UNIVERSIDADE TÉCNICA DO ATLÂNTICO
INSTITUTO DE ENGENHARIA E CIÊNCIAS DO MAR

WEST AFRICAN SCIENCE SERVICE CENTRE ON CLIMATE CHANGE AND ADAPTED LAND USE

Master Thesis

COMPARATIVE ANALYSIS OF POPULATION STRUCTURE OF *Gymnothorax vicinus* (MURAENIDAE) IN THE NORTHERN ISLANDS OF CABO VERDE THROUGH MORPHOLOGICAL FEATURES AND GENETIC MARKERS

Moussa Dothian KONATE

Master Research Program on Climate Change and Marine Sciences

São Vicente

2021

UNIVERSIDADE TÉCNICA DO ATLÂNTICO
INSTITUTO DE ENGENHARIA E CIÊNCIAS DO MAR

WEST AFRICAN SCIENCE SERVICE CENTRE ON CLIMATE CHANGE AND ADAPTED LAND USE

Master Thesis

COMPARATIVE ANALYSIS OF POPULATION STRUCTURE OF *Gymnothorax vicinus* (MURAENIDAE) IN THE NORTHERN ISLANDS OF CABO VERDE THROUGH MORPHOLOGICAL FEATURES AND GENETIC MARKERS

Moussa Dothian KONATE

Master Research Program on Climate Change and Marine Sciences

Supervisor | Dr. Heino Fock
Co-supervisor | Mr. Evandro P. Lopes

São Vicente
2021

UNIVERSIDADE TÉCNICA DO ATLÂNTICO
INSTITUTO DE ENGENHARIA E CIÊNCIAS DO MAR

WEST AFRICAN SCIENCE SERVICE CENTRE ON CLIMATE CHANGE AND ADAPTED LAND USE

Comparative analysis of population structure of *Gymnothorax vicinus* (Muraenidae) in the northern islands of Cabo Verde through morphological features and genetic markers

Moussa Dothian KONATE

Master's thesis presented to obtain the master's degree in Climate Change and Marine Sciences, by the Institute of Engineering and Marine Sciences, Atlantic Technical University in the framework of the West African Science Service Centre on Climate Change and Adapted Land Use

Supervisor

Dr. Heino Fock
Thünen Institute of Sea Fisheries
Bremerhaven, Germany

Co-supervisor

Mr. Evandro P. Lopes
Universidade Técnica do Atlântico
(UTA)

São Vicente
2021

UNIVERSIDADE TÉCNICA DO ATLÂNTICO
INSTITUTO DE ENGENHARIA E CIÊNCIAS DO MAR

WEST AFRICAN SCIENCE SERVICE CENTRE ON CLIMATE CHANGE AND ADAPTED LAND USE

Comparative analysis of population structure of *Gymnothorax vicinus* (Muraenidae) in the northern islands of Cabo Verde through morphological features and genetic markers.

Moussa Dothian KONATE

Panel defense

President

_____

Examiner 1

_____

Examiner 2

_____

São Vicente
2021

Financial support

The German Federal Ministry of Education and Research (BMBF) in the framework of the West African Science Service Centre on Climate Change and Adapted Land Use (WASCAL) through WASCAL Graduate Studies Programme in Climate Change and Marine Sciences at the Institute for Engineering and Marine Sciences, Atlantic Technical University, Cabo Verde.

Dedication

To my father, Mr. Dothian KONATE, and my mother, Djénéba Traoré, I dedicate this present thesis.

Acknowledgments

I thank Allahoussoubahana watallah, the almighty, the almightiest, and its Prophet Mohamed (PSL), who by his grace made me this way and showed me this so special day.

This work was done first at the Institute of Marine Research (IMAR), where I received the morphological data from 2002 to 2020 through the collaboration of the Director of the Institute, Dr. Vito Ramos, and the second part of the work has been done at the Laboratory of ISECMAR. I respectfully thank Dr. Vito Ramos for his creativity and enthusiasm. Thank you very much for welcoming me to your Institute, for giving me the data, and for giving me the opportunity to follow the sessions of activities so enriching and exciting in your laboratory. I would like to thank all the workers of this institution, especially Ivanice, Alciany, and Sandra, for their efforts.

I would like to thank especially the Director of the Wascal Cabo Verde program, Doctor Corrine Almeida, and all her collaborators for having treated me well during all my stays in Cabo Verde for this program.

I sincerely thank my supervisor Dr. Heino FOCK, Head of Marine Ecosystems section Thünen Institute of Sea Fisheries Bremerhaven, Germany, for introducing me to the concept of fish stock assessment management. May he find in this work a living tribute to his strong personality, words fail me to thank this great man who has always considered me as a friend, a filston, in short thank you for your availability even if I need him on Sunday, he is ready for me thank you for the quality of the education obtained.

I would like to thank my co-supervisor Mr. Evandro P. Lopes, Teacher Researcher at ISECMAR Cabo Verde, for his availability and his courage to show me how the genetic materials work; thank you for your kind encouragement.

I thank all the professors, mainly from Wascal Cabo Verde, as well as all the professors from Wascal in West Africa, without distinction for the quality of the training obtained.

I also thank the Director of Wascal Mali, Doctor Yacouba Diallo, and all his collaborators for their encouragement, and all my teachers in Mali; especially

Professor Amoro Coulibaly ext. Director of Wascal Mali; Doctor Siaka Dembele, Mrs. Gabriel Awa Coulibaly, Mr. Mahamadou Kelepily and Professor Famoussaba Dansoko.

I address a special thought to my aunt Founey Sylla in America for her unfailing support. I address a special thought to my uncle, my namesake Moussa KONATE Commanding officer of the Brigade de Gendarmerie of Yelimani, and his elements for his unfailing support.

I would like to thank all my friends and all the people who participated in the realization of this work.

Resumo

Dados de 10039 espécimes de *Gymnothorax vicinus* (Castelnau 1855) comummente conhecidos como moreias pretas foram amostradas nas ilhas de São Vicente, Santa Luzia e Santo Antão, entre 2002 a 2020. A serie de dados morfométricos foram fornecidos pelo Instituto de Investigação Marinha em Cabo Verde - IMAR, fruto dos 18 anos de amostragem biológica da espécie. Igualmente, um total de 26 espécimes foram coletados no presente estudo, analisados para fins morfométricos e sequenciados o Citocrome Oxidase I para caracterização genética. A análise das relações comprimento-peso mostrou uma diferença significativa entre as ilhas e os respectivos parâmetros da recta de correlação (a e b) foram 0,0019 e 2,95 para Santa Luzia, 0,0017 e 2,97 para São Vicente e 0,0026 e 2,88 para Santo Antão, respectivamente. A tendência para o factor de condição Fulton K revelou que foram encontrados valores relativamente elevados entre 2006 - 2007 e 2013 - 2014. Para distribuições de frequência de comprimentos foram identificadas diferenças significativas para amostras de menor tamanho bem como não significativas entre ilhas. Com o aumento do tamanho das amostras as diferenças entre as ilhas desapareceram. As diferenças nas relações comprimento-peso entre Santo Antão de um lado, e São Vicente e Santa Luzia do outro, não repercutiram nas análises dos dados genéticos. Dos 26 exemplares, alinhados com dados provenientes do Genbank, foram encontrados 18 haplótipos exclusivos de Cabo Verde. A diversidade haplotípica foi alta e nucleótidica foi baixa. A rede de haplótipos revelou a não existência de um padrão genético populacional no *G. vicinus* entre estas três ilhas, sendo os haplótipos em Cabo Verde muito semelhantes entre si. A única diferença significativa que foi observada nas relações comprimento-peso poderá estar relacionada com as influências ambientais dos habitats desses espécimes.

Palavras-chave: *Gymnothorax vicinus*, mtDNA, análise morfológica, ilhas de Cabo Verde

Abstract

Data from 10039 specimens of *Gymnothorax vicinus* (Castelnau 1855), commonly known as black moray eels, were sampled on the islands of São Vicente, Santa Luzia, and Santo Antão, between 2002 and 2020. The series of morphometric data were provided by the Institute of Marine Research in Cabo Verde - IMAR, as a result of 18 years of biological sampling of the species. Likewise, a total of 26 specimens were collected in the present study, analysed for genetic purposes, and sequenced the Cytochrome Oxidase I for genetic characterisation. The analysis of the length-weight relationships showed a significant difference between the islands, and the respective parameters of the correlation straight line (a and b) were 0.0019 and 2.95 for Santa Luzia, 0.0017 and 2.97 for São Vicente, and 0.0026 and 2.88 for Santo Antão, respectively. The trend for the Fulton K condition factor revealed that relatively high values were found between 2006 - 2007 and 2013 - 2014. For length-frequency distributions, significant differences were identified for smaller sample sizes as well as not significant between islands. With increasing sample size, the differences between islands disappeared. The differences in length-weight relationships between Santo Antão on one side and São Vicente and Santa Luzia on the other did not have repercussions on the analyses of genetic data. Of the 26 specimens aligned with data from Genbank, 18 haplotypes unique to Cabo Verde were found. Haplotype diversity was high, and nucleotide diversity was low. The haplotype network revealed the non-existence of a population genetic pattern in *G. vicinus* between these three islands, the haplotypes in Cabo Verde is very similar to each other. The only significant difference that was observed in the length-weight relationships could be related to the environmental influences of the habitats of these specimens.

Key-words: *Gymnothorax vicinus*, mtDNA, morphological analysis, Cabo Verde islands

Abbreviations and acronyms

ANOVA	Analysis of variance
BAH1	Bahia1
BEL1	Belize1
BOLD	Barcode of Life Data System
CLT	Central Limit Theorem
Cm	Centimeter
COI	Cytochrome oxidase subunit 1
CvM	Cramer Von-Mises test
DNA	Deoxyribonucleic acid
DnaSP5	Sequence polymorphism Deoxyribonucleic acid
EDF	Empirical distribution function
EEZ	Exclusive economic zone
Eq.1	Equation1
FAO	Food and agriculture organization
FN	Fundamental numbers
GPS	Geographic position system
GVIC	*Gymnothorax vicinus*
H. d	Haplotype diversity
IMAR	Marine Research Institute
ISECMAR	Institute of Marine Science and Engineering
Km^2	Kilometer square
Lat	Latitude
LFD	Length-Frequency Distribution
Lmax	Maximum size
Long	Longitude
LWR	Length-Weight Relationship
MEX1	Mexico1
Ml	Milliliters
mtDNA	Mitochondrial Deoxyribonucleic acid
PCR	Polymerase chain reaction
PLD	Pelagic larvae distribution
TL	Total length

USA1	United stated1
VA05	*Vicinus* from Santo Antão
VL01	*Vicinus* from Santa Luzia
VV04	*Vicinus* from São Vicente
µl	Microliters

General index

Figure index

Table index

.

Appendix index

1 Introduction

1.1 Background and Context

Muraenidae (Order Anguilliformes) are the main family of eels and one of the most complex fish families, with 219 species extant (Fricke et *al.*, 2020). They are very common in tropical and temperate coastal seas around the world and are major mesopredators in coral reefs and rocky environments, feeding on a wide variety of vertebrates (fish) and invertebrates (crustaceans and cephalopods) (Böhlke et *al.*, 1989). This family consists of 16 genera, separated into two subfamilies: Muraenidae, with 164 species, and Uropterygiinae, with 36 species. The genus *G. vicinus* is the most common, with 126 species known to date (Coluccia et *al.*, 2020) through ecological and genetic studies. The Anguilliformes, the "true eels," represent an ecologically complex group, mainly of marine origin, whose members can be easily recognized by their extremely stretched body, reduced sections, and the systematic absence of pelvic fins by their extremely elongated body, with reduced cross-sections, and by the universal absence of a pelvic fin (Peninal et *al.*, 2017).

The studies by (LIoris, 1991) provided the first summary of the ichthyofauna biogeography. The Macaronesian ichthyofauna was inferred based on an inventory of 913 species of fish, including data on fish from African sites (Morocco and Western Sahara), and proposed for the first time a hierarchy of biogeographic levels for Cabo Verde. In comparing Cabo Verde to other Macaronesian archipelagos, it was estimated that the zoogeographic composition of its coastal and coastal ichthyofauna is predominantly tropical, with Afro-tropical species largely dominant, followed by tropical and subtropical amphi-Atlantic species tropical and subtropical, then by Mediterranean species and Atlantic species, and then by Mediterranean species and circum-tropical (Wirtz et al., 2013). Almost half of the total diversity of cryptobenthic fish species in Cabo Verde consists of endemic species, some of which have recently been discovered (Freitas, 2014).

The ichthyofauna of the Cabo Verde islands has been examined by many researchers (Wirtz et al., 2013). Since then, several publications have noted the presence of additional species, designated new species, and or revised the genders and families of the region. However, coastal fish are defined here as those fish that are likely to be present in the first 60 meters of the water column from the shore (including species which generally live deeper but have been recorded in this depth range in Cabo Verde) and upper water species that come

close enough to shore (occasionally) to be seen by swimmers (Wirtz et al., 2013) that can move between islands of the Cabo Verde archipelago.

The new analysis of 185 coastal osteichthyans consisted of 61 families and 135 genera, dominated by morays (*Gymnothorax* spp and *Muraena* spp with six and four species respectively), followed by the damsels *Abudefduf* spp, and the *Diplodus* spp (sea bream), with four species each, five genera with three species, 26 genera with two species and 100 genera with two species only a single representant (Wirtz et al., 2013). The distinctive morphological features of family Uropterygiinae concern dorsal and anal fins limited to the tip of the tail in family Muraenidae, the dorsal fin usually appears near the opening of the gills, and the anal fin begins just behind the anus roughly in the middle of the body. Piscivorous morphology characterizes the larger genus of Muraenidae (e.g., *Gymnothorax* spp.) and most other genera; and is similar to the jaws of Uropterygiins, which do not have the same durophagous shape (Reece et al., 2010).

The combination of a cryptic lifestyle and the absence of commercial fishing has allowed moray eels, in general, to remain virtually undetected in standard underwater visual surveys (Gilbert et al., 2005). As nocturnal species, they often take refuge during the day, and the size of their populations and the resulting effects on their community are not studied yet. Then, nocturnal communities tend to have a higher relative abundance of predatory species than diurnal communities (Higgins & Mehta, 2017). The genetic analyses confirm the monophyly of Anguilliformes, but the phylogenetic relationships within the Order deduced from DNA analysis do not correspond to those established by morphological comparisons (Wang et al., 2003). They comprise 15 families, 141 genera, and 791 fish species (Vasconcelos & Molina, 2009). The diversity of the adult niche can also affect the genetic distinction between places within species (Reece et al., 2011). Most marine organisms are distributed exclusively through a pelagic larval stage, and the pelagic larval stage (PLD) affects the dispersal ability of these species.

The genetic analysis provides many benefits for fisheries monitoring, particularly by facilitating techniques for safely identifying specimens and assessing stand stability (Ward, 2000). In most situations, genetics can be the best way to decide whether a species is eligible for protection on the endangered list. Genetic homogeneity of extremely threatened species indicates a need for them to be repopulated by translocation (Teske et al., 2003), employing a program established based on the perception of quantitative genetics, life history, and fluctuation of the DNA.

Limitations of identification based on morphology and meristic characteristics require the use of molecular tools that are capable of identifying fish species. This technique based on molecular manipulation is known under the term “DNA barcoding,” which can identify species at different stages of evolution but also distinguish species that are morphologically similar or cryptic (Jaonalison, 2019). The unambiguous identification of fish eggs and larvae is an essential tool for fish ecology and perennial conservation (Becker et al., 2015).

The current state of many important fishery resources around the world can be characterized as seriously depleted or threatened with depletion due to poor management practices and poor management fishing pressure, for example. However, unsustainable fishing techniques, combined with a maximum level of investment in fishing capacity, have led to severe degradation and poor performance of reserves in developed countries, creating new threats to the resources of developing countries (Benchimol et al., 2009).

In the present study, DNA barcoding was adopted to better address the taxonomic concerns of *G. vicinus* at the species level. On the one hand, it is about a new observation with its distribution in the islands of Cabo Verde; on the other hand, it is a question of qualifying this species on the basis of morphological criteria and DNA barcoding and of correcting the current COI sequences of this species published in GenBank. However, the results will show the need to be careful when identifying moray eels and will promote fisheries management, biodiversity conservation, and sustainable management of this species.

1.2 Problem Statement

The difficulties related to the observation and collection of individuals free-living have prevented the establishment of mark-recapture studies, thus inhibiting the assessment of population size, population structure, population variability, and demographic traits. Then, these difficulties can be avoided by genetic markers (Ribout et al., 2018).

The species *G. vicinus* exhibits unique karyotypes, which nevertheless show the pattern of the family Muraenidae with a high number of acrocentric chromosomes, a rare condition in other Anguilliformes. One exception is the karyotype of G. miliaris (2n = 42) that presents the highest fundamental numbers (FN) reported so far in Muraenidae (FN = 84), likely due to pericentric inversions (Vasconcelos et al., 2009).

Structural complexity within one or more habitats not only serves to prevent resource overlap; but can also play a role in improving quality of life by allowing people to access

resources against predators, in particular smaller organisms that may be more vulnerable. Thus, the extent to which habitat complexity mitigates trophic interactions, possibly decreasing predation and competition, may depend on various factors, such as shelter availability, behavioral attributes of interacting organisms, and stages of development of predator or prey species during interaction (Higgins et al., 2017).

The *G. vicinus* has two different colors variation, and these also indicate that probably some genetic variability in their population is present: grey above, pale below; dorsal and fins edged in Black; pectoral fins dark grey or black distally, Figure 2 showing a greenish mottling on the back (FAO, 1981). Thus, DNA analysis can be an effective identification tool in eels or other fish, as shown for leptocephali (Tawa et al., 2012).

Jespersen in 1942, showed that the morphology of the species of anguillid leptocephali of the Indonesian seas was examined in detail, but no character was found, allowing to formally separate the species of anguillid in the region. This problem has been solved recently by the use of molecular genetic techniques, which allow identifying with certainty any species of leptocephali anguillid using methods that have been developed specifically for the identification of anguillid eels. Identification at the generic level of leptocephali in Muraenidae has been shown to be very complicated due to the complexity of muraenid taxonomy. The most important factor established so far for the determination of most types of leptocephali within the species is the total number of myomeres, which generally represents the total number of vertebrae in adults (Sam Wouthuyzen & Mochioka, 2014).

According to the inaccurate identification of species and a decrease in management, this sometimes results in a subsequent increase or decrease in the level of accepted species in the marine ecosystem (Sam Wouthuyzen et al.,2014). Changes in oceanographic conditions can lead to changes in the structure of the ichthyoplankton community due to the effect this change would have on the abundance and species composition associated with it (Souza et al., 2019).

1.3 Research Questions

Here are listed the main questions that will be answered in this study:

- Are there differences in size composition, genetic markers, and condition indices that could indicate the existence of local subpopulations?

- Do interannual changes in size composition indicate a fishing impact on the population? Knowing population structure will help to develop management concepts for this species to enable long-term sustainable fisheries.

1.4 Relevance and Importance of the Research

The stock concept as the backbone of fisheries management to prevent overfishing of a population is important to indicate the relevance of our research. Fished on all islands, it is important to understand whether all *G. vicinus* belong to one unit or if subunits exist on all islands, which have to be managed individually. Therefore, we conducted a series of morphometric analyses and genetic analyses.

1.5 Objectives of the work

The main objective is to study the population structure of *G. vicinus* (Castelnau, 1855) sampled in different locations on Cabo Verde Northwest islands from 2002 to 2020, based on morphometric characteristics and genetic markers.

1.5.1 Specific objectives

We have three specific objectives of this research:

- Analyse and compare length-frequency distributions of *G. vicinus* in the time series from 2002 to 2020 collected in the north of Cabo Verde provided by IMAR to indicate if different size and growth patterns exist.
- Analyse length-weight relationships from different places and sampling periods.
- Compare the genetic diversity between sampling places.

1.6 Structure of work

This thesis is organized into 6 main chapters. The background, the problem statement, objectives, as well as the questions this study is trying to answer, are presented in the first chapter; the second chapter describes the general concept of this study revealed in the literature; chapter three describes the detailed methods used to achieve the objectives of the study; the results of the study are presented in the fourth chapter, and discussed in chapter five. The last chapter concludes the study and gives recommendations based on the findings.

2 Literature review

2.1 General overview

The Anguilliformes, the "real eels," are a very diverse group from an environmental point of view, mainly of marine origin, whose individuals are easily recognizable by their particularly elongated body, the rounded cross-section, and the absence of pelvic fins (Peninal et al., 2017). The Muraenidae family, whose members are commonly referred to as moray eels, is one of the largest and most diverse in the Anguilliformes. They are found all over the world in tropical and subtropical areas. Their presence manifests itself from the coast to the edge of the continental shelf, and they occupy an important place from an ecosystem point of view in the habitat of coral reefs, where they reside in holes and crevices, coming out to feast on a great diversity of fish and invertebrates (Smith, 2012).

Muraenidae are well known from the Cabo Verde islands in the eastern Atlantic, where they occur regularly, according to unpublished data amongst others from the Institute of Marine Research Mindelo and published data, although not always species-level identification was provided (Brito et al., 2014). Despite the abundant documentation on the life cycle of the ichthyofauna and their related environments in neritic waters of coastal areas, information on the more sedentary and cryptic vertebrate species of the kelp forest and reef ecosystems is generally lacking and is biased towards better visible species (Higgins & Mehta, 2017).

2.2 Habitat, biology, and fisheries of the species

Morays are generally considered nocturnal predators due to their relatively small eyes and well-developed scent detection system. However, some moray eel species have been reported to forage during the day, relying on their eyes (Wang et al., 2011).

They usually inhabit shallow rock habitats, areas of coral reefs, and seagrass beds. They are found at depths often up to 40 meters and can be very aggressive and even dangerous when treated alive. They are very popular and consumed locally. Large specimens are potentially ciguatoxin, which means they carry certain toxins only known to reef fishes. The toxin ciguatoxin is produced by small unicellular algae which are consumed by herbivorous that are eaten by carnivores. Fishing is done by trawl, trap, and line (Kent E. Carpenter, 2002).

2.3 Species Distribution

The *G. vicinus* is widely distributed in the western Atlantic, from Bermuda to the Bahamas, off North Carolina, the Florida Keys, as well as the northern Gulf of Mexico, all of the Caribbean islands, and off the central coasts and northern parts of South America. It is present in the central-eastern Atlantic, off Madeira, the Cabo Verde archipelago, and the bay of Biafra (FAO zone 34). Off Brazil (FAO zone 41) and in the southeast Atlantic off Ascension island the southeast Atlantic off Ascension island (FAO zone 47) (Kent E. Carpenter, 2002).

In comparing Cabo Verde with other Macaronesian archipelagos, it was found that the zoogeographic distribution of its coastal ichthyofauna is generally tropical, with Afrotropical species largely predominant, followed by tropical and subtropical amphi-Atlantic species, then by Mediterranean species and circumtropical (Freitas, 2014). In turn, the other archipelagos of the Macaronesian system lack the tropical moray eels (Gonzalez et al., 2021). Thus, with regards to Macaronesian moray eels, the highest diversity is observed for Cabo Verde hosting 18 species, while the Canaries, Madeira, and the Azores have 10, 5, and 8 species resp (Gonzalez et al., 2021). However, studies made by (Brito et *al*. 2017) have shown that in the surrounding Canary islands, tropicalization mechanisms are important factors that help to enhance fish diversity. The presence of *G. vicinus* precisely in El Hierro island (Canary islands) (Falcón et *al*., 1996) allows us to suspect a recent colonization mechanism linked to the rise in temperature in Canary islands waters (tropicalization), as has been done in other cases of tropical fish with a strong swimming ability or highly specialized planktotrophic larvae (Brito et *al*., 2005).

In a comprehensive inventory of Cabo Verdean fish, Reiner (1996) listed 520 species of fish, an important publication with remarkable indications of Cabo Verdean fish but also containing many incorrect records. Several genera of moray eels are found in the Atlantic, and the widespread genus *G. vicinus* is found in all ocean basins. It remains unclear whether the presence of *G. vicinus* species in the Atlantic represents numerous invasions from the Indo-Pacific or a single invasion followed by speciation in this region (Reece et *al.*, 2010). Moray eels (Osteichthyes: Muraenidae) are diverse in terms of species, distributed worldwide in tropical and temperate seas, from intertidal waters to deep waters. Adults are benthic and most often reside in shallow water among rocks and coral heads. Of all the species, many are more

active at night and hide in holes, crevices, and other shady places during the day (González et *al.*, 2021). The *G. vicinus* (Castelnau 1855) inhabits coastal waters in the eastern and western Atlantic (Smith 2012) and is considered an amphi-Atlantic warm water species (Gonzalez et al. 2012). The currents in the tropical Atlantic and around Cabo Verde make it possible to exchange faunal elements both with the Canaries, the Azores and Madeira on the one side and the western Atlantic on the other side: the Canary current coming from the North and the North Equatorial Counter Current coming from the West. Thus, exchange and isolation processes (endemism) can drive processes to establish local faunas in the area.

The Cabo Verde archipelago has strong sparid endemism, attributed to multiple broadcasts, by populations with different trophic ecologies (Santini et *al.*, 2014). Cabo Verde has been mentioned several times in the area of pan-Atlantic marine connectivity, largely in relation to a distinct American (mainly Caribbean) footprint, expressed by the presence of algae from the West Atlantic, but also by fish species such as moray eels (Muraenidae) (Lopes et *al.*, 2021a). Only low levels of endemism exist in Macaronesia as a whole, which is probably related to the distance between these archipelagos and mainland Africa and Europe (Almeida et *al.*, 2017). It was newly identified in an area where coastal fish are frequently studied, possibly due to increasingly favorable environmental factors, which has led to an intensification of the population of *G. vicinus*. This mechanism had already been noticed in the Canaries among other species of thermophilic fish (Laguna, 2016). However, LIoris et al. (1991) believed that, from an ichthyological point of view, the ichthyological fauna from the Macaronesian archipelagos "do not present a coherent marine biogeographic region and cannot be considered as a province."

2.4 Identification methods

The studies made at the level of ichthyofauna communities have provided information on several biological characters of the species, making it possible to highlight the evolution of the spatial and temporal distribution of the species (Stratoudakis et al., 2003), also establishing preferred zones and seasons (Faria et al., 2006). Phylogenetic studies on muraenids are essentially based on morphometric analyzes and vertebral enumeration methods; very rarely have mitochondrial DNA sequences been analyzed (Loh et al., 2008).

Even among experienced taxonomists, the correct identification of species is often very complex, which is why most morphological descriptions are based only on adults (Almeida et *al.*, 2017). DNA barcoding based on the cytochrome oxidase1 (COI) gene sequence is

considered to be a commonly used molecular tool for species identification (Jaonalison, 2019). It is no wonder that poorly identified DNA barcodes from these species were found in GenBank, calling for correct identifications (Li et *al.*, 2018).

Recently, Tautz et *al*. (2002, 2003) justified the adoption of a taxonomic system based on DNA. Thus, DNA sequence analysis has been applied for about 30 years to aid species identification, but various sequences have been used for various taxonomic groups and in different laboratories. Hebert et *al*. (2003) estimated that a single species gene sequence would be effective in differentiating all animals, or at least the vast majority of them, proposed the use of the mitochondrial DNA cytochrome oxidase subunit I (cox1) gene as an overall animal bioidentification system.

3 Materials and Methods

3.1 Study area

The Republic of Cabo Verde is a country in West Africa comprised within the limits of parallels 17 ° 12 'and 14 ° 48' north latitude and meridians 22 ° 44 'and 25 ° 22' longitude and 25 ° 22 'west longitude (Government of Cabo Verde; Ministry of Environment and Rural Development, 2010). It is made up of 10 islands, on which nine are inhabited, and several uninhabited islets, separated into two groups by their arrangement in relation to the prevailing winds, the Barlavento (windward) group in the north, including from west to east the islands of Santo Antão, São Vicente, Santa Luzia (uninhabited), São Nicolau, Sal, Boavista, and the Sotavento group (leeward) in the south, formed from east to west and west to east, the group in the south, formed from east to west by Maio, Santiago, Fogo, and Brava.

The Cabo Verde belongs to the group of volcanic islands known as the Macaronesia region (Azores, Madeira, Canaries, Cabo Verde Salvages, and the North East coast of West Africa, from Morocco to Senegal) (Benchimol et al., 2009). The total area of the Cabo Verde archipelago is 4033 km². The islands are of volcanic origin, rising from a depth of at least 3 000 m, and the continental shelves, generally narrow and irregular, are limited to a total area of 5 394 km². The eastern islands Sal, Boavista, and Maio, form one system with a more extensive continental shelf compared to the other islands (Figure 1). However, the specific areas where the data were collected are:

- São Vicente (Portuguese for "Saint Vincent") is one of the Barlavento Islands, the northern group within the Cabo Verde archipelago in the Atlantic Ocean, off the west African coast. It is located between the islands of Santo Antão and Santa Luzia, with the Canal de São Vicente separating it from Santo Antão;
- Santa Luzia is an island located between São Nicolau and São Vicente. The channel of Santa Luzia separates the island of São Vicente and is 8 km wide;
- Santo Antão (Portuguese for "Saint Anthony") is the westernmost island of Cabo Verde. At 785 km2 (303 sq. mi), it is the largest of the Barlavento Islands group and the second-largest island of Cabo Verde. The nearest island is São Vicente to the southeast, separated by the sea channel Canal de São Vicente;
- Boa Vista is an island in Cabo Verde, off the coast of West Africa. It is known for its sand dunes and the lunar volcanic landscapes of the Viana Desert. The white sand

beaches include those of Tortuga and Santa Monica. The remains of a 19th-century fort stand on an islet on the other side of the town of Sal Rei;

- Maio is a small island in Cabo Verde, a volcanic archipelago nation off the coast of West Africa. It's known for the sandy beaches around its main town of Vila do Maio and north, near the village of Morrinho.

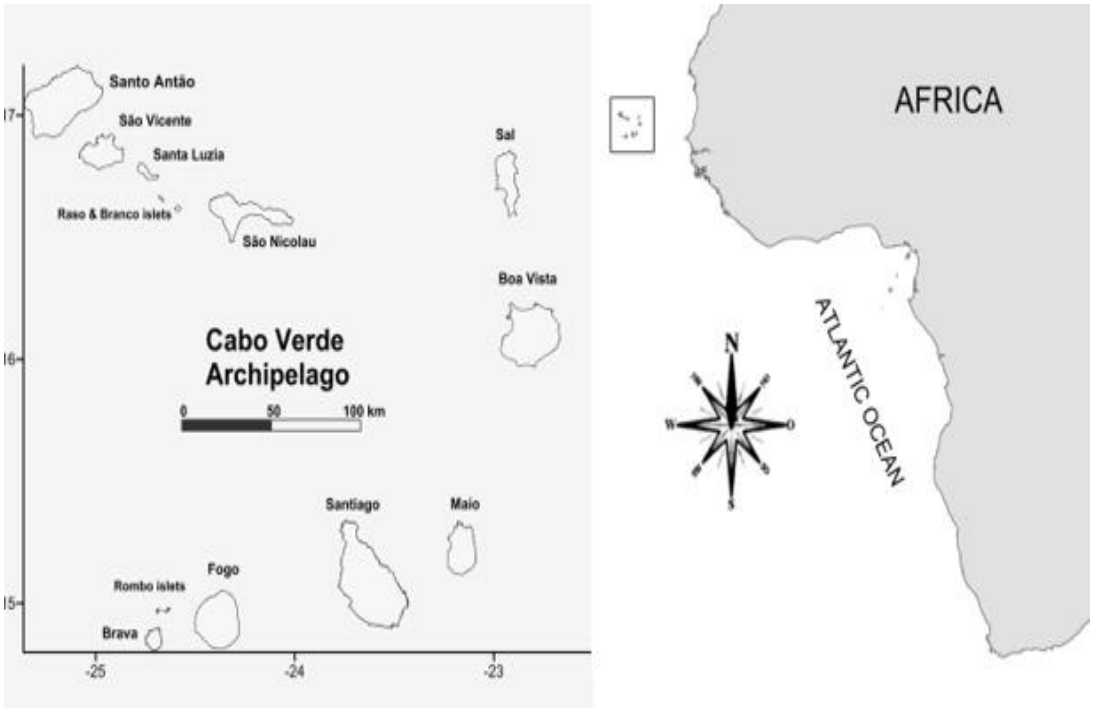

Figure 1: Geographic localization of Cabo Verde Archipelago map adapted from Lopes et al. 2019.

Although the continental shelf area is of limited area, the EEZ of Cabo Verde covers an area of about 789 400 km², much of which is not exploited by the national fisheries. The main Canary Current cues direction down to the Cabo Verde islands creates: (1) exposed rocky shores with some cliff's enclaves with major hydrodynamic forces at the north, northeast by trade winds, and (2) rocky-sandy shallow areas sheltered or moderate exposed in the opposite side of islands configurations but receiving energetic seasonal south swells. Reef fish assemblages in Cabo Verde are amongst the most important due to their relative abundance and biomass availability in coastal areas, being however balanced against the low catches, specifically for demersal species along the West African coast. On the integrative Catalog of the fishes of the Cabo Verde islands listed about 520 species but includes some old and erroneous records (Benchimol et al., 2009).

3.2 Material

3.2.1 Sample collection

As a relatively important component of Caboverdean fisheries, *G. vicinus* is caught throughout the year and all over the archipelago by artisanal fishermen. It is caught mainly by handline and sold in local markets. It has also been caught in experimental longline fisheries and sporadically in trawl fisheries (González et al., 2021). Artisanal fishermen have historically used the same size of hook (7-9 cm) for catching this and other demersal fish species. However, due to their importance to the artisanal fisheries in Cabo Verde, a biological sampling scheme was implemented for a group of demersal species, aiming to determine the exploitation status of the stocks in order to provide advice on their optimal exploitation and management (Pastor, 2002).

For this study, the samples were carried out at the local fishing port or at fish dealers. The fish were purchased when the fishermen landed their products. The survey was organized twice a week, corresponding to the different sampling locations. The sampling was performed during the first 15 days by month in a landing site. Individually 30 *G. vicinus* were randomly selected at the time of sampling for detailed biological sampling at the laboratory. The samples were bought if necessary in the different fish landing markets from different islands such as São Vicente, Maio, Santo Antão, and Calhau for the genetic analysis. Times series data on sizes, sexes, gonads, and weight were available from 2002 to 2020 by courtesy of IMAR (Dr. Vito Ramos, personal communication).

3.2.2 Genetic material sampling

The collection of specimens for DNA extraction took place in the three islands of the Cabo Verde archipelago (Santo Antão, São Vicente, and Santa Luzia). The samples from São Vicente (10) and Santa Luzia (10) were acquired from the tissue collection of ISECMAR, which were collected by IMAR in 2015. Fourteen additional specimens were obtained from samples collected at Santo Antão fish market (4) and São Vicente (10) for morphometric characterization, photography, and complete data set (Table 1). Muscle tissue samples were deposited in the newly established biodiversity reference collection at the ‘ISECMAR at Mindelo, Cabo Verde. The GPS was used to localize each of the sampling sites during the data collection.

Table 1: Geographical localization of samples collection (GPS)

COUNTRY	PLACE	COD	BOLD / GENBANK CODES	GPS	
				LAT	LONG
Brasil	Bahia	BAH1	MF999186	-12.93	-38.51
	Bahia	BAH2	HG3381	-13.8894	-38.9429
	Bahia - Ilha de Tinhare	BAH3	HG3795	-13.474	-38.914
	Bahia - Ilha de Tinhare	BAH4	HG3793	-13.475	-38.915
	Bahia - Ilha de Tinhare	BAH5	HG3794	-13.475	-38.915
Belize	Stann Creek District	BEL1	JQ841585 / BZLW8189	16.7598	-88.0761
	Reef crest in front of Carrie Bow Cay	BEL2	KUT 153	16.783	-88.067
Mexico	Othon P. Blanco	MEX1	MX188 / GU225293	18.271	-87.835
	Gulf of Mexico	MEX2	MH777674	-	-
	Gulf of Mexico	MEX3	MH777651	-	-
USA	USA	USA1	MF041617	-	-
Cabo Verde	Santo Antão	VA05	GVIC (01-04) VA05	17°12'17.59"N	25° 5'38.72"W
	São Vicente	VV04	GVIC (01-10) VV04	16°51'18.02"N	24°51'18.93"W
	Santa Luzia	VL01	GVIC (01-10) VL01	16°47'47.71"N	24°44'59.35"W
	Santa Luzia	VL06	GVIC (01-06) VL06	16°47'10.32"N	24°48'10.94"W

3.2.3 Methods

3.2.4 Species identification

The morphological identification (Figure 2) characters were based on the FAO guide such as Body strong, muscular, moderately compressed. Head with occipital region somewhat elevated; posterior nostril a simple pore without tube before the upper edge of eye; teeth shark-like in appearance, with strong serrations along both anterior and posterior margins, uniserial throughout. Dorsal fin originating on the head a short distance before gill opening, posterior to branchial lateral line pores. The total number of vertebrae ranges from 150 to 156.

Colour: overall color light yellowish to medium or dark brown (Figure 2), everywhere covered with fine spots and close-set on the head, larger on body and fins appearing dark in light individuals, light in dark ones; head noticeably dusky. Fins with light yellowish margins (FAO, 1981).

Figure 2: Morphological identification, *G. vicinus* light yellowish (A); *G. vicinus* Perplemouth dark brown (B).[1]

3.2.5 Morphometric analysis

The morphometric weight measurements have been taken at the level of each fish (Figure 3 (C)) using a balance. The ichthyometer (ruler) was used for the length measurement (Figure 3 (B)). These two measurements allowed us to determine the Weight-Length relation.

The dissection was done (Figure 3 (D)) to determine their sex, sex ratio (ratio of female to male), and the stage of maturity of their gonad. The species samples have all been counted and separated into different bags for identification. A digital camera was used to take pictures of the different samples in millimetric paper with a ruler. The morphometric measurement has been taken individually to measure the total length, and an ichthyometer was used for this measurement.

The length was expressed in terms of Total Length (TL) to the nearest centimetre, and the weight was measured in gram (g). The individual TL to the nearest centimetre, using 1-cm length groups, and the individual weights are recorded in Figure 3. Sex and maturity stages are determined macroscopically. Four stages of maturity are used: immature (1), maturing or developing (2), ripe, pre-spawning (3), and spawning, post-spawning (4). Finally, the gonad weight, in grams, is recorded.

[1] https://www.fishbase.se/summary/7548

Figure 3: Sampling of *G. vicinus* in the laboratory. A) Fish samples from São Vicente island; B) Determination of individual total length; C) Determination of individual weight; D) Dissection for the biological purpose: Specie information: weight 735.47g.

3.3 Condition factor and Length-Weight Relationship (LWR)

Weight-length relationships (WLR) are used to estimate the relative weight at a given length, and condition factors are intended to compare between locations or sampling periods the condition, fat, or well-being of fish, making the hypothesis that heavier fish of a given length are in better condition. It has been noted that the variation in weight at a given size in the same among species increases very much as the fish grows in length (Froese, 2006). The condition factor (K) of individuals within and across sites was analyzed. The condition factor (K) of the experimental fish was derived from the length-weight relationship:

$$W = aL^b \qquad \text{(Eq. 1)}$$

$$Log\ W = Log\ a + b\ \times b\ Log\ L \qquad \text{(Eq. 2)}$$

Where *b* ranges from 2.5 to 3.5 (Froese, 2006).

For Fulton K as condition index, we reduce the two unknown parameters to one unknown parameter and a fixed exponent of 3 for the length-weight relationship (Eq.1) and its logarithmic form (Eq. 2): Fulton K has obtained 100 times the weight of fish in grams divided the total length of the fish in centimeters to the power of 3 (Eq. 4):

$$W = aL^3 \qquad \text{(Eq. 3)}$$

$$K = a \times 100 = 100 \times \frac{W}{L^3} \qquad \text{(Eq. 4)}$$

Where W = weight of the fish in grams, L = the total length of the fish in centimeters, b = the exponent value obtained from the length-weight equation formula (Froese, 2006).

3.3.1 Maximum size (Lmax)

The maximum size was analysed from the time series as a fisheries indicator in terms of an indicator-based assessment. At a population level, the removal of large fish may be reflected in changes in mean length or weight of a population and/or in terms of maximum length (Shin et al., 2005).

3.3.2 Statistical Test

In this section, we will describe all the Tests that have been done in this thesis regarding the morphometric comparison. The first part will be talking about the Cramer Von-Mises test; in the second section, we will describe the Loess smoothing, and then in the third section, we will be talking about the significance test for means.

3.3.3 Cramer Von-Mises test

This is for testing differences in length-frequency distributions (LFDs) between two locations. The Cramer-von Mises two-sample test (CvM) is one of the best-known distribution-free two-sample tests and more accurate than the Kolmogorov–Smirnov test (Anderson, 1962). The Test is based on the difference between two empirical distributions function (EDF) of cumulative proportions x in size classes of two samples N, M (here L_cm in São Vicente and Santa Luzia). CvM is, as the EDF test, insensitive to changes in abundance in N, M but not to changes in distribution parameters.

First, we have created subsets for two locations, i.e., the place São Vicente and Santa Luzia. For comparisons, we could only choose comparable data, i.e., same month and year! We have used the Cramer Von-Mises test for two samples.

- **Loess smoothing**

The Loess smoothing has been used for the local regression, which is a non-parametric approach and allows us to analyze things that are not linear model skills. It was used for calculating the predicted values. Many nonparametric regression estimators have been deduced, and their practical effect is well established. Local regression, also known as moving regression, is a generalization of moving average and polynomial regression. Its most frequent methods, which were initially developed for point cloud smoothing, are LOESS (locally estimated scatterplot smoothing). LOESS combines much of the simplicity of linear least squares regression with the flexibility of nonlinear regression (Fox et al., 2018).

3.3.4 Significance tests for means

In the case of normality, Analysis of Variance (ANOVA) as a collection of statistical models and their associated estimation procedures (such as the "variation" among and between groups) is used to analyze the differences among means. In the case of non-normality distributed data, the Wilcoxon Rank Sum Test was applied.

QQ- plot, histogram plot, and Shapiro Wilks test have been applied to test for normality of the investigated variable. When N is relatively large, the sampling distribution will be approximately normal due to the Central Limit Theorem (CLT). Rank methods analyze ranks derived from quantitative data. Examples of nonparametric rank-based methods are Wilcoxon's test (cf. two-sample t-test), Kruskal-Wallis test (cf. one-way ANOVA), signed Wilcoxon test, and Spearman's correlation for classified data.

When the sample size N is large, the limited variability between the results derived from this set of methods suggests that the linear model is well specified. In the second set of robust approaches (e.g., winsorizing and fitting), non-normality due to outliers is taken as an indication of data contamination. This data contamination is addressed by modifying or eliminating extreme data points (Pek et *al.*, 2018). In this thesis, outliers in the analysis of length-weight relationships were removed because it was assumed that weighing errors and transcription mistakes could have taken place.

3.3.5 Genetic analysis

3.4 Extraction of DNA

DNA extraction was performed from samples collected in the tree islands, Santo Antão [four samples; collection code GVIC (01 to 04)VA05], São Vicente [10 samples, collection code GVIC (01 to 10)VV04] and Santa Luzia [16 samples, collection code GVIC(01 to 10)VL01; GVIC(01 to 06)VL06. The total genomic DNA of each individual was extracted from fragments of approximately 5 mg to 10 mg of tissue which has been stored in ethanol, using a salt method adapted from (Sambrook et al., 2001). Samples were incubated in 600µl of extraction buffer (0.05 M Tris, 0.1 M EDTA, 2% SDS) and 20 µl of Proteinase K (20mg/ml), for 4 hours at 65°C, in a closed 1.5ml tube. The solution containing the DNA was separated from the other cellular constituents with 200 µl 5M NaCl and centrifugation for 2 minutes at 11000rpm. The supernatant (+- 780 µl of extracted DNA) was transferred to a new 1.5ml tube, and the DNA was precipitated in 700µl of 100% alcohol ice-cold. Cleaning of the DNA was done by centrifugation for 15 minutes at 14000rpm (or maximum speed) and washing with 800 µl of 70% ethanol ice-cold. The pellets were dried at room temperature for 4 hours (minimum) and suspended in 150 µl ultra-pure H2O (Lopes et al., 2021b).

3.4.1 DNA amplification and sequencing

DNA amplification was achieved through polymerase chain reaction (PCR) for one mitochondrial fragment. Universal primers of (Ward et al., 2005) were used in PCRs to amplify about 600 bp of the mt cytochrome oxidase subunit I (COI) gene (FishF2: TCGACTAATCATAAAGATATCGGCAC; FishR2: ACTTCAGGGTGACCGAAGAATCAGAA. All PCR amplifications were conducted in 25-µl using Phusion High-Fidelity PCR Master Mix with HF Buffer, following the manufacturer's specifications, namely, 2 µL of genomic DNA, 1X PCR Master Mix, and 0.25 µM of each primer.

The cycling conditions started with an initial denaturation of 98°C for 30 s, followed by 45 cycles of 98°C for 30 s, followed by 45 cycles of 98°C for 30 s, 53°C for 30 s, 72°C for 30 s and a final extension at 72°C for 10 min. For quality check, 2.5 µL of each PCR product were visualized in agarose gel 1% stained with SERVA DNA Stain Clear G and observed in a transilluminator under UV light. Sequencing was performed on an automated Sanger

sequencer, and the results were compared with the sequences available in the GenBank using Geneious (Drummond et al., 2006).

3.4.2 Haplotypes (genetic) and nucleotides (π) diversity

The haplotype diversity index (Hd) indicates the probability that two haplotypes chosen at random in the sample set are different, and together with the variance were estimated from equations 5 of Nei, 1987. This value varies between 0 and 1, 1 being when haplotypes are 100% different. The values were determined using the DNASP 5.00.07 program (Rozas et al., 2003). The equations can be reduced to the following expressions:

$$h = \frac{n(1 - \sum xi^2)}{n - 1} \quad \text{(Nei, 1978)} \quad \text{(Eq. 5)}$$

Where n is the number of sequences, xi is the frequency of haplotype i in the sample, xj is the frequency of haplotype j in the sample.

The nucleotide diversity index (π) indicates the average number of nucleotide differences per site between pairs of haplotypes (Saitou & Nei, 1987) was estimated from equations 6 (Nei, 1987). The closer value to 0 indicates that if you catch two nucleotides at the same position, they have a high probability of being equal.

$$\pi = \frac{n}{(n-1)\sum xixj\pi ij} \quad \text{(Nei, 1978)} \quad \text{(Eq. 6)}$$

3.4.3 Sequence analysis

The edited sequences were compared with available sequences in Genbank by the BlastN tool (National Center for Biotechnology Information U.S. National) using default settings. The blast output was used to create a taxonomic classification. The translation of mitochondrial genes has been deduced using the mtDNA genetic code of vertebrates in Geneious 4.7 (Drummond et al.,2006), allowing inspection of the existence of stop codons and thus the presence of pseudogenes. The global alignment is also done using the Geneious. For phylogenetic analyses, COI sequences of species deposited in GenBank were included.

Haplotype genealogies using parsimony were reconstructed with TCS1.21 (Clement et *al.*, 2000), with the final manipulation in tcsBU (Múrias Dos Santos et *al.*, 2015), Pairwise distance values for all haplotypes were estimated above the 95% level to create the network. Genetic differentiation, the haplotype, and nucleotide diversity were done using DnaSP5

(Rozas et *al.*, 2003). For phylogenetic analyses, COI sequences of species deposited in GenBank and BOLD System were included.

4 Results

4.1 Data availability

4.1.1 Summary by island, location, and period

Most of the observations are coming from Santa Luzia (Table 2). In Table 2, islands names and locations on the islands are mixed. Boavista, Maio, and Santa Luzia are islands names. Locations names on São Vicente name are: Baía das Gatas Calhau, Salamansa, Lazareto, São Pedro, Canal de São Vicente; location names on Santo Antão: Norte de Santo Antão, Noroeste, Tarrafal de Monte Trigo; location names on Santa Luzia: Canalzinho, Colador de Santa Luzia.

Table 2: Number of total observations by locations

Sampling sites	Number of observations
Baía das Gatas	289
Boavista	60
Calador	31
Calhau	623
Canal de São Vicente.	29
Canalzinho	656
Colador Santa Luzia	8
Inconnu	29
Lazareto	4
Maio	180
Noroeste	21
Norte de Santo. Antão	30
Salamansa	651
Santa Luzia	6003
São Pedro	251
Tarrafal de Monte Trigo	1232

The sampling by islands shows us that most of the samples were collected from Santa Luzia, followed by São Vicente and Santo Antão, and the lowest number was observed in Boavista (Table 3).

Table 3: Numbers of total observation by islands

Sampling islands	Numbers of sampling
Boavista	60
Maio	180
Santa Luzia	6013
Santo Antão	1283
São Vicente	1847
São Vicente/Santa Luzia	656

Data were collected during all months by year, but most of the data were collected in October and November; the lower number of cases was collected in August (Table 4).

Table 4: Monthly data collection

Monthly sampling	Number of samples
1	553
2	752
3	884
4	952
5	824
6	877
7	808
8	412
9	854
10	1145
11	1118
12	918

Tabling data by island and year shows that the highest number of observations by means collection was registered in Santa Luzia island in 2004. Santa Luzia received the highest sampling effort in each year. The lowest number of cases was undertaken in Bovista, with almost null observations in most of the years except 60 observations in 2003 (Table 5).

Table 5: Number of observations by year and by islands

Years	Boavista	Maio	Santa Luzia	Santo Antão	São Vicente	São Vicente/Santa Luzia
2002	0	0	142	0	52	0
2003	60	150	331	0	262	25
2004	0	30	1278	144	558	15
2005	0	0	258	109	220	31
2006	0	0	85	102	98	0
2007	0	0	147	177	30	0
2008	0	0	263	30	117	30
2009	0	0	319	117	25	0
2010	0	0	330	60	30	60
2011	0	0	299	48	90	30
2012	0	0	232	144	30	60
2013	0	0	497	30	30	30
2014	0	0	380	30	87	0
2015	0	0	434	59	113	0
2016	0	0	203	58	58	179
2017	0	0	329	90	0	93
2018	0	0	90	0	0	58
2019	0	0	214	85	47	30
2020	0	0	182	0	0	15

4.1.2 Selection of data for monthly comparison of length composition

Table 5 shows that 3 islands (Santa Luzia, Santo Antão, and São Vicente) have a higher number of observations (n >100) in several years. For these 3 islands, the availability of monthly data is investigated. However, the main comparison is focused on these 3 islands, which we can not compare with those islands or locations that have observations lesser than 100.

4.1.3 Santa Luzia

The experiment shows that in this table 6, the data are not evenly distributed accordingly by month and by year. The data were well distributed all months for the years 2004 and 2015 (Table 6). However, in those years, such as 2002, 2005, 2006, 2007, and 2018, the number of observations varies from 2 to 5, which means that there are no more data available for these years.

Table 6: Distribution of data by year and month for Santa Luzia

Years	Months											
	1	2	3	4	5	6	7	8	9	10	11	12
2002	0	0	0	0	0	0	24	29	20	48	21	0
2003	37	0	0	30	29	90	30	30	22	63	0	0
2004	49	101	63	105	142	130	98	71	78	134	197	110
2005	0	62	59	30	0	0	0	0	17	60	30	0
2006	0	0	0	0	0	0	0	0	30	0	55	0
2007	0	0	0	0	0	0	0	0	30	58	59	0
2008	29	30	57	30	30	29	30	0	0	0	0	28
2009	29	29	58	0	29	26	0	28	0	60	30	30
2010	30	60	60	30	0	30	30	0	30	30	30	0
2011	0	0	30	59	0	0	60	0	30	60	30	30
2012	0	29	29	30	30	0	0	30	0	30	0	54
2013	30	59	60	28	60	59	30	30	0	30	51	60
2014	59	0	0	30	60	30	55	0	0	30	59	57
2015	29	60	30	25	30	20	30	30	30	60	30	60
2016	0	26	58	30	30	30	0	0	0	0	29	0
2017	30	0	30	60	60	0	30	30	30	0	30	29
2018	0	0	0	0	30	0	0	0	30	0	0	30
2019	0	30	0	60	30	30	0	0	0	23	18	23
2020	0	0	16	0	20	21	20	15	21	19	12	38

4.1.4 São Vicente

The table 7 tells us that the data were not well distributed by month and by year in São Vicente island, the only year well distributed monthly is 2004, and the lowest observation was seen respectively in years 2007, 2010, 2012, and 2013 which is 30 observations. No observations in 2020 (Table 7).

Table 7: Distribution of data by year and month for São Vicente.

Years	Months											
	1	2	3	4	5	6	7	8	9	10	11	12
2002	0	0	0	0	0	0	52	0	0	0	0	0
2003	0	0	0	0	0	57	57	30	62	0	26	30
2004	0	108	51	167	44	55	21	0	22	44	36	10
2005	20	0	0	0	30	60	30	0	0	0	29	51
2006	0	18	0	0	0	0	24	0	27	29	0	0
2007	30	0	0	0	0	0	0	0	0	0	0	0
2008	0	0	0	0	0	0	0	0	57	60	0	0
2009	0	0	0	25	0	0	0	0	0	0	0	0
2010	0	0	0	0	30	0	0	0	0	0	0	0
2011	30	30	0	0	0	0	0	0	0	0	30	0
2012	0	0	0	0	0	0	0	0	30	0	0	0
2013	0	0	0	30	0	0	0	0	0	0	0	0
2014	0	0	0	0	0	30	0	0	30	27	0	0
2015	0	0	24	30	0	0	0	0	30	0	29	0
2016	28	0	0	0	0	30	0	0	0	0	0	0
2019	0	0	30	0	0	0	0	0	0	0	17	0

4.1.5 Santo Antão

On the Table 8 are shows that there is no monthly coverage data from Santo Antão. Most of the observation is shown in 2007 with 5 observations; the years 2008, 2011, 2013, and 2014 have only 1 observation.

Table 8: Distribution of data by year and month for Santo Antão

Years	Months											
	1	2	3	4	5	6	7	8	9	10	11	12
2004	0	0	0	0	0	0	44	0	0	0	0	100
2005	14	28	0	0	25	0	0	0	42	0	0	0
2006	21	23	58	0	0	0	0	0	0	0	0	0
2007	0	29	55	0	37	0	26	0	0	0	0	30
2008	30	0	0	0	0	0	0	0	0	0	0	0
2009	0	0	0	30	0	0	28	0	0	0	30	29
2010	0	0	0	30	0	0	0	0	0	0	30	0
2011	0	0	0	0	48	0	0	0	0	0	0	0
2012	0	0	30	29	0	0	0	0	30	0	55	0
2013	0	0	0	0	0	0	0	0	30	0	0	0
2014	0	0	30	0	0	0	0	0	0	0	0	0
2015	0	0	0	0	0	30	29	0	0	0	0	0
2016	28	0	0	30	0	0	0	0	0	0	0	0
2017	0	0	0	0	0	30	0	0	0	30	0	30
2019	0	30	0	0	0	0	0	0	21	13	21	0

4.2 Comparison of Length-Frequency-Distributions (LFD)

Length-frequency distributions are compared pairwise for the 3 islands with sufficiently available data. However, to compare two islands or locations, they must have the same condition accordingly to the available data, same months and years.

4.2.1 Comparison for Santa Luzia and São Vicente in 2003, 2004 and 2015

The Cramer Von Mises-Test for Santa Luzia and São Vicente, respectively, in 2003, 2004, and 2015 is observed in Table 9. The results showed that the lowest number of observations has a P-value of p = 0.0002, which means that the Test is significantly different; the higher number of observations has a P-value of p = 0.840, by means, the test is not significantly different between these islands (Table 9).

Table 9: Cramer Von-Mises Test of Santa Luzia and São Vicente.

Years	Months selected	Number of observations São Vicente	Number of observations Santa Luzia	P-value (cvm-Test)
2003	6, 7, 8, 9	206	172	0.0002
2004	2, 3, 4, 5, 6, 7, 9, 11, 12	558	1158	0.840
2015	3, 4, 9, 11	113	115	0.002

The Cramer von Mises Test has concluded that there is no significant difference between Santa Luzia and São Vicente in 2004. Then, the size range was from 60 cm to 120 cm for São Vicente, whereas for Santa Luzia, two-class sizes were not observed, i.e., and 60 cm to 110 cm. For Santa Luzia, we have the dominated group at 80 and 90 cm, and for São Vicente at 65 to 80 cm. The different sizes in the histogram show that the data are not normally distributed. São Vicente has a complete size range populated, and Santa Luzia has one missing class size > 120 cm (Figure 4)

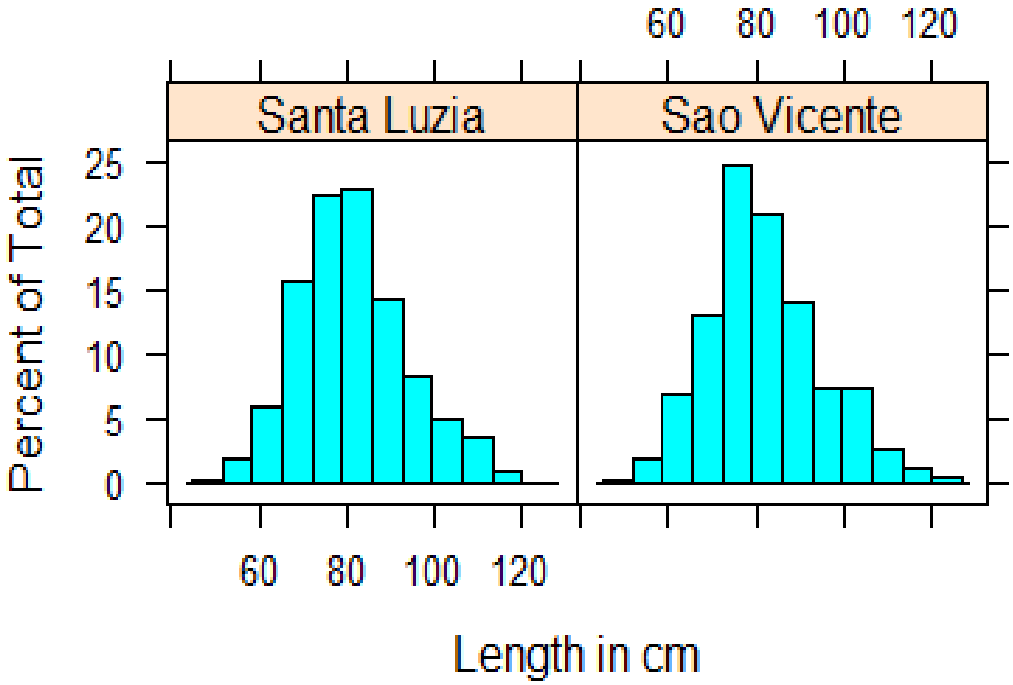

Figure 4: Comparison of length distribution between Santa Luzia and São Vicente in 2004.

The Cramer von Mises Test has shown that there is a significant difference between Santa Luzia and São Vicente in 2003, and Santa Luzia has a complete size range in the population, and the sizes range from 55 to 120 cm, and the maximum group size is at 80 cm. São Vicente has two population size classes missing, and the size ranges from 60 to 110 cm; the size class of about 80 cm is most abundant (Figure 5).

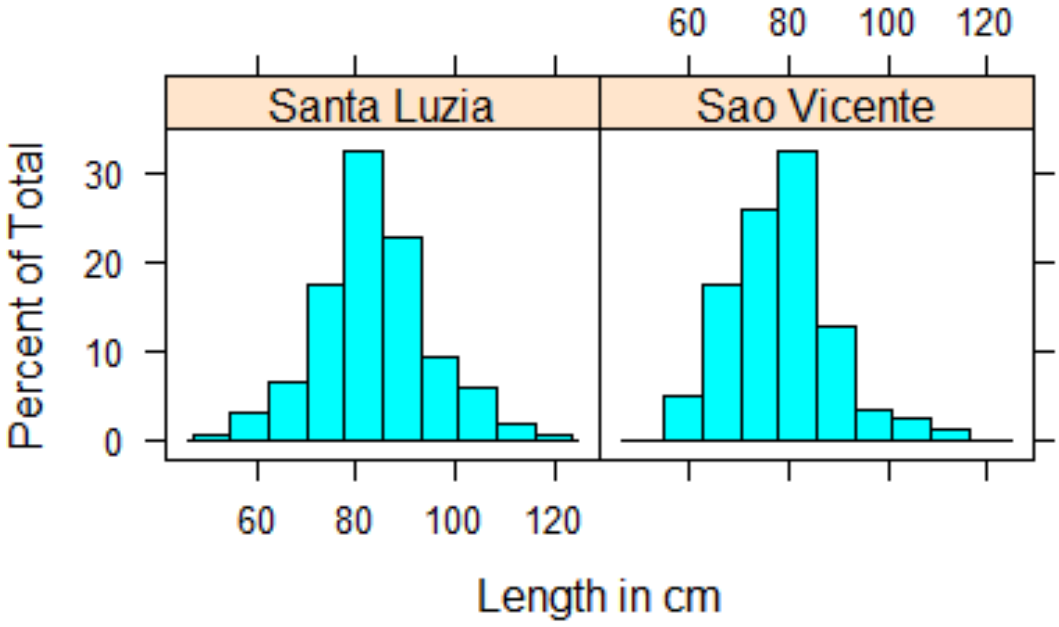

Figure 5: Comparison of length distribution between Santa Luzia and São Vicente in 2003.

The length-frequency distribution between São Vicente and Santa Luzia in 2015 looks quite different, and this is confirmed by testing the P-value. We see that data are not normally distributed; both have some missing data (Figure 6).

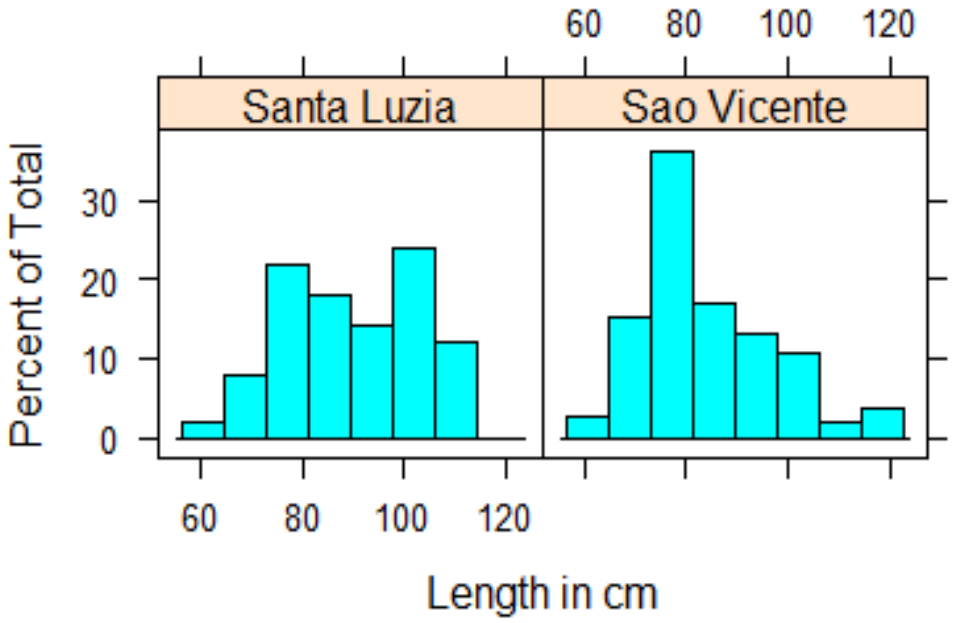

Figure 6: Comparison of Length Frequency Distribution between Santa Luzia and São Vicente in 2015

4.2.2 Comparison of length distributions for Santa Luzia and Santo Antão in 2004 and 2015

The Cramer Von-Mises Test between Santa Luzia and Santo Antão showed that the highest number of observations is always observed in Santa Luzia respectively in 2004 and 2015, and the lowest is observed in Santo Antão (Table 10).

Table 10: Cramer Von-Mises Test of Santa Luzia and Santo Antão

Years	Months selected	Number of observations Santa Luzia in selected months	Number of observations Santo Antão	P-value (cvm-Test)
2004	7; 12	208	144	0.0003
2015	6; 7	50	59	0.059

The length distribution for Santa Luzia and Santo Antão in 2004 looks very different, but the Cramer Von Mises Test shows a P-value of p = 0.0003, which means there is no significant difference. However, we can see that the data are not normally distributed, Santa Luzia has two missing populations, and the maximum group sizes are at 80 and 90 cm. Santo Antão has only a population missing, and its maximum size group is at 80 cm (Figure 7).

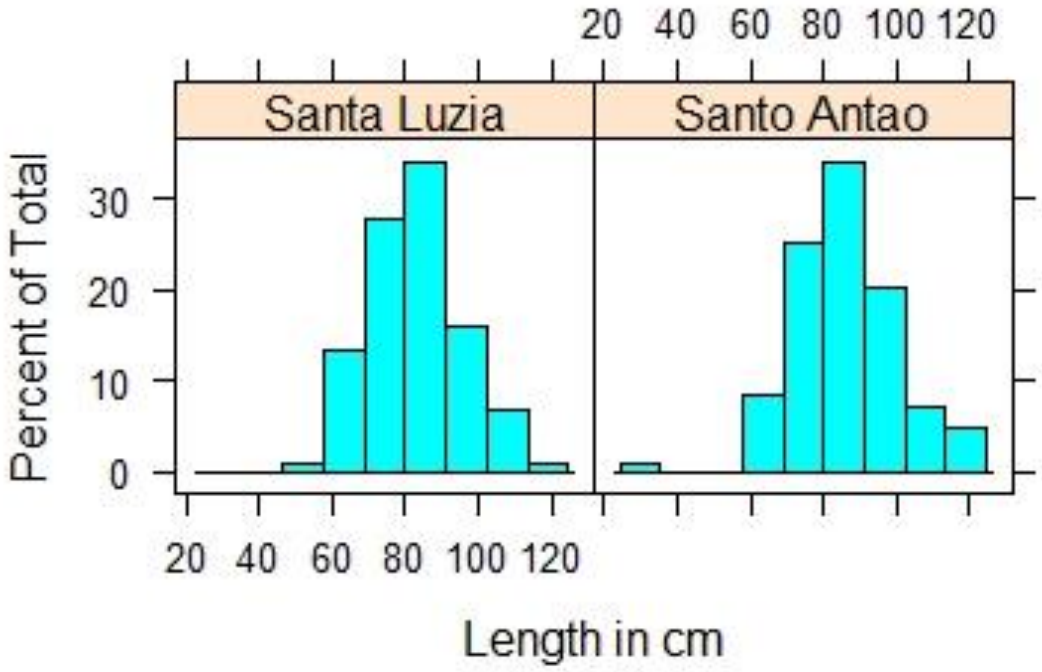

Figure 7: Comparison of Length Frequency Distribution between Santa Luzia and Santo Antão in 2004

They look quite different with regards to the smaller size classes, and the Test of Cramer Von-Mises proves that. The data are not normally distributed, and there is one more missing size class from Santo Antão than Santa Luzia (Figure 8). This indicates the presence of different age groups.

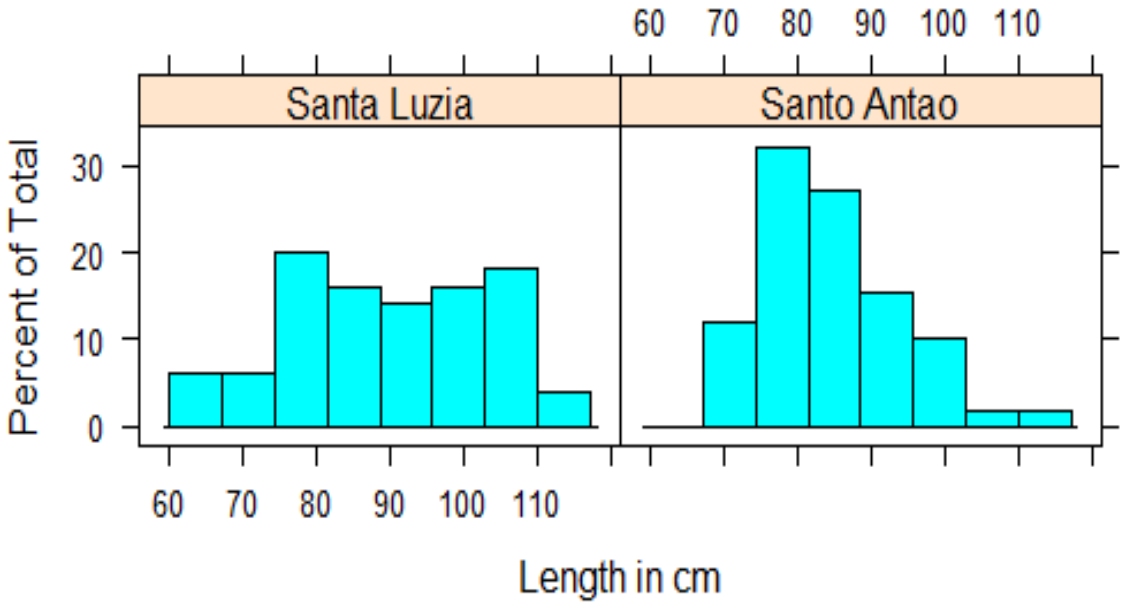

Figure 8: Comparison of Length Frequency Distribution between Santa Luzia and Santo Antão in 2015.

4.2.3 Comparison of length distributions for São Vicente and Santo Antão 2004 and 2019

The Cramer Von Mises Test showed the Test is very significantly different between São Vicente and Santo Antão in 2004 with a P-value of p = 0.002, but in 2019 there is no significant difference with a p = 0.262 (Table 11).

Table 11: Cramer Von-Mises Test of São Vicente and Santo Antão

Years	Months selected	Number of observations Santo Antão	Number of observations São Vicente	P-value (cvm-Test)
2004	7; 12	144	31	0.002
2019	11	21	17	0.262

The length-frequency distributions for São Vicente and Santo Antão in 2004 show that they are quite different, but the contrast is observed by doing the Test; they are not different. The data are unequally distributed. For Santo Antão, we have the dominated group at 80 cm

and 90 cm, one missing size class. The dominant group populated for São Vicente is at 70 cm and 80 cm, and we have also seen some missing values (Figure 9).

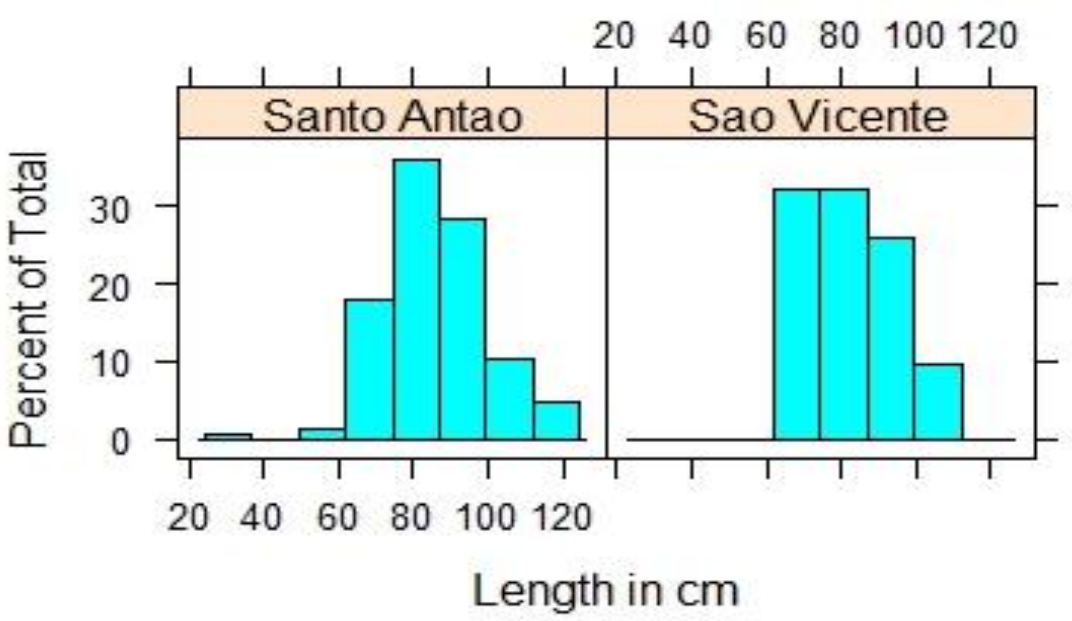

Figure 9: Comparison of Length Frequency Distribution between Santo Antão and São Vicente in 2004

The Test of Cramer von Mises confirmed that there is no significant difference in comparing the length-frequency distributions between São Vicente and Santo Antão in 2019 with a P-value of p = 0.262. Santo Antão has a maximum group size of 80 cm. São Vicente has the maximum population grouped at 70 cm and 90 cm. We are seen from both sides that the data are not normally distributed (Figure 10).

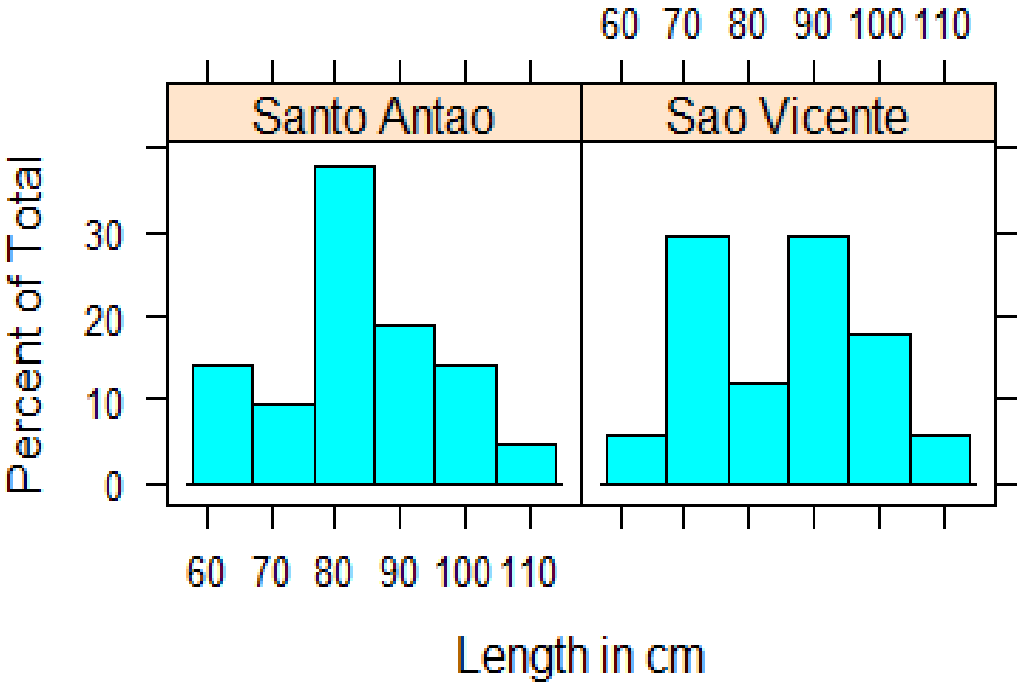

Figure 10: Comparison of Length Frequency Distribution between Santo Antão and São Vicente in 2019

4.2.4 Analysis of the trend of mean weight distribution

Analysis of the trend of mean weight distribution for São Vicente

The figure of mean weight distribution for São Vicente shows that the extreme highest distribution is observed in 2005, 2006, 2015, and 2020, and the extreme lowest distribution is seen in 2003, 2004, 2007, 2012, and 2017 (Figure 11).

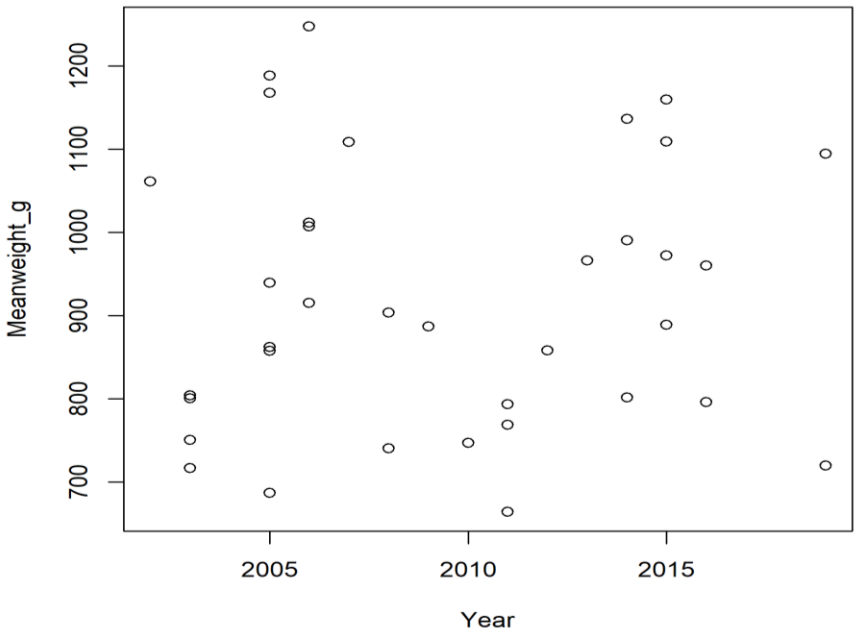

Figure 11: Mean weight distribution for São Vicente

Analysis of the trend of mean weight distribution for Santa Luzia

The mean weight distribution is observed in 2013, 2014, 2015, and 2016; the minimum weight distribution is observed in 2005, 2009, and 2020 (Figure 12).

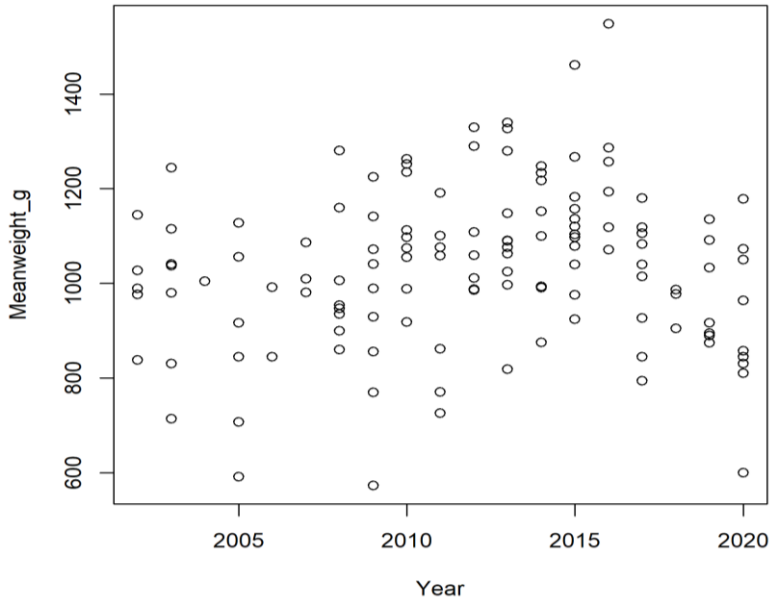

Figure 12: Mean weight distribution for Santa Luzia

Analysis of the trend of mean weight distribution for Santo Antão

The figure shows that the mean weight distribution is observed in 2005, 2006, 2007, 2013, 2017, and 2020, and the minimums are shown in 2002, 2009, 2010, and 2011 (Figure 13).

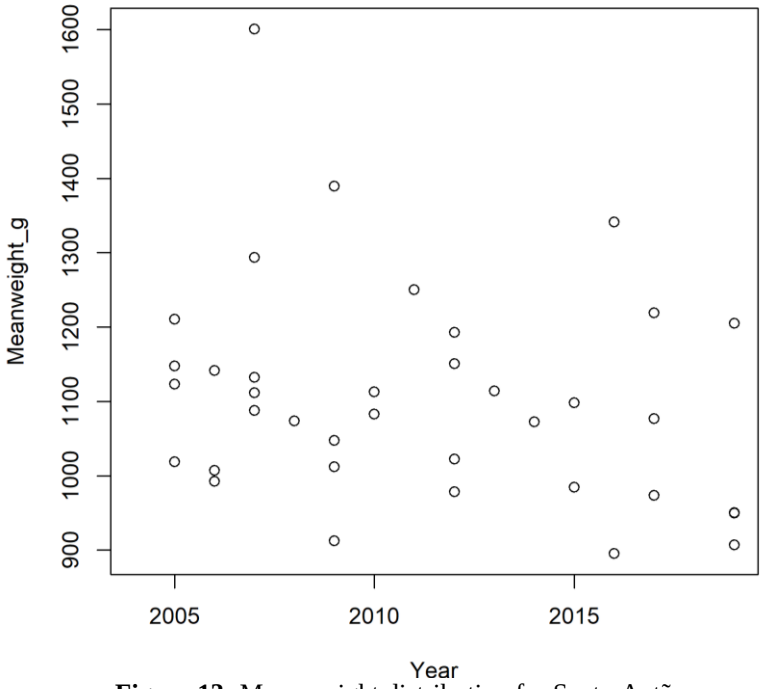

Figure 13: Mean weight distribution for Santo Antão

4.3 Analysis of Loess smoothing for mean average weight for the three islands (São Vicente, Santo Antão, and Santa Luzia)

The Figure shows that the predicted mean weight (red line) is slightly highest observed in 2016, 2015, and 2007; and slightly lowest in 2010, 2011, and 2019 (Figure 14). In the years with more data available, the predicted mean weight was higher, and also the years where the dataset had extremely low values, the predicted mean weight was also lowest (Figure 14).

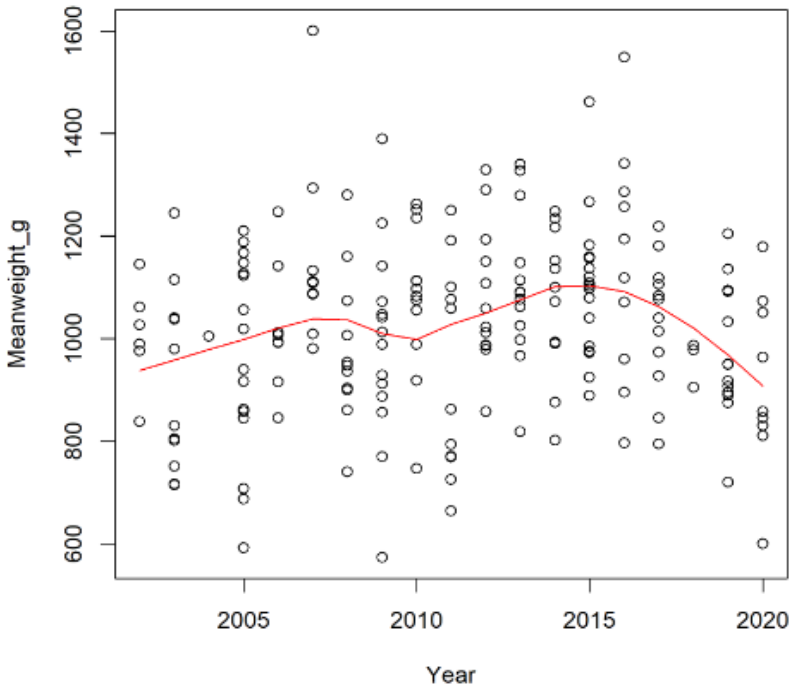

Figure 14: Analysis of Loess smoothing for mean average weight for the three islands (São Vicente, Santo Antão, and Santa Luzia)

4.3.1 Analysis of the trend of maximum size

Analysis of the trend of maximum size for São Vicente

The Figure 15 shows that the trend of maximum size is observed in 2004, 2005, 2008,2013, and 2015. The extreme lowest sizes are observed in 2002, 2007, 2012, 2017, and 2018 (Figure 15).

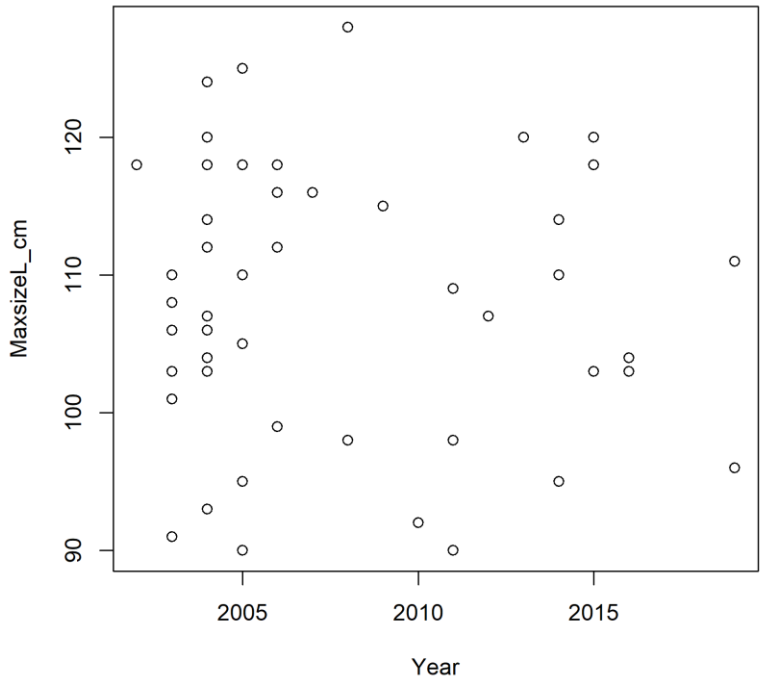

Figure 15: Analysis of the trend of maximum size for São Vicente

Analysis of the trend of maximum size for Santa Luzia

The extreme maximum size distributions are observed in 2013, 2014, 2015, and 2016; the extreme minimum size distributions are observed in 2005, 2009, and 2020 (Figure 16).

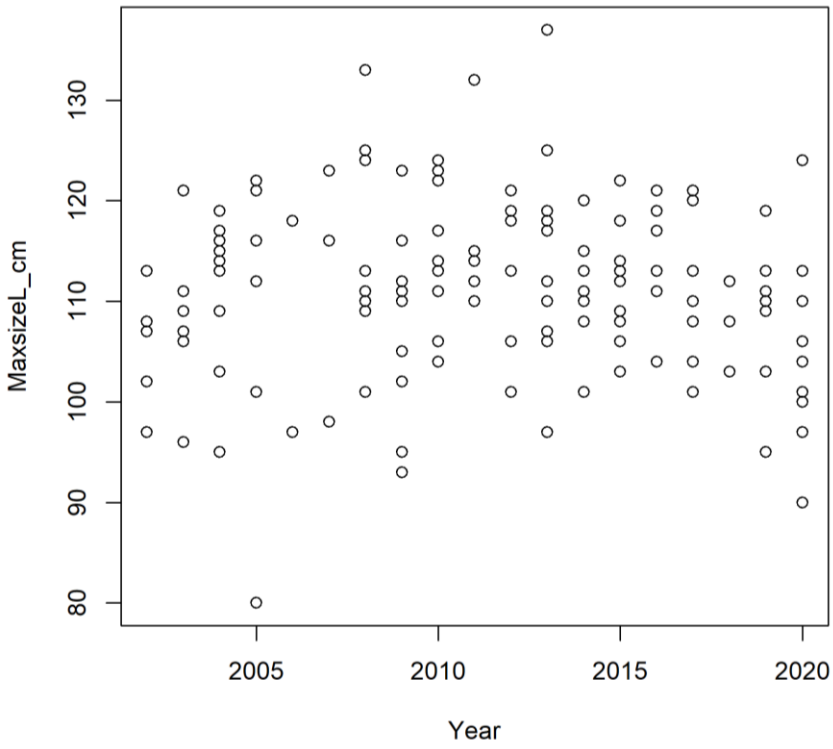

Figure 16: Analysis of the trend of maximum size for Santa Luzia

Analysis of the trend of maximum size for Santo Antão

The Figure shows that the extreme maximum size distributions are observed in 2005, 2006, 2007, 2013, 2017, and 2020, and the extreme minimums are shown in 2002, 2009, 2010, and 2011 (Figure 17).

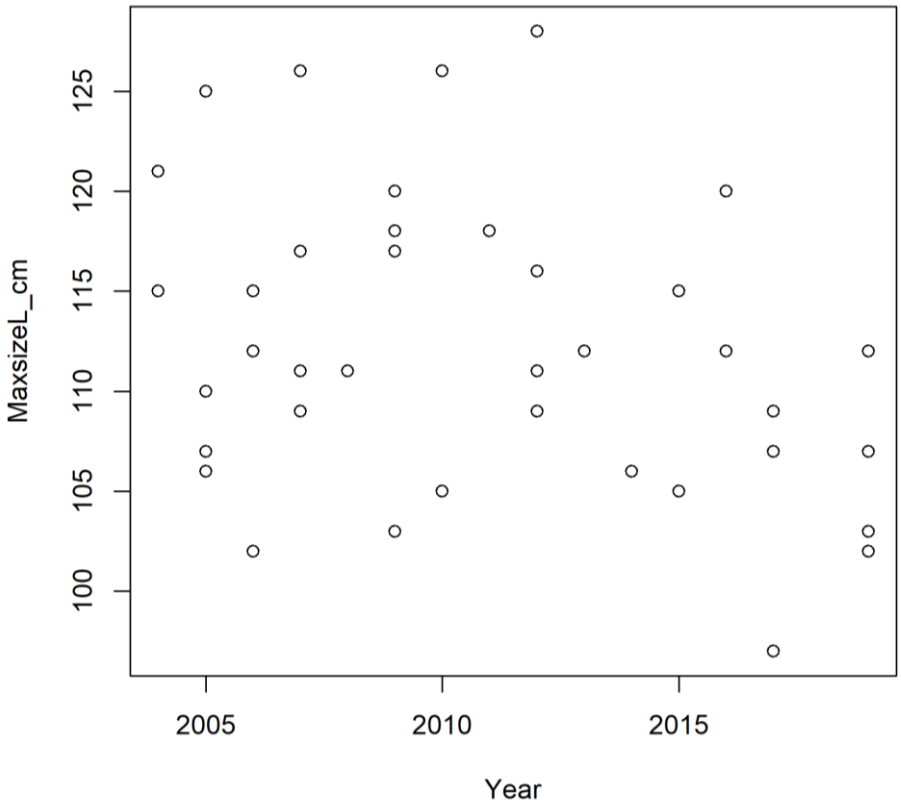

Figure 17: Analysis of the trend of maximum size for Santo Antão

4.3.2 Analysis of Loess smoothing for maximum size for the three islands (São Vicente, Santo Antão, and Santa Luzia)

The Figure shows that in the combined dataset of the three islands (São Vicente, Santa Luzia, and Santo Antão), the predicted maximum size (red line) is slightly higher and more or less constant from 2002 to 2008, and getting a slope from 2009 to 2012; and from 2013 to 2020 the maximum highest size become more or less constant, and also the minimum is shown from 2009 to 2013 and also decreasing from 2013 to 2020 (Figure 18).

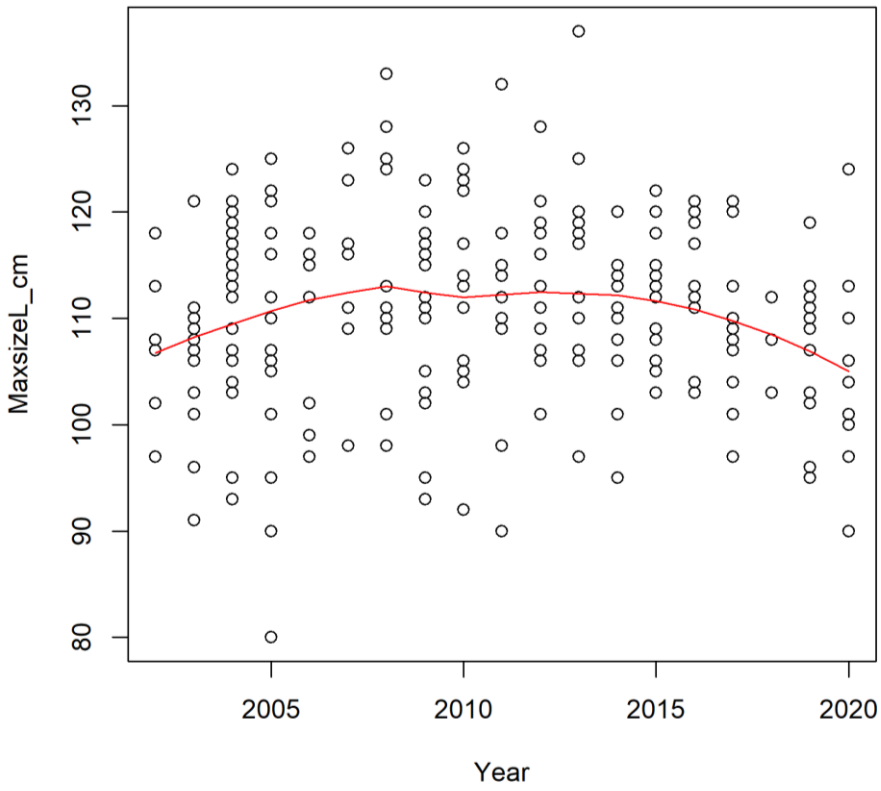

Figure 18: Analysis of Loess smoothing for maximum size for the three islands (São Vicente, Santo Antão, and Santa Luzia)

4.4 Analysis of Weight-Length relationships

For this analysis, we created a dataset for each island. After that, we created a linear model between the log-value of weight as y-variable and the log-value of length as x-variable for each corresponding dataset. For each x-value, a corresponding predicted value was calculated by using the antilog of the LWR model intercept (parameter a) times the corresponding length value to power the corresponding exponent (parameter b). Also, we calculated the residuals for each observation which is equal to the observed weight minus the calculated weight for the corresponding dataset, and then we plotted the residuals for each island.

The model shows that in Table 12, the length-weight relationship differs from one island to another island. So, for São Vicente the exponent is 2.97, Santa Luzia the exponent is 2.95, and the lowest exponent is observed in Santo Antão which is 2.88. Some of the residuals are large, which means that there is a big variability between weights at a given size. Santo Antão has the lowest degree of freedom which is 1300, and the larger number of degrees of freedom is seen in Santa Luzia with 5222; all three islands have the same P-value with $p < 0.001$, which means that the model is highly significant.

Table 12: Summary of the model of Weight-Length relationship for São Vicente, Santa Luzia, and Santo Antão

Parameters	São Vicente	Santo Antão	Santa Luzia
Exp (log a)	0.0017	0.0026	0.0019
Exponent / parameter b	2.97	2.88	2.95
Number of cases/degrees of freedom	1300	1197	5222
p-value of the model	< 0.001	< 0.001	< 0.001

4.4.1 Inspection of residuals from Weight-Length relationship for São Vicente, Santa Luzia, and Santo Antão

For this analysis, we created a dataset for each island. We created the dataset for each island, and we calculated the residuals for each island by using predicted values from the LWR model (see Table 12) for each corresponding island, such as the antilog (parameter a) and the exponent (parameter b), and then we plotted the residuals for each of the three islands. As shown in (Figures 19, 20, and 21), the residuals are centered around 0 and fairly normally distributed. However, it is assumed that larger residuals are to some degree due to weighing errors and transcription mistakes.

For São Vicente, knowing that the residual is the difference between observed and calculated weight, the model residuals of the length-weight-relationship (LWR) show that we have here the residuals that are more or less centered around 0. For example, we can see that at the observed weight of 2500 g, we have one observation where the residual is 1000, which means that the residual is about 40%. Altogether, the LWR appears to be a good model within most residuals centered around zero, but in some cases, a big difference between observed weight and predicted weight is evident (Figure 19).

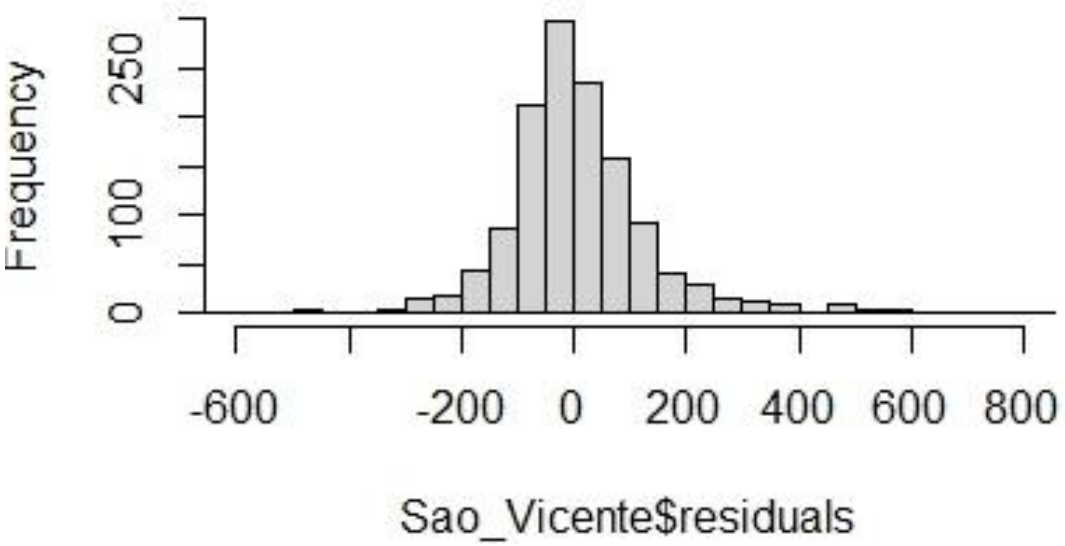

Figure 19: The residuals for the LWR model for São Vicente

For Santo Antão, the model residuals of the length-weight-relationship (LWR) show that the residuals are more or less centered around 0; and for example, we can see that at the observed weight of 300 g, we have one observation where the residual is 700, that means that altogether, the LWR can be accepted, however there some larger residuals present (Figure 20).

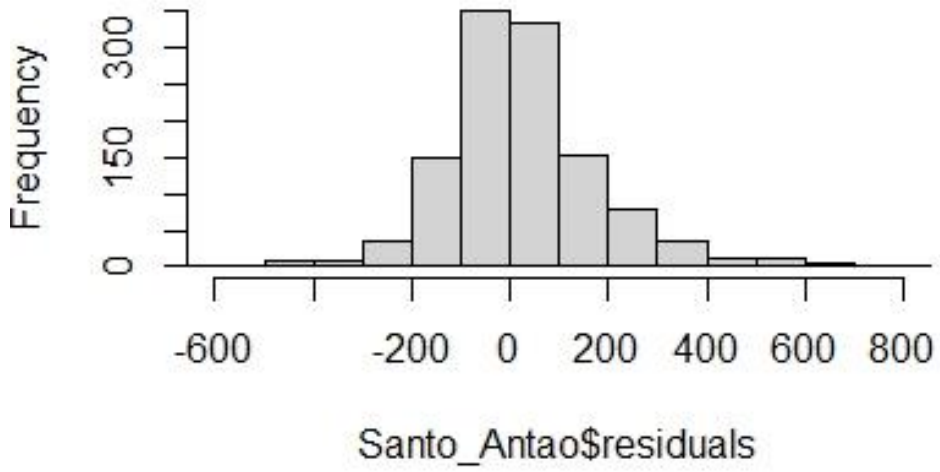

Figure 20*:* The residuals for the LWR model for Santo Antão

For Santa Luzia, the model residuals of the length-weight-relationship (LWR) show that the residuals are more or less centered around 0; and for example, we can see that at the observed weight of 1500 g, we have one observation where the residual is 800, that is mean

that the residual is about 53%. Then all together, the LWR model is a good model, however, with some larger residuals present (Figure 21).

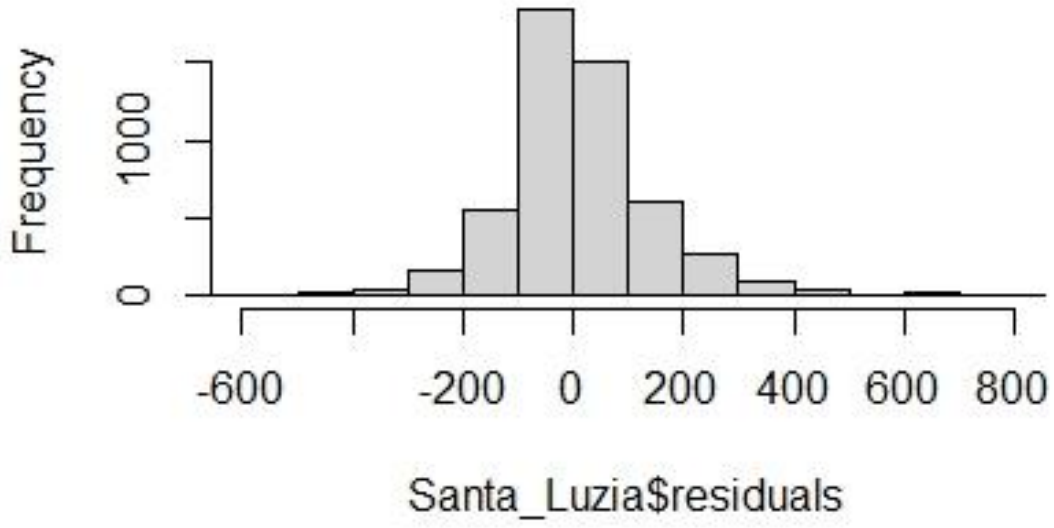

Figure 21: The residuals for the LWR model for Santa Luzia

4.4.2 ANOVA for São Vicente, Santo Antão and Santa Luzia

To analyse the difference between the three islands, the residuals of the combined model were analysed for São Vicente, Santo Antão, and Santa Luzia with an ANOVA.

The ANOVA is a statistical test that aims to assess how a quantitative dependent variable changes in relation to the degrees of one or more categorical independent variables. This test makes it possible to see if there is a difference between the classes at each level of the independent variable. We calculated the weights; accordingly, we use the antilog 'exp (log a)' for the coefficient 'a' from the equation W= a Lb (Eq. 1).

We calculated the residuals and plotted it to see if it is normally distributed with mean zero and checked with a box-wing plot of the three islands. We then checked the homogeneity of variances so that we applied an ANOVA. Homogeneity of variance is an important assumption in the analysis of variance ANOVA (Zuur et al., 2010). In regression-type models, verification of homogeneity of variances should be done by using the residuals of the model, i.e., by plotting residuals vs. fitted values and making a similar set of conditional boxplots for the residuals. This is shown in Figure 22.

It is possible to calculate the absolute coefficient of the residuals and use it as a response in a new one-way ANOVA with island as a factor. A significant variation would indicate an island effect. There may be other tests, but this one is fairly insensitive to deviations from non-normality and is simple to run (Asai, 1991). After removing residuals larger than absolute 500 (winsorizing), the ANOVA model was significant ($p<0.001$), indicating an island effect in the residuals and thus in LWR between islands. This figure of the residuals looks like the variability is more or less the same for the islands (Fig. 22). The box in the Box-whisker plot captures the data at the 25% and 75% quartile, between 25% and 75%, we have 50% of our values which is quite informative, and since the box looks the same for the three of the islands, then 50% of the values are in the same range. That is why we can use the box-whisker plot to investigate the homogeneity of the variance (Figure 22).

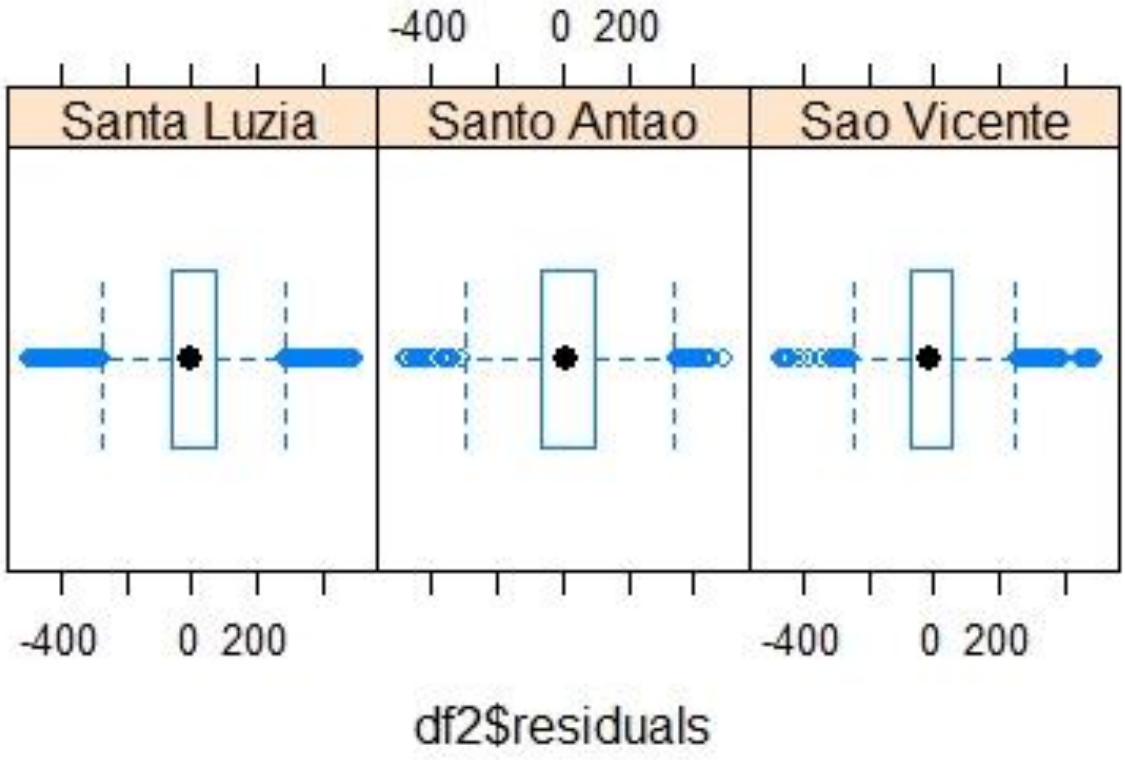

Figure 22: Box-whisker plot of residuals from the combined LWR analysis for São Vicente, Santo Antão, and Santa Luzia. The black dot represents the median for each series, the box describes the space between the 25%- and 75% quartile, and the whiskers describe the range beyond quartiles with a 1.5 interquartile range, the circles describe the remaining outliers truncated to plus or minus 500.

The entire model of the length-weight relationship of the combined dataset of the three islands shows in Table 12 that there is a similarity between the model for Santa Luzia parameters a and b from Table 12, given that most data were sampled from Santa Luzia.

Therefore, the P-value is very low with $p < 0.001$, which indicates that if the null hypothesis were indeed true, by means that there would be only one chance in 1000 of falsely rejecting the LWR model.

Table 13: Summary of the combined model of length-weight relationships for São Vicente, Santa Luzia, and Santo Antão

Parameters	Value
Exp(log a) / parameter a	0.0019
Exponent / parameter b	2.95
Degrees of freedom	7723
p-value of the model	< 0.001

4.4.3 Analysis of the trend of the condition factor (Fulton K)

We rearranged our data frame; after that, we calculated Fulton K by fixing the exponent at 3. For the trend of the condition factor (Fulton K), we create a new dataset called data frame 2, from which we bind the dataset for the three islands by using the function "rbind." We looked for the outliers for this dataset, and after running the model, we removed some outliers. We created a new model called model 2, from which we used the linear model function for the log that we linked length and weight from this dataset. We calculated Fulton K by fixing the exponent in the LWR model at 3.

We summarized model 2 to find out the coefficient a and the exponent. From this summary, we calculate the length-weight relationships by applying the equation: $W=aLb$, where w = weight of fish in grams, a = coefficient a, L = the total length of the fish in centimeters, b = coefficient b which as the exponent which was set to 3.

We also calculated the residuals from this dataset (df2) which is done by subtracting the observed weight minus the calculated weight. We removed those observations where the residuals were above 3000 g and also those below 500 g after running the model. We summarized this model and calculated the residuals from data frame 2, we computed the analysis of the variance, and we summarized the analysis of variance from which we concluded that there is a significant difference between islands.

The Figure 23 shows that the predicted values for condition factor K (red line) is slightly higher and seems more or less constant sinusoidally, the highest predicted values (red line)

are seen from 2005 to 2008, having a slope from 2009 to 2012, and became higher from 2013 to 2015, the lowest predicted values are observed from 2016 to 2020 (Figure 23).

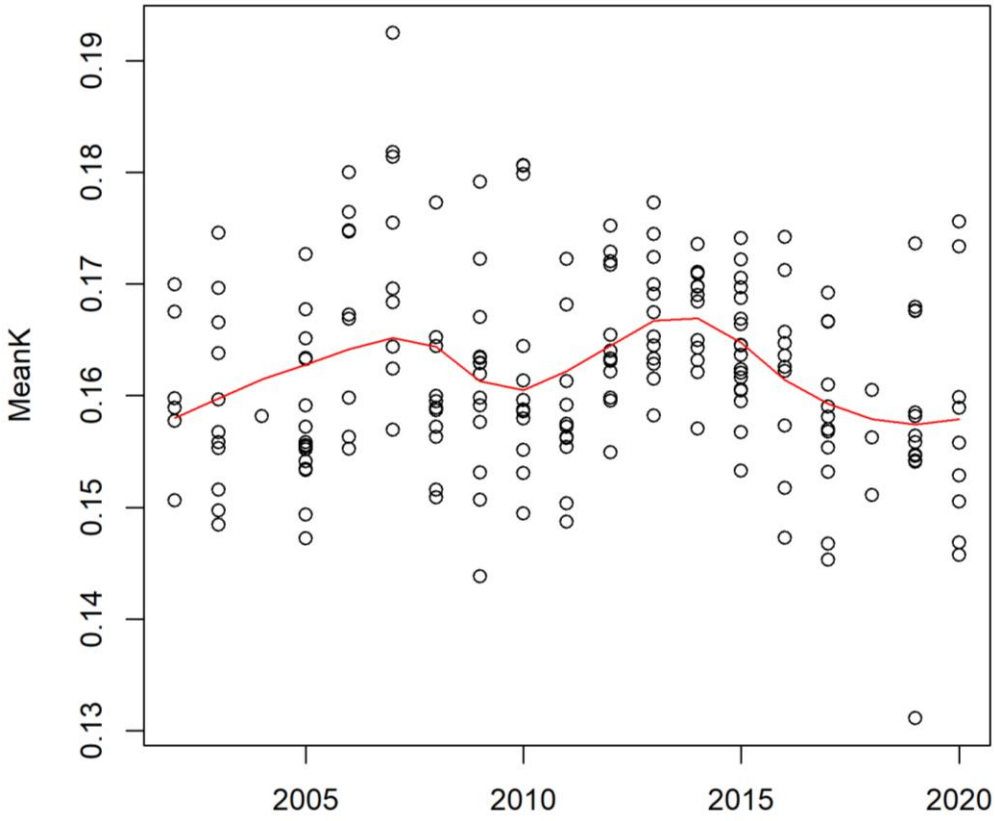

Figure 23: Analysis of the trend of the condition factor (Fulton K)

4.5 Genetic Results

Different from the analyses for morphometric data, species from outside the Cabo Verde archipelago are also included in the analysis of genetic similarity.

4.5.1 Haplotypes and nucleotides diversity

The number of individuals sampled totalled 28 specimens ranging from a minimum of 4 for Santo Antão to a maximum of 16 for Santa Luzia. A total of 26 were sequenced, where an alignment was obtained with 617 bp of the COI gene. A total of 18 haplotypes were obtained by joining the sequences from Cabo Verde with the sequences obtained from the online databases. The values of haplotype numbers, haplotype diversity, and nucleotide diversity are shown in Table 14.

Table 14: Parameters of genetic diversity. Number of sequenced individuals (N) per island, number of haplotypes (H), haplotype diversity (H.d), and nucleotide diversity (p) calculated in DNASP.

Countries	Locations	Samples	N	H	Hd	π
Cabo Verde	São Vicente	7	7	4	0.714	0.181
Cabo Verde	Santa Luzia	15	15	7	0.657	0.138
Cabo Verde	Santo Antão	4	4	2	0.5	0.265
Brasil	Bahia	2	2	2	1	0.5
Mexico	Othon P. Blanco	2	-	-	-	-
Mexico	Gulf of Mexico	2	2	2	1	0.5
USA	-	1	-	-		0.5
BRASIL	Bahia - Ilha de Tinhare	3	3	3	1	0.272

4.5.2 Sequences analysis and alignment

We realised that the haplotype diversity was relatively high, varying between 0.5 and 1 (mean = 0.75), and it seems more or less equal for the three islands (Santa Luzia, São Vicente, and Santo Antão), and also this haplotype diversity is unique for the Cabo Verde. The π is not higher for the three islands with respectively 0.138 for Santa Luzia, 0.181 for São Vicente, and 0.265 for Santo Antão. These results reflect no genetic structure difference between islands (Table 14).

No premature stop codons were seen in the COI alignment, and no gaps were considered. The pool of COI and 26 sequences analysed had a total length of 617 bp and 18 haplotypes unique to Cabo Verde islands, plus haplotypes from different geographic regions. Here is

presented the mitochondrial haplotype networks of *G. vicinus* for Cabo Verde and a Broad geographic region, and basically, we are seen that the haplotypes from the major geographic region are apart dispersed through the networks (Figure 24). From this Figure, the bubble size is describing the number of specimens that are very closely related, and we are seen that this closest relationship was observed between the *G. vicinus* specimens from the three Cabo Verde islands, as indicated by the colours of the central pie parts of Figure 24.

The mitochondrial haplotype networks of *G. vicinus* for Cabo Verde and a Broad geographic region show few mutational steps between the different zones. The mtDNA haplotype network (Figure 24) showed a star shape with the highest ancestral central-most common haplotype connected to many haplotypes. This similarity of haplotypes in Cabo Verde means there are no significant differences in haplotypes between *G. vicinus* in these three islands.

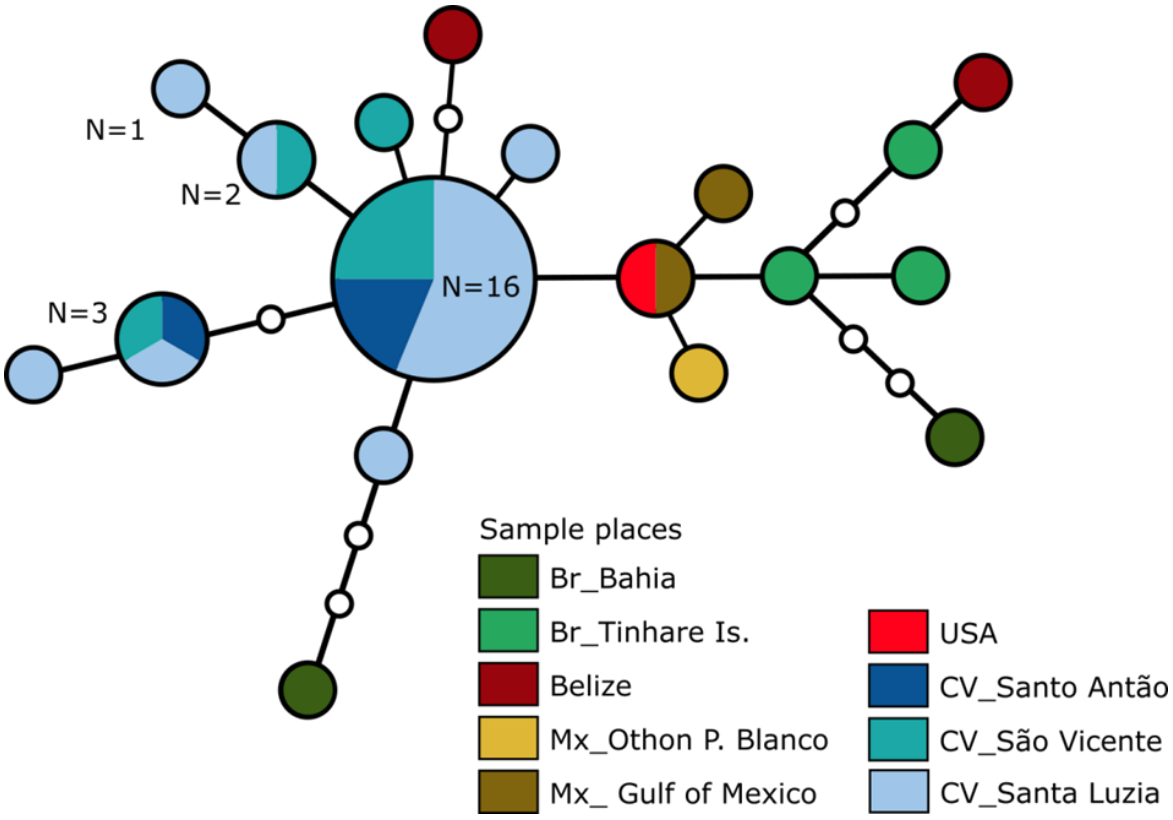

Figure 24: Statistical parsimony network of haplotypes (COI) of *G. vicinus* (95% connection limit); Colours represent the sampling side, and the line represents the mutation steps. The areas of the circle are proportional to the number of individuals sharing the respective haplotypes. The white dots represent the missing haplotypes, N represents the number frequency

From this haplotype network, it is clear that all haplotypes found in Cabo Verde are not found anywhere else. This is somehow very interesting because it has been saying that there

is no reproductive connectivity. However, with only 26 individuals, it is probably a bit speculative still, but on the other hand, looking at a gene that is quite conserved and normally only used to show differences at a species rather than at a population level, the results are quite conclusive.

5 Discussion

5.1 Data availability

Data coverage between islands and years and months was not the same for each of the islands in the Cabo Verde archipelago. Of the 10 islands, only data from 3 islands were available in larger numbers, and for pairwise comparisons between islands, months and years with overlapping data availability had to be selected. The effect of data availability is most evident in the analysis of length-weight relationships, when the combined model for Santa Luzia, Santo Antão, and São Vicente is almost equal to the model for Santa Luzia alone, given that Santa Luzia contributed more than 60 % percent of the data analysed (Tables 12 and 13).

The islands Santo Antão, Santa Luzia, and São Vicente are topographically close to each other, and it cannot be inferred from the data available whether other islands along the eastern or southern margin of the archipelago like Sal or Fogo would have the same morphometric properties in their respective *G. vicinus* populations. Even in the cases with overlapping data for pairwise comparisons, the number of observations is in some cases too small to infer differences between the local subpopulations. This will be explained concerning the analyses of length-frequency distributions.

5.1.1 Length-Frequency-Distributions (LFD)

Fock and Czudaj (2019) discussed that these types of analyses are prone to making type II statistical errors, i.e., making false negative statements or falsely accepting the null hypothesis of no difference. In this thesis, the type I statistical error to make a false positive statement or falsely accept the alternative hypothesis is minimized by the p-values of the statistical tests.

For length-frequency distributions, Fock & Czudaj (2019) applied the Cramer-van Mises test because it is robust. Based on a resampling procedure, the differences observed are dependent on the samples involved and thus on the respective sample sizes. In this thesis, significant and non-significant results were observed between islands and combinations of months in certain years. In particular, the LFD analysis for Santa Luzia and São Vicente (Table 9) shows that in the years 2003 and 2015 with joint sample sizes of 378 and 228, resp., each for 4 months, significant results were obtained, indicating that the average sample size per month was 38 for either island, which can be regarded a small sample size because of the size range of *G. vicinus* ranging from about 60 cm to larger than 120 cm. In turn, in Table 9

for 2004, with a total number of observations of 1716, no difference was observed. Thus, it can be concluded that with increasing sample size, the precision of the analysis to reveal the exact population LFD is better.

5.1.2 Trends in maximum size, mean weight and Fulton K, and length-weight relationships

The analysis of the maximum size and mean weight is in particular dependent on the sample sizes involved, the same as for LFDs. Fock and Czudaj (2019) used an inclusion index to indicate the degree of possible type II errors. Instead, the analysis of the length-weight relationships is rather dependent on covering the size range than on the size composition of the population, as evidenced by means of the LFDs. Thus, Froese (2006) proposes to include equal numbers of small, medium, and large specimens in order to obtain a robust LWR. The histograms (for instance, Fig. 4 for the comparison of the islands Santa Luzia and São Vicente) show that all size classes were included in the LWR analysis. This means that the LWR can be used to indicate differences at the population level between islands. The ANOVA showed that there was a significant difference in the residuals indicating that the group of specimens from Santo Antão was different from the specimens sampled at Santa Luzia and São Vicente.

Despite the possible bias in the analysis of the maximum size and mean weight, the plots for maximum size, mean weight, and the condition factor Fulton K all show a negative trend from 2013-2014. However, the trend is only slightly negative for maximum size (Fig. 18). The negative trend thus could be indicative of environmental changes in the Carbo Verde archipelago with a negative effect on G. vicinus. It is not clear if this trend is going on because the analysis of Fulton K in 2020 showed a very small upward trend again (Fig. 23). The environmental factors, for instance, could affect the food availability and growth conditions of the moray eels.

5.1.3 Genetic diversity

It is generally believed that benthic organisms with a long pelagic life stage will have a wide geographic distribution and poor genetic structure compared to those without free drift of the pelagic stage (Cunha et al., 2011). Here, we analysed the genetic structure of *G. vicinus* from three of the ten islands of the Cabo Verde archipelago, but we did not find evidence of any genetic structure differentiations. Nevertheless, being in the light of the absence of

genetic structure in all the islands analysed, there is no evidence of the existence of ecological transitions in the Cabo Verde habitat occupied by *G. vicinus*, indicating that *G. vicinus* would occupy different niches.

However, in another way, we can say that there was connectivity in the past, maybe at different time scales, explaining the single haplotypes in Brazil and Belize derived from the Caboverdean haplotype group. So, this could mean that at certain times in the past, the eastern Atlantic Lycodontis larvae crossed the Atlantic in a westward direction, using the North Equatorial Current and established new haplotype groups. Thus, the Cabo Verde archipelago, with its strong degree of sparid endemism, attributed to multiple broadcasts (Santini et al. 2014), also shows almost no signs of differentiation for *G. vicinus*. In turn, a close relationship to the Brazil specimens is in support of the pan-Atlantic marine connectivity, largely concerning to a distinct American (mainly Caribbean) footprint, as has been suggested for pelagic fishes (Lopes et al., 2021). Only low levels of endemism exist in Macaronesia as a whole, which is probably related to the distance between these archipelagos and mainland Africa and Europe (Almada et al., 2013).

6 Conclusion and recommendations

6.1 Conclusion

The results show that there is no significant overall genetic differentiation between specimens from the different islands of Cabo Verde, but rather that there are some significant distinctions also within the islands analysed. However, the difference between specimens from Cabo Verde and those from the eastern American coast is greater than the diversity within the Cabo Verde archipelago, and this difference is significant. Thus, the combination of current genetic data and morphological observations suggests that this taxonomic concept exclusively includes *G. vicinus* and a putative cryptic species or subspecies. On the other hand, a morphometric difference in terms of the length-weight relationships between the specimens from the islands of Santo Antão and those of São Vicente and Santa Luzia was indicated. Thus, given the genetic similarity, the morphological difference existing between the specimens of the different islands make is considered to be due to local environmental factors. On the one hand, it is about a new observation with its distribution in the islands of Cabo Verde; on the other hand, it is a question of qualifying this species on the based on morphological criteria and DNA barcoding and of correcting the current COI sequences of this species published in GenBank. The conclusion that local environmental factors play an important role in the morphometric parameters is supported by the fact that for average weight, maximum size, and condition factor Fulton K, similar temporal trends were indicated, presumably influenced by factors such as food availability and temperature.

6.2 Recommendations

The current threat of these activities is a challenge to fisheries conservation and sustainable exploitation of marine and coastal ecosystems in Cabo Verde. This study reinforces the importance of combining the morphological and genetic tools to solve in a general way the taxonomy and ecology of marine species. However, it is clear that in the case of *G. vicinus*, more sampling effort is required to have a better basis for comparisons between islands. For future work on this specific species, which is *G. vicinus*, more ecological also be need to see if there is an ecological impact on this species in all these islands. We will show in the future the need for more morphological and genetic data and to be careful when identifying moray eels and will promote fisheries management, biodiversity conservation, and sustainable management of this species.

7 References

Almeida, F. S., Same, W. F., & Diego, C. L. (2017). DNA barcoding is a useful tool for identifying non-native species of freshwater ichthyoplankton in the neotropics. Hydrobiologia. https://doi.org/10.1007/s10750-017-3443-5.

Anderson, T. W. (1962). On the Distribution of the Two-Sample Cramer-von Mises Criterion. *The Annals of Mathematical Statistics*, *33*(3), 1148–1159. https://doi.org/10.1214/aoms/1177704477.

Asai, H. (1991). Discipline Pascal with a descriptive environment; precise writing to learn to program and to avoid errors. *Computers and Education*, *16*(4), 323–335. https://doi.org/10.1016/0360-1315(91)90006-D.

Becker, R. A., Sales, N. G., Santos, G. M., Santos, G. B., & Carvalho, D. C. (2015). DNA barcoding and morphological identification of neotropical ichthyoplankton from the Upper Paraná and São Francisco. *Journal of Fish Biology*, *87*(1), 159–168. https://doi.org/10.1111/jfb.12707.

Benchimol, C., Francour, P., Lesourd, M., & Rouen-ledra, W. W. F. U. De. (2009). The preservation of marine biodiversity in West Africa, the Case of Cabo Verde Islands: Proposal of a new biodiversity policy management. *1st Cabo Verde Congress of Regional Development, Praia, Santiago Island, Cabo Verde.*, 1–22.

Brito, A., Moreno-Borges, S., Escánez, A., Falcón, J.M. & Herrera, R. (2017) New records of Actinopterygian fishes from the Canary Islands: tropicalization as the most important driving force increasing fish diversity. Revista de la Academia Canaria de Ciencias, 29, 31–44.

Brito, A., Dorta, P., & Falcon, J. M. (2014). First valid record of Gymnothorax vicinus (Pisces: MuraFenidae) for Macaronesian Ecoregion (Canary Islands). *Rev. Acad. Canar. Cienc.*, *26*(1), 71–78.

Brito, A., J.M. Falcón & R. Herrera. 2005. Sobre la tropicalización reciente de la ictiofauna litoral de las islas Canarias y su relación con cambios ambientales y actividades antrópicas. Vieraea, 33: 515-525.

Böhlke, E.B.; Mc Cosker, J.E.; Böhlke, J.E. Family Muraenidae. In Fishes of the Western North Atlantic; Böhlke, E.B., Ed.; Sears Foundation Marine Research: New Haven, CT, USA, 1989; pp. 104–206.

Clement et al., M. (2000). A pattern-matching algorithm for double-type characters. *Applied Mechanics and Materials*, *571–572*, 461–464. https://doi.org/10.4028/www.scientific.net/AMM.571-572.461.

Coluccia, E., Deidda, F., Lobina, C., Melis, R., Porcu, C., Agus, B., & Salvadori, S. (2020). Chromosome Mapping of 5S Ribosomal Genes in Indo-Pacific and Atlantic Muraenidae: Comparative Analysis by Dual Colour Fluorescence In Situ Hybridisation. *Genes*, *11*, 12. https://doi.org/doi:10.3390/genes11111319.

Cunha, R. L., Lopes, E. P., Reis, D. M., & Castilho, R. (2011). Genetic Structure of Brachidontes Puniceus Populations in Cabo Verde Archipelago shows Signature of Expansion during the Last Glacial Maximum. *Journal of Molluscan Studies*, *February*, 175–181. https://doi.org/10.1093/mollus/eyr001.

De Souza, C. S., & Mafalda Junior, P. O. (2019). Large-scale spatial and temporal variability of larval fish assemblages in the tropical atlantic ocean. *Anais Da Academia Brasileira de Ciencias*, *91*(1), 1–18. https://doi.org/10.1590/0001-3765201820170567.

Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. *PLoS Biology*, *4*(5), 699–710. https://doi.org/10.1371/journal.pbio.0040088.

FAO (1981). *Guide Eastern Centr Atl Fam Muraenesocidae Muraenidae.pdf* (p. 38).

Faria, A., Morais, P., & Chı, M. A. (2006). Ichthyoplankton dynamics in the Guadiana estuary and adjacent coastal area, South-East Portugal. *Estuarine, Coastal and Shelf Science*, *70*, 85–97. https://doi.org/10.1016/j.ecss.2006.05.032.

Falcón, J. M., S. A. Bortone, A. Brito & C. M. Bundrick. 1996. Structure of and

relationships within and between the littoral, rock-substrate fish communities off four islands in the Canarian Archipelago. Marine Biology, 125 (2): 215-231.

Fock, H. O., Czudaj, S., & Bartolino, V. (2019). Size structure changes of mesopelagic fishes and community biomass size spectra along a transect from the equator to the Bay of Biscay collected in 1966-1979 and 2014-2015. *ICES Journal of Marine Science*, *76*(3), 755–770. https://doi.org/10.1093/icesjms/fsy068.

Fox, J. . S. W. (2018). Nonparametric Regression. *Computational Statistics Handbook with MATLAB*, *2008*, 403–442. https://doi.org/10.1201/9781420035636-11.

Freitas, R. (2014). The coastal ichthyofauna of the Cabo Verde Islands: a summary and remarks on endemism. *Zoologia Caboverdiana*, *5*(1), 1–13.

Fricke, R., Eschmeyer, W.N.; Fong, J.D. Eschmeyer's Catalog of Fishes: Species by Family/Subfamily. Available online: http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.aspElectronicversion (accessed on 20 August 2020).

Froese, R. (2006). Cube law, condition factor, and weight-length relationships: History, meta-analysis, and recommendations. *Journal of Applied Ichthyology*, *22*(4), 241–253. https://doi.org/10.1111/j.1439-0426.2006.00805.x.

Gilbert, M., Rasmussen, J. B., & Kramer, D. L. (2005). Estimating the density and biomass of moray eels (Muraenidae) using a modified visual census method for hole-dwelling reef fauna. *Environmental Biology of Fishes*, *73*(4), 415–426. https://doi.org/10.1007/s10641-005-2228-2.

González, J. A., Correia, S., Jiménez, S., Monteiro, C. A., Delgado, J., Pinho, M. R., Lorenzo, J. M., & González-lorenzo, G. (2021). *The fish family Muraenidae : an ideal group for testing at small-scale the coherency of Macaronesia as a biogeographic unit , with the first report on separate fishery statistics*. *85*(September), 157–167.

Government of Cabo Verde; Ministry of Environment and Rural Development. (2010). Second National Communication on Climate Change of Cabo Verde. United Nations Framework Convention for Climate Change Ministry, October, 1–162. https://www.adaptation-undp.org/resources/assessments-and-background-documents/cape-verde-second-national-communication.

Higgins, B. A., & Mehta, R. S. (2017). Distribution and habitat associations of the California moray (Gymnothorax mordax) within Two Harbors, Santa Catalina Island, California. *Environmental Biology of Fishes*, *101*(1), 95–108. https://doi.org/10.1007/s10641-017-0684-0.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & de Waard, J. R. 2003 Biological identifications through DNA barcodes. Proc. R. Soc. B 270, 313–322. doi:10.1098/rspb.2002.2218.

https://www.fishbase.se/summary/7548.

Jaonalison, H. (2019). Les premiers stades de vie des poissons dans le SO de Madagascar: éléments pour une meilleure connaissance de la biodiversité et une meilleure gestion des ressources exploitées. *HAL*, 137.

Jespersen, P. (1942). Indo-Pacific leptocephalids of the genus Anguilla: systematic and biological studies. Dana-Report, 22.

Kent E. Carpenter. (2002). THE LIVING MARINE RESOURCES OF THE WESTERN CENTRAL ATLANTIC. *Asih Fao Ec*, *5*, 601–1374.

Laguna, L. (2016). *FIRST VALID RECORD OF Gymnothorax vicinus (PISCES : MURAENIDAE) FOR MACARONESIAN ECOREGION (CANARY ISLANDS): A PROCESS OF (PISCES : MURAENIDAE) FOR MACARONESIAN ECOREGION*. *December 2014*.

Li, Y., Zhang, L., Zhao, L., Feng, J., Loh, K., Zheng, X., & Lin, L. (2018). New identification of the moray eel *gymnothorax minor* (Temminck & schlegel, 1846) in China (Anguilliformes, Muraenidae). *ZooKeys*, *752*, 149–161. https://doi.org/10.3897/zookeys.752.24231.

LIoris D., Figueroa, J. R. and H. (1991). *Biogeography of the Macaronesian Ichthyofauna (The Azores, Madeira, The Canary Islands, Cabo Verde and The African Enclave* (p. 51).

Loh, K., Chen, I., Randall, J. E., & Chen, H. (2008). *A Review and molecular Phylogeny of the Moray eel Subfamily Uropterygiinae (Anguilliformes : Muraenidae) from Taiwan, with Description of a New Species*. *19*, 135–150.

Lopes, E. P., Monteiro, V., Martins, A., & Coelho, R. (2021a). Confirmed record of the roughear scad Decapterus table in the Cabo Verde Archipelago based on morphological and genetic data. *Sociedade Caboverdiana de Zoologia*, *9,1*, 3–13.

Lopes, E. P., Monteiro, V., Martins, A., & Coelho, R. (2021b). *Confirmed record of the roughear scad Decapterus table in the Cabo Verde Archipelago based on morphological and genetic data*. 3–13.

Múrias Dos Santos, A., Cabezas, M. P., Tavares, A. I., Xavier, R., & Branco, M. (2015). TcsBU: A tool to extend TCS network layout and visualization. *Bioinformatics*, *32*(4), 627–628. https://doi.org/10.1093/bioinformatics/btv636.

Pastor, O. T. (2002). *Life History and Stock Assessment of the African Hind (Cephalopholis Taeniops) (Valenciennes, 1828) in São Vicente- São Nicolau Insular Shelf of the Cabo Verde Archipelago*. 1–45.

Pek, J., Wong, O., & Wong, A. C. M. (2018). How to address non-normality: A taxonomy of approaches, reviewed and illustrated. *Frontiers in Psychology*, *9*(NOV), 1–17. https://doi.org/10.3389/fpsyg.2018.02104.

Peninal, S., Subramanian, J., Elavarasi, A., & Kalaiselvam, M. (2017). Genetic identification of marine eels through DNA barcoding from Parangipettai coastal waters Samuel. *Genomics Data*, *11*, 81–84. https://doi.org/10.1016/j.gdata.2016.12.002.

Reece, J. S., Bowen, B. W., Smith, D. G., & Larson, A. (2010). Molecular phylogenetics of moray eels (Muraenidae) demonstrates multiple origins of a shell-crushing jaw (Gymnomuraena, Echidna) and multiple colonizations of the Atlantic Ocean. *Molecular Phylogenetics and Evolution*, *57*(2), 829–835. https://doi.org/10.1016/j.ympev.2010.07.013

Reece, J. S., Bowen, B. W., Smith, D. G., & Larson, A. (2011). Comparative phylogeography of four Indo-Pacific moray eel species (Muraenidae) reveals comparable ocean-wide genetic connectivity despite five-fold differences in available adult habitat. *Marine Ecology Progress Series*, *437*(September 2011), 269–277. https://doi.org/10.3354/meps09248.

Reiner, F., 1996. Catálogo dos peixes do arquipélago de Cabo Verde. Publicações Avulsas do IPIMAR 2, Lisbon. 339 pp.

Ribout, C., Bech, N., Briand, M. J., Guyonnet, D., Letourneur, Y., Brischoux, F., & Bonnet, X. (2018). A lack of spatial genetic structure of *Gymnothorax chilospilus* (moray eel) suggests peculiar population functioning. *Biological Journal of the Linnean Society*, *125*(1), 142–151. https://doi.org/10.1093/BIOLINNEAN/BLY107.

Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X., & Rozas, R. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, *19*(18), 2496–2497. https://doi.org/10.1093/bioinformatics/btg359.

Saitou, N., & Nei, M. (1987). The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees1. *Mol. Biol. Evol*, *4*(4), 406–425.

Sam Wouthuyzen •, & Mochioka, N. (2014). Leptocephalus larvae of two moray eels (Anguilliformes; Muraenidae), Gymnothorax sagmacephalus, and Gymnothorax

albimarginatus, identified from morphometric and genetic evidence. *Ichthyological Research*, *61*(1), 32–41. https://doi.org/10.1007/s10228-013-0369-4.

Sambrook et al., J. (2001). Molecular cloning of Tupaia belangeri Chinensis neuropeptide Y and homology comparison with other analogs from primates. In *Dong wu xue yan jiu = Zoological research / "Dong wu xue yan jiu" bian Ji Wei yuan hui bian Ji* (Vol. 33, Issue 1). https://doi.org/10.3724/sp.j.1141.2012.01075.

Santini, F., Carnevale, G., & Sorenson, L. (2014). First multi-locus time tree of seabreams and porgies (Percomorpha: Sparidae). *Italian Journal of Zoology*, *81*(1), 55–71. https://doi.org/10.1080/11250003.2013.878960.

Shin, Y. J., Rochet, M. J., Jennings, S., Field, J. G., & Gislason, H. (2005). Using size-based indicators to evaluate the ecosystem effects of fishing. *ICES Journal of Marine Science*, *62*(3), 384–396. https://doi.org/10.1016/j.icesjms.2005.01.004.

Smith, D. G. (2012). A checklist of the moray eels of the world (Teleostei: Anguilliformes: Muraenidae). In *Zootaxa* (Vol. 64, Issue 3474).

https://doi.org/10.11646/zootaxa.3474.1.1

Stratoudakis, Y., Bernal, M., Borchers, D. L., & Borges, M. F. (2003). *Changes in the distribution of sardine eggs and larvae off Portugal, 1985–2000*. *2*(January 2001), 49–60.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. 2002 DNA points the way ahead in taxonomy. Nature 418, 479. doi:10.1038/418479a.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. 2003 A plea for DNA taxonomy. Trends Ecol. Evol. 18, 70–74. doi:10.1016/S0169-5347(02)00041-1.

Tawa, A., Kobayakawa, M., Yoshimura, T., & Mochioka, N. (2012). Identification of leptocephalus larvae of the tiger moray Scuticaria tigrina (Anguilliformes; Muraenidae) based on morphometric and genetic evidence. *Ichthyological Research*, *59*(4), 378–383. https://doi.org/10.1007/s10228-012-0292-0.

Teske, P. R., Cherry, M. I., & Matthee, C. A. (2003). Population genetics of the endangered Knysna seahorse, hippocampus capensis. *Molecular Ecology*, *12*(7), 1703–1715. https://doi.org/10.1046/j.1365-294X.2003.01852.x.

Vasconcelos, A. J. M., & Molina, W. F. (2009). Cytogenetical studies in five Atlantic Anguilliformes fishes. *Genetics and Molecular Biology*, *32*(1), 83–90. https://doi.org/10.1590/S1415-47572009005000015.

Wang, C. H., Kuo, C. H., Mok, H. K., & Lee, S. C. (2003). Molecular phylogeny of elopomorph fishes inferred from mitochondrial 125 ribosomal RNA sequences. *Zoologica Scripta*, *32*(3), 231–241. https://doi.org/10.1046/j.1463-6409.2003.00114.x.

Wang, F. Y., Tang, M. Y., & Yan, H. Y. (2011). A comparative study on the visual adaptations of four species of moray eel. *Vision Research*, *51*(9), 1099–1108. https://doi.org/10.1016/j.visres.2011.02.025.

Ward, R. D. (2000). Genetics in fisheries management. *Hydrobiologia*, *420*(1), 191–201. https://doi.org/10.1023/A:1003928327503.

Ward, Robert D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1847–1857. https://doi.org/10.1098/rstb.2005.1716.

Wirtz, P., Brito, A., Falcón, J. M., Freitas, R., Fricke, R., Monteiro, V., Reiner, F., & Tariche, O. (2013). The coastal fishes of the Cape Verde Islands - New records and an annotated checklist: (Pisces). *Spixiana*, *36*(1), 113–142.

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, *1*(1), 3–14. https://doi.org/10.1111/j.2041-210x.2009.00001.x.

8 Appendix

Appendix 1: Sampling locations and Islands

Locations	Variables					
	Total Count of W(g)	**Total Count of L (cm)**	**Total Count of SEXO**	**Total Count of Wgon (g)**	**Total Count of Wfig(g)**	**Total Count of EM**
Boavista	60	60	60	60	60	60
Maio	180	180	179	179	180	179
Santa Luzia	5229	6013	5225	5174	5176	5217
Calador	31	31	31	31	31	31
Colador/St. Luzia	8	8	8	8	8	8
Santo Antão	1199	1283	1199	1198	1198	1198
Banco Noroeste	21	21	21	21	21	21
Norte de S. Antão	30	30	30	30	30	30
Tarrafal de Monte Trigo	1148	1232	1148	1147	1147	1147
São Vicente	1302	1847	1301	1292	1301	1296
Baia das Gatas	264	289	264	264	264	264
Calhau	478	623	478	478	478	478
Canal de São Vicente.	29	29	29	29	29	29
Salamansa	374	651	374	370	373	374
São Pedro	157	251	156	151	157	151
São Vicente/Santa Luzia	641	656	641	641	641	641
Canalzinho	641	656	641	641	641	641
(blank)	58	58	58	58	58	58
inconnu	*29*	*29*	*29*	*29*	*29*	*29*
Total	**8640**	**10068**	**8643**	**8573**	**8585**	**8620**

Appendix 2: Estimate stages of maturity of gonads

Maturity Stage	Females	Males
I	Gonads small and string-like, sex cannot be determined with the naked eye	Gonads small and string-like, sex cannot be determined with the naked eye
1	Immature; gonads enlarged, slender, but the sex can be determined by eye	Immature; testicles extremely slender with ribbon-like appearance, but it is possible to distinguish the sex with the naked eye
2	First maturity; gonads larger, but eggs cannot be distinguished with the naked eye	Testicles larger; triangular in cross-section, no sperm in the central barrel.
3	Advanced maturity; larger gonads, eggs easily visible to the naked eye	In maturation, on squeezing the testicles, the sperm flow abundantly.
4	Maturation; ovaries much larger, translucent eggs were dislodging easily from follicle or loose in the ovarian lumen.	Maturation; testicles large, sperm were flowing abundantly.