Reconstruction and Visualization of Archaeological Networks

Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften

vorgelegt von

Jan Christoph Athenstädt

an der

Universität Konstanz

Mathematisch-Naturwissenschaftliche Sektion

Fachbereich Informatik und Informationswissenschaft

Konstanz, 2018

Tag der mündlichen Prüfung: 25.5.2018

1. Referent/Referentin: Prof. Dr. Ulrik Brandes

2. Referent/Referentin: PD Dr. Sabine Cornelsen

Reconstruction and Visualization of Archaeological Networks

Jan Christoph Athenstädt

Acknowledgements

I would like to thank my supervisor Ulrik Brandes and my referee Sabine Cornelsen for the academic advice and support while writing this thesis.

My thanks go to the *Algo*-group at the University of Konstanz for the mental support and the lively discussions during the coffee breaks. In particular I would like to thank my office mates Arlind Nocaj and Habiba for the good conversations, Christine Agorastos for the helpful support with all bureaucratic matters, Viviana Amati for the never-ending chocolate supply and her sympathetic ear, and Mereke van Garderen for taking the time to proofread this thesis and giving extensive feedback. I also want to thank the students whose theses I supervised, Jan Hildenbrand and Roman Lamsal.

In addition, I would like to thank the archaeology group from Leiden University for the insights into the fascinating field of archaeology. In particular, my thanks go to Corinne Hofman for giving me the opportunity to take part in excavations in the Caribbean, Lewis Borck for the collaboration and the "Great Road Trip of Polychrome and Plainware Guy", Katarina Jacobson for her patience in explaining pottery classification and her participation in the perception tests, and Emma Slayton for the collaboration with the canoe routes.

Last but not least, my special thanks go to my friends and family for being part of my life. Without their support I would not have finished this thesis.

This research has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 319209.

Deutsche Zusammenfassung

Gegenstand dieser Arbeit sind verschiedene Aspekte der Rekonstruktion und Visualisierung von Netzwerkdaten in einem archäologischen Kontext in der Karibik. Die vorgestellten Methoden können einen Beitrag leisten, Migrations- und Handelswege zu rekonstruieren und die resultierenden Daten zu veranschaulichen.

Die häufig sehr fragmentierten und unvollständigen Datensätze, basierend auf Ausgrabungen und der Analyse von Fundstücken, stellen große Herausforderungen an Archäologen, wenn es darum geht, menschliche Interaktionen in der Vergangenheit zu rekonstruieren. Insbesondere in der Region im Fokus dieser Arbeit, der Karibik, sind vor allem Artefakte aus organischem Material aufgrund der klimatischen Bedingungen kaum erhalten.

Teil 1 - Distanzen, basierend auf Kanurouten
In Teil 1 werden optimale Kanurouten zwischen den Inseln der Karibik rekonstruiert, unter Berücksichtigung von Umwelteinflüssen durch Strom, Wind und Wellen. Der Teil beginnt mit einer Analyse der Umwelteinflüsse und ihrer saisonalen Veränderungen mit einem Fokus auf den Strömungsdaten. Hierbei zeigen sich allerdings nur schwache Tendenzen, die von den täglichen Schwankungen überlagert werden.

Zwei existierende Ansätze zur optimalen Routenberechnung von Schiffen werden adaptiert, um der Situation in der Karibik mit vielen Inseln und Strömungsgeschwindigkeiten nahe an der Maximalgeschwindigkeit der Kanus gerecht zu werden. Im Vergleich der Methoden stellt sich die auf Zeitfronten basierende Methode als überlegen in Bezug auf die Rechenzeit heraus. Die Ergebnisse verschiedener Fallstudien mit dem Algorithmus werden vorgestellt.

Um zu evaluieren, inwieweit menschliche Akteure in der Lage sind, die optimalen Routen zu finden, wurde ein Paddelspiel entwickelt, dass es dem Spieler erlaubt, ein Kanu auf verschiedenen Routen zwischen den Inseln der Karibik zu paddeln. Im Rahmen eines zweiwöchigen Wettbewerbs wurden die Routen der Spieler aufgezeichnet und analysiert. Die Spieler waren zwar nicht in der Lage, die optimale Route zu finden; allerdings deutet eine Analyse der Spielzeiten der beiden aktivsten Spieler darauf hin, dass sich die Spieler in der Wahl der Routen kontinuierlich verbessert haben.

Teil 2 - Quantifizierung der Ähnlichkeitswahrnehmung von Tonscherben
In Teil 2 werden Methoden zur Quantifizierung menschlicher Wahrnehmung von Distanzen und Ähnlichkeiten entwickelt und evaluiert mit dem Ziel, diese Distanzmaße für eine Typisierug von Tonscherben einzusetzen. Die entwickelten Methoden werden mit Triadentests, einer etablierten Methode aus dem Bereich der Soziologie, verglichen.

In der ersten Methode (*two-dimensional arrangements*) werden die Teilnehmer gebeten, Objekte im zweidimensionalen Raum so anzuordnen, dass ihre Abstände den empfundenen Ähnlichkeiten entsprechen.

In der zweiten Methode (*multidimensional distance recording, MDR*) werden gemeinsam mit dem Teilnehmer mögliche Dimensionen zur Einordnung der Objekte herausgearbeitet und die Objekte in diesen Dimensionen auf kategorischen Skalen, ordinalen Skalen oder der Intervallskala angeordnet. Mit Hilfe der Gower-Distanz können Entfernungen zwischen den Positionen der Objekte berechnet werden.

Die Methoden werden in drei Fallstudien evaluiert:

- Die quantifizierten Distanzen eines Keramologen aus der *MDR*-Methode werden mit den Ergebnissen eines Triadentests verglichen. Die Ergebnisse zeigen, dass mit der *MDR*-Methode mit erheblich geringerem Zeitaufwand ähnlich gute Ergebnisse erzielt werden können, vorausgesetzt, die Zahl der gewählten Dimensionen ist ausreichend.
- In einer Online-Studie wird ein Triadentest mit 3D-Rekonstrukionen von Tontöpfen, die verschiedene Kombinationen aus Attributen vereinen, durchgeführt. Anhand der Ergebnisse wird aufgezeigt, wie ein solcher Test verwendet werden kann, um die wahrgenommene Relevanz der Attribute zu quantifizieren.
- In einem anthropologischen Experiment im Südwesten der USA werden Unterschiede in der Wahrnehmung von archäologischen Tonscherben zwischen (indigenen) Töpfern und Archäologen evaluiert. Die Ergebnisse in den Anordnungen der Scherben mittels der *two-dimensional arrangements*-Methode werden zwischen den Gruppen verglichen und die relevanten Attribute in den Anordnungen quantifiziert.

Teil 3 - Planare Einbettung überlappender Cluster
Im letzten Teil der Arbeit wird ein theoretisches, NP-vollständiges Problem aus dem Bereich des Graphenzeichnens betrachtet. Die zentrale Frage ist, ob es möglich ist, einen planaren Graphen mit – durch Regionen gekennzeichneten (überlappenden) Clustern – so zu zeichnen, dass

- die Region jedes Clusters genau die Knoten des Clusters enthält,
- eine Kante eine Clustergrenze maximal einmal schneiden darf,
- die Region eines Clusters, der eine Teilmenge eines anderen Clusters ist, innerhalb der Region des enthaltenden Clusters gezeichnet wird und
- jede zusammenhängende Region aus der Schnittmenge von zwei Clustern mindestens einen Knoten enthält.

Es wird gezeigt, dass das Problem NP-schwer bleibt, auch wenn die Clusterung aus zwei Partitionen zusammengesetzt ist. Im Gegenzug wird ein Algorithmus vorgestellt, der das Problem für zwei Partitionen in Linearzeit löst, wenn jeder Cluster und sein Komplement zusammenhängend sind. Für den generellen Fall (ohne die Bedingung der Partitionen) wird ein Polynomialzeit-Algorithmus präsentiert, der unter der Bedingung funktioniert, dass jeder Cluster einen zusammenhängenden Graphen induziert.

Contents

Introduction

The study of relational data as a means to model and understand complex systems has become increasingly popular over the recent years. Methods to reconstruct, analyze, and visualize networks helped providing new insights in a large variety of disciplines in both social and natural sciences as well as business and finance. This has led to the proclamation of *network science* as a new field of research whose aim is to "study of the collection, management, analysis, interpretation, and presentation of relational data." (Brandes et al. 2013, p.2)

In this thesis, we focus on various aspects of the reconstruction and visualization of network data in an archaeological context with a regional focus on the Caribbean. We present methods to

- analyze travel times in canoes, allowing to reconstruct a network of (potential) relationships between settlements on different islands,
- measure perceived distances and similarities between ceramic sherds and styles of pottery, opening up the possibility to establish typologies based on these distances, and to
- generate crossing-free embeddings of planar graphs with overlapping clusterings that can lead to a more readable representation of complex data.

Note that – despite the focus on archaeological data – this thesis is written from the perspective of a network and computer scientist. We will thus set our focus on the development of methods rather than their application to specific (archaeological) questions. Nevertheless we will conduct pilot studies to put our methods to the test whenever possible.

Archaeology

Archaeology as a field of research investigates past human activity and cultural practices through the study of material remains. It supplements history by providing an additional layer of information about the past. Material culture is used to validate or falsify written sources and to investigate prehistory[1]. Due to the large variety of research questions and methods, the field of archaeology is inherently cross-disciplinary. Methods from both social sciences and the humanities are used to analyze data collected in archaeological surveys and excavations.

[1]The time before the invention of writing systems

While early archaeologists in the 19th century mainly focused on finding interesting objects, putting little thought into the process of excavating, modern day excavations produce a significant amount of data. This includes the exact location where an object was found, its vertical layer in the ground, and information on the surrounding soil, to name just a few of the contextual attributes recorded for the objects. With the analysis of the finds in the lab, the amount of data increases even more. However, even with the advent of sophisticated technical and scientific methods, such as provenance studies based on isotopic compositions and DNA-analysis of the bacteria entrapped in centuries-old dental calculus, the nature of archaeological data remains inherently incomplete. In addition to this, a classical excavation only provides a glimpse into the life of people at one particular site during one particular time (Greene and Moore 2010; Henson 2012; Orton and Hughes 2013).

Looking for new ways to extend the analysis to more than just single sites, archaeologists recently adopted methods from network science to reconstruct relationships between multiple sites in an area (Brughmans 2010). By investigating interactions of ancient people(s) across multiple locations, archaeologists can infer information about migratory patterns and practices of exchanging goods and ideas. A classical case where it is possible to reconstruct pathways directly based on material evidence is the Roman road network whose remains can be found all across Europe (Forbes 1993). Unfortunately, in many other cases it is impossible to examine the direct evidence for ties between sites, even more so when the exchange relations involved seafaring. In these cases, archaeologists have to rely on proxy data to infer (possible) relationships.

This thesis focuses on the reconstruction and visualization of archaeological network data from the area of the Caribbean. In the following, we give a brief introduction to the region and the historical context of this thesis.

The Caribbean

The Caribbean is a very diverse region, both geographically and culturally. Being an archipelago, its geography ranges from flat, sandy limestone islands to high active volcanoes located in and around the Caribbean Sea (see Fig. 1). Due to the tropical climate, the islands are characterized by a rich flora and fauna, providing abundant food sources. Apart from the occasional occurrence of natural catastrophes, such as volcanic eruptions and hurricanes, the islands provided – and still provide – very favorable living conditions.

The first traces of human activity date back to finds from about 5500 BC on Trinidad, and in the following 1000 years, settlements were established on most of the islands in the chain all the way up to Cuba. The archaeological evidence shows continuous change of cultural practices during the seven millennia following the initial arrival of humans in the area. Different groups of people migrated into the area, bringing in new traditions which were partially blended into existing practices (Keegan et al. 2013).

What can be assumed to be the biggest disruption in the history of the Caribbean was triggered by the “discovery” of the new world by Columbus in 1492 (Mann

Figure 1: The Caribbean archipelago

2011). In the years succeeding the arrival of the first Europeans, the beginning of colonization of the area as well as the importation of the first slaves from Africa resulted in a "clash of cultures" that was unprecedented in human history. For the first time, people(s) from three different continents were living in the same (very confined) area of the Caribbean islands. The story did not end well for the Amerindians: Enslavement and diseases led to a rapid decline of the indigenous populations in the subsequent years. Since there are very few written sources about the life and culture of the indigenous people, and most of these sources are biased by a Eurocentric point of view, we have to rely on archaeological finds to tell the story of these people.

Early archaeologists studying the area followed the simplistic assumption that cultural change occurred sequentially, i.e., one culture was simply replaced by the next wave of immigrants or by the development of new practices and technologies. However, more recent archaeological evidence suggests that processes of change were much more complicated, including transculturation and accommodation (Keegan and Hofman 2017). The study of these complex processes is the aim of *NEXUS1492*.

NEXUS1492

This thesis was written in the context of the *NEXUS1492* project[2]. The goal of this ERC-Synergy research project is to "investigate the impacts of colonial encounters in the Caribbean" and to "address intercultural Amerindian-European-African dynamics at multiple temporal and spatial scales across the historical divide of 1492". With the aim to promote synergies between different fields of research, the international and interdisciplinary project brings together researchers from the fields of

[2] `http://www.nexus1492.eu/`

archaeology, geology, heritage management, and network science. In this thesis we will focus on the aspects of network reconstruction and visualization.

Drawing from this interdisciplinary approach, this thesis combines results from a large variety of fields, using methods from oceanography, meteorology, operations research, game design, cultural anthropology, statistics, and (theoretical) computer science. As archaeology is about the study of past human activity and behavior, one emphasis of this thesis is to validate findings and assumptions through perception and user studies.

Structure of the Thesis

The content of this thesis is organized into three parts, examining different aspects of the reconstruction of network data between sites and artifacts as well as the visualization of (clustered) networks. All three parts contribute to the goals of the *NEXUS1492* project and can be applied to analyze and visualize data from excavations and other studies within the project. At the same time, each part forms a cohesive unit that can be read independently from the rest of the thesis.

In Part 1, we will explore methods to analyze potential pathways of human migration and exchange across the Caribbean Sea by calculating optimal canoe routes. We analyze the environmental factors that had an influence on the speed and success of voyages in canoes. We examine historical and state-of-the art approaches in the calculation of optimal routing of ships. Based on this overview, we compare two different methods, adapting them to the special characteristics of canoes in an area with many obstacles (i.e., islands) and give an overview of the results of case studies conducted with the methods presented. Part 1 is concluded by a user study in which we examine the learning behavior of players of a canoe simulation game to see if human actors are capable of finding the optimal routes.

Part 2 of this thesis proposes an alternative approach to the classification of pottery that is based on quantitatively measured human perception instead of a subjective decision for associating a style. We begin by giving an overview of the analysis and classification of pottery and how ceramic sherds are recorded at Leiden University. We then present methods from the field of cultural domain analysis and propose our own methods of quantifying knowledge and perception tailored to research questions regarding archaeological pottery. We put our own as well as established methods to the test with three case studies: by comparing our methods to an established method from cultural domain analysis, by using an established method in an online survey to measure the perceived importance of attributes, and through a hands-on study with real sherds in the US-Soutwest where we compare the perceptions of archaeologists and (indigenous) potters.

In Part 3, we study a problem from the field of graph drawing, a research area in theoretical computer science that deals with optimal embeddings of clustered planar networks. We show that the problem is NP-complete, i.e., that it is unlikely that there is an efficient algorithm that can find an optimal embedding for all instances, even when the clusterings are restricted to two partitions. We then show that there are efficient solutions for other special cases. While being a rather theoretical

problem, the generation of such embeddings can help create meaningful and readable visual representations of archaeological data.

Publications and Presentations

Parts of this thesis have been published and/or presented at various conferences. Other parts are based on results from Bachelor and Master theses that I supervised at the University of Konstanz.

Part 1 The algorithms for the calculations of shortest paths in Section 1.4 have been developed in joint work with Jan Hildenbrand who implemented and compared them for his Master thesis (Hildenbrand 2015). The game in Section 1.5 has been developed by Roman Lamsal. He conducted the experiments and an analysis of the results as part of his Bachelor thesis (Lamsal 2017). We supported Emma Slayton's case studies on canoe routes by providing the necessary software and technical advice. The results of the case studies have been presented at the *XXVI Congress of the International Association for Caribbean Archaeology* in St. Martin (Slayton et al. 2015), at the *44th Annual Conference on Computer Applications and Quantitative Methods in Archaeology* in Oslo (Slayton et al. 2016), at the *Annual Conference of the European Association of Archaeologists* in Maastricht (Slayton et al. 2017a), and at the *45th Annual Conference on Computer Applications and Quantitative Methods in Archaeology* in Atlanta (Slayton et al. 2017b).

Part 2 The results of Section 2.5 have been presented at the *XXXVI Sunbelt Conference* in Newport Beach (Athenstädt et al. 2016). Methods for the reconstruction of similarity networks from assemblages based on a typology of sherds have been presented at the *2014 Connected Past* conference in Paris (Habiba et al. 2014) and will be published in the *Journal of Archaeological Science* (Habiba et al. 2018). The results in Section 2.7 have been presented at the *SAA's 83rd Annual Meeting* in Washington, DC (Athenstädt et al. 2018), and together with Lewis Borck, we are currently working on a journal publication about the study.

Part 3 The results in Section 3.3 are based on a presentation at the *22nd International Symposium on Graph Drawing* in Würzburg. The content is published in the proceedings of the conference (Athenstädt et al. 2014). The remaining results in Part 3 are published in the *Journal of Graph Algorithms and Applications* (Athenstädt and Cornelsen 2017).

Part 1

Distances at Sea: Canoe Routing

The chain of islands forming the Greater and Lesser Antilles extends over several thousand kilometers and the islands are separated by channels of open water with widths ranging from a few up to a hundred kilometers (see Fig. 1). Archaeological evidence shows that these channels were not only crossed once during the population of the islands, but that there must have been a regular exchange of material goods over large distances across the whole region.

Due to the dense vegetation on the islands and the lack of transportation infrastructure on land, canoe voyages across the channels (but also between coastal locations on the same island) have most likely been the main means of transportation in pre-Columbian times. This hypothesis is supported by the fact that there is often a higher similarity in material culture between settlements on different sides of a channel than between settlements on different sides of the same island (Hofman et al. 2010).

In this part of the thesis, we explore methods to analyze pathways for traveling by canoe between the islands in the Caribbean. We base our analysis on the character-

Figure 1.1: Replica of a Caribbean canoe (picture by Maria Neubauer, from Billard and Bérard (2011), reprinted with permission)

istics of pre-Columbian canoes and the environmental factors of currents, wind, and waves. We begin with a brief overview of the field of island archaeology and known facts about prehistoric canoe travel in the Caribbean (Section 1.1).

In Section 1.2, we will examine the available sources for environmental data. We start by describing the climatic circumstances and large-scale weather patterns around the Caribbean islands. We then focus on the models we use for obtaining the environmental data and give an overview of how the underlying methods work. An explorative analysis is carried out to find general trends in the data indicating seasonal patterns.

We then give an overview of the history and state of the art approaches to finding optimal routes for ships based on environmental factors and show how these approaches have been used in the field of archaeology (Section 1.3).

In Section 1.4 we evaluate methods for calculating optimal pathways for canoe voyages under the influence of the given environmental factors. We describe two different methods of calculating optimal pathways on anisotropic surfaces that we adapted to suit the characteristics of canoe travel in a region with many islands and compare them against each other. We also give a summary of the results of three case studies carried out with one of the methods, simulating voyages in different regions of the Caribbean during different seasons of the year based on the current and wind situation. For one of the simulated tracks we perform a more detailed analysis of the variability of travel times within and across different months of the year.

We conclude the part of the thesis with a user study to investigate to what extent it is possible for human actors to determine the optimal pathways by repeatedly paddling a certain route (Section 1.5). For the study, we developed a paddling game as a (mobile) application. The users can simulate paddling a canoe by swiping the touch screen on a mobile device or using the keyboard on a desktop machine. For the study, we set up three different missions, each with a static and a variable current scenario. We evaluate how users chose their paths under the influence of both constant and variable currents and compare the results to the optimal algorithmic results from the previous section.

1.1 Prehistoric Canoe Travel in the Caribbean

The archaeological evidence suggests that canoe travel was an essential part of human mobility and interaction within and between prehistoric Caribbean peoples (Hofman et al. 2010). As the majority of the geographic area is covered by water, travel by sea was the only way for people to move and exchange goods or ideas between different islands.

Early archaeologists often ignored the spaces of water between islands and saw islands as separate units in which goods were imported and exported without focusing on the details of this process (McNiven 2008). In recent years, together with the introduction of approaches from network science to the field of archaeology, the focus has been shifted to also examine the ties between the nodes, i.e., the interactions

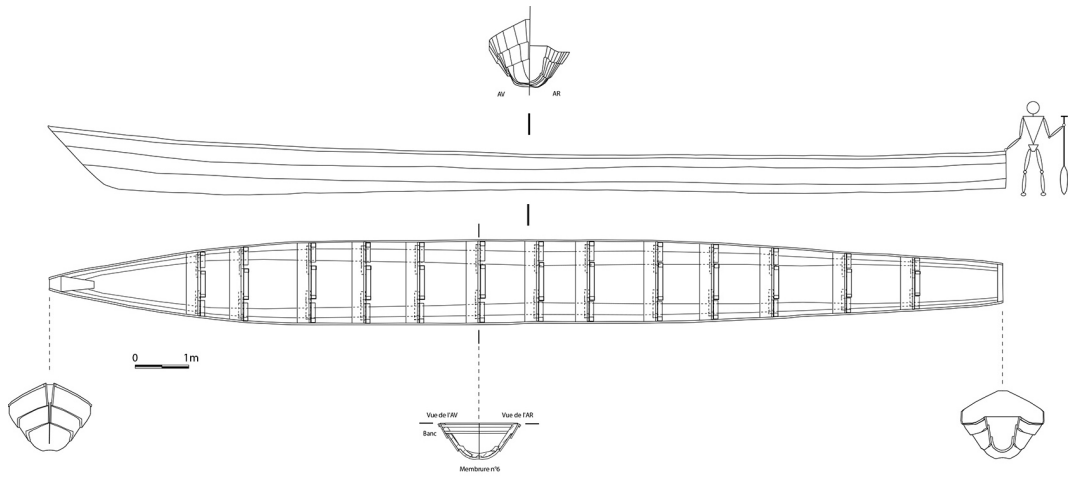


Figure 1.2: Drawing of a canoe used for experimental voyages by the Karisko project (image by Bérard et al. (2016), reprinted with permission)

between different communities. This has led to an increased interest in seascapes as the principle corridors of interaction among archaeologists (Boomert and Bright 2007; Hofman et al. 2007).

Yet the ease of use of these corridors heavily depends on the conditions of the sea surface, which in turn is influenced by the interplay of the environmental factors of currents, wind, and waves. Besides the use of the sea as means of transportation, food remains such as shells but also bones of larger fish indicate that the ocean was also a rich resource whose exploitation was part of everyday life.

While technological advances in archaeometry[1] greatly facilitated the analysis of the provenance of objects, this archaeological evidence is generally limited to reveal the origin of an object, adding this information to the knowledge about its place of deposition. The trajectories of an item and locations in which the item was used in between remain mostly unknown.

The study of possible canoe routes can help indicate potential points of contact along the direct route and give a general idea of average travel times and their seasonal variations. Of course there remain many uncertainties in the methods and models used, so findings will have to be compared against archaeological evidence and alternative theories.

Unfortunately there is little archaeological evidence of pre-Columbian canoes, as organic materials such as wood are decomposing very quickly in the climate of the Caribbean (Callaghan and Schwabe 2001). Most of the (few) fragments found stem from smaller vessels. However historical sources such as observations by Columbus himself suggest that sizes of vessels varied from small one-person canoes to large

[1]The application of scientific techniques to the analysis of archaeological materials

vessels carrying up to 45 persons (Hulme and Whitehead 1992). Some historical reports even mention canoes of lengths of up to 96 feet, carrying up to 80-100 people (Fitzpatrick 2013). However, these vessels must have been the exception and were likely to only have been used for very specific occasions (McKusick 1960).

Most sources agree that canoes were constructed from a single log that was dug out using fire and hand tools and spread open to achieve a wider beam. There is no textual evidence of an extension of the freeboard with the help of planks (Fitzpatrick 2013), yet experiments with reconstructions of canoes suggest that canoes without additional planks on the sides are very vulnerable to capsizing (Bérard et al. 2016).

Due to the lack of archaeological evidence, experimental archaeology plays a vital role in learning about construction techniques and the behavior of canoes at sea. The *Karisko* project[2] is an organization based in Martinique that reconstructed several voyaging canoes based on archaeological evidence and historical accounts from the Caribbean and the Amazon area (see Fig. 1.1 and 1.2). The project regularly organizes experimental canoe voyages with volunteers and thus can provide valuable information about the capabilities and limitations of large voyaging canoes and their behavior at sea (Bérard et al. 2009; Billard and Bérard 2011; Bérard et al. 2016).

For a more in-depth description of indigenous canoe travel, we refer to the PhD-thesis by Emma Slayton (2017).

1.2 Analysis of the Environmental Factors

Before we can analyze the pathways between the islands, it is crucial to get a general overview of the underlying data. We will first provide an introduction to the climatic situation in the Caribbean (Section 1.2.1). In this context, we explain the seasonal changes in weather patterns and how they influence wind, currents and waves. These are the environmental factors that must have had the most influence on prehistorical canoe voyages. We also briefly comment on potential changes in the climatic conditions between the times we are studying and today (Section 1.2.2). We then examine the predictive models that provide the environmental data for this study (Section 1.2.3). In the last step, we explore ways of aggregating the data from the model and finding patterns and trends (Section 1.2.4).

1.2.1 Overview of the Climatic Situation in the Caribbean

The weather in the Caribbean can be classified into two main seasons: The dry season (February to June), with only occasional rain showers, and the wet season (July to January) in which rain is more frequent and stronger. Temperatures stay mostly constant over the year at around 25° to 30° Celcius. Most relevant for our work are the changes in currents, wind and waves. We will now describe their general patterns in the area of the Carribbean Sea. In Section 1.3.3, we provide further

[2] `www.karisko.com`

details about the history of collecting environmental data and their aggregation into *pilot charts*.

We will use the following (established) notations: The direction of currents is referred to by the *direction the current is flowing towards*, whereas winds are named after the *direction the wind is blowing from*. So a northward current flows towards the north, a northerly wind comes from the north. A circular movement (of water or wind) is called *cyclonic* if it is turning in the same direction as the Earth's rotation. As the Caribbean is located on the northern hemisphere, a *cyclonic* movement is counterclockwise and an *anticyclonic* movement clockwise. The side of an object facing the wind is called *windward* side, the side facing away from the wind is referred to as *leeward* side.

A very detailed resource on the climatic conditions and weather patterns in the Caribbean can be found in the *Sailing Directions for the Caribbean Sea* (National Geospatial-Intelligence Service 2015).

Wind Patterns

The winds in the Caribbean Sea are dominated by the northeasterly *trade winds* which are present around the globe at the latitudes close to the equator. When reaching the Lesser Antilles, the winds tend to take a more easterly component, sometimes even southerly in the summer months. This has led to the naming of *Windward Islands* for the southern islands up to Martinique and *Leeward Islands* for the northern islands from Dominica on.[3]

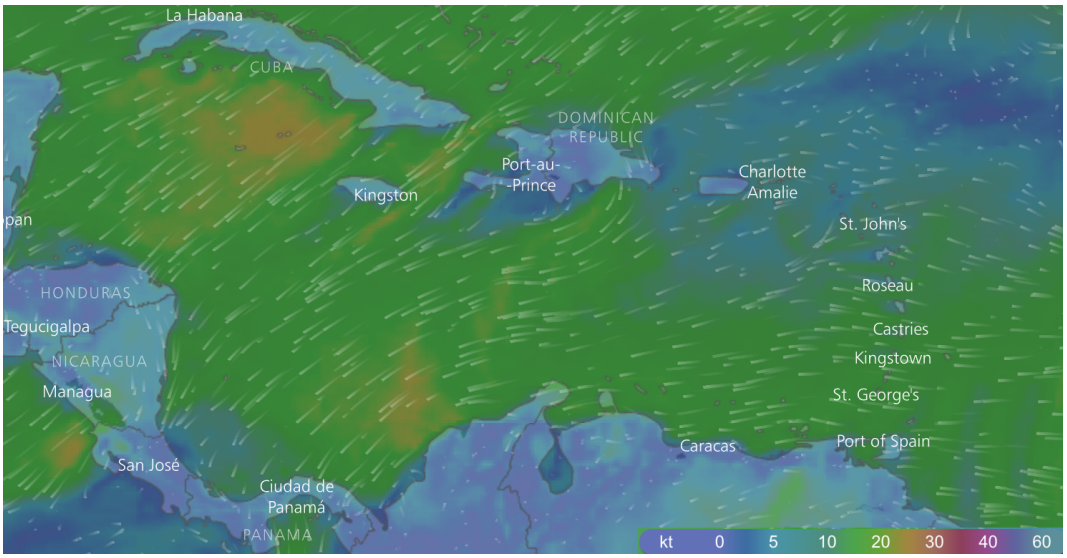


Figure 1.3: A snapshot of the wind situation in the Caribbean Sea on December 3rd, 2017 (screenshot from `www.windy.com`)

[3] In languages other than English, the whole of the Lesser Antilles are referred to as *windward* and the islands along the coast of Venezuela as *leeward* which reflects more the average wind patterns over the whole year

Despite these slight seasonal changes, the prevailing trade winds are very reliable in their direction and strength. For most locations, the wind comes from the same quadrant (90° sector) during 95% of the times observed. The average force of the wind in most places in the Caribbean is around 4 Beaufort, i.e., 5.5 -– 7.9 m/s (National Geospatial-Intelligence Service 2002). Figure 1.3 shows a typical wind situation in the Caribbean Sea during the winter months.

In the middle of the Caribbean Sea, the prevailing winds slightly diverge and become either more southerly, blowing towards the Gulf of Mexico, or northerly, blowing towards Panama in the south.

A note on Hurricanes One of the most infamous weather phenomena in the region are *hurricanes* (and their weaker version, the *tropical storms*). With sustained wind speeds up to 140 knots, these cyclonic storms can cause tremendous damage when reaching an island or the coast. The hurricane season is limited from June to November with the highest likelihood in August and September. There is no doubt that these extreme weather conditions had a major impact on past civilizations (Keegan et al. 2013). However, as these extreme events occur only once every few years at the same location, we do not give them a special treatment in our models. Yet, the environmental factors in the surrounding islands are influenced by hurricanes or tropical storms, which in turn is already incorporated in the data from the models we use.

Current Patterns

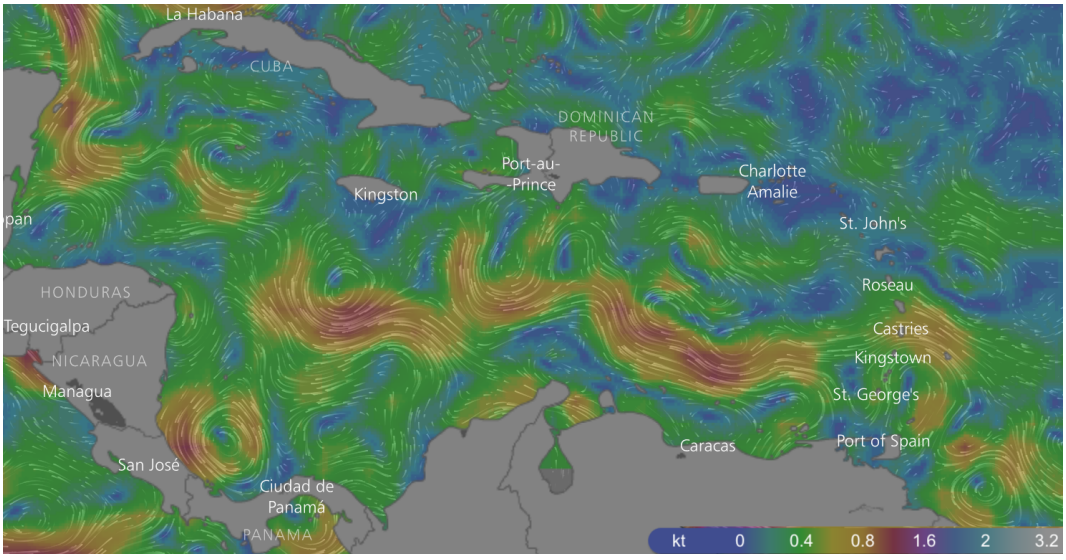


Figure 1.4: A snapshot of the surface current in the Caribbean on December 3rd, 2017 (screenshot from `www.windy.com`)

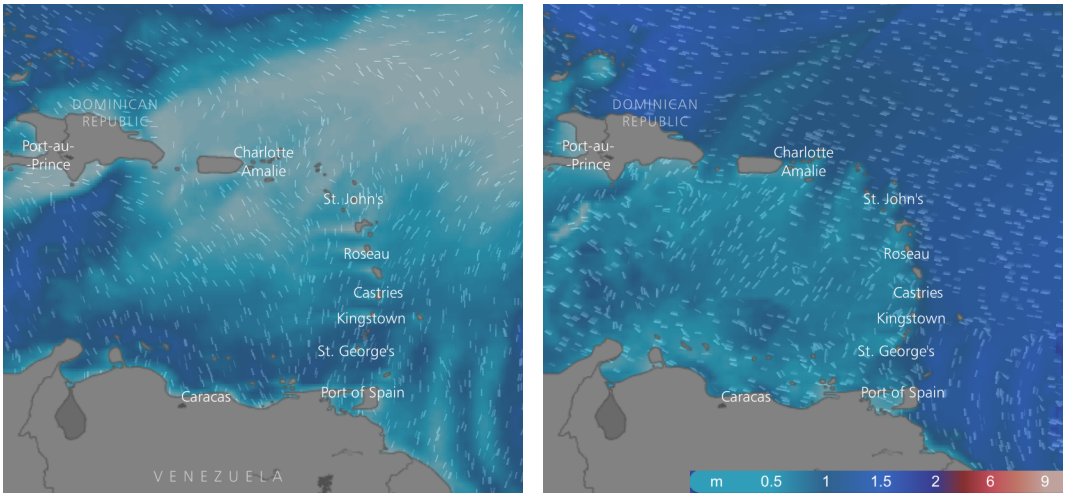
The currents flowing into the Caribbean Sea are mostly dominated by the North Equaritorial Current (NEC) which brings in warm surface water from the Atlantic

that is pushed westwards by the trade winds. This water mixes with freshwater outflow from the Amazon at the coast of South America and gets pushed through the chain of islands of the Lesser Antilles (mostly in the south of the chain). In irregular intervals, water from the South Atlantic resulting from vortices that break at the coast of Brazil is added to the flow. The islands limit the flow of water at deeper levels and cause turbulence, resulting in a series of cyclonic and anti-cyclonic vortices being pushed westward across the Caribbean Sea. Figure 1.4 shows a snapshot of a particular current situation of the surface layer. Note that the situation changes very quickly as vortices move westwards (Gyory et al. 2017; Zaron et al. 2015).

Most of the water leaves the Caribbean Sea through the Yucatan Channel into the Gulf of Mexico. There it forms the anticyclonic *Loop Current* which then exits the gulf between Cuba and Florida and finally becomes the Gulf Stream heading north along the East Coast of the United States and then eastwards towards Europe.

Wave Patterns

Waves are (with very few exceptions) created by wind blowing over a surface of water. The waves that are directly created by the wind are called *wind seas*. Their direction generally corresponds to the direction of the wind. Waves that indirectly result from wind blowing in an area further away or from past wind influence in an area are called *swell*. Figure 1.5 shows the swell and the wind seas at the same date. It is clearly visible how the wind seas are replicating the pattern of the winds (Fig. 1.3), while the swell is independent of the local wind situation. One can also observe that the swell is blocked by the chain of islands, leading to lower wave heights in the Caribbean Sea than in the open Atlantic.



(a) Wind waves **(b)** Swell

Figure 1.5: A visualization of the wave situation for the Caribbean Sea on December 3rd, 2017 (screenshot from `www.windy.com`)

While the wind seas are easy to predict by the prevailing wind patterns described above, the swell can result from winter storms in the North Atlantic or from passing

hurricanes in the summer months. They are thus highly variable and can cause *cross seas* if they come from a different direction than the wind seas, making paddling a canoe without capsizing hard or even impossible. (Benoît Bérard, personal communication, January, 2015).

Even in the relatively protected area of the Caribbean Sea, there are two peak periods in July and in the months of January through March in which significant wave heights above 8 feet occur more than 40% of the times observed (National Geospatial-Intelligence Service 2002), posing a non-negligible obstacle to safe canoe voyages in the area.

1.2.2 Modern vs. Ancient Climate

The question arises whether we can use simulations of the present-day climatic conditions to model ancient travel routes. There is no doubt that there have been changes in the climate of the Caribbean islands during the period of human presence: There is strong evidence for a sea level rise over the last 6,000 years of up to five meters. In addition, movements of the *Inter Tropical Convergence Zone (ITCZ)*[4] have certainly impacted the strengths of the prevailing winds and currents (Cooper and Peros 2010; Fritz et al. 2011). Giry et al. (2013) analyze the salinity of the Caribbean waters by coral records and suggest a weaker current inflow, mainly linked to the tradewinds that only slowly increased over the period of the Holocene[5] due to a southward shift in the ITCZ (Haug et al. 2001). On the other hand, when analyzing the case of the Mediterranean, Murray (1987) compares antique observations of wind patterns to modern simulations and finds general accordance.

Unfortunately, the available data sets for pre-Columbian climate data, such as the *Paleoclimate Modelling Intercomparison Project* (Meinshausen et al. 2011), are very limited both in spatial and temporal resolutions. In contrast, we do have snapshots of model runs for multiple times each day at spatial resolutions well below one degree for recent times. While there have been substantial changes during the Holocene, it is likely that the general patterns and trends we describe in this study have remained at least somewhat similar during the time of human activity in the area. Callaghan (2001) also observes that the bathymetry[6] has not changed significantly over the last 1000 years and concludes that the currents must not have changed much. Following the example of Warnking (2015) who analyzes Roman trade routes, we thus decided to use environmental data from modern times for our explorative analysis in this section and the simulations for the case studies in Section 1.4.3. All methods we describe would also work on ancient data if better simulations should become available in the future.

[4]This is the area around the equator where the highest proportion of the sun's energy gets absorbed. This causes high water temperatures and thus evaporation, heavily influencing the global climate

[5]The current geological epoch which started about 11,700 years ago

[6]The topography of the seabed

1.2.3 Data Sources

In this section, we are going to describe sources for current, wind, and wave data in the region of the Caribbean Sea. We will briefly explain the underlying models and what data is available over which periods of time.

Current Data

The current data we use is the *Amseas* data set[7] provided by the *Naval Oceanographic Office (NAVOCEANO)*. It comprises the entire region of the Gulf of Mexico and the Caribbean Sea. The development of this high-resolution nowcast/forecast system was triggered by the Deepwater Horizon oil spill event in 2010. The data set contains information on water temperature, salinity, and velocity at 40 different levels of depth. We use the northward and eastward velocity data at sea level. The data grid provides a spatial resolution of 1/30 degree of latitude and longitude ($\approx$ 3.7 km) and a temporal resolution of 3 hour intervals. Before April 2013, the spacial resolution was slightly different, and a different model for the boundary conditions was used. The data set is based on forecasts that are updated daily and the archive of the data goes back to 2010[8]. The data set is based on simulations with the *Navy Coastal Ocean Model (NCOM)* (Martin et al. 2009). An evaluation and detailed description of the data and the underlying model has been compiled by Zaron et al. (2015).

The data shows considerable variations in the surface currents not only when comparing monthly averages but also between two 3-hourly snapshots, probably due to tides and the influence of wind. In Section 1.2.4, we examine these variations.

Wind

The standard format for wind data as well as other weather (forecast) data is the GRIB (GRIdded Binary) format[9]. The data for wind forecasts is generally the result of computer simulations and is provided on a grid with a fixed resolution. There are multiple providers of wind data for the region of the Caribbean, however most of them focus on forecasts and charge for their services.

The *QuikSCAT Level 2B Ocean Wind Vectors*[10] provided by *NASA* is a rich data source based on satellite measurements. However, there exist only measurements when a satellite was currently above the given area, resulting in large gaps in the data. The same goes for the *CoastWatch Caribbean* project that provides *Near-Real Time Wind Data* based on measurements from satellites, drifters, and other sources. Even though the corresponding Java applet allows for an easy selection of

[7] `https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/navoceano-ncom-reg`

[8] `http://ecowatch.ncddc.noaa.gov/erddap/griddap/NCOM_amseas_20100509_to_20130404_3d.html`

[9] `http://www.wmo.int/pages/prog/www/DPS/FM92-GRIB2-11-2003.pdf`

[10] `http://podaac.jpl.nasa.gov/dataset/QSCAT_LEVEL_2B_OWV_COMP_12`

the relevant regions and data sets, each time snap only covers part of the region which makes it hard to interpolate in regions with no measurements within the next 200 kilometers.

We thus chose to obtain our data from the *Global Forecast System (GFS)*, provided by the *National Centers for Environmental Prediction*[11] instead. The resolution is lower, but there exists an archive of the model runs with a 0.5-degree resolution from October 2006 on[12]. The GFS is an atmospheric prediction model. It requires an initialization run in which data from weather stations, weather balloons, airports, and satellite observations around the globe are aggregated through finite difference approximations. The result is the 00h-forecast. From then on, the atmospheric movements based on differences in pressure, temperature, and the underlying terrain are calculated for the forecast periods in the future[13]. Obviously, the accuracy of the data declines with the forecast time. Since we are only concerned with the data from past model runs, we can always choose the 00h-forecast, thus working with the most accurate data the model provides.

Wave Data

The height of waves is described by the *significant wave height* H_S or $H_{1/3}$, indicating the mean of the highest third of the waves. Since the distribution of waves can be approximated by a Rayleigh distribution (see Papoulis and Pillai (2002)), it is not uncommon to encounter waves that are more than twice as high as H_S.

Besides their height, the period and direction are very important when studying the effect of waves on a vessel. The waves observed are often a combination of swell and wind seas with different heights, directions, and periods (see exemplary illustration in Fig.1.5).

A suitable data source for our simulations would be the wave data generated by the forecast system from NOAA using the *Wavewatch III* model (Tolman et al. 2014) on the NOAA weather predictions[14]. Wavewatch III takes into account the propagation of waves along great circles and local physics of how waves build up and decay. Based on the bathymetry, a full spectral wave field forecast is calculated for each grid point of the model.

The results of the model are provided as GRIB-files. The grid has a resolution of $\frac{1}{6}$ of a degree, i.e., roughly 18 kilometers between two grid points. In the data set, the full spectrum for each grid point has been reduced to height, direction and period of the wind-sea, and the highest and second highest swell fields.

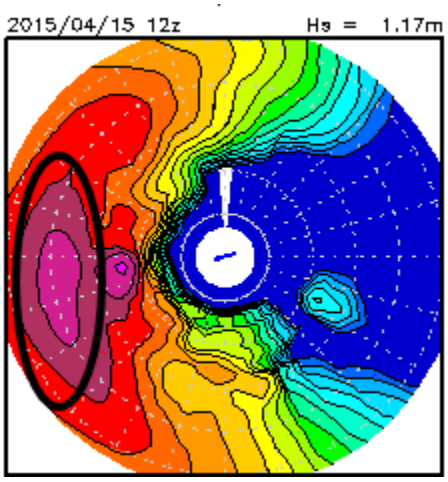
For specific locations, additional text bulletins and full spectral diagrams are provided that show the wave fields in full detail. For each hour of the day, the total significant wave height (Hs) in meters and the number of identified wave fields are

[11] `http://www.nco.ncep.noaa.gov/`

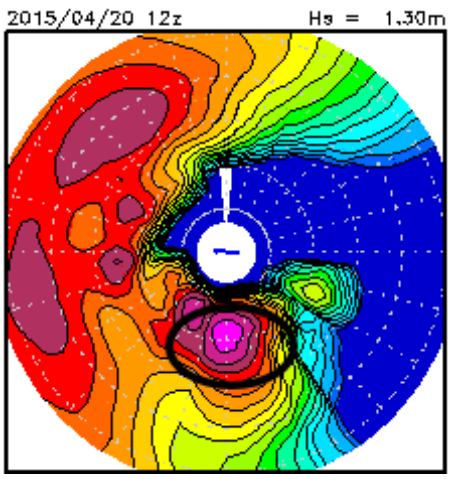
[12] `https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs`

[13] See `http://www.wcsu.edu/weather/0207.ppt` for a good introduction

[14] `http://polar.ncep.noaa.gov/waves/viewer.shtml?-multi_1-gmex-`



(a) Encircled area indicates waves with a short period and a rather irregular pattern in westerly direction



(b) Encircled area indicates waves with a long period and a more regular pattern in southerly direction

Figure 1.6: Example of a full spectral wave field forecast for a buoy located southeast of Puerto Rico for two different dates. The distance from the center marks the frequency, the direction *away from* the center the direction of the waves. The color scale shows the height of the waves: purple indicates the highest waves, blue the lowest. The significant wave height is given in the top right. (images created with *NWW3 Product Viewer*)

given. For each of these fields, the significant wave height, the peak period in seconds, and the mean direction in degrees is provided.

Two examples of such a spectral diagram for a wave forecast at two different times at one specific location are shown in Fig. 1.6. Each point in the diagram represents a (potential) wave that could be part of the wave spectrum for the location. The direction of the point *away* from the center indicates the direction of the wave. The distance of the point from the center indicates the frequency. The frequency near the center of the plot is 0.04 Hz (a wave period of 25 seconds, indicating longer waves), the one at the outside boundary 0.25 Hz (a wave period of 4 seconds, indicating shorter, steep waves). The color scale indicates the height of the wave: Purple are the highest waves, blue the lowest. The total significant wave height is indicated in the top right of the plot.

The example in Fig. 1.6a shows the typical wind seas caused by the easterly trade-winds running towards the west. They have high frequency, and the larger size of the peak (i.e., the encircled purple area) indicates a large variation in both frequency and direction. In Fig. 1.6b, we can see swell in southerly direction, potentially resulting from a storm in the north Atlantic. The period of these waves is longer and the direction more uniform as indicated by the smaller extent of the encircled purple area and its position closer to the center.

In order to capture the extreme events of hurricanes properly, the model is run twice, with the second run based on a high-resolution simulation of the wind fields within

the hurricane. The results are then blended together.

Two interesting tutorials on how to work with wave data from buoy-measurements[15] and the Wavewatch III model[16] provide an introduction to the potential but also the limitations of the wave data from the sources mentioned.

1.2.4 Examining Seasonal Patterns

In order to get a better understanding for the changes of the environmental factors over the year and whether there are certain seasons that are more suitable for voyaging, we developed a tool to plot the changes in currents at a fixed location over the available time period of six years. We focus on the currents here as they have the largest influence on the canoe's trajectory. The method could be easily adapted for the analysis of wind data and with some modifications also for wave data when limiting the analysis to the highest swell field.

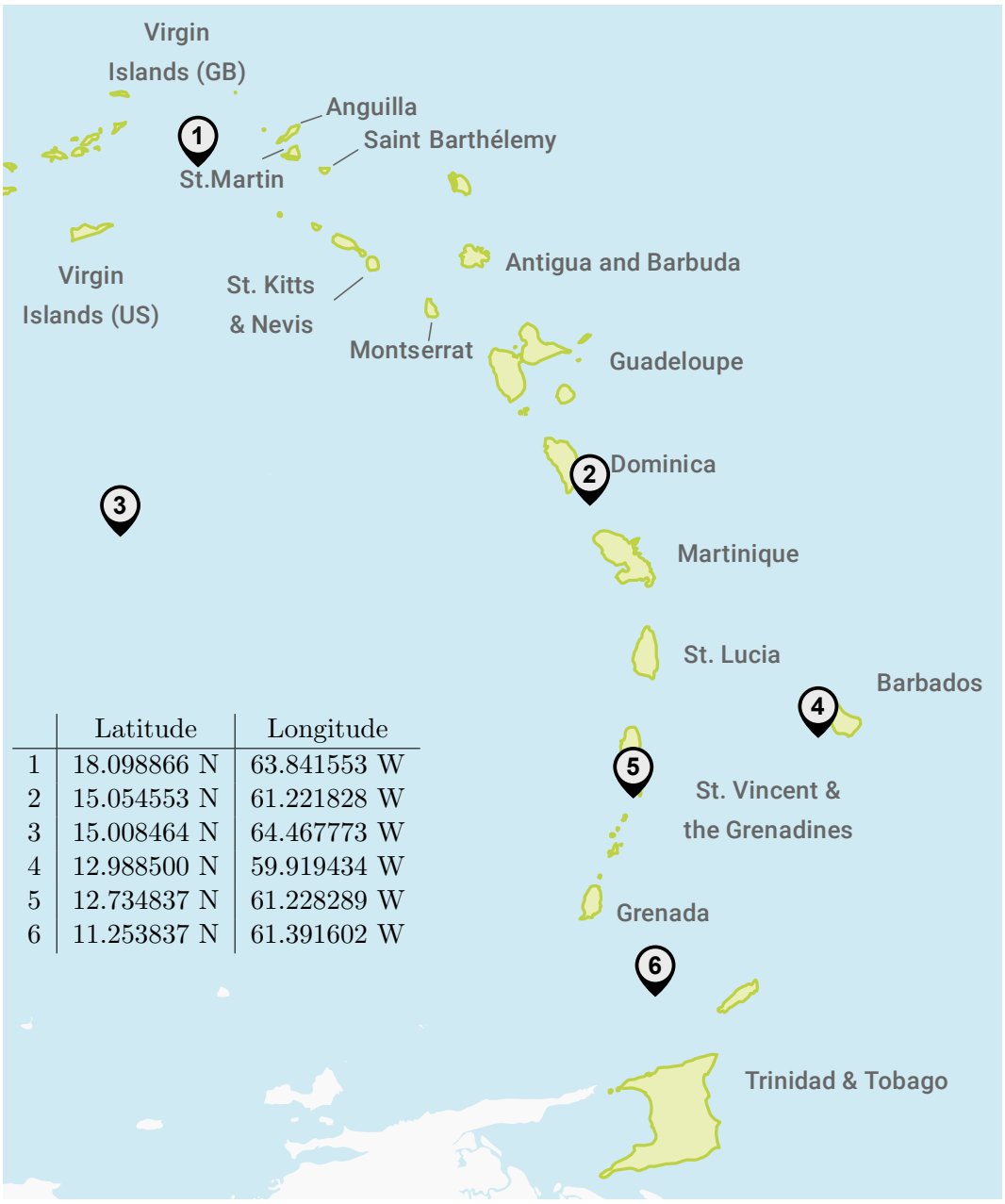
For the analysis of the current patterns, we chose six locations in and around the Lesser Antilles for a closer examination (Fig. 1.7). Location 1, 2, 5, and 6 are chosen in the middle of channels within the chain of islands, i.e., at places where there must have been significant canoe traffic. Location 4 is chosen close to Barbados which is an outlier of the chain and not visible from any other island. Yet, there was a significant indigenous population on the island and regular exchange with the other islands (Beckles 2006). Location 3 is set in the middle of the Caribbean Sea, and thus far from any routes that are likely to have been taken frequently. However, it has been hypothesized that the Greater Antilles might have been settled directly from the mainland of Venezuela (Callaghan 2003b). We thus included this point into our analysis as well.

Detailed Plots for Specific Locations

For our analysis, we first calculate the *main direction* of the currents by taking the average over all vectors. We then plot the current for each of the available data points in 3-hourly intervals. On the x-axis, we plot the time and on the y-axis the strength of the current. The color of the dots indicates the direction of the current in relation to the main direction, indicated by a grey dot. A red dot indicates a deviation to the left (counterclockwise), a green dot indicates a deviation to the right (clockwise), and blue dot indicates currents that are opposing the average direction.

As an example, Figure 1.8 shows the strength and direction of the currents at location 5 in the Grenadines for all available data points for the months of July through September in the years 2010 through 2016. The arrow at the top indicates that the prevailing direction for currents at this location is northwesterly. Quite regularly, there are changes towards the north (as indicated by the green dots), and occasionally and for shorter periods deviations towards the west or even southwest (as

[15] `http://www.meted.ucar.edu/oceans/ocean_swell/`
[16] `http://www.meted.ucar.edu/oceans/wavewatch3/`



	Latitude	Longitude
1	18.098866 N	63.841553 W
2	15.054553 N	61.221828 W
3	15.008464 N	64.467773 W
4	12.988500 N	59.919434 W
5	12.734837 N	61.228289 W
6	11.253837 N	61.391602 W

Figure 1.7: The selected locations for which we examined the seasonal variation of the currents

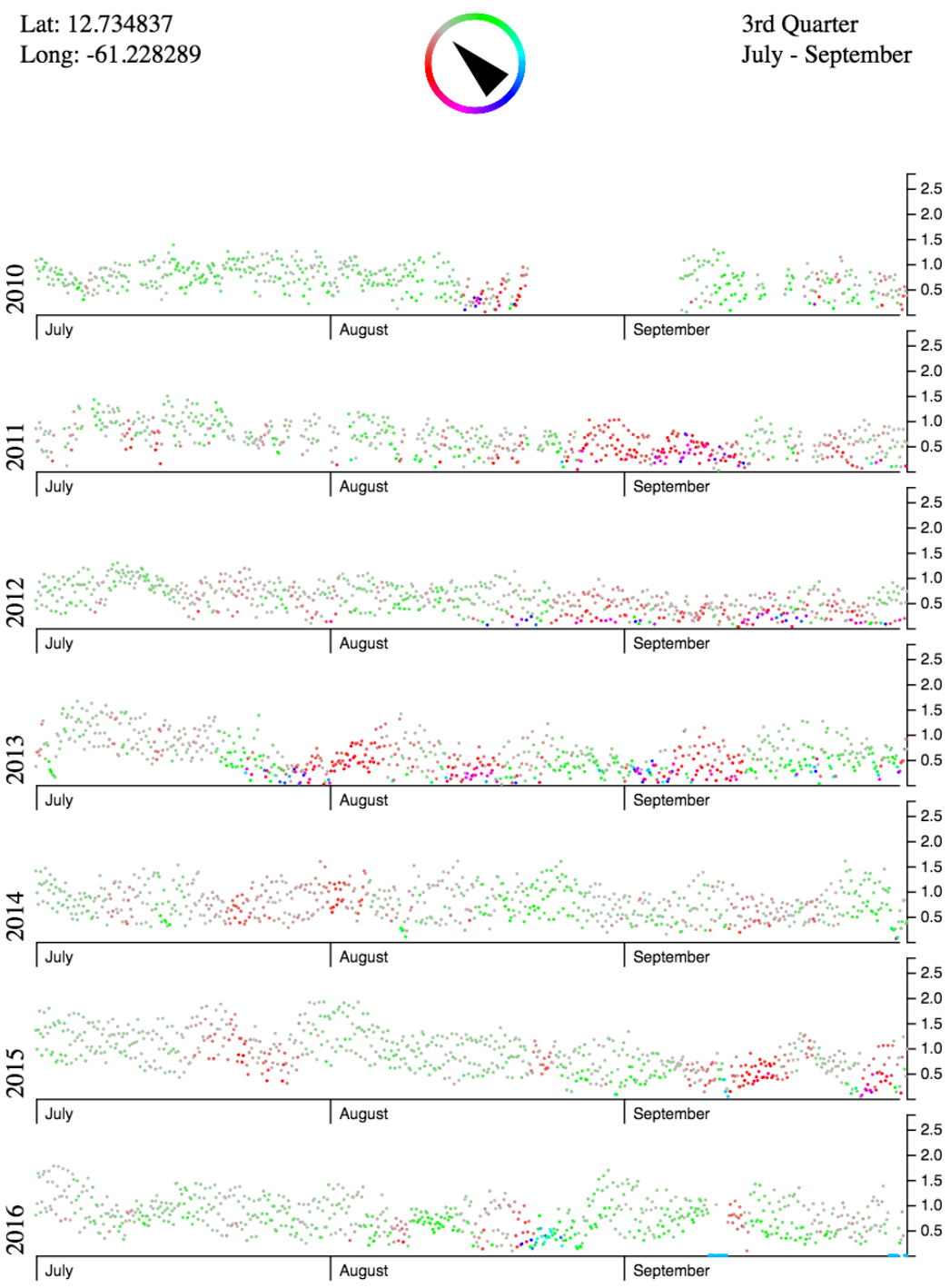


Figure 1.8: The plot of all available data points for currents in the Grenadines (Location 5) from July to September.

indicated by the red and purple dots, respectively). The strength of the currents varies from 0 to 2 m/s (i.e., close to 4 knots). Thus, it is crucial to take the currents into account when paddling a canoe at speeds around 3 knots in the area, as at certain times it will be impossible to paddle directly against the currents in order to reach certain destinations.

Checking for Periodicity

At some locations, we noticed a seemingly regular pattern in the current strength. In order to check whether there is indeed a periodicity (that could be due to tidal patterns but could also be an artifact of the model calculations), we fitted a quadratic function to the points indicating the current strength at regular intervals of length l: Given a starting time t_k, we chose the intervals $[t_k \dots t_{k+l-1}]$, $[t_{k+l} \dots t_{k+2l-1}]$, $[t_{k+2l} \dots t_{k+3l-1}]$ and so on.

We used period lengths from $l = 5$ to $l = 240$ datapoints (i.e., 30 days) and used all possible times $t_0 \leq t_k < t_l$ within the interval as starting times.

We then calculated the average coefficient of determination r^2 for each t_k in each interval. The idea is that the average r^2 values of the fitted functions should be significantly higher at a period length where there is a periodicity. Figure 1.9 shows the average as well as the maximum r^2 for each period lengths at location 2. As expected, r^2 is generally lower for longer periods (due to the higher number of values the quadratic function is fitted to). The larger number of starting times for longer periods explains the higher variability in the data (expressed by the larger difference between maximum and average value).

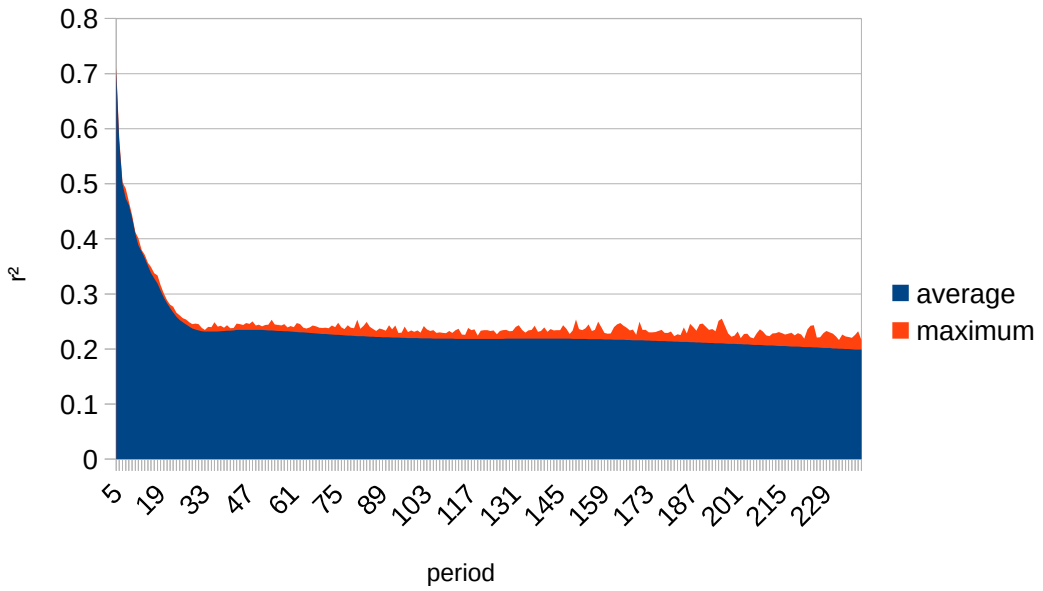


Figure 1.9: Average and maximum coefficient of determination (r^2) over all starting times per period on location 2

Unfortunately, at least for the selected locations, our analysis did not show any hints of a regular pattern in the current strengths. It is possible that such patterns only occur when examining shorter time frames than six years. Yet it appears that the currents are influenced much more by the chaotic system of global weather patterns than by regular phenomena such as tidal forces.

We will thus in the following focus on the larger patterns of seasonal variations to the current strengths and directions.

Exploring Seasonal Trends

The next question we wanted to answer is if there are any seasons in the year that are more suitable for voyages than others. In order to find seasonal variations in the currents, we first take the running average over a 30 and a 15-days period. For the strength of the currents, we average the strength over all data points in the period. For the directionality, we average the easterly and the northerly component of the vector separately. Note that this might not be stable in the case that opposing currents during a time period average out to a 0-vector.

Figure 1.10 shows the 30-days average of the currents at all six locations from Fig.1.7. The arrows indicate that the average directions of the currents are more eastward for the northern three locations and more northwestward for the southern locations. The strongest currents are in the Grenadines (Location 5) with average strengths of

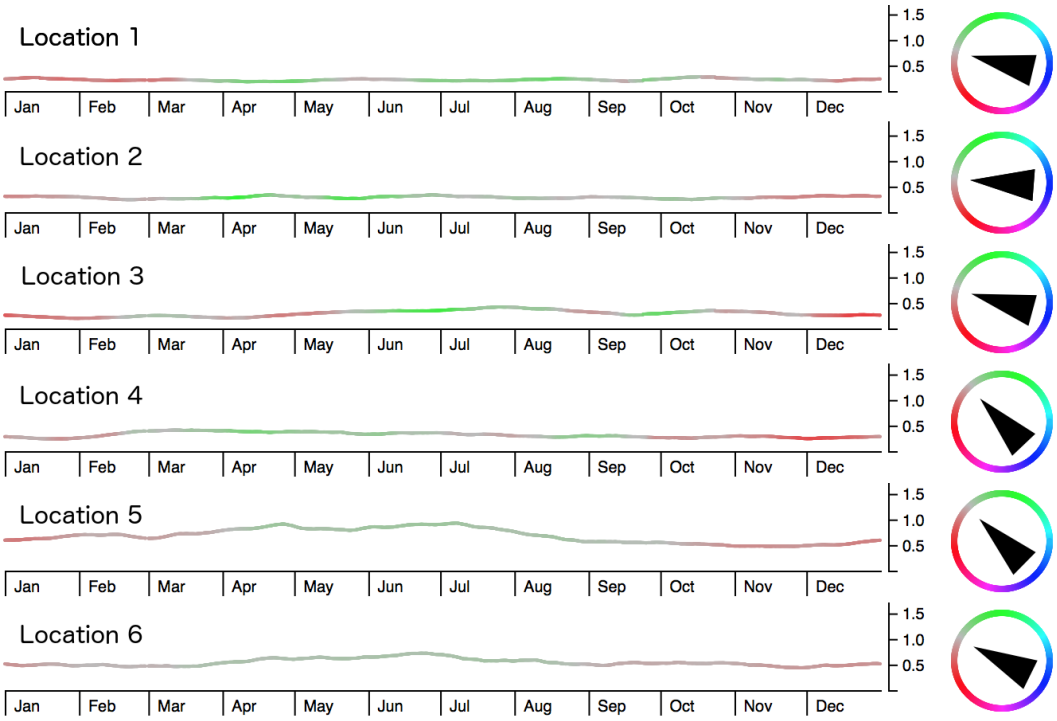


Figure 1.10: The average strength (m/s) and direction of currents at the chosen locations for the years 2010 through 2016 (30-days running average)

up to 1 m/s, followed by the channel north of Trinidad (Location 6) with average strengths up to 0.7 m/s. In these two locations, the strongest currents occur from the middle of the dry season in April through the beginning of the wet season in July and August. At this time, the currents are slightly more northward than the yearly average. During the rest of the year, the currents are slightly more westward, and the average strength decreases to about 0.5 m/s. In the northern locations, the average strength of the currents stays below 0.5 m/s throughout the whole year. Yet there are more variations regarding the average directions. There are periods of a few weeks in which the average direction of the otherwise westward flowing currents turns northward in the summer months and southward in the winter months.

In the case studies described in Section 1.4.3, we did a similar analysis for the results of simulated voyages and describe how the seasonal changes affect voyage times in the southern Lesser Antilles.

1.3 Overview: Optimal Vessel Routing

In this section, we give an overview of methods for calculating shortest paths on anisotropic surfaces where the speed of travel depends on the direction of movement. The methods can be applied to the routing of ships (where the speed over ground depends on the environmental factors of currents, winds, and waves) as well as airplanes (where winds are the determining factor).

According to the International Maritime Organization (2017), more than 90% of the worlds trade is carried by sea, and according to the International Air Transport Association (2016), the number of passengers in air-travel is going to nearly double in the coming 20 years. Thus, there is a huge economic and environmental potential in finding optimal routes for ships and aircraft that goes far beyond our application case in archaeology.

We motivate the search for optimal routes by an anecdotal example in Section 1.3.1. We then give an overview of the navigation technique of dead reckoning (Section 1.3.2) and the origins of optimal vessel routing (Section 1.3.3) as well as a number of state-of-the-art approaches (Section 1.3.4) to the topic. We also give examples where these methods have been applied to answer archaeological questions (Section 1.3.5).

1.3.1 Crossing a River: Duck vs. Dog

There is a thought experiment in popular culture that involves a dog and a duck that have to cross a river to reach some food on the other side. The dog as a land animal swims directly towards the goal. His naive (greedy) approach does not take the current into account and results in a longer and slower trajectory. The duck however, being accustomed to flowing water, swims at an angle against the current, resulting in a straight trajectory across the river that minimizes time and effort.

The sketch in Fig. 1.11 compares the trajectory of the optimal route from the duck, maximizing its speed towards the goal, to the naive approach of the dog who always heads into the direction of the goal. A formal scientific description of the two trajectories has been written by O'Shea (2010).

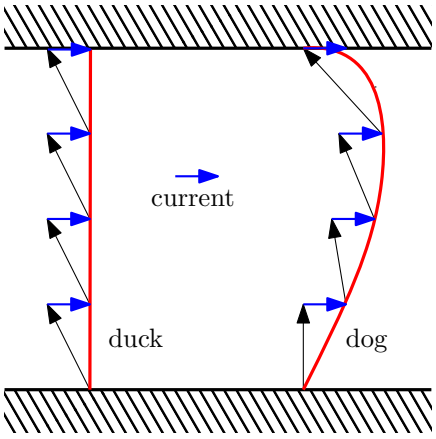


Figure 1.11: The trajectories of duck and dog

While this is a very simplified example, it illustrates well the principle of optimal routing under the influence of currents. If we want to reconstruct pre-Columbian canoe routes between the islands, the main question is whether people behaved more like the duck or more like the dog when choosing their routes.

As the channels between the islands were the only means of transportation at the time, we can assume that people were at least aware of the influence of currents. Especially in simple cases as the one described above, where a constant current runs perpendicular to the desired path, it is likely that people learned to counteract the current for a faster crossing. In Section 1.5 we try to evaluate to what extent human actors are capable of approximating more complex optimal routes such as the ones we will describe in the following sections.

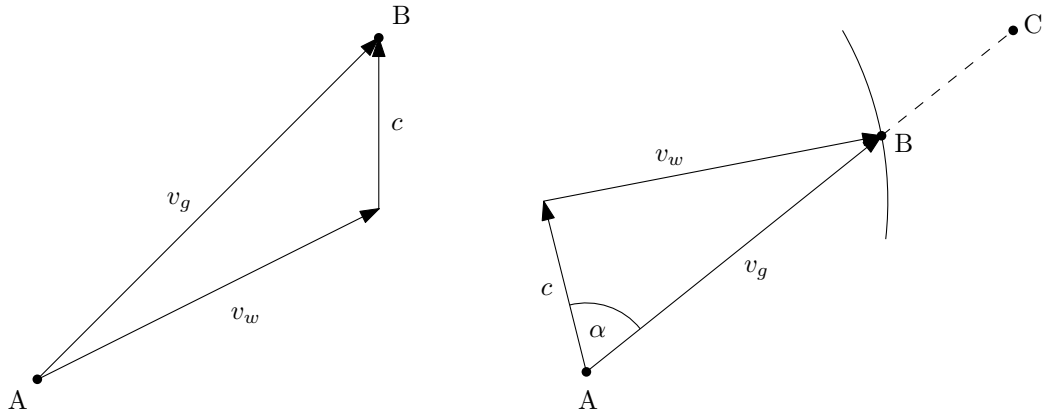
1.3.2 Dead Reckoning

The knowledge about the navigation of ships under the influence of wind, waves, and currents was a crucial skill for sailors since the early days of seafaring. While many of the early civilizations might have followed their intuition and indicative clues that have been passed on from generation to generation (Lusby et al. 2009), the mathematical side of the skill was developed with the advent of nautical charts and descriptions on current and wind patterns.

The basic principle is simple and referred to as *dead reckoning*. According to Bowditch (2002), the term describes the process of obtaining the approximate position of a ship (or airplane) based on the last known position, the speed and direction of the vessel, and the influence of environmental factors (usually currents in the case of a ship and wind in the case of an airplane).

The calculation on board of ships is usually performed directly on a nautical chart by manually adding the vector of the assumed influence by the currents to the vector indicating the course and speed through the water in regular intervals (usually every hour). The sum of the vectors then shows the assumed location B based on the known start point A and the estimated currents (Fig. 1.12a).

Besides calculating the position based on the given course and the influence of the currents, the method also allows to determine the necessary speed and course to end up at a specific location given the influence of the current.



(a) Determining position B based on v_w and c

(b) Determining the lenght of $|v_g|$ in direction of point C based on $|v_w|$, $|c|$, and the angle α

Figure 1.12: Principle of dead reckoning: the basic relationship between the current c (or other environmental factors), the steered course through the water v_w, and the course over ground v_g

In turn, most of the algorithms we describe in the following rely on the question of finding the maximum possible speed over ground $|v_g|$ in a given direction. Given the fixed speed through the water $|v_w|$ and a given current vector c, this problem can either be solved graphically with compasses on a chart (Fig. 1.12b) or via the law of cosine:

$$|v_w|^2 = |c|^2 + |v_g|^2 - 2|c||v_g|cos(\alpha)$$

The formula can be transformed to the following equation for finding the speed over ground $|v_g|$:

$$|v_g| = |c|cos(\alpha) + \sqrt{|v_w|^2 - |c|^2 - |c|^2cos^2(\alpha)}$$

All methods for optimal vessel routing we will present in the remainder of this section are based on this simple relationship.

1.3.3 The History of Vessel Routing

The optimal routing of ships based on environmental factors has been an area of interest for centuries. It is quite clear that all early civilizations, as soon as they had seaworthy boats, must have had at least a basic understanding of prevailing wind and/or current patterns that helped them to safely navigate to distant places (Warnking 2015). However, despite this general knowledge, the roots of systematical scientific research in this field were only laid in the mid-19th century. During the great age of trade under sail in the 19th century, the duration of a voyage started to have significant impact on the profit of the ships operator, and fast voyaging thus became a crucial factor to be competitive in the business of transportation.

The Origins: Pilot Charts

The pioneer of the systematical study of ocean currents and wind patterns was Matthew Fontaine Maury (1806-1873), lieutenant in the US navy, and later called the "Pathfinder of the Seas" (for his biographies, see Lewis (1980) and Hearn (2003)). Based on the log-books of thousands of sailing ships, he published his *Wind and Current Charts* in 1847, showing the ship's trajectories and observed wind and current strength and direction, color-coded by season[17] (see Fig. 1.13). Black dots indicate observations in winter, green in spring, red in summer, and blue in autumn. The "feathers" at the dots indicate the direction the wind was blowing to, the narrower the angle of the feathers the stronger the wind.

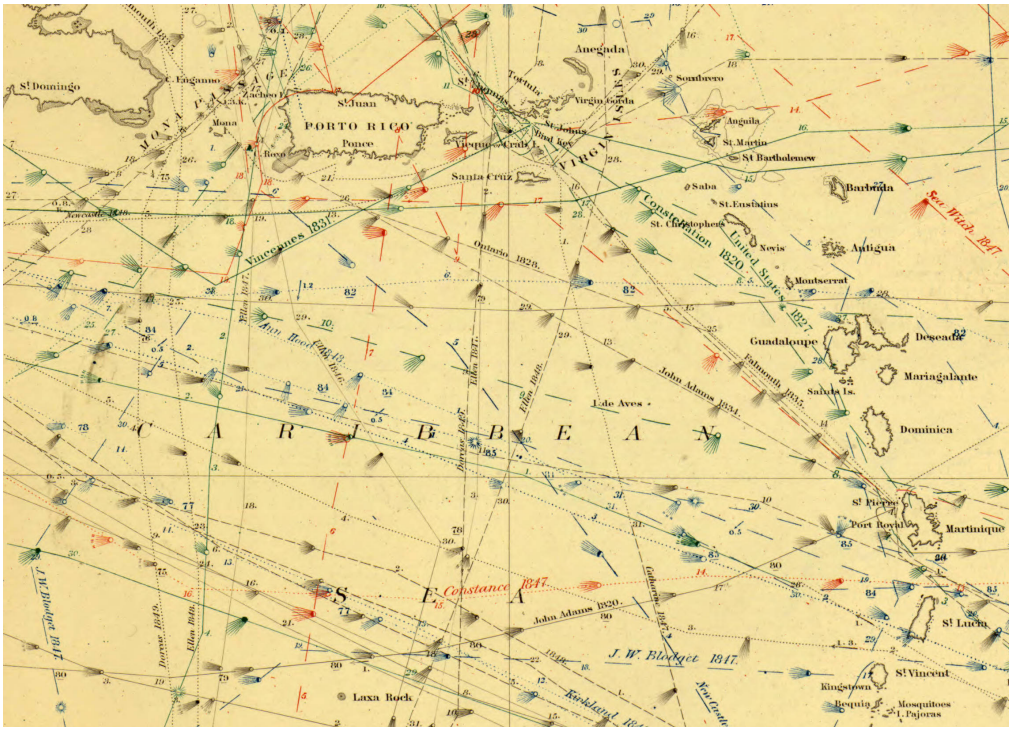


Figure 1.13: Excerpt of a *Wind and Current Chart* by Maury, summarizing the log book entries of all ships sailing in the area of the Caribbean

Maury later accumulated the data for grid cells of 5° latitude/longitude[18], showing the statistical frequency of certain winds for each month. Known as *Pilot Charts*, they quickly became famous among ship owners and captains, as they helped to considerably reduce travel times. For example, the average time for a trip from New York to California around Cape Horn went down from 183 to 139 days with the use of recommended seasonal routes (Bowditch 2002)). However, optimal routes based on these charts were mostly created by "eyeballing" paths that looked promising or were used by many ships before on fast voyages.

[17] The original chart in high resolution can be found at `http://collections.lib.uwm.edu/cdm/ref/collection/agdm/id/1717`

[18] In the vicinity of the equator, this corresponds to an area of approximately 300×300 nautical miles (≈ 555×555 kilometers)

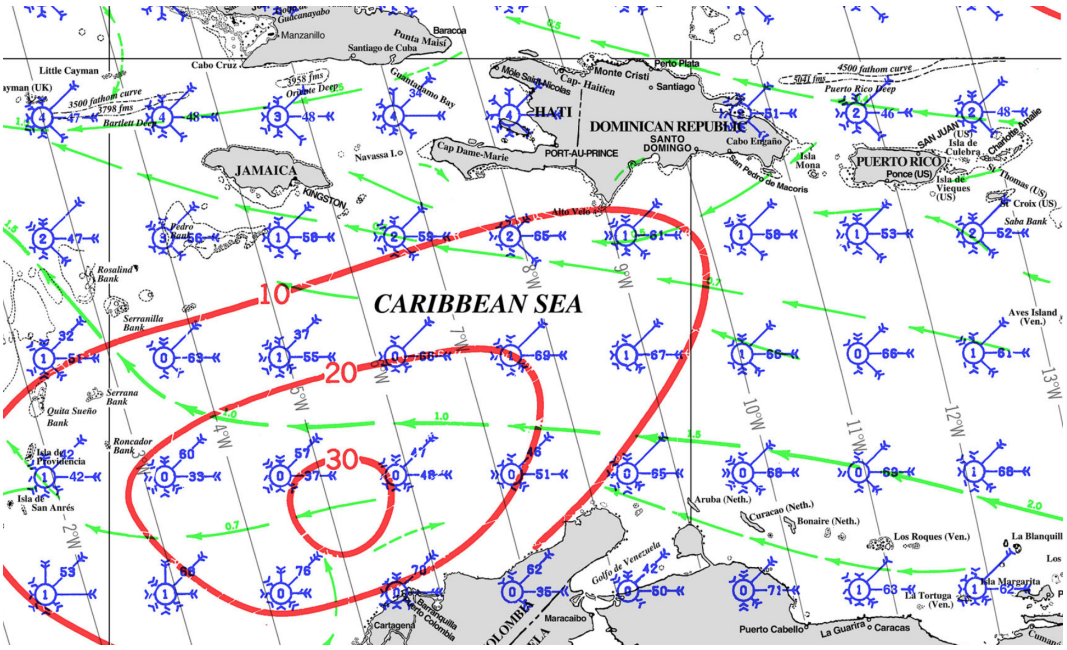


Figure 1.14: Excerpt of the 2002 edition of the pilot charts of the Caribbean Sea for the month of April (National Geospatial-Intelligence Service 2002) The blue wind roses indicate frequency and average strenght of wind from certain directions, the green arrows show the average currents, and the red lines enclose areas of a certain percentage of times that waves of over 8 feet were observed

Maury's studies were also the basis to his publication of *The Physical Geography of the Sea* (Maury 1855) which is considered to be the first book on physical oceanography. It includes the first hypothesis of the gulf stream being a major influence for the climate in the northern hemisphere.

Over the years, the *Pilot Charts* have been continuously improved. They are still used today for the planning of long ocean passages, in particular on board of sailing yachts. The current version offers a separate chart for each month of the year and – besides the prevailing winds and currents – also shows average wave heights, visibility, barometric pressures, sea surface temperatures, and ice limits (National Geospatial-Intelligence Service (2002), see Fig. 1.14)[19].

1.3.4 Algorithmic Approaches to Vessel Routing

In the 1950s, ocean weather forecasts became more reliable, allowing weather routing of ships based on forecasts instead of average seasonal patterns. This gave rise to several private vessel routing bureaus. The first algorithmic methods used by these agencies were designed to be executed by hand (James 1957; Hanssen and James 1960).

[19]The charts can be downloaded at `http://msi.nga.mil/NGAPortal/MSI.portal?_nfpb=true&_pageLabel=msi_portal_page_62&pubCode=0003`

With the increasing availability of computers and the rapid decline in costs for computation time, computer programs for finding optimal paths for vessels were developed and quickly gained popularity among ship operators.

Due to the ever-increasing accuracy of global weather forecasts (also a result of more accurate computational models), these approaches can calculate optimal routes based on long term weather forecasts, sometimes even in combination with the probability for different scenarios.

Environmental factors used as input for these models include current and wind strengths and directions as well as wave heights and their direction. The optimization goal is usually to minimize either the travel time given the speed of the vessel under certain conditions or the fuel consumption under the constraint of a given latest time of arrival.

With a few exceptions, there are three main approaches to optimal vessel routing: *calculus of variations*, *dynamic programming* and the *isochrone method*. Within these approaches, there is a further distinction between deterministic and stochastic methods. In deterministic approaches, only one scenario for environmental factors is considered, with potential updates from more recent forecasts that require a recalculation of the route from the vessel's current position. Stochastic approaches take the uncertainty of weather forecasts into account and can deal with multiple scenarios and their probabilities. Walther et al. (2016) provide a good overview of the various approaches to the problem.

Calculus of Variations In this approach, the problem is described by a functional (i.e., a function that maps a set of functions to a scalar) of which the minimum is calculated. The minimum is the function describing a path that minimizes the travel time or the energy spent.

Haltiner et al. (1962) presented the first analytical approach to the problem of optimal routing based on wave height and direction. They model a simple version of the problem with constant sea state using Euler equations. The model takes into account three different wave directions: head, beam and following waves. The method is tested on two generic scenarios, resulting in a significant reduction in travel times compared to great circle routes (i.e., the shortest distance along the curvature of the earth) as the wave height increases.

Faulkner (1962) presents a similar approach, but with the goal to find a route that minimizes the radioactive dose while traveling through a fallout area.

Papadakis and Perakis have worked on minimal time vessel routing problems, both in a deterministic (Papadakis and Perakis 1990) and in a time-dependent environment (Perakis and Papadakis 1989), using local optimality conditions to simplify the problem.

The general problem with these approaches is that it is hard to find functions that provide an accurate fit to the more and more fine-grained predictions of the environmental conditions.

Dynamic Programming Dynamic programming is a method to solve optimization problems that have an *optimal substructure*, i.e., an optimal solution to the problem can be constructed from an optimal solution of its subproblems. Shortest path problems in graphs are such problems: If node B lays on a shortest path p from A to C, then the subpaths of p from A to B and from B to C are also shortest paths.

Dynamic programming methods usually divide the ocean surface into a grid and generate a graph by connecting each grid point to its neighboring points. Then the cost (usually time) of moving from one point to its neighbors is calculated based on the characteristics of the ship and the environmental conditions between the points (Section 1.3.2). Note that in most cases, the costs are direction dependent, i.e., they are different when moving from A to B than when moving from B to A.

Some approaches use a grid that is aligned to the grid of latitudes and longitudes. In other methods, the grid is generated along the great circle route from the origin to the destination. Jensen et al. (2017) use grid points that are evenly spaced along the edges of quads (i.e., rectangular regions).

Given the directed graph with all costs, a dynamic programming method for calculating shortest paths is applied. Common methods are Dijkstra's algorithm (Dijkstra 1959) or the A*-method (Hart et al. 1968). Jensen et al. (2017) take advantage of the geometric properties of the quads, providing a more efficient solution than the A*-method.

The drawback of the dynamic programming method is that the number of neighboring points limits the choice of directions from a given location, resulting in potentially suboptimal routes. The problem can be reduced however by increasing the density of points along a potential route or using an approach like points along the edges of quads with a sufficient density.

The Isochrone Method This is the oldest systematic approach to weather routing and was originally developed for manual execution (James 1957; Hanssen and James 1960). The principle of this method is the calculation of all points that can be reached from the point of departure within a given time based on the environmental factors. The totality of all those points forms a line called isochrone. From points of this line, the new points that can be reached in the next timestep form the next isochrone.

In his PhD-Thesis, Hagiwara (1989) proposes a *modified isochrone method* that renders the isochrone method suitable for automatic computation. His method works as follows: From the start point, it calculates the position of the vessel for a predefined number of different headings after the first time step. This gives the first isochrone. From a fixed number of points on this isochrone, the calculation is repeated. The trick to avoid an exponential growth of the points to be evaluated is to pre-define sectors from the start point and only choose the point with the maximum distance from the origin for each sector to construct the next isochrone. Hagiwara presents both a deterministic and a stochastic version of his approach. In Section 1.4, we give a more detailed description of the method and describe how we extended it to

suit the situation in the Caribbean.

A version of the approach that also takes fuel consumption into account is presented by Klompstra et al. (1992). They create isopones, i.e., 3-dimensional fronts of equal fuel consumption. A similar method has been used by Lin et al. (2013), implementing a 3D version of Hagiwara's method.

There have been numerous changes and modifications to the isochrone method, adapting it for specific situations, e.g., by partitioning the area (Szlapczynska and Smierzchalski 2007) or using a pruning technique to limit the exponential growth of the paths as in the freeware weather routing software *vrtool*[20]. For this method, a virtual boat-wake is used, removing all points in the area between a certain angle to the left and right of the direction the boat comes from.

Unfortunately, for most commercial routing software, the details of the used algorithms remain undisclosed. We thus develop our own extensions to Hagiwara's method in Section 1.4 to adapt it to the situation of canoe travel.

Other Methods The problem has also been approached by a variety of methods inspired by processes from the field of biology. Tsou and Cheng (2013) use an ant colony algorithm to tackle the problem. Hinnenthal and Clauss (2010) and Maki et al. (2011), among others, use genetic algorithms to find a (possibly local) optimum of the routing problem.

1.3.5 Applications to Archaeological Questions

Some methods of creating optimal routes have already been applied in archaeological research for reconstructing ancient routes of travel and exchange. In these applications, the real-time calculation based on forecasts is not relevant. Thus the calculations can be purely based on observations or statistical assumptions about the environmental factors in the past.

With the rise of GIS-systems and least cost pathway analysis, shortest path calculations on the water became popular in the field of archaeology. In the following we give a brief overview of selected publications on the topic.

Vessel Routing with Archaeological Data

Levison et al. (1973) and Ward et al. (1973) were the first to simulate historical ocean voyages on a computer to study the settlement of remote islands in the Pacific. They simulated both drift voyages and navigated voyages in which a given course is held as far as the wind situation allowed. The environmental factors were assigned at random based on information about average wind and current data on a 5-degree-resolution.

[20] `www.tecepe.com.br/nav/vrtool/`

Irvin et al. (1990) ran simulations to explain the settlement of the islands in the South Pacific by sailing canoes. They base their model on wind data from monthly meteorological charts from the British Meteorological Office. The model does not only take into account trips in random direction but favors safe voyages. These are voyages that start against the wind, allowing for a safe return with the wind on a given latitude.

Leidwanger (2013) calculated travel distances of sailing ships around 600 BC in the Mediterranean based on wind data in the area. He uses wind data from the web service *windfinder.com* and the *Wind and Wave Atlas of the Mediterranean Sea* and interpolates the data for gridpoints on a 500m grid. Based on observations of the sailing performance of a replica of a 300-BC greek ship, he constructs an anisotropic surface in ArcGIS, making the travel speed dependent on the direction towards the wind.

A very detailed study on trade at sea during Roman times has recently been published by Warnking (2015). He uses the professional weather-routing tool *expedition*[21] – which is also used for yacht routing – for simulating roman sailing ships.

Archaeological Studies in the Caribbean

Callaghan (2003b) applied computer models to calculate the likelihood of different potential origins of the first settlers in the Greater Antilles between 4000 and 2000 BC. He claims that "three mainland areas had the potential for chance discovery of the Greater Antilles and that during the Ceramic period there were no technological or environmental barriers requiring navigation to follow a 'stepping stone' pattern from island to island" (Callaghan 2003b, p. 325). He thus hypothesizes that the Greater Antilles might have been settled by paddling directly from the mainland of Venezuela. Calaghan bases his calculations on the *United States Navy Marine Climatic Atlas* (Fleet Numerical Meteorology and Oceanography Center 1995) with a resolution of 1 degree. He calculates success rates of ending up on the Greater Antilles for drift voyages from different locations on the mainland towards the Greater Antilles.

His results indicate that the only start points for a drift voyage with a success rate over 1 percent are located on the coast of Venezuela, resulting in drifting times of more than a month. He does not comment on the problem of supplying the crew with food or water during a month long drift across the Caribbean Sea. He also describes the probability of success for intentional voyages. For this he assumes a constant speed of paddling and a constant heading towards the islands. Unsurprisingly, these voyages have a success rate of 100 percent when the start point was conveniently located. Unfortunately, his paper does not describe the exact algorithm he uses for his calculations.

In subsequent articles, Callaghan applies his method to the Pacific coast of Middle America (Callaghan 2003a) and the South Pacific (Callaghan and Fitzpatrick 2008).

[21] `http://www.expeditionmarine.com`

Another study on the Caribbean has been done by Altes (2011). He calculates the current flow from Venezuela and adds a constant paddling speed of 8 km per hour. This results in estimated travel time of 9 days from Venezuela to Cuba.

Cooper (2010) uses a cost function on land and sea to calculate distances among sites in Cuba within a broader GIS-based analysis of several excavated sites in a small scale region. He does not incorporate currents, though, but instead regards the sea as flat surface.

In her PhD-thesis, Slayton (2017) applied the models we will describe in the following section on three different case studies in the Caribbean. We will summarize her results in Section 1.4.3.

1.4 Shortest Path Calculations for Canoes

In the following we describe how we adapted two methods from Section 1.3 to the specific situation of canoe routes in the Caribbean (Section 1.4.1) and how they compare against each other.

We then discuss the difficulties of finding a proper weighting of the environmental factors acting on the canoe (Section 1.4.2).

In Section 1.4.3 we then take a closer look at the seasonal variations of the travel times for a specific route in the northern Lesser Antilles.

1.4.1 Finding a Suitable Method

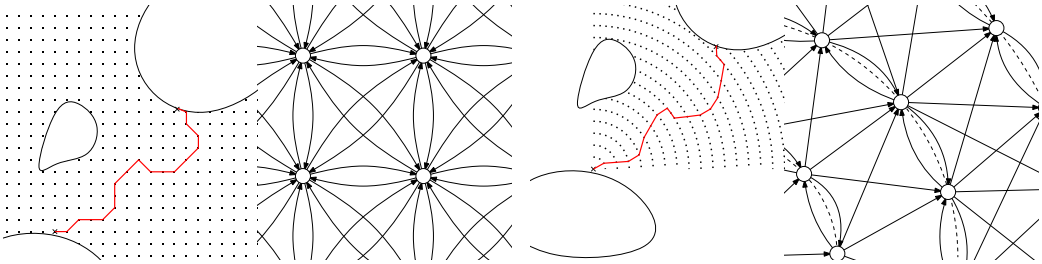
In the following we describe how we adapted a grid-based dynamic programming approach and Hagiwara's isochrone method to the situation of canoe routing in the Caribbean. The situation in our case differs to the routing of cargo-ships on long distance ocean routes for which the methods were originally developed: The distances are shorter and the paddling speeds are slower, leading to a greater relative influence of the environmental factors. Thus the optimal routes deviate more from the straight-line course. Additionally, there are often islands in the way of the straight-line route, requiring the algorithm to allow for sharp turns in order to reach destinations behind an island.

We will describe the basic principles of the algorithms and provide a high-level description of the techniques used. For a pseudocode and more details we refer to the master thesis by Jan Hildenbrand (2015).

This section is based on joint work with Jan Hildenbrand who wrote his master thesis (Hildenbrand 2015) on the comparison of the methods described here and who implemented the algorithms in the course of his master's project at the University of Konstanz.

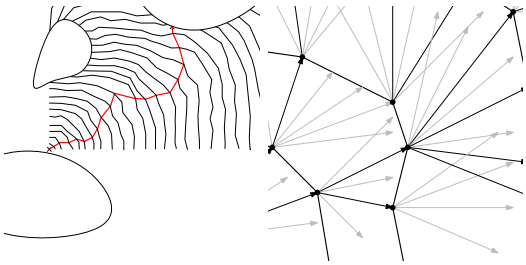
Isodistant Method (Dynamic Programming)

For this approach, we select points on the ocean surface as nodes and generate a weighted directed graph by connecting each point to its neighbors. We assign the weight to the edges according to the travel time from the start to end of the edge using the calculations described in Section 1.3.2. In order to get more accurate results, we divide each edge into smaller sections and interpolate the environmental factors for the center of each section. We then add up the travel times for the sections to get the weight of the edge. We use the A*-algorithm (Hart et al. 1968) to find the shortest path in the graph from the source to the target through the graph.



(a) Dynamic programming with regular grid

(b) Dynamic programming with adapted grid (isodistant method)



(c) Isochrone Method

Figure 1.15: Dynamic programming on two different grids in comparison to the isochrone method

The most straightforward approach would be to choose the grid points of the underlying data as nodes which are provided at equal distances along the axis of the world's coordinate system. We could then connect each grid point to its eight neighboring points, creating a directed graph. (Note that the distances between the grid points are not equidistant due to the curvature of the earth.) The clear advantage of this method would be that we only need to create the graph once and can then easily calculate shortest paths between any two nodes. The major downside is that the choice of directions from each grid point is very limited, resulting in "zig-zagging" around the direct route. Therefore the output will almost always be less than optimal. Also, the desired start and end points do not necessarily coincide with a grid point, so we have to choose a grid point in the vicinity. This leads to additional inaccuracies when the grid is coarse (see Fig. 1.15a).

In order to overcome these difficulties and to make the method comparable to the

isochrone approach, we instead decided to construct an *adapted grid* that overcomes most of the problems of the fixed grid (see Fig. 1.15b). We create points that lie on circular arcs around the start point at a given distance d_l. The points on each arc are located at a distance d_p from each other. We then introduce directed edges from each node to the nearest c_e nodes on the next circular arc. In order to also allow for perpendicular movements, we also connect each point on an arc to its neighbors on the same arc. While the start point now coincides with the first grid point, we still have to choose a point close to the destination as end point. The disadvantage of this method is that we have to construct a new grid for each start point.

Due to the fixed distance between the arcs, we call this method *isodistant method* in contrast to the isochrone method we will describe now.

Isochrone Method

We based our implementation of the isochrone method on the work of Hagiwara (1989). We introduced a few extensions to allow routing around obstacles such as islands or strong counter currents.

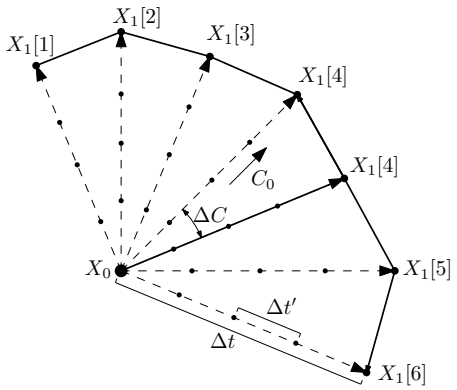
The original method from Hagiwara works as follows (see Fig. 1.15c and 1.16): Let X_0 be the departing point and t_0 the time of departure. We assume a constant speed through the water V_s of the vessel. The time interval Δt gives the temporal resolution of the isochrones. It is further divided into q subintervals $\Delta t'$, i.e., $\Delta t = q\Delta t'$.

Based on a chosen angular resolution ΔC, we start the simulation from X_0 with $2m + 1$ discretized headings $C_0 \pm i\Delta C, i = 1 \dots m$ around the direct course (great circle route) C_0 towards the goal. For each heading, we calculate how far the ship could move along the heading with the given environment factors every $\Delta t'$ time units (see Section 1.3.2 for details on this calculation) until we have reached the time $t_1 = t_0 + \Delta t'$. The points $X_1 = \{X_1[1], \dots X_1[2m + 1]\}$ of the resulting points define the first isochrone of time t_1.

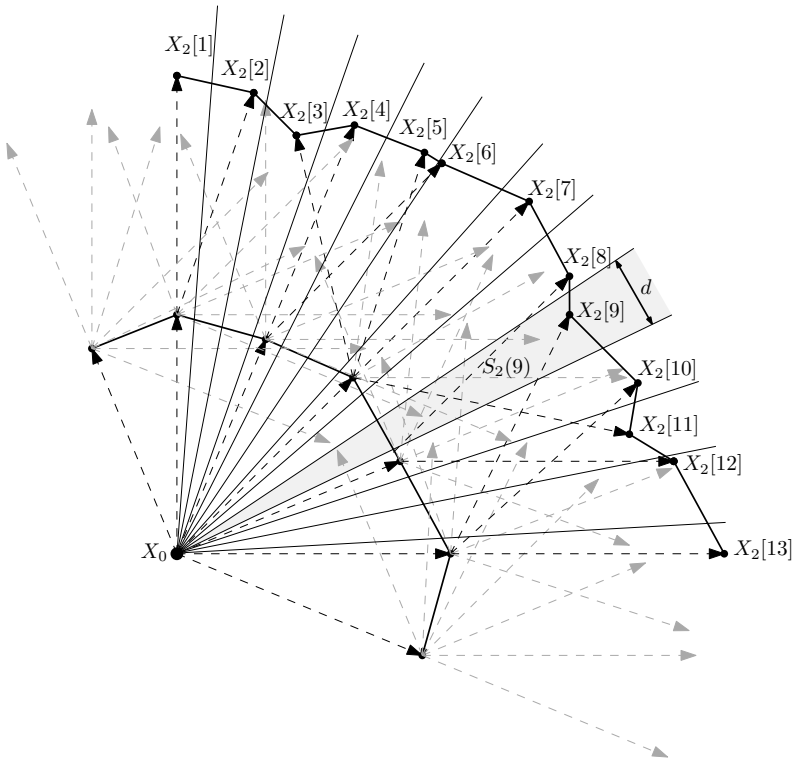
In the next step, we repeat the process with each of the points on X_1 as start point. However, now we have $(2m + 1)^2$ points in the set X_2. In order to avoid an exponential growth of the points, we choose a set of subsectors $S_2 = \{S_2(1), \dots S_2(2p)$ originating from the start point and only choose the point in each subsector that is *furthest* from X_0. The new isochrone $X_2 = \{X_2[1], \dots X_2[2p]\}$ is formed by resulting $2p$ points. The subsectors are centered around the direct route (great circle) and their angle is chosen such that the width of the sector is d at the expected distance $k\Delta tV_s$ of the isochrone X_k, leading to a similar resolution for each isochrone.

Extensions to Hagiwara's method

Hagiwara chooses the *furthest* point based on its great circle distance from X_0. This choice seems reasonable for open ocean courses when the course closely follows the great circle route and the velocity induced by the environmental factors in relation to the ship's velocity is rather low. In our case-study however, there are often islands in the way, and at times, the speed of (counter) currents comes close to (and sometimes

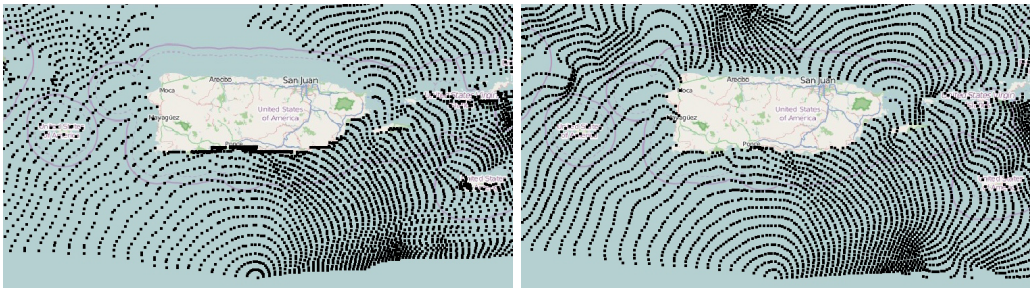


(a) First step: An initial isochrone is calculated



(b) Second step: Only the furthest point in each sector is selected for the new isochrone

Figure 1.16: Detailed illustration of the isochrone method by Hagiwara (1989)



(a) Hagiwara's original method **(b)** Our extensions

Figure 1.17: Comparison of Hagiwara's original method to our extensions: The original method fails to reach points behind the island (images by Hildenbrand (2015))

even exceeds) the speed of paddling. In these conditions, Hagiwara's method does not perform well, as the selection of the furthest point from the source for each sector prevents the choice of headings whose angle deviates too much from the straight-line course away from the start point. Points behind islands are thus never reachable, and large deviations from the straight-line route get more and more discouraged with increasing distance from the start point (see Fig. 1.17a).

Thus, we introduce the following extensions to Hagiwara's method:

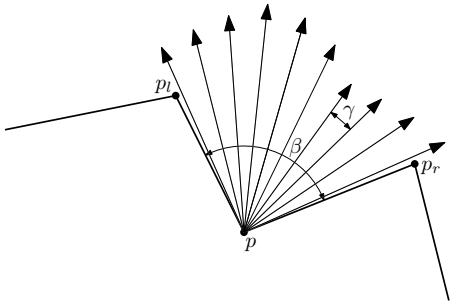
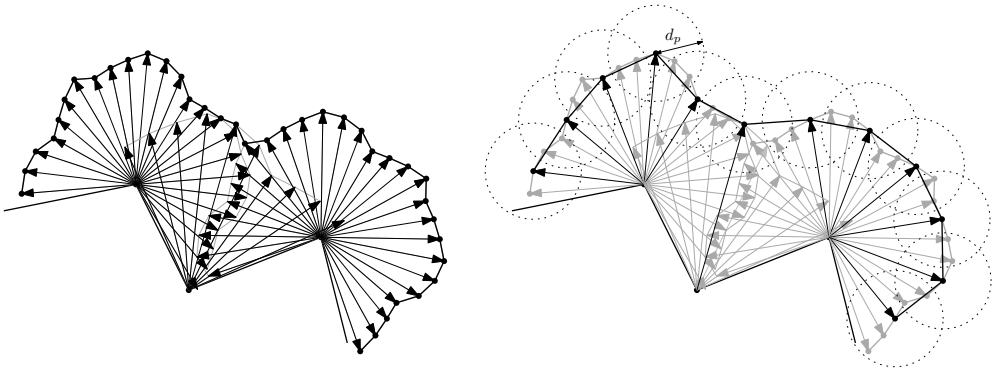


Figure 1.18: Neighbor aware heading selection

Neighbor Aware Heading Selection Hagiwara uses a fixed angle to the left and to the right of the straight course within which he calculates different headings from a start point p on an isochrone X_i. We decided instead to look at the two neighboring points p_l and p_r of p on X_i and fill the angle β between them with a fan of headings spaced at a given angle γ, starting from point p and leaving equal angles to the right and the left of the fan. For the start and end point of the isochrone, we choose a fixed offset angle β'.

Isochrone Candidate Selection In order to allow for isochrones that are not just (more or less) perpendicular to the direction from the start point, we use a more versatile approach than the selection of the furthest points from the origin. For the generation of the next isochrone, we generate a polyline for each start point

from the points of each heading from the neighbor aware heading selection. We then combine the polylines by intersecting them and by removing all points below the outer line (Fig. 1.19a). In the implementation of Hildenbrand (2015), we do this by checking if a point is contained within one of the triangles formed by the origin point and two neighboring headings from the next fan of headings. This turned out to be reasonably fast in practice. A faster approach would be to find intersections of the polylines via a variation of the Bentley-Ottman sweep line technique (Bentley and Ottmann 1979; Park et al. 2001), leading to run times of $\mathcal{O}((n+k)\log n)$ where n is the number of segments and k the number of intersections.



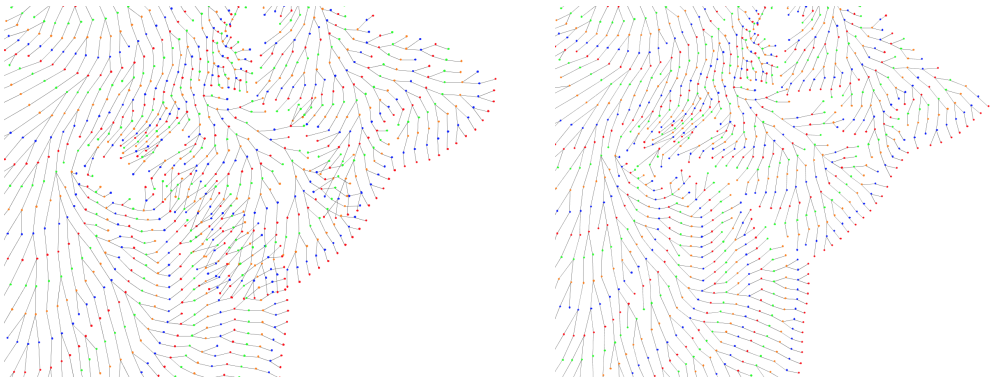
(a) Intersection of the polylines **(b)** Reducing the number of points

Figure 1.19: Isochrone candidate selection

The resulting polyline is a more exact approximation of the next isochrone. In order to reduce the complexity (i.e., the number of points on the isochrone), we start on one end and remove all subsequent points that are closer to the current point than a given distance d_p. We then keep doing this for the whole isochrone, starting at the next point that has not been removed (Fig. 1.19b).

Elimination of Loops In degenerate cases, for example when the isochrones meet again behind an island or if the velocity of the currents changes rapidly in an area, it is possible that the algorithm used for the isochrone candidate selection in (Hildenbrand 2015) produces self-intersecting isochrones. This can lead to undesired outcomes (see Fig. 1.20). We thus detect loops and remove them as follows: For a fast detection whether a loop is present, we follow the isochrone from point to point and add up the bending angles of the isochrone on each point (negative angles for left turns, positive angles for right turns). If the angle is greater than π or smaller than $-\pi$ at any time during the calculation, we have likely found a loop. We can now search the pair of intersecting line segments (e.g., with a sweep line technique (Bentley and Ottmann 1979) or by a simple pairwise check) and remove all points between them.

Figure 1.17b shows how our extensions affect the area covered by the isochrones. The extensions also cover places behind islands by allowing isochrones to deviate at larger angles from the default, which would be perpendicular to the straight-line course.



(a) Before elimination of loops **(b)** After elimination of loops

Figure 1.20: Loop elimination removes overlapping isochrones behind an island (points in the same color are on the same isochrone, image by Hildenbrand (2015))

Comparison of the Methods

Both the extended isochrone method and the dynamic programming approach with the adapted grid (isodistant method) work reasonably well in the tests performed by Hildenbrand (2015). Based on real-world current data, a total of 2,700 routes were calculated using different parameter settings on the extended isochrone method, Hagiwara's original isochrone method, and two versions of the isodistant method (with links to the neighbors on the same arc and without them).

The travel times of the calculated routes vary slightly depending on the choice of parameters, so there is no clear "winner" as to which algorithm produces "more optimal" routes. Although all approaches should in theory produce the same – optimal – result, existing differences can be explained by the limited number of (discrete) headings, points on the grid or the isochrone that are not exactly on the desired end point, as well as slightly different results of the interpolation of the currents due to different start- and end points of the mini-steps.

Interestingly, the comparison of the travel times of the routes calculated by Hildenbrand (2015) shows only relatively low advantages of the optimal methods to the naive approach of just heading towards the target (Section 1.3.1). On the longer routes, travel times are only about 5-10 percent faster with the optimal approach when the paddling speed was set to 1 m/s. At lower paddling speeds, the differences were larger, as the relative influence of the currents is stronger. For some of the shorter routes, the optimal routes are even slower than the naive, which can be explained by the fact that the naive routes end directly at the desired end point and there is thus no additional cost of travel from the nearest grid or isochrone point to the desired end point.

Regarding the practical run times, there is a clear winner though: The implementation of the isochrone method outperformed both variants of the isodistant method, being about three times faster. Interestingly, the extension of Hagiwara's method improved the run time even though the techniques used for the isochrone candidate selection are not asymptotically optimal (see Fig. 1.21).

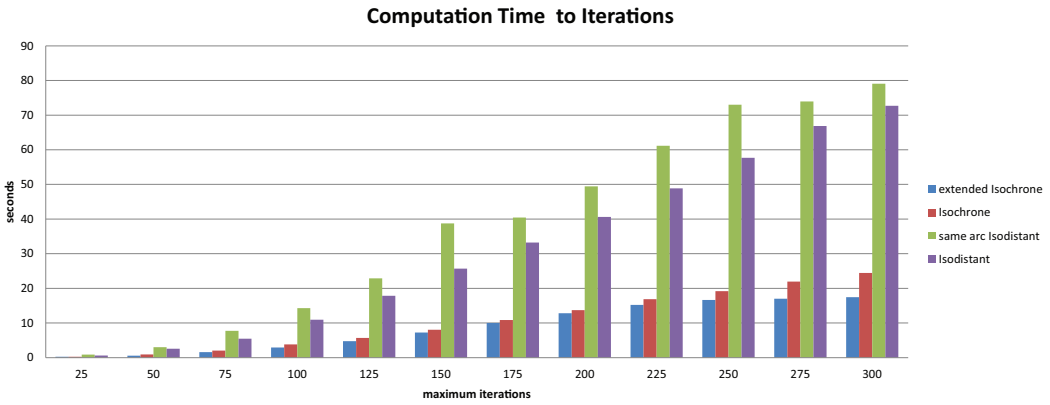


Figure 1.21: Comparison of the measured run times for the implementations of different methods. The extended isochrone method outperforms the other methods indepent of the number of iterations (image by Hildenbrand (2015))

We thus decided to use the extended isochrone method in the tool we implemented for Emma Slayton for the experiments in the course of her thesis (Slayton 2017). The routes used for the analysis of the seasonal variations in travel times in Section 1.4.3 are also calculated by this method.

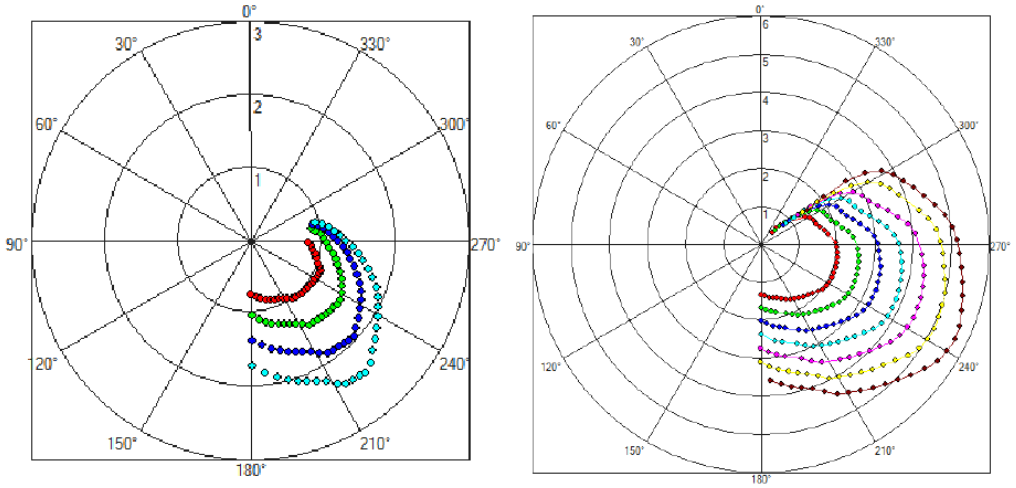
1.4.2 How to Weight the Environmental Factors

Given the three potential factors of influence (currents, wind, and waves), the question is how these factors influence the velocity of the canoe. The influence of currents is the most straightforward one: The velocity vector of the currents simply gets added to the velocity vector of the boat. (Since both speeds are rather low, it is safe to ignore a theoretical slow down due to air resistance.)

The influence of wind, also known as *leeway*, is much more challenging. If a floating object is subject to the influence of wind, its movement through the water depends on many factors, such as its shape above the waterline, its shape below the waterline, and the angle from which the wind force is applied, to name just a few examples. According to Hackett et al. (2004), due to the high number of possible parameters and the high degree of uncertainty, stochastic methods need to be used that are based on empirical observations of similar objects.

With the help of sophisticated simulation software or through extensive sea trials in controlled conditions, it would be possible to create a polar diagram of the canoe. Such a diagram shows the (drifting) speed of the boat through the water depending on the direction of the wind. Boeck et al. (2012) did such simulations for a polynesian voyaging canoe (see Fig. 1.22). The results are not applicable to our situation though, since these canoes were equipped with sails, and for the creation of the diagrams, an optimal sail trim for the given course is assumed.

Things get even more complicated when it comes to the influence of waves. From experimental voyages, we know that waves play an important role in the safe navigation of a large canoe (Benoît Bérard, personal communication, January, 2015



(a) U-shaped hull **(b)** V-shaped hull

Figure 1.22: Polar diagrams for two different hull shapes of polynesian voyaging canoes equipped with sails. The dots indicate the possible boat-speed depending on the angle to the wind direction (from the top) and strenghts (indicated by the colors). Note that with the V-shaped hull, even upwind sailing is possible (figure by Boeck et al. (2012), reprinted with permission)

and Emma Slayton, personal communication, April, 2016). The best strategy when dealing with waves of a short wavelength appears to be to either head directly into the waves or in the opposite direction in order not to capsize. In contrast, swell of greater wavelengths seems to be less of a problem and does not affect the operation of the boat as much.

In general, heading into waves can slow down the boat significantly, having waves from behind can increase the speed of the boat, yet only if the waves arrive at a certain frequency. It seems to be crucial to keep the boat at a certain speed when navigating in waves in order not to capsize. In addition, it takes a skilled paddler and a lot of attention to paddle in choppy waters since the depths of the paddle strokes has to be adjusted to the height of a wave at the entry point of the paddle.

Thus, a polar diagram for the influence of waves would have to simultaneously include multiple waves from different directions and of different frequencies. In addition, it would have to be recalculated depending on the current speed of the boat.

Unfortunately, it was not possible to get a proper hydrodynamic model of a typical pre-Columbian voyaging canoe, nor to perform sufficient tests with the replicas to estimate polar diagrams of the boat for the influence of wind and waves.

Based on the evaluation of a variety of sources on the matter, including finding suitable parameters for the description (Breivik et al. 2011), leeway drift characteristics of a (small) tropical Pacific island craft (Brushett et al. 2014) and a Hawaiian travel canoe (Polynesian Voyaging Society 2017) as well as reports from the US Coast Guard regarding leeway of floating persons and life rafts (Allen et al. 1999; Allen

2005), we decided to integrate a reasonable small, constant influence of the wind into our model and disregarded the influence of the waves for the following studies.

The user of the routing tool we developed thus has the option to set a percentage for the influence of the wind. This influence is implemented by simply adding the wind vector – multiplied by the set percentage – to the influence of the canoe by the current vector. From discussions with Benoît Bérard and Emma Slayton (personal communication, April, 2016) we determined an influence of two percent to be a reasonable value. Thus, if there is a 20-knot wind, we estimate a leeway drift of 0.4 knots in the wind direction that is added to the current vector. (Note that the current data is also generated with wind data as a factor, thus there is an additional indirect influence of the wind as well.)

While this method of weighing the factors is not exact, it is likely to be better than ignoring wind altogether and coming up with non-exact data for the waves. Should sufficient data about the hydrodynamics of the boat become available in the future, it would be easy to incorporate this data into the existing algorithm.

1.4.3 Modeling Pathways in the Caribbean

Emma Slayton used our routing tool extensively in her PhD-thesis (Slayton 2017), supplementing archaeological theories on movement patterns between known sites with the calculated routes and travel costs. She conducted three case studies:

1. Connections within the group of Leeward Islands from Anguila to Antigua including periphery connections to Guadeloupe (Fig. 1.23)

2. Connections crossing the Mona Passage and the Anegada Passage between the Greater and the Lesser Antilles (Fig. 1.24)

3. Connections between Guyana on the South American mainland to the Windward Islands of St. Vincent and Grenada (Fig. 1.25)

Due to the high variability of currents on a day-to-day basis, a large number of sample voyages were simulated for each of the examined months with starting times roughly every three hours. This allowed us to get an estimate of the variation in times throughout the month and to calculate the average travel times.

The paddling speed was set to 1.5 m/s, and winds were set to have an impact of 2% (see Section 1.4.2). As expected by the analysis of the current data (Section 1.2.4), there was no clear pattern of the variations of the travel times (except for the area of the third case study) since daily variations by far exceeded variations in monthly averages. In the following we will highlight some of the results of the case studies by Slayton (2017) and then more thoroughly examine the variations in travel times for an exemplary voyage.

Case Study 1: Lithic Exchange in the Northern Lesser Antilles

The study examines the connection between early sites from the Archaic Age, dating from 2000 BC to AD 100, for which a connection is indicated by common finds of archaeological material, mainly Long Island flint. As Long Island is the only major source of this raw material in the region, the island was likely the most important export site for flint to all islands in the area. The main extraction site for flint on the island was probably Flinty Bay. The archaeological evidence suggests that many of the tools were already pre-shaped close to the site, reducing cargo weight for the transportation in canoes (Knippenberg 2007).

The tracks of the routes modeled by Slayton (2017) can give a general hint as to which additional sites or islands might have been part of the exchange network as many of the optimal pathways from Long Island to other sites come close to these sites. For example, while most routes pass by St. Kitts, some of the optimal paths from Long Island to Saba deviate to the north, passing by the islands of Anguilla and St. Martin.

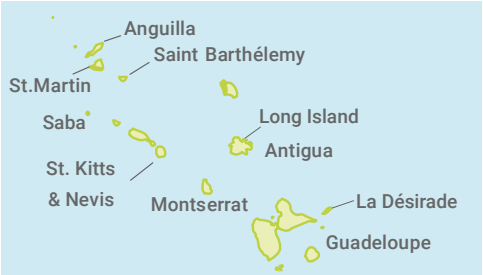


Figure 1.23: Area of Case Study 1

Regarding the travel times, the routes indicate no clear pattern as to a certain month or time of day that would be favorable for traveling. There is a slight tendency for less variation in travel times during the winter months though, which is also reflected in the exemplary analysis we perform at the end of this section.

Case Study 2: Crossing the Passage between the Greater and Lesser Antilles

The second case study by Slayton (2017) is set in the context of connections between the Dominican Republic and the northern Lesser Antilles that must have existed in the Late Ceramic Age (AD 1200-1500) (Hofman 1993). The optimal routes across the wider channels quite often pass in the vicinity of smaller islands in the channels, indicating that these might have been resting areas of a strategic importance. Routes were simulated for the months of January, April, July, and October.

Route trajectories across the Anegada Passage between Saba and Puerto Rico often pass either north or south of St. Croix, indicating the potential importance of the island as a stopover point after the long stretch of open water coming from the island of Saba (Faber Morse 1989).

Many route trajectories across the Mona Passage between Puerto Rico and the Dominican Republic and back pass closely to the north coast of Mona island, indicating the possible importance of the island at least as a navigational mark. There exist no landing opportunities on the north coast of the small island due to the presence of high cliffs (the only way to land on the island is from the South West). However,

Figure 1.24: Area of Case Study 2

caves in the cliffs show traces of human occupation and their location would allow an observation of passing canoes. They might thus have served as visual connection points from the sea as well (Samson et al. 2015).

As with the previous study, there are generally no significant seasonal changes in the travel times except for voyages from the Dominican Republic to the south west of Puerto Rico where average travel times are longer in October and shorter in April.

Case Study 3: Long-Distance Voyages between Guyana and the Lesser Antilles

Figure 1.25: Area of Case Study 3

Evidence in material culture suggests that there existed a recurrent exchange of objects between peoples on the mainland of South America (Guyana, Suriname, and French Guiana) and the Kalinago peoples on the islands of the southern Lesser Antilles (Keegan and Hofman 2017). Unfortunately there is no knowledge about connections between particular communities or settlements. The voyages in this case study are simulated in order to explore the possibility of direct connections from the mainland to the islands, requiring multi-day voyages in the open ocean. Average voyage times have been calculated for the months of January, April, August, and November.

The average time for direct trips from a point off the coast of Guyana to St. Vincent ranges from about 87 hours in April to around 113 hours in August. Thus in this case we have a clear seasonal component. This variation is in accordance with the plots for Location 5 in

Section 1.2.4 where the average currents are significantly stronger in the summer months.

The trajectory of the direct optimal routes varies between heading northwards offshore and then turning northwestwards towards St. Vincent or staying closer to the coast, passing between Trinidad and Tobago. An open question is whether paddlers actually chose the offshore-route despite its dangers and the logistics involved with staying at sea for almost 5 days or whether they rather chose to follow a slower, but safer route along the South-American coastline before crossing the channel from Trinidad to Grenada.

We will now perform a more in-depth analysis of the variation of travel times for a particular route from the first case study.

Examplary Voyage: Variability of Travel Times Over Time

The following analysis is based on data from Emma Slayton who made extensive use of our implementation of the proposed extensions of Hagiwara's method (see Section 1.4).

We will now have a closer look at the variability of the simulated travel times of the route between Long Island / Antigua (17.159° North, 61.752° West) and the Fort Bay Ferry Terminal on Saba (17.616° North, 63.252° West) from the first case study. We base this analysis on 15,335 runs of the model on current and wind data from the timespan of March 2010 until October 2015. The model runs were mostly evenly spaced with about 3 hours between the starting times with exceptions in a few months where the spacing was larger. Some periods have missing data, partially due to missing environmental data and partially due to failed model runs (most likely caused by a lack of sufficient memory). Figure 1.26 shows the travel times for all available model runs. We also calculated a moving average that takes into account all values two weeks prior and two weeks after the given date.

The average travel time was just above 28 hours and 35 minutes with a maximum time of 1 day, 18 hours and 23 minutes and a minimum time of 20 hours and 7 minutes. However, the extreme values are outliers, caused by extreme weather patterns (both the fastest and slowest travel times occur on August 21st 2011, the day when hurricane *Irene* hit the area). The variance of the travel times over the whole year is 3 hours and 8 minutes. The variance depends on the months and ranges from 1 hour and 8 minutes in March to 4 hours in October. Figure 1.27a shows box plots of the distribution of travel times for each month, color coded by the year.

The box plots indicate that there is more variation of the times in the summer months, and we thus compared the distribution of travel times of the dry season (February–June) to the wet season (July–January) in Fig. 1.27b. The distribution shows that indeed the variation is slightly higher and travel times are slightly shorter in the wet season. However, the absolute differences in travel times seem too small to have prevented voyaging during certain periods of the year.

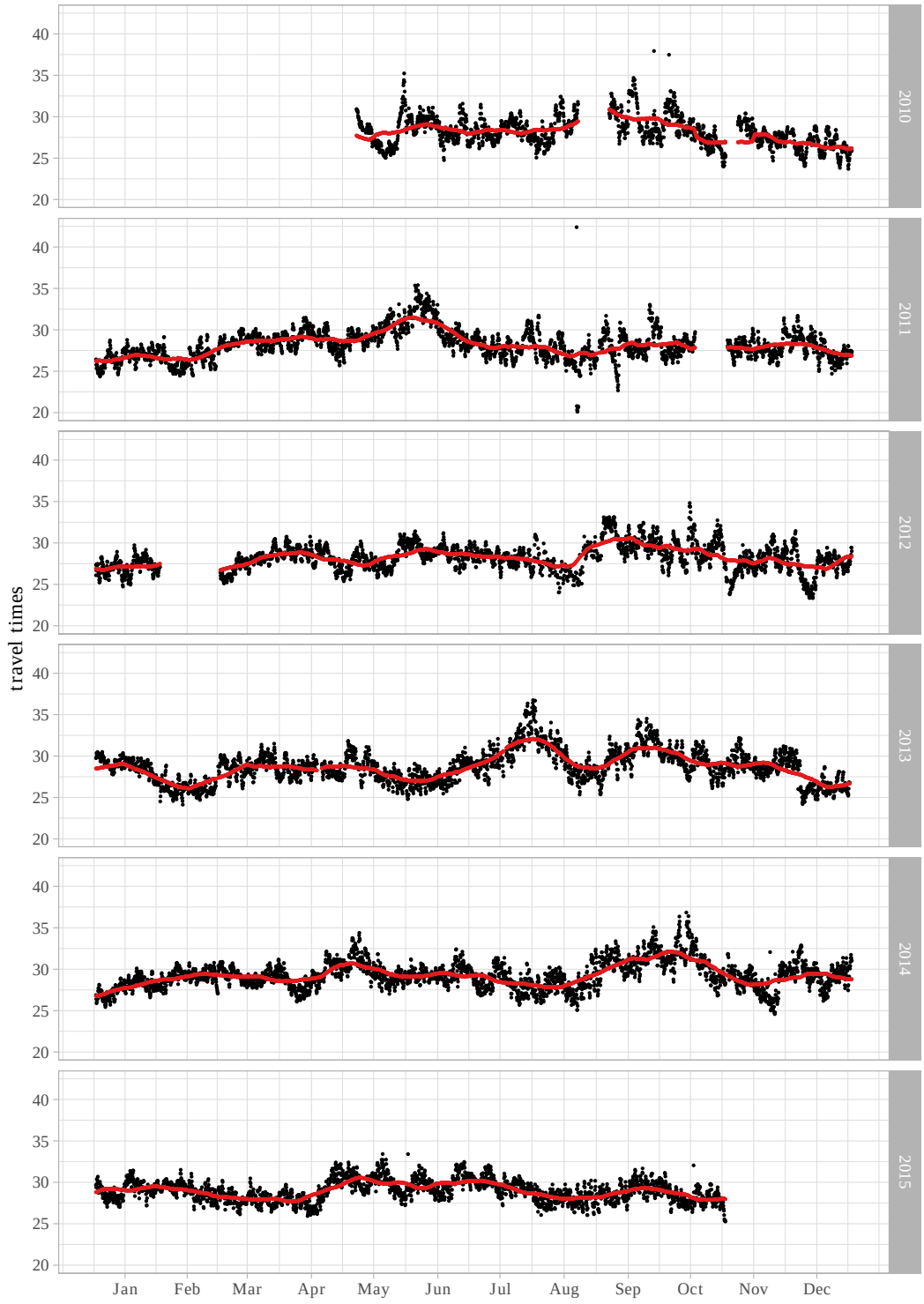
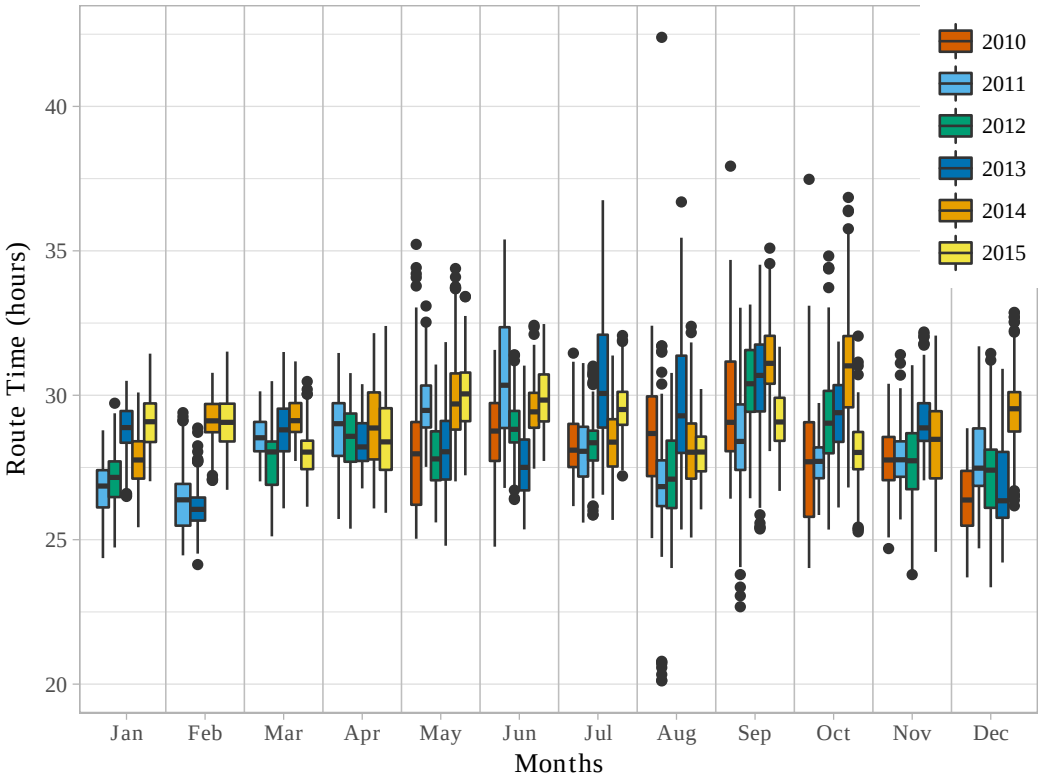
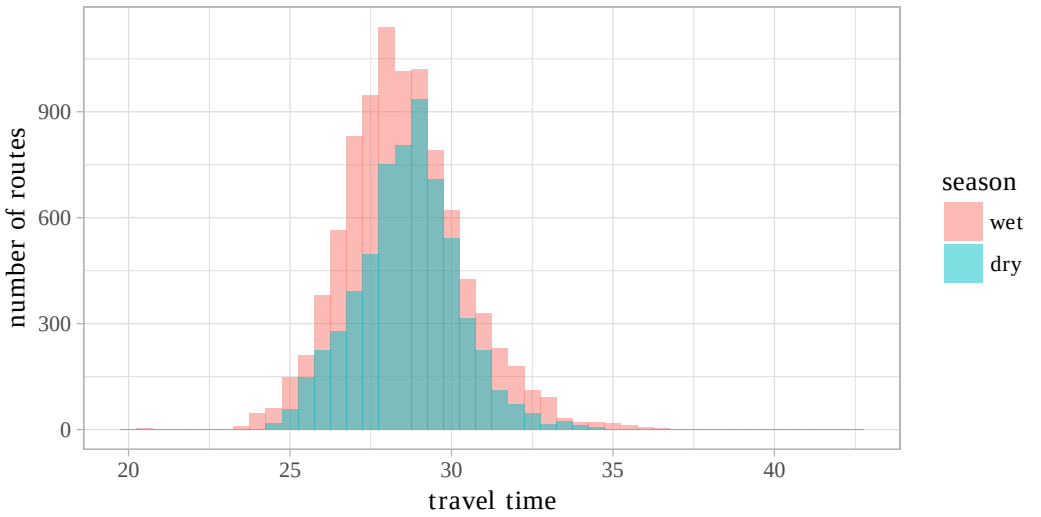


Figure 1.26: The travel times between Long Island and Saba for all available model runs (black points) and the running average of a 4-week window

After having focused on optimal travel routes until now, we will try to validate how realistic it is for a human actor to find routes that are (close to) optimal.



(a) Box plots of all model runs per month and year (May 2010 – October 2015)



(b) Comparison of the distribution of travel times of the dry season (February–June) to the wet season (July–Janary)

Figure 1.27: Travel times between Long Island and Saba

1.5 Comparing Human Agency to Optimal Pathways

The calculation of the optimal routes does not only require a substantial amount of computing power, but also a perfect knowledge of the currents on each possible location in the area. Thus, it is unlikely that humans were able to find these exact pathways. However, the seafaring and navigational capabilities of past civilizations have been often underestimated by modern researchers. For example polynesian sailors were able to sail hundreds of miles in the open ocean without loosing track of their position long before Europeans started trans-oceanic voyages. See (Lusby et al. 2009) for an introduction to polynesian navigation skills.

In order to get an idea of the capabilities of human actors to optimize their travel times by canoe, we developed a simulation game and hosted a competition for the fastest paddling times among the players. This allowed us to gather data while at the same time offering an enjoyable experience to the participants and contributing to the community outreach of the *NEXUS1492* project.

This section is joint work with Roman Lamsal, who implemented the game and organized the competition in the course of his bachelor thesis (Lamsal 2017). While parts of the following results originate in his thesis, we did a reanalysis of the data to better visualize the learning behavior of the players.

1.5.1 The Canoeing Game *RowYerBoat*

RowYerBoat is a Java based game programmed in *libGDX* that can be played on desktop computers as well as Android devices. The goal is to paddle a canoe on a map from a start point to an end point while passing a number of checkpoints on the way. An underlying grid of currents influences the boat's velocity. The map represents the islands of the Caribbean in the area between 9.5° and 19° north and 55° and 68° west, covering the Lesser Antilles including Puerto Rico and parts of the coast of the mainland of South America.

Missions can be added to the game depending on the situation (community outreach, user studies). A mission consists of a number of checkpoints which are implemented as circular areas on the map, highlighted by a red marker. As soon as the boat enters a checkpoint, the next checkpoint becomes active.

A paddling stroke on the left or right side of the boat is performed by pressing the left or right key (on a desktop machine) or swiping down on the left or right side of the screen (when playing on a mobile device). The stroke increases the velocity of the boat and introduces a momentum of turning towards the opposing side. The slower the boat, the stronger the turning momentum. The maximum velocity of the boat is 1.5 m/s. The velocity is reduced quickly over time, so it is crucial to keep paddling throughout the course of the game to keep the boat at maximum speed.

There is also the possibility to brake by holding the paddle into the water (up and down keys on a desktop and touch-hold on a mobile device). The boat slows down quickly and rotates towards the side of the paddle in the water. This can be a useful

action for performing a tight turn once a checkpoint is reached and a new course has to be set.

We tried to model the movement of the boat as close as possible to the behavior of a 4-person canoe we recorded during test sessions on Lake Constance.

In order to make the game interesting and keep the time required for a mission below 3 minutes, we had to introduce a factor of around 2,700 to scale the speed of the boat[22]. Thus, paddling for 10 seconds in the game corresponds to a real voyage time of approximately 7.5 hours. We also scaled the boat to a size at which the apparent paddling speed in relation to the size of the boat is fun to play and yet maintains the impression of realistic boat speeds.

The velocity vector of the boat is the sum of the vector generated by the players actions and the current-vectors interpolated from the underlying grid. Due to the scaling of the boat, the currents can be quite different at bow and stern. We decided to incorporate this difference by calculating the influence of the currents independently for both bow and stern and then rotate the boat according to the resulting direction vector and translating it according to the average translation vector.

All major Caribbean islands are represented on the map. Their shape is stylized due to memory restrictions, yet they can be distinguished by their height profile which is adopted from a digital elevation model.

The player sees the boat through a perspective camera from behind. The next checkpoint is marked by a red triangle. There is also a head-up display that indicates the direction of the next checkpoint as well as the forces of the currents acting on the boat at the moment. Figure 1.28 shows a screen shot of the game.

Note that in a real world situation, it would be impossible to determine the direction of the current from a boat in deep water when there are no visual clues such as land marks. We added the indicator feature as compensation for other clues on environmental factors that people in real canoes might have been able to experience which can not be conveyed through the game.

If a player hits the last checkpoint, the mission is successful. A mission is failed if the player hits an island, runs over the "edge" of the game world, or voluntarily aborts the game. After each round, a map of the world is displayed with a track of the players route. Note that this is also additional information that was not available to paddlers in the past.

Even though real-world current patterns can be added into the game, the differences to an actual trip in a canoe can not be neglected. While the game provides a very simplified environment where the only influence are the currents, the decision for a route in the real world probably depended on many factors such as the wind- and wave-situation, the time of day and year, the objective of the voyage and many more. The player of the game only has to decide whether he wants to set his route more to starbord or to port.

[22]We use this factor as well to account for the curvature of the earth, and thus it varies between 2,745 and 2,632, depending on the latitude in the game and the heading of the boat

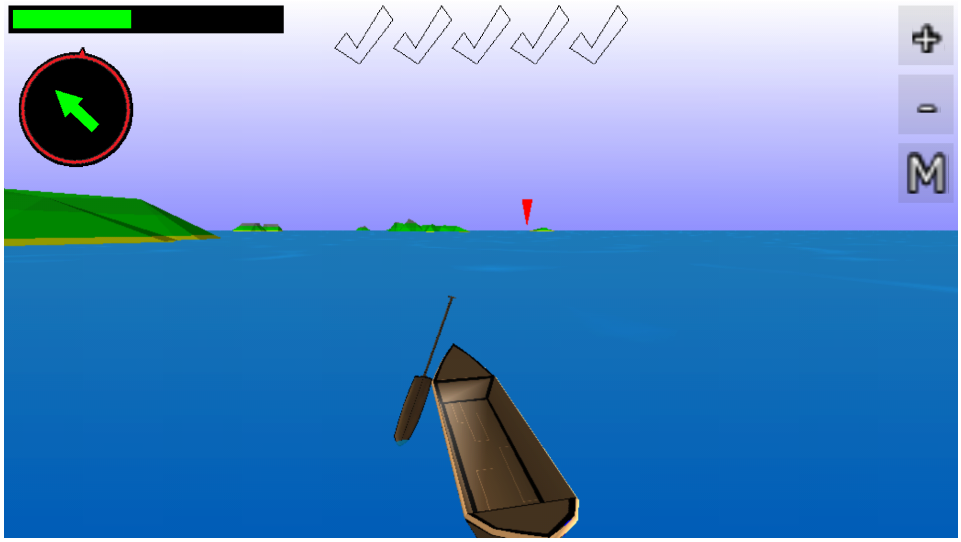


Figure 1.28: A screenshot of the game *RowYerBoat* on a mobile device. The next checkpoint is indicated by the triangle. The head-up display on the top left indicates the boat speed (green bar), the current direction (green arrow) as well as the direction towards the next checkpoint (little red arrow)

Yet, despite its arcade-like character, we would place the game in the scope of a *serious game* (Djaouti et al. 2011) since the prime motivation is its use to collect data on the player's behavior. In addition, the game is envisioned to be used for educational purposes in the Caribbean to teach children about the seafaring traditions of their ancestors.

1.5.2 The Competition

In order to answer our research question, we set up a competition right after the release of the game over the course of two weeks (June 30, 2017 through July 16, 2017). During that time, players could enter their e-mail address to be notified in case they won a prize. A total sum of 250 Euro was awarded to the players ranking at the top of the high scores.

For the competition, we implemented a tutorial mission in which the steering and the behavior of the boat in currents could be tested as well as three competitive missions. The missions were designed such that they covered the majority of islands in the Lesser Antilles. The first mission, *Pottery Acquisition* takes place in the northern part of the map. It starts west of Guadeloupe and the checkpoints follow the chain of islands, ending south of Puerto Rico (see Fig. 1.29a). The second mission, *Jaguar Teeth I* starts in Martinique, crosses over to Barbados, and then back to St. Vincent (checkpoints 1,2, and 3 on Fig. 1.29b). The third mission *Jaguar Theeth II* starts at the end of the previous mission and leads the player all the way to the north of Trinidad and then back to Martinique (checkpoints 3, 4, and 1 on Fig. 1.29a). During the course of the competition, we noticed that the major portion of the logs resulted from the *Pottery Acquisition* mission. In order to get a more even

(a) Checkpoints for the *Pottery Acquisition* mission

(b) Checkpoints for the *Jaguar Teeth* missions

Figure 1.29: Overview of the missions in *RowYerBoat*

distribution of the logs, we decided to release a new version of the game through the course of the competition in which the order of the missions was switched. This has led to a more even distribution (and even an over-compensation).

For each mission, there is a *static* and a *dynamic version*. In the static version, the currents remain the same every time the mission is played. In the dynamic version, a random current from a set of four different currents is chosen.

The current data used in the missions is based on the *Amseas* data set (see Section 1.2.3). We reduced the resolution used in the game to 1/10 degree instead of 1/30 degree in order to reduce the amount of data that has to be kept in memory.

We chose the currents at 00h00 UTC on the following dates based on their visual difference in current patterns: May 13, 2016 for the static missions and March 15, June 15, September 1 and November 15, 2016 for the dynamic missions. Unfortunately, during the competitions, there was an index-error in the read-in function of the current data. This led to a mix-up in the current data (see Fig. 1.30). We were still able to compare the track data and calculate optimal paths on the mixed up data, but the currents do not reflect the real world situation anymore. The problem has been resolved in newer releases of the game.

For each version of a mission, a global high-score is kept on a server. Whenever a player finishes a mission, the track gets uploaded (or, if the device is offline, uploaded once the internet connection is restored).

The competition was advertised through email newsletters within the *NEXUS1492* project as well as through flyers at the dining hall of the University of Konstanz. During the two weeks of the competition, we gathered a total of 2,920 logs with a total play time of 58 hours from 79 distinct players. However, most players played the game only a few times, so we decided to only analyze the data from players that had at least one hour of total play time (see Fig. 1.31a).

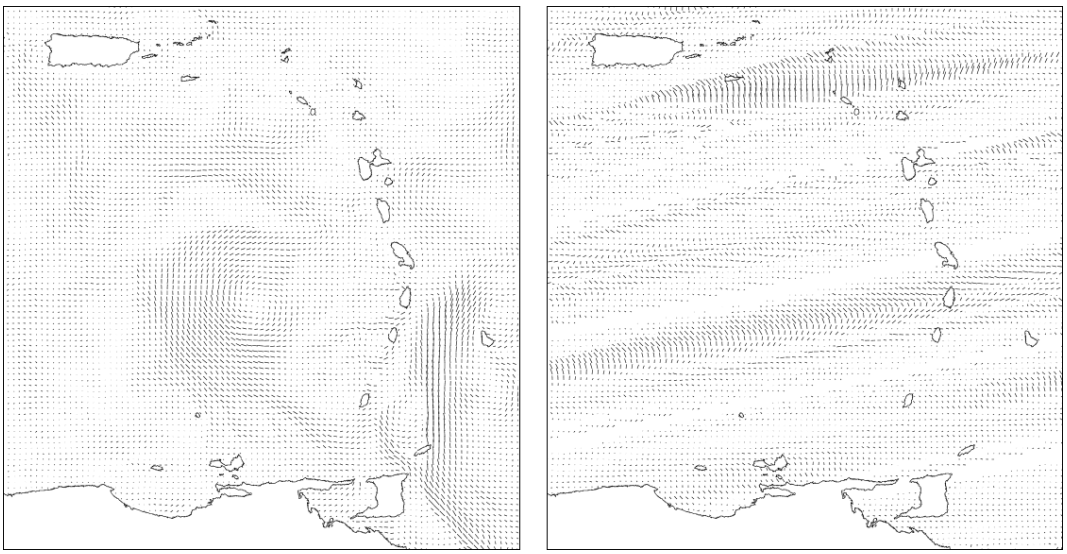


Figure 1.30: The actual current data and the current data in the game for May 13, resulting from an index-error (figure by Lamsal (2017))

We ordered the players by their play time and identified them as *Player A* through *Player L*. Table 1.31b shows that the players mostly engaged with the static missions. This is further emphasized by the observation that the dynamic missions were often quit within the first few seconds if the current did not seem suitable for the player (players confirmed this assumption in personal communications).

1.5.3 Analysis of the Tracks

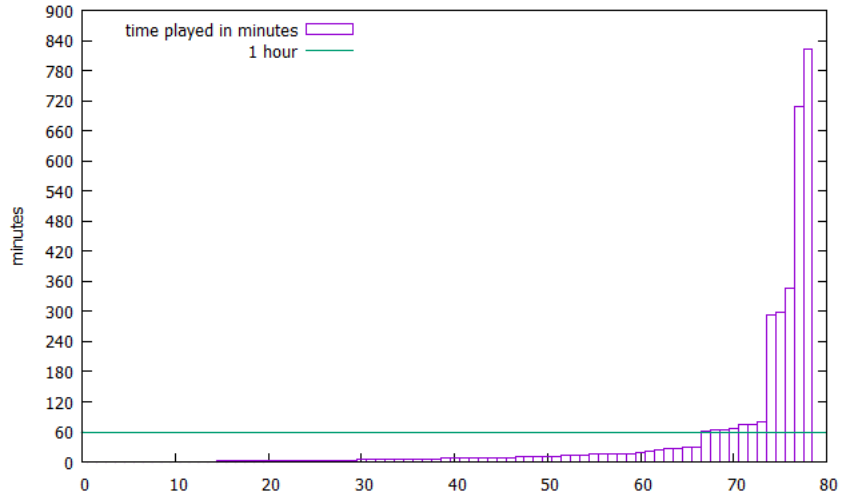
The results of the competition suggest that players are capable to improve their choice of routes. Yet, there is still a significant difference to the optimal path as computed by the algorithm in Section 1.4.

Since Player A and Player B are the players with by far the most rounds played, we will mainly focus on their learning behavior. A more extensive analysis of the results can be found in the master thesis by Roman (Lamsal 2017).

Similarity between Tracks

Each track of a player is logged on the server as a series of points recorded every ~ 0.2 seconds. In order to compare the tracks to the optimal route and among each other, we need a measure of similarity between tracks. Yet, there is no straightforward measure of similarity between trajectories, and many methods have been developed for specific tasks (Wang et al. 2013).

We decided to use the following two measures of similarity between tracks, depending on the use case:



(a) Distribution of the play times per player, Player A and B are on the right

PlayerID	jt1	jt2	pa	jt1Dyn	jt2Dyn	paDyn	Sum
Player A	340	168	156	111	15	45	835
Player B	109	29	109	107	57	18	429
Player C	126	17	47	5	23	21	239
Player D	39	16	16	32	5	27	135
Player E	29	1	11	12	48	27	128
Player F	8	4	10	7	6	2	37
Player G	12	10	13	0	0	1	36
Player H	10	12	10	7	2	3	44
Player I	23	8	3	2	3	1	40
Player J	2	0	3	0	0	1	6
Player K	5	6	8	10	7	1	37
Player L	3	1	14	2	0	0	20

(b) Total amount of attempts per player and mission, including unsuccessful rounds

Figure 1.31: Engagement of the players with the game during the competition (figures by Lamsal (2017))

(a) *Jaguar Teeth I* **(b)** *Jaguar Teeth II*

Figure 1.32: Comparison of similarity measures: the optimal pathway (green), and the closest tracks according to the Fréchet distance (red) and the simple Euclidean distance (blue)

Discrete Fréchet Distance The *discrete Fréchet distance* is the maximum length of a "leash" between two points of the polylines when taking into account their ordering, i.e., without the possibility of moving backwards (Eiter and Mannila 1994). The Fréchet distance of two points on a path is recursively defined as

$$D_f(p_i, q_j) = \max\{\min\{D_f(p_{i-1}, q_j), D_f(p_{i-1}, q_{j-1}), D_f(p_i, q_{j-1})\}, d_{p_i,q_j}\}.$$

where d_{p_i,q_j} is the Euclidean distance between two points. The Fréchet distance between two paths is defined as the Fréchet distance of their end points.

The measure attributes low distance to pairs of paths whose maximal distance is low and takes the ordering of the points into account. A potential drawback of this approach is that a track can be a very good fit in one part and then only locally deviate more than the maximum of another track which is deviating over large parts of the track. This can be undesired in our case since sometimes players performed very well at the beginning of a mission, only to deviate from the optimal track at the end. Figure 1.32a shows this issue: The maximum deviation of red track from the optimal route (green) is smaller than the one from the blue track, even though the red one is further apart from the optimal route most of the time.

Simple Euclidean Distance An alternative measure that focuses more on the average distance between the paths is what we call *simple Euclidean distance*. This is the average squared Euclidean distance of the points on one polyline to their nearest neighbor on the other polyline. Since the distance is not necessarily symmetrical, we take the larger value.

$$D_e(P,Q) = \max \left\{ \frac{\sum_{u \in P} |u - closest(u,Q)|}{|P|}, \frac{\sum_{v \in Q} |v - closest(v,P)|}{|Q|} \right\}$$

The advantage of this method is that it takes the average distance along the whole path and small local deviations are averaged out (see Fig. 1.32a). However, it does not take into account the ordering of the vertices. This can lead to a measurement of distance to the "wrong" vertices when a path has a turning point and returns close to itself. An example is the blue route in Fig. 1.32b: Both the blue and the red route deviate to the west in comparison to the optimal path (green). The blue route is a good fit according to the simple Euclidean distance but mostly follows return path of the optimal route in both directions. In this case, the Fréchet distance would be a more appropriate measure, even though the red track is still not aligning well with the optimal route.[23]

Clustering the Tracks

In order to visualize different strategies of the players, we found it helpful to aggregate similar tracks into a cluster. We used our own implementation of the PAM method for the k-medoids approach (Kaufman and Rousseeuw 1987, 1990), choosing the number of medoids k based on the visual distinctiveness of the clusters. For the *Pottery Acquisition* and the *Jaguar Teeth I* missions, $k = 5$ turned out to produce good results with the simple Euclidean distance as metric. In the case of the *Jaguar Teeth II* mission, we chose $k = 2$ and the Fréchet distance.

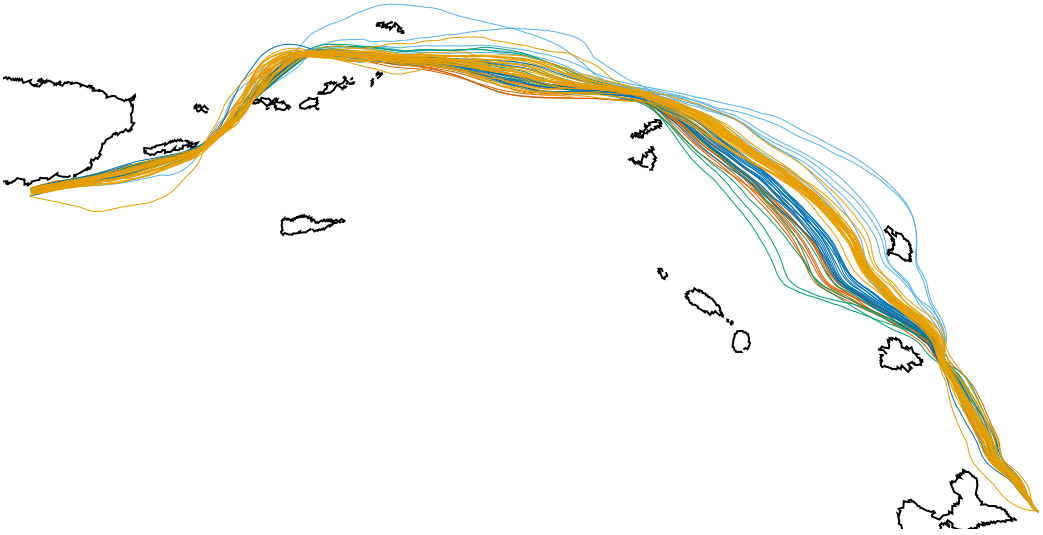
1.5.4 Results

In the following we will examine the learning behavior of selected players and analyze how closely the players were able to match the optimal routes. In addition to the analysis of the tracks, we were able to interview the two most active players (Player A and Player B) to verify our hypothesis about their behavior and strategies. (See the guidelines by Cote and Raz (2015) for such interviews). The transcript of the interviews can be found in Appendix A.

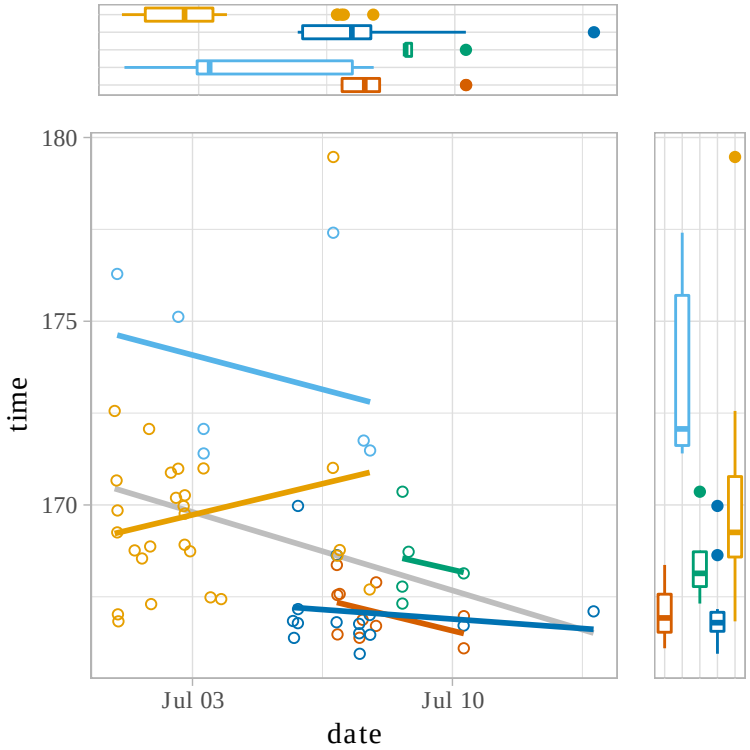
Learning Patterns of Selected Players

In the following we describe the learning patterns of Player A and Player B who were by far the most active players. From the interviews (Appendix A) we know that the players had a mental map of their current location and tried to optimize their routes via the map view at the end of the mission. Both players claim to have put most of the optimization efforts between the first two checkpoints. This can be confirmed for the *Jaguar Teeth I*-mission. Interestingly, the game tracks show that

[23] In Fig. 1.32b, the optimal pathway is routed through an island. This is an artefact caused by the index-error explained in Fig. 1.30



(a) Clustered trajectories



(b) Distribution of the route times depending on the date played. The linear regression lines indicate that the player got better over time

Figure 1.33: *Pottery Acquisition* Player A: Clustering with five clusters on the simple Euclidean distances. Each color represents a different cluster

for the *Pottery Acquisition* mission the most divergence for both players happens between checkpoints 2 and 3 and 3 and 4 (e.g., see Fig. 1.33a).

In order to analyze a player's progress and learning behavior, we created a scatter-plot of the player's tracks with the date of the log on the x-axis and the duration of the mission on the y-axis. A linear regression on the points can indicate a trend in the player's required time for completing the mission.

Due to the optimization attempts of the players, their times do not necessarily improve in a monotone way. We thus decided to cluster the tracks to visualize the learning behavior for each cluster separately. We also introduced box-plots for each cluster, indicating the range and distribution of times as well as dates for the tracks in the cluster.

Figure 1.33 shows the analysis of all successful tracks of Player A for the *Pottery Acquisition* mission. The tracks (Fig. 1.33a) have been clustered based on the Euclidean distance ($k = 5$). The time taken for each track is plotted in Fig. 1.33b in relation to the date the track was recorded. The linear regression of all tracks (gray line) indicates a trend of overall improvement of the player's times for the mission. With the exception of the orange cluster, the regression lines of all clusters indicate an improvement over time.

The development of the clusters over the dates played illustrates nicely the learning behavior of the player. While at first mostly taking the North-Easterly routes in the light-blue and orange cluster, around July 5th, the player discovers the quicker paths more to the South-West in the dark-blue and red cluster. Around July 8th, the player starts experimenting with even more South-Easterly paths (green cluster) but eventually sticks to the middle ground for the remaining few logs.

We refer to Appendix B for the remaining visualizations of Player A's and Player B's learning behavior on the static missions.

Comparison to Optimal Tracks

While players gradually learned to improve their tracks, none of the players managed to match the optimal route over the course of a mission. There are cases however where a player came close to the optimum during a part of the mission but deviated at a different part. Figure 1.32 shows the path which is closest to the optimal path according to the two measures used. The simple Euclidean distance is a good method to find such paths if there is no turning point such that paths come close to themselves (as in Fig. 1.32a). For the case that the path describes a back-and forth movement, the Fréchet Distance provides more meaningful results (as in Fig. 1.32b).

Figure 1.34 shows the fastest tracks of Player A and B in comparison to the calculated optimal routes for the *Jaguar Theeth 1* mission.

There are several reasons why players might not have been able to come close to the optimal paths. Firstly, the task is not an easy optimization problem. While there is an improvement over time for the players, it might well be that the number of attempts per mission were not enough to come close to the optimum in the large

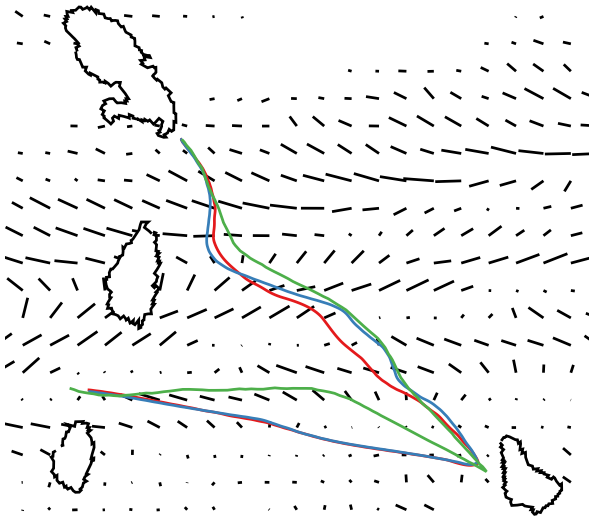


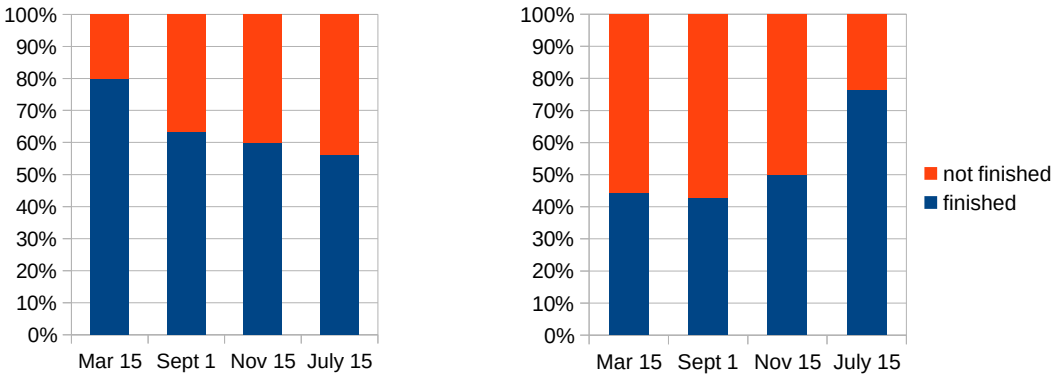
Figure 1.34: Jaguar Theeth 1: The fastest routes of Player A (red) and Player B (blue) in comparison to the optimal route (green). The black dashes indicate the underlying currents

space of possible routes. Secondly, just because one trajectory was determined to be optimal does not mean that there are no alternative trajectories that would result in times almost as good as the calculated optimum. Thirdly, there are practical differences between the model of the canoe used for the simulation of the optimal paths and the behavior of the boat in the game: The boat in the simulation always runs at the maximum speed, and changes in direction can be performed immediately and at arbitrary sharp angles. In contrast, the boat in the game takes some time to accelerate, and at high speeds the rate of turn is rather limited. Sharp turns can only be made by "breaking" with a paddle in the water, resulting in a loss of speed and the need to accelerate again. Additionally, the checkpoints in the game were implemented as a circular area, allowing a deviation of the routes in the area of the checkpoints.

Dynamic Missions

Since the majority of the tracks were recorded for the static missions, we focused our analysis of the learning behavior on these tracks.

However, it is worth noting that the dynamic missions led to a behavior of the players that we did not anticipate: Players learned which one of the set of four currents was the most favorable one and quit their missions if they noticed that they were paddling in another current. Especially Player B was very selective as to when he kept paddling and when he voluntarily quit: In both of the dynamic *Jaguar Teeth* missions, there is a current scenario for which the proportion of successful rounds is significantly higher than in the other three scenarios (see Fig. 1.35). In both cases, almost all of the non-finished attempts were voluntarily aborted. This strategy was confirmed by Player B in personal communication.



(a) *Jaguar Teeth I* **(b)** *Jaguar Teeth II*

Figure 1.35: Proportion of successful rounds for Player B in the dynamic missions, depending on the current scenario

We can assume that this behavior is similar to what people in the past might have done: only starting a voyage if certain clues indicate that it will be successful and reasonably fast.

1.6 Conclusion

In this part of the thesis, we described methods to analyze environmental patterns, to simulate optimal canoe routes, and to evaluate human ability to optimize pathways. When applied to archaeological research questions, these methods have the potential to provide additional insights into patterns of migration and exchange between island communities.

The methods for visualization of changes in speed and direction of current (and potentially wind and wave) data we presented in Section 1.2 allow archaeologists to explore seasonal variations and other patterns in the data. These trends can – in combination with provenance studies of archaeological material – indicate seasonal activities on certain islands. However, our analysis of the patterns in the example locations suggests that in the current data alone there is no simple answer to the question whether a certain season was more suitable for traveling or whether there are pathways that are not navigable by canoes during certain times.

In order to further examine the suitability of certain seasons for canoe voyages and to learn about the trajectories of such routes, we developed methods for simulating optimal canoe routes (Section 1.4). Since existing methods did not cater for the special circumstances of our study, we had to come up with extensions, allowing for current speeds approaching the speed of the boat and including areas behind islands. The results of these simulations – when averaged over multiple model runs during a certain time period – can be used as an indicator for the strength of ties between two sites. In addition, the pathways from the model runs of the case studies can indicate potential stopover locations and highlight the importance of sites located at strategic locations.

Since the pathways generated in the case studies are based on complete knowledge of the environmental factors, they are likely to deviate from the actual routes taken by the indigenous peoples. Through the development of the canoeing game *RowYerBoat*, we tried to evaluate to what degree human actors are capable of approximating these optimal routes. Our experiments in Section 1.5 suggest that human actors are generally capable of learning about current patterns even though the optimal pathways were not reached in our experiments.

Despite the differences between paddling in the game and the real paddling experience, the game can still provide useful insights to human behavior to an extent that would not be possible with traditional methods from (experimental) archaeology. We think that the use of serious games in archaeology has the potential to greatly facilitate the collection of data regarding (past) human behavior. As a side effect, such games provide a good way for community outreach and can raise awareness for archaeological projects.

An interesting project for the future would be the creation of an (interactive) map of Caribbean canoe routes based on simulations with our algorithms. Such a map could show the optimal routes between all known sites in the area depending on the selected season. The map could be combined with a community outreach project in which school children or the visitors of a museum could compete against each other in the *RowYerBoat* game and then see their tracks on the map in relation to the optimal routes.

Another interesting project would be to use the environmental factors to calculate pathways based on the characteristics of European sailing ships from the 16th century. A comparison to the routes of the canoes could lead to interesting insights that could supplement written accounts of colonial encounters.

Part 2

Similarities of Pottery: Quantifying Human Perception

Sherds of ceramic vessels are by far the most common material remains of the later indigenous populations in Caribbean islands. The vessels have been used for a variety of purposes, including storage, presentation, cooking, and rituals. The possibility to infer ancient cultural practices from the types of ceramic vessels found at a site makes ceramic sherds a highly interesting field of study for archaeologists (Greene and Moore 2010; Henson 2012; Bahn 2012). In the course of the *NEXUS1492* project, many different features of these ceramic remains are studied: The material composition, the provenance of the material, the way the pots were fired, coiling techniques, etc.

In this part of the thesis, we focus on the shape and other visible features of the vessels. Since the mid-1980s, all ceramic sherds found by the Caribbean Research Group of Leiden University have been recorded according to the same scheme. We will explore ways to establish a typology from the database of sherds based on the recorded attributes and how they are perceived by humans. The typology can later be used to compare sites based on their similarity in the frequencies of the types found (Habiba et al. 2018).

The goal of our studies is to explore new approaches for classification with the aim to establish a more objective typology than the existing ones. We still incorporate the knowledge of experts in the field while at the same time eliminating potential bias as much as possible. Our approaches record the knowledge of experts in a series of questionnaires, establishing distances that allow the application of clustering methods based on the recorded knowledge to any database of sherds from the region. The database only needs to contain very low-level information on the sherds, which can be recorded by untrained students or research assistants.

We begin this part by giving a brief overview of the history of ceramic analysis in archaeology and about existing classification schemes (Section 2.1).

In Section 2.2, we will explain the recording scheme used by researchers from Leiden University, present an attempt to digitally reconstruct pots from this information,

and give an outlook on possible extensions of the recording scheme based on our work.

We then give an overview of existing methods of cultural domain analysis used to quantify human perception (Section 2.3) before introducing two methods for the quantification of perceived similarities: two-dimensional arrangements and multidimensional distance recordings (Section 2.4).

To evaluate our new (as well as some existing) methods of cultural domain analysis in the context of pottery perception, we conduct three exemplary user studies:

- We compare our proposed method of multidimensional distance recording to the established method of triad tests (Section 2.5) based on test results from a ceramicist.
- We use three-dimensional reconstructions of pots in an online survey consisting of triad tests to measure the perceived relevance of specific attributes (Section 2.6).
- We perform a hands-on study with archaeologists and native American potters in the US-Soutwest where we document the interaction with real archaeological material and record perceived similarities of sherds (Section 2.7).

2.1 Overview of (Caribbean) Pottery Analysis

The production of pottery by firing clay is a cultural technique that was used by nearly all human cultures around the globe for the production of ceramic vessels for storage, cooking, or presentation of food as well as for the creation of decorative objects (Orton and Hughes 2013)

Even though ceramic pots are prone to breaking and have a limited lifespan when used as a cooking vessel, the remaining sherds are very durable especially when protected from erosion by a covering layer of soil.

Especially in the context of the Caribbean where organic materials decompose very quickly due to the climatic conditions, pottery makes up a major portion of the material remains found at sites from the ceramic age[1]. Due to the lack of organic material (and thus the opportunity to gather C14 dates[2]), the analysis of pottery styles became an important factor in determining a site's age.

Besides qualitative studies of form and function of pottery, a way for archaeologists to better understand patterns and styles in pottery creation is the *classification* of pottery into groups (and potentially subgroups) based on stylistic (e.g., form, shape, decoration) or technological features (e.g., temper, clay type, finishing technique). Such a *typology* of ceramics does not only facilitate the exchange of information among archaeologists (i.e., when talking about specific styles), but is also crucial for

[1]The ceramic age in the Caribbean began between 500 BC and 500 AD, depending on the island

[2]The C14 method (radiocarbon dating) is a way to determine the age of organic material by measuring the concentration of the carbon isotope ^{14}C

techniques in *seriation*[3] as well as the reconstruction of networks based on similarities (Habiba et al. 2018).

2.1.1 Pottery Classification and its History

The classification (or typology) of pottery is the establishment of groups of sherds that are more similar to each other than to sherds in other groups. In our case, similarity of two sherds means that they have a number of common or similar features. Classification is an important first step for the quantitative analysis of the pottery of a given site or region of archaeological interest. It allows the comparison of distributions of types between sites or within a site on both temporal and spatial levels.

A typology is the basis for techniques such as seriation (Petrie 1899; Brainerd 1951; Robinson 1951) as well as the reconstruction of archaeological networks based on similarity measures (Brughmans 2010; Mills et al. 2013; Habiba et al. 2018). Horejs et al. (2010) provide a good overview of existing approaches to classification, ranging from rather subjective decision making to the use of scientific methods to measure attributes.

Despite these use cases, the necessity for a classification scheme once was heavily debated among archaeologists. In the early days of quantitative archaeology, the introduction of methods such as the ones developed by Brainerd and Robinson (Brainerd 1951; Robinson 1951) led to a division among archaeologists. The most prominent discourse on the topic is know as the *Ford-Spaulding Debate*: Spaulding (1953) argued that new methods of seriation and typology provided a way for archaeologists to produce meaningful types and give the field more scientific credibility. Ford (1954) replied that the methods did not reflect how cultures evolved. He was arguing in favor of approaches to classification that encompass more the point of view of the creators of the pots. According to Ford, the dividing lines between types should be based on proven historical factors instead of the subjective views of archaeologists who want to establish a "nice" descriptive order. Over the years, a majority of archaeologists adopted Spaulding's quantitative approach while trying to use Ford's more organic process in their argumentation (Lyman et al. 1997).

In many archaeological areas, pottery typologies were established based on archaeological expertise (Horejs et al. 2010). Existing systems for classification often have been developed in the very beginning of the archaeological excavations in a region and therefore might not be suitable for the more recent findings in the area. This leads to sometimes ambiguous definitions of types and results in categories that are not clearly separable. There is also a debate between the philosophies of "lumpers" (people in favor of establishing a limited number of fairly general categories) and "splitters" (people who favor a more fine-grained system of categories). This issue has been discussed among archaeologists already in the 1950s (Wheat et al. 1958).

Another problem with such organically grown classification systems is that they rely

[3]Seriation is a technique to establish a chronological order of layers based on the types of objects found (Greene and Moore 2010)

on expert knowledge and expertise, which makes it more difficult to involve student helpers in the classification of sherds in the field.

Over the last decades, a number of approaches to automate the classification of pottery have been proposed. Already in the early 80s, Rice and Saffer (1982) used clustering techniques for provenance studies of pottery based on its chemical composition. Gilboa et al. (2004) analyze and classify sherds based on scanned drawings of profiles, while Hörr et al. (2008) use techniques from machine learning based on 3D-scans and Makridis and Daras (2013) use a similar approach based on photographs. Karasik and Smilansky (2011) cluster sherds based on the mathematical function describing their profile. Lucena et al. (2016) do the same with a polyline approximating the profile.

In contrast to theses approaches, we propose classification methods based on prerecorded features from a pottery database which we describe in the following section. The goal is to apply clustering methods to the entries of the database based on distances between the features that were recorded through structured questioning of experts. We do this by applying established methods from the field of *Cultural Domain Analysis* (Section 2.3) as well as our own methods to quantify perceived distances between and within attributes of ceramic sherds (Section 2.4). These distances can then in turn be used to cluster the entries of a database of sherds.

With these techniques, we hope to contribute to bringing together the ideas of Spaulding and Ford by quantifying the human perception of pottery such that a resulting classification reflects the similarities as perceived by human actors. Note that our goal in this study is not to replace existing typologies, but rather to provide means for a more data-driven analysis of pottery. In fact, when used in combination with existing typologies, clusterings based on the recorded perception of archaeologists have the potential to reveal and quantify their implicit assumptions in the classification process. This can help students and researchers to get a better understanding of the underlying thought processes of classification by experts.

2.1.2 Caribbean Pottery

Due to the lack of preserved organic material, the Caribbean was one of the early testing grounds for ceramic typology (Hofman et al. 2008). Thus, the debates about the need for a classification system did not leave the archaeology in the region untouched.

Enforced by the heterogeneity of the region and the different nationalities (and thus archaeological traditions) of archaeologists working in the area, a large variety of chronological and cultural typologies were introduced in the region from the 1940s until the end of the 1970s. We refer to (Hofman et al. 2008, pp. 2–5) for a list of relevant literature on the topic. Over time, most researchers in the area adopted the ceramic typological framework proposed by Rouse (1960, 1965) in which new ceramic styles are named after the location they were first found at (i.e., “Huecoid”, “Saladoid”, and “Troumassoid” for styles found in the Lesser Antillies, and “Ostionoid”, “Meillacoid”, and “Chicoid” for styles found in the Greater Antilles).

Recognizing the potential disconnect between the *etic*[4] view of archaeologists interpreting ceramics and the *emic*[5] view of the producers of those ceramics, Caribbean archaeologists tried to improve their interpretations of functionality of ceramics with ethnographic work (Barbotin 1974; Petitjean-Roget 1963). This is in accordance with the early arguments by Rouse (1960), stating that typological analyses are bifurcated between natural and artificial categories.

In Rouse's classification scheme, archaeologists construct types by selecting modes (groupings of artifact attributes) that they determine to be relevant. The modes that are selected, are – in Rouse's view – inherent to the archaeological material and thus to the culture that created that material. However – being established in the early days of Caribbean archaeology – the classification system has led to a rather rigid typology based on the (limited) knowledge about Caribbean ceramics in the 1960s. Even though the system has been extended over the years, it is still based on its early foundation. Rouse's system thus does not cope well with the integration of recent discoveries, in particular when multiple ceramic styles, previously thought to originate from clearly separated cultural backgrounds, are blended together into new (sub)types.

An additional problem with such a system is that, since there is no set of clearly defined rules for the classification, the archaeologist performing the classification task might be biased by previous experience or the lack thereof.

As a supplement to using Rouse's system for the classification of site assemblages, the researchers from Leiden University have recorded low-level attributes for each sherd they excavated and aggregated the information in a database. We will now describe the recording scheme used in this database.

2.2 Pottery Data Collected by Leiden University

Since the 1980s, the Caribbean Archaeology Group at Leiden University has been applying a specific system to describe pottery sherds. First introduced by Hofman (1993), it has been used for the analysis of ceramic material from around 30 different sites across the Caribbean.

In a typical excavation, a site is divided into squares of one-by-one meter, which are excavated in layers of 10 centimeters. For each layer in each square, sherds are counted and weighed, and the number of sherds with special features is noted. Special features can be e.g., red slip, specific decorations, or whether the sherd is part of the rim or the base of the vessel, or an appendage such as a handle or a modeled ornament.

Special attention is given to rim sherds (i.e., sherds stemming from a section of the top ledge of the vessel) larger than 5 centimeters, as these sherds are most indicative of the original shape and function of a vessel (see Fig. 2.1). Each of the rim sherds is classified based on the different features described in a code book. Due to this

[4] From the outside
[5] From within the group

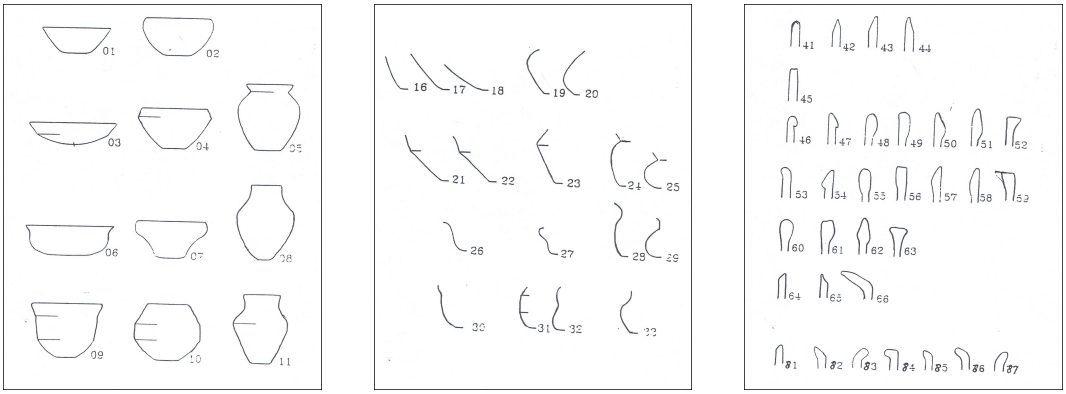
detailed information recorded on a per-sherd basis, we will set our focus entirely on rim sherds in the remainder of this work. In the next section we will describe the criteria which are recorded for each sherd.

(a) Outside **(b)** Inside

Figure 2.1: Example of a (relatively large) rim sherd

2.2.1 The Code Book of Caribbean Ceramics

The *Code Book of Caribbean Ceramics*[6] or *Code Book* from the the Caribbean Research Group at Leiden University (Hofman 1993) describes a number of *attributes* and their *domain of possible values*.



(a) Vessel shapes **(b)** Wall profiles **(c)** Lip shapes and rim profiles

Figure 2.2: Examples on the graphical representation of the possible values in the Code Book

The main focus of the Code Book lays on the classification of rim sherds. It contains a textual description for each possible value of an attribute and – for attributes describing shapes – a sketch illustrating the shape (Fig.2.2). There are also a number of codes reserved for *other* values of attributes. These codes can be assigned for

[6]This work is based on the "edition 2005" of the book

types that are site specific and can vary from excavation to excavation. Due to this variability, we decided to ignore them in this work. In the future it is envisioned to merge these codes in a unified database (see Section 2.2.4).

The following features are recorded for each rim sherd larger than 5 centimeters:

Vessel Shape The general shape of the vessel is recorded based on the characterization by Shepard (1956). The code numbers 01 to 15 are possible values for the attribute *vessel shape*, depending on the restrictedness of the vessel and the complexity of its contour. (See Fig. 2.2a. Codes 12–15 are not defined and reserved for *other* shapes.)

Wall Profile The *wall profile* allows for a more detailed description of the vessel shape. It takes into account information, e.g., on the height-diameter ratio of the vessel and where the largest diameter is located. It is a refinement of the vessel shape attribute in the sense that it is possible to infer the code of the vessel shape from the code of the wall profile. The code numbers 16 to 40 are possible values for the wall profile (codes 34–40 for *other*). See Fig. 2.2b.

Lip Shape The *lip shape* of a vessel describes the upper part of the rim. The lip can be shaped in a large variety of ways, helping to make the vessel more robust or facilitate the pouring of liquids. Lip shapes can be round, flat, inward, outward, or double thickened as well as bevelled, combined with different taperings or wedges. Code numbers 41-80 are possible values for the lip shape with 67–80 reserved for *other*. See Fig. 2.2c, upper part.

Rim Profile The angle between the lip and the neck/wall of the vessel is referred to as *rim profile*. It can be either straight or inward or outward flaring to various degrees. Codes are 81–90 (88–90 for *other*). See Fig. 2.2c, bottom.

Wall Thickness The *wall thickness* is measured two centimeters under the rim and recorded in millimeters.

Diameter The *diameter* of the vessel is estimated by placing the sherd on a template with circles of various sizes. It is recorded in centimeters. Together with the diameter, the percentage of the rim that is present is recorded, allowing to estimate the accuracy of the diameter.

Decoration Code numbers 91 to 110 are reserved for modes of *decoration*. However, these depend heavily on the site and period and have been changed over time. For now, this makes a comparison of sites based on this attribute impossible. For the future, it is planned to go through the reports and make the decoration modes consistent throughout the whole database.

Color The *color* of the sherd on both the inside and the outside is recorded by comparing the sherd to the soil color charts by Munsell (1912). Code numbers 111–140 correspond to different hues and color numbers in the Munsell charts (133–140 for *other*).

Firing Color With clippers, a fresh break is produced on an edge of the sherd and the combination of the core and sub-surface color is recorded by comparing it to the Musell charts. Codes 141–165 are used (160–165 for *other*).

Surface Finishing The finishing of the surface (e.g., smooth, crude, scratched...) is recorded for both inside and outside of the sherd. Possible code numbers are 166–175 (172–175 for *other*).

Red Slip Red slip is a special treatment to the surface, making it smooth and sealing it off. Code numbers 176–185 describe the location of red slip found on the sherd (e.g., outside, lip, inside and lip...) (183–185 are used for *other*).

During each excavation, student helpers analyze each rim sherd and record the codes on a separate sheet of paper for each layer of an excavated square. The sheets then get manually transferred into a database for the excavation. The combination of these databases into a unified *ceramics database* provides a detailed recording of all sherds found during the excavations of the Caribbean Research Group over the last 30 years.

In the next section we describe how we reconstruct the shape of a vessel based on the information of rim sherds from the Ceramics Database.

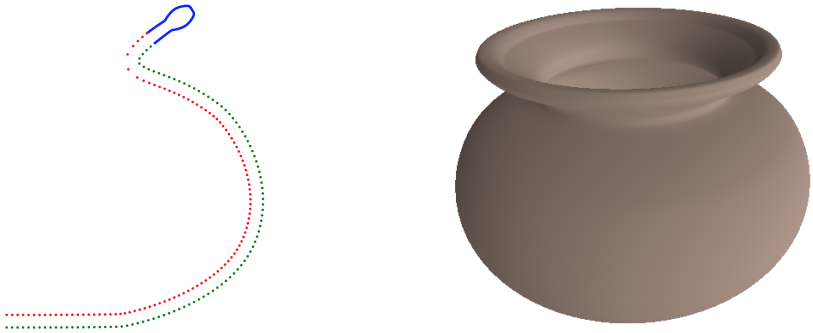
2.2.2 Reconstructing Pottery from the Database

As a basis for our online perception tests in Section 2.6 and as a way to visually represent entries of the Ceramics Database, we wanted to digitally reconstruct the recorded pots. We thus developed the browser-based software *PotBuilder* to visualize what a pot might have looked like based on the recorded attributes from the Code Book.

The software allows the user to select the possible values of the following attributes from the pottery database: wall profile, lip shape, wall thickness, diameter, and color. From this information, the drawing of a profile is automatically created. The profile can then be turned into a three-dimensional rotation object, and the user can interactively explore the resulting pot.

The reason we did not (yet) include the rim profile was an unclear description of the concrete portion of the pot where the rim starts. Also other features like the firing atmosphere, decoration, and the distinction between inside and outside color are omitted for now.

As underlying data we use SVG-paths for each wall profile and lip shape as well as a translation of the Munsell chart to RGB values. The paths for the wall profile are formatted in a way that their starting point (the center at the bottom of the vessel) is set to the origin of the drawing space, and that the normal at this point is vertical. The path from the lip shape is scaled dynamically in such a way that the distance between its two end points corresponds to the thickness parameter. The path for the wall profile is scaled dynamically such that the x-coordinate of the top point corresponds to half of the diameter of the pot.



(a) Points of the polyline for the profile

(b) Three-dimensional rendering of the rotation object

Figure 2.3: Reconstruction of a pot in the *PotBuilder* with the following settings: wall profile 25, lip shape 61, diameter 30 cm, thickness 10 mm, color 125 (see Section 2.2.1 for an explanation of the attributes)

When constructing a pot, we start sampling the path from its starting point, creating a polyline for the profile. Once we reach the point where the lip shape starts (which is determined by the height of the lip shape scaled by the thickness parameter), we start offsetting our sampling by the deviation of the path on the outside of the lip shape from the vertical line in the direction of the normal of the wall profile path. Since the amount of detail is highest for the lip shape, we sample it with a higher resolution.

Once we reached the end point of the path for the wall profile (which corresponds to the end top point of the path of the lip shape), we start sampling downwards again, now offsetting by the inside of the lip shape. When the end of the path for the lip shape is reached, we keep sampling the path for the wall profile, offset by the thickness parameter in the direction of the normal.

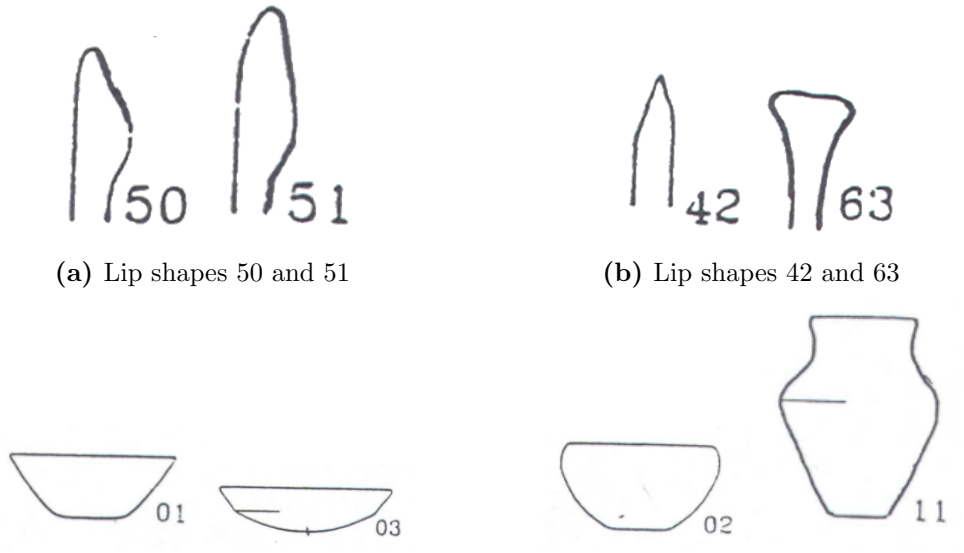
The resulting polyline can then be transformed into a three-dimensional rotation object with the *LatheGeometry* function from the JavaScript library *THREE.js*[7]. Figure 2.3 shows the profile as well as the 3D-view in the PotBuilder. The green dots show the sampling of the wall profile for the outside, blue is the sampling of the lip shape, and red the offset sampling of the wall profile for the inside.

2.2.3 The Problem of Classification

The recordings in the ceramics database are so detailed that it is hard to get an idea of the overall composition of an assemblage from the raw data. As a basis for seriation or the comparison of different sites to each other, we need a classification of the pottery into coarser categories. In the following sections we present our approach to this problem.

In order to create a clustering of sherds into groups, we need to establish a distance-function between sherds based on their attributes. An obvious distance function

[7] `https://threejs.org/`



(a) Lip shapes 50 and 51

(b) Lip shapes 42 and 63

(c) Vessel shapes 1 and 3

(d) Vessel shapes 2 and 11

Figure 2.4: The differences between attributes can range from small variations to major differences (drawings from the *Code Book of Caribbean Ceramics*)

between two sherds with nominal attributes would be the number of different attributes between two types or a more elaborate distance function that takes mixed attributes as input, such as the one from Gower (1971).

The drawback of this approach is that it gives equal weight to all attributes and, most importantly, that it does not take into account that some possible values of an attribute might be more similar to each other than to others. A different number in, for example, the lip-shape attribute can indicate a completely different pottery-style (i.e., if it is 42 instead of 63) or just a slight variation of the lip (i.e., 50 instead of 51, see Fig. 2.4a). In this case, the sherds might even be part of the same pot.[8]

A measure for the distances within a category is also crucial to detect possible errors in classification since we can assume for example that it is more likely that vessel shape 1 and 3 are confounded (Fig. 2.4c) than vessel shape 2 and 11 (Fig. 2.4d). The textual descriptions in the Code Book already give a hint on the distances between the codes: Vessel shape 1 is described as *unrestricted vessel with a simple contour*, vessel shape 2 as *restricted vessel with a simple contour*, vessel shape 3 as *unrestricted vessel with a composite contour*, and vessel shape 11 as *independent restricted vessel with a complex contour*. Yet, the textual descriptions do not provide a quantitative measurement of difference.

For the problem of weighting the attributes and their possible values, we propose methods to quantify human perception in Section 2.4 and apply them in case studies in Sections 2.5, 2.6, and 2.7.

[8]Note that the difference in the code numbers is not suitable as an indicator for the difference between the corresponding shapes

2.2.4 Outlook: The Future of the Code Book

The analysis of the pottery data sets entered by students at the Faculty of Archaeology at Leiden University has shown that the current process of recording the sherds is very prone to errors. In addition, the use of the *PotBuilder* has revealed that there are some wall profiles found in the Caribbean that are currently not described in the Code Book. Furthermore, in order to reconstruct a pot, there are some parameters missing such as the location at which the bend of the rim profile starts as well as the thickness of the lip shape.

These observations have led to an extension of the Code Book and the development of a web app that allows students to enter the data directly into the database by selecting the attributes on the touch screen interface of a mobile device. The *PotBuilder* will be extended and included in the system to provide students and researchers a live preview of what the original pot from a recorded sherd might have looked like.

Before we proceed to proposing our own methods to quantify human perception (Section 2.4), we will now give an overview of related methods from the area of cultural domain analysis.

2.3 Overview: Cultural Domain Analysis

In the field of cultural anthropology, *cultural domain analysis (CDA)* describes a set of methods with the goal to understand the semantic structure of cultural domains, i.e., the mental categories of people regarding a set of words, images, or other items (Borgatti 1994, 1999).

These methods have traditionally been applied in situations where perceptions of groups of people were studied or compared. One strength of these methods is that they allow to measure not only information about perceived attributes of an item (*monadic data*) but also the perceived relationships between items (*dyadic data*). In the following sections, we will apply methods from CDA to evaluate the perceived similarity of pottery types and sherds.

We now briefly introduce some common techniques for CDA. We refer to Borgatti (1999) for a more detailed description.

Freelists

Freelists are a techinque in which participants are simply asked to name all items of a specific domain. Besides being used to establish a set of items in the domain, the resulting lists can also be analyzed regarding the frequency of certain items or a ranking based on the order in which the items were named (monadic). Also, a network of the participants can be constructed based on the co-occurrence of certain items (dyadic).

Rank Order

In a *rank order* task, participants are asked to establish an ordering of the items based on a specific attribute. The resulting monadic attributes can be transformed into similarity values between the participants (dyadic), indicating whether there are groups with specific views. In the case study in Section 2.7 we use this technique for one of the tasks.

Pairwise Rating

The most straightforward way to get dyadic information about perceived similarity of items is a *pairwise rating*. In this interview technique, participants are asked for each pair of objects to rank on a scale how well they go together or how (dis)similar they are. Unfortunately, there are $\frac{n(n-1)}{2}$ different pairs for n items, leading to very time consuming tests, even for smaller n.

Pile Sorts

With the help of *pile sorts*, a similarity matrix can be constructed based on how participants group objects. A shuffled set of objects (usually words or images on cards, but real objects have also been used) is given to the participants. In a *free pile sort*, the task is to form piles, grouping similar objects together. From these clusters, a similarity matrix can be created by interpreting the similarity of two objects as the number of participants that put these objects on the same pile. Boster and Johnson (1989) use this technique to compare the perceived similarity of fish species between novices and experts. A common problem is that some people form only very few piles ("lumpers") while others make very fine-grained distinctions ("splitters"). One way to deal with this problem is to ask participants to successively split up or aggregate their groups. This procedure is called *successive pile sort* (Boster 1994).

Triad Tests

Triad tests are an alternative method to quantify distances between different items. They have been introduced by Torgerson (1951) for psychometric studies and later been adopted to anthropology by Romney and d'Andrade (1964). In a triad test, a participant is presented a series of 3-tuples (triads) of items and asked to determine the most different one. This choice is used as an indicator for a stronger similarity between the remaining two items.

Since each pair appears multiple times, triad tests generate a more nuanced similarity matrix than pile sorts. The drawback is that in a complete triad test, every possible 3-tuple has to be evaluated, leading to a total number of $\frac{n\cdot(n-1)\cdot(n-2)}{6}$ triads for n items. This proves to be impractical even for small numbers of items. As the number grows cubically, we already have 1,140 triads to examine for $n = 20$ which by far exceeds the reasonable amount of questions in a user study.

For this reason, Burton and Nerlove (1976) introduced the concept of *balanced incomplete block designs (BIBD)* to triad tests. A block is a subset of items presented to the user, i.e., in our case, the blocks are the triads. Such a design contains only a subset of the triads but with the property of being balanced, i.e., each pair of items occurs exactly the same amount of times. This leads to a reduction of the number of possible items. We will now give a formal definition of (balanced incomplete) block designs and available heuristics for their construction:

Block Designs An incomplete block design is defined by the following tuple of parameters: (v, b, r, k, λ), where

- v: number of items (often called treatments)
- b: total number of blocks, in our case the number of triads
- r: number of blocks in which each unique item appears
- k: number of items in each block, in the case of a triad test $k = 3$
- λ: number of times each pair appears in a block

An extensive overview of the topic has been compiled by Stinson (2004). A design is *balanced* if r is the same for each item (*first order balance*) and λ is the same for each pair of items (*second order balance*). In balanced block designs, not all parameters are independent, and v, k, and λ completely determine the design. The following equations hold:

$$b = \frac{v \cdot (v-1) \cdot \lambda}{2 \cdot k} \tag{2.1}$$

$$bk = vr \tag{2.2}$$

$$\lambda(v-1) = r(k-1) \tag{2.3}$$

A block design is called *simple*, if no block appears twice in the design. A block design with $k = 3$ is also referred as *triple system* $TS(v, \lambda)$ of order v and index λ (Colbourn and Lindner 1992) or as a *Steiner triple system* of order v when $\lambda = 1$ and $k = 3$.

Unfortunately, there is no general algorithmic approach to generating simple BIBDs. There are, however, software packages for non-simple BIBD's: the R-package *ibd*[9] and the Gendex Package *IBD*[10]. The algorithms used in these packages are based on a linear programming approach (Mandal 2015) and an iterative improvement strategy (Nguyen 1994). However, both algorithms do not necessarily generate simple block designs.

There exist some constructive proofs for special cases: Dehon (1983) proves that there exists a simple $(v, 3, \lambda)$ design if and only if $\lambda v(v-1) = 0 \mod 6$ and $\lambda(v-1) = 0 \mod 2$. He bases his proof on an earlier result from Teirlinck (1977) who shows among other cases that for $\lambda = 6$, the maximum number of mutually disjoint

[9] `https://cran.r-project.org/web/packages/ibd/index.html`
[10] `http://designcomputing.net/gendex/ibd/`

balanced simple BIBDs is $\frac{v-2}{6}$ for all $v = 2 \mod 6$. We use some of these results for the construction of the triad tests in the case studies in Section 2.5 and 2.6.

In the following section, we will introduce modifications to the methods described here that are tailored to the specific questions we have in the case of analyzing pottery and put them to the test in the following case studies.

2.4 Quantifying Perceived Distances through Spatial Arrangements

We will now present two methods to quantify knowledge or perception of distances between objects. The first method (Section 2.4.1) has the potential to record a large amount of (hidden) features from participants who do not necessarily need to have an in-depth knowledge of the field. The second method (Section 2.4.2) requires more in-depth knowledge of the participant and preparation and engagement during the process, yet provides more insight into the process of decision making.

2.4.1 Two-Dimensional Arrangements

The first method can be seen as an extension of the *rank order* task described in Section 2.3. Given a set of items (i.e., cards with words or images or real objects), the participants are asked to arrange them on a square surface (or a computer screen) such that "objects that are more similar are placed closer together and objects that are less similar are placed further apart". The method has been proposed before by Goldstone (1994) and is also known as "Spatial Arrangement Method" (SpAM).

The question can be asked in an undirected way or with a specific criterion of similarity, e.g., "made by similar groups of people" or "serving a similar purpose".

During the test it is important to tell the participant that there is no right or wrong answer, but that the goal is to find out their personal opinion on the relationships between the items.

Once the participant is satisfied with their arrangement, a picture is taken of the whole square surface. Using an image editor or similar tool, the pixel positions of the corners of the square surface as well as the locations of the center of each item are tagged and recorded. (In the case of the arrangement of items on a screen, the pixel coordinates of the items can directly be stored.)

Before we can start the analysis, all positions have to be projected into the unit square. If the positions are recorded from a photograph, we first need to apply a homography[11] to correct distortions of the arrangement induced by the angle of the camera. The necessary parameters for the transformation from the quadrilateral marked by the positions of the four corner points to the 1x1 square can be calculated based on the methods described by Criminisi et al. (1999).

[11] A homography is a projection of the vector space that preserves straight lines

Now all pairwise (Euclidean) distances of the items can be computed and normalized such that the largest distance between two objects is 1. The resulting distance matrix can be analyzed using the methods we describe in Section 2.4.3.

We use the method as an alternative to the triad tests which are often seen as tedious by participants and very time consuming. A two-dimensional arrangement exercise allows us to record the perceived distances of 30 or more items within a reasonable time (usually 10–30 minutes, depending on the participant) as opposed to a triad test that would encompass at least 290 questions (if a BIBD with $\lambda = 2$ was used) or a pairwise comparison consisting of 435 questions.

Even though two-dimensional arrangements have been critiqued for having caveats in comparison to asking for a pairwise evaluation of distances (Verheyen et al. 2016), they have been shown to produce reasonable good results within a fraction of the time required for pairwise evaluation (Hout et al. 2013; Hout and Goldinger 2016).

An obvious drawback of the method is that the placement of the objects is restricted by the two-dimensional space. The hope is that this leads participants to focus on the most important similarities and thus the intuitive construction of something similar to a multidimensional scaling (see Section 2.4.3). When averaging the results of a sufficient number of participants, the resulting distance matrix should provide a close representation of the (average) mental model of the distances of the participants.

For the case that the number of participants is limited (or just a single expert is interviewed) and the time is less restricted, we developed the following extension to the method that does not restrict the data to two dimensions.

2.4.2 Multidimensional Distance Recording (MDR)

The following method requires a close collaboration of the interviewer with the participant. It also requires the participant to have a basic understanding of interval scales and nominal and ordinal categories.

We begin with a free listing task (see Section 2.3) in which the participant is asked to list all relevant criteria according to which they think the items can be categorized. The specific formulation of the question depends on the context and the familiarity of the participant with the items as well as the goals of the test.

For each of the determined criteria, the interviewer and the participant determine together whether the criterion is best described on an interval scale or rather fits into k nominal or ordinal categories. If there is only a single participant whose knowledge is to be quantified, we can now directly proceed to the second part of the test. Otherwise the free listing answers of all participants should be collected first and then aggregated into a final list of all relevant criteria.

Now the participant is asked to arrange the items according to each criterion. This can either be done with the real objects, with cards with names or images of the objects, or with the help of a computer program in which digital representations of the items can be arranged via drag and drop. The instructions for the arrangement are as follows:

- If the criterion fits into k nominal categories, the participant is asked to form k piles and place each item on its corresponding pile.
- If the criterion fits into k ordinal categories, the participant is asked to form k piles, ordered from 1 to k and place the items accordingly.
- If the criterion can be described on an interval scale, the participant is asked to spatially arrange the items along a line according to the perceived value of the items on the scale. If the participant is confident enough, a two-dimensional arrangement can be used to record two different criteria in the same exercise.

The values are recorded and aggregated into a table. Nominal categories are labeled with roman letters, ordinal categories with integers, and intervals by normalized float values.[12]

In an optional third step, the participant can be asked to judge the relative importance of the criteria, for example by arranging the criteria on an interval scale.

From the table of the recorded values, we can create a distance matrix for the possible values for each attribute. However, there is no straightforward way to create distances between mixed attributes on nominal, ordinal, and interval scales. The most commonly used approach has been proposed by Gower (1971) whose metric calculates distances between variables with mixed attributes consisting of nominal, interval, and ranked attributes. It has already been used in an archaeological context in a provenience study by Rice and Saffer (1982). Gower's general formula for the distance G_{jk} between two objects j and k with mixed attribute vectors $x_{1j} \dots x_{nj}$ and $x_{1k} \dots x_{nk}$ is the following:

$$G_{jk} = \frac{\sum_{i=1}^{n} w_{ijk} s_{ijk}}{\sum_{i=1}^{n} w_{ijk}}$$

- $w_{ijk} = 0$ if object j or k is unknown, otherwise $w_{ijk} = 1$ (or the weight for attribute i in the case of weighted attributes)
- $s_{ijk} = 0$ if $x_{ij} \neq x_{ik}$ and $s_{ijk} = 1$ if $x_{ij} = x_{ik}$ for nominal and binary variables (an alternative measure sets w_{ijk} and s_{ijk} to 0 for binary variables that are both 0)
- $s_{ijk} = 1 - (\frac{\|x_{ij} - x_{ik}\|}{r_k})$ for variables on the interval or ratio scale, where r_k is the range of attribute k, i.e., the difference between its maximum and its minimum

We use the extension of Gower's method by Podani (1999) which handles ordinal attributes that are not required to be a complete ranking.[13] Ordinal attributes first get replaced by their rank r_{ij}, where the same values result in the same rank. If $r_{ij} = r_{ik}$, s_{ijk} in the formula above is set to 1. Otherwise,

$$s_{ijk} = 1 - \frac{\|r_{ij} - r_{ik}\| - (T_{ij} - 1)/2 - (T_{ik} - 1)/2}{max(r_i) - min(r_i) - (T_{i,max} - 1)/2 - (T_{i,min} - 1)/2}$$

[12] See Section 2.4.1 for details on how to project the tags from photographs into the unit square

[13] The functions are implemented in R as *daisy* and *gowdis* in the packages *cluster* and *FD*, respectively

where T_{ij} is the number of objects with the same rank score for i as object j, including j. The formula for s_{ijk} expresses the distance between two objects as the number of objects between them based on the ranking, normalized by the number of objects between the extreme values.

If data on the relative importance of the attributes was recorded, this information can be used to weight the criteria in the distance function accordingly.

2.4.3 Analyzing Spatial Arrangements

We will now give a brief overview of methods to analyze the spatial arrangements and distance matrices obtained from the methods described above. The case studies in the following sections give examples of how the methods can be used in practice.

(Classical) Multidimensional Scaling

After averaging the distance matrices from multiple participants in a two-dimensional arrangement exercise or from the calculation of the Gower-distance in an MDR-study, we end up with a (symmetric) matrix of distances between n objects. In most cases, it will not be possible to find an arrangement of the objects in the two-dimensional space that represents all distances.[14]

Multidimensional scaling (Torgerson 1951, 1958) is a method to project the n-dimensional space to the k "most relevant" dimensions. The most commonly used methods for generating MDS representations use the fact that the most relevant dimensions are spanned by the eigenvectors of the largest eigenvalues of the distance matrix. In order to create visual representations, we set $k = 2$, giving us two-dimensional positions of the n items that most closely represent their distances in the given distance matrix.

We use the function *cmdscale* from the R-package *stats* to calculate the MDS representations in the following sections.

Modularity

Modularity (Newman and Girvan 2004; Newman 2006) is a widely used concept from the field of network science to measure how well the structure of a network corresponds to a given division into clusters ("modules"). It is defined as the fraction of intra-cluster edges[15] minus the expected fraction of intra-cluster edges in a random network with the same degree distribution. For weighted graphs it is defined as follows:

$$Q = \frac{1}{2m} \sum_{ij} (A_{ij} - \frac{k_i k_j}{2m} \delta(c_i, c_j))$$

[14] For correctly representing the distances between n objects in the general case, we need $n - 1$ dimensions

[15] Edges connecting nodes that are in the same cluster

A_{ij} is the weight of the edge between node i and node j, k_i the weighted degree of node i (i.e., the sum of the edge weights of all edges incident to i), m is the sum of all edge weights, c_i is the cluster of node i, and $\delta(c_i, c_j)$ is a function that is 1 if $c_i = c_j$ and 0 otherwise.

In the following case studies we use the R-package *igraph* to calculate modularity on weighted similarity matrices.

While calculating the modularity of a given clustering is simple, finding a clustering that optimizes modularity is an NP-complete problem[16] (Brandes et al. 2007). However there are greedy methods that work reasonably well in practice such as the Louvain method (Blondel et al. 2008) that we use in Section 2.5.

Mantel Test

A common method used to determine the correlation between two distance matrices is the *Mantel Test*, first introduced by Mantel (1967). Since the values in a distance matrix are not independent, a simple correlation coefficient on the values of the matrix would not produce meaningful results for the significance of the test. The Mantel test is based on a high number of random permutations of the rows and columns of the matrices where the significance is the proportion of permutations that lead to a higher coefficient.

We use the function *mantel* from the R-package *vegan* to calculate correlations between distance matrices in the following sections.

Kendall's τ

The rank correlation coefficient *Kendall's* τ is used in statistics to measure the similarity of orderings (Kendall 1938). Its range is from 1 if the rankings are the same to -1 if the rankings are inverted.

It is defined as

$$\tau = \frac{p_c - p_d}{n(n-1)/2}$$

where p_c is the number of *concordant pairs*, i.e., pairs of values whose order is the same in both rankings, and p_d is the number of *discordant pairs*, i.e., pairs of values whose order differs in the rankings.

In the analysis for the case studies we use the function *cor* from the *stats* package with the "kendall" option.

Procrustes Analysis

Procrustes analysis is a method from statistics used to compare shapes of objects described by a set of points (Dean 2000). The point sets are superimposed by

[16] See Section 3.2.2 for an explanation of this concept

finding an optimal translation, rotation, and uniform scaling that minimizes the sum of squared distances between the points. The square root of the minimal sum of squared distances is called *Procrustes distance*.

In our case studies we use the *procrustes* function from the R-package *vegan* with the parameters *scale* and *symmetric* both set to *TRUE*.

2.5 Case Study: Multidimensional Distance Recording vs. Triad Tests

In this section our aim is to explore how the results from a perception study with the multidimensional distance recording (MDR) compare to the results of a triad test. As both methods are very time consuming and the MDR requires a deep understanding of the field, we limited this case study to one participant, Katarina Jacobson. Jacobson is a ceramic analyst within the *NEXUS1492* project with a special focus on production techniques of Caribbean pottery.

We used nine attributes from the Code Book and followed the procedure of the MDR test described in Section 2.4.2. For the comparison, we administered a triad test with the same participant for each attribute. Each triad consisted of three different possible values for the attribute. Note that due to technical reasons, we had to administer the MDR first. This might potentially have influenced the results of the answers of the triad test, as the participant was already conscious of the criteria she selected in the MDR-part of the test. However, there was a time span of several weeks between the tests. Also, since by the design of the triad test, an independent choice of which criterion is relevant has to be made for every single triad and the participant was an expert in the field who is very aware of the possible criteria, we do not think that this order had a great influence on the results of the tests.

This study is joint work with Katarina Jacobson (University of Leiden) who did not only take the (considerable amount of) time to record the tests but also provided valuable insights into the details of ceramic analysis and helped develop the methods used. Part of this work has been presented at the XXXVI Sunbelt Conference 2016 (Athenstädt et al. 2016).

2.5.1 Multidimensional Distance Recording (MDR)

For the MDR part of the test, we implemented a tool that allows the participant to arrange the possible values of an attribute in the two-dimensional space based on a given criterion. The criteria for each attribute were obtained by a (guided) free-listing exercise from the participant. For each criterion, it was decided in accordance with the participant whether it is best described on a continuous (interval) scale or as nominal or ordinal categories. In the latter case, the number of required categories was determined, and the tool was set up accordingly.

Figure 2.5 shows the user interface of our survey tool. The participant can drag and drop the shapes from the Code Book on the screen, arranging them into nominal

(a) Lip shapes classified into six nominal categories

(b) Vessel shapes classified into four ordinal categories

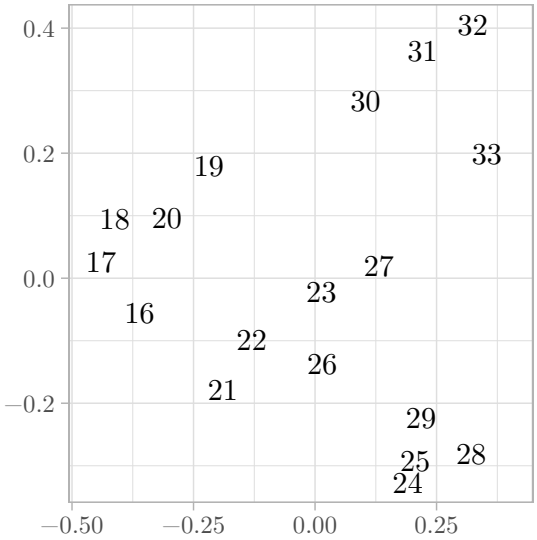
(c) Wall profiles arranged in two-dimensional interval space

Figure 2.5: The different types of scales in the MDR survey tool

(Fig. 2.5a) or ordinal (Fig. 2.5b) categories or in a two-dimensional interval space (Fig. 2.5c). Multiple repetitions based on different criteria are possible for the same attribute. It is crucial that the semantic of the criteria given by the participant is documented during the whole process, in particular the meaning of the two dimensions in the interval space. The test can only produce meaningful results if the participant is informed about and aware of how their choices will influence the outcome of the study.

	contour (ordinal)	proportions (ordinal)	orifice (nominal)
16	0	2	A
17	0	1	A
18	0	0	A
19	0	1	C
20	0	1	C
21	1	1	A
22	1	1	A
23	1	2	C
24	1	2	B
25	1	1	B
26	2	1	A
27	2	1	C
28	2	2	B
29	2	1	B
30	3	1	A
31	3	1	C
32	3	1	C
33	3	1	B

(a) Three of the recorded criteria



(b) MDS of the Gower distances (based on all criteria)

Figure 2.6: Table of three recorded criteria and the MDS from the Gower distance of all dimensions for the wall profiles

The results of the recordings are aggregated into a multidimensional position vector for each possible value of an attribute. Table 2.6a shows an example with the result of three recordings for each possibility of the wall profile.

With the help of the extended Gower distance (see Section 2.4.2), we can trans-

form the vector of attributes into a distance matrix. Figure 2.6b shows the results of an MDS on the distances generated for the wall profile attribute. We refer to Appendix C for the results of the other attributes.

Aggregation of Multiple Attributes

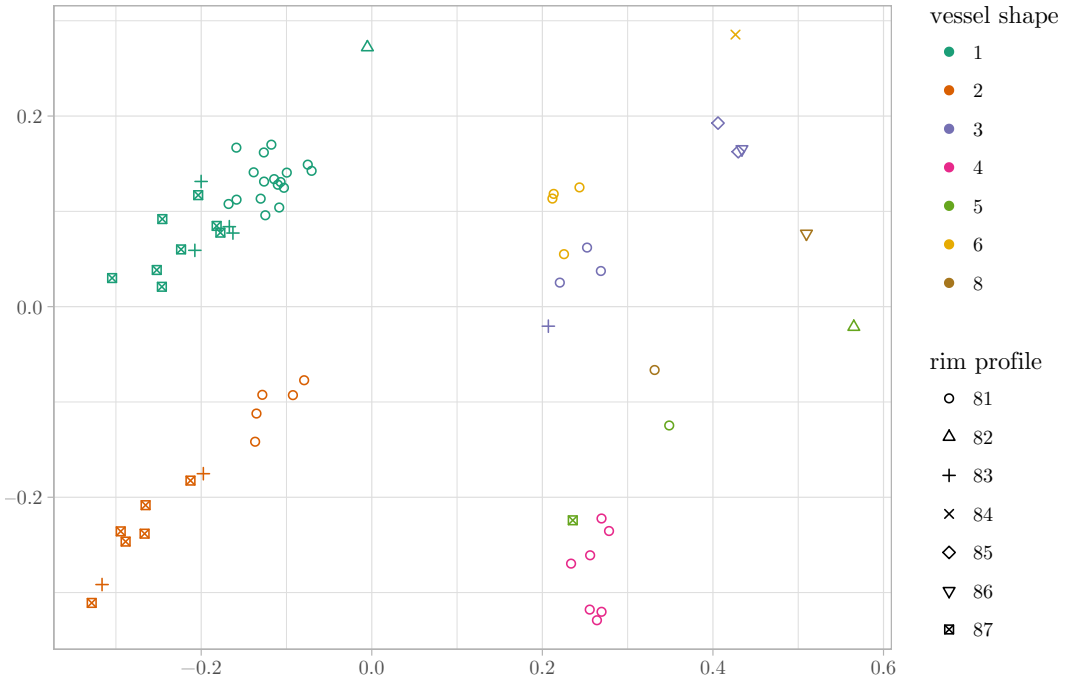


Figure 2.7: MDS of the Gower distances for the *Argyle* site from the combination of the four shape attributes from the MDR-recordings

As a first test on how the recorded distances can help in creating a typology, we calculated aggregated distances between variations of sherds from *Argyle*, a site on St. Vincent. In order to limit the amount of different combinations of attributes, we only incorporated the shape attributes, i.e., vessel shape, wall profile, lip shape and rim profile.

A total of 67 different combinations of these four attributes were recorded for the rim sherds of the site. We calculated the Gower distance for the combination of attributes with the following weights of the criteria: If an attribute had n recorded criteria, we weight each criterion by $\frac{1}{n}$ in the combined vector of criteria. This results in an equal weight for each attribute.[17]

Figure 2.7 shows an MDS of the Gower distances between the combinations of attributes from the *Argyle* site. In order to illustrate how the distances relate to the attributes, we used colors to indicate the vessel shape attribute and symbols to indicate the rim profile.

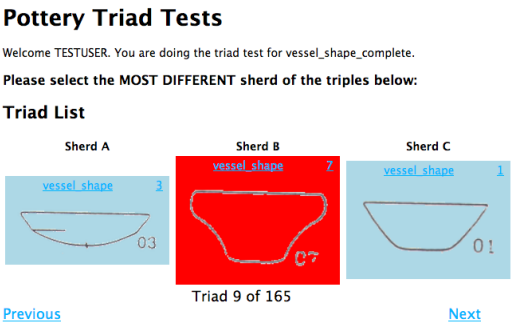
[17] We will explore ways to find a better weighting for the attributes in Section 2.6

In a next step, one could calculate a clustering on the distances and use this clustering as the basis for a typology. We leave this step for future work, as the recordings we use only come from a single participant and we think that more emphasis has to be put on the collection of data before producing a meaningful typology.

This section gave an example on how the MDR method can be used to record distances. To validate the method, we will compare the results to the triad tests we describe below.

2.5.2 Triad Tests

In order to create a benchmark for the MDR method, we asked the participant to take triad tests for all nominal categories of the Code Book. For attributes where the amount of triads was reasonable, we used the complete set, otherwise we constructed BIBDs (see Section 2.3): In the case of lip shape, we used the constructive proof from Teirlinck (1975), for the case of color and firing color, we repeatedly ran the Gendex Package IBD until we found a BIBD that was simple. The table in Fig. 2.8b shows the number of triads for each attribute b and whether a full test was administered or, if not, which λ was used for the construction of a BIBD.



(a) Interface of the triad test. Variation B has been selected as most different

	v	λ	b
vessel shape	11	full (9)	165
wall profile	18	full (16)	816
lip shape	26	6	650
rim profile	7	full (5)	35
decoration	11	full (9)	165
color	22	2	154
firing color	19	4	228
surface finishing	6	full (4)	20
red slip	7	full (5)	35

(b) Details on the triad tests performed for each attribute: Number of possible values v, λ if a BIBD was used, and the number of the resulting triads b

Figure 2.8: Triad tests were administered for all (non-interval) attributes

We developed an online survey tool to record the participant's choice of the most different value of the attribute (Fig. 2.8a). Before the test, we randomized both the attribution of treatments to the possible values of the attributes and the order of the treatments.

From the recorded results, we created similarity networks with weighted edges, where the weight of an edge indicates how often the pair was in a triad for which the third option was chosen as most different. Figure 2.9 shows the resulting network for the wall profile attribute. The layout is an MDS-layout with stress minimization as implemented in *visone* (Brandes and Wagner 2003). The desired length of the edges is set to be proportional to the inverse of the edge weights. The thickness of the

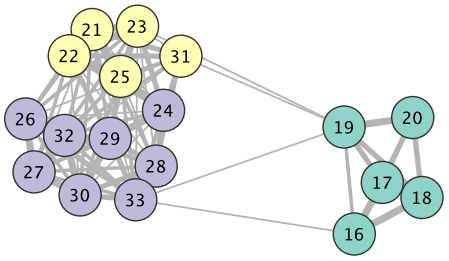


Figure 2.9: Network resulting from the triad test of the wall profile attribute

edges corresponds to the weight. The coloring of the nodes is based on a Louvain clustering on weighted edges. See Appendix C for the generated networks from the other attributes.

In order to also create a benchmark for the MDR of the combination of shapes attributes at the *Argyle* site (Fig. 2.7), we also combined the recordings of the triad tests for the site. Figure 2.10 shows the similarities of all combinations of the

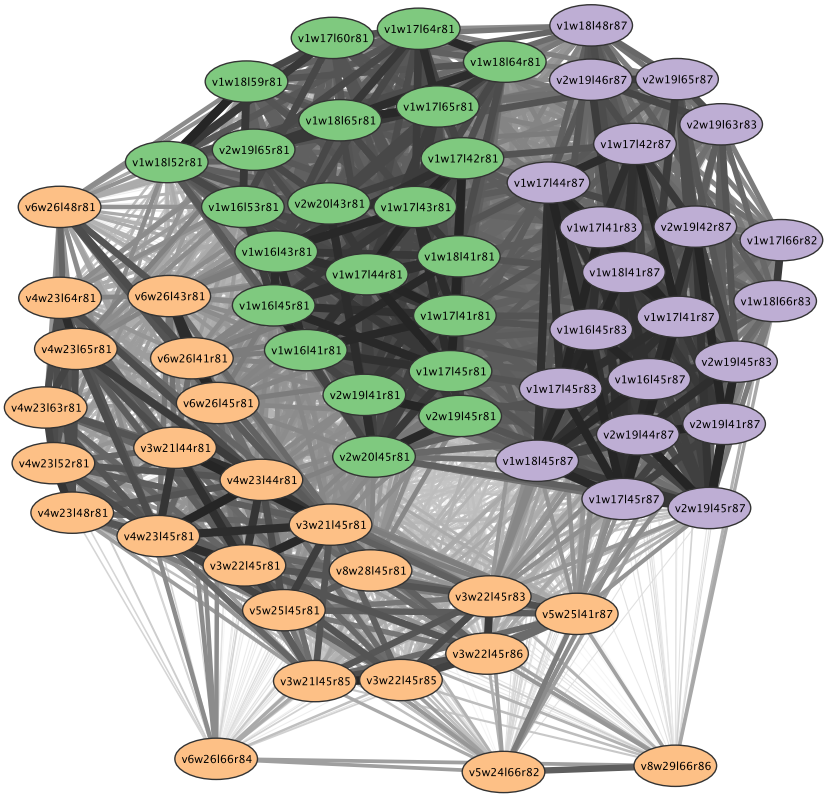
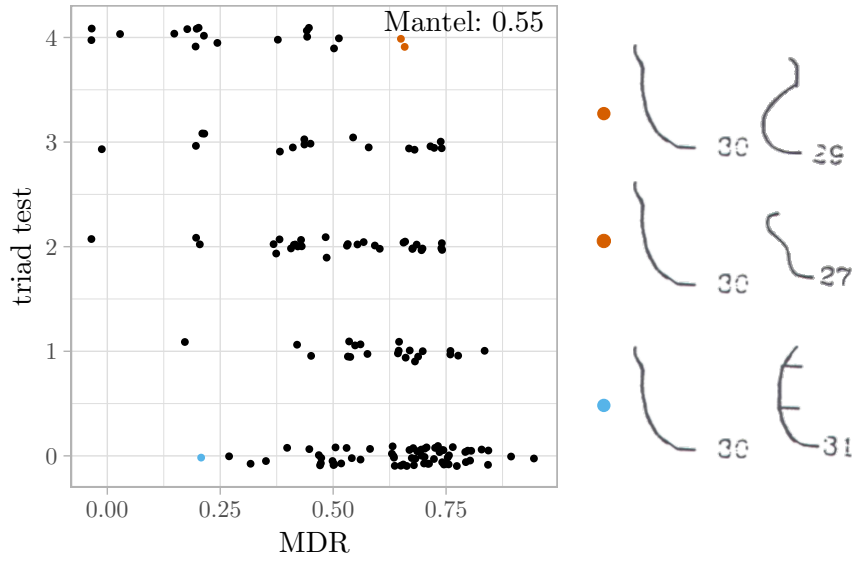


Figure 2.10: Resulting network from the combination of the similarities of the triad tests of the shape attributes for the *Argyle* site (stress minimization layout with overlap removal). Edge width and saturation are proportional to the similarity. Node labels describe combination of attributes, e.g., v2w17l43r81 corresponds to vessel shape 2, wall profile 17, lip shape 43, and rim profile 81

shape attributes found at the *Argyle* site. Each node represents one (or several) sherds found at the site with the combination of attributes indicated by the label. The similarities between the sherds are calculated by summing up the similarities for each of their attributes. The colors are based on a Louvain clustering on the (weighted) network. This would be a possible way to generate a typology based on triad tests of the possible values of the individual attributes.

2.5.3 Comparison of the Methods



(a) Each dot represents a pair of possible values **(b)** Outliers in detail

Figure 2.11: Comparison of the distances from the MDR to the similarities from the triad tests for the wall profile attribute (jitter has been applied to show overlapping points)

In order to compare the results of the two methods, we plotted the edge-weights of the triad tests against the pairwise Gower distances from the MDR. Figure 2.11 shows the plots for wall profile attribute. Since one of the methods produces a distance matrix and the other a similarity matrix, we can expect a negative correlation indicating agreement of the two methods.

The results of the Mantel test (with inverted values for the triad test) indicate that there is a (negative) correlation between the two methods for the wall profile attribute, but the images also reveal that there are cases where the two methods disagree.

Interestingly, the highlighted outliers all involve wall profile 30. The orange dots indicate the recorded relationship between wall profile 30 and wall profile 29 and 27, respectively. They have a maximal similarity of 4 according to the triad test but also a relatively large distance of 0.7 according to the MDR. In contrast, wall profile 30 and wall profile 31 have no similarity according to the triad test but a relatively

small MDR distance of 0.2. According to the recorded MDR, they only differ in one criterion, the *orifice* of the vessel, whereas wall profile 30 differs from 29 and 27 in three criteria, namely the *orifice*, the *contour*, and the *shape*.

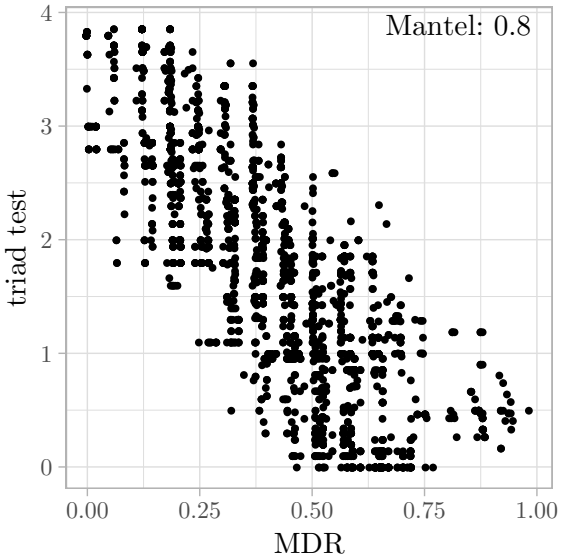


Figure 2.12: Comparison of the MDR recordings to the triad tests of the combination of the shape attributes for the *Argyle* site

Figure 2.12 shows the Gower distances of the combined shape attributes for the *Argyle* site in comparison to the combinations of the triad tests. The result of the Mantel test of (-)0.8 shows that through the combination of the attributes the differences between the methods get averaged out.

	Mantel correlation r	simulated significance p
vessel shape	0.6371	0.001
wall profile	0.5473	0.001
lip shape	0.6848	0.001
rim profile	0.1315	0.288
decoration	0.7873	0.001
color	0.2972	0.001
firing color	0.7526	0.001
surface finishing	0.8072	0.0027778
red slip	-1.335e-17	0.524

Table 2.1: Results of the Mantel tests for the attributes

Table 2.1 shows the results of the Mantel tests for all individual attributes. With the exception of rim profile and red slip, there is a clear correlation between the MDR distances and the (inverted) similarities resulting from the triad test. However, for some of the attributes, the correlation is rather weak. These are typically attributes for which it was difficult to determine multiple dimensions. One such example is color, where the participant only chose one dimension (the "tone" of the color). A

limited number of dimensions also is reflected in the MDS of an attribute when many values are mapped to the same position (see Appendix C). In these cases, triad tests seem to provide more meaningful results.

The advantage of the MDR method is that it allows to control which factors led to the distances by the choice of dimensions. In this way, the important criteria can be directly integrated into the process of recording the data. In the case of the Code Book for Caribbean Ceramics, instead of recording the codes for the attributes, one could instead directly record the relevant criteria such as the orifice or the contour of the vessel. However, since the dimensions have to be determined beforehand, there is the risk of missing important factors or overemphasizing ones that are not so important. Furthermore, it is not clear how to choose a proper weight for the dimensions.

Triad tests on the other hand do not require predetermined dimensions. This could deliver an implicit weighting of the factors but might also lead to a potential bias towards one factor while disregarding others. In general, triad tests are easier to understand for participants but are more time-consuming and often require the use of BIBDs which are not always easy to construct. The MDR method is faster in execution but requires explanation and guidance in determining which dimensions should be incorporated.

2.6 Case Study: Extracting Weights of the Factors

In this section, we describe a first pilot study in quantifying the perceived importance of the attributes in the Code Book with the help of triad tests. In order to keep the amount of questions manageable, we decided to limit the study to three indicative attributes (which we also implemented in the pot-builder): wall profile, lip shape and color.

Despite this limitation to just three attributes, combining all possible values of these attributes (see Section 2.2.1) would still lead to a total of 13,662 combinations to be tested. Even a reduction of the number of tests with the use of BIBD's would result in a triad test with at least a quadratic number of test questions, i.e., more than 180 million. Obviously this is not feasible.

We thus decided to only test a subset of the possible values per attribute based on the networks generated from the triad tests of the ceramic analyst in Section 2.5 (see Appendix C for the networks of lip shape and color). We chose the values of the attributes which were most distinct according to the networks by selecting an independent set (i.e., nodes that are not connected) of three values for each attribute with the additional requirement that one value from each cluster is present. For the wall profile attribute, we chose values 18, 25, and 26, for lip shape[18] we chose 43, 55, and 63, and for color, we chose values 114, 119 and 132 (see Fig. 2.13).

[18] We deliberately excluded the outlier 66, as it is currently discussed to remove this shape from the Code Book in the future since it resembles more a rim profile

(a) Wall profile **(b)** Wall profile **(c)** Color

Figure 2.13: The selected values for each attribute for the triad test

2.6.1 Design of the Survey

We generated a three-dimensional reconstruction for each possible combination of the values in the *PotBuilder* (Section 2.2.2), resulting in a total of 27 pots. Since we had the possibility to have the test executed online by a variety of different people, we decided to test the full set of 2,925 triads.

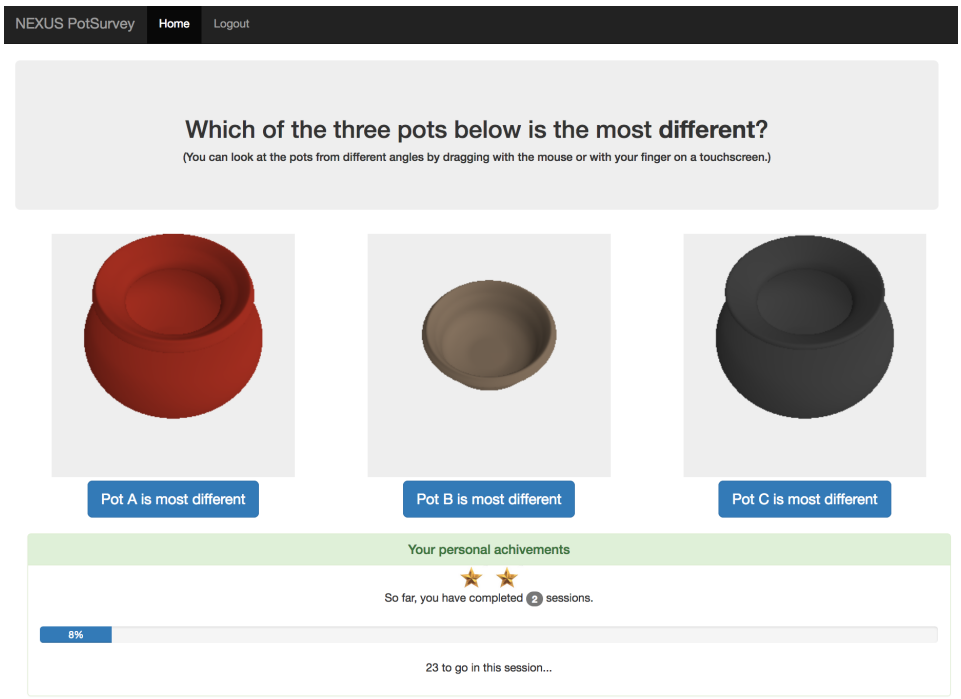


Figure 2.14: A screen shot of the *PotSurvey* web app

We created the web app *PotSurvey* for the test which allowed us to integrate interactive three-dimensional models of the pots into the online survey. After registering for the test[19], the participants could interactively inspect the three pots from the current triad by rotating and zooming them on the canvas in their browser. We asked the typical question for triad tests: *Which of the three pots below is the most different?* The most different pot could be selected by clicking on the button below the visualization. When the participant made a choice, it was stored together with the user's ID and a new triad was randomly selected from the list of 2,925 triads

[19]We required an email address, age, gender, and whether the subject had experience with archaeology and/or pottery

(we also randomized the order of the elements within the triad) and presented to the participant. In order to keep players motivated, we introduced a progress bar, showing how much of the current "set" of 25 triads was complete. Whenever a set was finished, the user got a "star" and a new set began. See Fig. 2.14.

Due to the random selection of triads, some triads were shown to multiple test subjects. This gives us a chance to compare the results and examine more closely the cases where the choices differ. In order to get all triads covered at least once, we occasionally removed triads from the survey which were already covered.

We advertised the survey through mailing lists and the social media channels of the *NEXUS1492* project. After about a month, all 2,925 triads were treated at least once. A total of 116 people participated in the study, 43 of which claimed to have no previous experience, 3 had experience with pottery but not archaeology, 8 had experience with archaeology but not pottery, and 62 stated that they had experience with both archaeology and pottery. The average age of the participants was 40 years.

2.6.2 Network Generation and Analysis

We constructed a weighted network from the results of the tests by increasing the weight of an edge between two pots A and B by 1 if the third pot C from a triad was chosen as the most different one. In the case that the triad was shown to multiple participants and there was no consensus between their answers, we distributed the value added according to the choices of the participants, e.g., if the same triad was shown to two participants and one participant chose pot A as most different and the other pot B as most different, we increased the weight of the edges between B and C as well as A and C by 0.5 each.

A stress minimization layout of the resulting (weighted) network is shown in Fig. 2.15. The result clearly indicates that the main criterion for grouping was the wall profile attribute. We also generated separate networks for the group of *experienced* participants (i.e., the ones who answered that they had experience both with pottery and with archaeology) and the group of *non-experienced* participants (i.e., the ones who answered that they had no experience in pottery or archaeology). Note that these networks are based on a subset of the complete triad test, and thus some pairs might be over- or underrepresented. However, due to the relatively large sample size per group, we can assume these differences are averaged out.

	wall profile	lip shape	color
all participants	0.3348	-0.0148	0.0036
only experienced	0.2769	0.0118	0.0313
only non-experienced	0.378	-0.03661	-0.0294

Table 2.2: Modularity of the attributes

As the drawings of the networks per group are very similar to the network for all the participants, we do not show them here. We can however compare their structural properties. Table 2.2 shows the modularity for the clusters induced by the attributes for each group. The results indicate that the experienced participants put slightly

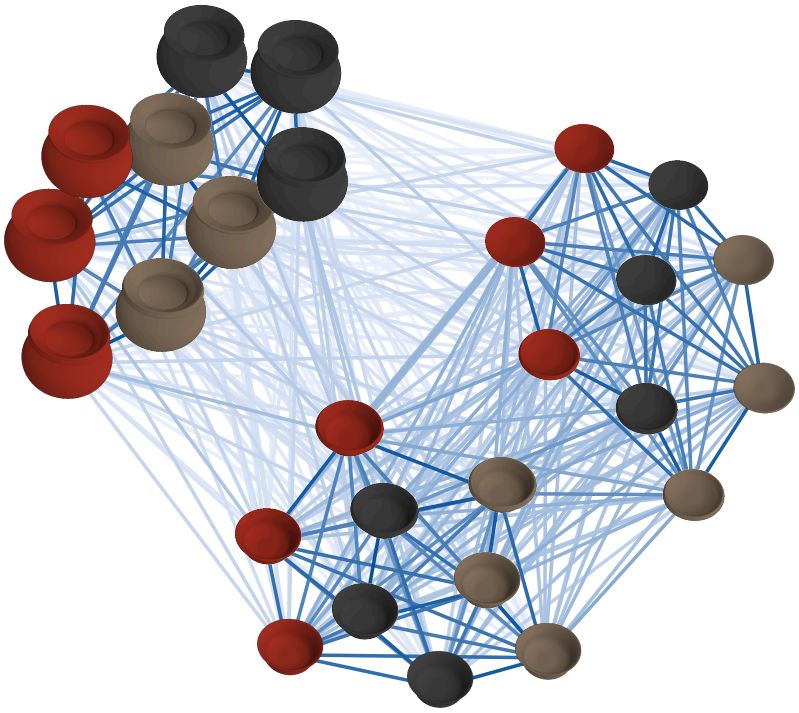
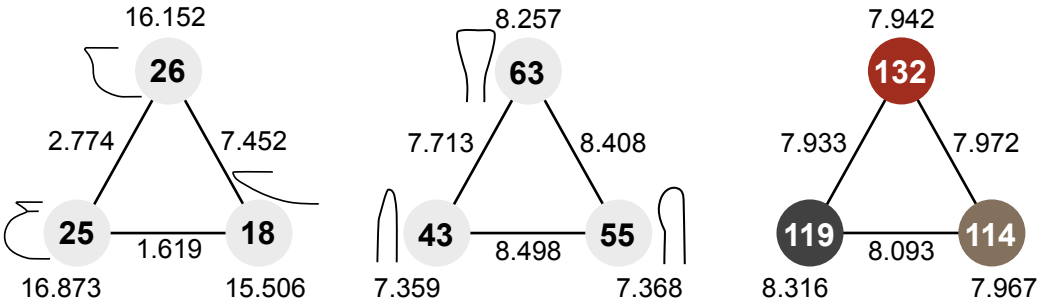


Figure 2.15: Weighted stress minimization layout of the network from the full triad test. Higher weighted edges have darker color

more emphasis on the attributes of lip shape and color than the non-experienced participants. Despite these differences, wall profile remains the most important factor for all groups.

In order to illustrate how each attribute influenced the choice of the participants in the network of all participants, we generated aggregations of the networks based on the different attributes (Fig. 2.16). Each node represents a cluster of all nodes in the network with this value of the attribute. The numbers on the edges show the average strengths of inter-cluster edges, the numbers next to the nodes the average strengths of intra-cluster edges. The aggregations show that – as expected from the drawing in Fig. 2.15 and the modularity scores in Table 2.2 – the average values of the intra-cluster edges are highest for the wall profile attribute (Fig. 2.16a). The value 25 was perceived to be the most different for this attribute. Both lip shape and color were considered much less important, with the average value for the intra-cluster edges for the color attribute only slightly higher than for the lip shapes.

Based on these results, a question that naturally arises is whether we can find a way to give an absolute measure of the strength of a particular attribute. One problem is that there is a dependency of the ties, e.g., if the wall profile attribute was not present, modularity for the other attributes would likely be much higher. Under the assumption that there are no other attributes than the three we incorporated in the test, we propose the following approach to calculate an absolute weighting of the attributes.



(a) Wall profiles **(b)** Lip shapes **(c)** Colors

Figure 2.16: The average strenghts of ties between and within the clusters for each attribute

We first have to establish the origin of the scale, i.e., the minimum possible strength of an attribute. In order to do this, we can calculate the minimum possible value for the intra-cluster strength as follows: In the full set of 2,925 triads, there exist 84 triads that contain only pots within the cluster of 9 pots with the same value for the attribute. Thus in each of these 84 cases, the weight of one of the intra-cluster edges has to be increased. In all other cases, it would be possible to increase the weight of an inter-cluster edge by choosing a possible value within the group as most different. There are 36 edges within the cluster, so if we divide the minimal (aggregated) inter-cluster weight of 84 by 36, we end up with a minimum possible value for the average intra-cluster strength of $\frac{7}{3}$.

Now we can establish an absolute "strength" of the attributes by substracting the minimum possible value from the average intra-cluster strength of the attributes. We end up with the following perceived weights of the participants for the attributes:

- wall profile: 13.844
- lip shape: 5.328
- color: 5.742

2.6.3 Consistency of the Answers

Due to the random assignments of triads to the participants, we were able to record more than one answer for some of the triads. Table 2.3 shows how many triads were treated how many times and how the consistency of the answers is distributed. The variation in the answers indicates that further testing with a greater number of participants would be required to create reliable results that can actually be used for the creation of a classification. Until quite recently it would have been impossible to get a sufficient number of participants for such a survey. This changed with the advent of platforms for citizen science projects such as *Zooniverse*[20] which open up possibilities to conduct online surveys with thousands of volunteers. Such a large scale user study would be an interesting project for future work.

[20] `https://www.zooniverse.org/`

number of treatments	1	2	3	4	5	6	7	8
triads	1,444	755	477	169	64	10	4	2
consistent answer	1,444	517	262	73	26	1	3	0
two different answers	-	238	200	87	35	8	1	2
three different answers	-	-	15	9	3	1	0	0

Table 2.3: Consistency in the answers of the participants for the triad test

2.7 Case Study: Arranging Real Sherds

In the previous sections we demonstrated methods for quantifying the knowledge of archaeologists (and perceptions of the general public) about ceramic attributes. A reconstruction of their mental model can help weighting attributes and their values when it comes to the categorization of sherds.

But how can we find out if the knowledge about pottery from the point of view of an archaeologist corresponds to the meaning that was attributed to the pot when it was made by a potter? In an attempt to quantify differences in perception, we started a pilot study in the US-Southwest. The goal was to measure if there are differences in the interaction with pottery between different groups of people. We were particularly interested in the comparison of (indigenous) potters to archaeologists, and thus we chose the participants for this pilot study from the following groups of people: archaeologists (both ceramic analysts and general archaeologists[21]), traditional technology potters (with both native and non-native background), as well as participants from the general public who do not have particular knowledge about archaeology or pottery.

The objective of this pilot study was to answer the following questions:

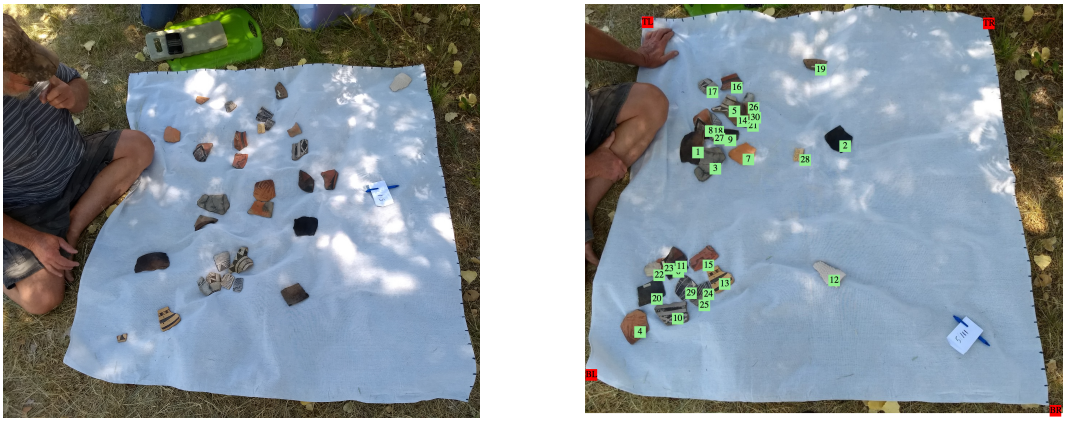
- How (if at all) does the perception of pottery differ within and between different groups of people?
- What are the implications of these differences for archaeological ceramic analysts (i.e., in recording attributes)?
- Can archaeologists use the results to help refine and improve their social interpretations of ceramic data (i.e., examining signals for group affiliation; the importance of low-visibility versus high-visibility attributes between groups)?

This section describes joint work with Lewis Borck (Athenstädt et al. 2018). I focused primarily on the design of the study and the quantitative analysis of the results. Lewis Borck focused on the qualitative analysis and the archaeological background.

2.7.1 Test Design

We presented a preselected set of 30 real sherds from a (mostly) prehistoric context to a number of individuals from different groups of people: archaeologists with or

[21] Archaeologists with a specialization other than Southwest pottery



(a) Participants arrange the sherds on a canvas

(b) Pixel positions of the sherds are tagged and transformed into the unit square

Figure 2.17: The two-dimensional arrangement test carried out in the US-Southwest

without a specialization in ceramics, indigenous potters, non-indigenous traditional technology potters and two people from the general public as a control group. We asked the participants to arrange the sherds on a canvas according to different criteria (Fig. 2.17). For the execution of the test we closely followed a written protocol in order to avoid bias induced by the formulation of the tasks. In the following we briefly describe the tasks for the test, the full protocol can be found in Appendix D.

Tasks

After general questions regarding demographics and familiarity with pottery and archaeology, participants were asked to arrange the sherds on a 5-foot by 5-foot canvas, according to given criteria. After each task, a photograph of the final layout was taken and the participants were asked for a selection of the sherds why they are placed in this way. The sherds were sorted by each participant according to the following directions:

Warmup Example In order to prepare the participants for the upcoming tasks, we started with a set of 11 *Lego-Bricks* of varied shapes and colors. The participants were asked to arrange the Legos such that *the Legos that are more similar to each other are closer together and those that are more different are farther apart.*

Task 1: Undirected two-dimensional arrangement For the first task participants are asked to *arrange the sherds so that the sherds that are most similar to each other are closer together and those that are more different are farther apart.*

Task 2: Two-dimensional arrangement by assumed origin In the second task, participants should *arrange the sherds so that the sherds they think were made by a similar group of people are closer together and sherds from different groups are further apart.*

Task 3: Two-dimensional arrangement by perceived function During the third task, the instructions are to *place sherds from pots with a similar function closer together and sherds from pots for different purposes farther apart.*

Task 4: Ranking by "hardness to make" In the last task, the participants are asked to place the sherds on a line, *ranking them from most difficult to make to easiest to make.*

Choice of Sherds

In order to keep the tasks within a reasonable time frame and to provide enough space on the canvas, we decided to limit the study to a total of 30 different sherds (Fig. 2.18). We tried to select the sherds in such a way that they provide a good sample of the pottery found in the US-Southwest region.

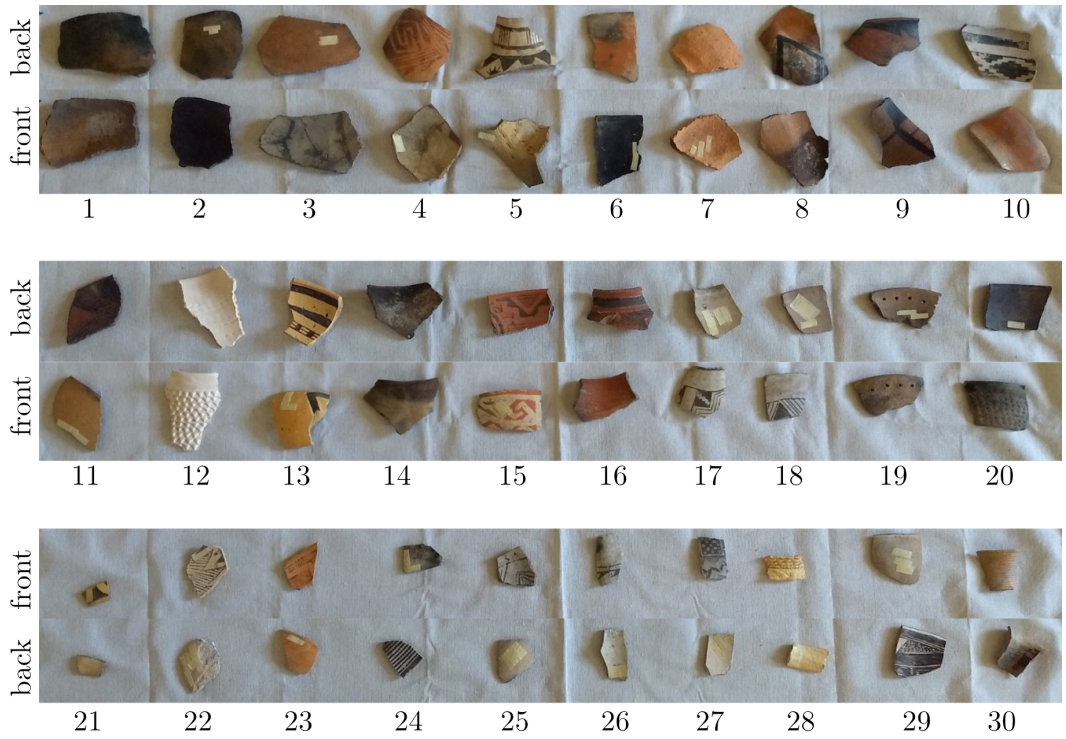


Figure 2.18: The 30 sherds used for this study

While it would have been desirable, it was not possible to obtain sherds of the same size from all types we wanted to incorporate, so there is some variation in the size. However this seems not to have a large influence on the results, as only one of the participants reported size of the sherds to be one of the (minor) factors influencing their arrangement during one task.

A large part of the sherds from the US-Southwest area are decorated, which is reflected in the choice of our sample by incorporating both painted and textured

sherds. We also added a few undecorated sherds of *brown* and *gray ware*. Whenever possible we chose rim sherds, as they allow a better estimate of the vessels shape. In order to limit bias based on taphonomic processes[22], we tried to choose sherds with as little corrosion as possible and made sure that each sherd had a fresh break to allow clues about the material composition and the firing process. The sherds were selected by two experts in Soutwest Archaeology (Lewis Borck and Leslie Aragon) and documented both photographically and according to established ceramic analysis techniques for Southwest pottery. In addition, we recorded the attributes according to the Code Book for Caribbean Ceramics (Section 2.2.1) whenever applicable.

We also included one modern sherd (sherd 12) which is from a pot that was made by an artist from the *Pueblo of Acoma*.

Participants

We chose the participants from five different groups of people: the general public (*gp*), general archaeologists without specialization in US Southwest pottery (*ga*), ceramic analysts with a specialization in the analysis of ceramics from the US Southwest (*ca*), non-indigenous traditional technology potters with a European ancestry (*nip*), and indigenous potters (*ip*). The goal for this pilot study was to administer the tests with at least five people from each group and more if possible. We have not reached this number for all groups yet, but the study is ongoing. So far we interviewed

- 6 indigenous potters (ip)
- 4 non-indigenous potters (nip)
- 5 ceramic analysts (ca)
- 4 general archaeologists (ga)
- 2 people from the general public (gp)

The following results are thus preliminary and will be extended to more subjects for the journal publication of this study.

2.7.2 Results

We analyzed the spatial arrangements in the first three tests by averaging the normalized distances of the sherds for each group of participants. Through plotting an MDS of the average distances for each group, we could get a first idea on general patterns for the groups. See Fig. 2.19 for an example.

We then went on to quantify to what extent each of the attributes of the sherds is reflected in the arrangements. In order to do this, we first transformed the distances into a network of similarities by inverting the values (i.e., the weight of an edge between two sherds is 1 divided by their normalized distance).

[22]The processes of corrosion or decay that occur after the deposition of the sherd

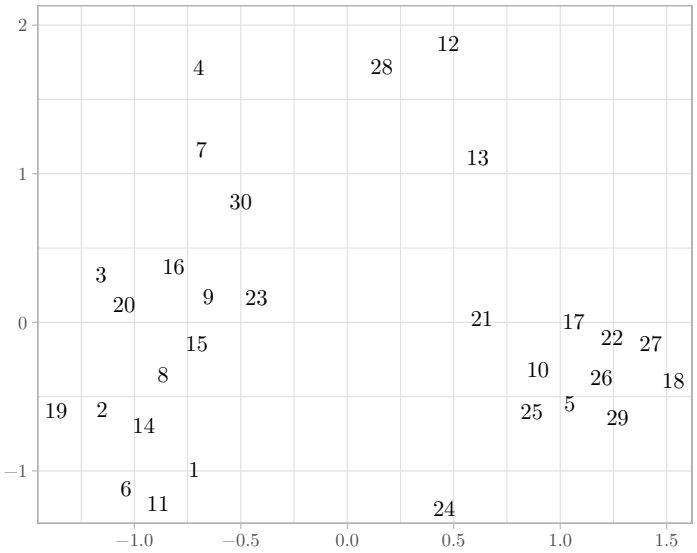
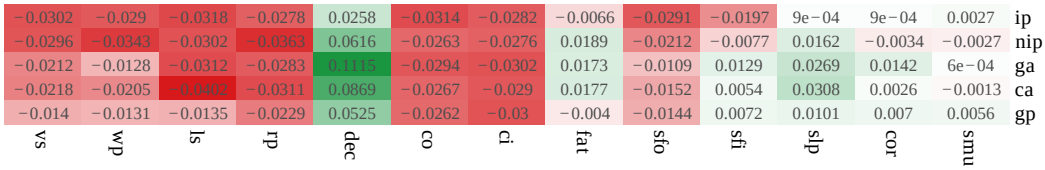


Figure 2.19: MDS of the average distances between the sherds in the recordings of the indigenous potters for Test 1

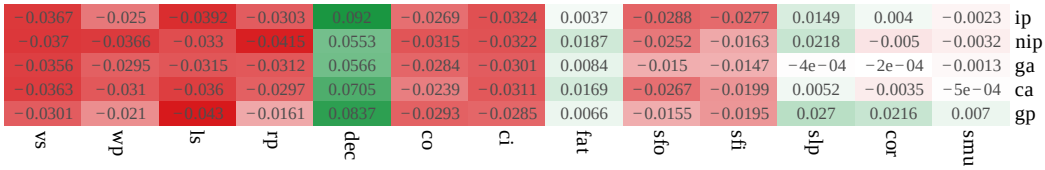
On the resulting networks, we calculated the weighted modularity of the clusters induced by the attributes. Figure 2.20 highlights positive (green) and negative (red) modularity values for selected attributes on the similarity matrices based on the aggregated distances per group. The attributes shown in the figures are the following (see Section 2.2.1 for a detailed description of the attributes):

- vessel shape (vs)
- wall profile (wp)
- lip shape (ls)
- rim profile (rp)
- decoration[23] (dec)
- color outside (co)
- color inside (ci)
- firing atmosphere (fat)
- surface finishing outside (sfo)
- surface finishing inside (sfi)
- slip (slp)
- corrugation (cor)
- smudging (smu)

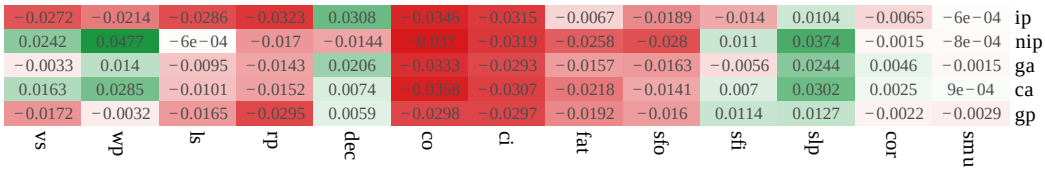
[23]The definitions for decoration are ambiguous for various sites. We classified the sherds according to the Code Book whenever possible and made use of the values for *other* to encode different kinds of decoration that are typical for pottery in the US Southwest



(a) Test 1: undirected arrangement



(b) Test 2: assumed origin



(c) Test 3: assumed function

Figure 2.20: Modularity of the clusterings induced by selected attributes on the inverted distance matrices

Corrugation and smudging are attributes that do not appear in the Code Book for Caribbean Ceramics but are commonly found in pottery from the US-Southwest. Corrugation is a technique in which the coils are pressed together with the finger during the production process, leaving a regular pattern of indentations on the surface (Pierce 1999). Smudging refers to the technique of producing a black film made of carbon or tar on the surface during or after firing (Steponaitis 2009).

Note that the absolute scores of modularity are quite low, which can be explained by the structure of the networks (they are complete networks with the weights resulting from the inverted distances) that makes modularity scores close to the (theoretical) maximum of 1 unlikely (or even impossible). However, the relative scores can be an indicator on how well a clustering is represented by the arrangement.

It turns out that in almost all arrangements, decoration (dec) is the most driving factor with the exception of the assumed function (Test 3) where the groups of non-indigenous potters (nip) and ceramic analysts (ca) considered both wall profile (wp) and the presence of slip (slp) more important.

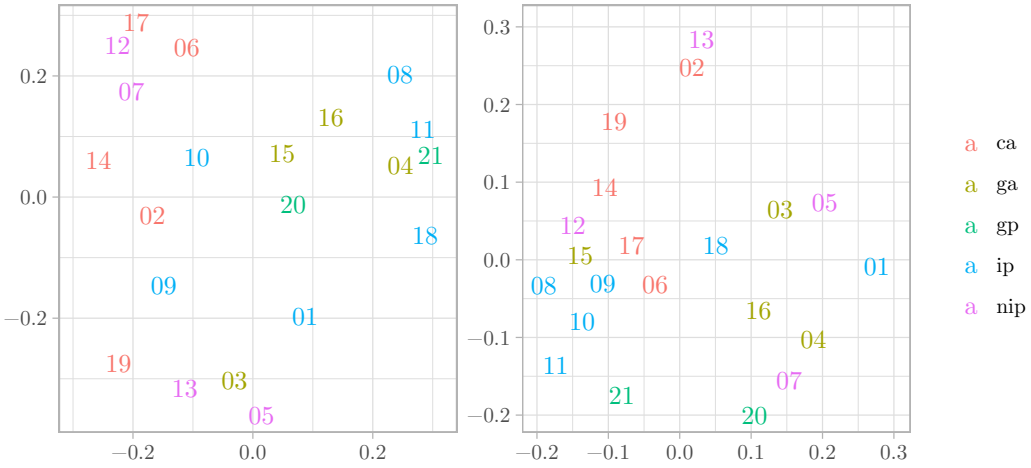
The table for the undirected arrangement in Test 1 (Fig. 2.20a) shows that even though all groups put an emphasis on decoration (dec) in their arrangements, the strengths of this emphasis varies. While the modularity is the highest of all tests and groups for general archaeologists (ga), indigenous potters (ip) seem to put (on average) a lot less emphasis on the criterion.

Interestingly, the table for the arrangement based on the assumed origin in Test 2

(Fig. 2.20b) shows a reverse picture: Here, the indigenous potters (ip) base their judgment more on decoration than the general archaeologists (ga).

In Test 3, asking for an arrangement by function (Fig. 2.20c), the groups of general archaeologists (ga) and indigenous potters (ip), diverging in the previous tests, have similar emphasis in their arrangements, keeping decoration (dec) as an important factor. Non-indigenous potters (nip) seem to agree more with the ceramic analysts as they put their highest emphasis on wall profile (wp) and slip (slp).

In Appendix E we provide additional tables of the modularity for each individual arrangement.



(a) Procrustes distances **(b)** Inverted Mantel correlation

Figure 2.21: MDS of two different distances for the participants for Test 2

In order to compare the individual arrangements of the sherds, we did a pairwise Procrustes analysis of the positions of the sherds in the individual arrangements, resulting in a matrix of the Procrustes distances between the participants' arrangements. In addition, we calculated pairwise Mantel tests on the distance matrices of the participants' arrangements and created a distance matrix of the inverted correlations between the participants by subtracting the Mantel correlation from 1.

Figure 2.21 shows the MDS from the resulting distances for Test 2. Participants are color-coded by the group they belong to. The quite significant differences between the two MDS representations can be explained by different properties of the distance measures. The most important difference being that the Procrustes analysis does not allow a mirroring of the arrangement: Two arrangements that are just flipped on one axis would result in the same distance matrix and thus have a distance of 0 on the inverted Mantel correlations. Yet, they would have a rather large Procrustes distance since they can not be mapped onto each other by just translation, scaling, and rotation.

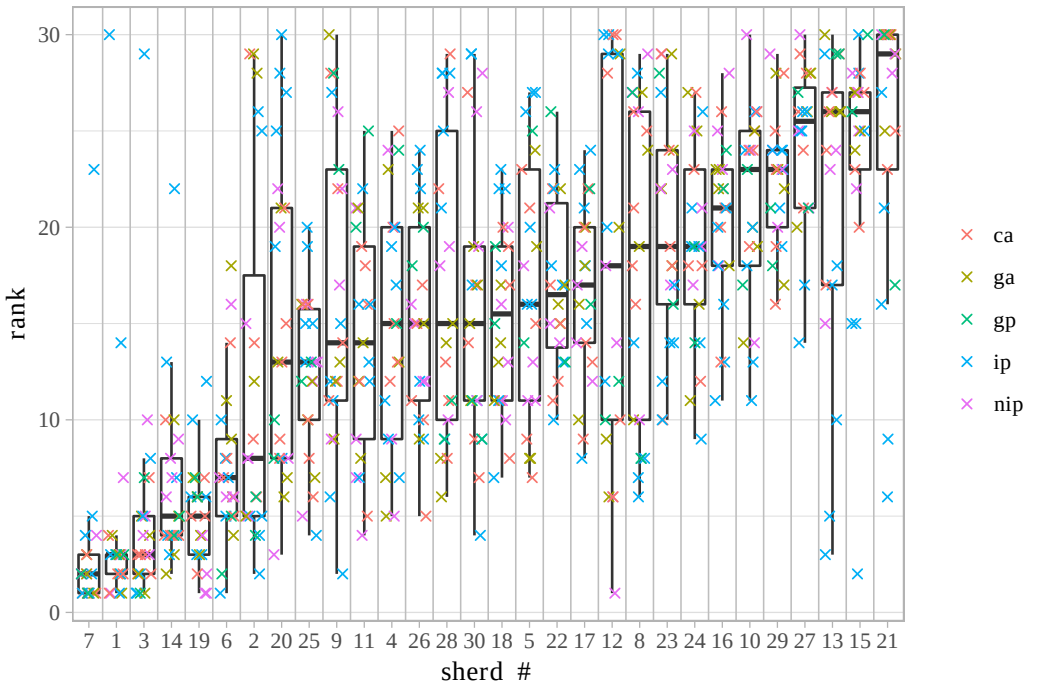


Figure 2.22: Boxplots of the participants' rankings of the sherds in Test 4, ordered by the mean value, overlaid by a scatter plot of the individual ranks colored by group

Test 4 - Ranking of the Sherds

The boxplot in Figure 2.22 shows the variability of the rankings for the different sherds ordered by the median value. The higher the ranking the more difficult it was to make the sherd according to the participants. While there is a relatively strong agreement about the sherds that were easiest to make (the undecorated sherds 1, 3, and 7), opinions diverge more when it comes to the sherds whose median ranks are between 10 and 20. The highest variation in the answers was observed for sherd 12, which is the modern sherd from the Acoma-potter. While considered "hard to make" by some of the participants, others rejected the sherd as being "fake" (yet not necessarily badly made). Other sherds for which opinions diverged include sherd 2 (which is very thin but not decorated), sherd 20 and 28 (both are corrugated), and sherd 8 and 9 (they are both polychrome).

The overlay of the scatter plot showing the individual rankings color coded by groups hints at an interesting pattern: Almost all outliers of the top-ranked sherds as well as the lowest-ranked sherds come from participants within the group of indigenous potters (ip).

In order to examine this trend more closely, we established distances between the participants based on their rank correlation. We calculated Kendall's τ on the rankings and generated a distance matrix based on the inverted values (i.e., $1 - \tau$). Figure 2.23 shows an MDS of the inverted correlations. Clearly the indigenous potters are spread out much more than the rest of the groups, potentially indicating a higher diversity in the notion of "hardness to make".

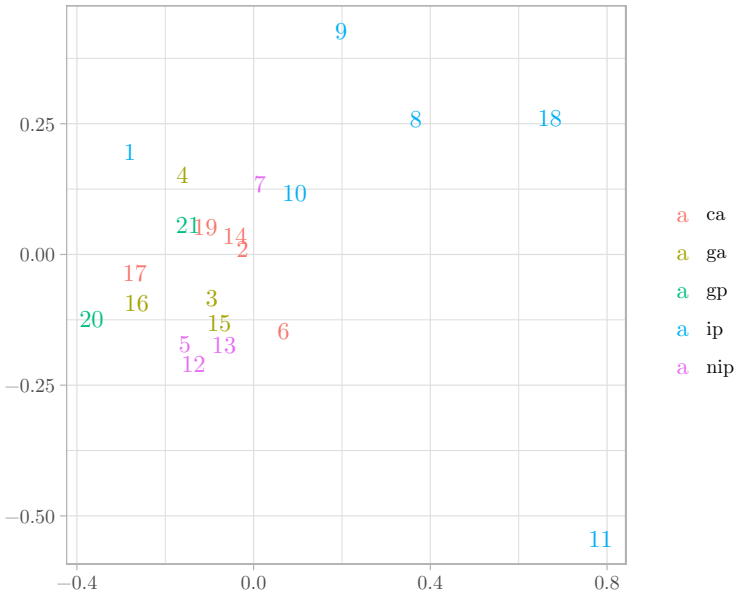


Figure 2.23: MDS of the inverted rank corellation calculated with Kendall's τ for Test 4

2.7.3 Outlook

Even though the sample size in our study is too small to make reliable statements about the differences in perception between different groups, the results indicate that such differences might exist and are measurable with the methods we propose.

Yet, in order to get quantifiable data, in particular when the results will be incorporated into classification systems, the number of participants would need to be greatly increased. Unlike the online studies in the previous section which could be scaled quite easily, the number of potential participants is limited, and the amount of time required to administer a test is substantial.

However, we think that the results of this study are promising enough to extend the survey. The methods presented here could also provide valuable insights in other areas of anthropology, where triad tests would prove too cumbersome and time intensive for participants.

2.8 Conclusion

In this section, we proposed ways to incorporate human perception and knowledge into the classification of sherds with the goal to create typologies that reflect the mental image of (indigenous) potters or ceramic analysts.

We started the section by giving an overview of classification systems for (Caribbean) pottery and described the attributes recorded in the pottery database used at Leiden University, giving a motivation for establishing a classification system that puts more emphasis on human perception. We went on by presenting a variety of existing

methods from cultural domain analysis and proposed two new methods to quantify knowledge and perception: two-dimensional arrangements and multidimensional distance recording (MDR).

In a two-dimensional arrangement test, the participant is asked to arrange the items on a surface such that the distance between the items reflects the (dis)similarity they perceived. In an MDR, the participant proposes a set of relevant dimensions for describing the items and organizes them on a nominal, ordinal, or interval scale for each dimension. The Gower distance can then be used to calculate a distance matrix between the items.

The proposed methods can reduce the time required to administer perception studies, rendering the recording of distances between attributes and between their values more efficiently.

We put the methods to the test by recording perceived distance (and similarity) matrices between the possible values of an attribute (Section 2.5), between attributes (Section 2.6), and between real objects (Section 2.7).

The comparison of the MDR method to the established method of triad tests in the first study (Section 2.5) indicates that the MDR method has the potential to generate similar results to the (much more time consuming) triad tests as long as a sufficient number of dimensions can be determined for arranging the items. The drawback of the method is that it requires a deeper understanding of the backgrounds of classification from the participant.

In the second study (Section 2.6), we used triad tests in an online survey to measure the perceived importance of attributes. We could show that triad tests are a valuable method for recording perceptions when applied to our data set. It has the potential to generate an absolute weighting of the attributes from the Code Book of Caribbean Ceramics. Unfortunately, an application of the method to a larger set of attributes would require a significant amount of effort and a high number of participants.

The study with real sherds in the US-Southwest (Section 2.7) gave us a chance to evaluate the two-dimensional arrangement method. We were able to record the participants' perceived distances between 30 different sherds, a number that would have been infeasible with triad tests. Despite the limitations imposed by the two-dimensional area, the recorded distances can indicate the most relevant attributes in the arrangements and highlight differences between the groups.

The goal of all three studies was the evaluation of the methods and not the generation of a reliable data set. If the data were to be used in a "production ready" classification system, the studies would need to be extended to more participants and more care would be required in controlling for external factors that might have influenced the participants.

The results of the case studies show that our methods have the potential to help establishing a new way of classifying pottery that (while still based on "hard" criteria such as the ones in the Code Book) incorporates human perception. If the knowledge of (indigenous) potters is incorporated, it can even provide a way to introduce the emic knowledge of the people who produce(d) the objects into the – otherwise

inherently etic – way for classification of material by archaeologists.

Such a classification would not only provide an insight into the development of styles from a more emic perspective. It could also help with the goals of reconstructing connections between sites. In addition to the travel distances we calculated in Part 1, similarities of site assemblages could provide another way to infer past relationships. We provide an overview of similarity measures between vectors of pottery types in an article in the *Journal of Archaeological Science* (Habiba et al. 2018).

In the long term, a step towards making the analysis more objective would be a system based on a database with 3D-scanned sherds. This could include automatic categorization based on the Code Book and a classification based on perception studies over the full range of attributes and for all sites excavated in the Caribbean.

Part 3

Overlapping Clustered Planarity and Unions of Partitions

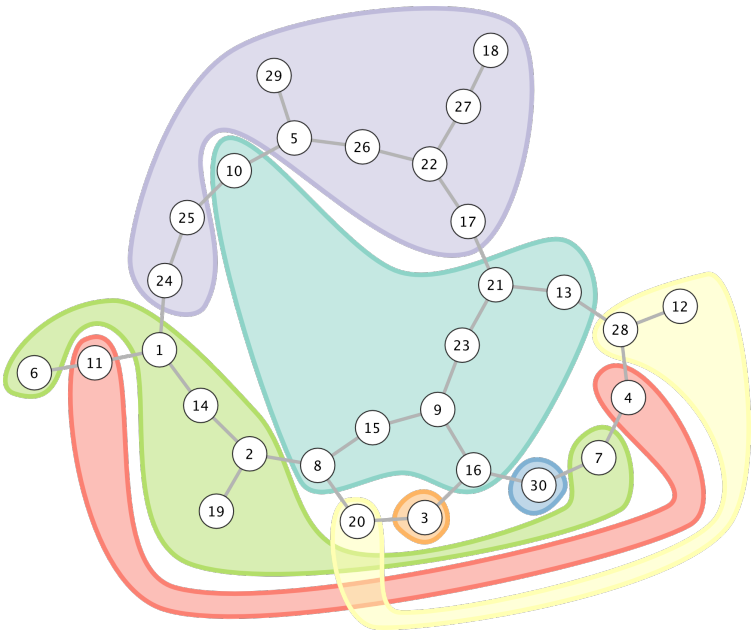
In this part of the thesis, we will focus on the theoretical aspects of visualization methods for clustered networks. The embeddings we create have the potential to improve readability of diagrams, not only for archaeological data, but also in other domains where the task is to visualize network data with additional (nominal) attributes.

We are considering the problem of overlapping clustered planarity: Given a graph and a set of clusters, i.e., subsets of the vertex set, the question is whether there exists a planar embedding of the graph in which the clusters are drawn as simple closed regions such that

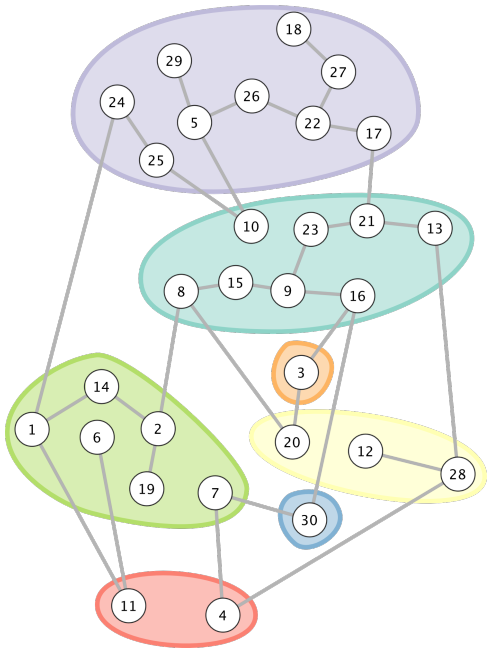
- each region contains exactly the vertices from the corresponding cluster,
- an edge may cross the boundary of a region at most once,
- if a cluster is a subset of another cluster, its corresponding region is a subset of the corresponding region of the other cluster, and
- each connected region resulting from the intersection of two regions must contain at least one vertex.

These requirements minimize the number of crossings and should therefore improve the readability of the resulting drawing.

In order to demonstrate the problem in an archaeological context, we generated the *relative neighborhood graph (RNG)* of the MDS layout of the sherd arrangements by the indigenous potters from Fig. 2.19 in Part 2. The RNG is a planar graph that can be calculated from a set of points in the plane by creating an edge between two points p and q if there is no point r that is closer to p or q than they are to each other. In order to keep the drawing readable, we chose a non-overlapping clustering induced by the decoration attribute of the sherds. Figure 3.1 shows two different embeddings of the RNG, demonstrating the properties we want to avoid: The embedding in Fig. 3.1a has an intersection of the yellow region with the red and green region, respectively, that do not contain a common vertex. In the embedding in



(a) Planar embedding with two intersections of cluster regions (yellow-green and yellow-red) that do not contain a common vertex



(b) Planar embedding with two intersections of the edge $(16, 30)$ with the yellow cluster

Figure 3.1: Two embeddings of the relative neighborhood graph of the MDS layout of the indigenous potters for Test 1 (Fig. 2.19) clustered by the decoration of the sherds

Fig. 3.1b, the edge (16,30) crosses the boundary of the yellow cluster twice. The goal is to find a planar embedding of the graph and its clusters without such intersections and crossings.

Unfortunately, the decision problem whether such an embedding exists is NP-hard in the general case (Didimo et al. 2008). Our contribution is to examine under which conditions the problem becomes hard and for which types of instances we can find a solution in polynomial or even linear time.

In Section 3.1, we give a formal description of the problem. We then examine related work in Section 3.2 and explain the data structures and techniques required for the following proofs.

In the subsequent sections, we examine the special case that the clusters are the union of two partitions. We show that the problem remains NP-hard under these conditions (Section 3.3), even if the underlying graph is 2-connected, each cluster contains at most two connected components, and the complement of each cluster contains at most three connected components. However, for the even more restricted case that each cluster, as well as its complement, induces a connected subgraph, we can provide a linear-time algorithm to check overlapping clustered planarity (Section 3.4).

We will then show that the problem can be solved in polynomial time for arbitrary overlapping clusterings (that are not necessarily the union of two partitions) if each cluster induces a connected subgraph. We first provide a method to solve the problem for 2-connected graphs (Section 3.5) and then extend the method to arbitrary graphs (Section 3.6).

The majority of this part of the thesis is joint work with Sabine Cornelsen and has been published in the Journal of Graph Algorithms and Applications *(Athenstädt and Cornelsen 2017). Section 3.3 is based on joint work with Tanja Hartmann and Martin Nöllenburg and has been presented at the* 22nd International Symposium on Graph Drawing *(Athenstädt et al. 2014).*

3.1 Introduction

An (undirected) *graph* $G = (V, E)$ consists of a set of *vertices* V and *edges* E where an edge $e \in E$ is a pair of vertices $e = \{u, v\}, u, v \in V$, connecting u and v. Vertices u and v are *incident* to edge $e = \{u, v\}$ and vice versa. Two vertices are *adjacent* if they are connected by an edge, and two edges are adjacent if they share a common vertex. A *path* from v_1 to v_n in G is an alternating sequence $v_1, e_1, v_2, \ldots, v_{n-1}, e_{n-1}, v_n$ of edges and vertices in G where $e_i = \{v_i, v_{i+1}\}$ and the vertices $v_2 \ldots v_n$ are distinct. The *length* of a path is the number of edges in the path. If $v_1 = v_n$, the path is a (simple) *cycle*. A graph is connected if there exists a path between any two vertices in V.

A graph is *simple* if it does not have any parallel (duplicate) edges or circles of length one (loops). Unless explicitly stated, we assume all graphs in this work to be simple.

An *(overlapping) clustered graph* $(G = (V, E), \mathcal{C})$ consists of an undirected graph G and a set $\mathcal{C}$ of subsets of the vertex set V. The elements of $\mathcal{C}$ are called *clusters*.

A vertex may be contained in several clusters. Moreover, clusters may overlap, i.e., there might be $C_1, C_2 \in \mathcal{C}$ with $C_1 \cap C_2 \neq \emptyset$, $C_1 \not\subseteq C_2$, and $C_2 \not\subseteq C_1$. Observe that the abstract Hasse diagram (Didimo et al. 2008) of $\mathcal{C} \cup \{\{v\}, v \in V\} \cup \{V\}$ is a succinct representation of an overlapping clustering $\mathcal{C}$. Actually, Didimo et al. (2008) define overlapping clustered graphs via the Hasse diagram.

Didimo et al. (2008) defined planarity for overlapping clustered graphs geometrically: An overlapping clustered graph $(G = (V, E), \mathcal{C})$ is *clustered planar* if the vertices can be represented by distinct points[1], each edge $e \in E$ by a curve $R(e)$ between its incident vertices, and each cluster $C \in \mathcal{C}$ by a simple closed region $R(C)$ in the plane such that for $X, Y \in E \cup \mathcal{C}$ we have that

(i) $X \subset R(X)$ and $(V \setminus X) \cap R(X) = \emptyset$,

(ii) $X \subseteq Y \Rightarrow R(X) \subseteq R(Y)$, and

(iii) every connected region of $R(X) \cap R(Y)$ contains a vertex.

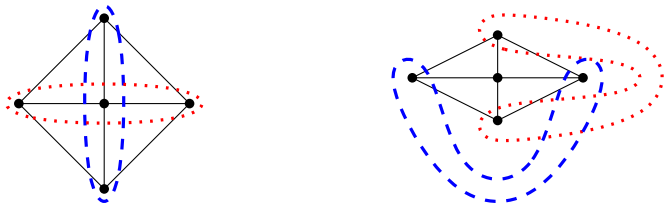
Recall that edges of an undirected graph are subsets of the vertex-set of size two.

Condition (i) means that edges may not contain non-incident vertices while clusters must enclose the vertices they contain and no others.

Condition (ii) implies that cluster boundaries may not intersect if one cluster is contained in the other and that intra-cluster edges must be routed within the cluster.

Condition (iii) ensures several things: (a) two edges intersect at most in a common incident vertex, (b) an edge crosses a cluster boundary at most once, (c) if two clusters do not share a vertex, their regions must be disjoint, and (d) the regions of clusters with common vertices might intersect in several connected components, yet each of them must contain at least one vertex.

Thus, the clustered graph in Fig. 3.2a is clustered planar while the clustered graph in Fig. 3.2b is not.



(a) Clustered planar **(b)** Not clustered planar

Figure 3.2: Two graphs with two clusters each (set of vertices enclosed by red dotted and blue dashed curve, respectively)

Clustered planarity is NP-complete in general as shown by Johnson and Pollak (1987), where the case with $E = \emptyset$ is examined. Didimo et al. (2008) showed that

[1]For the sake of simplicity, we identify each vertex with its image in the plane

clustered planarity can be solved in polynomial time for the special case where each cluster overlaps with at most one other cluster, each cluster as well as the intersection of any two clusters induce connected subgraphs, and some additional connectivity properties hold. They posed it as an open question whether clustered planarity is polynomial-time solvable for general overlapping clustered graphs under the single condition that each cluster induces a connected subgraph. We will answer this question in the affirmative.

If the clustering is hierarchical, i.e., if any two clusters in $\mathcal{C}$ are either disjoint or one is contained in the other, then clustered planarity is the classical problem of c-planarity as considered by Feng et al. (1995). One of the most important open problems in the field of graph drawing is the complexity of c-planarity of hierarchically clustered graphs. An overview on the classical c-planarity problem can be found in (Bläsius and Rutter 2016; Cortese et al. 2008; Maurizio Patrignani 2014). Dahlhaus (1998) and later Cortese et al. (2008) showed that c-planarity of hierarchically clustered graphs can be solved in linear time if each cluster induces a connected subgraph. Their approaches make use of the decomposition of the graph into 3-connected components as represented by BC- and SPQR-trees.

Angelini et al. (2015) defined drawings with region-region crossings of hierarchically clustered graphs. These are essentially representations by points and regions such that all conditions of clustered planarity are fulfilled except for Condition (iii) when X and Y are both clusters. As an example, Fig. 3.2b shows a drawing of a clustered graph with one region-region crossing. Observe that the intersection of the two regions does not contain a vertex as would be required by the definition of clustered planarity. Angelini et al. (2015) showed how to use SPQR-trees to test in polynomial time whether any hierarchically clustered graph with an underlying 2-connected graph has a drawing with region-region crossings.

If $E = \emptyset$, then clustered planarity is closely related to the NP-complete problem of hypergraph (vertex) planarity as defined by Johnson and Pollak (1987): Given a set $\mathcal{C}$ of subsets of a set V, is there a planar support, i.e., a planar graph $G = (V, E)$ such that each set in $\mathcal{C}$ induces a connected subgraph of G? Various subclasses of planar supports that directly imply clustered planarity – such as trees, cacti, and outerplanar supports – were considered (van Bevern et al. 2016; Bixby and Wagner 1988; Brandes et al. 2011, 2012; Buchin et al. 2011; Kaufmann et al. 2009). Hypergraph planarity remains NP-complete even if $\mathcal{C}$ is the union of two partitions (Athenstädt et al. 2014).

3.2 Preliminaries

For a subset $C \subseteq V$ of the vertices of an undirected graph $G = (V, E)$, we denote by $G[C]$ the subgraph of G induced by C, i.e. the graph with vertex set C and edge set $\{e \in E;\ e \subseteq C\}$. A C-path (C-cycle) is a path (simple cycle) in $G[C]$. For two subsets $S, T \subseteq V$, let $E(S, T)$ be the set of edges with one end vertex in S and the other end vertex in T. A *partition* of V is a set $\mathcal{P}$ of subsets of V such that each vertex in V is contained in exactly one set in $\mathcal{P}$. We call such a subset *block* in $\mathcal{P}$.

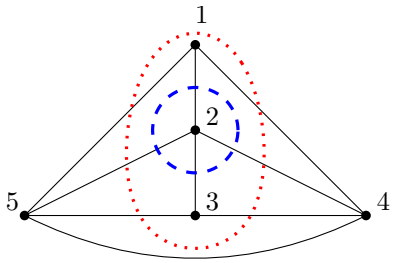


Figure 3.3: The blue (dashed) and red (dotted) lines induce two partitions $\mathcal{P}_B = \{\{2\},\{1,3,4,5\}\}$ and $\mathcal{P}_R = \{\{1,2,3\},\{4,5\}\}$. This leads to an intersection partition $\mathcal{P}_I = \{\{2\},\{1,3\},\{4,5\}\}$ and a connected intersection partition $\mathcal{P}'_I = \{\{1\},\{2\},\{3\},\{4,5\}\}$

For two partitions $\mathcal{P}_B = \{B_1,\dots,B_{\ell_B}\}$ and $\mathcal{P}_R = \{R_1,\dots,R_{\ell_R}\}$ of V, we define the *intersection partition* $\mathcal{P}_I = \{B_i \cap R_j; i = 1,\dots,\ell_B, j = 1,\dots,\ell_R\}$. (This is also known as *coarsest common refinement*.) The *connected intersection partition* $\mathcal{P}'_I$ of $\mathcal{P}_B$ and $\mathcal{P}_R$ is the partition induced by the connected components of $G[C], C \in \mathcal{P}_I$, i.e., $X \subseteq V$ is contained in $\mathcal{P}'_I$ if and only if X is the set of vertices of a connected component of one of the graphs $G[C], C \in \mathcal{P}_I$. See Fig. 3.3 for an example.

A *consecutive-ones ordering* of a binary matrix is a permutation of its columns such that in each row all of the 1s are consecutive, i.e., such that each row is of the form $0^*1^*0^*$. A binary matrix has the *consecutive-ones property* if and only if it has a consecutive-ones ordering. It can be tested in linear time whether a binary matrix has the consecutive-ones property, and a consecutive-ones ordering can be found in linear time if it exists (Booth and Lueker 1976).

3.2.1 Planarity of Overlapping Clustered Graphs

Let $(G = (V,E),\mathcal{C})$ be an overlapping clustered graph. The clustered graph $(G,\mathcal{C})$ is *c-connected* if $G[C]$ is connected for all $C \in \mathcal{C}$ and *c-co-connected* if both, $G[C]$ and $G[V \setminus C]$, are connected for all $C \in \mathcal{C}$.

If $(G,\mathcal{C})$ is c-connected, then a *c-planar embedding* of G for $\mathcal{C}$ is a planar embedding of G such that $V \setminus C$ is in the outer face of $G[C]$ for all $C \in \mathcal{C}$. A graph $G^+ = (V,E^+)$ is a *c-planar support* of a clustered graph $(G = (V,E),\mathcal{C})$ if $E \subseteq E^+$, $(G^+,\mathcal{C})$ is c-connected and there is a c-planar embedding of G^+ for $\mathcal{C}$.

It was shown that a c-connected overlapping clustered graph (Didimo et al. 2008) or a hierarchically clustered graph (Feng et al. 1995), respectively, is clustered planar in the sense of Didimo et al. (2008) if and only if it has a c-planar support.

In this part of the thesis, we define any clustered graph to be *c-planar* if and only if it has a c-planar support.

3.2.2 NP-complete Problems

In computational complexity, the class NP (*nondeterministic polynomial*) describes decision problems for which there exists a polynomial-time algorithm to verify whether a given solution (certificate) is correct. The class of NP-complete problems consists of problems that are "at least as hard" as all problems in NP in the sense that if there were a polynomial-time algorithm to solve an NP-complete problem, all other problems in NP could be solved in polynomial time as well.[2] Since Cook (1971) showed the problem SAT to be NP-complete, many more problems have been shown to belong to this class by a reduction of another NP-complete problem. A reduction is a polynomial time algorithm that transforms an instance x of an NP-complete decision problem A to an instance y of another decision problem B such that the output is the same for all instances. If B can be shown to be in NP, the reduction implies the NP-completeness of B.

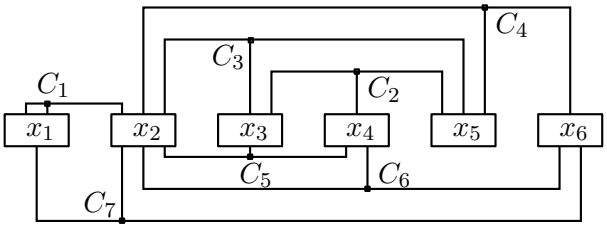
We will now give a brief overview of the chain of reductions of NP-complete problems leading to the NP-completeness of PLANAR-MONOTONE-3-SAT. In Section 3.3 we will then reduce PLANAR-MONOTONE-3-SAT to one of our embeddability problems, showing it's NP-completeness.

3-SAT An instance of 3-SAT consists of a set $\mathcal{U} = \{x_1, \ldots, x_n\}$ of n Boolean variables and a formula $\varphi = C_1 \wedge C_2 \wedge \ldots \wedge C_m$ in *conjunctive normal form* defined over $\mathcal{U}$. Each clause C_i, $1 < i < m$ is a disjunction of not more than three *literals* from $\mathcal{U}$. A literal is a single variable (*positive literal*) or its negation (*negative literal*). 3-SAT is one of "Karp's 21 NP-complete problems" (Karp 1972). It is thus NP-complete to decide whether there exists an assignment to the variables in $\mathcal{U}$ that fulfills all clauses in φ.

MONOTONE-3-SAT MONOTONE-3-SAT is a restriction of 3-SAT where all clauses in φ are *monotone*. A clause is monotone if it consists either of only positive or of only negative literals. Garey and Johnson (1979) showed that MONOTONE-3-SAT is NP-complete.

PLANAR-3-SAT An instance of PLANAR-3-SAT is an instance of 3-SAT whose *variable-clause graph* is planar. The variable-clause graph of a formula $\mathcal{C}$ in conjunctive normal form is defined as $\mathcal{G} = (\{C_1 \ldots C_m\} \cup \{x_1 \ldots x_n\}, \mathcal{E})$ where $\mathcal{E} = \{(x_i, C_j) \mid x_i \in C_j \vee \overline{x_i} \in C_j\}$. Lichtenstein (1982) proved that PLANAR-3-SAT is NP-complete as well. Knuth and Raghunathan (1992) showed that the problem stays NP-complete if it is limited to problems that have a *rectilinear representation*, i.e., the nodes corresponding to the variables are ordered, aligned horizontally and connected to their neighbors in the ordering. Clauses are located either above or below the line of variables and connected by rectangular edges without crossings. The graph in Fig. 3.4 is an example of a rectilinear representation.

[2] Whether such an algorithm exists (P = NP) or not (P ≠ NP) is one of the most famous open problems in theoretical computer science



$$C_1 = (x_1 \vee x_1 \vee x_2)$$
$$C_2 = (x_3 \vee x_4 \vee x_5)$$
$$C_3 = (x_2 \vee x_3 \vee x_5)$$
$$C_4 = (x_2 \vee x_5 \vee x_6)$$
$$C_5 = (\overline{x_2} \vee \overline{x_3} \vee \overline{x_4})$$
$$C_6 = (\overline{x_2} \vee \overline{x_4} \vee \overline{x_6})$$
$$C_7 = (\overline{x_1} \vee \overline{x_2} \vee \overline{x_6})$$

Figure 3.4: Monotone rectilinear representation of a formula $\varphi = C_1 \wedge C_2 \wedge \ldots \wedge C_7$

PLANAR-MONOTONE-3-SAT de Berg and Khosravi (2010) showed that Planar-3-Sat stays NP-complete if it is restricted to instances with only monotone clauses that have a *monotone rectilinear representation (MRR)* where all *negative clauses* (containing negative literals) are embedded on one side of the variables and all *positive clauses* (containing only positive literals) are embedded on the other side of the variables. This problem is called Planar-Monotone-3-Sat. The example in Fig. 3.4 together with the adjoining clauses forms an instance of Planar-Monotone-3-Sat: All clauses above the variables are positive and all clauses below the variables are negative. Note that clause C_1 consists only of two distinct variables. The double-occurrence of variables in clauses is not excluded in the definition of Planar-Monotone-3-Sat, and we will thus take this case into account in our proof in Section 3.3.

3.2.3 BC-Trees

A vertex v is a *cut vertex* of a connected graph G if the graph that results from G by deleting v and its incident edges is not connected. A connected graph is *2-connected* if it contains more than two vertices but no cut vertices. The *blocks* of a connected graph are the maximally 2-connected subgraphs and the subgraphs induced by bridges. The vertices of the *block–cut tree (BC-tree)* of a graph G are the blocks and the cut vertices of G. There is an edge in the block–cut tree between a block H and a cut vertex v if v is contained in H. See Fig. 3.21 for an illustration.

3.2.4 SPQR-Trees

A way of representing all possible embeddings of a 2-connected subgraph are SPQR-trees. Two vertices v and w are a *separation pair* of a 2-connected graph G if the graph that results from G by deleting v and w and their incident edges is not connected. A graph is *3-connected* if it contains more than three vertices but no separation pair. An *SPQR-tree* (Di Battista and Tamassia 1990) is a labeled tree that represents the decomposition of a 2-connected graph into 3-connected components. Each node ν of an SPQR-tree is labeled with a multi-graph skel(ν) – called the *skeleton* of ν. There are four different types of labels associated with the skeletons: *S-nodes* for simple cycles, *P-nodes* for three or more parallel edges, *R-nodes* for a simple 3-connected graph, and *Q-nodes* for two parallel edges.

The Q-nodes are the leaves of an SPQR-tree. Neither two S-nodes, nor two P-nodes

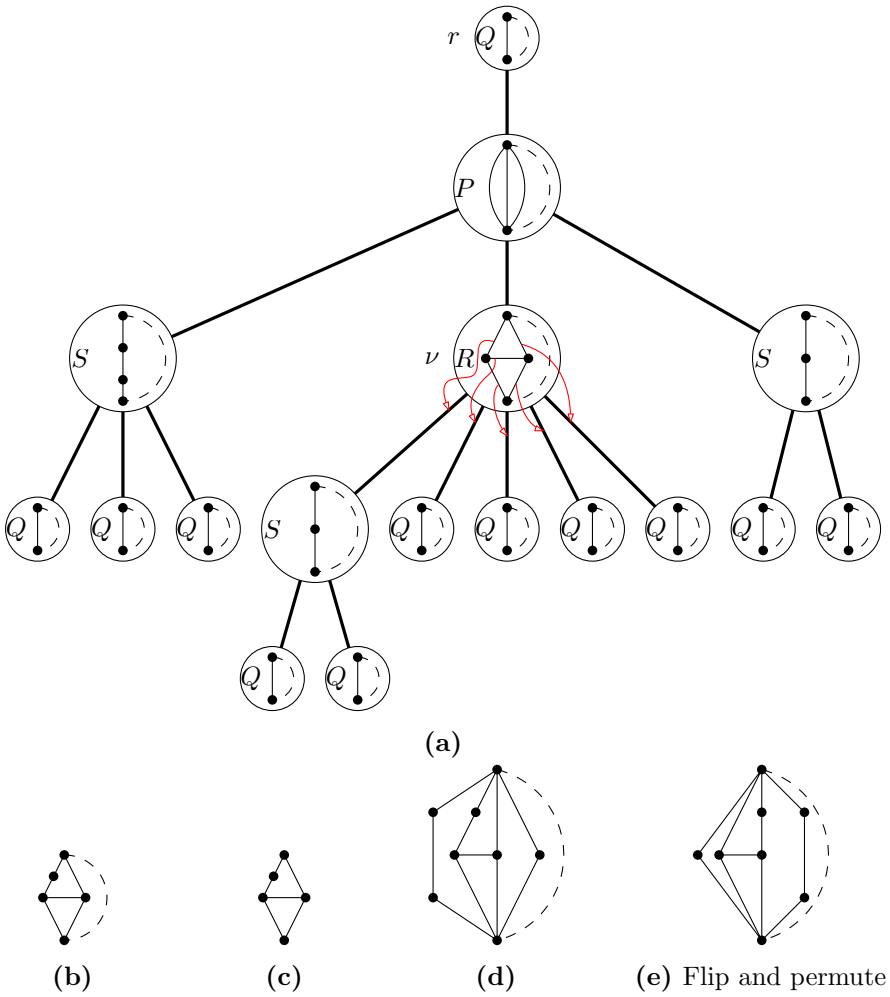


Figure 3.5: An SPQR-tree T rooted at a Q-node r and its represented graph $G_r(r)$. The root edge of each skeleton is dashed. The correspondence of the remaining edges of the skeleton of a node μ and the edges around μ is as follows. If μ is a P-node, we assume that the edges around μ and the edges around the topmost pole of skel(μ) are in the same clockwise order. If μ is an S-node, then we assume that the edges in skel(μ) from top to bottom correspond to the edges incident to μ from left to right. Finally, if μ is an R-node, we explicitly indicate the corresponding tree edges by little red arrows
(e) A different embedding of $G_r(r)$ with the root edge on the outer face can be obtained by flipping the R-node and permuting the children of the P-node

are adjacent in an SPQR-tree. For each node ν of an SPQR-tree there is a one-to-one correspondence of the edges of skel(ν) and the edges incident to ν (except for the Q-nodes where one of the two edges of the skeleton corresponds to the only incident edge of the Q-node). The edge of skel(ν) corresponding to the edge $\{\nu, \mu\}$ of the SPQR-tree is denoted by e_μ. We consider the edges of the skeletons oriented. For simplicity, we assume that the edges of the skeleton of an S-node are oriented as a directed cycle and the edges of the skeleton of a P-node are all oriented in parallel.

We consider the SPQR-tree T rooted at a Q-node r. Let ν be a node of T. The *root edge* of skel(ν) is the edge that corresponds to the parent edge of ν. The *poles* of skel(ν) (or node ν, respectively) are the end vertices of the root edge. Let skel$^-(\nu)$ be the skeleton of ν without the root edge.

Each node ν of the rooted SPQR-tree represents a (multi-)graph $G_r(\nu)$: The Q-nodes (excluding the root) represent a graph with two vertices connected by an edge and additionally by the root edge. Let ν be a non-leaf node of an SPQR-tree and let $\nu_1, \ldots, \nu_k$ be the children of ν. For $i = 1, \ldots, k$, remove the edge associated with $\{\nu, \nu_i\}$ from both skel(ν) and $G_r(\nu_i)$. Insert the remaining parts of $G_r(\nu_i)$ into skel(ν) identifying the poles of $G_r(\nu_i)$ with their counter parts in skel(ν). The *poles* of $G_r(\nu)$ are the poles of ν. Let $G_r^-(\nu)$ be $G_r(\nu)$ without the root edge of skel(ν). (Some papers refer to $G_r^-(\nu)$ as the *pertinent graph.*) $G_r(r)$ is the graph represented by the SPQR-tree T. The edges of $G_r(r)$ correspond to the Q-nodes of the SPQR-tree. See Fig. 3.5 for an example.

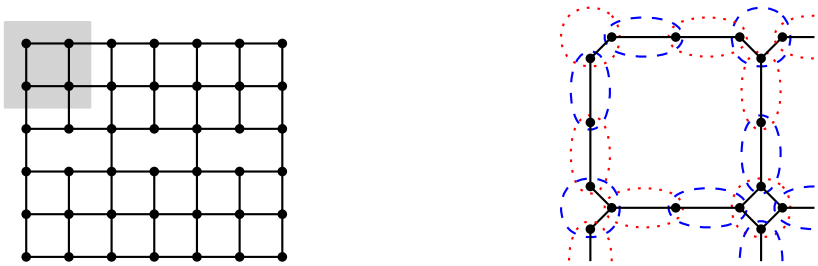
If all skeletons have a planar embedding with the root edge on the outer face, then the construction yields a planar embedding of $G_r(r)$ with the root edge on the outer face. On the other hand, any planar embedding of $G_r(r)$ with the root edge on the outer face can be obtained by permuting the ordering of the edges of the skeleton of a P-node or by flipping the skeleton of an R-node around its root edge.

Every 2-connected graph is represented by a unique SPQR-tree (up to the choice of the root), and the SPQR-tree of a 2-connected graph can be constructed in linear time (Gutwenger and Mutzel 2000).

3.3 NP-Completeness for the Case of two Partitions

In this section we show the NP-completeness of testing c-planarity for graphs whose clusters consist of the union of two partitions. The proof holds even if the underlying graph G is 2-connected, each cluster contains at most two connected components, and the complement of any cluster contains at most three connected components. The proof is based on a proof by Athenstädt et al. (2014), where we show the NP-completeness of a related problem from the field of hypergraph planarity. In contrast to the proof in (Athenstädt et al. 2014), we can now use edges which simplifies the construction of the grid.

Theorem 1. *Deciding c-planarity for a clustered graph is* NP*-complete (Johnson and Pollak 1987). This holds even if the set of clusters is the union of two partitions, the underlying graph G is 2-connected, each cluster contains at most two connected*



(a) Base graph $G_{2,3}$

(b) Detail of the clustered scaffold graph $(G^{\#}_{2,3}, \mathcal{P}^{\#}_{\mathcal{B}} \cup \mathcal{P}^{\#}_{\mathcal{R}})$

Figure 3.6: Extension of the base graph to the clustered scaffold graph (shown for the shaded area on the top left)

components, and the complement of any cluster contains at most three connected components.

We will prove the hardness of the decision problem by a reduction from the NP-complete problem MONOTONE PLANAR 3SAT (see Section 3.2.2). In the following we will construct a graph G^{φ} and a pair of partitions $\{\mathcal{P}^{\varphi}_{B}, \mathcal{P}^{\varphi}_{R}\}$ from a monotone planar 3SAT formula φ such that $(G^{\varphi}, \{\mathcal{P}^{\varphi}_{B} \cup \mathcal{P}^{\varphi}_{R}\})$ is c-planar if and only if φ is satisfiable.

As a prerequisite for the proof we introduce a special grid-like clustered *scaffold graph* $(G^{\#}_{m,n}, \mathcal{P}^{\#}_{\mathcal{R}} \cup \mathcal{P}^{\#}_{\mathcal{B}})$.

We start by constructing a *base graph* $G_{m,n}$ for two integers m and n. The graph $G_{m,n}$ is a grid with $mn + 1$ columns and $2m + 2$ rows of vertices with integer coordinates (i, j) for $0 \leq i \leq mn$ and $0 \leq j \leq 2m + 1$. Each vertex v with coordinates (i, j) is connected to the four vertices at coordinates $(i - 1, j), (i + 1, j), (i, j - 1), (i, j + 1)$ (if they exist). Between the middle rows m and $m + 1$ we remove all vertical edges except for those in columns $0, m, 2m, \ldots, nm$. This defines n larger grid cells of width m between these particular rows. Figure 3.6a shows an example.

We now extend $G_{m,n}$ to the scaffold graph $G^{\#}_{m,n}$ as follows: Each edge in $G_{m,n}$ is split into two edges, connected to a newly introduced *middle vertex*. Each vertex v in $G_{m,n}$ is split up into a set of vertices $S^{\#}(v)$, consisting of a distinct vertex for each adjacent edge. These vertices are connected to each other by an edge or a path of two edges or four-cycle of edges, depending on the degree of the original node. See Fig. 3.6b for an illustration of the construction.

In the next step, we create two partitions of the vertices of the graph. Let v be a vertex from the base graph with coordinates (i, j). If $i+j \pmod 2 = 0$ we aggregate the vertices in $S^{\#}(v)$ into a new block in $\mathcal{P}_{\mathcal{R}}$ and create a block in $\mathcal{P}_{\mathcal{B}}$ for each vertex in $S^{\#}(v)$ and its adjacent middle vertex. If $i + j \pmod 2 = 1$, we do the same but with the partitions reversed. Figure 3.6b shows the blocks created on the scaffold grid. It is easy to see that the resulting blocks in $\mathcal{P}^{\#}_{\mathcal{R}}$ and $\mathcal{P}^{\#}_{\mathcal{B}}$ form indeed each a partition of the vertex set in $G^{\#}_{m,n}$.

Note that the scaffold graph $G^{\#}_{m,n}$ is a subdivision of a 3-connected graph (assuming $n \geq 2$) and thus has a unique embedding in the plane (up to the choice of the outer face).

Now we have all the tools that we need to prove our main theorem in this section.

Proof. (of Theorem 1) First we show that the problem is in NP. According to our definition in Section 3.2.1, a clustered graph is c-planar if and only if it has a planar support. Thus we can guess a graph $G^+(V, E^+)$ with a fixed embedding and then test its planarity and support property in polynomial time. This shows membership in NP. It remains to describe the hardness reduction.

Let φ be a planar monotone 3SAT formula together with an MRR. First we construct the scaffold graph $(G^{\#}_{m,n}, \mathcal{P}^{\#}_{\mathcal{R}} \cup \mathcal{P}^{\#}_{\mathcal{B}})$ where m is the number of clauses of φ and n is the number of variables of φ. We call the n special faces that originated between rows m and $m+1$ of the base graph the *variable cells*.

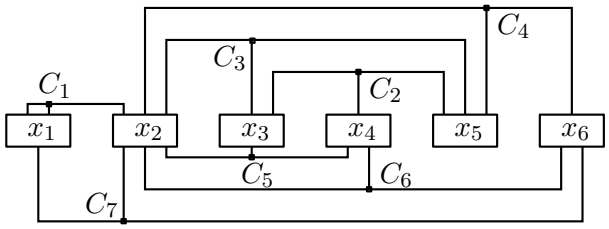
Next we augment the scaffold graph and the two partitions to a new graph $G^{\varphi}_{m,n}$ with partitions $\mathcal{P}^{\varphi}_{\mathcal{R}}$ and $\mathcal{P}^{\varphi}_{\mathcal{B}}$. For each clause we add two E-shaped *combs* to the grid, closely following the layout of the given MRR (see Fig. 3.7a). For the special case that a variable is contained twice in a clause (such as x_1 in C_1), we connect the variable cells by a path. We will explain the details of the augmentation later. For now we focus on the general layout of the paths[3].

Let $C_1, C_2, \ldots, C_l$ be the positive clauses of φ ordered so that if C_i is nested inside the E-shape of C_j in the given MRR then $i < j$. Analogously let $C_{l+1}, \ldots, C_m$ be the ordered negative clauses. We describe the definition of the combs P_i and P'_i for a positive clause C_i ($1 \leq i \leq l$, blue in Fig. 3.7b); combs for negative clauses are defined symmetrically with upper and lower half reversed ($l+1 \leq i \leq m$, red in Fig. 3.7b).

The comb P_i is assigned to the upper half of the base grid, running between rows $m-i$ and $m-i+1$. Furthermore, P_i is assigned to three columns leading towards the variable cells from the top. Let x_j be a variable contained in C_i and assume that C_i is the k-th positive clause from the right connecting to x_j in the embedding of the given MRR. Then P_i runs between columns $jm-k$ and $jm-k+1$. In the lower half of the base grid we translate and mirror the resulting E-shape, creating the comb P'_i as follows: We let P'_i occupy the cells between rows $2m+1-l+i-1$ and $2m+1-l+i$, and the three columns are shifted to the left by the number of occurrences of the respective variable in negative clauses (Fig. 3.7b, empty columns omitted). We thus place the end points of the combs from the top half of the grid to the rightmost possible position in the variable cell and the end points of the combs from the bottom half to the leftmost possible position. Since each variable cell is m columns wide, we can always assign each clause to a unique column of x_j in the top and bottom half of the grid in this way.

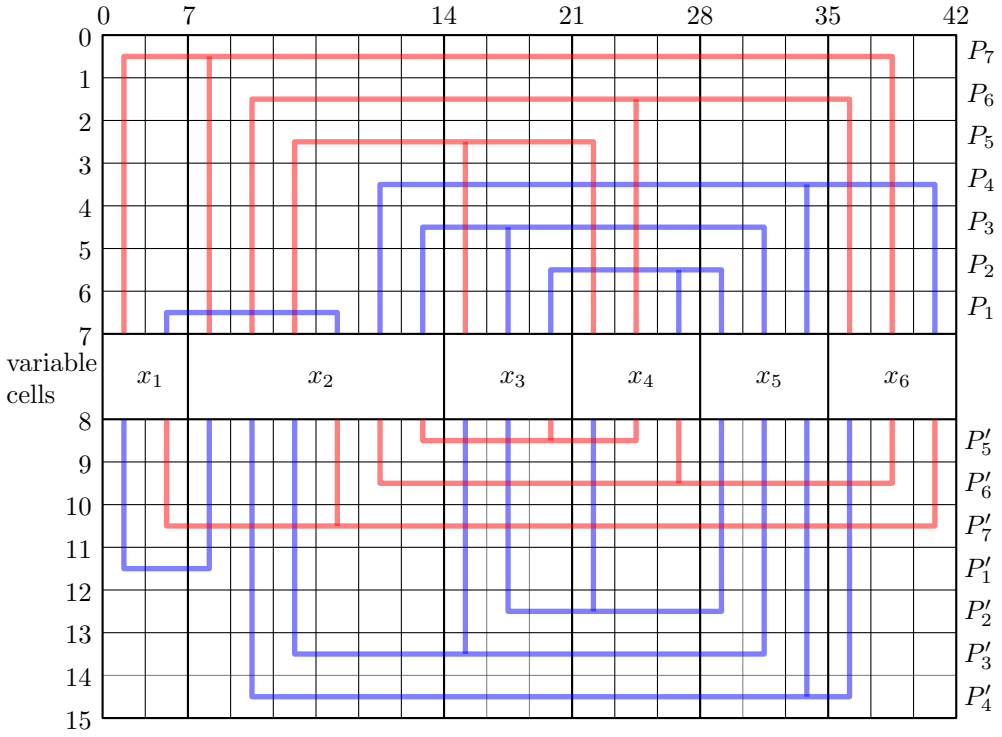
We fix P_i and P'_i to the scaffold graph by adding a *boundary vertex* at each crossing of the comb with the boundary of a grid cell (Fig. 3.8a). If C_i is a positive clause, we add the boundary vertices between the two vertices that share a block in $\mathcal{P}^{\varphi}_{R}$ and

[3]The idea of fixing paths to an underlying grid is inspired by Chaplick et al. (2012)



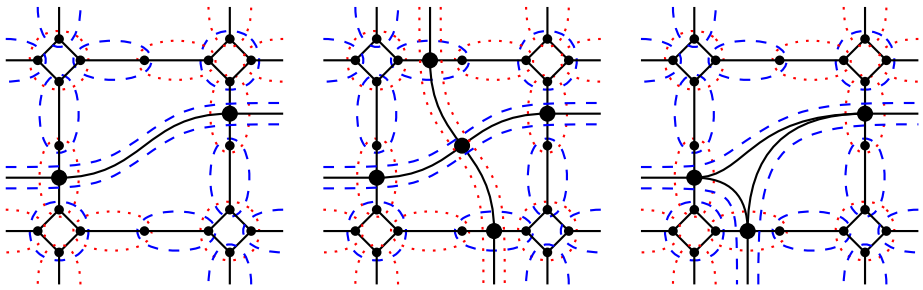
$$C_1 = (x_1 \vee x_1 \vee x_2)$$
$$C_2 = (x_3 \vee x_4 \vee x_5)$$
$$C_3 = (x_2 \vee x_3 \vee x_5)$$
$$C_4 = (x_2 \vee x_5 \vee x_6)$$
$$C_5 = (\overline{x_2} \vee \overline{x_3} \vee \overline{x_4})$$
$$C_6 = (\overline{x_2} \vee \overline{x_4} \vee \overline{x_6})$$
$$C_7 = (\overline{x_1} \vee \overline{x_2} \vee \overline{x_6})$$

(a) MRR of formula φ



(b) Simplified sketch of the layout of the clause blocks on top of scaffold graph (empty columns omitted, labels of the clause blocks on the right side of the corresponding rows)

Figure 3.7: Illustration of the NP-hardness reduction: Construction of $G^{\varphi}_{m,n}$ with partitions $\mathcal{P}^{\varphi}_{\mathcal{R}}$ and $\mathcal{P}^{\varphi}_{\mathcal{B}}$ from the MMR of a formula φ



(a) Part of a comb **(b)** Crossing of two combs **(c)** Fork

Figure 3.8: Augmentation of the scaffold graph (vertices introduced in G^{φ} drawn larger)

otherwise between the two vertices that share a block in $\mathcal{P}_B^\varphi$. The boundary vertex is added to the same block that already contains its neighboring vertices. If two combs cross in a cell we split each comb and add a *crossing vertex* in the center (Fig. 3.8b). For the forks of the combs, we add three edges in a cell, each one connecting a pair of boundary vertices (Fig. 3.8c).

We now create a common *clause block* B_i for P_i and P_i'. The block contains all vertices of the two combs and is added to $\mathcal{P}_B^\varphi$ if C_i is a positive clause and to $\mathcal{P}_R^\varphi$ otherwise.

The resulting graph G^φ is still a subdivision of a 3-connected planar graph and $\mathcal{P}_B^\varphi$ and $\mathcal{P}_R^\varphi$ each form a partition of the vertex set. Also note that G^φ is 2-connected, each cluster in $\mathcal{P}_B^\varphi \cup \mathcal{P}_R^\varphi$ contains at most two connected components (namely P_i and P_i'), and the complement of any cluster in $\mathcal{P}_B^\varphi \cup \mathcal{P}_R^\varphi$ contains at most three connected components (the outside as well as the two components inside the union of the two combs P_i and P_i').

Furthermore, G^φ is almost a planar support of $(G^\varphi, \mathcal{P}_B^\varphi \cup \mathcal{P}_R^\varphi)$. The only thing that is missing is at least one edge for each clause block B_i that connects its two combs P_i and P_i'.

Assume now that φ is a satisfiable formula and a satisfying variable assignment is given. Let x be a variable contained in a clause C_i. If x is set to *true* and C_i is a positive clause, we connect the end vertices of the combs P_i and P_i' in the corresponding variable cell of x. We do the same if C_i is a negative clause and x is set to false. If we do this for each variable cell, there must be at least one cell for each clause C_i in which the end points of combs P_i and P_i' are connected. These connections do not induce any crossings, since all introduced edges run in parallel either from top right to bottom left (if x is *true*) or from top left to bottom right (if x is *false*) through a cell. If we now remove again all but one of the introduced edges for each pair of combs, we end up with a planar support for $(G^\varphi, \mathcal{P}_B^\varphi \cup \mathcal{P}_R^\varphi)$.

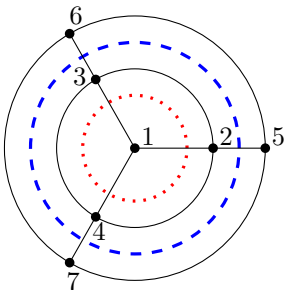
For the other direction let us assume that there is a planar support for $(G^\varphi, \mathcal{P}_B^\varphi \cup \mathcal{P}_R^\varphi)$. Thus, for every clause C_i, there must be an edge connecting the two combs P_i and P_i'. Due to the fixed embedding of G^φ, these connections can only be routed within the variable cells. Since all edges of positive clauses have to traverse the cell from top right to bottom left and all negative clauses from top left to bottom right, each cell can only be traversed either by edges from positive clauses or from negative clauses (otherwise the support would not be planar). If we now assign the variable for each cell accordingly, we end up with a satisfying variable assignment for φ. □

3.4 Two C-Co-Connected Partitions

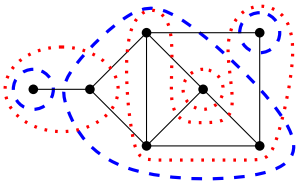
In this section we show that c-planarity can be tested in linear time for a c-co-connected clustered graph if the set of clusters is the union of two partitions. Observe that in contrast to the hierarchical case (Cornelsen and Wagner 2006), there are c-co-connected clustered graphs with an underlying planar graph that are not c-planar. See Fig. 3.9a for an example: The graph $G = (V, E)$ is 3-connected and, thus, has a

unique embedding up to the choice of the outer face. No matter which face we choose as the outer face, there is always at least one cluster C among the four clusters in $\mathcal{P}_B \cup \mathcal{P}_R$ such that $G[C]$ contains a simple cycle enclosing a vertex in $V \setminus C$.

The key for the linear-time algorithm to test c-planarity of c-co-connected clustered graphs is the following characterization.



(a) C-co-connected, but not c-planar



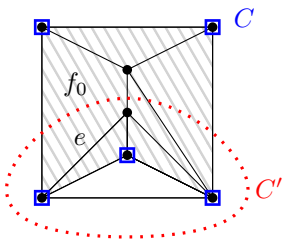
(b) $(G, \mathcal{P}_B)$, $(G, \mathcal{P}_R)$, $(G, \mathcal{P}'_I)$ c-planar, but $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ not

Figure 3.9: a) $\mathcal{P}_R = \{\{1\}, \{2,3,4,5,6,7\}\}, \mathcal{P}_B = \{\{1,2,3,4\}, \{5,6,7\}\}$, i.e., clusters are separated by the curves. $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is c-co-connected and G is planar but $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is not c-planar. b) Clusters are enclosed by the curves. $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is c-connected and $(G, \mathcal{P}_B)$, $(G, \mathcal{P}_R)$, and $(G, \mathcal{P}'_I)$ are c-planar, but $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is not

Theorem 2. *Let $G = (V, E)$ be a graph and let $\mathcal{P}_R$ and $\mathcal{P}_B$ be two partitions of V such that the clustered graph $(G, \mathcal{P}_R \cup \mathcal{P}_B)$ is c-co-connected. Let $\mathcal{P}'_I$ be the connected intersection partition of $\mathcal{P}_R$ and $\mathcal{P}_B$. Then $(G, \mathcal{P}_R \cup \mathcal{P}_B)$ is c-planar if and only if $(G, \mathcal{P}'_I)$ is c-planar.*

Proof. We first show that if $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is c-connected and c-planar, then $(G, \mathcal{P}'_I)$ is c-planar: Consider a c-planar embedding of G for $\mathcal{P}_R \cup \mathcal{P}_B$. Let $C \in \mathcal{P}'_I$, $R \in \mathcal{P}_R$, and $B \in \mathcal{P}_B$ with $C \subseteq B \cap R$. Assume that there is a vertex v in an inner face of $G[C]$. Observe that by c-planarity of $(G, \mathcal{P}_R \cup \mathcal{P}_B)$ all vertices in the inner faces of $G[C]$ are in $B \cap R$. Since $G[B]$ and $G[R]$ are connected, there must be a path from v to C that contains only vertices in one inner face of $G[C]$ and thus in $B \cap R$. Therefore, v and C are in the same connected component of $G[B \cap R]$ and hence $v \in C$. Thus no $v \in V \setminus C$ is in an inner face of $G[C]$ and therefore $(G, \mathcal{P}'_I)$ is c-planar.

We now show that if $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is c-co-connected and $(G, \mathcal{P}'_I)$ is c-planar, then $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is c-planar: Let $C \in \mathcal{P}_B \cup \mathcal{P}_R$. Being c-co-connected implies that $V \setminus C$ is contained in a single face of $G[C]$ in any planar drawing of G. We call C *bad* if $V \setminus C$ is contained in an inner face of $G[C]$. We have to show that there is an embedding that has no bad clusters. Among all planar embeddings of G that are c-planar for $\mathcal{P}'_I$ choose one that has the minimum number of bad clusters. Assume that there is a bad cluster $C \in \mathcal{P}_B \cup \mathcal{P}_R$. Without



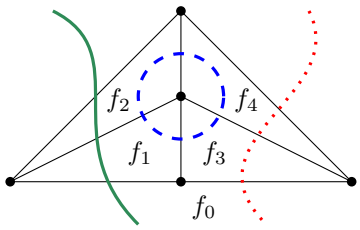


Figure 3.10: A c-co-connected clustered graph $(G, \mathcal{C})$ with 3 partions separated by the lines. In the respective coarsest common refinement, each connected component contains exactly one vertex. The graph is 3-connected, however, no choice of the outer face yields a c-planar embedding

loss of generality, let $C \in \mathcal{P}_B$ (squared blue vertices in the drawing). Let f be the face of $G[C]$ containing $V \setminus C$ (gray shaded area in the drawing).

If C is the union of some clusters in $\mathcal{P}_R$, choose a face f_0 of G inside f incident to a vertex of C as the outer face. Now C is not bad anymore and we do not create any new bad clusters. Thus, we decrease the number of bad clusters – contradicting that we started with an embedding with the minimum number of bad clusters.

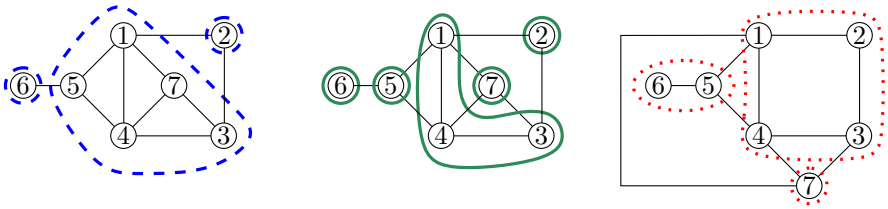
Otherwise, let $C' \in \mathcal{P}_R$ intersect C and $V \setminus C$. Since $G[C']$ is connected, $E(C' \cap C, C' \setminus C)$ is not empty. There must even be an edge $e \in E(C' \cap C, C' \setminus C)$ that is in the outer face of $G[C']$: Otherwise $G[C \cap C']$ would enclose $C' \setminus C$, i.e., there is a cycle c in $G[C \cap C']$ with a vertex in $C' \setminus C$ in its interior. Thus f is contained in the region bounded by c. However, c is contained in a connected component of $C \cap C'$. This contradicts the fact that $(G, \mathcal{P}'_I)$ is c-planar.

Let now f_0 be a face of G incident to e in the outer face of $G[C']$. Then f_0 is incident to a vertex of the outer face of both a graph induced by a cluster in $\mathcal{P}_R$ and a graph induced by cluster in $\mathcal{P}_B$. Thus, f_0 is not contained in any cluster. Choosing f_0 as the outer face decreases the number of bad clusters by the same argument as above. This contradicts that we have chosen a planar embedding minimizing the number of bad clusters. □

Note that this method does not work for more than two partitions. Consider the example in Fig. 3.10.

Since c-planarity of c-connected hierarchically clustered graphs can be tested in linear time (Cortese et al. 2008) and $(G, \mathcal{P}'_I)$ is c-connected and hierarchically clustered, it remains to show that $\mathcal{P}'_I$ can be constructed in linear time. Since connected components can be computed in linear time, it suffices to show that the intersection partition $\mathcal{P}_I$ of the two partitions $\mathcal{P}_B = \{B_1, \ldots, B_{\ell_B}\}$ and $\mathcal{P}_R = \{R_1, \ldots, R_{\ell_R}\}$ of V can be computed in linear time. This might be common knowledge, but for the sake of completeness we give a quick description here:

We introduce the following data structure: For $X \in \{B, R\}$, we use a vertex array with $X[v] = i$ for $v \in X_i$. We also initialize an array $S[1, \ldots, \ell_R]$ of stacks, where $S[i]$ will contain the vertices of R_i, $i = 1, \ldots, \ell_R$ in the order in which they appear in $B_1, \ldots, B_{\ell_B}$. We fill the stacks as follows: For $i = 1, \ldots, \ell_B$ and $v \in B_i$, we push



(a) $\mathcal{P}_B$ **(b)** $\mathcal{P}'_I = \mathcal{P}_I$ **(c)** $\mathcal{P}_R$

Figure 3.11: Illustration why in Fig. 3.9b, $(G, \mathcal{P}_B)$, $(G, \mathcal{P}_R)$, and $(G, \mathcal{P}'_I)$ are c-planar, but $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is not

v to $S[R[v]]$. Now, the sets in $\mathcal{P}_I$ can be obtained by going through the stacks and opening a new set whenever $B[v]$ changes. This concludes the proof of the following theorem.

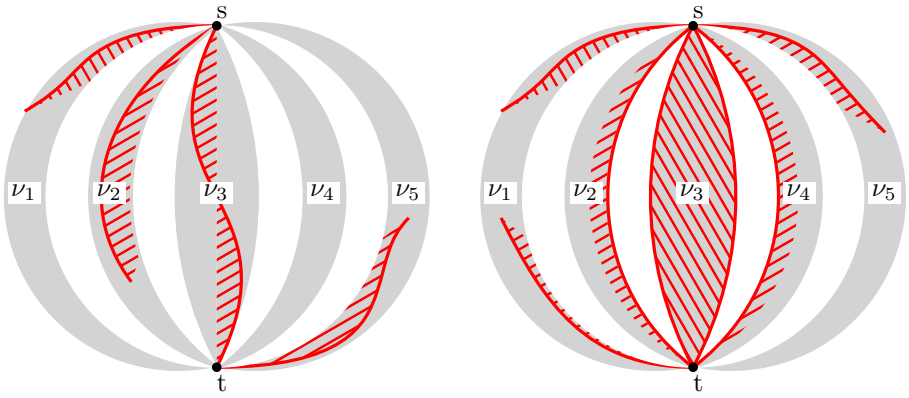
Theorem 3. *It can be tested in linear time whether a c-co-connected clustered graph is c-planar if the set of clusters is the union of two partitions of the vertex set.*

Observe that if $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is only c-connected, then $(G, P_B \cup \mathcal{P}_R)$ does not have to be c-planar even if $(G, \mathcal{P}_B)$, $(G, \mathcal{P}_R)$, and $(G, \mathcal{P}'_I)$ are. See the example in Fig. 3.9b: Let G be the graph and let $\mathcal{P}_R$ and $\mathcal{P}_B$, respectively, be the partition of the vertex set enclosed by the red dotted and blue dashed curves, respectively. Let $\mathcal{P}'_I$ be the connected intersection partition of $\mathcal{P}_R$ and $\mathcal{P}_B$. Then $(G, \mathcal{P}_B)$, $(G, \mathcal{P}_R)$, and $(G, \mathcal{P}'_I)$ are c-planar, but $(G, \mathcal{P}_B \cup \mathcal{P}_R)$ is not. This can be shown as follows:

The embedding in Fig. 3.9b is c-planar for $\mathcal{P}_B$ and $\mathcal{P}'_I$ – see also Fig. 3.11a and Fig. 3.11b. Fig. 3.11c shows an embedding that is c-planar for $\mathcal{P}_R$. Assume now that there would be an embedding that is c-planar for $(G, \mathcal{P}_B \cup \mathcal{P}_R)$. We use the vertex labeling indicated in Fig 3.11. Due to cluster $\{1, 2, 3, 4\}$ the interior of the cycle $c_R = \langle 1, 2, 3, 4 \rangle$ must be empty. Thus, vertex 5 must be drawn outside c_R. Due to the cluster $C = \{1, 3, 4, 5, 7\}$, vertex 2 and 6 must not be enclosed by the triangle $c_B = \langle 1, 4, 5 \rangle$. It follows that the edges connecting 5 to c_R must be drawn such that c_B does not enclose c_R and that 6 is outside c_B. Due to the edge $\{3, 7\}$, vertex 7 is not enclosed by c_B either. Thus, except for the edge $e = \{1, 7\}$, the embedding is as indicated in Fig. 3.11c. But no matter how we would add e in a planar embedding, we would either create a cycle in $G[C]$ enclosing vertex 2 or vertex 6 which are not in C.

3.5 C-Connected Clusterings on Blocks

We now present a polynomial-time method for testing c-planarity for a planar 2-connected graph G and a set of (overlapping) c-connected clusters $\mathcal{C}$. Throughout this section, let T be the SPQR-tree of G rooted at a Q-node r representing the edge e of G. We start with an informal description.



(a) Double-border **(b)** Inner and border

Figure 3.12: In (a) node ν_3 is double-border and all other nodes are outside. In (b) node ν_3 is inside, ν_2 and ν_4 are border, and ν_1 and ν_5 are outside. The red area sketches the vertices within the cluster

3.5.1 Informal Description

Recall that all planar embeddings of G with e on the outer face can be obtained by ordering the children of the P-nodes and by flipping the skeletons of the R-nodes. In order to find a suitable embedding, we will use a labeling scheme for each node ν in T and its corresponding edge in the skeleton of its parent node to capture which parts of skel(ν) are contained in a given cluster.

Assume for a moment that the SPQR-tree does not contain R-nodes. In order to find a c-planar embedding, we have to permute the children of each P-node in such a way that for any cluster C there is no C-cycle enclosing a vertex which is not in C. So consider a fixed P-node ν and a fixed cluster C. We call a C-cycle c of $G_r^-(\nu)$ *new* if c is not already contained in the graph represented by a child of ν. We distinguish four cases for a child ν_i of ν (see Fig. 3.12).

1. If $G_r^-(\nu_i)$ does not contain a C-path between its poles, it does not contribute to a new C-cycle, so we can put ν_i as one of the first or last children of ν and we call ν_i an *outside* node.

2. If $G_r^-(\nu_i)$ contains only vertices of C, then $G_r^-(\nu_i)$ can be contained in the interior of any C-cycle and we call ν_i an *inside* node.

3. If $G_r^-(\nu_i)$ contains a C-path between its poles, but also vertices not in C, we have to make sure that the latter are not enclosed by a C-cycle. We call ν_i *border* or *double-border*, depending on whether the vertices not in C can all be put on the same side of the C-paths, connecting the poles or not.

Now the children of ν must be permuted such that either there is one double-border node pre- and succeeded by arbitrary many outside nodes (see Fig. 3.12a) or there are arbitrary many consecutive inside nodes pre- and succeeded by at most one border node, respectively, and then arbitrary many outside nodes (see Fig. 3.12b).

Note that the method described in this section has some similarities with the al-

gorithm described by Angelini et al. (2015) for deciding whether a hierarchically clustered graph has a drawing with region-region crossings. They also use a labeling scheme for the nodes in the SPQR-tree. Their "full" corresponds to our "inside", their "spined" corresponds to our "non-outside". Observe, however, that their labels "side-spined" and "central-spined" depend on a given drawing while our labels "double-border" and "border" do not.

Depending on the labels, Angelini et al. (2015) permute the edges of the skeletons of each P-node and decide how to flip the children according to a 2-SAT formula. However, this requires that no two clusters determine different flips. This property does no longer hold if clusters overlap as in our case. We discuss this problem in more detail in Section 3.5.3. Thus, in our approach we find the permutations of the children of multiple P-nodes simultaneously. In addition, we demonstrate how our method can be extended to the case where the underlying graph is not 2-connected (see Section 3.6).

In the next section, we describe our labeling scheme more formally. In Section 3.5.3, we then show how to use the consecutive ones property to check for all P-nodes and for all clusters simultaneously whether the required permutation can be found and whether the R-nodes can be flipped accordingly.

3.5.2 Formal Description and Characterization

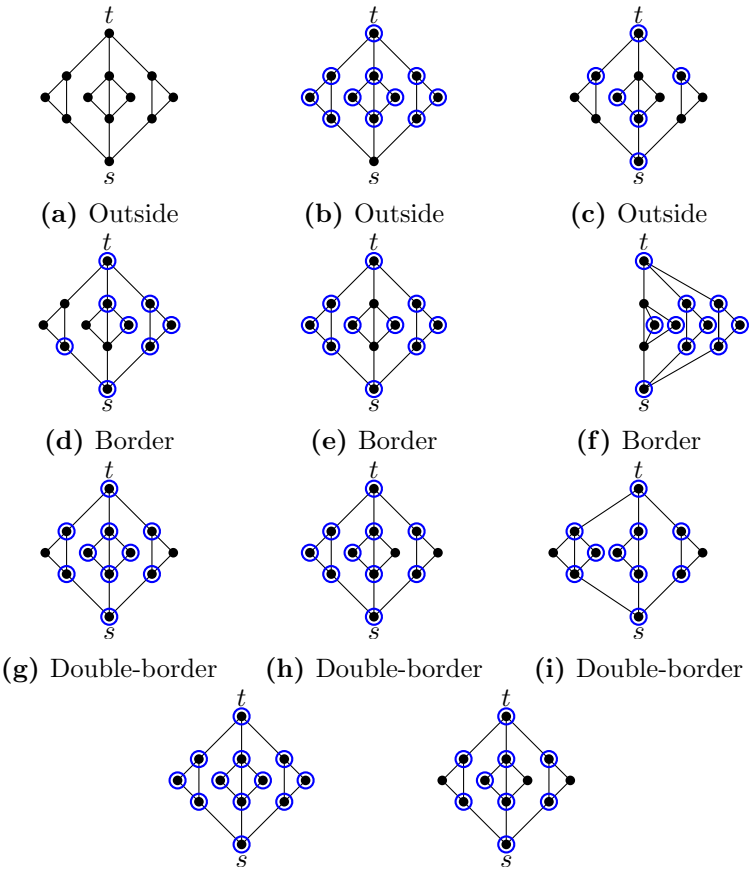
In the following let $C \in \mathcal{C}$ be a cluster such that $G[C]$ is connected and let $C_{\text{ext}} \subseteq C$ be a subset of vertices that we want to be incident to the outer face of $G[C]$. (We will need C_{ext} in Section 3.6 when extending the method to non-biconnected graphs. In that case, C_{ext} will be a subset of the set of cut-vertices. In the case of biconnected graphs, C_{ext} is empty.)

Let ν be a node in T and let s and t be the poles of $G_r(\nu)$. We call a planar embedding of $G_r^-(\nu)$ *appropriate* if (a) the poles are on the outer face and (b) no C-cycle encloses a vertex of $(V \setminus C) \cup C_{\text{ext}}$.

Given a planar embedding of $G_r(\nu)$, let f_1 and f_2 be the two faces of $G_r(\nu)$ that are incident to the root edge. An *outer s-t-path* is the s-t-path in $G_r^-(\nu)$ that is incident to f_1 or f_2, respectively. Observe that there is one outer s-t-path if ν is a Q-node or an S-node that has only Q-nodes as children. Otherwise there are exactly two outer s-t-paths. E.g., in Fig. 3.14 the two outer s-t-paths of $G_{\{1,2\}}(S_2)$ are $\langle 2, 8, 7, 6, 4, 1\rangle$ and $\langle 2, 6, 5, 1\rangle$. We call a vertex of an outer s-t-path *inner*, if it is not incident to both faces, f_1 and f_2. I.e., the poles s and t are never inner. Further, if ν is an S-node, then the poles of the children of ν are also not inner. E.g., vertex 5 is the only inner vertex of the outer s-t path $\langle 2, 6, 5, 1\rangle$ of $G_{\{1,2\}}(S_2)$ in Fig. 3.14.

Now, for a given cluster C, we label the node ν as

- *inside* if all vertices of $G_r(\nu)$ are contained in C and at most the poles s and t of $G_r(\nu)$ are in C_{ext}.
- *inappropriate* if $G_r^-(\nu)$ has no appropriate embedding.



(a) Outside **(b)** Outside **(c)** Outside

(d) Border **(e)** Border **(f)** Border

(g) Double-border **(h)** Double-border **(i)** Double-border

(j) Inside **(k)** Inappropriate

Figure 3.13: Illustration of the labeling of a P-node ν depending on various given clusters C. The pictures show the graph $G_r^-(\nu)$. Blue encircled vertices are in C. We always assume $C_{\text{ext}} = \emptyset$. The node ν is labeled border in 3.13e, since $G_r^-(\nu)$ has an embedding (shown in 3.13f) that is c-planar for $\{C\}$ in which only one of the outer s-t-paths contains vertices in C. Similarly, the embedding in 3.13i shows why ν is labeled double-border in 3.13h

- *outside* if it is not inappropriate and $G_r^-(\nu)$ contains no C-path between its poles.
- *border* if ν is neither inside nor outside and $G_r^-(\nu)$ has an appropriate embedding in which at most one of the outer s-t-paths contains inner vertices in $(V \setminus C) \cup C_{\text{ext}}$.
- *double-border* in all other cases.

See Fig. 3.13 for examples of the different labels in the case of a P-node and Fig. 3.14 for a more general example, both with $C_{\text{ext}} = \emptyset$. Let μ be a child of ν and let e_μ be the edge of skel(ν) corresponding to the edge $\{\nu, \mu\}$. Then e_μ gets the same label as μ. We refer to nodes (and their corresponding edges) of any label except inside as *non-inside* and of any label except outside as *non-outside*.

Without computing any particular embedding and only by traversing the SPQR-

tree T, we can decide for all nodes in linear time, whether they are inside, outside, border, double-border, or inappropriate for a given cluster C.

To this end, we use the following terminology given an *R-node* ν: Consider the unique embedding of $\text{skel}^-(\nu)$ with the poles s and t on the outer face. A *caging cycle* is a simple cycle of non-outside edges in $\text{skel}^-(\nu)$ that (a) contains double-border edges or (b) encloses vertices in C_{ext} or non-inside edges. The following remark is a direct consequence of the definitions.

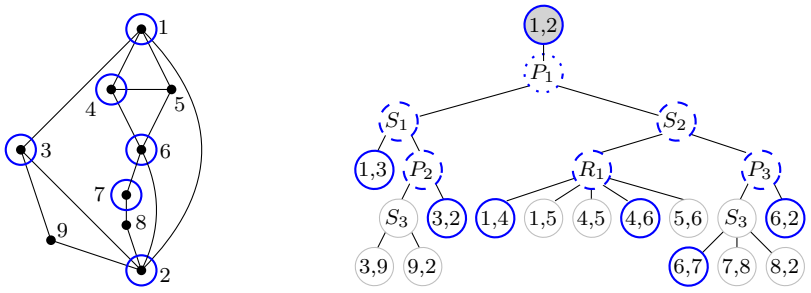
Remark 1. *An R-node ν is inappropriate if and only if one of its children is inappropriate or $\text{skel}^-(\nu)$ contains a caging cycle.*

Lemma 1. *Once the edges of the skeleton of an R-node ν without inappropriate children are labeled, it can be computed in time linear in the size of the skeleton, whether $\text{skel}^-(\nu)$ contains a caging cycle.*

Proof. First remove all outside edges of $\text{skel}^-(\nu)$. Remove all but the edges incident to the outer face of the resulting graph and also remove all bridges. The remainder S is a collection of edge-disjoint simple cycles each with empty interior (in S). If S contains any double-border edge then there is a caging cycle. Otherwise consider all edges and vertices of $\text{skel}^-(\nu)$ that were in the interior of one of the cycles of S. If among those there is a vertex in C_{ext} or an edge labeled other than inside then there is a caging cycle. Otherwise there is no caging cycle. □

Lemma 2. *For each cluster, the labels can be computed in time linear in the size of the input graph as summarized in Table 3.1.*

Proof. Since the size of an SPQR-tree (including the sum of the sizes of the skeletons) is linear in the size of the represented graph, the linear run time follows directly from the construction and the previous remark. □



(a) $(G, \{C\})$ **(b)** SPQR-tree T_r

Figure 3.14: (a) A graph G with cluster $C = \{1, 2, 3, 4, 6, 7\}$ containing the circled vertices and (b) the labeling of the nodes of its SPQR-tree: thick blue solid circled nodes are inside, dashed nodes border, and dotted nodes double-border. The root is inappropriate. The remaining nodes are outside. $C_{\text{ext}} = \emptyset$

An *external path* of a node ν is a path in G between the poles of $G_r(\nu)$ that does not contain any other vertices of $G_r(\nu)$. We label the root edge of $\text{skel}(\nu)$ *inside* for C if ν has an external C-path and *outside* otherwise. We say that an external path p of

Node ν is	**inside**	**outside**	**border**	**double-border**
		if ν has no inappropriate children and ...		
	if at most the poles are in C_{ext} and ...	ν not inside and ...	ν not inside, not outside, and ...	ν not inside, not outside, not border, and ...
Q-node (leaf)	both vertices are in C.	not both vertices are in C.		
P-node	all children are inside.	all children are outside.	ν has no double-border children and at most one border child.	ν has (a) two border children or (b) one double-border child and neither inside nor border children.
S-node	all children are inside.	ν has at least one outside child.	ν has no double-border children.	ν has at least one double-border child.
R-node	all children are inside.	skel(ν) contains no caging cycle and no s-t-path of non-outside edges.	skel(ν) contains no caging cycle and exactly one outer s-t-path contains non-inside edges or vertices from $C_{\text{ext}} \setminus \{s, t\}$.	skel(ν) contains no caging cycle.
Q-node (root)	the child of ν is inside.	the child of ν is outside and $s \notin C$ or $t \notin C$.	the child of ν is (a) border or (b) outside and $s, t \in C$.	the child of ν is double-border and $s \notin C$ or $t \notin C$.

Table 3.1: Computation of the labels of a node ν with respect to a cluster C and a subset $C_{\text{ext}} \subseteq C$, proceeding from the leaves to the root of the SPQR-tree. s and t are the poles. A node is inappropriate if and only if it is neither inside nor outside nor border nor double-border.

node ν is *to the right (left) of* ν with respect to the ordered pair (s, t) of its poles if the cycle that is induced by p in the graph that results from G by contracting $G_r(\nu)$ into a single vertex – and removing multi-edges and loops – is oriented (counter-) clockwise, assuming that p was oriented from t to s. Two external paths of ν are on the *same side* of ν if they are both to the right or both to the left of ν with respect to an arbitrary ordering of the poles of ν. Otherwise, they are on *different sides*.

The following lemma characterizes the c-planar embeddings with respect to a single cluster.

Lemma 3. *Let $G = (V, E)$ be a 2-connected graph, let $C \subset V$ be a cluster inducing a connected subgraph of G, let $C_{ext} \subseteq C$, let T be the SPQR-tree of G, and let r be a Q-node of T representing an edge e of G. A planar embedding of G with e on the outer face is c-planar for $\{C\}$ with C_{ext} incident to the outer face of $G[C]$, if and only if the following conditions are fulfilled for any non-inside node ν of its SPQR-tree T.*

1. *All external C-paths of ν are embedded on the same side of $G_r(\nu)$ – which is reflected by the side on which the root edge of skel(ν) is embedded.*
2. *skel(ν) contains no simple cycle of non-outside edges that encloses a non-inside edge or a vertex in C_{ext}.*

Proof. By definition, both conditions must be fulfilled for a c-planar embedding with C_{ext} on the outer face of the cluster. So assume now that both conditions are fulfilled. Let $v \in (V \setminus C) \cup C_{\text{ext}}$, let $e' \neq e$ be an edge incident to v, and let ν' be the Q-node representing e'. Assume that G contains a C-cycle c' enclosing v. By condition 1, there is no node ν on the ν'-r-path such that c' can be decomposed into two external C-paths of ν. So let ν be the first node on the ν'-r-path such that c' is contained in $G_r^-(\nu)$ or can be composed by a path in $G_r^-(\nu)$ and an external C-path of ν. Observe that c' induces a cycle c in skel(ν) that contains only non-outside edges. Let μ be the child of ν on the ν'-ν-path. By the choice of ν it follows that c does not contain the edge e_μ of skel(ν). Hence c encloses the edge e_μ. However, e_μ was either not inside or $v \in C_{\text{ext}}$ is an end vertex of e_μ not in c – contradicting condition 2. □

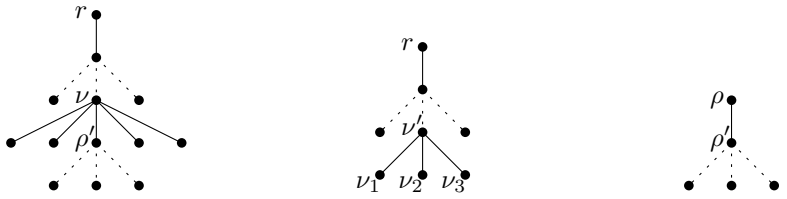
In Section 3.5.3, we construct a set of binary matrices from an initial embedding of T that have the consecutive-ones property if and only if there is a c-planar embedding for $\mathcal{C}$ with the fixed root edge on the outer face. The ordering of the columns of the matrices will correspond to the orderings of the children of the P-nodes and the flips of the R-nodes. The total size of the matrices will be in $\mathcal{O}(|V|^2|\mathcal{C}|)$.

3.5.3 Modeling by Consecutive-Ones Property

For each possible root r of T that is not inappropriate for any $C \in \mathcal{C}$, we start with a fixed embedding of T – including fixed flips of the R-nodes – and perform the following steps:

Splitting T

We split T at each R-node (see Fig. 3.15), removing the edges from the R-node to its children from T. Let T_r be the subtree containing r. For each former non-leaf child ρ' of an R-node ν we attach a new Q-node ρ to ρ'. We root the subtree containing ρ' at ρ and denote it by T_ρ. We label ρ inside for a cluster C if ρ' had an external C-path and outside otherwise. In the parent tree, we replace the R-node ν by a special P-node ν' with the same label and three Q-nodes ν_1, ν_2, ν_3 in this order as children. If the R-node ν was labeled border for a cluster C, we label ν_2 and exactly one among ν_1 and ν_3 as inside and the other as outside. More precisely, we label ν_1 as outside if and only if the left outer path of $\text{skel}^-(\nu)$ between its poles contains non-inside edges or vertices from C_{ext}. If the R-node ν was labeled double-border, we label ν_1 and ν_3 as border and ν_2 as inside. If the R-node was labeled inside or outside, we label all three children as inside or outside, respectively. We thus end up with a forest containing only S-, P-, and Q-nodes.



(a) SPQR tree T **(b)** T_r **(c)** T_ρ

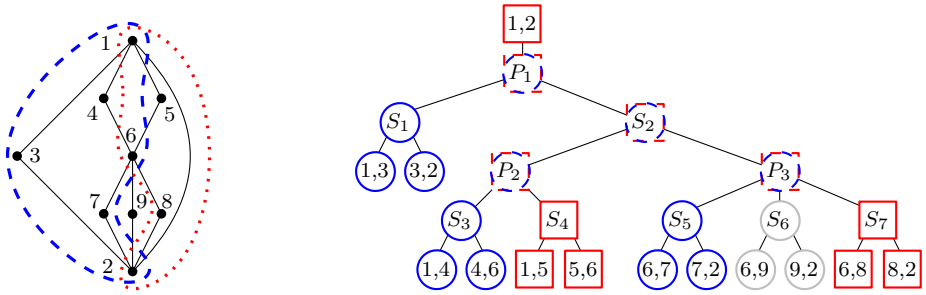
Figure 3.15: The tree T is split at every R node, resulting in a forest containing only S-, P-, and Q-nodes

Initializing the Matrices

For each root ρ of one of the subtrees, we create a new binary matrix M_ρ. A node in T_ρ is a *lowest-P-child*, if it is the child of a P-node and has no other P-nodes in its subtree. The embedding of T_ρ induces an ordering of the lowest-P-children from left to right. We initialize M_ρ with a column for each lowest-P-child in accordance with the ordering. For a node ν of T_ρ, we use $c(\nu)$ to refer to the set of its *corresponding columns* in M_ρ, i.e., the columns of the lowest-P-children in ν's subtree. For M_r, we create one additional *external column* $c(r)$. For $\rho \neq r$, we create two additional *external columns*, enclosing the rest of the matrix.

For each cluster C, one of the two external columns will represent the side of possible external C-paths of the child of ρ and will be denoted by $c_C(\rho)$.

We then create a row for each non-leaf node ν distinct from a lowest-P-child, adding 1s in the columns in $c(\nu)$ and 0s in all other columns. This ensures that in every permutation of the columns of M_ρ for which the 1s are consecutive in all rows, the columns of the lowest-P-children of each node remain adjacent, allowing a reconstruction of an embedding of T_ρ from the ordering of the columns in M_ρ. See Fig. 3.16. If $\rho \neq r$, we add two rows, having all 1s except for one 0 in the first or last external column, respectively.



(a) $(G, \mathcal{C})$ **(b)** SPQR-tree T_r

$c(S_1)\ c(S_3)\ c(S_4)\ c(S_5)\ c(S_6)\ c(S_7)\ c(1,2)$ $(=ext)$

P_1	initialization	1	1	1	1	1	1	0
S_2	initialization	0	1	1	1	1	1	0
P_2	initialization	0	1	1	0	0	0	0
P_3	initialization	0	0	0	1	1	1	0
$r_0(P_1, B)$		1	0	0	0	0	0	0
$r_1(P_1, B)$	P_1	1	1	1	1	1	1	0
$r_0(P_2, B)$	P_1	1	P_2	1	0	0	0	0
$r_0(P_3, B)$	P_1	1	S_2	1	1	P_3	1	0	0	0
$r_1(P_1, R)$		0	1	1	1	1	1	1
$r_0(P_2, R)$	P_1	0	P_2	0	1		1	1	1	1
$r_0(P_3, R)$	P_1	0	S_2	0	0	P_3	0	0	1	1

(c) M_r

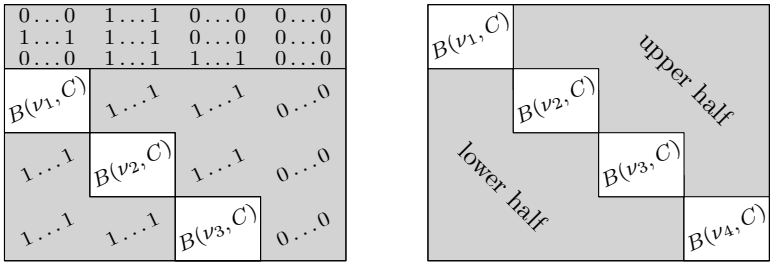
Figure 3.16: $\mathcal{C} = \{R, B\}$ with $R = \{1, 2, 5, 6, 8\}$, $B = \{1, 2, 3, 4, 6, 7\}$. Circled blue nodes in T_r are inside for B, squared red nodes inside for R. P_1, P_2, P_3, and S_2 are border for R and B. R and B have a different constraint in S_2, and thus different halves of the blocks are filled with 1s in M_r

In order to fill the matrix M_ρ, we traverse the tree T_ρ with a post-order DFS. For each cluster $C \in \mathcal{C}$ and each examined node, we add up to three rows to M_ρ. We define for each node ν and each cluster C a set $r(\nu, C)$ of *relevant rows*. For each lowest P-child ν, we set $r(\nu, C) = \emptyset$. The block $B(\nu, C)$ is the submatrix of M_ρ with entries in rows $r(\nu, C)$ and columns $c(\nu)$. When we create rows in M_ρ, the default entries are 0, and we explicitly mention when we set the entries to 1.

Handling P-nodes

For a P-node ν with children $\nu_1, \ldots, \nu_k$, we initialize $r(\nu, C)$ as $r(\nu_1, C) \cup \ldots \cup r(\nu_k, C)$. Since ν is not inappropriate, we have one among the following two cases: a) ν has only outside children, except for possibly one double-border child. Then the children of ν can be permuted arbitrarely. b) the children of ν must be permuted such that all inside children are consecutive and pre- and succeeded by at most one border child and arbitrary many outside children.

Thus, if ν is not outside and has no double-border child, we add up to 3 constraint-rows $r_0(\nu, C)$, $r_1(\nu, C)$, and $r_2(\nu, C)$ to $r(\nu, C)$. If ν has inside children, we add $r_0(\nu, C)$ with 1s in all columns in $c(\nu_i)$ where ν_i is an inside child of ν. This ensures that all inside children are placed in consecutive order. If ν has a child μ that is a border node, we add $r_1(\nu, C)$ with 1s in all columns in $c(\mu)$ and again with 1s in all columns in $c(\nu_i)$ where ν_i is an inside child of ν. We do the same for a potential second border child in a third row $r_2(\nu, C)$. This ensures that the border children are placed next to the inside children, with at most one border child on each side. Finally, let μ be a child of ν, let $i \in r(\mu, C)$, and let $j \in c(\nu) \setminus c(\mu)$. Then we set the entry in row i and column j to 1 if one or more of the rows in $r(\nu, C) \setminus r(\mu, C)$ contain a 1 in the same column. This is to make sure that the columns between the P-node and potential ones induced by external paths is filled with 1s. See Fig. 3.17a.



(a) P-node **(b)** S-node

Figure 3.17: (a) The block $B(\nu, C)$ for a P-node ν with an inside child ν_2, an outside child ν_4, and two border children ν_1 and ν_3. (b) The block $B(\nu, C)$ for an S-node ν with four children. If ν is not outer and has an external C-path, then the upper half, the lower half, or both are filled with 1s

Handling S-nodes

If an S-node ν with children $\nu_1, \ldots, \nu_k$ is outside, then $r(\nu, C) = \emptyset$, otherwise $r(\nu, C) = r(\nu_1, C) \cup \ldots \cup r(\nu_k, C)$.

Assume now that ν is not outside and has an external C-path. Observe that in this case ν cannot be double-border. Otherwise r would be inappropriate for C. If ν has two or more P-nodes as children, we have to make sure that the 1s in each P-node and the 1s in the external path can be made consecutive via additional 1s.

More precisely, let $\nu_1, \nu_2, \ldots, \nu_k$ be the children of ν that are P-nodes. The upper half for a cluster C are all entries in rows $r(\nu_i, C)$, $i = 1, \ldots, k$ and columns $c(\nu_j)$, $j = i + 1, \ldots, k$ while the lower half are all entries in rows $r(\nu_i, C)$, $i = 1, \ldots, k$ and columns $c(\nu_j)$, $j = 1, \ldots, i - 1$. We fill both, the upper and the lower half with 1s, if ν is inside and we fill either the upper or the lower half with 1s if ν is border. See Fig. 3.17b.

Recall that if ν is not inside, then the external C-paths must all be on the same side of $G_r(\nu)$ in a c-planar drawing of G. However, external C_1- and C_2-paths could be on different sides for distinct clusters C_1 and C_2. Hence, we cannot just always fill the upper half with 1s. To this end, we will define same and different constraints on

the clusters that are *critical* for ν, i.e., on the clusters C such that ν is border with respect to C and has an external C-path.

Let C_1 and C_2 be two clusters that are critical for ν. Then we say that there is a *same (different) constraint* between C_1 and C_2 if the external C_1-paths and the external C_2-paths must be on the same side of $G_r^-(\nu)$ in any c-planar embedding of G. The following remark implies that there is either a same or a different constraint between any two clusters that are critical for ν and describes how to decide the type of constraint.

Remark 2. *Let C_1 and C_2 be two clusters that are critical for ν. If ν has an external C_1-path that is also an external C_2-path, then there is a same constraint between C_1 and C_2. Otherwise, there is a different constraint between C_1 and C_2*

Proof. Assume that there is an external C_1-path p_1 and an external C_2-path p_2 of ν that are on the same side of $G_r(\nu)$ in a c-planar embedding of G.

Since ν is border, there is a C_i-path p_i^ν, $i = 1, 2$ in $G_r^-(\nu)$. Consider the cycles c_i, $i = 1, 2$, composed by p_i^ν and p_i. By c-planarity, each portion of p_1 that is inside c_2 must be in $G[C_2]$ and vice versa. Since p_1 and p_2 are on the same side of $G_r(\nu)$, there is an external path of ν that contains only edges of p_1 inside c_2, edges of p_2 inside c_1 and common edges of p_1 and p_2, i.e., only edges in $G[C_1 \cap C_2]$. □

Observe that for every S-node ν and every pair C_1, C_2 of clusters, we can decide in $\mathcal{O}(|V|)$ time whether ν has an external $(C_1 \cap C_2)$-path, and thus, whether there is a same-constraint between C_1 and C_2.

Fix now an arbitrary cluster C that is critical for ν and assign C the upper half. Assign to any other cluster C' that is critical for ν the upper half if there is a same constraint between C and C' and the lower half otherwise.

External Columns

If $\rho = r$ (i.e., the subtree has no parent), let ν be the unique child of r, and let e be the edge represented by the Q-node r. Then the external column is 1 for each row in $r(\nu, C)$ if the cluster C contains both end vertices of e.

If $\rho \neq r$, then the unique child ρ' of ρ was the child of an R-node ν. Consider the fixed embedding of $\text{skel}^-(\nu)$ with its poles s and t on the external face. Let C be a cluster for which ρ' is a border node and has an external C-path. We have to make sure that the parts of $G_r^-(\rho')$ that are not in $C \setminus C_{\text{ext}}$ are embedded such that they are not enclosed by a C-cycle in G that is composed of an external C-path of ρ' and a C-path in $G_r(\rho')$ between its poles s' and t'.

Consider first that $\text{skel}^-(\nu)$ contains a simple cycle c containing $e_{\rho'}$ and consisting only of non-outside edges. If c is (counter-)clockwise oriented when orienting $e_{\rho'}$ from s' to t', then we set $c_C(\rho)$ to be the (left) right external column.

Otherwise all external C-paths of ρ' must contain an external C-path of ν. Thus, ν is not double-border. Moreover, the set of vertices of $\text{skel}^-(\nu)$ that can be reached

from s, using only non-outside edges and not $e_{\rho'}$, induces an s-t-cut of $\text{skel}^-(\nu)$ that contains $e_{\rho'}$ and no other non-outside edges. It follows that $e_{\rho'}$ is on the left (right) outer s-t-path and all external C-paths of ν are to the left (right) of ν with respect to (s,t) in any c-planar embedding. Hence, if e is on the left (right) outer s-t-path, then we set $c_C(\rho)$ to be the left (right) external column. In both cases we set the entry in column $c_C(\rho)$ to 1 for each row in $r(\rho', C)$.

Clearly the number of columns is linear in the number of Q-nodes and R-nodes and thus linear in $|V|$ for planar graphs. During initialization, we add $\mathcal{O}(|V|)$ rows, namely one for each inner node. For a cluster C and a P-node ν, we add up to three rows but only if both poles are in C. Observe that at least one of the poles of a P-node ν is not a pole of another P-node ν' on the path from ν to the root. Hence, the number of rows added for P-nodes is bounded by $3 \cdot \sum_{C \in \mathcal{C}} |C| \in \mathcal{O}(|V| \cdot |\mathcal{C}|)$.

Applying the next theorem with $C_{\text{ext}} = \emptyset$ yields a characterization of c-connected overlapping clustered graphs with underlying 2-connected graphs.

Theorem 4. *A c-connected overlapping clustered graph $(G, \mathcal{C})$ with an underlying planar 2-connected graph G and sets $C_{ext} \subseteq C$, $C \in \mathcal{C}$ has a c-planar embedding in which C_{ext} is incident to the outer face of $G[C]$ for any $C \in \mathcal{C}$ if and only if the root of the SPQR-tree of G can be chosen such that it is not inappropriate for $C \in \mathcal{C}$ and all matrices M_ρ fulfill the consecutive-ones property.*

We prove the theorem in the following two subsections.

3.5.4 Proof of Sufficiency

Let the SPQR-tree T of G be rooted at the Q-node r, and let e be the edge represented by r. Assume that r is not inappropriate for any cluster and that the columns of all matrices M_ρ are permuted such that in each row the 1s are consecutive. We have to show that $(G, \mathcal{C})$ is c-planar.

We may assume without loss of generality that the external columns were not permuted. (Otherwise reverse the order of the columns.)

Starting from $\rho = r$, we traverse T and do the following at a non-leaf node ν:

If ν is a P-node, we permute the children $\nu_1, \ldots, \nu_k$ of ν according to the ordering of $c(\nu_1), \ldots, c(\nu_k)$ in the permuted matrix M_ρ.

If ν is an R-node, we fixed an embedding of $G_r(\nu)$ and replaced ν with a P-node and three incident Q-nodes ν_1, ν_2, ν_3 in this order. If ν was labeled inside or outside for all clusters, then we maintain the fixed flip of $G_r(\nu)$. Otherwise the labeling was such that $c(\nu_2)$ will remain between $c(\nu_1)$ and $c(\nu_3)$. We maintain the fixed embedding of $G_r(\nu)$ if $c(\nu_1)$ remains before $c(\nu_3)$ after the permutation and flip $G_r(\nu)$ otherwise. If we flip $G_r(\nu)$, we also reverse all matrices for all non-leaf nodes in the subtree rooted at ν that are children of an R-node. Finally, we embed e to the right of $G_r^-(r)$ if the external column of M_r is on the right hand side of M_r and to the left otherwise.

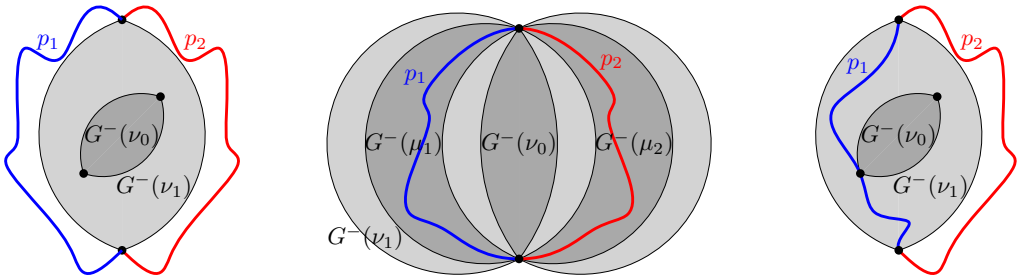


Figure 3.18: The three forbidden cases in the proof of Lemma 4

Lemma 4. *The thus constructed embedding of* $(G, \mathcal{C})$ *is c-planar.*

Proof. Let $C \in \mathcal{C}$ and let ν_1 be a non-inside node of T. We show by induction on the length of the ν_1-r-path that all external C-paths of ν_1 are on the same side and that no non-inside edge and no vertex in C_{ext} is enclosed by a simple cycle of non-outside edges in skel(ν_1) – provided that the root edge of skel(ν) is embedded on the same side as the external C-paths of ν_1. I.e., we have to consider the following three cases of pairs of C-paths between the poles of ν_1 (see Fig. 3.18).

(a) ν_1 **has two external** C**-paths** p_1 **and** p_2**.**

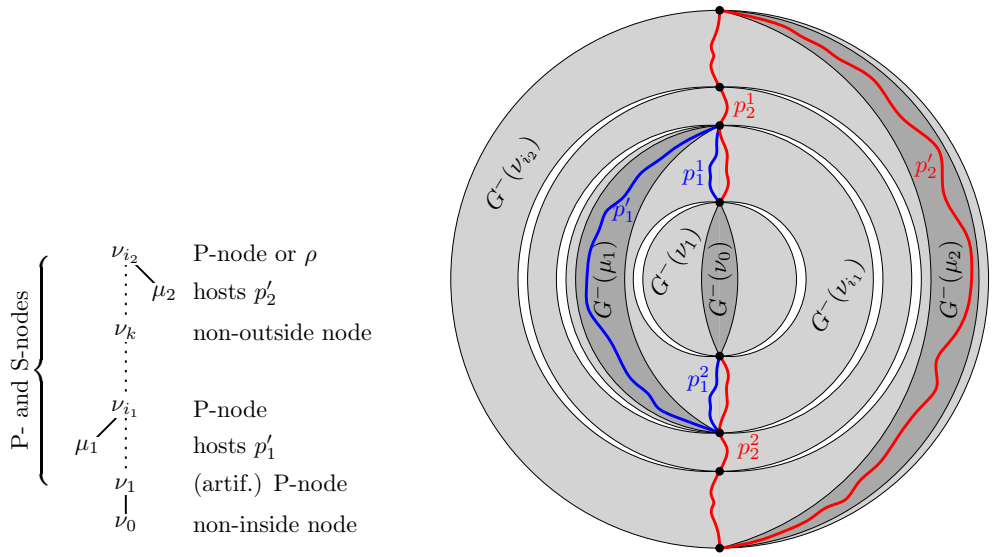
(b) **skel**$^-(\nu_1)$ **contains a cycle** c **of non-outside edges.** Observe that if c enclosed a non-inside edge or a vertex in C_{ext}, then ν_1 could not be an R-node: Otherwise c would be a caging cycle, and thus, ν_1 and all its ancestors, including the root, would be inappropriate. Thus, in this case we may assume that ν_1 is a P-node with two non-outside children μ_1, μ_2. Let p_j, $j = 1, 2$ be C-paths in $G_r^-(\mu_j)$ between its poles.

(c) **skel**(ν_1) **contains a cycle** c **of non-outside edges such that the root edge is contained in** c**.** Then there are the following two C-paths between the poles of ν_1: An external C-path p_2 of ν_1 and a path p_1 within $G_r^-(\nu_1)$ that goes exactly through the graphs represented by the children of ν_1 that correspond to the edges of c other than the root edge.

Let now e_0 be a non-inside edge of skel(ν_1), or let $v_0 \in C_{\text{ext}}$ be a vertex of skel(ν_1) other than the poles, and let e_0 be an edge of skel(ν_1) incident to v_0. Let ν_0 be the child of ν_1 corresponding to e_0. In all three cases, we have to show that the cycle composed by p_1 and p_2 does not enclose $G_r^-(\nu_0)$.

We will introduce a notation such that we can handle all three cases in one step. First, we will split the path p_j, $j = 1, 2$ into three segments, some of which might be empty. Figure 3.19 gives an illustration of the most general case.

Let $\nu_1, \dots, \nu_\ell = r$ be the ν_1-r-path. Let $j \in \{1, 2\}$. If p_j in $G^-(\nu_1)$, let $i_j = 1$. Otherwise, let $2 \leq i_j \leq \ell$ be minimum such that ν_{i_j} is an R-node or $G_r^-(\nu_{i_j})$ contains p_j. We may assume that $i_1 \leq i_2$. If ν_{i_j} is an R-node, we actually redefine ν_{i_j} to be the root ρ of the tree containing ν_{i_j-1}: We replace p_j by the respective path in $G_r(\nu_{i_j-1})$ through the root edge e_ρ of skel(ν_{i_j-1}). If ν_1 was an R-node, we



(a) Sub-SPQR-tree **(b)** Subgraph

Figure 3.19: Composition of the external paths in Lemma 4

redefine ν_1 to be the special P-node with which we replaced the R-node, and we redefine ν_0 to be one of the artificial non-inside Q-nodes we appended to ν_1.

Observe that ν_{i_j} is either ρ or a P-node and p_j is composed by two C-paths p_j^1 and p_j^2 connecting the poles of $G_r(\nu_1)$ with the poles of $G_r(\nu_{i_j})$ and a middle C-path p'_j. p_j^1 and p_j^2 are empty if $i_j = 1$. p'_j consists of the edge e_ρ if $\nu_{i_j} = \rho$. If ν_{i_j} is a P-node then it has a non-outside child $\mu_j \neq \nu_{i_j-1}$ such that p'_j is a path in $G_r^-(\mu_j)$ between its poles.

We distinguish some cases. (1) If $\nu_{i_1} = \rho$ or if $\nu_{i_1} = \nu_{i_2} \neq \rho$ and $\mu_1 = \mu_2$, then p_1 and p_2 are trivially on the same side of $G_r(\nu_1)$. (2) Assume that $\nu_{i_1} = \nu_{i_2} \neq \rho$ and $\mu_1 \neq \mu_2$. Since the 1s are consecutive in the rows inserted for ν_{i_j}, the two non-outside children μ_1 and μ_2 must be on the same side of the non-inside child ν_{i_j-1}. (3) Otherwise, observe that the C-paths p_2^1 and p_2^2, connecting the poles of $G_r(\nu)$ with the poles of $G_r(\nu_{i_2})$, must contain the poles of $G_r(\nu_k)$, $k = 1, \ldots, i_2$. (See Fig. 3.19b.) This implies especially that for each $k = i_1, \ldots, i_2$ the graphs $G_r^-(\nu_k)$ contain a C-path connecting their poles: Such a C-path can be composed by p'_1 and portions of p_2^1 and p_2^2. Hence, ν_k, $k = i_1, \ldots, i_2$ is not outside. Further, a subpath of p_2 is an external C-path of ν_{i_1}. Hence, ν_{i_1} cannot be double-border since otherwise the root would be inappropriate for C.

Since ν_{i_1-1} is non-inside and ν_{μ_1} is border or inside, there is a row κ inserted for ν_{i_1} that contains only 0s in $c(\nu_{i_1-1})$ and only 1s in $c(\mu_1)$. Further, when we handled ν_{i_2}, we added 1s in the row κ and the external column (if $\nu_{i_2} = r$) or the columns $c(\mu_2)$ (otherwise). Hence, since the 1s must be consecutive in κ, it follows that $c(\nu_{i_1-1})$ cannot be between $c(\mu_1)$ and the external column $c_C(\rho)$ or $c(\mu_2)$, respectively. Hence, p_1 and p_2 must be on the same side of $G_r(\nu_0)$.

Now, if $\nu_{i_2} \neq \rho$ we are done. Otherwise let ν be the parent R-node of ν_{i_2-1} in T. By

induction, we already know that all external C-paths of ν are on the same side and that $e_{\nu_{i_2-1}}$ is not enclosed by a simple cycle of non-outside edges in skel(ν). Hence, the external C-paths of ν_{i_2-1} are all on the same side, and by construction this is represented by the external column $c_C(\rho)$. □

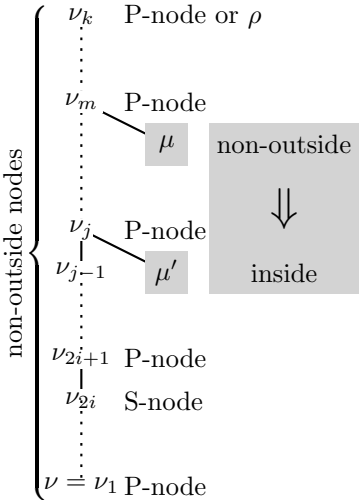
3.5.5 Proof of Necessity

Assume now that a c-planar embedding with e on the outer face is given in which C_{ext} is on the outer face of $G[C]$ for every cluster C. This yields a permutation of the children of the P-nodes of T and flips of the R-nodes. Permute the columns of the matrices accordingly. Let ρ be the root of a split off tree T_ρ, and let ρ' be the only child of ρ in T_ρ. The external columns of M_ρ are exchanged if on the ρ'-r-path there are an odd number of R-nodes that are flipped.

Lemma 5. *The thus constructed permutation of the columns is a consecutive-ones ordering up to a possible re-permutation due to the assignment of upper and lower halves for S-nodes.*

Proof. Recall that we have inserted up to three rows for each P-node and each cluster and no other rows into the matrices (except for the initialization). Let ν be a P-node in a subtree T_ρ, and let C be a cluster such that we have created a row κ for ν and C in M_ρ. Then ν has no double-border child. Due to c-planarity and the condition on all C_{ext}, the children of ν must be permuted such that all inside children are consecutive pre- and succeeded by at most one border child and arbitrary many outside children. It follows that the 1s in columns $c(\nu)$ must be consecutive.

Let $\nu = \nu_1, \dots, \nu_\ell = \rho$ be the path from ν to the root of T_ρ, and let $1 < k \leq \ell$ be maximum such that $\nu_1, \dots, \nu_k$ are not outside. If ν_1 was not a special P-node substituting an R-node, then ν_i is a P-node if i is odd and an S-node if i is even (otherwise it might be vice versa, but the situation is similar). ν_k is a P-node if $\nu_k \neq \rho$. Also observe that $c(\nu_{i-1}) \subseteq c(\nu_i)$, $i = 2, \dots, \ell$ and that for each $i = 1, \dots, \ell$ the columns in $c(\nu_i)$ are consecutive in the permuted matrix. If $k < \ell$, we have set $r(\nu_{k+1}, C) = \emptyset$. Hence, the entries in row κ are 0 in all columns in $c(\nu_{\ell-1}) \setminus c(\nu_k)$.



We consider first a P-node ν_i, $i = 3, \dots, k$ odd. Since ν_{i-1} is not outside, it follows that no child of ν_i other than ν_{i-1} can be double-border. Hence, for each non-outside child $\mu \neq \nu_{i-1}$ of ν_i there are 1s in row κ and all columns in $c(\mu)$.

Observe that due to c-planarity, the non-outside children of ν_i are consecutive. Moreover, if there are both, non-outside children of ν_i to the right and the left of ν_{i-1}, then ν_{i-1} is inside, and thus all columns in $c(\nu_1)$ as well as $c(\nu_j) \setminus c(\nu_{j-1})$ have entry 1 in row κ for all $3 \leq j < i$ odd.

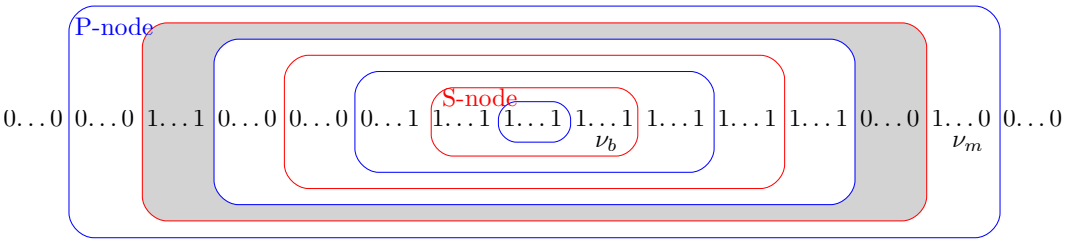


Figure 3.20: Possible ordering of the 0s and 1s in a row inserted for a P-node ν_1 according to a c-planar embedding

If $\nu_k = \rho$, let $m = k = \ell$ and assume that the external C-paths of $\nu_{\ell-1}$ are all to the right (left) of $\nu_{\ell-1}$, i.e., the column $c_C(\rho)$ is the right (left) external column. If $\nu_k \neq \rho$, let $m \leq k$ be maximum such that ν_m is a P-node that has a non-outside child other than ν_{m-1} (If no such P-node exists, then all entries in row κ other than in the columns $c(\nu)$ are zero and thus all 1s are consecutive.) Assume that ν_m has a non-outside child μ to the right (left) of ν_{m-1}. Assume now that there is a $1 \leq j < m$ odd such that the P-node ν_j has a child μ' to the right (left) of ν_{j-1}. I.e., the columns $c(\mu')$ are between the columns $c(\nu)$ and $c(\mu)$. If μ' was not inside, then $G_r(\mu')$ would contain a vertex in $V \setminus (C \setminus C_{\text{ext}})$ that would be enclosed by a C-cycle composed by the following four paths: (1) A C-path in $G_r^-(\nu_{j-1})$ between its poles, (2+3) two C-paths connecting the poles of $G_r(\nu_{m-1})$ with the poles of $G_r(\nu_j)$, and (4) either an external C-path of $\nu_{\ell-1}$, if $\nu_m = \rho$ or a C-path in $G_r(\mu)$ between its poles if ν_m is a P-node. Hence, the entries in $c(\mu')$ are all 1.

Consider now an S-node ν_i, $i = 2, \ldots, k$ even, that has an external C-path. By the choice of k, ν_i is not outside. Since the root is not inappropriate, ν_i is not double-border. Thus, we've set the entries in row κ and columns $c(\nu_i) \setminus c(\nu_{i-1})$ to 1 if ν_i is inside. Otherwise, we set the entries in $c(\nu_i) \setminus c(\nu_{i-1})$ that are to one side of $c(\nu_{i-1})$ to 1. Observe that an S-node ν_i has an external path if and only if $\ell = k$ or $i < m$.

Hence, row κ looks as follows: Assume without loss of generality that $c(\nu_m) \setminus c(\nu_{m-1})$ contains a 1 to the right of $c(\nu_{m-1})$. Then the entries in $c(\nu)$ are ordered such that all 0s (if any) are to the left and all 1s are to the right. Moreover, if ν is inside let $1 \leq b \leq k$ be maximal such that ν_b is inside. Then all entries in columns $c(\nu_b)$ are 1. Otherwise let $b = 1$.

For $i = b+1, \ldots, m-2$ odd all entries in $c(\nu_i) \setminus c(\nu_{i-1})$ that are on the right side of $c(\nu_{i-1})$ are 1. For $i = b+1, \ldots, m-1$ even, all entries in $c(\nu_i) \setminus c(\nu_{i-1})$ on one side of $c(\nu_{i-1})$ are 1 – however, for some i that could be the right-hand side and for others the left-hand side. Finally, the entries in $c(\nu_m) \setminus c(\nu_{m-1})$ to the right of $c(\nu_{m-1})$ are ordered such that the 1s are to the left and the 0s (if any) are to the right. See Fig. 3.20.

Hence, the 1s in row κ are consecutive up to maybe the wrong choice of the side for the 1s inserted for border S-nodes. However, observe that on one hand we could remove now the 1s from the wrong side and insert them on the right side and would thus obtain the 1s consecutive. We could obtain that for one cluster also by permuting the columns for the children of the S-node accordingly. On the other

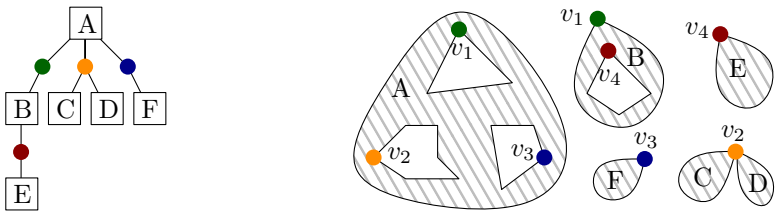
hand, the assignment to sides was forced by the same and different constraints – up to the choice for one cluster. Hence, if we do the permuting that works for one cluster, it will create the feasible assignment we would obtain if we had assigned the sides now that we know where the external paths are embedded. □

3.6 C-Connected Clusterings on General Graphs

We now show how to extend the method from the last section to work for an arbitrary planar graph. Let $(G, \mathcal{C})$ be a c-connected overlapping clustered graph with underlying planar graph G. If G is not connected, we can test each connected component separately since the c-connectivity limits each cluster to a single component.

It remains to show the case where G is connected but not 2-connected and can thus be represented by a BC-tree. We consider the BC-tree of G rooted at a block H_r. The interpretation of this choice is that H_r should contain an edge incident to the outer face of G in a planar drawing. Let H be a block of G. If $H \neq H_r$ then the *parent cut vertex* of H is the cut vertex of H on the path from H to H_r. H is a *child block* of its parent cut vertex. All other cut vertices of H are called *child cut vertices* of H. All cut vertices of H_r are *child cut vertices* of H_r.

Consider the SPQR-tree T of H. If $H = H_r$, any root of T is *suitable*. Otherwise the parent cut vertex of H must be on the outer face of H. Thus, a root of the SPQR-tree T is *suitable* if it corresponds to an edge incident to the parent cut vertex of H. See Fig. 3.21.



(a) Rooted BC-tree **(b)** The corresponding embedding

Figure 3.21: The hierarchy in the BC-tree is given by the choice of the root. The green vertex is the parent cut vertex of B while the red vertex is a child cut vertex of B. Note that the child blocks must be embedded with their parent cut vertices on the outer face

Let $H_1, \ldots, H_k$ be the child blocks of a cut vertex v and let V_i, $i = 1, \ldots, k$ be the set of vertices in the connected components of $G - v$ containing H_i. We call a cluster C *relevant* for a child block H_i if $v \in C$ and $V_i \not\subseteq C$. Let C_{ext} be the set of child cut vertices v of H such that C is relevant for a child block of v.

We use the algorithm for 2-connected graphs, restricting the roots for the SPQR-trees to be suitable, to test whether there is some c-planar embedding for each block H with the parent cut vertex on the outer face of H and C_{ext} on the outer face of $H[C]$. For each child cut vertex v of a block H and for each child block H_i of v, we test whether there is a free face, i.e., a face f of H incident to v such that

the boundary of f contains a vertex not in C for any cluster C that is relevant for H_i. If so, the c-planar embeddings of the blocks can be combined into a c-planar embedding of the whole graph. In the following, we show that otherwise there is no c-planar drawing for the whole graph with the given choices of the root of the BC-tree and the roots of the SPQR-trees.

Given a c-planar embedding, a face f is *free* with respect to a subset $\mathcal{C}' \subseteq \mathcal{C}$ of clusters if f is not enclosed by a C-cycle for any $C \in \mathcal{C}'$. Otherwise, f is *covered* by $\mathcal{C}'$. In the following we use C instead of $\{C\}$ if the context is clear.

Remark 3. *A face is covered by C if and only if its boundary is a C-cycle.*

A vertex v is *free* with respect to a subset $\mathcal{C}' \subseteq \mathcal{C}$ of clusters if one of its incident faces is free with respect to $\mathcal{C}'$ and *covered* by $\mathcal{C}'$ otherwise. Our goal is to prove the following lemma.

Lemma 6. *Let H be a block, v a vertex in H, and $\mathcal{C}' \subseteq \mathcal{C}$. If there is a c-planar embedding of H in which v is free with respect to $\mathcal{C}'$, then v is free with respect to $\mathcal{C}'$ in any c-planar embedding of H in which v is free with respect to C for all $C \in \mathcal{C}'$.*

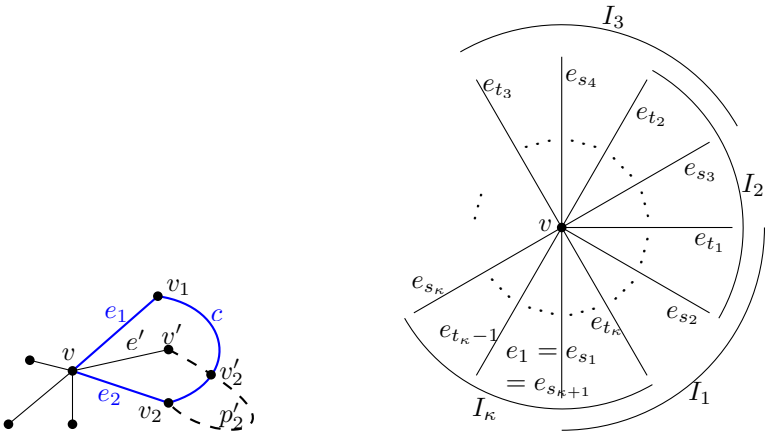
Given a vertex v in a block H, we call two incident edges e_1 and e_2 of v *equivalent* with respect to a set $\mathcal{C}'$ of clusters if they are in the same block of $H[\bigcap_{C\in\mathcal{C}'} C]$, i.e., if there is a simple cycle in H that is a C-cycle for any $C \in \mathcal{C}'$ and contains both, e_1 and e_2. A *$\mathcal{C}'$-equivalence class* around v is a maximal set of edges incident to v that are pairwise equivalent with respect to $\mathcal{C}'$.

Lemma 7. *Let v be a vertex of a block H and let $\mathcal{C}' \subseteq \mathcal{C}$. Then a $\mathcal{C}'$-equivalence class around v is a consecutive set of edges in the cyclic order around v in any c-planar embedding of H.*

Proof. Let $\mathcal{C}'$ be a set of clusters, v a vertex in block H, and let $e_i = \{v, v_i\}$, $i = 1, 2$ be two edges incident to v that are contained in a simple cycle c in $H[\bigcap_{C\in\mathcal{C}'} C]$. See Fig. 3.22a. Then all vertices that are enclosed by c are in $\bigcap_{C\in\mathcal{C}'} C$. Let $e' = \{v, v'\}$ be an edge enclosed by c. Let $i \in \{1, 2\}$. Since H is 2-connected, there must be a v'-v_i path p_i in H not containing v. Let v_i' be the first vertex of p_i on c. Let c_i' be the cycle formed by the v'-v_i'-subpath of p_i, the v_i'-v-subpath of c containing e_i, and the edge e'. Then c_i' is a simple cycle in $H[\bigcap_{C\in\mathcal{C}'} C]$ containing e' and e_i. □

Let v be a vertex that is free with respect to any $C \in \mathcal{C}'$. Given a c-planar embedding of a block, a *$\mathcal{C}'$-interval* around a vertex v is a maximal sequence of consecutive edges around v that are (a) equivalent with respect to $\mathcal{C}'$ and such that (b) the face between any two consecutive edges is covered by C for all $C \in \mathcal{C}'$. Note that there is a one-to-one correspondence between the $\mathcal{C}'$-equivalence classes and the $\mathcal{C}'$-intervals around v: The condition that v is free with respect to any cluster in $\mathcal{C}'$ guarantees that the $\mathcal{C}'$-intervals have a well defined start and end point. Also note that there might be several distinct $\mathcal{C}'$-intervals around v – even if $\mathcal{C}'$ contains only one cluster.

Proof. (of Lemma 6) Assume that there is a c-planar embedding of H in which v is free with respect to C for all $C \in \mathcal{C}'$ but v is not free with respect to $\mathcal{C}'$. Consider the cyclic order $e_1, \ldots, e_\ell$ of the edges around v. Since v is not covered by any $C \in \mathcal{C}'$,



(a) $e_1 \sim e_2 \Rightarrow e_1 \sim e' \sim e_2$ **(b)** Minimal set of intervals covering v

Figure 3.22: Illustration of the proofs of (a) Lemma 7 and (b) Lemma 6. (The symbol $\sim$ refers to equivalence with respect to the set $\mathcal{C}'$)

the C-intervals around v are well defined. Among all C-intervals for all $C \in \mathcal{C}'$, let $\mathcal{I}$ be a minimal set of intervals such that all faces around v are covered by at least one interval in $\mathcal{I}$. Let $I_i = \langle e_{s_i}, \dots, e_{t_i} \rangle, i = 1, \dots, \kappa$ be the intervals in $\mathcal{I}$ in cyclic order around v. See Fig. 3.22b. We assume that $s_1 = 1$, $s_i < t_i$ for $i = 1, \dots, \kappa - 1$, and $t_\kappa > s_\kappa$. For simplicity, we set $s_{\kappa+1} := s_1$. Let $C_i \in \mathcal{C}'$ be such that I_i is a C_i-interval.

Since all faces around v are covered, it holds that $s_{i+1} \leq t_i$. $\langle e_{s_{i+1}}, \dots, e_{t_i} \rangle$ is a $\{C_i, C_{i+1}\}$-interval. Let $s_{i+1} \leq j < t_i$. The face f between e_j and e_{j+1} is covered by both, C_i and C_{i+1}. Hence, the boundary of f is a both a C_i- and a C_{i+1}-cycle. Thus $\{e_{s_{i+1}}, \dots, e_{t_i}\}$ is consecutive in any c-planar embedding. Hence, in any c-planar embedding, the ordering of edges around v is a sequence of overlapping C-intervals for several $C \in \mathcal{C}'$. Hence, for any face f incident to v there is at least one $C \in \mathcal{C}'$ such that f is covered by C. Therefore v cannot be free with respect to $\mathcal{C}'$ in any c-planar embedding. □

We now apply Lemma 6 to any child cut vertex v of any block H and to the set $\mathcal{C}'$ of relevant clusters of any child block of v to obtain our main result. Observe that the particular choice of C_{ext} in the following theorem guarantees that the child cut vertices are free with respect to any relevant cluster.

We call a cut vertex v *free* for a child block H_i if v is free with respect to the set of clusters that are relevant for H_i.

Theorem 5. *A c-connected overlapping clustered graph $(G, \mathcal{C})$ is c-planar if and only if G is planar, and for each connected component of G, there is a root block of its BC-tree for which there exist suitable root nodes of the SPQR-tree of each block that are not inappropriate for any $C \in \mathcal{C}$ with $C_{ext} = \{v;\ v$ child cut vertex and C relevant for a child block of $v\}$ such that*

1. *all binary matrices fulfill the consecutive-ones property and*

2. *given an arbitrary consecutive-ones ordering of the binary matrices, each cut vertex is free for each of its child blocks in the corresponding embedding.*

Proof. Assume that the two conditions hold. We embed the blocks as in the proof of Theorem 4 and combine the embeddings of the blocks as follows. Let H be a block, let v be a child cut vertex of H, and let H_i be a child block of v. We place H_i into a face of H incident to v that is free with respect to the set of H_i's relevant clusters. This yields a c-planar embedding of G:

Otherwise, there must be a cluster C and a vertex $w \in V \setminus C$ such that w is enclosed by a C-cycle c. Let c be in block H. The first condition requires, that the embedding of H is c-planar (see Theorem 4). Hence, w cannot be a vertex of H. Let v' be the parent cut vertex of H and let V' be the union of the sets of vertices in the connected components of $G - v'$ not containing H. By the choice of the root of the BC-tree, V' must be drawn in the outer face of H. Hence $w \notin V'$.

Finally, let v be a child cut vertex of H, let H_i be a child block of v, let V_i be the set of vertices in the connected components of $G - v$ containing H_i, and assume that $w \in V_i$. Then v must be enclosed by c and thus, by c-planarity of H, $v \in C$. Since $w \notin C$, it follows that C is relevant for H_i. Since we embedded H_i into a face of H that was free with respect to H_i's relevant clusters, it follows that w cannot be enclosed by the C-cycle c.

For the other direction assume now that there is a c-planar embedding $\mathcal{E}$. Without loss of generality, we assume that G is connected. Let the root H_r of the BC-tree be a block with an edge that is incident to the outer face of G. Root each SPQR-tree at an edge incident to the outer face of the respective block and incident to the parent cut vertex.

In a c-planar drawing, a child cut vertex v of a block H is placed on the outer face of $H[C]$ for any relevant cluster C of any of v's child blocks. Thus, Theorem 4 implies that the roots are not inappropriate and Condition 1 is fulfilled.

Obviously any block must be inserted into a face that is free with respect to its relevant clusters in any c-planar embedding of G. Consider now a block H and an embedding $\mathcal{E}'$ of H corresponding to an arbitrary consecutive-ones ordering of the binary matrices. Let v be a child cut vertex of H, and let H_i be a child block of v. Let $\mathcal{C}'$ be the set of relevant clusters for H_i.

The labeling guarantees that v is on the outer face of $H[C]$ for any $C \in \mathcal{C}'$. Thus, $\mathcal{E}'$ is a c-planar embedding of H in which v is free with respect to each $C \in \mathcal{C}'$. We further know that $\mathcal{E}$ induces a c-planar embedding of H in which v is free with respect to $\mathcal{C}'$. Hence, Lemma 6 implies that v is free with respect to $\mathcal{C}'$ in $\mathcal{E}'$. □

The characterization in the previous theorem immediately yields the following corollary.

Corollary 1. *It can be tested in polynomial time whether a c-connected overlapping clustered graph is c-planar.*

Proof. We consider the connected components independently. For each choice of the root of the BC-tree and for each cluster C, we first proceed bottom up in the BC-tree in order to compute for each node ν of the BC-tree the sets $C_{\text{ext}} = \{v;\ v$ child cut vertex and C relevant for a child block of $v\}$ in overall $\mathcal{O}(|V|)$ time. Then, for any choice of a suitable root of the SPQR-tree of each block, we have to construct the respective matrices and check whether they have the consecutive-ones property. The size of the matrices is in $\mathcal{O}(|V|^2|\mathcal{C}|)$. The bottle-neck of constructing them is the handling of the S-nodes which can be done in $\mathcal{O}(|V| \cdot |\mathcal{C}|^2)$ time per S-node and thus in overall $\mathcal{O}(|V|^2|\mathcal{C}|^2)$ time. Finally, we can check whether the cut vertices are free, by simply traversing the incident faces.

Since there are $\mathcal{O}(|V|)$ choices of the root of the BC-tree and for each SPQR-tree $\mathcal{O}(|V|)$ choices of their roots, we can test in $\mathcal{O}(|V|^4|\mathcal{C}|^2)$ time whether an overlapping clustered graph is c-planar. □

3.7 Conclusion

In this part of the thesis, we showed that the problem of clustered planarity of overlapping clustered graphs remains NP-complete when the instances are limited to clusterings consisting of two partitions (Section 3.3). The proof even holds if the underlying graph is 2-connected, each cluster contains at most two connected components, and the complement of each cluster contains at most three connected components.

In contrast to this, we could show that the problem with two partitions is solvable in linear time if each cluster as well as its complement induce a connected subgraph (Section 3.4).

In Section 3.5 and 3.6, we showed that the problem with general (overlapping) clusterings can be solved in polynomial time if the graphs are c-connected, thus solving an open problem stated by Didimo et al. (2008).

Unfortunately, the asymptotic run time of our algorithm is far from being efficient. For special instances of the problem, there exist efficient algorithms, like the cases discussed by Didimo et al. (2008) or the cases of two c-co-connected partitions discussed in Section 3.4. These approaches are based on reducing c-planarity of special subclasses of overlapping clustered graphs to c-planarity of sufficiently connected hierarchically clustered graphs. Then an efficient algorithm for this case is used, e.g., the algorithm of Cortese et al. (2008).

There are two main reasons why we can not (yet) match the linear run time of this algorithm with our approach: (1) Only one choice of the root of the SPQR-tree has to be considered in the hierarchical case – namely an edge that is not contained in any cluster (other than the whole vertex set), and (2) the characterization of Cortese et al. (2008) can make use of the interplay between the SPQR-tree and the rooted tree representing the cluster hierarchy. It would be interesting whether these ideas could be extended to our case as an interplay between the SPQR-tree and the Hasse diagram.

As we have so far only provided a theoretical description of the algorithm in Section 3.6, a logical next step would be the implementation of the method for practical use in the context of a visualization package.

Even though the problem described in this part of the thesis is of a rather theoretical nature, it has a variety of potential applications to archaeological data as suggested by the example in the introduction. The understanding of such fundamental combinatorial problems sets the foundation for more applied visualization techniques such as the ones implemented in various visualization software. For an overview of potential applications of state-of-the-art visualization methods to archaeological data, we refer to the PhD thesis by Mereke van Garderen (2018).

Conclusion

In the course of this thesis, we explored a variety of methods for analyzing and visualizing archaeological network data. We presented approaches for the reconstruction of travel distances, the recording of similarities based on human perception, and the embedding of clustered networks. A reflection on the advantages and disadvantages of the specific methods can be found at the end of each of the three parts. Therefore, this general conclusion is limited to a more general discussion about the methodologies used in this work.

- In Part 1, we developed methods for the simulation of (optimal) canoe travel between the islands of the Caribbean based on environmental factors. We explored ways to analyze the temporal variation of current, wind, and waves, and presented and evaluated methods of finding the fastest routes based on these factors. In a user study, we examined the ability of human actors to learn about current patterns and approximate the optimal trajectories.

  The methods we developed have the potential to infer (possible) relationships between sites on different islands based on the ease (or difficulty) of access for indigenous populations. In addition, they can hint at possible stopover locations on pathways between locations, thus potentially indicating undiscovered sites on other islands.

- In Part 2, we explored possibilities to incorporate human perception into the classification of pottery. We developed two methods of knowledge quantification and put them – together with established methods – to the test in three case studies. In the first study, we compared two methods for recording expert knowledge about distances (or similarities) between the possible values of attributes in a pottery database. The second study consisted of an online survey with the goal to quantify the relative importance of different attributes. In the third study, we evaluated differences in the perception of sherds between archaeologists and (indigenous) potters.

  With the help of the methods from this part, archaeologists have the opportunity to create a typology of ceramics that incorporates human perception and knowledge. This knowledge could either come from (indigenous) potters, leading to an approach to classification that incorporates the point of view of the people who made the pots, or from experts in classification, leading to a way of classifying pottery that is reproducible also for non-experts.

  The resulting classifications can help understanding the evolution of cultural

practices or determining the age of styles through seriation. In addition, they can serve as a foundation to establishing networks of similarity between sites based on the number of finds of certain types (Habiba et al. 2018).

- In Part 3 of this thesis, we set the focus on the visualization of network data. We examined the NP-complete problem of overlapping clustered planarity. We show that the problem remains NP-complete, even when restricted to a clustering that is the union of two partitions. In contrast, if the clustering consists of two partitions in which each cluster and its complement induce connected subgraphs, we could provide a linear-time algorithm. For the case of general (overlapping) clusterings, we present a polynomial-time algorithm that works under the condition that each cluster induces a connected subgraph.

  Even though the problems we examined are of a rather theoretical nature, they have the potential to facilitate the interpretation and presentation of archaeological network data. This is relevant not only for researchers, but also for community outreach projects in heritage management, such as museum exhibitions or school books.

In this thesis, we combined approaches from a variety of fields to develop methods that can be applied to research questions in an archaeological context. As the thesis is written from a network science perspective, we set our focus more on the development and validation of the methodologies and less on their application in archaeological studies. Yet we still put them to the test through examples and case studies to evaluate their effectiveness and applicability to archaeological data.

Since archaeology is a field that deals with the study of (past) human behavior, we found it appropriate to include the "human factor" into our methods whenever possible. Even though there is no guarantee that humans in the past had the same cultural perceptions as modern humans, we still considered it important to integrate their knowledge and perception into archaeological studies instead of relying purely on a data-driven approach. The canoe game in Part 1 gave us the possibility to analyze human agency to an extent that would be very hard to replicate with traditional methods of experimental archaeology. The methods described in Part 2 provide a way to quantify perception and knowledge that can be used to create more human-centered typologies of objects. As the methods are not limited to archaeology, they can be a valuable addition to the toolbox of cultural domain analysis.

As mentioned at the end of each part, there are many open questions that would justify an extension of the proposed methods and provide opportunities for more extensive case studies. We hope that through this work we could inspire researchers from archaeology and other disciplines to integrate some of our ideas into their own work and that our methods can be of use in finding answers to their research questions.

Bibliography

Allen, A. A. (2005). Leeway divergence. Technical report, Coast Guard Research and Development Center, Gronton CT.

Allen, A. A., Robe, R. Q., and Morton, E. (1999). The Leeway of Persons-In-Water and Three Small Craft. Technical report, Analysis and Technology INC, North Stonington CT.

Altes, C. F. (2011). A Brief Note on Currents, Current Archaeologists, and Ancient Fiber-Tempered Pots. *The Florida Anthropologist*, 64(2):113–118.

Angelini, P., Da Lozzo, G., Di Battista, G., Frati, F., Patrignani, M., and Roselli, V. (2015). Relaxing the constraints of clustered planarity. *Computational Geometry: Theory and Applications*, 48(2):42–75.

Athenstädt, J. C., Borck, L., Brandes, U., and Hofman, C. L. (2018). Plainware and Polychrome: How Do Ceramic Perceptions Differ between Communities? Presentation at the 83rd Annual meeting of the Society for American Archaeology, Washington, D.C.

Athenstädt, J. C. and Cornelsen, S. (2017). Planarity of Overlapping Clusterings Including Unions of Two Partitions. *Journal of Graph Algorithms and Applications*, 21(6):1057–1089.

Athenstädt, J. C., Hartmann, T., and Nöllenburg, M. (2014). Simultaneous Embeddability of Two Partitions. In Duncan, C. and Symvonis, A., editors, *Proceedings of the 22nd International Symposium on Graph Drawing (GD 2014)*, volume 8871 of *Lecture Notes in Computer Science*, pages 64–75. Springer.

Athenstädt, J. C., Jacobson, K., Hofman, C. L., and Brandes, U. (2016). Pottery Classification with Triad Tests. Presentation at the XXXVI Sunbelt Conference, Newport Beach.

Bahn, P. (2012). *Archaeology: A Very Short Introduction*. OUP Oxford.

Barbotin, M. (1974). *Archéologie caraïbe et chroniqueurs*. M. Barbotin.

Beckles, H. (2006). *A History of Barbados: From Amerindian Settlement to Caribbean Single Market*. Cambridge University Press.

Bentley, J. L. and Ottmann, T. A. (1979). Algorithms for reporting and counting geometric intersections. *IEEE Transactions on computers*, C-28(9):643–647.

Bérard, B., Billard, J.-Y., L'Etang, T., Lalubie, G., Nicolizas, C., Ramstein, B., and Slayton, E. (2016). Approche expérimentale de la navigation précolombienne dans les Antilles. *Journal de la société des américanistes*, 102(102-2):171–204.

Bérard, B., Billard, J.-Y., and Ramstein, B. (2009). Ioumoulicou. In *XXIII Congress of the International Association for Caribbean Archaeology*, pages 577–589.

de Berg, M. and Khosravi, A. (2010). Optimal Binary Space Partitions in the Plane. In Thai, M. and Sahni, S., editors, *Computing and Combinatorics (COCOON'10)*, volume 6196 of *Lecture Notes in Computer Science*, pages 216–225. Springer-Verlag.

van Bevern, R., Kanj, I. A., Komusiewicz, C., Niedermeier, R., and Sorge, M. (2016). Twins in Subdivision Drawings of Hypergraphs. In Hu, Y. and Nöllenburg, M., editors, *Proceedings of the 24th International Symposium on Graph Drawing (GD 2016)*, volume 9801 of *Lecture Notes in Computer Science*, pages 67–80. Springer.

Billard, J.-Y. and Bérard, B. (2011). Apport de l'hydrostatique à l'archéologie expérimentale: Etude d'une pirogue de haute mer (Kanawa). *Congrès Français de Mécanique*.

Bixby, R. E. and Wagner, D. K. (1988). An Almost Linear-Time Algorithm for Graph Realization. *Mathematics of Operations Research*, 13(1):99–123.

Bläsius, T. and Rutter, I. (2016). A new perspective on clustered planarity as a combinatorial embedding problem. *Theoretical Computer Science*, 609(P2):306–315.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008.

Boeck, F., Hochkirch, K., Hansen, H., Norris, S., and Flay, R. (2012). Side Force Generation of Slender Hulls - Influencing Polynesian Canoe Performance. *4th High Performance Yacht Design Conference, At Auckland, New Zealand*.

Boomert, A. and Bright, A. J. (2007). Island archaeology: in search of a new horizon. *Island Studies Journal*, 2(1):3–26.

Booth, K. S. and Lueker, G. S. (1976). Testing for the Consecutive Ones Property, Interval Graphs, and Graph Planarity Using PQ-Tree Algorithms. *Journal of Computer and System Sciences*, 13:335–379.

Borgatti, S. P. (1994). Cultural domain analysis. *Journal of Quantitative Anthropology*, 4(4):261–278.

Borgatti, S. P. (1999). Elicitation techniques for cultural domain analysis. *Enhanced ethnographic methods*, 3:115–151.

Boster, J. (1994). The Successive Pile Sort. *Cultural Anthropology Methods*, 6(2):11–12.

Boster, J. S. and Johnson, J. C. (1989). Form or function: A comparison of expert and novice judgments of similarity among fish. *American Anthropologist*, 91(4):866–889.

Bowditch, N. (2002). *American practical navigator: an epitome of navigation*. National Imagery and Mapping Agency.

Brainerd, G. W. (1951). The place of chronological ordering in archaeological analysis. *American Antiquity*, 16:301–3.

Brandes, U., Cornelsen, S., Pampel, B., and Sallaberry, A. (2011). Blocks of Hypergraphs applied to Hypergraphs and Outerplanarity. In Iliopoulos, C. and Smyth, W., editors, *Proceedings of the 21st International Workshop on Combinatorial Algorithms (IWOCA 2010)*, volume 6460 of *Lecture Notes in Computer Science*, pages 201–211. Springer.

Brandes, U., Cornelsen, S., Pampel, B., and Sallaberry, A. (2012). Path-Based Supports for Hypergraphs. *Journal of Discrete Algorithms*, 14:248–261.

Brandes, U., Delling, D., Gaertler, M., Görke, R., Hoefer, M., Nikoloski, Z., and Wagner, D. (2007). On finding graph clusterings with maximum modularity. In *Graph-Theoretic Concepts in Computer Science*, pages 121–132. Springer.

Brandes, U., Robins, G., McCranie, A., and Wasserman, S. (2013). What is network science? *Network Science*, 1(1):1–15.

Brandes, U. and Wagner, D. (2003). Visone – Analysis and Visualization of Social Networks. In *Graph Drawing Software*, pages 321–340. Springer-Verlag.

Breivik, Ø., Allen, A. A., Maisondieu, C., and Roth, J. C. (2011). Wind-induced drift of objects at sea: The leeway field method. *Applied Ocean Research*, 33(2):100–109.

Brughmans, T. (2010). Connecting the dots: towards archaeological network analysis. *Oxford Journal of Archaeology*, 29(3):277–303.

Brushett, B. A., Allen, A. A., Futch, V. C., King, B. A., and Lemckert, C. J. (2014). Determining the leeway drift characteristics of tropical Pacific island craft. *Applied Ocean Research*, 44:92 – 101.

Buchin, K., van Kreveld, M., Meijer, H., Speckmann, B., and Verbeek, K. (2011). On Planar Supports for Hypergraphs. *Journal on Graph Algorithms and Applications*, 15(4):533–549.

Burton, M. L. and Nerlove, S. B. (1976). Balanced designs for triads tests: Two examples from English. *Social Science Research*, 5(3):247–267.

Callaghan, R. (2003a). Prehistoric trade between Ecuador and West Mexico: A computer simulation of coastal voyages. *Antiquity*, 77(298):796–804.

Callaghan, R. and Schwabe, S. J. (2001). Watercraft of the islands. In *Proceedings of the XVIII congress of the international association for Caribbean archaeology. IACA Martinique, St. Georges, Grenada*, volume 1999, pages 231–242.

Callaghan, R. T. (2001). Ceramic age seafaring and interaction potential in the Antilles: a computer simulation. *Current Anthropology*, 42(2):308–313.

Callaghan, R. T. (2003b). Comments on the Mainland Origins of the Preceramic Cultures of the Greater Antilles. *Latin American Antiquity*, 14(3):323–338.

Callaghan, R. T. and Fitzpatrick, S. M. (2008). Examining Prehistoric Migration Patterns in the Palauan Archipelago: A Computer Simulated Analysis of Drift Voyaging. *Asian Perspectives*, 47(1):28–44.

Chaplick, S., Jelínek, V., Kratochvíl, J., and Vyskočil, T. (2012). Bend-bounded path intersection graphs: Sausages, noodles, and waffles on a grill. In *Graph-Theoretic Concepts in Computer Science*, pages 274–285. Springer.

Colbourn, C. J. and Lindner, C. C. (1992). Support sizes of triple systems. *Journal of Combinatorial Theory, Series A*, 61(2):193–210.

Cook, S. A. (1971). The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, pages 151–158. ACM.

Cooper, J. (2010). Modeling Mobility and Exchange in pre-Columbian Cuba: GIS led Approaches to Identifying Pathways and Reconstructing Journeys from the Archaeological Record. *Journal of Caribbean Archaeology*, Special Publication #3:122–137.

Cooper, J. and Peros, M. (2010). The archaeology of climate change in the Caribbean. *Journal of Archaeological Science*, 37(6):1226 – 1232.

Cornelsen, S. and Wagner, D. (2006). Completely Connected Clustered Graphs. *Journal of Discrete Algorithms*, 4(2):313–323.

Cortese, P. F., Di Battista, G., Frati, F., Patrignani, M., and Pizzonia, M. (2008). C-Planarity of C-Connected Clustered Graphs. *Journal on Graph Algorithms and Applications*, 12(2):225–262.

Cote, A. and Raz, J. G. (2015). In-depth interviews for games research. In *Game Research Methods*, pages 93–116. ETC Press.

Criminisi, A., Reid, I., and Zisserman, A. (1999). A plane measuring device. *Image and Vision Computing*, 17(8):625 – 634.

Dahlhaus, E. (1998). A Linear Time Algorithm to Recognize Clustered Planar Graphs and its Parallelization. In Lucchesi, C. L. and Moura, A. V., editors, *Proceedings of the 3rd Latin American Symposium on Theoretical Informatics (LATIN '98)*, volume 1380 of *Lecture Notes in Computer Science*, pages 239–248. Springer.

Dean, D. (2000). *Statistical shape analysis*. Elsevier.

Dehon, M. (1983). On the existence of 2-designs S λ (2, 3, υ) without repeated blocks. *Discrete Mathematics*, 43(2):155–171.

Di Battista, G. and Tamassia, R. (1990). On-Line Graph Algorithms with SPQR-Trees. In Paterson, M., editor, *Proceedings of the 17th International Colloquium on Automata, Languages and Programming*, volume 443 of *Lecture Notes in Computer Science*, pages 598–611. Springer.

Didimo, W., Giordano, F., and Liotta, G. (2008). Overlapping Cluster Planarity. *Journal on Graph Algorithms and Applications*, 12(3):267–291.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271.

Djaouti, D., Alvarez, J., Jessel, J.-P., and Rampnoux, O. (2011). Origins of serious games. In *Serious games and edutainment applications*, pages 25–43. Springer.

Eiter, T. and Mannila, H. (1994). Computing discrete Fréchet distance. Technical report, CD-TR 94/64, Information Systems Department, Technical University of Vienna.

Faber Morse, B. (1989). Saladoid settlement patterns on St. Croix. *Early Ceramic Population Lifeways and Adaptive Strategies in the Caribbean, BAR International Series*, 506:29–42.

Faulkner, F. D. (1962). Determining Optimum Ship Routes. *Operations Research*, 10(6):799 – 807.

Feng, Q., Cohen, R. F., and Eades, P. (1995). Planarity for Clustered Graphs. In Spirakis, P., editor, *Proceedings of the 3rd European Symposium on Algorithms (ESA '95)*, volume 979 of *Lecture Notes in Computer Science*, pages 213–226. Springer.

Fitzpatrick, S. M. (2013). Seafaring capabilities in the pre-Columbian Caribbean. *Journal of Maritime Archaeology*, 8(1):101–138.

Fleet Numerical Meteorology and Oceanography Center (1995). U.S. Navy marine climatic atlas of the world : including Mediterranean severe weather port guides, Ver. 1.0 .

Forbes, R. (1993). *Studies in ancient technology 2*. E.J. Brill.

Ford, J. A. (1954). On the Concept of Types. *American Anthropologist*, 56(1):42–57.

Fritz, S. C., Björck, S., Rigsby, C. A., Baker, P. A., Calder-Church, A., and Conley, D. J. (2011). Caribbean hydrological variability during the Holocene as reconstructed from crater lakes on the island of Grenada. *Journal of Quaternary Science*, 26(8):829–838.

van Garderen, M. (2018). *Pictures of the Past – Visualization and visual analysis in archaeological context.* PhD thesis, University of Konstanz.

Garey, M. R. and Johnson, D. S. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman & Co.

Gilboa, A., Karasik, A., Sharon, I., and Smilansky, U. (2004). Towards computerized typology and classification of ceramics. *Journal of Archaeological Science*, 31(6):681 – 694.

Giry, C., Felis, T., Kölling, M., Wei, W., Lohmann, G., and Scheffers, S. (2013). Controls of Caribbean surface hydrology during the mid- to late Holocene: insights from monthly resolved coral records. *Climate of the Past*, 9(2):841–858.

Goldstone, R. (1994). An efficient method for obtaining similarity data. *Behavior Research Methods, Instruments, and Computers*, 26:381–386.

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*, pages 857–871.

Greene, K. and Moore, T. (2010). *Archaeology: An Introduction.* Taylor & Francis.

Gutwenger, C. and Mutzel, P. (2000). A Linear Time Implementation of SPQR-Trees. In Goodrich, M. T. and Kobourov, S. G., editors, *Proceedings of the 10th International Symposium on Graph Drawing (GD 2002)*, volume 2528 of *Lecture Notes in Computer Science*, pages 77–90. Springer.

Gyory, J., Mariano, A. J., and Ryan, E. H. (2017). The Caribbean Current. `http://oceancurrents.rsmas.miami.edu/caribbean/caribbean.html`. Accessed: 2017-01-17.

Habiba, Athenstädt, J. C., and Brandes, U. (2014). Inferring Social Dynamics from Spatio-Temporal Network Data in the US Southwest. Presentation at the Connected Past conference, Paris.

Habiba, Athenstädt, J. C., Mills, B. J., and Brandes, U. (2018). Social Networks and Similarity of Site Assemblages. *Journal of Archaeological Science.* To appear.

Hackett, B., Breivik, Ø., and Wettre, C. (2004). Forecasting the drift of things in the oceans. *Proceedings of the Second Symposium on the Global Ocean Data Assimilation Experiment.*

Hagiwara, H. (1989). *Weather routing of (sail-assisted) motor vessels.* PhD thesis, Technische Universiteit Delft.

Haltiner, G., Hamilton, H., and 'Arnason, G. (1962). Minimal-time ship routing. *Journal of Applied Meteorology*, 1(1):1–7.

Hanssen, G. L. and James, R. W. (1960). Optimum ship routing. *Journal of Navigation*, 13(03):253–272.

Hart, P. E., Nilsson, N. J., and Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107.

Haug, G. H., Hughen, K. A., Sigman, D. M., Peterson, L. C., and Röhl, U. (2001). Southward Migration of the Intertropical Convergence Zone Through the Holocene. *Science*, 293(5533):1304.

Hearn, C. G. (2003). *Tracks in the Sea: Matthew Fontaine Maury and the Mapping of the Oceans.* Mcgraw Hill Book Co.

Henson, D. (2012). *Doing Archaeology: A Subject Guide for Students.* Doing series. Routledge.

Hildenbrand, J. (2015). Shortest Path Calculation on Water Surfaces. Master's thesis, University of Konstanz, Germany.

Hinnenthal, J. and Clauss, G. (2010). Robust Pareto-optimum routing of ships utilising deterministic and ensemble weather forecasts. *Ships and Offshore Structures*, 5(2):105–114.

Hofman, C. L. (1993). *In search of the native population of pre-Columbian Saba (400-1450 A.D.)*. Rijksuniversiteit te Leiden.

Hofman, C. L., Bright, A. J., Boomert, A., and Knippenberg, S. (2007). Island rhythms: the web of social relationships and interaction networks in the Lesser Antillean archipelago between 400 BC and AD 1492. *Latin American Antiquity*, 18(3):243–268.

Hofman, C. L., Bright, A. J., and Ramos, R. R. (2010). Crossing the Caribbean Sea: towards a holistic view of pre-colonial mobility and exchange. *Journal of Caribbean Archaeology*, 10:1–18.

Hofman, C. L., Hoogland, M. L. P., and van Gijn, A. L., editors (2008). *Crossing the Borders: New Methods and Techniques in the Study of Archaeology Materials from the Caribbean*. Caribbean Archaeology and Ethnohistory. University of Alabama Press.

Horejs, B., Jung, R., and Pavúk, P. (2010). *Analysing Pottery: Processing, Classification, Publication*. Studia archaeologica et medievalia. Comenius University in Bratislava.

Hörr, C., Lindinger, E., and Brunnett, G. (2008). New Paradigms for Automated Classification of Pottery. In *Proceedings of the 36th CAA Conference*, pages 268–277.

Hout, M. C. and Goldinger, S. D. (2016). Spam is convenient but also satisfying: Reply to verheyen et al.(2016). *Journal of Experimental Psychology: General*.

Hout, M. C., Goldinger, S. D., and Ferguson, R. W. (2013). The versatility of spam: A fast, efficient, spatial method of data collection for multidimensional scaling. *Journal of Experimental Psychology: General*, 142(1):256.

Hulme, P. and Whitehead, N. L. (1992). *Wild majesty: Encounters with Caribs from Columbus to the present day: an anthology*. Oxford University Press, USA.

International Air Transport Association (2016). IATA Forecasts Passenger Demand to Double Over 20 Years . `http://www.iata.org/pressroom/pr/Pages/2016-10-18-02.aspx`. Accessed: 2017-11-10.

International Maritime Organization (2017). IMO profile. `https://business.un.org/en/entities/13`. Accessed: 2017-11-10.

Irvin, G., Bickler, S., and Quirke, P. (1990). Voyaging by canoe and computer: experiments in the settlement of the Pacific Ocean. *Antiquity*, 64(242):34–50.

James, R. W. (1957). *Application of Wave Forecasts to Marine Navigation*. U. S. Naval Oceanographic Office.

Jensen, C. K., Chiarandini, M., and Larsen, K. S. (2017). Flight Planning in Free Route Airspaces. In D'Angelo, G. and Dollevoet, T., editors, *17th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS 2017)*, volume 59 of *OpenAccess Series in Informatics (OASIcs)*, pages 14:1–14:14, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Johnson, D. S. and Pollak, H. O. (1987). Hypergraph Planarity and the Complexity of Drawing Venn Diagrams. *Journal of Graph Theory*, 11(3):309–325.

Karasik, A. and Smilansky, U. (2011). Computerized morphological classification of ceramics. *Journal of Archaeological Science*, 38(10):2644–2657.

Karp, R. M. (1972). Reducibility Among Combinatorial Problems. In *Complexity of Computer Computations*, pages 85–103. Plenum Press.

Kaufman, L. and Rousseeuw, P. (1987). *Clustering by means of medoids*. North-Holland.

Kaufman, L. and Rousseeuw, P. J. (1990). Partitioning around medoids (program pam). *Finding groups in data: an introduction to cluster analysis*, pages 68–125.

Kaufmann, M., van Kreveld, M., and Speckmann, B. (2009). Subdivision Drawings of Hypergraphs. In Tollis, I. G. and Patrignani, M., editors, *Proceedings of the 16th International Symposium on Graph Drawing (GD 2008)*, volume 5417 of *Lecture Notes in Computer Science*, pages 396—-407. Springer.

Keegan, W. F. and Hofman, C. L. (2017). *The Caribbean Before Columbus*. Oxford University Press.

Keegan, W. F., Hofman, C. L., and Ramos, R. R. (2013). *The Oxford Handbook of Caribbean Archaeology*. Oxford Authors Series. OUP USA.

Kendall, M. G. (1938). A New Measure of Rank Correlation. *Biometrika*, 30(1/2):81–93.

Klompstra, M. B., Olsder, G., and Van Brunschot, P. (1992). The isopone method in optimal control. *Dynamics and Control*, 2(3):281–301.

Knippenberg, S. (2007). *Stone artefact production and exchange among the Lesser Antilles*, volume 13 of *Archaeological Studies Leiden University (ASLU)*. Amsterdam University Press.

Knuth, D. E. and Raghunathan, A. (1992). The problem of compatible representatives. *SIAM Journal on Discrete Mathematics*, 5(3):422–427.

Lamsal, R. (2017). Evaluation of Human Ability to Find Optimal Canoe Routes. Bachelor's thesis, University of Konstanz, Germany.

Leidwanger, J. (2013). Modeling distance with time in ancient Mediterranean seafaring: a GIS application for the interpretation of maritime connectivity. *Journal of Archaeological Science*, 40(8):3302–3308.

Levison, M., Ward, R. G., and Webb, J. W. (1973). *The settlement of Polynesia: A computer simulation*. University of Minnesota Press.

Lewis, C. L. (1980). *Matthew Fontaine Maury*. Arno Pr.

Lichtenstein, D. (1982). Planar Formulae and Their Uses. *SIAM Journal on Computing*, 11(2):329–343.

Lin, Y.-H., Fang, M.-C., and Yeung, R. W. (2013). The optimization of ship weather-routing algorithm based on the composite influence of multi-dynamic elements. *Applied Ocean Research*, 43:184–194.

Lucena, M., Fuertes, J. M., Martinez-Carrillo, A. L., Ruiz, A., and Carrascosa, F. (2016). Efficient classification of Iberian ceramics using simplified curves. *Journal of Cultural Heritage*, 19:538 – 543.

Lusby, S., Hannah, R., and Knight, P. (2009). Navigation and Discovery in the Polynesian Oceanic Empire – Part One. *Hydrographic Journal*, 131:17–25.

Lyman, R., O'Brien, M., and Dunnell, R. (1997). *The Rise and Fall of Culture History*. Springer US.

Maki, A., Akimoto, Y., Nagata, Y., Kobayashi, S., Kobayashi, E., Shiotani, S., Ohsawa, T., and Umeda, N. (2011). A new weather-routing system that accounts for ship stability based on a real-coded genetic algorithm. *Journal of marine science and technology*, 16(3):311.

Makridis, M. and Daras, P. (2013). Automatic Classification of Archaeological Pottery Sherds. *Journal of Computational Cultural Heritage*, 5(4):15:1–15:21.

Mandal, B. (2015). Linear integer programming approach to construction of balanced incomplete block designs. *Communications in Statistics-Simulation and Computation*, 44(6):1405–1411.

Mann, C. (2011). *1493: Uncovering the New World Columbus Created*. Knopf Doubleday Publishing Group.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer research*, 27(2 Part 1):209–220.

Martin, P. J., Barron, C. N., Smedstad, L. F., Campbell, T. J., Wallcraft, A. J., Rhodes, R. C., Rowley, C., Townsend, T. L., and Carroll, S. N. (2009). User's Manual for the Navy Coastal Ocean Model (NCOM) version 4.0. Technical report, DTIC Document.

Maurizio Patrignani (2014). Planarity testing and embedding. In Tamassia, R., editor, *Handbook of Graph Drawing and Visualization*, pages 1–42. CRC Press.

Maury, M. F. (1855). *The Physical Geography of the Sea*. Harper & Brothers.

McKusick, M. B. (1960). *Aboriginal canoes in the West Indies*. Department of Anthropology, Yale University.

McNiven, I. J. (2008). Sentient sea: Seascapes as spiritscapes. *Handbook of landscape archaeology*, pages 149–157.

Meinshausen, M., Smith, S., Calvin, K., Daniel, J., Kainuma, M., Lamarque, J., Matsumoto, K., Montzka, S., Raper, S., Riahi, K., et al. (2011). The Paleoclimate Modeling Intercomparison Project contribution to CMIP5. *WCRP Coupled Model Intercomparison Project-Phase 5-CMIP5*, 16(56):15–10.

Mills, B. J., Clark, J. J., Peeples, M. A., Haas Jr., W. R., Roberts Jr., J. M., Hill, J. B., Huntley, D. L., Borck, L., Breiger, R. L., Clauset, A., and Shackley, M. S. (2013). Transformation of social networks in the late prehispanic US Southwest. *Proceedings of the National Academy of Sciences*.

Munsell, A. H. (1912). A Pigment Color System and Notation. *The American Journal of Psychology*, 23(2):236–244.

Murray, W. M. (1987). Do modern winds equal ancient winds? *Mediterranean Historical Review*, 2(2):139–167.

National Geospatial-Intelligence Service (2002). *Atlas of Pilot Charts North Atlantic Ocean (including Gulf of Mexico)*. Pub. 106.

National Geospatial-Intelligence Service (2015). *Sailing Directions Caribbean Sea*, volume 1+2. CreateSpace Independent Publishing Platform.

Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582.

Newman, M. E. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113.

Nguyen, N.-K. (1994). Construction of optimal block designs by computer. *Technometrics*, 36(3):300–307.

Orton, C. and Hughes, M. (2013). *Pottery in Archaeology*. Cambridge Manuals in Archaeology. Cambridge University Press.

O'Shea, M. J. (2010). Crossing a river in a canoe – how complicated can it get? *European Journal of Physics*, 31(4):857.

Papadakis, N. A. and Perakis, A. N. (1990). Deterministic Minimal Time Vessel Routing. *Operations Research*, 38(3):426–438.

Papoulis, A. and Pillai, S. (2002). *Probability, random variables, and stochastic processes*. McGraw-Hill electrical and electronic engineering series. McGraw-Hill.

Park, S. C., Shin, H., and Choi, B. K. (2001). A sweep line algorithm for polygonal chain intersection and its applications. In *Geometric Modelling*, pages 309–321. Springer.

Perakis, A. N. and Papadakis, N. A. (1989). Minimal Time Vessel Routing in a Time-Dependent Environment. *Transportation Science*, 23(4):266–276.

Petitjean-Roget, J. (1963). The Caribs as seen through the dictionary of the Reverend Father Breton. In *Proceedings of the First International Convention for the Study of Pre-Columbian Culture in the Lesser Antilles*, volume 1, pages 43–68.

Petrie, W. M. F. (1899). Sequences in Prehistoric Remains. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 29(3):295–301.

Pierce, C. (1999). *Explaining corrugated pottery in the American Southwest: an evolutionary approach*. PhD thesis. Department of Anthropology, University of Washington, Seattle (unpublished).

Podani, J. (1999). Extending Gower's general coefficient of similarity to ordinal characters. *Taxon*, pages 331–340.

Polynesian Voyaging Society (2017). Estimating Distance and Direction Traveled. `http://archive.hokulea.com/ike/hookele/estimating_distance_direction.html`. Accessed: 2017-09-26.

Rice, P. M. and Saffer, M. E. (1982). Cluster analysis of mixed-level data: pottery provenience as an example. *Journal of Archaeological Science*, 9(4):395–409.

Robinson, W. S. (1951). A method for chronologically ordering archaeological deposits. *American Antiquity*, 16(4):293–301.

Romney, A. K. and d'Andrade, R. G. (1964). Cognitive aspects of English kin terms. *American Anthropologist*, 66(3):146–170.

Rouse, I. (1960). The classification of artifacts in archaeology. *American Antiquity*, 25(3):313–323.

Rouse, I. (1965). Caribbean Ceramics: a study in method and in theory. *Ceramics and Man. Viking Fund Publications in Anthropology*, 41:88–103.

Samson, A. V., Cooper, J., Nieves, M. A., Wrapson, L. J., Redhouse, D., Vieten, R.-M., de Jesús Rullan, O., de Victoria, T. G. L., Gómez, A. P., Puigdoller, V. S., et al. (2015). Indigenous cave use, Isla de Mona, Puerto Rico. In *Proceedings of the 25th International Congress for Caribbean Archaeology*, pages 414–444.

Shepard, A. O. (1956). *Ceramics for the Archaeologist*. Carnegie Institution for Science.

Slayton, E. (2017). *Seascape Corridors: How modelling routes through the sea can illuminate early island culture*. PhD thesis, Leiden University.

Slayton, E., Athenstädt, J. C., and Hildenbrand, J. (2016). Uncovering routes to Grenada: Exploring possible routes between mainland South America and the Southern Lesser Antilles. Paper presented at the 44th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Oslo.

Slayton, E., Athenstädt, J. C., Hildenbrand, J., and Hofman, C. L. (2015). Modelling Caribbean Seascapes: New methods in computer modeling of prehistoric canoe routes in the Caribbean. Paper presented at the XXVI Congress of the International Association for Caribbean Archaeology, St. Martin.

Slayton, E., Athenstädt, J. C., and Hofman, C. L. (2017a). Evaluating Least Cost Canoe Routes from The Caribbean. Presentation at the Annual Conference of the European Association of Archaeologists, Maastricht.

Slayton, E., Athenstädt, J. C., Hofman, C. L., and Hildenbrand, J. (2017b). Seafaring with Mental Sign Posts: Connecting Materials with Canoe Routes in the Archaic Age Northern Lesser Antilles. Paper presented at the 45th Annual Conference on Computer Applications and Quantitative Methods in Archaeology, Atlanta.

Spaulding, A. C. (1953). Statistical Techniques for the Discovery of Artifact Types. *American Antiquity*, 18(4):305–313.

Steponaitis, V. (2009). *Ceramics, Chronology, and Community Patterns: An Archaeological Study at Moundville*. University of Alabama Press.

Stinson, D. R. (2004). *Combinatorial designs: constructions and analysis*. Springer Science & Business Media.

Szlapczynska, J. and Smierzchalski, R. (2007). Adopted isochrone method improving ship safety in weather routing with evolutionary approach. *International Journal of Reliability, Quality and Safety Engineering*, 14(06):635–645.

Teirlinck, L. (1975). On the maximum number of disjoint triple systems. *Journal of Geometry*, 6(1):93–96.

Teirlinck, L. (1977). *Combinatorial structures*. PhD thesis. Université libre de Bruxelles, Faculté des sciences, Bruxelles (unpublished).

Tolman, H., Accensi, M., Alves, H., Ardhuin, F., Bidlot, J., Booij, N., Bennis, A.-C., Campbell, T., Chalikov, D. V., Filipot, J.-F., et al. (2014). User manual and system documentation of Wavewatch III version 4.18. Technical report, Environmental Modeling Center, Marine Modeling and Analysis Branch.

Torgerson, W. S. (1951). A theoretical and empirical investigation of multidimensional scaling. *ETS Research Bulletin Series*, 1951(2):i–123.

Torgerson, W. S. (1958). Theory and methods of scaling.

Tsou, M.-C. and Cheng, H.-C. (2013). An Ant Colony Algorithm for efficient ship routing. *Polish Maritime Research*, 20(3):28–38.

Verheyen, S., Voorspoels, W., Vanpaemel, W., and Storms, G. (2016). Caveats for the spatial arrangement method: Comment on hout, goldinger, and ferguson (2013). *Journal of Experimental Psychology: General*.

Walther, L., Rizvanolli, A., Wendebourg, M., and Jahn, C. (2016). Modeling and Optimization Algorithms in Ship Weather Routing. *International Journal of e-Navigation and Maritime Economy*, 4:31–45.

Wang, H., Su, H., Zheng, K., Sadiq, S., and Zhou, X. (2013). An effectiveness study on trajectory similarity measures. In *Proceedings of the Twenty-Fourth Australasian Database Conference-Volume 137*, pages 13–22. Australian Computer Society, Inc.

Ward, R. G., Webb, J. W., and Levison, M. (1973). The settlement of the Polynesian outliers: a computer simulation. *The Journal of the Polynesian Society*, pages 330–342.

Warnking, P. (2015). *Der Römische Seehandel in seiner Blütezeit.* Pharos - Studien zur griechisch-römischen Antike. VML, Rahden/Westf.

Wheat, J. B., Gifford, J. C., and Wasley, W. W. (1958). Ceramic Variety, Type Cluster, and Ceramic System in Southwestern Pottery Analysis. *American Antiquity*, 24(1):34–47.

Zaron, E., Fitzpatrick, P., Cross, S., Harding, J., Bub, F., Wiggert, J., Ko, D., Lau, Y., Woodard, K., and Mooers, C. (2015). Initial evaluations of a Gulf of Mexico/Caribbean ocean forecast system in the context of the Deepwater Horizon disaster. *Frontiers of Earth Science*, 9(4):605–636.

Appendices

Appendix A

Transcripts of the Interviews for the Canoeing Game (Section 1.5)

The interview has been conducted and transcribed by Roman Lamsal during the course of his bachelor's thesis (Lamsal 2017)

Player A

1. *How would you describe your learning process?*
   Playing a lot, trying to optimize on a large scale and test different currents.

2. *Order of missions played?*
   I just tried to be the number one in each mission. When someone took my place, I played that mission until I was number one again.

3. *(How) Did you optimize the paths between checkpoints?*
   I played a lot and tried to remember where I was going in my last run between the checkpoints. Then I tried to change it up a bit here and there to see if I could optimize parts of it, but overall, the majority of optimization took place between the start and the first checkpoint of each mission. That was more convenient. In order to reconstruct the former routes, I orientated myself along the islands: When the view looked familiar, I kind of knew that I was going in the right direction.

4. *Did you optimize your path outside of the game itself?*
   No.

5. *Was it hard to master the controls?*
   The 180° curves were hard to do without losing too much speed. I played on Android and sometimes lost the right timing on when to lift my finger again. When that happens, the paddle is not swung but the boat brakes which is obviously not desired in moments where you want to go at full speed.

6. *Would intermediate feedback help in optimizing?*
   It would have helped, but if I had intermediate feedback like the current time, I would have quit even more rounds. The arrow in the compass was quite helpful; I always just tried to stir the boat directly towards the checkpoint against the currents.

7. *If able, did you exchange with other players about their paths?*
   Another player told me his route in the Jaguar Teeth (1/2) mission. I tried it, but was not able to achieve good results and thus stopped trying and sticked to my own routes.

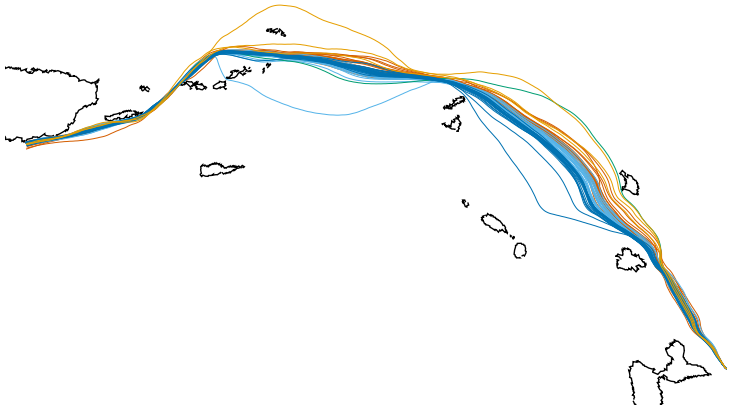
8. *Do you still play the game?*
   Not that much anymore, the competition kept me playing.

Player B

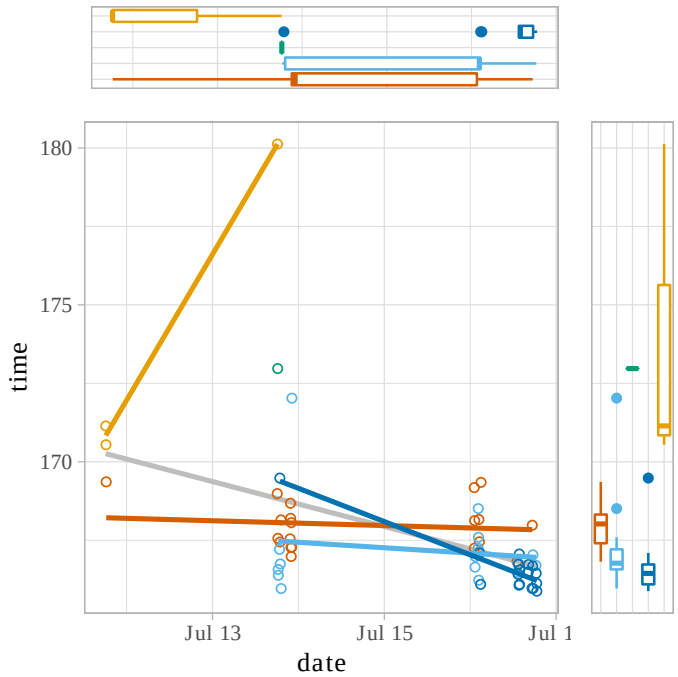
1. *How would you describe your learning process?*
   In general, I just tried to beat the record holder. I tried to use the currents as much as possible, because I mean, the game was about the currents?
2. *Order of missions played?*
   I just started with the first mission in the mission selection screen (remark: at this point, this was Jaguar Teeth (1/2), whereas PlayerA had Pottery Acquisition as first mission in the list) and tried to go with it. After I was in the range of the best player, I tried optimizing further until I just tried the other missions for good. The second Jaguar Teeth mission was not my thing. The mission was too long and to confusing with all the strong currents, so I tried some runs but pretty much gave up on it. I'd rather spend my time on other missions.
3. *(How) Did you optimize the paths between checkpoints?*
   I was lucky and found good routes pretty fast. But it was a hassle to optimize after the first checkpoint, depending on the mission, it took too long to get there and in the end you do not even not if you messed up before or after...
4. *Did you optimize your path outside of the game itself?*
   Yes I did. I was playing the desktop version of the game and took screenshots of my path after each round (remark: you can access the map representation of the round via the highscore screen), then I tried to alternate a little and compare.
5. *Was it hard to master the controls?*
   Sometimes I had a hard time to find the right rhythm for the paddling. I mean, sometimes I hit the button to early and then nothing happened, and in the end, I often times ended up just button mashing. I even had to stop at one point because my joints were hurting from all the paddling.
6. *Would intermediate feedback help in optimizing?*
   Greatly! But I guess then it would be too easy...
7. *If able, did you exchange with other players about their paths?*
   No, I was pretty content with my own routes.
8. *Do you still play the game?*
   No, the competition kept me playing. It is very time consuming to optimize paths and I feel like I have already done a great deal of optimization.

Appendix B

Routes for the Static Missions in the Canoeing Game (Section 1.5)

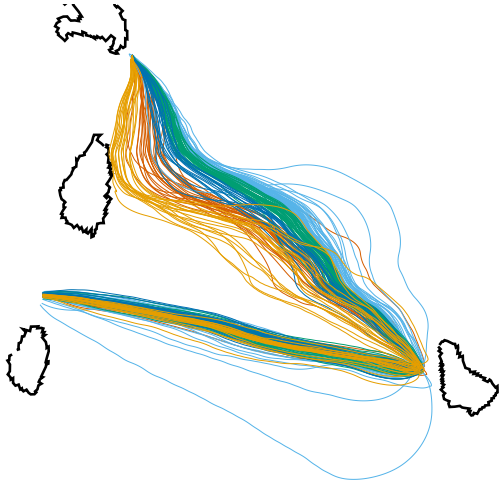


(a) Routes

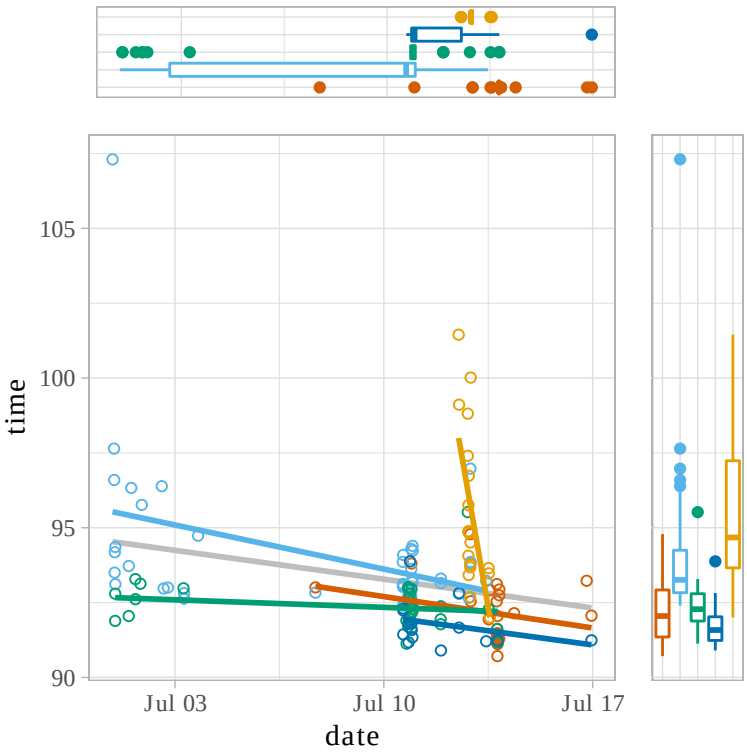


(b) Distribution

Figure B.1: *Pottery Acquisition*, Player B, five clusters based on the simple Euclidean distance

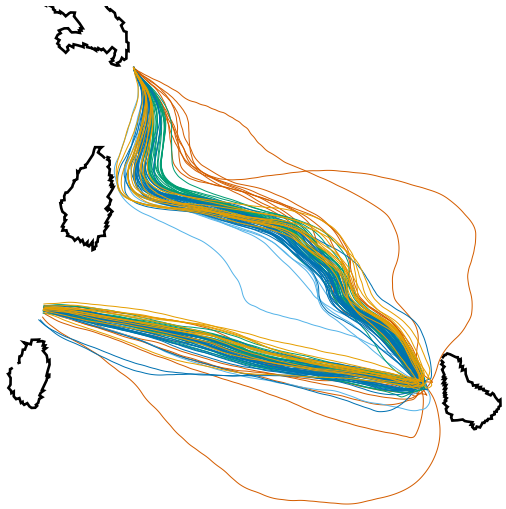


(a) Routes

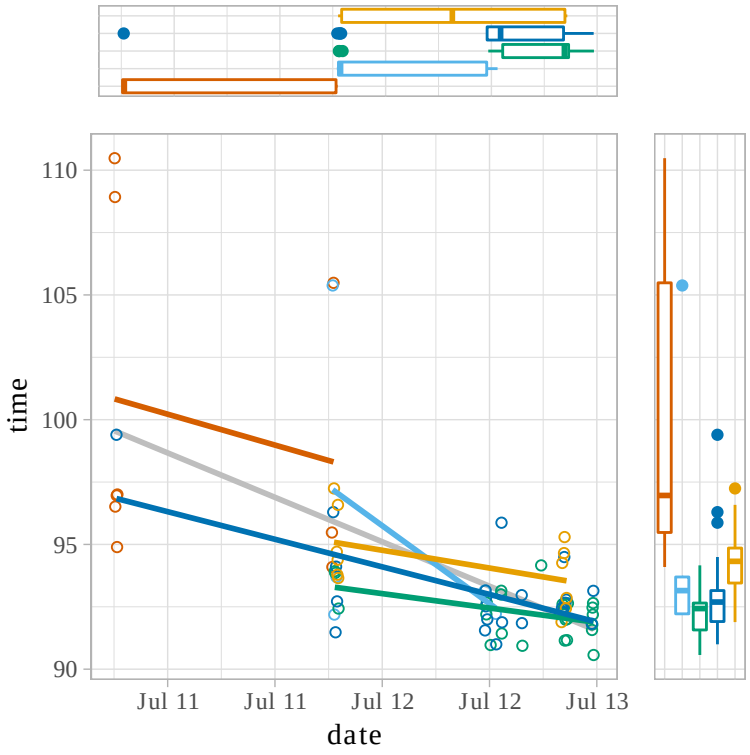


(b) Distribution

Figure B.2: *Jaguar Teeth I*, Player A, five clusters based on the simple Euclidean distance

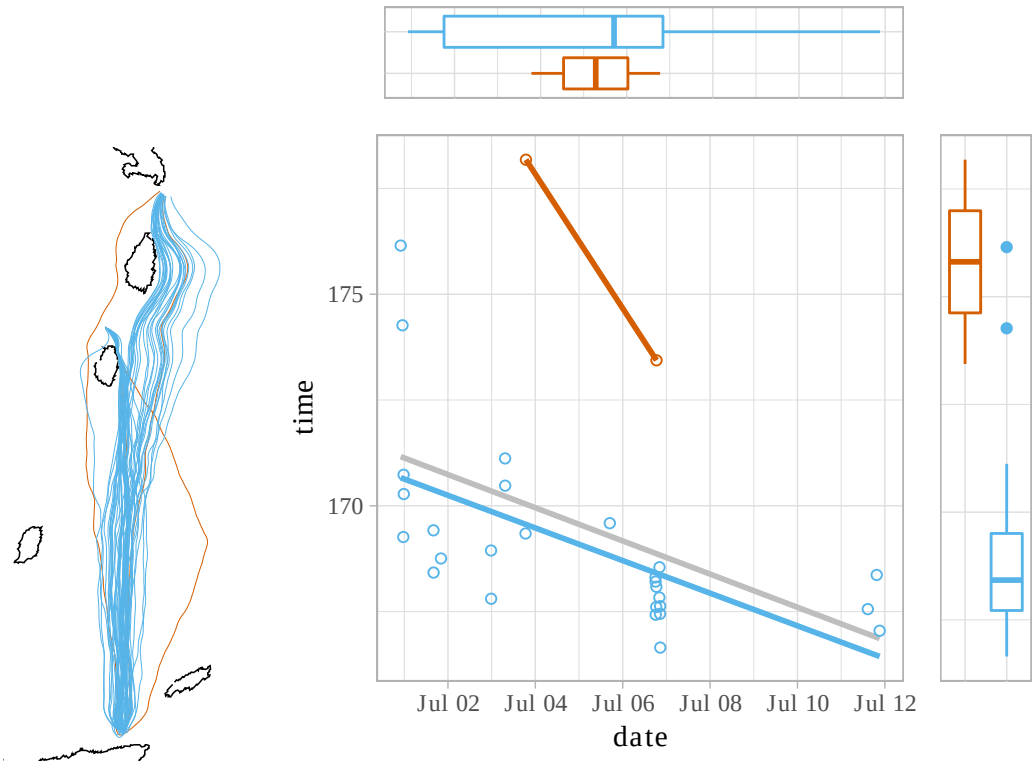


(a) Routes



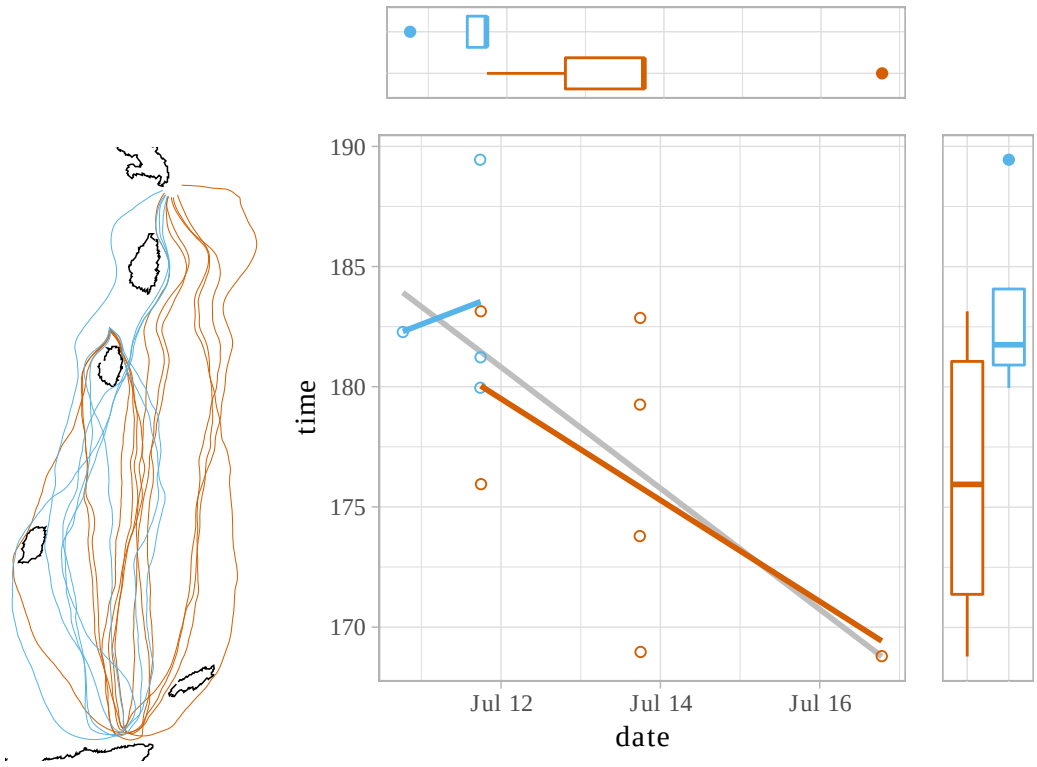
(b) Distribution

Figure B.3: *Jaguar Teeth I*, Player B, five clusters based on the simple Euclidean distance



(a) Routes **(b)** Distribution

Figure B.4: *Jaguar Teeth II*, Player A, two clusters based on Fréchet distance

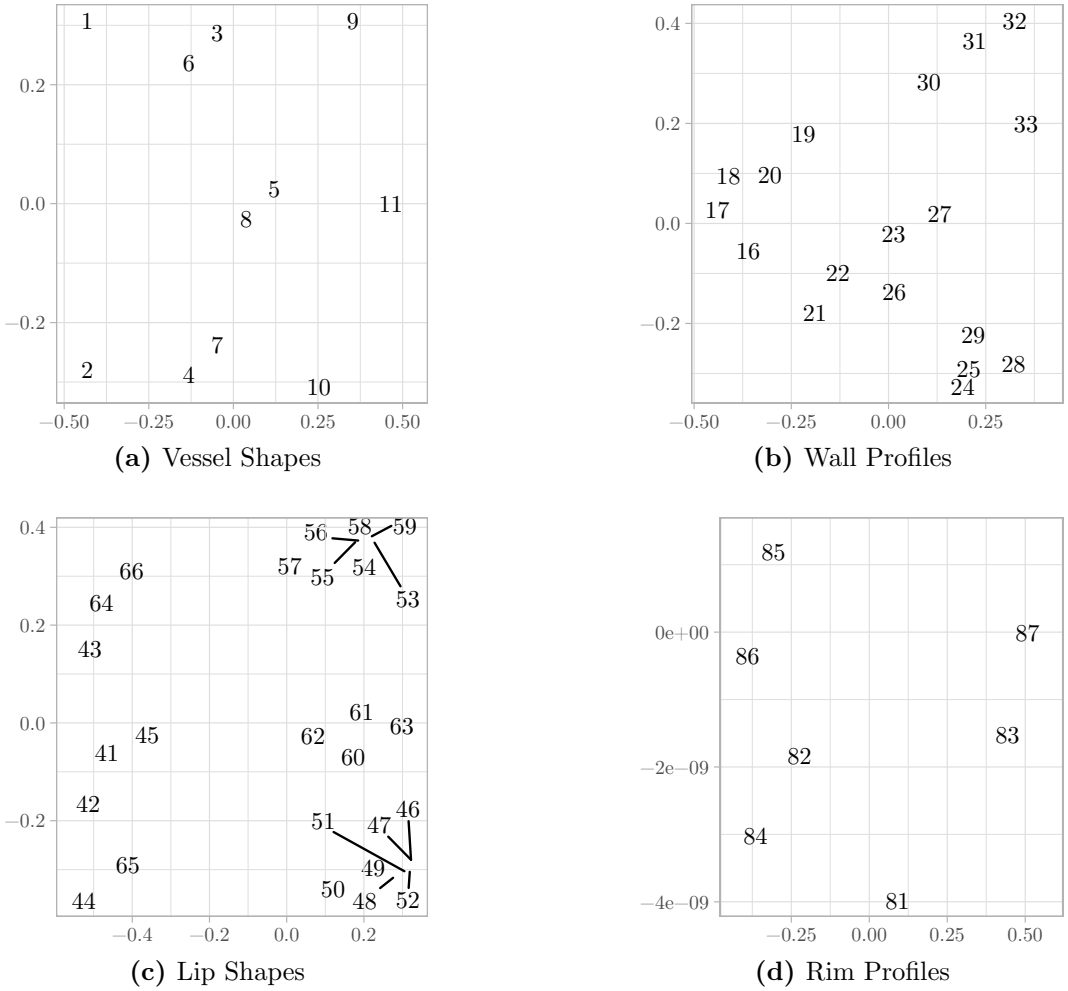


(a) Routes **(b)** Distribution

Figure B.5: *Jaguar Teeth II* Player B, two clusters based on Fréchet distance

Appendix C

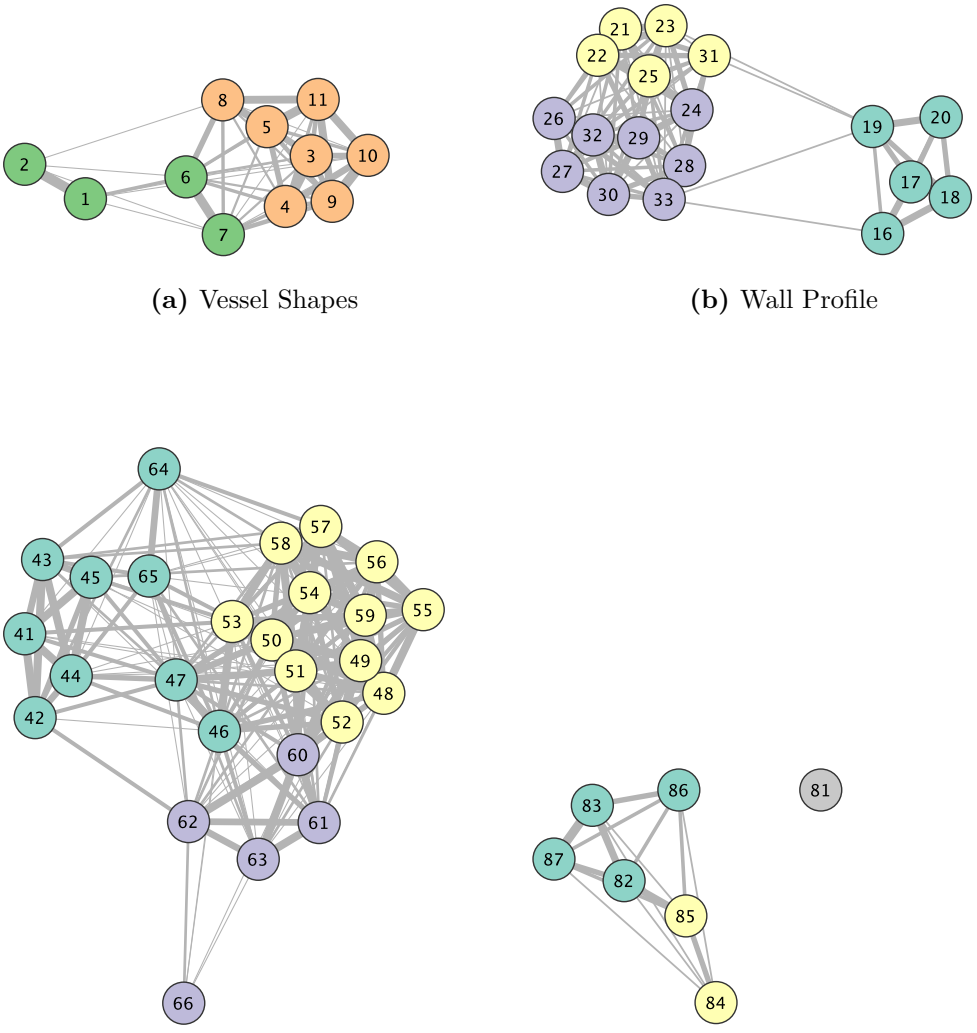
Remaining Results of the MDR and Triad Tests (Section 2.5)



(a) Vessel Shapes **(b)** Wall Profiles

(c) Lip Shapes **(d)** Rim Profiles

Figure C.1: MDS-plots based on the recorded distances for the shape attributes



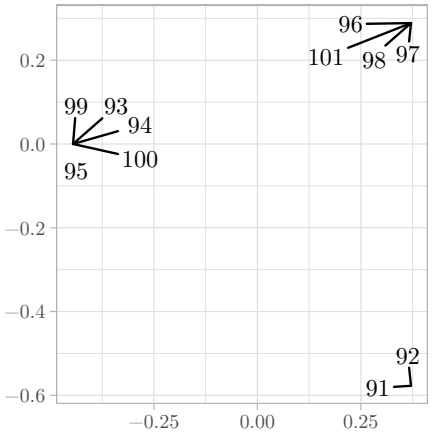
(a) Vessel Shapes

(b) Wall Profile

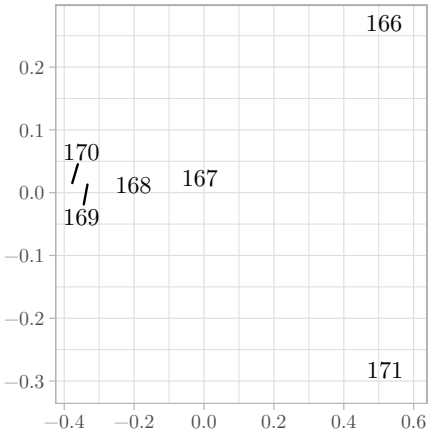
(c) Lip Shapes

(d) Rim Profile

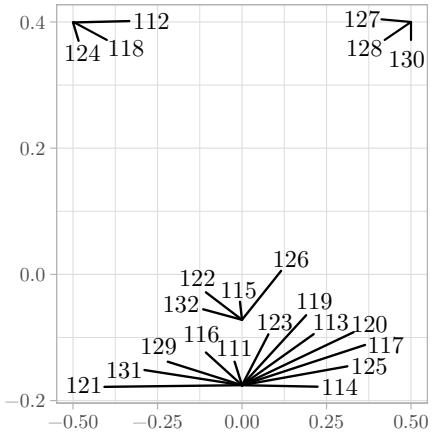
Figure C.2: The networks resulting from the triad tests on the shape attributes



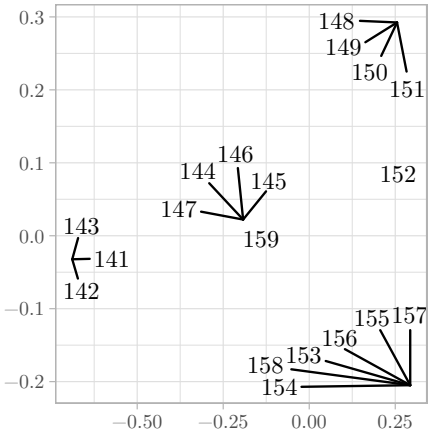
(a) Decoration



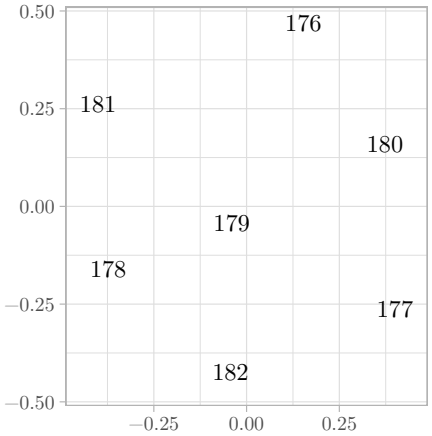
(b) Surface Finishing



(c) Color

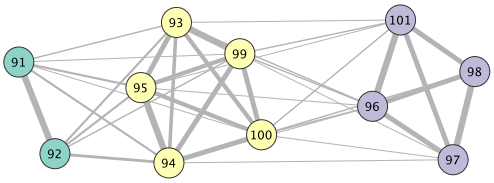


(d) Firing Color

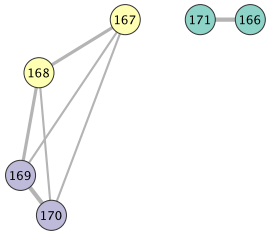


(e) Red Slip

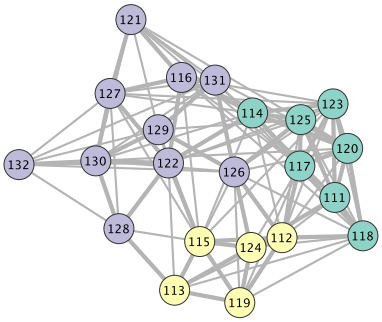
Figure C.3: MDS-plots based on the recorded distances for the non-shape attributes



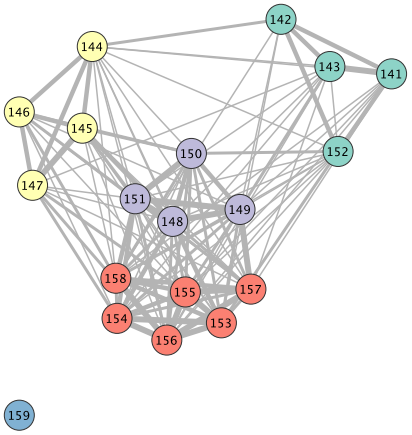
(a) Decoration



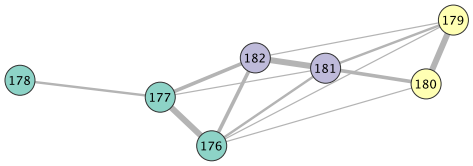
(b) Surface Finishing



(c) Color



(d) Firing Color



(e) Red Slip

Figure C.4: The networks resulting from the triad tests on the non-shape attributes

Appendix D

Protocol for the Pottery Perception Test (Section 2.7)

This test script has been developed in joint work with Lewis Borck

[TESTERS READ ITALICIZED SECTIONS.]

Introduction and General Questions

- [Hand out release form and discuss how we will anonymize.]
- [Before recording:] *Do you mind if we record this interview with an audio recording device? We will use this audio for personal research only. It just records quicker than we can write. If not, that is fine. Just let us know and we will take notes.*

[START RECORDING]

First, we want to make our research and intentions clear. Our names are Lewis Borck and Jan Athenstädt. Lewis is a lecturer and researcher in the Faculty of Archaeology at Universiteit Leiden in the Netherlands. He received his Ph.D. in Anthropology from the University of Arizona. He has worked as a Southwestern archaeologist for over a decade. Jan is a doctoral student at the Department of Computer Science and Information Science at Konstanz University in Germany. We are here to collect and study how individuals interact with, examine, and perceive archaeological ceramics to improve archaeological interpretations. This project arose as part of the research undertaken in the NWO Island Networks Project and the ERC Synergy Project NEXUS1492 *(*`www.nexus1492.eu`*) in the Caribbean. The project was transitioned from the Caribbean to the American Southwest as there are larger numbers of readily available archaeologists and traditional potters available to interview. We will create audio recordings of a range of interviewees. These interviewees will be anonymized to protect individual privacy. Transcripts from these anonymized interviews will be made available in edited form on the internet and might be embedded in a searchable and publicly accessible database of the* NEXUS1492 *project. Information from these interviews (and the associated tests) will be used in public and academic research publications. Edited interviews and photographs will bear the Creative Common attribution, non-commercial, sharealike license.*

- *Are you okay with us doing this interview and with us using the information you provide for research?*

If you want to stop this interview and test at any time, please do. You are under no obligation to continue and we will be more than happy to stop. Your consideration and comfort are of the highest importance to us.

When we write about the information you provided in this interview and test, your name will be anonymized.

This test is going to be a mix of hands-on movement and placement of archaeological pottery sherds with some directed and open-ended interview questions throughout. If it is okay with you, we will start with some of the questions.

- *Do you identify as Male/Female/Neither/Other(please describe if Other)? You can say "not applicable" if you do not feel comfortable answering.*
- *What is your highest level of schooling?*

- [if a degree] What is your degree?

- *What general area of the Greater American Southwest do you live in (e.g., Southeast Colorado)?*

- *Are you interested in archaeology?*
  - [if yes] *Are you a professional archaeologist?*
    - [if yes]
      - *How many years?*
      - *What is your job title?*
      - *What is your regional specialty?*
      - *What is your methodological specialty?*
      - *Do you consider yourself a ceramicist?*
    - [if no] *Are you an avocational archaeologist?*
      - [if yes] *How many years?*

- *Are you currently employed, unemployed, or retired?*
  - [if employed] What is your profession?
  - [if retired] What was your profession?
  - [if unemployed] What was your last profession?

- *Have you ever made pottery?*
  - [if yes]
    - *How many years have you been making pottery?*
    - *Have you ever sold pottery that you have made?*
    - *Do you consider yourself a professional potter?*

- *What, if any, is your ethnic/tribal affiliation? You can say "not applicable" if you do not feel comfortable answering.*

[FOR TESTERS ONLY]

- Is the participant:
  - General archaeologist
  - Archaeological ceramic analyst
  - Indigenous potter
  - Non-indigenous potter
  - General public

- Where is the test performed (e.g., outside, under a canopy, in a living room)?

[MANUAL PORTION BEGINS]

Everyone who takes this test is confused at the beginning. It should become more clear as we proceed, but please let us know if you do not understand any of our instructions. We will be very happy to explain or demonstrate.

Test 0 (familiarization)

[TESTERS PLACE LEGOS IN MIDDLE OF CANVAS]

We are going to start with a test run with 10 Lego bricks. Please arrange the legos so that the legos that are more similar to each other are closer together and those that are more different are farther apart. You are allowed to move the legos, once you have put them on the canvas.

[TESTERS RECORD LEGO PLACEMENT USING A DIGITAL CAMERA AND WRITE DOWN ANY PROBLEMS THE PARTICIPANT HAD.]

Test 1

[TESTERS REMOVE LEGOS AND PLACE ARCHAEOLOGICAL SHERDS IN A PILE IN THE MIDDLE OF THE CANVAS.]

For the first test please arrange the sherds so that the sherds that are most similar to each other are closer together and those that are more different are farther apart. You are allowed to move the sherds, once you have put them on the canvas.

TESTERS: If participants are having trouble starting, suggest randomly picking a sherd and placing it on the canvas and then placing the next sherd in relationship to the first sherd.

TESTERS: Please note whether the participant starts the test with groupings of sherds or if all sherds are more spaced out as individuals.

TESTERS: Photograph in plan-view the 2 meter x 2 meter canvas with the sherds placed by the participant.

Note the sherds that are farthest apart and ask the participant why they placed those so far apart.

Note the sherds that are closest together and ask the participant why they placed them so close together.

Find a group of sherds that are approximately in the middle for similarity distance and ask the participant why they placed those sherds that distance from each other.

Thank you for your help. We only have 3 more tests.

Test 2

[TESTERS PUSH ARCHAEOLOGICAL SHERDS BACK INTO A PILE IN THE MIDDLE OF THE CANVAS.]

For the second test, please arrange the sherds so that sherds you think were made by a similar group of people are closer together and sherds from different groups are further apart. You are allowed to move the sherds, once you have put them on the canvas.

TESTERS: If participants are having trouble starting, suggest randomly picking a sherd and placing it on the canvas and then placing the next sherd in relationship to the first sherd.

TESTERS: Please note whether the participant starts the test with groupings of sherds or if all sherds are more spaced out as individuals.

TESTERS: Photograph in plan-view the 2 meter x 2 meter canvas with the sherds placed by the participant.

Note the sherds that are farthest apart and ask the participant why they placed those so far apart.

Note the sherds that are closest together and ask the participant why they placed them so close together.

Find a group of sherds that are approximately in the middle for similarity distance and ask the participant why they placed those sherds that distance from each other.

Thank you for your help. We only have 2 more tests.

Test 3

[TESTERS PUSH ARCHAEOLOGICAL SHERDS BACK INTO A PILE IN THE MIDDLE OF THE CANVAS.]

For the third test, please arrange the sherds so that the sherds of which you think the pot had a similar function are closer together and those that come from pots that were used for different purposes are farther apart. You are allowed to move the sherds, once you have put them on the canvas.

TESTERS: If participants are having trouble starting, suggest randomly picking a sherd and placing it on the canvas and then placing the next sherd in relationship to the first sherd.

TESTERS: Please note whether the participant starts the test with groupings of sherds or if all sherds are more spaced out as individuals.

TESTERS: Photograph in planview the 2 meter x 2 meter canvas with the sherds placed by the participant.

Note the sherds that are farthest apart and ask the participant why they placed those so far apart.

Note the sherds that are closest together and ask the participant why they placed them so close together.

Find a group of sherds that are approximately in the middle for similarity distance and ask the participant why they placed those sherds that distance from each other.

Thank you for your help. We only have 1 more tests.

Test 4

[TESTERS PUSH ARCHAEOLOGICAL SHERDS BACK INTO A PILE IN THE MIDDLE OF THE CANVAS.]

For the fourth test, please arrange the sherds on a line so that the sherds that you think were the most difficult to make to the right and the sherds that were the easiest to make are on the left. You are allowed to move the sherds, once you have put them on the canvas.

TESTERS: If participants are having trouble starting, suggest randomly picking a sherd and placing it on the canvas and then placing the next sherd in relationship to the first sherd.

TESTERS: Please note whether the participant starts the test with groupings of sherds or if all sherds are more spaced out as individuals.

TESTERS: Photograph in plan-view the 2 meter x 2 meter canvas with the sherds placed by the participant.

Note the sherds that are farthest apart and ask the participant why they placed those so far apart.

Note the sherds that are closest together and ask the participant why they placed them so close together.

Find a group of sherds that are approximately in the middle for similarity distance and ask the participant why they placed those sherds that distance from each other.

Thank you for your help. That was the last test. Do you have any questions for us?

If you have any questions at a later date, you can contact either Lewis or Jan via our email addresses.

If at a later point in time you decide that you are not comfortable with the interview questions and you want us to remove you from the test and interview sample, just let us know and we will no longer use the information you provided. If this has already been published, we cannot retract that, but we can refrain from using it in the future. Finally, we will provide you with a copy of anything that is published as a result of this work.

Thank you!

[STOP RECORDING]

Appendix E

Detailed Results of the Southwest Perception Study (Section 2.7)

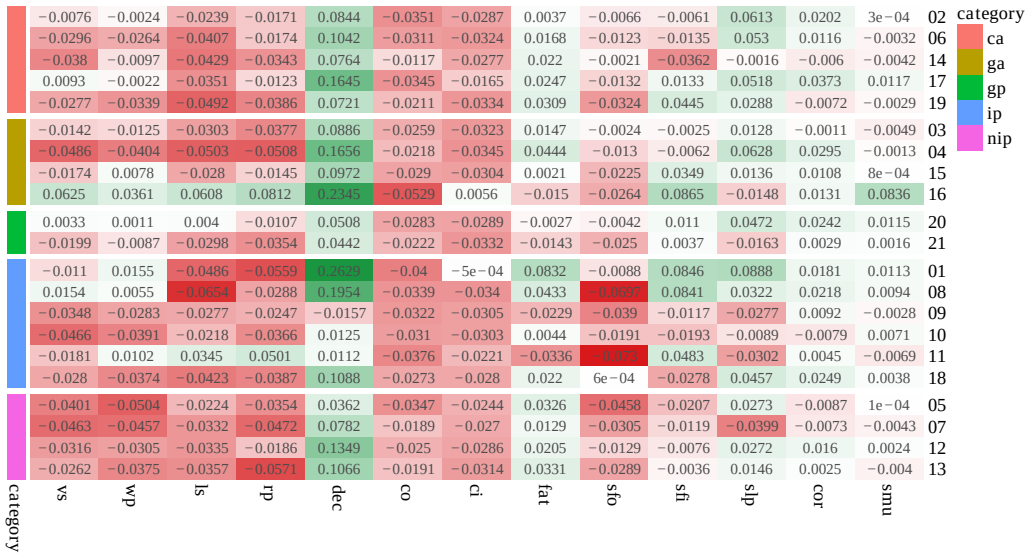
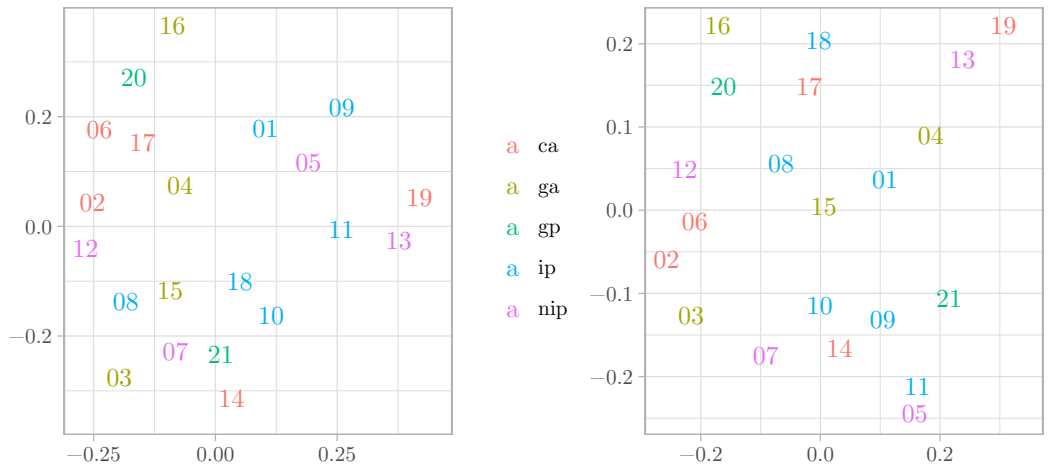


Figure E.1: Test 1: modularity of the attributes for each participant



(a) Procrustes distances **(b)** inverted Mantel correlation

Figure E.2: MDS of two different distances for the participants for Test 1

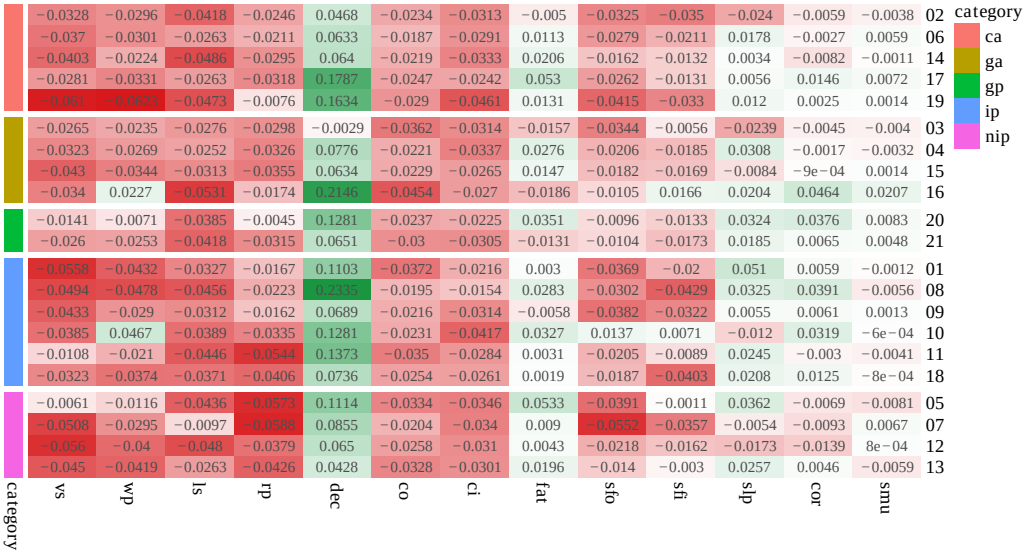
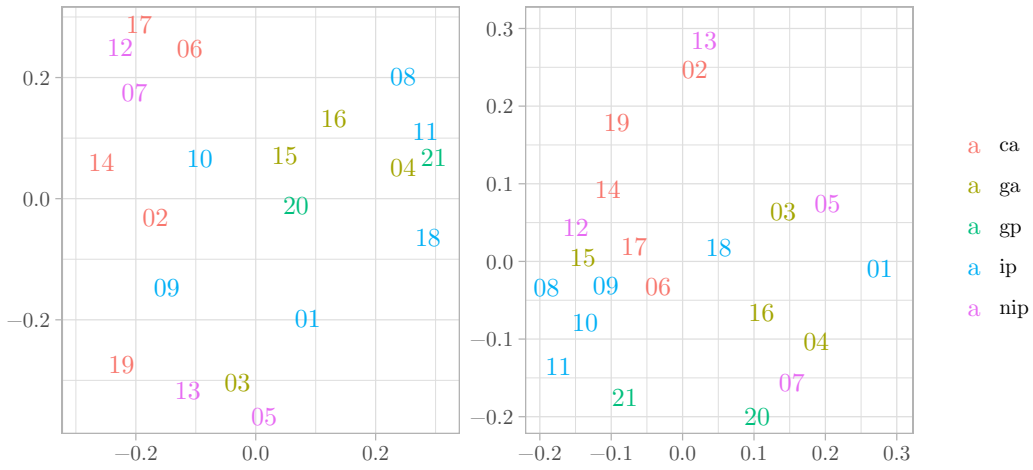


Figure E.3: Test 2: modularity of the attributes for each participant



(a) Procrustes distances **(b)** inverted Mantel correlation

Figure E.4: MDS of two different distances for the participants for Test 2

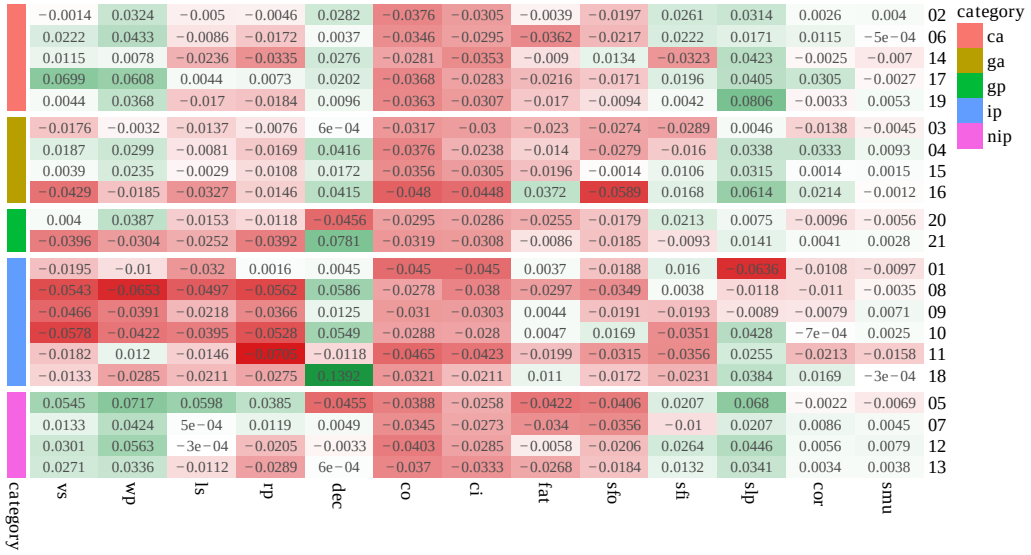
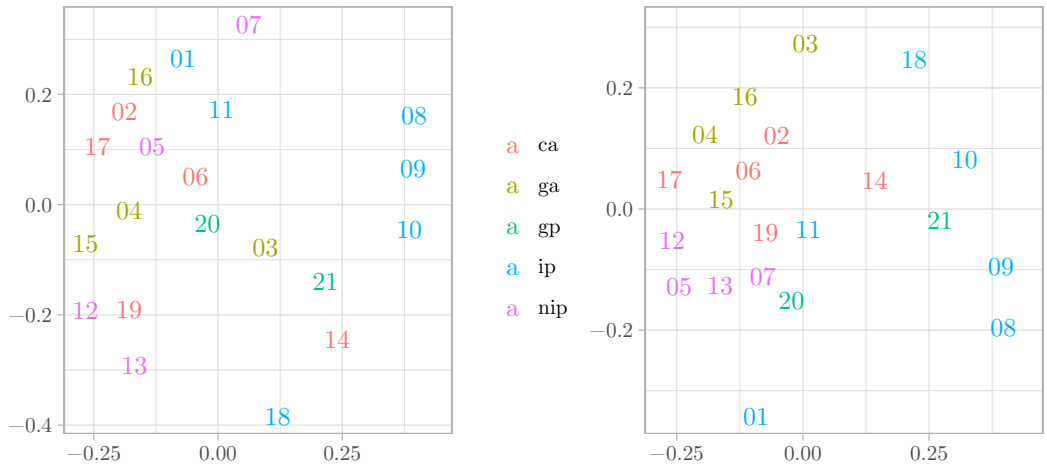


Figure E.5: Test 3: modularity of the attributes for each participant



(a) Procrustes distances **(b)** inverted Mantel correlation

Figure E.6: MDS of two different distances for the participants for Test 3