0.552	0.491	0.405	0.364	
110	Northerness	0.995	0.991	0.985	0.984	0.978	0.977	0.974	0.976	0.980
111	Percentile	Low Cardinality								
112	Percentile	Low Cardinality								
113	Percentile	Low Cardinality								
114	Percentile	Low Cardinality								
115	Plan Curvature	0.893	0.845	0.882	0.864	0.874	0.782	0.794	0.860	0.861
116	Plan Curvature					0.194			0.896	0.793
117	Plan Curvature	0.893	0.845	0.882	0.864	0.874	0.782	0.794	0.860	0.861
118	Plan Curvature	Complex Variable								
119	Plan Curvature	0.844	0.813	0.787	0.791	0.735				
120	Plan Curvature								0.677	0.592
121	Plan Curvature	0.929	0.890	0.870	0.857	0.824	0.727	0.784	0.842	0.846
122	Plan Curvature	0.893	0.845	0.882	0.864	0.874	0.782	0.794	0.860	0.861
123	Plan Curvature	0.916	0.893	0.889	0.899	0.914	0.924	0.907	0.910	0.894
124	Plan Curvature	0.758	0.711	0.701	0.697	0.809	0.652	0.778	0.833	0.614
125	Plan Curvature	0.916	0.893	0.889	0.899	0.914	0.924	0.907	0.910	0.894
126	Plan Curvature			0.412	0.318	0.297				
127	Plan Curvature	0.848	0.831	0.790	0.832	0.805		0.749	0.754	0.728
128	Plan Curvature	0.938	0.900	0.894	0.899	0.912		0.908	0.910	0.902
129	Plan Curvature	0.962	0.936	0.930	0.946	0.964		0.967	0.969	0.963
130	Plan Curvature	0.959	0.941	0.924	0.945	0.964		0.980	0.984	0.983
131	Plan Curvature	0.957	0.933	0.921	0.930	0.948		0.957	0.961	0.960
132	Plan Curvature		0.812	0.829		0.883				
133	Planar Curvature		0.812	0.829		0.883				
134	Planimetric-to-surface ratio	No Loadings								
135	Profile Curvature		0.841	0.869	0.870					
136	Profile Curvature	Complex Variable								
137	Profile Curvature		0.841	0.869	0.870					
138	Profile Curvature	0.619								
139	Profile Curvature		0.787	0.780	0.733	0.732		0.678		
140	Profile Curvature									0.682
141	Profile Curvature			0.861	0.828	0.807	0.702	0.793	0.813	0.806
142	Profile Curvature		0.841	0.869	0.870	0.829	0.729	0.771	0.848	0.845
143	Profile Curvature									
144	Profile Curvature							0.492	0.563	0.469
145	Profile Curvature	0.916	0.893	0.889	0.899	0.914	0.924	0.907	0.910	0.894
146	Profile Curvature			0.414	0.344	0.350				0.165
147	Profile Curvature	0.848	0.841	0.818	0.820	0.805	0.706	0.746	0.766	0.740
148	Profile Curvature	0.916	0.893	0.889	0.899	0.914	0.924	0.907	0.910	0.894
149	Profile Curvature	0.944	0.928	0.938	0.947	0.960	0.965	0.955	0.969	0.961
150	Profile Curvature	0.944	0.927	0.936	0.949	0.964	0.984	0.970	0.982	0.978
151	Profile Curvature	0.935	0.915	0.919	0.931	0.948	0.968	0.952	0.960	0.954
152	Profile Curvature	0.763	0.727	0.698	0.943	0.776	0.978	0.907		0.949

153	Profile Curvature	0.763	0.727	0.698	0.943	0.776	0.978	0.907	0.768	0.949
154	Range	0.964	0.961	0.948	0.949	0.941	0.910	0.924	0.931	0.975
155	Range									0.975
156	Real Area	0.965	0.935	0.943	0.929	0.854	0.781	0.762		
157	Relative deviation from mean	0.941	0.913	0.927	0.930	0.944	0.930	0.944	0.954	0.949
158	Representativeness			0.841	0.880	0.870				0.240
159	Residual at centre	0.312								
160	Roughness	0.964	0.961	0.948	0.949	0.941	0.910	0.924	0.931	0.975
161	Ruggedness Index	0.991	0.983	0.972	0.968	0.954	0.865	0.948	0.947	
162	Slope	0.982	0.972	0.982	0.972	0.962	0.944	0.899	0.997	0.995
163	Slope	0.937	0.874	0.866	0.995	0.989	0.989	0.985	0.996	0.994
164	Slope	0.982	0.972	0.982	0.972	0.962	0.944	0.899	0.997	0.995
165	Slope	0.971	0.938	0.943	0.950	0.925	0.889	0.818		
166	Slope	0.982	0.972	0.982	0.972	0.962	0.944	0.899	0.997	0.995
167	Slope	0.979	0.963	0.965	0.957	0.958	0.934	0.886	0.990	0.973
168	Slope	0.918	0.871	0.847	0.987	0.980	0.984	0.974	0.991	0.990
169	Slope	0.971	0.938	0.943	0.950	0.925	0.889	0.818		
170	Slope	0.982	0.972	0.982	0.972	0.962	0.944	0.899	0.997	0.995
171	Slope	0.937	0.874	0.866	0.995	0.989	0.989	0.985	0.996	0.994
172	Slope	0.982	0.972	0.982	0.972	0.962	0.944	0.899	0.997	0.995
173	Slope	0.971	0.938	0.943	0.950	0.925	0.889			
174	Slope			0.907						
175	Slope	0.971	0.938	0.943	0.950	0.925	0.889			
176	Slope	0.937	0.874	0.866	0.986	0.989	0.989	0.985	0.996	0.994
177	Slope	0.982	0.972	0.982	0.972	0.962	0.944	0.899	0.997	0.995
178	Slope	0.841				0.911		0.840	0.874	
179	Slope	0.843				0.911		0.840	0.874	
180	Slope	0.600			0.443	0.337		0.298	0.402	0.362
181	Slope	0.805	0.660					0.818		
182	Slope	0.862	0.735			0.920		0.431	0.841	0.808
183	Slope	0.862	0.735			0.920		0.431	0.841	0.808
184	Slope	0.860	0.741		0.500	0.905		0.394	0.858	0.786
185	Slope	0.863	0.740			0.958		0.441	0.948	0.908
186	Slope	0.952	0.820				0.670	0.670	0.744	0.725
187	Slope	0.982	0.972	0.982	0.972	0.962	0.944	0.899	0.997	0.995
188	Slope Variability	0.452	0.394	0.426						
189	Standard Deviation	0.993	0.986	0.985	0.987	0.978	0.961	0.966	0.962	0.960
190	Standard Deviation	0.993	0.986	0.985	0.987	0.978	0.961	0.966	0.962	0.960
191	Standard Deviation	0.965	0.950	0.950	0.960	0.918				
192	Standard Deviation	0.985	0.979	0.976	0.977	0.954	0.735	0.825	0.830	0.736
193	Standard Deviation	0.972	0.960	0.960	0.968	0.934		0.720	0.730	
194	Standard Deviation	0.965	0.950	0.950	0.960	0.918				
195	Standard Deviation of Slope	0.718	0.644				0.212		0.254	0.214
196	Surface Area	0.990	0.979	0.987	0.986	0.974	0.922	0.912	0.916	
197	Surface Area	0.990	0.980	0.982	0.981	0.917	0.869	0.878	0.873	0.815
198	Surface Area to Planar Surface	0.990	0.979	0.987	0.986	0.974	0.922	0.912	0.916	
199	Surface Ratio	0.990	0.979	0.987	0.981	0.974	0.922	0.912	0.916	
200	Surface Ratio	0.990	0.980	0.982	0.981	0.917	0.869	0.878	0.873	0.815
201	Surface Roughness Index	0.778	0.753							
202	Tangential Curvature	0.733							0.948	0.880
203	Tangential Curvature	0.811	0.823							
204	Tangential Curvature	0.553	0.830	0.809	0.706	0.823		0.588	0.747	0.728

205	Terrain Ruggedness (VRM)	0.778	0.753					0.660	0.890	0.603
206	Terrain Ruggedness Index			0.942	0.940	0.939	0.916	0.921	0.936	0.964
207	Terrain Ruggedness Index	0.986	0.974	0.957	0.946	0.937	0.823	0.918	0.918	
208	Terrain Ruggedness Index	0.962	0.939	0.938	0.955	0.930				
209	Terrain Ruggedness Index	0.996	0.992	0.974	0.968	0.954	0.806	0.928	0.932	
210	Terrain Ruggedness Index	0.976	0.961	0.953	0.961	0.941		0.769	0.784	
211	Terrain Ruggedness Index	0.962	0.939	0.938	0.955	0.930				
212	Topographic Position Index	0.798	0.807	0.854	0.870	0.881	0.844	0.884	0.884	0.881
213	Topographic Position Index	0.975	0.978	0.989	0.988	0.985	0.980	0.978	0.985	0.985
214	Topographic Position Index	0.878	0.881	0.952	0.942	0.948	0.894	0.907	0.973	0.982
215	Topographic Position Index	0.912	0.901	0.962	0.970	0.962	0.904	0.895	0.977	0.976
216	Topographic Position Index	0.902	0.906	0.960	0.952	0.957	0.921	0.932	0.978	0.984
217	Topographic Position Index	0.878	0.881	0.952	0.942	0.948	0.894	0.907	0.973	0.982
218	Topographic Ruggedness Index	0.991	0.983	0.972	0.968	0.954	0.865	0.948	0.947	
219	Total Curvature	0.838	0.796	0.847	0.814					
220	Total Curvature	0.654	0.615	0.770	0.854					
221	Value Range	0.916	0.882	0.876	0.895					
222	Value Range	0.916	0.882	0.876	0.895					
223	Value Range	0.916	0.882	0.876	0.895					
224	Value Range	0.916	0.882	0.876	0.895					
225	Variance	0.967	0.952	0.954	0.957	0.934	0.914			
226	Vector Ruggedness Measure	0.778	0.753					0.660	0.890	0.603
227	Vector Ruggedness Measure	0.990	0.985	0.981	0.973	0.961		0.820	0.929	0.718
228	Vector Ruggedness Measure	0.989	0.984	0.972	0.957	0.949		0.871	0.904	0.830
229	Vector Ruggedness Measure	0.994	0.995	0.997	0.989	0.990		0.888	0.978	0.845
230	Vector Ruggedness Measure	0.990	0.985	0.981	0.973	0.961		0.820	0.929	0.718
AVERAGE:		0.917	0.904	0.902	0.902	0.891	0.879	0.853	0.887	0.866

Table C.4: Average covariation ranking from VIF and MI. Legend: Values in italic correspond to low cardinality variables, complex variables, or variables with low communality. Attributes with lower values are least multicollinear.

ID	TERRAIN ATTRIBUTES	AVERAGE RANK OF EACH VARIABLE FOR EACH SURFACE								
		A0	A1	A2	A3	A4	A5	A6	A7	A8
1	Bathymetric Position Index	*72*	*90.5*	*101*	*106*	*106.5*	*110.5*	*112.5*	*114.5*	*121*
2	Center versus Neighbors Variability	*4*	*2.5*	*10*	*10.5*	*9*	*6*	*5.5*	*6*	*9*
3	Coefficient of variation	*38*	*30*	45	48	50.5	76.5	75	74	79.5
4	Convergence Index	44.5	*21.5*	40	40	50	*55*	*57*	*57.5*	*61.5*
5	Convergence Index	84.5	93.5	104	106.5	109.5	110.5	115.5	113.5	116.5
6	Cross-sectional Curvature	113	99	104	127.5	*111.5*	*128.5*	*112*	*106*	*114*
7	Cross-sectional Curvature	110	82.5	*73.5*	*107*	*106.5*	*81*	*75.5*	*63.5*	*111*
8	Curvature	132	133	137.5	141	146.5	142	143.5	138	144
9	Curvature	132	133	137.5	141	146.5	142	143.5	138	144
10	Curvature	160	141	141.5	162.5	141	142	142	164.5	168.5
11	Curvature	*101.5*	*105*	*108.5*	*101.5*	*98*	*96.5*	*98.5*	98.5	104.5
12	Curvature	152	152.5	155	153	156	156	152	149	155
13	Curvature	161	154.5	164	167.5	170	169	164.5	167.5	172.5
14	Curvature	182	182	182	182	182	182	182	182	186
15	Curvature	83	79.5	81	82.5	85.5	85	89	86.5	87.5
16	Curvature	182	182	182	182	182	182	182	182	186
17	Deviation from Mean Elevation	74	81	85.5	90.5	90.5	95.5	93	92.5	96
18	Deviation from Mean Value	78.5	80.5	91	95	105	108	110.5	110.5	114.5
19	Deviation from Mean Value	111.5	115.5	123.5	125	131	132.5	131	134	137
20	Deviation from Mean Value	118.5	123	129	131.5	133.5	135	135.5	134.5	137
21	Deviation from Mean Value	78.5	80.5	91	95	105	108	110.5	110.5	114.5
22	Difference from Mean Elevation	36.5	*44.5*	56.5	72.5	80.5	87	91	94	97
23	Difference from Mean Value	96.5	105.5	115	120.5	129	130	132.5	134	139
24	Difference from Mean Value	124	127.5	133	136	138	138	138.5	136.5	139.5
25	Difference from Mean Value	127	131	134.5	136	136.5	140.5	144	145.5	148.5
26	Difference from Mean Value	96.5	105.5	115	120.5	129	130	132.5	134	139
27	Easterness	177	177.5	173	175.5	128	179	179	176.5	153
28	Easterness	145	168.5	131	138.5	90	170	169.5	170.5	92
29	Easterness	91.5	91.5	180	180	130	95.5	178.5	180.5	132.5
30	Easterness	167	166	163	166	115	153.5	164.5	165	157.5
31	Easterness	164.5	165	135	133.5	164.5	164.5	152.5	152.5	166
32	Easterness	88.5	84.5	77	66.5	96	112	67.5	68.5	97.5
33	Easterness	145	154	144	146.5	148	131.5	149.5	119.5	165
34	Easterness	162.5	134.5	117	115	131	130	92	90.5	132.5
35	Easterness	128	101.5	168.5	135	54.5	75.5	127.5	128	58
36	Easterness	86.5	86.5	172	161.5	162.5	89	170	170	127.5
37	Easterness	117	102.5	79	76	135.5	113.5	74	100.5	84.5
38	Easterness	84.5	76	53.5	44	51	58	42	41.5	43.5
39	Easterness	117	102.5	79	76	135.5	113.5	74	100.5	84.5
40	Easterness	154.5	152	120	152	161	151.5	136.5	138	169
41	Easterness	140.5	164	152	159.5	174	163	149.5	156.5	174.5
42	Easterness	*40.5*	*23*	*27.5*	*14*	*17*	*16.5*	*11*	*24*	*20*
43	Easterness	64.5	69.5	64.5	65	67.5	73	71	*72.5*	*75.5*
44	Easterness	79	51	46.5	37	58.5	49	38.5	30.5	47
45	Easterness	59.5	50	24.5	23	37.5	33	16	4	33.5
46	Easterness	78.5	74	51	60	47.5	57	52	51	53.5

47	Easterness	78.5	74	51	60	47.5	57	52	51	53.5
48	Easterness	104	100.5	93.5	91	90.5	92	88	87.5	89
49	Easterness	118.5	116	111.5	112	106	105	104	105	107
50	Easterness	23	53.5	29	39	64	52.5	46	50.5	70
51	Easterness	91.5	91.5	180	180	130	95.5	178.5	180.5	132.5
52	Fractal Dimension	61.5	45.5	*43.5*	*64.5*	67	112.5	110	113.5	94
53	General Curvature	132	157.5	164	160.5	168	164	163	137	140.5
54	Gradient	152.5	151	153	154	156	155.5	155	*156*	*161.5*
55	Gradient	116.5	127	151	98	132	100.5	129	139.5	106.5
56	Gradient	157.5	157.5	148	153.5	130	152	156.5	161	160.5
57	Longitudinal Curvature	*98*	*117.5*	*121.5*	*108*	*128.5*	*113*	*130*	*112*	*106*
58	Longitudinal Curvature	*72.5*	*113.5*	*99.5*	*82*	*82.5*	*96*	*119.5*	*98.5*	*73.5*
59	Maximum	112	115	90.5	95	88.5	88.5	83.5	100.5	127
60	Maximum	*112*	*115*	*90.5*	*95*	*88.5*	*88.5*	*83.5*	*100.5*	120.5
61	Maximum Curvature	39.5	54	56	76	65.5	69.5	72.5	72	78
62	Maximum Curvature	86	56.5	65	55.5	65.5	70	69	102.5	66.5
63	Maximum Value	94	94	25	14.5	13	84	9.5	81	84.5
64	Maximum Value	94	94	25	14.5	13	84	9.5	81	84.5
65	Maximum Value	94	94	25	14.5	13	84	9.5	81	84.5
66	Maximum Value	94	94	25	14.5	13	84	9.5	81	84.5
67	Mean	88.5	93	114.5	100.5	107	95	101.5	80.5	92.5
68	Mean	*88.5*	*93*	*114.5*	*100.5*	*107*	*95*	*101.5*	*80.5*	128
69	Mean Curvature	152.5	156.5	163	140	158	153.5	155	154	159.5
70	Mean of Residuals	*6*	*6*	*5.5*	*5*	*3.5*	*4.5*	*8.5*	*5.5*	*4*
71	Mean Value	128	132.5	132	123	111.5	106	98.5	86	77
72	Mean Value	123	139	140	132.5	134.5	129.5	126.5	120.5	121
73	Mean Value	147.5	133	132	127	137	130.5	130.5	128	129
74	Mean Value	128	132.5	132	123	111.5	106	98.5	86	77
75	Median	100	104.5	103.5	96.5	98.5	92	86	80	78
76	Minimum	134	136.5	136	130	131	126	125	126.5	140.5
77	Minimum	134	136.5	136	130	131	126	125	126.5	140.5
78	Minimum Curvature	44.5	45.5	56.5	79.5	65	70.5	75	76	85
79	Minimum Curvature	56.5	85.5	88.5	87.5	90.5	96	97.5	65.5	100
80	Minimum Value	14	13.5	80	89.5	89.5	16	85.5	7	15.5
81	Minimum Value	14	13.5	80	89.5	89.5	16	85.5	7	15.5
82	Minimum Value	14	13.5	80	89.5	89.5	16	85.5	7	15.5
83	Minimum Value	14	13.5	80	89.5	89.5	16	85.5	7	15.5
84	Modified/local Melton Ruggedness Index	99.5	128.5	124	127.5	124.5	110	108.5	107.5	160.5
85	Morphometric Protection Index	42.5	55	74	78.5	89	93	91	90	95.5
86	Northerness	171	172	171.5	172.5	172	178	164	160.5	143
87	Northerness	143	94.5	124.5	90.5	170	152.5	90.5	90.5	93.5
88	Northerness	176	128	179	180	180	91	180.5	180.5	183.5
89	Northerness	158	141	161	161.5	164.5	163.5	164.5	152	166.5
90	Northerness	139.5	169	131	136	151.5	163	151	133	135.5
91	Northerness	72	84	76	91	73.5	76	89.5	91.5	92.5
92	Northerness	142	143.5	137.5	145	144	128.5	144	120	150.5
93	Northerness	121.5	149	117	116.5	91.5	161	122.5	110.5	109.5
94	Northerness	164.5	51	166	55	103	169	58	57.5	60.5
95	Northerness	166	125.5	167.5	172	161	85.5	172.5	176.5	176.5
96	Northerness	82	117.5	69	80.5	66.5	101	74.5	112.5	66.5
97	Northerness	60	69.5	39	56	38	54.5	45	44.5	41
98	Northerness	82	117.5	69	80.5	66.5	101	74.5	112.5	66.5

99	Northerness	91	155.5	115	162.5	149	110	161	164	168
100	Northerness	150	134	149.5	155	160.5	134.5	130	163	168
101	Northerness	13	26	15	28	17	29.5	29	27.5	27
102	Northerness	44.5	53.5	59	67.5	66.5	72.5	73.5	72.5	75.5
103	Northerness	46.5	34	52.5	57.5	48	52.5	52.5	47	50.5
104	Northerness	31	39	24	39.5	25.5	30	37	37	36
105	Northerness	60.5	56	49	46.5	48	50.5	42	40	39
106	Northerness	60.5	56	49	46.5	48	50.5	42	40	39
107	Northerness	96	92.5	88.5	89	84	82.5	85	80.5	77
108	Northerness	114	108.5	106.5	106.5	103	100.5	103.5	103.5	102.5
109	Northerness	17	33	16	31	13.5	27	37.5	37.5	41
110	Northerness	176	128	179	180	180	91	180.5	180.5	183.5
111	Percentile	41.5	63.5	75.5	81	89.5	95.5	101	101	103.5
112	Percentile	41.5	63.5	75.5	81	89.5	95.5	101	101	103.5
113	Percentile	41.5	63.5	75.5	81	89.5	95.5	101	101	103.5
114	Percentile	41.5	63.5	75.5	81	89.5	95.5	101	101	103.5
115	Plan Curvature	138	97.5	143	91.5	106.5	156	108	69.5	102.5
116	Plan Curvature	75	71	72.5	71	67.5	72.5	71.5	60	63
117	Plan Curvature	138	97.5	143	91.5	106.5	156	108	69.5	102.5
118	Plan Curvature	42.5	38	36.5	33.5	36.5	38	35.5	32.5	32
119	Plan Curvature	50.5	48.5	55	61.5	58	55.5	59.5	115	71
120	Plan Curvature	26	29	23.5	22	23	25.5	21.5	16.5	16.5
121	Plan Curvature	131	116	120	122.5	117	129.5	123.5	128	128.5
122	Plan Curvature	96	103.5	74.5	120.5	122	81.5	115	156.5	109.5
123	Plan Curvature	182	182	182	182	182	182	182	182	186
124	Plan Curvature	45	36	28.5	32	29	29	43	32.5	31.5
125	Plan Curvature	182	182	182	182	182	182	182	182	186
126	Plan Curvature	10.5	6	17	8.5	7.5	6	21	15	5.5
127	Plan Curvature	21	19.5	32.5	42.5	17	43.5	18.5	14	13
128	Plan Curvature	19.5	25	9.5	15	30.5	21	31	30	30
129	Plan Curvature	68	69	68	83.5	67.5	85	67.5	87	65
130	Plan Curvature	71	73.5	71	80.5	82	85	88.5	89	87.5
131	Plan Curvature	93	94	88	96	93	100.5	101.5	104	104
132	Plan Curvature	59	105.5	138	76.5	98.5	151	119.5	153.5	124.5
133	Planar Curvature	139	85.5	60	141	123	68.5	113.5	100.5	135.5
134	Planimetric-to-surface ratio	4	5	5	3	4	4	4.5	5	6
135	Profile Curvature	98.5	134.5	101	49	70	152	77	75	80
136	Profile Curvature	61	61	59.5	59.5	56.5	64	56	59	61
137	Profile Curvature	98.5	134.5	101	49	70	152	77	75	80
138	Profile Curvature	10	9	6	5.5	8.5	4	12	9.5	3.5
139	Profile Curvature	40	46.5	57	46	54	62	54.5	69	109
140	Profile Curvature	29	25	24.5	21.5	24.5	21	23	19	16.5
141	Profile Curvature	110.5	125.5	122	114.5	123.5	121.5	130.5	124	131.5
142	Profile Curvature	106	56.5	105.5	146.5	148.5	74	156	154	134.5
143	Profile Curvature	182	182	182	182	182	182	182	182	186
144	Profile Curvature	26.5	26.5	31.5	30.5	35	35.5	29.5	37	40
145	Profile Curvature	182	182	182	182	182	182	182	182	186
146	Profile Curvature	18	17.5	6	17	19	21.5	11.5	21.5	24
147	Profile Curvature	34.5	40	34	13	43	12.5	44	45	47.5
148	Profile Curvature	20.5	14	65.5	30.5	14	29	11.5	19.5	23.5
149	Profile Curvature	72	71.5	32.5	67	80.5	66.5	84.5	66.5	88.5
150	Profile Curvature	84	85	82.5	84	87.5	87	88	89	90.5

151	Profile Curvature	102	102	101	101	103	104.5	104.5	106	108.5
152	Profile Curvature	24.5	68	67.5	128.5	71.5	80	133.5	28	74
153	Profile Curvature	122.5	81	79.5	29	79	70	30.5	137	85.5
154	Range	155	116.5	131.5	106	101.5	133	131.5	102	141
155	Range	155	116.5	131.5	106	101.5	133	131.5	102	94
156	Real Area	125.5	126	121.5	114.5	117	121.5	113.5	113.5	108
157	Relative deviation from mean	93	99.5	108	110	111.5	114	114.5	115	120
158	Representativeness	9.5	9	17.5	15.5	16.5	17.5	17	16.5	19
159	Residual at centre	3	64	64.5	64	64	70	66	67.5	74
160	Roughness	18	112.5	109	113.5	111	91	87	83	157.5
161	Ruggedness Index	110.5	100	97.5	93	116	116	96.5	117.5	99
162	Slope	169	172.5	173.5	169	170	166.5	134	160.5	128.5
163	Slope	161	161	159.5	159.5	157	140	160.5	137.5	140.5
164	Slope	169	172.5	173.5	169	170	166.5	134	160.5	168.5
165	Slope	165.5	164.5	161	162.5	156.5	121	156	153.5	155.5
166	Slope	169	172.5	173.5	169	170	166.5	134	160.5	128.5
167	Slope	114.5	108	93	77	79.5	74.5	69.5	57	62
168	Slope	45	51	19	123	119.5	76	40.5	42	126
169	Slope	151.5	151.5	100.5	88	137.5	107.5	94	74	131
170	Slope	143.5	150.5	131.5	130	135.5	152.5	131	147.5	156.5
171	Slope	152.5	152.5	140.5	119.5	148.5	147	146	146.5	155.5
172	Slope	161	156	161	156	158.5	145	157.5	106	154.5
173	Slope	101	97	131.5	139.5	77	134.5	109	133	56
174	Slope	82.5	87.5	78	68.5	65.5	53	50	48	35
175	Slope	101	97	131.5	139.5	77	134.5	109	133	56
176	Slope	131	98	108.5	134	66.5	115.5	101	118	88.5
177	Slope	132	123.5	135	123.5	138.5	111	134.5	134.5	135.5
178	Slope	59	47	49	43.5	47.5	44	48	50	50.5
179	Slope	100.5	102.5	113	113.5	118	120	121	122.5	122
180	Slope	62	41.5	57	18	27	20	26.5	32.5	29.5
181	Slope	48.5	36	27	17	17	17.5	16.5	19	17.5
182	Slope	67.5	64.5	52.5	43.5	39	36.5	33.5	34.5	29
183	Slope	67.5	64.5	52.5	43.5	39	36.5	33.5	34.5	29
184	Slope	103.5	98	87.5	75	72	68	60	62.5	57
185	Slope	126.5	122.5	115	107	100	94	88	85.5	81.5
186	Slope	63.5	38.5	52	42	51	45	43	51	54
186	Slope	169	172.5	173.5	169	170	166.5	134	160.5	168.5
188	Slope Variability	8.5	8.5	12.5	10.5	19	19.5	14	19.5	21
189	Standard Deviation	140.5	140.5	136	117.5	115.5	115.5	113.5	127.5	129
190	Standard Deviation	140.5	140.5	136	117.5	115.5	115.5	113.5	127.5	129
191	Standard Deviation	95.5	96	86.5	82	75	71	66.5	63.5	60.5
192	Standard Deviation	128	128	118.5	116	109	107.5	104	100.5	98.5
193	Standard Deviation	126.5	125	118.5	117	112	110.5	108.5	107.5	102.5
194	Standard Deviation	128	128	118.5	116	109	107.5	104	100.5	98.5
195	Standard Deviation of Slope	35	36	29	17.5	13	12	12.5	9	9.5
196	Surface Area	126.5	141.5	106.5	116	91.5	107.5	83.5	106	106
197	Surface Area	141	141.5	141	139.5	139	131.5	135.5	134.5	100.5
198	Surface Area to Planar Surface	126.5	141.5	106.5	116	91.5	107.5	83.5	106	106
199	Surface Ratio	156.5	134.5	147	120	136.5	111	129.5	102.5	95.5
200	Surface Ratio	141	141.5	141	139.5	139	131.5	135.5	134.5	100.5
201	Surface Roughness Index	114.5	89	96	97.5	144	96	95	141.5	94.5
202	Tangential Curvature	93	89.5	86.5	91.5	78	75.5	83.5	105.5	103

203	Tangential Curvature	84	86	*80.5*	*88*	*85*	*82*	*85*	*87.5*	*86.5*
204	Tangential Curvature	*20.5*	30.5	21	18.5	30	*39.5*	*18.5*	19	41
205	Terrain Ruggedness (VRM)	24.5	81	*66.5*	*73*	*26.5*	*74.5*	73.5	25.5	72
206	Terrain Ruggedness Index	*59.5*	*72.5*	48	68.5	65.5	71	62	94.5	51
207	Terrain Ruggedness Index	103	103.5	79.5	61.5	56.5	59.5	60	58	*60*
208	Terrain Ruggedness Index	78.5	82	60.5	52.5	46	*44.5*	*50.5*	*52*	*47*
209	Terrain Ruggedness Index	123	118	109	106.5	107	99	99	101	*100*
210	Terrain Ruggedness Index	117.5	119.5	104.5	104	100.5	*96.5*	94	95	*94.5*
211	Terrain Ruggedness Index	78.5	82	60.5	52.5	46	*44.5*	*50.5*	*52*	*47*
212	Topographic Position Index	50.5	63.5	78	84	89.5	86.5	92.5	92	95.5
213	Topographic Position Index	127.5	132	137.5	148	151.5	153.5	157.5	159.5	165
214	Topographic Position Index	96.5	105.5	115	120.5	129	130	132.5	134	139
215	Topographic Position Index	155.5	163.5	164	160.5	161	161.5	163	162	166.5
216	Topographic Position Index	127	131	134.5	136	136.5	140.5	144	145.5	148.5
217	Topographic Position Index	96.5	105.5	115	120.5	129	130	132.5	134	139
218	Topographic Ruggedness Index	150.5	148.5	143	140	117	120.5	136.5	115	*141.5*
219	Total Curvature	25	27.5	26.5	28.5	*5.5*	*6.5*	*15*	*8.5*	*8.5*
220	Total Curvature	14	15.5	8.5	15	*29*	*29.5*	*31*	*32*	*32*
221	Value Range	50.5	47	33.5	41	*42*	*40*	*31*	*40*	*32*
222	Value Range	50.5	47	33.5	41	*42*	*40*	*31*	*40*	*32*
223	Value Range	50.5	47	33.5	41	*42*	*40*	*31*	*40*	*32*
224	Value Range	50.5	47	33.5	41	*42*	*40*	*31*	*40*	*32*
225	Variance	104.5	104	99	118	115	97.5	*113.5*	*95.5*	*87*
226	Vector Ruggedness Measure	145.5	138	*159.5*	*159*	*158.5*	*159*	158.5	158.5	162
227	Vector Ruggedness Measure	11	6	10	5	13	*12.5*	9	11.5	11.5
228	Vector Ruggedness Measure	51.5	48.5	45.5	49	46.5	*44*	46.5	42.5	38.5
229	Vector Ruggedness Measure	86.5	85.5	85.5	87.5	85.5	*83*	82.5	81	74.5
230	Vector Ruggedness Measure	11	6	10	5	13	*12.5*	9	11.5	11.5

Table C.5: Final PCA solution (rotated component matrix) for surface A0.

			Components										
ID	**Terrain Attributes**	**Software**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
213	Topographic Position Index	Quantum GIS	.954										
12	General Curvature	SAGA GIS	.952										
215	Topographic Position Index	SAGA GIS	.928										
9	General Curvature	ArcGIS with Python	.928										
13	General Curvature	SAGA GIS	.928										
24	Deviation from Mean Value	SAGA GIS	.924										
117	Plan Curvature	ArcGIS with Python	.911										
122	Plan Curvature	SAGA GIS	.911										
69	Mean Curvature	Landserf	.906										
53	General Curvature	DEM Surface Tools	.906										
10	General Curvature	SAGA GIS	.906										
157	Relative Deviation from Mean	ArcGIS with Python	-.905								.337		
121	Plan Curvature	SAGA GIS	.901										
19	Deviation from Mean Value	SAGA GIS	.900										
15	General Curvature	SAGA GIS	.899										
216	Topographic Position Index	SAGA GIS	.894										
5	Convergence Index	SAGA GIS	.890										
214	Topographic Position Index	SAGA GIS	.878										
20	Deviation from Mean Value	SAGA GIS	.871										
17	Deviation from Mean Elevation	Whitebox GAT	.869								-.363		
6	Cross-Sectional Curvature	DEM Surface Tools	.864										
119	Plan Curvature	SAGA GIS	.859										
18	Deviation from Mean Value	SAGA GIS	.854										
78	Minimum Curvature	Idrisi	-.851										
61	Maximum Curvature	Idrisi	-.849										
7	Cross-Sectional Curvature	Landserf	.845										
203	Tangential Curvature	uDig	-.827										
85	Morphometric Position Index	SAGA GIS	-.820										
62	Maximum Curvature	Landserf	.813										
124	Plan Curvature	SAGA GIS	.812										
79	Minimum Curvature	Landserf	.807										
212	Topographic Position Index	ArcGIS with Python	.796										
202	Tangential Curvature	DEM Surface Tools	.772										
4	Convergence Index	SAGA GIS	.740										

159	Residual at Centre	Landserf	.515										
172	Slope	Quantum GIS		.955									
177	Slope	SAGA GIS		.955									
166	Slope	ArcGIS with Python		.955									
56	Gradient	uDig		.955									
170	Slope	Quantum GIS		.955									
167	Slope	Diva-GIS		.954									
173	Slope	SAGA GIS		.948									
165	Slope	DEM Surface Tools		.948									
169	Slope	Landserf		.948									
54	Gradient	uDig		.948									
171	Slope	Quantum GIS		.928									
163	Slope	DEM Surface Tools		.928									
55	Gradient	uDig		.928									
176	Slope	SAGA GIS		.928									
168	Slope	Idrisi		.918									
185	Slope	TNTmips		.882									
182	Slope	TNTmips		.881									
184	Slope	TNTmips		.881									
181	Slope	TNTmips		.846									
52	Fractal Dimension	Idrisi		.828			.503						
226	Vector Ruggedness Measure	ArcGIS with Python		-.816									
205	Vector Ruggedness Measure	Benthic Terrain Modeler		-.816									
201	Surface Roughness Index	ArcGIS with Python		.816									
195	Standard Deviation of Slope	ArcGIS with Python		-.781									
180	Slope	SAGA GIS		.722									
179	Slope	SAGA GIS		.716							-.408		
178	Slope	SAGA GIS		.716							-.406		
31	Easterness	ArcGIS with Python			.998								
36	Easterness	Quantum GIS			.998								
29	Easterness	DEM Surface Tools			.998								
27	Easterness	Benthic Terrain Modeler			.998								
34	Easterness	Quantum GIS			.998								
41	Easterness	SAGA GIS			.998								
33	Easterness	Landserf			.997								
37	Easterness	SAGA GIS			.997								
30	Easterness	DEM Surface Tools			.997								
40	Easterness	SAGA GIS			.996								

35	Easterness	Quantum GIS			.996								
28	Easterness	DEM Surface Tools			.996								
38	Easterness	SAGA GIS			.996								
32	Easterness	Idrisi			.996								
44	Easterness	SAGA GIS			.996								
48	Easterness	TNTmips			.994								
49	Easterness	TNTmips			.994								
46	Easterness	TNTmips			.994								
45	Easterness	TNTmips			.991								
43	Easterness	SAGA GIS			.983								
50	Easterness	uDig			.976								
22	Difference from Mean Elevation	Whitebox GAT	.475		.702								
100	Northerness	SAGA GIS				.996							
93	Northerness	Quantum GIS				.996							
95	Northerness	Quantum GIS				.996							
88	Northerness	DEM Surface Tools				.996							
86	Northerness	Benthic Terrain Modeler				.996							
90	Northerness	ArcGIS with Python				.996							
92	Northerness	Landserf				.995							
89	Northerness	DEM Surface Tools				.995							
96	Northerness	SAGA GIS				.995							
87	Northerness	DEM Surface Tools				.992							
94	Northerness	Quantum GIS				.992							
99	Northerness	SAGA GIS				.992							
97	Northerness	SAGA GIS				.989							
91	Northerness	Idrisi				.989							
103	Northerness	SAGA GIS				.982							
108	Northerness	TNTmips				.964							
105	Northerness	TNTmips				.964							
107	Northerness	TNTmips				.963							
104	Northerness	TNTmips				.956							
102	Northerness	SAGA GIS				.926							
109	Northerness	uDig				.857							
225	Variance	Benthic Terrain Modeler		.368			.908						
209	Terrain Ruggedness Index	SAGA GIS		.498			.858						
199	Surface Ratio	DEM Surface Tools		.500			.856						
198	Planimetric-to-Surface Ratio	Benthic Terrain Modeler		.500			.856						
210	Terrain Ruggedness Index	SAGA GIS		.485			.853						

218	Terrain Ruggedness Index	Whitebox GAT		.504			.853						
161	Ruggedness Index	Quantum GIS		.504			.853						
207	Terrain Ruggedness Index	Quantum GIS		.504			.850						
192	Standard Deviation	SAGA GIS		.504			.848						
208	Terrain Ruggedness Index	SAGA GIS		.478			.848						
190	Standard Deviation	ArcGIS with Python		.520			.846						
193	Standard Deviation	SAGA GIS		.498			.843						
191	Standard Deviation	SAGA GIS		.496			.840						
200	Surface Ratio	ArcGIS with Python		.529			.839						
84	Melton Ruggedness Index	ArcGIS with Python		.510			.832						
160	Roughness	Quantum GIS		.510			.832						
154	Range	ArcGIS with Python		.510			.832						
156	Real Area	SAGA GIS		.524			.827						
221	Value Range	SAGA GIS		.471			.819						
186	Slope	uDig		.483			.773						
71	Mean Value	SAGA GIS						.987					
73	Mean Value	SAGA GIS						.987					
67	Minimum Value	ArcGIS with Python						.987					
72	Mean Value	SAGA GIS						.986					
80	Minimum Value	SAGA GIS						.986					
76	Mean	ArcGIS with Python						.986					
75	Median	Diva-GIS						.985					
59	Maximum Value	ArcGIS with Python						.985					
63	Maximum Value	SAGA GIS						.982					
150	Profile Curvature	TNTmips							.912				
149	Profile Curvature	TNTmips							.910				
151	Profile Curvature	TNTmips							.909				
148	Profile Curvature	TNTmips							.902				
147	Profile Curvature	TNTmips							.835				
131	Plan Curvature	TNTmips	.366							.897			
130	Plan Curvature	TNTmips	.364							.896			
129	Plan Curvature	TNTmips	.366							.895			
128	Plan Curvature	TNTmips	.365							.889			
127	Plan Curvature	TNTmips	.340							.814			
153	Profile Curvature	Whitebox GAT									.855		
152	Profile Curvature	uDig									.855		
138	Profile Curvature	Landserf									-.766		

204	Tangential Curvature	Whitebox GAT									.699		
229	Vector Ruggedness Measure	SAGA GIS										.941	
227	Vector Ruggedness Measure	SAGA GIS										.940	
228	Vector Ruggedness Measure	SAGA GIS										.937	
219	Total Curvature	DEM Surface Tools											.913
220	Total Curvature	Whitebox GAT											.806
188	Slope	ArcGIS with Python											.666

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 7 iterations.

Table C.6: Final PCA solution (rotated component matrix) for surface A1.

			Components										
ID	**Terrain Attributes**	**Software**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
12	General Curvature	SAGA GIS	.968										
213	Topographic Position Index	Quantum GIS	.968										
24	Deviation from Mean Value	SAGA GIS	.946										
215	Topographic Position Index	SAGA GIS	.933										
13	General Curvature	SAGA GIS	.933										
9	General Curvature	ArcGIS with Python	.933										
157	Relative Deviation from Mean	ArcGIS with Python	-.924										
53	General Curvature	DEM Surface Tools	.921										
10	General Curvature	SAGA GIS	.921										
69	Mean Curvature	Landserf	.921										
19	Deviation from Mean Value	SAGA GIS	.920										
216	Topographic Position Index	SAGA GIS	.918										
214	Topographic Position Index	SAGA GIS	.902										
5	Convergence Index	SAGA GIS	.895										
20	Deviation from Mean Value	SAGA GIS	.892										
15	General Curvature	SAGA GIS	.881										
117	Plan Curvature	ArcGIS with Python	.875										
122	Plan Curvature	SAGA GIS	.875										
18	Deviation from Mean Value	SAGA GIS	.874										
78	Minimum Curvature	Idrisi	-.871										
17	Deviation from Mean Elevation	Whitebox GAT	.867								-.341		
61	Maximum Curvature	Idrisi	-.866										
121	Plan Curvature	SAGA GIS	.862							.322			
85	Morphometric Position Index	SAGA GIS	-.858										
212	Topographic Position Index	ArcGIS with Python	.852										
119	Plan Curvature	SAGA GIS	.829										
6	Cross-Sectional Curvature	DEM Surface Tools	.826							.346			
62	Maximum Curvature	Landserf	.823										
79	Minimum Curvature	Landserf	.820										
124	Plan Curvature	SAGA GIS	.797										
7	Cross-Sectional Curvature	Landserf	.795							.379			
142	Profile Curvature	SAGA GIS	.772						.448				
137	Profile Curvature	ArcGIS with Python	-.772						-.448				
203	Tangential Curvature	uDig	-.729								.504		

139	Profile Curvature	SAGA GIS	.721						.496				
177	Slope	SAGA GIS		.948									
170	Slope	Quantum GIS		.948									
56	Gradient	uDig		.948									
166	Slope	ArcGIS with Python		.948									
172	Slope	Quantum GIS		.948									
167	Slope	Diva-GIS		.943									
173	Slope	SAGA GIS		.926									
165	Slope	DEM Surface Tools		.926									
54	Gradient	uDig		.926									
169	Slope	Landserf		.926									
176	Slope	SAGA GIS		.899									
171	Slope	Quantum GIS		.899									
163	Slope	DEM Surface Tools		.899									
55	Gradient	uDig		.899									
168	Slope	Idrisi		.895									
52	Fractal Dimension	Idrisi		.829			.473						
184	Slope	TNTmips		.802									
185	Slope	TNTmips		.800									
182	Slope	TNTmips		.796									
181	Slope	TNTmips		.752									
205	Vector Ruggedness Measure	Benthic Terrain Modeler		-.707									.397
226	Vector Ruggedness Measure	ArcGIS with Python		-.707									.397
201	Surface Roughness Index	ArcGIS with Python		.707									-.397
195	Standard Deviation of Slope	ArcGIS with Python		-.660			-.367						
36	Easterness	Quantum GIS			.994								
41	Easterness	SAGA GIS			.994								
29	Easterness	DEM Surface Tools			.994								
27	Easterness	Benthic Terrain Modeler			.994								
34	Easterness	Quantum GIS			.994								
31	Easterness	ArcGIS with Python			.994								
30	Easterness	DEM Surface Tools			.992								
37	Easterness	SAGA GIS			.992								
33	Easterness	Landserf			.992								
40	Easterness	SAGA GIS			.988								
28	Easterness	DEM Surface Tools			.988								
35	Easterness	Quantum GIS			.988								
32	Easterness	Idrisi			.988								

38	Easterness	SAGA GIS			.987								
49	Easterness	TNTmips			.980								
48	Easterness	TNTmips			.979								
46	Easterness	TNTmips			.979								
45	Easterness	TNTmips			.970								
50	Easterness	uDig			.959								
44	Easterness	SAGA GIS			.954								
43	Easterness	SAGA GIS			.953								
100	Northerness	SAGA GIS				.990							
86	Northerness	Benthic Terrain Modeler				.990							
90	Northerness	ArcGIS with Python				.990							
95	Northerness	Quantum GIS				.990							
88	Northerness	DEM Surface Tools				.990							
93	Northerness	Quantum GIS				.990							
92	Northerness	Landserf				.986							
96	Northerness	SAGA GIS				.986							
89	Northerness	DEM Surface Tools				.986							
87	Northerness	DEM Surface Tools				.978							
94	Northerness	Quantum GIS				.978							
99	Northerness	SAGA GIS				.978							
97	Northerness	SAGA GIS				.975							
91	Northerness	Idrisi				.975							
107	Northerness	TNTmips				.928							
108	Northerness	TNTmips				.928							
105	Northerness	TNTmips				.927							
104	Northerness	TNTmips				.909							
103	Northerness	SAGA GIS				.888							
102	Northerness	SAGA GIS				.880							
109	Northerness	uDig				.829							
225	Variance	Benthic Terrain Modeler		.327			.914						
209	Terrain Ruggedness Index	SAGA GIS		.446			.877						
199	Surface Ratio	DEM Surface Tools		.449			.875						
198	Planimetric-to-Surface Ratio	Benthic Terrain Modeler		.449			.875						
207	Terrain Ruggedness Index	Quantum GIS		.453			.869						
218	Terrain Ruggedness Index	Whitebox GAT		.457			.869						
161	Ruggedness Index	Quantum GIS		.457			.869						
210	Terrain Ruggedness Index	SAGA GIS		.423			.868						
192	Standard Deviation	SAGA GIS		.453			.866						

190	Standard Deviation	ArcGIS with Python		.481			.862						
208	Terrain Ruggedness Index	SAGA GIS		.412			.858						
193	Standard Deviation	SAGA GIS		.442			.858						
154	Range	ArcGIS with Python		.464			.855						
84	Melton Ruggedness Index	ArcGIS with Python		.464			.855						
160	Roughness	Quantum GIS		.464			.855						
191	Standard Deviation	SAGA GIS		.438			.853						
200	Surface Ratio	ArcGIS with Python		.508			.844						
156	Real Area	SAGA GIS		.504			.821						
221	Value Range	SAGA GIS		.410			.819						
186	Slope	uDig		.333			.688	.339					
80	Minimum Value	SAGA GIS						.979					
67	Minimum Value	ArcGIS with Python						.975					
71	Mean Value	SAGA GIS						.968					
73	Mean Value	SAGA GIS						.968					
72	Mean Value	SAGA GIS						.968					
76	Mean	ArcGIS with Python						.968					
75	Median	Diva-GIS						.967					
59	Maximum Value	ArcGIS with Python						.960					
63	Maximum Value	SAGA GIS						.953					
150	Profile Curvature	TNTmips							.904				
149	Profile Curvature	TNTmips							.902				
151	Profile Curvature	TNTmips							.901				
148	Profile Curvature	TNTmips							.892				
147	Profile Curvature	TNTmips							.822				
130	Plan Curvature	TNTmips								.906			
131	Plan Curvature	TNTmips								.905			
129	Plan Curvature	TNTmips								.903			
128	Plan Curvature	TNTmips								.892			
127	Plan Curvature	TNTmips								.812			
204	Tangential Curvature	Whitebox GAT									.892		
153	Profile Curvature	Whitebox GAT									.832		
152	Profile Curvature	uDig									.832		
132	Plan Curvature	Whitebox GAT	-.414								.794		
133	Planar Curvature	uDig	-.414								.794		
229	Vector Ruggedness Measure	SAGA GIS										.983	
227	Vector Ruggedness Measure	SAGA GIS										.978	

228	Vector Ruggedness Measure	SAGA GIS										.976	
219	Total Curvature	DEM Surface Tools											.888
220	Total Curvature	Whitebox GAT											.781
188	Slope	ArcGIS with Python											.604

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

Table C.7: Final PCA solution (rotated component matrix) for surface A2.

			Components												
ID	Terrain Attributes	Software	1	2	3	4	5	6	7	8	9	10	11	12	13
213	Topographic Position Index	Quantum GIS	.979												
12	General Curvature	SAGA GIS	.978												
24	Deviation from Mean Value	SAGA GIS	.955												
215	Topographic Position Index	SAGA GIS	.948												
13	General Curvature	SAGA GIS	.948												
9	General Curvature	ArcGIS with Python	.948												
157	Relative Deviation from Mean	ArcGIS with Python	-.945												
19	Deviation from Mean Value	SAGA GIS	.943												
216	Topographic Position Index	SAGA GIS	.925												
5	Convergence Index	SAGA GIS	.924												
69	Mean Curvature	Landserf	.922												
53	General Curvature	DEM Surface Tools	.922												
10	General Curvature	SAGA GIS	.922												
20	Deviation from Mean Value	SAGA GIS	.914												
214	Topographic Position Index	SAGA GIS	.907												
18	Deviation from Mean Value	SAGA GIS	.895											.331	
212	Topographic Position Index	ArcGIS with Python	.894												
17	Deviation from Mean Elevation	Whitebox GAT	.892												
85	Morphometric Position Index	SAGA GIS	-.890												
78	Minimum Curvature	Idrisi	-.889												
61	Maximum Curvature	Idrisi	-.889												
15	General Curvature	SAGA GIS	.884												
117	Plan Curvature	ArcGIS with Python	.855												
122	Plan Curvature	SAGA GIS	.855												
142	Profile Curvature	SAGA GIS	.833						.329						
137	Profile Curvature	ArcGIS with Python	-.833						-.329						
121	Plan Curvature	SAGA GIS	.831							.375					
62	Maximum Curvature	Landserf	.821												
79	Minimum Curvature	Landserf	.819												
124	Plan Curvature	SAGA GIS	.808												
141	Profile Curvature	SAGA GIS	.797						.414						
119	Plan Curvature	SAGA GIS	.797							.346					
6	Cross-Sectional Curvature	DEM Surface Tools	.786							.386					
139	Profile Curvature	SAGA GIS	.765						.389						

4	Convergence Index	SAGA GIS	.750											.471	
22	Difference from Mean Elevation	Whitebox GAT	.736		.351									.327	
225	Variance	Benthic Terrain Modeler		.955											
209	Terrain Ruggedness Index	SAGA GIS		.942											
192	Standard Deviation	SAGA GIS		.941											
198	Planimetric-to-Surface Ratio	Benthic Terrain Modeler		.939											
199	Surface Ratio	DEM Surface Tools		.939											
190	Standard Deviation	ArcGIS with Python		.938											
207	Terrain Ruggedness Index	Quantum GIS		.928											
161	Ruggedness Index	Quantum GIS		.926											
218	Terrain Ruggedness Index	Whitebox GAT		.926											
154	Range	ArcGIS with Python		.926											
84	Melton Ruggedness Index	ArcGIS with Python		.926											
160	Roughness	Quantum GIS		.926											
193	Standard Deviation	SAGA GIS		.920											
210	Terrain Ruggedness Index	SAGA GIS		.918											
191	Standard Deviation	SAGA GIS		.908											
200	Surface Ratio	ArcGIS with Python		.905			.366								
208	Terrain Ruggedness Index	SAGA GIS		.894											
156	Real Area	SAGA GIS		.856			.364								
221	Value Range	SAGA GIS		.846											-.344
174	Slope	SAGA GIS		.793			.514								
206	Terrain Ruggedness Index	ArcGIS with Python		.762			.325	.498							
3	Coefficient of variation	Diva-GIS		.750				-.527							
41	Easterness	SAGA GIS			.993										
34	Easterness	Quantum GIS			.993										
36	Easterness	Quantum GIS			.993										
27	Easterness	Benthic Terrain Modeler			.993										
29	Easterness	DEM Surface Tools			.993										
31	Easterness	ArcGIS with Python			.993										
30	Easterness	DEM Surface Tools			.986										
37	Easterness	SAGA GIS			.986										
33	Easterness	Landserf			.986										
28	Easterness	DEM Surface Tools			.976										
35	Easterness	Quantum GIS			.976										
40	Easterness	SAGA GIS			.976										
32	Easterness	Idrisi			.976										
38	Easterness	SAGA GIS			.972										

44	Easterness	SAGA GIS			.967										
48	Easterness	TNTmips			.957										
49	Easterness	TNTmips			.957										
46	Easterness	TNTmips			.955										
50	Easterness	uDig			.935										
45	Easterness	TNTmips			.929										
43	Easterness	SAGA GIS			.895										
88	Northerness	DEM Surface Tools				.991									
86	Northerness	Benthic Terrain Modeler				.991									
95	Northerness	Quantum GIS				.991									
90	Northerness	ArcGIS with Python				.991									
93	Northerness	Quantum GIS				.991									
100	Northerness	SAGA GIS				.991									
92	Northerness	Landserf				.982									
89	Northerness	DEM Surface Tools				.982									
96	Northerness	SAGA GIS				.982									
94	Northerness	Quantum GIS				.970									
87	Northerness	DEM Surface Tools				.970									
99	Northerness	SAGA GIS				.970									
91	Northerness	Idrisi				.969									
97	Northerness	SAGA GIS				.961									
103	Northerness	SAGA GIS				.955									
107	Northerness	TNTmips				.883									
108	Northerness	TNTmips				.882									
105	Northerness	TNTmips				.880									
102	Northerness	SAGA GIS				.854									
104	Northerness	TNTmips				.853									
109	Northerness	uDig				.812									
170	Slope	Quantum GIS		.413			.898								
56	Gradient	uDig		.413			.898								
166	Slope	ArcGIS with Python		.413			.898								
172	Slope	Quantum GIS		.413			.898								
177	Slope	SAGA GIS		.413			.898								
167	Slope	Diva-GIS		.414			.888								
173	Slope	SAGA GIS		.429			.862								
54	Gradient	uDig		.429			.862								
165	Slope	DEM Surface Tools		.429			.862								

169	Slope	Landserf		.429			.862								
176	Slope	SAGA GIS		.394			.794								
163	Slope	DEM Surface Tools		.394			.794								
171	Slope	Quantum GIS		.394			.794								
55	Gradient	uDig		.394			.794								
168	Slope	Idrisi		.389			.785								
63	Maximum Value	SAGA GIS						.992							
71	Mean Value	SAGA GIS						.992							
73	Mean Value	SAGA GIS						.991							
59	Maximum Value	ArcGIS with Python						.991							
72	Mean Value	SAGA GIS						.991							
76	Mean	ArcGIS with Python						.990							
75	Median	Diva-GIS						.989							
67	Minimum Value	ArcGIS with Python						.987							
80	Minimum Value	SAGA GIS						.986							
150	Profile Curvature	TNTmips							.906						
149	Profile Curvature	TNTmips							.902						
151	Profile Curvature	TNTmips							.900						
148	Profile Curvature	TNTmips							.886						
147	Profile Curvature	TNTmips							.799						
131	Plan Curvature	TNTmips								.892					
130	Plan Curvature	TNTmips								.892					
129	Plan Curvature	TNTmips								.891					
128	Plan Curvature	TNTmips								.880					
127	Plan Curvature	TNTmips								.781					
204	Tangential Curvature	Whitebox GAT									.878				
152	Profile Curvature	uDig									.807				
153	Profile Curvature	Whitebox GAT									.807				
133	Planar Curvature	uDig	-.486								.746				
132	Plan Curvature	Whitebox GAT	-.486								.746				
229	Vector Ruggedness Measure	SAGA GIS										.971			
227	Vector Ruggedness Measure	SAGA GIS										.963			
228	Vector Ruggedness Measure	SAGA GIS										.960			
219	Total Curvature	DEM Surface Tools											.903		
220	Total Curvature	Whitebox GAT											.829		
188	Slope	ArcGIS with Python											.548		
126	Plan Curvature	SAGA GIS												.573	

146	Profile Curvature	SAGA GIS											.569	
158	Representativeness	SAGA GIS					-.323							-.845
70	Mean of Residuals	Landserf												.365

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

Table C.8: Final PCA solution (rotated component matrix) for surface A3.

			Components													
ID	Terrain Attributes	Software	1	2	3	4	5	6	7	8	9	10	11	12	13	14
213	Topographic Position Index	Quantum GIS	.982													
12	General Curvature	SAGA GIS	.980													
157	Relative Deviation from Mean	ArcGIS with Python	-.958													
24	Deviation from Mean Value	SAGA GIS	.958													
19	Deviation from Mean Value	SAGA GIS	.956													
215	Topographic Position Index	SAGA GIS	.949													
13	General Curvature	SAGA GIS	.949													
9	General Curvature	ArcGIS with Python	.949													
5	Convergence Index	SAGA GIS	.935													
216	Topographic Position Index	SAGA GIS	.927													
20	Deviation from Mean Value	SAGA GIS	.927													
212	Topographic Position Index	ArcGIS with Python	.914													
10	General Curvature	SAGA GIS	.912													
53	General Curvature	DEM Surface Tools	.912													
69	Mean Curvature	Landserf	.912													
214	Topographic Position Index	SAGA GIS	.909													
18	Deviation from Mean Value	SAGA GIS	.907													
17	Deviation from Mean Elevation	Whitebox GAT	.907													
85	Morphometric Position Index	SAGA GIS	-.905													
61	Maximum Curvature	Idrisi	-.895													
78	Minimum Curvature	Idrisi	-.894													
15	General Curvature	SAGA GIS	.876													
122	Plan Curvature	SAGA GIS	.857													
117	Plan Curvature	ArcGIS with Python	.857													
137	Profile Curvature	ArcGIS with Python	-.837													
142	Profile Curvature	SAGA GIS	.837													
121	Plan Curvature	SAGA GIS	.828							.368						
79	Minimum Curvature	Landserf	.812													
62	Maximum Curvature	Landserf	.809													
124	Plan Curvature	SAGA GIS	.809													
22	Difference from Mean Elevation	Whitebox GAT	.808												.327	
141	Profile Curvature	SAGA GIS	.801						.363							
119	Plan Curvature	SAGA GIS	.781							.353						
6	Cross-Sectional Curvature	DEM Surface Tools	.772							.388						

139	Profile Curvature	SAGA GIS	.757						.340							
4	Convergence Index	SAGA GIS	.725												.478	
225	Variance	Benthic Terrain Modeler		.954												
190	Standard Deviation	ArcGIS with Python		.945												
198	Planimetric-to-Surface Ratio	Benthic Terrain Modeler		.945												
199	Surface Ratio	DEM Surface Tools		.945												
209	Terrain Ruggedness Index	SAGA GIS		.935												
154	Range	ArcGIS with Python		.931												
84	Melton Ruggedness Index	ArcGIS with Python		.931												
160	Roughness	Quantum GIS		.931												
192	Standard Deviation	SAGA GIS		.927												
218	Terrain Ruggedness Index	Whitebox GAT		.927												
161	Ruggedness Index	Quantum GIS		.927												
207	Terrain Ruggedness Index	Quantum GIS		.923												
193	Standard Deviation	SAGA GIS		.894												
200	Surface Ratio	ArcGIS with Python		.894			.347									
210	Terrain Ruggedness Index	SAGA GIS		.893												
191	Standard Deviation	SAGA GIS		.876												-.338
208	Terrain Ruggedness Index	SAGA GIS		.857												-.364
156	Real Area	SAGA GIS		.840												
221	Value Range	SAGA GIS		.796												-.428
3	Coefficient of variation	Diva-GIS		.780				-.502								
206	Terrain Ruggedness Index	ArcGIS with Python		.778				.492								
34	Easterness	Quantum GIS			.991											
31	Easterness	ArcGIS with Python			.991											
36	Easterness	Quantum GIS			.991											
29	Easterness	DEM Surface Tools			.991											
41	Easterness	SAGA GIS			.991											
27	Easterness	Benthic Terrain Modeler			.991											
33	Easterness	Landserf			.982											
37	Easterness	SAGA GIS			.982											
30	Easterness	DEM Surface Tools			.982											
32	Easterness	Idrisi			.965											
40	Easterness	SAGA GIS			.965											
28	Easterness	DEM Surface Tools			.965											
35	Easterness	Quantum GIS			.965											
38	Easterness	SAGA GIS			.959											
49	Easterness	TNTmips			.939											

48	Easterness	TNTmips			.939											
46	Easterness	TNTmips			.937											
44	Easterness	SAGA GIS			.904											
45	Easterness	TNTmips			.902											
50	Easterness	uDig			.901											
43	Easterness	SAGA GIS			.853											
93	Northerness	Quantum GIS				.990										
100	Northerness	SAGA GIS				.990										
86	Northerness	Benthic Terrain Modeler				.990										
88	Northerness	DEM Surface Tools				.990										
95	Northerness	Quantum GIS				.990										
90	Northerness	ArcGIS with Python				.990										
92	Northerness	Landserf				.979										
89	Northerness	DEM Surface Tools				.979										
96	Northerness	SAGA GIS				.979										
87	Northerness	DEM Surface Tools				.966										
94	Northerness	Quantum GIS				.966										
99	Northerness	SAGA GIS				.966										
91	Northerness	Idrisi				.966										
97	Northerness	SAGA GIS				.960										
103	Northerness	SAGA GIS				.928										
107	Northerness	TNTmips				.892										
108	Northerness	TNTmips				.891										
105	Northerness	TNTmips				.888										
104	Northerness	TNTmips				.861										
102	Northerness	SAGA GIS				.858										
109	Northerness	uDig				.754										
166	Slope	ArcGIS with Python		.342			.911									
172	Slope	Quantum GIS		.342			.911									
56	Gradient	uDig		.342			.911									
170	Slope	Quantum GIS		.342			.911									
177	Slope	SAGA GIS		.342			.911									
173	Slope	SAGA GIS		.359			.905									
169	Slope	Landserf		.359			.905									
165	Slope	DEM Surface Tools		.359			.905									
54	Gradient	uDig		.359			.905									
167	Slope	Diva-GIS		.341			.902									

184	Slope	TNTmips		.367			.589									
71	Mean Value	SAGA GIS						.994								
73	Mean Value	SAGA GIS						.994								
72	Mean Value	SAGA GIS						.993								
59	Maximum Value	ArcGIS with Python						.991								
76	Mean	ArcGIS with Python						.991								
63	Maximum Value	SAGA GIS						.989								
75	Median	Diva-GIS						.989								
67	Minimum Value	ArcGIS with Python						.986								
80	Minimum Value	SAGA GIS						.981								
180	Slope	SAGA GIS						.606								
150	Profile Curvature	TNTmips							.927							
149	Profile Curvature	TNTmips							.922							
151	Profile Curvature	TNTmips							.921							
148	Profile Curvature	TNTmips							.907							
147	Profile Curvature	TNTmips							.808							
130	Plan Curvature	TNTmips								.912						
129	Plan Curvature	TNTmips								.907						
131	Plan Curvature	TNTmips								.907						
128	Plan Curvature	TNTmips								.893						
127	Plan Curvature	TNTmips								.805						
171	Slope	Quantum GIS		.364			.518				.758					
163	Slope	DEM Surface Tools		.364			.518				.758					
55	Gradient	uDig		.364			.518				.758					
176	Slope	SAGA GIS		.364			.518				.758					
168	Slope	Idrisi		.366			.515				.753					
229	Vector Ruggedness Measure	SAGA GIS										.936				
227	Vector Ruggedness Measure	SAGA GIS										.932				
228	Vector Ruggedness Measure	SAGA GIS										.911				
152	Profile Curvature	uDig											.957			
153	Profile Curvature	Whitebox GAT											.957			
204	Tangential Curvature	Whitebox GAT											.821			
220	Total Curvature	Whitebox GAT												.885		
219	Total Curvature	DEM Surface Tools												.874		
126	Plan Curvature	SAGA GIS													.517	
146	Profile Curvature	SAGA GIS													.441	
158	Representativeness	SAGA GIS														-.879

70	Mean of Residuals	Landserf														.395

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

Table C.9: Final PCA solution (rotated component matrix) for surface A4.

			Components														
ID	**Terrain Attributes**	**Software**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
213	Topographic Position Index	Quantum GIS	.979														
12	General Curvature	SAGA GIS	.978														
24	Deviation from Mean Value	SAGA GIS	.963														
19	Deviation from Mean Value	SAGA GIS	.962														
157	Relative Deviation from Mean	ArcGIS with Python	-.959														
5	Convergence Index	SAGA GIS	.945														
215	Topographic Position Index	SAGA GIS	.943														
13	General Curvature	SAGA GIS	.943														
9	General Curvature	ArcGIS with Python	.943														
20	Deviation from Mean Value	SAGA GIS	.937														
216	Topographic Position Index	SAGA GIS	.935														
18	Deviation from Mean Value	SAGA GIS	.919														
214	Topographic Position Index	SAGA GIS	.918														
85	Morphometric Position Index	SAGA GIS	-.917														
212	Topographic Position Index	ArcGIS with Python	.915														
69	Mean Curvature	Landserf	.904														
10	General Curvature	SAGA GIS	.904														
53	General Curvature	DEM Surface Tools	.904														
17	Deviation from Mean Elevation	Whitebox GAT	.902														
61	Maximum Curvature	Idrisi	-.895														
78	Minimum Curvature	Idrisi	-.887														
15	General Curvature	SAGA GIS	.879														
122	Plan Curvature	SAGA GIS	.860														
117	Plan Curvature	ArcGIS with Python	.860														
22	Difference from Mean Elevation	Whitebox GAT	.842											.338			
142	Profile Curvature	SAGA GIS	.836														
121	Plan Curvature	SAGA GIS	.829								.325						
124	Plan Curvature	SAGA GIS	.811												-.357		
79	Minimum Curvature	Landserf	.807														
141	Profile Curvature	SAGA GIS	.800							.347							
62	Maximum Curvature	Landserf	.797														
119	Plan Curvature	SAGA GIS	.773														
4	Convergence Index	SAGA GIS	.751											.468			
139	Profile Curvature	SAGA GIS	.747							.347							

116	Plan Curvature	DEM Surface Tools	.325														
34	Easterness	Quantum GIS		.989													
31	Easterness	ArcGIS with Python		.989													
36	Easterness	Quantum GIS		.989													
29	Easterness	DEM Surface Tools		.989													
27	Easterness	Benthic Terrain Modeler		.989													
41	Easterness	SAGA GIS		.989													
30	Easterness	DEM Surface Tools		.975													
33	Easterness	Landserf		.975													
37	Easterness	SAGA GIS		.975													
35	Easterness	Quantum GIS		.957													
28	Easterness	DEM Surface Tools		.957													
40	Easterness	SAGA GIS		.957													
32	Easterness	Idrisi		.956													
38	Easterness	SAGA GIS		.944													
44	Easterness	SAGA GIS		.923													
48	Easterness	TNTmips		.918													
49	Easterness	TNTmips		.917													
46	Easterness	TNTmips		.914													
50	Easterness	uDig		.897													
45	Easterness	TNTmips		.871													
43	Easterness	SAGA GIS		.828													
88	Northerness	DEM Surface Tools			.987												
86	Northerness	Benthic Terrain Modeler			.987												
95	Northerness	Quantum GIS			.987												
90	Northerness	ArcGIS with Python			.987												
100	Northerness	SAGA GIS			.987												
93	Northerness	Quantum GIS			.987												
92	Northerness	Landserf			.972												
96	Northerness	SAGA GIS			.972												
89	Northerness	DEM Surface Tools			.972												
99	Northerness	SAGA GIS			.958												
94	Northerness	Quantum GIS			.958												
87	Northerness	DEM Surface Tools			.958												
91	Northerness	Idrisi			.957												
97	Northerness	SAGA GIS			.939												
103	Northerness	SAGA GIS			.928												
107	Northerness	TNTmips			.845												

108	Northerness	TNTmips			.843												
105	Northerness	TNTmips			.838												
102	Northerness	SAGA GIS			.815												
104	Northerness	TNTmips			.812												
109	Northerness	uDig			.737												
198	Planimetric-to-Surface Ratio	Benthic Terrain Modeler				.939											
199	Surface Ratio	DEM Surface Tools				.939											
209	Terrain Ruggedness Index	SAGA GIS				.929											
225	Variance	Benthic Terrain Modeler				.929											
190	Standard Deviation	ArcGIS with Python				.927											
218	Terrain Ruggedness Index	Whitebox GAT				.926											
161	Ruggedness Index	Quantum GIS				.926											
207	Terrain Ruggedness Index	Quantum GIS				.914											
154	Range	ArcGIS with Python				.908											
84	Melton Ruggedness Index	ArcGIS with Python				.908											
160	Roughness	Quantum GIS				.908											
192	Standard Deviation	SAGA GIS				.883											
210	Terrain Ruggedness Index	SAGA GIS				.865											
200	Surface Ratio	ArcGIS with Python				.834	.355										
193	Standard Deviation	SAGA GIS				.833											.376
208	Terrain Ruggedness Index	SAGA GIS				.808											.429
191	Standard Deviation	SAGA GIS				.803											.412
3	Coefficient of variation	Diva-GIS				.776		-.503									
206	Terrain Ruggedness Index	ArcGIS with Python				.769		.477									
156	Real Area	SAGA GIS				.762			.356								
169	Slope	Landserf				.330	.894										
165	Slope	DEM Surface Tools				.330	.894										
54	Gradient	uDig				.330	.894										
173	Slope	SAGA GIS				.330	.894										
56	Gradient	uDig					.893										
170	Slope	Quantum GIS					.893										
166	Slope	ArcGIS with Python					.893										
172	Slope	Quantum GIS					.893										
177	Slope	SAGA GIS					.893										
167	Slope	Diva-GIS					.891										
71	Mean Value	SAGA GIS						.993									
73	Mean Value	SAGA GIS						.992									

72	Mean Value	SAGA GIS						.991									
59	Maximum Value	ArcGIS with Python						.988									
76	Mean	ArcGIS with Python						.988									
63	Maximum Value	SAGA GIS						.987									
75	Median	Diva-GIS						.987									
67	Minimum Value	ArcGIS with Python						.983									
80	Minimum Value	SAGA GIS						.982									
171	Slope	Quantum GIS					.334		.887								
163	Slope	DEM Surface Tools					.334		.887								
55	Gradient	uDig					.334		.887								
176	Slope	SAGA GIS					.334		.887								
168	Slope	Idrisi					.330		.884								
52	Fractal Dimension	Idrisi				.453	.395		.754								
180	Slope	SAGA GIS							.343								
150	Profile Curvature	TNTmips								.930							
151	Profile Curvature	TNTmips								.925							
149	Profile Curvature	TNTmips								.923							
148	Profile Curvature	TNTmips								.910							
147	Profile Curvature	TNTmips								.795							
130	Plan Curvature	TNTmips									.925						
129	Plan Curvature	TNTmips									.920						
131	Plan Curvature	TNTmips									.919						
128	Plan Curvature	TNTmips									.903						
127	Plan Curvature	TNTmips									.790						
204	Tangential Curvature	Whitebox GAT										.898					
153	Profile Curvature	Whitebox GAT										.868					
152	Profile Curvature	uDig										.868					
133	Planar Curvature	uDig	-.493									.787					
132	Plan Curvature	Whitebox GAT	-.493									.787					
229	Vector Ruggedness Measure	SAGA GIS											.961				
227	Vector Ruggedness Measure	SAGA GIS											.949				
228	Vector Ruggedness Measure	SAGA GIS											.935				
146	Profile Curvature	SAGA GIS												.500			
126	Plan Curvature	SAGA GIS												.490			
178	Slope	SAGA GIS	.491												.763		
179	Slope	SAGA GIS	.491												.763		
185	Slope	TNTmips				.326	.470									.777	

182	Slope	TNTmips				.330	.460									.759	
184	Slope	TNTmips				.326	.464									.741	
158	Representativeness	SAGA GIS															.859
70	Mean of Residuals	Landserf															-.453

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

Table C.10: Final PCA solution (rotated component matrix) for surface A5.

			Components								
ID	**Terrain Attributes**	**Software**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
213	Topographic Position Index	Quantum GIS	.986								
12	General Curvature	SAGA GIS	.984								
24	Deviation from Mean Value	SAGA GIS	.971								
19	Deviation from Mean Value	SAGA GIS	.968								
157	Relative Deviation from Mean	ArcGIS with Python	-.963								
215	Topographic Position Index	SAGA GIS	.949								
9	General Curvature	ArcGIS with Python	.949								
13	General Curvature	SAGA GIS	.949								
5	Convergence Index	SAGA GIS	.947								
20	Deviation from Mean Value	SAGA GIS	.944								
216	Topographic Position Index	SAGA GIS	.943								
18	Deviation from Mean Value	SAGA GIS	.926								
214	Topographic Position Index	SAGA GIS	.926								
85	Morphometric Position Index	SAGA GIS	-.923								
212	Topographic Position Index	ArcGIS with Python	.914								
17	Deviation from Mean Elevation	Whitebox GAT	.908								
61	Maximum Curvature	Idrisi	-.903								
78	Minimum Curvature	Idrisi	-.900								
53	General Curvature	DEM Surface Tools	.897								
10	General Curvature	SAGA GIS	.897								
69	Mean Curvature	Landserf	.897								
122	Plan Curvature	SAGA GIS	.880								
117	Plan Curvature	ArcGIS with Python	.880								
15	General Curvature	SAGA GIS	.872								
22	Difference from Mean Elevation	Whitebox GAT	.867								
121	Plan Curvature	SAGA GIS	.849								
142	Profile Curvature	SAGA GIS	.839								
124	Plan Curvature	SAGA GIS	.801								
141	Profile Curvature	SAGA GIS	.800								
79	Minimum Curvature	Landserf	.798								
62	Maximum Curvature	Landserf	.793								
186	Slope	uDig	.680			.359					
36	Easterness	Quantum GIS		.989							
29	Easterness	DEM Surface Tools		.989							

27	Easterness	Benthic Terrain Modeler		.989							
31	Easterness	ArcGIS with Python		.989							
41	Easterness	SAGA GIS		.989							
34	Easterness	Quantum GIS		.989							
30	Easterness	DEM Surface Tools		.971							
37	Easterness	SAGA GIS		.971							
33	Easterness	Landserf		.971							
32	Easterness	Idrisi		.946							
28	Easterness	DEM Surface Tools		.946							
35	Easterness	Quantum GIS		.946							
40	Easterness	SAGA GIS		.946							
38	Easterness	SAGA GIS		.928							
48	Easterness	TNTmips		.915							
49	Easterness	TNTmips		.914							
46	Easterness	TNTmips		.909							
44	Easterness	SAGA GIS		.900							
50	Easterness	uDig		.883							
45	Easterness	TNTmips		.856							
43	Easterness	SAGA GIS		.810							
93	Northerness	Quantum GIS			.988						
100	Northerness	SAGA GIS			.988						
90	Northerness	ArcGIS with Python			.988						
95	Northerness	Quantum GIS			.988						
88	Northerness	DEM Surface Tools			.988						
86	Northerness	Benthic Terrain Modeler			.988						
92	Northerness	Landserf			.967						
89	Northerness	DEM Surface Tools			.967						
96	Northerness	SAGA GIS			.967						
87	Northerness	DEM Surface Tools			.952						
94	Northerness	Quantum GIS			.952						
99	Northerness	SAGA GIS			.952						
91	Northerness	Idrisi			.951						
97	Northerness	SAGA GIS			.928						
103	Northerness	SAGA GIS			.900						
107	Northerness	TNTmips			.812						
108	Northerness	TNTmips			.809						
102	Northerness	SAGA GIS			.808						
104	Northerness	TNTmips			.772						

109	Northerness	uDig			.698						
199	Surface Ratio	DEM Surface Tools				.947					
198	Planimetric-to-Surface Ratio	Benthic Terrain Modeler				.947					
190	Standard Deviation	ArcGIS with Python				.934					
225	Variance	Benthic Terrain Modeler				.932					
218	Terrain Ruggedness Index	Whitebox GAT				.920					
161	Ruggedness Index	Quantum GIS				.920					
160	Roughness	Quantum GIS				.912					
84	Melton Ruggedness Index	ArcGIS with Python				.912					
154	Range	ArcGIS with Python				.912					
207	Terrain Ruggedness Index	Quantum GIS				.901					
209	Terrain Ruggedness Index	SAGA GIS				.890					
3	Coefficient of variation	Diva-GIS				.842		-.398			
192	Standard Deviation	SAGA GIS				.820					
206	Terrain Ruggedness Index	ArcGIS with Python				.818		.400			
200	Surface Ratio	ArcGIS with Python				.811	.393				
156	Real Area	SAGA GIS				.734			.408		
172	Slope	Quantum GIS					.896				
166	Slope	ArcGIS with Python					.896				
56	Gradient	uDig					.896				
170	Slope	Quantum GIS					.896				
177	Slope	SAGA GIS					.896				
167	Slope	Diva-GIS					.892				
165	Slope	DEM Surface Tools					.892				
54	Gradient	uDig					.892				
169	Slope	Landserf					.892				
173	Slope	SAGA GIS					.892				
195	Standard Deviation of Slope	ArcGIS with Python					-.362				
72	Mean Value	SAGA GIS						.994			
73	Mean Value	SAGA GIS						.994			
71	Mean Value	SAGA GIS						.994			
76	Mean	ArcGIS with Python						.990			
75	Median	Diva-GIS						.986			
59	Maximum Value	ArcGIS with Python						.983			
67	Minimum Value	ArcGIS with Python						.982			
63	Maximum Value	SAGA GIS						.971			
80	Minimum Value	SAGA GIS						.970			

168	Slope	Idrisi							.901		
176	Slope	SAGA GIS							.900		
171	Slope	Quantum GIS							.900		
163	Slope	DEM Surface Tools							.900		
55	Gradient	uDig							.900		
52	Fractal Dimension	Idrisi				.408	.363		.810		
150	Profile Curvature	TNTmips								.952	
151	Profile Curvature	TNTmips								.945	
149	Profile Curvature	TNTmips								.942	
148	Profile Curvature	TNTmips								.924	
147	Profile Curvature	TNTmips								.791	
153	Profile Curvature	Whitebox GAT									.973
152	Profile Curvature	uDig									.973

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

Table C.11: Final PCA solution (rotated component matrix) for surface A6.

			Components											
ID	Terrain Attributes	Software	1	2	3	4	5	6	7	8	9	10	11	12
213	Topographic Position Index	Quantum GIS	.983											
12	General Curvature	SAGA GIS	.981											
19	Deviation from Mean Value	SAGA GIS	.972											
24	Deviation from Mean Value	SAGA GIS	.971											
157	Relative Deviation from Mean	ArcGIS with Python	-.967											
20	Deviation from Mean Value	SAGA GIS	.949											
5	Convergence Index	SAGA GIS	.945											
216	Topographic Position Index	SAGA GIS	.945											
215	Topographic Position Index	SAGA GIS	.942											
9	General Curvature	ArcGIS with Python	.942											
13	General Curvature	SAGA GIS	.942											
18	Deviation from Mean Value	SAGA GIS	.931											
214	Topographic Position Index	SAGA GIS	.928											
212	Topographic Position Index	ArcGIS with Python	.927											
85	Morphometric Position Index	SAGA GIS	-.923											
17	Deviation from Mean Elevation	Whitebox GAT	.907											
61	Maximum Curvature	Idrisi	-.896											
78	Minimum Curvature	Idrisi	-.895											
69	Mean Curvature	Landserf	.891											
10	General Curvature	SAGA GIS	.891											
53	General Curvature	DEM Surface Tools	.891											
22	Difference from Mean Elevation	Whitebox GAT	.875											
15	General Curvature	SAGA GIS	.865											
117	Plan Curvature	ArcGIS with Python	.857											
122	Plan Curvature	SAGA GIS	.857											
142	Profile Curvature	SAGA GIS	.848											
141	Profile Curvature	SAGA GIS	.826											
121	Plan Curvature	SAGA GIS	.822											
79	Minimum Curvature	Landserf	.793											
124	Plan Curvature	SAGA GIS	.791										-.372	
62	Maximum Curvature	Landserf	.785											
139	Profile Curvature	SAGA GIS	.750											
4	Convergence Index	SAGA GIS	.724											
186	Slope	uDig	.692											

144	Profile Curvature	SAGA GIS	.602											
41	Easterness	SAGA GIS		.987										
31	Easterness	ArcGIS with Python		.987										
36	Easterness	Quantum GIS		.987										
29	Easterness	DEM Surface Tools		.987										
27	Easterness	Benthic Terrain Modeler		.987										
34	Easterness	Quantum GIS		.987										
30	Easterness	DEM Surface Tools		.965										
37	Easterness	SAGA GIS		.965										
33	Easterness	Landserf		.965										
28	Easterness	DEM Surface Tools		.938										
35	Easterness	Quantum GIS		.938										
40	Easterness	SAGA GIS		.938										
32	Easterness	Idrisi		.938										
38	Easterness	SAGA GIS		.920										
49	Easterness	TNTmips		.901										
48	Easterness	TNTmips		.901										
46	Easterness	TNTmips		.897										
50	Easterness	uDig		.869										
44	Easterness	SAGA GIS		.869										
45	Easterness	TNTmips		.829										
43	Easterness	SAGA GIS		.785										
90	Northerness	ArcGIS with Python			.986									
86	Northerness	Benthic Terrain Modeler			.986									
88	Northerness	DEM Surface Tools			.986									
95	Northerness	Quantum GIS			.986									
100	Northerness	SAGA GIS			.986									
93	Northerness	Quantum GIS			.986									
89	Northerness	DEM Surface Tools			.963									
96	Northerness	SAGA GIS			.963									
92	Northerness	Landserf			.963									
87	Northerness	DEM Surface Tools			.942									
94	Northerness	Quantum GIS			.942									
99	Northerness	SAGA GIS			.942									
91	Northerness	Idrisi			.942									
97	Northerness	SAGA GIS			.915									
103	Northerness	SAGA GIS			.872									
107	Northerness	TNTmips			.805									

108	Northerness	TNTmips			.802									
102	Northerness	SAGA GIS			.797									
104	Northerness	TNTmips			.759									
109	Northerness	uDig			.632									
199	Surface Ratio	DEM Surface Tools				.930								
198	Planimetric-to-Surface Ratio	Benthic Terrain Modeler				.930								
161	Ruggedness Index	Quantum GIS				.908								
218	Terrain Ruggedness Index	Whitebox GAT				.908								
190	Standard Deviation	ArcGIS with Python				.903								
209	Terrain Ruggedness Index	SAGA GIS				.884							-.349	
207	Terrain Ruggedness Index	Quantum GIS				.883							-.334	
160	Roughness	Quantum GIS				.879								
84	Melton Ruggedness Index	ArcGIS with Python				.879								
154	Range	ArcGIS with Python				.879								
3	Coefficient of variation	Diva-GIS				.805		-.419						
192	Standard Deviation	SAGA GIS				.789								
206	Terrain Ruggedness Index	ArcGIS with Python				.783		.407						
210	Terrain Ruggedness Index	SAGA GIS				.775							-.386	
200	Surface Ratio	ArcGIS with Python				.734	.420							
193	Standard Deviation	SAGA GIS				.701						-.392		
156	Real Area	SAGA GIS				.654			.433					
56	Gradient	uDig					.907							
172	Slope	Quantum GIS					.907							
166	Slope	ArcGIS with Python					.907							
177	Slope	SAGA GIS					.907							
170	Slope	Quantum GIS					.907							
167	Slope	Diva-GIS					.899							
173	Slope	SAGA GIS					.875							
54	Gradient	uDig					.875							
165	Slope	DEM Surface Tools					.875							
169	Slope	Landserf					.875							
185	Slope	TNTmips					.615							
182	Slope	TNTmips					.601							
184	Slope	TNTmips					.579							
72	Mean Value	SAGA GIS						.991						
73	Mean Value	SAGA GIS						.990						
71	Mean Value	SAGA GIS						.988						

76	Mean	ArcGIS with Python						.985						
59	Maximum Value	ArcGIS with Python						.979						
75	Median	Diva-GIS						.977						
67	Minimum Value	ArcGIS with Python						.963						
63	Maximum Value	SAGA GIS						.957						
80	Minimum Value	SAGA GIS						.946						
180	Slope	SAGA GIS						.460						
176	Slope	SAGA GIS							.931					
55	Gradient	uDig							.931					
163	Slope	DEM Surface Tools							.931					
171	Slope	Quantum GIS							.931					
168	Slope	Idrisi							.928					
52	Fractal Dimension	Idrisi					.339		.853					
150	Profile Curvature	TNTmips								.938				
151	Profile Curvature	TNTmips								.933				
149	Profile Curvature	TNTmips								.924				
148	Profile Curvature	TNTmips								.912				
147	Profile Curvature	TNTmips								.770				
130	Plan Curvature	TNTmips									.938			
131	Plan Curvature	TNTmips									.931			
129	Plan Curvature	TNTmips									.927			
128	Plan Curvature	TNTmips									.910			
127	Plan Curvature	TNTmips									.764			
229	Vector Ruggedness Measure	SAGA GIS										.932		
228	Vector Ruggedness Measure	SAGA GIS										.920		
227	Vector Ruggedness Measure	SAGA GIS										.894		
205	Vector Ruggedness Measure	Benthic Terrain Modeler										.653		
226	Vector Ruggedness Measure	ArcGIS with Python										.653		
178	Slope	SAGA GIS	.486										.743	
179	Slope	SAGA GIS	.486										.743	
153	Profile Curvature	Whitebox GAT												.944
152	Profile Curvature	uDig												.944
204	Tangential Curvature	Whitebox GAT												.749

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

Table C.12: Final PCA solution (rotated component matrix) for surface A7.

			Components															
ID	**Terrain Attributes**	**Software**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
213	Topographic Position Index	Quantum GIS	.983															
12	General Curvature	SAGA GIS	.977															
157	Relative Deviation from Mean	ArcGIS with Python	-.966															
19	Deviation from Mean Value	SAGA GIS	.964															
24	Deviation from Mean Value	SAGA GIS	.962															
215	Topographic Position Index	SAGA GIS	.959															
13	General Curvature	SAGA GIS	.959															
9	General Curvature	ArcGIS with Python	.959															
5	Convergence Index	SAGA GIS	.945															
20	Deviation from Mean Value	SAGA GIS	.936															
216	Topographic Position Index	SAGA GIS	.931															
212	Topographic Position Index	ArcGIS with Python	.925															
17	Deviation from Mean Elevation	Whitebox GAT	.924															
61	Maximum Curvature	Idrisi	-.917															
18	Deviation from Mean Value	SAGA GIS	.915															
85	Morphometric Position Index	SAGA GIS	-.912															
214	Topographic Position Index	SAGA GIS	.912															
78	Minimum Curvature	Idrisi	-.908															
122	Plan Curvature	SAGA GIS	.875															
117	Plan Curvature	ArcGIS with Python	.875															
142	Profile Curvature	SAGA GIS	.868															
22	Difference from Mean Elevation	Whitebox GAT	.865												.338			
15	General Curvature	SAGA GIS	.861										-.340					
69	Mean Curvature	Landserf	.852															
10	General Curvature	SAGA GIS	.852															
53	General Curvature	DEM Surface Tools	.852															
121	Plan Curvature	SAGA GIS	.814															
141	Profile Curvature	SAGA GIS	.814															
124	Plan Curvature	SAGA GIS	.784										-.437					
79	Minimum Curvature	Landserf	.755															
62	Maximum Curvature	Landserf	.750															
186	Slope	uDig	.673				.394											
144	Profile Curvature	SAGA GIS	.595													.370		

31	Easterness	ArcGIS with Python		.986														
36	Easterness	Quantum GIS		.986														
27	Easterness	Benthic Terrain Modeler		.986														
29	Easterness	DEM Surface Tools		.986														
34	Easterness	Quantum GIS		.986														
41	Easterness	SAGA GIS		.986														
37	Easterness	SAGA GIS		.962														
33	Easterness	Landserf		.962														
30	Easterness	DEM Surface Tools		.962														
32	Easterness	Idrisi		.932														
40	Easterness	SAGA GIS		.932														
28	Easterness	DEM Surface Tools		.932														
35	Easterness	Quantum GIS		.932														
38	Easterness	SAGA GIS		.906														
48	Easterness	TNTmips		.897														
49	Easterness	TNTmips		.897														
46	Easterness	TNTmips		.891														
50	Easterness	uDig		.860														
44	Easterness	SAGA GIS		.846														
45	Easterness	TNTmips		.815														
43	Easterness	SAGA GIS		.773														
93	Northerness	Quantum GIS			.987													
95	Northerness	Quantum GIS			.987													
100	Northerness	SAGA GIS			.987													
88	Northerness	DEM Surface Tools			.987													
90	Northerness	ArcGIS with Python			.987													
86	Northerness	Benthic Terrain Modeler			.987													
89	Northerness	DEM Surface Tools			.958													
92	Northerness	Landserf			.958													
96	Northerness	SAGA GIS			.958													
91	Northerness	Idrisi			.939													
99	Northerness	SAGA GIS			.939													
87	Northerness	DEM Surface Tools			.939													
94	Northerness	Quantum GIS			.939													
97	Northerness	SAGA GIS			.908													
103	Northerness	SAGA GIS			.853													
102	Northerness	SAGA GIS			.774													
104	Northerness	TNTmips			.691													

109	Northerness	uDig			.588													
198	Planimetric-to-Surface Ratio	Benthic Terrain Modeler				.936												
199	Surface Ratio	DEM Surface Tools				.936												
218	Terrain Ruggedness Index	Whitebox GAT				.915												
161	Ruggedness Index	Quantum GIS				.915												
190	Standard Deviation	ArcGIS with Python				.914												
154	Range	ArcGIS with Python				.894												
84	Melton Ruggedness Index	ArcGIS with Python				.894												
160	Roughness	Quantum GIS				.894												
209	Terrain Ruggedness Index	SAGA GIS				.887							-.349					
207	Terrain Ruggedness Index	Quantum GIS				.883							-.345					
3	Coefficient of variation	Diva-GIS				.817	-.410											
206	Terrain Ruggedness Index	ArcGIS with Python				.793	.441											
192	Standard Deviation	SAGA GIS				.787												
210	Terrain Ruggedness Index	SAGA GIS				.780							-.376					
200	Surface Ratio	ArcGIS with Python				.716		.328								-.365		
193	Standard Deviation	SAGA GIS				.695												
73	Mean Value	SAGA GIS					.992											
72	Mean Value	SAGA GIS					.991											
71	Mean Value	SAGA GIS					.990											
76	Mean	ArcGIS with Python					.974											
75	Median	Diva-GIS					.960											
67	Minimum Value	ArcGIS with Python					.954											
59	Maximum Value	ArcGIS with Python					.938											
80	Minimum Value	SAGA GIS					.928											
63	Maximum Value	SAGA GIS					.916											
180	Slope	SAGA GIS					.524											
170	Slope	Quantum GIS						.936										
177	Slope	SAGA GIS						.936										
56	Gradient	uDig						.936										
166	Slope	ArcGIS with Python						.936										
172	Slope	Quantum GIS						.936										
167	Slope	Diva-GIS						.932										
195	Standard Deviation of Slope	ArcGIS with Python						-.339										
168	Slope	Idrisi							.951									
171	Slope	Quantum GIS							.949									
163	Slope	DEM Surface Tools							.949									

55	Gradient	uDig							.949									
176	Slope	SAGA GIS							.949									
52	Fractal Dimension	Idrisi							.891									
150	Profile Curvature	TNTmips								.945								
151	Profile Curvature	TNTmips								.937								
149	Profile Curvature	TNTmips								.933								
148	Profile Curvature	TNTmips								.914								
147	Profile Curvature	TNTmips								.769	.326							
130	Plan Curvature	TNTmips									.943							
131	Plan Curvature	TNTmips									.935							
129	Plan Curvature	TNTmips									.931							
128	Plan Curvature	TNTmips									.911							
127	Plan Curvature	TNTmips								.332	.760							
229	Vector Ruggedness Measure	SAGA GIS										.975						
227	Vector Ruggedness Measure	SAGA GIS										.959						
228	Vector Ruggedness Measure	SAGA GIS										.906						
178	Slope	SAGA GIS	.480										.776					
179	Slope	SAGA GIS	.480										.776					
185	Slope	TNTmips						.326						.884				
182	Slope	TNTmips												.845				
184	Slope	TNTmips						.364						.814				
11	General Curvature	SAGA GIS	.488												.720			
120	Plan Curvature	SAGA GIS	.383												.639			
226	Vector Ruggedness Measure	ArcGIS with Python										.342				.810		
205	Vector Ruggedness Measure	Benthic Terrain Modeler										.342				.810		
116	Plan Curvature	DEM Surface Tools	.349														.877	
202	Tangential Curvature	DEM Surface Tools	.445														.860	
153	Profile Curvature	Whitebox GAT																.859
204	Tangential Curvature	Whitebox GAT																.805

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

Table C.13: Final PCA solution (rotated component matrix) for surface A8.

			Components													
ID	Terrain Attributes	Software	1	2	3	4	5	6	7	8	9	10	11	12	13	14
213	Topographic Position Index	Quantum GIS	.985													
12	General Curvature	SAGA GIS	.976													
157	Relative Deviation from Mean	ArcGIS with Python	-.970													
215	Topographic Position Index	SAGA GIS	.967													
13	General Curvature	SAGA GIS	.967													
9	General Curvature	ArcGIS with Python	.967													
19	Deviation from Mean Value	SAGA GIS	.963													
24	Deviation from Mean Value	SAGA GIS	.961													
5	Convergence Index	SAGA GIS	.949													
20	Deviation from Mean Value	SAGA GIS	.934													
17	Deviation from Mean Elevation	Whitebox GAT	.932													
212	Topographic Position Index	ArcGIS with Python	.930													
216	Topographic Position Index	SAGA GIS	.928													
61	Maximum Curvature	Idrisi	-.923													
78	Minimum Curvature	Idrisi	-.921													
18	Deviation from Mean Value	SAGA GIS	.913													
85	Morphometric Position Index	SAGA GIS	-.912													
214	Topographic Position Index	SAGA GIS	.908										.334			
117	Plan Curvature	ArcGIS with Python	.889													
122	Plan Curvature	SAGA GIS	.889													
142	Profile Curvature	SAGA GIS	.878													
22	Difference from Mean Elevation	Whitebox GAT	.873										.369			
15	General Curvature	SAGA GIS	.857													
69	Mean Curvature	Landserf	.826													
53	General Curvature	DEM Surface Tools	.826													
10	General Curvature	SAGA GIS	.826													
121	Plan Curvature	SAGA GIS	.821													
141	Profile Curvature	SAGA GIS	.815													
124	Plan Curvature	SAGA GIS	.775													
62	Maximum Curvature	Landserf	.728													
186	Slope	uDig	.726			.323										
79	Minimum Curvature	Landserf	.725													
144	Profile Curvature	SAGA GIS	.612													
31	Easterness	ArcGIS with Python		.984												

27	Easterness	Benthic Terrain Modeler		.984												
34	Easterness	Quantum GIS		.984												
36	Easterness	Quantum GIS		.984												
29	Easterness	DEM Surface Tools		.984												
41	Easterness	SAGA GIS		.984												
33	Easterness	Landserf		.955												
30	Easterness	DEM Surface Tools		.955												
37	Easterness	SAGA GIS		.955												
28	Easterness	DEM Surface Tools		.917												
35	Easterness	Quantum GIS		.917												
40	Easterness	SAGA GIS		.917												
32	Easterness	Idrisi		.917												
38	Easterness	SAGA GIS		.892												
48	Easterness	TNTmips		.888												
49	Easterness	TNTmips		.888												
46	Easterness	TNTmips		.881												
50	Easterness	uDig		.849												
44	Easterness	SAGA GIS		.846												
45	Easterness	TNTmips		.793												
43	Easterness	SAGA GIS		.753												
93	Northerness	Quantum GIS			.990											
100	Northerness	SAGA GIS			.990											
95	Northerness	Quantum GIS			.990											
88	Northerness	DEM Surface Tools			.990											
90	Northerness	ArcGIS with Python			.990											
86	Northerness	Benthic Terrain Modeler			.990											
89	Northerness	DEM Surface Tools			.957											
96	Northerness	SAGA GIS			.957											
92	Northerness	Landserf			.957											
91	Northerness	Idrisi			.924											
99	Northerness	SAGA GIS			.924											
87	Northerness	DEM Surface Tools			.924											
94	Northerness	Quantum GIS			.924											
97	Northerness	SAGA GIS			.888											
103	Northerness	SAGA GIS			.839											
102	Northerness	SAGA GIS			.754											
72	Mean Value	SAGA GIS				.993										
73	Mean Value	SAGA GIS				.991										

71	Mean Value	SAGA GIS				.985										
76	Mean	ArcGIS with Python				.970										
68	Mean	Diva-GIS				.970										
75	Median	Diva-GIS				.948										
59	Maximum Value	ArcGIS with Python				.934										
60	Maximum Value	Diva-GIS				.934										
67	Minimum Value	ArcGIS with Python				.918										
63	Maximum Value	SAGA GIS				.893										
80	Minimum Value	SAGA GIS				.884										
180	Slope	SAGA GIS				.451										
146	Profile Curvature	SAGA GIS				.353										
160	Roughness	Quantum GIS					.966									
155	Range	Diva-GIS					.966									
84	Melton Ruggedness Index	ArcGIS with Python					.966									
154	Range	ArcGIS with Python					.966									
190	Standard Deviation	ArcGIS with Python					.946									
206	Terrain Ruggedness Index	ArcGIS with Python				.372	.881									
3	Coefficient of variation	Diva-GIS				-.346	.877									
192	Standard Deviation	SAGA GIS					.805									
200	Surface Ratio	ArcGIS with Python					.751	.374								
158	Representativeness	SAGA GIS					-.332									
166	Slope	ArcGIS with Python						.955								
172	Slope	Quantum GIS						.955								
187	Slope	Whitebox GAT						.955								
56	Gradient	uDig						.955								
177	Slope	SAGA GIS						.955								
170	Slope	Quantum GIS						.955								
167	Slope	Diva-GIS						.945								
195	Standard Deviation of Slope	ArcGIS with Python						-.348								
163	Slope	DEM Surface Tools							.966							
55	Gradient	uDig							.966							
171	Slope	Quantum GIS							.966							
176	Slope	SAGA GIS							.966							
168	Slope	Idrisi							.964							
52	Fractal Dimension	Idrisi							.901							
150	Profile Curvature	TNTmips								.939						
151	Profile Curvature	TNTmips								.931						

149	Profile Curvature	TNTmips								.926						
148	Profile Curvature	TNTmips								.902						
147	Profile Curvature	TNTmips								.755						
130	Plan Curvature	TNTmips									.945					
131	Plan Curvature	TNTmips									.940					
129	Plan Curvature	TNTmips									.928					
128	Plan Curvature	TNTmips									.914					
127	Plan Curvature	TNTmips								.376	.718					
229	Vector Ruggedness Measure	SAGA GIS										.917				
228	Vector Ruggedness Measure	SAGA GIS										.908				
227	Vector Ruggedness Measure	SAGA GIS										.840				
226	Vector Ruggedness Measure	ArcGIS with Python										.692				
205	Vector Ruggedness Measure	Benthic Terrain Modeler										.692				
11	General Curvature	SAGA GIS	.432										.783			
140	Profile Curvature	SAGA GIS	.350										.658			
120	Plan Curvature	SAGA GIS	.331										.577			
153	Profile Curvature	Whitebox GAT												.972		
152	Profile Curvature	uDig												.972		
204	Tangential Curvature	Whitebox GAT												.799		
185	Slope	TNTmips						.330							.870	
182	Slope	TNTmips													.839	
184	Slope	TNTmips						.418							.755	
202	Tangential Curvature	DEM Surface Tools	.437													.825
116	Plan Curvature	DEM Surface Tools	.323													.823

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

Appendix D: Additional Information on Maps and Models Performance (Chapter 4)

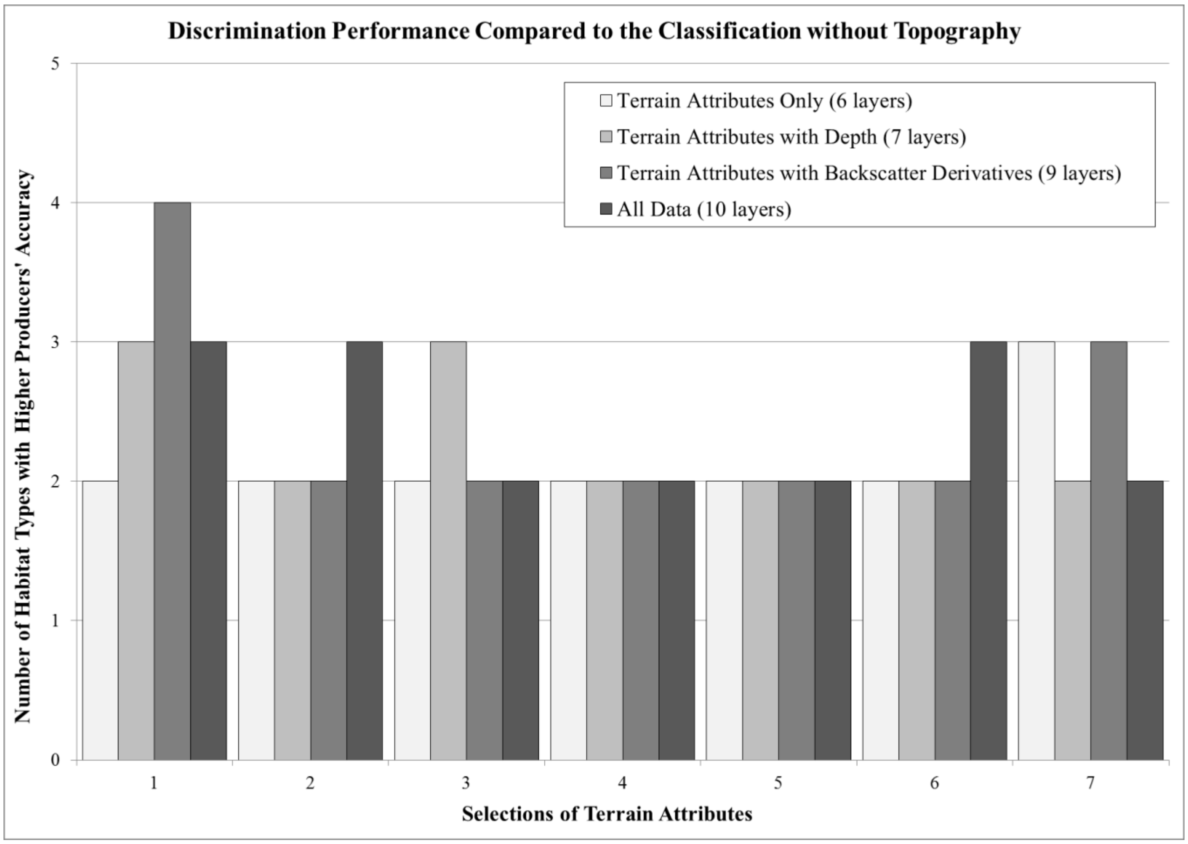

Figure D.1: Comparison of the discrimination ability of the computed classifications with that of the classification computed using only bathymetry and the backscatter derivatives, based on the number of bottom types (maximum possible of 5) that were better discriminated.

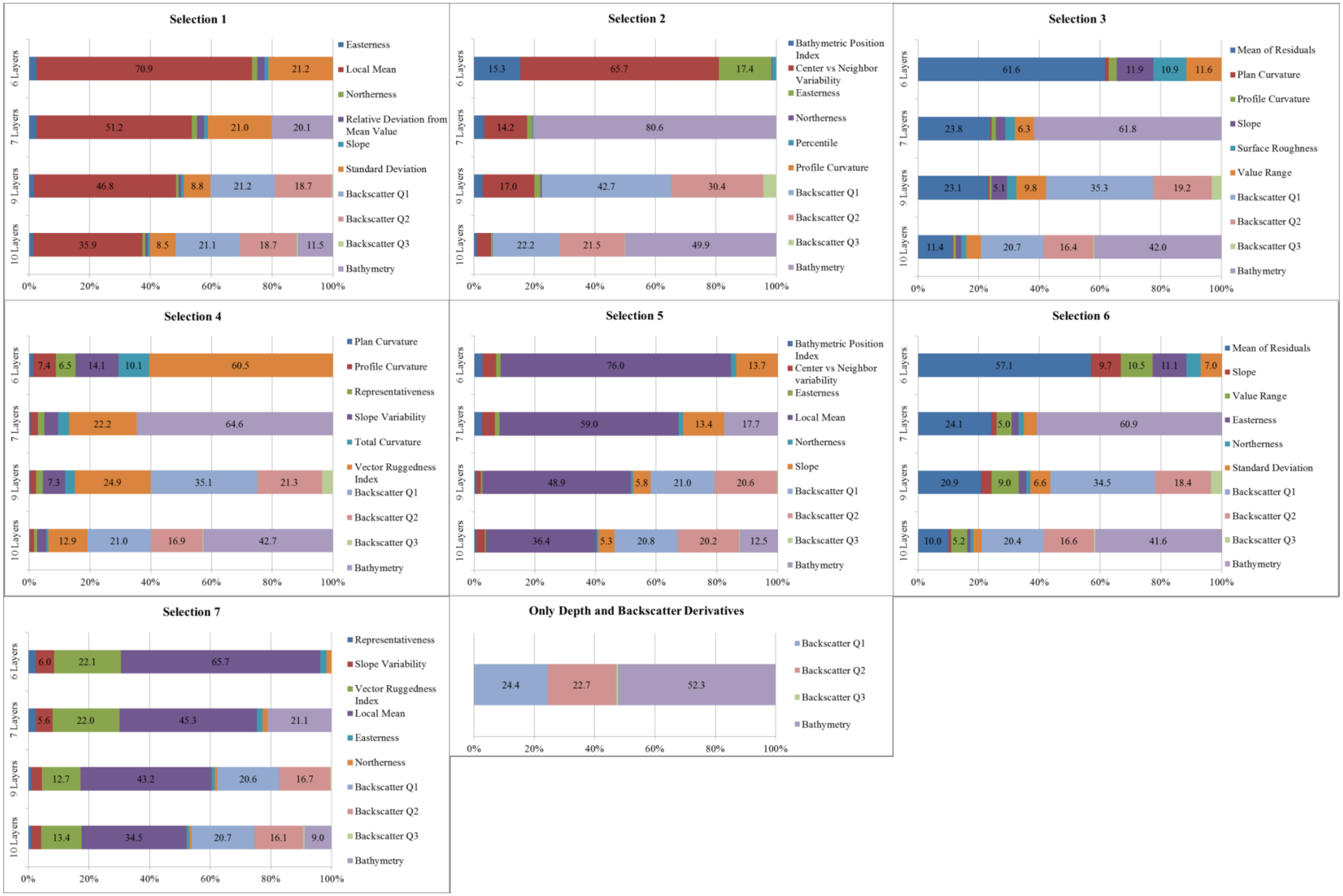

Figure D.2: Percentage of variable contribution for the 29 MaxEnt models. Only contributions greater than 5% are labeled.

Appendix E: Detailed Results (Chapter 5)

E.1 Spatial Similarity

E.1.1 Full Extent

When looking at average correlations between the bathymetric and terrain attributes surfaces altered by heave and their respective reference surfaces, average correlation coefficients indicated that finer-scale altered surfaces are more dissimilar to their reference surfaces than coarser-scale altered surfaces. Topographic position had the lowest correlation coefficients, followed by easterness, northerness, rugosity, slope, bathymetry and topographic mean. A summary of the modelling results for spatial similarity is shown in Table E.1, and proportions of the variance explained by the significant models are presented in Table E.2. The modelling of the change in correlation as a function of level of induced heave artefact yielded significant models only for the five scales of easterness and topographic position, for the four finest scales of northerness and slope, and for the two finest scales of rugosity (Table E.1). The equations modelled adequately the measured relationships (average r^2 of 0.866, Table E.2), and all described a loss in spatial similarity with increasing heave amplitude.

Table E.1: Summary of the modelling results for spatial similarity. Grey cells indicate that the models were not significant based on the F statistic. Grey cells with a symbol (*) indicate that the equations were significant based on the F statistic, but that the rate of change was not significant based on the t test. White cells indicate significant model, and the sign within it indicate whether the equations had a positive or negative rate of change.

	Bathymetry					Topo. Mean					Easterness					Northerness					Slope					Rugosity					Topo. Position				
	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100
Heave											-	-	-	-	-	-	-	-	-		-	-	-	-	*	-	-	*		*	-	-	-	-	-
Pitch	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Roll	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Time		-	-	-	-						-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table E.2: Range and average of r^2 values for all significant relationships that were modelled for the full extent.

		Heave		Pitch		Roll		Time	
		Range	Average	Range	Average	Range	Average	Range	Average
Spatial Similarity	Surfaces	0.644-0.959	0.866	0.699-0.997	0.953	0.747-0.999	0.868	0.737-0.957	0.846
Range	Surfaces			0.605-0.989	0.842	0.629-0.976	0.918	0.733-0.947	0.843
	Error	0.621-1.000	0.896	0.593-0.943	0.798	0.697-0.989	0.925	0.657-0.748	0.703
Mean	Surfaces			0.635-0.998	0.958	0.570-1.000	0.886	0.651-0.987	0.882
	Error	0.866-0.994	0.957	0.721-0.992	0.962	0.694-0.981	0.882	0.701-0.890	0.791
Standard Deviation	Surfaces	0.843	0.843	0.848-0.998	0.970	0.652-1.000	0.912	0.585-0.993	0.830
	Error	0.578-0.993	0.959	0.590-0.990	0.951	0.610-0.968	0.838	0.674-0.912	0.768
Skewness	Surfaces	0.869	0.869	0.541-0.987	0.879	0.586-0.999	0.863	0.662-0.927	0.793
	Error	0.848-0.999	0.947	0.596-0.918	0.833	0.675-0.995	0.761		
Kurtosis	Surfaces			0.746-0.999	0.933	0.555-1.000	0.851	0.529-0.888	0.651
	Error	0.904-0.995	0.962	0.641-0.871	0.799	0.581-0.996	0.723		
Spatial Autocorrelation	Surfaces	0.661-0.679	0.670	0.767-0.995	0.952	0.646-0.999	0.851	0.699-0.983	0.878
	Error	0.585-0.971	0.880	0.660-0.907	0.803	0.608-0.959	0.725	0.598-0.894	0.805
Fractal Dimension	Surfaces			0.776-0.978	0.912	0.891-0.922	0.902	0.576	0.576

For pitch artefacts, the correlation coefficients were on average lower for finer scales than for coarser scales, and lowest for topographic position, followed by northerness, easterness, slope, rugosity, bathymetry and topographic mean. Models that were all significant, explained well the relationships (average r^2 of 0.953) and all indicated a loss in spatial similarity with increasing artefact amplitude. The rates of change followed the same ranking as for the average correlations, meaning that topographic position had the quickest loss in association with the reference surfaces per degree of pitch, and topographic mean had the slowest. In all cases, finer scales changed more rapidly than coarser scales.

When looking at average correlations between the bathymetric and terrain attribute surfaces altered by roll and their respective reference surfaces, northerness generally showed the lowest correlations, followed by topographic position, easterness, slope, rugosity, bathymetry and topographic mean. Finer scales also showed lower correlations than coarser scales. Like pitch, models were all significant, explained adequately the relationships (average r^2 of 0.868, Table E.2) and confirmed a decrease in correlation with increasing roll. The rates of change showed that slope was the most impacted group of surfaces, *i.e.* becoming least similar to the reference surfaces as roll increases, followed by rugosity, topographic position, northerness, easterness, bathymetry and topographic mean. In general, finest scales had greatest rates of change than coarser scales.

Finally, the patterns of average correlation coefficients for time artefacts were the same as for pitch artefacts: correlations were weaker at finer scales, and lowest for topographic position followed by northerness, easterness, slope, rugosity, bathymetry and topographic mean. Models for topographic mean and the model for bathymetry at 10 m resolution were not significant. The remaining models all had a negative rate of change and an average r^2 of 0.846.

E.1.2 Sub-Areas

Average correlation coefficients for heave for the shallower sub-area followed the same patterns in term of scale and rankings than the full extent. For the deeper sub-area, correlation coefficients of topographic position were lowest, followed by northerness, easterness, slope, rugosity, bathymetry and topographic mean. In general, coefficients for finer scales were lower than those for coarser scales, except for easterness that did not have such a clear pattern. When looking at the shallower sub-area in terms of pitch artefacts, correlation coefficients followed in average the same patterns than for the full extent both for scales and ranking of surfaces, except for the inversion of slope and rugosity in the ranking. For the deeper sub-area, correlation coefficients for the analysis of pitch were always lower for finer scales, and were lowest for topographic position, followed by slope, northerness, rugosity, easterness, bathymetry and topographic mean. For the shallower sub-area and roll, average correlation coefficients followed the same pattern as for the full extent, except for the inversion of easterness and rugosity in the ranking. The deeper sub-area showed a different pattern. First, bathymetry and topographic mean behaved the same way as seen before with scale, *i.e.* finer scales had lower correlation coefficients than coarser spatial resolutions. However, the opposite pattern was observed with easterness, northerness, slope and rugosity, where coarser scales had generally lower correlation coefficients in average than finer scales. Topographic position did not show any particular pattern. Average correlation coefficients for time showed the same scale and ranking patterns than the full extent, except for the inversion of slope and rugosity in the ranking for the shallower sub-area. In general, when looking at the effect of roll and pitch on the correlations between altered and reference surfaces, correlation coefficients of the shallower sub-area were much higher than those of the deeper sub-area. The opposite pattern was observed for heave artefacts, and no clear relationship was found for time artefacts, likely due to the high correlation coefficients measured.

In terms of modelling of spatial similarity with artefact for the sub-areas, a summary of models significance is presented in Table E.3 below. Only two models were significant for heave: easterness at 25 and 50 m resolutions in the deeper sub-area. These two models showed that altered surfaces become more similar to the reference surfaces as heave level increases.

Table E.3: Summary of the modelling results for spatial similarity of the two sub-areas. See Table E.1 for legend.

		Bathymetry					Topo. Mean					Easterness					Northerness					Slope					Rugosity					Topo. Position				
		10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100
Shallower, High-Density Sub-Area	Heave		*	*	*	*			*	*		*	*		*	*	*	*		*	*	*	*	*		*	*	*	*		*	*		*		
	Pitch	-	-	-	-	-	-	-	-	-	-	-	-				*	-	-	-		-	-	-	-	-	-	-	-	-	*	*	-			*
	Roll	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Time	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Deeper, Low-Density Sub-Area	Heave	*		*	*		*	*	*	*	*	*	+	+	*			*				*	*	*		*			*			*		*		
	Pitch	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		-	-	-	-
	Roll	-	-	-	-	-	-	-	-	-	-				*	*	-	-	-	-	-	-	-	-	-		-			-		-	-	-	-	-
	Time	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*	-	-	-	-	-

For pitch, a few models were not significant for the shallower sub-area, and only one was not for the deeper sub-area (Table E.3). All significant models were negative, indicating decreasing similarity with increasing artefacts level (average r^2 of 0.919 for the shallower sub-area and 0.968 for the deeper one). The only generalizable observations regarding rates of change that could be made based on pitch models of the shallower sub-area were that bathymetry and topographic mean were the least impacted by pitch, and that correlations for bathymetry, slope and rugosity changed faster at finer scales. For the deeper sub-area, the greatest rate of change was observed in equations of topographic position, followed by slope, northerness, rugosity, easterness, bathymetry and topographic mean. The rates of change were also more important at finer scales than at coarser scales, except for easterness from which no distinct pattern could be identified.

Models for roll were all significant and negative for the shallower sub-area (average r^2 of 0.927). Unlike the full extent, it was not possible to generalize which terrain attribute was most impacted for that area since it changed with scales. For instance, at 10 and 25 m resolution, slope had the steepest rate of change with roll while at 50 m it was rugosity, at 75 m it was topographic position, and at 100 m it was northerness. Only bathymetry and easterness had a pattern with scale where finer resolutions were more impacted than coarser resolutions. In the deeper sub-area, fewer models were significant (Table E.3): easterness equations and some of rugosity were not significant. Significant models had an average r^2 value of 0.718 and negative rates of change. These models enabled identifying that topographic mean had the greatest rate of change, followed by bathymetry, topographic position, northerness, slope and rugosity. Finer scales changed more rapidly than coarser scales for northerness, easterness, rugosity and slope, but coarser scales changed more rapidly for topographic mean.

Finally, most models were significant in both areas for time artefacts with average r^2 values of 0.902 (shallower) and 0.893 (deeper) and indicated a loss in spatial similarity with increasing level of artefact induced. Finer scales always had a greater rate of change than coarser scales. The pattern for the deeper sub-area in terms of which surface was more impacted based on the rate of change followed that of the full extent. However, the shallower sub-area presented a slightly different pattern with topographic position as the most impacted, followed by easterness, northerness, rugosity, slope, bathymetry and topographic mean.

E.2 Error Modelling

E.2.1 Full Extent

Skewness, kurtosis and spatial autocorrelation of errors were not measured at 10 m resolution for the full extent analysis due to limitations in computer power.

When looking at their statistical distribution, errors induced by heave generally had a right-skewed, leptokurtic distribution, except for topographic mean (50 and 75 m resolution) for which errors followed a normal distribution, and topographic mean (25 and 100 m resolution) and bathymetry (75 m) that had a symmetrical leptokurtic distribution. The distribution of errors induced by pitch also had right-skewed, leptokurtic distributions, except for topographic mean (25 to 75 m resolutions) with symmetrical leptokurtic distributions. For roll, the error was in average normally distributed, except for rugosity, slope and topographic position at 25 m resolution that were leptokurtic right-skewed, and for rugosity (50 m and 75 m), topographic position (50 to 100 m), slope (50 m) and topographic mean (100 m) that were symmetrical leptokurtic. Finally, the distributions of errors caused by time were all leptokurtic right-skewed. Under the influence of certain artefacts, the distribution of errors for some terrain attributes became more symmetrical with finer scales: topographic mean (roll), easterness (heave, pitch and time), northerness (all four artefacts), and topographic position (heave, pitch and roll). In other cases, errors at coarser scales became more symmetrical (less skewed), namely topographic mean (pitch), rugosity (roll) and topographic position (roll).

Assuming that a Moran's I value of 0.5 and higher indicates positive autocorrelation, the errors induced by heave artefacts were autocorrelated for topographic mean and rugosity at 10 m resolution. Autocorrelation of errors for other surfaces were in average positive but low. The same pattern was observed for errors induced by pitch, although the bathymetric surfaces that had positive values of pitch induced also had spatially autocorrelated errors. More terrain attributes altered by roll showed autocorrelated errors. These included bathymetry, topographic mean, easterness and rugosity. For time, only the error of topographic mean was spatially autocorrelated. In general, errors from finer-scale surfaces were more autocorrelated than those of broader-scale surfaces.

A summary of results from the modelling of the error as a function of amplitude of artefact are presented in Table E.4 and a summary of the proportion of the total variance they explained is provided in Table E.2. As for the models previously introduced, r^2 values showed that the relationships were adequately modelled. The mean error introduced by each of the four types of artefact significantly increased with the amplitude of artefact for all of the bathymetric and terrain attributes surfaces at all scales, except for topographic position at 10 m resolution for time artefacts, for which the rate of change was not deemed significant by the t test. Other than the mean error, heave artefacts did not have any other generalizable impact on bathymetry. However, it increased the range of error values for topographic mean, and increased the standard deviation of error for the five other terrain attributes. For these five – easterness, northerness, slope, rugosity and topographic position – heave also modified the distribution of error (*i.e.* skewness and kurtosis) as it increased: the more important the artefact, the more the statistical distribution of error tend towards a normal distribution. In general, the error also became more autocorrelated with greater heave artefacts, except for topographic position where it decreased. Pitch artefacts had a very similar impact on the errors than heave. In this case however, like for the other terrain attributes, the standard deviation of error for bathymetry and topographic mean also increased. Also, patterns of spatial autocorrelation for errors in easterness and northerness were not significant. Roll artefacts had a clear impact on errors of all surfaces. The range of all of them but easterness, northerness and topographic position increased with the level of roll induced. Mean error and standard deviation of errors increased with roll for all the seven types of surfaces and the statistical distribution of errors became more symmetrical for all but bathymetry and topographic mean, for which no significant changes were recorded. The spatial autocorrelation of errors increased for bathymetry and rugosity but decreased for easterness and topographic position. Finally, errors caused by time artefacts only changed with amplitude in terms of mean error, standard deviation of error, and spatial autocorrelation. Generally, the two first statistics increased with greater time delays, except the standard deviations of topographic position that were not significant. Spatial autocorrelation of errors increased for bathymetry and topographic mean and decreased for northerness.

Table E.4: Summary of the modelling results for the descriptive statistics of artefact-induced errors. See Table E.1 for legend.

		Bathymetry					Topo. Mean					Easterness					Northerness					Slope					Rugosity					Topo. Position				
		10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100
Heave	Range	*	-	*	*	-	*	-	+	+	+			*		*					*		-					-				*	+			
	Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Standard Deviation	+										+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Skewness												-	-	-	-		-	-	-	-		-	-		-		-	-	-	-		-	-	-	-
	Kurtosis												-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-
	Spatial Autocorrelation												+	+	+	+		+	+	+	+			+	+	+		*	+	+	+		-	-	-	
Pitch	Range				*	+	+		+	+	+									+					+	+				+	+					
	Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Standard Deviation	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Skewness		*	*	*	*		*	*	*	*		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-
	Kurtosis												-	-	-	-		-	-	-	-			-	-			-	-	-			-	-	-	-
	Spatial Autocorrelation		*	*	*	*		*	*	*			*	*	*	+		*	*	*	*		+	+	+	+		*		+			*	*	*	*
Roll	Range	+		+	+	+	+	+	+	+	+					+						+	+	+	+	+	+	+	+	+	+	*				
	Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Standard Deviation	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	Skewness			*	*	*		*	*	*	-		-	-	-	-		-	-	-	-		-	-	-	-			+	-	-		-	-	-	-
	Kurtosis			-	*	-		*	-	*	-		-	-	-	-		-	-	-	-			*	-	-			-	-			-	-	-	-
	Spatial Autocorrelation		+	+	+	+			+	*	*			-	-	-		+	*	*	*				*	*		+		+	+		-	-	-	*
Time	Range							+							-																					
	Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	*	+	+	+	+
	Standard Deviation	+	+	+	+	+	+	+	+	+	+	*	+	+	+	+	*	*	+	+	+	+	+	+	+	+	*	*	+	+			*	*	*	*
	Skewness		*					*																*												
	Kurtosis																							*												
	Spatial Autocorrelation		+	+	+	+		+	+	+			*	-	*	*		*	-	-	-							-					*	-	*	*

Once again, finer scales were generally more impacted than coarser scales according to the measured rates of change. This was the case of the mean error for bathymetry (roll and time), topographic mean (pitch and time), easterness (pitch and time), northerness (pitch and time), slope (all four artefacts), and topographic position (pitch, roll and time). The standard deviations of error for bathymetry (heave, roll and time), topographic mean (heave, roll and time), easterness (pitch and time), northerness (pitch and time), slope (pitch, roll and time), rugosity (time) and topographic position (pitch and time) were also more impacted at finer scales. Some terrain attributes showed the opposite pattern, where coarser scales were more impacted than finer scales, for instance the mean error of bathymetry (heave), rugosity (roll) and topographic position (heave) and the standard deviation of error of easterness (heave), northerness (heave and roll), rugosity (roll) and topographic position (roll).

Table E.5 shows the spatial correlations among error surfaces of different levels of artefacts for each bathymetric and terrain attributes groups. On average, heave error surfaces were the most correlated with each other, followed by those of roll, time and pitch. The errors caused by roll on bathymetry and topographic mean were highly correlated, and so were the errors caused by heave on the five other types of surfaces. The correlation of errors caused by roll on easterness and northerness were particularly low.

Table E.5: Range and average of correlation coefficients recorded between error surfaces for the full extent area.

		Heave		Pitch		Roll		Time	
		Range	Average	Range	Average	Range	Average	Range	Average
Bathymetry	10	-0.766-0.998	0.103	0.041-0.947	0.379	0.946-0.980	0.958	0.214-0.893	0.516
	25	-0.836-0.998	0.050	0.029-0.959	0.383	0.962-0.986	0.971	0.223-0.903	0.551
	50	-0.896-0.998	0.004	0.004-0.963	0.375	0.967-0.988	0.975	0.219-0.877	0.537
	75	-0.916-0.998	-0.012	0.000-0.963	0.372	0.970-0.989	0.977	0.226-0.890	0.552
	100	-0.862-0.998	0.020	-0.002-0.963	0.369	0.971-0.989	0.978	0.230-0.872	0.529
Topographic Mean	10	-0.940-0.999	-0.053	-0.066-0.985	0.386	0.979-0.997	0.990	0.183-0.963	0.605
	25	-0.954-0.999	-0.069	-0.098-0.988	0.376	0.984-0.998	0.993	0.144-0.959	0.597
	50	-0.975-0.999	-0.082	-0.124-0.989	0.360	0.986-0.998	0.994	0.143-0.935	0.548
	75	-0.983-0.999	-0.086	-0.135-0.989	0.349	0.987-0.998	0.994	0.164-0.923	0.528
	100	-0.973-0.999	-0.082	-0.147-0.989	0.341	0.986-0.998	0.993	0.179-0.898	0.495
Easterness	10	0.997-1.000	0.999	0.200-0.797	0.428	-0.372-0.905	0.189	0.201-0.833	0.439
	25	0.997-1.000	0.999	0.234-0.866	0.498	-0.419-0.946	0.198	0.280-0.870	0.525
	50	0.995-1.000	0.998	0.286-0.879	0.534	-0.422-0.962	0.213	0.328-0.871	0.567
	75	0.993-1.000	0.997	0.315-0.889	0.551	-0.402-0.967	0.233	0.350-0.866	0.581
	100	0.189-1.000	0.712	0.333-0.896	0.559	-0.373-0.968	0.252	0.362-0.859	0.589
Northerness	10	0.997-1.000	0.999	0.232-0.794	0.417	-0.430-0.908	0.151	0.182-0.836	0.427
	25	0.998-1.000	0.999	0.251-0.866	0.478	-0.486-0.947	0.155	0.247-0.861	0.501
	50	0.997-1.000	0.999	0.276-0.887	0.514	-0.468-0.961	0.181	0.285-0.859	0.532
	75	0.997-1.000	0.999	0.302-0.893	0.536	-0.418-0.964	0.216	0.300-0.857	0.544
	100	0.997-1.000	0.998	0.322-0.894	0.549	-0.381-0.963	0.242	0.316-0.855	0.554
Slope	10	0.998-1.000	0.999	0.288-0.850	0.484	0.399-0.819	0.551	0.226-0.887	0.517
	25	0.997-1.000	0.999	0.284-0.895	0.521	0.288-0.907	0.543	0.203-0.908	0.541
	50	0.991-1.000	0.997	0.247-0.895	0.517	0.299-0.963	0.603	0.188-0.878	0.508
	75	0.987-1.000	0.996	0.235-0.888	0.508	0.424-0.980	0.694	0.195-0.889	0.505
	100	0.985-1.000	0.995	0.259-0.885	0.514	0.493-0.987	0.748	0.203-0.866	0.487
Rugosity	10	0.997-1.000	0.999	0.355-0.829	0.527	0.480-0.837	0.640	0.249-0.838	0.526
	25	0.997-1.000	0.999	0.375-0.880	0.565	0.314-0.886	0.590	0.237-0.869	0.550
	50	0.992-1.000	0.997	0.286-0.873	0.510	0.178-0.940	0.537	0.194-0.817	0.479
	75	0.987-1.000	0.996	0.270-0.870	0.498	0.242-0.968	0.597	0.207-0.869	0.502
	100	0.989-1.000	0.997	0.346-0.872	0.538	0.306-0.980	0.659	0.198-0.830	0.494
Topographic Position	10	0.998-1.000	0.999	0.186-0.709	0.360	0.304-0.792	0.468	0.136-0.805	0.375
	25	0.998-1.000	0.999	0.228-0.783	0.429	0.334-0.860	0.517	0.214-0.827	0.458
	50	0.998-1.000	0.999	0.234-0.833	0.461	0.316-0.925	0.546	0.250-0.818	0.490
	75	0.998-1.000	0.999	0.260-0.841	0.475	0.271-0.949	0.549	0.262-0.811	0.498
	100	0.997-1.000	0.999	0.281-0.842	0.483	0.236-0.958	0.546	0.277-0.803	0.505

E.2.2 Sub-Areas

When looking at the statistical distribution of errors for the sub-areas, a lower number of terrain attributes had right-skewed error distributions compared to the full extent area. In the shallower sub-area, the errors for five scales of topographic mean, three scales of bathymetry (10 m, 25 m and 100 m) and slope (100 m) were normally distributed. The two remaining scale of bathymetry, rugosity at 25 m, 50 m and 100 m, slope at 25 m, 50 m, and 75 m, and topographic position (10 m) had symmetrical leptokurtic distributions. Topographic mean was tending towards being more heavy-tailed and right-skewed with finer scales, while topographic position tended to become more skewed with coarser scales. A similar pattern was observed in the deeper sub-area, where all five scales of bathymetry and topographic mean errors were normally distributed, in addition to rugosity and slope at 100 m resolution. The errors of other terrain attributes were right-skewed, except for rugosity (75 m) and slope (50 and 75 m) that were symmetrical leptokurtic. Errors of finer scales easterness were usually more symmetrically distributed as opposed to errors in slope that were more symmetrical at coarser scales. The distributions of errors caused by roll also differed from the full extent. In the shallower sub-area, errors in topographic mean at each of the five scales were normally distributed, and most scales of bathymetry, rugosity and slope were symmetrical leptokurtic. Errors in northerness were more symmetrical at finer scales, and like for heave, errors in slope were more symmetrically distributed at coarser scales. In the deeper sub-area, errors in bathymetry (75 and 100 m resolution) and topographic mean were normally distributed, and errors in bathymetry (10 to 50 m), rugosity (100 m), northerness (10 m) and topographic position (10-25 m) were symmetrically distributed and heavy-tailed. Easterness, northerness and topographic position errors were more symmetrical at finer scales, and those of topographic mean and rugosity were more symmetrical at coarser scales. Patterns in statistical distribution of roll errors were very similar to the full extent for the two sub-areas. In the shallower sub-area, errors in bathymetry and northerness were more symmetrical at finer scales, but in the deeper sub-area, errors in topographic mean and northerness were more symmetrical at coarser scales. Time errors also had similar distribution than in the full extent analysis.

In general, errors caused by pitch values were more symmetrically distributed in the shallower sub-area for bathymetry (pitch and time), topographic mean (pitch and time), northerness (time), slope (pitch and time) and rugosity (pitch, roll and time), while bathymetry (roll), topographic mean (roll), easterness (pitch and roll), northerness (pitch and roll) and topographic position (pitch and time) errors were more symmetrically distributed in the deeper sub-area. Errors in slope caused by roll were more symmetrical in the deeper sub-area at coarser scales, and more symmetrical in the shallower sub-area at finer scales. A similar thing happened with errors in easterness caused by time artefacts, where at finer scales they were more symmetrical in the shallower sub-area and at coarser scales in the deeper sub-area. Patterns in spatial autocorrelation of errors with scale were generally not as clear as for the full extent.

In terms of spatial autocorrelation, heave and pitch errors in the shallower sub-area showed similar pattern than for the full extent. In the deeper sub-area however, errors in bathymetry were also autocorrelated, in addition to those of topographic mean. Errors in roll also did not demonstrate any major difference in spatial autocorrelation in the two sub-areas than for the full extent. Finally, time errors were in general less autocorrelated in the two sub-areas than in the full study area. Compared to the shallower sub-area, errors were more autocorrelated in the deeper sub-area for bathymetry (all four artefacts), topographic mean (all four artefacts), easterness (pitch and roll), northerness (pitch and roll), slope (roll and time) and rugosity (pitch, roll and time). The opposite was true of easterness (heave), rugosity (heave) and topographic position (pitch, roll and time). Time errors of easterness and northerness were more autocorrelated in the deeper sub-area at coarser scales but more autocorrelated in the shallower sub-area at finer scales.

Summaries of the proportion of the variance explained by the error models for the sub-areas and their significance are presented in Table E.6 and Table E.7. Average r^2 values were high. For heave, the patterns that were observed for the full extent were not seen in the sub-areas, except for the increase in mean error. Two additional patterns were however observed for the shallower sub-area: the range of error values increased and the skewness of error values decreased with more heave artefacts. For pitch, not as many trends were deciphered for the shallower sub-areas than for the full extent but those that were found were consistent with what was described for the full extent. For the deeper sub-area, the same trends were found but also the ranges of error values for bathymetry, topographic mean, slope and rugosity increased with greater pitch artefacts. The only difference in terms of modelling of roll error for the shallower sub-area compared to the full extent was that the spatial autocorrelation of easterness error increased with more artefacts instead of decreasing. The only difference for the deeper sub-area was that the skewness and kurtosis values increased for rugosity instead of decreasing. Little difference was observed in terms of error modelling of time artefact, although the range of error values increased for bathymetry in the shallower sub-areas and for topographic mean in both sub-areas.

Table E.6: Range and average of r^2 values for all significant relationships that were modelled for the two sub-areas.

			Heave		Pitch		Roll		Time	
			Range	Average	Range	Average	Range	Average	Range	Average
Shallower, High-Density Sub-Area	Spatial Similarity	Surfaces			0.771-0.998	0.919	0.735-1.000	0.927	0.712-0.993	0.902
	Fractal Dimension	Surfaces			0.715-0.796	0.748	0.846-0.997	0.937	0.770-0.938	0.854
	Range	Surfaces			0.738-0.934	0.806	0.589-0.988	0.902	0.652-0.835	0.758
		Error	0.634-0.939	0.805	0.718-0.918	0.836	0.616-0.986	0.910	0.587-0.940	0.828
	Mean	Surfaces			0.505-0.996	0.877	0.688-1.000	0.946	0.547-0.901	0.733
		Error	0.629-0.967	0.878	0.682-0.987	0.907	0.787-0.994	0.924	0.675-0.877	0.775
	Standard Deviation	Surfaces			0.552-0.968	0.825	0.700-1.000	0.940	0.581-0.960	0.804
		Error	0.611-0.823	0.691	0.762-0.988	0.889	0.636-0.984	0.889	0.658-0.925	0.798
	Skewness	Surfaces			0.761-0.800	0.781	0.662-0.969	0.848	0.550-0.944	0.739
		Error	0.613-0.876	0.734	0.744-0.956	0.888	0.581-0.929	0.748	0.623	0.623
	Kurtosis	Surfaces			0.660-0.758	0.709	0.675-0.992	0.875	0.736-0.959	0.856
		Error	0.621-0.850	0.690	0.612-0.932	0.842	0.589-0.910	0.687	0.630	0.630
	Spatial Autocorrelation	Surfaces			0.589-0.955	0.851	0.573-0.998	0.865	0.744-0.935	0.835
		Error	0.586-0.854	0.718	0.699-0.974	0.887	0.598-0.941	0.749		
Deeper, Low-Density Sub-Area	Spatial Similarity	Surfaces	0.721-0.721	0.721	0.868-0.994	0.968	0.582-0.800	0.718	0.752-0.992	0.893
	Fractal Dimension	Surfaces			0.851-0.992	0.945			0.688-0.916	0.822
	Range	Surfaces	0.580-0.805	0.693	0.814-0.993	0.937	0.608-0.977	0.914	0.669-0.973	0.797
		Error	0.628-0.892	0.778	0.577-0.987	0.875	0.735-0.966	0.940	0.700-0.933	0.806
	Mean	Surfaces			0.560-1.000	0.955	0.528-1.000	0.903	0.587-0.946	0.857
		Error	0.641-0.940	0.858	0.815-0.993	0.940	0.583-0.973	0.878	0.719-0.881	0.785
	Standard Deviation	Surfaces			0.779-1.000	0.959	0.601-0.971	0.922	0.540-0.967	0.807
		Error	0.638-0.867	0.739	0.710-0.973	0.922	0.634-0.966	0.889	0.709-0.913	0.787
	Skewness	Surfaces			0.620-0.994	0.897	0.585-0.965	0.752	0.621-0.912	0.809
		Error	0.583-0.907	0.757	0.611-0.897	0.764	0.601-0.945	0.723		
	Kurtosis	Surfaces			0.675-0.999	0.912	0.530-0.834	0.683	0.580-0.961	0.790
		Error	0.689-0.945	0.791	0.576-0.775	0.636	0.591-0.924	0.735	0.661	0.661
	Spatial Autocorrelation	Surfaces			0.583-1.000	0.913	0.538-0.869	0.606	0.564-0.960	0.810
		Error	0.617-0.679	0.657	0.592-0.742	0.679	0.591-0.929	0.712	0.952-0.972	0.959

Table E.7: Summary of the modelling results for the descriptive statistics of artefact-induced errors for the sub-areas. See Table E.1 for legend.

			Bathymetry					Topo. Mean					Easterness					Northerness					Slope					Rugosity					Topo. Position				
			10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100
Heave	Shallower, High-Density Sub-Area	Range	+	+	+	+									*	-	*	*	*			*		+									-		*		
		Mean	+	+	+	+	+	+	+	+	+	+		*					*					-			-						*	-			-
		Standard Deviation										-	+				*	+			*		+					*	*								-
		Skewness	-	-	-	-	*	*	*	*	*	*					*	*			+	*	*	*		-		*	+		-		-		*		
		Kurtosis			+												*	*			+		*			-	-	*	*		-		-				
		Spatial Autocorrelation											+			*		+	-	*					-	+					-		-	+			-
	Deeper, Low-Density Sub-Area	Range	+	+										-			-		*	+	-		-	+	-		+	*	-	+			-		+		
		Mean	+	+	+	+	+	+	+	+	+	+	+	*		+			+							*		-	+		*					-	
		Standard Deviation											+	-	-	+		-					*			*		*		+	*	-		+			
		Skewness							*	*	*	*		-		+	*	*			-						+		-	+		+				*	
		Kurtosis												-		+	*				-				*		+		-	*			*			*	
		Spatial Autocorrelation																		+	+										-			*	*		
Pitch	Shallower, High-Density Sub-Area	Range	+	*	+		+	+	+	+	*	+											+									+			*		
		Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	*	+	+	+	+	*	+	+	+	+	+	+	+	+	+	*	*	+	+	+	+
		Standard Deviation	+	+	+	+	+	+	+	+	+	+	+	+				*	+	+	*		+	+	+	+	*	+	+	+	+	*	*	*			*
		Skewness											-	-	-			-	-				*		*			*					*	-	-		
		Kurtosis											-	-	-			*	-	*			*					*				*	*	*	-		
		Spatial Autocorrelation	+	*	+	+	+	*	*	*			*	*	*			*	*	*			*	*	*	-		*	*	*	*		*	*			
	Deeper, Low-Density Sub-Area	Range	+	+	+	+	+	+	+	+	+	+					+						+	+	+	+	+	+	+	+	+	+	+			+	+
		Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+
		Standard Deviation	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+
		Skewness	-	-	-	-	-	-	-	*	*	*	-	-	-	-	-	-	-	-	-	-	*					-				*		-	-	-	
		Kurtosis				-	-		-	-				-	-	-	-		-	-	-	-		-										-	-		
		Spatial Autocorrelation					+																		*		+	+	*	*	*	*	-		*	*	
Roll	Shallower, High-Density Sub-Area	Range	+	+	+	+	+	+	+	+	+	+					+						+	+	+	+	+	+	+		+	+			+	+	
		Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+
		Standard Deviation	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+
		Skewness			-			*	*				-	-	-	-	-	-	-	-	-	-		-	-		*		-	*	-	-	-	-	-	-	-
		Kurtosis						-					-	-			-				-		-	-	-	-	*			+	-	-		-	-	-	
		Spatial Autocorrelation	+	+	*			*	*		*		+	+			+	+	*	*		+		-		*			+	*	+	+	-	-	-	-	*
	Deeper, Low-Density Sub-Area	Range	+	+	+	+	+	+	+	+	+	+					*						+	+	+	+	+	+	+	+	+	+				*	
		Mean	+	+	+	+	+	+	+	+	+	+	+			*	*	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+	+
		Standard Deviation	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+						+	+	+	+	+	+	+	+	+	+		+	+	+	*
		Skewness		-	*		*	*	*	*	*	*				*	*	-	-	-	-	*	-	-					+		+	+		-		-	-
		Kurtosis	-	-			*	*	*		*	-						-	-		*		-	-					+		+	+		-			
		Spatial Autocorrelation	+	*	*	*	*	+	*	+	*	*						+	*	*	*	+		+	*	-	-				-	-	-	-	-	*	
Time	Shallower, High-Density Sub-Area	Range	+	*	+		+	+	+	+	+	+											*	+	+		+		+	+		+					
		Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	*	+	+	+	+	*	+	+	+	+
		Standard Deviation	+	+	+	+	+	+	+	+	+	+	+	+	*	+	+		*	+			+	+	+	+	+	+	+	+	+	+		*	+	*	
		Skewness														*					*																-
		Kurtosis														*					*																-
		Spatial Autocorrelation	*					*																													
	Deeper, Low-Density Sub-Area	Range	*	*	*		+	+	+	+	+	+					*					*	+	+	*		+	*	*	*	+	*		*	*	+	
		Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	*	*	+	+	+	*	*	+	+	+	+
		Standard Deviation	+	+	+	+	*	+	+	+	+	+	*	+	+	*	*	*	+	+	*	*	+	+	+	+	*	+	+	+	+	*	*	+	+	+	*
		Skewness				*																														*	
		Kurtosis			-																									*						*	
		Spatial Autocorrelation	+	+	*			+	*		*			*				*													*						

Patterns for scale were similar to those observed for the full extent. Additional patterns included the mean error caused by pitch and time that was impacting finer scales easterness and topographic position more than coarser scales in the shallower sub-area, and the standard deviation of errors caused by pitch that was also greater for finer scales bathymetry. In the deeper sub-area, the mean errors of topographic mean (roll), rugosity (time) and topographic position (roll, pitch and time) were more impacted at finer scales, like the standard deviation of errors caused by pitch on bathymetry and topographic mean. Coarser scales topographic mean were more impacted by heave than finer scales in terms of mean error in the shallower sub-area. Pitch had a stronger impact in the deeper sub-area, where mean errors of bathymetry, topographic mean and rugosity, in addition to standard deviations of error of easterness and rugosity were all more impacted at coarser scales.

In general, based on the rates of change extracted from the modelling equations, heave had a greater impact in the shallower sub-area for bathymetry, topographic mean, easterness and northerness. Pitch, however, had more influence on errors of all seven types of surfaces in the deeper sub-area, while time had more influence on errors of all seven types of surfaces in the shallower sub-area. Roll had greater rates of change in the shallower sub-area for easterness, northerness and topographic position, and greater rates in the deeper sub-area for bathymetry, topographic mean, slope, and rugosity.

Finally, spatial correlation of errors, presented in Table E.8, behaved similarly to the full extent. Heave errors were slightly more correlated for the sub-areas than for the full extent, except for bathymetry and topographic mean. For pitch, correlations were higher for the full extent, except for easterness, northerness and topographic position in the shallower sub-area. Correlations in the shallower sub-area were consistently higher than in the deeper sub-area, except for broader scales slope and rugosity errors (50 to 100 m resolution). The correlations of errors caused by roll were generally higher in the deeper sub-area, except for easterness and broader-scale northerness (50 to 100 m resolution). Correlations in the deeper sub-areas were generally higher than for the full extent, and those for the shallower sub-area were generally lower than for the full extent. Finally, correlations of time errors were higher for the full extent, followed by the deeper sub-area. The only exception to this pattern was for northerness, where correlations in errors were higher in the shallower sub-area.

Table E.8: Range and average of correlation coefficients recorded among error surfaces for the sub-areas.

		Heave				Pitch				Roll				Time			
		Shallow		Deep		Shallow		Deep		Shallow		Deep		Shallow		Deep	
		Range	Average	Range	Average	Range	Average	Range	Average	Range	Average	Range	Average	Range	Average	Range	Average
Bathymetry	10	-1.000-1.000	-0.097	-1.000-1.000	-0.094	-0.524-0.866	0.117	-0.716-0.962	0.058	0.825-0.952	0.871	0.962-0.988	0.972	0.060-0.780	0.338	0.100-0.923	0.414
	25	-1.000-1.000	-0.099	-1.000-1.000	-0.096	-0.552-0.870	0.107	-0.743-0.970	0.054	0.841-0.952	0.879	0.974-0.993	0.982	0.074-0.787	0.368	0.092-0.933	0.453
	50	-1.000-1.000	-0.100	-1.000-1.000	-0.095	-0.565-0.876	0.100	-0.766-0.974	0.051	0.819-0.954	0.873	0.979-0.996	0.987	0.072-0.773	0.375	0.127-0.942	0.449
	75	-1.000-1.000	-0.097	-1.000-1.000	-0.098	-0.605-0.865	0.073	-0.824-0.972	0.019	0.770-0.960	0.848	0.979-0.996	0.988	0.145-0.727	0.391	0.189-0.924	0.492
	100	-1.000-1.000	-0.105	-1.000-1.000	-0.100	-0.553-0.857	0.086	-0.844-0.971	0.015	0.749-0.941	0.839	0.985-0.998	0.991	0.123-0.775	0.416	0.185-0.922	0.495
Topographic Mean	10	-1.000-1.000	-0.110	-1.000-1.000	-0.102	-0.734-0.970	0.122	-0.842-0.993	0.029	0.913-0.991	0.958	0.984-0.999	0.994	-0.047-0.912	0.393	0.032-0.976	0.546
	25	-1.000-1.000	-0.111	-1.000-1.000	-0.104	-0.769-0.963	0.093	-0.921-0.996	-0.009	0.888-0.982	0.938	0.987-0.999	0.996	-0.054-0.849	0.315	-0.022-0.963	0.469
	50	-1.000-1.000	-0.111	-1.000-1.000	-0.106	-0.726-0.961	0.102	-0.956-0.996	-0.033	0.803-0.965	0.879	0.985-0.999	0.995	-0.010-0.815	0.298	-0.059-0.934	0.351
	75	-1.000-1.000	-0.111	-1.000-1.000	-0.107	-0.695-0.954	0.068	-0.969-0.997	-0.050	0.772-0.978	0.886	0.984-0.999	0.995	-0.020-0.754	0.262	-0.003-0.923	0.355
	100	-1.000-1.000	-0.111	-1.000-1.000	-0.107	-0.720-0.950	0.096	-0.978-0.997	-0.052	0.829-0.979	0.905	0.987-1.000	0.996	0.001-0.717	0.255	0.073-0.914	0.346
Easterness	10	1.000-1.000	1.000	1.000-1.000	1.000	0.314-0.836	0.582	0.085-0.747	0.304	0.070-0.793	0.338	-0.837-0.942	-0.039	0.116-0.797	0.381	0.144-0.888	0.410
	25	1.000-1.000	1.000	1.000-1.000	1.000	0.393-0.878	0.652	0.084-0.805	0.375	0.171-0.840	0.408	-0.939-0.982	-0.071	0.218-0.829	0.469	0.212-0.930	0.489
	50	1.000-1.000	1.000	1.000-1.000	1.000	0.499-0.881	0.699	0.086-0.849	0.388	0.137-0.885	0.417	-0.957-0.990	-0.079	0.317-0.873	0.558	0.175-0.934	0.472
	75	1.000-1.000	1.000	1.000-1.000	1.000	0.417-0.872	0.616	0.138-0.887	0.447	0.089-0.873	0.403	-0.956-0.993	-0.073	0.320-0.868	0.566	0.173-0.929	0.490
	100	1.000-1.000	1.000	1.000-1.000	1.000	0.420-0.847	0.633	0.226-0.887	0.477	0.107-0.862	0.451	-0.959-0.997	-0.063	0.138-0.853	0.489	0.248-0.940	0.541
Northerness	10	1.000-1.000	1.000	1.000-1.000	1.000	0.338-0.779	0.551	0.030-0.740	0.253	-0.661-0.937	0.091	-0.280-0.841	0.152	0.201-0.705	0.432	0.085-0.908	0.374
	25	1.000-1.000	1.000	1.000-1.000	1.000	0.411-0.853	0.647	0.053-0.862	0.348	-0.609-0.955	0.156	-0.355-0.925	0.168	0.271-0.814	0.505	0.160-0.936	0.447
	50	1.000-1.000	1.000	1.000-1.000	1.000	0.389-0.863	0.653	0.051-0.876	0.376	-0.320-0.951	0.308	-0.411-0.965	0.201	0.272-0.792	0.487	0.112-0.935	0.461
	75	1.000-1.000	1.000	1.000-1.000	1.000	0.428-0.903	0.643	0.105-0.891	0.424	-0.222-0.944	0.355	-0.397-0.975	0.224	0.231-0.841	0.549	0.179-0.939	0.533
	100	1.000-1.000	1.000	1.000-1.000	1.000	0.590-0.944	0.780	0.167-0.902	0.476	-0.175-0.932	0.393	-0.381-0.986	0.230	0.268-0.897	0.666	0.374-0.954	0.589
Slope	10	1.000-1.000	1.000	1.000-1.000	1.000	0.108-0.784	0.420	0.103-0.763	0.314	0.355-0.950	0.608	0.320-0.800	0.525	0.028-0.764	0.296	0.086-0.916	0.388
	25	1.000-1.000	1.000	1.000-1.000	1.000	0.102-0.781	0.432	0.171-0.877	0.415	0.353-0.971	0.617	0.449-0.944	0.734	0.018-0.768	0.289	0.057-0.943	0.442
	50	1.000-1.000	1.000	1.000-1.000	1.000	0.049-0.759	0.394	0.139-0.909	0.472	0.325-0.945	0.603	0.337-0.984	0.791	0.014-0.734	0.266	0.066-0.934	0.395
	75	1.000-1.000	1.000	1.000-1.000	1.000	0.018-0.759	0.392	0.073-0.905	0.491	0.295-0.932	0.632	0.251-0.990	0.810	0.005-0.668	0.250	0.140-0.902	0.368
	100	1.000-1.000	1.000	1.000-1.000	1.000	0.147-0.750	0.431	0.069-0.932	0.542	0.396-0.947	0.662	0.414-0.997	0.863	0.109-0.702	0.312	0.148-0.890	0.372
Rugosity	10	1.000-1.000	1.000	1.000-1.000	1.000	0.124-0.723	0.405	0.086-0.734	0.294	0.352-0.931	0.607	0.299-0.760	0.525	0.073-0.701	0.298	0.096-0.871	0.374
	25	1.000-1.000	1.000	1.000-1.000	1.000	0.110-0.743	0.423	0.120-0.865	0.379	0.096-0.956	0.512	0.532-0.930	0.760	0.031-0.726	0.295	0.084-0.915	0.396
	50	1.000-1.000	1.000	1.000-1.000	1.000	0.037-0.752	0.384	0.124-0.914	0.437	0.095-0.958	0.478	0.351-0.978	0.799	-0.010-0.692	0.249	0.081-0.926	0.362
	75	1.000-1.000	1.000	1.000-1.000	1.000	-0.001-0.757	0.342	0.069-0.917	0.444	0.108-0.932	0.457	0.250-0.987	0.808	-0.071-0.646	0.227	0.148-0.896	0.377
	100	1.000-1.000	1.000	1.000-1.000	1.000	0.174-0.773	0.429	0.161-0.901	0.478	0.058-0.928	0.460	0.333-0.995	0.844	0.061-0.651	0.263	0.190-0.907	0.415
Topographic Position	10	1.000-1.000	1.000	1.000-1.000	1.000	0.218-0.764	0.443	0.110-0.582	0.263	0.181-0.847	0.406	0.231-0.778	0.460	0.096-0.667	0.317	0.082-0.877	0.338
	25	1.000-1.000	1.000	1.000-1.000	1.000	0.245-0.759	0.497	0.092-0.739	0.289	0.070-0.924	0.434	0.351-0.874	0.565	0.127-0.741	0.380	0.134-0.912	0.413
	50	1.000-1.000	1.000	1.000-1.000	1.000	0.350-0.805	0.584	0.064-0.820	0.321	0.035-0.953	0.492	0.454-0.946	0.655	0.176-0.767	0.475	0.148-0.923	0.429
	75	1.000-1.000	1.000	1.000-1.000	1.000	0.344-0.848	0.584	0.035-0.857	0.342	0.181-0.956	0.561	0.460-0.973	0.666	0.213-0.754	0.473	0.180-0.917	0.467
	100	1.000-1.000	1.000	1.000-1.000	1.000	0.215-0.815	0.542	0.111-0.834	0.372	0.190-0.920	0.574	0.345-0.991	0.641	0.241-0.787	0.451	0.165-0.909	0.462

E.3 Changes in Surfaces

E.3.1 Full Extent

Skewness, kurtosis and spatial autocorrelation of surfaces could not be measured at 10 m resolution because of limited computer power. A summary of results from the modelling of the change in descriptive statistics of the altered bathymetric and terrain attribute surfaces are presented in Table E.9 and a summary of the proportion of the total variance they explained is provided in Table E.2. As shown by the high average r^2 values, significant equations explained adequately the relationships.

Results showed that increasing heave artefact did not significantly change bathymetric and terrain attributes surfaces. Only four models indicated significant change, thus preventing any pattern generalization. The spatial autocorrelation of depth values decreased with increasing artefacts at 25 m resolution, and so did value of topographic mean at 75 m resolution. Then, the standard deviation of slope values increased with amplitude of heave artefact at 100 m resolution. Finally, the skewness of the distribution of slope values at 75 m resolution decreased with more artefacts.

Table E.9: Summary of the modelling results for changes in surfaces. See Table E.1 for legend.

		Bathymetry					Topo. Mean					Easterness					Northerness					Slope					Rugosity					Topo. Position				
		10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100
Heave	Range																																			
	Mean	*	*	*	*	*		*	*	*	*																									
	Standard Deviation															+					*					*										
	Skewness		*					*										*	*	*	*				-											
	Kurtosis														*						*															
	Spatial Autocorrelation		-							-																										
	Fractal Dimension										*																									
Pitch	Range			-	-	-	-	-	-	-	-														-						+		*	-		
	Mean	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	-	*	-	+	+	*	-	-	+	+	+	+	-	-	-	-	+	
	Standard Deviation	-	-	-	-	-		-	-	-	-		-	-	-	-	+	+	+	+	+	-	-	-	-	-	+	-	-	-	-	+	+	+	+	+
	Skewness		-	*	*	*		-	*	*	*		-	-	-	-		+	+	+	+				*	*			+	+	+		+	+	+	
	Kurtosis		*	+	*	*		*	+	*	*		*	*	*	*		-	-	-	-				*	+				+	+		-	-	-	-
	Spatial Autocorrelation			-	-	-				-			-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-		-	-	-	-
	Fractal Dimension		+	+	+	+		+																												
Roll	Range	+	+	+	+	+	+	+	+	+	+											+	+	+	+	+	+	+	+	+	+		-			
	Mean	-	-	-	-	-		-	-	-	-	+					+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-			
	Standard Deviation	+	+	+	+	+		+	+	+	+		-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	-		+	+
	Skewness		-	-	-	-		-	-	-	-							-			-		-	-	-	-		+	-	-	-		+			
	Kurtosis		+	+	+	+		+	+	+	+		+	+	+	+		*	+	+	+			-	-	-				-	-		+	+		-
	Spatial Autocorrelation		-	-	-	-		-	-	-	-				-	-			-	-	-		-	-	-	-		-	-	-	-		+	+		
	Fractal Dimension		+	+	+	+		+	+	+	+																									
Time	Range				+			*	+	+	+														-											
	Mean	-		-				-	-	-					*		+	+				-	-	-	-	*		-	-	-	-	-	-	+		*
	Standard Deviation	-		*		*			*	*	+			-	-		+	*	+			-	-		-				*			+	-	-	-	-
	Skewness			*											+																		+	+		*
	Kurtosis			-		+					+					*		-	-	-													-			
	Spatial Autocorrelation				+								+	+	+			+	+	+			+		*	*		*						+	+	+
	Fractal Dimension		*	*	*	*		*	*	*	+																									

Bathymetric and topographic mean surfaces were most impacted by pitch and roll artefacts, although in opposite ways. While pitch artefacts produced a decrease in range and standard deviation values and an increase in mean values, roll artefacts produced an increase of range and standard deviation and a decrease in mean values. Based on the lack of significance of the models, pitch did not change the distribution of these two types of surfaces, while roll decreased skewness and increased kurtosis, making the distribution of depth values more leptokurtic and symmetrical. Both types of artefacts induced an increase in fractal dimension, and roll generated a higher rate of change in fractal dimension for coarser scales than for finer scales. Both pitch and roll decreased spatial autocorrelation values of bathymetry and topographic mean, although generalization could not be made regarding spatial autocorrelation of the latter caused by pitch since only the 75 m resolution model was significant. It was also challenging to generalize any impact from time artefact based on the lack of significance of many models. The only patterns that appeared were that time artefact increased the range of values of topographic mean and decreased their mean. The range of values of topographic mean was changing more rapidly with increasing amplitude of artefact at finer scales than at broader scales.

No artefact changed significantly the range of easterness and northerness values, likely due to the fact that these terrain attributes are the only ones with a limited range of values (-1 to 1), which is usually fully represented within an area. For these terrain attributes, pitch seemed to have been the artefact with the most impact. For easterness, pitch generally increased the mean values and decreased the standard deviation, skewness and spatial autocorrelation values. For northerness, it increased standard deviation and skewness and decreased kurtosis and spatial autocorrelation. It also altered the mean of northerness values but whether its impact was positive or negative depended on scale. In terms of roll, it decreased spatial autocorrelation and increased kurtosis for both easterness and northerness. While roll decreased standard deviation values for easterness, it increased those of northerness together with its mean values. Finally, the most noticeable impact of time artefacts on easterness and northerness was that it increased the spatial autocorrelation of values.

More statistics of slope and rugosity were altered by roll artefacts than by the other types of artefacts. Roll increased the ranges in slope and rugosity values, the means and the standard deviations. Roll also decreased skewness, kurtosis and spatial autocorrelation of these two terrain attributes. In general, pitch lowered the standard deviation and spatial autocorrelation values of slope and rugosity, and increased the mean, skewness and kurtosis of rugosity. The impact of pitch on mean values of slope was dependent on scale. The rate of change of mean values of slope and rugosity was greater at finer scales than coarser scales. Finally, time artefacts only decreased the mean values of these two terrain attributes.

Topographic position was the terrain attribute from which it was most difficult to generalize impacts from artefacts because a lot of the broader scales models were not significant. Pitch decreased kurtosis and spatial autocorrelation of this terrain attribute and increased its standard deviation and skewness. No generalization could be made from roll impacts other than spatial autocorrelation increased at finer scales and standard deviation decreased at finer scales and increased at broader scales. Time artefacts increased spatial autocorrelation and decreased standard deviation values.

Some patterns in relation to scale were observed based on the different rates of change (*i.e.* the first term of each quadratic equation). Usually, finer scales were more impacted than coarser scales. This was true of (1) the impact of pitch on the standard deviations of easterness, northerness and topographic position, and the means of slope and rugosity; (2) the impact of roll on the mean of slope and standard deviations of bathymetry and topographic mean, and (3) the impact of time on means of topographic mean and slope. The loss of autocorrelation with increasing level of artefact was also greater at finer scales when pitch was affecting easterness, northerness, topographic position, slope and rugosity, and when roll was affecting topographic position. However, in some cases, the impacts were greater for coarser scales. This was for instance observed for (1) the impact of pitch on the standard deviation of topographic mean, (2) the impact of roll on the standard deviation of topographic position and on the mean values of northerness and rugosity, and (3) the loss of autocorrelation caused by roll on bathymetry, easterness, topographic mean and northerness.

E.3.2 Sub-Areas

A summary of the models for the sub-areas is presented in Table E.10 and the proportion of the variance they explain is presented in Table E.2. Again, the significant models explained well the modelled relationships with high average r^2 values. Results were more heterogeneous for the sub-areas than for the full extent, *i.e.* not necessarily following clear trends. To remain succinct, we focus here on how the observed patterns agree or disagree with those observed in the analysis of the full extent area.

Table E.10: Summary of the modelling results for changes in surfaces for the two sub-areas. See Table E.1 for legend.

			Bathymetry					Topo. Mean					Easterness					Northerness					Slope					Rugosity					Topo. Position				
			10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100	10	25	50	75	100
Heave	Shallower, High-Density Sub-Area	Range																																			
		Mean	*	*	*	*	*	*	*	*	*	*																									
		Standard Deviation																																			
		Skewness																																			
		Kurtosis																																			
		Spatial Autocorrelation																																			
		Fractal Dimension																																			
	Deeper, Low-Density Sub-Area	Range															-			+																	
		Mean	*	*	*	*	*	*	*	*	*	*																									
		Standard Deviation																																			
		Skewness																																			
		Kurtosis																																			
		Spatial Autocorrelation																																			
		Fractal Dimension																																			
Pitch	Shallower, High-Density Sub-Area	Range				-		-	-	-	-	-											*		*	*	*			*	*	*				*	
		Mean	+	+	+	+	+	+	+	+	+	+			-	-	*			*			-	-	-	-	-	*	-	-	-	-	-				
		Standard Deviation	-	-	-	-	-	-	-	-	-	-	*	-	-	*	-	*	+	+	*	+	-	-	-	-		*	*	-	-	*	*	*		*	*
		Skewness	*	*	*	+	*	*	*	*	*	*	+					*					*	*		*		*	*	*				*	*		*
		Kurtosis					*		*	*			*	+			*	*		-			*	*		*		*								*	
		Spatial Autocorrelation	-	-		+		-			*		-			*		*		*	*	*	-	*		*		-	*		*		-	*			
		Fractal Dimension	+	+		*		+		*	*	*																									
	Deeper, Low-Density Sub-Area	Range	+	+	+	+	+	+	*	*	+	*	*										*	*	*	+	+	*	*	*	*		+		*		+
		Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	*	*
		Standard Deviation	+	+	+	+	+	+	+	+	*	+	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
		Skewness	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	*	+	-	-	-	*	-
		Kurtosis	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+	*	-	-	-	*		-	-	-	-	-
		Spatial Autocorrelation	-	-	-	-	-	*	*	+	-	+	-	-	+	*	-	-	-	-	-	-	-	-		*	+		*	+	*	+	-	-	-	-	
		Fractal Dimension	+	+	+	+	+	+	+	+	+	+																									
Roll	Shallower, High-Density Sub-Area	Range	+	+	+	+	+	+	+	+	+	+		-		-							+	+	+	+	+	+	+	+	+	+				*	
		Mean	+	+	+	+	+	+	+	+	+	+		-	-	*	*	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		-	*	+	*
		Standard Deviation	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+
		Skewness	-	-	-	+	*	-	-	+	+	*	+	+	+	*	*	-	-	-	-	-	+	-	-	-	*	+	-	-	-	*				+	
		Kurtosis	+	+	+	*	+	+	+	*	-	+	+	+	+	+	+	-	-	-	-	-	+	-	-	*	+	+	-	-	-	+	+	-	-	-	-
		Spatial Autocorrelation	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-		-	-	+	+	-	-	-
		Fractal Dimension	+	+	+	+	+	+	+	+	+	+																									
	Deeper, Low-Density Sub-Area	Range	+	+	+	+	+	+	+	+	+	+											+	+	+	+	+	+	+	+	+	+					-
		Mean	-	-	-	-	-	-	-	-	-	-						+	+	+	+		+	+	+	+	+	+	+	+	+	+	+				*
		Standard Deviation	+	+	+	+	+	+	+	+	+	+						-					+	+	+	+	+	+	+	+	+	+	-				
		Skewness																-															-	-	-		
		Kurtosis	-	-	-	-	-	-	-	-	-	-						+		-	-										+	+	+	+	+		
		Spatial Autocorrelation	+	+	+					+		+											-		*								-		+		
		Fractal Dimension																																			
Time	Shallower, High-Density Sub-Area	Range				-	*	-	*		+	*			*								-			*	*									*	-
		Mean	-		*			-	-	*			*	*		*					*		-	-				-	-	*						-	*
		Standard Deviation	-	-				-	-	-			-	-		*	-	+	+		*	+	*	*		*		+	+	*	-			-			-
		Skewness	*	*	-			*	*		-	*											+	*				+	*		+			-			
		Kurtosis	*	*		-		*	*	*	*	*	+	+		*	+	-	-	-	*		+	*				+	*		+					*	
		Spatial Autocorrelation	+	+	*			+	*	*			+	+				*	*	*			+	+				+			-			+			
		Fractal Dimension	*		+			-																													
	Deeper, Low-Density Sub-Area	Range		+	+	+	+		*	+	*	+											+	*	*	*	-	-	-	+	*	*			*		-
		Mean	*	*	*			*	*	*	*			+	-	+	+				*	-	-	-	-	-	*		-	-		-			-	-	+
		Standard Deviation	*	-	-	-	-	*	-	-	-	-	+	+	-	+	+	-		+	*	-	*	-	-	*	-		-	-	-	*	+	-	*	+	-
		Skewness	*			+		*		*	+	*		-		-			*			*	+	-	*	-	*	+		-	*	*	+	*	*		+
		Kurtosis	*	*	+		*	*	*	+		*		-		-		+	*			+				*	*						-				
		Spatial Autocorrelation	*	*	*	*		+	*	-	*		-	*	*	*			+	-		-	-	*	+	*		-			-	*	-			-	+
		Fractal Dimension	*	*	*	-	*	-	-	-	*	*																									

Like for the full extent, results show that heave did not significantly change the characteristics of the bathymetric and terrain attributes surfaces. Only two models were significant: the ranges of easterness at 100 m and northerness at 50 m resolution, both for the deeper sub-area.

In comparison with the results from the full study area, pitch artefacts had a very similar impact on bathymetry and topographic mean in the shallower sub-area. However, the deeper sub-area presented a different pattern: the range of depth and the standard deviation both increased with artefacts and the distribution of depth values became more leptokurtic and symmetrical. In terms of topographic mean, the mean values and standard deviation increased with pitch level. For roll, both sub-areas shared similarities and differences with the full extent in terms of bathymetry. For the shallower sub-area, the mean values increased with roll instead of decreasing. For the deeper sub-area, spatial autocorrelation increased with more roll and the distribution of values became more platykurtic. Finally, while no particular impact of time was found on the statistics of the full area, the range and standard deviation of values for the deeper sub-area respectively increased and decreased with greater time artefacts.

No major differences were found for easterness and northerness compared to the full extent, except that in the shallower sub-area, the mean decreased with increasing pitch. Also, no change in easterness values distribution (*i.e.* skewness and kurtosis) was found for roll in the deeper sub-area. The different descriptive statistics for slope also behaved the same as the full extent for both sub-areas. One difference was observed for rugosity, in the shallower sub-area: the mean values decreased with pitch artefacts, which was not the case for the full extent. Finally, topographic position behaved differently in the deeper sub-area, where mean values increased with pitch artefacts.

In general, based on the rates of change from the models, the following terrain attributes and statistics combinations were more impacted by pitch in the deeper sub-area than in the shallower sub-area: bathymetry (standard deviation and spatial autocorrelation), topographic mean (standard deviation), easterness (spatial autocorrelation), northerness (standard deviation), slope (standard deviation and spatial autocorrelation), rugosity (standard deviation). Roll artefacts impacted the shallower sub-area more than the deeper one for bathymetry (mean and spatial autocorrelation) and topographic mean (mean and spatial autocorrelation). The opposite was found for bathymetry (standard deviation), topographic mean (standard deviation), easterness (standard deviation), northerness (mean), slope (mean and standard deviation), and rugosity (mean, standard deviation). Finally, time artefacts impacted topographic mean (standard deviation and spatial autocorrelation), easterness (spatial autocorrelation), slope (spatial autocorrelation) and topographic position (standard deviation) more in the deeper sub-area, while it impacted more northerness (standard deviation) in the shallower sub-area.

Finally, in terms of scale, results from the shallower sub-area indicated that the standard deviation of bathymetry and easterness decreased faster at coarser scales and that the standard deviation of northerness increased faster at coarser scales. In the deeper sub-area, easterness was also more impacted at coarser scales but topographic mean and slope were more impacted at finer scales. Roll had a more consistent impact on terrain attributes. The standard deviation of finer-scale bathymetry, slope and rugosity surfaces were more impacted than coarser-scale surfaces in the shallower sub-area. The same pattern was observed for bathymetry and topographic mean in the deeper sub-area, but this time rugosity was more impacted at broader scales. In terms of skewness for the shallower sub-are, pitch artefacts impacted more the coarser scales of bathymetry and topographic position and the finer scales of easterness and northerness. Roll also impacted spatial autocorrelation of easterness (shallower sub-area) and bathymetry (deeper sub-area) more at finer scales.

Appendix F: Overall Accuracies of the Habitat Maps (Chapter 6)